Problèmes d’interfaces et couplages singuliers dans les systèmes hyperboliques : analyse et analyse numérique

Nina Aguillon

THÈSE

présentée pour obtenir

LE GRADE DE DOCTEUR EN SCIENCES
DE L'UNIVERSITÉ PARIS SUD

Spécialité : Mathématiques
École Doctorale 142 : mathématiques de la région Paris-Sud

par

Nina AGUILLON

Problèmes d'interfaces et couplages singuliers dans les systèmes hyperboliques : analyse et analyse numérique

soutenue le 29 septembre 2014 devant la commission d'examen :

Directeur de thèse	M. Frédéric LAGOUTIÈRE
Rapporteurs	M. François BOUCHUT M. Franck BOYER
Examinateurs	Mme Sylvie BENZONI-GAVAGE Mme Claire CHAINAIS-HILLAIRET M. Frédéric ROUSSET

Nina AGUILLON
nina.aguillon@math.u-psud.fr
http://www.math.u-psud.fr/~aguillon

Thèse préparée au

Département de Mathématiques d'Orsay
Laboratoire de Mathématiques (UMR 8628), Bât. 425
Université Paris-Sud XI
F-91 405 Orsay CEDEX

À mes grands-parents Jean, Nicole, Agnès et Raoul,

en souvenir des repas du mardi soir.

Problèmes d'interfaces et couplages singuliers dans les systèmes hyperboliques : analyse et analyse numérique

DANS CE TRAVAIL, nous nous intéressons à deux problèmes de la théorie des systèmes hyperboliques faisant intervenir des interfaces. Le premier concerne des modèles de couplages entre un fluide compressible et une particule ponctuelle et le second concerne la capture numérique précise des chocs, ces discontinuités qui apparaissent dans les solutions des systèmes hyperboliques.

Sur la première thématique, nous commençons par introduire les différents modèles, dans lesquels la particule et le fluide interagissent à travers une force de frottement qui tend à rapprocher leurs vitesses. Le couplage est singulier car il fait intervenir le produit d'une fonction discontinue par une mesure de Dirac. On peut toutefois définir précisément le système en voyant la particule comme une interface à travers laquelle des relations liant les propriétés du fluide et celle de la particule sont imposées. Lorsque le fluide suit une équation de Burgers, nous démontrons la convergence d'une classe de schéma numérique, et nous obtenons l'existence d'une solution au problème de Cauchy pour une donnée initiale à variation totale bornée. Dans le cas plus complexe où le fluide est décrit par les équations d'Euler isothermes, on prouve l'existence et l'unicité d'une solution autosemblable au problème de Riemann lorsque la particule est immobile. Des simulations numériques sont également présentées.

La dernière partie de la thèse est consacrée à la construction de schémas non diffusifs pour les systèmes hyperboliques. Ces schémas, de type volumes finis, sont construits pour être exact lorsque la donnée initiale est un choc isolé. Ils sont basé sur une reconstruction discontinue de la solution au début de chaque itération en temps, dans le but de reconstituer des chocs à l'intérieur de certaines cellules du maillage. Cette stratégie mène à des schémas très peu diffusifs qui, lorsque l'opérateur de reconstruction est bien choisi, approchent correctement les solutions de cas tests problématiques (chocs lents, chocs forts, réflexions pour la dynamique des gaz, chocs non classiques pour les systèmes qui ne sont pas vraiment non linéaires).

Mots-clefs : système hyperbolique, équations d'Euler, couplage fluide-particule, produit non conservatif, schéma de volumes finis, diffusion numérique, choc non classique.

Classification AMS : 35L65, 65M08, 35L67, 76N10.

Interfaces Problems and Singular Couplings in Hyperbolic Systems: Analysis and Numerical Analysis

In this work, we study two problems concerning hyperbolic systems involving interfaces. The first one concerns the study of models of coupling between a compressible fluid and a pointwise particle. The second one deals with the sharp numerical approximation of shocks, which are discontinuities that appear in the solutions of hyperbolic systems.

In the first two parts of the manuscript, we introduce different models of fluid-particle couplings. The fluid and the particle interact on each other through a drag force, which brings their velocities closer to one another. The coupling is singular because it can be written as the product of a discontinuous function by a Dirac measure. However, the system can be precisely defined as follows. The particle is seen as an interface through which interface conditions linking the properties of the fluid with those of the particle are imposed. When the fluid follows the compressible Burgers equations, we prove the convergence of a family of finite volume schemes and obtain the existence of a solution when the initial data has total bounded variation. In the more difficult case where the fluid is described by the isothermal Euler equations, we prove the existence and uniqueness of a selfsimilar solution to the Riemann problem, when the particle is motionless. Numerical experiments are also presented.

In the last part of this work, we build non diffusive numerical schemes for different hyperbolic systems. These finite volume schemes are build to be exact when the initial data is an isolated shock. They are based on a discontinuous reconstruction of the solution at the beginning of each time step, in order to reconstruct shocks inside some specific cells of the mesh. The schemes we present have a very low numerical diffusion and, when the reconstruction operator is well chosen, they are able to correctly approximate the solution on various problematic test cases. These cases include slowly moving shocks, strong shocks and shock reflections for gas dynamics, as well as the apparition of nonclassical shocks for systems that are not truely nonlinear.

Key words and phrases: hyperbolic system, Euler equations, fluid-particle coupling, nonconservative product, finite volume scheme, numerical diffusion, nonclassical shock.

AMS classification: 35L65, 65M08, 35L67, 76N10.

Remerciements

Mes premiers et mes plus chaleureux remerciements vont à mon directeur de thèse Frédéric Lagoutière. Durant cette thèse et mon stage de M2, il fut un directeur fantastique. Je le remercie tout particulièrement pour son enthousiasme et sa disponibilité sans faille durant ces trois années, et pour la pédagogie avec laquelle il partage sa grande culture mathématique et ses nombreuses idées.

Je suis très honorée que François Bouchut et Franck Boyer aient accepté de rapporter mon manuscrit et je les remercie vivement pour leur travail détaillé et la grande attention qu'ils ont portée à mon travail. Merci également à Sylvie Benzoni-Gavage, Claire Chainais-Hillairet et Frédéric Rousset d'avoir montré leur intérêt pour cette thèse en acceptant de faire partie de mon jury.

Durant ces quelques années, j'ai eu la chance et le plaisir de travailler avec Christophe Chalons, Nicolas Seguin, Boris Andreianov, Raul Borsche, Clément Cancès et Pauline Lafitte. Merci à eux de partager mon intérêt pour certaines questions mathématiques : j'ai beaucoup appris à leurs côtés, et j'espère que cela va continuer. Ici encore, je remercie vivement Frédéric qui m'a encouragée à travailler avec les uns et les autres ; ce fut chaque fois une expérience très enrichissante.

Je tiens maintenant à exprimer ma gratitude envers trois personnes sans qui cette thèse aurait probablement été fort différente. Merci à Dominique Hulin pour son tout dernier cours d'équations différentielles en L3 qui a éveillé mon intérêt pour l'analyse numérique, merci à Patrick Gérard de m'avoir convaincue de faire mon M2 à Orsay et merci à Bertrand Maury de m'avoir incitée à faire ma thèse avec Frédéric.

Merci à la direction du laboratoire de mathématiques d'Orsay de m'avoir permis d'effectuer ce travail dans les meilleures conditions et merci à Valérie Blandin-Lavigne et à Catherine Poupon (dont le soutien fut souvent bien plus qu'administratif) d'avoir toujours su répondre promptement à mes questions et de m'avoir aidée à maintes reprises. Je remercie également Pascale Starck et Nathalie Carrièrre pour leur compétence assortie de gentillesse. Je suis également reconnaissante à l'égard de l'Institut Élie Cartan de Lorraine pour m'avoir accueillie régulièrement lors de ces deux dernières années.

J'adresse un grand merci aux membres de l'équipe ANEDP qui ont partagé mon quotidien ces dernières années. La bonne ambiance au travail et la confiance avec laquelle j'ai pu discuter de mathématiques doit beaucoup à leur accueil sympathique. J'ai une pensée particulière pour les doctorants ayant partagé mon bureau ou le bureau voisin et, par conséquent, les pauses et moments de détente des journées de travail. Merci à Tony, Jean, Guilhem, Justine, Anthony, Loïc, Maxime, Alpar, Antonin, Christèle, Yueyuan, Pierre, Nicolas, Benoit, Adeline et Aude, ainsi qu'aux petits nouveaux Clémentine, Fatima et Paul. Je n'oublie pas les autres doctorants du laboratoire (dont je n'oserai pas faire la liste) et j'ai une pensée particulière pour Olivier, Laure, Aurélien et Caroline, qui ont su me mettre à l'aise lorsque je n'étais qu'une timide stagiaire de M2.

Merci à mes amis qui m'ont permis de m'évader de temps en temps. Beaucoup d'entre eux sont matheux également, et c'est tout à leur mérite de me distraire et de me parler d'autre chose lorsque je les vois. Un grand merci à Aurelia qui fut un soutien très important cette année.

Enfin, merci à ma famille, à ma belle-famille, et à Jérémie (qui rentre dans de nombreuses catégories de ces remerciements) pour leur soutien inconditionnel et leur persévérance à s'informer de l'avancée de mes travaux.

Table des matières

1 Introduction

Table des matières

Dans cette thèse sont présentées quelques contributions à l'étude théorique et numérique de problèmes d'interfaces dans les systèmes hyperboliques. Ce type de problèmes apparaît dans la modélisation de nombreux phénomènes physiques, comme le refroidissement des barres de combustibles par de l'eau pressurisée dans les centrales nucléaires ou la transition de phase entre deux arrangements atomiques dans un solide. Les solutions de ce type d'équations peuvent être discontinues même lorsque la donnée initiale est très régulière. Les interfaces considérées seront de deux natures différentes. Dans les parties I et II, elles sont dues à des particules ponctuelles modifiant l'écoulement d'un fluide (liquide ou gaz). Le fluide et la particule interagissent à travers une force de *frottement* qui tend à rapprocher leurs vitesses respectives. Cela modélise, par exemple, l'écoulement d'un fluide à travers un grillage ou la chute d'une bille dans un tube rempli d'air. Dans la partie III, les interfaces sont les discontinuités présentes dans le fluide. Dans chacun des cas, les états de part et d'autre de l'interface sont liés par des relations qui gardent la trace de petites échelles physiques sous-jacentes au modèle.

Les systèmes hyperboliques étudiés dans ce manuscrit correspondent à des systèmes d'équations où des paramètres physiques petits sont supprimés des équations. Par exemple, les équations d'Euler compressibles dont il sera question tout au long de ce mémoire correspondent aux équations de Navier–Stokes compressibles avec une viscosité nulle. Lorsque le paramètre est très petit, on s'attend à ce que les solutions du système hyperbolique limite et du système initial soient qualitativement très proches les unes des autres. Cependant, en négligeant ces termes, on perd certaines propriétés de la solution. Les équations d'Euler n'admettent par exemple pas une unique solution faible. Pour résoudre ce problème, on peut ajouter au système des contraintes issues du système augmenté, qui auraient été oubliées lorsqu'on annule le petit paramètre. Pour les équations d'Euler, c'est une inégalité d'entropie, qui traduit qu'une certaine énergie du système limite est dissipée.

L'ajout d'une inégalité d'entropie dans les équations d'Euler permet de sélectionner une unique solution faible, mais ce n'est pas toujours vrai dans des cas plus généraux. Lorsque deux petites échelles physiques rentrent en compétition, par exemple lorsqu'on a un équilibre entre les effets de diffusion et de capillarité, on dispose d'une seule inégalité d'entropie « physique » qui ne règle pas le problème de la non-unicité des solutions. Il peut être nécessaire d'imposer une relation supplémentaire qui garde une trace de l'équilibre entre ces deux contributions, une *relation cinétique*. Un choc sélectionné par une telle relation est dit non classique car il ne vérifie pas les critères usuels d'Oleinik (dans le cas d'une loi de conservation) ou de Liu (dans le cas d'un système). Il sera question de la capture numérique de tels chocs dans la partie III.

Le même type de questions se pose dans le cadre du couplage entre fluide et particule présenté dans les parties I et II, où le petit paramètre est la taille de la particule. Nous nous intéressons au système limite où la particule est ponctuelle (de taille nulle). Ce système limite est mal défini : un terme source singulier s'écrivant comme le produit d'une fonction discontinue et d'une masse de Dirac apparaît. Pour définir ce terme source, on revient à l'étude du problème avec une particule épaissie (de taille non nulle), ce qui fournit des relations d'interfaces universelles (i.e. qui ne dépendent pas de la taille de la particule) que l'on impose ensuite au travers de la particule ponctuelle. Cela permet de prouver l'existence et l'unicité au problème de Riemann (cas d'une donnée initiale en marche d'escalier) lorsque la force de frottement entre le fluide et la particule possède de bonnes propriétés. Ces résultats sont exposés dans le chapitre 4.

Numériquement, considérer les systèmes limites permet de ne pas résoudre les petites échelles physiques sous-jacentes, ce qui peut se révéler très coûteux puisque le pas d'espace doit être d'autant plus petit que le paramètre (taille de particule, viscosité) rentrant en jeu l'est. La question est de savoir si les relations de saut ajoutées aux systèmes (conditions d'interface au travers de la particule, relation cinétique pour les chocs non classiques) peuvent être imposées numériquement. Pour le couplage entre fluide et particule, la résolution explicite du problème de Riemann permet de construire des schémas qui semblent converger ; ces schémas sont présentés dans le chapitre 5. Dans le cas où le fluide suit l'équation plus simple de Burgers, la convergence d'une classe plus large de schémas est démontrée dans le chapitre 3.

Le cas de la capture des chocs non classiques est plus ardu, et cette fois-ci les méthodes basées sur le solveur de Riemann exact échouent. En effet, les schémas numériques usuels contiennent une phase de projection pendant laquelle une moyenne sur chaque maille est effectuée. Cette diffusion numérique détruit l'équilibre entre les effets de dispersion et de diffusion dont la trace était gardée par la relation cinétique, et les schémas numériques correspondant convergent vers la solution classique (qui correspond au cas où la viscosité domine). Une autre anomalie liée à la diffusion numérique est l'apparition d'oscillations parasites de la quantité de mouvement lorsqu'on approche un choc qui se déplace lentement par rapport à la vitesse du son.

Pour réussir à capturer correctement ces chocs (non classiques ou lents), un schéma de volumes finis doit donc être très peu, voire pas du tout, diffusif pour ne pas perturber l'équilibre entre les petites échelles physiques. Or les schémas d'ordre élevé le sont dans les zones où la solution est régulière. Autour d'un choc, pour éviter l'apparition d'oscillations parasites, ils dégénèrent vers des schémas d'ordre 1 et sont donc diffusifs. Les schémas présentés dans la partie III de cette thèse ne sont pas conçus pour être précis dans les zones régulières. En revanche, ils sont exacts lorsque la donnée initiale est un choc isolé. L'idée est de reconstruire chaque valeur moyenne sur une maille comme la moyenne d'un choc situé quelque part à l'intérieur de la maille. Ces schémas, dits de reconstruction discontinue, permettent de capturer les chocs non classiques de l'élastodynamique non linéaire, et d'approcher correctement les chocs lents dans les équations d'Euler.

1.1 Particules ponctuelles dans des fluides compressibles sans viscosité

Les parties I et II de ce manuscrit sont consacrées à l'étude d'une classe de modèles de couplage entre un fluide compressible et une particule ponctuelle, qui interagissent à travers une force de rappel, ou de frottement, qui tend à rapprocher leurs vitesses. La quantité de mouvement du système total (fluide et particule) est conservée, mais un échange s'effectue, au niveau de la particule, entre la quantité de mouvement du fluide et celle de la particule. Par rapport à la vaste littérature s'intéressant au couplage entre les équations de Navier–Stokes et un solide (on pourra consulter par exemple [Hil07a] pour une discussion sur différents paradoxes concernant l'apparition de collisions), ces modèles ont deux originalités : d'une part le fluide considéré est non visqueux, et d'autre part la vitesse de la particule diffère de celle du fluide. Voici quelques situations qui peuvent être décrites par ces modèles : de l'air passant au travers d'une grille, de l'eau pressurisée traversant des barres de combustible dans une centrale nucléaire ou encore une bille chutant dans un tube rempli d'air. Dans chacun de ces cas, considérer l'obstacle comme ponctuel est une simplification qui nous évite d'avoir à en faire une modélisation détaillée. De même, le fait que le fluide et la particule ne partagent pas la même vitesse est douteux quand on considère la situation réelle ; mais « vu de loin », les propriétés du fluide peuvent être différentes à l'avant et à l'arrière de l'obstacle. Lorsque la particule est mobile, son caractère ponctuel nous permet d'écrire très simplement son équation de mouvement grâce au principe fondamental de la dynamique de Newton.

Le fluide considéré étant non visqueux, sa vitesse et sa densité n'ont aucune raison d'être continues au point où se situe la particule (et d'ailleurs elles ne le seront pas). Ainsi on ne peut pas les évaluer en ce point, et toute une partie de notre travail consiste à remplacer des expressions du type « la vitesse du fluide au point où se situe la particule » par « une certaine fonction des caractéristiques du fluide juste devant la particule et de celles juste après ». Pour cela, on étudie un modèle où la particule est épaissie. On remarque alors que la vitesse et la densité du fluide à l'entrée de la particule et à sa sortie sont indépendantes de la manière dont on a épaissi la particule. Cela permet de définir le modèle comme un problème d'interface, à travers laquelle ces relations, qui gardent en mémoire ce qui se passe à l'échelle de la particule, sont imposées.

Ces modèles sont suffisamment simples pour bien étudier l'interaction entre les chocs et la particule, et suffisamment riches pour décrire différentes situations d'intérêt physique (on trouvera plus loin des simulations d'aspirateur bouché, de billes tombant dans des tubes et d'air passant à travers des grillages plus ou moins épais). Nous pouvons étudier entièrement le problème de Riemann, c'est-à-dire le cas où les caractéristiques du fluide sont initialement constantes de part et d'autre de la particule, qu'on astreint à avoir une vitesse constante (dans ce cas le couplage est incomplet, seule la particule a une influence sur le fluide). Notre contribution principale sur ce sujet est un critère d'existence et d'unicité selon la forme du frottement entre le fluide et la particule. Ce résultat est décrit dans le chapitre 4.

Une autre partie de cette thèse concerne l'approximation numérique de ces systèmes, et on peut construire des schémas ayant pour brique élémentaire la résolution du problème de Riemann. Numériquement, la gestion de la particule est délicate, car il faut à la fois imposer des conditions d'interface au travers de l'obstacle et gérer son mouvement. L'obtention de schémas qui se passent de la résolution du problème de Riemann est difficile, nous décrirons dans le chapitre 5 plusieurs tentatives en ce sens. Une fois que les schémas sont construits, il faut démontrer leur convergence vers une solution du problème initial. On est ainsi assuré d'approcher quelque chose de correct, et on obtient au passage l'existence d'une solution. Le chapitre 3 est consacré à une telle preuve lorsque les équations d'Euler sont remplacées par l'équation de Burgers et où la particule est mobile, ce qui constitue une nouveauté par rapport. C'est un travail en commun avec Frédéric Lagoutière et Nicolas Seguin.

1.1.1 Le modèle

Considérons d'une part un gaz dont la densité au point d'espace $x \in \mathbb{R}$ et au temps $t \in \mathbb{R}_+$ est $\rho(t,x)$ et dont la vitesse est $u(t,x)$, et d'autre part une particule ponctuelle, de masse m, dont la position, la vitesse et l'accélération à l'instant t sont notées $h(t)$, $h'(t)$ et $h''(t)$. Dans nos modèles, l'interaction entre le fluide et la particule se fait à travers une force de frottement, de rappel ou de traînée) D. La particule étant ponctuelle, le principe fondamental de la dynamique donne

$$mh''(t) = D(\rho,\alpha)(t,h(t)), \tag{1.1}$$

où α est une quantité de mouvement calculée dans le repère de la particule :

$$\alpha(t,x) = \rho(t,x)\left(u(t,x) - h'(t)\right).$$

Le frottement D dépend de la densité du fluide au point où se situe la particule, et de la quantité α qui est de même signe que la différence entre la vitesse du fluide et celle de la particule. Ainsi si D a le même signe que $u - h'$, ce que nous supposerons dans toute la suite, la particule accélère si elle va moins vite que le fluide (auquel cas $u(t, h(t)) > h'(t)$) et décélère sinon (auquel cas $u(t, h(t)) \leq h'(t)$). Il reste à préciser comment évoluent la densité et la vitesse du fluide. D'après le principe d'action et de réaction (troisième loi de Newton), la particule exerce sur le fluide une force $-D$. Cette force ne s'applique qu'au point où se situe la particule. Par ailleurs, nous nous plaçons dans le cas d'un fluide isotherme, c'est-à-dire pour lequel la pression p est égale à $c^2\rho$, où c est la vitesse du son. L'avantage de ce choix de pression est que le vide n'apparaît pas. Au final, le modèle s'écrit

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p(\rho)) = -D(\rho, \alpha)\delta_{h(t)}(x), \\ mh''(t) = D(\rho, \alpha)(t, h(t)). \end{cases} \tag{1.2}$$

La première équation traduit la conservation de la masse du fluide. La seconde est le principe fondamental de la dynamique appliqué au fluide.

Le système (1.2) est un cas très particulier du modèle de spray (c'est-à-dire d'une suspension diluée de gouttelettes dans un gaz) étudié dans [BD06], lorsque la fonction de distribution de la phase diluée devient une masse de Dirac en espace et en vitesse, c'est-à-dire lorsqu'il n'y a qu'une seule gouttelette ponctuelle. On trouvera dans cet article une discussion et des références bibliographiques sur la forme de la force de frottement D. Les expressions usuelles sont $D(\rho, \rho(u-h')) = C(u-h')$ et $D(\rho, \rho(u-h')) = \rho(u-h')$, où C est une constante positive.

On peut également s'intéresser au modèle où la particule a une vitesse constante fixée $v \in \mathbb{R}$:

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p(\rho)) = -D(\rho, \alpha)\delta_{vt}(x). \end{cases} \tag{1.3}$$

Le système (1.2) (ou (1.3) lorsque la particule est immobile) est une généralisation du couplage introduit pour la première fois dans [LST08]. Dans cet article, le « fluide » est décrit par l'équation de Burgers, et le système étudié s'écrit :

$$\begin{cases} \partial_t u + \partial_x \frac{u^2}{2} = -\lambda(u - h'(t))\delta_{h(t)}(x), \\ mh''(t) = \lambda(u(t, h(t)) - h'(t)), \end{cases} \tag{1.4}$$

où λ est un réel positif représentant un paramètre de friction. Dans la série d'articles [ALST10], [AS12], et [ALST13], les auteurs démontrent que le problème de Cauchy associé à ce système est bien posé. Dans l'article [BCG14], on trouve un autre modèle de couplage entre une particule ponctuelle et les équations d'Euler compressibles, obtenu à partir de considérations sur la conservation totale de la masse totale, de l'impulsion totale et de l'énergie totale. La particularité de ces trois modèles est que la particule et le fluide ne partagent pas la même vitesse, ce qui distingue ces travaux de la large littérature concernant le couplage de fluides et de solides avec conditions d'adhérence sur les parois du solide. En particulier, ces modèles autorisent les collisions entre particules n'ayant pas la même vitesse.

1.1.2 Un système non conservatif, non hyperbolique et non caractéristique

Concentrons-nous d'abord sur le système (1.3) où la particule a une vitesse constante fixée $v \in \mathbb{R}$. Le second membre $D(\rho, \alpha)\delta_{vt}(x)$ est mal défini si D n'est pas continue en $x = vt$. Or des discontinuités sont susceptibles de se créer à l'intérieur du fluide et de heurter la particule à tout moment, et ρ et u seront typiquement discontinues à travers la particule.

Pour illustrer la nécessité de définir le terme source $D(\rho, \alpha)\delta_{vt}(x)$, considérons le cas $v = 0$ et $D(\rho, \alpha) = \lambda\alpha$, où λ est un paramètre de friction positif. La conservation de la masse du fluide au travers de la particule implique que α est constant à travers la particule. Ainsi, avec ce choix de frottement, le terme $\alpha\delta_0$ a un sens. Malgré cela, si on essaye d'approcher directement le système (1.3), on ne parvient

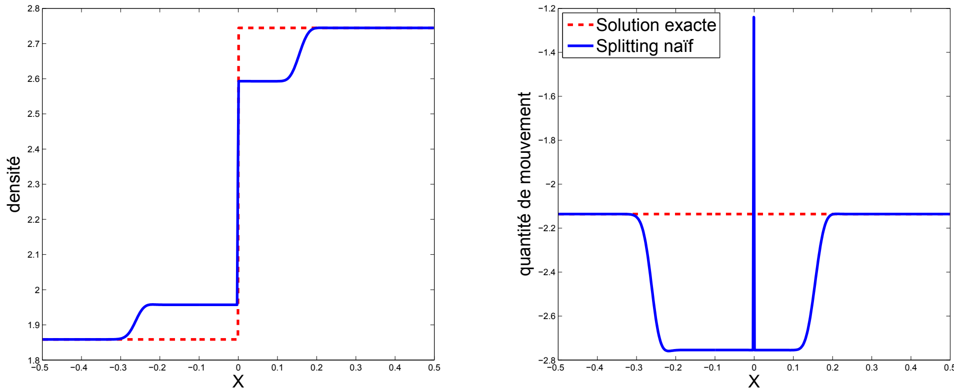


FIGURE 1.1. *Un schéma de splitting naïf entre le fluide et le terme source ne parvient pas à capturer la bonne solution.*

pas à capturer la bonne solution. Par exemple, le schéma de splitting naïf, où le terme d'interaction est traité séparément de la partie fluide, ne converge pas vers la bonne solution. On peut le constater sur la figure 1.1. La solution exacte représentée sur cette figure sera définie dans les parties suivantes.

Le produit non conservatif $D(\rho,\alpha)\delta_{vt}(x)$ (non conservatif car ne pouvant pas être écrit comme une dérivée) peut toutefois être défini de manière précise et intrinsèque comme un problème d'interface grâce à l'étude du système où la particule est épaissie, c'est-à-dire où l'on remplace la masse de Dirac par une de ses régularisations. On trouve alors que la vitesse et la densité en entrée et en sortie de la particule sont toujours liées par les mêmes relations, qui ne dépendent pas de la régularisation choisie. En toute généralité, la définition d'un produit non conservatif dépend du choix d'une régularisation (voir [DMLM95]), et il est remarquable que ça ne soit pas le cas ici.

On introduit la fonction de Heaviside $H(x) = \mathbf{1}_{x\geq 0}$ et on considère $w(t,x) = H(x-vt)$ comme une nouvelle inconnue. Le système se réécrit sous la forme

$$\begin{cases} \partial_t\rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p(\rho)) + D(\rho,\alpha)\partial_x w = 0, \\ \partial_t w + v\partial_x w = 0. \end{cases}$$

Sa forme complètement non conservative est

$$\partial_t \begin{pmatrix} \rho \\ q \\ w \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ c^2-u^2 & 2u & D(\rho,\alpha) \\ 0 & 0 & v \end{pmatrix} \partial_x \begin{pmatrix} \rho \\ q \\ w \end{pmatrix} = 0. \tag{1.5}$$

Un calcul simple montre que les valeurs propres de la matrice

$$\begin{pmatrix} 0 & 1 & 0 \\ c^2-u^2 & 2u & D(\rho,\alpha) \\ 0 & 0 & v \end{pmatrix}$$

sont $u+c$, v et $u-c$, et qu'elle n'est pas diagonalisable lorsque $|u-v| = c$. Ainsi nous avons affaire à un système qui n'est pas hyperbolique. Lorsque deux valeurs propres s'identifient, les ondes associées ne sont plus séparées et le système devient résonnant : deux ondes correspondant à deux valeurs propres différentes peuvent avoir la même vitesse et interagir pendant quelque temps (et pas seulement ponctuellement comme dans le cas des systèmes strictement hyperboliques). L'existence et l'unicité au problème de Riemann ont été étudiées dans [IT92] dans le cas conservatif, et étendu dans [GL04] dans le cas non conservatif, à chaque fois pour des données petites. Les résultats présentés plus loin s'affranchissent de cette hypothèse de petitesse.

Dans le cas du système entièrement couplé (1.2), le fait que D soit discontinue en $x = h(t)$ fait que l'EDO (1.1) décrivant la trajectoire de la particule est, elle aussi, mal définie. Formellement, l'impulsion totale

$$t \mapsto mh'(t) + \int_{\mathbb{R}} (\rho u)(t,x)dx$$

est conservée par les solutions du système (1.2). Un calcul plus rigoureux montre que pour conserver cette propriété, la trajectoire de la particule doit vérifier :

$$mh''(t) = c^2(\rho_-(t) - \rho_+(t)) \left(1 - \frac{(u_-(t) - h'(t))(u_+(t) - h'(t))}{c^2}\right). \tag{1.6}$$

Ici, les indices $+$ et $-$ dénotent les traces à gauche et à droite de la particule respectivement :

$$\rho_-(t) = \lim_{x \to h(t), x<h(t)} \rho(t,x) \quad \text{et} \quad \rho_+(t) = \lim_{x \to h(t), x>h(t)} \rho(t,x)$$

Dans [CG10], R. Colombo et G. Guerra démontrent un résultat d'existence local pour des données initiales petites pour les systèmes de lois de conservation avec des conditions de bord en $x = 0$. Cette analyse est étendue dans [BCG10] et [BCG12] au cas du couplage entre un système hyperbolique et une EDO décrivant le mouvement d'une interface mobile, autour de laquelle sont imposées des conditions d'interfaces. Elle a été appliquée au cas de la dynamique des gaz dans l'article [BCG14] (la modélisation diffère de celle proposée dans cette thèse). Dans chacun des cas, une hypothèse clé est que l'interface n'est pas caractéristique, c'est-à-dire que sa vitesse est strictement séparée de celle des ondes. Cette hypothèse n'est pas remplie lorsque $|u| = c$.

La nouveauté des résultats présentés dans le chapitre 4 est qu'aucune hypothèse n'est faite sur les données initiales : les données seront arbitrairement grandes et la solution pourra être résonnante (toutefois par rapport à [CG10], on se restreint aux données initiales de Riemann).

1.1.3 Définition des solutions et résultats d'existence et d'unicité

Dans cette section sont décrits les principaux résultats que nous avons obtenus sur le couplage fluide-particule (1.2). Comme souligné précédemment, il est nécessaire de donner une définition du membre de droite de (1.2). À cet effet, considérons le cas où la particule a une vitesse constante v et remplaçons la mesure de Dirac par une de ses régularisations. Plus précisément, soit ε un réel strictement positif et H^ε une régularisation de la fonction de Heaviside H telle que :

- H^ε est de classe $\mathcal{C}^\infty$ et croissante ;
- $H^\varepsilon(x) = 0$ si $x < -\varepsilon/2$ et $H^\varepsilon(x) = 1$ si $x > \varepsilon/2$.

On s'intéresse au système avec particule épaissie

$$\begin{cases} \partial_t \rho^\varepsilon + \partial_x q^\varepsilon = 0, \\ \partial_t q^\varepsilon + \partial_x \left(\frac{(q^\varepsilon)^2}{\rho^\varepsilon} + c^2 \rho^\varepsilon\right) = -D(\rho^\varepsilon, \alpha^\varepsilon)(H^\varepsilon)'(x - vt). \end{cases} \tag{1.7}$$

Les propriétés du fluide à gauche et à droite de la particule dans (1.7) sont indépendantes de sa taille ε et de sa « forme » H^ε. Elles sont liées par quelques relations décrites ci-dessous, dans le cas le plus simple où $D(\rho, \alpha) = \lambda\alpha$ avec $\lambda \geq 0$.

Définition 1.1. *On appelle* germe à vitesse *v l'ensemble $\mathcal{G}_D(v)$ contenant toutes les paires d'états $((\rho_-, q_-)$, $(\rho_+, q_+))$ de $(\mathbb{R}_+^* \times \mathbb{R})^2$ qui vérifient les relations suivantes :*

1. *d'une part, $\alpha_- = \alpha_+ := \alpha$;*

2. *d'autre part,* $\left(\frac{\alpha^2}{\rho_-} + c^2\rho_-\right) - \left(\frac{\alpha^2}{\rho_+} + c^2\rho_+\right) = \lambda\alpha$;
3. *enfin, si* $0 \leq u_- - v \leq c$, *alors* $0 \leq u_+ - v \leq c$, *et si* $-c \leq u_+ - v \leq 0$, *alors* $-c \leq u_- - v \leq 0$.

Dans le cas d'une force de frottement D quelconque, les premier et troisième points restent inchangés. Le second point est en revanche modifié. Le théorème suivant précise dans quelle mesure le germe $\mathcal{G}_D(v)$ est lié à (1.7) mais universel par rapport au choix de la régularisation H_ε.

Théorème 1.1. *Soit $v \in \mathbb{R}$ une vitesse de particule fixée. Si $((\rho_-, q_-), (\rho_+, q_+))$ appartient au germe $\mathcal{G}_D(v)$, alors pour tout $\varepsilon > 0$ et pour tout H^ε croissante, de classe $\mathcal{C}^\infty$ et ne variant que sur l'intervalle $[-\varepsilon/2, \varepsilon/2]$, il existe une solution de* (1.7) *ne dépendant que de $\xi = x - vt$, de classe $\mathcal{C}^1$ par morceaux, dont les éventuelles discontinuités sont des chocs entropiques pour les équations d'Euler, et telle que*

$$(\rho^\varepsilon(\xi = -\varepsilon/2), q^\varepsilon(\xi = -\varepsilon/2)) = (\rho_-, q_-) \quad \text{et} \quad (\rho^\varepsilon(\xi = \varepsilon/2), q^\varepsilon(\xi = \varepsilon/2)) = (\rho_+, q_+).$$

Réciproquement, si $(\rho^\varepsilon, q^\varepsilon)$ est une telle solution de (1.7), *alors*

$$((\rho^\varepsilon(\xi = -\varepsilon/2), q^\varepsilon(\xi = -\varepsilon/2)), (\rho^\varepsilon(\xi = \varepsilon/2), q^\varepsilon(\xi = \varepsilon/2))) \in \mathcal{G}_D(v).$$

Ainsi, la vitesse et la densité de part et d'autre de la particule sont indépendantes du choix de ε et de H_ε dans (1.7). Reprenant les trois points de la définition 1.1, on a les propriétés suivantes :

1. la quantité de mouvement α calculée dans le repère de la particule est constante au travers de la particule ;
2. la charge (calculée dans le repère de la particule) $\frac{\alpha^2}{\rho} + c^2\rho$ décroit proportionnellement à α au travers de la particule ;
3. si la vitesse du fluide (calculée dans le repère de la particule) à l'entrée de la particule est subsonique, elle le reste.

Cela nous mène à la définition suivante des solutions, où la particule est vue comme une interface au travers de laquelle sont imposées ces trois relations.

Définition 1.2. *Le couplet $(\rho, q) \in L^\infty(\mathbb{R}_+ \times \mathbb{R}) \times L^\infty(\mathbb{R}_+ \times \mathbb{R})$ est une solution de* (1.2) *pour $h \in W^{2,\infty}(\mathbb{R}_+)$ fixé si pour presque tout temps t, ρ et q admettent des traces $(\rho_-(t), q_-(t))$ et $(\rho_+(t), q_+(t))$ à gauche et à droite de la particule, qui de plus appartiennent au germe :*

$$\text{pour presque tout temps } t,\ ((\rho_-(t), q_-(t)), (\rho_+(t), q_+(t))) \in \mathcal{G}_D(h'(t)),$$

et si (ρ, q) est une solution faible entropique des équations d'Euler sur les ensembles $\{(t, x) : x < h(t)\}$ et $\{(t, x) : x > h(t)\}$. Si la trajectoire de la particule h n'est pas fixée, l'équation différentielle ordinaire (1.6) *doit de plus être vérifiée pour presque tout temps.*

Une étape importante lorsqu'on étudie un système hyperbolique est la résolution du problème de Riemann, c'est-à-dire du cas où la donnée initiale est en marche d'escalier. Dans le cas d'une particule à vitesse fixe v, ce problème de Riemann s'écrit :

$$\begin{cases} \partial_t \rho + \partial_x q = 0, \\ \partial_t q + \partial_x \left(\frac{q^2}{\rho} + c^2\rho\right) = -D(\rho, \alpha)\delta_{vt}, \\ \rho(0, x) = \rho_L \mathbf{1}_{x<0} + \rho_R \mathbf{1}_{x>0}, \\ q(0, x) = q_L \mathbf{1}_{x<0} + q_R \mathbf{1}_{x>0}. \end{cases} \tag{1.8}$$

Le fait que la trajectoire soit une ligne droite permet de chercher une solution autosemblable (c'est-à-dire qui ne dépend que de x/t) et d'utiliser la connaissance des problèmes de Riemann pour les équations d'Euler sans terme source. Le théorème suivant donne des conditions suffisantes pour que le problème de Riemann soit bien posé.

Théorème 1.2. *Soit*

$$D : \begin{array}{ccc} \mathbb{R}_+^* \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ (\rho, \alpha) & \mapsto & D(\rho, \alpha) \end{array}$$

une force de frottement de classe $\mathcal{C}^1$, ayant le même signe que α (qui est de même signe que $u - h'$) et s'annulant pour $\alpha = 0$, croissante en α, et dont la valeur absolue est une fonction décroissante de ρ. Alors, pour tout (ρ_L, q_L) et (ρ_R, q_R) dans $\mathbb{R}_+^* \times \mathbb{R}$, et pour toute vitesse de particule v de $\mathbb{R}$, le problème de Riemann (1.8) *a une unique solution autosemblable.*

La structure du problème de Riemann est plus complexe qu'en l'absence de particule. La solution peut contenir jusqu'à cinq ondes, deux de chaque côté de la particule plus une onde-particule à travers laquelle les relations de la définition 1.1 sont vérifiées. De plus, l'onde-particule peut interagir avec une onde-fluide, comme c'est le cas sur la Figure 1.2 où une onde de détente s'étend jusqu'à la particule.

