hyperbolic models and numerical analysis for shallow water flows

Emmanuel Audusse

UNIVERSITÉ PARIS VI

PIERRE ET MARIE CURIE

- Laboratoire Jacques-Louis Lions -

THÈSE

pour obtenir le titre de

DOCTEUR

Spécialité : **Mathématiques appliquées**

présentée par

EMMANUEL AUDUSSE

Modélisation hyperbolique et analyse numérique pour les écoulements en eaux peu profondes

Directeur de thèse : Benoit PERTHAME

---

soutenue le 14 Septembre 2004 devant le jury composé de :

M.	**Bruno DESPRES**	Président
M.	**Rémi ABGRALL**	Rapporteur
M.	**Sebastian NOELLE**	Rapporteur
M.	**Benoit PERTHAME**	Examinateur
Mme.	**Marie-Odile BRISTEAU**	Examinateur
M.	**Jean-Frédéric GERBEAU**	Examinateur
M.	**Jean-Michel HERVOUET**	Examinateur

Des histoires comme celles de nos barrages, il n'y en a pas deux.
Marguerite Duras [Un barrage contre le pacifique]

Remerciements

En premier lieu, je tiens à remercier Benoit Perthame,
Qui m'a orienté vers le sujet de mon mémoire,
Un travail autour des équations de Saint-Venant,
Avant de me guider pendant trois ans dans ma
Thèse, proposant sans cesse de nouvelles pistes.
Il est aussi très important pour moi de remercier Marie
Odile Bristeau, qui a quotidiennement répondu à mes
Nombreuses questions, demandes et interrogations,
Soient plus de 1086 interventions en autant de jours...

Dire toute ma gratitude, également, à Bruno Despres, président du jury,
Et à Rémi Abgrall et Sebastian Noelle, rapporteurs de cette thèse.

Sans eux, certains points de ce travail seraient sans doute restés obscurs :
Alors un grand merci à Francois Bouchut, Jean-Frédéric Gerbeau et Rupert Klein.
Idem pour Jean-Michel Hervouet, avec qui la collaboration fut fructueuse.
Ne pas omettre non plus trois professeurs qui m'ont particulièrement marqué,
Tout mon parcours leur doit beaucoup : Mrs. Deyris, Le Bris et Balabane.

Viennent enfin tous ceux que j'ai cotoyés à l'INRIA,
Elise, Chiara, Iria, Cecile, Anka, Paola, Claire, Astrid,
Nuno, Mohamed, Miguel, Loic, Boris, Frédéric, Gabriel,
Au revoir et merci, à eux, ainsi qu'aux autres.
Nouvelles seront les eaux éclairant nos lendemains,
Tel ciel y offrira aussi d'autres perspectives.

Modélisation hyperbolique et analyse numérique pour les écoulements en eaux peu profondes

Résumé : Nous étudions dans cette thèse différentes lois de conservation hyperboliques associées à la modélisation des écoulements en eaux peu profondes.
Nous nous consacrons d'abord à l'analyse numérique du système de Saint-Venant avec termes sources. Nous présentons un schéma volumes finis bidimensionnel d'ordre 2, conservatif et consistant, qui s'appuie sur une interprétation cinétique du système et une méthode de reconstruction hydrostatique des variables aux interfaces. Ce schéma préserve la positivité de la hauteur d'eau et l'état stationnaire associé au lac au repos.
Nous étendons ensuite l'interprétation cinétique au couplage du système avec une équation de transport. Nous construisons un schéma volumes finis à deux pas de temps, qui permet de prendre en compte les différentes vitesses de propagation de l'information présentes dans le problème. Cette approche préserve les propriétes de stabilité du système et réduit sensiblement la diffusion numérique et les temps de calcul.
Nous proposons également un nouveau modèle de Saint-Venant multicouche, qui permet de retrouver des profils de vitesse non constants, tout en préservant le caractère invariant et bidimensionnel du domaine de définition. Nous présentons sa dérivation à partir des équations de Navier-Stokes et une étude de stabilité - énergie, hyperbolicité. Nous étudions également ses relations avec d'autres modèles fluides et sa mise en oeuvre numérique.
Enfin nous établissons un théorème d'unicité pour les lois de conservation scalaires avec flux discontinus. La preuve est basée sur l'utilisation d'une nouvelle famille d'entropies, qui constituent une adaptation naturelle des entropies de Kruzkov classiques au cas discontinu. Cette méthode permet de lever certaines hypothèses classiques sur le flux - convéxité, existence de bornes BV, nombre fini de discontinuités - et ne nécessite pas l'introduction d'une condition d'interface.

Mots clés : système de Saint-Venant, termes sources, modèle multicouche, équation de transport, loi de conservation scalaire, flux discontinus, volumes finis, schéma équilibre, interprétation cinétique, reconstruction hydrostatique, entropies de Kruzkov.

Hyperbolic Models and Numerical Analysis for Shallow Water Flows

Abstract : In this work we study some hyperbolic conservation laws related to shallow water flows.
First we consider the Saint-Venant system with source terms and we develop a second order bidimensional well-balanced finite volumes scheme that is based on a kinetic interpretation of the system and on a hydrostatic reconstruction of the interfaces values. The scheme is consistent and conservative and it preserves the non-negativity of the water height.
Then we extend the kinetic interpretation to the coupling with a transport equation. We construct a two time steps scheme that takes into account all the eigenvalues of the problem. This approach preserves the stability properties of the system and reduces the numerical diffusion and the computational cost.
We also present a new multilayer Saint-Venant system that allows us to obtain non constant vertical velocity profiles while preserving an invariant two dimensional domain of calculation. We present the derivation of the system and we study its stability - energy, hyperbolicity. We also investigate its relation with other fluid models and we perform its numerical implementation.
Finally we prove a uniqueness theorem for scalar conservation laws with discontinuous flux. Our proof uses a new family of entropies that are a natural way to adapt classical Kruzkov's entropies to the discontinuous case. This new method avoids the making of some classical hypothesis on the flux such as convexity, BV bounds or finite number of dicontinuities and does not need the definition of some interface condistion.

Key words : shallow water equations, source terms, multilayer model, transport equation, scalar conservation law, discontinuous flux, finite volumes, well-balanced scheme, kinetic interpretation, hydrostatic reconstruction, Kruzkov's entropies.

Table des matières

Chapitre 1

Présentation des travaux et Principaux résultats

1.1 Introduction

Nous étudions dans cette thèse différentes questions liées au système de Saint-Venant - ou *shallow water equations* - avec termes sources. Plus généralement, nous considérons son couplage avec une équation de transport ainsi que la dérivation et l'étude de nouveaux modèles de type Saint-Venant multicouche. La plupart des questions abordent des sujets fondamentaux de l'analyse numérique. Néanmoins elles ont également conduit à de nouvelles implémentations dans le code TELEMAC d'EDF.

Le système de Saint-Venant est un système hyperbolique, introduit à la fin du dix neuvième siècle, qui permet de modéliser les fluides géophysiques à surface libre en écoulements "peu profonds". Du fait de sa validité expérimentale et de son efficacité numérique largement reconnues, il est aujourd'hui très utilisé pour la simulation de nombreux phénomènes d'actualité : pollution environnementale, catastrophes naturelles, évolution climatique... Enoncer ces thématiques, c'est déjà dire que toute expérimentation en vraie grandeur est impossible à réaliser. On comprend alors pourquoi ces dernières années ont vu un engouement sans précédent pour la résolution numérique de ces équations, suscitant ainsi l'apparition de nombreuses questions nouvelles. Parallèllement, des progrès ont également été accomplis en amont : d'une part, dans l'étude mathématique du système de Saint-Venant proprement dit et, d'autre part, dans l'étude des relations entre les modèles de Saint-Venant et de Navier-Stokes.

La première partie de ce travail est consacrée à l'étude numérique des équations de Saint-Venant. Si les difficultés liées à l'aspect partiellement conservatif du système connaissent maintenant des réponses classiques - nous privilégierons ici une approche volumes finis - d'autres points restent plus ouverts. Une première question, la plus récurrente, déjà abordée dans quelques travaux théoriques à la fin des années 80 et qui a donné naissance au milieu des années 90 à la notion de schéma équilibre - ou *well balanced scheme* - est liée à la présence, dans le système de Saint-Venant, d'un terme source topographique. Ce terme est en effet à l'origine de difficultés numériques liées à l'existence d'états stationnaires non triviaux. Une deuxième question, moins souvent abordée mais tout aussi fondamentale, est la préservation au niveau numérique de deux propriétés de stabilité du système continu : la positivité de la hauteur d'eau dans l'écoulement pour tout temps et l'existence d'une inégalité d'entropie. Nous commencons ici par répondre à la deuxième question en adaptant au système de Saint-Venant les schémas cinétiques développés dans le cadre des équations d'Euler. Puis nous proposons pour le premier problème une solution originale, basée sur une reconstruction hydrostatique des variables et qui préserve les acquis de stabilité.
Dans de nombreuses applications, notamment les phénomènes de pollution, une équation de transport est ajoutée aux équations de Saint-Venant. Des problèmes identiques aux précédents se posent naturellement : conservation de la quantité de traceur, existence d'un domaine invariant dans lequel évolue la concentration du traceur... Mais il ap-

paraît aussi une nouvelle problématique liée à la différence de vitesse de propagation de l'information entre la partie hydrodynamique et la partie transport. Cette différence n'est pas préjudiciable à la stabilité du processus, mais, si elle n'est pas prise en compte, peut être à l'origine d'une importante diffusion numérique sur les quantités liées au traceur. A l'inverse, des techniques de discretization adaptées pourront, outre la résolution de ce problème, permettre de précieux gains en temps de calcul et en capacité de stockage. Là encore, dans la méthode que nous proposons, nous obtenons les propriétés de stabilité en étendant les schémas cinétiques à ce nouveau problème. Nous abordons ensuite les problèmes liés à la différence des célérités de l'information en introduisant une nouvelle méthode à deux pas de temps, préservant elle aussi la stabilité des schémas.

Dans la deuxième partie, nous abordons une question relative à l'étude mathématique des lois de conservation scalaire avec flux discontinus. Une des motivations provient du système de Saint-Venant et de son couplage avec une équation de transport. Nous nous placons ici dans le cadre d'un écoulement stationnaire monodimensionnel - dont l'existence est facile à établir - et nous nous intéressons au couplage avec une équation de transport. Le problème posé est alors celui de la résolution d'une équation hyperbolique scalaire dont le flux dépend explicitement de la variable d'espace, éventuellement de manière discontinue - puisque les solutions du système de Saint-Venant peuvent présenter des discontinuités. Dans le cas continu, Kruzkov, en utilisant la notion d'entropie mathématique, a démontré l'existence et l'unicité d'une solution à ce problème. Plus récemment, de nombreux travaux ont été consacrés au cas discontinu et des progrès notables sont acquis. Mais les résultats d'existence et d'unicité de solution sont encore parcellaires. Nous proposons ici une preuve d'unicité, utilisant une méthode naturellement adaptée au cas d'un flux discontinu, basée sur une adaptation des entropies de Kruzkov. Ceci nous conduit à considérer de nouvelles conditions sur la dépendance du flux par rapport aux inconnues du problème - moins contraignantes cependant que les classiques hypothèses de convexité - et nous permet de lever de nombreuses restrictions sur la dépendance spatiale du flux - existence de bornes BV, nombre fini de discontinuités - et sur la forme des solutions - existence de traces aux points de discontinuité du flux.

Enfin nous consacrons la troisième et dernière partie de cette thèse à la dérivation, puis à l'analyse mathématique et à la mise en oeuvre numérique, d'un modèle de Saint-Venant à plusieurs couches. Cette approche est l'occasion d'apporter une nouvelle contribution à l'étude des relations entre les systèmes de Saint-Venant et de Navier-Stokes incompressible à surface libre. Mais notre but est avant tout de proposer une alternative à la résolution des équations de Navier-Stokes tridimensionnelles qui soit plus précise que le système de Saint-Venant classique, ou, plus exactement, qui permette de retrouver des profils de vitesse - selon la verticale - proches de ceux des solutions de Navier-Stokes. En l'absence de viscosité, la dérivation formelle des équations de Saint-Venant à partir des équations de Navier-Stokes est très classique. Très récemment, deux travaux ont

étendus cette dérivation au cas visqueux. En nous inspirant des techniques qui y sont utilisées, mais en introduisant, dans le modèle et non dans la méthode numérique, une discretization dans la direction verticale, nous établissons ici un nouveau modèle de Saint-Venant multicouche qui préserve les acquis du système de Saint-Venant classique - réduction de la dimension du problème, invariance du domaine de calcul - tout en s'affranchissant de sa principale limitation, à savoir l'absence de dépendance verticale de la vitesse horizontale. Après avoir présenté les principales caractéristiques de ce nouveau modèle - energie, questions relatives à l'hyperbolicité... - nous montrons qu'il est possible d'étendre les schémas cinétiques à un tel problème. Des tests numériques monodimensionnels indiquent une bonne adéquation avec les solutions des équations de Navier-Stokes.

Avant de détailler chacune de ces études, nous voudrions souligner le fait que tous les développements numériques consacrés au système de Saint-Venant et à son couplage avec une équation de transport ont d'ores et déjà été intégrés dans le code TELEMAC, développé et géré par le Laboratoire Nationale d'Hydraulique et d'Environnement de EDF. Une version bidimensionnelle du modèle multicouche, en vue d'une intégration dans ce logiciel, est également à l'étude. Nous avons donc accordé une attention toute particulière à concilier la rigueur mathématique des schémas et leur robustesse avec la préservation de la simplicité des algorithmes utilisés. Dans la même optique, nous avons également validé nos schémas à la fois sur des cas-tests académiques pour lesquels une étude d'erreur sérieuse est possible et sur des cas-tests réels, provenant d'études réalisées par le LNHE, afin de vérifier que ces schémas pouvaient être utilisés de manière "industrielle" et avec des temps de calcul raisonnables.

1.2 Système de Saint-Venant et Volumes finis

1.2.1 Présentation et domaine de validité

Le système d'équations aux dérivées partielles qui nous intéressera tout au long de ce travail a été introduit en 1871 dans un Compte Rendu à l'Académie des Sciences rédigé par l'ingénieur des Ponts et Chaussées Adhémar Jean-Claude Barré de Saint-Venant [115]. Dans sa version monodimensionnelle initiale, il décrit l'écoulement de l'eau dans un canal rectangulaire à fond horizontal par l'intermédiaire de la hauteur d'eau $h(t,x) \geq 0$ et de la vitesse moyenne $u(t,x) \in \mathbb{R}$

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) = 0, \tag{1.2.1}$$

$$\frac{\partial hu}{\partial t} + \frac{\partial}{\partial x}\left(hu^2 + \frac{gh^2}{2}\right) = 0, \tag{1.2.2}$$

où g désigne la gravité. La seconde variable conservative est ici $q(t,x) = hu$, qui désigne la quantité de mouvement, ou débit.
Ici nous nous intéresserons à une version bidimensionnelle qui peut inclure plusieurs termes sources : termes moteurs de l'écoulement dûs à la topographie, termes dissipatifs de friction... Ce système décrit donc les écoulements dans des baies peu profondes ou dans des rivières quand les effets bidimensionnels - horizontaux - ne peuvent être négligés, toujours par l'intermédiaire de la hauteur d'eau $h(t,x,y) \geq 0$ et de la vitesse moyenne $\mathbf{u}(t,x,y) \in \mathbb{R}^2$

$$\frac{\partial h}{\partial t} + \mathrm{div}(h\mathbf{u}) = 0, \tag{1.2.3}$$

$$\frac{\partial h\mathbf{u}}{\partial t} + \mathrm{div}(h\mathbf{u} \otimes \mathbf{u}) + \nabla(\frac{g}{2}h^2) = -gh\nabla Z - \kappa\mathbf{u}, \tag{1.2.4}$$

où g désigne toujours la gravité, κ est un coefficient de friction et $Z(x,y)$ représente le profil du fond du canal - $h+Z$ est donc la cote de la surface libre de l'écoulement. Notons que la forme du terme de frottement présentée ici découle d'une étude asymptotique que nous détaillerons ultérieurement - voir Chapitre 6 - mais que ce n'est pas la forme qui est privilégiée par les ingénieurs. Ils lui préfèreront plutôt les coefficients empiriques dits de Manning, de Chézy ou de Strickler [66]. Notons également que l'introduction de termes plus spécifiques permet de prendre en compte d'autres phénomènes physiques - érosion/sédimentation si Z dépend du temps, introduction de termes visqueux - ou d'étendre la validité du système à d'autres types d'écoulements - force de Coriolis pour les écoulements océaniques et atmosphériques [51, 132], termes de frottements spécifiques pour des écoulements "solides" tels les avalanches [22, 61], introduction du module d'Young des parois dans les termes de pression pour les écoulements sanguins [47].
Précisons enfin que, dans toute la partie consacrée à l'étude numérique du système, nous ne considèrerons que le seul terme source topographique - voir ci-dessous et les Chapitres 2 et 3. A l'inverse une attention toute particulière sera portée aux différents termes dissipatifs - viscosité et frottement - lorsque nous dériverons, au Chapitre 6, un modèle de Saint-Venant multicouche.

L'intérêt majeur d'une approche de type Saint-Venant est de permettre, grâce à l'utilisation de la vitesse moyenne de l'écoulement et à l'introduction explicite de la hauteur d'eau comme inconnue, d'aborder des problèmes physiques tridimensionnels et instationnaires, posés sur des domaines mobiles, au travers de l'étude d'un système posé sur un domaine bi- (voir mono-) dimensionnel et invariant en temps.
Se pose néanmoins la question de la pertinence des informations apportées par une telle approche. Autrement dit : Quel est le domaine de validité du système de Saint-Venant ? Quand le système de Saint-Venant est-il d'une précision suffisante et quand faut-il recourir à des modèles plus précis - équations de Navier-Stokes par exemple ? Les limitations principales inhérentes au système de Saint-Venant classique sont au nombre

de deux : une perte d'information évidente due à l'utilisation de la vitesse moyenne de l'écoulement et, plus difficile à évaluer, le manque de pertinence des solutions dans le cas d'écoulements particulièrement hétérogènes.
L'expérience des ingénieurs apporte un premier élément de réponse à ces remarques puisque les solutions des équations de Saint-Venant sont en très bon accord avec les données expérimentales sur de nombreux exemples et suffisent souvent à fournir les informations désirées. Cette approche empirique ne peut néanmoins pas dispenser d'une étude théorique plus poussée et plusieurs auteurs ont déjà abordé la question des relations entre le système de Saint-Venant et les équations de Navier-Stokes - voir les récents travaux [50, 45]. Une des possibilités offertes par de telles études est de pouvoir, au travers d'une meilleure compréhension des systèmes, dériver de nouveaux modèles, qui conjuguent les avantages des différentes approches. Cette question fait l'objet du Chapitre 6.

1.2.2 Hyperbolicité, Stabilité, Equilibres

Nous présentons ici les principales propriétés du système de Saint-Venant. Les avoir a l'esprit sera utile par la suite puisque leur préservation guidera l'ensemble de notre démarche numérique.

Hyperbolicité. En dehors des zones sèches, le système de Saint-Venant (1.2.3)-(1.2.4) est un système strictement hyperbolique de lois de conservation du premier ordre avec termes sources. Il peut en effet s'écrire

$$\frac{\partial U}{\partial t} + \mathrm{div} F(U) = B(U), \tag{1.2.5}$$

où $U(t,x,y) = (h, q_x, q_y)^T$ est le vecteur des variables conservatives et où

$$F(U) = \begin{bmatrix} q_x & q_y \\ \frac{q_x^2}{h} + \frac{gh^2}{2} & \frac{q_x q_y}{h} \\ \frac{q_x q_y}{h} & \frac{q_y^2}{h} + \frac{gh^2}{2} \end{bmatrix} \quad \text{et} \quad S(U) = \begin{pmatrix} 0 \\ -gh\partial_x Z \\ -gh\partial_y Z \end{pmatrix}$$

désignent respectivement le flux et les termes sources - nous omettons ici les termes de frottement. Partant, des calculs classiques montrent que le système s'écrit encore

$$\frac{\partial U}{\partial t} + DF_x \frac{\partial U}{\partial x} + DF_y \frac{\partial U}{\partial y} = S(U), \tag{1.2.6}$$

où DF_x et DF_y sont les matrices jacobiennes du flux

$$DF_x = \begin{bmatrix} 0 & 1 & 0 \\ -\frac{q_x^2}{h^2} + gh & 2\frac{q_x}{h} & 0 \\ -\frac{q_x q_y}{h^2} & \frac{q_y}{h} & \frac{q_x}{h} \end{bmatrix} \quad \text{et} \quad DF_y = \begin{bmatrix} 0 & 0 & 1 \\ -\frac{q_x q_y}{h^2} & \frac{q_y}{h} & \frac{q_x}{h} \\ -\frac{q_y^2}{h^2} + gh & 0 & 2\frac{q_y}{h} \end{bmatrix}$$

Suivant les techniques habituelles [118], on introduit maintenant un vecteur $\xi \in \mathbb{R}^2$ et on définit $DF(\xi) = \xi_x DF_x + \xi_y DF_y$. Il vient que, pour tout $\xi \in \mathbb{R}^2$, la matrice $DF(\xi)$ possède trois valeurs propres définies par

$$\lambda_1(\xi) = u_\xi - c, \qquad \lambda_2(\xi) = u_\xi \qquad \text{et} \qquad \lambda_3(\xi) = u_\xi + c$$

où $u_\xi = \xi_x \frac{q_x}{h} + \xi_y \frac{q_y}{h}$ est la vitesse de l'écoulement dans la direction ξ et $c = \sqrt{gh}$ représente la célérité de l'information dans l'écoulement. Si la hauteur d'eau est non nulle, il apparaît clairement que, pour tout ξ non nul, $\lambda_1(\xi) < \lambda_2(\xi) < \lambda_3(\xi)$, ce qui démontre bien la stricte hyperbolicité du système.
L'étude des systèmes hyperboliques du premier ordre est actuellement un des grands champs d'investigation de l'étude des équations aux dérivées partielles. Ils apparaissent en effet dans la modélisation de nombreux phénomènes physiques : électromagnétisme, dynamique des matériaux hyperélastiques, phénomènes de transports et, bien sûr, dynamique des gaz en variables eulériennes, modélisée par les célèbres équations d'Euler - pour plus de détails, se référer à [118]. L'étude du système de Saint-Venant, en tant que système hyperbolique de lois de conservation, pourra donc bénéficier de l'expérience acquise dans l'étude de ces autres problèmes. En particulier, l'analogie évidente qui peut être faite avec les équations d'Euler, et notamment avec leur version isentropique, dont le système de Saint-Venant homogène n'est qu'un cas particulier, est à l'origine de plusieurs options adoptées dans ce travail. A titre d'exemple, notons le fait que les méthodes numériques utilisées devront être adaptées aux éventuelles discontinuités que les solutions du système de Saint-Venant, à l'instar de toute solution d'un système hyperbolique, pourront développer.
Cependant, le problème de la non hyperbolicité du système en présence de zones sèches est ici d'une importance capitale puisque de telles zones apparaissent dans nombre d'applications - marées, crues... La stabilité des méthodes numériques dans la résolution des zones sèches sera donc primordiale.

Stabilité. Nous avons déjà mentionné que les variables conservatives du système de Saint-Venant évoluaient dans un demi-espace inclus dans $\mathbb{R}^3$, puisque la hauteur d'eau reste positive - par la suite, comme ici, nous entendrons toujours ce terme au sens de positive ou nulle. Le fait que le système de Saint-Venant soit une loi de conservation assure également l'existence de bornes de type L^1, au moins pour la hauteur d'eau. Nous pouvons préciser cette assertion en exhibant une propriété de décroissance sur l'énergie du système. En effet le système de Saint-Venant (1.2.3)-(1.2.4) admet l'énergie suivante

$$E(t,x,y) = \frac{h|\mathbf{u}|^2}{2} + \frac{gh^2}{2} + ghZ \tag{1.2.7}$$

et des calculs classiques - par exemple la méthode de viscosité, voir [118] - montrent qu'elle vérifie l'inégalité suivante

$$\frac{\partial E}{\partial t} + \operatorname{div}\left[\left(E + \frac{gh^2}{2}\right)\mathbf{u}\right] \leq 0.$$

Nous rappelons que, en dimension un, cette inégalité devient une égalité pour des solutions suffisamment régulières et que, dans le cas du système homogène, l'énergie n'est qu'une des entropies qu'il convient d'associer au système pour que le problème soit bien posé - voir à ce sujet [118, 37] et [95, 96] pour le cas particulier des équations d'Euler isentropiques. Si cette seule inégalité n'est donc pas suffisante pour une étude mathématique rigoureuse, elle assure néanmoins la présence d'une borne supplémentaire sur la norme L^1 des inconnues du système et fournit des informations sur le choix d'une solution physique du problème.
Notons que, là encore, la présence éventuelle de zones sèches dans l'écoulement rend ces propriétés de stabilité, et notamment la positivité de la hauteur d'eau, primordiales. Cette question sera abordée au Chapitre 3.

Equilibres. L'existence d'états stationnaires non triviaux - entendez pour lesquels les inconnues ne sont pas constantes sur le domaine - est le principal sujet de la majorité des publications consacrées à Saint-Venant dans les dix dernières années - voir par exemple, mais la liste n'est pas exhaustive, les travaux [13, 62, 12, 91, 49, 84, 114, 33, 71, 108, 15, 59, 56, 7]. Du fait de la présence du terme source topographique, le système de Saint-Venant présente en effet la particularité de posséder des états stationnaires complexes. En dimension un, et pour des solutions régulières, ils sont par exemple caractérisés par les relations suivantes

$$\frac{\partial hu}{\partial x} = 0, \qquad \frac{\partial H}{\partial x} = 0 \tag{1.2.8}$$

où $H(t,x) = \frac{u^2}{2} + g(h+Z)$ désigne la charge hydraulique. Si les solutions présentent des discontinuités, il convient, au niveau des chocs, de remplacer ces relations par les relations de Rankine-Hugoniot correspondant à un choc stationnaire - voir [118].
La préservation numérique de ces états stationnaires n'a rien d'évident. Ils correspondent en effet à un équilibre entre termes de flux et termes sources, dont les discretizations sont habituellement décoréllées. Or dans le cas du système de Saint-Venant, certains d'entre eux présentent un intérêt majeur. A titre d'exemple, notons le fait que la préservation d'une zone au repos - cas particulier des relations (1.2.8) avec $u = 0$ - est importante pour au moins trois raisons : la crédibilité des résultats, la capture correcte des phénomènes de faible amplitude et le fait qu'une onde incidente n'interagit pas de la même manière avec une zone au repos ou une zone en mouvement.
En dimension deux, toujours pour des solutions régulières, les états stationnaires sont caractérisés par les relations suivantes

$$\mathrm{div}(h\mathbf{u}) = 0, \qquad \nabla H - \mathrm{curl}\mathbf{u} \begin{pmatrix} v \\ u \end{pmatrix} = 0, \tag{1.2.9}$$

où $H(x,y) = \frac{|\mathbf{u}|^2}{2} + g(h+Z)$ désigne toujours la charge hydraulique et où $\mathbf{u} = (u,v)^T$. On retrouve alors les équilibres monodimensionnels (1.2.8) lorsque l'on suit les lignes

de courant. Or celles-ci, pour des écoulements en géométrie complexe, ne peuvent être établies *a priori* et le maillage n'a donc aucune raison d'être correctement adapté. Il apparaît donc impossible de préserver numériquement les équilibres bidimensionnels. Excepté celui, déjà mentionné, qui correspond à une zone au repos, et dont la caractérisation est indépendante de la dimension considérée

$$h + Z = C_0, \qquad \mathbf{u} = 0. \tag{1.2.10}$$

La question de la préservation des états stationnaires est abordée aux Chapitres 2 et 3.

1.2.3 Les volumes finis

Les deux sections précédentes présentent deux visions différentes du système de Saint-Venant : sa dérivation et ses applications physiques mettent en évidence ses relations avec les équations de Navier-Stokes ; ses propriétés mathématiques soulignent l'analogie avec les équations d'Euler. Privilégier l'une ou l'autre, c'est aussi choisir telle ou telle voie de discretization. Ainsi le champ d'application commun qui existe avec les équations de Navier-Stokes incite plutôt à l'utilisation des éléments finis quand l'analogie avec les équations d'Euler conduit plutôt vers les volumes finis - nous ne mentionnons pas ici les différences finies, qui nécessitent l'utilisation de grilles structurées, et sont donc mal adaptées aux écoulements en géométries complexes.

Les éléments finis ont été abondamment utilisés dans l'étude du système de Saint-Venant. Plusieurs raisons à cela : leur antériorité - voir par exemple l'historique en introduction de la thèse de Proft [110] - ou leur utilisation dans des logiciels d'hydrodynamique résolvant à la fois Navier-Stokes et Saint-Venant - voir par exemple [66] pour leur utilisation dans le code TELEMAC. Nous référons également à [40, 78] pour des résultats plus récents. Abondamment validée, cette méthode fournit aujourd'hui encore des algorithmes rapides et des résultats souvent satisfaisants.
Néanmoins - et là encore, c'est l'analogie avec Navier-Stokes qui prévaut - la discretization par éléments finis est intimement liée à une formulation hauteur-vitesse du système de Saint-Venant

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) = 0, \tag{1.2.11}$$

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + g\frac{\partial h}{\partial x} = -g\frac{\partial Z}{\partial x}, \tag{1.2.12}$$

qui n'est équivalente à la formulation hauteur-débit habituelle (1.2.1)-(1.2.2) que pour des solutions suffisamment régulières - les conditions de Rankine-Hugoniot n'étant pas identiques pour les deux systèmes - ce qui crée des difficultés dans la résolution de phénomènes présentant des discontinuités - rupture de barrage, par exemple. De plus les méthodes classiques d'éléments finis ne sont naturellement très bien adaptées ni

au caractère discontinu des solutions, ni au caractère conservatif du système. Il est également difficile d'assurer certaines propriétés de stabilité, notamment la positivité de la hauteur d'eau.

La méthodes des volumes finis, que nous privilégierons ici, est elle intimement liée au caractère hyperbolique du système de Saint-Venant et au fait qu'il s'agit - au moins partiellement - d'une loi de conservation. En effet cette méthode, développée au préalable dans le cadre du traitement numérique des équations d'Euler, présente le grand intérêt d'être intrinséquement conservative et s'adapte très bien à l'aspect discontinu des solutions. Il s'agit ici de découper l'espace en cellules - construites par exemple à partir d'un maillage éléments finis classique - puis d'intégrer le système considéré sur chaque cellule et sur un pas de temps. Apparaissent alors les moyennes des solutions sur chaque cellule - d'où des solutions constantes par morceaux et donc naturellement discontinues - et des termes de bords, autrement dit les flux échangés entre les cellules au niveau de leur frontière, ou interface - ainsi le flux sortant d'une cellule est égal à celui qui rentre dans la cellule voisine, d'où un algorithme conservatif. En une dimension d'espace, un schéma volume fini explicite peut donc s'écrire sous la forme générale suivante

$$U_i^{n+1} - U_i^n + \sigma_i^n \left(F_{i+1/2}^n - F_{i-1/2}^n \right) = \Delta t^n S_i^n, \tag{1.2.13}$$

où $\sigma_i^n = \Delta t^n / \Delta x_i$ est le rapport entre les pas de temps et d'espace, U_i^n (resp. S_i^n) représente la moyenne de la solution (resp. du terme source) sur la cellule C_i au temps t_n, et oú $F_{i+1/2}^n$ représente le flux numérique à l'interface entre les cellules C_i et C_{i+1}. Ces flux aux interfaces font naturellement intervenir les valeurs des variables aux interfaces, qui ne sont pas connues. Développer une méthode volumes finis consiste donc à définir, à partir des valeurs des variables aux noeuds du maillage, un flux numérique qui soit consistant avec le système continu, et, si possible, préserve ses propriétés de stabilité. Nous nous restreindrons ici aux schémas dits à trois points. Les flux aux interfaces prennent alors la forme

$$F_{i+1/2}^n = \mathcal{F}(U_i^n, U_i^{n+1}). \tag{1.2.14}$$

Notons malgré tout deux inconvénients de cette méthode. Le premier, de portée générale, est que la plupart des schémas sont liés à un traitement explicite du problème, ce qui nécessite, via l'introduction d'une condition de CFL et afin d'obtenir un schéma stable, de considérer des pas de temps souvent petits et augmente donc le temps de calcul. Le second, plus spécifique au cas des systèmes non homogènes, est que la méthode des volumes finis est trés bien adaptée au traitement des termes de flux, mais que son extension aux termes sources, notamment dans le but de préserver des états stationnaires du système, est plus délicate. Nous mentionnons donc que certains auteurs [49, 58] préfèrent considérer un système de Saint-Venant homogène - en ajoutant une équation triviale sur l'évolution de la topographie - mais non conservatif - à cause du terme lié à la topographie - ce qui évite le problème de l'équilibre termes de flux - termes sources, mais soulève de nouvelles questions, liées à l'aparition de produits non

conservatifs. Puisqu'il apparaît dans la dérivation du système de Saint-Venant à partir des équations de Navier-Stokes [45] que celle-ci n'est valable que si le terme source topographique reste petit, nous conserverons ici l'écriture habituelle (1.2.1)-(1.2.2), ce qui revient à considérer le terme topographique comme une perturbation du système homogène.
Pour plus de détails sur les différents schémas existants, nous renvoyons le lecteur aux ouvrages [44, 54, 90, 124]. La mise en oeuvre de la méthode des volumes finis, dans le cadre bidimensionnel, est rappelée dans le Chapitre 3.

1.3 Un schéma équilibre positif d'ordre 2 pour le système de Saint-Venant avec termes sources sur maillages non structurés : Analyse et mise en oeuvre numérique

Dans la section précédente nous avons décrit le cadre discret dans lequel nous avons choisi de nous placer et nous avons pointé les questions principales qui se posent lorsque l'on aborde la discretization du système de Saint-Venant : l'hyperbolicité, le problème du vide, la positivité de la hauteur d'eau, la décroissance de l'énergie et la préservation d'équilibres non-triviaux. Dans la première partie de cette thèse, nous proposons, grâce à l'adaptation de travaux existants - les schémas cinétiques - et par l'introduction d'idées nouvelles - la reconstruction hydrostatique - un schéma volumes finis qui satisfait à toutes ces exigences.

1.3.1 Système de Saint-Venant homogène : Etat de l'art

Nous considérons ici le système de Saint-Venant homogène. Puisqu'il s'agit alors d'un système hyperbolique de lois de conservation du premier ordre classique, de nombreux schémas numériques peuvent être utilisés. Tous ceux que nous citons ici ont connu des extensions destinées à construire des schémas équilibres, et sur lesquelles nous reviendrons ultérieurement. Commencons par le plus ancien, le schéma de Godunov [55], qui repose sur une résolution exacte du problème de Riemann à chaque interface du maillage - le problème de Riemann est un problème de Cauchy pour lequel la solution initiale est constante par morceaux et ne contient qu'une seule discontinuité - et qui possède donc toutes les propriétes de stabilité voulues, mais qui nécessite des temps de calcul importants. Citons également les solveurs de Riemann approchés - parmi eux, le plus célèbre est sans doute le schéma proposé par Roe dans [112] - qui utilisent la résolution exacte de problèmes de Riemann plus simples et constituent donc une tentative pour corriger les défauts du schéma de Godunov, mais échouent bien souvent à préserver les résultats de stabilité - voir les travaux récents de Benkhaldoun et al. [101] sur ce problème. Pour une utilisation de ces schémas dans le cadre du système

de Saint-Venant homogène, voir par exemple [3]. Evoquons enfin une autre approche, la méthode dite de *flux vector splitting* - voir [128, 119] - qui permet de généraliser la notion de schéma décentré et dont les schémas cinétiques sont un exemple particulier. Bien d'autres schémas existent - pour un plus large apercu nous renvoyons le lecteur à [44, 54, 90, 124]. Néanmoins nous allons montrer que les schémas cinétiques constituent - et c'est tout particulièrement vrai dans le cas de la résolution du système de Saint-Venant - un remarquable compromis entre rapidité, précision et stabilité.

1.3.2 Système de Saint-Venant homogène : Les schémas cinétiques

Les premiers schémas cinétiques ont été introduits dès les années 70 par Sanders et Prendergast [116]. Dans les années 80, plusieurs versions nouvelles ont été proposées - voir références dans [106]. C'est néanmoins en 1990 qu'un pas décisif est franchi, quand Perthame [104] montre qu'on peut considérer, dans l'interprétation cinétique des équations, un équilibre de Gibbs à support compact - voir ci-dessous pour une définition précise - en lieu et place de la maxwellienne habituellement utilisée. Ceci va permettre d'obtenir de nouvelles propriétés de stabilité discrètes et d'améliorer les résultats numériques tout en réduisant le temps de calcul. Dès lors, les schémas cinétiques ont été abondamment utilisés, notamment dans l'étude des équations d'Euler - voir par exemple [79, 111]. Sur la base de l'analogie existant entre ces équations et le système de Saint-Venant, nous avons choisi de les utiliser ici. Notons de plus que, dans la résolution des équations d'Euler, le principal inconvénient des schémas cinétiques est leur manque de précision dans la résolution des discontinuités de contact - voir encore [106]. Or le système de Saint-Venant homogène, contrairement aux équations d'Euler, ne donne pas naissance à de telles ondes.

L'idée de base des schémas cinétiques est d'utiliser les liens qui unissent la mécanique des fluides macroscopique - équations d'Euler ou de Navier-Stokes, système de Saint-Venant - et les équations cinétiques de type Boltzmann - voir par exemple [31] pour plus de détails sur ces équations. Tout le processus qui va être décrit ici repose donc sur la définition d'un équilibre de Gibbs - autrement dit d'une densité microscopique de particule d'une forme particulière

$$M(t,x,y,\xi) = M(h(t,x,y), \xi - \mathbf{u}(t,x,y)) = \frac{h(t,x,y)}{\tilde{c}^2}\ \chi(\frac{\xi - \mathbf{u}(t,x,y)}{\tilde{c}}), \qquad (1.3.1)$$

où $\sqrt{2}\tilde{c} = \sqrt{gh}$ est la célérité de l'information dans l'écoulement et où la fonction χ est une probabilité à double symmétrie axiale dont les deuxièmes moments vérifient certaines propriétés particulières. Cette définition permet d'établir un lien entre les niveaux cinétiques et macroscopiques, via le théorème suivant, que nous établissons au Chapitre 3

Theorème 1.3.1 (Chapitre 3, Page 80)
Les fonctions $(h, \mathbf{q})$ sont solutions du système de Saint-Venant avec termes sources (1.2.3)-(1.2.4) si et seulement si $M(t,x,y,\xi)$ est solution de l'équation cinétique

$$\frac{\partial M}{\partial t} + \mathrm{div}(\xi M) - g\nabla Z \cdot \nabla_\xi M = Q(t,x,y,\xi), \tag{1.3.2}$$

pour un terme de collision $Q(t,x,y,\xi)$ qui satisfait, pour presque tout (t,x,y),

$$\int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ \xi \end{pmatrix} Q \; d\xi = 0. \tag{1.3.3}$$

L'idée de la preuve est que le système de Saint-Venant peut être vu comme l'intégrale en ξ, et contre le vecteur $(1,\xi)^T$, de l'équation cinétique (1.3.2).
Nous avons mentionné ici les termes sources afin de montrer qu'il peuvent être intégrés dans l'interprétation cinétique, mais, comme cela avait été le cas au début de ce chapitre, nous nous restreignons maintenant de nouveau au cas du système de Saint-Venant homogène. Ceci permet de dissocier les problèmes liés à la stabilité de ceux relatifs à la préservation des équilibres, cette dernière question n'ayant plus lieu d'être. Pour ξ fixé, et si on ne tient pas compte de son second membre, l'équation (1.3.2) est alors une simple équation de transport linéaire. La question de sa discretization dans le cadre des volumes finis est classique et de nombreux schémas numériques sont à notre disposition. Nous choisirons ici un schéma décentré.

$$f_i^{n+1,-} - M_i^n + \sigma_i^n \xi_+ (M_i^n - M_{i-1}^n) + \sigma_i^n \xi_- (M_i^n - M_{i+1}^n) = 0 \tag{1.3.4}$$

Il suffit alors d'intégrer ce schéma numérique, toujours en ξ et contre le vecteur $(1,\xi)^T$, pour obtenir un schéma numérique consistant avec le système de Saint-Venant homogène. Comme annoncé au préalable, le schéma résultant fait partie de la famille des *flux vector splitting schemes*, puisque le flux à l'interface (1.2.14) peut s'écrire

$$F_{i+\frac{1}{2}}^n = F^+(U_i^n) + F^-(U_{i+1}^n),$$

où l'interprétation cinétique apparaît au niveau de l'écriture de chacun des demi-flux

$$F^\pm(U_i^n) = \int_{\substack{\xi \geq 0 \\ \xi \leq 0}} \xi \begin{pmatrix} 1 \\ \xi \end{pmatrix} \; M_i^n(\xi)\, d\xi. \tag{1.3.5}$$

Notons que le second membre de (1.3.2), qui n'est pas pris explicitement en compte dans le schéma, intervient implicitement dans l'étape de projection qui a lieu au temps t^{n+1}. En effet, la densité de particules $f_i^{n+1,-}$ obtenue par le schéma (1.3.4) n'a aucune raison d'être un équilibre de Gibbs, mais permet tout de même de recouvrir une quantité macroscopique U_i^{n+1} consistante. En déduire alors un nouvel équilibre de Gibbs

M_i^{n+1} par la formule (1.3.1) revient à prendre en compte, en une seule fois, toutes les collisions contenues dans le second membre.

Les conditions posées sur l'équilibre de Gibbs autorisent de nombreux choix possibles. Ici nous imposerons une condition supplémentaire, à savoir que la probabilité χ, dans la lignée de l'idée introduite par Perthame dans [104], sera prise à support compact. Cette dernière condition rend le processus brièvement décrit ici particulièrement intéressant car il permet alors d'allier précision et rapidité avec des propriétés de stabilité, ce que nous établissons dans les propositions et théorèmes suivants

Theorème 1.3.2 (Chapitre 3, Page 87)
Le schéma cinétique est consistant et conservatif. Il assure la positivité de la hauteur d'eau sous une condition de CFL. Cette condition de CFL est liée, entre autres, à la taille du support de l'équilibre de Gibbs choisi et assure la stabilité du schéma décentré au niveau cinétique.

Theorème 1.3.3 (Chapitre 3, Page 88)
De plus, pour un choix particulier de l'équilibre de Gibbs, le schéma cinétique satisfait une inégalité d'entropie discrète associée à l'énergie (1.2.7).

Proposition 1.3.1 (Chapitre 3, Page 85)
Pour certains choix de l'équilibre de Gibbs, l'interprétation cinétique est numériquement "transparente" : les intégrales en ξ (1.3.5) peuvent être calculées analytiquement et la vitesse cinétique ξ n'apparaît donc pas explicitement dans le schéma cinétique. L'équilibre de Gibbs du théorème précédent vérifie cette propriété.

En une dimension d'espace, ou en deux dimensions d'espace, mais pour le choix particulier de l'équilibre de Gibbs mentionné au Théorème 1.3.3 - qui se trouve être invariant par rotation - les preuves de la positivité de la hauteur d'eau et de l'existence d'une inégalité d'entropie discrète peuvent être établies au niveau cinétique. La positivité de la hauteur d'eau résulte alors du fait que, grâce à la limitation du pas de temps via la condition de CFL, la densité de particules $f_i^{n+1,-}$ est obtenue par une simple combinaison convexe de densités de particules prises au temps t^n, et donc, par récurrence, des densités de particules initiales, supposées positives. L'existence d'une inégalité d'entropie est liée au choix d'un équilibre de Gibbs qui, parmi toutes les densités de particules qui permettent de recouvrer les bonnes quantités macroscopiques, minimise l'énergie du système. Ceci assure que la décroissance de l'entropie, vérifiée par le schéma décentré (1.3.4), est encore vraie lors de l'étape de projection qui consiste à remplacer la densité $f_i^{n+1,-}$ par l'équilibre M_i^{n+1}. Pour des raisons de rapidité des calculs et de précision des résultats, nous présentons une implémentation bidimensionnelle légèrement différente, qui oblige à démontrer la positivité de la hauteur d'eau au niveau macroscopique. Les calculs sont alors un peu plus techniques et sont présentés au Chapitre 3.

Pour clore cette section, notons que l'interprétation cinétique n'est pas seulement un moyen de définir des schémas performants. En effet, une version plus complète de l'interprétation cinétique des modéles fluides a permis à Lions et al. [96, 95] d'établir en 1996 la première démonstration de l'existence d'une solution des équations d'Euler isentropiques - et donc du système de Saint-Venant homogène (1.2.1)-(1.2.2) - après apparition de singularités.

1.3.3 Décentrement des sources aux interfaces : Etat de l'art

Revenons maintenant au cas général du système de Saint-Venant non-homogène. Nous avons vu que la préservation des équilibres, et tout particulièrement de l'équilibre au repos, devenait alors un enjeu crucial. Plusieurs voies ont été suivies pour aboutir à ce résultat.
En 1993, dans [13], Bermudez et Vazquez proposent une méthode pour rendre compatible la discretization des termes de flux et des termes sources. Partant d'un schéma de Roe classique [112], il s'agit d'étendre la notion de décentrage aux termes sources, en les projetant sur les vecteurs propres de la matrice jacobienne du flux. Ils obtiennent ainsi la préservation de l'équilibre au repos. Ce travail a été étendu au cas bidimensionnel dans [12] puis la méthode a été appliquée au traitement du terme de frottement dans [129]. Chacon Rebollo et al. [34, 33] ont montré que la méthode pouvait être appliquée à d'autres flux que celui de Roe - par exemple aux schémas basés sur un *flux vector splitting* - et obtenu de bons résultats numériques. Néanmoins, dans tous ses travaux, les algorithmes de calcul du terme source sont complexes et la positivité de la hauteur d'eau n'apparaît pas clairement. Une autre manière d'adapter les flux de Roe ou de Godunov au cas non homogène a été proposée par Jin [71] : il s'agit cette fois d'utiliser dans la discretization du terme source les valeurs des variables aux interfaces fournies par le solveur lors de la résolution du problème de Riemann homogène. Le traitement du terme source est alors très simple, mais la méthode nécessite le calcul explicite de la solution du problème de Riemann à l'interface et conserve les inconvénients du solveur initial.
A la suite des travaux publiés à partir de 1996 par Greenberg et Leroux sur une adaptation du schéma de Godunov aux lois de conservations scalaires avec termes sources [62, 63], et plus généralement dans la lignée des travaux de Isaacson et Temple sur les problèmes résonnants [69], d'autres auteurs ont préféré construire des schémas basés sur l'étude du problème de Riemann pour la forme homogène non-conservative du système de Saint-Venant, évoquée dans la Section 1.2.3. Cette approche a d'abord été développée dans le cadre scalaire [59, 56]. Elle a ensuite été utilisée dans le traitement numérique de la gravité dans les équations d'Euler [26], avant d'être adaptée au système de Saint-Venant avec termes sources : Gosse [58] décrit ainsi une méthode non-conservative basée sur un schéma de Roe et Gallouët et. al. [49] proposent un schéma VFRoe-ncv. Gosse [57] a ensuite appliqué cette idée aux schémas basés sur

un *flux vector splitting*. Toutes ces approches préservent tous les équilibres monodimensionels (1.2.8). Néanmoins certaines ne peuvent pas être appliquées dans le cas de termes sources raides et, bien que des tests avec apparition de zones sèches soient présentés, la positivité de la hauteur d'eau n'est pas clairement établie. D'autre part, là encore, les algorithmes sont souvent complexes, du fait de la présence de produits non-conservatifs. Enfin, dans le contexte particulier des *flux vector splitting schemes*, l'extension bidimensionnelle n'est présentée que sur des grilles cartésiennes. Notons que, très récemment, Castro et Pares [29] ont unifié les deux "écoles" précédentes en construisant un schéma de Roe pour le système homogène non conservatif, puis en l'identifiant au schéma introduit dans [13].
De l'autre côté de l'Atlantique, R.J. LeVeque s'est lui aussi penché sur la question du traitement numérique des termes sources dans les lois de conservation hyperboliques depuis de nombreuses années. Dès 1990, il s'intéressait au cas des lois de conservation scalaires en présence de réactions chimiques et montrait dans [92] que les méthodes classiques n'étaient pas satisfaisantes en présence de termes sources raides. Son intérêt pour les équations de Saint-Venant est plus récent. Dans la lignée de travaux effectués dans le cadre homogène [93], il propose dans [91] une adaptation au cas non homogène : dans l'expression des flux, les variables classiques - moyennes sur les cellules - sont remplacées par des variables reconstruites de manière à préserver les équilibres. Nous verrons que cette approche présente des similitudes avec la reconstruction hydrostatique que nous détaillerons plus loin. Néanmoins, outre le fait que le schéma de base est de type Godunov et donc nécessite des temps de calcul importants, plusieurs limitations apparaissent, notamment le fait que la reconstruction des variables est très simple et ne permet en aucun cas de préserver les propriétés de stabilité du solveur homogène, notamment la positivité de la hauteur d'eau. Notons aussi le fait que la méthode est développée sur des grilles cartésiennes et que la topographie est donnée aux interfaces, information dont nous ne disposons pas lorsqu'on travaille en deux dimensions avec des grilles non-structurées. Ces travaux ont été repris récemment dans [11]. Notons également que plusieurs auteurs ont adapté la notion de *central scheme* au système de Saint-Venant avec termes sources - voir [84, 114] - mais que, là encore, une extension bidimensionnelle nécessite l'utilisation de grilles cartésiennes.
Mentionnons enfin que, dans le cadre scalaire, Perthame et Simeoni [109] ont introduit la méthode *upwind interface source* et ont établi un résultat de convergence. Ils ont aussi montré que certains des travaux précédents [71, 91] pouvaient être considérés comme des cas particuliers de cette méthode.

Même si la liste précédente n'est sans doute pas exhaustive, il apparaît donc qu'il existe de nombreuses méthodes adaptées au cas du système de Saint-Venant avec termes sources, certaines permettant même de préserver non seulement les états stationnaires associés au lac au repos (1.2.10), mais également tous les états stationnaires monodimensionnels (1.2.8) du système. Néanmoins il apparaît également que ces méthodes font souvent appel à des algorithmes complexes et, surtout, que certaines conditions de sta-

bilité, notamment la positivité de la hauteur d'eau, ne sont pas clairement démontrées. L'absence d'intérêt pour cette propriété pourtant importantes s'explique, entre autres, par le fait que nombre de ces méthodes sont basées sur l'utilisation des flux de Roe, qui, même dans le cas homogène, n'assurent pas la positivité de la hauteur d'eau. L'intérêt potentiel des schémas cinétiques apparaît donc ici très nettement : réunissant les propriétés de stabilité des schémas de Godunov et de simplicité des schémas de Roe, ils pourraient être à la base de schémas équilibres adaptés au cas non homogènes et qui préserveraient à la fois la positivité de la hauteur d'eau et la simplicité des algorithmes utilisés.

Une première possibilité consiste à intégrer le terme source topographique dans l'interprétation cinétique discrète - nous avons vu au Théorème 1.3.1 que cela était possible pour le système continu. Dans [108], Perthame et Simeoni étudient ainsi la possibilité, pour une "particule" cinétique, de passer, ou non, d'une cellule à l'autre, compte tenu du rapport existant entre son énergie cinétique propre et l'énergie potentielle liée à la différence de topographie entre les deux cellules. Cette idée permet de développer un schéma monodimensionnel positif et entropique et préservant l'équilibre du lac au repos (1.2.10). Cependant ce schéma utilise des algorithmes complexes, difficiles à mettre en oeuvre en dimension deux. De plus la Proposition 1.3.1 n'est plus vérifiée. Le calcul des solutions nécessite donc plusieurs intégrations numériques, ce qui induit des temps de calcul extrêmement longs. L'idée d'introduire la variation de topographie dans l'équation cinétique est également à la base des travaux de Xu [131], mais là encore, les algorithmes présentés s'avèrent fort complexes. Notons enfin que, dans le cas scalaire, une interprétation cinétique du terme source a permis à Botchorishvili et al. [15] de formuler un schéma qui préserve tous les équilibres et vérifie toutes les entropies discrètes, et pour lequel une preuve de convergence peut être établie.

1.3.4 Décentrement des sources aux interfaces : La reconstruction hydrostatique

Au vu des intéressantes propriétés de stabilité des schémas cinétiques mais devant la compléxité algorithmique qui résulte d'une approche cinétique du terme source, nous avons choisi ici de suivre une autre voie : puisque nous disposions déjà d'un solveur homogène stable, notre idée a été de développer une méthode, indépendante du solveur homogène choisi, mais qui adapterait ce dernier au cas non-homogène en assurant la préservation de certains états stationnaires (1.2.9) du système continu, et plus particulièrement de l'équilibre associé au lac au repos (1.2.10), tout en conservant intactes les propriétés de stabilité obtenues au préalable dans le cas homogène.
La méthode que nous proposons, dite de "reconstruction hydrostatique", repose sur l'idée suivante, présentée ici en une dimension d'espace, pour plus de simplicité. Les écoulements proches de l'équilibre (1.2.10) sont presque hydrostatiques, i.e. caractérisés par la relation $u \ll \sqrt{gh}$. A la limite, la hauteur d'eau doit donc être telle que le flux

de quantité de mouvement et le terme source topographique s'équilibrent

$$\partial_x \left(\frac{g\underline{h}^2}{2} \right) = -\underline{h}\, g z_x \,. \tag{1.3.6}$$

Après intégration sur une cellule nous obtenons donc une approximation cohérente du terme source topographique

$$-\int_{x_{i-1/2}}^{x_{i+1/2}} \underline{h}\, g z_x \; dx = \frac{g}{2}\underline{h}^{*2}_{i+1/2-} - \frac{g}{2}\underline{h}^{*2}_{i-1/2+} \,, \tag{1.3.7}$$

qui nécessite la définition de hauteurs d'eau de part et d'autre des interfaces. Un moyen simple de les obtenir est de considérer la forme ponctuelle (1.2.10) des équilibres qui nous intéressent, ce qui mène à

$$\underline{h}^*_{i+1/2-} = h_i + z_i - z^*_{i+1/2}, \qquad \underline{h}^*_{i+1/2+} = h_{i+1} + z_{i+1} - z^*_{i+1/2}. \tag{1.3.8}$$

Ne reste plus alors qu'à définir une topographie à l'interface. Nous considérons ici la définition suivante

$$z^*_{i+\frac{1}{2}} = max(z_i, z_{i+1}), \tag{1.3.9}$$

qui permet d'obtenir des bornes supérieures pour les hauteurs d'eau reconstruites. Nous introduisons également des bornes inférieures en utilisant dans notre schéma les parties positives des hauteurs d'eau reconstruites

$$h^*_{i+1/2\pm} = max(\underline{h}^*_{i+1/2\pm}, 0). \tag{1.3.10}$$

Nous considérons alors la discretization du terme source (1.3.7) et nous construisons les flux (1.2.14) à partir des valeurs reconstruites aux interfaces (1.3.10) - notons que, dans tout le processus de reconstruction, seule la hauteur d'eau est affectée, la vitesse restant inchangée. Cette construction trouve son aboutissement - et sa justification - dans le théorème

Theorème 1.3.4 (Chapitre 2, Page 60)
Considérant un flux numérique consistant avec le problème homogène, qui préserve la positivité de la hauteur d'eau et satisfait une inégalité d'entropie discrète associée à l'énergie (1.2.7), le schéma volumes finis semi-discret basé sur la reconstruction hydrostatique des variables (1.3.8)-(1.3.10) et sur la prise en compte du terme source sous la forme (1.3.7)
(i) préserve la positivité de la hauteur d'eau,
(ii) préserve les états stationnaires associés au lac au repos (1.2.10),
(iii) est consistant avec le système de Saint-Venant (1.2.3)-(1.2.4),
(iv) satisfait une inégalité d'entropie semi-discrète associée à l'énergie (1.2.7).

Le point clé de la démonstration de la préservation de la positivité de la hauteur d'eau est que les définitions proposées pour la topographie et les hauteurs d'eau aux interfaces assurent

$$0 \leq min(h^*_{i-1/2+}, h^*_{i+1/2-}) \leq max(h^*_{i-1/2+}, h^*_{i+1/2-}) \leq h_i. \tag{1.3.11}$$

La preuve de la préservation des états stationnaires liés au repos est dûe au fait que, par construction,

$$h_i + z_i = h_{i+1} + z_{i+1} \Rightarrow h^*_{i+1/2-} = h^*_{i+1/2+}, \tag{1.3.12}$$

ce qui entraîne l'annulation des flux de masse et l'égalité entre flux de quantité de mouvement et termes sources.
Les preuves de la consistance et de l'existence d'une inégalité d'entropie semi-discrète utilisent des techniques initialement développées dans [18] et sont détaillées au Chapitre 2. Notons néanmoins que l'établissement de l'inégalité d'entropie est intimement liée à la nature semi-discrète du schéma. Ainsi le passage au cas discret conduit à une éventuelle perte de l'inégalité d'entropie, quelque soit le pas de temps choisi. On établit ainsi la proposition

Proposition 1.3.2 (Chapitre 2, Page 62)
Considérant un flux numérique consistant avec le problème homogène et qui préserve la positivité de la hauteur d'eau, le schéma volumes finis basé sur la reconstruction hydrostatique des variables
(i) préserve la positivité de la hauteur d'eau sous la même condition de CFL que le schéma initial,
(ii) préserve les états stationnaires associés au lac au repos (1.2.10),
(iii) est consistant avec le système de Saint-Venant (1.2.3)-(1.2.4).

1.3.5 Le schéma d'ordre 2

Afin d'améliorer la précision des résultats, nous avons développé une extension formelle du schéma précédent à l'ordre 2. La méthode utilisée repose sur des notions classiques - approche MUSCL [127]. Le caractère conservatif de la montée en ordre - propriété triviale en dimension 1, mais plus délicate à obtenir en dimensions supérieures - fait appel à des techniques développées dans [107].
Il est ici nécessaire que la reconstruction d'ordre 2 préserve également l'équilibre du lac au repos et la positivité de la hauteur d'eau. Une autre question est liée au fait que les variables h, z et $h + z$ ne peuvent être reconstruites séparément. Nous démontrons que la positivité de la hauteur d'eau, ainsi que la préservation de l'équilibre du lac au repos, sont vérifiées si on choisit de reconstruire les variables h - dont on veut préserver la positivité - et $h + z$ - liée à l'équilibre - puis d'en déduire une reconstruction de la topographie du fond z. En effet, dans le cas d'une interface sec / mouillé, tout autre choix conduit à la négation d'un des points requis. Une fois cette reconstruction liée à la montée en ordre effectuée, une deuxième reconstruction, la reconstruction

hydrostatique, est appliquée, de manière identique à ce qui a été présenté dans la section précédente.
Enfin, il est nécessaire que le schéma d'ordre 2 soit consistant avec le système de Saint-Venant. Or si les techniques classiques fournissent cette consistance pour les termes de flux, il n'en va pas de même pour les termes sources. Ainsi, dans le cas présent, le terme source décentré, issu du schéma d'ordre 1, est toujours indispensable à la préservation des équilibres, mais n'assure plus la consistance du schéma. Il est donc nécessaire d'introduire un second terme source, centré sur la cellule, construit à partir des reconstructions d'ordre 2, et présenté ici en dimension un

$$S_{ci}^n = g\frac{h_{i-\frac{1}{2}}^n + h_{i+\frac{1}{2}}^n}{2}(z_{i-\frac{1}{2}}^n - z_{i+\frac{1}{2}}^n) \tag{1.3.13}$$

Finalement le schéma d'ordre 2 que nous proposons s'écrit en une dimension d'espace

$$U_i^{n+1} - U_i^n + \sigma_i^n \left(F^+(U_{i+\frac{1}{2},-}^{*n}) + F^+(U_{i+\frac{1}{2},+}^{*n}) - F^+(U_{i-\frac{1}{2},-}^{*n}) - F^+(U_{i-\frac{1}{2},+}^{*n}) \right) = \Delta t^n S_i^n \quad + \quad \Delta t^n S_{ci}^n,$$

où S_i^n et S_{ci}^n désignent respectivement le terme source décentré (1.3.7), basé sur les reconstructions d'ordre 2 hydrostatiques, et le terme source centré (1.3.13), basé sur les reconstructions d'ordre 2. Les variables $U_{i\pm\frac{1}{2},\pm}^{*n}$ dénotent les reconstructions d'ordre 2 hydrostatiques qui interviennent dans la définition des flux (1.3.5). Il est alors possible d'établir le théorème

Theorème 1.3.5 (Chapitre 3, Page 94)
Le schéma d'ordre 2 basé sur le solveur cinétique et sur la reconstruction hydrostatique
(i) préserve la positivité de la hauteur d'eau sous une condition de CFL,
(ii) préserve les états stationnaires associés au lac au repos (1.2.10),
(iii) est consistant avec le système de Saint-Venant (1.2.3)-(1.2.4).