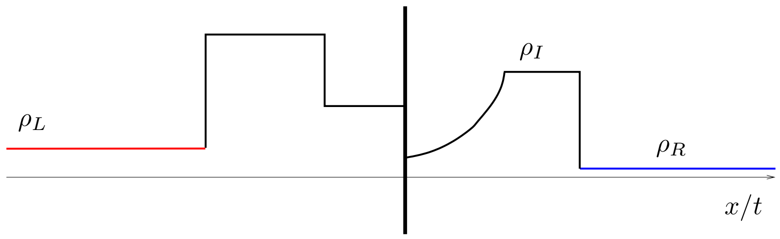


FIGURE 1.2. *Structure de la solution au problème de Riemann pour le couplage Euler-particule.*

Suivons [GL04] et intéressons-nous à la manière dont l'état intermédiaire à droite de la particule (ρ_I, q_I) est atteint. D'une part, il est tout simplement sur la courbe usuelle de 2-onde partant de (ρ_R, q_R). Cette courbe est représentée en bleu sur la figure 1.3. D'autre part, on peut l'atteindre à partir de (ρ_L, q_L) par une succession d'une 1-onde et d'une 2-onde se déplaçant moins vite que la particule, puis d'une onde à travers la particule, et enfin d'une 1-onde allant plus vite que la particule. La courbe regroupant tous les états qui peuvent être joint à (ρ_L, q_L) par une telle succession d'onde est représentée en rouge sur la figure 1.3. Si les hypothèses du théorème 1.2 ne sont pas vérifiées, cette courbe peut avoir une forme de Z et le problème de Riemann peut avoir jusqu'à trois solutions.

Obtenir des résultats plus généraux d'existence et d'unicité est une tâche ardue. Le cas le plus simple semble être d'étendre le théorème 1.2 au cas où la particule bouge sous l'influence du fluide, c'est-à-dire d'étudier le problème de Riemann suivant :

$$\begin{cases} \partial_t \rho + \partial_x q = 0, \\ \partial_t q + \partial_x \left(\frac{q^2}{\rho} + c^2 \rho \right) = -D(\rho, \alpha) \delta_{h(t)}, \\ m h''(t) = (\rho_-(t) - \rho_+(t)) \left(c^2 - (u_-(t) - h'(t))(u_+(t) - h'(t)) \right), \\ \rho(0, x) = \rho_L \mathbf{1}_{x<0} + \rho_R \mathbf{1}_{x>0}, \\ q(0, x) = q_L \mathbf{1}_{x<0} + q_R \mathbf{1}_{x>0}, \\ h(0) = h^{\text{ini}}, h'(0) = v^{\text{ini}}. \end{cases} \tag{1.9}$$

Toutefois, la trajectoire de la particule n'est alors plus une droite et il n'y a plus aucune chance que la solution soit autosemblable. De plus, contrairement au système (1.4) de [LST08] où le fluide suit une équation de Burgers, les traces de part et d'autre de la particule changent à tout instant. Les méthodes développées dans [CG10], [BCG10], [BCG12] et [BCG14] permettent d'obtenir une solution quand les données initiales sont quelconques, mais petites et subsoniques, et que les traces à l'instant initial sont dans le germe :

$$(\rho(0, h_-^{\text{ini}}), q(0, h_-^{\text{ini}}), (\rho(0, h_+^{\text{ini}}), q(0, h_+^{\text{ini}}))) \in \mathcal{G}_D(v^{\text{ini}}).$$

Dans la proposition suivante, on s'affranchit de l'hypothèse de petitesse sur la donnée mais, en contrepartie, on se restreint aux données initiales de Riemann.

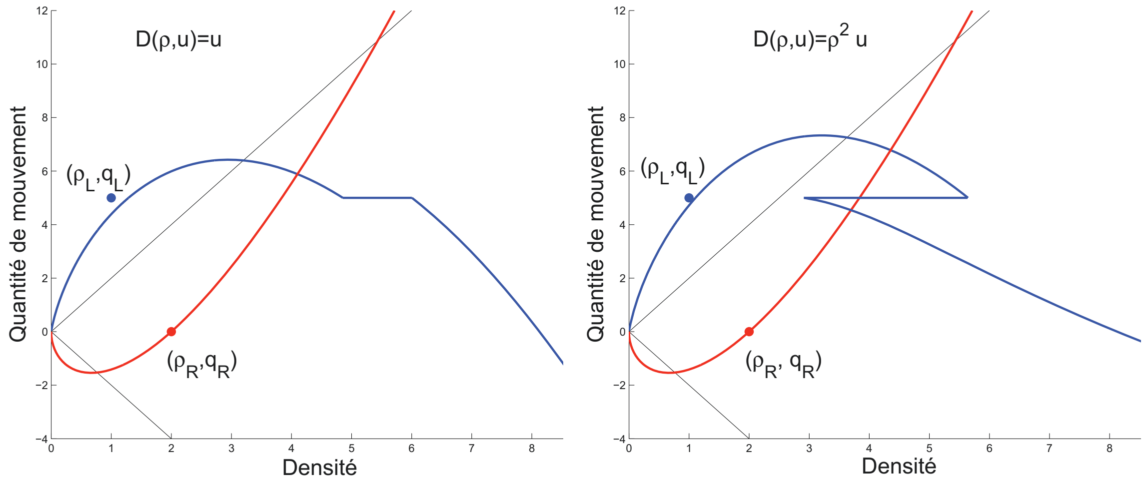


FIGURE 1.3. *Selon la forme de la courbe rouge, le problème de Riemann a une unique solution (à gauche, sous les hypothèses du théorème 1.2) ou jusqu'à trois solutions (à droite, $|D|$ est une fonction croissante de ρ).*

Proposition 1.3. *Soit h^{ini} et v^{ini} deux réels. Si (ρ_L, q_L) et (ρ_R, q_R) sont deux éléments de $\mathbb{R}_+^* \times \mathbb{R}$ tels que $|u_L - v^{\text{ini}}| < c$ et $|u_R - v^{\text{ini}}| < c$, et si*

$$((\rho_L, q_L), (\rho_R, q_R)) \in \mathcal{G}_D(v^{\text{ini}}),$$

alors le système (1.9) *admet une solution en temps court lorsque la force de frottement est de la forme $D(\rho, \alpha) = \lambda\alpha$, où λ est un paramètre de friction positif.*

1.2 Capture de chocs dans les systèmes hyperboliques

Lorsqu'on approche un système hyperbolique avec un schéma de volumes finis, on introduit une petite viscosité numérique qui dépend du schéma choisi. La conséquence principale est que les chocs sont lissés sur plusieurs mailles. Mais dans certains cas, cette diffusion numérique a des effets bien plus spectaculaires. Dans le cas de la mécanique des gaz, plusieurs anomalies numériques sont connues et répertoriées par exemple dans [Qui94] et [Zai12]. L'une d'entre elles est l'apparition d'un pic non physique dans la quantité de mouvement lorsqu'on approche un choc qui se déplace lentement par rapport à la vitesse du son. Le pic oscille de manière quasi périodique en temps, et est donc une source constante d'erreur dans la solution (bien que cela n'empêche pas les schémas de converger pour la norme L^1). Une autre anomalie est l'élévation artificielle de la température lorsqu'un choc se réfléchit sur une paroi, décrite dans [Noh87].

C'est aussi cette viscosité numérique qui explique la difficulté à capturer les chocs non classiques. Ces chocs vérifient une inégalité d'entropie, mais ne vérifient pas les critères classiques d'admissibilité des chocs ([Ole57] dans le cas scalaire, [Lax57] ou [Liu74] pour les systèmes). On impose en plus de cette inégalité d'entropie une relation cinétique, qui garde une trace de l'équilibre entre, par exemple, les effets de diffusion et de dispersion (voir par exemple la monographie [LeF02]). La diffusion numérique détruit cet équilibre et rend la diffusion dominante, et les schémas ne parviennent pas à capturer les chocs non classiques. Dans cette thèse, nous étendons aux systèmes de lois de conservation les schémas non dissipatifs introduits dans [Lag08], [LagXX] et [BCLL08]. L'idée est de reconstruire à l'intérieur de chaque maille une discontinuité d'intérêt, qui aurait été moyennée par le schéma.

1.2.1 Chocs non classiques

Intéressons-nous aux solutions de la loi de conservation à flux cubique

$$\partial_t u + \partial_x u^3 = 0, \tag{1.10}$$

qui sont limites, quand $\varepsilon > 0$ tend vers 0, de l'équation de Burgers–Korteweg de Vries modifiée

$$\partial_t u_\varepsilon + \partial_x u_\varepsilon^3 = \varepsilon\beta\partial_{xx}u_\varepsilon + \varepsilon^2\gamma\partial_{xxx}u_\varepsilon, \tag{1.11}$$

où β est positif. Dans [JMS95], les auteurs montrent que lorsque γ est positif, le profil des ondes progressives de (1.11) peut tendre, lorsque ε tend vers 0, vers un choc sous-compressif qui ne vérifie pas la condition d'Oleinik [Ole57].

Numériquement, on peut capturer de tels *chocs non classiques* en utilisant un schéma où les termes de diffusion et de dispersion de (1.11) sont discrétisés, et où la discrétisation du flux $\frac{u^3}{3}$ est discrétisé à un ordre suffisament élevé pour ne pas introduire de diffusion et de dispersion supplémentaires dans (1.11). Comme on peut le voir sur la figure 1.4, le schéma de Godunov appliqué à (1.10) ne parvient pas à capturer le choc non classique qui apparaît dans la solution de (1.11). Le pas de temps est $\Delta t = \Delta x/300$ et le pas d'espace est $\Delta x = 0.005$, ce qui correspond à $\varepsilon = 0.025$ dans (1.11). On compare au temps $T = 0.02$ la solution obtenue grâce à la discrétisation en espace d'ordre 4 de (1.11) proposée dans [HL98] et un schéma de point-milieu en temps avec le schéma de Godunov non classique pour (1.10). La donnée initiale est

$$u^0(x) = 6 * \mathbf{1}_{x<0} - 3 * \mathbf{1}_{x\geq 0},$$

et on a pris $\beta = 10$ et $\gamma = 19.7$. Puisque toutes les ondes partent vers la droite dans (1.10), le schéma de Godunov s'écrit comme un simple schéma décentré à gauche. Par conséquent, il coïncide avec le schéma de Godunov classique, qui converge vers la solution d'Oleinik, ici un choc.

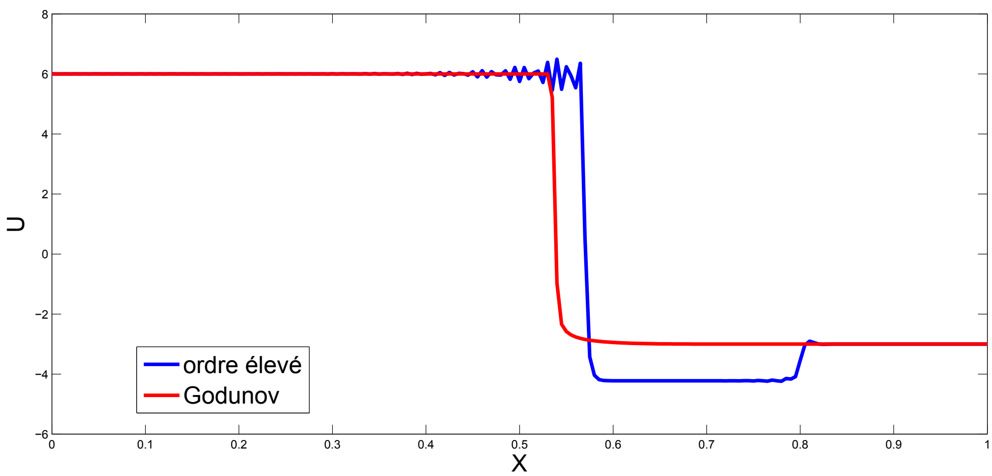


FIGURE 1.4. *Apparition de chocs non classiques. Le schéma de Godunov appliqué à l'équation limite où $\varepsilon = 0$ n'est pas capable de capturer ces chocs.*

Nous présentons maintenant un petit rappel du cadre général pour la théorie des chocs non classiques. Considérons le problème de Cauchy pour un système de lois de conservation

$$\begin{cases} \partial_t U + \partial_x f(U) = 0, \\ U(t=0) = U^0(x). \end{cases} \tag{1.12}$$

Le vecteur U, défini sur un ouvert Ω de $\mathbb{R}^n$ ($n \geq 1$) que l'on supposera convexe, regroupe les variables conservées, et $f : \mathbb{R}^n \to \mathbb{R}^n$ est le flux. Lorsque la matrice jacobienne $Df(U)$ de f en U est diagonalisable pour tout U de Ω, avec des valeurs propres toutes distinctes

$$\lambda_1(U) < \lambda_2(U) < \cdots < \lambda_n(U),$$

on dit que le système (1.12) est strictement hyperbolique. On notera $r_i(U)$ un vecteur propre associé à la i-ème valeur propre :

$$Df(U)r_i(U) = \lambda_i(U)r_i(U).$$

On s'intéresse aux solutions faibles de (1.12) qui vérifient de plus (au sens faible) l'inégalité d'entropie

$$\partial_t E(U) + \partial_x G(U) \leq 0. \tag{1.13}$$

Dans la suite, on supposera que l'entropie E est strictement convexe. Lorsque chaque champ du système (1.12) est vraiment non linéaire, c'est-à-dire lorsque pour tout i dans $\{1, \cdots, n\}$,

$$\nabla\lambda_i(U) \cdot r_i(U) \neq 0 \quad \forall U \in \Omega, \tag{1.14}$$

l'ajout de l'inégalité d'entropie (1.12) permet de récupérer l'existence et l'unicité locale au problème de Riemann. Dans ce cas, le fait qu'un choc est entropique est équivalent au critère de Lax. Ce critère, présenté dans [Lax57], stipule qu'un j-choc entre deux états U_L et U_R est admissible si sa vitesse s vérifie

$$\lambda_j(U_R) < s < \lambda_j(U_L) \quad \text{et} \quad \lambda_{j-1}(U_L) < s < \lambda_{j+1}(U_R). \tag{1.15}$$

Ce critère a été étendu au cas où (1.14) n'est pas vérifié dans [Liu74]. Le critère proposé est équivalent à celui de Lax dans le cas vraiment non linéaire. Il est également prouvé que, sous ce critère, il existe une unique solution au problème de Riemann. Cependant, lorsque (1.14) n'est pas vérifié, il existe des chocs qui sont entropiques sans vérifier le critère de Liu. De plus, de tels chocs peuvent apparaître comme les solutions limites d'un système augmenté où les effets de viscosité et de dispersion sont pris en compte (voir par exemple [JMS95]). Il apparaît cependant que si on admet tous les chocs entropiques, même s'ils ne sont pas admissibles au sens de Liu, certains problèmes de Riemann admettent une infinité de solutions (voir [LeF02], [HL97] et [HL98] entre autres). En un sens, l'inégalité d'entropie (1.13) oublie trop d'informations présentes dans le système augmenté (avec viscosité et dispersion évanescentes). Pour résoudre ce problème, on ajoute en plus de (1.13) une relation cinétique qui sélectionne certains chocs non classiques en précisant, par exemple, la dissipation d'entropie dans (1.13) comme une fonction de la vitesse de propagation du choc.

Un exemple de système hyperbolique dans lequel apparaissent des chocs non classiques est celui de l'élastodynamique non linéaire

$$\begin{cases} \partial_t v - \partial_x(w^3 + mw) = 0, \\ \partial_t w - \partial_x v = 0. \end{cases} \tag{1.16}$$

Ce système est strictement hyperbolique si $m > 0$ mais les champs associés à chacune des valeurs propres ne sont pas vraiment non linéaires :

$$\nabla\lambda_i(v, w) \cdot r_i(v, w) = 0 \ \text{ si et seulement si } \ w = 0.$$

On s'intéresse aux solutions qui sont limites du modèle diffusif-dispersif

$$\begin{cases} \partial_t v - \partial_x(w^3 + mw) = \varepsilon\partial_{xx} v - \alpha\varepsilon^2\partial_{xxx} w, \\ \partial_t w - \partial_x v = 0, \end{cases} \tag{1.17}$$

où α est une constante positive fixée et le paramètre positif ε tend vers 0. De (1.17) on tire l'inégalité d'entropie

$$\partial_t \left(\frac{v^2}{2} + \frac{w^4}{4} + m^2\frac{w^2}{2}\right) - \partial_x \left(-v(w^3 + mw)\right) \leq 0,$$

qui n'est pas suffisante pour sélectionner une unique solution (voir [HL98] et l'étude détaillée du problème de Riemann dans [LT01]). La figure 1.5 illustre la convergence du schéma de Godunov vers la solution classique, alors que le solveur de Riemann utilisé calcule la solution *non classique* exacte. Comme dans le cas scalaire illustré par (1.10), la phase de projection sur le maillage détruit l'équilibre entre effets de viscosité et de capillarité, et le schéma converge vers la solution classique.

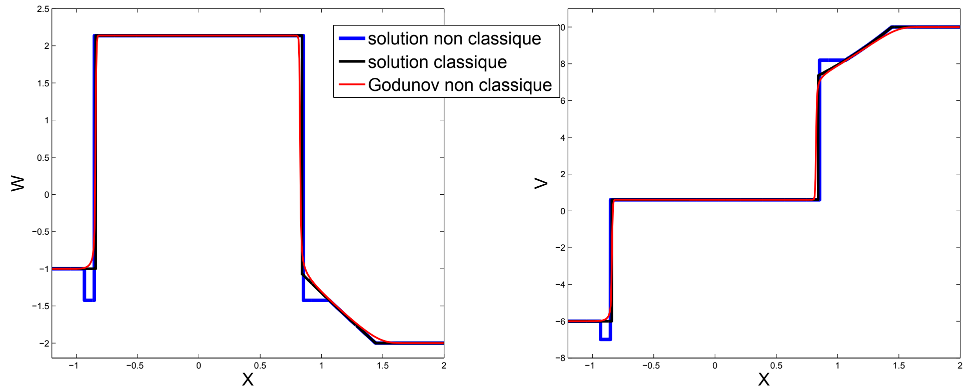


FIGURE 1.5. *Le schéma de Godunov basé sur le solveur de Riemann non classique converge vers la solution classique (cas de l'élastodynamique non linéaire).*

Les schémas de volumes finis développés dans la partie III sont capables d'approcher correctement les chocs non classiques. Avant de décrire leur construction, nous présentons une autre catégorie de chocs qui dépendent finement des petites échelles physiques, et qui sont difficiles à approcher numériquement.

1.2.2 Anomalies numériques dans la dynamique des gaz

Pour approcher la solution du problème de Cauchy pour les équations d'Euler isothermes

$$\begin{cases} \partial_t \rho + \partial_x q = 0, \\ \partial_t q + \partial_x \left(\frac{q^2}{\rho} + c^2 \rho \right) = 0, \end{cases} \tag{1.18}$$

on peut par exemple utiliser le schéma de Lax–Friedrichs :

$$\begin{cases} \rho_j^{n+1} = \frac{\rho_{j-1}^n + \rho_{j+1}^n}{2} - \frac{\Delta t}{2\Delta x} \left(q_{j+1}^n - q_{j-1}^n \right), \\ q_j^{n+1} = \frac{q_{j-1}^n + q_{j+1}^n}{2} - \frac{\Delta t}{2\Delta x} \left(\frac{(q_{j+1}^n)^2}{\rho_{j+1}^n} + c^2 \rho_{j+1}^n - \frac{(q_{j-1}^n)^2}{\rho_{j-1}^n} - c^2 \rho_{j-1}^n \right). \end{cases} \tag{1.19}$$

Un développement limité montre que ce schéma est consistant à l'ordre 2 avec le système

$$\partial_t \begin{pmatrix} \rho \\ q \end{pmatrix} + \partial_x \begin{pmatrix} q \\ \frac{q^2}{\rho} + c^2 \rho \end{pmatrix} = \Delta t \partial_x \left[\begin{pmatrix} \frac{\Delta x^2}{\Delta t^2} - (c^2 - u^2) & -2u \\ -2u(c^2 - u^2) & \frac{\Delta x^2}{\Delta t^2} - (3u^2 + c^2) \end{pmatrix} \partial_x \begin{pmatrix} \rho \\ q \end{pmatrix} \right].$$

Les termes diagonaux $\frac{\Delta x^2}{\Delta t^2}$ proviennent de la discrétisation en espace et les autres de la discrétisation en temps. La matrice

$$D_{LF}(\rho, q) = \begin{pmatrix} \frac{\Delta x^2}{\Delta t^2} - (c^2 - u^2) & -2u \\ -2u(c^2 - u^2) & \frac{\Delta x^2}{\Delta t^2} - (3u^2 + c^2) \end{pmatrix}$$

est définie positive si et seulement si la condition de Courant–Friedrichs–Lewy $\Delta t \leq \frac{\Delta x}{2(|u|+c)}$ est vérifiée. On appelle le terme correctif

$$\partial_x(D_{LF}(\rho, q)\partial_x(\rho, q))$$

viscosité numérique (ou *diffusion numérique*) par analogie avec l'équation de la chaleur.

L'effet le plus connu et le plus répandu de cette diffusion numérique est de lisser les discontinuités sur plusieurs mailles. On peut le vérifier sur la figure 1.6 ci-dessus. On y compare deux schémas d'ordre 1, les schémas de Godunov et de Lax–Friedrichs, avec leurs extensions d'ordre 2, respectivement les schémas MUSCL [vL97] et de Nessyahu–Tadmor [NT90].

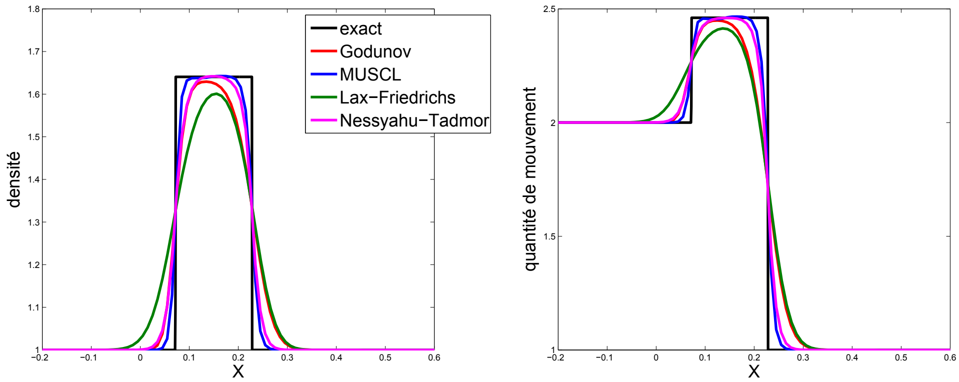


FIGURE 1.6. *Diffusion des chocs avec différents schémas de volumes finis. Le temps final est* 0.1 *et la vitesse du son est* 1.

La donnée initiale est un problème de Riemann dont la solution ne contient que des chocs. Sur la figure 1.7, la donnée initiale est un choc dont la vitesse est petite devant celle du son (ici cinq fois plus petite, la vitesse du son est fixée à 0.5). On observe l'apparition d'oscillations parasites dans la quantité de mouvement à l'avant du choc (en fait ces oscillations sont aussi présentes dans la densité, mais leur amplitude est beaucoup plus faible). Le pic apparaît dans les premières itérations en temps. Les schémas de volumes finis étant conservatifs, il est compensé par l'apparition d'un creux. Le pic oscille de manière presque périodique en temps, la période correspondant au temps nécessaire à un choc pour traverser une maille. Ainsi, même si les oscillations sont vite amorties par la diffusion numérique, c'est une source constante d'erreur en temps.

Dans l'article [JL96], les auteurs étudient les ondes progressives (les solutions ne dépendant que de $\frac{x-st}{\varepsilon}$, où s est la vitesse de l'onde) de

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = \varepsilon \partial_{xx}\rho, \\ \partial_t(\rho u) + \partial_x\left(\rho u^2 + c^2\rho\right) = \varepsilon\partial_{xx}(\rho u). \end{cases} \tag{1.20}$$

La quantité de mouvement ρu n'est pas monotone lorsque s est petit (choc lent), alors qu'elle l'est dans les solutions ondes progressives du système de Navier–Stokes

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x\left(\rho u^2 + c^2\rho\right) = \varepsilon\partial_{xx} u. \end{cases} \tag{1.21}$$

Cela peut expliquer que les schémas ayant une viscosité ressemblant à (1.20) font apparaître un pic dans la quantité de mouvement. Pour le faire disparaitre, le point crucial est qu'il n'y ait pas de viscosité sur la variable ρ, autrement dit que la viscosité numérique se rapproche de la viscosité physique de Navier–Stokes (1.21).

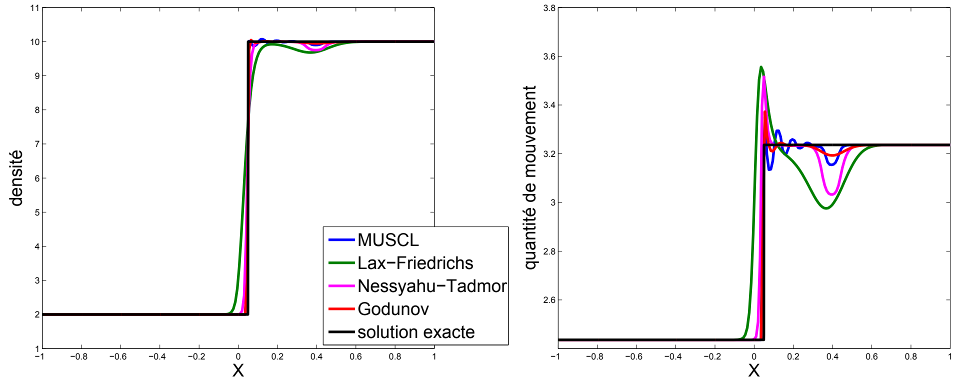


FIGURE 1.7. *Apparitions d'oscillations devant un choc lent au temps* 0.5. *Les schémas d'ordre* 2 *(MUSCL, Nessyahu–Tadmor) peuvent être moins bons que les schémas d'ordre* 1 *(Godunov, Lax–Friedrichs).*

1.2.3 Schémas de reconstruction discontinue pour les systèmes hyperboliques

La classe de schémas de volumes finis présentée dans la partie III est construite pour être exacte lorsque la donnée initiale est un choc (classique ou non classique selon les cas) ou une discontinuité de contact. Considérons le cas où la donnée initiale U^0 est un choc :

$$U^0(x) = U_L \mathbf{1}_{x<0} + U_R \mathbf{1}_{x>0}$$

où les états gauche et droit sont liés par les relations de Rankine–Hugoniot :

$$\exists s \in \mathbb{R} : \quad s(U_L - U_R) = f(U_L) - f(U_R). \tag{1.22}$$

Dans ce cas, la discontinuité initiale se propage à vitesse s et la solution exacte est

$$U^{\text{exa}}(x,t) = U_L \mathbf{1}_{x<st} + U_R \mathbf{1}_{x>st}.$$

Si le choc est initialement placé sur une interface du maillage, après une itération en temps il se situera à une distance $s\Delta t$ de celle-ci. Supposons que s est positif et que $s\Delta t < \Delta x$, et numérotons 1 la maille dans laquelle le choc avance. Un schéma de volumes finis fournit des solutions constantes sur chaque maille et est conservatif. S'il est exact à la première itération, on obtient après le phase de projection

$$\begin{cases} U_j^1 = U_L & \text{si } j \le 0, \\ U_1^1 = \frac{s\Delta t}{\Delta x} U_L + \frac{\Delta x - s\Delta t}{\Delta x} U_R, & \\ U_j^1 = U_R & \text{si } j \ge 2. \end{cases} \tag{1.23}$$

Ainsi, une nouvelle valeur intermédiaire est créée, qui ne correspond pas à une valeur ponctuelle de la solution. L'idée des schémas de reconstruction discontinue est d'inverser l'étape de projection des schémas de volumes finis en reconstruisant le choc à l'intérieur de la maille 1. Elle apparaît dans [LagXX] et [BCLL08]. Plus précisément, la valeur U_j^n fournie par le schéma est vue, dès que possible, comme la valeur moyenne d'un choc entre deux états $U_{j,L}^n$ et $U_{j,R}^n$. La discontinuité est placée à une distance $d_j^n \in [0, \Delta x]$ de l'interface de gauche. On calcule ensuite la solution exacte au temps Δt lorsque la donnée initiale est la fonction reconstruie

$$U_{\text{rec}}^n(x) = \sum_j U_{j,L}^n \mathbf{1}_{x_{j-1/2}^n < x < x_{j-1/2}^n + d_j^n}(x) + U_{j,R}^n \mathbf{1}_{x_{j-1/2}^n + d_j^n < x < x_{j+1/2}^n}(x)$$

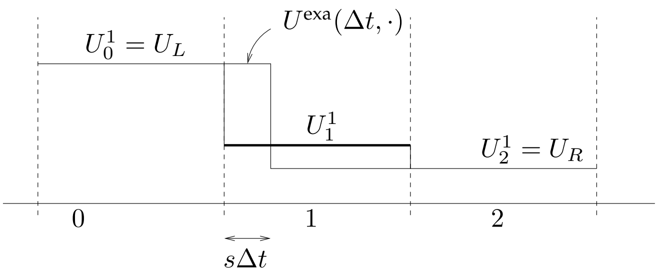


FIGURE 1.8. *Création d'une valeur intermédiaire dans l'approximation par un schéma de volumes finis d'un choc isolé.*

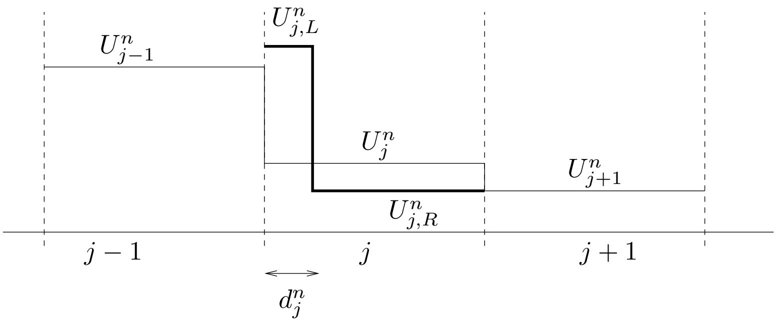


FIGURE 1.9. *Reconstruction d'un choc dans la maille j. La discontinuité est placée de sorte que U soit conservée dans cette maille.*

et U_j^{n+1} est défini comme la valeur moyenne de cette solution sur la maille j au temps Δt. Les questions qui se posent sont les suivantes.

- Comment choisir les mailles où appliquer la procédure de reconstruction ?
- Comment choisir les états reconstruits à gauche et à droite $U_{j,L}^n$ et $U_{j,R}^n$?
- Comment choisir la distance à l'interface de gauche d_j^n ?
- Peut-on calculer facilement la moyenne sur chaque maille de la solution à l'instant Δt avec la donnée initiale U_{rec}^n ?

Il est clair que dans le cas du choc isolé (1.23) décrit par la figure 1.8, la procédure de reconstruction doit être appliquée dans la maille 1 *et nulle par ailleurs*, de sorte que U_{rec}^1 coïncide avec $U_{exa}(\Delta t, \cdot)$. Cela impose

$$U_{1,L}^1 = U_0^1,\ U_{1,R}^1 = U_2^1 \text{ et } d_1^{1,U} = s\Delta t.$$

La difficulté nouvelle par rapport au cas scalaire est que U_{j-1}^n et U_{j+1}^n ne sont en général pas liés par un choc. Sauf exception, la solution au problème de Riemann entre ces deux états contiendra n ondes, chocs, contacts ou détentes. Le choc à reconstruire est donné par un opérateur de reconstruction.

Définition 1.3. *Un opérateur de reconstruction est une fonction*

$$\begin{array}{llcl} \mathcal{R} : & \Omega \times \Omega \times \Omega & \to & \Omega \times \Omega \\ & (U_L, U_M, U_R) & \mapsto & (U_{M,L}, U_{M,R}) \end{array}$$

tels que les états $U_{M,L}$ et $U_{M,R}$ sont liés par un choc entropique ou un contact. On note $s(U_{M,L}, U_{M,R})$ la vitesse associée à cette discontinuité, donnée par les relations de Rankine–Hugoniot (1.22). *Lorsque $U_{M,L} = U_{M,R}$, on fixe arbitrairement $s(U_{M,L}, U_{M,R}) = \lambda_1(U_{M,L})$.*

Quant à la distance d_j^n, elle est calculée de manière à ce que les variables conservées U soit bien conservées dans la maille :

$$\Delta x U_j^n = d_j^n \times U_{j,L}^n + (\Delta x - d_j^n) \times U_{j,R}^n.$$

Cela donne a priori des distances différentes pour chaque quantité conservée : d_j^n est un vecteur et $\times$ indique le produit composante par composante. Les notations sont rappelées sur la figure 1.9. Les flux numériques sont calculés en advectant chacun des chocs reconstruits à la vitesse $s(U_{M,L}, U_{M,R})$ pendant Δt. L'utilisation d'un maillage mobile se déplaçant plus vite que les ondes permet de calculer ces flux simplement, sans résoudre d'interactions d'ondes.

La principale difficulté est de définir l'opérateur $\mathcal{R}$ de manière à ce que le schéma numérique soit capable d'approcher exactement les chocs isolés. Le théorème ci-dessous donne un critère pour que ce soit le cas.

Théorème 1.4. *Soient U_L et U_R deux états liés par un choc ou une discontinuité de contact. Si pour tout $\alpha \in (0,1)$, on a*

$$\begin{cases} \mathcal{R}(U_L, \alpha U_L + (1-\alpha)U_R, U_R) = (U_L, U_R) \\ \mathcal{R}(U_L, U_L, \alpha U_L + (1-\alpha)U_R) = (U_L, U_L) \\ \mathcal{R}(\alpha U_L + (1-\alpha)U_R, U_R, U_R) = (U_R, U_R) \end{cases}$$

alors sous une condition de Courant–Friedrichs–Lewy classique, le schéma de reconstruction discontinue est exact lorsque la donnée initiale est un choc isolé : U_j^n est la moyenne de la solution exacte au temps t^n sur la maille j.

De plus, l'opérateur de reconstruction doit être suffisamment souple pour que cette bonne propriété soit héritée dans des cas plus généraux. Par exemple, si le problème de Riemann entre U_L et U_R est presque un choc (c'est-à-dire si les autres ondes sont d'amplitudes faibles), comme sur la figure 1.9, on souhaite que ce choc soit reconstruit. La figure 1.10 illustre cette difficulté. On a utilisé deux opérateurs de reconstruction vérifiant les hypothèses du théorème précédent, mais un seul permet de capturer correctement le choc lent apparaîssant dans le problème de Riemann.

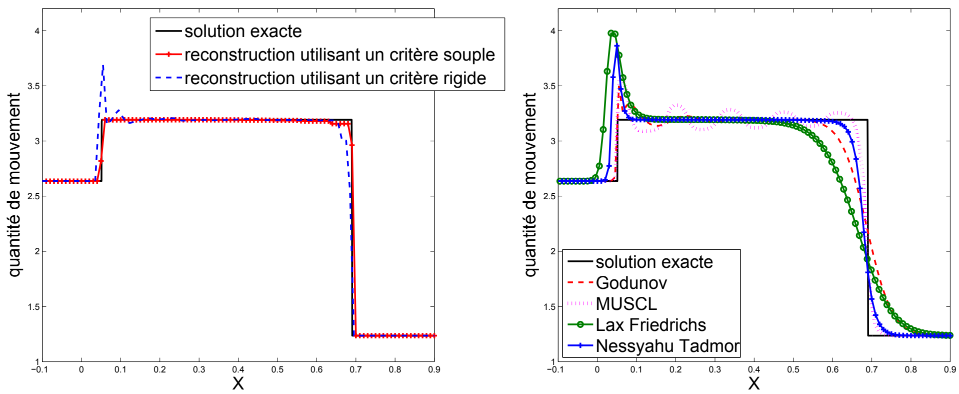


FIGURE 1.10. *Pour approcher les chocs lents correctement, l'opérateur de reconstruction doit être assez souple.*

Les opérateurs de reconstruction proposés dans les sections 6.4, 6.5 et 6.8 s'appuient sur la résolution exacte de problèmes de Riemann autour des mailles où l'on a préalablement détecté un choc. Se passer d'un solveur de Riemann exact est fondamental tant du point de vue du coût de calcul que de la généralisation du schéma à des cas plus généraux, où l'on pourrait ne pas disposer d'un tel solveur. Toutefois, le solveur approché doit fournir la solution exacte lorsque la donnée est un choc pur, ce qui élimine d'emblée presque tous les solveurs de Riemann approchés disponibles dans la littérature. Il existe toutefois un tel solveur, introduit dans [CC13], qui est couplé avec la stratégie de reconstruction dans la section 6.6.

L'utilisation de ce solveur a deux avantages : d'une part, la détection des chocs est déjà contenue dans le solveur, et d'autre part le temps de calcul est fortement réduit. Cela permet de faire des simulations en 2D (en utilisant un splitting directionnel) en un temps raisonnable. La figure 1.11 montre l'évolution de la hauteur d'eau (en coordonnées lagrangienne) dans un bassin carré. La donnée initiale est une colonne d'eau située au milieu du bassin, et la solution est représentée après de multiples rebonds sur le bord du bassin. Cette figure illustre le caractère peu diffusif (aux alentours des discontinuités du moins) des schémas de reconstruction discontinues : les chocs sont bien plus raides que ceux obtenus avec le schéma de Godunov et de nombreux détails apparaîssent.

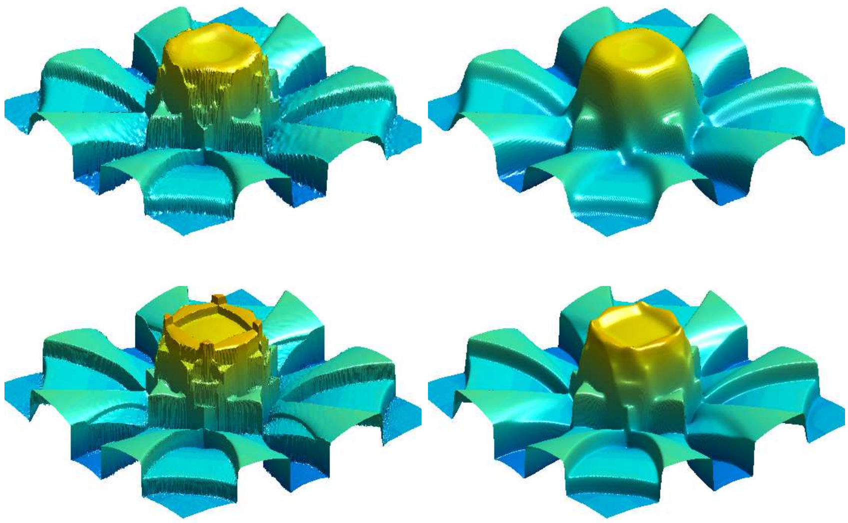


FIGURE 1.11. *Simulation de la chute d'une colonne d'eau dans un bassin carré. À gauche, avec le schéma de reconstruction discontinue, à droite avec un schéma de relaxation, en haut avec un maillage* 200×200*, en bas avec un maillage* 800×800.

I

Couplage entre l'équation de Burgers et une particule

Présentation générale 2

Dans ce chapitre, nous présentons un modèle de couplage entre l'équation de Burgers et une particule ponctuelle. Le système proposé contient un terme source qui s'écrit comme le produit d'une fonction discontinue par une masse de Dirac. Dans un premier temps, nous définissons précisément ce terme et donnons plusieurs définitions équivalentes des solutions du système. Nous présentons ensuite quelques solutions particulières faisant intervenir des collisions entre particules. Enfin, nous rappelons les résultats d'existence et d'unicité obtenus sur ce système, et nous présentons différents types de schémas numériques convergents.

Table des matières

Nous présentons les résultats existant sur le problème de Cauchy pour le modèle de couplage fluide-particule :

$$\begin{cases} \partial_t u + \partial_x \frac{u^2}{2} = -\lambda(u - h'(t))\delta_{h(t)}(x), \\ m_p h''(t) = \lambda(u(t, h(t)) - h'(t)), \\ u_{|t=0} = u^0, h(0) = h^0, h'(0) = v^0. \end{cases} \tag{2.1}$$

Ce modèle a été introduit par Lagoutière, Seguin et Takahashi dans [LST08]. Ce système décrit l'évolution de la trajectoire h d'une particule ponctuelle de masse m_p qui bouge sous l'influence d'un fluide, dont la vitesse u est décrite par l'équation de Burgers. L'interaction entre le fluide et la particule est décrite par un terme force de rappel entre leurs vitesses respectives : le fluide exerce la force

$$D(u - h'(t)) = \lambda(u - h'(t))$$

sur la particule, et la particule exerce la force $-D$ sur le fluide. Cette force s'applique au point où se situe la particule, ce que traduit la multiplication par une mesure de Dirac dans le terme source

$$-\lambda(u - h'(t))\delta_{h(t)}(x). \tag{2.2}$$

Le paramètre λ est un coefficient de friction positif, de sorte que l'EDO sur la particule

$$m_p h''(t) = \lambda(u(t, h(t)) - h'(t)) \tag{2.3}$$

a le bon signe : si la particule va moins vite que le fluide, $u(t, h(t)) > h'(t)$ et la particule accélère. Le cas où la particule a une vitesse fixée $v \in \mathbb{R}$ est intéressant en lui-même, et c'est également une étape théorique naturelle dans l'étude du système entièrement couple (2.1). Dans ce cas, l'équation sur la particule disparaît et on obtient le système

$$\begin{cases} \partial_t u + \partial_x \frac{u^2}{2} = -\lambda(u - v)\delta_{vt}(x); \\ u_{|t=0} = u^0. \end{cases} \tag{2.4}$$

Une nouveauté de ce modèle est de décrire le fluide par une équation non visqueuse. L'équation de Burgers

$$\partial_t u + \partial_x \frac{u^2}{2} = 0$$

admet des solutions discontinues même si la donnée initiale est régulière. De plus, on ne demande pas que le fluide et la particule aient la même vitesse. Cela contraste avec le modèle Burgers-particule étudié par Vázquez et Zuazua dans [VZ03]. D'autre part, le système (2.1) est un cas dégénéré du modèle de spray introduit dans [BD06], lorsque la densité du spray est un Dirac en vitesse et en position.

Pour étudier le système (2.1), il est nécessaire de reformuler le problème pour donner un sens au produit non conservatif (2.2) et à l'équation différentielle (2.3). Grâce à l'étude d'un système régularisé où la particule est épaissie, le système (2.1) est redéfini comme un problème d'interface où des relations entre la vitesse de la particule $h'(t)$ et la vitesse du fluide de part et d'autre de la particule

$$u(t, h(t)^-) = \lim_{x \to h(t), x < h(t)} u(t, x) \quad \text{et} \quad u(t, h(t)^+) = \lim_{x \to h(t), x > h(t)} u(t, x)$$

sont imposées. Cette réécriture rentre dans le cadre de la théorie des germes L^1 dissipatifs développée dans [AKR10] et [AKR11] pour les lois de conservation scalaires à flux discontinus, et permet de donner plusieurs définitions équivalentes des solutions. Les questions d'existence et d'unicité des solutions de (2.1) sont abordées dans la série d'articles [LST08] (existence et unicité au problème de Riemann pour une particule immobile ; existence au problème de Riemann couplé), [AS12] (le problème de Cauchy lorsque la particule est immobile est bien posé dans L^∞) et [ALST10] et [ALST13] (cas du couplage complet (2.1)).

2.1 Définition des solutions

Le premier objectif est de définir le terme source $-\lambda(u - h')\delta_h$. Pour cela, on suppose que la vitesse de la particule est égale à une constante v et on épaissit la particule. On trouve alors que les vitesses à l'entrée et à la sortie de la particule épaissie appartiennent toujours au même ensemble, indépendant de la taille de la particule et de sa forme. On appelle cet ensemble le *germe*.

Définition 2.1 (Lagoutière, Seguin, Takahashi)**.** *Le* germe à vitesse v *est l'ensemble*

$$\mathcal{G}_\lambda(v) := \mathcal{G}_\lambda^1 \cup \mathcal{G}_\lambda^2(v) \cup \mathcal{G}_\lambda^3(v),$$

où

$$\mathcal{G}_\lambda^1 = \left\{(u_-, u_+) \in \mathbb{R}^2 : u_- = u_+ + \lambda\right\},$$

$$\mathcal{G}_\lambda^2(v) = \left\{(u_-, u_+) \in \mathbb{R}^2 : v \leq u_- \leq v + \lambda,\ v - \lambda \leq u_+ \leq v \quad \textit{et} \quad u_- - u_+ < \lambda\right\},$$

et

$$\mathcal{G}_\lambda^3(v) = \left\{(u_-, u_+) \in \mathbb{R}^2 : 2v - \lambda \leq u_+ + u_- \leq 2v + \lambda \quad \textit{et} \quad u_- - u_+ > \lambda\right\}.$$

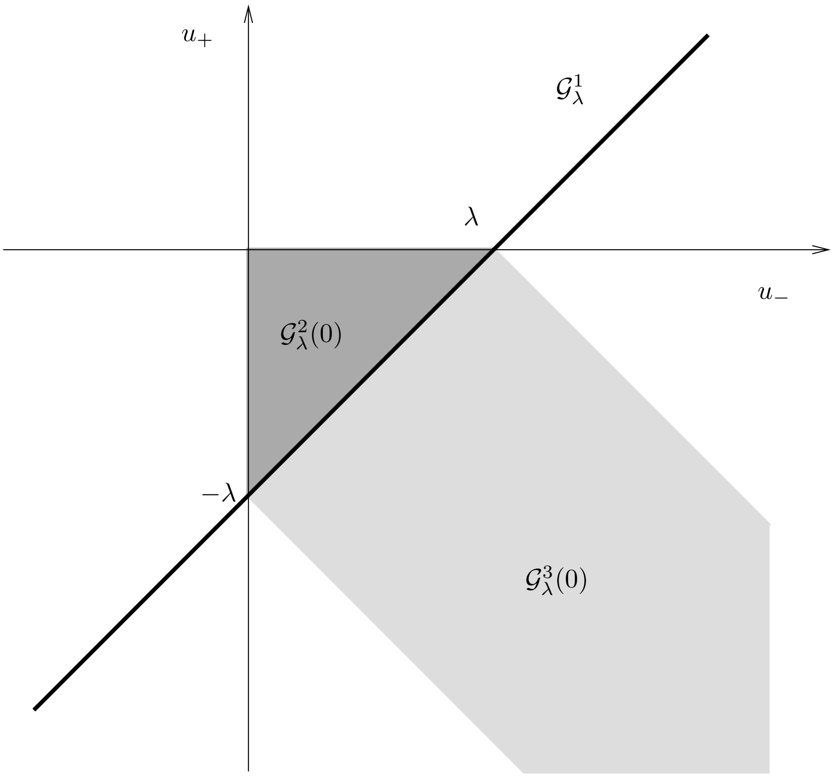


FIGURE 2.1. *Le germe pour une particule immobile.*

Le théorème ci-dessous précise le lien entre le germe et l'épaississement de la particule.

Théorème 2.1 (Lagoutière, Seguin, Takahashi)**.** *Soit H_ε une régularisation de classe $\mathcal{C}^\infty$ de la fonction de Heaviside, croissante et telle que*

$$H_\varepsilon(x) = \begin{cases} 0 & \text{si } x < -\varepsilon/2, \\ 1 & \text{si } x > \varepsilon/2. \end{cases}$$

Si $u_\varepsilon(t,x) = U_\varepsilon(x - vt)$ est une solution de classe $\mathcal{C}^1$ par morceaux de l'équation

$$\partial_t u_\varepsilon + \partial_x \frac{(u_\varepsilon)^2}{2} = -\lambda(u_\varepsilon - v) H'_\varepsilon(x - vt) \tag{2.5}$$

dont les discontinuités sont décroissantes, alors $(U_\varepsilon(-\varepsilon/2), U_\varepsilon(\varepsilon/2))$ appartient à $\mathcal{G}_\lambda(v)$.

Réciproquement, si (u_-, u_+) est un élément de $\mathcal{G}_\lambda(v)$, alors il existe une fonction U_ε de classe $\mathcal{C}^1$ par morceaux et dont les discontinuités sont décroissantes, telle que $(t,x) \mapsto U_\varepsilon(x - vt)$ est une solution de (2.5)*,*

$$u_- = U_\varepsilon(-\varepsilon/2) \quad \textit{et} \quad u_+ = U_\varepsilon(\varepsilon/2).$$

On peut maintenant donner une définition du système (2.1) comme un problème d'interface. L'équation différentielle ordinaire sur la particule est reformulée de sorte que l'impulsion totale soit conservée au cours du temps.

Définition 2.2 (Lagoutière, Seguin, Takahashi)**.** *La paire $(u, h) \in L^\infty(\mathbb{R}_+ \times \mathbb{R}) \times W^{2,\infty}_{loc}(\mathbb{R}_+)$ est une solution de* (2.1) *si les conditions suivantes sont vérifiées.*

- *La fonction u est une solution entropique de l'équation de Burgers sur les ensembles $\{(t,x)\ x < h(t)\}$ et $\{(t,x) :\ x > h(t)\}$, avec condition initiale u^0.*

- *Pour presque tout temps $t > 0$, les traces autour de la particule appartiennent au germe :*

$$(u(t, h(t)_-), u(t, h(t)_+)) \in \mathcal{G}_\lambda(h'(t)).$$

- *La trajectoire de la particule vérifie, pour presque tout temps $t > 0$,*

$$mh''(t) = (u(t, h(t)_-) - u(t, h(t)_+)) \left(\frac{u(t, h(t)_-) + u(t, h(t)_+)}{2} - h'(t) \right).$$

Remarque 1. Cette définition utilise l'existence de traces des solutions de Burgers sur des demi-espaces, démontrée par exemple dans [Pan07].

L'existence d'une solution est obtenue en démontrant la convergence d'un schéma numérique vers une solution de (2.1). La définition précédente utilise des traces ponctuelles autour de la particule. Or on peut tout à fait envisager des schémas créant des couches limites autour de la particule : ils convergent vers la bonne solution, mais les traces numériques autour de la particule ne convergent pas vers les traces de la solution exacte. Heureusement, on dispose d'une définition équivalente qui ne fait pas intervenir les traces.

Proposition 2.2 (Andreianov, Seguin)**.** *La définition 2.2 est équivalente à ce que $(h, u) \in L^\infty(\mathbb{R}_+ \times \mathbb{R}) \times W^{2,\infty}_{loc}(\mathbb{R}_+)$ vérifie les propositions suivantes.*

- *La fonction u est une solution entropique de l'équation de Burgers sur les ensembles $\{(t, x)\ x < h(t)\}$ et $\{(t, x) :\ x > h(t)\}$, avec donnée initiale u^0.*
- *Pour tout $(c_-, c_+) \in \mathbb{R}^2$ et presque tout $t_0 > 0$, il existe un voisinage $(t_0 - \delta, t_0 + \delta)$ de t_0 et une constante $A = A(||u^0||_\infty, \lambda, c, ||h'||_\infty)$ tels que, pour toute fonction test positive $\varphi \in \mathcal{C}^\infty_0((t_0 - \delta, t_0 + \delta) \times \mathbb{R})$,*

$$\begin{aligned} \int_{\mathbb{R}_+} \int_{\mathbb{R}} \big(|u - c|(t, x)\partial_t\varphi(t, x - h(t)) + \Phi_{h(t)}(u, c)(t, x)\partial_x\varphi(t, x - h(t)) \big)\, dx\, dt \\ \geq -A \int_{\mathbb{R}_+} \operatorname{dist}_1((c_-, c_+), \mathcal{G}_\lambda(h'(t)))\varphi(t, 0)\, dt, \end{aligned} \tag{2.6}$$

où $c(t, x) = \mathbf{1}_{x<h(t)}c_- + \mathbf{1}_{x>h(t)}c_+$ et dist_1 est la distance ensembliste usuelle pour la norme L^1 de $\mathbb{R}^2$.

- *Pour tout $\xi \in \mathcal{C}^\infty_0([0, T))$ et pour tout $\psi \in \mathcal{C}^\infty_0(\mathbb{R})$ tel que $\psi(0) = 1$,*

$$\begin{aligned} -\int_0^T m_p h'(t)\xi'(t)\, dt = m_p v^0 \xi(0) &+ \int_{\mathbb{R}} \int_0^T \frac{u^2}{2}(s, x)\xi(s)\psi'(x - h(s))\, ds\, dx \\ &+ \int_{\mathbb{R}} \int_0^T u(s, x)[\xi'(s) - h'(s)\psi'(x - h(s))]\, ds\, dx \\ &+ \int_{\mathbb{R}} u^0(x)\psi(x - h(0))\xi(0)\, dx, \end{aligned}$$

et $h(0) = h^0$, $h'(0) = v^0$.

Le reformulation (2.6) utilise de manière fondamentale le fait que le germe est L^1-dissipatif (selon la terminologie de [AKR11]).

2.2 Collisions entre particules

Cette section illustre le fait que le système (2.1) autorise des collisions entre particules. C'est une nouveauté de ce modèle par rapport à celui proposé par Vázquez et Zuazua (voir [VZ06]). On étend aisément le système au cas de deux particules tant qu'elles ne rentrent pas en collision. On note h_i

la position de la i-ème particule et, pour simplifier, on suppose que les deux particules ont la même masse m et sont modélisées par le même paramètre de frottement λ. Le problème de Cauchy s'écrit

$$\begin{cases} \partial_t u + \partial_x \frac{u^2}{2} = -\lambda(u - h_1'(t))\delta_{h_1(t)}(x) - \lambda(u - h_2'(t))\delta_{h_2(t)}(x), \\ m h_1''(t) = \lambda(u(t, h_1(t)) - h_1'(t)), \\ m h_2''(t) = \lambda(u(t, h_2(t)) - h_2'(t)), \\ u_{|t=0} = u^0, \; h_1(0) = h_1^0, \; h_1'(0) = v_1^0, \; h_2(0) = h_2^0, \; h_2'(0) = v_2^0. \end{cases} \tag{2.7}$$

Le terme source est naturellement défini quand les particules sont à des positions différentes avec la définition 2.1. Dans cette section, on montre l'apparition de collisions franches, c'est-à-dire d'un temps T pour lequel $h_1(T) = h_2(T)$ et $h_1'(T) \neq h_2'(T)$. À l'instant T, une modélisation supplémentaire est nécessaire pour définir le produit non conservatif

$$-\lambda(u - h_1')\delta_{h_1(T)} - \lambda(u - h_2')\delta_{h_2(T)},$$

c'est pourquoi nous ne dirons rien sur ce qui se passe après la collision. Nous donnons trois exemples de collisions.

2.2.1 Collision avec une particule au repos

Considérons le cas où le fluide est au repos ($u^0(x) = 0$), où $h_1^0 < h_2^0$ et où $v_1^0 > 0 = v_2^0$. Dans ce cas, tant que le fluide autour de la seconde particule est au repos, tout se passe comme si cette particule n'était pas là. Or, dans ce cas, la particule 1 ne perturbe le fluide que derrière elle (voir les lemmes 5.6 et 5.7 de [LST08]). Ainsi, si on prend h_1^0 suffisamment proche de h_2^0, on aura une collision franche. La figure 2.2 montre l'allure de la solution à deux temps successifs. Le trait fin représente la vitesse du fluide en fonction de la position et les disques représentent les particules, avec en abscisse leur position et en ordonnées leur vitesse.

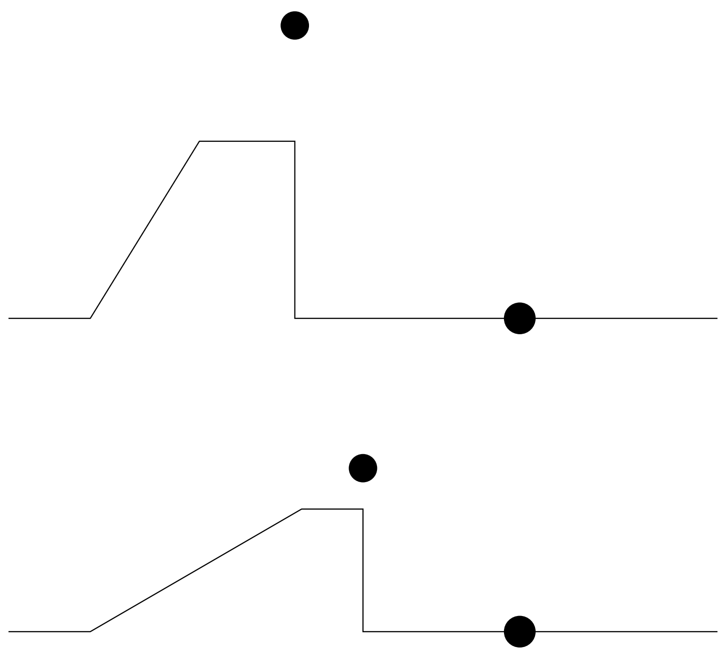


FIGURE 2.2. *Allure de la solution à deux temps successifs dans le cas où $v_1^0 > \lambda$.*

2.2.2 Collision entre particules voyageant dans des chocs

Le système (2.1) admet des solutions où la particule se déplace à l'intérieur d'un choc (voir le lemme 5.5 de [LST08]). D'après (2.3), la vitesse de la particule est de plus constante si la vitesse de la particule est $\frac{u_- + u_+}{2}$. On en déduit une classe de données initiales menant à des collisions entre particules, en recollant deux tels chocs. La solution est décrite par la figure 2.3. Pour tout $h_1^0 < h_2^0$, pour tout v, si les

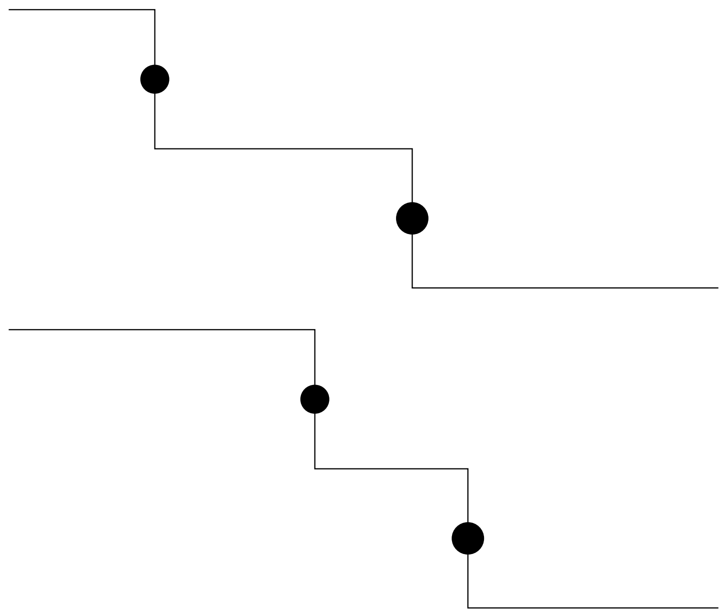


FIGURE 2.3. *Deux particules se déplaçant dans des chocs. Le choc de gauche est plus rapide que celui de droite.*

vitesses initiales sont $v_1^0 = v$ et $v_2^0 = v - \lambda$, la solution au problème de Cauchy (2.7) avec

$$u^0(x) = \left(v + \frac{\lambda}{2}\right) \mathbf{1}_{x<h_1^0} + \left(v - \frac{\lambda}{2}\right) \mathbf{1}_{h_1^0 \le x < h_2^0} + \left(v - \frac{3\lambda}{2}\right) \mathbf{1}_{h_2^0 < x}$$

est donnée par

$$u^0(t,x) = \left(v^0 + \frac{\lambda}{2}\right) \mathbf{1}_{x<h_1^0+vt} + \left(v - \frac{\lambda}{2}\right) \mathbf{1}_{h_1^0+vt \le x < h_2^0+(v-\lambda)t} + \left(v - \frac{3\lambda}{2}\right) \mathbf{1}_{h_1^0+(v-\lambda t)<x},$$

jusqu'au temps $\frac{h_2^0 - h_1^0}{\lambda}$ où les particules rentrent en collision.

2.2.3 Phénomène d'entrainement-collision-dépassement

Dans cette partie nous montrons que deux particules lancées avec la même vitesse dans un fluide au repos peuvent rentrer en collision : la particule de devant entraine le fluide derrière elle, ce qui accélère la particule de derrière. C'est le phénomène d'entrainement-collision-dépassement bien connu des cyclistes (appelé drafting-kissing-tumbling en anglais). Nous notons v_0 la vitesse initiale des particules, h_1 et h_2 leur position et nous les numérotons de sorte que $0 = h_1^0 < h_2^0$. On choisit $0 < v_0 < \lambda$ de sorte que les particules se déplacent plus vite que le fluide. C'est la particule 1 qui a une chance de rattraper la particule 2. Au début, les ondes créées par les deux particules ne se rencontrent pas. La solution est alors donnée par

$$u(t,x) = \begin{cases} 0 & \text{si } x < h_1^0 \\ x/t & \text{si } h_1^0 < x \leq h_1(t) \\ 0 & \text{si } h_1(t) < x \leq h_2^0 \\ (x-h_2^0)/t & \text{si } h_2^0 < x \leq h_2(t) \\ 0 & \text{si } h_2(t) < x \end{cases}$$

et, pour $i \in \{1,2\}$,

$$m_i h_i''(t) = \left(\frac{h_i(t)-h_i^0}{t}\right)\left(\frac{h_i(t)-h_i^0}{2t} - h_i'(t)\right),$$

(voir [LST08, lemme 5.6]).

Lemme 2.3. *Soit v_0 un réel strictement positif. Il existe une unique solution maximale de classe $\mathcal{C}^2$ au problème de Cauchy*

$$\begin{cases} mh''(t) = \frac{h(t)}{t}\left(\frac{h(t)}{2t} - h'(t)\right), \\ (h(0), h'(0)) = (0, v_0). \end{cases} \tag{2.8}$$

Cette solution est définie sur $\mathbb{R}_+$ tout entier, concave, et h' ainsi que $t \mapsto \frac{h(t)}{t}$ décroissent de v_0 à 0.

Preuve du lemme : On démontre l'existence et l'unicité locale en adaptant la preuve du théorème de Cauchy–Lipschitz sur un espace fonctionnel intégrant la condition initiale $h(0) = 0$. Montrons que h est strictement concave sur son intervalle de définition. Supposons par l'absurde que la dérivée seconde de h s'annule et posons $T_1 = \inf\{t > 0, h''(t) = 0\}$. Comme

$$\lim_{t\to 0^+} \frac{h(t)}{t} = v_0 \quad \text{et} \quad \lim_{t\to 0^+} \frac{h(t)}{2t} - h'(t) = -\frac{1}{2}v_0,$$

on a $h''(0) = -\frac{1}{2m}v_0^2 < 0$ et donc $T_1 > 0$. Ainsi h est strictement concave sur l'intervalle $[0, T_1]$. Deux cas sont alors possibles.

1. Soit $\frac{h(T_1)}{T_1} = 0$ et $\frac{h(T_1)}{2T_1} - h'(T_1) \leq 0$. Dans ce cas $h'(T_1) \geq 0$, et par stricte concavité de h sur $[0, T_1]$, $\frac{h(T_1)}{T_1} > h'(T_1) \geq 0$, ce qui contredit $\frac{h(T_1)}{T_1} = 0$.
2. Soit $\frac{h(T_1)}{T_1} > 0$ et $\frac{h(T_1)}{2T_1} - h'(T_1) = 0$. Posons alors $\gamma(t) = \frac{h(t)}{2t} - h'(t)$. On a $\gamma(T_1) = 0$ et
$$\gamma'(T_1) = \frac{h'(T_1)}{2T_1} - \frac{h(T_1)}{2T_1^2} - h''(T_1) = \frac{1}{2T_1}\left(h'(T_1) - \frac{h(T_1)}{T_1}\right) < 0,$$
ce qui contredit que γ est strictement négative sur $[0, T_1[$.

Les deux termes apparaissant dans (2.8) sont de signe constant, donc

$$0 \leq \frac{h(t)}{t} \leq h'(t) \leq v_0$$

et la solution est globale. Les deux quantités $t \mapsto \frac{h(t)}{t}$ et h' décroissent vers la même limite. En effet, soit $l = \lim_{t\to+\infty} h'(t)$, soit $\epsilon > 0$ et soit t_0 tel que pour tout $t > t_0$, $|h'(t) - l| < \frac{\epsilon}{2}$. Alors pour tout $t > \max(t_0, \frac{2(v_0-l)t_0}{\epsilon})$ on a

$$\left|\frac{h(t)}{t} - l\right| \leq \frac{\int_0^t (h'(s)-l)ds}{t} \leq \frac{t_0(v_0-l)}{t} + \frac{t-t_0}{t}\frac{\epsilon}{2} \leq \epsilon.$$

Il s'ensuit que h'' a également une limite, égale à $-\frac{l^2}{2m}$, qui est nécessairement nulle puisque h' a une limite finie. □

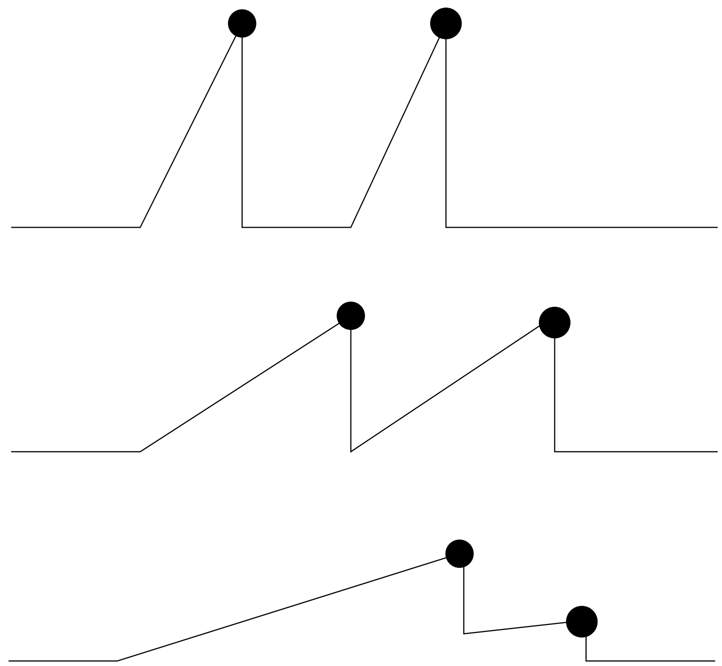


FIGURE 2.4. *Phénomène d'entrainement-dépassement-collision. Jusqu'au temps t_1, les deux particules « ne se voient pas » (en haut). À l'instant t_1, au milieu, la particule de gauche rencontre l'onde créée par celle de droite. Par la suite, la particule de gauche rencontre un fluide qui va à vitesse positive, ce qui l'accélère (en bas).*

Comme $v_0 > 0$, on peut choisir h_2^0 assez proche de h_1^0 de sorte qu'il existe un temps t_1 tel que $h_1(t_1) = h_2^0$. À partir de cet instant, la première particule rencontre sur sa droite l'onde créée par la particule 2 et donc une vitesse de fluide plus grande que u_0. L'allure de la solution avant le temps t_1 est donnée par la figure 2.4, en haut. Une solution est alors donnée, tant que $h_1(t) < h_2(t)$, par

$$\begin{cases} u_0 & \text{si } x < 0, \\ x/t & \text{si } 0 \le x \le h_1(t), \\ (x - h_2^0)/t & \text{si } h_1(t) < x \le h_2(t), \\ u_0 & \text{si } h_2(t) < x. \end{cases}$$

On a alors

$$u(t, h_1(t)_-) = \frac{h_1(t)}{t},\ u(t, h_1(t)_+) = \frac{h_1(t) - h_2^0}{t},\ u(t, h_2(t)_-) = \frac{h_2(t) - h_2^0}{t},\ u(t, h_2(t)_+) = 0$$

et on en déduit les équations différentielles que doivent vérifier h_1 et h_2 :

$$m_1 h_1''(t) = \frac{h_1(t_1)}{t} \left(\frac{h_1(t)}{t} - \frac{h_2^0}{2t} - h_1'(t) \right) \tag{2.9}$$

et

$$m_2 h_2''(t) = \left(\frac{h_2(t) - h_2^0}{t} - u_0 \right) \left(\frac{h_2(t) - h_2^0}{2t} + \frac{u_0}{2} - h_2'(t) \right).$$

L'équation régissant la trajectoire de la seconde particule reste inchangée donc, d'après le lemme 2.3, sa vitesse tend vers 0. On peut montrer que h_2' décroît vers 0. L'équation (2.9) se réécrit

$$m_1 h_1''(t) = -h_2^0 \left(\frac{h_1(t) - \frac{1}{2} h_1(t_1)}{t} \right)'.$$

La fonction $g(t) = h_1(t) - \frac{1}{2}h_1(t_1)$ vérifie l'équation différentielle ordinaire

$$\begin{cases} g'(t) + \alpha \frac{g(t)}{t} = \beta \\ g(t_1) = \gamma \end{cases}$$

avec $\gamma = \frac{h_1(t_1)}{2}$, $\alpha = \frac{h_1(t_1)}{m_1}$ et $\beta = h_1'(t_1) + \frac{h_2^0 \gamma}{t_1 m_1}$. Par conséquent,

$$g(t) = \frac{\beta}{\alpha+1} t + \left(\gamma t_1^\alpha - \frac{\beta t_1^{\alpha+1}}{\alpha+1} \right) t^{-\alpha}$$

et on a

$$g'(t) \xrightarrow{t \to +\infty} \frac{\beta}{\alpha+1}.$$

On en déduit que la vitesse h_1' de la particule de gauche reste plus grande que $u_0 + \frac{h_2^0}{2t_1(h_2^0+m_1)}$ qui est strictement positive. Il reste à vérifier que pour presque tout temps,

$$\left(\frac{h_1(t)}{t}, \frac{h_1(t) - h_2^0}{t} \right) \in \mathcal{G}_\lambda(h_1'(t)).$$

On a

$$\begin{aligned} \frac{g(t)}{t} - g'(t) &= (1+\alpha) \left(\gamma t_1^\alpha - \frac{\beta t_1^{\alpha+1}}{\alpha+1} \right) t^{-\alpha-1} \\ &= (\gamma - h_1'(t_1)) t_1^\alpha t^{-\alpha-1}. \end{aligned}$$

Ainsi, $t \mapsto \frac{g(t)}{t} - g'(t)$ est monotone et tend vers 0. Supposons dans un premier temps que $\gamma - h_1'(t_1)$ est positif. La fonction $t \mapsto \frac{g(t)}{t} - g'(t)$ est décroissante donc

$$\frac{h_1(t) - \frac{1}{2}h_1(t_1)}{t} - h_1'(t) \geq 0$$

et en particulier, $\frac{h_1(t)}{t} \geq h_1'(t)$. D'autre part, on a

$$\begin{aligned} \frac{h_1(t)}{t} - h_1'(t) &= \frac{g(t)}{t} - g'(t) + \frac{h_1(t_1)}{2t} \\ &\leq \frac{g(t_1)}{t_1} - g'(t_1) + \frac{h_1(t_1)}{2t_1} \\ &\leq \frac{h_1(t_1)}{t_1} \leq v_0 \leq \lambda. \end{aligned}$$

Montrons maintenant que

$$\frac{h_1(t) - h_1(t_1)}{t} \in \left[2h_1'(t) - \frac{h_1(t)}{t} - \lambda, h_1'(t) \right].$$

D'une part, $\frac{h_1(t) - h_2^0}{t} \leq h'(t)$ si et seulement si $\frac{g(t)}{t} - g'(t) \leq \frac{h_1(t_1)}{2t}$. On le vérifie facilement :

$$t \left(\frac{g(t)}{t} - g'(t) \right) = (\gamma - h_1'(t_1)) t_1^\alpha t^{-\alpha} \leq (\gamma - h_1'(t_1)) \leq \gamma = \frac{h_1(t_1)}{2}.$$

D'autre part, la seconde inégalité est vérifiée si et seulement si $\frac{g(t)}{t} - g'(t) \geq -\frac{\lambda}{2}$ ce qui est le cas puisque cette dernière fonction décroît vers 0.

Supposons maintenant que $\gamma - h_1'(t_1)$ est négatif, autrement dit que $t \mapsto \frac{g(t)}{t} - g'(t)$ croît vers 0. En particulier, cette quantité est bien plus petite que $\frac{h_1(t_1)}{2t}$. De plus, d'après le lemme 2.8,

$$\frac{g(t)}{t} - g'(t) \geq \frac{g(t_1)}{t_1} - g'(t_1) \geq \frac{h(t_1)}{2t_1} - h'(t_1) \geq -\frac{h'(t_1)}{2} \geq \frac{-\lambda}{2}.$$

Il nous reste à montrer que $\frac{h_1(t)}{t} \in [h_1'(t), h_1'(t) + \lambda]$. Le lemme 2.8 donne

$$h_1'(t) + \lambda - \frac{h_1(t)}{t} = g'(t) + \lambda - \frac{g(t)}{t} - \frac{h_1(t_1)}{2t} \geq \lambda - \frac{h_1(t_1)}{2t_1} \geq \lambda - \frac{v_0}{2} \geq 0$$

et, pour conclure,

$$\frac{h_1(t) - h_1(t_1)}{t} - h_1'(t) = \frac{g(t)}{t} - g'(t) - \frac{h_1(t_1)}{2t_1} \leq -\frac{h_1(t_1)}{2t_1} \leq 0.$$

2.3 Résultats d'existence et d'unicité

Les premiers résultats d'existence et d'unicité, obtenus dans [LST08], concernent le problème de Riemann. Par *solution*, on entend solution au sens de la définition 2.2 (sans le troisième point si la trajectoire de la particule est donnée, comme dans le théorème ci-dessous).

Théorème 2.4 (Lagoutière, Seguin, Takahashi). *Pour tout u_L, u_R et v dans $\mathbb{R}$, il existe une unique solution autosemblable (i.e qui ne dépend que de x/t) au problème de Riemann lorsque la particule a une vitesse fixée $v \in \mathbb{R}$*

$$\begin{cases} \partial_t u + \partial_x \frac{u^2}{2} = -\lambda(u - v)\delta_{vt}(x), \\ u_{|t=0} = u_L \mathbf{1}_{x<0} + u_R \mathbf{1}_{x\geq 0}. \end{cases} \tag{2.10}$$

D'autre part, il existe au moins une solution au problème de Riemann pour le problème entièrement couplé (2.1) *avec*

$$u_{|t=0} = u_L \mathbf{1}_{x<0} + u_R \mathbf{1}_{x\geq 0},\ h^0 = 0 \quad \textit{et} \quad v^0 = v.$$

La construction de la solution est explicite. Lorsque la particule a une vitesse fixe, la solution est entièrement décrite par la donnée des traces autour de la particule $u_- = u(t, vt^-)$ et $u_+ = u(t, vt^+)$. La figure 2.5 décrit l'endroit où se situe (u_-, u_+) sur le germe, en fonction de la donnée initiale (u_L, u_R).

La connaissance de la solution dans le cas où la particule a une vitesse fixée permet de construire une solution pour (2.1), car les traces autour de la particule ne changent pas à tout instant. Dans [AS12], les auteurs démontrent que le problème de Cauchy pour une particule immobile est bien posé.

Théorème 2.5 (Andreianov, Seguin). *Soit u^0 une fonction de $L^\infty(\mathbb{R})$. Le système* (2.4) *a une unique solution dans $L^\infty(\mathbb{R}_+ \times \mathbb{R})$.*

La preuve de l'existence repose sur l'étude de convergence d'un schéma numérique qui sera présenté dans la prochaine section. Les estimations sur la variation totale de ce schéma sont obtenues de manière plus simples dans [ALST13] grâce à un changement de variable astucieux. Cela permet d'extraire une sous-suite convergente, dont on montre que la limite est solution en obtenant une formulation discrète de (2.6). Dans le chapitre 3, la preuve est étendue au cas où la particule bouge librement sous l'influence du fluide, et on obtient l'existence d'une solution de (2.1) quand la donnée initiale est à variation totale bornée. L'unicité découle d'une inégalité de Kato. Lorsque la particule n'est pas immobile, obtenir l'unicité est plus délicat car il faut comparer des solutions avec deux trajectoires de particules différentes. Une telle estimation, dans l'espace des fonctions à variation totale bornée, est donnée dans [ALST13]. De plus, une méthode de point fixe permet de prouver l'existence d'une solution dans L^∞.