Nous présentons sur la Figure 1.3.1 une illustration de la méthode sur un cas test "industriel" : la simulation de l'écoulement lié à la rupture du barrage de Malpasset, dans le sud de la France, en 1959. Cet exemple met en jeu une géométrie bidimensionnelle complexe et une topographie très irrégulière. De plus, il présente de nombreuses zones de recouvrement-découvrement de zones sèches - phénomènes liés à l'innondation qui a suivi la rupture du barrage - ainsi qu'une zone au repos - la mer, en haut, à droite, qui, à l'instant choisi, n'a pas encore été atteinte par lécoulement. La connaissance du temps d'arrivée de l'écoulement sur certaines zones stratégiques - villes, transformateurs - permet en outre de comparer les résultats numériquesaux relevés de l'époque. D'autres exemples numériques sont présentés aux Chapitres 2 et 3.

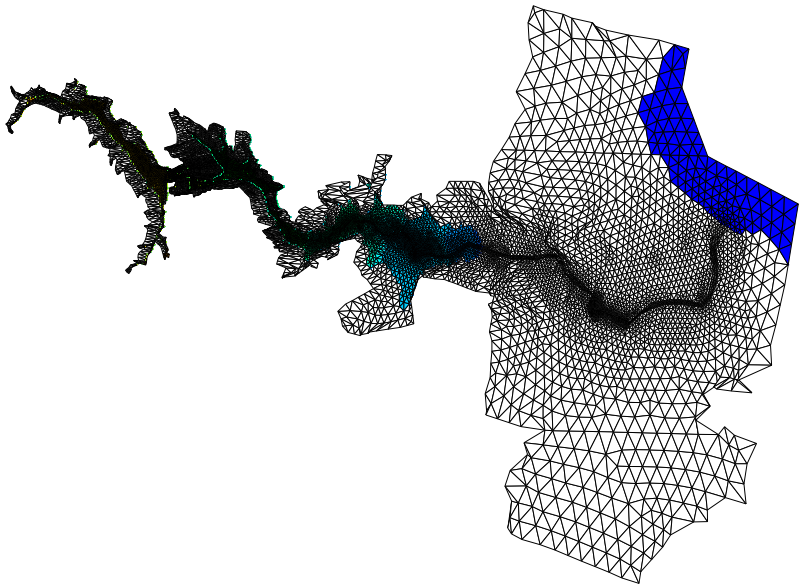


FIG. 1.3.1: Rupture du barrage de Malpasset - Surface libre

1.4 Un schéma à deux pas de temps pour le système Saint-Venant / transport : Analyse et mise en oeuvre numérique

Dans de nombreuses applications, au premier rang desquelles on trouve tous les problèmes liés à la pollution des eaux, le système de Saint-Venant est couplé avec une équation de transport d'un traceur passif. En 1d, on obtient alors le système suivant

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) = 0, \tag{1.4.1}$$

$$\frac{\partial hu}{\partial t} + \frac{\partial}{\partial x}\left(hu^2 + \frac{gh^2}{2}\right) = -gh\frac{\partial Z}{\partial x}, \tag{1.4.2}$$

$$\frac{\partial}{\partial t}(hT) + \frac{\partial}{\partial x}(huT) = 0, \tag{1.4.3}$$

où $T(x,t)$ désigne la concentration du traceur dans l'écoulement. Afin de mieux rendre compte de la réalité et de construire un véritable modèle de qualité d'eau, ce modèle peut évidemment être affiné, notamment par l'introduction d'une diffusion sur le traceur ou par la prise en compte de réactions (bio)chimiques ayant lieu dans l'écoulement - voir la thèse de Martin [99]. Plusieurs codes commerciaux proposent ainsi un très large panel de modules, que ce soit par l'intermédiaire d'une approche éléments finis - par exemple le code SUBIEF développé par EDF-LNHE - ou d'une approche volumes finis - voir par exemple la notice d'utilisation du code DELFT-3D-WAQ [41].

1.4.1 Etat de l'art

L'abondance de logiciels ne se retrouve pas dans les travaux académiques. Or, si ces codes couvrent un nombre incalculable de cas particuliers, leur but n'est pas de mener une étude précise et détaillée du système (1.4.1)-(1.4.3), qui est pourtant à la base de tout calcul. Parmi les travaux qui s'intéressent à cette question, citons néanmoins ceux de Dawson et Proft [39, 40] - dans un cadre éléments finis - et ceux de Kurganov et al. [35, 36, 85] ainsi que des travaux déjà menés à l'INRIA [24] - pour une approche volumes finis.

L'extension des schémas cinétiques à un tel problème a déjà été réalisée [24]. Il suffit pour cela de considérer le nouvel équilibre de Gibbs $\tilde{M}$, défini par

$$\tilde{M}(t,x,y) = \begin{pmatrix} 1 \\ T(t,x,y) \end{pmatrix} M(t,x,y),$$

Le système (1.4.1)-(1.4.3) peut alors être vu comme l'intégrale de l'équation cinétique (1.3.2) contre la matrice

$$K(\xi) = \begin{pmatrix} 1 & \xi & 0 \\ 0 & 0 & 1 \end{pmatrix}^T.$$

Comme dans le cas du système de Saint-Venant seul - et en utilisant exactement les mêmes méthodes - considérer le système (1.4.1)-(1.4.3) sous l'angle cinétique permet d'obtenir très naturellement un schéma numérique qui vérifie certaines des propriétés essentielles de l'équation continue, notamment la conservation de la masse totale de traceur ainsi que des bornes supérieures et inférieures sur la concentration du traceur - en sus de celles déjà exposées au préalable pour la partie hydrodynamique. Or la préservation de ces propriétés est, d'un point de vue pratique, peut-être encore plus importante pour ce nouveau problème, notamment parce que le problème du vide apparaît dans toutes les applications - il existe presque toujours des zones où la concentration du polluant est nulle - mais également parce que l'expérience des ingénieurs tend à prouver que les problèmes numériques qui résultent du non-respect de ces nouvelles propriétés liées au traceur sont plus importants.

Ce point de vue cinétique global ne tient pas compte d'une donnée importante du problème : tant que l'on considère un traceur passif, le système (1.4.1)-(1.4.3) est en réalité découplé, la partie transport n'ayant pas d'influence sur la partie hydrodynamique. Il en résulte que, comme toutes les méthodes considérant le système couplé, le schéma cinétique présente un inconvénient majeur : il dissipe fortement les discontinuités présentes dans la concentration du traceur, notamment pour les écoulements à faible nombre de Froude

$$Fr = \frac{|u|}{\sqrt{gh}} \tag{1.4.4}$$

- voir [24]. Ce phénomène est dû au fait que la stabilité du schéma cinétique nécessite l'adaptation du pas de temps, via la condition de CFL, à la plus grande vitesse de propagation de l'information, qui est toujours liée à la partie hydrodynamique. En effet, nous comparons alors des quantités de l'ordre de $|u| + \sqrt{gh}$ - célérité des informations de la partie hydrodynamique - et de l'ordre de $|u|$ - célérité des informations de la partie transport. Le rapport de ces célérités est clairement lié au nombre de Froude, et il apparaît que la condition de CFL de l'hydrodynamique est mal adaptée au problème de transport quand ce nombre est petit : la variable traceur est alors remise à jour trop fréquemment. Or il est bien connu que, dans les méthodes volumes finis, ce sont les étapes de projection liées à la mise à jour des variables qui sont la principale cause de la diffusion numérique.
De plus, cette inadaptation du pas de temps n'a pas pour seule conséquence une perte de précision numérique. Elle est également la cause d'autres difficultés, liées à des questions de temps de calcul et de stockage, tout aussi importantes d'un point de vue "industriel". En effet, si le couplage ne porte pas sur le transport d'une espèce, mais sur l'advection de plusieurs dizaines de traceurs, le surcoût, en temps de calcul, lié aux mises à jour inutiles peut devenir très important. D'autre part et puisque le caractère découplé du problème nous le permet, on peut choisir de stocker les informations hydrodynamiques nécessaires à la simulation de différents problèmes de transport sur le même écoulement. Il faut alors stocker les flux pour chaque interface et à chaque pas de temps, ce qui nécessite des capacités de stockage très importantes. Or ces deux problèmes sont très courants. D'une part, une rivière contient toujours quantités d'espèces, qui souvent réagissent entre elles, et un modèle de qualité d'eau doit donc toutes les intégrer. D'autre part, dans les études d'impact environnemental, il est également habituel de tester divers phénomènes de pollution, correspondant à diverses concentrations initiales de traceur, sur un même écoulement hydrodynamique stationnaire.

Il s'avère donc indispensable d'utiliser le caractère découplé du système. Une première possibilité est de résoudre le système de Saint-Venant seul puis d'utiliser le champ de vitesse obtenu pour résoudre l'équation de transport dans sa forme non-conservative équivalente

$$\frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} = 0. \tag{1.4.5}$$

Pour transposer cette idée dans une résolution numérique, il suffit de considérer un solveur classique du système de Saint-Venant - voir la section précédente - et de lui ajouter un solveur adapté à l'équation de transport non conservative (1.4.5), les résultats du premier servant de données d'entrée au second. A titre d'exemple, Kurganov et al. [35, 36] couplent un solveur *central scheme* avec une méthode particulaire et améliorent nettement les résultats par rapport à ceux obtenus en appliquant un solveur *central scheme* au système couplé [85]. Cette méthode présente plusieurs avantages : chaque équation est résolue avec un solveur spécialement adapté et le solveur Saint-Venant

peut-être utilisé comme une boîte noire puisqu'aucune modification n'est nécessaire. Néanmoins, le passage de la forme (1.4.3) à la forme (1.4.5) de l'équation de transport entraîne une perte d'information importante puisque la propriété de conservation de la masse totale du traceur n'est plus explicite et n'a donc aucune raison d'être vérifiée dans la résolution numérique. Peut également se poser le problème de la complexité des algorithmes : si la méthode particulaire est très précise, son coût reste élevé.

1.4.2 Le schéma à deux pas de temps

Fort de ces expériences nous avons donc cherché à coupler l'approche cinétique globale et l'approche découplée. Pour cela, nous conservons l'écriture (1.4.3) de l'équation de transport ainsi que le formalisme cinétique et le découplage est introduit dans la résolution numérique, non pas en considérant deux méthodes de résolution différentes, mais en introduisant deux pas de temps différents : il s'agit de ne remettre à jour la variable traceur que lorsque cela est nécessaire, afin de minimiser la diffusion numérique, tout en préservant les acquis de stabilité des schémas cinétiques.
Puisque nous avons choisi une écriture conservative des équations - et que nous appliquons une méthode numérique conservative - le problème de la conservation de la masse du traceur ne se pose plus. Le point important de cette méthode est donc de trouver une manière de définir ces nouveaux pas de temps, adaptée au phénomène de transport du traceur, mais qui préserve un principe du maximum - et un principe du minimum - sur la concentration du traceur - puisque c'est cette propriété qui apparaît maintenant comme la plus fondamentale. Notons aussi que, pour des problèmes de compatibilité, il est indispensable que les instants de mise à jour du traceur correspondent à des instants où les variables hydrodynamiques sont elles aussi mises à jour.

Nous avons pour cela mis au point un algorithme basé sur une double condition de CFL. La condition de CFL classique est préservée et la résolution de la partie hydrodynamique reste inchangée, ce qui permet de préserver les propriétés de stabilité démontrée dans la section précédente. Mais, à chaque étape de mise à jour des variables hydrodynamiques, la mise à jour du traceur n'est plus automatique : un test, relié à une condition de CFL différente, est effectué, afin de savoir si il convient, ou non, de mettre à jour la variable traceur. Si le résultat du test est négatif pour toutes les cellules du maillage, les flux reliés à la variable traceur - c'est à dire les flux de masse, puisque nous résolvons la forme conservative (1.4.3) de l'équation du traceur - sont simplement incrémentés, dans l'attente du test suivant. Ainsi le nouveau schéma volumes finis pour le calcul du traceur aura la forme suivante

$$(hT)_i^{N+1} - (hT)_i^N + \frac{(\Delta t_T)^N}{\Delta x_i}\left[(T_i^N(F_{T,i+\frac{1}{2}}^N)_+ - T_{i+1}^N(F_{T,i+\frac{1}{2}}^N)_-) - (T_{i-1}^N(F_{T,i-\frac{1}{2}}^N)_+ - T_i^N(F_{T,i-\frac{1}{2}}^N)_-)\right] = 0, \tag{1.4.6}$$

où l'indice N est différent de l'indice n présent dans le schéma volumes finis résolvant la partie hydrodynamique (1.2.13), puisque le pas de temps $(\Delta t_T)^N$ est une somme de pas de temps hydrodynamiques Δt^k. Les flux associés au traceur $F^N_{T,i+\frac{1}{2}}$ sont quant à eux le résultat de l'agrégation de plusieurs flux de masse.
Dans la version que nous présentons, la nouvelle condition de CFL traduit simplement le fait que, lors d'une itération sur le traceur, on ne peut pas enlever d'une cellule plus de traceur qu'il n'y en a au début de l'itération - seul les flux sortants sont ici pris en compte, on ne regarde pas les éventuels apports de traceur provenant des autres cellules. Après mise en facteur de la concentration du traceur sur la cellule, cela revient en fait à s'assurer qu'on ne peut pas enlever plus d'eau qu'il n'y en a initialement dans la cellule, d'où la présence des flux de masse dans les formules précédentes.

La stabilité de la méthode - donc la préservation des propriétés du système - est assurée par la manière dont nous avons défini la nouvelle condition de CFL. Elle assure en effet qu'en fin d'itération la quantité d'eau reste positive dans toutes les cellules, ce qui est suffisant pour assurer des principes du maximum et du minimum discrets. Nous pouvons alors établir le théorème suivant

Theorème 1.4.1 (Chapitre 4, Pages 112, 113 et 117)
Le schéma cinétique à deux pas de temps est conservatif et préserve l'invariance du domaine d'évolution de la concentration du traceur, ce qui se traduit par un principe du maximum (et du minimum) discret local

$$\forall n \quad \forall i \quad min(T^n_{i-1}, T^n_i, T^n_{i+1}) \leq T^{n+1}_i \leq max(T^n_{i-1}, T^n_i, T^n_{i+1}). \tag{1.4.7}$$

Abordons maintenant le problème de la diffusion numérique de la concentration du traceur. La nouvelle condition de CFL compare le flux de masse sortant d'une cellule à la masse de cette cellule. Elle est donc reliée à la vitesse de l'écoulement, puisqu'elle est le résultat du ratio d'un débit par une hauteur. Ainsi cette condition de CFL se trouve décorréllée des célérités liées à la partie hydrodynamique et est compatible avec la célérité liée au transport. Dans le cas d'écoulements à faible nombre de Froude, donc lorsque ces célérités sont très différentes, la longueur des pas de temps $(\Delta t_T)^N$ liés au transport s'en trouve fortement augmentée. Nous exhibons ainsi, sur un exemple numérique effectué sur une géométrie complexe en deux dimensions, et pour un nombre de Froude proche de 0,1 - caractéristique d'un écoulement classique en rivière - que le pas de temps lié au traceur peut être cent fois plus grand que le pas de temps hydrodynamique. Cette réduction du nombre de mises à jour s'accompagne fort logiquement, dans tous les tests effectués, d'une diminution notable de la diffusion numérique, le phénomène étant particulièrement marqué pour des maillages réguliers. Une illustration liée à cet exemple est présentée sur la Figure 1.4.1, accompagnée d'un tableau indiquant le nombre de pas de temps effectués.

Venons en enfin au problème du nombre de calculs et des capacités de stockage. L'allongement du pas de temps se traduit automatiquement par une diminution du nombre de

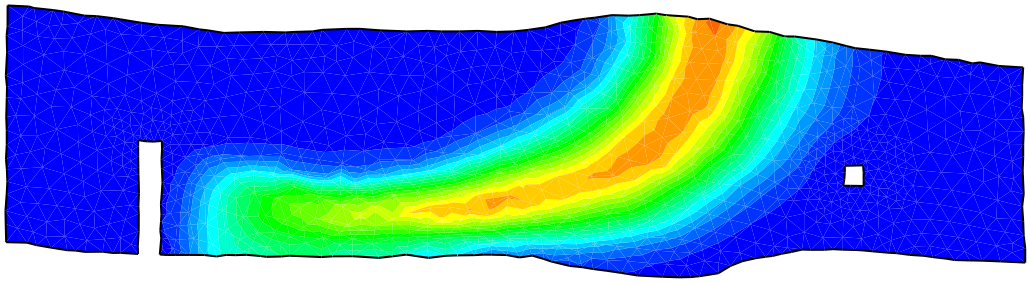


FIG. 1.4.1: Concentration du traceur - Schéma cinétique à deux pas de temps

Froude number	Transport Steps	Hydrodynamic Steps
0.08	320	45637

Table 1 : Nombre de pas de temps hydrodynamiques et nombre de pas de temps transports pour la simulation, par le schéma à deux pas de temps, de l'écoulement en rivière présenté sur la Figure 1.4.1.

calculs et du nombre de flux à stocker. Encore faut-il pouvoir effectuer ces stockages... Autrement dit, encore faut-il pouvoir effectuer le calcul hydrodynamique et stocker les flux nécessaires en amont, indépendamment de toute connaissance sur la distribution de la concentration des traceurs. C'est ici qu'intervient un autre des points essentiels de la méthode proposée, à savoir que la définition de la nouvelle condition de CFL, même si elle est adaptée au phénomène de transport, ne prend en compte que des données hydrodynamiques et est indépendante de toutes données sur le traceur. Ceci est du au choix de ne pas tenir compte, dans la définition de la nouvelle condition de CFL, des flux de traceur entrant, qui, si ils étaient pris en compte, transporteraient des concentrations de traceur différentes de celle de la cellule considérée. La mise en facteur évoquée précédemment ne pourrait être effectuée et le test ne pourrait donc pas porter que sur des quantités purement hydrodynamiques.
Ici, on peut donc stocker, lors d'un calcul hydrodynamique, les flux $F^N_{T,i+\frac{1}{2}}$ nécessaires aux futurs calculs des phénomènes de transport associés - soit, dans l'exemple évoqué plus haut, une seule quantité par interface pour cent pas de temps hydrodynamiques. On peut ensuite appliquer le schéma (1.4.6) avec diverses concentrations initiales - ou diverses espèces - de traceurs. Le pas de temps de ce second calcul numérique est alors bien évidemment le pas de temps $(\Delta t_T)^N$ associé au phénomène de transport, le pas de temps Δt^n associé à l'hydrodynamique ayant totalement disparu.

1.5 Lois de conservation scalaires avec flux discontinu : Un théorème d'unicité

1.5.1 Position du problème

Les problèmes hyperboliques de lois de conservation auxquels nous nous intéressons dans cette thèse présentent une grande compléxité mathématique, liée notamment au caractère singulier des solutions. Pour les systèmes, le problème de l'existence et de l'unicité des solutions reste encore très largement ouvert. Les seuls résultats connus concernent les systèmes 2×2. Nous avons, par exemple, déjà signalé qu'il existait, pour le système de Saint-Venant homogène, une démonstration de l'existence d'une solution après apparition de singularités. Néanmoins ce résultat [95] reste très récent et est limitée à un problème posé en une dimension d'espace. De plus, aucune démonstration d'unicité n'existe à ce jour. Le cas des lois de conservation scalaire est tout de même un peu mieux connu. Ainsi, Kruzkov [83] a démontré, en 1970, l'existence et l'unicité d'une solution entropique à un problème du type

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} F(x,u) = 0. \tag{1.5.1}$$

pour lequel la dépendance spatiale du flux est suffisamment régulière. La preuve repose principalement sur une l'utilisation d'une famille d'entropies particulières, les entropies de Kruzkov. Les problèmes qui se posent ici sont donc ceux de l'existence et de l'unicité des solutions d'une telle équation quand la dépendance spatiale du flux présente des singularités.

Cette question trouve des motivations dans les applications liées à la modélisation des écoulements en eaux peu profondes par le système de Saint-Venant. Evoquons ici deux d'entre elles. La première est directement reliée au chapitre précédent puisqu'elle concerne le couplage du système de Saint-Venant avec une équation de transport. Considérons ainsi un écoulement stationnaire unidimensionnel pour lequel la vitesse présente des discontinuités - voir par exemple les exemples numériques présentés dans le Chapitre 3. Il apparaît clairement que léquation de transport associée (1.4.3) peut alors s'écrire sous la forme (1.5.1) avec $F(x,u) = v(x)u$ où $\tilde{v}$ désigne la vitesse de l'écoulement et u la quantité de traceur. La deuxième application concerne l'utilisation du système de Saint-Venant pour la modélisation des écoulements sanguins - voir [47, 48]. Le module d'Young des parois veineuses apparait alors explicitement dans les termes de pression de l'équation sur la quantité de mouvement. Modéliser une veine dans laquelle on a posé un cateter conduit donc à considèrer des flux à coefficients discontinus, même si il ne s'agit pas ici d'une loi de conservation scalaire.

1.5.2 Etat de l'art

Dans la dernière décennie, plusieurs travaux ont été consacrés au problème des lois de conservation scalaires à coefficients discontinus. Puisque nous présentons ici un théorème d'unicité, nous insisterons plus particulièrement sur ce point, mais, parmi les travaux que nous citerons, nombreux sont ceux qui abordent également les problèmes d'existence. Notons aussi que certains n'abordent pas seulement le cas hyperbolicue, mais aussi les problèmes paraboliques dégénérés.
Commencons par le cas d'une équation de transport, puisque nous l'avons évoqué plus haut. Le flux est alors de la forme $F(x,u) = a(x)u$. Le résultat d'unicité le plus général a été établi par Bouchut et James, dans [20], où il démontre l'unicité de la solution de l'équation (1.5.1) associée à la condition initiale $u(0,x) = 0$ sous la seule condition que $a \in L^\infty(\mathbb{R})$ et que u vérifie la majoration suivante $u \leq \tilde{v}$ où v est une solution positive de (1.5.1).
Dans le cas non linéaire, les travaux sont plus nombreux. Les hypothèses sur le flux et sur les solutions, en vue de démontrer l'unicité des solutions, le sont également. La plupart des travaux supposent le flux convexe par rapport à la variable u [81, 125, 80, 82, 117]. Récemment Karlsen, Risebro et Towers [75] et Adimurthi, Jaffre et Veraapa [2] sont parvenus à lever cette hypothèse. Néanmoins les seconds considèrent que le flux n'a qu'un point de minimum et les premiers, quoiqu'autorisant une forme plus générale, doivent également introduire une hypothèse restrictive, dite *crossing condition*. D'autre part, dans tous ces articles, la dépendance spatiale du flux est soumise à des bornes BV, le nombre de points de discontinuités est supposé fini et le flux est supposé régulier - souvent C^1 en dehors de ces points. De plus une hypothèse est également nécessaire sur les solutions : l'existence de traces au niveau des discontinuités du flux. Dans les deux travaux les plus récents et les plus complets [75, 2], la dpreuve d'unicité est basée sur l'utilisation des entropies de Kruzkov loin des discontinuités et sur l'utilisation d'une condition d'interface, définie à partir des traces de la solution, aux points de discontinuité. Notons que les conditions d'interfaces introduites dans [75] et [2] sont différentes et que pour certains flux particuliers, elles selectionnent des solutions différentes.

1.5.3 Les entropies de Kruzkov partiellement adaptées

Nous présentons ici une nouvelle preuve d'unicité pour les solutions de (1.5.1), qui peut s'appliquer à la fois pour des problèmes linéaires et non linéaires. Cette méthode suppose elle aussi des hypothèses restrictives sur le flux, concernant sa dépendance par rapport à la variable u. Ainsi nous ne supposons pas le flux convexe, mais localement Lipschitz et nous n'autorisons au plus qu'un point de minimum. De plus la valeur atteinte en ce point minimum doit être la même, quelquesoit le point d'espace considéré. Par contre, les hypothèses sur la discontinuité spatiale du flux sont très générales : nous supposons simplement que le flux est L^∞ et continu en dehors d'un ensemble négligeable.

Dans [125, 75, 117, 2], nous avons dit que la démonstration d'unicité repose sur l'utilisation des entropies de Kruzkov classiques couplées avec l'introduction d'une condition d'interface, afin d'adapter la preuve au cas d'un flux discontinu. Ici, nous choisissons d'adapater la définition des entropies de Kruzkov au cas discontinu, ce qui nous permet ensuite d'utiliser une preuve très proche de celle de Kruzkov. Notamment, aucune condition d'interface n'est plus nécessaire - et la notion de traces sur les solutions n'est donc plus requise.
Le point clé de notre méthode est donc l'introduction des entropies de Kruzkov partiellement adaptées. Les entropies de Kruzkov classiques sont définies par

$$E_k(u) = |u - k|,$$

où $k \in \mathbb{R}$. Ici nous introduisons des entropies de Kruzkov partiellement adaptées, de la forme

$$E_\alpha(x, u) = |u - k_\alpha(x)|, \tag{1.5.2}$$

où $k_\alpha(x)$ est défini par la relation

$$F(x, k_\alpha(x)) = \alpha \quad \text{for a.e. } x \in \mathbb{R}. \tag{1.5.3}$$

La définition des $k_\alpha(x)$, via la relation (1.5.3), est possible grâce aux hypothèses faites sur le flux.
Comme dans le cas classique, on peut montrer que si une solution vérifie une inégalité d'entropie pour toute cette famille d'entropie, alors c'est une solution entropique. Mais le grand intérêt de cette nouvelle définition est qu'elle étend au cas discontinu, la relation vérifiée par les entropies de Kruzkov classiques dans le cas où le flux n'a pas de dépendance spatiale. Ainsi la relation

$$\frac{\partial}{\partial x} F(k) = 0$$

devient maintenant

$$\frac{\partial}{\partial x} F(x, k_\alpha(x)) = 0, \tag{1.5.4}$$

alors que, dans le cas discontinu, on a clairement

$$\frac{\partial}{\partial x} F(x, k) \neq 0.$$

Ainsi, grâce à la propriété (1.5.4), les inégalités d'entropie utilisant les entropies de Kruzkov partiellement adatées (1.5.2) ont la forme habituelle

$$\frac{\partial}{\partial t}|u - k_\alpha(x)| + \frac{\partial}{\partial x}\left[(F(x, u) - F(x, k_\alpha(x)))\ \mathrm{sgn}(u - k_\alpha(x))\right] \leq 0. \tag{1.5.5}$$

1.5.4 Le théorème d'unicité

Les hypothèses annoncées sur les flux prennent la forme suivante

- (H1) $A(x,u)$ est continu en tous points de $\mathbb{R}\backslash\mathcal{N}\times\mathbb{R}$ où $\mathcal{N}$ est un ensemble de mesure nulle,
- (H2) $\exists(f,g)\in\left(C^0(\mathbb{R})\right)^2$ telles que $\forall x\in\mathbb{R}\quad f(u)\leq|A(x,u)|\leq g(u)$. On suppose aussi que $|f(\pm\infty)|=+\infty$.
- (H3) Pour $x\in\mathbb{R}\backslash\mathcal{N}$, $A(x,.)$ est une fonction localement lipschitzienne et bijective de $\mathbb{R}$ dans $\mathbb{R}$.
- (H3') Pour $x\in\mathbb{R}\backslash\mathcal{N}$, $A(x,.)$ est une fonction localement lipschitzienne et bijective de $-\infty,u_M(x)]$ et $[u_M(x),+\infty]$ dans $\mathbb{R}^+$.

On peut alors établir un principe de contraction L^1

Theorème 1.5.1 (Chapitre 5, Page 145)
Soient u et $v\in L^\infty([0,T],\mathbb{R})\cap C^0([0,T],L^1_{loc}(\mathbb{R}))$ une sous- et une sur-solution entropiques du problème de Cauchy associé à (1.5.1) avec données initiales $u_0,v_0\in L^\infty(\mathbb{R})$. Sous les hypothèses (H1)-(H2)-(H3) - ou (H1)-(H2)-(H3') - sur le flux, on a pour presque tout $t\in[0,T]$

$$\int_a^b\left(u(x,t)-v(x,t)\right)_+dx\leq\int_{a-Mt}^{b+Mt}\left(u_0(x)-v_0(x)\right)_+dx \tag{1.5.6}$$

La preuve repose sur la technique habituelle de dédoublement des variables. Nous considérons ainsi deux solutions u et v de (1.5.1) et nous définissons deux fonctions $\tilde{u}$ et $\tilde{v}$ telles que

$$\begin{array}{l}F(y,\tilde{u}(t,x,y))=F(x,u(t,x))\\F(x,\tilde{v}(s,y,x))=F(y,v(s,y))\end{array}\quad\text{for a.e. } t,s,x,y$$

et nous les introduisons respectivement dans les inégalités d'entropie (1.5.5) sur v et sur u. L'addition de ces inégalités et leur intégration contre une fonction test $\psi(t,x,s,y)\in C_0^\infty$ de la forme $\psi_{\eta\epsilon}(t,x,s,y)=\phi(t,x)\xi_\eta(x-y)\rho_\epsilon(t-s)$, où ξ_η et ρ_ϵ sont des approximations de la masse de Dirac, conduit à l'inégalité suivante

$$\text{(I)} \qquad \int_{Q^2} (u(t,x) - \tilde{v}(s,y,x))_+ \partial_t \phi(x,t) \rho_\epsilon(t-s) \xi_\eta(x-y) dy ds dx dt$$

$$\text{(II)} \qquad -\int_{Q^2} ((u(t,x) - \tilde{v}(s,y,x))_+ - (v(s,y) - \tilde{u}(t,x,y))_-) \phi(x,t) \rho_\epsilon{}'(t-s) \xi_\eta(x-y) dy ds dx dt$$

$$\text{(III)} \qquad +\int_{Q^2} (A(x,u(t,x)) - A(x,\tilde{v}(s,y,x))) \partial_x \phi(x,t) \rho_\epsilon(t-s) \xi_\eta(x-y) (\mathrm{sgn}_+(u(t,x) - \tilde{v}(s,y,x)) dy ds dx dt$$

$$\text{(IV)} \qquad -\int_{Q^2} (A(x,u(t,x)) - A(y,v(s,y))) \phi(x,t) \rho_\epsilon(t-s) \xi_\eta{}'(x-y) (\mathrm{sgn}_+(u(t,x) - \tilde{v}(s,y,x)) + \mathrm{sgn}_-(v(s,y) - \tilde{u}(t,x,y))) dy ds dx dt$$

$$\text{(V)} \qquad +\int_{Q \times \mathbb{R}} (u_0(x) - \tilde{v}(s,y,x))_+ \phi(x,0) \rho_\epsilon(-s) \xi_\eta(x-y) dy ds dx$$

$$\text{(VI)} \qquad +\int_{Q \times \mathbb{R}} (v_0(y) - \tilde{u}(t,x,y))_- \phi(x,t) \rho_\epsilon(t) \xi_\eta(x-y) dy dx dt \geq 0.$$

Les termes (I), (III) et (V) sont classiques et conduisent, à la limite en η et ϵ, au principe de contraction L^1. Le terme (VI) est nul. Les termes (II) et (IV) sont reliés au caractère discontinu du flux. Nous montrons que le terme (IV) est toujours nul, quelquesoient η et ϵ - les deux fonctions signes sont opposées. Nous montrons ensuite, qu'à ϵ fixé, la limite en η du terme (II) est nulle.

Nous terminons ce chapitre en appliquant notre méthode à un cas particulier, afin de souligner un point intéressant de la méthode présentée. Nous avons dit que, dans les travaux existants [75, 2], et pour certains flux, différentes conditions d'interfaces pouvaient conduire à sélectionner différentes solutions. Or, si notre méthode nécessite des hypothèses plus restrictives sur la dépendance en u du flux, elle permet par contre d'éviter l'introduction de telles conditions d'interfaces. Au travers de comparaisons des solutions sélectionnées par les différentes méthodes - avec ou sans conditions d'interfaces - dans des cas où elles sont toutes applicables, elle peut donc fournir un moyen pour effectuer un choix parmi les différentes conditions d'interfaces. Ainsi nous montrons que, dans l'exemple présenté, la solution sélectionnée par notre méthode est identique à celle de [2], mais diffère de celle de [75].

1.6 Un modèle Saint-Venant multicouche : Dérivation et Analyse du modèle, Mise en oeuvre numérique

1.6.1 Un modèle intermédiaire entre Saint-Venant et Navier-Stokes

Le système de Saint-Venant (1.2.1)-(1.2.2) a initialement été introduit de manière indépendante, au travers d'une analyse des forces exercées sur une section du fluide [115]. Ses relations avec les autres modèles fluides, et plus particulièrement avec les équations de Navier-Stokes, ont néanmoins été étudiées depuis longtemps - voir par exemple l'appendice de Friedrichs à un article de Stoker [120] ou les ouvrages de Stoker [121] ou Whitham [130]. Grossièrement, le système de Saint-Venant peut en effet être vu comme le résultat d'une moyenne selon la verticale des équations de Navier-Stokes. Notons à ce propos que ces deux systèmes sont en fait étroitement liés depuis l'origine puisque, selon certaines sources [4], Saint-Venant aurait aussi été le premier, deux ans avant Stokes [122], a donné la forme correcte des équations que Navier [102] avait introduites en 1823. Reste maintenant à établir quand le processus d'homogénéisation verticale est valable, ou, autrement dit, quand, pour des conditions initiales compatibles, les solutions des systèmes de Saint-Venant et de Navier-Stokes restent compatibles.
Cette question a été reprise récemment dans deux articles [50, 45]. Dans [50], Gerbeau et Perthame établissent grâce à une étude asymptotique formelle que le système de Saint-Venant est une approximation au premier ordre des équations de Navier-Stokes sous l'hypothèse d'eau peu profonde. Outre la petitesse du rapport d'une hauteur à une longueur caractéristiques de l'écoulement, cette dérivation suppose que la viscosité et le coefficient de frottement sont du même ordre de grandeur que ce rapport et que le gradient de surface libre reste borné. Dans [45], Saleri et Ferrari étendent cette étude au cas bidimensionnel, en présence d'une bathymétrie non triviale et d'un gradient de pression atmosphérique. Il faut alors supposer que les gradients de fond et de pression atmosphérique sont eux aussi petits.
Outre le fait qu'ils demeurent formels, l'impact direct de tels travaux dans l'étude du domaine de validité du système de Saint-Venant reste limité : ils fournissent quelques indications, établissent des tendances asymptotiques, mais les ingénieurs ont établi depuis longtemps que le système de Saint-Venant fournit des solutions raisonnables dans des situations qui violent une ou plusieurs des hypothèses énoncées ci-dessus - citons par exemple le cas des ruptures de barrage, pour lesquelles le gradient de surface libre initial est infini. Néanmoins, l'intérêt principal de ces travaux est ailleurs : en effet de telles études, précises et détaillées, fournissent des outils pour construire de nouveaux modèles fluides. A titre d'exemple, dans [50], les auteurs, en présentant une approximation au deuxième ordre, dérivent un nouveau système de type Saint-Venant, dit *système de Saint-Venant visqueux*, dont la nature laisse à penser qu'il possède un

domaine de validité plus grand que celui du système classique - pour une compléxité identique. Plusieurs résultats numériques viennent corroborer cette assertion.

Notre objectif, dans le travail présenté ici, est également de dériver un nouveau modèle fluide, "intermédiaire" entre Saint-Venant et Navier-Stokes. Notre but n'est cependant plus seulement d'étendre le domaine de validité du système de Saint-Venant, mais de lever une des limitations a priori intrinsèques à ce type de modèle, à savoir le fait que toute la section verticale d'eau est supposée avoir la même vitesse. En effet, puisque le système de Saint-Venant est issu des équations de Navier-Stokes par l'intermédiaire d'une moyenne selon la verticale, tous les profils verticaux, en particulier celui de la vitesse horizontale, sont forcément constants. Hors cette limitation est rédhibitoire pour bon nombre d'applications : recirculation dans un lac sous l'effet du vent - l'eau est entrainée en surface et recircule en profondeur - ou bien calcul du frottement exercé par le fluide - dont la valeur est évidemment liée à la vitesse horizontale du fluide à l'interface fluide / solide et non à sa vitesse moyenne - sur le fond d'une rivière - phénomènes d'érosion - ou sur les parois veineuses - détérioration des tissus. Nous voulons donc établir un modèle qui préserve l'efficacité du système de Saint-Venant - réduction de la dimension du problème, domaine de calcul fixe - mais permet d'obtenir un profil vertical de vitesse horizontale non constant. Dans la suite, nous parlerons de modèle de Saint-Venant multicouche.

1.6.2 Etat de l'art

Quelques travaux existants partagent une partie de ces préoccupations.
Citons notamment le travail de Lazzaroni [86], sous la direction de Saleri, qui présente l'étude numérique, par une méthode d'éléments finis, d'un modèle multicouche particulier. Outre les différences entre les deux modèles, les approches numériques sont donc également très dissemblables. De plus l'accent est mis sur les techniques numériques et ni la dérivation du modèle, ni son étude, ne sont présentées.
Citons également les travaux de Castro, Pares et al. [27, 28, 29], concernant un modèle de Saint-Venant bi-fluide - et non bi-couche car il s'agit ici de deux fluides de densités différentes - et sa discretization par une méthode de volumes finis - la principale motivation est ici l'application du modèle au détroit de Gibraltar [28]. Certains résultats établis dans cette dernière série de travaux - notamment à propos de l'hyperbolicité des modèles - seront repris, commentés et adaptés à notre problème.
Notons aussi que de nombreux modèles Navier-Stokes simplifiés existent, mais qu'il s'agit toujours de problèmes 3d, souvent posés sur des domaines mobiles, donc hors du champ de notre étude.

1.6.3 Dérivation et analyse du système multicouche

Nous reprenons ici le formalisme développé dans [50] et, pour plus de simplicité, nous ne considérons qu'une dimension horizontale et nous nous placons dans le cas d'un fond plat. Partant des équations de Navier-Stokes incompressibles à surface libre, nous introduisons la mise à l'échelle associée à l'hypothèse d'eau peu profonde, puis, en éliminant les termes d'ordres élevés dans le système adimensionnel obtenu, nous dérivons des modèles de Navier-Stokes simplifiés. A partir de ces modèles simplifiés, grâce à un processus de moyenne sur la verticale, nous obtenons des modèles du type Saint-Venant. Néanmoins, grâce à l'introduction dans le modèle d'une discretization verticale, nous pouvons ici définir des couches de fluide possédant chacune leur propre vitesse horizontale. Nous obtenons ainsi un système de type Saint-Venant, mais à plusieurs couches, dans lequel l'influence de la viscosité apparaît explicitement, et nous établissons qu'il s'agit d'une approximation au premier ordre des équations de Navier-Stokes. Pour une couche générique - l'écriture est légèrement modifiée pour les couches extrêmes, inférieure et supérieure - le système revêt la forme suivante

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) = 0, \tag{1.6.1}$$

$$\frac{\partial hu}{\partial t} + \frac{\partial}{\partial x}\left(hu^2\right) + gh\frac{\partial H}{\partial x} = 2\mu\left(\frac{u_+ - u}{h_+ + h} - \frac{u - u_-}{h + h_-}\right), \tag{1.6.2}$$

où h et u sont respectivement la hauteur et la vitesse de la couche considérée, H est la hauteur totale de l'écoulement et où les indices $+$ et $-$ se réfèrent respectivement aux couches immédiatement supérieure et inférieure à la couche considéré. g est la gravité et μ désigne la viscosité de lécoulement.

Nous démontrons que ce système de Saint-Venant multicouche, comme le système de Saint-Venant et les équations de Navier-Stokes, possède une énergie - propriété de stabilité du modèle. Nous prouvons aussi que pour certaines données initiales et dans le cas d'un frottement nul, le système de Saint-Venant homogène (1.2.1)-(1.2.2) fournit une solution à ce nouveau système multicouche - propriété de compatibilité du modèle. Néanmoins nous établissons également que ce système n'est ni conservatif, ni hyperbolique, la perte d'hyperbolicité étant liée à la définition arbitraire - et donc artificielle - des interfaces. Il est établi dans [27] que cette perte d'hyperbolicité est la cause d'importantes instabilités numériques.

Nous proposons donc plusieurs autres écritures possibles du système, toutes possédant un membre de gauche conservatif et hyperbolique - le traitement numérique différencié de cette partie hyperbolique et des termes sources non conservatifs ainsi isolés assurera alors la stabilité des solutions. Grâce à l'étude des valeurs propres de la partie hyperbolique, nous exhibons une écriture particulière, qui possède, parmi ses valeurs propres, deux valeurs propres, dites barotropiques et associées au déplacement de la

surface libre, qui sont compatibles avec les valeurs propres du système de Saint-Venant homogène, ce qui assure la cohérence du modèle quant au déplacement de la surface libre. De plus nous prouvons que, pour cette écriture et dans les applications qui nous intéressent, les termes non conservatifs restent petits et peuvent donc être considérés comme des termes sources correctifs, dont l'influence reste limitée, ce qui nous permettra de privilégier la stabilité de la discretization sur sa précision. Ce système s'écrit de la manière suivante - nous présentons, dans l'ordre, l'écriture pour la couche inférieure, pour une couche générique, puis pour la couche supérieure.

$$\frac{\partial h_1}{\partial t} + \frac{\partial h_1 U_1}{\partial x} = 0, \tag{1.6.3}$$

$$\frac{\partial h_1 U_1}{\partial t} + \frac{\partial}{\partial x}\left(h_1 U_1^2 + g\frac{h_1\left(\sum_{\beta=1}^{M} h_\beta\right)}{2}\right) = \frac{g\left(\sum_{\beta=1}^{M} h_\beta\right)^2}{2}\frac{\partial}{\partial x}\left(\frac{h_1}{\sum_{\beta=1}^{M} h_\beta}\right) + 2\mu\frac{U_2 - U_1}{h_2 + h_1} - \kappa U_1, \tag{1.6.4}$$

$$\frac{\partial h_\alpha}{\partial t} + \frac{\partial h_\alpha U_\alpha}{\partial x} = 0, \tag{1.6.5}$$

$$\frac{\partial h_\alpha U_\alpha}{\partial t} + \frac{\partial}{\partial x}\left(h_\alpha U_\alpha^2 + g\frac{h_\alpha\left(\sum_{\beta=1}^{M} h_\beta\right)}{2}\right) = \frac{g\left(\sum_{\beta=1}^{M} h_\beta\right)^2}{2}\frac{\partial}{\partial x}\left(\frac{h_\alpha}{\sum_{\beta=1}^{M} h_\beta}\right) + 2\mu\left(\frac{U_{\alpha+1} - U_\alpha}{h_{\alpha+1} + h_\alpha} - \frac{U_\alpha - U_{\alpha-1}}{h_\alpha + h_{\alpha-1}}\right) \tag{1.6.6}$$

$$for \quad \alpha = 2...M-1,$$

$$\frac{\partial h_M}{\partial t} + \frac{\partial h_M U_M}{\partial x} = 0, \tag{1.6.7}$$

$$\frac{\partial h_M U_M}{\partial t} + \frac{\partial}{\partial x}\left(h_M U_M^2 + g\frac{h_M\left(\sum_{\beta=1}^{M} h_\beta\right)}{2}\right) = \frac{g\left(\sum_{\beta=1}^{M} h_\beta\right)^2}{2}\frac{\partial}{\partial x}\left(\frac{h_M}{\sum_{\beta=1}^{M} h_\beta}\right) - 2\mu\frac{U_M - U_{M-1}}{h_M + h_{M-1}}. \tag{1.6.8}$$

où M désigne le nombre total de couches, les indices $1, \alpha, \beta, M$ désignant différentes couches. Les variables h et u désignent encore respectivement la hauteur et la vitesse

d'une couche. Le terme $\sum_{\beta=1}^{M} h_\beta$ dénote la hauteur totale de l'écoulement, μ est toujours la viscosité de l'écoulement et κ est un coefficient de friction.

Theorème 1.6.1 (Chapitre 6, Pages 163, 165, 166 et 168)
Le système de Saint-Venant multicouche avec viscosité et friction (1.6.3)-(1.6.8) est le résultat d'une approximation asymptotique formelle en $O(\epsilon)$ - où ϵ désigne le ratio entre hauteur et longueur caractéristiques - couplée avec une discrétistation verticale du modèle hydrostatique et donc des équations de Navier-Stokes.
De plus ce système de Saint-Venant multicouche vérifie les propriétés suivantes
(P1) Le système admet une énergie.
(P2) Le système obtenu en remplacant le membre de droite par 0 est hyperbolique.
(P3) Les valeurs propres barotropiques du système sont compatibles avec les valeurs propres du système de Saint-Venant classique.
(P4) Pour des données initiales correctes et sous certaines conditions sur les coefficients de frottement et de viscosité, le membre de droite reste "petit".
(P5) La somme sur toutes les couches des membres de gauche et de droite permet de retrouver une approximation au premier ordre des membres de gauche et de droite du système de Saint-Venant classique, respectivement.
(P6) Pour certaines données initiales et certaines valeurs du coefficient de frottement, le système de Saint-Venant classique fournit une solution au modèle hydrostatique et au système de Saint-Venant multicouche.
(P7) Le système de Saint-Venant multicouche préserve certaines discretizations liées à l'état stationnaire correspondant à un lac au repos.

1.6.4 Etude numérique du système multicouche

Le seul travail autour du traitement numérique d'un système multicouche, et dans le cadre des volumes finis, est, à notre connaissance, dû à Castro et Pares et al. [27, 29]. Or le cadre diffère du nôtre puisqu'il s'agit d'un problème bi-fluide et, si l'écriture des systèmes est très semblable, la différence de densité entre les deux couches permet de retrouver l'hyperbolicité du système. De plus, puisqu'il s'agit de deux fluides différents la position de l'interface est importante et les auteurs sont donc conduits à considérer une écriture du système proche de (1.6.1)-(1.6.2), non conservative mais homogène, puis à développer un schéma numérique construit sur le système couplé, basé sur une généralisation du schéma de Roe à des systèmes non-conservatifs.
Cette approche est, dans le cadre du modèle multicouche que nous considérons, et du fait de sa non-hyperbolicité, impossible. Elle conduirait en effet, de par sa nature même, au développement d'instabilités au niveau de l'interface [27]. Une autre donnée s'oppose aussi à l'emploi d'une telle méthode : les schémas construits sur le système couplé sont adaptés à un nombre de couches donné. Or nous souhaitons ici pouvoir travailler avec un nombre de couches arbitraire. Cette remarque nous incite donc à nous tourner vers l'emploi de schémas découplés, autrement dit, à considérer le modèle

multicouche comme autant de systèmes de Saint-Venant avec termes sources. La forme du système que nous avons retenue (1.6.3)-(1.6.8) nous conforte en partie dans cette optique puisque, pour chaque couche, l'analogie avec le système de Saint-Venant est forte : similitude des termes de flux et petitesse des termes sources. Les termes de flux sont néanmoins légèrement différents de ceux du système de Saint-Venant classique (1.2.1)-(1.2.2). En effet, afin de préserver la propriété (P3) - et celles qui en découlent - il s'est avéré nécessaire d'introduire dans le système un certain couplage à l'intérieur même des termes de flux, qui se traduit par la présence dans ces termes de la hauteur d'eau totale de l'écoulement. Dès lors, il devient nécessaire d'adapter les travaux effectués autour du système de Saint-Venant classique, sans toutefois remettre en cause le principe d'une résolution couche par couche.

L'interprétation cinétique offre une nouvelle fois une solution simple et élégante. En effet, l'équation cinétique qui sous-tend ces systèmes de Saint-Venant modifiés est la même que l'équation cinétique classique - version 1d de (1.3.2). Seule la définition de l'équilibre de Gibbs est - très légèrement - modifiée par l'introduction de la hauteur totale de l'écoulement : ainsi dans la définition (1.3.1)

$$M(t,x,\xi) = M(h(t,x), \xi - u(t,x)) = \frac{h(t,x)}{\tilde{c}^2}\ \chi(\frac{\xi - u(t,x)}{\tilde{c}}),$$

le terme $\tilde{c}$ doit rester lié à la célérité des ondes dans l'écoulement et vérifie donc maintenant la relation $\sqrt{2}\tilde{c} = \sqrt{gH}$, où H est la hauteur totale de l'écoulement. Il s'ensuit que l'adaptation des algorithmes est immédiate et que les propriétés de stabilité des schémas cinétiques peuvent être étendues sans difficulté à ce nouveau problème, ce qui nous permet d'établir la proposition suivante

Proposition 1.6.1 (Chapitre 6, Pages 173)
Le schéma cinétique associé au système de Saint-Venant multicouche vérifie les proporiétés suivantes
(i) il préserve la positivité de la hauteur d'eau sous une condition de CFL,
(ii) il est consistant avec le système (1.6.3)-(1.6.8)
(iii) pour des données initiales adéquates, il préserve l'état stationnaire du lac au repos,
(iv) dans le cas d'un frottement nul et pour des conditions initiales adéquates, les solutions obtenues sont compatibles avec les solutions du schéma cinétique classique.

D'autre part le traitement numérique des termes sources s'appuie sur ce que nous avons dit plus haut et fait donc appel à des méthodes très classiques et très stables : résolution explicite - avec introduction de limiteurs - des termes sources faisant intervenir les hauteurs d'eau et traitement implicite des termes sources visqueux, ce qui conduit à la résolution d'un système linéaire tridiagonal, pour lequel nous établissons la proposition suivante

Proposition 1.6.2 (Chapitre 6, Pages 172)
Le système linéaire issu du traitement implicite des termes visqueux est toujours inversible.

Le coût de la résolution du système multicouche est clairement lié au nombre de couches, et, plus précisément, il est légèrement supérieur au produit du nombre de couches et du temps nécessaire à la résolution d'un problème de Saint-Venant classique. Néanmoins, et quelquesoit le nombre de couches, il préserve le caractère uni- ou bidimensionnel et l'invariance des maillages. De plus l'algorithme présenté est aisément parallelélisable, puisque les calculs volumes finis, sur chaque couche, peuvent être effectués indépendamment et simultanément. Il en va de même pour l'inversion du système linéaire, en chaque noeud du maillage.

Nous pouvons maintenant revenir à notre motivation première et comparer les solutions obtenues par cette méthode avec, d'une part, les solutions du système de Saint-Venant, et, d'autre part, les solutions des équations de Navier-Stokes. Nous établissons d'abord que le modèle multicouche agrandit le domaine de validité du système de Saint-Venant classique. Mais, surtout, nous montrons, sur plusieurs tests numériques, que les profils verticaux de vitesse horizontale issus du modèle multicouche fournissent une très bonne approximation des profils issus de calculs Navier-Stokes, et ce pour une très large gamme de coefficients de frottement. Un exemple de comparaison des profils verticaux est présenté sur la Figure 1.6.1

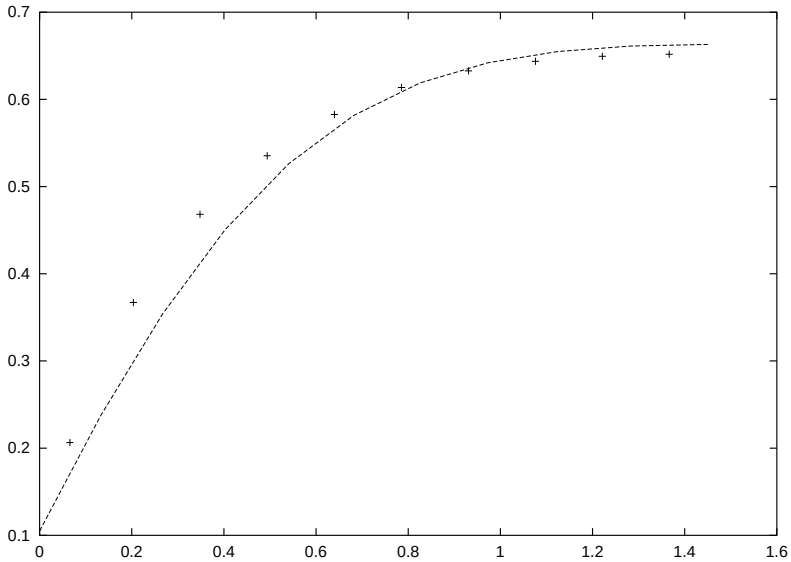


FIG. 1.6.1: VITESSE (Profil vertical) - Modèles Navier-Stokes et Saint-Venant multicouche - Dix couches - Coefficient de frottement = .1

1.7 Conclusions et Perspectives

Nous avons introduit dans cette thèse plusieurs idées nouvelles relatives à la modélisation hyperbolique des écoulements en eaux peu profondes et à l'analyse numérique des

modèles obtenus. Nous voudrions ici pointé les nouveaux développements que connaîssent actuellement ces outils.
Les schémas cinétiques sont connus depuis longtemps, mais leur application au système de Saint-Venant est récente [9] et leur adaptation au système de Saint-Venant multicouche est une des nouveautés de cette thèse. Notons qu'ils ont également été utilisés dans la discretization d'un système de Saint-Venant particulier, modélisant les avalanches [98].
Partant d'un solveur homogène positif, la reconstruction hydrostatique nous a permis de construire un schéma positif adapté au système de Saint-Venant avec termes sources, qui préserve l'état stationnaire associé au lac au repos. Des travaux en cours étudient la possibilité d'étendre la propriété de préservation des équilibres à tous les états stationnaires monodimensionnels (1.2.8) fluviaux - i.e. tels que le nombre de Froude (1.4.4) soit inférieur à un. D'autre part, l'idée de la méthode a été transposée avec succès au traitement d'un autre terme source : en présence de coefficients de frottement importants, une méthode de reconstruction des variables permet de rendre les solutions numériques du système de Saint-Venant cohérentes avec celles des équations de Darcy, qui sont la limite naturelle du système continu quand le coefficient de frottement tend vers l'infini. Enfin, l'idée d'une discretization conservative du terme de Coriolis, apparaissant dans les écoulements océaniques ou atmosphériques, est elle aussi à l'étude.
Le système de Saint-Venant multicouche introduit dans cette thèse doit encore faire l'objet de validations numériques. Puisqu'une des étapes de la dérivation du modèle multicouche est l'obtention d'un modèle hydrostatique, une comparaison avec les solutions des équations de Navier-Stokes hydrostatiques serait à ce titre particulièrement intéressante. Notons également que plusieurs extensions sont envisagées : développer un schéma équilibre adapté au système multicouche, développer un modèle multicouche axisymmétrique adapté aux écoulements sanguins...
Enfin de nouveaux développements de la méthode des entropies de Kruzkov partiellement adaptées, permettant de lever certaines des hypothèses nécessaires à la démonstration du théorème d'unicité, sont en cours.

Chapitre 2

A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows

Le travail présenté dans ce chapitre a été effectué en collaboration avec Francois Bouchut, Marie-Odile Bristeau, Rupert Klein et Benoit Perthame. Il a été publié dans *SIAM Journal of Scientific Computing*, Vol. **25** (2004), no. 6, pp 2050–2065.

2.1 Introduction

The classical Saint-Venant system for shallow water has been widely validated. It assumes a slowly varying topography $z(x)$ (x denotes a coordinate in the horizontal direction) and describes the height of water $h(t,x)$, and the water velocity $u(t,x)$ in the direction parallel to the bottom. It uses the following equations in one space dimension,

$$\begin{cases} \partial_t h + \partial_x(hu) = 0, \\ \partial_t(hu) + \partial_x(hu^2 + gh^2/2) = -hgz_x, \end{cases} \tag{2.1.1}$$

where $g > 0$ denotes the gravity constant. For future reference we denote the flux by $F(U) = (hu, hu^2 + gh^2/2)$, with $U = (h, hu)$. This model is very robust, being hyperbolic and admitting an entropy inequality (related to the physical energy)

$$\partial_t \widetilde{\eta}(U,z) + \partial_x \widetilde{G}(U,z) \leq 0, \tag{2.1.2}$$

where

$$\begin{aligned} \eta(U) &= hu^2/2 + \frac{g}{2}h^2, \qquad & G(U) &= \Big(hu^2/2 + gh^2\Big)u, \\ \widetilde{\eta}(U,z) &= \eta(U) + hgz, \qquad & \widetilde{G}(U,z) &= G(U) + hgzu. \end{aligned} \tag{2.1.3}$$

Another nice property is that it preserves the steady state of a lake at rest

$$h + z = Cst, \qquad u = 0. \tag{2.1.4}$$

When solving numerically (2.1.1), it is very important to be able to preserve these steady states at the discrete level and to accurately compute the evolution of small deviations from them, because the majority of real-life applications resides in this flow regime. Other steady states with non vanishing velocity can also be considered, but we shall not do so in the present work.

Since the early works of Leroux and coauthors [59], [62], schemes satisfying such a property are called well-balanced. Several schemes have been proposed that satisfy this property (exactly or at least at second-order), [91], [71], [57], [49], [131], [129], [11]. But the difficulty is then to get schemes that also satisfy very natural properties such as conservativity of the water height h, nonnegativity of h, the ability to compute dry states $h = 0$ and transcritical flows when the jacobian matrix F' of the flux function becomes singular (this difficulty is related to resonance, and theoretical studies can be found for instance in [97], [70]), and eventually to satisfy a discrete entropy inequality. This last property ensures the admissibility of shocks, and gives overall the nonlinear stability of the scheme. Theoretically, the exact Godunov scheme satisfies these requirements [87], but it is in practice computationally too expensive, and not easily adaptable to more complex systems, such as for example the models proposed in [22]. The first attempt to derive an approximate solver satisfying all the requirements was performed in [15] for a scalar equation (in this case, only the ability to treat transcritical flows with an entropy inequality is meaningful, together with the well-balanced property).

A generalization to the case of the Saint-Venant system was obtained in [108], and another method by relaxation is also proposed in [18]. However, these approximate solver methods are still quite heavy in practice. The aim of this paper is to explain how it is possible by a very flexible approach involving a hydrostatic reconstruction, to obtain a well-balanced scheme satisfying all the above requirements, and that is computationally inexpensive. The present approach unifies and generalizes ideas developed independently in [16, 17] for nearly hydrostatic, multi-dimensional compressible flow, and in [9] for the Saint-Venant shallow water model. In contrast with the above mentioned methods [108], [18], it is generic, in the sense that it can be used in conjunction with any given numerical flux for the homogeneous (i.e. with constant topography) Saint Venant problem.

2.2 Well-balanced scheme with hydrostatic reconstruction

2.2.1 Semi-discrete scheme

Finite volume schemes for hyperbolic systems consist in using an upwinding of the fluxes. In the semi-discrete case they provide a discrete version of (2.1.1) under the form

$$\Delta x_i \frac{d}{dt} U_i(t) + F_{i+1/2} - F_{i-1/2} = S_i, \tag{2.2.1}$$

where Δx_i denotes a possibly variable mesh size $\Delta x_i = x_{i+1/2} - x_{i-1/2}$, and the cell-centered vector of discrete unknowns is

$$U_i(t) = \begin{pmatrix} h_i(t) \\ h_i(t)u_i(t) \end{pmatrix}. \tag{2.2.2}$$

In a basic first-order accurate scheme, the fluxes are classically computed as $F_{i+1/2} = \mathcal{F}(U_i(t), U_{i+1}(t))$ with a numerical flux $\mathcal{F}$ that is computed via an approximate resolution of the Riemann problem (a so-called *solver*), which provides stability of the method. We refer to [54] for the description of the most well-known solvers : Godunov, Roe, Kinetic...It is known since [59],[62] that cell-centered evaluations of the source term in (2.2.1) will generally not be able to maintain in time steady states of a lake at rest, which are characterized by

$$h_i + z_i = Cst, \qquad u_i = 0. \tag{2.2.3}$$

Following [9], [16, 17], we propose and analyze finite volume schemes according to (2.2.1) with flux functions

$$F_{i+1/2} = \mathcal{F}(U_{i+1/2-}, U_{i+1/2+}), \tag{2.2.4}$$

where the interface values $U_{i+1/2-}, U_{i+1/2+}$ are derived from a local hydrostatic reconstruction to be described shortly, which is similar to second-order reconstructions in higher-order methods. The source term is discretized as

$$S_i = \begin{pmatrix} 0 \\ \frac{g}{2} h_{i+1/2-}^2 - \frac{g}{2} h_{i-1/2+}^2 \end{pmatrix} . \tag{2.2.5}$$

This ansatz is motivated by a balancing requirement, as follows. For nearly hydrostatic flows one has $u \ll \sqrt{gh}$. In the associated asymptotic limit the leading order water height $\underline{h}$ adjusts so as to satisfy the balance of momentum flux and momentum source terms, i.e.

$$\partial_x \left(\frac{g\underline{h}^2}{2} \right) = -\underline{h}\, g z_x \, . \tag{2.2.6}$$

Integrating over, say, the ith grid cell we obtain an approximation to the net source term as

$$- \int_{x_{i-1/2}}^{x_{i+1/2}} \underline{h}\, g z_x \, dx = \frac{g}{2} \underline{h}_{i+1/2-}^2 - \frac{g}{2} \underline{h}_{i-1/2+}^2 \, . \tag{2.2.7}$$

Thus we are able to locally represent the cell-averaged source term as the discrete gradient of the hydrostatic momentum flux, and this motivates the source term discretization in (2.2.5).

It is obvious now that any hydrostatic state is maintained exactly if, for such a state, the momentum fluxes in (2.2.1) and the locally reconstructed heights satisfy $F_{i+1/2}^{hu} = \frac{1}{2} g h_{i+1/2-}^2 = \frac{1}{2} g h_{i+1/2+}^2$. This is the motivation for (2.2.4), which gives this property if for hydrostatic states we have $U_{i+1/2-} = U_{i+1/2+} = (h_{i+1/2-}, 0) = (h_{i+1/2+}, 0)$.

The hydrostatic balance in (2.2.6) is equivalent to the "lake at rest" equation (2.1.4), so that the reconstruction of the leading order heights is straightforward,

$$\underline{h}_{i+1/2-} = h_i + z_i - z_{i+1/2}, \qquad \underline{h}_{i+1/2+} = h_{i+1} + z_{i+1} - z_{i+1/2}. \tag{2.2.8}$$

An important challenge is to design a scheme that robustly captures dry regions where $h \equiv 0$. In order to ensure nonnegativity of the water height even when cells begin to "dry out", we need first to perform a truncation of the leading order heights in (2.2.8), $h_{i+1/2\pm} = \max(0, \underline{h}_{i+1/2\pm})$. Next, the evaluation of the cell interface height $z_{i+1/2}$ has to be done in a quite subtle way. Our construction, combined with a centered value of $z_{i+1/2}$, is not stable. We rather take an *upwind* evaluation of the form

$$z_{i+1/2} = \max(z_i, z_{i+1}). \tag{2.2.9}$$

With these choices, we ensure that $0 \le h_{i+1/2-} \le h_i$ and $0 \le h_{i+1/2+} \le h_{i+1}$, and we prove below that this property ensures the nonnegativity requirement.

With these rules in place we can now summarize our first-order well-balanced finite volume scheme by

$$\Delta x_i \frac{d}{dt} U_i(t) + F_{i+1/2} - F_{i-1/2} = S_i, \tag{2.2.10}$$

where

$$F_{i+1/2} = \mathcal{F}(U_{i+1/2-}, U_{i+1/2+}), \tag{2.2.11}$$

$$U_{i+1/2-} = \begin{pmatrix} h_{i+1/2-} \\ h_{i+1/2-}\, u_i \end{pmatrix}, \qquad U_{i+1/2+} = \begin{pmatrix} h_{i+1/2+} \\ h_{i+1/2+}\, u_{i+1} \end{pmatrix}, \tag{2.2.12}$$

$$h_{i+1/2-} = \max(0, h_i + z_i - z_{i+1/2}), \qquad h_{i+1/2+} = \max(0, h_{i+1} + z_{i+1} - z_{i+1/2}), \tag{2.2.13}$$

and

$$S_i = S_{i+1/2-} + S_{i-1/2+} \equiv \begin{pmatrix} 0 \\ \frac{g}{2} h_{i+1/2-}^2 - \frac{g}{2} h_i^2 \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{g}{2} h_i^2 - \frac{g}{2} h_{i-1/2+}^2 \end{pmatrix}. \tag{2.2.14}$$

The latter expression for the source is equivalent to the earlier (2.2.5), it shows that the source may be considered as being distributed to the cell interfaces. With this re-interpretation in mind, we may also rewrite the scheme as

$$\Delta x_i \frac{d}{dt} U_i(t) + \mathcal{F}_l(U_i, U_{i+1}, z_i, z_{i+1}) - \mathcal{F}_r(U_{i-1}, U_i, z_{i-1}, z_i) = 0, \tag{2.2.15}$$

with left and right numerical fluxes

$$\begin{aligned} \mathcal{F}_l(U_i, U_{i+1}, z_i, z_{i+1}) &= F_{i+1/2} - S_{i+1/2-} \\ &= \mathcal{F}(U_{i+1/2-}, U_{i+1/2+}) + \begin{pmatrix} 0 \\ \frac{g}{2} h_i^2 - \frac{g}{2} h_{i+1/2-}^2 \end{pmatrix}, \\ \mathcal{F}_r(U_i, U_{i+1}, z_i, z_{i+1}) &= F_{i+1/2} + S_{i+1/2+} \\ &= \mathcal{F}(U_{i+1/2-}, U_{i+1/2+}) + \begin{pmatrix} 0 \\ \frac{g}{2} h_{i+1}^2 - \frac{g}{2} h_{i+1/2+}^2 \end{pmatrix}. \end{aligned} \tag{2.2.16}$$

Our construction is reminiscent of the formulas proposed in [62], [59], [57] using the full steady state equations to compute intermediate states at which the numerical flux is evaluated. The difference is that here, (2.2.12), (2.2.13) mean that we try to impose interface values satisfying some modified steady equations $h_{i+1/2-} + z_{i+1/2} = h_i + z_i$, $u_{i+1/2-} = u_i$, $h_{i+1/2+} + z_{i+1/2} = h_{i+1} + z_{i+1}$, $u_{i+1/2+} = u_{i+1}$, i.e. $h + z = cst$, $u = cst$ instead of Bernoulli's law $u^2/2 + g(h + z) = cst$, $hu = cst$. The advantage of these new relations is that now we have no singularity at critical points (observe that the numerical fluxes $\mathcal{F}_l$, $\mathcal{F}_r$ depend continuously on the data), while these relations coincide with the exact ones in the case of interest $u = 0$ corresponding to (2.2.3). Strikingly, this modification does not affect the consistency of the scheme, even in far from steady situations.

Theorem 2.2.1 *Consider a consistent numerical flux $\mathcal{F}$ for the homogeneous problem that preserves nonnegativity of $h_i(t)$ and satisfies an in-cell entropy inequality corresponding to the entropy η in (2.1.3). Then the finite volume scheme (2.2.9)-(2.2.14)*
(i) preserves the nonnegativity of $h_i(t)$,
(ii) is well-balanced, i.e. it preserves the steady state of a lake at rest (2.2.3),
(iii) is consistent with the Saint-Venant system (2.1.1),
(iv) satisfies an in-cell entropy inequality associated to the entropy $\widetilde{\eta}$ in (2.1.3),

$$\Delta x_i \frac{d}{dt}\widetilde{\eta}(U_i(t), z_i) + \widetilde{G}_{i+1/2} - \widetilde{G}_{i-1/2} \leq 0. \tag{2.2.17}$$

The statement that $\mathcal{F}$ preserves the nonnegativity of $h_i(t)$ means exactly that $\mathcal{F}^h(h_i = 0, u_i, h_{i+1}, u_{i+1}) - \mathcal{F}^h(h_{i-1}, u_{i-1}, h_i = 0, u_i) \leq 0$, for all choices of the other arguments. Since the sources in (2.2.14) have no contribution to the first component, $h_i(t)$ in our scheme satisfies a conservative equation with flux $\mathcal{F}^h(U_{i+1/2-}, U_{i+1/2+})$. Therefore we need to check that $\mathcal{F}^h(U_{i+1/2-}, U_{i+1/2+}) - \mathcal{F}^h(U_{i-1/2-}, U_{i-1/2+}) \leq 0$ whenever $h_i = 0$. As mentioned above, our construction (2.2.9), (2.2.13), ensures that $h_{i+1/2-} \leq h_i$ and $h_{i+1/2+} \leq h_{i+1}$, thus $h_{i+1/2-} = h_{i-1/2+} = 0$ when $h_i = 0$, and this gives (i).

Then we prove statement (ii). On a steady state of a lake at rest, we have $h_{i+1/2-} = h_{i+1/2+}$, $u_{i+1} = u_i = 0$, thus $U_{i+1/2-} = U_{i+1/2+}$ and by consistency of $\mathcal{F}$

$$F_{i+1/2} = F(U_{i+1/2-}) = F(U_{i+1/2+}) = \begin{pmatrix} 0 \\ \frac{g}{2}h^2_{i+1/2-} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{g}{2}h^2_{i+1/2+} \end{pmatrix}. \tag{2.2.18}$$

Together with the expression of the source terms in (2.2.14), we get $\mathcal{F}_l = F_{i+1/2} - S_{i+1/2-} = F(U_i)$, $\mathcal{F}_r = F_{i+1/2} + S_{i+1/2+} = F(U_{i+1})$, and this proves (ii).

To prove (iii), we apply the criterion in [109], [18], and we need to check two properties related to the consistency with the exact flux F and the consistency with the source. The consistency with the exact flux $\mathcal{F}_l(U, U, z, z) = \mathcal{F}_r(U, U, z, z) = F(U)$ is obvious since $U_{i+1/2-} = U_i$ and $U_{i+1/2+} = U_{i+1}$ whenever $z_{i+1} = z_i$. For consistency with the source, the criterion becomes for the Saint-Venant system

$$\mathcal{F}^{hu}_r(U_i, U_{i+1}, z_i, z_{i+1}) - \mathcal{F}^{hu}_l(U_i, U_{i+1}, z_i, z_{i+1}) = -hg\Delta z_{i+1/2} + o(\Delta z_{i+1/2}), \tag{2.2.19}$$

as $U_i, U_{i+1} \to U$ and $\Delta z_{i+1/2} \to 0$, where $\Delta z_{i+1/2} = z_{i+1} - z_i$. In our case,

$$\mathcal{F}_r - \mathcal{F}_l = S_{i+1/2-} + S_{i+1/2+} = \begin{pmatrix} 0 \\ \frac{g}{2}h^2_{i+1/2-} - \frac{g}{2}h^2_i + \frac{g}{2}h^2_{i+1} - \frac{g}{2}h^2_{i+1/2+} \end{pmatrix}. \tag{2.2.20}$$

Now, assuming $h > 0$, the maxima in (2.2.13) play no role if $h_i - h$, $h_{i+1} - h$ and $\Delta z_{i+1/2}$ are small enough. Thus we have $h^2_{i+1/2-}/2 - h^2_i/2 = h(z_i - z_{i+1/2}) + o(\Delta z_{i+1/2})$, $h^2_{i+1/2+}/2 - h^2_{i+1}/2 = h(z_{i+1} - z_{i+1/2}) + o(\Delta z_{i+1/2})$, which gives (2.2.19). In the special

case $h = 0$, the maxima in (2.2.13) can play a role only when $h_i = O(\Delta z_{i+1/2})$, and we conclude that (2.2.19) always holds, proving (iii).

In order to prove (iv), we first write that the original numerical flux $\mathcal{F}$ satisfies a semi-discrete entropy inequality. According to [18], this means that we can find a numerical entropy flux $\mathcal{G}$ such that

$$\begin{aligned} &G(U_{i+1}) + \eta'(U_{i+1})(\mathcal{F}(U_i, U_{i+1}) - F(U_{i+1})) \\ &\leq \mathcal{G}(U_i, U_{i+1}) \leq G(U_i) + \eta'(U_i)(\mathcal{F}(U_i, U_{i+1}) - F(U_i)), \end{aligned} \tag{2.2.21}$$

where η' is the derivative of η with respect to $U = (h, hu)$, $\eta'(U) = (gh - u^2/2, u)$. Similarly, having an entropy inequality (2.2.17) for (2.1.1) with $\widetilde{G}_{i+1/2} = \widetilde{\mathcal{G}}(U_i, U_{i+1}, z_i, z_{i+1})$ is equivalent to finding some numerical entropy flux $\widetilde{\mathcal{G}}$ such that

$$\begin{aligned} &\widetilde{G}(U_{i+1}, z_{i+1}) + \widetilde{\eta}'(U_{i+1}, z_{i+1})(\mathcal{F}_r(U_i, U_{i+1}, z_i, z_{i+1}) - F(U_{i+1})) \\ &\leq \widetilde{\mathcal{G}}(U_i, U_{i+1}, z_i, z_{i+1}) \leq \widetilde{G}(U_i, z_i) + \widetilde{\eta}'(U_i, z_i)(\mathcal{F}_l(U_i, U_{i+1}, z_i, z_{i+1}) - F(U_i)). \end{aligned} \tag{2.2.22}$$

Let us prove that (2.2.22) holds with

$$\widetilde{\mathcal{G}}(U_i, U_{i+1}, z_i, z_{i+1}) = \mathcal{G}(U_{i+1/2-}, U_{i+1/2+}) + \mathcal{F}^h(U_{i+1/2-}, U_{i+1/2+}) g z_{i+1/2}. \tag{2.2.23}$$

Since both inequalities are obtained by the same type of estimates, let us prove only the upper inequality involving $\mathcal{F}_l$ in (2.2.22). By comparison to (2.2.21), it is enough to prove that

$$\begin{aligned} &G(U_{i+1/2-}) + \eta'(U_{i+1/2-})\big(\mathcal{F}(U_{i+1/2-}, U_{i+1/2+}) - F(U_{i+1/2-})\big) \\ &\quad + \mathcal{F}^h(U_{i+1/2-}, U_{i+1/2+}) g z_{i+1/2} \\ &\leq G(U_i) + \eta'(U_i)(\mathcal{F}_l - F(U_i)) + \mathcal{F}^h(U_{i+1/2-}, U_{i+1/2+}) g z_i. \end{aligned} \tag{2.2.24}$$

This inequality can be written, by denoting $\mathcal{F} = (\mathcal{F}^h, \mathcal{F}^{hu}) = \mathcal{F}(U_{i+1/2-}, U_{i+1/2+})$,

$$\begin{aligned} &(u_i^2/2 + g h_{i+1/2-}) h_{i+1/2-} u_i + (g h_{i+1/2-} - u_i^2/2)(\mathcal{F}^h - h_{i+1/2-} u_i) \\ &\quad + u_i(\mathcal{F}^{hu} - h_{i+1/2-} u_i^2 - g h_{i+1/2-}^2/2) + \mathcal{F}^h g(z_{i+1/2} - z_i) \\ &\leq (u_i^2/2 + g h_i) h_i u_i + (g h_i - u_i^2/2)(\mathcal{F}^h - h_i u_i) + u_i(\mathcal{F}_l^{hu} - h_i u_i^2 - g h_i^2/2), \end{aligned} \tag{2.2.25}$$

or after simplification

$$u_i(\mathcal{F}^{hu} - g h_{i+1/2-}^2/2) + \mathcal{F}^h g(h_{i+1/2-} - h_i + z_{i+1/2} - z_i) \leq u_i(\mathcal{F}_l^{hu} - g h_i^2/2). \tag{2.2.26}$$

Since $\mathcal{F}_l^{hu} - g h_i^2/2 = \mathcal{F}^{hu} - g h_{i+1/2-}^2/2$ by definition of $\mathcal{F}_l$ in (2.2.16), our inequality finally reduces to

$$\mathcal{F}^h(U_{i+1/2-}, U_{i+1/2+})(h_{i+1/2-} - h_i + z_{i+1/2} - z_i) \leq 0. \tag{2.2.27}$$

Now, according to (2.2.13), when this quantity is nonzero, we have $h_{i+1/2-} = 0$ and the expression between parentheses is nonnegative. But since $\mathcal{F}$ preserves nonnegativity, we have $\mathcal{F}^h(h_{i+1/2-} = 0, u_i, h_{i+1/2+}, u_{i+1}) \leq 0$ and we conclude that (2.2.27) always holds. This completes the proof of (iv). □

2.2.2 Fully discrete scheme and CFL condition

When using the time-space fully discrete scheme

$$U_i^{n+1} - U_i^n + \frac{\Delta t}{\Delta x_i}\Big(\mathcal{F}_l(U_i, U_{i+1}, z_i, z_{i+1}) - \mathcal{F}_r(U_{i-1}, U_i, z_{i-1}, z_i)\Big) = 0, \tag{2.2.28}$$

the consistency and the well-balanced property are of course still valid. The question is then to obtain a CFL condition that guarantees stability.

One can prove that our hydrostatic reconstruction scheme does not satisfy a fully discrete entropy inequality. Indeed there exist some data with $h_i + z_i = cst$, $u_i = cst \neq 0$ such that for any $\Delta t > 0$, the fully discrete entropy inequality $\widetilde{\eta}(U_i^{n+1}, z_i) - \widetilde{\eta}(U_i^n, z_i) + \frac{\Delta t}{\Delta x_i}(\widetilde{G}_{i+1/2} - \widetilde{G}_{i-1/2}) \leq 0$ is violated. However these data are not preserved by the scheme. The consequence is that in practice we do not observe instabilities, as long as the water height h_i remains nonnegative.

In order to preserve the nonnegativity of h_i, the CFL condition that needs to be used is not more restrictive than that of the homogeneous solver.

Proposition 2.2.1 *Assume that the homogeneous flux $\mathcal{F}$ preserves the nonnegativity of h by interface with a numerical speed $\sigma(U_i, U_{i+1}) \geq 0$, which means that whenever the CFL condition*

$$\sigma(U_i, U_{i+1})\Delta t \leq \min(\Delta x_i, \Delta x_{i+1}) \tag{2.2.29}$$

holds, we have

$$\begin{aligned} &h_i - \frac{\Delta t}{\Delta x_i}(\mathcal{F}^h(U_i, U_{i+1}) - h_i u_i) \geq 0, \\ &h_{i+1} - \frac{\Delta t}{\Delta x_{i+1}}(h_{i+1}u_{i+1} - \mathcal{F}^h(U_i, U_{i+1})) \geq 0. \end{aligned} \tag{2.2.30}$$

Then the fully discrete hydrostatic reconstruction scheme (2.2.28) also preserves the nonnegativity of h by interface,

$$\begin{aligned} &h_i - \frac{\Delta t}{\Delta x_i}(\mathcal{F}^h(U_{i+1/2-}, U_{i+1/2+}) - h_i u_i) \geq 0, \\ &h_{i+1} - \frac{\Delta t}{\Delta x_{i+1}}(h_{i+1}u_{i+1} - \mathcal{F}^h(U_{i+1/2-}, U_{i+1/2+})) \geq 0, \end{aligned} \tag{2.2.31}$$

under the CFL condition

$$\sigma(U_{i+1/2-}, U_{i+1/2+})\Delta t \leq \min(\Delta x_i, \Delta x_{i+1}). \tag{2.2.32}$$

Under the CFL condition (2.2.32), we have

$$\begin{aligned} &h_{i+1/2-} - \frac{\Delta t}{\Delta x_i}(\mathcal{F}^h(U_{i+1/2-}, U_{i+1/2+}) - h_{i+1/2-}u_{i+1/2-}) \geq 0, \\ &h_{i+1/2+} - \frac{\Delta t}{\Delta x_{i+1}}(h_{i+1/2+}u_{i+1/2+} - \mathcal{F}^h(U_{i+1/2-}, U_{i+1/2+})) \geq 0. \end{aligned} \tag{2.2.33}$$

As previously mentioned, with the choice (2.2.9), (2.2.13), we have $h_{i+1/2-} \leq h_i$ and $h_{i+1/2+} \leq h_{i+1}$. Thus we deduce that (2.2.31) holds as soon as $1 + u_i \Delta t/\Delta x_i \geq 0$ and $1 - u_{i+1}\Delta t/\Delta x_{i+1} \geq 0$, which is necessarily the case from (2.2.32). □

2.3 Second-order extension

Starting from a given first-order method, a common way to obtain a second-order extension is, as it was already mentioned before for the analogy with our hydrostatic reconstruction, to compute the fluxes from limited reconstructed values on both sides of each interface rather than cell-centered values, see [54], [90] or [124]. These new values are classically obtained with three ingredients : prediction of the gradients in each cell, linear extrapolation, and limitation procedure.

In the presence of a source and in the context of well-balanced schemes, this approach needs to be described in more detail. In particular, according to [76], [77], [18], since not only the reconstructed values $U_{i,r}$ at $i+1/2-$, and $U_{i+1,l}$ at $i+1/2+$ need be defined but also $z_{i,r}$, $z_{i+1,l}$, a cell-centered source term S_{ci} must by added to preserve the consistency. We remark that even if z_i do not depend on time, the reconstructed values $z_{i,l}$, $z_{i,r}$ could depend on time via a coupling with U_i in the reconstruction step. Once these second-order reconstructed values are known, we apply the hydrostatic reconstruction scheme exposed in the previous section at each interface. This gives the second-order well-balanced scheme

$$\Delta x_i \frac{d}{dt} U_i(t) + F_{i+1/2} - F_{i-1/2} = S_i + S_{ci}, \tag{2.3.1}$$

where

$$F_{i+1/2} = \mathcal{F}(U_{i+1/2-}, U_{i+1/2+}), \tag{2.3.2}$$

$$U_{i+1/2-} = \begin{pmatrix} h_{i+1/2-} \\ h_{i+1/2-}\ u_{i,r} \end{pmatrix}, \qquad U_{i+1/2+} = \begin{pmatrix} h_{i+1/2+} \\ h_{i+1/2+}\ u_{i+1,l} \end{pmatrix}, \tag{2.3.3}$$

and the hydrostatic reconstruction is now

$$h_{i+1/2-} = \max(0, h_{i,r} + z_{i,r} - z_{i+1/2}), \qquad h_{i+1/2+} = \max(0, h_{i+1,l} + z_{i+1,l} - z_{i+1/2}), \tag{2.3.4}$$

with

$$z_{i+1/2} = \max(z_{i,r}, z_{i+1,l}). \tag{2.3.5}$$

The source term is distributed as before at the interfaces,

$$S_i = S_{i+1/2-} + S_{i-1/2+}, \tag{2.3.6}$$

$$S_{i+1/2-} = \begin{pmatrix} 0 \\ \frac{g}{2} h_{i+1/2-}^2 - \frac{g}{2} h_{i,r}^2 \end{pmatrix}, \quad S_{i-1/2+} = \begin{pmatrix} 0 \\ \frac{g}{2} h_{i,l}^2 - \frac{g}{2} h_{i-1/2+}^2 \end{pmatrix}. \tag{2.3.7}$$

A simple well-balanced choice for the centered source term S_{ci} is

$$S_{ci} = \begin{pmatrix} 0 \\ g\frac{h_{i,l}+h_{i,r}}{2}(z_{i,l} - z_{i,r}) \end{pmatrix}. \tag{2.3.8}$$

Using the definitions of the left and right numerical fluxes $\mathcal{F}_l$, $\mathcal{F}_r$ in (2.2.16), a compact formulation of the scheme is

$$\Delta x_i \frac{d}{dt} U_i(t) + \mathcal{F}_l(U_{i,r}, U_{i+1,l}, z_{i,r}, z_{i+1,l}) - \mathcal{F}_r(U_{i-1,r}, U_{i,l}, z_{i-1,r}, z_{i,l}) = S_{ci}. \tag{2.3.9}$$

This formulation ensures that the second-order scheme inherits the stability properties of the first-order one.

Theorem 2.3.1 *Consider a consistent numerical flux $\mathcal{F}$ for the homogeneous problem that preserves nonnegativity of $h_i(t)$. Assume that the second-order reconstruction gives nonnegative values $h_{i,l}$, $h_{i,r}$, is well-balanced and is second-order-centered in z, which means by definition that whenever the sequences (U_i) and (z_i) are the cell averages of smooth functions $U(x)$, $z(x)$, we have*

$$\begin{aligned} z_{i+1,l} - z_{i,r} &= O\left((\Delta x_i + \Delta x_{i+1})^3\right), \\ \frac{z_{i,r} - z_{i,l}}{\Delta x_i} &= z_x(x_i) + O\left((\Delta x_{i-1} + \Delta x_i + \Delta x_{i+1})^2\right). \end{aligned} \tag{2.3.10}$$

Then the finite volume scheme (2.3.1)-(2.3.8) preserves the nonnegativity of $h_i(t)$, is well-balanced, i.e. it preserves the steady states of a lake at rest (2.2.3), and is second-order accurate.

It is well known that the second-order reconstruction strategy preserves the nonnegativity of the water height (under a half CFL condition in the fully discrete case). Here only the centered source term S_{ci} in (2.3.9) could cause difficulties, but it does not since its first component vanishes.

The preservation of the lake at rest steady states can be checked easily from the property of the second-order reconstruction to be well-balanced, which means by definition that if $u_i = 0$ and $h_i + z_i = h_{i+1} + z_{i+1}$ for all i, then $u_{i,l} = u_{i,r} = 0$ and $h_{i,l} + z_{i,l} = h_{i,r} + z_{i,r} = h_i + z_i$ for all i. Indeed we just have to notice that for a steady state, $S_{ci} = (0, g(h_{i,r}^2 - h_{i,l}^2)/2)$.

In order to prove the second-order accuracy, let us assume that (U_i) and (z_i) are realized as the cell averages of smooth functions $U(x)$ and $z(x)$, and denote by $\hbar$ the mesh size. Then, since we assumed implicitly that the second-order reconstruction is second-order, we have that $U_{i,r} = U(x_{i+1/2}) + O(\hbar^2)$, $U_{i+1,l} = U(x_{i+1/2}) + O(\hbar^2)$, $z_{i,r} = z(x_{i+1/2}) + O(\hbar^2)$, $z_{i+1,l} = z(x_{i+1/2}) + O(\hbar^2)$. It follows from (2.3.3)-(2.3.5) that $U_{i+1/2\pm} = U(x_{i+1/2}) + O(\hbar^2)$, thus by (2.3.2) $F_{i+1/2} = F(U(x_{i+1/2})) + O(\hbar^2)$. This proves the second-order accuracy in the weak sense of the flux difference in (2.3.1) since this part is in conservative form. For the right-hand side, there is no

such cancellation thus we can only allow errors in $O(\Delta x_i \hbar^2)$ in (2.3.1). We have $(h_{i,l} + h_{i,r})/2 = h(x_i) + O(\hbar^2)$, and the second expansion in (2.3.10) yields with (2.3.8) that $S_{ci} = \big(0, -gh(x_i)z_x(x_i)\Delta x_i + O(\Delta x_i \hbar^2)\big) = \int_{x_{i-1/2}}^{x_{i+1/2}} (0, -gh(x)z_x(x))\, dx + O(\Delta x_i \hbar^2)$. Since $S_{i+1/2\pm} = O(z_{i+1,l} - z_{i,r}) = O(\hbar^3)$ by the first expansion in (2.3.10), this gives that $S_i = O(\hbar^3)$ and concludes the proof in the "regular" case when $\hbar = O(\Delta x_i)$, by just considering S_i as an error in (2.3.1). In the general case, we have to introduce a slightly different interpretation of the scheme via a weighted average flux

$$\widetilde{F}_{i+1/2} = \frac{\Delta x_{i+1}\mathcal{F}_l + \Delta x_i \mathcal{F}_r}{\Delta x_i + \Delta x_{i+1}} = F_{i+1/2} + \frac{\Delta x_i S_{i+1/2+} - \Delta x_{i+1} S_{i+1/2-}}{\Delta x_i + \Delta x_{i+1}}. \tag{2.3.11}$$

By the first line in (2.3.10), we have $\widetilde{F}_{i+1/2} = F_{i+1/2} - S_{i+1/2-} + O(\Delta x_i \hbar^2)$ and also $\widetilde{F}_{i+1/2} = F_{i+1/2} + S_{i+1/2+} + O(\Delta x_{i+1} \hbar^2)$. Therefore,

$$\begin{aligned} \widetilde{F}_{i+1/2} - \widetilde{F}_{i-1/2} &= F_{i+1/2} - F_{i-1/2} - S_{i+1/2-} - S_{i-1/2+} + O(\Delta x_i \hbar^2) \\ &= F_{i+1/2} - F_{i-1/2} - S_i + O(\Delta x_i \hbar^2), \end{aligned} \tag{2.3.12}$$

and (2.3.1) can be rewritten as

$$\Delta x_i \frac{d}{dt} U_i(t) + \widetilde{F}_{i+1/2} - \widetilde{F}_{i-1/2} = S_{ci} + O(\Delta x_i \hbar^2), \tag{2.3.13}$$

which proves the second-order accuracy. □

Some important features arising in the second-order reconstruction must now be specified. First, the cell by cell reconstruction preserves the mass conservation property of the finite volume method. Second, the limitation procedure ensures the nonnegativity of the second-order reconstructed water heights. The third important point is that the second-order reconstruction has to preserve the lake at rest steady state. To ensure this property we reconstruct also the bottom topography $z(x)$ although it is a data. The idea to do so is not so new, see [88], [49], but here we give details on the more stable way to do it. Indeed only two of the three quantities $h, z, h+z$ need be explicitly reconstructed, the last being necessarily a combination of the other two. A critical test to make the right choice is given by considering a lake at rest with non vertical shores, that is nothing else but considering an interface between a wet cell and a dry cell in the case where the bottom of the dry cell is higher than the free surface in the wet cell and where the fluid is at rest in the wet cell. As it appears in Figure 2.3.1, for the minmod reconstruction, the only choice which preserves the steady state *and* the nonnegativity of the water height at a wet-dry interface, is to work with the quantities h and $h+z$. Notice that it follows that in some respect the bottom topography changes at each timestep. This choice is consistent with the strategy for second-order extensions of a well-balanced Godunov-type scheme to multi-dimensional compressible flow under gravity in [16, 17], in that the *deviations from the non-constant steady state* form the basis for reconstruction and slope limiting (even if this rule does not exclude the worst

choice in the context of a wet-dry interface, namely to reconstruct z and $h+z$). It is obvious then that the chosen second-order reconstruction preserves also the steady state in the classical case of wet-wet interfaces since we explicitly reconstruct the quantity $h+z$. The second-order-centered condition (2.3.10) can be realized with a second-order ENO reconstruction for example, but in practice we shall not do so because it becomes too complicate for 2d unstructured meshes, even if it necessarily means a slight loss of accuracy. Second-order accuracy in time can be obtained as usual by a convex two-step integration of (2.3.9), and the CFL condition need not be modified.

2.4 Numerical results

All numerical tests are performed with a kinetic solver for the homogeneous problem. This solver is based on the kinetic theory developed in [105] and has the advantage to keep the water height nonnegative, to verify a discrete in-cell entropy inequality and to be able to compute problems with shocks or vacuum. First and second-order in space computations are proposed, but only first-order in time is used.

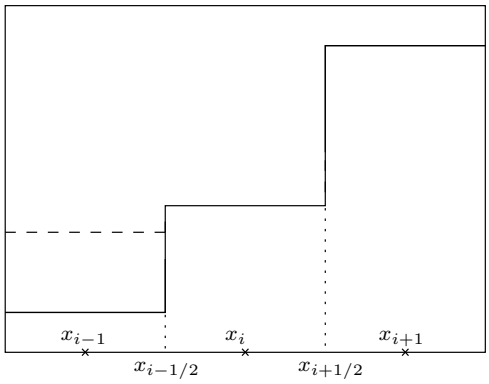
2.4.1 1d assessments

We first illustrate that the hydrostatic reconstruction does not affect the robustness of the homogeneous solver. We present a very classical numerical test of a constant discharge transcritical flow with shock over a bump, we refer to [60] for a complete presentation. In Figures 2.4.1 and 2.4.2, where 101 points are used, we observe good first and second-order results for this test, the stiffness of which is well-known. As we are far from a hydrostatic steady state the results of the well-balanced and standard schemes are quite similar. Notice however that the well-balanced version is less affected, whatever is the order of resolution, than the standard scheme where the derivative of the bottom topography presents strong variations.

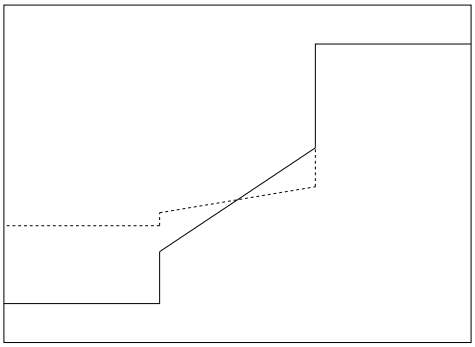
To exhibit the improvement due to the hydrostatic reconstruction we present now a quasi stationary case first proposed by Leveque in [91] which consists in computing small perturbations of the steady state of a lake at rest with a varying bottom topography,

$$\begin{aligned} z(x) &= \left(0.25\left(\cos\left(\pi(x-1.5)/0.1\right)+1\right)\right)_+, \\ h(0,x) &= 1. + 0.001 \mathbb{1}_{1.1\le x\le 1.2}. \end{aligned}$$

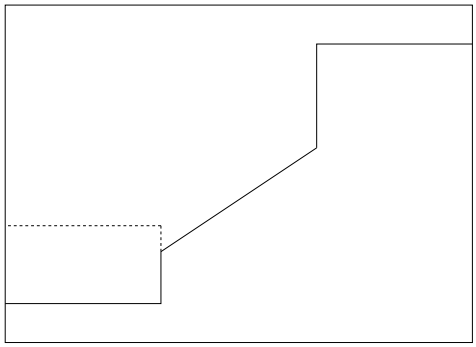
As we can see by considering linearized equations, the small perturbation simply moves to the right with a speed equal to $\sqrt{h(t,x)}$, i.e. $\sqrt{1-z(x)}$ at first-order approximation (gravity is equal to one). We present in Figure 2.4.3 the results obtained at $t = 0.7s$



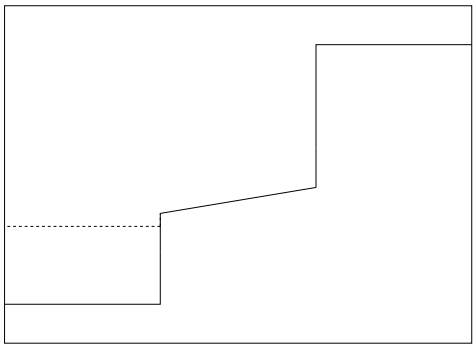
First-order solution



Reconstructed quantities : z and h+z



Reconstructed quantities : z and h



Reconstructed quantities : h and h+z

FIG. 2.3.1: Second-order reconstruction strategy

Free surface (dotted line)
Bottom topography (continuous line)

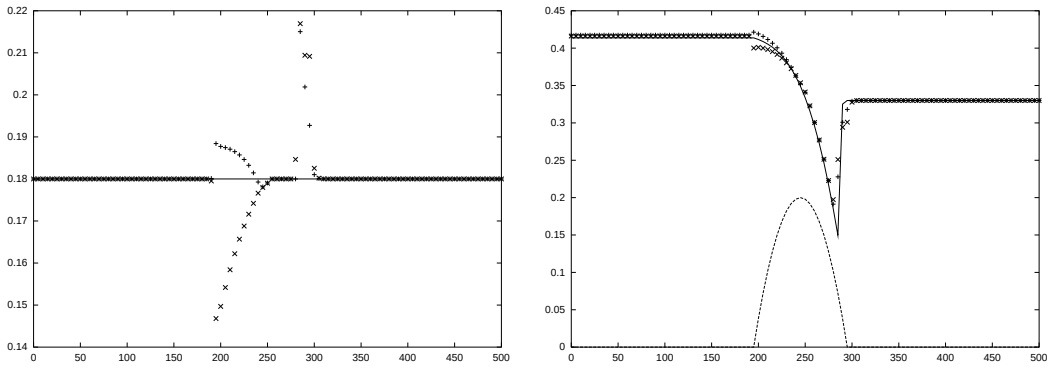


FIG. 2.4.1: Constant discharge problem with shock - Discharge and water height

First-order standard scheme (times crosses)
First-order well-balanced scheme (plus crosses)
Exact solution and bottom topography (continuous and dotted lines)

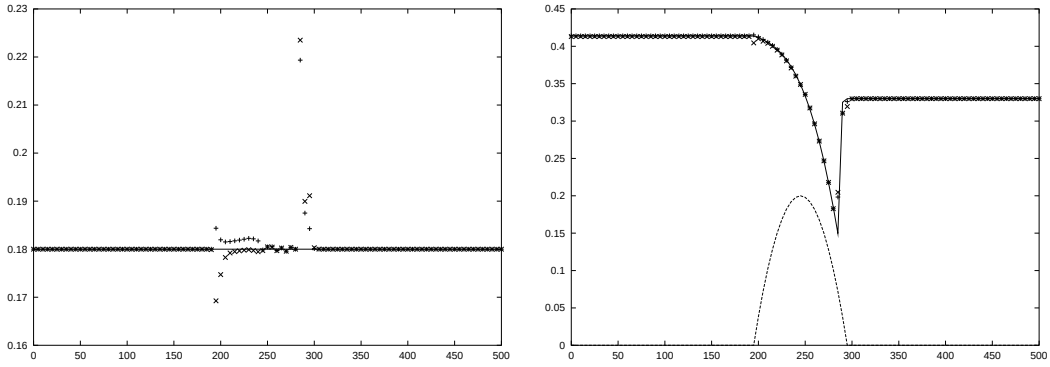
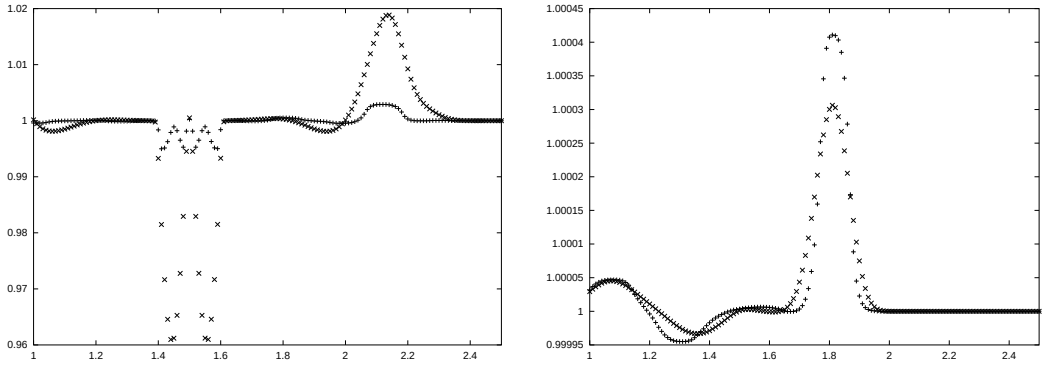


FIG. 2.4.2: Constant discharge problem with shock - Discharge and water height

Second-order standard scheme (times crosses)
Second-order well-balanced scheme (plus crosses)
Exact solution and bottom topography (continuous and dotted lines)

and with 150 points, with the well-balanced scheme on the right, and with the standard one on the left. Notice that the scale is not the same on both graphs. It appears that even for the second-order computation the unphysical perturbations induced by the standard scheme are larger than the initial perturbation of the free surface. Moreover, the standard scheme induces not only a perturbation of the bump, but also a perturbation which moves to the right at the same speed as the initial perturbation, but which is more than one order of magnitude greater than the initial perturbation. On the contrary, the results obtained with the well-balanced scheme are good, even for the first-order solution.



Standard scheme Well-balanced scheme

FIG. 2.4.3: Quasi stationary problem with small perturbation

First-order scheme (times crosses)
Second-order scheme (plus crosses)

To finish our assessment of scheme performance in one space dimension, we present a test case which is indicative of the robustness of a solver, as it involves vacuum conditions. It exhibits very clearly the improvement due to the second-order extension. We are interested in the case of an oscillating lake with a non flat bottom and non vertical shores. The lake is initially at rest but a small sinusoidal perturbation affects the free surface,

$$z(x) = .5(1 - .5(\cos(\pi(x - .5)/.5) + 1)),$$
$$h(0, x) = \max(0, .4 - z(x) + .04\sin((x - .5)/.25) - \max(0, -.4 + z(x))).$$

Then the flow oscillates, and at each timestep we have to treat an interface between a wet cell and a dry cell on each shore of the lake. We present in Figure 2.4.4 the results obtained with the well-balanced scheme with 200 points at $t = 19.87s$, because

it corresponds to a time where the flow reaches its higher level on the left shore. Both first and second-order well-balanced schemes are robust but the first-order scheme damps the oscillations much faster, fifty oscillations are enough to get back to rest. On the other hand the second-order well-balanced scheme keeps the periodic regime up to the machine accuracy.

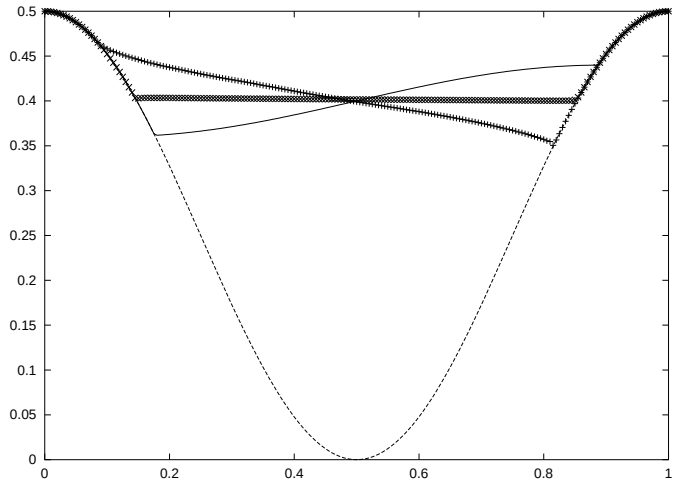


FIG. 2.4.4: Oscillating lake - Well-balanced scheme

First-order scheme (times crosses)
Second-order scheme (plus crosses)
Initial solution and bottom topography (continuous and dotted lines)

2.4.2 2d assessments

As the extension to second-order accuracy, the extension to the bidimensional case does not modify the idea of the method. The one-dimensional solver is used at each cell interface, and we have a numerical flux on each side of the interface, which are computed by $\mathcal{F}_l$ and $\mathcal{F}_r$ in (2.2.16) after an appropriate rotation. In this way, a piece of the source term is naturally discretized at the interface. The scheme is automatically well-balanced in the sense that lake at rest steady states $u = 0$, $h + z = cst$ are exactly preserved, and the water height remains nonnegative. However, some specific problems arise, especially in the construction of a 2d well-balanced second-order scheme. We refer to [9], [10] for a detailed description, especially for the explanation of the 2d hydrostatic second-order reconstruction.

We first present the academic case of a dam break on a dry bed but containing a wet zone which consists in a small lake at rest. This case involves the vacuum and it

allows to exhibit the effect of the hydrostatic reconstruction to preserve the initially at rest area. The first subfigure in Figure 2.4.5 presents the mesh and the bottom topography. The bottom topography of the lake we can see on the right is hemispheric. On the second subfigure we can see the initial water height : we see the dam in the middle and the small lake at rest on the right. The free surface level in the lake coincides with the reference level of the bottom topography of the river, i.e. $(h + z)(0, x, y) = 0$ everywhere on the right of the dam. On the third subfigure we can see the rarefaction wave. Since it does not yet reach the lake, the steady state is preserved. Then on the fourth subfigure the rarefaction wave reaches the lake and the water begins to move. On the last subfigure is presented the free surface level at this final time. We can notice strong variations on the lake area which lead to the formation of a hole in the left part of the lake, in blue on the subfigure, and a bump in the right part, in green in the subfigure.

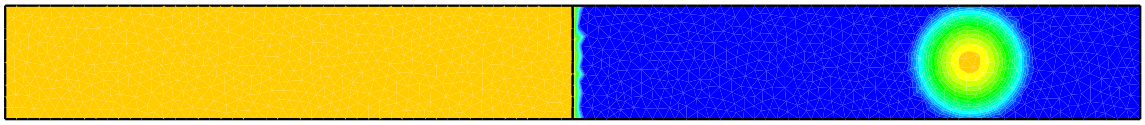
Then we present in Figure 2.4.6 another 2d numerical test corresponding to the filling up of a river. This test still involves vacuum but also deals with complex realistic geometry and bottom topography since it takes into account a jetty in the transversal direction, in the upper part of the figures ; a bridge pillar, the square on the lower part ; and a small bump in the bottom topography. We start with an empty river and we prescribe a given water level as inflow condition. On the first subfigure are presented the mesh and the associated bottom topography. Then we can notice that the strong variations in the bottom topography due to the jetty or the pillar bridge does not affect the robustness of the computation. On the third and fourth subfigures we can see the bump since the water skirts it.

More results can be found in [9], [23], [108], [98], and in [21] with Coriolis force.

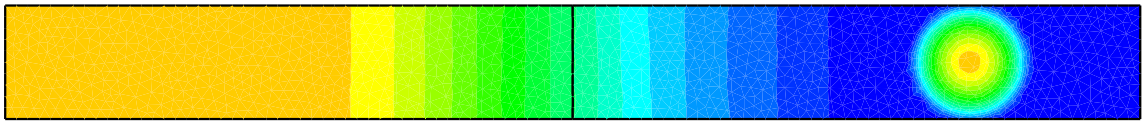
Acknowledgements

This work was partially supported by ACI Modélisation de processus hydrauliques à surface libre en présence de singularités (http ://www-rocq.inria.fr/m3n/CatNat/), by HYKE European programme HPRN-CT-2002-00282 (http ://www.hyke.org), by EDF/LNHE (E.A., F.B., M.-O.B., B.P.), and by grant KL 611/6 of the Deutsche Forschungsgemeinschaft (R.K.).

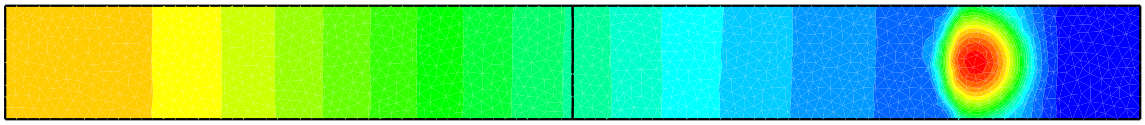
Mesh and bottom topography



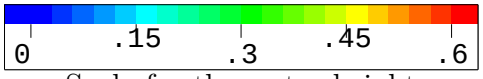
Water height - Initial Solution



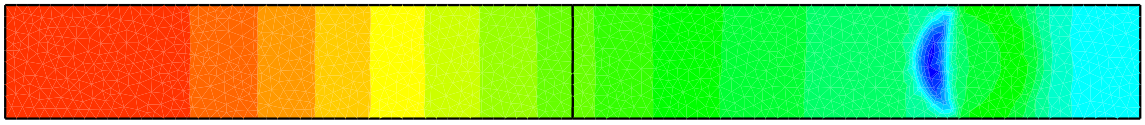
Water height - Solution at t=.08s



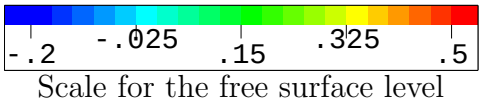
Water height - Solution at t=.16s



Scale for the water height



Free surface - Solution at t=.16s



Scale for the free surface level

FIG. 2.4.5: 2D dam break on dry bed with a lake at rest area

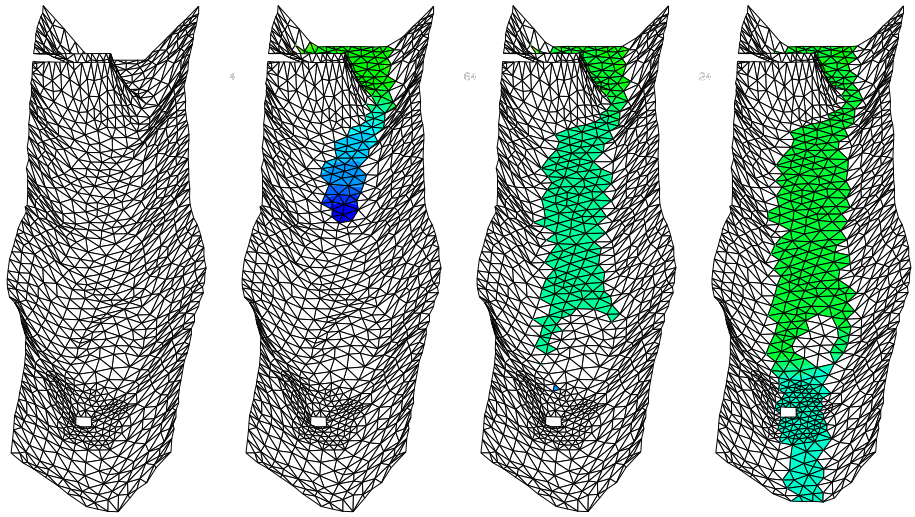


FIG. 2.4.6: Filling up of a river - Bottom topography and free surface at different times

Chapitre 3

A second order well-balanced positivity preserving scheme for the Saint-Venant system on unstructured grids

Le travail présenté dans ce chapitre a été effectué en collaboration avec Marie-Odile Bristeau. Il a été soumis pour publication.

3.1 Introduction

We consider in this article the 2D Saint-Venant system with topographic source term. This system introduced in [115] is very commonly used for the numerical simulation of various geophysical shallow-water flows, such as rivers, lakes or coastal areas, or even oceans, atmosphere or avalanches [22, 61] when completed with appropriate terms. It can be derived as a formal first order approximation of the three-dimensional free surface incompressible Navier-Stokes equations, using the so-called shallow water assumption [45, 50]. Usually, several other terms are added in order to take into account frictions on the bottom and the surface and other physical features. One can also describe the evolution of a temperature (or a concentration of a pollutant) advected by the flow by adding a third equation to the system [8].

The difficulty to define accurate numerical schemes is related to the deep mathematical structure of such hyperbolic systems ; the first existence proof of weak solutions after shocks in the large is due to [95] in 1995. It is based on the kinetic interpretation of the system which is also a method to derive numerical schemes with good properties.

Even if some authors use a finite element framework [66, 78], a classical approach for solving hyperbolic systems of conservation laws consists in using finite volume schemes (see [54, 89, 18]). In the context of the discretization of the Saint-Venant system, it is important to get schemes that satisfy very natural properties such as positivity of the water depth, ability to compute dry areas, preservation of steady states such as still water equilibrium (see [7] and the references therein), and eventually satisfy a discrete entropy inequality [108]. For hyperbolic systems with source terms, the difficulty to build stable schemes which preserve equilibria was pointed out by several authors, see Bermudez and Vasquez [13], Greenberg and Leroux [62], LeVeque [91] and led to the notion of *well-balanced* schemes.

The difficulty to allow h to vanish in a Roe solver is treated in [101]. Different approaches to satisfy the well-balanced property have been proposed. The Roe solver [112] has been modified in order to preserve steady states in Bermudez and Vasquez [13]. A two-dimensional extension is performed in Bermudez et al [12] and recent extensions to other types of homogeneous solvers can be found in [33, 34]. In [71] Jin proposes an other way to adapt the Roe solver to the non homogeneous case. Following the idea of Leroux et al. [62] for the scalar case, Gosse [57, 58] or Gallouët et al. [49] construct numerical schemes based on the solution of the Riemann problem associated with a larger system where a third equation on the variable describing the bottom topography is added. Another approach by LeVeque [91] is based on using the Godunov scheme for reconstructed variables. Approaches based on central schemes are used in Kurganov and Levy [84] or Russo [114]. Notice also that for the one-dimensional system, Perthame and Simeoni propose in [108] a kinetic method that includes the source term in the kinetic formulation and so that allows to ensure a discrete in-cell entropy inequality. For the simpler case of a scalar conservation law, a kinetic scheme that preserves equilibrium and which is proved to be convergent, is presented in Botchorishvili et al

[15]. Excepted in the two last references, preservation of steady states and positivity of the water depth are never treated at the same time.

In this paper we describe a possible solution method for the 2D Saint-Venant system using a *kinetic solver* and a *hydrostatic reconstruction* procedure. We first present the kinetic formulation of the Saint-Venant system and how to deduce a macroscopic scheme - for the homogeneous case - with good stability properties as the built-in preservation of the water depth positivity even when applications with dry areas are considered and existence of a discrete in-cell entropy inequality. We refer to Perthame [105] for a survey of the theoretical properties of the kinetic schemes. Then we present a *well-balanced* extension that is based on a *hydrostatic reconstruction* of the water depth and preserves the still water equilibrium while satisfying the stability properties of the homogeneous solver. We finally present and detail a formally second order extension which increases the accuracy of the results while also preserving the equilibrium and stability properties of the scheme.

The outline of the paper is the following. After recalling the 2D shallow water equations and their main properties in Section 2, we introduce the kinetic representation of the system in Section 3. In Section 4 we deduce a standard kinetic scheme, then in Section 5 we introduce the hydrostatic reconstruction and we deduce a *well-balanced* kinetic scheme which is adapted to the still water equilibrium. In Section 6, a second order extension is developed. The Section 7 illustrates the possibilities of the method by numerical results.

3.2 The Saint-Venant system

3.2.1 Equations

We consider the 2D Saint-Venant system, written in its physical conservative form

$$\frac{\partial h}{\partial t} + \mathrm{div}(h\mathbf{u}) = 0, \tag{3.2.1}$$

$$\frac{\partial h\mathbf{u}}{\partial t} + \mathrm{div}(h\mathbf{u}\otimes\mathbf{u}) + \nabla(\frac{g}{2}h^2) + gh\nabla Z = 0, \tag{3.2.2}$$

where we denote $h(t,x,y) \geq 0$, the water depth, $\mathbf{u}(t,x,y) = (u,v)^T$ the flow velocity, g the acceleration due to gravity intensity and $Z(x,y)$ the bottom depth, and therefore $h+Z$ is the water surface level (see Fig. 3.2.1). We denote also $\mathbf{q}(t,x,y) = (q_x,q_y)^T = h(t,x,y)\mathbf{u}(t,x,y)$ the flux of water.

To obtain a well-posed problem we add to this system some initial conditions

$$h(0,x,y) = h^0(x,y), \qquad \mathbf{u}(0,x,y) = \mathbf{u}^0(x,y), \tag{3.2.3}$$

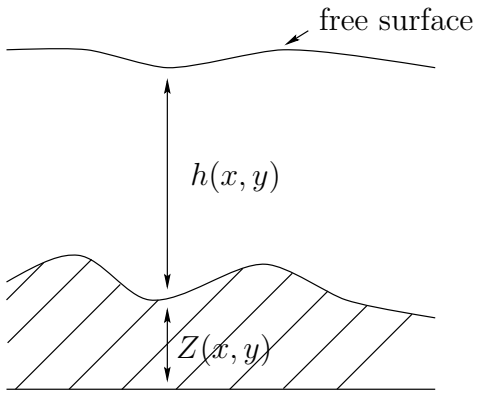


FIG. 3.2.1:

and boundary conditions. In this paper we consider only the following types of boundaries :

– Solid walls on which we prescribe a slip condition

$$\mathbf{u} \cdot \mathbf{n} = 0 \tag{3.2.4}$$

with $\mathbf{n}$ the unit outward normal to the boundary,

– Fluid boundaries on which we prescribe zero, one or two of the following conditions depending of the type of the flow (fluvial or torrential)
–Water level $h_g + Z$ given,
–Flux $\mathbf{q}_g$ given.

3.2.2 Properties of the system

The system (3.2.1)-(3.2.2) is a first order conservation laws system and can be written in the general form

$$\frac{\partial \mathbf{U}}{\partial t} + \text{div } \mathbf{F}(\mathbf{U}) = \mathbf{B}(\mathbf{U}), \tag{3.2.5}$$

with $\mathbf{U} = (h, q_x, q_y)^T$ and

$$\mathbf{F}(\mathbf{U}) = \begin{pmatrix} q_x & q_y \\ \frac{q_x^2}{h} + \frac{g}{2}h^2 & \frac{q_x q_y}{h} \\ \frac{q_x q_y}{h} & \frac{q_y^2}{h} + \frac{g}{2}h^2 \end{pmatrix}, \qquad \mathbf{B}(\mathbf{U}) = \begin{pmatrix} 0 \\ -gh\partial_x Z \\ -gh\partial_y Z \end{pmatrix}. \tag{3.2.6}$$

This system is *strictly hyperbolic* for $h > 0$ (see [23]). It admits an *invariant region* $h(t,x) \geq 0$, the water depth h can indeed vanish (flooding zones, dry regions, tidal

flats) and the system loses hyperbolicity at $h = 0$ which generates theoretical and numerical difficulties.

The system is also concerned with a fundamental *entropy property* which is given in the following theorem.

Theorem 3.2.1 *The system (3.2.1)–(3.2.2) admits a mathematical entropy (which is also the energy)*

$$E(h, u, Z) = h\frac{|\mathbf{u}|^2}{2} + \frac{gh^2}{2} + ghZ, \tag{3.2.7}$$

which satisfies

$$\frac{\partial}{\partial t}(E) + \mathrm{div}[\mathbf{u}\,(E + \frac{gh^2}{2})] \leq 0. \tag{3.2.8}$$

We do not prove this theorem which relies on the classical theory of hyperbolic equations and simple algebraic calculation, see Serre [118], Dafermos [37]. We just recall that for smooth solutions the inequality in (3.2.8) is an equality.

Another important property is that the Saint-Venant system admits *steady states*. They are characterized by

$$\begin{aligned} \mathrm{div}(h\mathbf{u}) &= 0, \\ \nabla P - \mathrm{curl}\ \mathbf{u} \begin{pmatrix} v \\ u \end{pmatrix} &= 0, \end{aligned}$$

where

$$P(x, y) = \frac{|\mathbf{u}|^2}{2} + g(h + Z).$$

It follows in particular that P is constant along streamlines and in the irrotationnal areas. For 2D problems, we are interested particularly by the so-called *lake at rest steady state*

$$\mathbf{u} = 0, \qquad h + Z = H, \tag{3.2.9}$$

where H is a constant. It will be important that the numerical scheme also satisfies this property.

For 1D flows, steady states with non vanishing velocity are also characterized by simple relations

$$hu = C_1, \qquad \frac{|u|^2}{2} + g(h + Z) = C_2, \tag{3.2.10}$$

where C_1 and C_2 are two constants. These relations (3.2.10) are used to compute the exact solution of the 1D flows in a 2D channel with a bump at the bottom, which are presented in Sec. 3.7.

3.3 Kinetic representation

We introduce a kinetic approach to system (3.2.1)–(3.2.2) and in the next section we deduce from the discretization of the corresponding kinetic equation, a kinetic scheme for this system.