Théorème 2.6 (Andreianov, Lagoutière, Seguin, Takahashi). *Pour tout u^0 de $L^\infty(\mathbb{R})$, il existe au moins une solution de* (2.1). *De plus, si u^0 est à variation totale bornée, cette solution est à variation totale bornée pour tout temps et est unique dans $L^\infty(\mathbb{R}_+ \times \mathbb{R})$.*

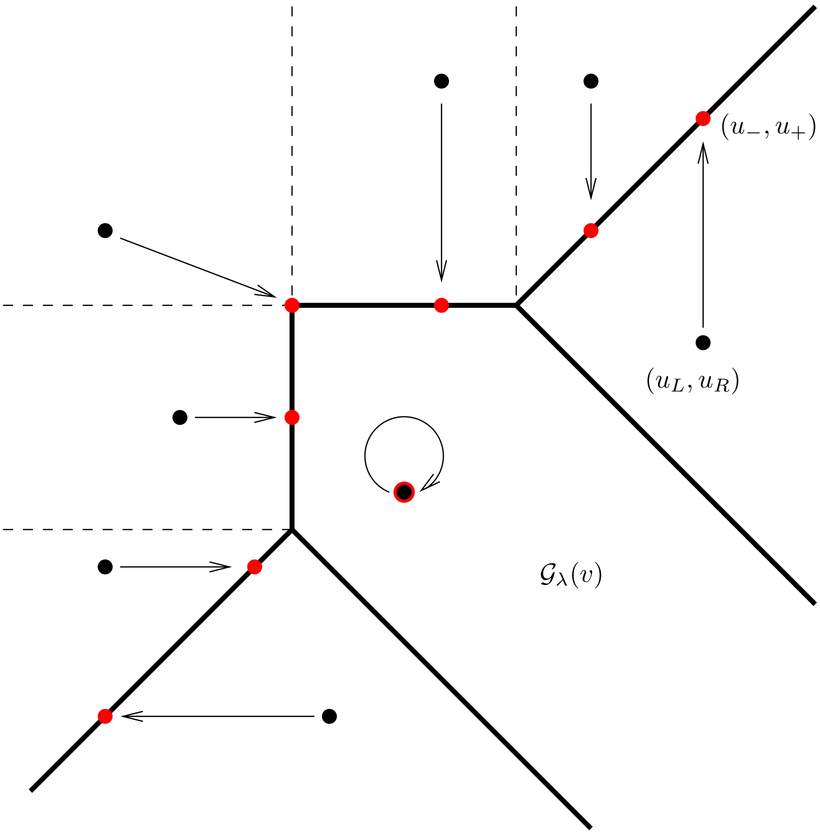


FIGURE 2.5. *Les traces autours de la particule (u_-, u_+) (en gris) en fonction de la donnée initiale (u_L, u_R) (en noir).*

2.4 Schémas numériques pour une particule immobile

Les difficultés numériques pour approcher les solutions du système (2.1) sont de deux ordres. D'une part, il est nécessaire de considérer le système comme un problème d'interface et de prendre en compte numériquement la définition 2.1 du germe. Ainsi la particule doit être vue comme une interface, et elle doit être positionnée sur une interface du maillage à chaque pas de temps. Ce problème du mouvement de la particule sera abordé dans le chapitre 5.

2.4.1 Splitting fluide-particule

Tout d'abord, considérons un schéma naïf pour le système (2.4), où l'évolution du fluide et l'interaction avec la particule sont traités séparément. Pour calculer la solution au temps Δt de (2.4), on calcule d'abord la solution u^* au temps Δt du système sans particule

$$\begin{cases} \partial_t u + \partial_x \frac{u^2}{2} = 0, \\ u_{|t=0} = u^0. \end{cases} \tag{2.11}$$

La solution u est approchée par la solution à l'instant Δt de la phase d'interaction fluide-particule pure

$$\begin{cases} \partial_t u = -\lambda(u - v)\delta_{vt}(x), \\ u_{|t=0} = u^*. \end{cases} \tag{2.12}$$

Le schéma numérique correspondant est le suivant. On suppose ici que $v = 0$ et que $x_0 = 0$, de sorte que la particule se situe toujours dans la maille numérotée 0. Le schéma est initialisé avec une moyenne exacte de la solution :

$$u_j^0 = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} u^0(x)\, dx.$$

On effectue ensuite un splitting de Lie entre la partie fluide et la partie interaction fluide-particule. La partie fluide est obtenue grâce à un schéma de volumes finis :

$$u_j^{n+1/2} = u_j^n - \frac{t^{n+1} - t^n}{\Delta x}(f_{j+1/2}^n - f_{j-1/2}^n).$$

L'interaction fluide-particule est locale, et n'impacte que la cellule 0. Finalement, la solution au temps t^{n+1} est donnée par

$$u_j^n = \begin{cases} u_j^{n+1/2} & \text{si } j \neq 0, \\ u_0^{n+1/2} - \frac{t^{n+1}-t^n}{\Delta x}\lambda u_0^{n+1/2} & \text{si } j = 0. \end{cases}$$

Comme on peut le voir sur la Figure 2.6, ce schéma ne converge pas vers la bonne solution. On a pris $\lambda = 1$ et la donnée initiale

$$u^0(x) = 3 * \mathbf{1}_{x<0} + 2 * \mathbf{1}_{x\geq 0},$$

qui est un état d'équilibre du système : $(3, 2) \in \mathcal{G}_1(0)$. La solution exacte est $u(t, x) = u^0(x)$. Le schéma fait apparaître un palier intermédiaire, dont la hauteur dépend du rapport entre le pas d'espace et le pas de temps. Ici, la partie fluide est résolue par une méthode de Godunov. L'intervalle d'espace $[-1, 1]$ est discrétisé avec $4\,000$ mailles et on a pris à chaque itération en temps

$$t^{n+1} - t^n = \frac{0.2\Delta x}{\max_{j\in\mathbb{Z}} |u_j^n| + \lambda}.$$

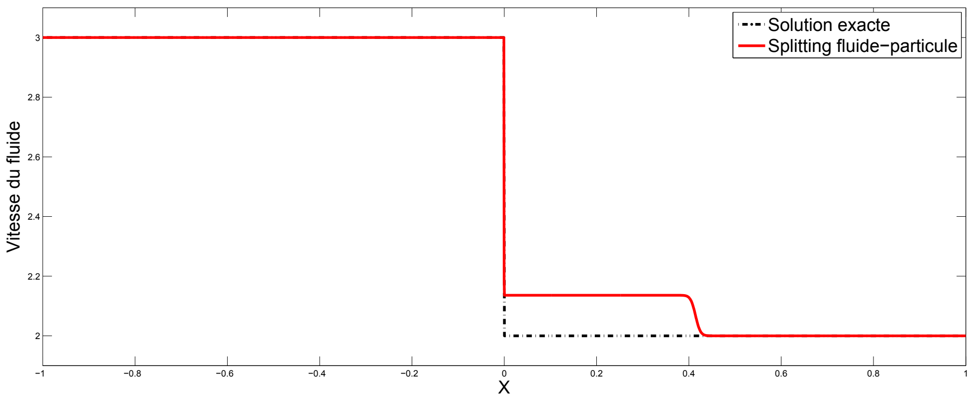


FIGURE 2.6. *Non-convergence du schéma de splitting naïf.*

L'exemple précédent illustre la nécessité de considérer la particule comme une interface. Dorénavant, on la place sur l'interface $x_{1/2}$ du maillage, et on se concentre sur le cas où la particule est immobile, repoussant les questions sur la gestion du mouvement de la particule à plus tard. En dehors des mailles 0 et 1 de part de d'autre de la particule, le fluide suit une équation de Burgers sans terme source et peut donc être approché numériquement par n'importe quel schéma de volumes finis. La question ici est de définir les flux à l'interface $1/2$. La quantité u n'étant pas conservée au travers de la particule, on calcule un flux à gauche de la particule et un flux à droite. Finalement, les schémas numériques qui nous intéressent s'écrivent

$$\begin{cases} u_j^{n+1} = u_j^n - \frac{t^{n+1}-t^n}{\Delta x}(g(u_j^n, u_{j+1}^n) - g(u_{j-1}^n, u_j^n)) & \text{si } j \notin \{0,1\}, \\ u_0^{n+1} = u_0^n - \frac{t^{n+1}-t^n}{\Delta x}(g_\lambda^+(u_0^n, u_1^n) - g(u_{-1}^n, u_0^n)) & \text{si } j = 0, \\ u_1^{n+1} = u_1^n - \frac{t^{n+1}-t^n}{\Delta x}(g(u_1^n, u_2^n) - g_\lambda^-(u_0^n, u_1^n)) & \text{si } j = 1. \end{cases} \tag{2.13}$$

On supposera que le flux numérique g vérifie les hypothèses usuelles : consistance avec le flux de Burgers, croissance selon la première variable et décroissance selon la seconde.

2.4.2 Méthodes de mailles fantômes

Ces méthodes consistent à introduire deux mailles supplémentaires de part et d'autre de la particule, et à déterminer la vitesse du fluide $u^n_{\text{part},-}$ et $u^n_{\text{part},+}$ en fonction de u_0^n et u_1^n dans ces mailles. On peut alors utiliser un schéma de volumes finis usuel pour actualiser la vitesse du fluide, y compris dans les mailles 0 et -1 autour de la particule :

$$\begin{cases} g_\lambda^-(u_0^n, u_1^n) & = g(u_0^n, u^n_{\text{part},-}), \\ g_\lambda^+(u_0^n, u_1^n) & = g(u^n_{\text{part},+}, u_1^n). \end{cases} \tag{2.14}$$

Une première possibilité est d'utiliser la solution exacte du problème de Riemann avec particule pour déterminer $u^n_{\text{part},-}$ et $u^n_{\text{part},+}$. On note $U(x/t)$ la solution (autosemblable) de (2.10) avec $u_L = u_0^n$, $u_R = u_1^n$ et $v = 0$, et on pose

$$\begin{cases} u^n_{\text{part},-} = \lim_{\xi\to 0, \xi<0} U(\xi), \\ u^n_{\text{part},+} = \lim_{\xi\to 0, \xi>0} U(\xi). \end{cases}$$

Selon la terminologie de [GL96], ce choix fournit un schéma équilibre qui conserve tout le germe $\mathcal{G}_\lambda(0)$: si $u_j^0 = u_L$ pour tout $j \le 0$ et $u_j^0 = u_R$ pour tout $j \ge 1$, avec $(u_L, u_R) \in \mathcal{G}_\lambda(0)$, alors pour tout n positif, on a également $u_j^n = u_L$ pour tout $j \le 0$ et $u_j^n = u_R$ pour tout $j \ge 1$. Notons de plus que si le flux numérique g est le flux de Godunov, le schéma (2.13) avec (2.14) n'est rien d'autre que le schéma de Godunov classique avec prise en compte de la particule dans les mailles 0 et 1. Dans [AS12], les auteurs proposent un schéma qui ne conserve pas tous les états d'équilibre de $\mathcal{G}_\lambda(0)$, mais seulement ceux de $\mathcal{G}_\lambda^1$. L'idée est de déterminer $u^n_{\text{part},+}$ comme une trace à droite de la particule si la vitesse du fluide à gauche de la particule est u_0^n. Au vu de la définition 2.1, il peut exister plusieurs états $u^n_{\text{part},+}$ tels que $(u_0^n, u^n_{\text{part},+})$ appartienne à $\mathcal{G}_\lambda(v)$. En revanche, si on demande que $(u_0^n, u^n_{\text{part},+})$ appartienne à $\mathcal{G}_\lambda^1(v)$ il n'y a plus qu'une possibilité, donnée par $u^n_{\text{part},+} = u_0^n - \lambda$. On raisonne de même pour déterminer $u^n_{\text{part},-}$ à partir de u_1^n, et on trouve :

$$\begin{cases} u^n_{\text{part},-} = u_1^n + \lambda, \\ u^n_{\text{part},+} = u_0^n - \lambda. \end{cases} \tag{2.15}$$

Le convergence du schéma correspondant est prouvée dans [AS12].

Théorème 2.7 (Andreianov, Seguin). *Supposons que le flux g vérifie les hypothèses classiques, et que la fonction*

$$(a, b) \mapsto \partial_a g(a, b) + \partial_b g(a, b) \tag{2.16}$$

est croissante selon ses deux variables. Alors, sous la conditions CFL

$$\forall n,\ 2\max\{|g'(u)|, u \in [\min u^0 - \lambda, \min u^0 + \lambda]\}(t^{n+1} - t^n) \le \Delta x,$$

le schéma numérique (2.13–2.15) converge vers la solution de (2.4).

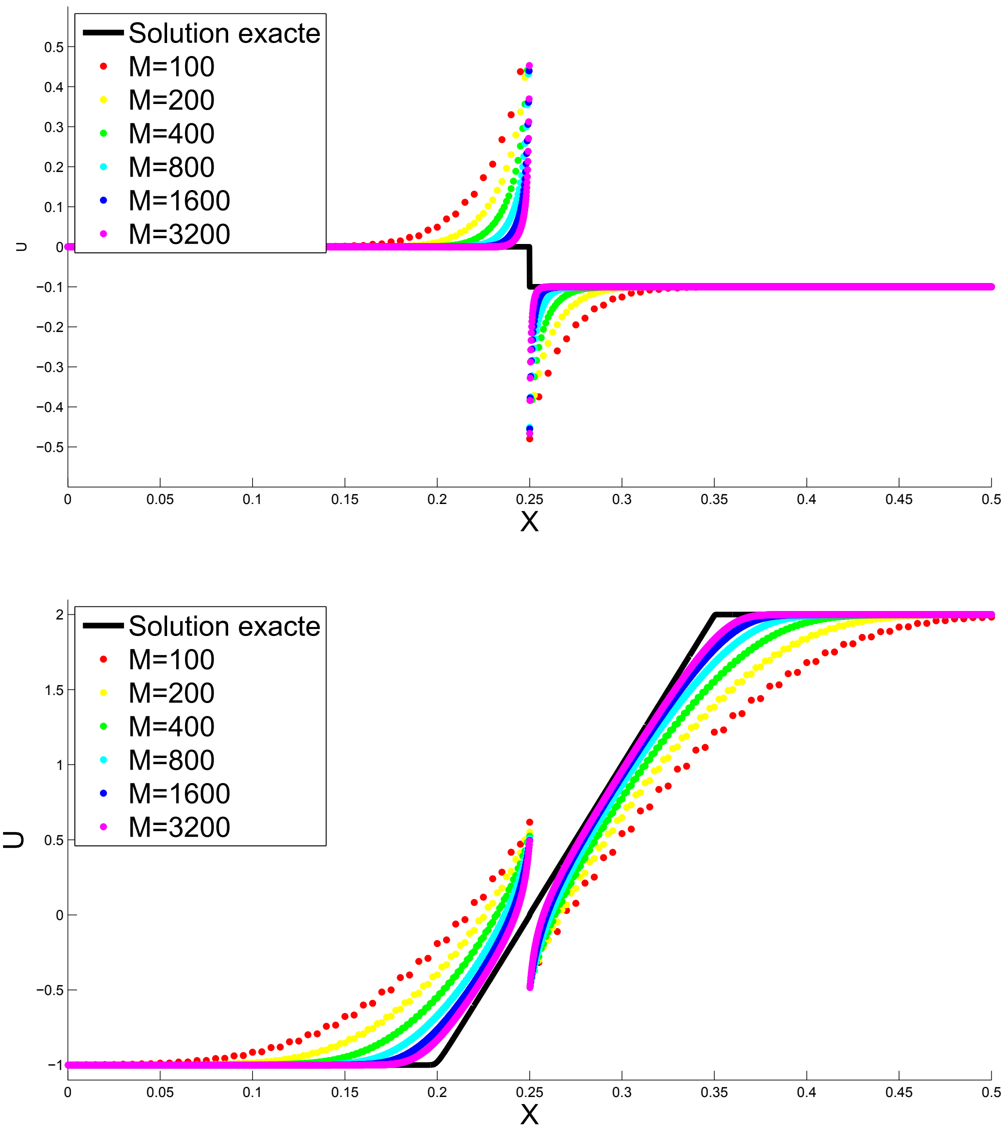


FIGURE 2.7. *Apparition de couches limites avec le schéma* (2.15). *La particule est immobile en* $x = 0.25$. *La hauteur des couches limites est égale au paramètre de friction* $\lambda = 1$. *Le flux numérique* g *est celui de Lax–Friedrichs, et le temps final est* $T = 0.1$ *en haut et* $T = 0.05$ *en bas.*

L'inconvénient de ce schéma est qu'il fait apparaître des couches limites autour de la particule. Cela est illustré par la figure 2.7, où l'on voit également que ces couches limites n'empêchent pas le schéma de converger. Le chapitre 3 étend ce résultat au cas où la particule bouge sous l'influence du fluide (on utilise alors un maillage qui suit la particule). Dans ce cas, la condition technique (2.16) assure que l'accélération numérique de la particule a le bon signe.

2.4.3 Schéma de relaxation

Les schémas de relaxation s'appuient sur la résolution exacte du problème de Riemann d'un système augmenté, où les flux non linéaires sont remplacées par de nouvelles variables, qu'on force à être proches du flux initial grâce à un terme source très raide (voir par exemple [JX95] et [CLL94]). Par exemple, si $(u^\varepsilon, \eta^\varepsilon)$ est solution du système relaxé

$$\begin{cases} \partial_t u_\varepsilon + \partial_x \eta_\varepsilon = 0, \\ \partial_t \eta_\varepsilon + A^2 \partial_x u_= \frac{1}{\varepsilon}\left(\frac{u_\varepsilon^2}{2} - \eta_\varepsilon\right), \\ u_(t=0,x) = u^0(x), \eta_\varepsilon(t=0,x) = \frac{(u^0(x))^2}{2}, \end{cases}$$

alors u^ε est moralement proche, quand le paramètre ε positif tend vers 0, de la solution de

$$\begin{cases} \partial_t u + \partial_x \left(\frac{u^2}{2}\right) = 0, \\ u(t=0,x) = u^0(x). \end{cases}$$

En effet, quand ε devient petit, le terme source explose si η^ε est différent de $\frac{u^2}{2}$, et on retrouve l'équation de Burgers sur la première ligne lorsque $\eta = \frac{u^2}{2}$. Faisons maintenant un développement de Chapman–Enskog à l'ordre 1 : $\eta_\varepsilon = \frac{u_\varepsilon^2}{2} + \varepsilon w_\varepsilon + O(\varepsilon^2)$. On obtient

$$\begin{cases} \partial_t u_\varepsilon + \partial_x \frac{(u_\varepsilon)^2}{2} = O(\varepsilon), \\ \partial_t \frac{u_\varepsilon^2}{2} + A^2 \partial_x u_\varepsilon + w_\varepsilon = O(\varepsilon), \\ u_\varepsilon(t=0,x) = u^0(x), \eta_\varepsilon(t=0,x) = \eta^0(x). \end{cases}$$

En éliminant formellement la dérivée en temps, il vient $w_\varepsilon = -(A^2 - u_\varepsilon^2)\partial_x u_\varepsilon$, et le système hyperbolique vérifié par u_ε à l'ordre 1 est

$$\partial_t u_\varepsilon + \partial_x \frac{u_\varepsilon^2}{2} = -\varepsilon \partial_x w_\varepsilon = \varepsilon \partial_x((A^2 - u_\varepsilon^2)\partial_x u_\varepsilon).$$

Le paramètre A doit donc vérifier la condition sous-caractéristique

$$|A| > \max |u_\varepsilon|.$$

Pour le système avec particule (2.4), le système de relaxation s'écrit

$$\begin{cases} \partial_t u + \partial_x \eta = -\lambda u \delta_0, \\ \partial_t \eta + A^2 \partial_x u = \frac{1}{\varepsilon}\left(\frac{u^2}{2} - v\right) - A^2 \mu \delta_0, \\ u(t=0,x) = u^0(x), v(t=0,x) = \eta^0(x). \end{cases} \tag{2.17}$$

On a introduit un terme source $-A^2\mu\delta_0$ sur l'équation en η pour autoriser la variable u a être discontinue au travers de la particule. Dans [CC13], cette même stratégie est adoptée pour fournir un schéma capable d'approcher exactement les chocs isolés dans les équations d'Euler en coordonnées lagrangiennes. On approche la solution de (2.4) au temps Δt comme la solution de (2.17) au temps Δt avec des données initiales à l'équilibre $\eta^0 = \frac{(u^0)^2}{2}$. Cela revient à effectuer un splitting entre la relaxation, qu'on suppose infiniment rapide, et le système linéarisé sans terme source. La solution au problème de Riemann

$$\begin{cases} \partial_t u + \partial_x \eta = -\lambda u \delta_0, \\ \partial_t \eta + A^2 \partial_x u = -A^2 \mu \delta_0, \\ u(t=0,x) = u_L \mathbf{1}_{x<0} + u_R \mathbf{1}_{x\geq 0}, \\ v(t=0,x) = \frac{u_L^2}{2} \mathbf{1}_{x<0} + \frac{u_R^2}{2} \mathbf{1}_{x\geq 0}. \end{cases}$$

contient trois ondes se déplaçant aux vitesses respectives $-A$, 0 et A. L'onde 0 est définie comme auparavant grâce à une particule épaissie. On notera (u_-, v_-) l'état à sa gauche et (u_+, v_+) l'état à sa droite. On introduit H_ε une régularisation de la fonction de Heaviside, croissant de 0 à 1 sur l'intervalle $[-\varepsilon/2, \varepsilon/2]$ et on étudie les solutions stationnaires sur $[-\varepsilon/2, \varepsilon/2]$ de

$$\begin{cases} \partial_t u + \partial_x \eta = -\lambda u H_\varepsilon', \\ \partial_t \eta + A^2 \partial_x u = A^2 \mu H_\varepsilon'. \end{cases}$$

Les solutions de ce système n'admettent pas de discontinuités et on obtient $u' = -\mu H_\varepsilon'$ et $\eta' = -\lambda u H_\varepsilon'$. Finalement quel que soit le choix de H_ε, les valeurs $u_- = -u(-\varepsilon/2)$ et $\eta_- = \eta(-\varepsilon/2)$ à gauche de la particule sont liés aux valeurs $u_+ = u(\varepsilon)$ et $\eta_+ = \eta(\varepsilon)$ à droite de la particule par les relations

$$\begin{cases} u_- - u_+ = \mu, \\ \eta_- - \eta_+ = \lambda \left(u_- - \frac{\mu}{2} \right). \end{cases} \tag{2.18}$$

Ces deux relations seront imposées à travers l'onde 0. Les ondes à vitesses $-A$ et A ne font pas intervenir le second membre, ce sont des ondes fluides classiques. Les relations de Rankine–Hugoniot sont les mêmes pour les deux lignes du système, et on obtient

$$\begin{cases} -A(u_L - u_-) = (\eta_L - \eta_-), \\ A(u_R - u_+) = (\eta_R - \eta_+). \end{cases} \tag{2.19}$$

On rappelle que $\eta_L = \frac{u_L^2}{2}$ et $\eta_R = \frac{u_R^2}{2}$. Les solutions du système (2.18–2.19) sont

$$\begin{cases} u_- = \dfrac{A(u_R + u_L) + (\eta_L - \eta_R) + A\mu + \lambda\mu/2}{\lambda + 2A}, \\ u_+ = u_- - \mu, \\ \eta_- = \eta_L + A(u_L - u_-), \\ \eta_+ = \eta_R - A(u_R - u_+). \end{cases} \tag{2.20}$$

Finalement, les flux numériques correspondant à cette relaxation sont

$$g_\lambda^-(u_L, u_R) = \eta_- \quad \text{et} \quad g_\lambda^+(u_L, u_R) = \eta_+,$$

où η_- et η_+ sont donnés par (2.20). Il reste à préciser comment choisir μ.

Proposition 2.8. *Si $\mu = \lambda$, et si (u_L, u_R) appartient à $\mathcal{G}_\lambda^1$, on a $g_\lambda^-(u_L, u_R) = \frac{u_L^2}{2}$ et $g_\lambda^+(u_L, u_R) = \frac{u_R^2}{2}$. Autrement dit, ce schéma conserve tout les états d'équilibre de $\mathcal{G}_\lambda^1$. De plus, si*

$$|A| > max(|u_L|, |u_R|) + \lambda, \tag{2.21}$$

alors g_λ^- et g_λ^+ sont croissantes selon u_L et décroissantes selon u_R.

Preuve. On vérifie ces deux propriétés par un calcul simple. □

Théorème 2.9. *Sous la condition sous-caractéristique* (2.21) *et sous la condition CFL*

$$\forall n, \quad t^{n+1} - t^n \leq \frac{\Delta x}{2 \max_{x \in \mathbb{R}} |u^0(x)| + \lambda},$$

le schéma numérique de relaxation converge vers la solution du problème (2.4).

Preuve. Ce schéma, tout comme celui de la partie précédente, conserve la partie $\mathcal{G}_\lambda^1$ du germe. Pour adapter la preuve de convergence de [AS12], le point clé est de démontrer que si la donnée initiale est $u_L \mathbf{1}_{x<0} + u_R \mathbf{1}_{x \geq 0}$, alors pour tout n, u_j^n est dans $[0, \lambda]$ pour j strictement négatif et dans $[-\lambda, 0]$ pour j positif. Cela permet de démontrer la convergence du schéma lorsque la donnée initiale est un état stationnaire dans $\mathcal{G}_\lambda^1$ (Proposition 6 de [AS12]), ce qui permet d'obtenir la convergence pour n'importe quelle donnée initiale à variations bornées. Pour montrer ce fait, on démontre que $(u_j^n)_{j \in \mathbb{Z}}$ est croissante et comprise entre 0 et λ pour j strictement négatif, et croissante et comprise entre $-\lambda$ et 0 pour j positif. Une fois qu'on a prouvé que

$$g_\lambda^-(u_L, u_R) \geq g(u_L, u_R + \lambda) \quad \text{si} \quad u_R \geq -\lambda$$

et que

$$g_\lambda^+(u_L, u_R) \leq g(u_L - \lambda, u_R) \quad \text{si} \quad u_L \leq \lambda,$$

la preuve se fait grâce aux propriétés de monotonie de g (voir la preuve du lemme 3.18 dans le chapitre 3). □

Convergence of finite volumes schemes for the coupling between the inviscid Burgers equation and a particle

3

Ce chapitre est consacré à l'analyse numérique de schémas de volumes finis pour le couplage *complet* entre l'équation de Burgers et une particule ponctuelle. Les schémas étudiés dans ce chapitre ne conservent pas tous les états d'équilibre du couplage fluide-particule, mais seulement ceux qui correspondent à une diminution d'un facteur λ (le coefficient de frottement) de la vitesse du fluide à travers la particule. Nous étendons au cas où la particule bouge sous l'influence du fluide le résultat de convergence obtenu dans [AS12] pour une particule immobile. Les deux difficultés principales sont, d'une part, de traiter le couplage entre l'équation différentielle ordinaire décrivant le mouvement de la particule et l'équation aux dérivées partielles décrivant la vitesse du fluide et, d'autre part, de prendre en compte des conditions d'interface qui dépendent du temps.

Ce travail correspond à l'article en préparation [ALS14], en collaboration avec Frédéric Lagoutière et Nicolas Seguin.

Contents

3.1 Introduction

We study the numerical convergence of finite volume schemes for the Cauchy problem

$$\begin{cases} \partial_t u + \partial_x \frac{u^2}{2} = -\lambda(u - h'(t))\delta_{h(t)}(x), \\ m_p h''(t) = \lambda(u(t,h(t)) - h'(t)), \\ u_{|t=0} = u^0, h(0) = h^0, h'(0) = v^0. \end{cases} \tag{3.1}$$

It models the behavior of a pointwise particle of position h, velocity h' and acceleration h'' with mass m_p, immersed into a "fluid," whose velocity at time t and point x is $u(t,x)$. The velocity of the fluid is assumed to follow the inviscid Burgers equation. This system is fully coupled: the fluid exerts a drag force $D = \lambda(u(t, h(t)) - h'(t))$ on the particle, where λ is a positive friction parameter. By the action–reaction principle, the particle exerts the force $-D$ on the particle. The interaction is local: it applies only at the point where the particle is. This friction force tends to bring the velocities of the fluid and the particle closer to each other: as λ is positive, the particle accelerates if $u(t, h(t))$ is larger than $h'(t)$ and vice-versa. This toy model was introduced in [LST08]. In contrast with the model studied in [VZ03] and [VZ06], the particle and the fluid do not share the same velocity and the fluid is inviscid. In particular the fluid velocity is typically discontinuous through the particle. It yields to issues to define correctly the product $(u - h')\delta_h$ and the ODE for the particle in system (3.1). To do so, the idea is to regularize the Dirac measure in (3.1), and to remark that the values of the fluid velocity on both sides of this thickened particle are independent of the regularization. It allows to reformulate System (3.1) as an interface problem, where the traces around the particle $u_-(t) = \lim_{x\to h(t)^-} u(t, x)$ and $u_+(t) = \lim_{x\to h(t)^+} u(t, x)$ must belong to a set $\mathcal{G}_\lambda(h'(t))$, which takes into account the interface conditions. This study was done in details in [LST08]. The germ is defined as follow.

Definition 3.1. *For any given speed $v \in \mathbb{R}$, the germ at speed v, $\mathcal{G}_\lambda(v)$, is the set of all (u_-, u_+) in $\mathbb{R}^2$ such that*

$$(u_-, u_+) \in \mathcal{G}^1_\lambda(v) \cup \mathcal{G}^2_\lambda(v) \cup \mathcal{G}^3_\lambda(v),$$

where

$$\mathcal{G}^1_\lambda(v) = \left\{(u_-, u_+) \in \mathbb{R}^2 : u_- = u_+ + \lambda\right\},$$

$$\mathcal{G}^2_\lambda(v) = \left\{(u_-, u_+) \in \mathbb{R}^2 : v \le u_- \le v + \lambda,\ v - \lambda \le u_+ \le v \ \textit{and} \ u_- - u_+ < \lambda\right\},$$

and

$$\mathcal{G}^3_\lambda(v) = \left\{(u_-, u_+) \in \mathbb{R}^2 : -\lambda \le u_+ + u_- - 2v \le \lambda \ \textit{and} \ u_- - u_+ > \lambda\right\}.$$

The germ $\mathcal{G}_\lambda(0)$ and its partition are depicted on Figure 3.1 on the left (note that the germ $\mathcal{G}_\lambda(v)$ is the translation of $\mathcal{G}_\lambda(0)$ by the vector (v, v)). Here, we choose a slightly different partition of the germ than in [AS12] and [ALST13], which is depicted on the right of Figure (3.1). The reason is that we are able to find a class of finite volume scheme which are consistent with $\mathcal{G}^1_\lambda(0) \cup \mathcal{G}^2_\lambda(0)$ with this choice, but not with the original partition. However, the essential property that $\mathcal{G}^1_\lambda(0) \cup \mathcal{G}^2_\lambda(0)$ is *maximal* part of the germ still holds true with the partition of Definition 3.1 (more details are given in Definition 3.3 and Proposition 3.14). Once the germ has been defined, System (3.1) is defined as an interface problem.

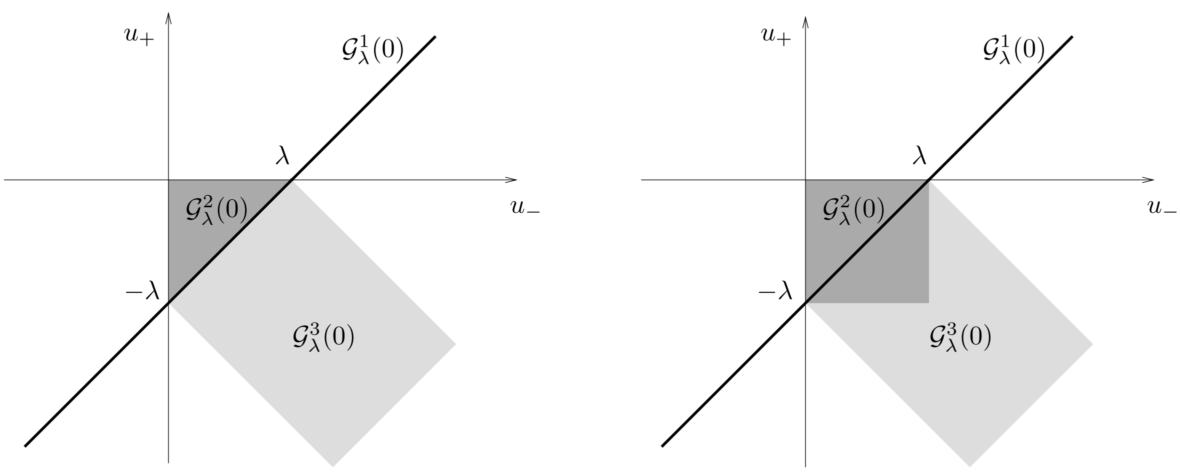


Figure 3.1. *The germ for a motionless particle and its partitions. Left: the partition used in this work. Right: the partition used in [AS12] and [ALST13].*

The equation on the particle is reformulated to keep the conservation of total momentum

$$m_p h'(t) + \int_{\mathbb{R}} u(t,x)dx$$

which holds formally in (3.1). In [LST08], an entropy inequality that takes into account the particle is also derived.

Definition 3.2. *A pair (u,h) of functions in $L^\infty(\mathbb{R}_+ \times \mathbb{R}) \times W^{2,\infty}(\mathbb{R}_+)$ is a solution of* (3.1) *with initial data u^0 in $L^\infty(\mathbb{R})$ and $(h^0, v^0) \in \mathbb{R}^2$ if:*

- *the function u is an entropy weak solution of the Burgers equations on the sets $\{(t,x),\ x < h(t)\}$ and $\{(t,x),\ x > h(t)\}$,*
- *for almost every positive time t,*

$$m_p h''(t) = (u_-(t) - u_+(t)) \left(\frac{u_-(t) + u_+(t)}{2} - h'(t) \right) \tag{3.2}$$

and

$$(u_-(t), u_+(t)) \in \mathcal{G}_\lambda(h'(t)).$$

This definition requires the existence of traces along the particle's trajectory h. It follows from the works of Panov [Pan07] and Vasseur [Vas01]. When the particle is motionless, well-posedness in the BV setting was proved in [AS12], while for the fully coupled system (3.1), it is proved in [ALST10] and [ALST13].

Remark that Definition 3.2 is not suitable to prove convergence of finite volume schemes in a general framework. Indeed, a scheme can create a numerical boundary layer near the particle, of several cells width. It does not prevent the scheme from converging in, say, L^∞_{loc} in time and L^1 in space; but in that case we cannot expect the numerical traces to converge to their correct values. Nevertheless we will prove the convergence of some schemes that create such boundary layers. The key point is to use, instead of Definition 3.1, an equivalent definition which does not contains the traces of u. We begin with some properties useful to decide if a pair (c_-, c_+) belongs to the germ $\mathcal{G}_\lambda(v)$. We adopted the vocabulary of the theory of conservation law with discontinuous flux function of [AKR10] and [AKR11].

In the sequel, we denote by Φ_v the so-called Kruzhkov entropy flux associated with $f_v(u) = \frac{u^2}{2} - vu$:

$$\Phi_v : \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (a,b) & \longmapsto & \operatorname{sign}(a-b)\left(\left(\frac{a^2}{2} - va\right) - \left(\frac{b^2}{2} - vb\right)\right) \end{array}$$

and we define

$$\Xi_v : \begin{array}{ccc} \mathbb{R}^2 \times \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ ((a_-, a_+), (b_-, b_+)) & \longmapsto & \Phi_v(a_-, b_-) - \Phi_v(a_+, b_+) \end{array}$$

Proposition 3.1. *If both (a_-, a_+) and (b_-, b_+) belong to $\mathcal{G}_\lambda(v)$, then*

$$\Xi_v((a_-, a_+), (b_-, b_+)) \geq 0.$$

Conversely, if (a_-, a_+) is such that

$$\forall (b_-, b_+) \in \mathcal{G}_\lambda(v), \quad \Xi_v((a_-, a_+), (b_-, b_+)) \geq 0,$$

then (a_-, a_+) belongs to the germ.

Definition 3.3. *A subset $\mathcal{H}_\lambda(v)$ of $\mathcal{G}_\lambda(v)$ is said to be maximal if any (a_-, a_+) that satisfies*

$$\forall (b_-, b_+) \in \mathcal{H}_\lambda(v), \quad \Xi_v((a_-, a_+), (b_-, b_+)) \geq 0 \tag{3.3}$$

belongs to the germ $\mathcal{G}_\lambda(v)$.