Let $\chi(w)$ be a positive, even function defined on $\mathbb{R}^2$ i.e.

$$\chi(w) = \chi(-w) \geq 0 \tag{3.3.1}$$

and satisfying

$$\int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ w_i w_j \end{pmatrix} \chi(w)\ dw = \begin{pmatrix} 1 \\ \delta_{ij} \end{pmatrix} \tag{3.3.2}$$

with δ_{ij} the Kronecker symbol.
In addition we assume that $\chi(w)$ is compactly supported, i.e.

$$\exists w_M \in \mathbb{R},\ \text{such that} \quad \chi(w) = 0 \text{ for } \|w\| \geq w_M. \tag{3.3.3}$$

An example of function χ satisfying these properties is

$$\chi(w) = \frac{1}{12} \mathbb{1}_{|w_i| \leq \sqrt{3}},\ i = 1, 2. \tag{3.3.4}$$

We introduce a microscopic density of particles $M(t, x, y, \xi)$ defined by a so-called *Gibbs equilibrium*

$$M(t, x, y, \xi) = M(h, \xi - \mathbf{u}) = \frac{h(t, x, y)}{\tilde{c}^2}\ \chi(\frac{\xi - \mathbf{u}(t, x, y)}{\tilde{c}}), \tag{3.3.5}$$

with $\tilde{c}$ defined by

$$\tilde{c}^2 = \frac{g\, h}{2}. \tag{3.3.6}$$

With these definitions we can write a kinetic interpretation of the system (3.2.1)–(3.2.2) and we have the following statement :

Theorem 3.3.1 *The functions $(h, \mathbf{q})$ are strong solutions to the system (3.2.1)–(3.2.2) or (3.2.5)–(3.2.6) if and only if $M(t, x, y, \xi)$ satisfies the kinetic equation*

$$\frac{\partial M}{\partial t} + \xi \cdot \nabla_{\mathbf{x}} M - g \nabla Z \cdot \nabla_{\xi} M = Q(t, x, y, \xi), \tag{3.3.7}$$

for some "collision term" $Q(t, x, y, \xi)$ which satisfies for a.e. $(t,\ x, y)$,

$$\int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ \xi \end{pmatrix} Q\ d\xi = 0. \tag{3.3.8}$$

Proof. The proof relies on a very obvious computation. The two Saint-Venant equations are equivalent with the equation (3.3.7) once integrated in ξ against 1 and ξ. These are consequences of the usual relations deduced from the properties of χ and from (3.3.8) :

$$\begin{pmatrix} h \\ \mathbf{q} \\ \dfrac{\mathbf{q}\otimes\mathbf{q}}{h} + \dfrac{g}{2}\, h^2\mathbf{Id} \end{pmatrix} = \int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ \xi \\ \xi\otimes\xi \end{pmatrix} M(\xi)\, d\xi, \tag{3.3.9}$$

and

$$\begin{pmatrix} 0 \\ h \end{pmatrix} = -\int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ \xi \end{pmatrix} \nabla_\xi M(\xi)\, d\xi. \tag{3.3.10}$$

□

This theorem produces a very useful consequence : the non-linear system (3.2.1)–(3.2.2) can be viewed as a linear transport equation on a nonlinear quantity M, for which it is easier to find a simple numerical scheme with good theoretical properties.

The mathematical entropy property, for the energy, can also be considered in terms of the kinetic approach. Indeed if the "collision term" $Q(t,x,y,\xi)$ satisfies also for a.e. $(t,\ x, y)$,

$$\int_{\mathbb{R}^2} |\xi|^2 Q\ d\xi \leq 0,$$

the entropy inequality (3.2.8) is equivalent with the equation (3.3.7) once integrated in ξ against $|\xi|^2/2 + gZ$. It is a consequence of the following relation between the kinetic density M and the macroscopic energy E

$$E(t,x,y) = \int_{\mathbb{R}^2} \left(\frac{|\xi|^2}{2} + gZ\right)\ M(t,x,y,\xi)\ d\xi. \tag{3.3.11}$$

Remark 3.3.1 *The relation between M and E is more complex in the one-dimensional case since it requires to take into account the transverse translational energy through a cubic term [108].*

Among the functions that verify (3.3.1)-(3.3.3) one of them can be computed through a minimization problem. In particular it justifies the interpretation of such a density $M(t,x,y,\xi)$ as the microscopic equilibrium of the system.

Proposition 3.3.1 *The minimum of the energy*

$$\epsilon(f) = \int_{\mathbb{R}^2} \left(\frac{|\xi|^2}{2} + gZ\right) f(\xi) d\xi,$$

under the constraints

$$0 \leq f \leq \frac{1}{4\pi}, \qquad \int_{\mathbb{R}^2} f(\xi) d\xi = h, \qquad \int_{\mathbb{R}^2} \xi f(\xi) d\xi = h\mathbf{u},$$

is achieved by the function $M(t,x,y,\xi)$ defined by (3.3.5) with χ given by

$$\chi(w) = \frac{1}{4\pi}\mathbb{1}_{\|w\|\leq 2}. \tag{3.3.12}$$

Proof. The form of the constraints and of the functional that we minimize leads to search solutions of the form

$$M(t,x,y,\xi) = A(t,x,y)\mathbb{1}_{\|\xi-\mathbf{u}\|\leq B(t,x,y)}$$

and the results follows obviously. □

Remark 3.3.2 *The minimum $M(t,x,y,\xi)$ computed in Proposition 3.3.1 is of the form (3.3.1)-(3.3.3), (3.3.5) and then it verifies (3.3.11). It follows that $\epsilon(M(t,x,y,\xi)) = E(t,x,y)$ where $E(t,x,y)$ is the macroscopic energy of the Saint-Venant system defined in (3.2.7). This justifies the choice of the constants in the constraints in Proposition 3.3.1.*

Remark 3.3.3 *The function χ defined by (3.3.12) is also the only choice such that $M(t,x,y,\xi)$ defined by (3.3.5) preserves the microscopic lake at rest steady state, i.e. satisfies the kinetic equation (3.3.7) with $Q(t,x,y,\xi) = 0$ on all macroscopic steady states associated to a lake at rest defined by (3.2.9). For the one-dimensional Saint-Venant system, the authors of [108] use the equivalent 1d property to derive a well-balanced kinetic scheme by integrating the topographic source term in the kinetic fluxes.*

3.4 Finite volumes / Kinetic solver

A classical approach for solving hyperbolic systems consists in using finite volume schemes (see [54]) which are defined by the fluxes computed at the control volume interfaces. We show in Section 3.4.2 how the fluxes of the kinetic scheme are deduced from the discretization of the kinetic equation (3.3.7). In this section we do not take into account the topographic source term, i.e. we first consider the homogeneous system of equations (3.2.5) with $\mathbf{B} = 0$. In Section 3.5 we adapt the scheme to the non flat bottom case.

3.4.1 Finite volume formalism

We recall here the general formalism of finite volumes. Let Ω denote the computational domain with boundary Γ, which we assume polygonal. Let $\mathcal{T}_h$ be a triangulation of Ω which vertices are denoted P_i with $\mathbf{S_i}$ the set of interior nodes and $\mathbf{G_i}$ the set of boundary nodes. The dual cells C_i are obtained by joining the centers of mass of the triangles surrounding each vertex P_i. We use the following notations (see Fig. 3.4.1) :

- K_i, set of subscripts of nodes P_j surrounding P_i,
- $|C_i|$, area of C_i,
- Γ_{ij}, boundary edge between the cells C_i and C_j,
- L_{ij}, length of Γ_{ij},
- $\mathbf{n}_{ij}$, unit normal to Γ_{ij}, outward to C_i ($\mathbf{n}_{ji}$=-$\mathbf{n}_{ij}$).

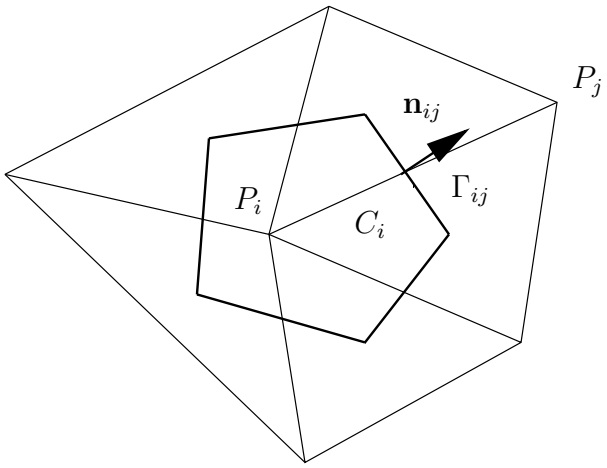


FIG. 3.4.1: Dual cell C_i.

If P_i is a node belonging to the boundary Γ, we join the centers of mass of the triangles adjacent to the boundary to the middle of the edge belonging to Γ (see Fig. 3.4.2) and we denote

- Γ_i, the two edges of C_i belonging to Γ,
- L_i, length of Γ_i (for sake of simplicity we assume in the following that $L_i = 0$ if P_i does not belong to Γ),
- $\mathbf{n}_i$, the unit outward normal defined by averaging the two adjacent normals.

Let Δt be the timestep, we set $t^n = n\ \Delta t$. We denote by $\mathbf{U}_i^n$ the approximation of the cell average of the exact solution at time t^n

$$\mathbf{U}_i^n \simeq \frac{1}{|C_i|}\int_{C_i} \mathbf{U}(t^n, \mathbf{x})\ d\mathbf{x}. \tag{3.4.1}$$

We integrate in space and time the equation (3.2.5) on the set $C_i \times (t^n, t^{n+1})$, and, integrating by parts the divergence term, we obtain

$$\int_{C_i} \mathbf{U}(t^{n+1}, \mathbf{x})\ d\mathbf{x} - \int_{C_i} \mathbf{U}(t^n, \mathbf{x})\ d\mathbf{x} + \int_{t^n}^{t^{n+1}} \int_{\partial C_i} \mathbf{F}(\mathbf{U}) \cdot \mathbf{n}\ d\mathbf{x}dt = 0. \tag{3.4.2}$$

So we can write

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \sum_{j \in K_i} \sigma_{ij} \mathcal{F}(\mathbf{U}_i^n, \mathbf{U}_j^n, \mathbf{n}_{ij}) - \sigma_i \mathcal{F}(\mathbf{U}_i^n, \mathbf{U}_{e,i}^n, \mathbf{n}_i) \tag{3.4.3}$$

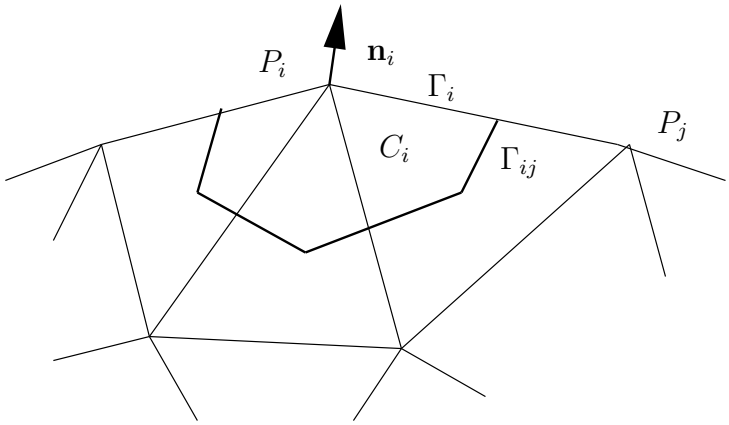


FIG. 3.4.2: Boundary cell C_i.

with

$$\sigma_{ij} = \frac{\Delta t\, L_{ij}}{|C_i|}, \qquad \sigma_i = \frac{\Delta t\, L_i}{|C_i|}, \tag{3.4.4}$$

In (3.4.3) the term $\mathcal{F}(\mathbf{U}_i, \mathbf{U}_j, \mathbf{n}_{ij})$ denotes an interpolation of the normal component of the flux $\mathbf{F}(\mathbf{U}) \cdot \mathbf{n}_{ij}$ along the edge Γ_{ij}. This interpolation is usually performed using a one-dimensional solver since locally the problem looks like a planar discontinuity. In the next section we define $\mathcal{F}(\mathbf{U}_i, \mathbf{U}_j, \mathbf{n}_{ij})$ using the kinetic interpretation of the system. The computation of the value $\mathbf{U}_{e,i}$ which denotes a value outside C_i defined such that the boundary conditions are satisfied and the definition of the boundary flux $\mathcal{F}(\mathbf{U}_i, \mathbf{U}_{e,i}, \mathbf{n}_i)$ are detailed in Section 3.4.5.

3.4.2 Kinetic solver

In the following of the present section we assume that P_i is an interior point. Being given the solution $\mathbf{U}_i^n$ at time t^n for each cell, we compute $\mathbf{U}_i^{n+1}$ by the following algorithm with three steps :

- We define $M_i^n = M(h_i^n, \xi - \mathbf{u}_i^n)$ with M defined by (3.3.5).
- We use the microscopic equation (3.3.7). Since this equation is linear, we can apply a simple upwind scheme [54] which defines a density function $f_i^{n+1}(\xi)$

$$f_i^{n+1}(\xi) - M_i^n(\xi) + \frac{\Delta t}{|C_i|} \sum_{j \in K_i} L_{ij}\ \xi \cdot \mathbf{n}_{ij}\ M_{ij}^n(\xi) = 0, \tag{3.4.5}$$

  with the fluxes $M_{ij}^n(\xi)$ computed by the upwind formula

$$M_{ij}^n(\xi) = \begin{cases} M_i^n(\xi) & \text{for } \xi \cdot \mathbf{n}_{ij} \geq 0, \\ M_j^n(\xi) & \text{for } \xi \cdot \mathbf{n}_{ij} \leq 0. \end{cases}$$

Notice however that the density function $f(\xi)$ is not an equilibrium (see remark (3.4.1)).

– Nevertheless, by analogy with the computations in the proof of Theorem (3.3.1), we can recover the macroscopic quantities $\mathbf{U}_i^{n+1}$ at time t^{n+1} by integration

$$\mathbf{U}_i^{n+1} = \int_{\mathbb{R}^2} \begin{pmatrix} 1 \\ \xi \end{pmatrix} f_i^{n+1}(\xi)\, d\xi. \tag{3.4.6}$$

Remark 3.4.1 *The interpretation is that, as usual, the collision term Q, which forces the relaxation of f to Gibbs equilibrium M, is neglected in the advection scheme (3.4.5). And at each timestep we deduce $M_i^{n+1}(\xi)$ from $\mathbf{U}_i^{n+1}$ which is a way to perform all collisions at once and to recover the Gibbs equilibrium without computing them explicitely.*

The numerical consistency of the kinetic solver relies on the fact that if we consider the exact solution of the homogeneous kinetic transport equation - (3.3.7) with $\nabla Z = 0$ and $Q = 0$ - we can prove that the macroscopic quantities obtained through the integration process describes previously are first order approximations in time of the solutions of the Saint-Venant system - see [54].

The numerical feasability of the kinetic solver relies on the possibility to write directly a finite volume formula, which therefore avoids using the extra variable ξ in the actual implementation. Indeed, the equation (3.4.6) can be written with the form (3.4.3) with

$$\mathcal{F}(\mathbf{U}_i, \mathbf{U}_j, \mathbf{n}_{ij}) = \mathbf{F}^+(\mathbf{U}_i, \mathbf{n}_{ij}) + \mathbf{F}^-(\mathbf{U}_j, \mathbf{n}_{ij}), \tag{3.4.7}$$

and

$$\mathbf{F}^+(\mathbf{U}_i, \mathbf{n}_{ij}) = \int_{\xi\cdot\mathbf{n}_{ij}\geq 0} \xi\cdot\mathbf{n}_{ij} \begin{pmatrix} 1 \\ \xi \end{pmatrix} M_i(\xi)\, d\xi, \tag{3.4.8}$$

$$\mathbf{F}^-(\mathbf{U}_j, \mathbf{n}_{ij}) = \int_{\xi\cdot\mathbf{n}_{ij}\leq 0} \xi\cdot\mathbf{n}_{ij} \begin{pmatrix} 1 \\ \xi \end{pmatrix} M_j(\xi)\, d\xi. \tag{3.4.9}$$

Notice that (3.4.8)-(3.4.9) imply

$$\mathbf{F}_h^+(\mathbf{U}_i, \mathbf{n}_{ij}) \geq 0, \qquad \mathbf{F}_h^-(\mathbf{U}_j, \mathbf{n}_{ij}) \leq 0. \tag{3.4.10}$$

3.4.3 Numerical implementation

We give here some details on the implementation of the kinetic scheme defined by (3.4.3), (3.4.7)-(3.4.9). For the efficiency of the method, we code in fact a variant where the choice of the function χ depends on the interface under consideration. For an interface with unit normal $\mathbf{n} = (n_x, n_y)^T$, we define a local basis (n, τ) associated to the normal direction and to the tangential one. We denote $\hat{\mathbf{U}}_{\mathbf{n}} = (h, q_n, q_\tau)^T$, the vector

deduced from $\mathbf{U}$ by the rotation in this new basis and $\hat{\mathbf{u}} = (u_n, u_\tau)^T = (\frac{q_n}{h}, \frac{q_\tau}{h})^T$. So we have $\hat{\mathbf{U}}_\mathbf{n}$ defined by

$$\hat{\mathbf{U}}_\mathbf{n} = \mathbf{R}_\mathbf{n}\ \mathbf{U} \quad \text{with} \quad \mathbf{R}_\mathbf{n} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & n_x & n_y \\ 0 & -n_y & n_x \end{pmatrix} \tag{3.4.11}$$

and

$$\mathbf{F}^+(\mathbf{U}, \mathbf{n}) = \mathbf{R}_\mathbf{n}^{-1} \hat{\mathbf{F}}^+(\hat{\mathbf{U}}_\mathbf{n}). \tag{3.4.12}$$

Using (3.4.8), we give the detailed expression of $\hat{\mathbf{F}}^+(\hat{\mathbf{U}}_i)$ related to the interface Γ_{ij}

$$\hat{\mathbf{F}}^+(\hat{\mathbf{U}}_{i,\mathbf{n}_{ij}}) = \frac{h_i}{\tilde{c}_i^2} \int_{\{\xi_n \geq 0\} \times \mathbb{R}} \xi_n \begin{pmatrix} 1 \\ \xi \end{pmatrix} \chi\left(\frac{\xi - \hat{\mathbf{u}}_i}{\tilde{c}_i}\right) d\xi \tag{3.4.13}$$

or, after the change of variables $w = \dfrac{\xi - \hat{\mathbf{u}}_i}{\tilde{c}_i}$,

$$\hat{\mathbf{F}}^+(\hat{\mathbf{U}}_{i,\mathbf{n}_{ij}}) = h_i \int_{\{w_n \geq \frac{-u_{i,n}}{\tilde{c}_i}\} \times \mathbb{R}} (u_{i,n} + w_n\, \tilde{c}_i) \begin{pmatrix} 1 \\ u_{i,n} + w_n\, \tilde{c}_i \\ u_{i,\tau} + w_\tau\, \tilde{c}_i \end{pmatrix} \chi(w)\, dw, \tag{3.4.14}$$

we have dropped here the subscript ij for the components u_n, u_τ.

Due to the fact that $\chi(w)$ is even, the term with w_τ disappears in (3.4.14) and we obtain the simpler formula

$$\hat{F}^+_{u_\tau}(\hat{\mathbf{U}}_{i,\mathbf{n}_{ij}}) = \hat{u}_{i,\tau}\ \hat{F}^+_h(\hat{\mathbf{U}}_{i,\mathbf{n}_{ij}}). \tag{3.4.15}$$

We obtain $\mathbf{F}^-$ by an analogous computation and so we have the same property for $\hat{\mathcal{F}}$. We will use this property to deduce a modified scheme with better accuracy.

3.4.4 Upwind kinetic scheme

In order to reduce the diffusion of the scheme, we modify the computation of the flux related to the tangential component. For the computation of $u_{i,\tau}^{n+1}$ we replace the expression of $\hat{\mathcal{F}}$ by the following :

$$\hat{\mathcal{F}}_{u_\tau}(\hat{\mathbf{U}}_{i,\mathbf{n}_{ij}}, \hat{\mathbf{U}}_{j,\mathbf{n}_{ij}}) = u_{ij,\tau}\ \hat{\mathcal{F}}_h(\hat{\mathbf{U}}_{i,\mathbf{n}_{ij}}, \hat{\mathbf{U}}_{j,\mathbf{n}_{ij}}) \tag{3.4.16}$$

with

$$u_{ij,\tau} = \begin{cases} u_{i,\tau} & \text{for } \hat{\mathcal{F}}_h \geq 0, \\ u_{j,\tau} & \text{for } \hat{\mathcal{F}}_h \leq 0. \end{cases} \tag{3.4.17}$$

Formula (3.4.17) introduces some upwinding depending on the sign of the total flux. We give in [24] a numerical result showing the efficiency of (3.4.16)-(3.4.17).

3.4.5 Boundary conditions

The treatment of the boundary conditions is presented in details in [23]. Here we just recall some main features about the computation of the boundary flux $\mathcal{F}(\mathbf{U}_i^n, \mathbf{U}_{e,i}^n, \mathbf{n}_i)$ appearing in (3.4.3). Notice first that the variable $\mathbf{U}_{e,i}^n$ can be interpreted as an approximation of the solution in a "fictitious" cell adjacent to the boundary. As before we introduce the local coordinates and define $\hat{\mathbf{U}}_{e,i}^n = \left(h_{e,i}^n, q_{e,i,n}^n, q_{e,i,\tau}^n\right)^T$. Then we can use the local flux vector splitting form associated to the kinetic formulation

$$\hat{\mathcal{F}}(\hat{\mathbf{U}}_{i,\mathbf{n}_i}^n, \hat{\mathbf{U}}_{e,i}^n) = \hat{F}^+(\hat{\mathbf{U}}_{i,\mathbf{n}_i}^n) + \hat{F}^-(\hat{\mathbf{U}}_{e,i}^n).$$

On the solid wall we prescribe a continuous slip condition - see Section 3.2. In the numerical scheme we prescribe it weakly by defining $\hat{\mathbf{U}}_{e,i}^n = \left(h_i^n, -q_{i,n}^n, q_{i,\tau}^n\right)^T$. It follows that finally

$$\hat{\mathcal{F}}(\hat{\mathbf{U}}_{i,\mathbf{n}_i}^n, \hat{\mathbf{U}}_{e,i}^n) = (0,\ \frac{g h_i^{n2}}{2},\ 0)^T. \tag{3.4.18}$$

On the fluid boundaries, the type of the flow and then the number of boundary conditions depend on the Froude number. Here we consider a local Froude number associated to the normal component of the velocity. For the fluvial cases, we define completely $\mathbf{U}_e$ by adding to the given boundary condition, the assumption that the *Riemann invariant* related to the outgoing characteristic is constant along this characteristic (see [23]).

3.4.6 Properties of the scheme

It is clear from (3.4.3), (3.4.7)-(3.4.9) that the scheme is consistent and conservative. The CFL condition for the explicit scheme (3.4.5) applied to the linear microscopic equation writes

$$\Delta t \leq \min \frac{|C_i|}{(L_i + \sum_{j\in K_i} L_{ij})\ (|\mathbf{u}_i^n| + w_M \tilde{c}_i^n)}. \tag{3.4.19}$$

Besides we have the following stability theorem :

Theorem 3.4.1 *Under the CFL condition (3.4.19), the kinetic scheme (3.4.3), (3.4.7), (3.4.12), (3.4.14) preserves the water depth positivity.*

Proof. Suppose that we have $h_i^n \geq 0$. It follows from the definitions (3.4.3), (3.4.7), (3.4.12), (3.4.14) that - here and all along the paper, when superscripts are omitted, the quantities have to be taken at time t^n

$$h_i^{n+1} = h_i^n - \sum_{j\in K_i} \sigma_{ij}(\hat{F}_h^+(\hat{\mathbf{U}}_{i,\mathbf{n}_{ij}}^n) + \hat{F}_h^-(\hat{\mathbf{U}}_{j,\mathbf{n}_{ij}}^n)) - \ \sigma_i(\hat{F}_h^+(\hat{\mathbf{U}}_{i,\mathbf{n}_i}^n) + \hat{F}_h^-(\hat{\mathbf{U}}_{e,i}^n)). \tag{3.4.20}$$

The relations (3.4.10) and (3.4.12) imply

$$\hat{F}_h^-(\hat{\mathbf{U}}_{j,\mathbf{n}_{ij}}^n) \leq 0, \qquad \hat{F}_h^-(\hat{\mathbf{U}}_{e,i}^n) \leq 0, \tag{3.4.21}$$

and, using the detailed expression of the flux (3.4.14),we have

$$h_i^{n+1} \geq h_i^n \quad \left(1 - \sum_{j\in K_i} \sigma_{ij} \int_{\{w_n \geq \frac{-u_{i,nij}}{\tilde{c}_i}\}\times\mathbb{R}} (u_{i,nij} + w_n\,\tilde{c}_i)\,\chi(w)\,dw \right.$$
$$\left. - \sigma_i \int_{\{w_n \geq \frac{-u_{i,ni}}{\tilde{c}_i}\}\times\mathbb{R}} (u_{i,ni} + w_n\,\tilde{c}_i)\,\chi(w)\,dw \right). \qquad (3.4.22)$$

Since χ satisfies (3.3.1)-(3.3.2), we have for each n

$$\int_{\{w_n \geq \frac{-u_{i,n}}{\tilde{c}_i}\}\times\mathbb{R}} (u_{i,n} + w_n\,\tilde{c}_i)\,\chi(w)\,dw \leq |u_{i,n}| + \tilde{c}_i \int_{\{w_n\geq 0\}\times\mathbb{R}} w_n\chi(w)\,dw, \qquad (3.4.23)$$

and from (3.3.1)-(3.3.3) we deduce

$$\int_{\{w_n\geq 0\}\times\mathbb{R}} w_n\chi(w)\,dw \leq \int_{\{0\leq w_n\leq 1\}\times\mathbb{R}} \chi(w)\,dw + \int_{\{w_n\geq 1\}\times\mathbb{R}} w_n^2\chi(w)\,dw \leq 1. \qquad (3.4.24)$$

Finally using (3.4.23)-(3.4.24) in (3.4.22) we obtain

$$h_i^{n+1} \geq h_i^n \left(1 - \frac{\Delta t}{|C_i|}[\sum_{j\in K_i} L_{ij}\,(|u_{i,nij}| + \tilde{c}_i) + L_i\,(|u_{i,ni}| + \tilde{c}_i)]\right),$$

and it follows the positivity of h_i^{n+1} under the CFL condition (3.4.19) (from (3.3.2) we have $w_M \geq 1$). □

In the particular case where the function χ is defined through the minimization problem of Proposition 3.3.1, the deduced kinetic scheme satisfies another stability property as shown in the following theorem.

Theorem 3.4.2 *Under the CFL condition (3.4.19) and with the particular choice where χ is defined by (3.3.12) the kinetic scheme (3.4.3), (3.4.7), (3.4.12), (3.4.14) satisfies a conservative in-cell entropy inequality.*

Proof. In this case χ is invariant by rotation and then is the same one whatever is the interface under consideration. It follows that we can establish the proof at the kinetic level. Integrating the microscopic scheme (3.4.5) in ξ against $\xi^2/2$ we obtain - with the notations of Proposition 3.3.1

$$\epsilon\left(f_i^{n+1}\right) - E_i^n + \frac{\Delta t}{|C_i|}\sum_{j\in K_i} L_{ij} \int_{\mathbb{R}^2} \frac{|\xi|^2}{2}\,\xi\cdot\mathbf{n}_{ij}\; M_{ij}^n(\xi)d\xi = 0.$$

It follows from Proposition 3.3.1 - with $Z = 0$ - that the particular choice of χ implies

$$E_i^{n+1} - E_i^n + \sum_{j\in K_i} \sigma_{ij}\;\eta(\mathbf{U}_i^n, \mathbf{U}_j^n, \mathbf{n}_{ij}) \leq 0,$$

where $\eta(\mathbf{U}_i^n, \mathbf{U}_j^n, \mathbf{n}_{ij})$ denotes the entropy flux defined by

$$\eta(\mathbf{U}_i^n, \mathbf{U}_j^n, \mathbf{n}_{ij}) = \int_{\mathbb{R}^2} \frac{|\xi|^2}{2}\, \xi \cdot \mathbf{n}_{ij}\; M_{ij}^n(\xi) d\xi.$$

This concludes the proof. □

3.5 Well-balanced scheme

In order to be able to compute realistic flows we consider now the case $\nabla Z \neq 0$. As motivated in the introduction the *well-balanced* requirement is to preserve a discrete equivalent of the continuous lake at rest steady state (3.2.9). Then we modify the original scheme in order to satisfy the following discrete local lake at rest steady state

$$\left(\begin{array}{ll} \forall j \in K_i & \begin{array}{l} h_j^n + Z_j = h_i^n + Z_i = H \\ \mathbf{u}_j^n = \mathbf{u}_i^n = 0 \end{array} \end{array}\right) \quad \Rightarrow \quad \begin{array}{l} h_i^{n+1} + Z_i = H \\ \mathbf{u}_i^{n+1} = 0 \end{array} . \tag{3.5.1}$$

The well-balanced scheme that we present in this section is an adaptation to the two-dimensional flows of the *interface hydrostatic reconstruction* method developed by Audusse et al. in [7] in the 1d framework.
We first construct a piecewise constant approximation of the bottom topography $Z(x)$

$$Z_i = \frac{1}{|C_i|}\int_{C_i} Z(x)dx.$$

We define an interface topography

$$Z_{ij}^* = Z_{ji}^* = max(Z_i, Z_j), \tag{3.5.2}$$

and then we define new interface values by $\mathbf{U}_{ij}^* = (h_{ij}^*, h_{ij}^* \mathbf{u}_i)^T$ where h_{ij}^* is the *hydrostatic reconstructed* water depth

$$h_{ij}^* = (h_i + Z_i - Z_{ij}^*)_+. \tag{3.5.3}$$

We write the well-balanced scheme with the form

$$\begin{aligned} \mathbf{U}_i^{n+1} = \mathbf{U}_i^n \;\; &- \sum_{j \in K_i} \sigma_{ij} \mathcal{F}(\mathbf{U}_{ij}^{*,n}, \mathbf{U}_{ji}^{*,n}, \mathbf{n}_{ij}) - \sigma_i \mathcal{F}(\mathbf{U}_i^n, \mathbf{U}_{e,i}^n, \mathbf{n}_i) \\ &+ \sum_{j \in K_i} \sigma_{ij} S(\mathbf{U}_i^n, \mathbf{U}_{ij}^{*,n}, \mathbf{n}_{ij}), \end{aligned} \tag{3.5.4}$$

where

$$S(\mathbf{U}_i, \mathbf{U}_{ij}^*, \mathbf{n}_{ij}) = \begin{pmatrix} 0 \\ \frac{g}{2}(h_{ij}^{*2} - h_i^2)\mathbf{n}_{ij} \end{pmatrix}. \tag{3.5.5}$$

Theorem 3.5.1 *The scheme defined by (3.5.4)–(3.5.5) with (3.4.7)– (3.4.9) satisfies the following properties*

(i) it preserves the water depth positivity under the CFL condition

$$\Delta t \leq \min \frac{|C_i|}{\sum_{j\in K_i}\left[L_{ij}\ (|\mathbf{u}_i^n| + w_M \tilde{c}_{ij}^{*,n})\right] + L_i\ (|\mathbf{u}_i^n| + w_M \tilde{c}_i^n)}, \tag{3.5.6}$$

a fortiori if Δt satisfies (3.4.19),

(ii) it preserves the steady state of a lake at rest.

Proof. For (i), first we prove as in Theorem 3.4.1 that the water depth positivity is preserved under the CFL condition (3.5.6). Then as the definition (3.5.3) implies that $h_{ij}^* \leq h_i$, we deduce that the CFL condition (3.4.19) is more restrictive than (3.5.6).

To prove (ii), we assume that the solution at time n satisfies (3.5.1), then we have

$$\sum_{j\in K_i} \sigma_{ij} \mathcal{F}(\mathbf{U}_{ij}^{*,n}, \mathbf{U}_{ji}^{*,n}, \mathbf{n}_{ij}) = \sum_{j\in K_i} \sigma_{ij} \begin{pmatrix} 0 \\ \frac{g}{2}\ h_{ij}^{*2}\mathbf{n}_{ij} \end{pmatrix}. \tag{3.5.7}$$

Concerning the boundary term, we assume also that the boundary conditions will preserve the steady state, so they can be either a slip condition, a given flux $\mathbf{q} = 0$ or a water depth given $h + Z = H$. Following the treatment of the boundary conditions developed in [23] the boundary term reduces to

$$\sigma_i \mathcal{F}(\mathbf{U}_i^n, \mathbf{U}_{e,i}^n, \mathbf{n}_i) = \begin{pmatrix} 0 \\ \frac{g}{2}\ h_i^2\ \mathbf{n}_i \end{pmatrix}. \tag{3.5.8}$$

From (3.5.4)–(3.5.5), we obtain finally

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \sum_{j\in K_i} \sigma_{ij} \begin{pmatrix} 0 \\ \frac{g}{2} h_i^2\ \mathbf{n}_{ij} \end{pmatrix} - \sigma_i \begin{pmatrix} 0 \\ \frac{g}{2} h_i^2\ \mathbf{n}_i \end{pmatrix} \tag{3.5.9}$$

and using the property

$$\sum_{j\in K_i} L_{ij}\mathbf{n}_{ij} + L_i\mathbf{n}_i = 0,$$

this proves the preservation of the steady state. □

Remark 3.5.1 *From the definitions (3.5.2)–(3.5.5), it is obvious that for $Z = Cst$, we recover the original scheme.*

Remark 3.5.2 *It appears in the proof of Theorem 3.5.1 that to construct the scheme on the interface values $\mathbf{U}_{ij}^*$ instead of the cell values $\mathbf{U}_i$ allows to numerically preserve at each interface the balance between the hydrostatic pressure and the influence of the topographic source terms in a lake at rest. This explains the name interface hydrostatic reconstruction method. For further details refer to [7].*

Remark 3.5.3 *By constrast with the homogeneous case we can not prove that the fully discrete kinetic scheme (3.5.4)–(3.5.5) satisfies an entropy inequality, whatever is the time step. Nevertheless it is proved in [7] that the semi-discrete version of the scheme (3.5.4)–(3.5.5) satisfies a conservative in-cell entropy inequality.*

3.6 Second order extension

The first order scheme defined in Sections 3.4-3.5 can be extended to a "formally" second order one using a MUSCL like extension (see [127]). In Section 3.6.1, we define limited reconstructed variables and in Section 3.6.2, we introduce a "second order" well-balanced scheme.

3.6.1 Second order reconstructions

In the definition of the flux (3.4.7), we replace the piecewise constant values $\mathbf{U}_i, \mathbf{U}_j$ by more accurate reconstructions deduced from piecewise linear approximations, namely the values $\mathbf{U}_{ij}, \mathbf{U}_{ji}$ reconstructed on both sides of the interface. More precisely, we are looking for piecewise linear approximation of the primitive variable $\hat{\mathbf{W}} = (h, u_n, u_\tau)^T$, actually the detailed expression of the flux given in (3.4.14) uses the primitive variables.

We divide each cell C_i in subtriangles obtained by joining each edge Γ_{ij} to the node P_i,

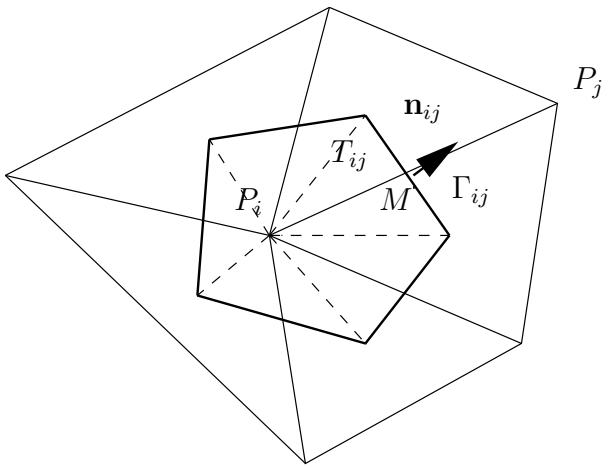


FIG. 3.6.1: Subcells T_{ij}.

we denote T_{ij} the subtriangle related to Γ_{ij} (see Fig.(3.6.1)). We denote $|C_{ij}|$ the area of T_{ij}. Let M be the middle point of the interface Γ_{ij}, we define $\hat{\mathbf{W}}_{ij} = (h_{ij}, u_{ij,n}, u_{ij,\tau})^T$ as an approximation of $\hat{\mathbf{W}}$ at point M, deduced from a piecewise linear reconstruction

on the subtriangle T_{ij} :

$$\hat{\mathbf{W}}_{ij} = \hat{\mathbf{W}}_i + \overrightarrow{P_iM} \cdot \nabla \hat{\mathbf{W}}_{ij} \tag{3.6.1}$$

with $\nabla \hat{\mathbf{W}}_{ij}$ defined here as follows (see [64]).

If the point M belongs to the triangle T_k, we denote $\nabla \hat{\mathbf{W}}_M = \nabla \hat{\mathbf{W}}|_{T_k}$ where $\nabla \hat{\mathbf{W}}|_{T_k}$ is the constant gradient of $\hat{\mathbf{W}}$ deduced from a P1 approximation on the triangle T_k. We denote by $\nabla \hat{\mathbf{W}}_i$ an approximate gradient at node P_i computed by a weighted average of the gradients on the surrounding triangles

$$\nabla \hat{\mathbf{W}}_i = \frac{\sum_{k \in \mathcal{T}_i} |C_k| \ \nabla \hat{\mathbf{W}}|_{T_k}}{\sum_{k \in \mathcal{T}_i} |C_k|} \tag{3.6.2}$$

and

$$\nabla \hat{\mathbf{W}}_{mi} = (1+\beta) \ \nabla \hat{\mathbf{W}}_i - \beta \ \nabla \hat{\mathbf{W}}_M, \quad 0 \leq \beta \leq 1, \tag{3.6.3}$$

where $\mathcal{T}_i$ is the set of triangles surrounding the node P_i.

Then we use an appropriate slope limiter to deduce $\nabla \hat{\mathbf{W}}_{ij}$

$$\nabla \hat{\mathbf{W}}_{ij} = \mathrm{Lim}(\nabla \hat{\mathbf{W}}_M, \nabla \hat{\mathbf{W}}_{mi}). \tag{3.6.4}$$

In the following computations we have used either the minmod limiter defined by

$$\mathrm{Lim}(a,b) = \begin{cases} 0 & \text{if } \mathrm{sign}(a) \neq \mathrm{sign}(b) \\ \mathrm{sign}(a)\min(|a|,|b|) & \text{otherwise} \end{cases}$$

or the Van Albada limiter defined by

$$\mathrm{Lim}(a,b) = \begin{cases} 0 & \text{if } \mathrm{sign}(a) \neq \mathrm{sign}(b) \\ \dfrac{a(b^2+\varepsilon) + b(a^2+\varepsilon)}{a^2+b^2+2\varepsilon} & \text{otherwise} \end{cases}$$

with $0 \leq \varepsilon \ll 1$.

Once the water depth is computed by (3.6.4), it is corrected (see [107]) to ensure the conservation of the reconstructed values, i.e.

$$\sum_{j \in K_i} |C_{ij}| h_{ij} = |C_i| h_i. \tag{3.6.5}$$

In the case where $\mathbf{B} = 0$, the "second order" scheme is obtained by replacing (3.4.3) by

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \sum_{j \in K_i} \sigma_{ij} \mathcal{F}(\mathbf{U}_{ij}^n, \mathbf{U}_{ji}^n, \mathbf{n}_{ij}) - \sigma_i \mathcal{F}(\mathbf{U}_i^n, \mathbf{U}_{e,i}^n, \mathbf{n}_i). \tag{3.6.6}$$

3.6.2 Second order well-balanced scheme

For the cases where $\mathbf{B} \neq 0$ we consider also piecewise linear approximation of the variable Z and we reconstruct values Z_{ij}, Z_{ji} on both sides of the interface as it is done before for the primitive variables. In fact, so that the second order scheme preserves the *well-balanced* property, it is necessary that the second order reconstruction preserves an interface equilibrium. It means that if

$$h_i + Z_i = h_j + Z_j = H, \qquad \mathbf{u}_i = \mathbf{u}_j = 0, \tag{3.6.7}$$

then the second order reconstructed values have to satisfy

$$h_{ij} + Z_{ij} = h_{ji} + Z_{ji} = H, \qquad \mathbf{u}_{ij} = \mathbf{u}_{ji} = 0. \tag{3.6.8}$$

The velocity part is obvious but - as we require also that the second order reconstruction preserves the positivity of the water depth - it is enounced in [7] that the right way to build a *well-balanced* second order scheme is to reconstruct and correct the variables $h+Z$ and h and then to deduce the interface values for Z - see [7] for further explanations, especially for the case of dry/wet interface.
Then we adapt the first order strategy departing from the reconstructed interface values instead of the cell values. We define

$$Z^*_{ij} = Z^*_{ji} = \max(Z_{ij}, Z_{ji}) \tag{3.6.9}$$

then we define new interface values by $\mathbf{U}^*_{ij} = (h^*_{ij},\ h^*_{ij}\mathbf{u}_{ij})^T$ with

$$h^*_{ij} = (h_{ij} + Z_{ij} - Z^*_{ij})_+. \tag{3.6.10}$$

We write the second order well-balanced scheme with the form

$$\begin{aligned} \mathbf{U}_i^{n+1} = \mathbf{U}_i^n \quad &- \quad \sum_{j\in K_i} \sigma_{ij}\mathcal{F}(\mathbf{U}_{ij}^{*,n}, \mathbf{U}_{ji}^{*,n}, \mathbf{n}_{ij}) - \sigma_i \mathcal{F}(\mathbf{U}_i^n, \mathbf{U}_{e,i}^n, \mathbf{n}_i) \\ &+ \quad \sum_{j\in K_i} \sigma_{ij}[S(\mathbf{U}_{ij}^n, \mathbf{U}_{ij}^{*,n}, \mathbf{n}_{ij}) + S^c(\mathbf{U}_i^n, \mathbf{U}_{ij}^n, Z_i, Z_{ij}, \mathbf{n}_{ij})] \end{aligned} \tag{3.6.11}$$

where

$$S(\mathbf{U}_{ij}, \mathbf{U}^*_{ij}, \mathbf{n}_{ij}) = \begin{pmatrix} 0 \\ \dfrac{g}{2}(h^{*2}_{ij} - h^2_{ij})\mathbf{n}_{ij} \end{pmatrix} \tag{3.6.12}$$

and

$$S^c(\mathbf{U}_i^n, \mathbf{U}_{ij}^n, Z_i, Z_{ij}, \mathbf{n}_{ij}) = \begin{pmatrix} 0 \\ -\dfrac{g}{2}(h_{ij} + h_i)(Z_{ij} - Z_i)\mathbf{n}_{ij} \end{pmatrix}. \tag{3.6.13}$$

Remark 3.6.1 *By contrast to the first order scheme, we have to introduce here a centered term (3.6.13) to satisfy the consistency with the source terms.*

Theorem 3.6.1 *The formally second order scheme defined by (3.6.11)–(3.6.13) with (3.4.7)- (3.4.9) satisfies the following properties*

(i) it preserves the water depth positivity under the CFL condition

$$\Delta t \leq \min\left[\min_{i\in\mathbf{S_i}}\min_{j\in K_i}\frac{|C_{ij}|}{L_{ij}\ (|\mathbf{u}_{ij}^n|+w_M\tilde{c}_{ij}^{*,n})},\right.$$
$$\left.\min_{i\in\mathbf{G_i}}\max_{0\leq\alpha\leq 1}\left(\alpha\frac{|C_i|}{L_i\ (|\mathbf{u}_i^n|+w_M\tilde{c}_i^n)},\ (1-\alpha)\min_{j\in K_i}\frac{|C_{ij}|}{L_{ij}\ (|\mathbf{u}_{ij}^n|+w_M\tilde{c}_{ij}^{*,n})}\right)\right], \tag{3.6.14}$$

(ii) it preserves the steady state of a lake at rest.

Proof. We follow the idea developed in [18] for the 1D problem. First we assume that P_i is an interior node. Using (3.4.7) and the definition (3.4.4) of σ_{ij}, the scheme (3.6.11) defines h_i^{n+1} by

$$h_i^{n+1}=\frac{1}{|C_i|}[|C_i|h_i^n-\frac{1}{\Delta t}\sum_{j\in K_i}L_{ij}(\mathbf{F}_h^+(\mathbf{U}_{ij}^{*,n},\mathbf{n}_{ij})+\mathbf{F}_h^-(\mathbf{U}_{ji}^{*,n},\mathbf{n}_{ij}))]. \tag{3.6.15}$$

Using (3.6.5), we have

$$h_i^{n+1}=\frac{1}{|C_i|}\sum_{j\in K_i}[\ |C_{ij}|h_{ij}^n-\frac{L_{ij}}{\Delta t}(\mathbf{F}_h^+(\mathbf{U}_{ij}^{*,n},\mathbf{n}_{ij})+\mathbf{F}_h^-(\mathbf{U}_{ji}^{*,n},\mathbf{n}_{ij}))]. \tag{3.6.16}$$

To verify the positivity of h_i^{n+1}, it is sufficient to have

$$|C_{ij}|h_{ij}^n-\frac{L_{ij}}{\Delta t}\mathbf{F}_h^+(\mathbf{U}_{ij}^{*,n},\mathbf{n}_{ij})\geq 0\quad \text{for j}\in \mathrm{K_i}, \tag{3.6.17}$$

now using the relation $h_{ij}^*\leq h_{ij}$ and following the proof of theorem (3.4.1), we obtain that the inequality (3.6.17) is satisfied under the condition (3.6.14).
If P_i is a boundary node, we have

$$\begin{aligned}h_i^{n+1}=\frac{1}{|C_i|}[|C_i|h_i^n &- \frac{1}{\Delta t}\sum_{j\in K_i}L_{ij}(\mathbf{F}_h^+(\mathbf{U}_{ij}^{*,n},\mathbf{n}_{ij})+\mathbf{F}_h^-(\mathbf{U}_{ji}^{*,n},\mathbf{n}_{ij}))\\ &- \frac{L_i}{\Delta t}(\mathbf{F}_h^+(\mathbf{U}_i^n,\mathbf{n}_i)+\mathbf{F}_h^-(\mathbf{U}_{e,i}^n,\mathbf{n}_i))].\end{aligned} \tag{3.6.18}$$

Then using (3.6.5) we can write

$$|C_i|h_i=\alpha|C_i|h_i+(1-\alpha)\sum_{j\in K_i}|C_{ij}|h_{ij}),\quad ,0\leq\alpha\leq 1, \tag{3.6.19}$$

and with the same arguments as previously, we obtain the positivity of the waterdepth under the condition (3.6.14). □

Remark 3.6.2 *To obtain the second order accuracy it is necessary to consider more sophisticated reconstruction technics as the second order ENO reconstruction (see [1] for more details about the 2d ENO method and [76] for error estimates in the 1d scalar case with source terms). They are not used here because they are very tricky to apply on 2d unstructured meshes. Nevertheless the numerical results presented in the next section show an obvious accuracy improvement obtained with the formally second order scheme developed in this paper. Second order accuracy in time can be obtained as usual by a Runge-Kutta method and the CFL condition need not be modified.*

3.7 Numerical results

We present here the numerical results of different test problems. We begin with the two-dimensional version of a classical academic test problem extracted from [60] and commonly used (see, e.g. [33, 49]) : a stationary flow over a parabolic bump for which an exact solution can be computed. We condider a rectangular channel with length 20. and width 2. (we assume an adimensionalized problem), the bottom is defined by

$$Z(x,y) = \begin{cases} 0.2 - 0.05(x-10.)^2 & \text{if } 8. \leq x \leq 12., \quad \forall y, \\ 0. \quad \text{else.} \end{cases}$$

Depending on the values of the boundary conditions, we compute three different flows defined as follows :

- fluvial flow
  inflow : $\mathbf{q}_g = (4.42, 0)^T$, outflow $h_g = 2.$,
- transcritical without shock (torrential outflow)
  inflow : $\mathbf{q}_g = (1.53, 0)^T$, initial water depth $h^0 = 0.66$,
- transcritical with shock
  inflow : $\mathbf{q}_g = (0.18, 0)^T$, outflow $h_g = 0.33$

The given discharge is prescribed for each node of the inflow boundary. The initial solution is given by $\mathbf{q}^0 = \mathbf{q}_g, h^0 = h_g$. The three flows are computed on a rather coarse unstructured mesh of 510 nodes and 886 triangles (60 edges on the length and 6 edges on the width). In Figures 3.7.1-(a), 3.7.1-(c) and 3.7.1-(e), the free surface profiles computed with the first order and second order schemes are compared to the exact solution. We have plotted only the points on the line $y = 0$ but we claim that the two-dimensional effects are negligible. Results are quite good for such a coarse grid and the improvement due to the second order extension appears to be noticeable for all the cases, even the ones involving a discontinuity. Note also that the presence of a sonic point in the two last test cases does not need a special treatment.

Even if we have no error estimations for our problem, it appears interesting to look at the convergence rate of the error versus the space discretization for the three above problems. We have plotted in Figures 3.7.1-(b), 3.7.1-(d) and 3.7.1-(f), the Log (L^1-error) of the water depth versus Log (h_{a_0}/h_a) for the first and the second order scheme

and they are compared to the theoretical order (we denote by h_a the average of edge length and h_{a_0} the average edge length of the coarser mesh). These errors are computed on five meshes with respectively 10, 20, 30, 40 and 60 edges on the length of the channel. These meshes are very coarse, nevertheless, it appears that the computed convergence rate are not far from the theoretical ones, the formally second order scheme provides an effective convergence up to the second order when the flow is sufficiently smooth and according to the estimations, the second order scheme reduces to first order in the presence of a discontinuity.

The second test problem is one of the tests of the Telemac code developed at EDF/LNHE [66], it concerns a water drop in a basin and we look at the solution after some reflections on the walls. The basin is a 20.x20. square box with flat bottom, the initial solution shown in Fig. 3.7.2 (a), is defined by

$$h = 2.4(1. + e^{-0.25[(x-10.05)^2+(y-10.05)^2]}), \quad \mathbf{u} = 0.$$

The solutions at $t = 1., 2., 3., 4.$s obtained with the second order approximation are given in Fig. 3.7.2 (b)-(e) while the solution at $t = 4.$s, damped by the first order scheme is shown in Fig. 3.7.2 (f). This result also shows the accuracy improvement due to the second order scheme.

The third test problem is a real life application, it concerns the Malpasset dam break. All the details on the data and a reference solution computed with the Telemac code are given in [67]. We present here in Fig. 3.7.3-3.7.5, the initial solution and the second order solutions at $t = 1000$ s and $t = 2500$ s. These solutions are in good agreement with solutions obtained by other methods in [67]. The computation of this problem allows to test, among others, the ability of the method to treat the still water (the sea area before the wave reaches it) and the wet-dry interfaces.

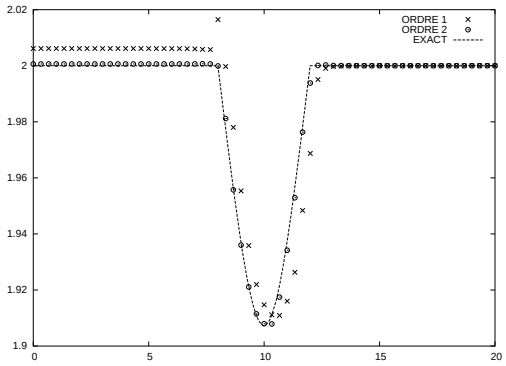
3.8 Conclusion and outlook

In this article we have introduced on one hand a stable homogeneous kinetic solver and on the other hand a hydrostatic reconstruction method to compute the source term while preserving the stability properties of the homogeneous solver. We have also presented a second order compatible extension. Thanks to these three ingredients we finally derived a positivity preserving well-balanced second order scheme. According to the simplicity of the presented algorithms, we emphasize that this solution method seems to be a good compromise between efficiency, stability and accuracy. These properties are experimentally verified by using the algorithm to reproduce complex physical phenomena.

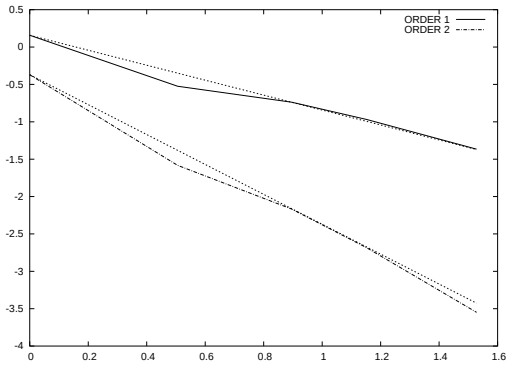
Moreover let us notice some extensions that can be derived. The stability properties of the kinetic solver can be used to derive stable schemes for avalanche flows [98] and can be extended to a multilayer Saint-Venant model [6]. An extension of the hydrostatic reconstruction that preserves all the one-dimensional subsonic steady states is

under investigation and the same idea can be adapted to take into account the relation between the Darcy equation and the Saint-Venant system with strong friction coefficient.

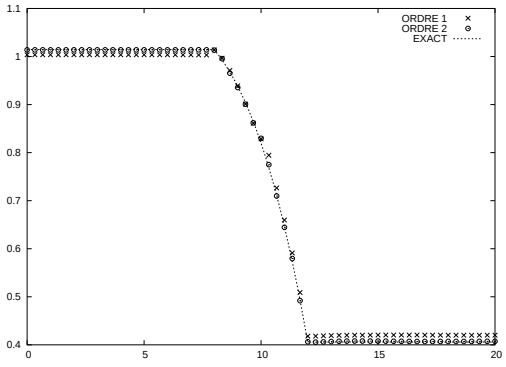
Acknowledgements. The authors thank F. Bouchut, J.M. Hervouet, R. Klein and B. Perthame for fruitful discussions and helpful comments. This work was partially supported by EDF/LNHE and by HYKE European programme HPRN-CT-2002-00282 (http ://www.hyke.org).



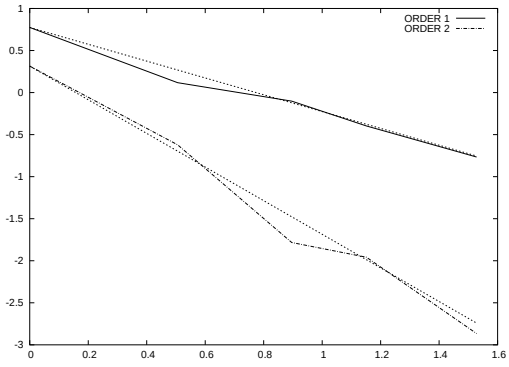
(a) Fluvial flow - Free Surface



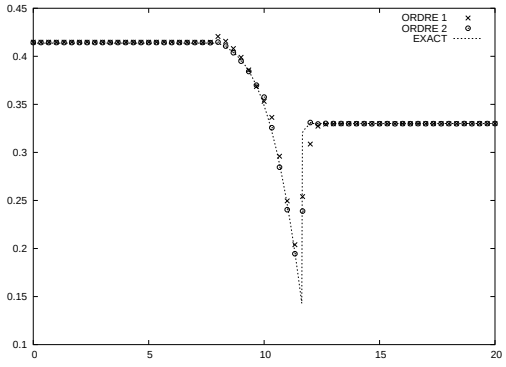
(b) Fluvial flow - Convergence rate



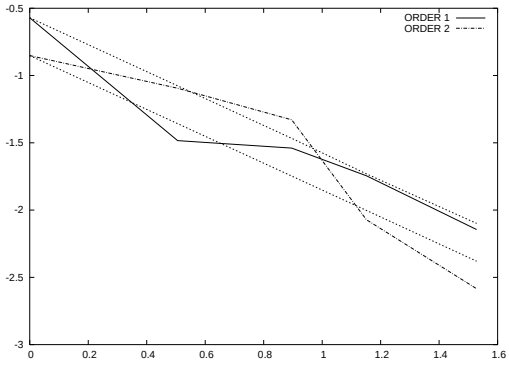
(c) Transcritical flow - Free Surface



(d) Transcritical flow - Convergence rate

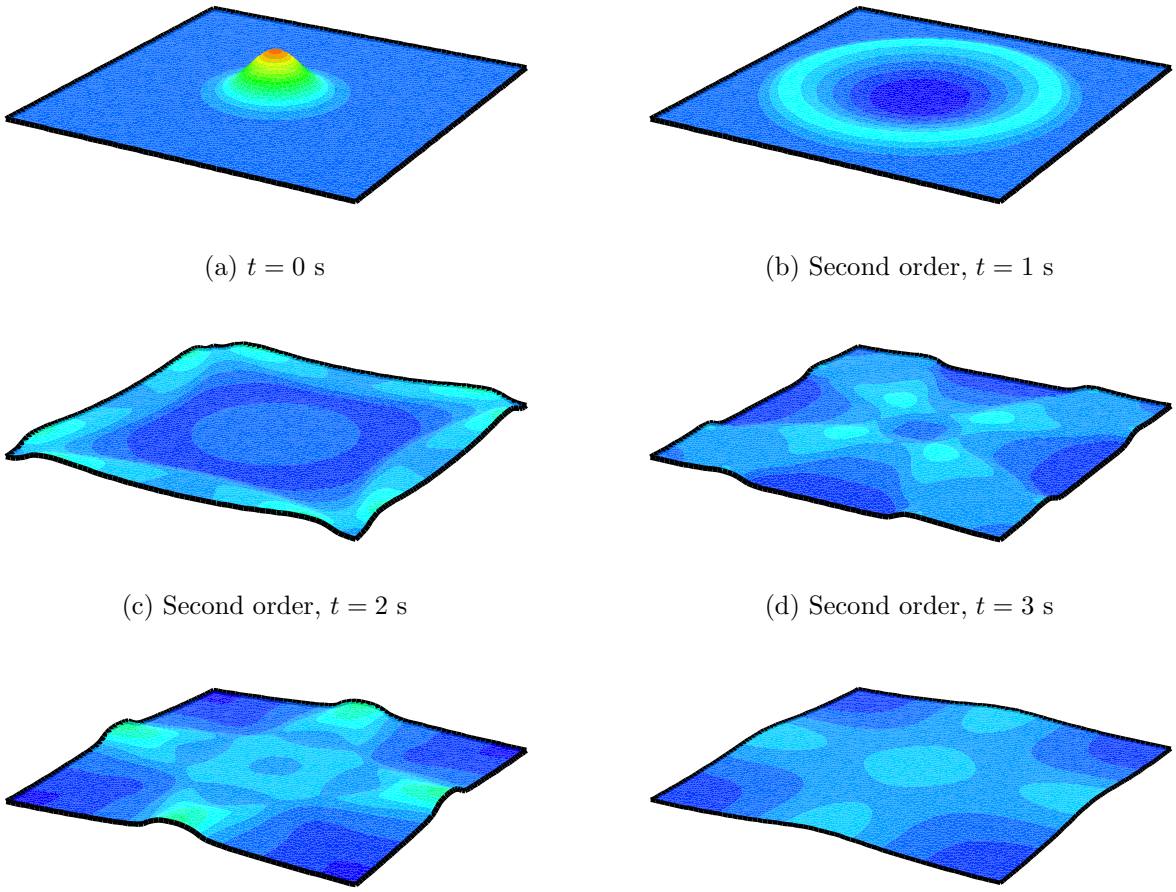


(e) Transcritical flow with shock - Free Surface



(f) Transcritical with shock - Convergence

FIG. 3.7.1: Stationary flows over a bump.



(a) $t = 0$ s

(b) Second order, $t = 1$ s

(c) Second order, $t = 2$ s

(d) Second order, $t = 3$ s

(e) Second order, $t = 4$ s

(f) First order, $t = 4$ s

FIG. 3.7.2: Water drop in a basin.

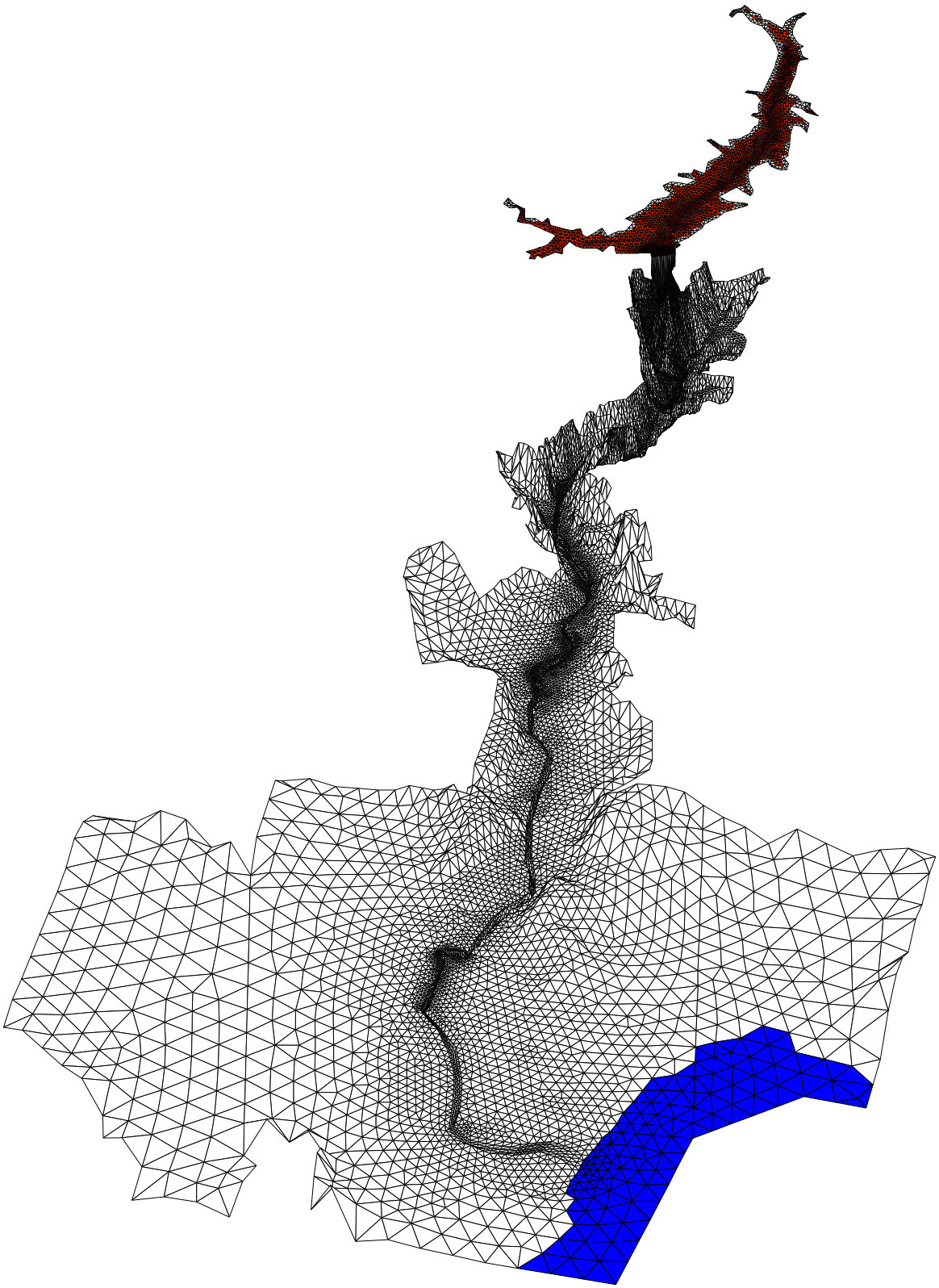


FIG. 3.7.3: Malpasset dam break - Time : 0 s

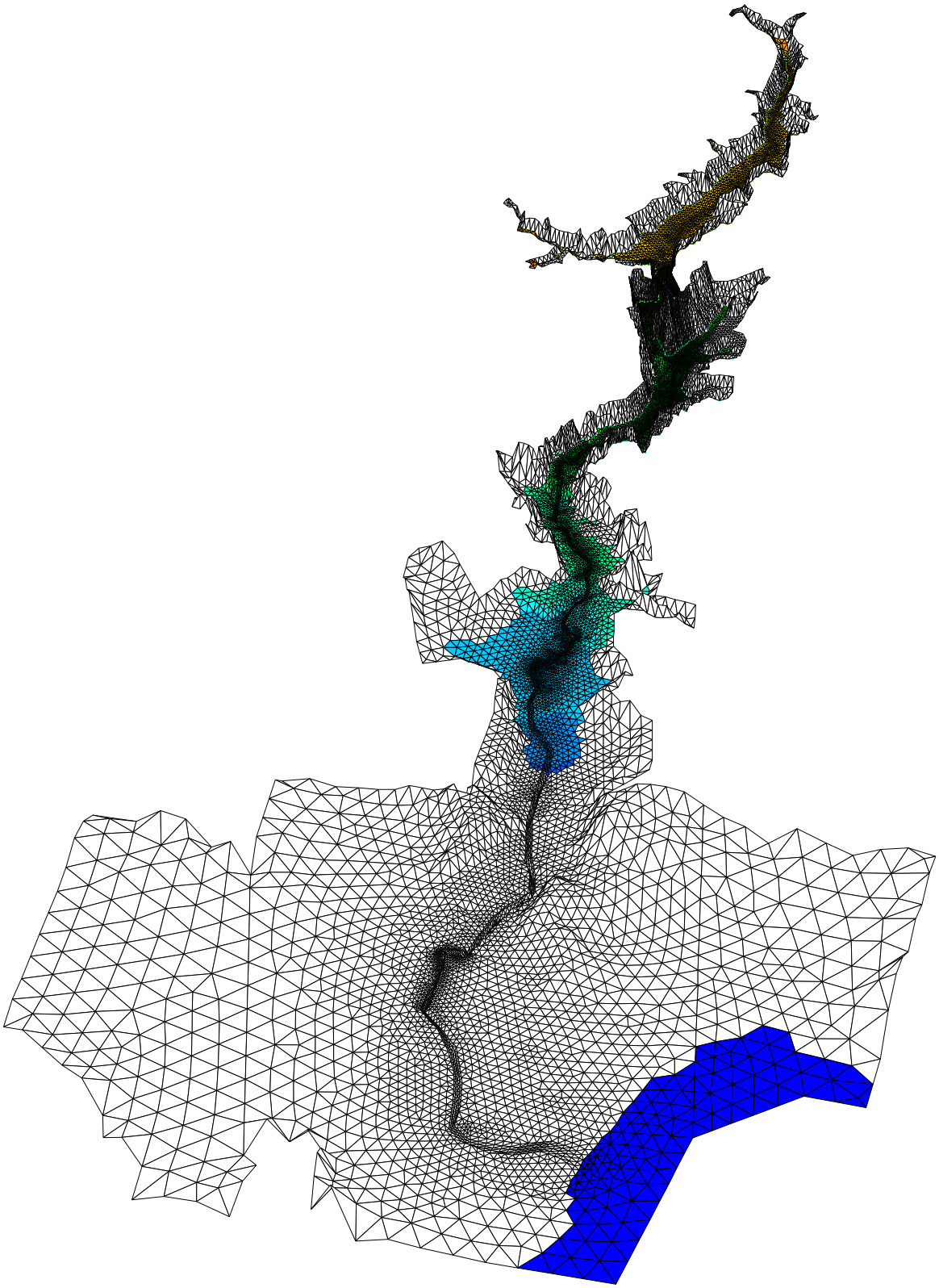


FIG. 3.7.4: Malpasset dam break - Time : 1000 s

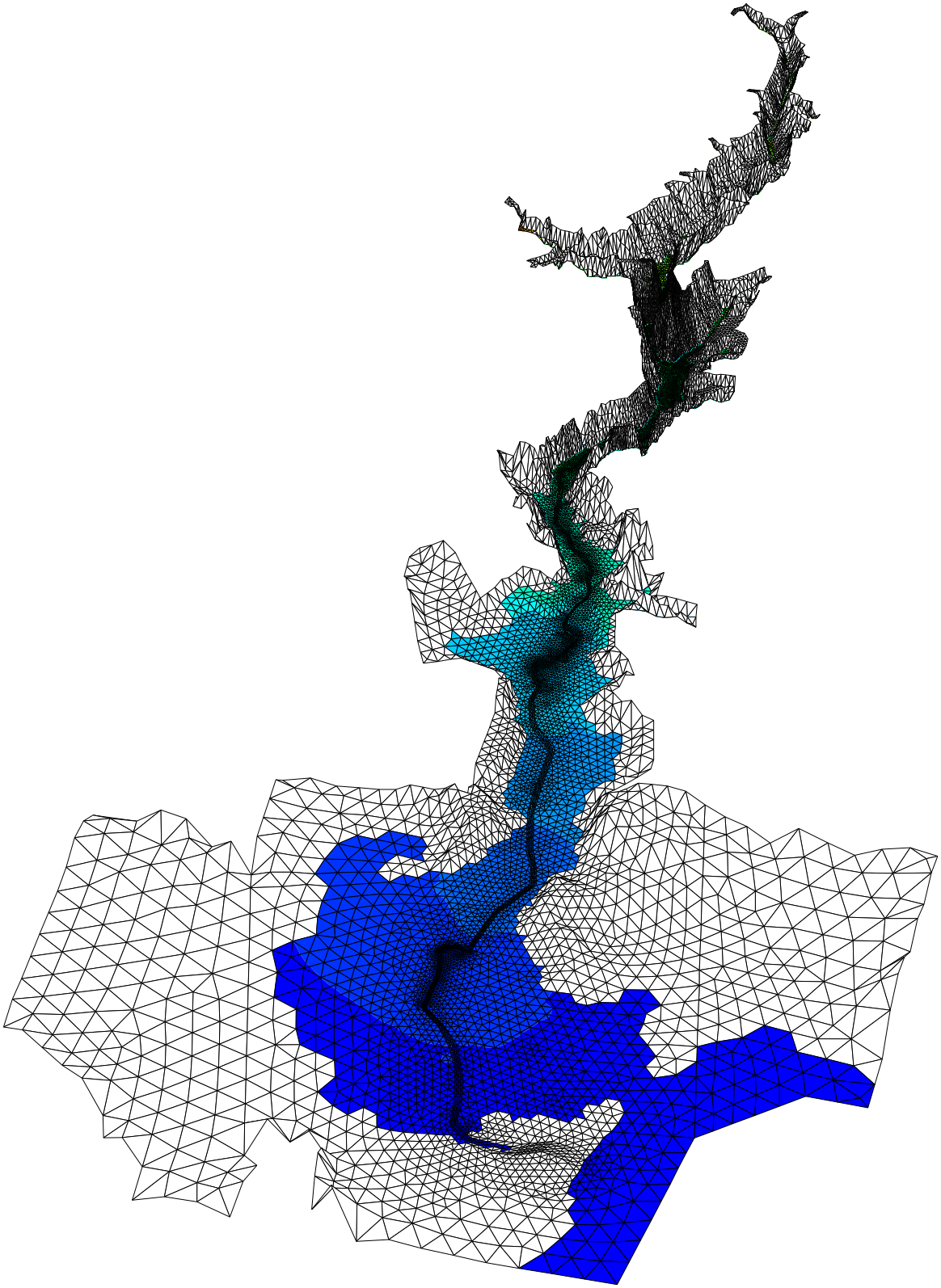


FIG. 3.7.5: Malpasset dam break - Time : 2500 s

Chapitre 4

Transport of pollutant in shallow water flows : A two times step kinetic method

Le travail présenté dans ce chapitre a été effectué en collaboration avec Marie-Odile Bristeau. Il a été publié dans *Mathematical Models and Numerical Analysis*, Vol. **37** (2003), no. 2, pp 389–416.

4.1 Introduction

The shallow water equations are an usual model to describe the flows in rivers or coastal areas. The conservative form is written as a first order hyperbolic system with source terms coming from the bottom topography or the friction on the bed river. Research on solution methods for these equations has received considerable attention in the past two decades and a great number of finite-volume schemes have been developed - refer to [54] or [89] for a detailed presentation. The finite volume method can be used on general triangular grids with a finite element data structure and preserves the conservativity property of the equations. It requires to compute the fluxes at the control volume interface and its stability requires some upwinding in the interpolation of the fluxes and a CFL condition on time steps - refer to [44] for a survey of its properties.

A well-known difficulty of the Saint-Venant system is the preservation of nontrivial equilibria due to the presence of the source terms. In the last years many authors treated this question along with an early idea of Roe [113] to upwind the source terms at the interfaces - see for instance [13], [49], [59], [71] or [91]. On an other hand the nonnegativity of the water height, especially when applications with dry areas are considered, is still a problem for several schemes.

Here we use a kinetic scheme initiated in [9] and developed in [108] and [23]. This scheme is based on a kinetic theory exposed in [105] that allows to link the shallow water equations to a kinetic equation at the mesoscopic level. In [9] only the homogeneous Saint-Venant system has a kinetic interpretation and the source terms are upwind at the macroscopic level. In [108] the source terms are also included in the kinetic formulation. The deduced schemes are conservative and we can prove analytically the nonnegativity of the water height and the conservation of the equilibrium for the steady state of a lake at rest. Moreover with the second scheme the numerical solutions satisfy a discrete entropy inequality. For proofs and numerical examples, see the papers cited above.

More specifically in this paper we are concerned with addressing several questions related to the advection of a pollutant in the Saint-Venant system. It is introduced with a classical transport equation on the concentration of pollutant. The pollution phenomena have today a very large audience and many industrial applications - see for instance [66] for the French electricity group EDF. Then several studies were initiated to understand better the chemical and biological background or the physical mechanisms - see for instance the study about the river Seine near Paris in [99]. But we do not find many works about the well adapted mathematical and numerical treatments. However it is an important point because the transport equation presents many properties which would be interesting to satisfy at the discrete level. Indeed as for the water height we must ensure the nonnegativity of the concentration of pollutant but also a

maximum principle while keeping the conservation property for the pollutant. On an other hand as we do not introduce diffusion in the model we want also to preserve the steady state equilibria related to a lake at rest. Finally we use in our scheme the fact that the characteristic velocities of the hydrodynamic and transport phenomena can be very different in order to disconnect the two time discretizations. It allows a larger time step for pollutant transport based upon the CFL condition

$$|u|\Delta t \leq \Delta x$$

which does not take into account the speed of sound by opposition to the hydrodynamic CFL. It is particularly useful when numerous pollutants are computed or when different pollutant problems are computed with the same hydrodynamic background as it appears often in water quality questions.

The outline of this paper is the following. After presenting the equations in Section 2, we briefly recall the kinetic formulation, the deduced scheme and its extension to the advection equation in section 3. Then we prove in Section 4 the nonnegativity properties, the maximum principle for the pollutant and the preservation of equilibria. Finally in Section 5 we present a two time steps scheme and the associated time discretization. In Section 6 we present some numerical results to illustrate the improvements of the method. Then in Section 7 we discuss an extension to the two dimensional case and we present some numerical results on realistic geometries.

4.2 Equations

The one dimensional shallow water system allows to describe the flow in an ideal rectangular river, at time $t \in \mathbb{R}_+$ and at the point $x \in \mathbb{R}$, through the water height $h(t,x) \in \mathbb{R}_+$ and the velocity $u(t,x) \in \mathbb{R}$, by the hyperbolic system

$$\frac{\partial h}{\partial t} + \frac{\partial (hu)}{\partial x} = S(t,x), \tag{4.2.1}$$

$$\frac{\partial (hu)}{\partial t} + \frac{\partial}{\partial x}(hu^2 + \frac{gh^2}{2}) + gh\frac{\partial z}{\partial x} = 0, \tag{4.2.2}$$

with g the gravity acceleration and $z(x) \in \mathbb{R}$ the bottom topography and where $S(t,x) \in \mathbb{R}$ denotes the sources of water (in $m \times s^{-1}$). Therefore $h + z$ is the level of water surface and in the following we denote $q = hu$ the discharge.
These equations were originally written by Saint-Venant in [115]. Gerbeau and Perthame present in [50] a derivation from the Navier Stokes system. The system (4.2.1)-(4.2.2) corresponds to a very simple case. Other terms can be added in the right hand side, in order to take into account frictions on the bottom or other phenomena.

To perform the analysis of transport of pollutant in this ideal river, we add a third equation

$$\frac{\partial(hT)}{\partial t} + \frac{\partial(huT)}{\partial x} = T_S S, \tag{4.2.3}$$

where $T(t,x) \in \mathbb{R}_+$ is the concentration of pollutant and T_S are the given values of the concentration of pollutant at the sources S. In the following we denote $e = hT$ the quantity of pollutant in the flow.
We can also write this equation on the nonconservative form where 'monotonicity' is better seen

$$\frac{\partial T}{\partial t} + u\frac{\partial T}{\partial x} = \frac{(T_S - T)S}{h}. \tag{4.2.4}$$

It is a classical transport equation. Here we suppose that the pollutant is passive and does not interact with the flow. But in some cases, other phenomena like sedimentation, erosion, birth or death of particles have to be considered.

We can write this system in the conservative and compact form

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} = B(U),$$

with

$$U = \begin{pmatrix} h \\ q \\ e \end{pmatrix} \quad , \quad F(U) = \begin{pmatrix} q \\ \frac{q^2}{h} + \frac{gh^2}{2} \\ \frac{qe}{h} \end{pmatrix} \quad , \quad B(U) = \begin{pmatrix} S \\ -gh\frac{\partial z}{\partial x} \\ T_S S \end{pmatrix}.$$

4.3 The kinetic scheme

4.3.1 Kinetic interpretation of the shallow water equations

We introduce here a kinetic approach to system (4.2.1)-(4.2.3) written with S=0. Then we can deduce from the discretization of the kinetic equation a "kinetic scheme" for the macroscopic system.
Let $\chi(w)$ be an even and compactly supported probability defined on $\mathbb{R}$ satisfying that its second moment is equal to one. Then we introduce two microscopic densities of particles $M(t,x,\xi)$ and $N(t,x,\xi)$ defined by a Gibbs equilibrium

$$M(t,x,\xi) = \frac{h(t,x)}{c(t,x)}\chi(\frac{\xi - u(t,x)}{c(t,x)}), \tag{4.3.1}$$

$$N(t,x,\xi) = \frac{e(t,x)}{c(t,x)}\chi(\frac{\xi - u(t,x)}{c(t,x)}), \tag{4.3.2}$$

with

$$c(t,x)^2 = \frac{gh(t,x)}{2}.$$

We denote

$$G(t,x,\xi) = \begin{pmatrix} M(t,x,\xi) \\ N(t,x,\xi) \end{pmatrix}. \tag{4.3.3}$$

Theorem 4.3.1 *The functions (h,q,e)(t,x) are strong solutions of the shallow water system (4.2.1)-(4.2.3) if and only if $G(t,x,\xi)$ is solution of the kinetic equation*

$$\frac{\partial G}{\partial t} + \xi \frac{\partial G}{\partial x} - g \frac{\partial z}{\partial x} \frac{\partial G}{\partial \xi} = Q(t,x,\xi), \tag{4.3.4}$$

where $Q(t,x,\xi)$ is a "collision term" equal to zero at the macroscopic level.

Proof 4.3.1 *The result is obtained by a simple integration in ξ of the equation (4.3.4) against the matrix $K(\xi)$*

$$K(\xi) = \begin{pmatrix} 1 & 0 \\ \xi & 0 \\ 0 & 1 \end{pmatrix}.$$

The non-linear Saint-Venant system is now reduced to a linear transport system on nonlinear quantities M and N, for which it is easier to find a simple numerical scheme with good theoretical properties. For a detailed proof of the hydrodynamic part of the kinetic interpretation refer to [9] and for the treatment of the source term at this microscopic level see [108].

4.3.2 The kinetic scheme

We now describe the kinetic scheme without taking into account the source term due to the bottom topography. Indeed we are more interested here with the treatment of the pollutant and for the simplicity of the purpose we restrict ourselves to the flat bottom case. For a complete hydrodynamic presentation we refer to [9] or [108].

To approximate the solution $U(t,x)$, $x \in X \in \mathbb{R}$, $t \geq 0$, of the shallow water equations with transport of pollutant by discrete values U_i^n, $i \in I \subset \mathbb{Z}$, $n \in \mathbb{N}$ we consider as usual a grid of points $x_{i+1/2}$, $i \in I$,

$$\cdots < x_{i-1/2} < x_{i+1/2} < x_{i+3/2} < \ldots,$$

and we define the cells (or finite volumes) and their lengths

$$C_i =]x_{i-1/2}, x_{i+1/2}[, \qquad \Delta x_i = x_{i+1/2} - x_{i-1/2} > 0.$$

We shall denote also $x_i = (x_{i-1/2} + x_{i+1/2})/2$ and we consider a timestep $\Delta t^n > 0$ - that will be specified later - and define the discrete times by

$$t^n = \sum_0^{n-1} \Delta t^k, \quad n \in \mathbb{N}^*.$$

Now being given a piecewise constant approximation of the initial data, we must find a formula of the form

$$U_i^{n+1} = U_i^n + \sigma_i^n (F_{i-\frac{1}{2}}^n - F_{i+\frac{1}{2}}^n), \tag{4.3.5}$$

where

$$U_i^n = \begin{pmatrix} h_i^n \\ q_i^n \\ e_i^n \end{pmatrix}$$

and U_i^{n+1} are the piecewise constant approximations at times t^n and t^{n+1} on the cell C_i, where σ_i^n is defined by

$$\sigma_i^n = \frac{\Delta t^n}{\Delta x_i},$$

and where the fluxes $F_{i-\frac{1}{2}}^n$ and $F_{i+\frac{1}{2}}^n$ must be specified.
Here we use the kinetic interpretation to precise the formula (4.3.5). First we define two discrete densities of particles $M_i^n(\xi)$ and $N_i^n(\xi)$ by

$$\begin{aligned} M_i^n(\xi) &= \frac{h_i^n}{c_i^n} \chi(\frac{\xi - u_i^n}{c_i^n}), \\ N_i^n(\xi) &= \frac{e_i^n}{c_i^n} \chi(\frac{\xi - u_i^n}{c_i^n}), \end{aligned}$$

and the corresponding quantity $G_i^n(\xi)$. Then we define a new density function $G_i^{n+1}(\xi)$ at time t^{n+1} with applying a simple upwind scheme for the discrete version of the kinetic equation (4.3.4) for every ξ and without taking account the right hand side

$$G_i^{n+1}(\xi) = G_i^n(\xi) - \xi \sigma_i^n \left(G_{i+\frac{1}{2}}^n(\xi) - G_{i-\frac{1}{2}}^n(\xi) \right),$$

with

$$G_{i+\frac{1}{2}}^n(\xi) = \begin{cases} G_i^n(\xi) & if \quad \xi \geq 0 \\ G_{i+1}^n(\xi) & if \quad \xi < 0 \end{cases}$$

This new density function is not an equilibrium but thanks to the property of the right hand side of (4.3.4) we can recover the macroscopic quantity at time t^{n+1} by a simple integration in ξ. Finally we can precise the macroscopic formula (4.3.5) with

$$F_{i+\frac{1}{2}}^n = F(U_i^n, U_{i+1}^n) = F^+(U_i^n) + F^-(U_{i+1}^n), \tag{4.3.6}$$

where

$$F^-(U_i^n) = \int_{\xi \in \mathbb{R}_-} \xi K(\xi)\ G_i^n(\xi)\ d\xi, \tag{4.3.7}$$

$$F^+(U_i^n) = \int_{\xi \in \mathbb{R}_+} \xi K(\xi)\ G_i^n(\xi)\ d\xi. \tag{4.3.8}$$

A detailed expression of $F^+(U_i)$ can also be written

$$F^+(U_i) = \begin{pmatrix} F_h^+(U_i) \\ F_q^+(U_i) \\ F_e^+(U_i) \end{pmatrix} = h_i \int_{w \geq \frac{-u_i}{c_i}} (u_i + wc_i) \begin{pmatrix} 1 \\ u_i + wc_i \\ T_i \end{pmatrix} \chi(w) dw, \tag{4.3.9}$$

This kinetic method is interesting because it gives a very simple and natural way to propose a numerical flux through the kinetic interpretation and, if we can perform analytically the integration in (4.3.9), it is numerically powerful because the kinetic level disappears and the scheme is written directly as a macroscopic scheme.

Notice also that according to the above derivation, the pollutant flux $F_e^+(U_i)$ that we refer to as the classical kinetic scheme and that we denote $F_{k,e}^+(U_i)$ in the following, has a very simple expression

$$F_{k,e}^{\pm}(U_i) = T_i F_h^{\pm}(U_i). \tag{4.3.10}$$

Nevertheless we will not use it in practice because of a lack of accuracy as we will show in the next Subsection.

4.3.3 Preservation of the equilibria

We do not treat here the problem of the hydrodynamic equilibria because in the case of a flat bottom they are trivially preserved. We refer to [9] or [108] where the authors prove analytically the preservation of the lake at rest equilibrium whatever is the bottom topography.

We are now interested with the problem of the conservation of equilibria for the pollutant. Consider the situation of a lake at rest - with a flat bottom or not - and a numerical scheme that preserves it. Then the numerical discharge vanishes and the water height is constant. So

$$F_h(U_j, U_{j+1}) = 0 \quad \forall\ j \in \mathbb{Z}, \tag{4.3.11}$$

and so due to (4.3.6)

$$F_h^+(U_j) = -F_h^-(U_{j+1}) \quad \forall\ j \in \mathbb{Z},$$

but - from (4.3.9) - these two fluxes are not equal to zero.
Now put some pollutant with concentration equal to one at the node i of the mesh and zero everywhere else. As there is no diffusion in the model this situation is an equilibrium. But numerically - from (4.3.10) - the classical kinetic scheme induces

$$\begin{aligned} &F^{\pm}_{k,e}(U_i) \neq 0, \\ &F^{\pm}_{k,e}(U_j) = 0 \qquad \forall\ j \neq i, \end{aligned}$$

and then (4.3.6) shows immediately that the equilibrium is not preserved for the concentration of pollutant. For long time integration the classical kinetic scheme will create very large diffusion for the pollutant and thus we discard it.

We rather introduce some upwinding in the transport advection depending on the sign of the total mass flux. Then the pollutant flux vanishes with the total mass flux. It is done with the introduction of the new pollutant flux $F_{uk,e}(U_i, U_{i+1})$ defined in the following formula that replaces (4.3.10)

$$F_{uk,e}(U_i, U_{i+1}) = T_{i+\frac{1}{2}} F_h(U_i, U_{i+1}) \tag{4.3.12}$$

where

$$T_{i+\frac{1}{2}} = \left(\begin{array}{cc} T_i & for \quad F_h \geq 0, \\ T_{i+1} & for \quad F_h \leq 0. \end{array} \right) \tag{4.3.13}$$

In the following we call this new scheme the upwind kinetic scheme.

Theorem 4.3.2 *With the pollutant flux (4.3.12)-(4.3.13) the upwind kinetic scheme (4.3.5)-(4.3.8) preserves the pollutant equilibrium in a lake at rest.*

Proof 4.3.2 *The hydrodynamic computation is unchanged and so (4.3.11) is always true. Then from (4.3.12)-(4.3.13) we have immediately*

$$F_{uk,e}(U_j, U_{j+1}) = 0 \qquad \forall\ j \in \mathbb{Z}.$$

4.4 Properties of the scheme

4.4.1 Positivity of the water height

In addition to the preservation of the hydrodynamic and transport equilibria in a lake at rest the upwind kinetic scheme has numerous good properties. As we are interested here with the coupling between the shallow water system and a transport equation, the problem of the source term is not treated. As for the hydrodynamic part of the proofs we restrict ourselves to the case of a flat bottom. But notice that these results also hold true with a non-flat bottom, cf [9] or [108] for complete proofs on

the hydrodynamic part. In Sections 6 and 7 some numerical examples with a non-flat bottom are presented using our scheme.

Theorem 4.4.1 *The scheme is consistent and conservative. It ensures the nonnegativity of the water height under the CFL condition*

$$\Delta t^n \leq min\left(\frac{\Delta x_i}{|u_i^n| + w_M c_i^n}\right). \tag{4.4.1}$$

where $2w_M$ is the size of the compact support of χ.

Proof 4.4.1 *Consistency and conservativity are easily deduced by (4.3.5)-(4.3.8). To prove the positivity of the water height let us go back to the microscopic level. Suppose we have $h_i^n \geq 0\ \forall i \in \mathbb{Z}$. From the definition of the function M in (4.3.1) and the positivity of the function χ, we immediately deduce*

$$M_i^n(\xi) \geq 0 \qquad \forall i.$$

We now introduce the quantities ξ_+, ξ_- defined by

$$\xi_+ = \max(0, \xi), \quad \xi_- = \max(0, -\xi), \tag{4.4.2}$$

and so we write the upwind microscopic scheme deduced by (4.3.4) in the form

$$f_i^{n+1}(\xi) = (1 - \sigma_i^n|\xi|)M_i^n(\xi) + \sigma_{i-1}^n \xi_+ M_{i-1}^n(\xi) + \sigma_{i+1}^n \xi_- M_{i+1}^n(\xi). \tag{4.4.3}$$

Then, for each j, either the value of ξ is such that

$$|\xi - u_j^n| \geq w_M c_j^n,$$

and then from the definitions of w_M and of function M we have

$$M_j^n(\xi) = 0,$$

or the value of ξ is such that

$$|\xi - u_j^n| \leq w_M c_j^n,$$

which implies that $|\xi| \leq (|u_j^n| + w_M c_j^n)$ and then, using the CFL condition (4.4.1), we obtain

$$\sigma_j^n |\xi| \leq 1.$$

Therefore in the relation (4.4.3), $f_i^{n+1}(\xi)$ is a convex combination of nonnegative quantities and thus

$$f_i^{n+1}(\xi) \geq 0.$$

With a simple integration in ξ, we obtain

$$h_i^{n+1} \geq 0.$$

4.4.2 Positivity of the concentration of pollutant

Theorem 4.4.2 *The upwind kinetic scheme (4.3.5)-(4.3.8) and (4.3.12)-(4.3.13) preserves the positivity of the concentration of pollutant.*

Proof 4.4.2 *We present first another equivalent formula for the pollutant flux (4.3.12)-(4.3.13)*

$$F_{uk,e}(U_i, U_{i+1}) = T_i F_h(U_i, U_{i+1})_+ - T_{i+1} F_h(U_i, U_{i+1})_-, \tag{4.4.4}$$

where we refer to (4.4.2) for the definitions of the positive and negative parts.
We assume that the concentration of pollutant is nonnegative at time t^n and we drop the superscripts n for simplicity. Then

$$(hT)_i^{n+1} = (hT)_i + \sigma_i \left(F_{uk,e}(U_{i-1}, U_i) - F_{uk,e}(U_i, U_{i+1}) \right),$$

that is exactly with (4.4.4)

$$\begin{aligned} (hT)_i^{n+1} &= (hT)_i \\ &+ \sigma_i [T_{i-1} F_h(U_{i-1}, U_i)_+ - T_i F_h(U_{i-1}, U_i)_- \\ &\quad - T_i F_h(U_i, U_{i+1})_+ + T_{i+1} F_h(U_i, U_{i+1})_-]. \end{aligned} \tag{4.4.5}$$

Thanks to the nonnegativity of the concentration of pollutant at time t^n we can write the inequality

$$(hT)_i^{n+1} \geq T_i \left(h_i - \sigma_i \left(F_h(U_{i-1}, U_i)_- + F_h(U_i, U_{i+1})_+ \right) \right). \tag{4.4.6}$$

Now we remark from (4.3.7)-(4.3.8) and due to the positivity of function M that

$$\begin{aligned} F^-(U_j) \leq 0 \quad & \forall\ j \in \mathbb{Z}, \\ F^+(U_j) \geq 0 \quad & \forall\ j \in \mathbb{Z}, \end{aligned}$$

and so we can write, using (4.3.6),

$$(hT)_i^{n+1} \geq T_i \left(h_i - \sigma_i \left(F_h^+(U_i) - F_h^-(U_i) \right) \right).$$

Until now we have not used the fact that we are on a flat bottom and if we forget the concentrations of the pollutant in the precedent expression we recognize a step of the proof of the nonnegativity of the water height. We can conclude the two proofs are linked and so if the scheme ensures the nonnegativity of the water height then it ensures automatically the nonnegativity of the concentration of pollutant - whatever is the bottom topography. More precisely in the case of a flat bottom the end of the proof is to use the detailed expression of the fluxes to minimize the right hand side. Indeed

from (4.3.9) and thanks to the property of χ

$$\begin{aligned} F_h^+(U_i) - F_h^-(U_i) &= h_i \int_{w \geq \frac{-u_i}{c_i}} (u_i + wc_i)\chi(w)dw \\ &\quad -h_i \int_{w \leq \frac{-u_i}{c_i}} (u_i + wc_i)\chi(w)dw \\ &\leq h_i \int_{w \in \mathbb{R}} (u_i + |w|c_i)\chi(w)dw \\ &\leq h_i(u_i + w_M c_i). \end{aligned}$$

Therefore, we deduce

$$(hT)_i^{n+1} \geq T_i h_i \left(1 - \sigma_i(u_i + w_M c_i)\right)$$

and the CFL condition (4.4.1) ensures also the nonnegativity of concentration of pollutant.

4.4.3 Maximum principle for the concentration of pollutant

Theorem 4.4.3 *The upwind kinetic scheme (4.3.5)-(4.3.8) and (4.3.12)-(4.3.13) ensures a maximum principle for the concentration of pollutant. Indeed it satisfies*

$$\forall n \quad \forall i \quad T_i^{n+1} \leq max(T_{i-1}^n, T_i^n, T_{i+1}^n). \tag{4.4.7}$$

Proof 4.4.3 *We rewrite (4.4.5) in another form - still dropping the superscripts n for more readability*

$$\begin{aligned} (hT)_i^{n+1} &= T_i \left(h_i - \sigma_i \left(F_h(U_{i-1}, U_i)_- + F_h(U_i, U_{i+1})_+\right)\right) \\ &\quad + T_{i-1}\sigma_i F_h(U_{i-1}, U_i)_+ + T_{i+1}\sigma_i F_h(U_i, U_{i+1})_-. \end{aligned} \tag{4.4.8}$$

The two last parts of the right hand side are clearly nonnegative and we established in the precedent proof it is true for the first part too. So we can maximize the left quantity

$$\begin{aligned} (hT)_i^{n+1} \leq \quad & max\left(T_{i-1}, T_i, T_{i+1}\right) \\ & \left(h_i + \sigma_i(-F_h(U_{i-1}, U_i)_- - F_h(U_i, U_{i+1})_+ \right. \\ & \left. + F_h(U_{i-1}, U_i)_+ + F_h(U_i, U_{i+1})_-)\right). \end{aligned}$$

A simple reorganization of the right hand side gives with the formulae (4.3.5) and (4.3.6)

$$(hT)_i^{n+1} \leq h_i^{n+1} max\left(T_{i-1}, T_i, T_{i+1}\right).$$

4.5 Larger time steps for the pollutant

4.5.1 Motivation

It is well known that the two eigenvalues of the Saint-Venant system are related to the velocity of the flow and to the water height by the formulae $u+\sqrt{gh}$ and $u-\sqrt{gh}$. The eigenvalue of the transport equation is simply equal to the velocity of the flow. So it appears in the cases where the Froude number defined by

$$Fr = \frac{u}{\sqrt{gh}}$$

is small - which is almost always the case for a classical river - that the characteristic time for information transfer is very different for the hydrodynamic and for the transport parts. Then if the CFL number is relevant for the hydrodynamic computation it is not connected to the physical background of the transport phenomenon. Therefore it could be interesting to create an adaptive numerical scheme that allows to disconnect the two time discretizations. Especially in some realistic applications, when we can have to treat twenty or thirty different pollutants - and some of them are modeled by more complicated equations than (4.2.3) - or when we want to test different phenomena of pollution in the same hydrodynamic background. So we would like to solve only the relevant transport states and to store only the useful global hydrodynamic informations while ensuring all the properties presented in the previous section.

The way to reach our goal is to disconnect the transport solution from the CFL condition (4.4.1). Indeed this time step condition is strictly related to the eigenvalues of the hydrodynamic process but is not necessary to ensure the transport properties. In fact it appears that the only theoretical time step restriction is given by the positivity of the right hand side of (4.4.6) - the nonnegativity property that we can then deduce is also sufficient to prove the maximum principle. It is clear that the CFL condition (4.4.1) is more restrictive. On an other hand notice that the positivity of the right hand side of (4.4.6) is the less restrictive formula that we can obtain, if we impose that the nonnegativity condition for the concentration of pollutant depends only of the hydrodynamic process.
At last notice that the new condition we introduce exhibits a term which is analogous to a velocity and so it is in accordance with the remark on the eigenvalues that we make first. Furthermore in the very simple case where we consider a stationary flow on a flat bottom - constant discharge and water height - the nonnegativity of the right hand side of (4.4.6) is reduced to

$$\Delta t^n \leq min\left(\frac{\Delta x_i}{|u_i^n|}\right).$$

4.5.2 Algorithm

We now consider two different time step conditions. Notice that the CFL condition (4.4.1) is an a priori condition - we need only to know the hydrodynamic state at time t^n to compute the time step - while the transport time step condition (4.4.6) is an a posteriori condition - we must compute the mass fluxes before we compute the time step condition. On an other hand it is convenient to manage with the hydrodynamic and the pollutant quantities at the same time and so the transport time step must be the sum of some hydrodynamic time steps.

So we propose the following algorithm : starting from a given state we perform the hydrodynamic computation - time step and fluxes - then we compute the transport time step condition (4.4.6) with this time step and these fluxes. If (4.4.6) is satisfied we perform another hydrodynamic computation, we sum the hydrodynamic time steps and fluxes and we compute (4.4.6) with these sums. We continue this algorithm until (4.4.6) is not satisfied. Then we compute the new pollutant state.

Before we present precisely the algorithm let us make some useful remarks. First the hydrodynamic computation is exactly the same as in the case of the upwind solution. It means that our strategy preserves the fact that the pollutant has no influence on the hydrodynamic phenomena and it ensures that the positivity property is preserved. Second the transport time step is computed as a sum of hydrodynamic time steps. It means that we can not know a priori the transport time step but it allows us to be very adaptive and the transport time step to be as large as possible in relation to the hydrodynamic conditions. Third the reference to the transport time step condition ensures the preservation of the properties of the precedent section.

Let us now present the details of the algorithm :

Initialization

- Start from the state h_i^n, q_i^n, e_i^n.
- Initialization of the transport time step and total fluxes :

$$\begin{aligned} \Delta t^n &= 0, \\ G_h^n(U_{i-1}, U_i) &= 0, \\ G_h^n(U_i, U_{i+1}) &= 0. \end{aligned}$$

- Initialization of the step indicator :

$$k = 0.$$

- Initialization of the small time step hydrodynamic computation :

$$\begin{aligned} h_i^{n,0} &= h_i^n, \\ q_i^{n,0} &= q_i^n. \end{aligned}$$

Computation at the step n,k

1. Computation of hydrodynamic time step $\Delta t^{n,k}$ deduced by the CFL condition (4.4.1) related to $h_i^{n,k}$ and $Q_i^{n,k}$.
2. Computation of the partial fluxes $F^+(U_i)^{n,k}$ and $F^-(U_i)^{n,k}$ with the formula (4.3.9).
3. Computation of the total interfaces fluxes $F^{n,k}(U_i, U_{i+1})$ and $F^{n,k}(U_{i-1}, U_i)$ with the formula (4.3.6).
4. Test based on an extension of the nonnegativity condition (4.4.6) :

$$\begin{aligned} h_i^n &- \frac{1}{\Delta x_i}\left(\sum_0^k \Delta t^{n,j} F_h^{n,j}(U_i, U_{i+1})\right)_+ \\ &- \frac{1}{\Delta x_i}\left(\sum_0^k \Delta t^{n,j} F_h^{n,j}(U_{i-1}, U_i)\right)_- \geq 0. \end{aligned} \tag{4.5.1}$$

Updating of the solution

– If (4.5.1) is true then update the hydrodynamic quantities

$$h_i^{n,k+1} = h_i^{n,k} - \frac{\Delta t^{n,k}}{\Delta x_i}\left(F_h^{n,k}(U_i, U_{i+1}) - F_h^{n,j}(U_{i-1}, U_i)\right),$$
$$Q_i^{n,k+1} = Q_i^{n,k} - \frac{\Delta t^{n,k}}{\Delta x_i}\left(F_Q^{n,k}(U_i, U_{i+1}) - F_Q^{n,j}(U_{i-1}, U_i)\right),$$

the transport quantities

$$G_h^n(U_{i-1}, U_i) = G_h^n(U_{i-1}, U_i) + \Delta t^{n,k} F_h^{n,j}(U_{i-1}, U_i),$$
$$G_h^n(U_i, U_{i+1}) = G_h^n(U_i, U_{i+1}) + \Delta t^{n,k} F_h^{n,k}(U_i, U_{i+1}),$$
$$\Delta t^n = \Delta t^n + \Delta t^{n,k},$$

and the step indicator

$$k = k + 1.$$

Then go to the **Computation** step.

– If (4.5.1) is false then update the hydrodynamic quantities

$$h_i^{n+1} = h_i^{n,k},$$
$$Q_i^{n+1} = Q_i^{n,k}.$$

and the new concentration of pollutant

$$\begin{aligned}(hT)_i^{n+1} &= (hT)_i^n \\ &+ \frac{1}{\Delta x_i}\left(T_{i-1}\left(G_h^n(U_{i-1}, U_i)\right)_+ + T_i\left(G_h^n(U_{i-1}, U_i)\right)_-\right) \\ &- \frac{1}{\Delta x_i}\left(T_i\left(G_h^n(U_i, U_{i+1})\right)_+ + T_{i+1}\left(G_h^n(U_i, U_{i+1})\right)_-\right).\end{aligned}$$

Then go to the **Initialization** step.

4.5.3 Consistency, conservativity, positivity, maximum principle and preservation of equilibria

The hydrodynamic computation is unchanged and so the scheme is still consistent and conservative for its hydrodynamic part and the nonnegativity of the water height is still ensured by the CFL condition introduced at each hydrodynamic time step.

On an other hand the consistency and the conservativity are also preserved for the transport part thanks to the form of the algorithm and the nonnegativity of the concentration of pollutant is ensured at each transport time step by the test (4.5.1). The

maximum principle is satisfied too because the essential point - that is the nonnegativity of each term in equation (4.4.8) - is also ensured by the same test (4.5.1).

Finally the preservation of the different equilibria in a lake at rest is always true because even if the scheme contains now two different time discretizations its global form for hydrodynamic part on one hand and for transport part on the other hand is not modified.

4.6 Numerical results

4.6.1 Transport of pollutant in a flat bottom channel with constant discharge

We consider a flat bottom channel and the associated stationary solution of the shallow water equations with constant discharge and water height. Taking into account this hydrodynamic background and the form of the transport equation (4.2.4) a polluted area will be simply transported with the constant speed u_c of the flow

$$\frac{\partial T}{\partial t} + u_c \frac{\partial T}{\partial x} = 0.$$

The interest of this very simple case is to clearly exhibit the influence of the Froude number and the diffusion of the schemes by comparing the numerical results with an exact solution.

The data of the numerical test are

$$\begin{aligned}
&\text{Spatial domain : } x \in [0, 500] \\
&\text{Uniform mesh with 101 nodes.} \\
&\text{Water height : 1 meter.} \\
&\text{Initial concentration of pollutant : } \begin{cases} 1. & \text{if } \ x \in [20, 70] \\ 0. & \text{if not} \end{cases}
\end{aligned}$$

Then we compute the solution for different Froude numbers Fr - the simulation time is equal to $\frac{100}{Fr}$ - and we present in the following table the number of necessary transport and hydrodynamic time steps. We do not compute the case of the lake at rest since we established that the hydrodynamic and transport equilibria were preserved by both schemes.

Froude number	Simulation Time	Transport Steps	Hydrodynamic Steps
10.	10	71	71
1.	100	70	140
0.1	1000	64	830
0.01	10000	63	7735

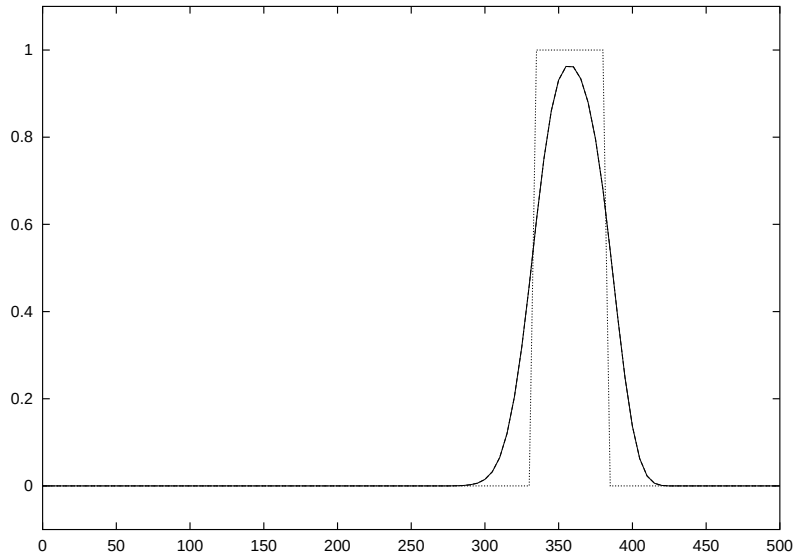
Table 1 : Comparison between numbers of hydrodynamic and transport time steps

First of all these results exhibit the autoadaptive aspect of the two time steps method. Indeed the ratio between the number of transport and hydrodynamic time steps appears to be a function of the Froude number. Then the results imply clearly that in the cases where the Froude number is small the computation cost economy could be very important because the gain in term of number of time steps is very large. It will authorize also to apply easily different pollution models on the same hydrodynamic background since we need to store only one global hydrodynamic information - which corresponds to the global mass flux at each interface - for each transport time step.

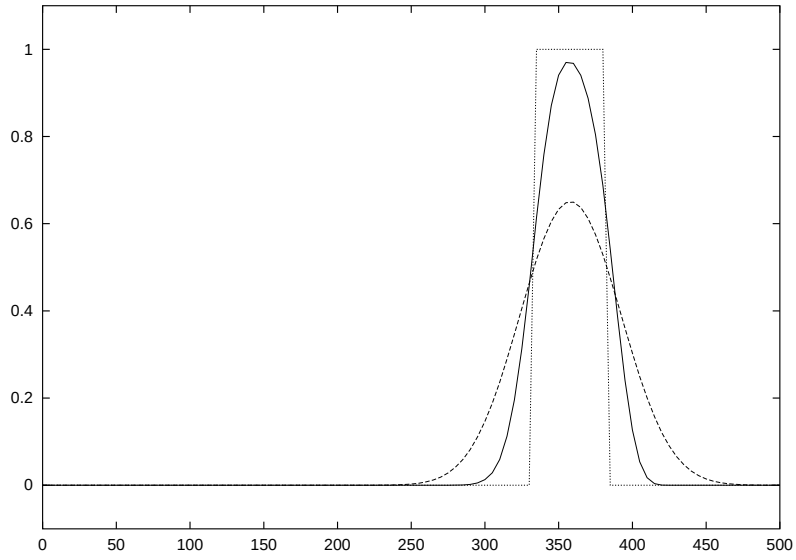
We are interested now with the numerical results. We just mention that the scheme preserves the hydrodynamic stationarity and we present in Figures 4.6.1 and 4.6.2 the numerical results for the pollutant concentration. We indicate the exact solution and the results for the upwind kinetic scheme and the two time steps scheme. First notice that the numerical results are in accordance with the different properties of the schemes established in the precedent sections. Then it appears that the two time steps method does not only improve the computation time but also the precision of the numerical results. Indeed the two time steps solution is always closer to the exact solution than the upwind kinetic one. And lower is the Froude number larger is the difference between the two schemes : if the results are the same for the case where the Froude number is equal to 10. - we see also on Table 1 that the numbers of time steps are the same in this case - the upwind kinetic results are worse and worse when the Froude number decreases - because it corresponds to a large increasing of the simulation time and then of the number of hydrodynamic time steps - while the two time steps results are better and better - because here the number of transport time steps is very stable. To quantify this analysis we present on Table 2 the discrete relative L^1 error on the pollutant concentration for both schemes and for the different cases.

	Fr = 10.	Fr = 1.	Fr = .1	Fr = .01
Upwind kinetic scheme	0.427	0.906	1.099	1.125
Two time steps scheme	0.427	0.412	0.192	0.110

Table 2 : Discrete relative L^1 error



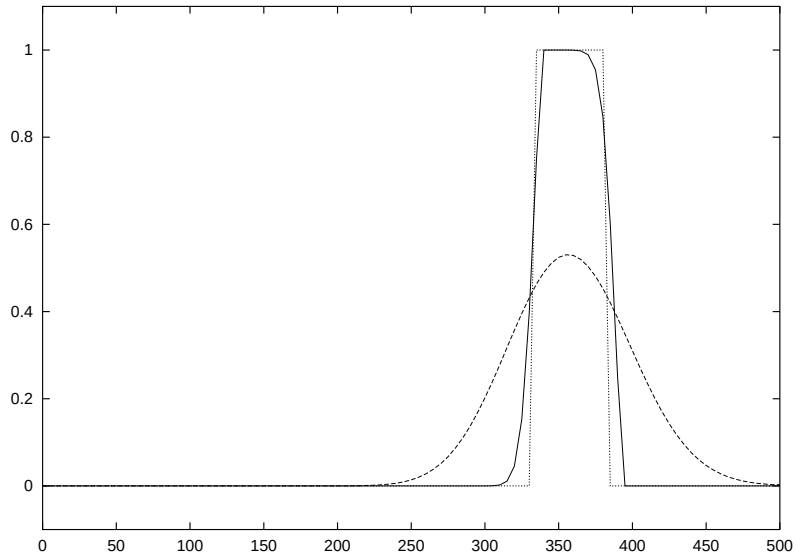
Fr = 10. - t = 10 s



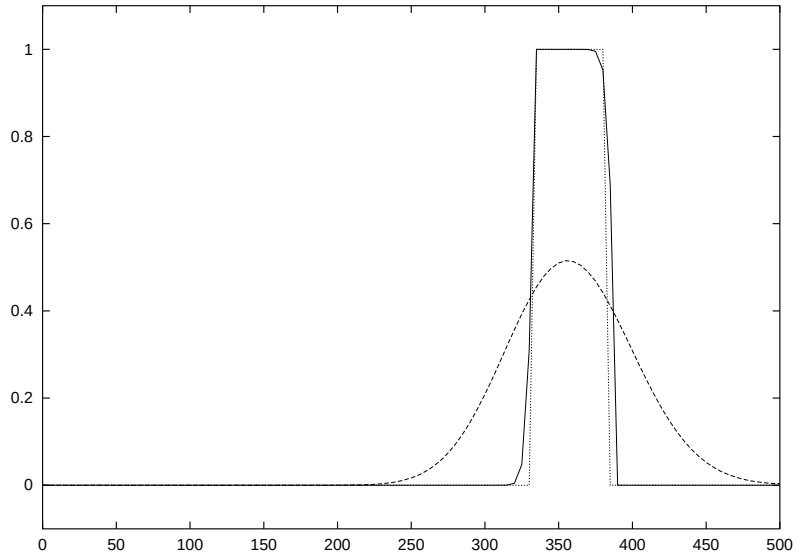
Fr = 1. - t = 100 s

FIG. 4.6.1: Concentration of pollutant for some constant discharge problems

Exact solution (dotted line)
Upwind scheme (dash line)
Two time steps scheme (continuous line)



Fr = .1 - t = 1000 s



Fr = .01 - t = 10000 s

FIG. 4.6.2: Concentration of pollutant for some constant discharge problems

Exact solution (dotted line)
Upwind scheme (dash line)
Two time steps scheme (continuous line)

4.6.2 Dam break

We now consider the very classical case of a dam break on a flat bottom in which the concentration of pollutant is different on each side of the dam.
The interest of this second test is to present a more complex - and then more interesting - hydrodynamic background but preserving the existence of an exact solution for both hydrodynamic and pollutant problems.

The geometrical and time data are the following :

Spatial domain : $x \in [-1000, 1000]$
Uniform mesh with 101 nodes.
Physical time : $240s$.

We present three different cases by modifying only the water height on the right side of the dam. The hydrodynamic and pollutant initial data are - with the classical notations and with the subscripts l for the negative values of x and r for the positive ones

$$\begin{array}{ll} h_l = 1. & h_r = \{.95\ ;\ .8\ ;\ .2\} \\ Q_l = 0. & Q_r = 0. \\ T_l = .7 & T_r = .5 \end{array}$$

The analytic hydrodynamic solution of this problem is given in [120]. It is composed of three flat zones - the two original inactive zones at the two ends and an intermediate one - separated by two simple waves - a rarefaction wave going to the left and a shock wave going to the right. And the jump in the concentration is just transported at the speed u_i of the intermediate flat zone which is given by the relation

$$-2u_i^2 \left(c_l - \frac{u_i}{2}\right)^2 c_r^2 + \left(\left(c_l - \frac{u_i}{2}\right)^2 - c_r^2\right)\left(\left(c_l - \frac{u_i}{2}\right)^4 - c_r^4\right) = 0, \qquad (4.6.1)$$

where c_l and c_r are defined by

$$c_l^2 = gh_l, \quad c_r^2 = gh_r.$$

We present in Figure 4.6.3 the water height and the discharge profiles for the large jump - $h_r = .2$. It appears that the velocity of the shock and the water height of the intermediate flat zone are well captured.

Then as for the first test case we exhibit the autoadaptive aspect of the method in terms of ratio between the number of hydrodynamic time steps and the number of transport time steps.

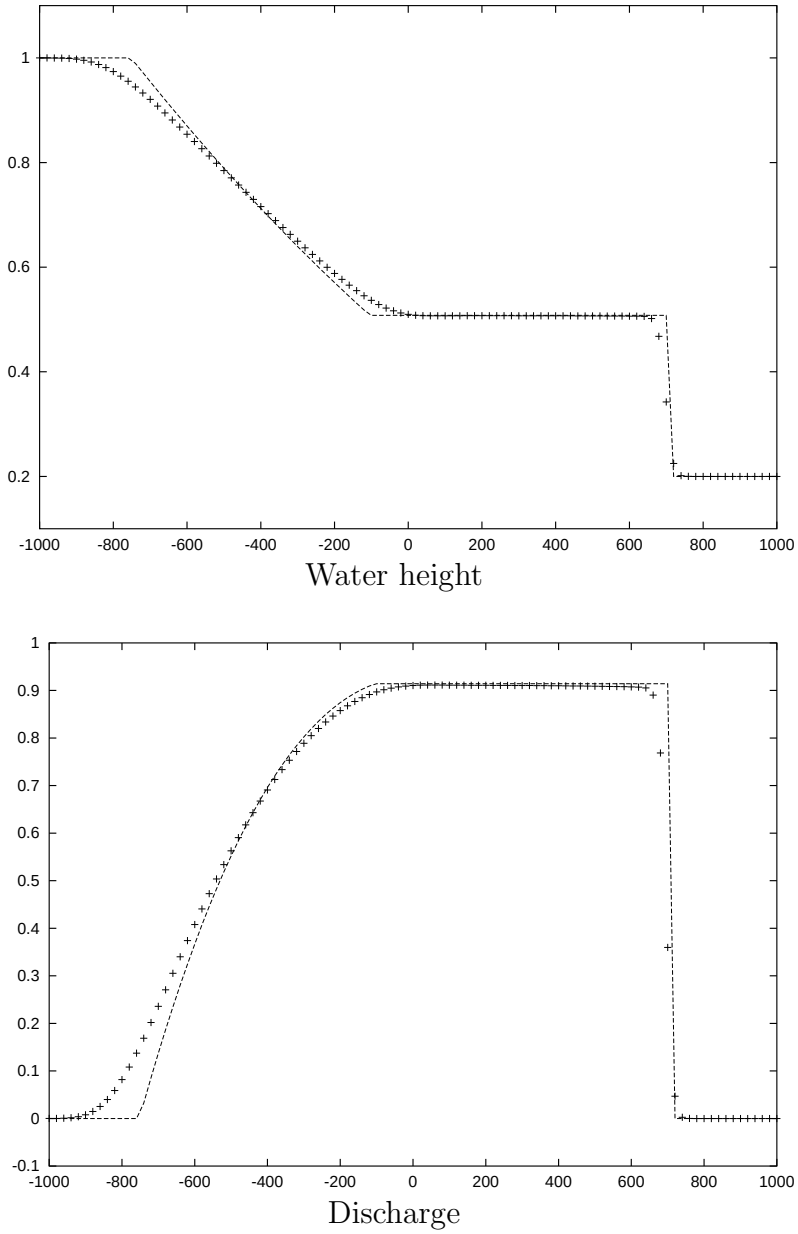
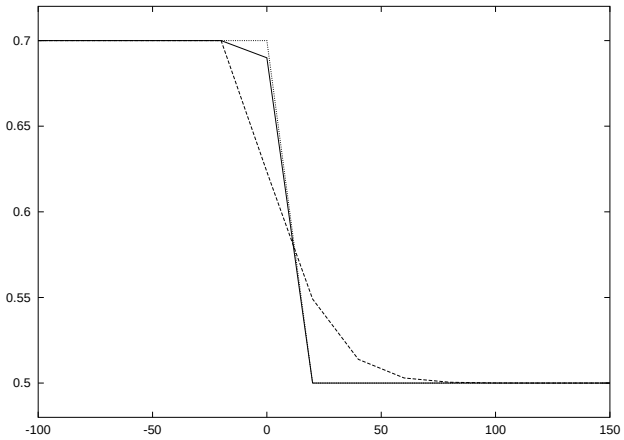
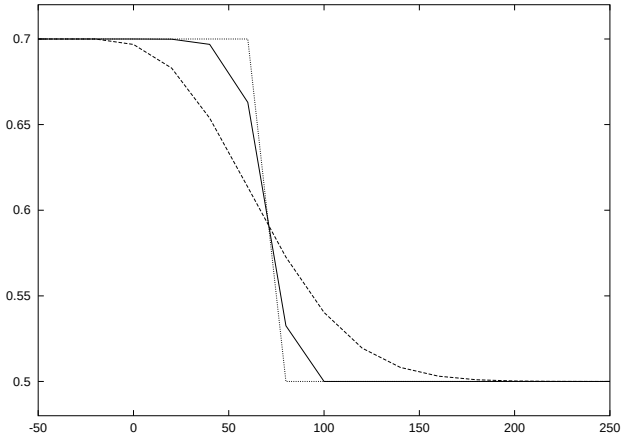


FIG. 4.6.3: Dam break on a flat bottom : Hydrodynamic results

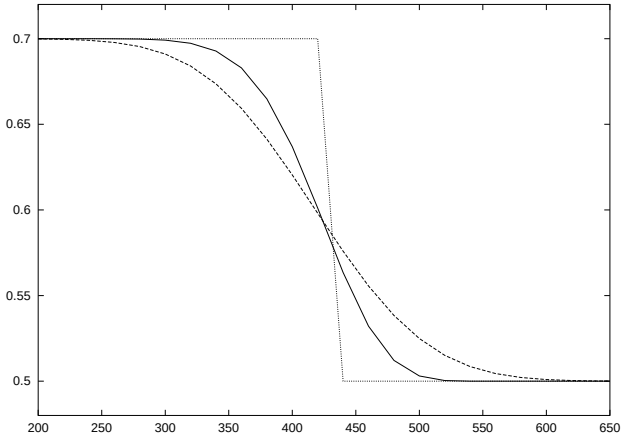
Exact solution (dash line)
Kinetic scheme (crosses)



$h_l = 1. \quad - \quad h_r = .95$



$h_l = 1. \quad - \quad h_r = .8$



$h_l = 1. \quad - \quad h_r = .2$

FIG. 4.6.4: Concentration of pollutant for some dam break problems

Exact solution (dotted line)
Upwind scheme (dash line)
Two time steps scheme (continuous line)

h_l	h_r	Froude number	Transport Steps	Hydrodynamic Steps
1.	.95	0.026	1	47
1.	.8	0.112	5	48
1.	.2	0.804	27	54

Table 3 : Comparison between numbers of hydrodynamic and transport time steps

Finally we present in Figure 4.6.4 the zoom view of the discontinuous area for the concentration of pollutant for the three cases.
Let us notice some remarks : First the pollutant shock front is well captured by both schemes and its speed grows with the size of the water height jump as it is given by the formula (4.6.1). Second - as for the first test case - the two time steps solution is always better than the upwind one and the difference is more important when the jump is smaller - i.e. when the Froude number is smaller.

4.6.3 Peak in the concentration of pollutant

The hydrodynamic initialization is still a dam break. The concentration of pollutant is still different on each side of the dam. But now there is a peak in the concentration of pollutant just before the dam.
In this third test case we mix the first and the second one. The hydrodynamic background is quite complex and the initial concentration of pollutant present some oscillations. The 1D results presented in this section will be later compared with a 2D simulation.