We will prove in Proposition 3.14 that $\mathcal{G}^1_\lambda(v) \cup \mathcal{G}^2_\lambda(v)$ is maximal. In the sequel $\mathcal{H}_\lambda(v)$ always denotes a maximal part of $\mathcal{G}_\lambda(v)$. We now focus on alternative traceless characterizations of entropy solutions. For all (c_-, c_+) we denote by c the piecewise constant function

$$c(t,x) = c_- \mathbf{1}_{x<h(t)} + c_+ \mathbf{1}_{x \geq h(t)},$$

and by $\operatorname{dist}_1(a, X)$ the L^1-distance of a point $a := (a_-, a_+)$ of $\mathbb{R}^2$ to a set X included in $\mathbb{R}^2$:

$$\operatorname{dist}_1((a_-, a_+), X) = \inf_{(x_-,x_+)\in X} |a_- - x_-| + |a_+ - x_+|.$$

Proposition 3.2. *Let h be a function of $W^{2,\infty}_{loc}(\mathbb{R}_+)$ and let u be a function of $L^\infty_{loc}(\mathbb{R}_+ \times \mathbb{R})$, which is an entropy solution of the Burgers equation on the sets $\{(t,x),\ x < h(t)\}$ and $\{(t,x),\ x > h(t)\}$. The following assertions are equivalent.*

- *For almost every time $t > 0$, $(u_-(t), u_+(t))$ belongs to $\mathcal{G}_\lambda(h'(t))$.*
- *For almost every time $t > 0$, for all $(c_-, c_+) \in \mathbb{R}^2$, there exist $\delta \in (0, t)$ and a constant A depending only on $\|u^0\|_\infty$, λ, (c_-, c_+) and $\|h'\|_\infty$ such that for every nonnegative function φ in $\mathcal{C}^\infty_0((t-\delta, t+\delta) \times \mathbb{R})$,*

$$\begin{aligned}\int_{\mathbb{R}_+}\int_{\mathbb{R}} &|u-c|(s,x)\partial_t\varphi(s,x-h(s)) + \Phi_{h'(t)}(u,c)(s,x)\partial_x\varphi(s,x-h(s))dx\,ds \\ &\geq -A\int_{\mathbb{R}_+} \operatorname{dist}_1((c_-,c_+), \mathcal{H}_\lambda(h'(s)))\varphi(s,0)\,ds.\end{aligned} \tag{3.4}$$

Proof. For the sake of completeness we reproduce here the main ingredients of the proof that can be found in [AS12]. Let φ be in $\mathcal{C}^\infty_0((t-\delta, t+\delta)\times\mathbb{R})$, where δ belongs to $(0,t)$. For positive ε, we introduce the function

$$\zeta_\varepsilon(z) = 1 - \min(1, |z|/\varepsilon),$$

whose support is $(-\varepsilon, \varepsilon)$. The support of the function

$$\psi_\varepsilon(t,x) = (1-\zeta_\varepsilon)\varphi(t, x-h(t))$$

is included in $\{(t,x),\ t>0,\ x \neq h(t)\}$. The function u is a entropy solution of the Burgers equation on the sets $\{(t,x),\ x<h(t)\}$ and $\{(t,x),\ x>h(t)\}$, thus for all real κ,

$$\iint_{\mathbb{R}_+\times\mathbb{R}} |u(s,x)-\kappa|\partial_s\psi_\varepsilon(s,x) + \Phi_0(u(s,x),\kappa)\partial_x\psi_\varepsilon dx\,ds \geq 0.$$

But $\partial_s\psi_\varepsilon(s,x) = \partial_s((1-\zeta_\varepsilon)\varphi)(s,x-h(s)) - h'(s)\partial_x((1-\zeta_\varepsilon)\varphi)(s,x-h(s))$, and we using the fact that

$$\Phi_v(a,b) = \Phi_0(a,b) - v|a-b|,$$

we obtain

$$\iint_{\mathbb{R}_+\times\mathbb{R}} |u-c|(s,x)(\partial_s(1-\zeta_\varepsilon)\varphi)(s,x-h(s)) + \Phi_0(u,c)(s,x)\partial_x((1-\zeta_\varepsilon)\varphi)(s,x-h(s))dx\,ds \geq 0.$$

Thus we have

$$\begin{aligned}
&\iint_{\mathbb{R}_+\times\mathbb{R}} |u-c|(s,x)\partial_t\varphi(s,x-h(s)) + \Phi_{h'(t)}(u,c)(s,x)\partial_x\varphi(s,x-h(s))dx\,ds \\
&\geq \liminf_{\varepsilon\to 0} \iint_{\mathbb{R}_+\times\mathbb{R}} |u-c|(s,x)(\partial_t(\zeta_\varepsilon\varphi))(s,x-h(s)) + \Phi_{h'(t)}(u,c)(s,x)(\partial_x(\zeta_\varepsilon\varphi))(s,x-h(s))dx\,ds \\
&= \int_{\mathbb{R}_+} \Phi_{h'(s)}(u_-(s),c_-) - \Phi_{h'(s)}(u_+(s),c_+)\varphi(s,0)ds \\
&= \int_{\mathbb{R}_+} \Xi_{h'(s)}((u_-(s),u_+(s)),(c_-,c_+))\varphi(s,0)ds
\end{aligned}$$

For all s for which the pair $(u_-(s),u_+(s))$ exists and belongs to $\mathcal{G}_\lambda(h'(s))$, we denote by $(\tilde{c}_-(s),\tilde{c}_+(s))$ a L^1-projection of (c_-,c_+) on $\mathcal{H}_\lambda(h'(s))$. We have

$$\begin{aligned}
\Xi_{h'(s)}((u_-(s),u_+(s)),(c_-(s),c_+(s))) &\geq \Xi_{h'(s)}((u_-(s),u_+(s)),(\tilde{c}_-(s),\tilde{c}_+(s))) \\
&\quad - |\Xi_{h'(s)}((u_-(s),u_+(s)),(c_-(s),c_+(s))) - \Xi_{h'(s)}((u_-(s),u_+(s)),(\tilde{c}_-(s),\tilde{c}_+(s)))|.
\end{aligned}$$

Since $(\tilde{c}_-(s),\tilde{c}_+(s))$ belongs to $\mathcal{H}_\lambda(h'(s))$, Proposition 3.1 yields

$$\int_{\mathbb{R}_+} \Xi_{h'(s)}((u_-(s),u_+(s)),(\tilde{c}_-(s),\tilde{c}_+(s)))\varphi(s,0)ds \geq 0.$$

On the other hand

$$\begin{aligned}
&|\Xi_{h'(s)}((u_-(s),u_+(s)),(c_-(s),c_+(s))) - \Xi_{h'(s)}((u_-(s),u_+(s)),(\tilde{c}_-(s),\tilde{c}_+(s)))| \\
&\quad \leq |\Phi_{h'(s)}(u_-(s),c_-(s)) - \Phi_{h'(s)}(u_-(s),\tilde{c}_-(s))| + |\Phi_{h'(s)}(u_+(s),c_+(s)) - \Phi_{h'(s)}(u_+(s),\tilde{c}_+(s))|
\end{aligned}$$

which is smaller than a constant depending only on $||h'||_\infty$, $||u||_\infty$, c and λ (since $c\mapsto\tilde{c}$ depends on λ), multiplied by the L^1-distance between (c_-,c_+) and $(\tilde{c}_-(s),\tilde{c}_+(s))$, and we obtain the result.

Conversely, using a sequence of test functions φ concentrating at a time t for which u has traces in Proposition 3.4, we obtain that for all (c_-,c_+) in $\mathcal{H}_\lambda(h'(t))$,

$$\Xi_{h'(t)}((u_-(t),u_+(t)),(c_-,c_+)) \geq 0,$$

and thus by Proposition 3.1, $(u_-(t),u_+(t))$ belongs to the germ $\mathcal{G}_\lambda(h'(t))$. ☐

Proposition 3.3. *Let u in $L^\infty_{loc}(\mathbb{R}_+\times\mathbb{R})$ be a solution of the Burgers equation on the sets $\{(t,x),\ x<h(t)\}$ and $\{(t,x),\ x>h(t)\}$. Consider a function h in $W^{2,\infty}_{loc}(\mathbb{R}_+)$ which verifies* (3.2) *with initial data $h(0)=h^0$ and $h'(0)=v^0$ almost everywhere if and only if for all $\xi\in\mathcal{C}_0^\infty([0,T))$ and for all $\psi\in\mathcal{C}_0^\infty(\mathbb{R})$ such that $\psi(0)=1$*

$$\begin{aligned}
-\int_0^T m_p h'(t)\xi'(t)dt = m_p v^0\xi(0) &+ \int_{\mathbb{R}}\int_0^T \frac{u^2}{2}(s,x)\xi(s)\psi'(x-h(s))ds\,dx \\
&+ \int_{\mathbb{R}}\int_0^T u(s,x)[\xi'(s)-h'(s)\psi'(x-h(s))]ds\,dx \qquad (3.5)\\
&+ \int_{\mathbb{R}} u^0(x)\psi(x-h(0))\xi(0)dx.
\end{aligned}$$

Proof. This characterization were proved in [ALST10]. It follows from the application of the Green–Gauss theorem and the fact that u is an entropy solution of the Burgers equation away from the particle:

$$
\begin{aligned}
&\int_0^T \int_{\mathbb{R}} \frac{u^2}{2}(s,x)\xi(s)\psi'(x-h(s)) + u(s,x)[\xi'(s) - h'(s)\psi'(x-h(s))]ds\,dx \\
&\quad = \int_0^T \int_{\mathbb{R}} \frac{u^2}{2}(s,x)\partial_x(\xi\psi(x-h(s))) + u(s,x)\partial_s(\xi\psi(x-h(s)))ds\,dx \\
&\quad = -\int_0^T \int_{x\neq h(t)} \left(\partial_x \frac{u^2}{2} + \partial_t u\right)(\xi\psi)dx\,ds - \int_{\mathbb{R}} u^0(x)\psi(x-h(0))\xi(0)dx \\
&\qquad + \int_0^T \xi(s)\left(\left(\frac{u_-^2(s)}{2} - h'(s)u_-(s)\right) - \left(\frac{u_+^2(s)}{2} - h'(s)u_+(s)\right)\right)ds \\
&\quad = \int_0^T m_p h''(s)\xi(s)ds - \int_{\mathbb{R}} u^0(x)\psi(x-h(0))\xi(0)dx.
\end{aligned}
$$

□

We now present the family of finite volume schemes for which we prove convergence. The proof follows the guidelines of the Lax–Wendroff theorem. In Section 3.2, we obtain a TVD bound on the fluid velocity and a $W^{2,\infty}$ bound on the particle's trajectory that allows to extract convergent subsequences in $L^1_{loc}(\mathbb{R}_+ \times \mathbb{R})$ and $W^{1,\infty}_{loc}(\mathbb{R}_+)$. The difficulties are to treat numerically the interface conditions enclosed in the germ and the coupling between an ODE and a PDE. More precisely:

- First, we have to take into account at the numerical level the interface condition of Definition 3.1. We will use schemes that preserves a "sufficiently large" part of the germ.
- Second, to deal with a moving particle. We use a mesh that tracks the particle and we update the particle's velocity by conservation of total impulsion.

Let us fix a time step Δt and a space step Δx. In the sequel we suppose that the time step and the space step are proportional, and we denote by $\mu = \frac{\Delta t}{\Delta x}$ their ratio. We propose to approximate the solution of (3.1) with a finite volume scheme. We use a mesh that follows the particle, which is placed between the cells numbered 0 and 1. The speed of the particle is approximated by a piecewise constant $(v^n)_{n\in\mathbb{N}}$. Given the solution a time $n\Delta t$: we consider that the particle has constant velocity v^n on the whole time step $(n\Delta t, (n+1)\Delta t)$ to update the fluid velocity, then we update v^n by conservation of the total impulsion. The interface $1/2$ where the particle lies is special, and we have to use appropriate fluxes at this interface. Due to the source term, the equation is not conservative around the particle, thus we have two different fluxes $f^{n,-}_{1/2}$ and $f^{n,+}_{1/2}$ on the left and on the right of the particle respectively. Away from the particle, Equation (3.1) writes as a scalar conservation law, and we can use any standard flux for the Burgers equation. The scheme is initialized with

$$\forall j \in \mathbb{Z},\ u_j^0 = \frac{1}{\Delta x}\int_{x^0_{j-1/2}}^{x^0_{j+1/2}} u^0(x)\,dx.$$

From the integration of the first equation of (3.1) on the space time cell

$$\mathcal{C}_j^n = \{(n\Delta t + s, x^n_{j-1/2} + y + sv^n), 0 \le s < \Delta t, 0 \le y < \Delta x\},$$

we obtain the finite volume scheme

$$\begin{cases} u_j^{n+1} &= u_j^n - \mu(f_{j+1/2}^n(v^n) - f_{j-1/2}^n(v^n)) \text{ for } j \in \mathbb{Z}, j \notin \{0,1\}, \\ u_0^{n+1} &= u_0^n - \mu(f_{1/2,-}^n(v^n) - f_{-1/2}^n(v^n)), \\ u_1^{n+1} &= u_1^n - \mu(f_{3/2}^n(v^n) - f_{1/2,+}^n(v^n)), \\ v^{n+1} &= v^n + \frac{\Delta t}{m_p}(f_{1/2,-}^n(v^n) - f_{1/2,+}^n(v^n)), \\ x_j^{n+1} &= x_j^n + v^n \Delta t. \end{cases} \tag{3.6}$$

Here we emphasized the dependency of the flux on the particle's velocity. In the sequel we denote by $u_{\Delta t}$ the constant by cell function

$$u_{\Delta t}(t,x) = u_j^n \quad \text{if} \quad (t,x) \in \mathcal{C}_j^n. \tag{3.7}$$

and by $v_{\Delta t}$ and $h_{\Delta t}$ the constant and linear by cell functions:

$$\begin{cases} v_{\Delta t}(t) &= v^n \quad \text{if } n\Delta t \le t < (n+1)\Delta t, \\ h_{\Delta t}(t) &= h^0 + \Delta t \sum_{m=0}^{n-1} v^m + v^n(t - n\Delta t) \quad \text{if } n\Delta t \le t < (n+1)\Delta t. \end{cases} \tag{3.8}$$

Another way to proceed is to performed the change of variable

$$\tilde{u}(t,x) = u(t, x + h(t))$$

in (3.1). This function verifies the PDE

$$\partial_t \tilde{u} + \partial_x \left(\frac{\tilde{u}^2}{2} - h'(t)\tilde{u} \right) = -\lambda(\tilde{u} - h')\delta_0(x) \tag{3.9}$$

The particle is now motionless but the flux depends on time. We denote by $f_v(u) = \frac{u^2}{2} - vu$. Integrating (3.9) on $[n\Delta t, (n+1)\Delta t] \times [x_{j-1/2}^0, x_{j+1/2}^0]$, and using special flux around the particle (still placed at interface 1/2), we obtain the finite volume scheme

$$\begin{cases} \tilde{u}_j^{n+1} &= \tilde{u}_j^n - \mu(f_{j+1/2}^{v^n,n} - f_{j-1/2}^{v^n,n}) \text{ for } j \in \mathbb{Z} \setminus \{0,1\}, \\ \tilde{u}_0^{n+1} &= \tilde{u}_0^n - \mu(f_{1/2}^{v^n,n-} - f_{-1/2}^{v^n,n}), \\ \tilde{u}_1^{n+1} &= \tilde{u}_1^n - \mu(f_{3/2}^{v^n,n} - f_{1/2}^{v^n,n+}), \\ v^{n+1} &= v^n + \frac{\Delta t}{m_p}(f_{1/2}^{v^n,n-} - f_{1/2}^{v^n,n+}). \end{cases} \tag{3.10}$$

The two points of view are illustrated on Figure 3.2.

The fluxes $f_{j-1/2}^n(v^n)$ with $j \neq 1/2$ (or $f_{1/2}^{n\pm}(v^n)$ if $j = 1/2$) are strongly related to the fluxes $f_{j-1/2}^{v^n,n}$: in (3.6), $f_{j+1/2}^n(v^n)$ is an approximation of

$$\frac{1}{\Delta t} \int_{n\Delta t}^{(n+1)\Delta t} (f^0(u) - v^n u)(t, x_{j+1/2}^n + v^n t) dt$$

while in (3.10), $f_{-1/2}^{v^n,n}$ is an approximation of

$$\frac{1}{\Delta t} \int_{n\Delta t}^{(n+1)\Delta t} f^{v^n}(\tilde{u})(t, x_{j+1/2}^0) dt$$

In the following we prove the convergence of Scheme (3.6) under a set of assumptions on the fluxes $f_{j+1/2}^n$, $f_{1/2,-}^n$ and $f_{j+1/2,+}^n$ and a Courant-Friedrichs-Lewy condition. We restrict the study to two-points fluxes

$$f_{j+1/2}^n = g(u_j^n, u_{j+1}^n, v^n) \quad \text{and} \quad f_{1/2,\pm}^n = g_\lambda^\pm(u_j^n, u_{j+1}^n, v^n).$$

The assumptions on the flux $f_{j+1/2}^n$ away from the particle are the classical ones:

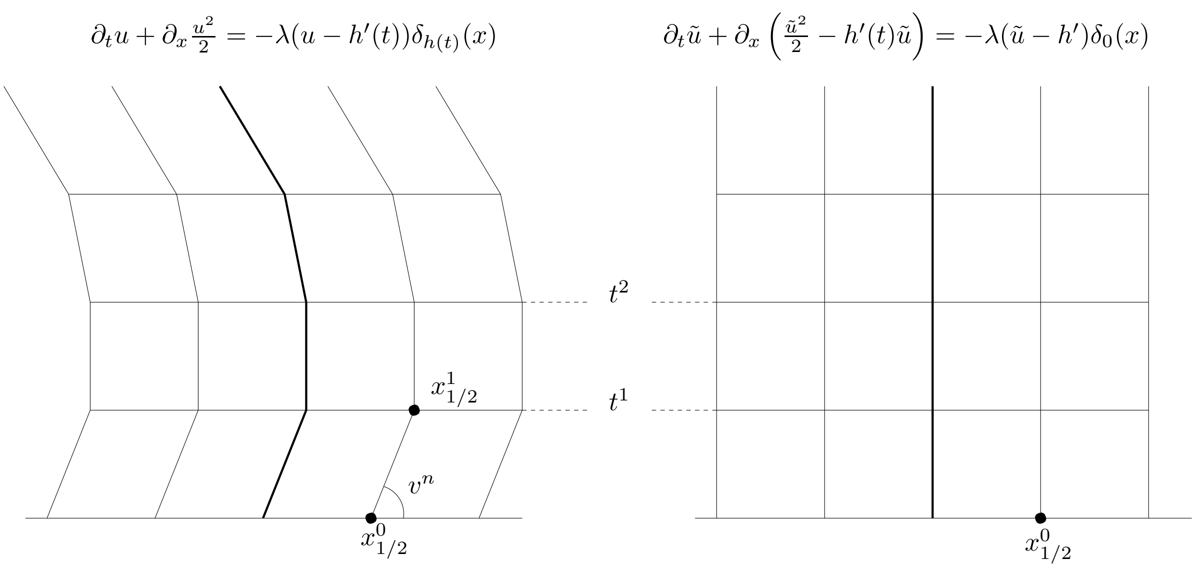


Figure 3.2. *To approximate the solution of* (3.1), *we can either use a mesh that follows the particle (on the left) or straighten the particle's trajectory and approximate the solution of* (3.9). *In both case, the particle's trajectory is the bold line.*

- consistency with the modified Burgers equation:

$$\forall a \in \mathbb{R},\ \forall v \in \mathbb{R},\ g(a, a, v) = \frac{a^2}{2} - va, \tag{3.11}$$

- monotonicity with respect to the first two arguments:

$$\forall (a, b) \in \mathbb{R}^2,\ \forall v \in \mathbb{R},\ \partial_1 g(a, b, v) \geq 0 \quad \text{and} \quad \partial_2 g(a, b, v) \leq 0. \tag{3.12}$$

- g is locally Lipschitz-continuous; (3.13)

they ensure convergence of the scheme to an entropy solution of the Burgers equation away from the particle.

The assumptions on the fluxes around the particle are the following. We first have some consistency assumptions, which ensure that some particular solutions corresponding to a large enough part of the germ are exactly preserved by the numerical scheme. We do not ask the flux to preserve the whole germ though, but only, in Section 3.3 with a *maximal* part of the germ, and in Section 3.4, with $\mathcal{G}^1_\lambda$. More precisely, the hypothesis on the fluxes $g^\pm_\lambda$ are:

- consistency the part $\mathcal{G}^1_\lambda$ of the germ:

$$\forall v \in \mathbb{R},\ \forall (a, b) \in \mathcal{G}^1_\lambda(v),\ g^-_\lambda(a, b, v) = \frac{a^2}{2} - va \quad \text{and} \quad g^+_\lambda(a, b, v) = \frac{b^2}{2} - vb. \tag{3.14}$$

  In Section 3.4, we make the stronger assumption that g is consistent with a *maximal* subset $\mathcal{H}_\lambda$ of $\mathcal{G}_\lambda$ (see Definition 3.3)

$$\forall v \in \mathbb{R},\ \forall (a, b) \in \mathcal{H}_\lambda(v),\ g^-_\lambda(a, b, v) = \frac{a^2}{2} - va \quad \text{and} \quad g^+_\lambda(a, b, v) = \frac{b^2}{2} - vb. \tag{3.15}$$

Hypothesis (3.14) will be used to prove TVD estimates on the fluid part $(u^n_j)_{j\in\mathbb{Z}, n\in\mathbb{N}}$. We also assume that

- if the particle has the same velocity than the fluid, its velocity does not change:

$$\forall v \in \mathbb{R},\ g_\lambda^-(v,v,v) = g_\lambda^+(v,v,v). \tag{3.16}$$

This hypothesis be used to prove a L^∞ bound on the particle velocity $(v^n)_{n\in\mathbb{N}}$. We add two classical conditions of regularity and monotonicity, also used to prove the TVD bound on $(u_j^n)_{j\in\mathbb{Z},n\in\mathbb{N}}$. We assume that:

- both g_λ^- and g_λ^+ are locally Lipschitz-continuous; (3.17)
- g_λ^- and g_λ^+ are nondecreasing with respect to their first arguments, and nonincreasing with respect to their second arguments. (3.18)

Just like in [AS12], we need a dissipativity property to prove discrete entropy inequalities. Moreover, it will also be a key assumption to prove the bounds on the particle's velocity.

- The function $g_\lambda^- - g_\lambda^+$ is nondecreasing with respect to its first two arguments. (3.19)

For this family of finite volume schemes, we are able to prove the following convergence theorem.

Theorem 3.4. *Consider a finite volume scheme of the form* (3.6) *that satisfies the set of hypothesis (3.11–3.14) and (3.16–3.19), and* (3.15) *in Section 3.4. Suppose that u^0 belongs to $BV(\mathbb{R}) \cap L^1(\mathbb{R})$. Let us denote by L the largest Lipschitz constant of g, g^+ and g^- on the set $[m,M]^2 \times [\underline{v}, \bar{v}]$, where*

$$\begin{cases} m = \min\{\operatorname{essinf}_{\mathbb{R}^-} u^0 - \lambda, \operatorname{essinf}_{\mathbb{R}^+} u^0\}, \\ M = \max\{\operatorname{esssup}_{\mathbb{R}^-} u^0, \operatorname{esssup}_{\mathbb{R}^+} u^0 + \lambda\}, \\ \underline{v} = \min(m, v^0), \\ \bar{v} = \max(M, v^0). \end{cases}$$

Then, under the Courant-Friedrichs-Lewy condition

$$L\mu \leq \frac{1}{2}, \tag{3.20}$$

the sequence $(u_{\Delta t})$ converges in $L^1_{loc}(\mathbb{R}_+ \times \mathbb{R})$ toward u and the sequence $(h_{\Delta t})$ converges in $W^{1,\infty}_{loc}(\mathbb{R}_+)$ toward h when Δt tends to 0, where (h,u) is the solution of (3.1).

The next three Sections are devoted to the proof. In Section 3.2, we prove bounds on the total variation of the fluid and on the acceleration of the particle, which permit us to extract converging subsequences. Then in Section 3.3, we prove Theorem 3.4 under Hypothesis (3.15), which is sufficient to obtain a discrete version of (3.4). In Section 3.4, we drop hypothesis (3.15) and prove the convergence of the family of schemes such that

$$\begin{cases} g_\lambda^-(a,b,v) &= g(a, b+\lambda, v), \\ g_\lambda^+(a,b,v) &= g(a-\lambda, b, v), \end{cases}$$

where g satisfies assuptions (3.11–3.13). This type of schemes was introduced in [AS12]. They only preserve the part $\mathcal{G}_\lambda^1(v)$ of the germ, in the sense that if (a,b) belongs to $\mathcal{G}_\lambda^1(v)$, then

$$g_\lambda^-(a,b,v) = f^v(a) \quad \text{and} \quad g_\lambda^-(a,b,v) = f^v(b).$$

We recall that $\mathcal{G}_\lambda^1(v)$ is not a maximal subset of $\mathcal{G}_\lambda(v)$. Under the set of assumptions specified above (except (3.15)) we extend the proof of convergence of [AS12] to the fully coupled case (3.1).

3.2 A priori bounds

In the sequel we suppose that u^0 belongs to $L^1(\mathbb{R}) \cap BV(\mathbb{R})$, that Hypothesis (3.11), (3.12) and (3.13) on the flux g are fulfilled, and that the monotonicity and regularity assumptions (3.18) and (3.17) on $g^\pm$ are verified. We will specify the consistency hypothesis on $g^\pm$ along the way. We first consider the uncoupled problem where $(v^n)_{n\in\mathbb{N}}$ is fixed.

Proposition 3.5. *Let u^0 be in $BV(\mathbb{R}) \cap L^1(\mathbb{R})$. Let $(v^n)_{n\in\mathbb{N}}$ be given and $\underline{v}$ and $\bar{v}$ in $\mathbb{R}$ such that*

$$\forall n \in \mathbb{N},\ \underline{v} \leq v^n \leq \bar{v}.$$

Consider the finite volume scheme

$$\begin{cases} u_j^{n+1} &= u_j^n - \mu(g(u_j^n, u_{j+1}^n, v^n) - g(u_{j-1}^n, u_j^n, v^n)) \text{ for } j \in \mathbb{Z} \setminus \{0,1\}, \\ u_0^{n+1} &= u_0^n - \mu(g_\lambda^-(u_0^n, u_1^n, v^n) - g(u_{-1}^n, u_0^n, v^n)), \\ u_1^{n+1} &= u_1^n - \mu(g(u_1^n, u_2^n, v^n) - g_\lambda^+(u_0^n, u_1^n, v^n)). \end{cases}$$

Suppose that the fluxes $g^\pm$ verify (3.14) and that the CFL condition (3.20) holds. Then we have the following L^∞ and BV estimates in space on $u_{\Delta t}$:

$$\forall n \geq 0, \forall j \in \mathbb{Z},\ m \leq u_j^{n+1} \leq M. \tag{3.21}$$

and

$$\forall n \in \mathbb{N},\ \sum_{j\in\mathbb{Z}} |u_j^n - u_{j-1}^n| \leq \sum_{j\in\mathbb{Z}} |u_j^0 - u_{j-1}^0| + 2\lambda. \tag{3.22}$$

Proof. Due to the presence of the particle, the maximum and the total variation of the exact solution u of (3.1) can increase through time. For example if u^0 is constant equals to 0 and if $v^0 > \lambda$, then $||u(0^+,\cdot)||_{L^\infty(\mathbb{R})} = ||u^0||_{L^\infty(\mathbb{R})} + \lambda$ and $||u(0^+,\cdot)||_{BV(\mathbb{R})} = ||u^0||_{BV(\mathbb{R})} + 2\lambda$ (see [LST08], Lemma 5.7). This prevents us for applying the Le Roux and Harten lemma (see [Har84] and [LeR77]) directly to $(u_j^n)_{j\in\mathbb{Z},\, n\in\mathbb{N}}$. Yet it can be applied to the sequence $(w_j^n)_{j\in\mathbb{Z},\, n\in\mathbb{N}}$ defined by

$$w_j^n = \begin{cases} u_j^n - \frac{\lambda}{2} & \text{if } j \leq 0, \\ u_j^n + \frac{\lambda}{2} & \text{if } j \geq 1. \end{cases}$$

Let us prove that there exists two families of real $(C_{j+1/2}^n)_{j\in\mathbb{Z},n\in\mathbb{N}}$ and $(D_{j+1/2}^n)_{j\in\mathbb{Z},n\in\mathbb{N}}$ such that for all j in $\mathbb{Z}$, for all n in $\mathbb{N}$,

$$w_j^{n+1} = w_j^n + C_{j+1/2}^n(w_{j+1}^n - w_j^n) - D_{j-1/2}^n(w_j^n - w_{j-1}^n), \tag{3.23}$$

and

$$0 \leq 1 - C_{j+1/2}^n - D_{j+1/2}^n \leq 1,\ 0 \leq C_{j+1/2}^n \leq 1 \text{ and } 0 \leq D_{j+1/2}^n \leq 1.$$

In other words, w_j^{n+1} writes as a convex combination of w_{j-1}^n, w_j^n and w_{j+1}^n and therefore,

$$\forall n \geq 0,\ \min_k w_k^n \leq w_j^{n+1} \leq \max_k w_k^n.$$

As a consequence, for all $n \in \mathbb{N}$ and for $j \leq 0$,

$$\min_k w_k^0 + \lambda/2 \leq u_j^n \leq \max_k w_k^0 + \lambda/2$$

which rewrites

$$\min\{\cdots, u_0^0, u_1^0 + \lambda, \cdots\} \leq u_j^n \leq \max\{\cdots, u_0^0, u_1^0 + \lambda, \cdots\}.$$

Similarly, for all $n \in \mathbb{N}$ and for all $j \geq 1$,

$$\min\{\cdots, u_0^0 - \lambda, u_1^0, \cdots\} \leq u_j^n \leq \max\{\cdots, u_0^0 - \lambda, u_1^0, \cdots\},$$

hence the L^∞ bound (3.21) is proven. Moreover, the Leroux and Harten lemma yields

$$\forall n \in \mathbb{N}, \sum_{j\in\mathbb{Z}} |w_j^{n+1} - w_{j-1}^{n+1}| \leq \sum_{j\in\mathbb{Z}} |w_j^n - w_{j-1}^n|,$$

and thus (3.22).

Let us go back to the existence of $C_{j+1/2}^n$ and $D_{j-1/2}^n$. In the sequel we denote by $|a,b|$ the interval $[\min(a,b), \max(a,b)]$. Suppose first that (3.23) holds for some $n \in \mathbb{N}$. Then for every $j \leq -1$, there exists $\tilde{w}_{j-1/2}^n \in |w_{j-1}^n, w_j^n|$ and $\bar{w}_{j+1/2}^n \in |w_j^n, w_{j+1}^n|$

$$\begin{aligned}
w_j^{n+1} &= w_j^n - \mu\left(g(u_j^n, u_{j+1}^n, v^n) - g(u_{j-1}^n, u_j^n, v^n)\right) \\
&= w_j^n - \mu\left(g_\lambda\left(w_j^n + \frac{\lambda}{2}, w_{j+1}^n + \frac{\lambda}{2}, v^n\right) - g_\lambda\left(w_{j-1}^n + \frac{\lambda}{2}, w_j^n + \frac{\lambda}{2}, v^n\right)\right) \\
&= w_j^n - \mu\left(\partial_1 g_\lambda\left(\tilde{w}_{j-1/2}^n + \frac{\lambda}{2}, w_{j+1}^n + \frac{\lambda}{2}, v^n\right)(w_j^n - w_{j-1}^n)\right. \\
&\qquad \left. + \partial_2 g_\lambda\left(w_{j-1}^n + \frac{\lambda}{2}, \bar{w}_{j+1/2}^n + \frac{\lambda}{2}, v^n\right)(w_{j+1} - w_j)\right)
\end{aligned}$$

Both triplets $\left(\tilde{w}_{j-1/2}^n + \frac{\lambda}{2}, w_{j+1}^n + \frac{\lambda}{2}, v^n\right)$ and $\left(w_{j-1}^n + \frac{\lambda}{2}, \bar{w}_{j+1/2}^n + \frac{\lambda}{2}, v^n\right)$ belong to $[m, M]^2 \times [\underline{v}, \bar{v}]$. The CFL condition (3.20), and the fact that $\partial_1 g \geq 0$ and $\partial_2 g \leq 0$, yield the result with

$$\begin{cases}
D_{j-1/2}^n &= \mu \partial_1 g_\lambda \left(\tilde{w}_{j-1/2}^n + \frac{\lambda}{2}, w_{j+1}^n + \frac{\lambda}{2}, v^n\right), \\
C_{j+1/2}^n &= -\mu \partial_2 g_\lambda \left(w_{j-1}^n + \frac{\lambda}{2}, \bar{w}_{j+1/2}^n + \frac{\lambda}{2}, v^n\right).
\end{cases}$$

The case $j \geq 2$ can be treated in the exact same way. We now turn to the trickier case $j = 0$. The facts that g_λ^- is consistent with $\mathcal{G}_\lambda^1$ and that g is consistent (Hypothesis (3.14) and (3.11)) imply that

$$g_\lambda^-\left(w_0^n + \frac{\lambda}{2}, w_0^n - \frac{\lambda}{2}, v^n\right) = g_\lambda\left(w_0^n + \frac{\lambda}{2}, w_0^n + \frac{\lambda}{2}, v^n\right),$$

allow us to write

$$\begin{aligned}
w_0^{n+1} &= w_0^n - \mu\left(g_\lambda^-(u_0^n, u_1^n, v^n) - g(u_{-1}^n, u_0^n, v^n)\right) \\
&= w_0^n - \mu\left(g_\lambda^-\left(w_0^n + \frac{\lambda}{2}, w_1^n - \frac{\lambda}{2}, v^n\right) - g_\lambda\left(w_{-1}^n + \frac{\lambda}{2}, w_0^n + \frac{\lambda}{2}, v^n\right)\right) \\
&= w_0^n - \mu\left(g_\lambda^-\left(w_0^n + \frac{\lambda}{2}, w_1^n - \frac{\lambda}{2}, v^n\right) - g_\lambda^-\left(w_0^n + \frac{\lambda}{2}, w_0^n - \frac{\lambda}{2}, v^n\right)\right. \\
&\qquad \left. + g_\lambda\left(w_0^n + \frac{\lambda}{2}, w_0^n + \frac{\lambda}{2}, v^n\right) - g_\lambda\left(w_{-1}^n + \frac{\lambda}{2}, w_0^n + \frac{\lambda}{2}, v^n\right)\right)
\end{aligned}$$

Thus, there exists $\tilde{w}_{-1/2}^n \in |w_{-1}^n, w_0^n|$ and $\bar{w}_{1/2}^n \in |w_0^n, w_1^n|$ such that

$$\begin{aligned}
w_0^{n+1} = w_0^n - \mu &\left(\partial_2 g_\lambda^-\left(w_0^n + \frac{\lambda}{2}, \bar{w}_{1/2}^n - \frac{\lambda}{2}, v^n\right)(w_1^n - w_0^n)\right. \\
&\left. + \partial_1 g_\lambda\left(w_0^n + \frac{\lambda}{2}, \tilde{w}_{-1/2}^n + \frac{\lambda}{2}, v^n\right)(w_0^n - w_{-1}^n)\right)
\end{aligned}$$

Once again, both triplets $\left(w_0^n + \frac{\lambda}{2}, \bar{w}_{1/2}^n - \frac{\lambda}{2}, v^n\right)$ and $\left(w_0^n + \frac{\lambda}{2}, \tilde{w}_{-1/2}^n + \frac{\lambda}{2}, v^n\right)$ belong to $[m, M]^2 \times [\underline{v}, \bar{v}]$. The monotonicity on g and g_λ^- allow to conclude with

$$\begin{cases}
D_{-1/2}^n &= \mu \partial_1 g_\lambda \left(w_0^n + \frac{\lambda}{2}, \tilde{w}_{-1/2}^n + \frac{\lambda}{2}, v^n\right), \\
C_{1/2}^n &= -\mu \partial_2 g_\lambda^- \left(w_0^n + \frac{\lambda}{2}, \bar{w}_{1/2}^n - \frac{\lambda}{2}, v^n\right).
\end{cases}$$

The case $j = 1$ can be treated in the exact same way, using the consistency assumption

$$g_\lambda^+\left(w_1^n + \frac{\lambda}{2}, w_1^n - \frac{\lambda}{2}, v^n\right) = g_\lambda\left(w_1^n - \frac{\lambda}{2}, w_1^n - \frac{\lambda}{2}, v^n\right). \qquad \square$$

We now turn to the case where the particle's velocity is updated from time to time, and focus on the estimates on the velocity and acceleration of the particle.

Proposition 3.6. *Let us denote by*

$$\underline{v} = \min(m, v^0) \quad \textit{and} \quad \bar{v} = \max(M, v^0),$$

where m and M are defined as in Theorem 3.4. Suppose that the fluxes $g^\pm$ verify (3.16), (3.19) *and* (3.20) *that the time step verifies*

$$\frac{4L}{m_p}\Delta t \le 1 \tag{3.24}$$

Then, the sequence $(u_j^n)_{j\in\mathbb{Z},n\in\mathbb{N}}$ (defined by (3.6)*) verifies Estimates* (3.21) *and* (3.22)*, while $(v^n)_{n\in\mathbb{N}}$ verifies the following estimates:*

$$\forall n \in \mathbb{N}, \quad \underline{v} \le v^n \le \bar{v}. \tag{3.25}$$

and

$$\forall n \in \mathbb{N}, \quad \left|\frac{v^{n+1} - v^n}{\Delta t}\right| \le \frac{2L}{m_p}(||u^0||_\infty + \lambda + ||v||_\infty). \tag{3.26}$$

Proof. We proceed by induction. Let us first remark that if the estimate (3.25) on v^n is fulfiled at time t^n, the proof of Proposition 3.5 yields the L^∞ and TVD estimates on $(u_j^{n+1})_{j\in\mathbb{Z}}$. Therefore, we focus on the estimate on v^{n+1}. Using Hypothesis (3.16), we introduce the null quantity $g_\lambda^-(v^n, v^n, v^n) - g_\lambda^+(v^n, v^n, v^n)$ and write

$$\begin{aligned} v^{n+1} &= v^n + \frac{\Delta t}{m_p}(g_\lambda^-(u_0^n, u_1^n, v^n) - g_\lambda^+(u_0^n, u_1^n, v^n)) \\ &= v^n + \frac{\Delta t}{m_p}\left(\int_0^1 \partial_s(g_\lambda^-(v^n + s(u_0^n - v^n), v^n + s(u_1^n - v^n), v^n))ds \right.\\ &\qquad \left. - \int_0^1 \partial_s(g_\lambda^+(v^n + s(u_0^n - v^n), v^n + s(u_1^n - v^n), v^n))ds\right), \end{aligned}$$

and we obtain

$$\begin{aligned} v^{n+1} = v^n + \frac{\Delta t}{m_p}&\left(\int_0^1 (u_0^n - v^n)\partial_1(g^- - g^+)(v^n + s(u_0^n - v^n), v^n + s(u_1^n - v^n), v^n)ds \right.\\ &\left. + \int_0^1 (u_1^n - v^n)\partial_2(g^- - g^+)(v^n + s(u_0^n - v^n), v^n + s(u_1^n - v^n), v^n)ds\right) \end{aligned} \tag{3.27}$$

Suppose now that $v^n \le \min(u_0^n, u_1^n)$. Then both $(u_1^n - v^n)$ and $(u_0^n - v^n)$ are nonnegative. Moreover, the dissipativity assumption (3.19) implies that $\partial_1(g^- - g^+)$ and $\partial_2(g^- - g^+)$ are also nonnegative. Hence we have $v^{n+1} \ge v^n$ and Hypothesis (3.24) yields

$$\begin{aligned} v^{n+1} &\le v^n + 2L\frac{\Delta t}{m_p}(u_0^n - v^n + u_1^n - v^n) \\ &\le \left(1 - \frac{4L\Delta t}{m_p}\right)v^n + \frac{4L\Delta t}{m_p}\max(u_0^n, u_1^n) \le \bar{v}. \end{aligned}$$

We now treat the case $u_0^n \le v^n \le u_1^n$. The only difference is that $u_0^n - v^n$ is now negative. The integral form (3.27) of v^{n+1} and Hypothesis (3.25) yield

$$m \le v^n - 2L\frac{\Delta t}{m_p}(v^n - u_0^n) \le v^{n+1} \le v^n + 2L\frac{\Delta t}{m_p}(u_1^n - v^n) \le M.$$

Once the L^∞ bounds on $(u_j^n)_{j\in\mathbb{Z},n\in\mathbb{N}}$ and $(v^n)_{n\in\mathbb{N}}$ are proven, the bound of the particle's acceleration (3.26) is an easy consequence of the integral form of v^{n+1}. □

We are now in position to extract converging subsequences of (u_Δ) and (h_Δ) (defined in (3.7) and (3.8)). In Section 3.3, we will prove that their limits are solutions of the Cauchy problem (3.1) for the fully coupled problem.

Proposition 3.7. *Assume that u^0 belongs to $BV(\mathbb{R})\cap L^1(\mathbb{R})$, and that Hypothesis (3.11-3.14) and (3.16-3.19) are verified. Moreover, suppose that the CFL condition* (3.20) *holds. Then there exists u in $BV_{loc}(\mathbb{R}_+\times\mathbb{R})$ and h in $W^{2,\infty}_{loc}(\mathbb{R}_+)$ such that, up to a subsequence, the sequence $(u_{\Delta t})$ converges in $L^1_{loc}(\mathbb{R}_+\times\mathbb{R})$ toward u and the sequence $(h_{\Delta t})$ converges in $W^{1,\infty}_{loc}(\mathbb{R})$ toward h as Δt tends to* 0.