The numerical values are the same as in the first example of the previous subsection excepted the initial value for the concentration of pollutant which is

$$\begin{aligned} -1000. \leq x < -100. \quad &: \quad T(x,0) = T_l = 0.7 \\ -100. \leq x \leq 0. \quad &: \quad T(x,0) = T_i = 0.9 \\ 0. < x \leq 1000. \quad &: \quad T(x,0) = T_r = 0.5 \end{aligned}$$

We do not have an analytical solution for this case. But we know the initial value at the node $x = 0.$ will be simply transported with the speed u_i given by (4.6.1). So as we prove on an other hand a maximum principle for the concentration of pollutant we know its maximum value will stay equal to T_i.

On the first plot of the Figure 4.6.5 are presented the two numerical solutions after 250 seconds. We indicate the initial solution too. As for the second case we observe

a shock front and as for the first case the large diffusion of the upwind kinetic scheme conducts to loose the exact maximum value of the concentration of pollutant while the two time steps scheme computes it very precisely. Informations about the comparison between the numbers of time steps are given in a following subsection.

4.6.4 Emission of pollutant in a non flat bottom channel

Here we want to test the introduction of a source of pollutant in a stationary flow on a non flat bottom. It is the academic 1D version of a 2D problem that will be presented later and that models for instance the emission of waste water in a river.

We consider an academic parabolic bottom profile with a length of 500 meters

$$Z(x) = \left(0.2 - 0.05(x - 250)^2\right)_+ ,$$

and we use a uniform mesh with 101 points of discretization.
Then there exist four different hydrodynamic stationary cases following the different fluid states - fluvial or torrential flows. See [9] or [108] for more details. Here we choose to work with the fluvial flow problem where the improvement due to the two time steps scheme is very clear because of a small Froude number.

$$h(x, 0) = H = 2.$$
$$Q(x, 0) = Q = 1.$$

We introduce between times t_b and t_e and at the node I_s a source of water S_s with a concentration of pollutant T_s. Then we follow the evolution of this pollutant layer. Notice that the source modifies locally in time and space the hydrodynamic computation.

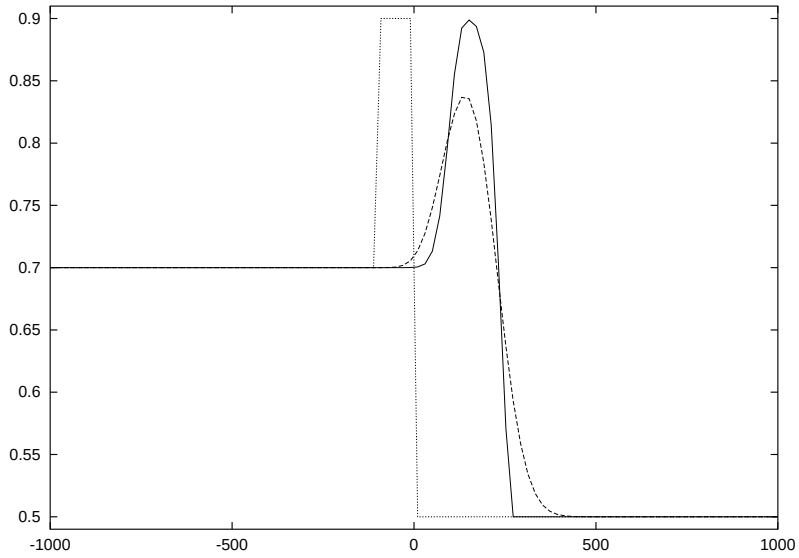
$$I_s = 10, \quad S_s = .01m * s^{-1}, \quad T_s = 10.$$
$$t_b = 100s, \quad t_e = 300s$$

Numerical results are presented in Figure 4.6.6 and informations about the number of time steps are given in the next subsection.

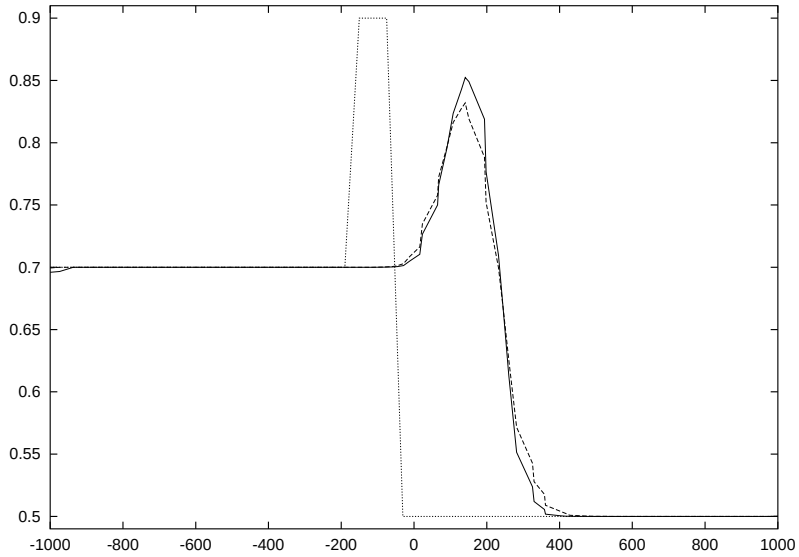
4.6.5 With a non uniform mesh

We mention before that the two time steps method allows an important improvement in term of accuracy. Here we want to exhibit - through some numerical examples - that this improvement is very related to the regularity of the mesh. We introduce a non-uniform mesh - which is a controlled random perturbation of the uniform one - and we characterize it through its minimum and maximum space steps

$$\Delta x_{min} = 0.33316$$
$$\Delta x_{max} = 1.81694$$



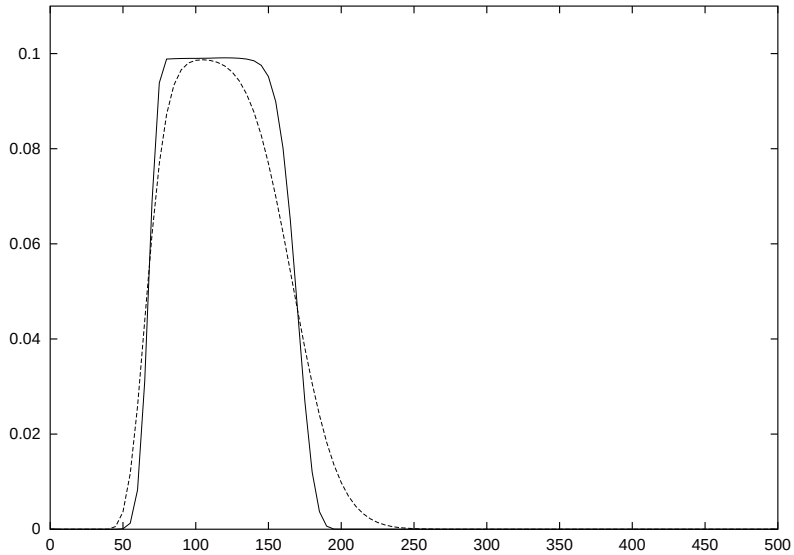
Uniform mesh



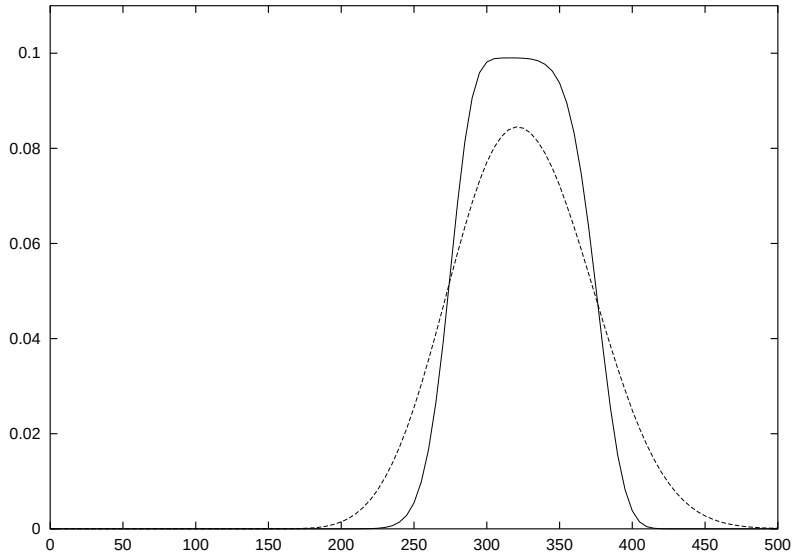
Non uniform mesh

FIG. 4.6.5: Concentration of pollutant for the peak problem

Initial data (dotted line)
Upwind scheme (dash line)
Two time steps scheme (continuous line)



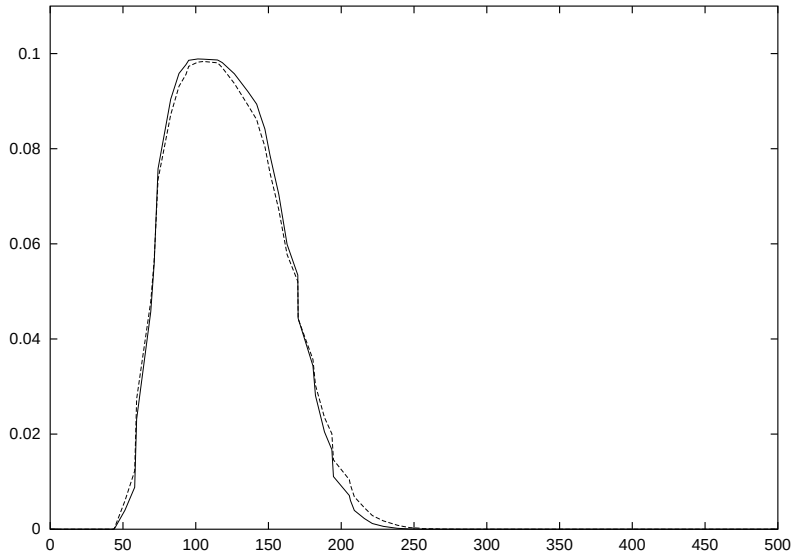
Uniform mesh - t = 350 s



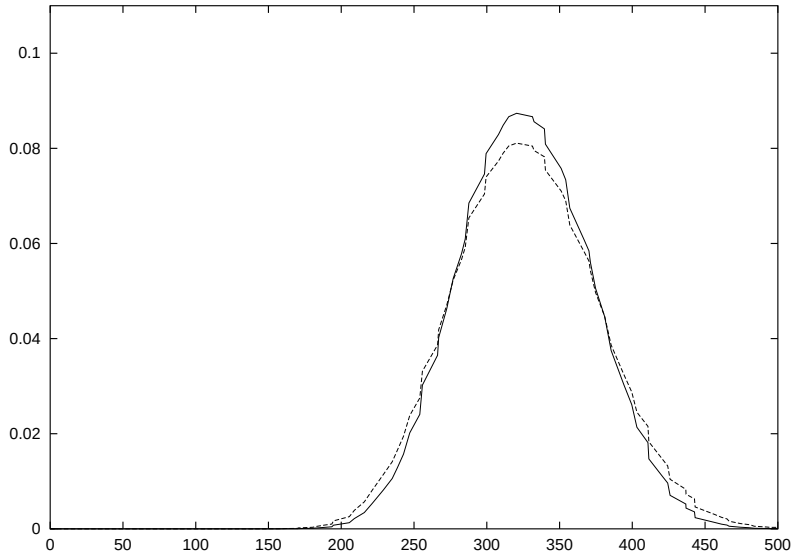
Uniform mesh - t = 750 s

FIG. 4.6.6: Concentration of pollutant - Emission of pollutant in a fluvial flow over a bump

Upwind scheme (dash line)
Two time steps scheme (continuous line)



Non uniform mesh - t = 350s



Non uniform mesh - t = 750s

FIG. 4.6.7: Concentration of pollutant - Emission of pollutant in a fluvial flow over a bump

Upwind scheme (dash line)
Two time steps scheme (continuous line)

Then we come back on the two last test cases. We begin with the dam break problem with a peak in the concentration of pollutant. We compute this test on the new mesh and with the two methods - the upwind kinetic scheme and the two time steps kinetic scheme - and we present the results on the second plot of the Figure 4.6.5. We first remark that more irregular is the mesh more the CFL condition is hard since it is related to the lower space step. So the number of time steps and then the numerical diffusion increase with the irregularity of the mesh and the precision of the results decreases. On an other hand it appears clearly that the difference in the precision of the results between the upwind and the two time steps methods decreases when the mesh is very irregular. However the two time steps method is always more accurate than the upwind one and - that is the essential point - we show on the table below that the ratio between the number of transport and hydrodynamic time steps is independent of the regularity of the mesh. So there is always an important improvement on the computation time and on the storage.

Mesh	Transport Steps	Hydrodynamic Steps
Regular mesh	13	53
Semi-random mesh	30	148

Table 4 : Comparison between different meshes for a dam break problem with a peak in the concentration of pollutant

Then we compute the emission of pollutant problem of the precedent subsection. We apply the upwind and the two time steps methods on our new mesh and we present the results for two times - 350 seconds and 750 seconds - in the Figure 4.6.7. We present the informations about the number of time steps in the table below too. The conclusions are the same as for the dam break problem.

Mesh	Physical time	Transport Steps	Hydrodynamic Steps
Regular mesh	350	42	416
Regular mesh	750	89	890
Semi-random mesh	350	114	1246
Semi-random mesh	750	243	2669

Table 5 : Comparison between different meshes for an emission of pollutant problem

4.7 Extension to the 2D case

We now want to apply our new method to two-dimensional problems. We do not want to make here a complete presentation of the two dimensional finite volume method

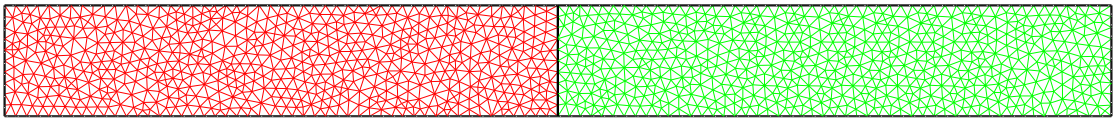
on a general triangular grid or of the general kinetic theory in 2D. We just mention that starting from a triangulation of $\mathbb{R}^2$ the dual cells C_i are obtained by joining the centers of mass of the triangles surrounding each vertex P_i. Then the general method is close to the 1D finite volume method. Indeed the fluxes which appear in the scheme are interpolations of the normal component of the fluxes on the edge of each cell. So locally the problem can be treated as a planar discontinuity and the interpolation can be performed using a one dimensional solver. For a complete presentation of the kinetic interpretation of the hydrodynamic part in 2D refer to [9] and for a presentation that includes the transport theory see [24]. Before we present the numerical results let us insist on the fact that as the 2D computation is based on a 1D strategy the properties we proved in 1D are also true for the 2D schemes. We do not reproduce the proofs here because they are easy extensions of those in Section 4.

We perform two numerical tests : first a dam break problem in a rectangular channel with a peak in the initial concentration of pollutant - see the subsection 6.2 in 1D - and then an emission of pollutant problem in a realistic river geometry.
For the dam break problem we use two meshes : a uniform one and an unstructured one. Then as in the 1D case we can compare the results based on the mesh regularity. The uniform mesh has 1111 nodes and 2000 elements and the unstructured one has 1347 nodes and 2472 elements. As the dam break case is essentially a 1D problem we can also make comparisons between 1D and 2D solutions.
The realistic river case is managed only with an unstructured mesh and exhibits that our method is well adapted to treat complex geometry since the gain on the number of time steps is even more important than for the academic problems.

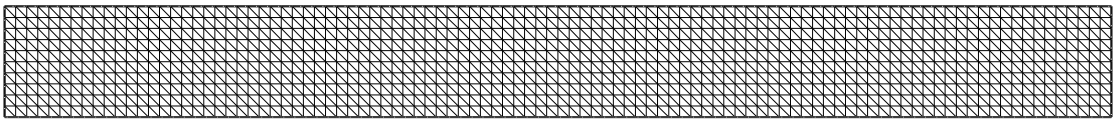
4.7.1 A 2D dam break problem

We begin with the dam break problem. The data are the same as for the problem performed in the Subsection 6.2 except the length of the channel which is now one meter - and so the final time is only one second. The width is chosen equal to 0.1 meter.
The results are presented in Figures 4.7.1 and 4.7.2. The data are invariant in y-coordinates and so the solution is very close to a 1D solution. For each figure the blue color indicates the minimum level and the red color the maximum one.
In Figure 4.7.1 we first present the two meshes. Then we present the water height in the channel. On the left we recognize the rarefaction wave and on the right the shock wave. Then on the last plot the Froude number is presented. The profile is very close to the discharge profile in 1D.
In Figure 4.7.2 is first presented the initial data for the pollutant concentration and then on the four last plots are presented the concentration of pollutant performed with the upwind kinetic scheme and with the two time steps scheme on the unstructured

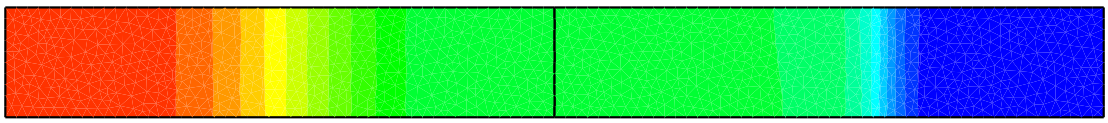
mesh and then the same results with the uniform one. The profiles are very similar to the 1D profile. With the unstructured mesh the two profiles are very close - even if the maximum of the concentration is a little larger with the two time steps method as in the 1D computation when the mesh is non-uniform. With the uniform mesh the profiles are more different. With the upwind kinetic scheme the results are worse than on the unstructured mesh - notice that the uniform mesh has less nodes - but with the two time steps scheme the results are better than on the unstructured mesh. So as in the 1D computation the improvement on the accuracy of the results is more significant when the mesh is more regular.



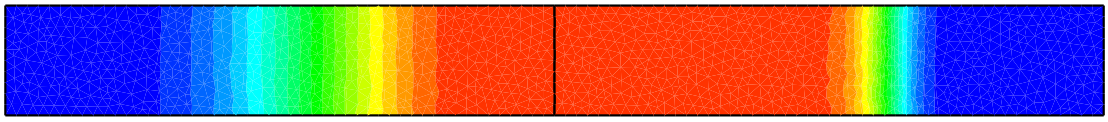
Unstructured mesh



Uniform mesh



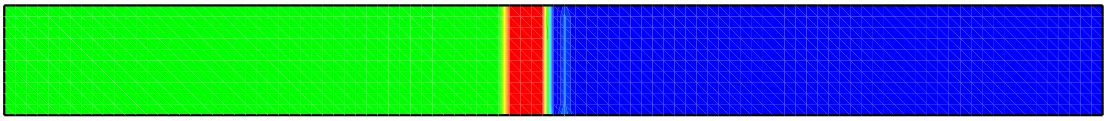
Water height



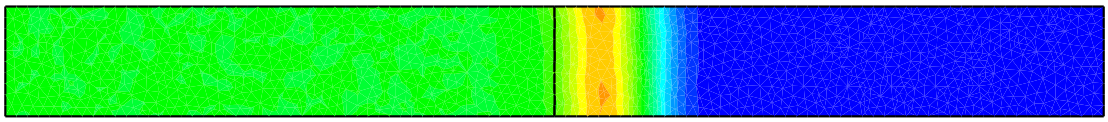
Froude Number

FIG. 4.7.1: 2D dam break- Meshes and hydrodynamic results

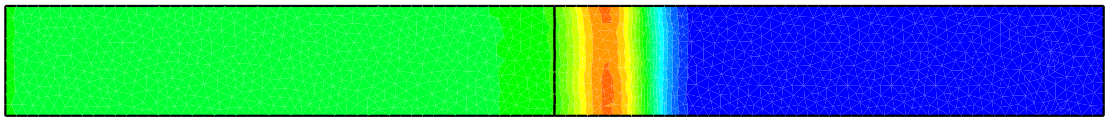
Before to end with this problem we make a comparison between the 1D and the 2D results. So we perform the computations with the same initial conditions and we consider the 2D results on the line y=0.5. Then we present in Figure 4.7.3 the results on the two 1D meshes and on the two 2D meshes. We can see that the regularity of the mesh has the same effects in 1D and in 2D - even if the 1D uniform mesh induces a smaller diffusion than the 2D uniform mesh.



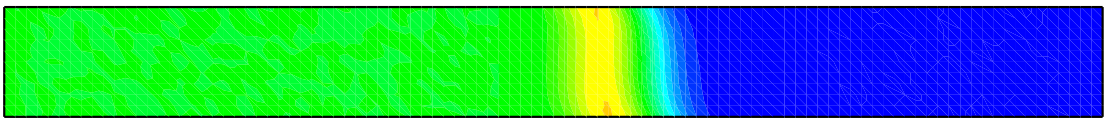
Initial solution



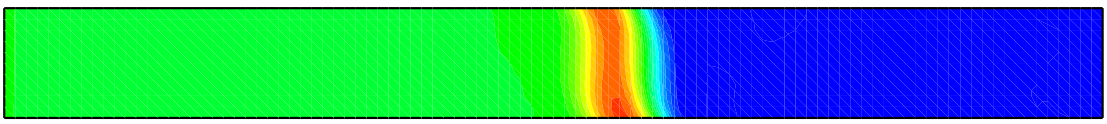
Unstructured mesh - Upwind kinetic scheme



Unstructured mesh - Two time steps kinetic scheme



Uniform mesh - Upwind kinetic scheme



Uniform mesh - Two time steps kinetic scheme

Scale for concentration of pollutant values

FIG. 4.7.2: 2D dam break - Concentration of pollutant

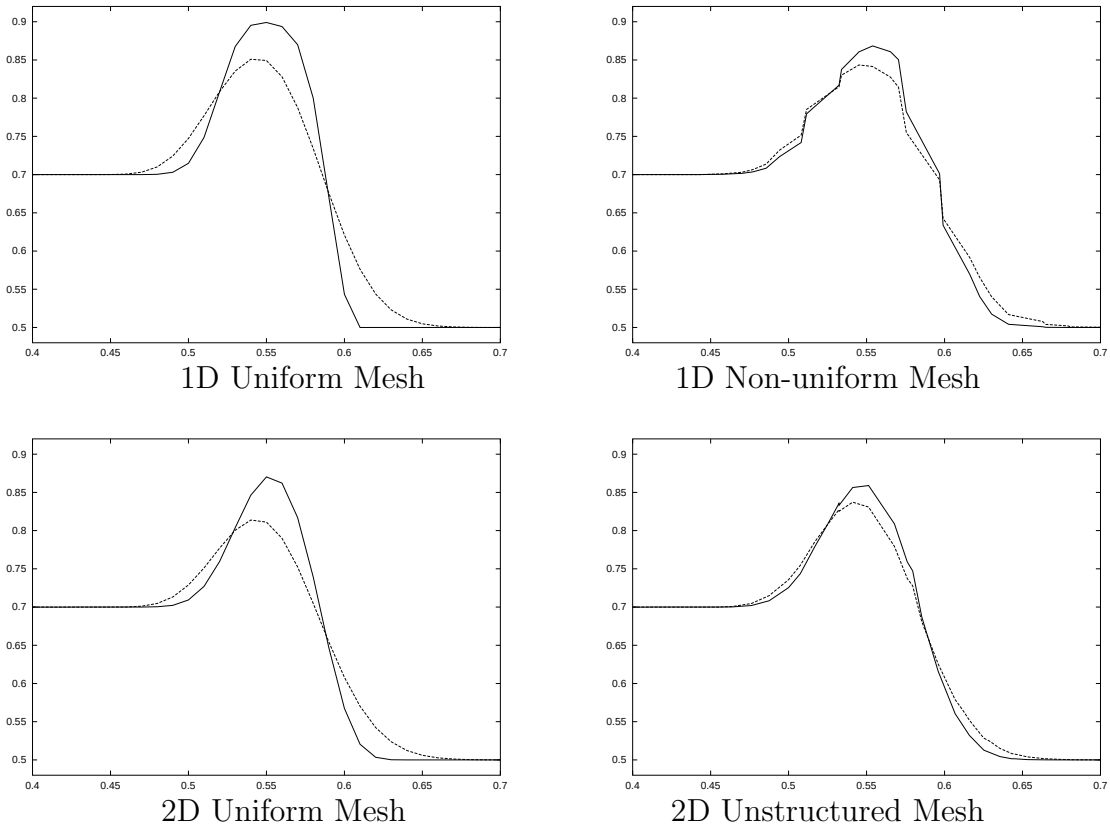


FIG. 4.7.3: Concentration of pollutant for a peak problem
Comparison between the 1D problem and the 2D problem (plane y=0.5)

Upwind scheme (dash line)
Two time steps scheme (continuous line)

4.7.2 Emission of pollutant in a realistic river

Then we perform an emission of pollutant problem in a river. The geometric data include a jetty in the transversal direction and a bridge pillar. We introduce a source of pollutant at one node of the mesh and for a given time - from 1000 seconds to 1800 seconds - and we follow the evolution of the pollutant layer.
We use a second order scheme to compute the hydrodynamic part because the first order solution does not show the recirculation after the jetty which is very important to compute a realistic profile of the concentration of pollutant. We keep a first order solution for the transport part not to mix the improvement due to high order schemes and those due to the applied method. Results are presented in Figure 4.7.5. Here the mesh is unstructured and so the concentration of pollutant profiles are very close. But as in the other cases the two time steps scheme is a little bit better.

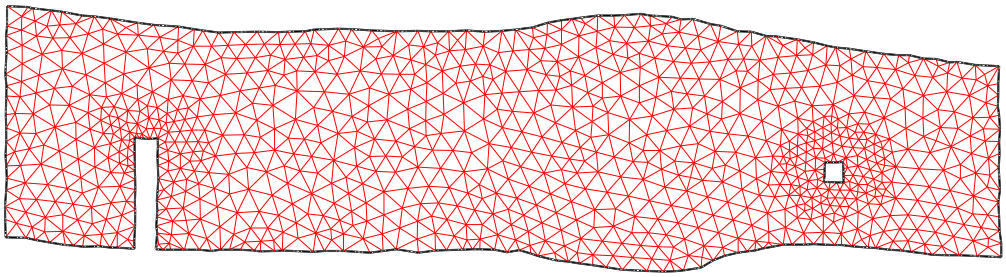
4.8 Conclusion

Thanks to a precise analysis of the upwind kinetic scheme we have deduced a two time steps kinetic scheme that preserves the theoretical properties of the upwind kinetic scheme. By opposition to the hydrodynamic CFL condition, the new transport time step condition links automatically the transport time step, the space step and the fluid velocity, ignoring the sound speed.
The hydrodynamic part of the computation remains unchanged. The whole hydrodynamic information that is useful for the pollutant transport computation can be stored in global interface fluxes at each transport time step.
The developed method is very interesting for the small Froude number flows. Indeed the two time steps are very close when the speed of the flow is large compared with the sound speed but they are very different in the other case. As we can see on the table below, in a 2D realistic geometry and for a Froude number close to 0.1 - which is a usual order of magnitude for rivers, the new transport time step is around hundred and fifty times larger than the hydrodynamic time step issued from the CFL condition.

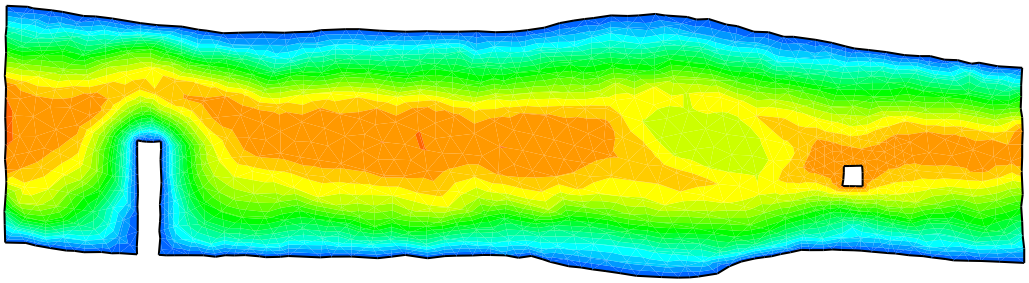
	Froude number	Transport Steps	Hydrodynamic Steps
Unstructured channel	0.35	28	385
Uniform channel	0.37	14	170
Realistic river	0.08	320	45637

Table 6 : Comparison between numbers of hydrodynamic and transport time steps for the 2D test cases

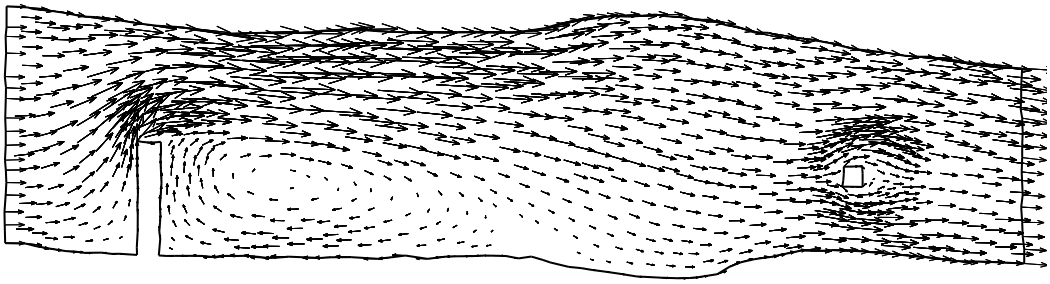
The improvement is proved to be optimal if we want to preserve a priori the nonnegativity properties of the upwind kinetic scheme.



The mesh

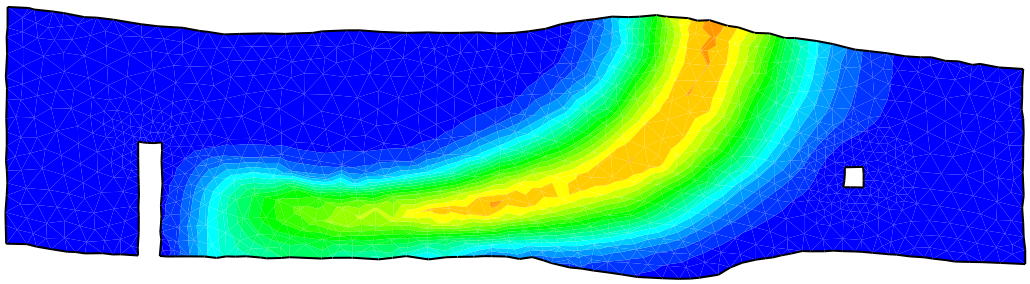


Water height

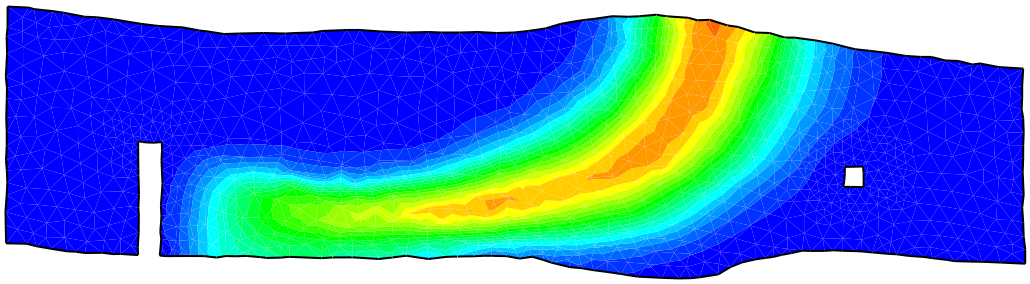


Velocity

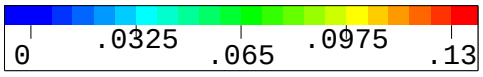
FIG. 4.7.4: River with emission of pollutant - Mesh and hydrodynamic results



Concentration of pollutant - Upwind kinetic scheme



Concentration of pollutant - Two time steps kinetic scheme



Scale for concentration of pollutant values

FIG. 4.7.5: River with emission of pollutant - Concentration of pollutant

As the new time step condition is specifically adapted to the transport equation the numerical diffusion of the two time steps kinetic scheme is lower than the numerical diffusion of the upwind kinetic scheme and we therefore improve the accuracy of the results.

Acknowledgment. This research has been partially supported by EDF/LNHE. The authors thank Benoit Perthame and Jean-Michel Hervouet for helpful discussions.

Chapitre 5

Uniqueness for a scalar conservation law with discontinuous flux via adapted entropies

Le travail présenté dans ce chapitre a été effectué en collaboration avec Benoit Perthame. Il a été soumis pour publication.

5.1 Introduction

We consider the Cauchy problem associated with a scalar conservation law where the flux depends discontinuously on space

$$\begin{cases} \partial_t u + \partial_x \left[A(x,u)\right] = 0 \qquad x \in \mathbb{R},\ t \in \mathbb{R}_+, \\ u(0,x) = u_0(x) \in L^\infty(\mathbb{R}). \end{cases} \tag{5.1.1}$$

We propose a new method to prove a L^1 contraction principle for a class of solutions to this equation when the space dependence of the flux is discontinuous. As in most of the recent papers that deal with this subject [75, 2] our proof is based on Kruzkov's framework. But our idea is to adapt the definition of Kruzkov entropies to the discontinuous case and thus to avoid a special treatment of the interface.

When the space dependence of the flux is sufficiently smooth, this scalar equation is quite well known. In particular, Kruzkov's theory applies and provides existence and uniqueness of a weak solution to (5.1.1) that satisfies Kruzkov entropy inequalities - see [37, 83, 118].
We consider in this article the case where the flux is a discontinuous function of x, not necessarily of bounded variation. First existence results for such a problem were obtained through an analogy with 2×2 hyperbolic systems and the study of the related Riemann problem. Indeed if we assume that the flux is on the form $A(x,u) = f(\gamma(x),u)$, the equation (5.1.1) can be written as a system in (u,γ) by adding the trivial equation $\partial_t \gamma = 0$. In the 80's and considering particular forms of the flux, [123, 52, 53] established the global existence of a solution for the corresponding Cauchy problems by proving the convergence of different numerical methods. Later on existence results were extended to more general fluxes by using convergence of numerical schemes [81, 63, 125, 126, 82, 73, 72, 75, 25, 2] or regularization of the coefficients [80, 103, 117, 74].
Here we are interested in the problem of the uniqueness of solutions. First results about this topic were obtained in the middle of the 90's and different methods have been investigated. In [43], Diehl considered a flux on the form $A(x,u) = H(x)f(u)+(1-H(x))g(u)$ where $H(x)$ is the Heavyside function and proved existence and uniqueness locally in time by introducing a coupling condition Γ at the interface. In the same year, in [81], Klingenberg and Risebro considered a multiplicative flux $A(x,u) = k(x)f(u)$ such that f is a convex function that satisfies $f(0) = f(1) = 0$ and $k(x) \geq k > 0$ is a BV piecewise smooth function with a finite number of discontinuity points; they proved uniqueness for a solution that satisfies a wave entropy condition - see also [80] where the authors proved continuous dependence on the coefficient k and on the initial data for the same problem. In [63], Greenberg et al. considered a convex additive flux $A(x,u) = f(u) + a(x)$ where f is even and convex and a is piecewise constant and proved a contraction principle for the solution that they constructed by solving Riemann problems and studying interactions of waves in a right way. In [103] Ostrov

proposed an other approach : he proved uniqueness of a solution of the Hamilton Jacobi equation obtained as the limit of viscosity solutions for regularized coefficients cases. Then he concluded for (5.1.1) by using the equivalency between Hamilton Jacobi equations and scalar conservation laws. He extended uniqueness result to fluxes on the form $A(t,x,u) = f(k(t,x),u)$ where f is convex in u and satisfies a superlinear growth condition and k is bounded and discontinuous along a finite number of curves and is Lipschitz continuous away from these curves. Towers [125] came back to the multiplicative case and established a L^1 contraction principle for a class of solutions that satisfies Kruzkov type entropy inequalities - it means that the solution satisfies classical Kruzkov entropy inequalities away from the discontinuities of the flux and satisfies a geometric condition at the discontinuity points that can be interpreted as an interface entropy condition. Note that to give meaning to this new entropy condition he needed to assume some additional regularity conditions on the solution, namely that u is piecewise C^1 and possesses traces on the discontinuities of k. In this earlier work the flux was assumed to be convex in u, but this approach was further investigated by Karlsen, Risebro and Towers in [75] and the uniqueness result has been extended to non convex fluxes on the form $A(x,u) = f(k(x),u)$ where k is a piecewise C^1 BV function with a finite number of discontinuities and f is Lipschitz continuous in u and k and satisfies a given crossing condition. In that paper and in [117] the existence of traces for u is proved for particular additive / multiplicative fluxes but should be assumed in the general case. Very recently, Adimurthi, Jaffre and Veerappa Gowda [2] introduced an other interface entropy condition - still coupled with classical Kruzkov entropy inequalities away from the discontinuity - and proved also a L^1 contraction principle for this new class of solutions. They considered a Heavyside flux type where f and g have only one global minimum and no local minimum and assumed the existence of traces on the discontinuity. For particular fluxes and initial data it can be proved that the interface conditions in [75] and [2] do not select the same solution. We prove in Section 5 that our *interface condition free* method selects the solution derived from the interface condition of [2].

In [125, 117, 75, 2] the uniqueness proof is based on the use of classical Kruzkov entropies which leads to the following entropies inequalities

$$\partial_t|u-k| + \partial_x\left[(A(x,u)-A(x,k))\ \mathrm{sgn}(u-k)\right] + \mathrm{sgn}(u-k)\partial_x A(x,k) \leq 0. \qquad (5.1.2)$$

Thus an interface entropy condition has to be introduced by the authors to deal with the discontinuities of the flux and to give meaning to the last term of the left hand side. Here we propose to adapt the definition of Kruzkov entropies to the discontinuous case by introducing partially adapted Kruzkov entropies

$$E_\alpha(x,u) = |u - k_\alpha(x)|,$$

where $k_\alpha(x)$ satisfies

$$A(x,k_\alpha(x)) = \alpha.$$

This new definition allows us to remove the problematic term in the entropy inequalities (5.1.2) since we obtain

$$\partial_t|u-k_\alpha(x)|+\partial_x\left[(A(x,u)-A(x,k_\alpha(x)))\ \mathrm{sgn}(u-k_\alpha(x))\right]\leq 0. \tag{5.1.3}$$

Thus the interface does not need a special treatment and no interface entropy condition is needed. Uniqueness then follows from arguments very close to Kruzkov's original proof and the main difficulty is now to deal with the family(ies) of functions $k_\alpha(x)$. This new method allows us to remove the hypothesis about the traces of the solution on the discontinuities of the flux and the BV bounds on the space dependence of the flux and on the initial data. Also we can deal with an infinite number of discontinuity points and we do not need convexity assumptions or crossing conditions. However we need some other hypothesis on the flux - and more particurlarly on the u dependence of the flux - to be able to define our partially adapted Kruzkov entropies.

The outline of the paper is the following. In Section 2 we list the hypothesis on the flux and we comment them with some examples. In Section 3 we define the partially adapted Kruzkov entropies and in Section 4 we prove the L^1 contraction principle. Finally in Section 5, and for a particular flux, we study the selected solution and we compare it with the existing results [2, 75].

5.2 Hypothesis on the flux

In this work we assume the following hypothesis on the flux A

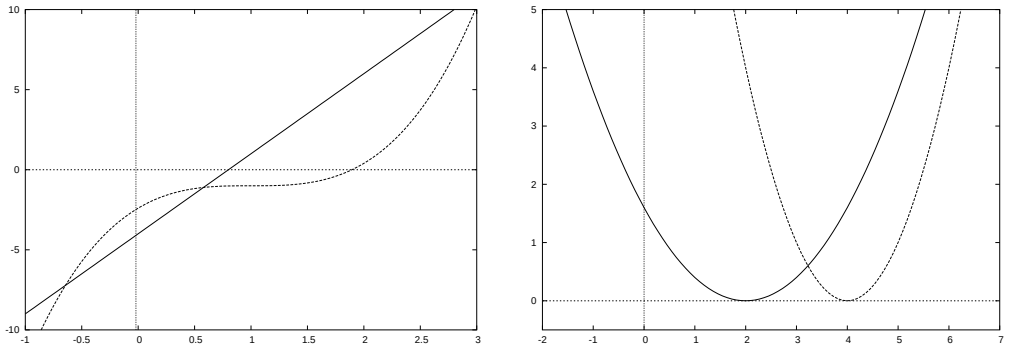
• (H1) $A(x,u)$ is continuous at all points of $\mathbb{R}\backslash\mathcal{N}\times\mathbb{R}$ where $\mathcal{N}$ is a closed zero measure set,

• (H2) $\exists(f,g)\in\left(C^0(\mathbb{R})\right)^2$ such that $\forall x\in\mathbb{R}\quad f(u)\leq|A(x,u)|\leq g(u)$. We assume that $|f(\pm\infty)|=+\infty$.

• (H3) For $x\in\mathbb{R}\backslash\mathcal{N}$, $A(x,.)$ is a locally Lipschitz one to one function from $\mathbb{R}$ to $\mathbb{R}$.

All along the article we also consider an alternative case by replacing the hypothesis (H3) by

• (H3') There is a function $u_M(x)$ from $\mathbb{R}$ to $\mathbb{R}$ such that for $x\in\mathbb{R}\backslash\mathcal{N}$, $A(x,.)$ is a locally Lipschitz one to one function from $[-\infty,u_M(x)]$ and $[u_M(x),+\infty]$ to $[0,+\infty]$ that satisfies $A(x,u_M(x))=0$.



(a) Hypothesis (H3) (b) Hypothesis (H3')

FIG. 5.2.1: Admissible fluxes (Heavyside type)

Two examples of Heavyside type fluxes $A(x,u) = H(x)f(u) + (1 - H(x))g(u)$ that satisfy hypothesis (H3) and (H3') are presented in Figure 5.2.1-(a) and Figure 5.2.1-(b) respectively. For the case of (H3), it is enough that f anf g are increasing one-to-one functions.

An example of application - with hypothesis (H3) - is the classical transport equation but with a (positive) discontinuous coefficient $S(x)$ in the flux

$$\partial_t u + \partial_x \left[S(x)u\right] = 0. \tag{5.2.1}$$

However notice that in this case our result is less general than those in [20].

The alternative case involving hypothesis (H3') is obviously related to discontinuous Burger-Hopf's equation - here also $S(x)$ is supposed to be positive

$$\partial_t u + \partial_x \left[S(x)u^2\right] = 0. \tag{5.2.2}$$

In both examples (5.2.1) and (5.2.2), hypothesis (H3) (resp. (H3')) is satisfied, hypothesis (H1) gives the admissible discontinuous form of S and hypothesis (H2) is equivalent to

$$\exists (m_S, M_S) \quad \text{for a.e. } x \in \mathbb{R} \quad 0 < m_S \leq S(x) \leq M_S < +\infty.$$

But hypothesis (H3') is also able to cover more general crossing convex fluxes on the form $A(x,u) = k_\pm(u - \beta_\pm)^2$ - see Figure 5.2.1-b ans Section 5.

The discontinuous flux case is concerned with numerous applications : sedimentation process, two phase flow in porous media, road traffic... Let us also mention that the discontinuous flux case has natural links with Saint-Venant models : the modelization

of blood flow with the Saint-Venant system exhibits the Young modulus of arteries - that can be discontinuous after chirurgical acts - as a coefficient in the pressure flux - see [48] ; for a stationary flow, the coupled transport equation is of the form $A(x,u) = a(x)u$ where $a(x)$ is the velocity of the flow, that can be discontinuous - see [8].

5.3 Partially Adapted Kruzkov Entropies

Given $\alpha \in \mathbb{R}$, an immediate consequence of hypothesis (H1) and (H3) is the existence and the uniqueness of a function k_α from $\mathbb{R}$ to $\mathbb{R}$ such that

$$A\left(x, k_\alpha(x)\right) = \alpha \qquad \text{for a.e. } x \in \mathbb{R}. \tag{5.3.1}$$

The alternative case (H3') leads to similar conclusions : given $\alpha \in [0, +\infty]$ and $x \in \mathbb{R}\backslash\mathcal{N}$, there are two unique real numbers $k_\alpha^+(x) \in [u_M(x), +\infty]$ and $k_\alpha^-(x) \in [-\infty, u_M(x)]$ such that

$$A\left(x, k_\alpha^\pm(x)\right) = \alpha. \tag{5.3.2}$$

In the following, when relations are valid under hypothesis (H3) or (H3'), and to avoid unuseful repetitions, $k_\alpha(x)$ will denote either $k_\alpha(x)$ or $k_\alpha^\pm(x)$.

These definitions allow us to introduce partially adapted Kruzkov entropies, which are a natural way to extend classical Kruzkov entropies to the discontinuous flux case.

Definition 5.3.1 *Let u (resp. v) $\in L^\infty([0,T]\times\mathbb{R}) \cap C^0([0,T], L^1_{loc}(\mathbb{R}))$. We say it is an entropy subsolution -resp. supersolution- of (5.1.1) if and only if for all $\alpha \in \mathbb{R}$ (or $\mathbb{R}^+$ under hypothesis (H3'))*

$$\partial_t (u - k_\alpha(x))_+ + \partial_x \left[\left(A(x,u) - A(x,k_\alpha(x)) \right) \operatorname{sgn}_+(u - k_\alpha(x)) \right] \leq 0, \tag{5.3.3}$$

-resp.

$$\partial_t (v - k_\alpha(x))_- + \partial_x \left[\left(A(x,v) - A(x,k_\alpha(x)) \right) \operatorname{sgn}_-(v - k_\alpha(x)) \right] \geq 0. \tag{5.3.4}$$

Our motivation to introduce these adapted entropies comes from the contraction property which still holds true under the form (5.1.3). It is natural to state Kruzkov entropy with the steady state solution

$$\frac{\partial}{\partial_t} k_\alpha(x) + \frac{\partial}{\partial_x} A\left(x, k_\alpha(x)\right) = 0, \tag{5.3.5}$$

but not with constants.

In a future work, we will prove that this condition can be derived from the vanishing viscosity method and thus is a natural entropy condition.

5.4 Uniqueness Theorem

Theorem 5.4.1 *Let u and $v \in L^\infty([0,T],\mathbb{R}) \cap C^0([0,T],L^1_{loc}(\mathbb{R}))$ be respectively an entropy sub- and supersolution to the initial value problem (5.1.1) with initial data $u_0, v_0 \in L^\infty(\mathbb{R})$. Assume hypothesis (H1)-(H2)-(H3) or (H1)-(H2)-(H3') on the flux are true. Then for a.e. $t \in [0,T]$*

$$\int_a^b (u(x,t)-v(x,t))_+ \, dx \leq \int_{a-Mt}^{b+Mt} (u_0(x)-v_0(x))_+ \, dx. \tag{5.4.1}$$

Proof of Theorem 5.4.1. We denote by $Q = \mathbb{R}\times[0,T]$. Proving the theorem is equivalent to establishing the following inequality for all $\phi \in C_0^\infty(Q)$ - see [37, 118]

$$\begin{aligned}\int_Q (u(t,x)-v(t,x))_+ \partial_t\phi dxdt \\ + \int_Q \big(A(x,u(t,x))-A(x,v(t,x))\big)\mathrm{sgn}_+(u(t,x)-v(t,x))\partial_x\phi dxdt \\ + \int_{\mathbb{R}} (u_0(x)-v_0(x))_+\phi(x,0)dx \geq 0.\end{aligned} \tag{5.4.2}$$

Since $A(.,u)$ is continuous for $x \in \mathbb{R}\backslash\mathcal{N}$, we can define for a.e. $(x,y,s,t) \in Q^2$ two functions $\tilde{u}(t,x,y)$ and $\tilde{v}(s,y,x)$ from $[0,T]\times\mathbb{R}^2$ to $\mathbb{R}$ such that

$$\begin{aligned} A(y,\tilde{u}(t,x,y)) &= A(x,u(t,x)), \\ A(x,\tilde{v}(s,y,x)) &= A(y,v(s,y)). \end{aligned} \tag{5.4.3}$$

According to the notations of Section 5.3, and under hypothesis (H3), it is equivalent to

$$\begin{aligned} \tilde{u}(t,x,y) &= k_{A(x,u(t,x))}(y), \\ \tilde{v}(s,y,x) &= k_{A(y,v(s,y))}(x). \end{aligned} \tag{5.4.4}$$

In the case (H3'), we impose also that

$$\begin{aligned} \mathrm{sgn}(\tilde{u}(t,x,y)-u_M(y)) &= \mathrm{sgn}(u(t,x)-u_M(x)), \\ \mathrm{sgn}(\tilde{v}(s,y,x)-u_M(x)) &= \mathrm{sgn}(v(s,y)-u_M(y)). \end{aligned}$$

We denote the new sign function by $\widetilde{\mathrm{sgn}}(x,u) = \mathrm{sgn}(u-u_M(x))$. According to the previous notations it means that

$$\begin{aligned} \tilde{u}(t,x,y) &= k^+_{A(x,u(t,x))}(y)\ \widetilde{\mathrm{sgn}}_+(x,u(t,x)) + k^-_{A(x,u(t,x))}(y)\ \widetilde{\mathrm{sgn}}_-(x,u(t,x)), \\ \tilde{v}(s,y,x) &= k^+_{A(y,v(s,y))}(x)\ \widetilde{\mathrm{sgn}}_+(y,v(s,y)) + k^-_{A(y,v(s,y))}(x)\ \widetilde{\mathrm{sgn}}_-(y,v(s,y)). \end{aligned} \tag{5.4.5}$$

Now we write the entropy condition (5.3.3) for $u(t,x)$ with $\alpha = A(y, v(s,y))$

$$\begin{aligned}&\partial_t (u(t,x) - k_{A(y,v(s,y))}(x))_+ \\ &\quad +\partial_x \left[\big(A(x,u(t,x)) - A(x, k_{A(y,v(s,y))}(x))\big) \operatorname{sgn}_+\big(u(t,x) - k_{A(y,v(s,y))}(x)\big)\right] \leq 0.\end{aligned}$$

It leads to

$$\begin{aligned}&\partial_t \left(u(t,x) - \tilde{v}(s,y,x)\right)_+ \\ &\quad +\partial_x \left[\big(A(x,u(t,x)) - A(y, v(s,y))\big) \operatorname{sgn}_+\big(u(t,x) - \tilde{v}(s,y,x)\big)\right] \leq 0. \quad (5.4.6)\end{aligned}$$

We obtain a similar inequality when we write partially adapted Kruzkov entropy relation (5.3.4) for $v(s,y)$ with $\alpha = A(x, u(t,x))$

$$\begin{aligned}&\partial_s \left(v(s,y) - \tilde{u}(t,x,y)\right)_- \\ &\quad +\partial_y \left[\big(A(y,v(s,y)) - A(x, u(t,x))\big) \operatorname{sgn}_-\big(v(s,y) - \tilde{u}(t,x,y)\big)\right] \leq 0. \quad (5.4.7)\end{aligned}$$

Now for $\epsilon > 0, \eta > 0$, we introduce two positive functions $\rho, \xi \in C_0^\infty(\mathbb{R})$ such that

$$\int_{\mathbb{R}} \rho(z)dz = \int_{\mathbb{R}} \xi(z)dz = 1, \quad (5.4.8)$$

and, for $\eta, \epsilon > 0$, we define two families of functions $\rho_\epsilon, \xi_\eta \in C_0^\infty(\mathbb{R})$ such that

$$\xi_\eta(z) = \frac{1}{\eta}\xi(\frac{z}{\eta}), \qquad \rho_\epsilon(z) = \frac{1}{\epsilon}\rho(\frac{z}{\epsilon}),$$

which provide two approximations of the Dirac mass δ_0. Moreover we impose that the support of ρ is included in $]-2,-1[$. Then we add (5.4.7) to (5.4.6) and we integrate in y, s, x, t against a function $\Phi_{\eta\epsilon}(x,t,y,s) \in C_0^\infty(Q^2)$ with $\Phi_{\eta\epsilon}(x,t,y,s) =$

$\phi(x,t)\rho_\epsilon(t-s)\xi_\eta(x-y)$. Finally we obtain

$$
\begin{aligned}
&\text{(I)} \quad \int_{Q^2}(u(t,x)-\tilde{v}(s,y,x))_+\partial_t\phi(x,t)\rho_\epsilon(t-s)\xi_\eta(x-y)dydsdxdt\\
&\text{(II)} \quad -\int_{Q^2}((u(t,x)-\tilde{v}(s,y,x))_+-(v(s,y)-\tilde{u}(t,x,y))_-)\\
&\qquad\qquad \phi(x,t)\rho_\epsilon{}'(t-s)\xi_\eta(x-y)dydsdxdt\\
&\text{(III)} \quad +\int_{Q^2}(A(x,u(t,x))-A(x,\tilde{v}(s,y,x)))\partial_x\phi(x,t)\rho_\epsilon(t-s)\xi_\eta(x-y)\\
&\qquad\qquad (\mathrm{sgn}_+(u(t,x)-\tilde{v}(s,y,x))dydsdxdt\\
&\text{(IV)} \quad -\int_{Q^2}(A(x,u(t,x))-A(y,v(s,y)))\phi(x,t)\rho_\epsilon(t-s)\xi_\eta{}'(x-y)\\
&\qquad (\mathrm{sgn}_+(u(t,x)-\tilde{v}(s,y,x))+\mathrm{sgn}_-(v(s,y)-\tilde{u}(t,x,y)))dydsdxdt\\
&\text{(V)} \quad +\int_{Q\times\mathbb{R}}(u_0(x)-\tilde{v}(s,y,x))_+\phi(x,0)\rho_\epsilon(-s)\xi_\eta(x-y)dydsdx\\
&\text{(VI)} \quad +\int_{Q\times\mathbb{R}}(v_0(y)-\tilde{u}(t,x,y))_-\phi(x,t)\rho_\epsilon(t)\xi_\eta(x-y)dydxdt\geq 0.
\end{aligned}
\tag{5.4.9}
$$

The main difference with classical Kruzkov's proof - see [83, 37, 118] - is the Terms (II) and (IV). Notice that the derivatives that appear in these terms are derivatives of functions that tend to Dirac masses. We will prove that Term (IV) is equal to zero for all (η,ϵ). For Term (II), the main idea of the proof is to first consider the limit when η tends to zero - with a fixed ϵ - and to show that this limit is equal to zero for all ϵ.

Let us first establish that $\tilde{v}$ and $\tilde{u}$ belong to $L^\infty([0,T]\times\mathbb{R}^2)$. We give the proof for $\tilde{v}$. By hypothesis $v\in L^\infty([0,T]\times\mathbb{R})$. It follows from (H2) that for a.e. s,y

$$|A(y,v(s,y))|\leq \max_{-\|v\|_{L^\infty}\leq\sigma\leq\|v\|_{L^\infty}} g(\sigma)=M.$$

Since $A(x,\tilde{v}(s,y,x))=A(y,v(s,y))$ and using (H2) again we obtain

$$f(\tilde{v}(s,y,x))\leq M.$$

Finally, since $|f(\pm\infty)|=+\infty$, we conclude that $\tilde{v}\in L^\infty\left([0,T]\times\mathbb{R}^2\right)$.

Term (IV). We now treat the part involving the sign functions. We prove that

$$\mathrm{sgn}(u(t,x)-\tilde{v}(s,y,x))=\mathrm{sgn}(\tilde{u}(t,x,y)-v(s,y)) \quad \text{for a.e. } t,x,s,y\in Q^2. \tag{5.4.10}$$

By definition of $\tilde{u}$ and $\tilde{v}$ in (5.4.3), we have for a.e. $t,x,s,y\in Q^2$

$$A(x,u(t,x))-A(x,\tilde{v}(s,y,x))=A(y,\tilde{u}(t,x,y))-A(y,v(s,y)). \tag{5.4.11}$$

Under the hypothesis (H3), $A(x,.)$ is monotonous and therefore (5.4.11) implies the result (5.4.10). Under hypothesis (H3'), it follows from (5.4.5) that

$$\widetilde{\mathrm{sgn}}(x,u) - \widetilde{\mathrm{sgn}}(y,\tilde{u}) = \widetilde{\mathrm{sgn}}(y,v) - \widetilde{\mathrm{sgn}}(x,\tilde{v}) = 0. \tag{5.4.12}$$

The case $\widetilde{\mathrm{sgn}}(x,u) = \widetilde{\mathrm{sgn}}(y,v)$ reduces to hypothesis (H3) since $A(x,.)$ is monotonous on each semi-space $[-\infty, u_M(x)]$ and $[u_M(x), +\infty]$. If $\widetilde{\mathrm{sgn}}(x,u) \neq \widetilde{\mathrm{sgn}}(y,v)$, the result (5.4.10) is an immediate consequence of (5.4.12).
Then from (5.4.10) we deduce that for a.e. $t,x,s,y \in Q^2$

$$(A(x,u) - A(y,v))\phi(x,t)\rho_\epsilon(t-s)\xi_\eta{}'(x-y)(\mathrm{sgn}_+(u-\tilde{v}) + \mathrm{sgn}_-(v-\tilde{u})) = 0.$$

Since $u, v, \tilde{u}, \tilde{v} \in L^\infty$ and, for $\eta, \epsilon > 0$, $\phi, \xi_\eta, \rho_\epsilon \in C_0^\infty$, we can apply Lebesgue's theorem and conclude that, for every $\eta, \epsilon > 0$, Term (IV) is equal to zero.

Term(II). We first observe that

$$\begin{aligned} &|(u(t,x) - \tilde{v}(s,y,x))_+ - (v(s,y) - \tilde{u}(t,x,y))_-| \\ &\qquad \leq |u(t,x) - \tilde{u}(t,x,y)| + |v(s,y) - \tilde{v}(s,y,x)|, \end{aligned}$$

and then it is sufficient to prove that

$$\int_{\mathbb{R}} |v(s,y) - \tilde{v}(s,y,x)|\xi_\eta(x-y)dx \xrightarrow[\eta\to 0]{} 0 \quad \text{for a.e. } s,y \in Q, \tag{5.4.13}$$

to establish that, for every $\epsilon > 0$, the limit in η of Term (II) is equal to zero. Indeed, once we have (5.4.13), and since, for $\epsilon > 0$, all functions are bounded, we can apply dominated convergence to conclude that the integral in s,y,t,x tends to zero. The results for $|u(t,x) - \tilde{u}(t,x,y)|$ is obviously similar. Thus the absolute value of Term (II) is bounded by an expression that vanishes with η.
In order to prove (5.4.13) we now establish that

$$\tilde{v}(s,y,x) \xrightarrow[x\to y]{} \tilde{v}(s,y,y) = v(s,y) \quad \text{for a.e. } s,y \in Q. \tag{5.4.14}$$

Here we use the assumtion (H3) or (H3'), i.e. that A is continuous outside a negligeable set. Then for $y \in \mathbb{R}\backslash\mathcal{N}$

$$A(x,\tilde{v}(s,y,y)) \xrightarrow[x\to y]{} A(y,\tilde{v}(s,y,y)).$$

On an other hand, we have by hypothesis

$$A(y,\tilde{v}(s,y,y)) = A(y,v(s,y)) = A(x,\tilde{v}(s,y,x)) \quad \text{for a.e. } s,y,x \in Q \times \mathbb{R}.$$

Thus

$$A(x,\tilde{v}(s,y,x)) - A(x,\tilde{v}(s,y,y)) \xrightarrow[x\to y]{} 0,$$

and (5.4.14) is a consequence of the fact that $A(x,.)$ is a one to one function.
Now we claim that the integral in (5.4.13) can be written

$$\int_{\mathbb{R}} |v(s,y) - \tilde{v}(s,y,y+\eta z)|\xi(z)dz,$$

and then, since all functions are bounded and since the support of ξ is bounded also, we can use the result (5.4.14) and dominated convergence to conclude (5.4.13).

Terms (I) and (III) are more classical. The only key point is that we must deal with "tilda functions", but we will use the result (5.4.13) to recover classical Kruzkov's proof.