Proof. Let us first fix a time $T>0$ and a constant $A>0$ and prove the convergence in $L^1([0,T]\times[-A,A])$ and $W^{1,\infty}([0,T])$. By Proposition 3.5, we can use Helly's theorem to prove the convergence in $L^1([0,T]\times[-A,A])$ of $(u_{\Delta t})$, toward a function u in $BV(([0,T]\times[-A,A]))$. Similarly Proposition 3.6 allows us to apply Arzelà-Ascoli's theorem to prove convergence in $W^{1,\infty}([0,T])$ of $(h_{\Delta t})$ to a function h belonging to $W^{2,\infty}([0,T])$. The result is extended to the whole time-space $\mathbb{R}_+\times\mathbb{R}$ thanks to the Cantor diagonal extraction argument. □

Remark 2. In particular, up to the same subsequence, $(v_{\Delta t})$ converges toward h' in L^1_{loc}. Moreover the sequence of functions $(c_{\Delta t})$ defined by

$$c_{\Delta t}(t,x)=\begin{cases} c_- & \text{if } t<h_{\Delta x}(t),\\ c_+ & \text{if } t>h_{\Delta x}(t),\end{cases}$$

converges in L^1_{loc} toward

$$c(t,x)=\begin{cases} c_- & \text{if } t<h(t),\\ c_+ & \text{if } t>h(t).\end{cases}$$

Indeed, we have

$$\int_{-A}^{A}\int_0^T |c_{\Delta t}(t,x)-c(t,x)|dtdx \le |c_+-c_-|\int_0^T |h_{\Delta t}(t)-h(t)|dt \le 2LT\Delta t.$$

3.3 Convergence of schemes consistent with a maximal part of the germ

For now on, we suppose that all the hypotheses of Proposition 3.7 are fulfilled, and that both Conditions (3.20) and (3.24) are verified. The aim of this section is to prove Theorem 3.4. To that purpose, we prove that under Condition (3.15), which states that the fluxes $g^{\pm}_{\lambda}$ around the particle are consistent with a maximal subset $\mathcal{H}_\lambda$ of the germ (see Definition 3.3), the limit (u,h) of the scheme is the solution of (3.1).

The fact that the Cauchy problem (3.1) is well posed in $BV(\mathbb{R})$ is proven in [ALST13]. Once we know that Scheme (3.6) converges toward a solution of (3.1), the uniqueness of the solution yields that the whole sequence $(u_{\Delta t},h_{\Delta t})$ converges. Theorem 3.4 gives a different way to prove the existence of a solution (but not the uniqueness).

3.3.1 Convergence of the fluid's part

The aim of this subsection is to prove that the limit u of $(u_{\Delta t})$ verifies (3.4). We prove in Proposition 3.9 that $(u_j^n)_{j\in\mathbb{Z},n\in\mathbb{N}}$ verifies a discrete version of (3.4). In the sequel, for all reals number a and b we denote by

$$a\top b=\max(a,b) \text{ and by } a\bot b=\min(a,b).$$

In the following proposition, we establish a discrete entropy inequality.

Proposition 3.8. *Assume that Hypothesis (3.11-3.19) hold (included* (3.15)*) and that the CFL condition* (3.20) *is fulfilled. Then for all (c_-, c_+) in $\mathbb{R}^2$, there exists a constant A, depending only on λ, $||u^0||_\infty$, $||v||_\infty$ and (c_-, c_+), such that for all $j \in \mathbb{Z}$, for all $n \in \mathbb{N}$, the following inequality holds:*

$$\frac{|u_j^{n+1} - c_j| - |u_j^n - c_j|}{\Delta t} + \frac{G_{j+1/2,-}^n - G_{j-1/2,+}^n}{\Delta x} \leq \varepsilon_j \frac{A}{\Delta x} \operatorname{dist}_1((c_-, c_+), \mathcal{H}_\lambda(v^n)), \tag{3.28}$$

where

$$\forall j \neq 0,\ G_{j+1/2,-}^n = G_{j+1/2,+}^n = G_{j+1/2}^n,$$

with

$$G_{j+1/2}^n = g(u_j^n \top c_j, u_{j+1}^n \top c_{j+1}, v^n) - g(u_j^n \bot c_j, u_{j+1}^n \bot c_{j+1}, v^n),$$

and

$$G_{1/2,\pm}^n = g_\lambda^\pm(u_0^n \top c_0, u_1^n \top c_1, v^n) - g_\lambda^\pm(u_0^n \bot c_0, u_1^n \bot c_1, v^n),$$

and

$$c_j = \begin{cases} c_- & \text{if } j \leq 0, \\ c_+ & \text{if } j \geq 1, \end{cases} \quad \text{and} \quad \varepsilon_j = \begin{cases} 1 & \text{if } j \in \{0,1\}, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. We follow the guidelines of proofs of classical entropy inequalities. They rely on the identity

$$|u_j^{n+1} - c_j| = u_j^{n+1} \top c_j - u_j^{n+1} \bot c_j.$$

For $j \in \mathbb{Z} \setminus \{0,1\}$, we use the condensed notation $u_j^{n+1} = H(u_{j-1}^n, u_j^n, u_{j+1}^n, v^n)$. Hypothesis (3.12) on the monotonicity of the fluxes and the CFL condition (3.20) ensure that for every v, H is increasing with respect to its first three arguments. Moreover if $j \in \mathbb{Z} \setminus \{0,1\}$, $c_{j-1} = c_j = c_{j+1}$ and we use the consistency of the flux away from the particle (3.11) to write $c_j = H(c_{j-1}, c_j, c_{j+1}, v^n)$. It follows that

$$\begin{aligned} u_j^{n+1} \top c_j &= H(u_{j-1}^n, u_j^n, u_{j+1}^n, v^n) \top H(c_{j-1}, c_j, c_{j+1}, v^n) \\ &\leq H(u_{j-1}^n \top c_{j-1}, u_j^n \top c_j, u_{j+1}^n \top c_{j+1}, v^n) \\ u_j^{n+1} \bot c_j &= H(u_{j-1}^n, u_j^n, u_{j+1}^n, v^n) \bot H(c_{j-1}, c_j, c_{j+1}, v^n) \\ &\geq H(u_{j-1}^n \bot c_{j-1}, u_j^n \bot c_j, u_{j+1}^n \bot c_{j+1}, v^n) \end{aligned}$$

and that

$$\begin{aligned} |u_j^{n+1} - c_j| &\leq H(u_{j-1}^n \top c_{j-1}, u_j^n \top c_j, u_{j+1}^n \top c_{j+1}, v^n) - H(u_{j-1}^n \bot c_{j-1}, u_j^n \bot c_j, u_{j+1}^n \bot c_{j+1}, v^n) \\ &\leq u_j^n \top c_j - u_j^n \bot c_j - \mu(G_{j+1/2}^n - G_{j-1/2}^n) \\ &\leq |u_j^n - c_j| - \mu(G_{j+1/2}^n - G_{j-1/2}^n) \end{aligned}$$

Let us now focus on the more complicated case $j = 0$ (the case $j = 1$ can be treated in the exact same way). We denote by $(\tilde{c}_0^n, \tilde{c}_1^n)$ a projection of $(c_-, c_+) = (c_0^n, c_1^n)$ on $\mathcal{H}_\lambda(v^n)$ for the L^1-norm, and by $(\tilde{c}_j^n)_{j \in \mathbb{Z}, n \in \mathbb{N}}$ and $(\tilde{G}_{j+1/2}^n)_{j \in \mathbb{Z}, n \in \mathbb{N}}$ the analogues of $(c_j)_{j \in \mathbb{Z}}$ and $(G_{j+1/2}^n)_{j \in \mathbb{Z}, n \in \mathbb{N}}$ constructed with $\tilde{c}$:

$$\forall j \neq 0,\ \tilde{G}_{j+1/2,-}^n = \tilde{G}_{j+1/2,+}^n = \tilde{G}_{j+1/2}^n = g(u_j^n \top \tilde{c}_j, u_{j+1}^n \top \tilde{c}_{j+1}, v^n) - g(u_j^n \bot \tilde{c}_j, u_{j+1}^n \bot \tilde{c}_{j+1}, v^n),$$

$$\tilde{G}_{1/2,\pm}^n = g_\lambda^\pm(u_0^n \top \tilde{c}_0, u_1^n \top \tilde{c}_1, v^n) - g_\lambda^\pm(u_0^n \bot \tilde{c}_0, u_1^n \bot \tilde{c}_1, v^n).$$

Let us first remark that

$$\begin{aligned} |u_0^{n+1} - c_0| - |u_0^n - c_0| &\leq |u_0^{n+1} - \tilde{c}_0^n| + |\tilde{c}_0^n - c_0| - \big| |u_0^n - \tilde{c}_-^n| - |\tilde{c}_0^n - c_0| \big| \\ &\leq |u_0^{n+1} - \tilde{c}_0^n| - |u_0^n - \tilde{c}_-^n| + 2|\tilde{c}_0^n - c_0|. \end{aligned}$$

Thus we have

$$\begin{aligned}&\frac{|u_0^{n+1}-c_0|-|u_0^n-c_0|}{\Delta t}+\frac{G^n_{1/2,-}-G^n_{-1/2}}{\Delta x}\\&\quad\le\frac{|u_0^{n+1}-\tilde c_0^n|-|u_0^n-\tilde c_0^n|}{\Delta t}+\frac{G^n_{1/2,-}-G^n_{-1/2}}{\Delta x}+\frac{2}{\Delta t}\operatorname{dist}_1((c_-,c_+),\mathcal{H}_\lambda(v^n))\\&\quad\le\frac{G^n_{1/2,-}-G^n_{-1/2}}{\Delta x}-\frac{\tilde G^n_{1/2,-}-\tilde G^n_{-1/2}}{\Delta x}+\frac{2}{\Delta t}\operatorname{dist}_1((c_-,c_+),\mathcal{H}_\lambda(v^n)).\end{aligned}$$

Indeed, as $(\tilde c_0^n,\tilde c_1^n)$ belongs to $\mathcal{H}_\lambda(v^n)$, Hypothesis (3.15) yields that $\tilde c_0^n = H_\lambda(\tilde c_{-1}^n,\tilde c_0^n,\tilde c_1^n,v^n)$, and we obtain as before

$$|u_0^{n+1}-\tilde c_j|\le|u_j^n-\tilde c_j|-\mu(\tilde G^n_{j+1/2}-\tilde G^n_{j-1/2}).$$

We now attempt to bound

$$\begin{aligned}G^n_{1/2,-}-\tilde G^n_{1/2,-}&=g_\lambda^-(u_0^n\top c_0,u_1^n\top c_1,v^n)-g_\lambda^-(u_0^n\bot c_0,u_1^n\bot c_1,v^n)\\&\quad-g_\lambda^-(u_0^n\top\tilde c_0^n,u_1^n\top\tilde c_1^n,v^n)+g_\lambda^-(u_0^n\bot\tilde c_0^n,u_1^n\bot\tilde c_1^n,v^n).\end{aligned}$$

As $(v^n)_{n\in\mathbb{Z}}$ is bounded (Proposition 3.6), the maximum and minimum over n of $\tilde c_\pm^n$ is a bounded function of (c_-,c_+) and $||v||_\infty$. Thus the set

$$[\min(m,c_-,c_+,\tilde c_-^n,\tilde c_+^n),\max(M,c_-,c_+,\tilde c_-^n,\tilde c_+^n)]^2\times[\underline{v},\bar v].$$

is compact. Therefore, with L_c the Lipschitz constant of g_λ^- over this set, we have

$$\begin{aligned}&|g_\lambda^-(u_0^n\top c_0,u_1^n\top c_1,v^n)-g_\lambda^-(u_0^n\top\tilde c_0^n,u_1^n\top\tilde c_1^n,v^n)|\\&\quad\le|g_\lambda^-(u_0^n\top c_0,u_1^n\top c_1,v^n)-g_\lambda^-(u_0^n\top\tilde c_0^n,u_1^n\top c_1,v^n)|\\&\qquad+|g_\lambda^-(u_0^n\top\tilde c_0^n,u_1^n\top c_1,v^n)-g_\lambda^-(u_0^n\top\tilde c_0^n,u_1^n\top\tilde c_1^n,v^n)|\\&\quad\le L_c\operatorname{dist}_1((c_-,c_+),\mathcal{H}_\lambda(v^n)),\end{aligned}$$

and similarly

$$|g_\lambda^-(u_0^n\bot c_0,u_1^n\bot c_1,v^n)-g_\lambda^-(u_0^n\top\tilde c_0^n,u_1^n\top\tilde c_1^n,v^n)|\le L_c\operatorname{dist}_1((c_-,c_+),\mathcal{H}_\lambda(v^n)),$$

which concludes the proof with $A=2L_c+\frac{2\Delta x}{\Delta t}$. □

We are now in position to obtain a discrete version of (3.4).

Proposition 3.9. *Let $(\varphi_j^n)_{j\in\mathbb{Z},n\in\mathbb{N}}$ be a sequence of nonnegative reals. If* (3.28) *holds for all n in $\mathbb{N}$ and j in $\mathbb{Z}$, then*

$$\begin{aligned}&\Delta t\Delta x\sum_{j\in\mathbb{Z},n\in\mathbb{N}}|u_j^{n+1}-c_j|\frac{\varphi_j^{n+1}-\varphi_j^n}{\Delta t}+\Delta x\sum_{i\in\mathbb{Z}}|u_j^0-c_j|\varphi_j^0+\Delta t\Delta x\sum_{j\in\mathbb{Z}^*,n\in\mathbb{N}}G^n_{j+1/2}\frac{\varphi^n_{j+1}-\varphi^n_j}{\Delta x}\\&\quad+\Delta t\Delta x\sum_{n\in\mathbb{N}}G^n_{j+1/2,+}\frac{\varphi_1^n-\varphi_0^n}{\Delta x}\ge-A\Delta t\sum_{n\in\mathbb{N}}\operatorname{dist}_1(c,\mathcal{H}_\lambda(v^n))(\varphi_0^n+\varphi_1^n).\end{aligned}\tag{3.29}$$

Proof. Classically, the starting point is to multiply Equation (3.28) by φ_j^n and to sum over $j\in\mathbb{Z}$ and $n\in\mathbb{N}$. Then the different terms are rearranged to bring out discrete time and space derivatives of φ. However, this is not straightforward around the particle, because two different fluxes are used on its left and on its right. The first term of (3.28) yields

$$\sum_{j\in\mathbb{Z},n\in\mathbb{N}}\frac{|u_j^{n+1}-c_j|-|u_j^n-c_j|}{\Delta t}\varphi_j^n=\sum_{j\in\mathbb{Z},n\in\mathbb{N}}|u_j^{n+1}-c_j|\frac{\varphi_j^n-\varphi_j^{n+1}}{\Delta t}-\frac{1}{\Delta t}\sum_{j\in\mathbb{Z}}|u_j^0-c_j|\varphi_j^0,$$

and the second term yields

$$\sum_{j\in\mathbb{Z},n\in\mathbb{N}} \frac{G^n_{j+1/2,-} - G^n_{j-1/2,+}}{\Delta x}\varphi^n_j = \sum_{j\in\mathbb{Z}^*,n\in\mathbb{N}} G^n_{j+1/2}\frac{\varphi^n_j - \varphi^n_{j+1}}{\Delta x} + \sum_{n\in\mathbb{N}} \frac{\varphi^n_0}{\Delta x}G^n_{1/2,-} - \frac{\varphi^n_1}{\Delta x}G^n_{1/2,+}$$
$$= \sum_{j\in\mathbb{Z}^*,n\in\mathbb{N}} G^n_{j+1/2}\frac{\varphi^n_j - \varphi^n_{j+1}}{\Delta x} + \sum_{n\in\mathbb{N}} \frac{\varphi^n_0}{\Delta x}(G^n_{1/2,-} - G^n_{1/2,+})$$
$$+ \sum_{n\in\mathbb{N}} \frac{\varphi^n_0 - \varphi^n_1}{\Delta x}G^n_{1/2,+}.$$

We almost have a discrete version of (3.4). The following lemma ensures that the corrective term $\sum_{n\in\mathbb{N}} \frac{\varphi^n_0}{\Delta x}(G^n_{1/2,-} - G^n_{1/2,+})$ has the correct sign.

Lemma 3.10. *If $g^-_\lambda - g^+_\lambda$ is nondecreasing with respect to its first two arguments then we have the dissipativity property*

$$G^n_{1/2,-} - G^n_{1/2,+} \geq 0.$$

Proof of Lemma 3.10. Let us denote by $a = u^n_0 \top c_0$, $\tilde{a} = u^n_0 \perp c_0$, $b = u^n_1 \top c_1$ and $\tilde{b} = u^n_1 \perp c_1$, such that $a \geq \tilde{a}$ and $b \geq \tilde{b}$. The dissipativity property holds if and only if

$$g^-_\lambda(a, b, v^n) - g^-_\lambda(\tilde{a}, \tilde{b}, v^n)) \geq g^+_\lambda(a, b, v^n) - g^+_\lambda(\tilde{a}, \tilde{b}, v^n),$$

which is a straightforward consequence of the monotonicity of $g^-_\lambda - g^+_\lambda$ with respect to its two first variables. □

Let us go back to the proof of Proposition 3.9. Hypothesis (3.19) exactly says that $g^-_\lambda - g^+_\lambda$ is nondecreasing with respect to its two first arguments. Thus we can apply Lemma 3.10 to obtain

$$\sum_{j\in\mathbb{Z},n\in\mathbb{N}} \frac{G^n_{j+1/2,-} - G^n_{j-1/2,+}}{\Delta x}\varphi^n_j \geq \sum_{j\in\mathbb{Z}^*,n\in\mathbb{N}} G^n_{j+1/2}\frac{\varphi^n_j - \varphi^n_{j+1}}{\Delta x} + \sum_{n\in\mathbb{N}} \frac{\varphi^n_0 - \varphi^n_1}{\Delta x}G^n_{1/2,+}.$$

Eventually, we have

$$\sum_{j\in\mathbb{Z},n\in\mathbb{N}} \varepsilon_j \frac{A}{\Delta x} \operatorname{dist}_1((c_-, c_+), \mathcal{H}_\lambda(v^n))\varphi^n_j = \frac{A}{\Delta x}\sum_{n\in\mathbb{N}} \operatorname{dist}_1((c_-, c_+), \mathcal{H}_\lambda(v^n))(\varphi^n_0 + \varphi^n_1)$$

and (3.29) is obtained by regrouping all the terms and changing their signs, and multiplying by $\Delta t \Delta x$. □

Passing to the limit $\Delta t \to 0$ in Equation (3.29), we obtain the following proposition.

Proposition 3.11. *If u^0 belongs to $BV(\mathbb{R}) \cap L^1(\mathbb{R})$, if the CFL condition* (3.20) *holds and if Hypothesis (3.11-3.19), (*included (3.15)*), are fulfilled, then the limit u of $(u_{\Delta t})$ verifies Inequality* (3.4) *for any nonnegative function φ in $\mathcal{C}^\infty_0(\mathbb{R}_+ \times \mathbb{R})$.*

Proof. For small enough Δt, Condition (3.24) is verified. Let us fix (c_-, c_+) in $\mathbb{R}^2$, and prove that for every nonnegative φ in $\mathcal{C}^\infty_0$, the discrete inequality (3.29) converges to the continuous entropy inequality (3.4), where the sequence $(\varphi^n_j)_{j\in\mathbb{Z},n\in\mathbb{N}}$ is defined by $\varphi^n_j = \varphi(n\Delta t, x^n_j - h^n)$. We recall that $\mathcal{C}^n_j$ is the space-time cell

$$\mathcal{C}^n_j = \{(n\Delta t + s, x^n_{j-1/2} + y + sv^n), s \in [0, \Delta t), y \in [0, \Delta x)\},$$

that h^n is the discrete position of the particle's trajectory deduced from its velocity:

$$h^{n+1} = h^n + v^n \Delta t$$

and that the mesh is moving with the particle: $x_j^{n+1} = x_j^n + v^n \Delta t$. We first treat the first term of (3.29). The sequence of piecewise constant functions $(\zeta_{\Delta t})$ defined by

$$\zeta_{\Delta t}(t,x) = \frac{\varphi_j^{n+1} - \varphi_j^n}{\Delta t} \quad \text{if } (t,x) \in {}_j^{n+1}$$

converges uniformly to the function $(t,x) \mapsto (\partial_t \varphi)(t, x - h(t))$. Indeed, for every $(t,x) \in {}_j^{n+1}$, there exists $\tilde{t} \in [n\Delta t, (n+1)\Delta t]$ such that

$$\begin{aligned}
|\zeta_{\Delta x}(t,x) - (\partial_t\varphi)(t, x-h(t))| &= \left| \frac{\varphi((n+1)\Delta t, x_j^{n+1} - h^{n+1}) - \varphi(n\Delta t, x_j^n - h^n)}{\Delta t} - (\partial_t \varphi)(t, x - h(t)) \right| \\
&= |(\partial_t\varphi)(\tilde{t}, x_j^n - h^n) - (\partial_t\varphi)(t, x - h(t))| \\
&\leq C(|\tilde{t} - t| + |x - x_j^n| + |h^n - h(t)|) \\
&\leq C(\Delta t + \Delta x + ||h_{\Delta t} - h||_\infty)
\end{aligned}$$

We used the fact that $x_j^{n+1} - h^{n+1} = x_j^n - h^n$. We conclude thanks to Remark 2 :

$$\begin{aligned}
\Delta t \Delta x \sum_{j\in\mathbb{Z}, n\in\mathbb{N}} |u_j^{n+1} - c_j| \frac{\varphi_j^{n+1} - \varphi_j^n}{\Delta t} &= \sum_{j\in\mathbb{Z}, n\in\mathbb{N}} \int_{{}_j^{n+1}} |u_{\Delta t} - c_{\Delta t}| \zeta_{\Delta t} dt\, dx \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}_+} \mathbf{1}_{t\geq \Delta t} |u_{\Delta t} - c_{\Delta t}| \zeta_{\Delta t} dt\, dx \\
&\longrightarrow \int_{\mathbb{R}} \int_{\mathbb{R}_+} |u - c| (\partial_t \varphi)(t, x - h(t)) dt\, dx
\end{aligned}$$

On the other hand,

$$\Delta t \Delta x \sum_{j<0, n\in\mathbb{N}} G_{j+1/2}^n \frac{\varphi_{j+1}^n - \varphi_j^n}{\Delta x} = \int_{x < -\frac{\Delta x}{2}} \int_{\mathbb{R}_+} G_{\Delta t} \xi_{\Delta t} dt\, dx$$

where for every (t,x) in ${}_{j+1/2}^n = \{(n\Delta t + s, x_j + y + v^n s), 0 \leq s < \Delta t, 0 \leq y < \Delta x\}$,

$$\begin{aligned}
G_{\Delta t}(t,x) = G_{j+1/2} =& g_\lambda \left(u_{\Delta t}\left(t, x - \frac{\Delta x}{2}\right) \top c_-, u_{\Delta t}\left(t, x + \frac{\Delta x}{2}\right) \top c_-, v_{\Delta t}(t) \right) \\
&- g_\lambda \left(u_{\Delta t}\left(t, x - \frac{\Delta x}{2}\right) \perp c_-, u_{\Delta t}\left(t, x + \frac{\Delta x}{2}\right) \perp c_-, v_{\Delta t}(t) \right)
\end{aligned}$$

and for every (t,x) in ${}_{j+1/2}^n$,

$$\xi_{\Delta t}(t,x) = \frac{\varphi_{j+1}^n - \varphi_j^n}{\Delta x}.$$

The sequence $(\xi_{\Delta t})$ converges uniformly to $(t,x) \mapsto \partial_x \varphi(t, x - h(t))$. By continuity of translations in L^1, the sequences $(u_{\Delta t}(t, \cdot + \frac{\Delta x}{2}))_{\Delta t}$ and $(u_{\Delta t}(t, \cdot - \frac{\Delta x}{2}))_{\Delta t}$ converge in L^1_{loc}, and therefore up to extraction almost everywhere, toward u. On the other hand, $(v_{\Delta t})$ converges almost everywhere toward h'. The consistency of the germ implies that $G_{\Delta t}$ converges almost everywhere to

$$g(u \top c_-, u \top c_-, h') - g(u \perp c_-, u \perp c_-, h') = \operatorname{sign}(u - c_-) \left(\left(\frac{u^2}{2} - h'u \right) - \left(\frac{c_-^2}{2} - h'c_- \right) \right).$$

As $(u_{\Delta t})$ and $(v_{\Delta t})$ are uniformly bounded in L^∞, the dominated convergence theorem yields

$$\Delta t \Delta x \sum_{j<0, n\in\mathbb{N}} G_{j+1/2}^n \frac{\varphi_{j+1}^n - \varphi_j^n}{\Delta x} \longrightarrow \int_{\mathbb{R}_-} \int_{\mathbb{R}_+} \Phi_{h'(t)}(u(t,x), c_-) \partial_x \varphi(t, x - h(t)) dt\, dx.$$

The second and fourth terms of (3.29) are easily treated:

$$\Delta x \sum_{i\in\mathbb{Z}} |u_j^0 - c_j|\varphi_j^0 \longrightarrow \int_{\mathbb{R}} |u^0 - c|\varphi(0,x)dx$$

and

$$\Delta t \Delta x \sum_{n\in\mathbb{N}} G^n_{j+1/2,+} \frac{\varphi_1^n - \varphi_0^n}{\Delta x} \longrightarrow 0.$$

Eventually, we study the convergence of

$$\Delta t \sum_{n\in\mathbb{N}} \mathrm{dist}_1(c, \mathcal{H}_\lambda(v^n))(\varphi_0^n + \varphi_1^n) = 2\int_{\mathbb{R}_+} \mathrm{dist}_1(c, \mathcal{H}_\lambda(v_{\Delta t}))\frac{\varphi_{\Delta t}(t, -\frac{\Delta x}{2}) + \varphi_{\Delta t}(t, \frac{\Delta x}{2})}{2} dt.$$

Clearly, $\frac{\varphi_{\Delta x}(t,-\frac{\Delta x}{2})+\varphi_{\Delta x}(t,\frac{\Delta x}{2})}{2}$ converges uniformly to $\varphi(\cdot, 0)$. Moreover,

$$\begin{aligned}
|\,\mathrm{dist}_1(c, \mathcal{H}_\lambda(v_{\Delta t})) - \mathrm{dist}_1(c, \mathcal{H}_\lambda(h'))| &= |\,\mathrm{dist}_1(c, (v_{\Delta t} - h', v_{\Delta t} - h') + \mathcal{H}_\lambda(h')) - \mathrm{dist}_1(c, \mathcal{H}_\lambda(h'))| \\
&= |\,\mathrm{dist}_1(c - (v_{\Delta t} - h', v_{\Delta t} - h'), \mathcal{H}_\lambda(h')) - \mathrm{dist}_1(c, \mathcal{H}_\lambda(h'))| \\
&\le |v_{\Delta t} - h'|
\end{aligned}$$

and

$$\Delta t \sum_{n\in\mathbb{N}} \mathrm{dist}_1(c, \mathcal{H}_\lambda(v^n))(\varphi_0^n + \varphi_1^n) \longrightarrow 2\int_{\mathbb{R}_+} \mathrm{dist}_1(c, \mathcal{H}_\lambda(h'))\varphi(t,0)dt,$$

which concludes the proof. □

3.3.2 Convergence of the particle's part

We now prove that the limit h of $(h_{\Delta t})$ verifies (3.5). To begin with, we prove that a discrete version of (3.5) holds.

Proposition 3.12. *Let $(u_j^n)_{n\in\mathbb{N},j\in\mathbb{Z}}$ and $(v^n)_{n\in\mathbb{N}}$ be given by Scheme* (3.6). *Then, for every sequences $(\xi^n)_{n\in\mathbb{N}}$ and $(\psi_j^n)_{n\in\mathbb{N},j\in\mathbb{Z}}$ such that $\psi_0^n = \psi_1^n = 1$ for all integrer n,*

$$\begin{aligned}
-m\Delta t \sum_{n\in\mathbb{N}^*} v^n \frac{\xi^n - \xi^{n-1}}{\Delta t} &= mv^0\xi^0 + \Delta x \Delta t \sum_{n\in\mathbb{N}^*, j\in\mathbb{Z}} u_j^n \frac{\psi_j^n \xi^n - \psi_j^{n-1}\xi^{n-1}}{\Delta t} \\
&\quad + \Delta x \sum_{j\in\mathbb{Z}} u_j^0 \xi^0 \psi_j + \Delta t \Delta x \sum_{n\in\mathbb{N}, j\neq 0} f^n_{j+1/2}\xi^n \frac{\psi^n_{j+1} - \psi^n_j}{\Delta x}.
\end{aligned} \tag{3.30}$$

Proof. We write

$$\begin{aligned}
m\sum_{n\in\mathbb{N}} v^{n+1}\xi^n = m\sum_{n\in\mathbb{N}} v^n\xi^n &+ \Delta t \sum_{n\in\mathbb{N}} (f^n_{1/2,-} - f^n_{1/2,+})\xi^n \\
&+ \Delta x \sum_{n\in\mathbb{N}} \sum_{j\notin\{0,1\}} \left[(u_j^n - u_j^{n+1}) - \mu(f^n_{j+1/2} - f^n_{j-1/2})\right]\xi^n\psi_j^n \\
&+ \Delta x \sum_{n\in\mathbb{N}} \left[(u_0^n - u_0^{n+1}) - \mu(f^n_{1/2,-} - f^n_{-1/2})\right]\xi^n \\
&+ \Delta x \sum_{n\in\mathbb{N}} \left[(u_1^n - u_1^{n+1}) - \mu(f^n_{3/2} - f^n_{1/2,+})\right]\xi^n.
\end{aligned}$$

This comes from the fact that the sum of the last three lines is zero. We now rearrange the different terms. On the one hand we have:

$$\sum_{n\in\mathbb{N}, j\le -1} (f^n_{j+1/2} - f^n_{j-1/2})\xi^n\psi_j^n = \sum_{n\in\mathbb{N}, j\le -1} f^n_{j+1/2}\xi^n(\psi_j^n - \psi_{j+1}^n) + \sum_{n\in\mathbb{N}} \xi^n f^n_{-1/2},$$

and on the other hand we have:

$$\sum_{n\in\mathbb{N},j\geq 2}(f^n_{j+1/2}-f^n_{j-1/2})\xi^n\psi^n_j=\sum_{n\in\mathbb{N},j\geq 1}f^n_{j+1/2}\xi^n(\psi^n_j-\psi^n_{j+1})-\sum_{n\in\mathbb{N}}\xi^n f^n_{3/2}.$$

It follows that

$$m\sum_{n\in\mathbb{N}}v^{n+1}\xi^n=m\sum_{n\in\mathbb{N}}v^n\xi^n+\Delta x\sum_{n\in\mathbb{N},j\in\mathbb{Z}}(u^n_j-u^{n+1}_j)\xi^n\psi^n_j-\Delta t\sum_{n\in\mathbb{N},j\neq 0}f^n_{j+1/2}\xi^n(\psi^n_j-\psi^n_{j+1}).$$

To conclude, we just have to rearrange the sum over n. Being careful with $n=0$ we obtain

$$\sum_{n\in\mathbb{N}}(v^{n+1}-v^n)\xi^n=\sum_{n\in\mathbb{N}^*}v^n(\xi^{n-1}-\xi^n)-v^0\xi^0$$

and

$$\sum_{n\in\mathbb{N},j\in\mathbb{Z}}(u^n_j-u^{n+1}_j)\xi^n\psi^n_j=\sum_{n\in\mathbb{N}^*,j\in\mathbb{Z}}u^n_j(\psi^n_j\xi^n-\psi^{n-1}_j\xi^{n-1})+\sum_{j\in\mathbb{Z}}u^0_j\xi^0\psi^0_j,$$

and the result follows by regrouping all the terms. $\square$

We can now pass to the limit $\Delta t\to 0$ in Proposition 3.12 to prove that h verifies (3.5).

Proposition 3.13. *Suppose that Hypothesis (3.11-3.19) hold, and that the CFL condition* (3.20) *is fulfilled. For all test functions ξ and ψ such that $\psi(0)=1$, the limit h of $(h_{\Delta t})$ verifies Inequality* (3.5).

Proof. Define

$$\psi^n_j=\psi(x^n_j-h^n)\quad\text{and}\quad\xi^n=\xi(n\Delta t).$$

Proposition 3.12 applies if $\psi^n_0=\psi^n_1=1$. Here, we only have

$$\forall j\in\{0,1\},\ \left|\psi^n_j-1\right|\leq C\Delta x.$$

The equality (3.30) holds up to the following corrections appearing in the left hand side:

$$\Delta x\Delta t\sum_{n\in\mathbb{N}^*,j\in\{0,1\}}u^n_j\frac{(1-\psi^n_j)\xi^n-(1-\psi^{n-1}_j)\xi^{n-1}}{\Delta t}+\Delta x\sum_{j\in\{0,1\}}u^0_j\xi^0(1-\psi^0_j)$$
$$+\Delta x\Delta t\sum_{n\in\mathbb{N}}\left(f^n_{-1/2}\xi^n\frac{(1-\psi^n_0)}{\Delta x}-f^n_{1/2}\xi^n\frac{(1-\psi^n_1)}{\Delta x}\right),$$

which all tends to zero since $\psi^n_0-1=O(\Delta x)$ and $\psi^n_1-1=O(\Delta x)$. The sequence

$$\zeta_{\Delta t}(t,x)=\frac{\psi^n_j\xi^n-\psi^{n-1}_j\xi^{n-1}}{\Delta t}\quad\text{if }(t,x)\in\mathcal{C}^n_j$$

converges uniformly to the function $(t,x)\mapsto\psi\xi'$. Indeed, by definition of the moving mesh, $x^n_j-h^n=x^{n-1}_j-h^{n-1}$. Therefore, $\psi^n_j=\psi^{n-1}_j$ and

$$\frac{\psi^n_j\xi^n-\psi^{n-1}_j\xi^{n-1}}{\Delta t}=\psi^n_j\frac{\xi^n-\xi^{n-1}}{\Delta t}$$

which converges uniformly toward the expected function. Now, define $F_{\Delta t}$ by

$$F_{\Delta t}(t,x)=g_\lambda\left(u_{\Delta t}\left(t,x-\frac{\Delta x}{2}\right),u_{\Delta t}\left(t,x+\frac{\Delta x}{2}\right),v_{\Delta t}(t)\right)$$

in such a way that for all (t,x) in ${}^{n}_{j+1/2}$,

$$F_{\Delta t}(t,x) = f^{n}_{j+1/2}.$$

By continuity of translations in L^1, the sequences $(u_{\Delta t}(t, \cdot + \frac{\Delta x}{2}))_{\Delta t}$ and $(u_{\Delta t}(t, \cdot - \frac{\Delta x}{2}))_{\Delta t}$ converge in L^1_{loc}, and therefore up to extraction, almost everywhere, toward u. On the other hand, $(v_{\Delta t})$ converges almost everywhere toward h'. The consistency of the flux (3.11) implies that $F_{\Delta t}$ converges almost everywhere to

$$g(u,u,h') = \frac{u^2}{2} - h'u.$$

□

3.3.3 A family of scheme consistent with a maximal part of the germ

In this section we exhibit a family of schemes that verifies the set of Assumptions (3.11-3.19). Let us clarify which maximal subset of $\mathcal{G}_\lambda$ is used.

Proposition 3.14. *The part $\mathcal{H}_\lambda(v) = \mathcal{G}^1_\lambda(v) \cup \mathcal{G}^2_\lambda(v)$ is a maximal subset of the germ.*

Proof. Following [AS12] (see Equations (13) and (14) in this reference), it suffices to show that if

$$\Xi_v((u_-,u_+),(v_-,v_+)) \geq 0 \quad \text{for any } (v_-,v_+) \in \mathcal{G}^2_\lambda(v), \tag{3.31}$$

then the stronger following property holds

$$\Xi_v((u_-,u_+),(v_-,v_+)) \geq 0 \quad \text{for any } (v_-,v_+) \in \mathcal{G}^1_\lambda(v) \cup \mathcal{G}^2_\lambda(v).$$

In the sequel we suppose that $v=0$. The general case follows by translation. The two main arguments are first, that Proposition 3.3 implies that this is automatically verified if (u_-,u_+) belongs to the germ, and second, that for all (v_-,v_+) in $\mathcal{G}^2_\lambda$, $\left|\frac{v_-^2 - v_+^2}{2}\right| \leq \frac{\lambda^2}{2}$. In the sequel, (v_-,v_+) always denotes an element of $\mathcal{G}^2_\lambda$. We proceed by a tedious, but not difficult, disjunction of cases.

- If $u_- \geq \lambda$ and $u_+ \geq 0$, then we want to prove that

  $$\frac{u_-^2 - v_-^2}{2} - \frac{u_+^2 - v_+^2}{2} \geq 0.$$

  If we apply Equation (3.31) to $(\lambda, 0)$, we obtain that

  $$\frac{u_-^2 - u_+^2}{2} \geq \frac{\lambda^2}{2}$$

  and the result follows.

- If $0 \leq u_- \leq \lambda$ and $u_+ \geq 0$, then (u_-,u_+) belongs to the germ. Indeed, Equation (3.31) applied to $(u_-,0)$ yields $-\frac{u_+^2}{2} \geq 0$ and therefore, $u_+ = 0$.

- If $u_- \leq 0$ and $u_+ \geq 0$, then (u_-,u_+) belongs to the germ. Indeed, Equation (3.31) applied to $(0,0)$ yields

  $$-\frac{u_-^2}{2} - \frac{u_+^2}{2} \geq 0$$

  and therefore, $u_- = u_+ = 0$.

- If $u_- \leq 0$ and $-\lambda \leq u_+ \leq 0$, then (u_-,u_+) belongs to the germ. Indeed, Equation (3.31) applied to $(0,u_+)$ yields $-\frac{u_-^2}{2} \geq 0$ and therefore, $u_- = 0$.

- If $u_- \leq 0$ and $\leq u_+ \leq -\lambda$, then we want to prove that
$$-\frac{u_-^2 - v_-^2}{2} + \frac{u_+^2 - v_+^2}{2} \geq 0.$$
If we apply Equation (3.31) to $(0, -\lambda)$, we obtain
$$-\frac{u_-^2}{2} + \frac{u_+^2 - \lambda^2}{2} \geq 0.$$
and the result follows.