Term (I). We first observe that

$$\begin{aligned}\Big| \int_{Q^2} (u(t,x) - \tilde{v}(s,y,x))_+ \partial_t\phi(x,t)\rho_\epsilon(t-s)\xi_\eta(x-y)dydsdxdt \\ - \int_{Q^2} (u(t,x) - v(s,y))_+ \partial_t\phi(x,t)\rho_\epsilon(t-s)\xi_\eta(x-y)dydsdxdt\Big| \\ \leq \int_{Q^2} |\tilde{v}(s,y,x) - v(s,y)|\, \partial_t\phi(x,t)\rho_\epsilon(t-s)\xi_\eta(x-y)dydsdxdt,\end{aligned}$$

and we use the previous computation - see (5.4.13) - to claim that the limit in η, ϵ of term (I) is the same as the limit of

$$\int_{Q^2} (u(t,x) - v(s,y))_+ \partial_t\phi(x,t)\rho_\epsilon(t-s)\xi_\eta(x-y)dydsdxdt.$$

Now we claim it is enough to prove that

$$\int_{Q^2} |v(t,x) - v(s,y)|\partial_t\phi(x,t)\rho_\epsilon(t-s)\xi_\eta(x-y)dydsdxdt \xrightarrow[\eta\to 0,\ \epsilon\to 0]{} 0, \qquad (5.4.15)$$

to conclude that the limit of Term (I), when η and ϵ tend to zero, is

$$\int_Q (u(t,x) - v(t,x))_+ \partial_t\phi(x,t)dtdx.$$

The proof of (5.4.15) is also a crucial step of the uniqueness proof when the flux does not depend on the space variable and we refer to [37, 118] for the details.

Term (III). Finally we consider the term that contains the fluxes. We define

$$G(x,u,w) = (A(x,u) - A(x,w))\mathrm{sgn}(u-w).$$

Hypothesis (H2) implies that G is a locally Lipschitz function of the third variable. Since $\tilde{v} \in L^\infty([0,T] \times \mathbb{R} \times \mathbb{R})$, it follows that

$$\begin{aligned}|G(x,u(t,x),\tilde{v}(s,y,x)) - G(x,u(t,x),\tilde{v}(s,y,y))| &\leq C|\tilde{v}(s,y,x) - \tilde{v}(s,y,y)| \\ &= C|\tilde{v}(s,y,x) - v(s,y)|,\end{aligned}$$

and it follows from (5.4.13) that the limit of Term (III) is the same as the limit of

$$\int_{Q^2} (A(x,u(t,x)) - A(x,v(s,y)))\ \mathrm{sgn}(u(t,x) - v(s,y)) \\ \partial_x\phi(x,t)\rho_\epsilon(t-s)\xi_\eta(x-y)dydsdxdt.$$

We use a second time the Lipschitz property on G. Now $v \in L^\infty([0,T] \times \mathbb{R})$ and

$$|G(x,u(t,x),v(s,y)) - G(x,u(t,x),v(t,x))| \leq C|v(s,y) - v(t,x)|.$$

Thus it is sufficient to prove that

$$\int_{Q^2} |v(t,x) - v(s,y)|\partial_x\phi(x,t)\rho_\epsilon(t-s)\xi_\eta(x-y)dydsdxdt \xrightarrow[\eta\to 0,\ \epsilon\to 0]{} 0, \tag{5.4.16}$$

to conclude that Term (III) tends to

$$\int_Q \big(A(x,u(t,x)) - A(x,v(t,x))\big)\mathrm{sgn}_+(u(t,x) - v(t,x))\partial_x\phi dxdt.$$

The integral in (5.4.16) appears in classical Kruzkov's proof. It is very similar to the one in (5.4.15) and same arguments lead to the result (5.4.16).

The computation of Terms (V) and (VI) are classical. Thanks to the hypothesis on the support of ρ, Term (VI) is equal to zero. For Term (V), we claim that the result (5.4.13) allows us to consider only

$$\int_{Q\times\mathbb{R}} (u_0(x) - v(s,y))_+\phi(x,0)\rho_\epsilon(-s)\xi_\eta(x-y)dydsdx \to \int_{\mathbb{R}} (u_0(x) - v_0(x))_+\phi(x,0)dx,$$

and then the end of the proof is standard - see [37, 118]. □

5.5 Application : Discontinuous convex flux

We noticed at the end of the introduction that different interface conditions can select different unique solutions. Since our method does not require an interface condition it can be used to discriminate the existing interface conditions, at least for the cases where our theory can be applied.

Here we propose to study a particular Heavyside type flux $A(x,u) = H(x)f_+(u) + (1 - H(x))f_-(u)$ where $f_\pm(u) = k_\pm(u - \beta_\pm)^2$ for which the interface conditions in [75, 2] do not select the same solution. This convex flux satisfies hypothesis (H3') and thus we can exhibit the solution that is selected by our method.
We study the Riemann problem associated with the very simple discontinuous convex flux of Heavyside type

$$A(x,u) = H(x)\frac{u^2}{2} + (1 - H(x))\frac{(u-1)^2}{2}, \tag{5.5.1}$$

and with the constant initial data

$$u_0(x) = \frac{1}{2}. \tag{5.5.2}$$

It is obvious that

$$u(t,x) = \frac{1}{2}, \tag{5.5.3}$$

is a weak solution of the Riemann problem (5.5.1)-(5.5.2). But more generally, for $u_i \in [0, \frac{1}{2}]$ the function defined by

$$u(t,x) = \begin{cases} \dfrac{1}{2} & x \le -\dfrac{t}{2} \\ 1 + \dfrac{x}{t} & -\dfrac{t}{2} < x \le -u_i t \\ 1 - u_i & -u_i t < x \le 0 \\ u_i & 0 < x \le u_i t \\ \dfrac{x}{t} & u_i t < x \le \dfrac{t}{2} \\ \dfrac{1}{2} & \dfrac{t}{2} < x \end{cases}, \tag{5.5.4}$$

is also a weak solution of the Riemann problem (5.5.1)-(5.5.2).
Now let us apply our entropy theory to this case. It follows from the definition (5.3.2) that

$$k_\alpha^\pm(x) = \pm\sqrt{2\alpha} + 1 - H(x). \tag{5.5.5}$$

Thus the entropy inequality (5.3.3) becomes

$$\begin{aligned} &\partial_t |u - (\pm\sqrt{2\alpha} + 1 - H(x))| \\ &+\partial_x \left[(H(x)\frac{u^2}{2} + (1 - H(x))\frac{(u-1)^2}{2} - \alpha)\ \mathrm{sgn}(u - (\pm\sqrt{2\alpha} + 1 - H(x))) \right] \le 0. \end{aligned} \tag{5.5.6}$$

Let us choose a solution on the form (5.5.4) with $u_i = 0$

$$u(t,x) = \begin{cases} \frac{1}{2} & x \leq -\frac{t}{2} \\ 1 + \frac{x}{t} & -\frac{t}{2} < x \leq 0 \\ \frac{x}{t} & 0 < x \leq \frac{t}{2} \\ \frac{1}{2} & \frac{t}{2} < x \end{cases} . \tag{5.5.7}$$

For (x,t) such that $2x \leq -t$ or $2x > t$ the entropy inequality (5.5.6) is obviously satisfied. Now for (x,t) such that $2x \in]-t,t]$, the solution can be denoted by

$$u(t,x) = \frac{x}{t} + 1 - H(x).$$

The entropy inequality (5.5.6) becomes

$$\partial_t \left| \frac{x}{t} \pm \sqrt{2\alpha} \right| + \frac{1}{2}\partial_x \left[\left(\left(\frac{x}{t} \right)^2 - 2\alpha \right) \operatorname{sgn}(\frac{x}{t} \pm \sqrt{2\alpha}) \right] \leq 0,$$

and is also obviously satisfied. Thus the solution (5.5.7) is the entropy solution of the Riemann problem (5.5.1)-(5.5.2).

Notice that for this particular crossing convex flux, the interface condition in [75] selects the constant solution (5.5.3) whereas the interface condition in [2] selects the solution (5.5.7).

Chapitre 6

A multilayer Saint-Venant model

Le travail présenté dans ce chapitre a été accepté pour publication dans *Discrete and Continuous Dynamical Systems - Serie B*.

6.1 Introduction

In this paper we are interested in modeling the so-called shallow water flows. It covers a very large range of applications in the domain of the geophysical flows - rivers, lakes, costal areas, oceans, atmosphere, avalanches... In order to provide a new modelization step between the complexity of the full Navier-Stokes equations for incompressible flows with free surface and the loss of generality of the classical Saint-Venant system, we introduce in this paper a new multilayer Saint-Venant type model which consists in a set of coupled Saint-Venant systems

$$\frac{\partial h_\alpha}{\partial t} + \frac{\partial h_\alpha U_\alpha}{\partial x} = 0, \tag{6.1.1}$$

$$\begin{aligned}\frac{\partial h_\alpha U_\alpha}{\partial t} + \frac{\partial}{\partial x}\left(h_\alpha U_\alpha^2 + g\frac{h_\alpha \sum_{\beta=1}^{M} h_\beta}{2}\right) \\ = \frac{g\left(\sum_{\beta=1}^{M} h_\beta\right)^2}{2}\frac{\partial}{\partial x}\frac{h_\alpha}{\sum_{\beta=1}^{M} h_\beta} + 2\mu\frac{U_{\alpha+1}-U_\alpha}{h_{\alpha+1}+h_\alpha} - 2\mu\frac{U_\alpha - U_{\alpha-1}}{h_\alpha + h_{\alpha-1}}.\end{aligned} \tag{6.1.2}$$

where $(h_\alpha, h_\alpha U_\alpha)(t,x)$ is the vector of the conservative variables - water height and discharge, thus $U_\alpha(t,x)$ is the velocity - and where the subscript α is related to the considered layer - thus $\alpha \pm 1$ are related to the neighbouring layers, above and below. M is the total number of layers. μ denotes the viscosity coefficient.

The main interest of such a formulation is to access to the vertical profile of the horizontal velocity while preserving the computational efficiency of the Saint-Venant model. Furthermore the model verifies some stability properties - existence of an energy, positivity of the water heights - and some physical properties - agreement with the classical Saint-Venant solution for compatible initial conditions, preservation of stationary states. One of the goals of the present paper is also to extend these properties to numerical simulation.

The classical Saint-Venant system [115] is a well known approximation of the free surface Navier-Stokes equations for shallow water flows, which has been widely validated for various geophysical flows, such as rivers or coastal areas [67], or even ocean and atmosphere dynamic or avalanches flows [22] when completed with appropriate terms. The derivation of the Saint-Venant system from the Navier-Stokes equations for shallow incompressible flows with a free moving boundary is now classical when the viscosity is neglected [130]. But this does not allow to justify that the right jumps - that appear in dam breaks or hydraulic jumps - are those obtained when using the momentum - and not the velocity - as the conservative variable. In [50], Gerbeau and Perthame propose a full derivation based on an asymptotic analysis of the dimensionless Navier-Stokes

equations. In particular they study the influence of the viscous term and its relation with the friction term. Then the classical Saint-Venant system with a friction term turns out to be a zeroth order approximation of the Navier-Stokes equations for incompressible flows with a free moving boundary under the shallow water assumption. They also derive a viscous Saint-Venant system through a first order investigation. In a recent work [45], Saleri and Ferrari extended this approach to the 2D case including a slowly varying topography and an atmospheric pressure term.

Direct free surface Navier-Stokes computations are possible [68] but are still expensive and require sophisticated algorithms. It is necessary to adapt a 3D mesh to the movement of the free surface and to the variation of the bottom topography in sedimentation cases. It requires also to deal with the incompressibility condition. The Saint-Venant computations are much more efficient since they deal with unvarying 2D meshes. But the background of this computational simplification is also the main limitation of the Saint-Venant approach : it uses only integrated quantities. The classical Saint-Venant system is related to a "motion by slice" approximation, *i.e.* the horizontal velocity does not depend upon the vertical coordinate. Therefore some information is lost and the practical range of validity of the system is sometimes very limited : it fails to reproduce the correct vertical profile of the horizontal velocity - and even the correct averaged horizontal velocity - when the friction on the bottom is not small enough, also it is not able to reproduce such phenomena as the recirculation due to the surface wind in a closed lake.
For these reasons, intermediate models have been further used. Most of them are *simplified Navier-Stokes models*. In [30, 46, 100], Quarteroni, Saleri *et al.* consider this set of approximations of the 3D Navier-Stokes equations and they derive different approximate models. In particular, in [46, 100], they work on the so-called hydrostatic approximation in a way that is a first step toward the Saint-Venant model since they consider the integrated continuity equation to rise the water height as an explicit variable. Nevertheless the background of the modelization is still the classical Navier-Stokes formalism, at least for the horizontal momentum equation : unknowns are velocities and the continuous problem is considered on a 3D domain. The explicitation of the 2D+1D form of the problem is performed in a second step through the choice of the vertical discretization : the domain is divided into several layers on which the same 2D finite element approximation is used to describe the horizontal velocity. The layers can be fixed horizontal strata as in [46, 100] or based on the so-called sigma transformation as in [68].

Our approach can be seen as a *augmented Saint-Venant model*, i.e. a discretization of the 3D Navier-Stokes equations where the vertical grid (h_α) is evolved "à la Saint-Venant" and some viscous effects are kept. In this it goes one step further in the analysis of [50] and as for the classical Saint-Venant system we exhibit an energy and we characterize some equilibrium states. As one can see, the set of Saint-Venant sytems

(6.1.1)-(6.1.2), being 2D, is able to preserve computational efficiency. The conservative form of the equations on water heights (h_α) leads to extend finite volume schemes to (6.1.1)-(6.1.2). This motivates our choice of writing (6.1.2) as an usual conservative left hand side plus non conservative corrections - on the right hand side. Indeed the challenge here would be to perform an algorithm able to keep (i) stability - $h_\alpha \geq 0$ and entropy inequality (ii) conservation of heights and total momentum $\sum h_\alpha u_\alpha$ when $\nu = 0$ (iii) the particular states $h_\alpha = h/M$, $u_\alpha = u$ gives a consistent approximation of the classical monolayer Saint-Venant system - when $\nu = 0$ (iv) the steady states of a lake at rest are preserved. Our work provides a first step in this direction and can be related to a work of Saleri and Lazzaroni [86] motivated by the same applications. But the approaches are quite different since [86] presents non conservative Saint-Venant systems discretized through a finite element framework. In [27, 28] Pares *et al.* present a bi-fluid shallow water problem and the system they consider presents also common parts with the two layers version of (6.1.1)-(6.1.2). As in [86] they consider a non conservative version and they exhibit that the system is no more hyperbolic when the densities of the fluids are too close. We prove in the paper that the left hand side of the system (6.1.1)-(6.1.2) is hyperbolic which is another motivation for our formulation.

The outline of the paper is as follows. In Section 2 we apply the shallow water scaling to the free surface Navier-Stokes equations and we present and analyse two hydrostatic models - as indicated in Figure 6.1.1. Then in Section 3 we derive the multilayer Saint-Venant model (6.1.1)-(6.1.2) and we discuss its essential properties. In Section 4 we establish the discrete version of the model using the finite volume framework. In Section 5 we present some numerical results and comparisons between the monolayer Saint-Venant and Navier-Stokes systems which validate the model. Finally in Section 6 we conclude and present some perspectives.

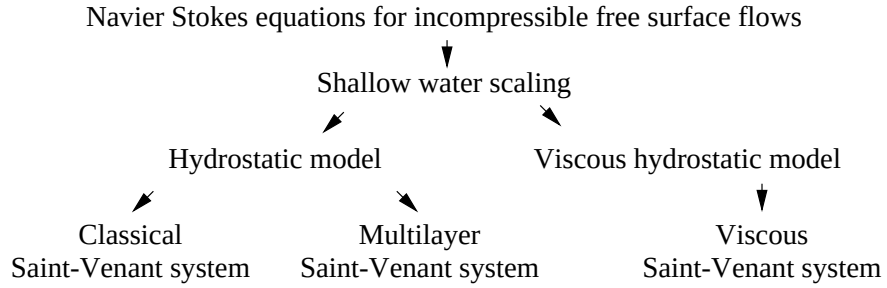


FIG. 6.1.1: From Navier-Stokes equations to Saint-Venant type models

6.2 Navier-Stokes equations and hydrostatic approximations

Following [50] we present here the first steps of the derivation of the Saint-Venant type models from the Navier-Stokes equations, *i.e.* the derivation of the simplified Navier-Stokes models - see Figure 6.1.1. We introduce a new focus on the question of the dissipation of the energy, whcich is an essential property of the Navier-Stokes equations. In particular it allows us to discriminate an *energy compatible* form of the viscous hydrostatic model introduced in Subsection 6.2.1. Moreover we clearly distinguish the basic effects of the shallow water assumption - Subsection 6.2.1 - from the ones of further assumptions about viscosity and friction - Subsection 6.2.2.
We consider the classical free surface Navier-Stokes equations

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0, \tag{6.2.1}$$

$$\frac{\partial u}{\partial t} + \frac{\partial u^2}{\partial x} + \frac{\partial uw}{\partial z} + \frac{\partial p}{\partial x} = 2\mu\frac{\partial^2 u}{\partial x^2} + \mu\,\frac{\partial^2 u}{\partial z^2} + \mu\frac{\partial^2 w}{\partial x\,\partial z}, \tag{6.2.2}$$

$$\frac{\partial w}{\partial t} + \frac{\partial uw}{\partial x} + \frac{\partial w^2}{\partial z} + \frac{\partial p}{\partial z} = -g + \mu\,\frac{\partial^2 w}{\partial x^2} + \mu\frac{\partial^2 u}{\partial x\,\partial z} + 2\mu\frac{\partial^2 w}{\partial z^2}, \tag{6.2.3}$$

with

$$t > 0, \quad x \in \mathbb{R}, \quad 0 \le z \le h(t,x),$$

where $u(t,x,z)$ is the horizontal velocity, $w(t,x,z)$ is the vertical velocity, $p(t,x,z)$ the pressure and where $h(t,x)$ denotes the water height, g the gravity and μ the viscosity. Notice that for simplicity we consider the flat bottom case. In [45] Saleri and Ferrari extended the analysis of Gerbeau and Perthame [50] concerning the classical Saint-Venant model to a slowly varying bottom. The present work does not introduce new contribution on this particular topic and the same extension can be done for the multilayer model that we introduce in Section 6.3. Notice also that we assume that the free surface is a function of (x,t).
On the bottom we consider a no penetration condition and we estimate the friction through a coefficient κ and a Navier condition

$$w(t,x,0) = 0, \quad \mu\,\frac{\partial u}{\partial z}(t,x,0) = \kappa u(t,x,0). \tag{6.2.4}$$

On the free surface we consider a no stress condition

$$\mu\,\frac{\partial u}{\partial z} + p\,\frac{\partial h}{\partial x} - 2\mu\frac{\partial u}{\partial x}\,\frac{\partial h}{\partial x} + \mu\frac{\partial w}{\partial x} = 0 \quad \text{on} \quad z = h(t,x), \tag{6.2.5}$$

$$p - 2\mu\,\frac{\partial w}{\partial z} + \mu\frac{\partial u}{\partial z}\,\frac{\partial h}{\partial x} + \mu\frac{\partial w}{\partial x}\,\frac{\partial h}{\partial x} = 0 \quad \text{on} \quad z = h(t,x), \tag{6.2.6}$$

and the kinematic boundary condition

$$\frac{\partial h}{\partial t} + u(t,x,z=h(t,x))\frac{\partial h}{\partial x} - w(t,x,z=h(t,x)) = 0. \tag{6.2.7}$$

We recall that (6.2.1), (6.2.4) and (6.2.7) lead to the following mass equation

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}\int_0^h u dz = 0.$$

We recall also that a fundamental stability property is related to the fact that the Navier-Stokes equations admit an energy

$$e_{ns}(t,x,z) = e_c(t,x,z) + e_p(t,x,z) = \frac{u^2+w^2}{2}(t,x,z) + gz.$$

and that the following equation holds

$$\begin{aligned}\frac{\partial}{\partial t}\int_0^h e_{ns}\, dz \;+\; \frac{\partial}{\partial x}\int_0^h \left(u(e_{ns}+p) - \mu\,\left(2u\,\frac{\partial u}{\partial x} + w\,\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right)\right)\right) dz \\ = -\,\kappa u^2(t,x,0) \;-\; 2\mu\int_0^h \left[\left(\frac{\partial u}{\partial x}\right)^2 + \frac{1}{2}\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial z}\right)^2\right] dz.\end{aligned}$$

For further analysis of the full Navier-Stokes equations we refer to [94].

We are now interested in the shallow water flows. As usual we introduce two characteristic dimensions H and L in the vertical and horizontal directions respectively. The shallow water flows are then characterized by the fact that H is very small compared with L. Thus we can make the so-called shallow water assumption, i.e. we introduce a "small parameter" $\epsilon = \frac{H}{L}$. Then after rescaling we can write (6.2.1)-(6.2.3) as a dimensionless Navier-Stokes system

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0, \tag{6.2.8}$$

$$\frac{\partial u}{\partial t} + \frac{\partial u^2}{\partial x} + \frac{\partial uw}{\partial z} + \frac{\partial p}{\partial x} = 2\nu\frac{\partial^2 u}{\partial x^2} + \frac{\nu}{\epsilon^2}\frac{\partial^2 u}{\partial z^2} + \nu\frac{\partial^2 w}{\partial x\,\partial z}, \tag{6.2.9}$$

$$\epsilon^2\left(\frac{\partial w}{\partial t} + \frac{\partial uw}{\partial x} + \frac{\partial w^2}{\partial z}\right) + \frac{\partial p}{\partial z} = -g + \epsilon^2\nu\,\frac{\partial^2 w}{\partial x^2} + \nu\frac{\partial^2 u}{\partial x\,\partial z} + 2\nu\frac{\partial^2 w}{\partial z^2}. \tag{6.2.10}$$

Notice that $\nu = \mu/(UL)$ is the dimensionless form of the viscosity coefficient and that here and in all the dimensionless equations g does not denote the gravity but the Froude number $g/(UL)$.
The associated boundary conditions on the bottom (6.2.4) are rescaled as

$$w(t,x,0) = 0, \quad \frac{\nu}{\epsilon}\,\frac{\partial u}{\partial z}(t,x,0) = \gamma u(t,x,0), \tag{6.2.11}$$

where $\gamma = \kappa/U$ is the dimensionless form of the friction coefficient. The boundary conditions on the free surface (6.2.5)-(6.2.6) are rescaled as

$$\frac{\nu}{\epsilon^2}\,\frac{\partial u}{\partial z} \;+\; p\,\frac{\partial h}{\partial x} \;-\; 2\nu\frac{\partial u}{\partial x}\,\frac{\partial h}{\partial x} \;+\; \nu\frac{\partial w}{\partial x} = 0 \quad \text{on} \quad z = h(t,x), \tag{6.2.12}$$

$$p \;-\; 2\nu\,\frac{\partial w}{\partial z} \;+\; \nu\frac{\partial u}{\partial z}\,\frac{\partial h}{\partial x} \;+\; \epsilon^2\nu\frac{\partial w}{\partial x}\,\frac{\partial h}{\partial x} = 0 \quad \text{on} \quad z = h(t,x). \tag{6.2.13}$$

The kinematic boundary condition (6.2.7) is unchanged.

6.2.1 A viscous hydrostatic model

We simplify the system (6.2.1)-(6.2.3) by retaining the zero and first order terms. We obtain a viscous hydrostatic model

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0, \tag{6.2.14}$$

$$\frac{\partial u}{\partial t} + \frac{\partial u^2}{\partial x} + \frac{\partial uw}{\partial z} + \frac{\partial p}{\partial x} = 2\nu\frac{\partial^2 u}{\partial x^2} + \frac{\nu}{\epsilon^2}\,\frac{\partial^2 u}{\partial z^2} + \nu\frac{\partial^2 w}{\partial x\,\partial z}, \tag{6.2.15}$$

$$\frac{\partial p}{\partial z} = -g + \nu\frac{\partial^2 u}{\partial x\,\partial z} + 2\nu\frac{\partial^2 w}{\partial z^2} + \epsilon^2\nu\,\frac{\partial^2 w}{\partial x^2}, \tag{6.2.16}$$

where

$$t > 0, \quad x \in \mathbb{R}, \quad 0 \le z \le h(t,x).$$

The associated boundary conditions are the same as the ones corresponding to the full dimensionless Navier-Stokes equations.

Remark 6.2.1 *Our motivation for keeping the second order term in the right hand side of the Navier-Stokes equations and in the free surface boundary conditions is that it is necessary for the energy dissipation as it is exhibited below. As it is an essential property of the Navier-Stokes equations we privilege this stability requirement to a strict first order approximation.*

Remark 6.2.2 *In (6.2.15)-(6.2.16) we have kept the symmetric form of the tensor - even though it can be simplified in Laplace terms - because of its natural relation to the boundary conditions (6.2.12)-(6.2.13).*

Now we check that the viscous hydrostatic approximation shares with the full Navier-Stokes equation on energy structure. We introduce the hydrostatic energy

$$e_h(t,x,z) = e_c(t,x,z) + e_p(t,x,z) = \frac{u^2}{2}(t,x,z) + gz. \tag{6.2.17}$$

Lemma 6.2.1 *The following energy equation holds for (6.2.14)-(6.2.16)*

$$\frac{\partial}{\partial t}\int_0^h e_h\,dz + \frac{\partial}{\partial x}\int_0^h \left(u(e_h+p) - \nu\,\left(2u\,\frac{\partial u}{\partial x} + w\,\frac{\partial u}{\partial z} + \epsilon^2 w\,\frac{\partial w}{\partial x}\right)\right)dz$$
$$= -\gamma u^2(t,x,0) - 2\nu\int_0^h \left[\left(\frac{\partial u}{\partial x}\right)^2 + \frac{1}{2}\left(\frac{1}{\epsilon}\,\frac{\partial u}{\partial z} + \epsilon\,\frac{\partial w}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial z}\right)^2\right]dz.$$

Proof. The proof uses classical computations. Following the purpose of Remark 6.2.1, let us notice the importance of the epsilon square term : it allows to include the term containing $\partial_z u \partial_x w$ in a square term and then to exhibit the decreasing of the energy. □

The computational complexity and cost of this viscous hydrostatic model remain similar to the ones of the full Navier-Stokes equations. In particular we do not recover the classical hydrostatic approximation. One way to go further in the simplification is to introduce a scaling in the viscosity and friction coefficients. It seems to be in accordance with the physical background : indeed the viscosity coefficient can be related to the turbulent viscosity of the flow which is in particular related to the geometric datas and thus to the shallow water assumption. Nevertheless there is many ways to define this scaling : it is possible to consider anisotropic viscosity coefficients and also to introduce different powers of ϵ - see [14]. Here we shall suppose that we are in the following regime

$$\nu = \epsilon \nu_0, \qquad \text{and} \qquad \gamma = \epsilon \gamma_0. \tag{6.2.18}$$

6.2.2 A classical hydrostatic model

We consider the system (6.2.8)-(6.2.10) with the particular form of the viscosity and friction coefficients (6.2.18) and one retains only the zero order terms. We obtain the very classical hydrostatic model

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0, \tag{6.2.19}$$

$$\frac{\partial u}{\partial t} + \frac{\partial u^2}{\partial x} + \frac{\partial uw}{\partial z} + \frac{\partial p}{\partial x} = \frac{\nu_0}{\epsilon} \frac{\partial^2 u}{\partial z^2}, \tag{6.2.20}$$

$$\frac{\partial p}{\partial z} = -g, \tag{6.2.21}$$

where

$$t > 0, \quad x \in \mathbb{R}, \quad 0 \leq z \leq h(t,x).$$

After simplifying the terms of corresponding order, the boundary conditions (6.2.11)-(6.2.13) become

$$w(t,x,0) = 0, \tag{6.2.22}$$

$$\frac{\nu_0}{\epsilon} \frac{\partial u}{\partial z}(t,x,0) = \gamma_0 u(t,x,0), \quad \frac{\partial u}{\partial z}(t,x,h(t,x)) = 0, \tag{6.2.23}$$

$$p(t,x,h(t,x)) = 0. \tag{6.2.24}$$

The system is still associated with the kinematic boundary condition (6.2.7). Taking into account the pressure boundary condition on the free surface (6.2.24), (6.2.21) is equivalent to

$$p(t,x,z) = g(h(t,x) - z). \tag{6.2.25}$$

It is now possible to check that the hydrostatic approximation shares with the full Navier-Stokes equation on energy structure. Namely considering the hydrostatic energy (6.2.17),

Lemma 6.2.2 *The following equality holds for (6.2.19)-(6.2.24)*

$$\frac{\partial}{\partial t}\int_0^h e_h\ dz + \frac{\partial}{\partial x}\int_0^h u(e_h+p)dz \ = \ -\gamma_0 u^2(t,x,0) \ - \ \frac{\nu_0}{\epsilon}\int_0^h \left(\frac{\partial u}{\partial z}\right)^2 dz.$$

Proof. The proof uses on classical computations. □

This hydrostatic model - or some variants with horizontal viscosity or other specific terms - is very used in geophysical flows studies. For further references see for example [14, 32, 65].

6.3 The Multilayer Saint-Venant System

We can now derive the Saint-Venant type models - see Figure 6.1.1. To do this we perform an asymptotic analysis of the hydrostatic models obtained in the previous section.
In [50] the authors present how to properly derive monolayer Saint-Venant models. Introducing an averaged velocity

$$U(t,x) = \frac{1}{h(t,x)}\int_0^h u(t,x,z)dz.$$

and departing respectively from the hydrostatic model (6.2.19)-(6.2.21) and from the viscous hydrostatic model (6.2.14)-(6.2.16) - under the hypothesis (6.2.18), they deduce the classical Saint-Venant system with friction

$$\frac{\partial h}{\partial t} + \frac{\partial hU}{\partial x} \ = \ 0, \tag{6.3.1}$$

$$\frac{\partial hU}{\partial t} + \frac{\partial}{\partial x}\left(hU^2 + \frac{gh^2}{2}\right) \ = \ -\kappa U. \tag{6.3.2}$$

and a viscous Saint-Venant system with friction

$$\frac{\partial h}{\partial t} + \frac{\partial hU}{\partial x} \ = \ 0, \tag{6.3.3}$$

$$\frac{\partial hU}{\partial t} + \frac{\partial}{\partial x}\left(hU^2 + \frac{gh^2}{2}\right) \ = \ -\frac{\kappa}{1+\frac{\kappa h}{3\mu}}U + 4\mu\frac{\partial}{\partial x}\left(h\frac{\partial U}{\partial x}\right). \tag{6.3.4}$$

as formal asymptotic approximations in $O(\epsilon)$ - respectively $O(\epsilon^2)$ - of the Navier-Stokes equations. More details about these derivations can be found in [45, 50].

These two Saint-Venant models are associated with a dissipation of energy. Indeed introducing

$$hE = hE_c + hE_p = \frac{hU^2}{2} + \frac{gh^2}{2},$$

we establish for the classical Saint-Venant model (6.3.1)-(6.3.2)

$$\frac{\partial hE}{\partial t} + \frac{\partial}{\partial x}\left(U\left(hE + \frac{gh^2}{2}\right)\right) = -\kappa U^2,$$

and for the viscous ones (6.3.3)-(6.3.4)

$$\frac{\partial hE}{\partial t} + \frac{\partial}{\partial x}\left(U\left(hE + \frac{gh^2}{2}\right) - 4\mu\frac{\partial}{\partial x}\left(hU\frac{\partial U}{\partial x}\right)\right) = -\frac{\kappa}{1+\frac{\kappa h}{3\mu}}U^2 - 4\mu h\left(\frac{\partial U}{\partial x}\right)^2.$$

Remark 6.3.1 *Notice that the classical Saint-Venant system (6.3.1)-(6.3.2) provides an exact solution to the hydrostatic system (6.2.19)-(6.2.21) when there is no friction on the bottom - i.e.* $\gamma_0 = 0$. *Indeed we can choose*

$$u(t,x,z) = U(t,x), \quad w(t,x,z) = -z\frac{\partial U}{\partial x}, \quad p(t,x,z) = g(h(t,x) - z),$$

where $(h,U)(t,x)$ *is a solution of the classical Saint-Venant system (6.3.1)-(6.3.2). It is in general no more the case for the viscous models. Indeed a solution* $(h_v, U_v)(t,x)$ *of the viscous Saint-Venant system (6.3.3)-(6.3.4) - with no friction - provides a solution of the viscous hydrostatic system (6.2.14)-(6.2.16) if and only if it verifies also the following equality*

$$4\frac{\partial}{\partial x}\left(h_v\frac{\partial U_v}{\partial x}\right) = 3h_v\frac{\partial^2 U_v}{\partial x^2}.$$

We now derive a more precise approximation. Especially we wish to keep some information on the vertical structure of the horizontal velocity as motivated in the introduction.

6.3.1 The Multilayer Saint-Venant model

We consider the hydrostatic model (6.2.19)-(6.2.21). We first introduce a discretization in the variable z - see Figure 6.3.1. Then for some $M \in \mathbb{N}$ we define M intermediate water heights $H_\alpha(t,x)$ such that

$$0 = H_0(t,x) \leq H_1(t,x) \leq H_2(t,x) \leq ... \leq H_{M-1}(t,x) \leq H_M(t,x) = h(t,x).$$

We characterize M layers through the definition of the indicator functions $\phi_\alpha(t,x,z)$

$$\forall \alpha \in \{1,M\} \quad \phi_\alpha(t,x,z) = \begin{cases} 1 & if \quad H_{\alpha-1}(t,x) \leq z \leq H_\alpha(t,x) \\ 0 & \text{otherwise,} \end{cases}$$

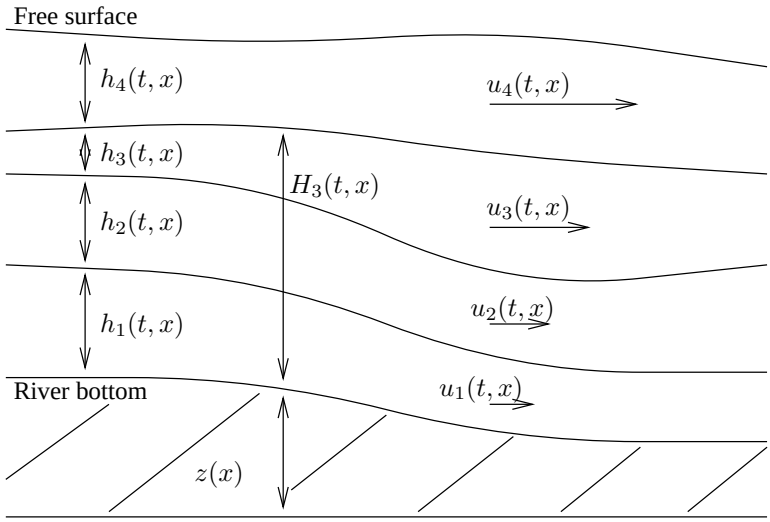


FIG. 6.3.1: A multilayer approach

that are advected by the flow

$$\frac{\partial \phi_\alpha}{\partial t} + \frac{\partial \phi_\alpha u}{\partial x} + \frac{\partial \phi_\alpha w}{\partial z} = 0. \tag{6.3.5}$$

Then for each layer we define its water height $h_\alpha(t,x)$ by

$$\forall \alpha \in \{1, M\} \qquad h_\alpha(t,x) = H_\alpha(t,x) - H_{\alpha-1}(t,x),$$

and an averaged velocitiy $U_\alpha(t,x)$

$$\forall \alpha \in \{1, M\} \qquad U_\alpha(t,x) = \frac{1}{h_\alpha(t,x)} \int_{H_{\alpha-1}}^{H_\alpha} u(t,x,z) dz.$$

We claim that

Theorem 6.3.1 *The multilayer Saint-Venant system with friction defined by*

$$\frac{\partial h_1}{\partial t} + \frac{\partial h_1 U_1}{\partial x} = 0, \tag{6.3.6}$$

$$\frac{\partial h_1 U_1}{\partial t} + \frac{\partial}{\partial x}\left(h_1 U_1^2\right) + g h_1 \frac{\partial}{\partial x}\sum_{\beta=1}^{M} h_\beta = 2\mu \frac{U_2 - U_1}{h_2 + h_1} - \kappa U_1, \tag{6.3.7}$$

$$\frac{\partial h_\alpha}{\partial t} + \frac{\partial h_\alpha U_\alpha}{\partial x} = 0, \tag{6.3.8}$$

$$\frac{\partial h_\alpha U_\alpha}{\partial t} + \frac{\partial}{\partial x}\left(h_\alpha U_\alpha^2\right) + g h_\alpha \frac{\partial}{\partial x}\sum_{\beta=1}^{M} h_\beta = 2\mu \frac{U_{\alpha+1} - U_\alpha}{h_{\alpha+1} + h_\alpha} - 2\mu \frac{U_\alpha - U_{\alpha-1}}{h_\alpha + h_{\alpha-1}}, \tag{6.3.9}$$

$$for \quad \alpha = 2...M-1$$

$$\frac{\partial h_M}{\partial t} + \frac{\partial h_M U_M}{\partial x} = 0, \tag{6.3.10}$$

$$\frac{\partial h_M U_M}{\partial t} + \frac{\partial}{\partial x}\left(h_M U_M^2\right) + g h_M \frac{\partial}{\partial x}\sum_{\beta=1}^{M} h_\beta = -2\mu \frac{U_M - U_{M-1}}{h_M + h_{M-1}}, \tag{6.3.11}$$

results from a formal asymptotic approximation in $O(\epsilon)$ coupled with a vertical discretization of the hydrostatic model and therefore of the Navier-Stokes equations.

Proof. We obtain by integration of (6.3.5) in the vertical direction M mass balance equations

$$\frac{\partial h_\alpha}{\partial t} + \frac{\partial}{\partial x}\int_{H_{\alpha-1}}^{H_\alpha} u dz = 0, \tag{6.3.12}$$

and M kinematic boundary conditions at the interfaces

$$\frac{\partial H_\alpha}{\partial t} + u(t,x,z=H_\alpha(t,x))\frac{\partial H_\alpha}{\partial x} - w(t,x,z=H_\alpha(t,x)) = 0. \tag{6.3.13}$$

Using the averaged velocities U_α we write the mass balance equations (6.3.12) in the Saint-Venant formalism

$$\frac{\partial h_\alpha}{\partial t} + \frac{\partial}{\partial x} h_\alpha U_\alpha = 0. \tag{6.3.14}$$

We now integrate on each layer the equation (6.2.20). The form of the hydrostatic pressure (6.2.25) leads to

$$\int_{H_{\alpha-1}}^{H_\alpha} \frac{\partial p}{\partial x} dz = \int_{H_{\alpha-1}}^{H_\alpha} g \frac{\partial h}{\partial x} dz = g \frac{\partial h}{\partial x}\int_{H_{\alpha-1}}^{H_\alpha} dz = g h_\alpha \frac{\partial h}{\partial x},$$

and thus we obtain - taking into account the kinematic boundary conditions (6.3.13) - M momentum equations

$$\frac{\partial}{\partial t}\int_{H_{\alpha-1}}^{H_\alpha} u dz + \frac{\partial}{\partial x}\int_{H_{\alpha-1}}^{H_\alpha} u^2 dz + g h_\alpha \frac{\partial h}{\partial x} = \frac{\nu_0}{\epsilon}\frac{\partial u}{\partial z}(H_\alpha(t,x)) - \frac{\nu_0}{\epsilon}\frac{\partial u}{\partial z}(H_{\alpha-1}(t,x)). \tag{6.3.15}$$

Now equations (6.2.20) and (6.2.23) implie

$$\frac{\partial^2 u}{\partial z^2} = O(\epsilon), \qquad \frac{\partial u}{\partial z}|_{z=0} = O(\epsilon), \qquad \frac{\partial u}{\partial z}|_{z=h} = 0,$$

and therefore

$$u(t,x,z) = U(t,x) + O(\epsilon) \qquad \forall z \quad 0 \le z \le h(t,x), \tag{6.3.16}$$

what implies

$$\frac{1}{h_\alpha(t,x)} \int_{H_{\alpha-1}}^{H_\alpha} u^2(t,x,z)dz = U_\alpha^2(t,x) + O(\epsilon).$$

Hence we can write the momentum equations (6.3.15) in the Saint-Venant formalism

$$\frac{\partial}{\partial t} h_\alpha U_\alpha + \frac{\partial}{\partial x} h_\alpha U_\alpha^2 + g h_\alpha \frac{\partial h}{\partial x} = \frac{\nu_0}{\epsilon} \frac{\partial u}{\partial z}(H_\alpha(t,x)) - \frac{\nu_0}{\epsilon} \frac{\partial u}{\partial z}(H_{\alpha-1}(t,x)) + O(\epsilon). \tag{6.3.17}$$

We drop the $O(\epsilon)$ and multiply (6.3.14) and (6.3.17) by HU^2/L in order to recover the variables with dimension. Finally we apply a finite difference method in the vertical direction to the right hand side of the momentum equations when it is concerned with an interface. □

6.3.2 Properties of the Multilayer Saint-Venant System

The multilayer Saint-Venant systems presents two fundamental stability properties.

Proposition 6.3.1 *The multilayer Saint-Venant system (6.3.23)-(6.3.28) preserves the positivity of the water height in each layer. It is also associated with an energy inequality. Namely denoting* $h = \sum_{\beta=1}^{M} h_\beta$ *and introducing the energy*

$$E = \sum E_\alpha = \sum \left(\frac{h_\alpha U_\alpha^2}{2h} + \frac{g h_\alpha}{2} \right),$$

we establish the following equality

$$\frac{\partial}{\partial t} hE + \frac{\partial}{\partial x} \left(h \sum U_\alpha E_\alpha + h \sum U_\alpha \frac{g h_\alpha}{2} \right) = -\kappa U_1^2 - 2\mu \sum_2^M \frac{(u_\alpha - U_{\alpha-1})^2}{h_\alpha + h_{\alpha-1}}.$$

Proof. The development of the proof is very classical since it mimicks what is done for the classical Saint-Venant system. The new dissipative term in the energy equation is due to the presence of the same viscous friction term at the interface $\alpha - 1/2$ in the momentum equations for the layers α and $\alpha - 1$. □

The multilayer Saint-Venant systems ensures also some relations with the classical Saint-Venant system

Proposition 6.3.2 *The multilayer Saint-Venant system (6.3.23)-(6.3.28) preserves the so-called lake at rest equilibrium*

$$\sum h_\alpha(0,x) = H, \qquad U_\alpha(0,x) = 0 \quad \forall \alpha = 1...M \tag{6.3.18}$$

Also when there is no friction on the bottom the classical Saint-Venant system (6.3.1)-(6.3.2) provides a solution for the multilayer case. Indeed if we choose the initial conditions $U_\alpha(t=0,x) = U_0(x)$ and $h_\alpha(t=0,x) = c_\alpha H_0(x)$ (such that $\sum c_\alpha = 1$), a multilayer solution is given by

$$\forall \alpha = 1...M \qquad \left\{ \begin{array}{l} h_\alpha(t,x) = c_\alpha h(t,x), \\ U_\alpha(t,x) = U(t,x), \end{array} \right.$$

where (h,U)(t,x) is a solution of the classical Saint-Venant system with initial values $(H_0, U_0)(t,x)$.

Proof. Proofs are obvious. Notice that the preservation of some steady states is a fundamental specific property of the Saint-Venant type models [13, 62]. The second point is is in agreement with Remark 6.3.1 since the multilayer system is an intermediate model between the hydrostatic system and the classical Saint-Venant system. □

6.3.3 Non conservativity and non hyperbolicity of the Multilayer Saint-Venant System

The multilayer Saint-Venant system (6.3.6)-(6.3.11) has two main drawbacks. First as opposed to the monolayer Saint-Venant models the pressure terms are not in a conservative form and thus their definition is not obvious when shocks occur. Nevertheless many works have been devoted to this question and some recent works propose numerical ways to choose a "right" solution - see [38] or [29] in the context of the Saint-Venant computations.
The second point is much more problematic. For the simplicity of the purpose let us consider the two layers version of the system (6.3.6)-(6.3.11). Following [27] we write the two layers system under the compact form

$$\frac{\partial X}{\partial t} + A(X)\frac{\partial X}{\partial x} = S(X), \tag{6.3.19}$$

where

$$X = \begin{pmatrix} h_1 \\ h_1 U_1 \\ h_2 \\ h_2 U_2 \end{pmatrix}, \qquad S(X) = \begin{pmatrix} 0 \\ 2\mu \dfrac{U_2 - U_1}{h_2 + h_1} - \kappa U_1 \\ 0 \\ 2\mu \dfrac{U_1 - U_2}{h_2 + h_1} \end{pmatrix},$$

$$A(X) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -U_1^2 + gh_1 & 2U_1 & gh_1 & 0 \\ 0 & 0 & 0 & 1 \\ gh_2 & 0 & -U_2^2 + gh_2 & 2U_2 \end{bmatrix}.$$

As the vertical profile of the velocity is given by the relation (6.3.16) we have

$$U_\alpha(t,x) = U(t,x) + O(\epsilon) \qquad \forall \alpha = 1,2. \tag{6.3.20}$$

Proposition 6.3.3 *Under the assumption (6.3.20) the non conservative two layers Saint-Venant system (6.3.19) is not hyperbolic. Moreover when $U_1(t,x) \neq U_2(t,x)$ the eigenvalues of the matrix $A(X)$ have a non vanishing imaginary part.*

Proof. The eigenvalues of the matrix $A(X)$ are solutions of

$$\left(\lambda^2 - 2U_1\lambda + U_1^2 - gh_1\right)\left(\lambda^2 - 2U_2\lambda + U_2^2 - gh_2\right) = g^2 h_1 h_2.$$

Therefore under the assumption (6.3.20) first order approximations in $U_1 - U_2$ of the eigenvalues are

$$\lambda_{ext}^{\pm} = U_m \pm \sqrt{g(h_1+h_2)}, \tag{6.3.21}$$

$$\lambda_{int}^{\pm} = U_c \pm i\frac{U_1 - U_2}{2}\sqrt{1 - \left(\frac{h_1 - h_2}{h_1 + h_2}\right)^2}, \tag{6.3.22}$$

where

$$U_m = \frac{h_1U_1 + h_2U_2}{h_1 + h_2}, \qquad U_c = \frac{h_1U_2 + h_2U_1}{h_1 + h_2}.$$

Notice that in the case $U_1 = U_2 = U$ the eigenvalues become real but the system is still not hyperbolic since U is a double eigenvalue associated with a one dimensional eigenspace. □

The eigenvalues $\lambda_{int}^{\pm}$ characterize the baroclinic part of the flow and are related to the celerity of a signal at the interface between the two layers [27, 51]. Their imaginary parts are related to the existence of instabilities at the interface which lead to a mixing between the two layers. Hence it is natural that the baroclinic eigenvalues of the multilayer system have a non vanishing imaginary part since the hypothesis (6.3.5) about the immiscibility of the layers is not linked to the physical background. But it follows also that the precise value of $\lambda_{int}^{\pm}$ is not relevant since we are not interested in the evolution of these "virtual" internal interfaces.
The two other eigenvalues $\lambda_{ext}^{\pm}$ characterize the barotropic part of the flow and are related to the celerity of a signal at the free surface. Therefore the multilayer Saint-Venant model is in agreement with the classical results about the monolayer Saint-Venant models for which the eigenvalues - and thus the celerities of a signal at the free surface - are $U \pm \sqrt{gh}$. Since these celerities are well-known to be in good accordance with the experiences, the agreement between the two models and thus the value of $\lambda_{ext}^{\pm}$ is an essential property of the multilayer system.

6.3.4 Conservative Form of the Multilayer Saint-Venant Model

To overcome the difficulties due to these non conservative structure and non hyperbolic nature of the multilayer Saint-Venant system (6.3.6)-(6.3.11) and therefore to avoid apparition of instabilities in the numerical simulations [27] we propose to consider a slightly different system

$$\frac{\partial h_1}{\partial t} + \frac{\partial h_1 U_1}{\partial x} = 0, \tag{6.3.23}$$

$$\frac{\partial h_1 U_1}{\partial t} + \frac{\partial}{\partial x}\left(h_1 U_1^2 + g\frac{h_1\left(\sum_{\beta=1}^{M} h_\beta\right)}{2}\right) = \frac{g\left(\sum_{\beta=1}^{M} h_\beta\right)^2}{2}\frac{\partial}{\partial x}\left(\frac{h_1}{\sum_{\beta=1}^{M} h_\beta}\right) + 2\mu\frac{U_2 - U_1}{h_2 + h_1} - \kappa U_1, \tag{6.3.24}$$

$$\frac{\partial h_\alpha}{\partial t} + \frac{\partial h_\alpha U_\alpha}{\partial x} = 0, \tag{6.3.25}$$

$$\frac{\partial h_\alpha U_\alpha}{\partial t} + \frac{\partial}{\partial x}\left(h_\alpha U_\alpha^2 + g\frac{h_\alpha\left(\sum_{\beta=1}^{M} h_\beta\right)}{2}\right) = \frac{g\left(\sum_{\beta=1}^{M} h_\beta\right)^2}{2}\frac{\partial}{\partial x}\left(\frac{h_\alpha}{\sum_{\beta=1}^{M} h_\beta}\right) + 2\mu\left(\frac{U_{\alpha+1} - U_\alpha}{h_{\alpha+1} + h_\alpha} - \frac{U_\alpha - U_{\alpha-1}}{h_\alpha + h_{\alpha-1}}\right) \tag{6.3.26}$$

$$for \quad \alpha = 2...M-1,$$

$$\frac{\partial h_M}{\partial t} + \frac{\partial h_M U_M}{\partial x} = 0, \tag{6.3.27}$$

$$\frac{\partial h_M U_M}{\partial t} + \frac{\partial}{\partial x}\left(h_M U_M^2 + g\frac{h_M\left(\sum_{\beta=1}^{M} h_\beta\right)}{2}\right) = \frac{g\left(\sum_{\beta=1}^{M} h_\beta\right)^2}{2}\frac{\partial}{\partial x}\left(\frac{h_M}{\sum_{\beta=1}^{M} h_\beta}\right) - 2\mu\frac{U_M - U_{M-1}}{h_M + h_{M-1}}. \tag{6.3.28}$$

Proposition 6.3.4 *Under the assumption (6.3.16) this new set-up of the same system has the following properties*

(P1) The system obtained by replacing the right-hand side by zero is hyperbolic.
(P2) The barotropic eigenvalues $\lambda^{\pm}_{ext}$ (6.3.21) are still first order approximations in $U_1 - U_2$ of two of its eigenvalues.
(P3) For suitable initial water height data, the non conservative terms in the right hand side are "small" - they vanish in the Saint-Venant approximation $h_\alpha = h/M$, $U_\alpha = U$.
(P4) The sum on all the layers of the equations that describe the evolution of the water height and of the discharge in each layer is a first order approximation of the classical Saint-Venant system - left and right hand side considered separately.

Proof. As for the non conservative model we consider for simplicity the two layers case. Hence the eigenvalues of the matrix on the left hand side are solutions of

$$\left(\lambda^2 - 2U_1\lambda + U_1^2 - gh_1 - \frac{gh_2}{2}\right)\left(\lambda^2 - 2U_2\lambda + U_2^2 - gh_2 - \frac{gh_1}{2}\right) = \frac{g^2h_1h_2}{4}.$$

Therefore under the assumption (6.3.20) first order approximations of the eigenvalues are

$$\lambda^{\pm}_{ext} = U_m \pm \sqrt{g(h_1+h_2)},$$
$$\lambda^{\pm}_{int} = U_c \pm \sqrt{\frac{g(h_1+h_2)}{2}},$$

where

$$U_m = \frac{h_1U_1 + h_2U_2}{h_1+h_2}, \qquad U_c = \frac{h_1U_2 + h_2U_1}{h_1+h_2}.$$

Properties (P1) and (P2) follow immediately. Properties (P3) and (P4) are consequences of obvious computations - notice that "suitable initial water height data" means that the initial non conservative pressure source term is small. The second part of (P3) means that in the zero friction case, and for suitable initial data, the non conservative pressure source term P and the viscous source term V vanish for each time.

$$P = \frac{g}{2}\left(\sum_{\beta=1}^{M} h_\beta\right)^2 \partial_x\left(\frac{h}{\sum_{\beta=1}^{M} h_\beta}\right), \quad V = 2\mu\left[\frac{(U_{\alpha+1}-U_\alpha)}{(h_{\alpha+1}+h_\alpha)} - \frac{(U_\alpha - U_{\alpha-1})}{(h_\alpha + h_{\alpha-1})}\right]$$

A consequence of (P4) is that the free surface and the total discharge of the multilayer system are first order approximation of the classical Saint-Venant results, plus corrections due to the more sophisticated vertical profile of the flow velocity. □

Remark 6.3.2 *Outside from the dry areas the conservative part of the two layers system is strictly hyperbolic. Nevertheless it is no more the case when we consider a multilayer system with three or more layers since the multiplicity of the eigenvalues $\lambda^{\pm}_{int}$ is equal to the number of interfaces - at least for the case $U_\alpha = U$. The system is still hyperbolic since the dimensions of the associated eigenspaces are also equal to the number of interfaces - see Serre [118], Dafermos [37] for precise definitions.*

Remark 6.3.3 *The most intuitive form for a conservative multilayer model is the one which mimicks exactly the flux of the classical Saint-Venant system (6.3.1)-(6.3.2) by introducing $\partial_x g h_\alpha^2$ as conservative pressure term. The non conservative part is then $g h_\alpha \partial_x \sum_{\beta \neq \alpha} h_\beta$ and can be interpreted as a topographic source term. The conservative part is hyperbolic since the eigenvalues are - for each layer - deduced from those of the classical Saint-Venant system. Also the strict analogy with the classical Saint-Venant system seems to be convenient for computational use. Nevertheless this method is in fact defective since it does not respect any of the three properties (P2)-(P4) and especially the eigenvalues of the hyperbolic system of the left hand side are far from the barotropic eigenvalues $\lambda_{ext}^{\pm}$ (6.3.21).*

Remark 6.3.4 *Both set-ups of the multilayer model preserve the lake at rest steady state. But if we have in mind the numerical treatment of this equilibrium state it is important to notice that, if for the non-conservative model and at the lake at rest equilibrium state both fluxes and source terms vanish, on the contratry, for the conservative model and always at the lake at rest equilibrium state fluxes and source terms are equal but different from zero. It follows that the situation for the conservative case is very similar to the preservation of the classical (monolayer) Saint-Venant lake at rest equilibrium in the presence of a non zero topographic term. And it is well known that the preservation of this equilibrium at the discrete level is far from being obvious [13, 62].*

6.4 The discrete multilayer scheme

We now present the discrete version of the conservative multilayer Saint-Venant system (6.3.23)-(6.3.28). The conservative multilayer Saint-Venant system (6.3.23)-(6.3.28) presents some new terms when compared with the classical Saint-Venant system (6.3.1)-(6.3.2) - in the source terms and even in the conservative part. Several strategies are possible and for instance Pares et al. [27] prefer to consider the full system and built a specific solver for the two layers case. Since here we wish to treat as many layers as possible we follow another strategy that can be seen as a modified layer-by-layer approach - with respect to the terminology introduced in [27]. Indeed we consider the multilayer system as M coupled modified Saint-Venant systems and we choose to adapt an existing Saint-Venant finite volume solver to the multilayer case - for a general introduction to the finite volume methods refer to [54, 90]. Especially the discretization technics have to preserve the properties of the continuous multilayer model stated in the Propositions 6.3.1 and 6.3.2.

To approximate the solution $(h_\alpha(t,x), U_\alpha(t,x)), \alpha = 1..M$ of the multi-layer Saint-Venant system by discrete values $(h_{\alpha j}^n, U_{\alpha j}^n), \alpha = 1..M, j \in \mathbb{Z}, n \in \mathbb{N}$ we introduce as usual a space-time discretization based on a grid of points $x_{j+1/2}$ with space steps $\Delta x_j = x_{j+1/2} - x_{j-1/2}$ and on a grid of points t^n defined by $t^n = \sum_{k \leq n} \Delta t^k$ where the time steps Δt^k will be precised later through a CFL condition. Then we use the finite

volume framework.
We choose to include explicitely the pressure source term in the finite volume solver. This follows usual ideas for bottom topographies - refer to [7]. The viscosity source term can be interpreted as a friction term between the layer we are considering and the two neighbouring layers. As usual we treat this friction term implicitly. This leads to solve a linear system.
For the simplicity of the purpose we now divide the computation in two steps : explicit computation of the fluxes taking into account the pressure source term and implicit computation of the friction source terms.

6.4.1 The finite volume solver

To perform the explicit step we use a finite volume kinetic scheme. The general form of a finite volume method is

$$X_{\alpha j}^{n+\frac{1}{2}} - X_{\alpha j}^{n} + \sigma_j^n [F_{\alpha,j+\frac{1}{2}}^{n} - F_{\alpha,j-\frac{1}{2}}^{n}] - \Delta t^n \; S_{\alpha j}^{n} = 0,$$

where $X_{\alpha j}^{n} = \left(h_{\alpha j}^{n}, q_{\alpha j}^{n} = h_{\alpha j}^{n} U_{\alpha j}^{n}\right)$ is the vector of the unknowns, $\sigma_j^n = \Delta t^n / \Delta x_j$ is the ratio between space and time steps, and where the discrete flux $F_{\alpha,j+\frac{1}{2}}^{n}$ is an approximation of the exact flux estimated at point $x_{j+\frac{1}{2}}$.
The kinetic scheme is a particular way to construct the numerical flux. For the case of the Saint-Venant system it presents a very good compromise between stability and accuracy. Thanks to a microscopic interpretation of the equations it allows to construct macroscopic schemes that preserve some essential properties of the continuous model : positivity of the water height, entropy inequality, ability to treat vacuum areas... We refer to [105] for a general survey of the kinetic theory and to [9, 108] for further details about kinetic schemes for the Saint-Venant system.
The multilayer conservative pressure term does not fit exactly in the usual frame work of the Saint-Venant type models. Thus it is not obvious that all the finite volume schemes that have been developped to solve the classical Saint-Venant system can be extend to the multilayer case. We claim that the extension of the kinetic scheme to the multilayer case is very natural. It is another argument for its choice. In fact the only difference with the classical case arises in the definition of the Gibbs equilibrium that appears in the microscopic interpretation of the Saint-Venant system. Indeed the modified Gibbs equilibrium involves at the same time the water height of the layer and the total water height of the flow. More precisely and with respect to the notations in [9] the Gibbs equilibrium for the layer α is now

$$M_\alpha(t,x,\xi) = \frac{h_\alpha(t,x)}{c(t,x)} \chi\left(\frac{\xi - U_\alpha(t,x)}{c(t,x)}\right),$$

where

$$c(t,x) = \sqrt{\frac{g \sum_{\beta=1}^{M} h_\beta(t,x)}{2}}.$$

The new non conservative pressure source term is discretized through a minmod limiter to ensure the robustness of the scheme

$$S_{\alpha j}^{n} = \begin{pmatrix} 0 \\ \frac{g\left(\sum_{\beta=1}^{M} h_{\beta j}^{n}\right)^{2}}{2} \text{minmod}\left(\frac{h_{\alpha,j+1}^{n}}{\sum_{\beta=1}^{M} h_{\beta,j+1}^{n}} - \frac{h_{\alpha j}^{n}}{\sum_{\beta=1}^{M} h_{\beta j}^{n}}, \frac{h_{\alpha j}^{n}}{\sum_{\beta=1}^{M} h_{\beta j}^{n}} - \frac{h_{\alpha,j-1}^{n}}{\sum_{\beta=1}^{M} h_{\beta,j-1}^{n}} \right) \end{pmatrix}. \tag{6.4.1}$$

6.4.2 The implicit computation

This implicit step does not affect the discrete water heights. Therefore

$$h_{\alpha j}^{n+1} = h_{\alpha j}^{n+\frac{1}{2}}. \tag{6.4.2}$$

The computation of the new velocities $U_{\alpha j}^{n+1}$ leads to the computation of a tridiagonal $M \times M$ linear system

$$\begin{bmatrix} a_{1j} & b_{1j} & 0 & \cdots & 0 \\ c_{2j} & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & b_{M-1,j} \\ 0 & \cdots & 0 & c_{Mj} & a_{Mj} \end{bmatrix} \begin{bmatrix} U_{1j}^{n+1} \\ \vdots \\ \vdots \\ \vdots \\ U_{Mj}^{n+1} \end{bmatrix} = \begin{bmatrix} q_{1j}^{n+\frac{1}{2}} \\ \vdots \\ \vdots \\ \vdots \\ q_{Mj}^{n+\frac{1}{2}} \end{bmatrix}, \tag{6.4.3}$$

where

$$a_{1j} = h_{1j}^{n+1} + \frac{2\mu\Delta t^{n}}{h_{1j}^{n+1} + h_{2j}^{n+1}} + \kappa\Delta t^{n},$$

$$a_{\alpha j} = h_{\alpha j}^{n+1} + 2\mu\Delta t^{n}\left(\frac{1}{h_{\alpha j}^{n+1} + h_{\alpha+1,j}^{n+1}} + \frac{1}{h_{\alpha j}^{n+1} + h_{\alpha-1,j}^{n+1}}\right) \qquad \alpha = 2...M-1,$$

$$a_{Mj} = h_{Mj}^{n+1} + \frac{2\mu\Delta t^{n}}{h_{Mj}^{n+1} + h_{M-1,j}^{n+1}}, \tag{6.4.4}$$

$$b_{\alpha j} = -\frac{2\mu\Delta t^{n}}{h_{\alpha j}^{n+1} + h_{\alpha+1,j}^{n+1}} \qquad \alpha = 1...M-1, \tag{6.4.5}$$

$$c_{\alpha j} = -\frac{2\mu\Delta t^{n}}{h_{\alpha j}^{n+1} + h_{\alpha-1,j}^{n+1}} \qquad \alpha = 2...M. \tag{6.4.6}$$

We notice that this matrix is a M-matrix i.e. the diagonal terms are strictly positive, the other ones are strictly negative and the diagonal terms are strictly dominant. Therefore the matrix is invertible.