- If $0 \leq u_- \leq \lambda$ and $u_+ \leq -\lambda$, let us first suppose that $u_- \geq v_-$. We have to prove that
$$\frac{u_-^2 - v_-^2}{2} + \frac{u_+^2 - v_+^2}{2} \geq 0.$$
But $0 \leq v_- \leq u_-$ and $0 \geq v_+ \geq u_+$, and we have the result:
$$\frac{v_-^2 + v_+^2}{2} \leq \frac{u_-^2 + v_+^2}{2} \leq \frac{u_-^2 + u_+^2}{2}.$$
We now suppose that $u_- \leq v_-$. We want to prove that
$$-\frac{u_-^2 - v_-^2}{2} + \frac{u_+^2 - v_+^2}{2} \geq 0.$$
Moreover, (u_-, u_+) does not belong to the germ $\mathcal{G}_\lambda$, and therefore $u_+ \leq -u_- - \lambda$ and
$$\frac{u_+^2 - u_-^2}{2} \geq \frac{2u_-\lambda + \lambda^2}{2} \geq \frac{\lambda^2}{2} \geq \frac{v_+^2 - v_-^2}{2}$$

- If $\lambda \leq u_-$ and $u_+ \leq -\lambda$, the result
$$\frac{u_-^2 - v_-^2}{2} + \frac{u_+^2 - v_+^2}{2} \geq 0$$
is a straightforward consequence of
$$\frac{u_-^2 + u_+^2}{2} \geq \lambda^2 \geq \frac{v_-^2 + v_+^2}{2}.$$

- Eventually, if $\lambda \leq u_-$ and $-\lambda \leq u_+ \leq 0$, let us first suppose that $u_+ \leq v_+$ and prove
$$\frac{u_-^2 - v_-^2}{2} + \frac{u_+^2 - v_+^2}{2} \geq 0.$$
It follows from
$$\frac{v_+^2 + v_-^2}{2} \leq \frac{u_+^2 + v_-^2}{2} \leq \frac{u_+^2 + u_-^2}{2}.$$
Suppose now that $u_+ > v_+$ and $u_+ \geq -u_- + \lambda$. The result
$$\frac{u_-^2 - v_-^2}{2} - \frac{u_+^2 - v_+^2}{2} \geq 0$$
comes from
$$\frac{u_-^2 - u_+^2}{2} \geq \frac{-2\lambda u_+ + \lambda^2}{2} \geq \frac{\lambda^2}{2} \geq \frac{v_-^2 - v_+^2}{2}.$$

□

It is possible to find fluxes that verifies (3.15) with $\mathcal{H}_\lambda = \mathcal{G}^1_\lambda \cup \mathcal{G}^2_\lambda$ and (3.18).

Proposition 3.15. *The family of finite volume schemes defined by*

$$\begin{cases} g^-_\lambda(u_-, u_+, v) = g(u_-, \min(u_+ + \lambda, \max(u_-, v)), v) \\ g^+_\lambda(u_-, u_+, v) = g(\max(u_- - \lambda, \min(u_+, v)), u_+, v) \end{cases} \tag{3.32}$$

is consistent with $\mathcal{G}^1_\lambda(v) \cup \mathcal{G}^2_\lambda(v)$ *and verifies the monotonicity assumptions* $\partial_1 g^\pm_\lambda \geq 0$ *and* $\partial_2 g^\pm_\lambda \leq 0$.

Proof. The proof consists in a simple verification. We first check that for all u_- and u_+ in $\mathbb{R}$,

$$g^-_\lambda(u_-, u_- - \lambda, v) = g(u_-, \min(u_-, \max(u_-, v)), v) = g(u_-, u_-, v)$$

and

$$g^+_\lambda(u_+ + \lambda, u_+, v) = g(\max(u_+, \min(u_+, v)), u_+, v) = g(u_+, u_+, v).$$

Then, we verify that for all u_+ in $[v - \lambda, v]$,

$$g^-_\lambda(v, u_+, v) = g(u_-, \min(u_+ + \lambda, \max(v, v)), v) = g(v, v, v)$$

and

$$g^+_\lambda(v, u_+, v) = g(\max(v - \lambda, \min(u_+, v)), u_+, v) = g(u_+, u_+, v)$$

while for every u_- in $[v, v + \lambda]$,

$$g^-_\lambda(u_-, v, v) = g(u_-, \min(v + \lambda, \max(u_-, v)), v) = g(u_-, u_-, v)$$

and

$$g^+_\lambda(u_-, v, v) = g(\max(u_- - \lambda, \min(v, v)), v, v) = g(v, v, v).$$

Eventually, the monotonicity properties are implied by those on g as soon as soon as the first component is not u_+ and the second is not u_-. But if the first component is u_+, then $u_+ < v$ and $\partial_2 g^+_\lambda = u_+ - v \leq 0$, while if the second component is u_-, then $u_- > v$ and $\partial_2 g^-_\lambda = u_- - v \geq 0$. □

It remains to prove that Assumption (3.19) holds. This is not the case for every choice of flux g (a counterexample can be found in [AS12]), but we can check it for three classical fluxes.

Proposition 3.16. *The family of finite volume schemes* (3.32) *verifies that* $g^-_\lambda - g^+_\lambda$ *is nondecreasing with respect to its two first variables if* g *is the Godunov, the Rusanov or the Enguish–Osher numerical flux.*

Proof. Let us divide the phase space (u_-, u_+) in six zones, depending on which values are taken by g^- and g^+:

$$g^-_\lambda(u_-, u_+, v) = \begin{cases} g(u_-, u_-, v) & \text{if } v \leq u_- \leq u_+ + \lambda & \text{zone I,} \\ g(u_-, v, v) & \text{if } u_- \leq v \leq u_+ + \lambda & \text{zone II,} \\ g(u_-, u_+ + \lambda, v) & \text{if } u_+ + \lambda \leq \max(u_-, v) & \text{zone III,} \end{cases}$$

while

$$g^+_\lambda(u_-, u_+, v) = \begin{cases} g(u_+, u_+, v) & \text{if } u_- - \lambda \leq u_+ \leq v & \text{zone 1,} \\ g(v, u_+, v) & \text{if } u_- - \lambda \leq v \leq u_+ & \text{zone 2,} \\ g(u_- - \lambda, u_+, v) & \text{if } \min(u_+, v) \leq u_- - \lambda & \text{zone 3.} \end{cases}$$

These zones are depicted on Figure 3.3. If u_+ belongs to zones 1 or 2, g^+ does not depends on u_- and $g^-_\lambda - g^+_\lambda$ is nondecreasing with respect to its first argument. Similarly, if u_- belongs to zones I or II, $g^-_\lambda - g^+_\lambda$ is nondecreasing towards its second argument. We focus on the case where u_- belongs to zone III or u_+ belongs to zone 3. Let us first remark that the case where u_- belongs to zone III and u_+ is in zone 3 reduces to the choice of flux studied in [AS12], where the monotonicity property has

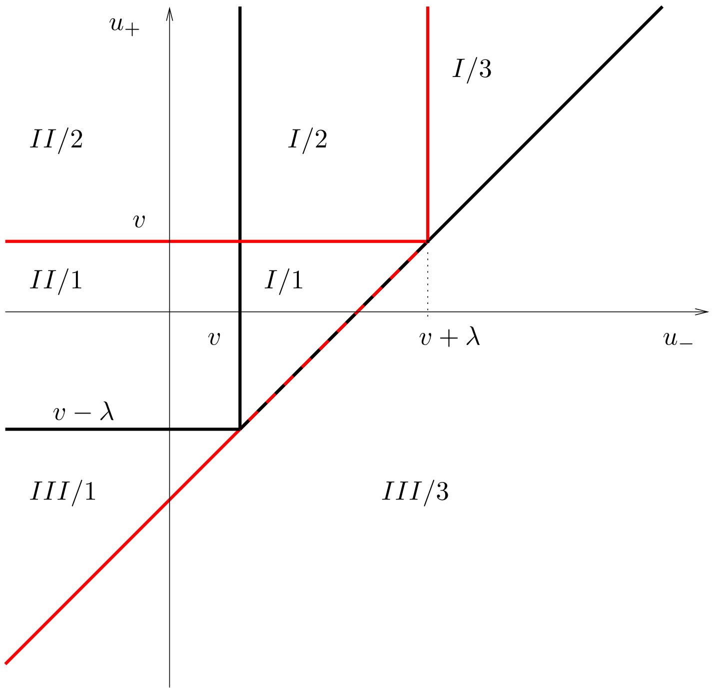


Figure 3.3. *Choice of the fluxes in the family of finite volume schemes* (3.32).

been proven for the Godunov, Rusanov and Enguist–Osher scheme. Suppose that case u_- is in zone I and u_+ is in zone 3. Then we have

$$(g_\lambda^- - g_\lambda^+)(u_-, u_+, v) = g(u_-, u_-, v) - g(u_- - \lambda, u_+, v).$$

For the sake of simplicity we assume that $v = 0$.

- If g is the Godunov flux, as $u_+ + \lambda \geq u_- \geq \lambda$, the Riemann problem between $u_- - \lambda$ and u_+ is a shock traveling faster than v. It follows that

$$(g_\lambda^- - g_\lambda^+)(u_-, u_+, 0) = \frac{(u_-)^2}{2} - \frac{(u_- - \lambda)^2}{2} = \lambda u_- - \frac{\lambda^2}{2}$$

  is nondecreasing toward its first two arguments.

- If g is the Rusanov's flux,

$$(g_\lambda^- - g_\lambda^+)(u_-, u_+, 0) = \frac{(u_-)^2}{2} - \left(\frac{(u_- - \lambda)^2 + u_+^2}{4} - (u_- - \lambda)\frac{u_+ - (u_- - \lambda)}{2}\right)$$

  and we have

$$\begin{aligned}\partial_1(g_\lambda^- - g_\lambda^+)(u_-, u_+, 0) &= u_- - \left(\frac{u_- - \lambda}{2} - \frac{u_+ - (u_- - \lambda)}{2} + \frac{u_- - \lambda}{2}\right) \\ &= \frac{-u_- + 3\lambda + u_+}{2}.\end{aligned}$$

  As u_- belongs to zone I, $u_+ + \lambda \geq u_-$, and the last quantity is larger than λ. On the other hand,

$$\partial_2(g_\lambda^- - g_\lambda^+)(u_-, u_+, 0) = -\frac{u_+ - (u_- - \lambda)}{2}$$

  and this last quantity is nonnegative because u_+ belongs to zone 3.

- Eventually, if g is the Enguist–Osher scheme, the fact that $0 \leq u_- - \lambda \leq u_+$ implies that

$$(g_\lambda^- - g_\lambda^+)(u_-, u_+, 0) = \frac{(u_-)^2}{2} - \frac{(u_- - \lambda)^2}{2} = \lambda u_- - \frac{\lambda^2}{2}$$

is once again nondecreasing with respect to its first two arguments. The case where u_- is in zone III while u_+ is in zone 1 can be treated in a symmetrical way.

□

3.4 Convergence of schemes only consistent with $\mathcal{G}_\lambda^1$

In this section, we no longer require Hypothesis (3.15) to be fulfilled, and prove convergence of a family of finite volume schemes that verifies only (3.14). The difficulty is that $\mathcal{G}_\lambda^1$ is not a maximal part of the germ, and we cannot prove a discrete version of (3.34) directly. The key point is to study the convergence of the solution of Scheme (3.6) for initial datum in the maximal subset of the germ $\mathcal{G}_\lambda^1 \cup \mathcal{G}_\lambda^2$. We then extend the comparison argument of [AS12] to prove convergence on arbitrary initial data.

3.4.1 Proof of convergence

Let us now focus on fluxes that do not preserve a maximal part of the germ (in the sense of Hypothesis (3.15)), but only the straight line $\mathcal{G}_\lambda^1(v)$, i.e. that verifies (3.14) but not (3.15). Our aim is to prove the following theorem.

Theorem 3.17. *If the numerical fluxes around the particle are given by*

$$\begin{cases} f_{1/2,-}^n(u_0^n, u_1^n, v^n) = g(u_0^n, u_1^n + \lambda, v^n), \\ f_{1/2,+}^n(u_0^n, u_1^n, v^n) = g(u_0^n - \lambda, u_1^n, v^n), \end{cases} \tag{3.33}$$

where g is a numerical flux verifying (3.11-3.14) and (3.16-3.19), and if the CFL condition (3.20) *holds, Scheme* (3.6) *converges toward the solution of* (3.1).

Proof. Let us first remark that Proposition 3.5 and Proposition 3.6 did not use Hypothesis (3.15), thus we can extract converging subsequences as we did in the previous Section. Now, consider a test function φ supported in $\{x < 0\}$ or $\{x > 0\}$, we have that $\varphi_0^n = \varphi_1^n = 0$ for small enough Δx. We easily obtain, as in Proposition 3.8, that for all c in $\mathbb{R}$, for all $j \leq -1$,

$$\frac{|u_j^{n-1} - c| - |u_j^n - c|}{\Delta t} + \frac{G_{j+1/2}^n - G_{j-1/2}^n}{\Delta x} \leq 0.$$

Multiplying by $\Delta t \Delta x \varphi_j^n$ and summing over $n \in \mathbb{N}$ and $j \leq -1$, we obtain as in Proposition 3.9

$$\Delta t \Delta x \sum_{j \in \mathbb{Z}, n \leq -1} |u_j^{n+1} - c| \frac{\varphi_j^{n+1} - \varphi_j^n}{\Delta t} + \Delta x \sum_{i \in \mathbb{Z}} |u_j^0 - c| \varphi_j^0 + \Delta t \Delta x \sum_{j \in \mathbb{Z}^*, n \leq -1} G_{j+1/2}^n \frac{\varphi_{j+1}^n - \varphi_j^n}{\Delta x} \geq 0$$

and we straightforwardly obtain that the limit u of the scheme is an entropy solution of the Burgers equation on the sets $\{x < h\}$ (and similarly on $\{x > h\}$). It remains to prove that the traces around the particle belong to the germ for almost every time. Let us fix a time t_0 such that h' and the traces $u_-(t_0)$ and $u_+(t_0)$ exist. Fix (c_-, c_+) in $\mathcal{H}_\lambda(h'(t_0))$. Our aim is to prove a discrete version of (3.4). Let us first suppose that (c_-, c_+) belongs to the straight line $\mathcal{G}_\lambda^1(h'(t_0))$ but not to the closed square $\overline{\mathcal{G}_\lambda^2(h'(t_0))}$. By continuity of h', there exists $\delta > 0$ such that,

$$\forall t \in (t_0 - \delta, t_0 + \delta),\ \operatorname{dist}_1((c_-, c_+), \mathcal{G}_\lambda^1(h'(t))) = \operatorname{dist}_1((c_-, c_+), \mathcal{H}_\lambda^1(h'(t)))$$

(see Figure 3.1). Up to taking a smaller δ, this equality is also true at the numerical level for small enough Δt, since from Proposition 3.7, $(v_n)_{n\in\mathbb{N}}$ converges. Therefore, passing to the limit in (3.29) with φ supported in time in $(t_0-\delta, t_0+\delta)$, we directly obtain (3.4).

We now treat the case where (c_-, c_+) belongs to the interior of $\mathcal{G}^2_\lambda(h'(t_0))$. The principle of the proof is to compare the numerical solution with another one, for which the initial data is much simpler as it corresponds to an element of $\mathcal{G}^2_\lambda(h'(t_0))$. Since h' is continuous, there exists δ such that

$$\forall t \in (t_0-\delta, t_0+\delta),\ (c_-, c_+) \in \mathcal{G}^2_\lambda(h'(t))$$

and on the time interval $(t_0-\delta, t_0+\delta)$, (3.4) becomes

$$\int_{\mathbb{R}_+}\int_{\mathbb{R}} |u-c|(s,x)\partial_t\varphi(s,x-h(s)) + \Phi_{h'(t)}(u,c)(s,x)\partial_x\varphi(s,x-h(s))dx\,ds \geq 0. \tag{3.34}$$

Up to reducing δ and for small enough Δt, this is also true at the numerical level. Now, for $(u_j^n)_{j\in\mathbb{Z},n\in\mathbb{N}}$ and $(v^n)_{n\in\mathbb{N}}$ given by the fully coupled scheme (3.6), consider $(c_j^n)_{j\in\mathbb{Z},n\in\mathbb{N}^*}$ the sequence given by the scheme

$$\begin{cases} c_j^{n+1} &= c_j^n - \mu(g(c_j^n, c_{j+1}^n, v^n) - g(c_{j-1}^n, c_j^n, v^n)) \text{ for } j \notin \{0,1\}, \\ c_0^{n+1} &= c_0^n - \mu(g(c_0^n, c_1^n+\lambda, v^n) - g(c_{-1}^n, c_0^n, v^n)), \\ c_1^{n+1} &= c_1^n - \mu(g(c_1^n, c_2^n, v^n) - g(c_0^n-\lambda, c_1^n, v^n)), \end{cases} \tag{3.35}$$

with initial data

$$c_j^0 = \begin{cases} c_- & \text{if } j \leq 0, \\ c_+ & \text{if } j > 0. \end{cases} \tag{3.36}$$

We recall that (c_-, c_+) belongs to $\mathcal{G}^2_\lambda(h'(t_0))$. Simple modifications of Propositions 3.8 and 3.9 yield

$$\begin{aligned} &\Delta t\Delta x \sum_{j\in\mathbb{Z},n\in\mathbb{N}} |u_j^{n+1} - c_j^{n+1}| \frac{\varphi_j^{n+1}-\varphi_j^n}{\Delta t} + \Delta x \sum_{i\in\mathbb{Z}} |u_j^0 - c_j^0|\varphi_j^0 \\ &+ \Delta t\Delta x \sum_{j\in\mathbb{Z}^*,n\in\mathbb{N}} G^n_{j+1/2}\frac{\varphi_{j+1}^n-\varphi_j^n}{\Delta x} + \Delta t\Delta x\sum_{n\in\mathbb{N}} G^n_{j+1/2,+}\frac{\varphi_1^n-\varphi_0^n}{\Delta x} \geq 0. \end{aligned}$$

Suppose that $(c_j^n)_{j\in\mathbb{Z},n\in\mathbb{N}}$ converges to $c(t,x) = c_-\mathbf{1}_{x<h(t)} + c_+\mathbf{1}_{x>h(t)}$ on the interval $(t_0-\delta, t_0+\delta)$. Then with $\varphi_j^n = \varphi(t^n, x_j^n)$ where φ is a test function supported in $(t_0-\delta, t_0+\delta)$, we obtain (3.34) by passing to the limit. We now study this convergence.

Lemma 3.18. *Suppose that at iteration n, the sequence $(c_j^n)_{j\in\mathbb{Z}}$ given by the scheme* (3.35) *is nondecreasing on $j \leq 0$ and on $j \geq 1$, and such that*

$$\forall j \leq 0,\ c_- \leq c_j^n \leq c_- + \lambda \ \text{ and } \ \forall j \geq 1,\ c_+ - \lambda \leq c_j^n \leq c_+$$

and

$$c_0^n - c_1^n \leq \lambda,$$

then the same holds at iteration $n+1$.

Proof. The monotonicity of $(c_j^{n+1})_{j\leq 0}$ follows from the monotonicity of H_λ under the CFL condition (3.20). For $j \leq -2$, we have

$$c_j^{n+1} = H_\lambda(c_{j-1}^n, c_j^n, c_{j+1}^n) \leq H_\lambda(c_j^n, c_{j+1}^n, c_{j+2}^n) = c_{j+1}^{n+1}.$$

As $c_0^n \leq c_1^n + \lambda$, we also have

$$c_{-1}^{n+1} = H_\lambda(c_{-2}^n, c_{-1}^n, c_0^n) \leq H_\lambda(c_{-1}^n, c_0^n, c_1^n+\lambda) = c_0^{n+1}.$$

Moreover, for $j \le -1$, both c_{j-1}^n, c_j^n and c_{j+1}^n are between c_- and $c_- + \lambda$, thus the same holds at iteration $n+1$. For $j = 0$, as $c_+ \le c_-$ (because (c_-, c_+) belongs to $\mathcal{G}_\lambda^2(h'(t_0))$), we conclude by remarking that

$$c_- \le c_0^n \le c_1^n + \lambda \le c_+ + \lambda \le c_- + \lambda.$$

The results for positive integers j are obtained in a similar way. Let us now prove that $u_0^{n+1} - u_1^{n+1} \le \lambda$. We have

$$\begin{aligned} c_0^{n+1} - c_1^{n+1} &= H_\lambda(c_{-1}^n, c_0^n, c_1^n + \lambda) - H_\lambda(c_0^n - \lambda, c_1^n, c_2^n) \\ &\le H_\lambda(c_0^n, c_0^n, c_1^n + \lambda) - H_\lambda(c_0^n - \lambda, c_1^n, c_1^n) \\ &\le c_0^n + \mu L |c_0^n - (c_1^n + \lambda)| - c_1^n + \mu L |(c_0^n - \lambda) - c_1^n| \\ &\le c_0^n - c_1^n + (c_1^n + \lambda - c_0^n) \\ &\le \lambda. \end{aligned}$$

□

For (c_-, c_+) in the *open* subset $\mathcal{G}_\lambda^2(h'(t_0))$, there exists a positive δ such that $h'(t)$ stays in the interval (c_+, c_-) on the time interval $(t_0 - \delta, t_0 + \delta)$. For small enough Δt, it is also true at the numerical level. Up to reducing slightly δ, (c_-, c_+) belongs to $\mathcal{G}_\lambda^2(v^n)$ for small enough Δt and for all iteration in time such that t^n belongs to $(t_0 - \delta, t_0 + \delta)$, and in particular $c_+ \ge v^n \ge c_-$.

Thus the limit c of the scheme (3.35) with initial data (3.36) at time $t_0 - \delta$ is such that c is larger than h' on $x < h$ and smaller on $x > h$. It allows to prove that c is, on $\{(t, x) : x < h(t)\}$, the solution of

$$\begin{cases} \partial_t u + \partial_x \frac{u^2}{2} = 0 & \forall t \in (t_0 - \delta, t_0 + \delta), \forall x < h(t), \\ u(t_0 - \delta, x) = c_- & \forall x < h(0) \\ u(t, h(t)) = h'(t) & \forall t \in (t_0 - \delta, t_0 + \delta). \end{cases} \tag{3.37}$$

As c_- is larger than h' on the whole time interval, the boundary condition is inactive and the solution is $u = c_-$. Let us recall the definition given by Bardos, LeRoux and Nedelec in [BLN79] of this conservation law on a bounded domain. A function u in L^∞ is a solution of

$$\begin{cases} \partial_t u + \partial_x f(u) = 0 & \forall t > 0, \forall x < h(t), \\ u(t = 0, x) = u^0(x) & \forall x < h(0), \\ u(t, h(t)) = u_b(t) & \forall t > 0, \end{cases}$$

if for all real κ and for all nonnegative function $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}_+ \times \mathbb{R})$, the following inequality holds:

$$\begin{aligned} &\int_{t>0} \int_{x<h(t)} |u(t,x) - \kappa| \partial_t \varphi(t, x - h(t)) + \Phi_{h'(t)}(u(t,x), \kappa) \partial_x \varphi(t, x - h(t)) dx\, dt \\ &\quad + \int_{x<h(0)} |u^0(x) - \kappa| \varphi(0, x) dx + \int_{t>0} \operatorname{sign}(\kappa - u_b(t)) \{ f(u(t, h(t)^-)) - f(\kappa) \} \varphi(t, 0) \ge 0. \end{aligned} \tag{3.38}$$

The convergence of finite volume schemes for scalar conservation laws in a bounded domain has been proven in [Vov02]. We are here in a favorable case: we can obtain a discrete version of (3.38) by summing (3.28) multiplied by $\Delta t\, \Delta x\, \varphi_j^n$ over $n \ge 0$ and $j \le -1$. We obtain

$$\begin{aligned} \Delta t \Delta x \sum_{n \ge 0, j \le -1} |c_j^{n+1} - \kappa| \frac{\varphi_j^{n+1} - \varphi_j^n}{\Delta t} + \Delta x \sum_{j \le -1} |c_j^0 - \kappa| \varphi_j^0 \\ \Delta t \Delta x \sum_{n \ge 0, j \le -1} G_{j+1/2}^n \frac{\varphi_{j+1}^n - \varphi_j^n}{\Delta x} - \Delta t \sum_{n \ge 0} G_{-1/2}^n \varphi_0^n \le 0. \end{aligned}$$

Passing to the limit yields

$$\int_{t>0}\int_{x<h(t)} |c(t,x)-\kappa|\partial_t\varphi(t,x-h(t)) + \Phi_{h'(t)}(c(t,x),\kappa)\partial_x\varphi(t,x-h(t))dx\,dt \\ + \int_{x<h(0)} |c_- - \kappa|\varphi(0,x)dx + \int_{t>0} \mathrm{sign}(\kappa - c(t,h(t)))\{f(c(t,h(t)^-)) - f(\kappa)\}\varphi(t,0) \geq 0.$$

To conclude we check that

$$\mathrm{sign}(\kappa - h'(t))\{f(c(t,h(t)^-)) - f(\kappa)\} \geq -\,\mathrm{sign}(c(t,h(t)-\kappa))\{f(c(t,h(t)^-)) - f(\kappa)\}.$$

This relies strongly on the fact that c remains larger than h'.

- If $h' \leq \kappa \leq c$, the inequality reduces to

  $$\{f(c(t,h(t)^-)) - f(\kappa)\} \geq -\{f(c(t,h(t)^-)) - f(\kappa)\}$$

  which holds because f is increasing on $(0,+\infty)$.

- If $h' \leq c \leq \kappa$ or $\kappa \leq h' \leq c$ the inequality reduces to

  $$\{f(c(t,h(t)^-)) - f(\kappa)\} \geq \{f(c(t,h(t)^-)) - f(\kappa)\}$$

  or

  $$-\{f(c(t,h(t)^-)) - f(\kappa)\} \geq -\{f(c(t,h(t)^-)) - f(\kappa)\},$$

  which are both trivial.

□

Remark 3. Of course, Theorem 3.17 applies when the initial data is

$$u^0(x) = c_-\mathbf{1}_{x<0} + c_+\mathbf{1}_{x\geq 0},$$

with $(c_-, c_+) \in \mathcal{G}^2_\lambda(v^0)$. In the next subsection, we prove the convergence for this specific initial data directly, without using the local in time comparison with the one-way scheme (3.35) in which the velocity of the particle is fixed.

3.4.2 Detailed analysis when the initial data belongs to $\mathcal{G}^2_\lambda(v^0)$

Our aim in this section is to prove directly that if

$$u^0(x) = u_-\mathbf{1}_{x<0} + u_+\mathbf{1}_{x\geq 0} \tag{3.39}$$

with (u_-, u_+) in $\mathcal{G}^2_\lambda(v^0)$, Scheme (3.6) converges toward the exact solution, which in that case is given by

$$\begin{cases} h(t) = \frac{u_-+u_+}{2}t + \left(v^0 - \frac{u_-+u_+}{2}\right)\frac{m_p}{u_- - u_+}\left(1 - e^{-\frac{u_- - u_+}{m_p}t}\right), \\ u(t,x) = u_-\mathbf{1}_{x<h(t)} + u_+\mathbf{1}_{x\geq h(t)}. \end{cases}$$

In this section and for technical reasons, we consider System (3.1) on the space interval $[-a,a]$, with periodic boundary condition

$$u(t,-a) = u(t,a) \quad \forall t > 0.$$

Numerically, the interval $[-a, a]$ is discretized with $2M_c$ cells ($\Delta x = \frac{a}{M_c}$) and we prove the convergence of the scheme

$$\begin{cases} u_j^{n+1} &= u_j^n - \mu(g(u_j^n, u_{j+1}^n, v^n) - f(u_{j-1}^n, u_j^n, v^n)) \text{ for } j \in [[-M_c+1, M_c]], j \notin \{0,1\}, \\ u_0^{n+1} &= u_0^n - \mu(g(u_0^n, u_1^n + \lambda, v^n) - f(u_{-1}^n, u_0^n, v^n)), \\ u_1^{n+1} &= u_1^n - \mu(g(u_1^n, u_2^n, v^n) - g(u_0^n - \lambda, u_1^n, v^n)), \\ u_{-M_c}^n &= u_{M_c}^n \quad \text{and} \quad u_{M_c+1}^n = u_{-M_c+1}^n, \\ v^{n+1} &= v^n + \frac{\Delta t}{m_p}(g(u_0^n, u_1^n + \lambda, v^n) - g(u_0^n - \lambda, u_1^n, v^n), \\ x_j^{n+1} &= x_j^n + v^n \Delta t. \end{cases} \tag{3.40}$$

We prove the convergence on the time interval $[0, T]$, and we chose a large enough so that the restriction of the periodic solution on $[-a/2, a/2]$ coincides with the solution with the initial data (3.39) on this interval. In particular for small enough Δt, Scheme (3.6) and Scheme (3.40) coincide for $j \in [[-M_c/2, M_c/2]]$.

Proposition 3.19. *Suppose that the numerical flux g verifies (3.11-3.14) and (3.16-3.19), that*

$$\forall A \in \mathbb{R}, \forall B \in \mathbb{R}, \quad g(v - A, v - B, v) = g(v + B, v + A, v), \tag{3.41}$$

and that $\partial_3 g$ is decreasing with respect to its first two arguments. Under Condition (3.20) *and for the initial data* (3.39)*, Scheme* (3.40) *converges toward the solution of* (3.1) *(with periodic boundary conditions).*

Proof. We prove as we did in Section 3.4 that $(u_j^n)_{-M_c/2 \le j \le 0}$ converges toward the solution of (3.37), with a Dirichlet boundary condition on the left of the particle. The key point is to prove that v^n remains smaller than c_- on the whole time interval, in which case the boundary condition is inactive and we obtain the result. Similarly on the right of the particle, the boundary condition is inactive if v^n remains larger than c_+.

To prove that $c_+ \le v^n \le c_-$, we apply the Crandall–Tartar lemma [CT80] to the application

$$T: \begin{array}{ccc} \mathcal{S} & \longrightarrow & \mathcal{S} \\ ((u_j^0)_{-M_c+1 \le j \le M_c}, v^0) & \longmapsto & ((u_j^{n+1})_{-M_c+1 \le j \le M_c}, v^{n+1}). \end{array}$$

where

$$\mathcal{S} = \{((b_j)_{j \in [[-M_c+1, M_c]]}, v) : \ b_1 \le b_2 \le \cdots \le b_{M_c} \le b_{-M_c+1} \le \cdots \le b_{-1} \le b_0 \le b_1 + \lambda\}.$$

Lemma 3.20 (Crandall–Tartar)**.** *Let (Ω, μ) be a measured space, and let $\mathcal{S}$ be a subset of $L^1(\Omega)$ stable by sup:*

$$\forall (u, v) \in \mathcal{S}^2, \quad \max(u, v) \in \mathcal{S}.$$

Consider a function $T : \mathcal{S} \to \mathcal{S}$ such that

$$\forall u \in \mathcal{S}, ||T(u)||_{L^1} = ||u||_{L^1}$$

Then, if T is order preserving,

$$||T(u) - T(v)||_{L^1} \le ||u - v||_{L^1}$$

In our case, $\Omega = \mathbb{R}^{2M_c} \times \mathbb{R}$ and

$$||(b_j)_{j \in [[-M_c+1, M_c]]}, v)||_{L^1} = \Delta x \sum_{j=-M_c+1}^{M_c} |b_j| + m|v|.$$

The fact that T takes its values in $\mathcal{S}$ is proven exacty as in the proof of Lemma 3.18. We prove in Lemma 3.22 that T is order preserving. Applying the Crandall–Tartar lemma to $((u_j^0), v^0)$ and $(\bar{u}_j^0, \bar{v}) = ((u_j^0), \frac{c_- + c_+}{2})$, we obtain

$$\Delta x \sum_{j=-M_c+1}^{M_c} |\bar{u_j}^{n+1} - u_j^{n+1}| + m|\bar{v}^{n+1} - v^{n+1}| \le m \left| v^0 - \frac{u_- + u_+}{2} \right|.$$

The result follows since $\bar{v}^{n+1} = \frac{u_-+u_+}{2}$ (see Lemma 3.21 below).

$\square$

Lemma 3.21. *If g verifies* (3.41) *and if the initial data is*

$$\begin{cases} u_j^0 = u_- & \text{for } j \leq 0, \\ u_j^0 = u_+ & \text{for } j \geq 1, \\ v^0 = \frac{u_-+u_+}{2}, \end{cases}$$

then Scheme (3.40) *verifies $v^n = v^0$ for all integer n.*

Proof. We prove by induction the following stronger result:

$$\forall n \in \mathbb{N}, \forall j \leq 0,\ v^n = \frac{u_- + u_+}{2} \quad \text{and} \quad u_{-j}^n - v^n = v^n - u_{j+1}^n.$$

The symmetry of the initial data ensures that this is verified for $n = 0$. Suppose that this is verified for some $n \geq 0$. Hypothesis (3.41) on the flux and the induction hypothesis yield

$$\begin{aligned} g(u_0^n, u_1^n + \lambda, v^n) &= g(v^n - (u_1^n + \lambda - v^n), v^n - (u_0^n - v^n), v^n) \\ &= g(u_0^n - \lambda, u_1^n, v^n). \end{aligned}$$

Hence, the velocity remains constant. A similar reasoning can be applied to the fluid velocity. Let us give some details for $j \leq -1$:

$$\begin{aligned} u_{-j}^{n+1} &= u_{-j}^n - \mu(g(u_{-j}^n, u_{-(j-1)}^n, v^n) - g(u_{-(j+1)}^n, u_{-j}^n, v^n)) \\ &= 2v^n - u_{j+1}^n - \mu\Big[g(v^n - (v^n - u_{-j}^n), (v^n - (v^n - u_{-(j-1)}^n), v^n)) \\ &\qquad\qquad -g(v^n - (v^n - u_{-(j+1)}^n), v^n - (v^n - u_{-j}^n), v^n)\Big] \\ &= 2v^n - \Big[u_{j+1}^n + \mu[g(2v^n - u_{-(j-1)}^n, 2v^n - u_{-j}^n, v^n) \\ &\qquad\qquad -g(2v^n - u_{-j}^n, 2v^n - u_{-(j+1)}^n, v^n)\Big] \\ &= 2v^n - \left(u_{j+1}^n - \mu\left[g(u_{j+1}^n, u_{j+2}^n, v^n) - g(u_j^n, u_{j+1}^n, v^n)\right]\right) \\ &= 2v^{n+1} - u_{j+1}^{n+1}, \end{aligned}$$

and for $j = 0$:

$$\begin{aligned} u_0^{n+1} &= u_0^n - \mu(g(u_0^n, u_0^n - \lambda, v^n) - g(u_{-1}^n, u_0^n, v^n) \\ &= 2v^n - u_1^n - \mu\left[g(v^n - (v^n - u_0^n), (v^n - (v^n - u_0^n + \lambda), v^n)) \right.\\ &\qquad\qquad \left. -g(v^n - (v^n - u_{-1}^n), v^n - (v^n - u_0^n), v^n)\right] \\ &= 2v^n - [u_1^n + \mu[g(2v^n - u_0^n + \lambda, 2v^n - u_0^n, v^n) \\ &\qquad\qquad -g(2v^n - u_0^n, 2v^n - u_{-1}^n, v^n)] \\ &= 2v^n - (u_1^n - \mu\left[g(u_1^n, u_2^n, v^n) - g(u_1^n + \lambda, u_1^n, v^n)\right]) \\ &= 2v^{n+1} - u_1^{n+1}. \end{aligned}$$

$\square$

Lemma 3.22. *Suppose that $\partial_3 g$ is decreasing with respect to its first two arguments, and that* (3.12), (3.20) *and* (3.19) *hold. Then, if two initial data are ordered, this order is conserved after one iteration of the scheme. More precisely, if $[(u_j^n)_{j\in\mathbb{Z}}, v^n]$ and $[(\bar{u}_j^n)_{j\in\mathbb{Z}}, \bar{v}^n]$ are two elements of $\mathcal{S}$ are such that*

$$\forall j \in \mathbb{Z},\ u_j^n \leq \bar{u}_j^n \quad \text{and} \quad v^n \leq \bar{v}^n,$$

then, if $\partial_3 g$ is decreasing with respect to its first two arguments and if

$$\frac{2\Delta t}{m_p} \max |\partial_3 g| < 1, \tag{3.42}$$

then

$$\forall j \in \mathbb{Z},\ u_j^{n+1} \leq \bar{u}_j^{n+1} \quad \textit{and} \quad v^{n+1} \leq \bar{v}^{n+1}.$$

Proof. The case where v^n is equal to $\bar{v}^n$ is a straightforward. On the one hand the monotonicity assumption (3.12) on g and CFL condition (3.20) yield as usual

$$u_j^{n+1} = H_\lambda(u_{j-1}^n, u_j^n, u_{j+1}^n, v^n) \leq H_\lambda(\bar{u}_{j-1}^n, \bar{u}_j^n, \bar{u}_{j+1}^n, v^n) = \bar{u}_j^{n+1}.$$

On the other hand,

$$\bar{v}^{n+1} - v^{n+1} = \frac{\Delta t}{m_p}\big((g_\lambda^- - g_\lambda^+)(\bar{u}_0^n, \bar{u}_1^n, v^n) - (g_\lambda^- - g_\lambda^+)(u_0^n, u_1^n, v^n)\big)$$

is nonnegative by Hypothesis (3.19).