6.4.3 Properties of the discrete multilayer kinetic scheme

We exhibit in this section that the discrete multilayer kinetic scheme ensures three essential properties of the continuous multilayer model, *i.e.* the positivity of the water height, the preservation of the lake at rest equilibrium and the agreement between the solutions of the monolayer and multilayer Saint-Venant models when there is no friction on the bottom. The proofs are very similar to the classical monolayer case and for the details, we refer the reader to [9].

Theorem 6.4.1 *Under the CFL condition*

$$\Delta t \leq \min_j \left(\frac{\Delta x_j}{\max_\alpha \left(|U_{\alpha,j}| + w_\chi \left(\frac{g h_j}{2} \right)^{1/2} \right)} \right), \tag{6.4.7}$$

where w_χ is related to the size of the support of χ, the discrete kinetic multilayer scheme preserves the non negativity of the water height for each layer.

Proof. As the viscous source term takes place in the momentum equation the implicit step does not concern the water heights. Hence it is enough to prove that the finite volume solver preserves the positivity of the water height.
The idea of the proof is then to exhibit that the discrete microscopic density of particle at time $t^{n+\frac{1}{2}}$ is, for all ξ, a convex combination of the Gibbs equilibriums at time t^n and for the neighbouring points. We refer to [9] for the details. The adaptation to the multilayer case is easy. Let us notice that the CFL condition (6.4.7) is a little more restrictive that the classical monolayer CFL condition - see [9] - since it contains the velocity of the quickest layer. □

Theorem 6.4.2 *The discrete kinetic multilayer system preserves the stationary states associated to the lake at rest*

$$\forall \alpha = 1...M \qquad \left\{ \begin{array}{l} h_{\alpha j} = h_\alpha, \\ U_{\alpha j} = 0. \end{array} \right. \tag{6.4.8}$$

Also in the zero friction case with initial data $U^0_{\alpha j} = U^0_j$ and $h^0_{\alpha j} = c_\alpha H^0_j$ (such that $\sum c_\alpha = 1$), the solution of the discrete kinetic multilayer scheme is

$$\forall \alpha = 1...M \qquad \left\{ \begin{array}{l} h^n_{\alpha j} = c_\alpha H^n_j, \\ U^n_{\alpha j} = U^n_j, \end{array} \right.$$

where (U^n_j, H^n_j) is the solution of the classical Saint-Venant kinetic scheme - refer to [9] - with initial data (U^0_j, H^0_j).

Proof. The proof for the preservation of the lake at rest equilibrium is obvious. Nevertheless notice that the discrete lake at rest (6.4.8) is only a particular case of the continuous lake at rest steady state (6.3.18). Indeed the continuous lake at rest equilibrium is the result of a balance between two terms while the discrete multilayer scheme preserves this equilibrium only if these two terms vanish separately.
The proof of the second result follows from two facts (i) the source term vanish at each time step for both schemes, (ii) the fluxes of the multilayer scheme are obtained by multiplying the flux of the classical Saint-Venant kinetic scheme by the constants c_α. □

6.5 Numerical assessment : a dam break problem

We consider the case of a dam break on a flat bottom. This example is used in [50] to compare the classical Saint-Venant system (6.3.1)-(6.3.2), the viscous Saint-Venant system (6.3.3)-(6.3.4) and the Navier-Stokes equations (6.2.1)-(6.2.3). We include in this comparison our new multilayer Saint-Venant model (6.3.23)-(6.3.28). We also mention the solution of the homogeneous Saint-Venant system - *i.e.* the classical Saint-Venant system (6.3.1)-(6.3.2) with a zero friction coefficient - since it is very classical in the Saint-Venant litterature.

The main parameters of the computation are the following : viscosity $\mu = 0.01$, gravity $g = 2.0$. The dam is located at the middle of the computational domain. The water is initially at rest and the left and right water heights are respectively equal to two and one meter. We discretize the horizontal domain with two hundred points and we perform the computation with three, six or ten layers in the multilayer Saint-Venant case, ten layers in the Navier-Stokes case. The simulation time is equal to 14 seconds.

6.5.1 The zero friction case

Firstly we check that in the zero friction case and for convenient initial data the multilayer model coincides with the monolayer Saint-Venant models (classical or viscous) as we announced in the previous Section. The result is obtained with an error of the order of the computer accuracy. On an other hand we can observe in Figure 6.5.1 that even in the case of a zero friction coefficient and with a constant initial velocity, the horizontal Navier-Stokes velocity is no more constant along the vertical at the end of the computation. This is due to non-hydrostatic effects which we have to keep in mind when we will compare results later.

6.5.2 Comparison with monolayer Saint-Venant models

Second we consider a case with a non zero friction coefficient and we compare the total waterheight - in Figure 6.5.2 - and the mean velocity - in Figure 6.5.3 - of the multilayer flow with the waterheight and the velocity of the other Saint-Venant models

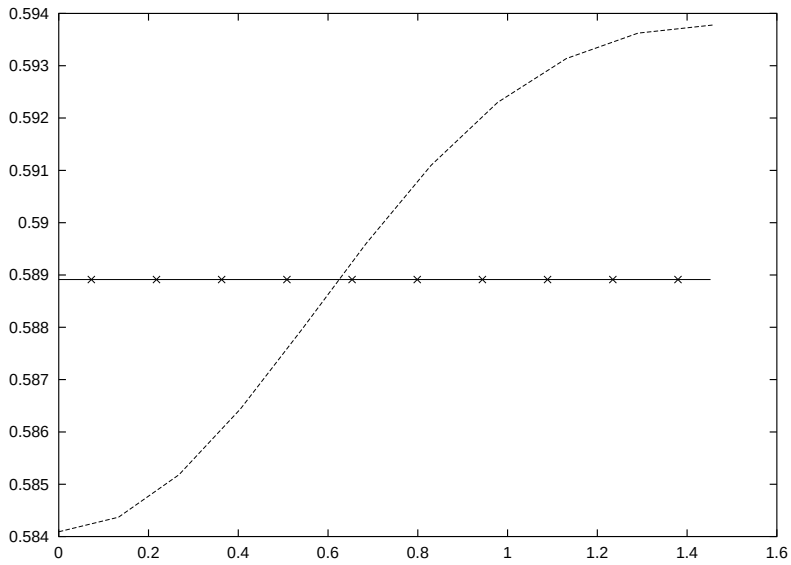


FIG. 6.5.1: VELOCITY (Vertical profiles) - Navier-Stokes and different Saint-Venant models - No Friction - Ten layers

(homogeneous, classical and viscous).
We first verify that the friction on the bottom has for consequence the decrease of the mean velocity of the flow - since the greatest value corresponds to the homogeneous case. Second it appears that the free surface and velocitiy longitudinal profiles are very different for the classical and for the viscous Saint-Venant systems. Since the viscous system is a more precise approximation of the Navier-Stokes equations it follows that in this case the classical Saint-Venant system fails to correctly characterize the flow. On the other hand the results from the viscous and multilayer Saint-Venant sytems are in very good agreement. It follows that in practice the mean quantities resulting from the multilayer computation are first order approximations of the Navier-Stokes results.

6.5.3 Multilayer aspect of the model

We now investigate the multilayer model by studying the evolution of each layer. We first present the velocity of each layer in Figure 6.5.4. The results are in accordance with what we could expect : the friction on the bottom makes the lowest layers slower compared to the velocity of the viscous Saint-Venant model - and the uppermost layers faster since the mean velocity of the multilayer flow is equal to the velocity of the viscous Saint-Venant model . Notice that the scheme appears to be able to compute large gaps between bottom and free surface velocities. In the case we have considered, the computed velocity for the uppermost layer is four times larger than the velocity computed for the lowest layer. Considering in addition the water heights - Figure 6.5.5

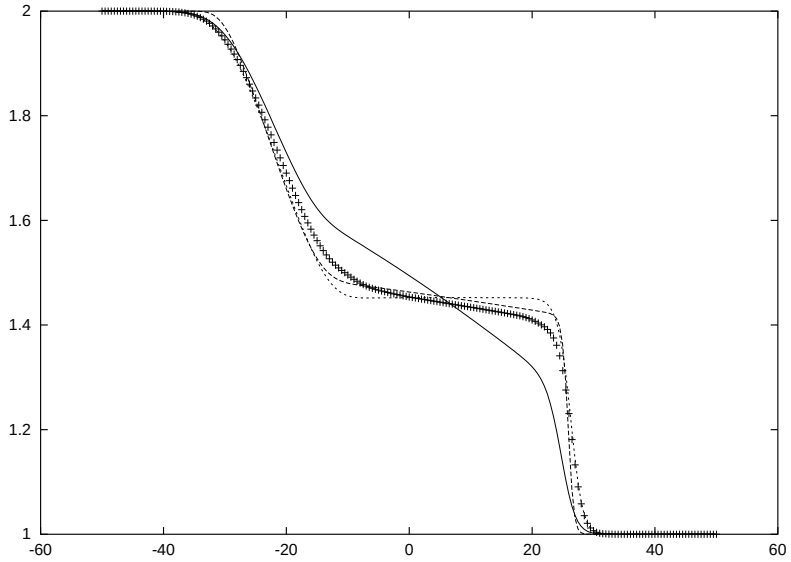


FIG. 6.5.2: FREE SURFACE (Longitudinal profiles) - Different Saint-Venant models - Ten layers - Friction coefficient = .1

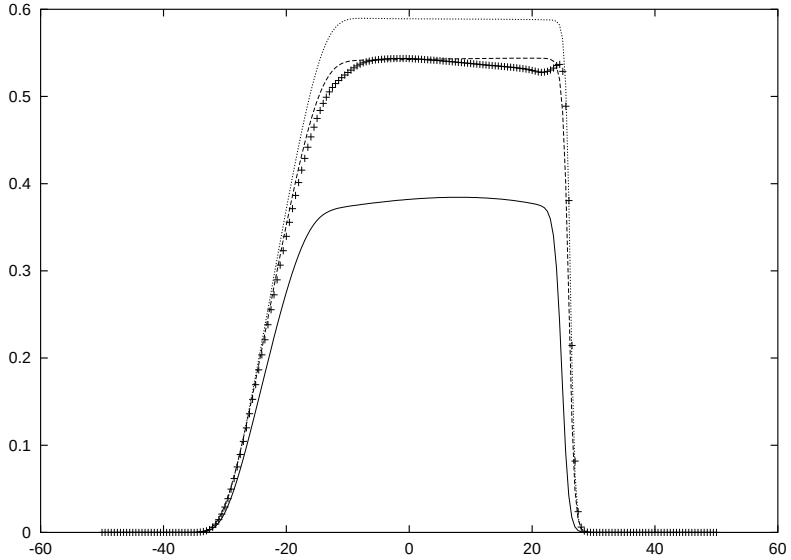


FIG. 6.5.3: VELOCITY (Longitudinal profiles) - Different Saint-Venant models - Ten layers - Friction coefficient = .1

- it appears also that - except for the lowest layers - the longitudinal profiles of the velocities and the water heights are smooth - even in the neighbourhood of the shock.

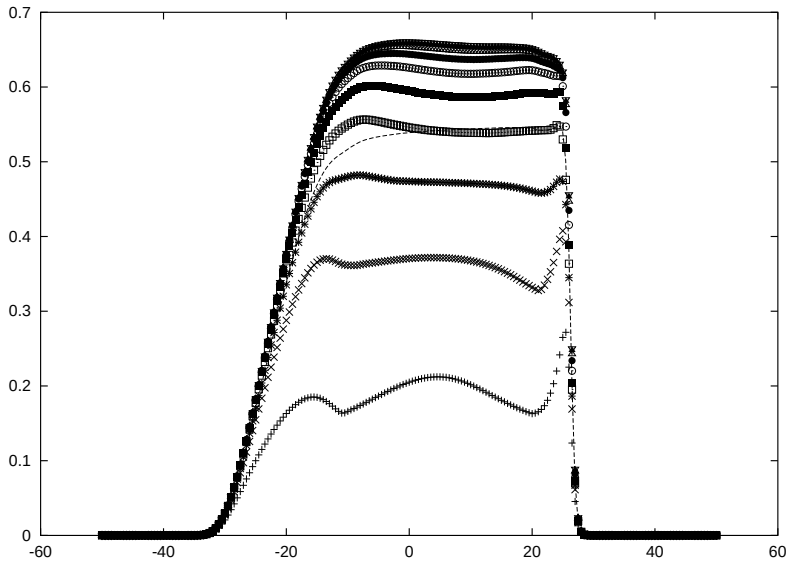


FIG. 6.5.4: LAYER VELOCITIES (Longitudinal profiles) - Mulitlayer Saint-Venant model (each type of cross corresponds to a different layer) - Ten layers - Friction coefficient = .1

6.5.4 Comparisons with Navier-Stokes velocity profiles

We now compare the two 2D models, *i.e.* the Navier-Stokes equations and the multilayer Saint-Venant system - we recall that we consider in this work only one horizontal direction. First we compare the two computed free surfaces of the flow. In Figure 6.5.6 we see that the two profiles are very close. In particular the speed of the shock wave and the slope of the free surface between the rarefaction and the shock waves are well computed.
We can now compare the two vertical profiles for the horizontal velocity. The Navier-Stokes profile (along the vertical) is continuous and piecewise linear. The multilayer Saint-Venant profile is constant on each layer and thus discontinuous. We choose to present - for each layer - the velocity at the middle of the layer - considering the vertical component. The multilayer and Navier-Stokes computations are managed with ten layers in the vertical direction. We choose an arbitrary vertical section included in the central zone - $x = 8m$, but what we observe is also valid for any section included in this region.
The results are presented in Figure 6.5.7. The two vertical profiles of the horizontal velocity are in good agreement. In particular notice that both values of the velocity at the bottom and at the free surface are well captured - even though the gap between

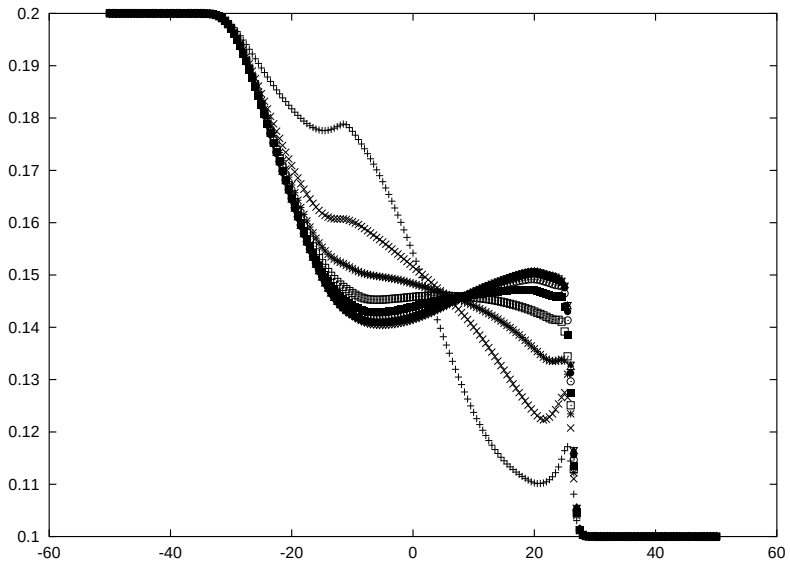


FIG. 6.5.5: LAYER HEIGHTS (Longitudinal profiles) - Multilayer Saint-Venant model (each type of cross corresponds to a different layer) - Ten layers - Friction coefficient = .1

these two values is large. Between these two extrema the curvature of the multilayer profile seems to be a bit larger than those of the Navier-Stokes profile but the amplitude of this difference is very limited.

6.5.5 Computational cost

Mono- and multilayer Saint-Venant systems are computed using the same finite volume algorithm. Thus the ratio between their respective computational costs is obviously linked to the number of layers. But the multilayer algorithm implies some supplementary computations - viscous effects. Moreover the multilayer CFL condition (6.4.7) is linked to the velocity of the fastest layer and thus it is a little more restrictive in the multi- than in the monolayer case. For example, in the considered case, with ten layers, the multilayer computation is about fifteen times more time consuming than the monolayer Saint-Venant one.
The Navier-Stokes algorithm that we used is based on an implicit ALE method with moving meshes. Thus the algebraic computations are much time consuming. Furthermore we consider an instationnary test case. It follows that, for this test case and for the same number of layers, the Navier-Stokes computation is about twenty times more time consuming than the multilayer Saint-Venant one.

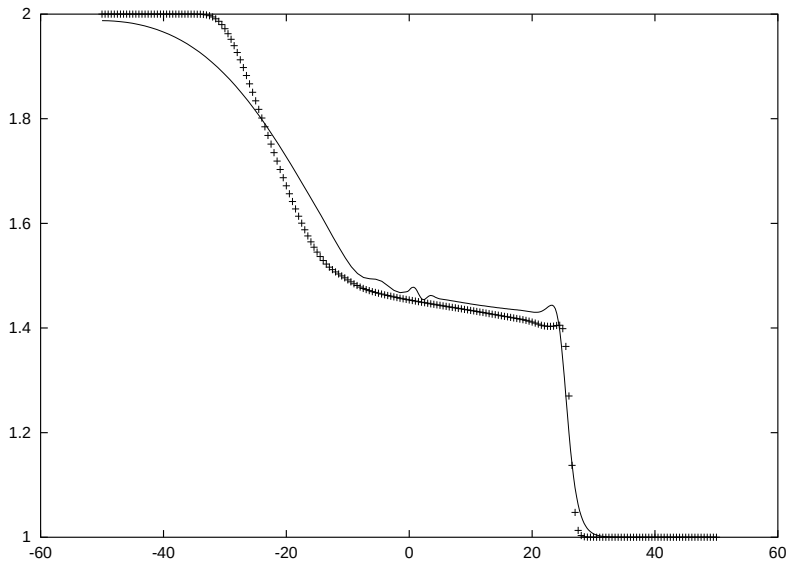


FIG. 6.5.6: FREE SURFACE (Longitudinal profiles) - Navier-Stokes and Multilayer Saint-Venant models - Ten layers - Friction coefficient = .1

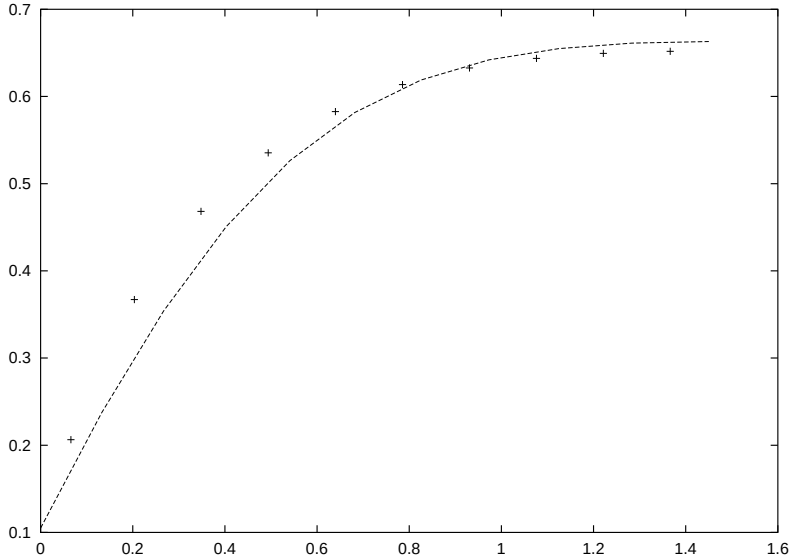


FIG. 6.5.7: VELOCITY (Vertical profiles) - Navier-Stokes and Multilayer Saint-Venant models - Ten layers - Friction coefficient = .1

6.5.6 Influence of the number of layers

Now we test the influence of the number of layers on the computed multilayer solution. We consider the same test case as before and we compute it with three and six layers. We present in Figure 6.5.8 the vertical velocity profiles obtained with the Navier-Stokes computation and with the multilayer one for three, six and ten layers. It appears that the multilayer results are in very good accordance with the Navier-Stokes ones as soon as we manage a really multilayer computation, even with a low number of layers. Indeed the computed vertical velocity profile for the three layers case is very close to the Navier-Stokes one. When we increase the number of layers the computed vertical profile seems to tend to a limit profile that is almost but not exactly the Navier-Stokes one - maybe due to the non hydrostatic effects.

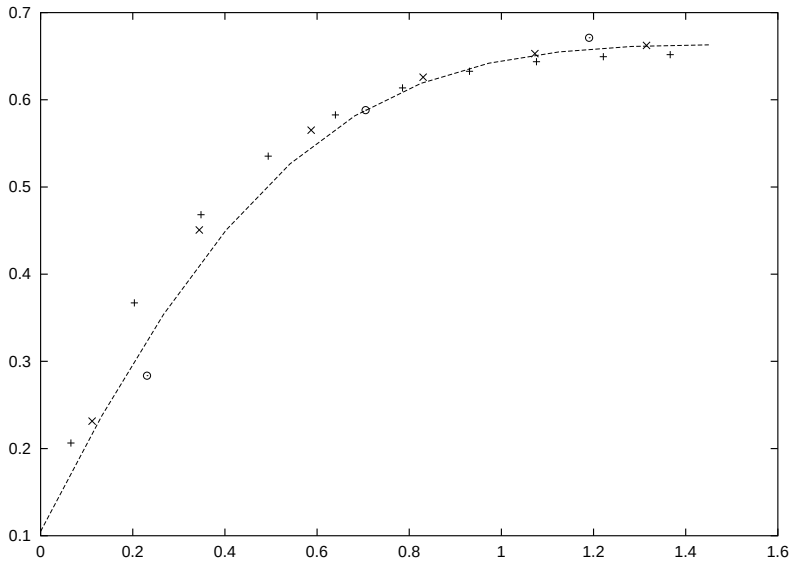


FIG. 6.5.8: VELOCITY (Vertical profiles) - Navier-Stokes and Multilayer Saint-Venant (with different number of layers) models - Friction coefficient = .1

6.5.7 Some other friction coefficients

To end this numerical validation we come back to the ten layers test case but changing the friction coefficient. We present the computed vertical velocity profiles with both Navier-Stokes and multilayer Saint-Venant approaches. In Figure 6.5.9 we present a small friction coefficient case. The difference between the two profiles is not neglectable and in particular the velocity at the bottom is not very well computed. This phenomenon can be related to the small non hydrostatic deviation that exists for the Navier-Stokes profile in the no friction case - Figure 6.5.1. Nevertheless let us observe that the difference between the top and the bottom velocities does not exceed

five per cent and then conclude that this cases are less interesting since for the small coefficients the classical Saint-Venant system is already a good approximation of the Navier-Stokes equations.
In Figure 6.5.10 we present the no-slip condition case. Notice that for this second example there is no notion of friction coefficient in the Navier-Stokes formalism since we impose that the velocity is equal to zero at the bottom. We extend this boundary condition to the multilayer computation. It follows that the velocity of the lowest layer is equal to zero. Like for the first test case the results are in quite good accordance. This result exhibits the robustness of the model since we are not at all in the asymptotic case from which we deduce the multilayer system since the friction coefficient is infinitely large. It follows that the multilayer Saint-Venant model has a quite large range of validity.

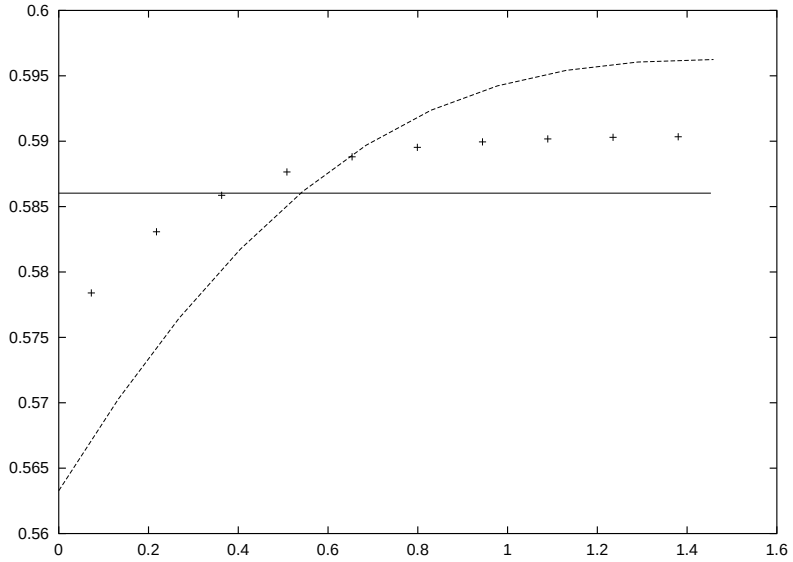


FIG. 6.5.9: VELOCITY (Vertical profiles) - Navier-Stokes and Multilayer Saint-Venant models - Ten layers - Friction coefficient = .001

6.5.8 Robustness of the scheme

We now discuss the robustness of our modified layer-by-layer approach (6.3.23)-(6.3.28) when the initial hypothesis of Property (P3) in Proposition 6.3.4 is violated. The worst case we consider here is when the initial ratio between the water height of one layer and the total water height exhibits a spatial discontinuity.
In [27] such a case occurs. The authors introduce a Q-scheme to study a bi-fluid 1D shallow water system. In particular they compare global upwinding and uncoupled upwinding approachs. If we assume that the densities of the two fluids are equal,

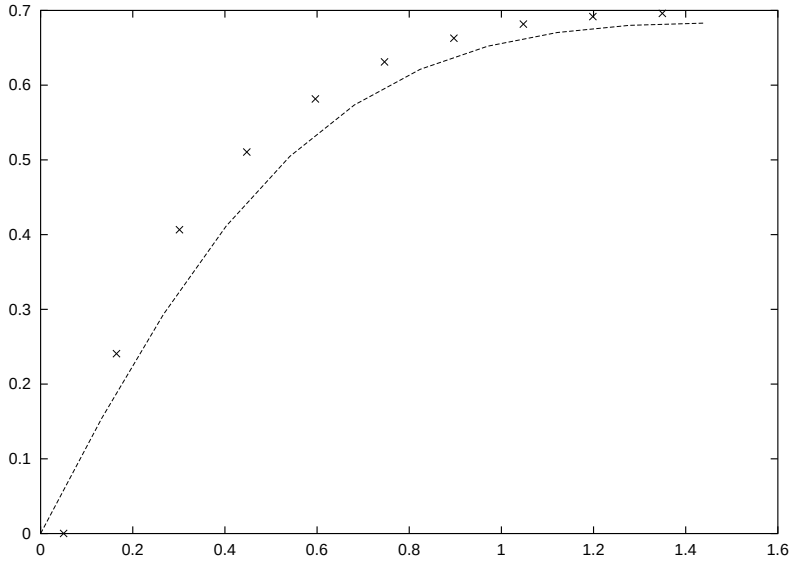


FIG. 6.5.10: VELOCITY (Vertical profiles) - Navier-Stokes and Multilayer Saint-Venant models - Ten layers - No slip

the "global upwinding" consists in dealing with the 4×4 non conservative system (6.3.6)-(6.3.11) whereas the "uncoupled upwinding" corresponds to a basic layer-by-layer approach - see Remark 6.3.3. A numerical example, that involves a discontinuity in the initial interface, exhibits that the "uncoupled upwinding" strategy leads to large oscillations in the solution - for both interface and free surface.
Here we are interested in studying our numerical method for such a test case. Thus we consider the same initial data (except that in our simulations the two densities are equal) and we do not consider any viscous stabilizing effects (i.e. that viscosity and friction coefficients are equal to zero). Then we compare in Figure 6.5.11 the results obtained with the basic layer-by-layer approach and with our modified layer-by-layer approach - see Section 6.4. It appears that our approach allows to avoid such oscillations. In spite of a large number of test cases we have performed, none of them exhibited any numerical instabilities despite the non-hyperbolicity of the system (6.3.6)-(6.3.11).

6.6 Conclusion and perspectives

Thanks to a formal asymptotic analysis of the full Navier-Stokes equations for incompressible flows with a free moving boundary under the shallow water assumption, we derive a multilayer Saint-Venant model which allows a non constant vertical profile while preserving the computational efficiency of the classical Saint-Venant model. The model has the same range of validity as the Navier-Stokes hydrostatic model.
This multilayer model verifies some essential stability properties - positivity of the

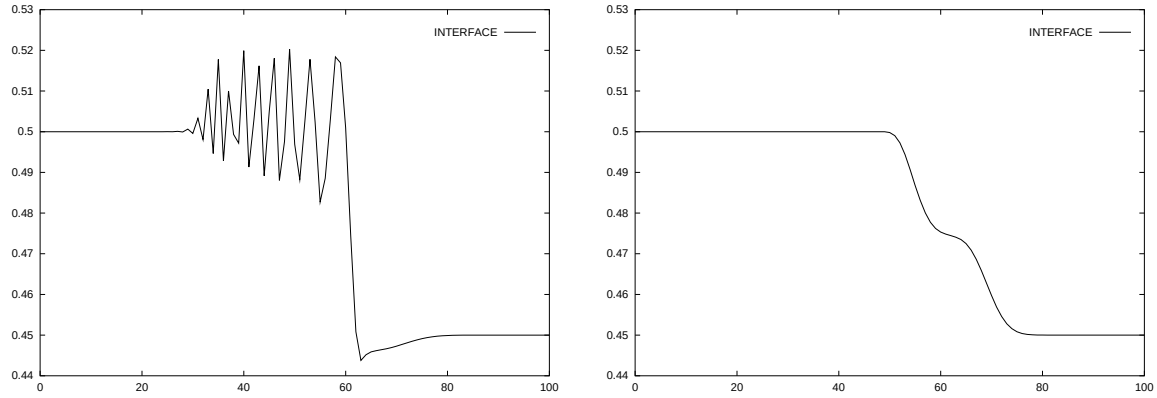


FIG. 6.5.11: Interface profiles for a bilayer test case - The basic (left) and the modified (right) layer-by-layer approachs

water height, existence of a non increasing energy, hyperbolicity of the conservative part. It ensures also the preservation of some physical requirements as the preservation of steady states or the accordance with the Saint-Venant equation and/or with the Navier-Stokes hydrostatic model for some particular flows.
Numerical comparisons with classical Saint-Venant and full Navier-Stokes computations on a dam break problem validate the model. In particular it appears that the vertical velocity profiles issued from the multilayer Saint-Venant and full Navier-Stokes computations are very close.
Some questions remain open : accordance with the vertical profile of the velocity issued from the Navier-Stokes hydrostatic model - to discriminate the influences of the hydrostatic approximation and of the discretization "à la Saint-Venant", preservation of steady states with topographic terms - this specific problem is well-known to be an essential property of the classical Saint-Venant solvers [7, 13, 62].

This work was partially supported by EDF/LNHE, by HYKE European programme HPRN-CT-2002-00282 (http ://www.hyke.org) and by the ACI Modélisation de processus hydrauliques à surface libre en présence de singularités
(http ://www-rocq.inria.fr/m3n/CatNat/).

The author thanks Marie-Odile Bristeau, Jean-Frédéric Gerbeau and Benoit Perthame for helpful discussions. Navier-Stokes results have been computed by Jean-Frédéric Gerbeau.

Bibliographie

[1] Abgrall R., On essentially non-oscillatory schemes on unstructured meshes : analysis and implementation, *J. Comput. Phys.*, 114 (1994), no. 1, 45–58.

[2] Adimurthi, Jaffre J. & Gowda G.D. Veerappa, Godunov-type methods for conservation laws with a flux function discontinuous in space, *SIAM J. Num. Anal.*, 42 (2004), no.1, 179–208.

[3] Alcrudo F. & Garcia-Navarro P., A High-Resolution Godunov-type Scheme in Finite Volumes for the 2D Shallow-water Equations, *Int. J. for Numerical Methods in Fluids*, 16 (1993), 489–505 .

[4] Anderson J.D., *A History of Aerodynamics*, Cambridge (1997) [cited in http ://www-gap.dcs.st-and.ac.uk/ history/Mathematicians/Saint-Venant.html].

[5] Angrand F., Dervieux A., Boulard V., Periaux J. & Vijayasundaram G., Transonic Euler simulation by means of finite element explicit schemes,*AIAA-83* (1984).

[6] Audusse E., A multilayer Saint-Venant model, *to appear in DCDS-B* (2004).

[7] Audusse E., Bouchut F., Bristeau M.O., Klein R. & Perthame B., A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows, *SIAM J. Sc. Comp.*, 25 (2004), no.6, 2050-2065.

[8] Audusse E. & Bristeau M.O., Transport of pollutant in shallow water, a two time steps kinetic method, *M2AN Math. Mod. Num. Anal.*, 37 (2003), no.2, 389–416.

[9] Audusse E., Bristeau M.O. & Perthame B., Kinetic schemes for solving Saint-Venant equations with source terms, *Inria report*, RR-3989 (2000), http ://www.inria.fr/RRRT/RR-3989.html

[10] Audusse E. & Bristeau M.O, A well-balanced positivity preserving second order scheme for shallow water flows on unstructured grids, *submitted*.

[11] Bale D.S., Leveque R.J., Mitran S. & Rossmanith J.A., A wave propagation method for conservation laws and balance laws with spatially varying flux functions, *SIAM J. Sc. Comp.*, 24 (2002), no.3, 955–978.

[12] Bermudez A., Dervieux A., Desideri J.A. & Vazquez M.E., Upwind schemes for the two-dimensional shallow water equations with variable depth using unstructured meshes, *Comput. Methods Appl. Mech. Engrg.*, 155 (1998), no. 1-2, 49–72.

[13] Bermudez A. & Vazquez M.E., Upwind methods for hyperbolic conservation laws with source terms, *Comput. Fluids*, 23 (1994), no.8, 1049–1071.

[14] Besson O. & Laydi M.R., Some estimates for the anisotropic Navier-Stokes equations and for the hydrostatic approximation, *RAIRO M2AN*, 26 (1992), no.7, 855–865.

[15] Botchorishvili R., Perthame B. & Vasseur A., Equilibrium schemes for scalar conservation laws with stiff sources, *Math. Comp.*, 72 (2003), no.241, 131–157.

[16] Botta N., Klein R., Langenberg S. & Lützenkirchen S., Well-balanced finite volume methods for nearly hydrostatic flows, *J. Comp. Phys.*, 196 (2004), no.2, 539–565.

[17] Botta N., Klein R. & Owinoh A., Distinguished Limits, Multiple Scales Asymptotics, and Numerics for Atmospheric Flows, *Amer. Meteorological Society, 13th Intl. Conf. on Atmosphere-Ocean Fluid Dynamics*, Breckenridge, Colorado, July (2001).

[18] Bouchut F., *Nonlinear stability of finite volume methods for hyperbolic conservation laws, and well-balanced schemes for sources*, Frontiers in Mathematics, Birkhauser (2004).

[19] Bouchut F., Construction of BGK models with a family of kinetic entropy equations for a given system of conservation laws. *J. Stat. Phys.*, 95 (1999), no. 1-2, 113–170.

[20] Bouchut F. & James F., One dimensional transport equation with discontinuous coefficient, *Nonlinear Anal.*, 32 (1998), no.7, 891–933

[21] Bouchut F., LeSommer J. & Zeitlin V., Frontal geostrophic adjustment and nonlinear-wave phenomena in one dimensional rotating shallow water ; Part 2 : high resolution numerical investigation, *J. Flu. Mech.*, 514 (2004), 35–63.

[22] Bouchut F., Mangeney-Castelnau A., Perthame B. & Vilotte J.P., A new model of Saint-Venant and Savage-Hutter type for gravity driven shallow water flows, *C. R. Math. Acad. Sci. Paris*, 336 (2003), no.6, 531–536.

[23] Bristeau M.O. & Coussin, B., Boundary Conditions for the Shallow Water Equations solved by Kinetic Schemes, *INRIA Report*, RR-4282 (2001), http ://www.inria.fr/RRRT/RR-4282.html

[24] Bristeau M. O. & Perthame B., Transport of Pollutant in Shallow Water using Kinetic schemes, *ESAIM Proceedings, 10*, CEMRACS 1999, 9–21, http ://www.emath.fr/Maths/Proc/Vol.10.

[25] Burger R., Karlsen K.H., Klingenberg C. & Risebro N.H., A front tracking approach to a model of continuous sedimentation in ideal clarifier-thickener units, *Nonlinear Anal. Real World Appl.*, 4 (2003), no. 3, 457–481.

[26] Cargo P. & LeRoux A.Y., Un schéma équilibre adapté au modèle d'atmosphère avec termes de gravité [A well-balanced scheme for a model of an atmosphere with gravity], *C. R. Acad. Sci. Paris Sr. I Math.*, 318 (1994), no.1, 73–76 [in french].

[27] Castro M.J., Macias J. & Pares 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*, 35 (2001), no.1, 107–127.

[28] Castro M.J., Garcia-Rodriguez J.A., Gonzanlez-Vida J.M., Macias J., Pares C. & Vazquez-Cendon E., Numerical Simulation of two layer shallow water flows through channels with irregular geometry, *JCP*, 195 (2004), no.1, 202–235.

[29] Castro M.J. & Pares C. On the well balanced property of Roe's method for non-conservative hyperbolic systems. Applications to shallow water systems, *submitted.*

[30] Causin P., Miglio E. & Saleri F., Algebraic factorizations for 3D non-hydrostatic free surface flows, *Comput. Vis. Sci.*, 5 (2002), no.2, 85–94.

[31] Cercignani C., *The Boltzmann equation and its applications*, Springer Verlag (1994).

[32] Chacon Rebollo T. & Guillen Gonzalez F., An intrinsic analysis of existence of solutions for the hydrostatic approximation of Navier-Stokes equations, *C. R. Acad. Sci. Paris, Serie 1*, 330 (2000), 841–846.

[33] Chacon Rebollo T., Delgado A.D. & Nieto E.D.F., An entropy-correction free solver for non homogeneous shallow water equations, *M2AN*, 37 (2003) 363–390.

[34] Chacon Rebollo T., Delgado A.D. & Nieto E.D.F., A family of stable numerical solvers for the shallow water equations with source terms, *Comp. Meth. Appl. Math. Engin.*, 192 (2003) 203–225.

[35] Chertock A. & Kurganov A., On a hybrid final-volume-particle method, *to appear in M2AN* (2003).

[36] Chertock A., Kurganov A. & Petrova G., Finite-volume-particle methods for models of transport of pollutant in shallow water, *to appear in J. Sci. Comput.* (2003).

[37] Dafermos C.M., *Hyperbolic conservation laws in continuum physics*, Springer Verlag, GM 325 (1999).

[38] Dal Maso G., Lefloch P. & Murat F., Definition and weak stability of nonconservative products, *J. Math. Pures Appl.*, (9) 74 (1995), no. 6, 483–548.

[39] Dawson C.N. & Proft J., Coupling of continuous and discontinuous Galerkin methods for transport problems, *Computer Methods in Applied Mechanics and Engineering*, 191 (2002), no. 29-30, 3213–3231.

[40] Dawson C.N. & Proft J., Discontinuous and coupled continuous/discontinuous Galerkin methods for the shallow water equations, *Computer Methods in Applied Mechanics and Engineering*, 191 (2002), no. 41-42, 4721–4746.

[41] DELFT-3D-WAQ, *User's Manual*, available at http ://www.wldelft.nl/soft/d3d/support/doc/index.html (2003).

[42] Delis A.I. & Katsaounis T., Relaxation schemes for the shallow water equations, *Internat. J. Numer. Methods Fluids*, 41 (2003), no. 7, 695–719.

[43] Diehl S., On Scalar conservation laws with point source and discontinuous flux functions, *SIAM J. Math. Anal.*, 26 (1995), no. 6, 1425–1451.

[44] Eymard R., Gallouët T. & Herbin R., *Finite volume methods, Handbook of numerical analysis*, vol VIII, P.G. Ciarlet and J.L. Lions editors, Amsterdam, North-Holland, (2000).

[45] Ferrari S. & Saleri F., A new two dimmensionnal shallow water model, *M2AN*, 38 (2004), no.2, 211–234.

[46] Fontana L., Miglio E., Quarteroni A. & Saleri F., A finite element method for 3D hydrostatic water flows, *Comput. Vis. Sci.*, 2 (1999), no.2, 85–93.

[47] Formaggia L., Lamponi D. & Quarteroni A., *One dimensional models for blood flow in arteries*, MOX Report 08 (2002), http ://mox.polimi.it/it/progetti/pubblicazioni/quaderni/mox08.pdf.

[48] Formaggia L., Nobile F. & Quarteroni A., A one dimensional model for blood flow : application to vascular prosthesis, *Mathematical modeling and numerical simulation in continuum mechanics (Yamaguchi, 2000)*, Lect. Notes Comput. Sci. Eng., Springer, Berlin, 19 (2002), 137–153.

[49] Gallouët T., Hérard J.M. & Seguin N., Some approximate Godunov schemes to compute shallow-water equations with topography, *Comput. and Fluids*, 32 (2003), 479–513.

[50] Gerbeau J.-F. & Perthame B., Derivation of Viscous Saint-Venant System for Laminar Shallow Water ; Numerical Validation, *Discrete Cont. Dyn. Syst. Ser. B,*, 1 (2001), no.1, 89–102.

[51] Gill A.E., *Atmosphere Ocean Dynamics*, International Geophysics Series Vol. 30, Academic Press Inc. (1982).

[52] Gimse T. & Risebro N.H., Riemann problems with a discontinuous flux function, *Third International Conference on Hyperbolic Problems, Vol. I, II (Uppsala, 1990)*, Studentlitteratur, Lund, (1991) 488–502.

[53] Gimse T. & Risebro N.H., Solution of the Cauchy problem for a conservation law with a discontinuous flux function, *SIAM J. Math. Anal.*, 23 (1992), no. 3, 635–648.

[54] Godlewski E. & Raviart P.-A., *Numerical approximations of hyperbolic systems of conservation laws*, Applied Mathematical Sciences 118, Springer-Verlag, New York (1996).

[55] Godunov S.K., A difference method for numerical calculation of dicontinuous equations of hydrodynamics, *Mat. Sb.*, (1959), 271–300 [in russian].

[56] Gosse L., A priori error estimate for a well-balanced scheme designed for inhomogeneous scalar conservation laws, *C.R. Acad. Sc. Paris*, 327 (1998), 467 – 472.

[57] Gosse L., A well-balanced flux-vector splitting scheme designed for hyperbolic systems of conservation laws with source terms, *Comp. Math. Appl.*, 39 (2000), 135–159.

[58] Gosse L., A well balanced scheme using nonconservative products designes for hyperbolic systems of conservation laws with source terms, *Math. Mod. Meth. Appl. Sci.*, 11 (2001), no. 2, 339–365.

[59] Gosse L. & Leroux A.-Y., A well-balanced scheme designed for inhomogeneous scalar conservation laws, *C. R. Acad. Sc., Paris, Sér. I* , 323 (1996), 543–546.

[60] Goutal N. & Maurel F. Proceedings of the 2nd workshop on dam-break simulation, Note technique EDF, HE-43/97/016/B (1997).

[61] Gray J.M.N.T., Tai Y.C. & Noelle S., Shock waves, dead zones and particle-free regions in rapid granular free-surface flows, *J. Fluid Mech.*, 491 (2003), 161–181.

[62] Greenberg J.M. & Leroux A.-Y., A well balanced scheme for the numerical processing of source terms in hyperbolic equations, *SIAM J. Num. Anal.*, 33 (1996), no.1, 1–16.

[63] Greenberg J.M., Leroux A.-Y., Baraille R. & Noussair A., Analysis and approximation of conservation laws with source terms,*SIAM J. Num. Anal.*, 34 (1997), no.5, 1980–2007.

[64] Guillard H. & Abgrall R.,*Modélisation numérique des fluides compressibles*, Series in Applied Mathematics, Gauthier-Villars, (2001) [in french].

[65] Guillen Gonzalez F., Masmoudi N. & Bellido R., Anisotropic estimates and strong solutions of the primitive equations, *Diff. Int. Equ.*, 14 (2001), no.11, 1381–1408.

[66] Hervouet J.M., *Hydrodynamique des écoulements à surface libre, modélisation numérique avec la méthode des éléments finis*, Presses des Ponts et Chaussées (2003) [in french].

[67] Hervouet J.M., A high resolution 2D dam break model using parallelization, *Hydrological Processes*, 14 (2000), no. 13, 2221–2230.

[68] Hervouet J.M., Janin J.M., Lepeintre F. & Pechon P., TELEMAC 3D : a finite element software to solve the 3D free surface flow problems, *International Conference on Hydroscience and Engineering*, Washington, USA, (1993).

[69] Isaacson E. & Temple B., Nonlinear reosnnance in systems of conservation laws, *SIAM J. Appl. Math.*, 52 (1992), no. 5, 1260–1278.

[70] Isaacson E. & Temple B., Convergence of the 2×2 Godunov method for a general resonant nonlinear balance law, *SIAM J. Appl. Math.*, 55 (1995), no. 3, 625–640.

[71] Jin S., A steady state capturing method for hyperbolic systems with geometrical source terms, *M2AN*, 35 (2001), 631–645.

[72] Karlsen K.H., Klingenberg C. & Risebro N.H., A relaxation scheme for conservation laws with a discontinuous coefficient, *Math. Comp.*, 73 (2004), no. 247, 1235–1259.

[73] Karlsen K.H., Risebro N.H. & Towers J.D., Upwind difference approximations for degenerate parabolic convection-diffusion equations with a discontinuous coefficient, *IMA J. Numer. Anal.*, 22 (2002), no. 4, 623–664.

[74] Karlsen K.H., Risebro N.H. & Towers J.D., On a nonlinear degenerate parabolic transport-diffusion equation with a discontinuous coefficient, *Electron. J. Differential Equations*, (2002), no.93, 1–23.

[75] Karlsen K.H., Risebro N.H. & Towers J., L^1 Stability for entropy solutions of non linear degenerate parabolic convection diffusion equations with discontinuous coefficient, *Skr. K. Nor. Vid. Selsk*, (2003) 49 pp.

[76] Katsaounis T. & Simeoni C., First and second order error estimates for the Upwind Interface Source method. *Math. Comp.*, 74 (2005), 103–122.

[77] Katsaounis T., Perthame B. & Simeoni C., Upwinding Sources at Interfaces in conservation laws, *to appear in Appl. Math. Lett.*, (2004).

[78] Katsaounis T. & Makridakis C., Relaxation Models and Finite Element Schemes for the Shallow Water Equations, *Hyperbolic Problems : Theory, Numerics, Applications*, Springer (2003), 621–631.

[79] Khobalatte B., Résolution numérique des équations de la mécanique des fluides par des méthides cinétiques, *PhD. Thesis, Université P. & M. Curie (Paris 6)*, (1993) [in french].

[80] Klausen R.A. & Risebro N.H., Stability of conservation laws with discontinuous coefficients, *J. Diff. Equ.*, 157 (1999), no.1, 41–60.

[81] Klingenberg C. & Risebro N.H., Convex conservation laws with discontinuous coefficients : Existence, uniqueness and asymptotic behavior, *Comm. Part. Diff. Equ.*, 20 (1995), no. 11-12, 1959–1990.

[82] Klingenberg C. & Risebro N.H., Stability of a resonnant system of conservation laws modeling polymer flow with gravitation, *J. Diff. Equ.*, 170 (2002), no. 2, 344–380.

[83] Kruzkov S.N., First order quasilinear equations in several independent variables, *Mat. Sb.*, 10 (1970), 217–243.

[84] Kurganov A. & Levy D., Central-upwind schemes for the Saint-Venant system, *M2AN*, 36 (2002), 397–425.

[85] Kurganov A. & Petrova G., Central schemes and contact discontinuities, *M2AN*, 34 (2000), no. 6, 1259–1275.

[86] Lazzaroni E., Approssimazione numerica di modelli multistrato per fluidi a superficia libera, *Tesi di laurea, Universita degli studi di Milano*, (2002) [in italian].

[87] LeRoux A.Y., Riemann solvers for some hyperbolic problems with a source term, *ESAIM proc.*, 6 (1999), 75–90.

[88] LeRoux A.Y., Discrétisation des termes sources raides dans les problèmes hyperboliques. In : Systèmes hyperboliques : Nouveaux schémas et nouvelles applications. *Ecoles CEA-EDF-INRIA 'problèmes non linéaires appliqués'*, INRIA Rocquencourt (France), March 1998 [in french]. Available from http ://www-gm3.univ-mrs.fr/~leroux/publications/ay.le_roux.html.

[89] LeVeque R.J., *Numerical Methods for Conservation Laws, Lectures in Mathematics*, ETH Zurich, Birkhauser (1992).

[90] LeVeque R.J., *Finite Volume Methods for Hyperbolic Problems*, Cambridge University Press (2002).

[91] LeVeque R.J., Balancing source terms and flux gradients in high-resolution Godunov methods : the quasi-steady wave-propagation algorithm. *J. Comput. Phys.*, 146 (1998), no. 1, 346–365.

[92] LeVeque R.J. & Yee H.C., A study of numerical methods for hyperbolic conservations laws with stiff source terms, *J. Comp. Phys.*, 86 (1990), 187–210.

[93] LeVeque R.J., Wave propagation algorithms for multi-dimensional hyperbolic systems, *J. Comp. Phys.*, 131 (1997), 327–353.

[94] Lions P.L., *Mathematical Topics in Fluid Mechanics Vol. 1*, Oxford University Press (1996).

[95] Lions P.L., Perthame B. & Souganidis P.E., Existence of entropy solutions for the hyperbolic systems of isentropic gas dynamics in Eulerian and Lagrangian coordinates, *Comm. Pure Appl. Math.*, 49 (1996), no. 6, 599–638.

[96] Lions P.L., Perthame B. & Tadmor E., Existence of entropy solutions to isentropic gas dynamics system in Eulerian and Lagrangian variables, *Comm. Math. Phys.* , 163 (1994), 415–431.

[97] Liu T.P., Nonlinear resonance for quasilinear hyperbolic equation, *J. Math. Phys.*, 28 (1987), 2593–2602.

[98] Mangeney A., Vilotte J.P., Bristeau M.O., Perthame B., Simeoni C. & Yernini S., Numerical modeling of avalanches based on Saint-Venant equations using a kinetic scheme, *submitted* (2004).

[99] Martin L., Fonctionnement écologique de la Seine à l'aval de la station d'épuration d'Achères : données expérimentales et modélisation bidimensionnelle, *Phd thesis, Ecole des Mines de Paris* (2001) [in french].

[100] Miglio E., Quarteroni A. & Saleri F., Finite element approximation of quasi-3D shallow water equations, *Comput. Methods Appl. Mech. Engrg.*, 174 (1999), no. 3-4, 355–369.

[101] Monthe L.A., Benkhaldoun F. & Elmahi I., Positivity preserving finite volume Roe schemes for transport-diffusion equations, *Comput. Methods Appl. Mech. Engrg.*, 178 (1999), 215–232.

[102] Navier C., Mémoire sur les lois du mouvement des fluides, *C. R. Acad. Sc., Paris*, 6 (1823), 389–416 [in french].

[103] Ostrov D.N., Solutions of Hamilton-Jacobi equations and scalar conservation laws with discontinuous space-time dependence, *J. Differential Equations*, 182 (2002), no. 1, 51–77.

[104] Perthame B., Boltzmann type schemes and the entropy condition, *SIAM J. on Num. Anal.*, 27 (1990), no.6, 1405–1421.

[105] Perthame B., *Kinetic formulations of conservation laws*, Oxford University Press (2002).

[106] Perthame B., An introduction to kinetic schemes for gas dynamics, *An introduction to recent developments in theory and numerics for conservation laws (Freiburg/Littenweiler, 1997)*, Lect. Notes Comput. Sci. Eng., 5, Springer, Berlin, (1999), 1–27.

[107] Perthame B. & Qiu Y., A variant of Van Leer's method for multidimensional systems of conservation laws, *J. Comp. Phys.*, 112 (1994), no.2, 370–381.

[108] Perthame, B. & Simeoni, C., A kinetic scheme for the Saint-Venant system with a source term, *Calcolo*, 38 (2001), no.4, 201–231.

[109] Perthame B. & Simeoni C., Convergence of the upwind Interface source method for hyperbolic conservation laws, *Proceedings of Hyp2002, T. Hou and E. Tadmor editors*, Springer, (2003).

[110] Proft J., Multi-Algorithmic Numerical Strategies for the Solution of Shallow Water Models, *PhD Thesis, TICAM Report 02-41* (2002).

[111] Qiu Y., Etude des équations d'Euler et de Boltzman et de leur couplage. Application à la simulation numérique d'écoulements hypersoniques de gaz raréfiés, *PhD. Thesis, Université P. & M. Curie (Paris 6)* (1993) [in french].

[112] Roe P.L., Approximate Riemann solvers, parameter vectors and difference schemes, *J. Comp. Phys.*, 43 (1981), 357–372.

[113] Roe P.L., Upwind differencing schemes for hyperbolic conservation laws with source terms, *Nonlinear Hyperbolic Problems*, C.Carasso, P.A.Raviart and D.Serre editors, Lecture Notes in Math., vol. 1270, Berlin, Springer-Verlag, (1987), 41–51.

[114] Russo G., Central schemes for balance laws, *Hyperbolic problems : theory, numerics, applications, Vol. I, II (Magdebourg, 2000)*, Internat. Ser. Numer. Math., 140-141, Birkhauser, Basel (2001), 821–829.

[115] de Saint Venant A.J.C., Théorie du mouvement non-permanent des eaux, avec application aux crues des rivières et à l'introduction des marées dans leur lit, *C. R. Acad. Sc., Paris*, 73 (1871), 147–154 [in french].

[116] Sanders R. & Prendergast K.H., The possible relations of the three kiloparsec arm to explosions in the galactic nucleus, *Astrophysical Journal*, 188 (1974).

[117] Seguin N. & Vovelle J., Analysis and aproximation of a scalar conservation law with a flux function with discon tinuous coefficients, *Math. Mod. Meth. App. Sci.*, 13 (2003), no. 2, 221–257.

[118] Serre D., *Systèmes hyperboliques de lois de conservation, Parties I et II.*, Diderot, Paris (1996).

[119] Steger J. & Warming R.F., Flux vector splitting of the inviscid gas dynamics equation with application to finite difference methods, *J. Comp. Phys.*, 40 (1981), 263–293.

[120] Stoker J.J., The formation of breakers and bores, *Communication on Applied Mathematics*, 1 (1948), no. 1, 1–87.

[121] Stoker J.J., *Water waves, the mathematical theory with applications*, Wiley, (1958).

[122] Stokes G., On the theories of the internal friction of fluids motion and of the equilibrium and motion of elastic solids, *Trans. Cambridge Phil. Soc.*, 8 (1845), 287–305.

[123] Temple B., Global solution of the Cauchy problem for a class of 2×2 nonstrictly hyperbolic conservation laws, *Adv. in Appl. Math.*, 3 (1982), no. 3, 335–375.

[124] Toro E.F., *Riemann solvers and numerical methods for fluid dynamics. A practical introduction*, Second edition, Springer-Verlag, Berlin (1999).

[125] Towers J., Convergence of a difference scheme for conservation laws with a discontinuous flux, *SIAM J. Num. Anal.*, 38 (2000), no.2, 681–698.

[126] Towers J., A difference scheme for conservation laws with a discontinuous flux : The non-convex case, *SIAM J. Num. Anal.*, 39 (2001), no.4, 1197–1218.

[127] Van Leer B., Towards the Ultimate Conservative Difference Schemes. V. A Second Order Sequel to the Godunov's Method, *J. Comp. Phys.*, 32 (1979), 101–136.

[128] Van Leer B., Flux Vector Splitting for the Euler equations, *Proc. 8th International Conference on numerical methods in fluids dynamics*, Berlin, Springer-Verlag (1982), 507–512.

[129] Vazquez-Cendon M.E., Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry, *J. Comput. Phys.*, 148 (1999), 497–526.

[130] Whitham G.B., *Linear and non linear waves*, John Wiley and Sons Inc., New York (1999).

[131] Xu K., A well-balanced gas-kinetic scheme for the shallow water equations with source terms, *J. Comput. Phys*, 178 (2002), no.2, 533–562.

[132] Zeitlin V., Medvedev S.B. & Plougonven R., Frontal geostrophic adjustment, slow manifold and nonlinear wave phenomena in 1d rotating shallow water. Part 1. Theory, *Journal of Fluid Mechanics*, 481 (2003), 269–290.

Modélisation hyperbolique et analyse numérique pour les écoulements en eaux peu profondes

Résumé : Nous étudions dans cette thèse différentes lois de conservation hyperboliques associées à la modélisation des écoulements en eaux peu profondes.
Nous débutons par l'analyse numérique du système de Saint-Venant avec termes sources et présentons un schéma volumes finis bidimensionnel d'ordre 2, conservatif et consistant, qui s'appuie sur une interprétation cinétique du système et une reconstruction hydrostatique par interface, et préserve la positivité de la hauteur d'eau et l'état stationnaire du lac au repos. Une extension au couplage avec une équation de transport, utilisant deux pas de temps différents, est décrite.
Nous établissons ensuite un nouveau modèle de Saint-Venant multicouche. Dérivation, étude de stabilité, relations avec les autres modèles fluides et mise en oeuvre numérique sont présentées.
Nous terminons sur un résultat général d'unicité pour les lois de conservation scalaires avec flux discontinus. La preuve, basée sur l'utilisation d'une nouvelle famille d'entropies, permet de lever certaines hypothèses classiques sur le flux et ne nécessite pas l'introduction d'une condition d'interface.

Mots clés : système de Saint-Venant, termes sources, modèle multicouche, équation de transport, loi de conservation scalaire, flux discontinus, volumes finis, schéma équilibre, interprétation cinétique, reconstruction hydrostatique, entropies de Kruzkov.

Hyperbolic Models and Numerical Analysis for Shallow Water Flows

Abstract : In this work we study some hyperbolic conservation laws related to shallow water flows.
First we consider the Saint-Venant system with source terms and we present a second order bidimensional finite volume scheme that is based on a kinetic interpretation of the system and on a interface hydrostatic reconstruction. The scheme is consistent and conservative and it preserves the non-negativity of the water height and the lake at rest steady state. An extension with two different times steps is used to treat the coupling with a transport equation.
Then we study a new multilayer Saint-Venant system. We present its derivation and a stability analysis. We also investigate its relation with other fluid models and we perform its numerical implementation.
Finally we prove a uniqueness theorem for scalar conservation laws with discontinuous flux. Our proof uses a new family of entropies and avoids the making of some classical hypothesis on the flux and the introduction of an interface condition.

Key words : shallow water equations, source terms, multilayer model, transport equation, scalar conservation law, discontinuous flux, finite volumes, well-balanced scheme, kinetic interpretation, hydrostatic reconstruction, Kruzkov's entropies.

Mathematics Subject Classification (2000) : 35L65, 35Q35, 65M12, 76M12.