We now focus on the case where $(u_j^n)_{j\in\mathbb{Z}}$ is equal to $(\bar{u}_j^n)_{j\in\mathbb{Z}}$ and $v^n \leq \bar{v}^n$. For $j \leq -1$ and $j \geq 2$, a straightforward computation gives that there exists $a_{j+1/2}^n \in [u_j^n, u_{j+1}^n]$ and $b_{j-1/2}^n \in [u_{j-1}^n, u_j^n]$

$$\begin{aligned}
u_j^{n+1} - \bar{u}_j^{n+1} &= \mu \int_0^1 \partial_t g(u_j^n, u_{j+1}^n, v^n + t(\bar{v}^n - v^n)) - \partial_t g(u_{j-1}^n, u_j^n, v^n + t(\bar{v}^n - v^n)) dt \\
&= \mu \int_0^1 (\bar{v}^n - v^n) \left[\partial_3 g(u_j^n, u_{j+1}^n, v^n + t(\bar{v}^n - v^n)) \right. \\
&\qquad\qquad \left. - \partial_3 g(u_{j-1}^n, u_j^n, v^n + t(\bar{v}^n - v^n))\right] dt \\
&= \mu \int_0^1 (\bar{v}^n - v^n) \Big[\partial_{23} g(u_j^n, a_{j+1/2}^n, v^n + t(\bar{v}^n - v^n))(u_{j+1}^n - u_j^n) \\
&\qquad\qquad + \partial_{13} g(b_{j-1/2}^n, u_j^n, v^n + t(\bar{v}^n - v^n))(u_j^n - u_{j-1}^n)\Big]
\end{aligned}$$

Moreover, $u_{j-1}^n \leq u_j^n \leq u_{j+1}^n$ because we are considering elements of $\mathcal{S}$, thus if $\partial_3 g$ is decreasing toward its first two variables, $u_j^{n+1} \leq \bar{u}_j^{n+1}$. The same reasoning extend to $j \in \{0, 1\}$ because $u_0^n - u_1^n \leq \lambda$. Eventually,

$$\begin{aligned}
\bar{v}^{n+1} - v^{n+1} &= \bar{v}^n - v^n + \frac{\Delta t}{m_p}(g(u_0^n, u_1^n + \lambda, \bar{v}^n) - g(u_0^n, u_1^n + \lambda, v^n)) \\
&\qquad - \frac{\Delta t}{m_p}(g(u_0^n - \lambda, u_1^n, \bar{v}^n) - g(u_0^n - \lambda, u_1^n, v^n)) \\
&\geq \left(1 - \frac{2\Delta t}{m_p} \max |\partial_3 g|\right)(\bar{v}^n - v^n),
\end{aligned}$$

which is non negative if (3.42) holds. □

II

Couplage entre les équations d'Euler et une particule

Existence and uniqueness results

4

Dans ce chapitre, nous étudions un modèle plus réaliste de couplage entre un fluide, décrit ici par les équations d'Euler isothermes, et une particule ponctuelle. Comme dans le cas Burgers-particule, nous commençons par définir précisément le terme de frottement entre le fluide et la particule qui s'écrit comme un produit non conservatif. Le modèle est redéfini comme un problème d'interface, où des conditions de saut sont imposées au travers de la particule. Lorsque la force de frottement a de bonnes propriétés et lorsque la particule a une vitesse constante, nous montrons que ces conditions d'interface mènent à l'existence et l'unicité d'une solution autosemblable au problème de Riemann. Nous proposons ensuite une étude détaillée du cas où l'unicité est perdue.

Les parties 4.1 à 4.5 correspondent à l'article [Agu14c], à paraître dans le journal *Mathematical Models and Methods in Applied Sciences*. La dernière partie, qui ne figure pas dans cet article, est consacrée à la preuve de quelques résultats qualitatifs sur le système entièrement couplé, où le mouvement de la particule est décrit par une EDO.

Contents

4.1 Introduction

In this chapter we introduce and study a simple one-dimensional model of fluid-structure interaction. We consider a compressible and inviscid fluid, which will be governed by the isothermal Euler equations. At point x and time t, it has velocity $u(t,x)$ and density $\rho(t,x)$. The particle is assumed to be pointwise, of mass m, and we denote by $h(t)$, $h'(t)$ and $h''(t)$ its position, velocity and acceleration at time t . The interaction between the fluid and the particle is achieved through a drag force D, reflecting the fact that both tend to share the same velocity. Our model writes

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, & (4.1a) \\ \partial_t(\rho u) + \partial_x\left(\rho u^2 + p(\rho)\right) = -D(\rho, \rho(u - h'(t)))\delta_{h(t)}(x), & (4.1b) \\ mh''(t) = D(\rho(t,h(t)), \rho(t,h(t))[u(t,h(t)) - h'(t)]), & (4.1c) \\ (\rho(0,x), u(0,x)) = (\rho_0(x), u_0(x)), \\ (h(0), h'(0)) = (0, v_0), \end{cases}$$

where p is the pressure law. We suppose that D has the same sign as $u - h'$ (and therefore than $\rho(u - h')$), which is natural for a drag force. As a result, Equation (4.1c) shows formally that the particle accelerates (respectively decelerates) when the fluid's velocity $u(t, h(t))$ at its position is larger (respectively smaller) than its own velocity $h'(t)$. We choose the isothermal pressure law $p(\rho) = c^2\rho$ to avoid vacuum related issues, but the results could probably be extended without major difficulty in the adiabatic case $p(\rho) = \rho^\gamma, 1 < \gamma \leq 3$. This model is a generalization of the one first introduced in [LST08], and later deeply investigate in [AS12], [ALST10] and [ALST13], in which the fluid followed a scalar Burgers equation.

In the past few years, the interaction between an incompressible and viscous fluid and rigid bodies has been widely studied. Many papers deal with the existence of weak or strong solutions ([SMST02], [GLS00], [Fei03b], [DE99] and [VZ03]). The matter of collision has also been extensively investigated ([VZ06], [Sta04], [Hil07b], [HT09]). Some other works consider compressible fluid or elastic bodies ([BST12], [CDEG05], [DE00] and [Fei03a]). The model we study strongly differs from those works as the equation governing the fluid is inviscid. In this framework, the d'Alembert's paradox states that in an incompressible and inviscid fluid, with vanishing circulation, no drag force exerts on a body moving at a constant speed. Birds should therefore not be able to fly. An answer to this paradox is that, even at very high Reynolds number, the effect of viscosity cannot be neglected in a thin layer around the body. In our model, this paradox is somehow ignored, as we directly impose a drag force D between the fluid and the particle. According to Newton's law, the particle follows the ordinary differential equation (4.1c)

$$mh''(t) = D.$$

The action-reaction principle is taken into account in equation (4.1b) on the momentum of the Euler equation: the particle applies the force $-D$ on the fluid. We suppose that the interaction is local: it applies only at point $h(t)$. Equation (4.1a) ensures that the fluid mass is conserved. This approach proved to be successful in the toy model [LST08] with a Burgers' fluid. In particular, it allows collisions between two particles having different velocities, unlike in the viscous case [VZ06]. For example, a particle trapped in a shock (case V of Lemma 5.5 in [LST08]) will collide with a particle placed in front of the first particle and sharing the velocity of the fluid. In [ALST10], the reader can find numerical simulations of the drafting kissing tumbling phenomenon. This work is strongly related to [BCG14], in which a coupling between a particle and the Euler equations is presented. However, the modelization is quite different, the particle being taken into account through conservation of mass and energy, while we enforce in the present work a drag force. Moreover, the nature of the theoretical results are different and complementary. In [BCG14], a local in time existence to the Cauchy problem for the fully coupled system is proved for small subsonic data. In the present work, we consider the Riemann problem for a particle having constant velocity, without making any assumption on the data.

We emphasize the fact that in model (4.1), the particle and the fluid do not share the same velocity. We do not impose any no slip condition as in works presented above. It can be justified by saying that the particle is porous and allows some fluid to pass through. It constitutes the main difficulty of this model. Indeed, as shocks develop in finite time in hyperbolic systems like (4.1a)-(4.1b), even with $D = 0$, the velocity u and the density ρ of the fluid have no reason to be continuous along the particle trajectory h. A first consequence is that the source term in (4.1b) is not well defined. A second one is that the ODE (4.1c) the particle satisfies must be considered in a weak sense. This paper focuses on the Riemann problem for the uncoupled problem

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x\left(\rho u^2 + p(\rho)\right) = -D(\rho, \rho(u - h'(t)))\delta_{vt}, \\ \rho(0,x) = \rho_L \mathbf{1}_{x<0} + \rho_R \mathbf{1}_{x>0}, \\ u(0,x) = u_L \mathbf{1}_{x<0} + u_R \mathbf{1}_{x>0}, \end{cases} \tag{4.2}$$

where the particle has a constant speed v. The difficulties arising from the coupling between an ODE and a PDE disappear, but the key point of the nonconservative source term remains. Another main difficulty in the analysis of (4.2) is that the Dirac measure in the source term corresponds to a linearly degenerate field and our system is not hyperbolic. This may lead to resonance phenomena when two families of waves interact. Near resonance, Riemann problems with such source terms have been investigated in a conservative framework in [IT92] and later extended in the nonconservative framework in [GL04]. Away from resonance, the particle trajectory can be treated as a noncharacteristic boundary (see [BCG10], [BCG12] and [BCG14]). Our contribution is that, unlike in those frameworks, we solve the Riemann problem for all choices of parameters (ρ_L, u_L), (ρ_R, u_R) and v, without making any assumptions on their smallness or their resonant character.

Let us outline the organization of the paper and sketch the main results. The first section is entirely devoted to the definition of the solutions of (4.1). We exhibit an entropy condition that takes into account the particle. Then we give a rigorous definition of the nonconservative product

$$D(\rho, \rho(u - h'(t)))\delta_{h(t)}$$

based on a thickening of the particle. Replacing the Dirac measure by one of its approximation, it appears that the density and velocity of the fluid at the entry of the particle and at its exit are always linked by the same relations. These relations are independent of the size and the shape of the thickened particle. This allows us to see the particle as an interface, through which those relations are imposed. They link the quantities $u(t, h(t)^{\pm})$, $\rho(t, h(t)^{\pm})$ and $h'(t)$. A first relation states that the quantity $\alpha := \rho(u - h')$ is constant across the particle (this is why we express D as a function of ρ and α). This is equivalent to the conservation of the fluid's mass through the particle. Another relation describes the influence of the particle on the flow and depends on D. When the drag force is $D(\rho, \alpha) = \alpha$ and the particle is motionless, it expresses that the loss of charge $\rho u^2 + c^2 \rho$ through the particle is proportional to the mass flow ρu. The second section is devoted to the solution of the Riemann problem (4.2) for a particle moving at constant speed v. In Theorem 4.6, we exhibit a two conditions on the drag force D that imply that (4.2) has a unique self-similar entropy solution. The case of subsonic and supersonic initial datum are treated separately in Propositions 4.13 and 4.15. In Subsection 4.3.4 we describe the two natural asymptotics when the drag force vanishes or becomes very large. Eventually in Section 4.4, we discuss the case where the hypothesis of Theorem 4.6 are not fulfilled. We recover, in some particular cases, the existence of up to three solutions, as shown in a general framework in [IT92] and [GL04], and well known for fluid in a nozzle with discontinuous cross-section [LT07] and for the shallow water equation with discontinuous decreasing topography [LT03].

4.2 Definition of the solutions

This section is devoted to the definition of the solution of the coupled system (4.1). The isothermal Euler equations are inviscid, so ρ and u can be discontinuous along the particle's trajectory h and the product $D(\rho, \rho(h' - u))\delta_h(x)$ does not make sense. Following [LST08], we consider two different regularizations in Sections 4.2.1 and 4.2.2. The first one consists in adding a vanishing viscosity to the equation. Passing to the non-regularized limit we deduce an entropy inequality for (4.1). The second regularization is a thickening of the particle, which yields to an intrinsic definition of the non-conservative source term as an interface.

4.2.1 Entropy inequality

Let us first focus on the following classical regularization of problem (4.1), where we add a vanishing viscosity to the Euler equation:

$$\begin{cases} \partial_t \rho^\varepsilon + \partial_x q^\varepsilon = \varepsilon \partial_{xx} \rho^\varepsilon, \\ \partial_t q + \partial_x \left(\frac{(q^\varepsilon)^2}{\rho^\varepsilon} + c^2 \rho^\varepsilon \right) = -D(\rho^\varepsilon, \alpha^\varepsilon) \delta_{h(t)} + \varepsilon \partial_{xx} q^\varepsilon, \\ m h''(t) = D(\rho^\varepsilon(t, h(t)), \alpha^\varepsilon(t, h(t))), \\ (\rho^\varepsilon(0,x), q^\varepsilon(0,x)) = (\rho_0^\varepsilon(x), q_0^\varepsilon(x)), \\ (h(0), h'(0)) = (h_0, v_0). \end{cases} \tag{4.3}$$

Here, $q = \rho u$ denotes the momentum of the fluid, α denotes the quantity $\rho(u - h')$ and we only assume that the drag force D has the same sign than α (and hence than $u - h'$). In [LST08] and [Dom02] it is proven that the system

$$\begin{cases} \partial_t u^\varepsilon + \partial_x \frac{(u^\varepsilon)^2}{2} = \varepsilon \partial_{xx} u^\varepsilon - \lambda (u^\varepsilon - h') \delta_{h(t)}(x) \\ m h''(t) = \lambda (u^\varepsilon - h'), \\ u^\varepsilon(0,x) = u^0(x), \\ (h(0), h'(0)) = (h_0, v_0). \end{cases}$$

admits regular a solution when u^0 is regular. In order to derive an entropy inequality for our fluid particle coupling (4.1), we assume that if the initial data (ρ^0, u^0) are smooth and that the solutions of its regularization (4.3) are also smooth. Let $E(\rho, q)$ and $G(\rho, q)$ be a flux-entropy flux pair, with E convex, such that $\partial_q E$ is a function of $u = \frac{q}{\rho}$ denoted by J.

Example 1. The usual entropy-entropy flux pair

$$E(\rho, q) = \frac{q^2}{2\rho} + c^2 \rho \log(\rho) \quad \text{and} \quad G(\rho, q) = \frac{q}{\rho}(E(\rho, q) + c^2 \rho)$$

fulfills this assumption: we have $J(u) = u$.

For the sake of simplicity we introduce

$$U_0^\varepsilon = (\rho_0^\varepsilon, q_0^\varepsilon), \;\; U^\varepsilon = (\rho^\varepsilon, q^\varepsilon) \quad \text{and} \quad F(U^\varepsilon) = \left(q^\varepsilon, \frac{(q^\varepsilon)^2}{\rho^\varepsilon} + c^2 \rho^\varepsilon \right).$$

Let $\Phi \in \mathcal{C}_0^\infty(\mathbb{R}_+ \times \mathbb{R})$ be a *non-negative* smooth function, and multiply the first equation of (4.3) by $\Phi \partial_\rho E$ and the second equation by $\Phi \partial_q E = \Phi J$. Let us add the two equations and integrate over $\mathbb{R}_+ \times \mathbb{R}$. We obtain

$$\begin{aligned} &\iint_{\mathbb{R}_+ \times \mathbb{R}} \Phi \nabla_{(\rho,q)} E(U^\varepsilon) \cdot \partial_t U^\varepsilon \, dt\, dx + \iint_{\mathbb{R}_+ \times \mathbb{R}} \Phi \nabla_{(\rho,q)} E(U^\varepsilon) \cdot \partial_x F(U^\varepsilon) \, dt\, dx \\ &-\varepsilon \iint_{\mathbb{R}_+ \times \mathbb{R}} \Phi \nabla_{(\rho,q)} E(U^\varepsilon) \cdot \Delta U^\varepsilon \, dt\, dx = -\int_{\mathbb{R}_+} \left[D(\rho^\varepsilon, \alpha^\varepsilon) \Phi J(u^\varepsilon) \right] (t, h(t)) dt. \end{aligned} \tag{4.4}$$

We first treat the left hand side of (4.4) by integrating by parts. The first term gives

$$\begin{aligned} \iint_{\mathbb{R}_+ \times \mathbb{R}} \Phi \nabla_{(\rho,q)} E(U^\varepsilon) \cdot \partial_t U^\varepsilon \, dt\, dx &= \iint_{\mathbb{R}_+ \times \mathbb{R}} \Phi \partial_t E(U^\varepsilon) \, dt\, dx \\ &= -\iint_{\mathbb{R}_+ \times \mathbb{R}} E(U^\varepsilon) \partial_t \Phi \, dt\, dx - \int_{\mathbb{R}} \Phi(0, \cdot) E(U_0^\varepsilon) dx \end{aligned}$$

and the second term yields

$$\begin{aligned}\iint_{\mathbb{R}_+\times\mathbb{R}} \Phi\nabla_{(\rho,q)}E(U^\varepsilon)\cdot\partial_x F(U^\varepsilon)\,dt\,dx &= \iint_{\mathbb{R}_+\times\mathbb{R}} \Phi\nabla_{(\rho,q)}E(U^\varepsilon)\cdot[DF(U^\varepsilon)\partial_x U^\varepsilon]\;dt\,dx\\ &= \iint_{\mathbb{R}_+\times\mathbb{R}} \Phi\nabla_{(\rho,q)}G(U^\varepsilon)\cdot\partial_x U^\varepsilon\,dt\,dx\\ &= \iint_{\mathbb{R}_+\times\mathbb{R}} \Phi\partial_x G(U^\varepsilon)\,dt\,dx\\ &= -\iint_{\mathbb{R}_+\times\mathbb{R}} G(U^\varepsilon)\partial_x\Phi\,dt\,dx.\end{aligned}$$

Let us now tackle the third term. We have

$$\begin{aligned}\nabla_{(\rho,q)}E(U^\varepsilon)\cdot\Delta U^\varepsilon &= \sum_{i=1}^{2}\partial_i E(U^\varepsilon)\partial_{xx}U_i^\varepsilon\\ &= \sum_{i=1}^{2}\left[\partial_x(\partial_i E(U^\varepsilon)\partial_x U_i^\varepsilon) - \left(\sum_{j=1}^{2}\partial_{ji}E(U^\varepsilon)\partial_x U_j^\varepsilon\right)\partial_x U_i^\varepsilon\right]\\ &= \partial_{xx}E(U^\varepsilon) - \sum_{i=1}^{2}\sum_{j=1}^{2}(\partial_x U_j^\varepsilon)(\partial_{ji}E(U^\varepsilon))(\partial_x U_i^\varepsilon),\end{aligned}$$

thus we obtain

$$\begin{aligned}-\varepsilon\iint_{\mathbb{R}_+\times\mathbb{R}} \Phi\nabla_{(\rho,q)}E(U^\varepsilon)\cdot\Delta U^\varepsilon\,dt\,dx =& -\varepsilon\iint_{\mathbb{R}_+\times\mathbb{R}} \Phi E(U^\varepsilon)\partial_{xx}\Phi\,dt\,dx\\ &+\varepsilon\iint_{\mathbb{R}_+\times\mathbb{R}} \Phi\sum_{i=1}^{2}\sum_{j=1}^{2}(\partial_x U_j^\varepsilon)\partial_{ij}E(U^\varepsilon)(\partial_x U_i^\varepsilon)\,dt\,dx.\end{aligned}$$

Remark that as Φ is non-negative and E is convex, the last term is non-negative. We now focus on the right hand side of (4.15). Let us multiply the ODE in the third equation of (4.3) by $J(h'(t))\Phi(t,h(t))$. We have

$$\int_{\mathbb{R}_+} -mh''(t)J(h'(t))\Phi(t,h(t)) + D(\rho^\varepsilon,\alpha^\varepsilon)(t,h(t))J(h'(t))\Phi(t,h(t))dt = 0,$$

which reads, with P an antiderivative of J,

$$\int_{\mathbb{R}_+} -m[P(h'(t))]'\Phi(t,h(t)) + D(\rho^\varepsilon,\alpha^\varepsilon)(t,h(t))J(h'(t))\Phi(t,h(t))dt = 0.$$

We can therefore replace the right hand side of (4.15) by

$$\int_{\mathbb{R}_+} mP(h'(t))\partial_t(\Phi(t,h(t)))dt + \int_{\mathbb{R}_+} D(\rho^\varepsilon,\alpha^\varepsilon)\left(J(h'(t)) - J(u^\varepsilon)\right)\Phi(h)dt + mP(v_0)\Phi(0,h_0).$$

The function $J(u) = \partial_q E(1,u)$ is nondecreasing as the restriction of E to the line $\rho = 1$ is convex. Moreover, D has the same sign as $u^\varepsilon - h'$. Thus the second term is non-positive. To conclude, we add the different terms and drop the two non-positive ones to obtain

$$\begin{aligned}\iint_{\mathbb{R}_+\times\mathbb{R}} E(U^\varepsilon)\partial_t\Phi\,dt\,dx + \iint_{\mathbb{R}_+\times\mathbb{R}} G(U^\varepsilon)\partial_x\Phi\,dt\,dx + \int_{\mathbb{R}_+} mP(h'(t))\partial_t(\Phi(t,h(t)))dt&\\ + \int_{\mathbb{R}} \Phi(0,\cdot)E(U_0^\varepsilon)\,dx + mP(v_0)\Phi(0,h_0) \geq -\varepsilon\iint_{\mathbb{R}_+\times\mathbb{R}} \Phi E(U^\varepsilon)\partial_{xx}\Phi\,dt\,dx.&\end{aligned}$$

Last, we formally pass to the limit as $\varepsilon \longrightarrow 0$ and get the following entropy inequality for the coupled problem (4.1):

$$\begin{aligned} \iint_{\mathbb{R}_+\times\mathbb{R}} E(U)\partial_t\Phi\,dt\,dx + \iint_{\mathbb{R}_+\times\mathbb{R}} G(U)\partial_x\Phi\,dt\,dx + \int_{\mathbb{R}_+} mP(h'(t))\partial_t(\Phi(t,h(t)))dt \\ + \int_{\mathbb{R}} \Phi(0,\cdot)E(U_0)\,dx + mP(v_0)\Phi(0,h_0) \geq 0. \end{aligned} \tag{4.5}$$

Remark 4. When the test function Φ is supported on $\{(t,x),\, x > h(t)\}$ or on $\{(t,x),\, x < h(t)\}$, the inequality (4.5) reduces to the classical entropy inequality for the isothermal Euler equation without source term.

4.2.2 How to handle the nonconservative product

In this section we assume that the particle trajectory h is given, and more precisely that it moves at constant speed v: $h(t) = vt$. We denote by H the Heaviside function $H(x) = \mathbf{1}_{x>0}$. Introducing the momentum $q = \rho u$ and the new unknown $w(t,x) = H(x - vt)$ allows us to rewrite the system (4.1a)-(4.1b) in the framework of hyperbolic equation,

$$\begin{cases} \partial_t\rho + \partial_x q = 0, \\ \partial_t q + \partial_x\left(\frac{q^2}{\rho} + c^2\rho\right) + D(\rho,\alpha)\partial_x w = 0, \\ \partial_t w + v\partial_x w = 0. \end{cases} \tag{4.6}$$

Its quasilinear form is

$$\partial_t \begin{pmatrix} \rho \\ q \\ w \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 \\ c^2 - u^2 & 2u & D(\rho,\alpha) \\ 0 & 0 & v \end{pmatrix} \partial_x \begin{pmatrix} \rho \\ q \\ w \end{pmatrix} = 0. \tag{4.7}$$

The eigenvalues of the Jacobian matrix are $u+c$, $u-c$ and v. This system is strictly hyperbolic whenever $u \neq v \pm c$. In the resonant case $u \pm c = v$, the matrix cannot be put in a diagonal form. Such resonant systems have been studied in [GL04] and [IT92]. However, with this source term, our system does not fall neither in the framework of [IT92], because it is not conservative, neither in the framework of [GL04], because one of the hypothesis on the source term (namely 1.7) is not satisfied when the drag force depends only on α. Following [LST08], [GL04], [IT92], [SV03] and [CLS04], we use a thickening of the particle to define the nonconservative product. Let H^ε be an approximation of the Heaviside function such that:

- $H^\varepsilon \in \mathcal{C}^0(\mathbb{R}) \cap \mathcal{C}^1((-\varepsilon/2, \varepsilon/2))$;
- H^ε is nondecreasing;
- $H^\varepsilon(x) = 0$ if $x \leq -\varepsilon/2$ and $H^\varepsilon(x) = 1$ if $x \geq \varepsilon/2$.

We replace the Dirac measure by its regularization $(H^\varepsilon)'$ to obtain the regularized system

$$\begin{cases} \partial_t\rho^\varepsilon + \partial_x q^\varepsilon & = 0; \\ \partial_t q^\varepsilon + \partial_x\left(\frac{(q^\varepsilon)^2}{\rho^\varepsilon} + c^2\rho^\varepsilon\right) & = -D(\rho^\varepsilon,\alpha^\varepsilon)(H^\varepsilon)'(x - vt). \end{cases} \tag{4.8}$$

We are interesting in what is happening inside the particle. In the spirit of traveling waves, we look for solution only depending on $x - vt$. With such a regularized source term, the values of the solutions at the extremities of the particle depend neither on the size of the thickened particle ε, nor on the choice of the regularization H^ε (satisfying the three hypotheses above). This allows us to define the source term $D(\rho,\alpha)\delta_{h(t)}$ as an interface condition.

In the sequel, for $\alpha \neq 0$, we denote by F_α the function

$$F_\alpha(\rho) = \int_{\frac{|\alpha|}{c}}^{\rho} \frac{1}{|D(r,\alpha)|} \left(-\frac{\alpha^2}{r} + c^2 \right) dr.$$

Remark that F_α decreasing on $\left(0, \frac{|\alpha|}{c}\right)$ and increasing on $\left(\frac{|\alpha|}{c}, +\infty\right)$.

Lemma 4.1. *Let $(\rho^\varepsilon, q^\varepsilon)$ be a piecewise $\mathcal{C}^1$ solution of* (4.8) *that only depends on $\xi = x - vt$ and defined for ξ in $[-\varepsilon/2, \varepsilon/2]$. Then, on every interval $\mathcal{I}$ where the solution is smooth, the quantity $\alpha^\varepsilon = q^\varepsilon - v\rho^\varepsilon$ remains constant. If $\alpha^\varepsilon = 0$ on $\mathcal{I}$, the density also remains constant, while if $\alpha^\varepsilon \neq 0$, the evolution of ρ^ε on $\mathcal{I}$ is given by*

$$(F_{\alpha^\varepsilon}(\rho^\varepsilon(\xi)))' = -\operatorname{sign}(\alpha^\varepsilon)(H^\varepsilon)'. \tag{4.9}$$

If the solution is discontinuous at a point $\xi^0 \in (-\varepsilon/2, \varepsilon/2)$, then

$$\begin{cases} \alpha^\varepsilon(\xi^0_-) = \alpha^\varepsilon(\xi^0_+) := \alpha^\varepsilon(\xi^0)\,; \\ \left(\dfrac{(\alpha^\varepsilon(\xi^0))^2}{\rho^\varepsilon(\xi^0_-)} + c^2\rho^\varepsilon(\xi^0_-) \right) - \left(\dfrac{(\alpha^\varepsilon(\xi^0))^2}{\rho^\varepsilon(\xi^0_+)} + c^2\rho^\varepsilon(\xi^0_+) \right) = 0 \end{cases} \tag{4.10}$$

Proof. Let $\rho^\varepsilon(x - vt)$ and $q^\varepsilon(x - vt)$ be a piecewise $\mathcal{C}^1$ solution of (4.8), only depending on $\xi = x - vt$. If the solution is smooth on the interval $\mathcal{I}$, it satisfies the following equations:

$$\begin{cases} -v(\rho^\varepsilon)' + (q^\varepsilon)' & = 0, \\ -v(q^\varepsilon)' + \left(\frac{(q^\varepsilon)^2}{\rho^\varepsilon} + c^2\rho^\varepsilon \right)' & = -D(\rho^\varepsilon, \alpha^\varepsilon)(H^\varepsilon)'(x - vt). \end{cases} \tag{4.11}$$

The first equation of (4.11) directly gives that α^ε remains constant on $\mathcal{I}$. Replacing q^ε by $\alpha^\varepsilon + v\rho^\varepsilon$ in the second line of (4.11) yields

$$-v^2(\rho^\varepsilon)' + \left(\frac{(\alpha^\varepsilon)^2 + 2\alpha^\varepsilon v\rho^\varepsilon + v^2(\rho^\varepsilon)^2}{\rho^\varepsilon} + c^2\rho^\varepsilon \right)' = -D(\rho^\varepsilon, \alpha^\varepsilon)(H^\varepsilon)'.$$

As α^ε and v are constant, this expression simplifies in

$$\left(-\frac{(\alpha^\varepsilon)^2}{(\rho^\varepsilon)^2} + c^2 \right)(\rho^\varepsilon)' = -D(\rho^\varepsilon, \alpha^\varepsilon)(H^\varepsilon)',$$

which rewrites, by definition of F_α,

$$[F_{\alpha^\varepsilon}(\rho^\varepsilon(\xi))]' = \frac{1}{|D(\rho^\varepsilon, \alpha^\varepsilon)|} \left(-\frac{(\alpha^\varepsilon)^2}{(\rho^\varepsilon)^2} + c^2 \right)(\rho^\varepsilon)' = -\operatorname{sign}(\alpha^\varepsilon)(H^\varepsilon)'.$$

On the other hand, if $(\rho^\varepsilon, q^\varepsilon)$ has a discontinuity at a point $\xi^0 \in (-\varepsilon/2, \varepsilon/2)$, it verifies the two relations:

$$\begin{cases} q^\varepsilon(\xi^0_+) - q^\varepsilon(\xi^0_-) = v(\rho^\varepsilon(\xi^0_+) - \rho^\varepsilon(\xi^0_-)), \\ \left(\dfrac{(q^\varepsilon(\xi^0_+))^2}{\rho^\varepsilon(\xi^0_+)} + c^2\rho^\varepsilon(\xi^0_+) \right) - \left(\dfrac{(q^\varepsilon(\xi^0_-))^2}{\rho^\varepsilon(\xi^0_-)} + c^2\rho^\varepsilon(\xi^0_-) \right) = v(q^\varepsilon(\xi^0_+) - q^\varepsilon(\xi^0_-)), \end{cases}$$

and we obtain the result (4.10) by introducing the conserved quantity $\alpha(\xi^0) = q^\varepsilon(\xi^0_-) - v\rho^\varepsilon(\xi^0_-) = q^\varepsilon(\xi^0_+) - v\rho^\varepsilon(\xi^0_+)$. □

Remark 5. The relations (4.10) are nothing but the Rankine–Hugoniot relations for a shock having speed v in the isothermal Euler equations. The lemma below states that entropy shocks only link supersonic states to subsonic states (from left to right if $\alpha > 0$, from right to left if $\alpha < 0$).

Lemma 4.2. *The shock corresponding to the Rankine–Hugoniot relations* (4.10) *is an entropy satisfying shock in the Euler equations (without source term) for the entropy-entropy flux pair of Example 1 if and only if* $\alpha^\varepsilon(\xi^0) > 0$ *and* $\alpha^\varepsilon(\xi^0) > c\rho^\varepsilon(\xi^0_-)$ *or if* $\alpha^\varepsilon(\xi^0) < 0$ *and* $\alpha^\varepsilon(\xi^0) < -c\rho^\varepsilon(\xi^0_+)$.

Proof. Suppose that $\alpha^\varepsilon(\xi^0)$ is positive. If $c\rho^\varepsilon(\xi^0_-) > \alpha^\varepsilon$ then $\rho^\varepsilon(\xi^0_-) > \rho^\varepsilon(\xi^0_+)$. If the shock was a Lax shock, it should be a 2-shock. Therefore we should have

$$v = \frac{q^\varepsilon(\xi^0_+)}{\rho^\varepsilon(\xi^0_+)} + c\sqrt{\frac{\rho^\varepsilon(\xi^0_-)}{\rho^\varepsilon(\xi^0_+)}},$$

which rewrites

$$c\sqrt{\frac{\rho^\varepsilon(\xi^0_-)}{\rho^\varepsilon(\xi^0_+)}} = -\frac{\alpha^\varepsilon}{\rho^\varepsilon(\xi^0_+)},$$

and contradicts the fact that $\alpha^\varepsilon > 0$. On the other hand if $c\rho^\varepsilon(\xi^0_-) < \alpha^\varepsilon$, we have that $\rho^\varepsilon(\xi^0_-) < \rho^\varepsilon(\xi^0_+)$, and the discontinuity should be a 1-shock. We obtain

$$c\sqrt{\frac{\rho^\varepsilon(\xi^0_-)}{\rho^\varepsilon(\xi^0_+)}} = \frac{\alpha^\varepsilon}{\rho^\varepsilon(\xi^0_+)},$$

which does not contradict the sign of α^ε. Similarly, we obtain that if α is negative, the jump is an entropy satisfying shock if and only if $c\rho^\varepsilon(\xi^0_+) < |\alpha^\varepsilon|$.

□

Definition 4.1. *The germ at speed* v *is the set* $\mathcal{G}_D(v)$ *containing all the pairs* $((\rho_-, q_-), (\rho_+, q_+))$ *of* $(\mathbb{R}^*_+ \times \mathbb{R})^2$ *verifying the two following relations.*

1. *First,*

$$q_- - v\rho_- = q_+ - v\rho_+.$$

We denote by α *this quantity.*

2. *Second,*

- *either* $\alpha = 0$, $\rho_- = \rho_+$ *and* $q_- = q_+$;
- *or* $0 < \alpha$, $(\frac{\alpha}{c} - \rho_+)(\frac{\alpha}{c} - \rho_-) \geq 0$ *and*

$$F_\alpha(\rho_-) - F_\alpha(\rho_+) = 1;$$

- *or* $\alpha < 0$, $(\frac{|\alpha|}{c} - \rho_+)(\frac{|\alpha|}{c} - \rho_-) \geq 0$ *and*

$$F_\alpha(\rho_+) - F_\alpha(\rho_-) = 1;$$

- *or* $c\rho_- < \alpha$, $\rho_+ \geq \frac{\alpha}{c}$, *and there exists* $\rho \in (\rho_-, \frac{\alpha}{c})$ *and* $\theta \in [0, 1]$ *such that*

$$\begin{cases} F_\alpha(\rho_-) - F_\alpha(\rho) & = \theta, \\ F_\alpha(\frac{\alpha^2}{c^2\rho}) - F_\alpha(\rho_+) & = (1-\theta); \end{cases}$$

- *or* $\alpha < -c\rho_+$, $\rho_- \geq \frac{|\alpha|}{c}$ *and there exists* $\rho \in (\rho_+, \frac{|\alpha|}{c})$ *and* $\theta \in [0, 1]$ *such that*

$$\begin{cases} F_\alpha(\rho_+) - F_\alpha(\rho) & = \theta, \\ F_\alpha(\frac{\alpha^2}{c^2\rho}) - F_\alpha(\rho_-) & = (1-\theta). \end{cases}$$

Theorem 4.3. *Suppose that $((\rho_-, q_-), (\rho_+, q_+))$ belongs to $\mathcal{G}_D(v)$. Then for all positive ε and for all regularization H^ε fulfilling the hypothesis of* (4.8)*, there exists a piecewise $\mathcal{C}^1$ entropy solution of* (4.8) *only depending on $\xi = x - vt$, such that $(q^\varepsilon(-\varepsilon/2), \rho^\varepsilon(-\varepsilon/2)) = (q_-, \rho_-)$ and $(q^\varepsilon(\varepsilon/2), \rho^\varepsilon(\varepsilon/2)) = (q_+, \rho_+)$. By entropy solution, we mean that each discontinuity in the solution corresponds to a entropy satisfying shock in the Euler equations.*

Conversly, if (ρ_-, q_-) and (ρ_+, q_+) are the values in $-\varepsilon/2$ and $\varepsilon/2$ of such a solution of (4.8)*, then they verify the two relations stated above.*

Proof. Let $(\rho^\varepsilon, q^\varepsilon)$ be a solution of (4.8) which depends only on $\xi = x - vt$, is piecewise $\mathcal{C}^1$ and whose discontinuities are entropy shocks. A straightforward consequence of Lemma 4.1 is that the quantity α^ε (equals to $\rho^\varepsilon(u^\varepsilon - v)$) is constant on the whole interval $[-\varepsilon/2, \varepsilon/2]$. In the sequel we suppose that α^ε is positive. In that case the fluid's velocity u^ε is everywhere larger than the particle's velocity v and the "entry" of the particle is on its right at $\xi = -\varepsilon/2$. We fix a state $(\rho_-, q_-) \in \mathbb{R}_+^* \times \mathbb{R}$ at the entry of the particle, and look for the $(\rho_+, q_+) \in \mathbb{R}_+^* \times \mathbb{R}$ at its exit. The reasoning is the same with $\alpha^\varepsilon < 0$ (and trivial if $\alpha^\varepsilon = 0$), but the entry of the particle is on its right and it is more convenient to fix the state (ρ_+, q_+).

There is only one solution which is smooth on the entire interval $[-\varepsilon/2, \varepsilon/2]$. We integrate (4.9) on this interval to obtain

$$F_{\alpha^\varepsilon}(\rho_-) - F_{\alpha^\varepsilon}(\rho_+) = 1.$$

As depicted on Figure 4.1, the function F_{α^ε} decreases on $(0, \frac{\alpha^\varepsilon}{c})$ and increases on $(\frac{\alpha^\varepsilon}{c}, +\infty)$. Its minimum is reached for $\frac{\alpha^\varepsilon}{c}$. Moreover, the regularization H^ε of the Heaviside function is increasing. As a consequence, $\xi \mapsto F_{\alpha^\varepsilon}(\rho^\varepsilon(\xi))$ decreases, and ρ^ε cannot cross continuously α^ε/c. On its interval of smoothness, a solution of (4.8) is always subsonic (i.e $|u^\varepsilon - v| \leq c$ or equivalently $c\rho^\varepsilon \geq \alpha^\varepsilon$) or always supersonic (i.e $|u^\varepsilon - v| \geq c$ or equivalently $c\rho^\varepsilon \leq \alpha^\varepsilon$). This explains the condition $(\frac{\alpha}{c} - \rho_+)(\frac{\alpha}{c} - \rho_-) \geq 0$ in the second point of 2 in Definition 4.1.

On the other hand by Lemma 4.2, a discontinuity at a point ξ^0 is entropy satisfying if and only if $\alpha^\varepsilon(\xi^0) > c\rho^\varepsilon(\xi^0_-)$. Therefore a solution has no discontinuity if $\alpha_- \leq c\rho_-$ and has at most one discontinuity if $\alpha_- > c\rho_-$. We focus on this last case. The solution is smooth on $(-\varepsilon/2, \xi^0)$ and (4.9) yields:

$$F_\alpha(\rho_-) - F_\alpha(\rho^\varepsilon(\xi^0_-)) = H^\varepsilon(\xi^0).$$

There is a shock in ξ^0, and Rankine–Hugoniot relations (4.10) imply that

$$\rho^\varepsilon(\xi^0_-)\rho^\varepsilon(\xi^0_+) = \frac{(\alpha^\varepsilon)^2}{c^2}.$$

In particular, $c\rho^\varepsilon(\xi^0_+) \geq \alpha^\varepsilon$ and there is no shock after ξ^0. We integrate (4.9) on $(\xi^0, \varepsilon/2)$ to get

$$F_\alpha(\rho^\varepsilon(\xi^0_+)) - F_\alpha(\rho^\varepsilon_+) = 1 - H^\varepsilon(\xi^0).$$

We obtain the third point with $\rho = \rho^\varepsilon(\xi^0_-)$ and $\theta = H^\varepsilon(\xi^0)$. The two types of solutions, continuous everywhere or containing a single entropy shock, are described on Figure 4.1. □

In the following Corollary we extract properties of the germ $\mathcal{G}_D(v)$.

Corollary 4.4. *If the pair $((\rho_-, q_-), (\rho_+, q_+))$ belongs to the germ $\mathcal{G}_D(v)$, then we necessary have:*

- $\alpha_- = \alpha_+ := \alpha$;
- *if $\alpha > 0$ and $c\rho_- \geq \alpha$, then $c\rho_- \geq c\rho_+ \geq \alpha$;*
- *if $\alpha < 0$ and $c\rho_+ \geq |\alpha|$, then $c\rho_+ \geq c\rho_- \geq |\alpha|$;*

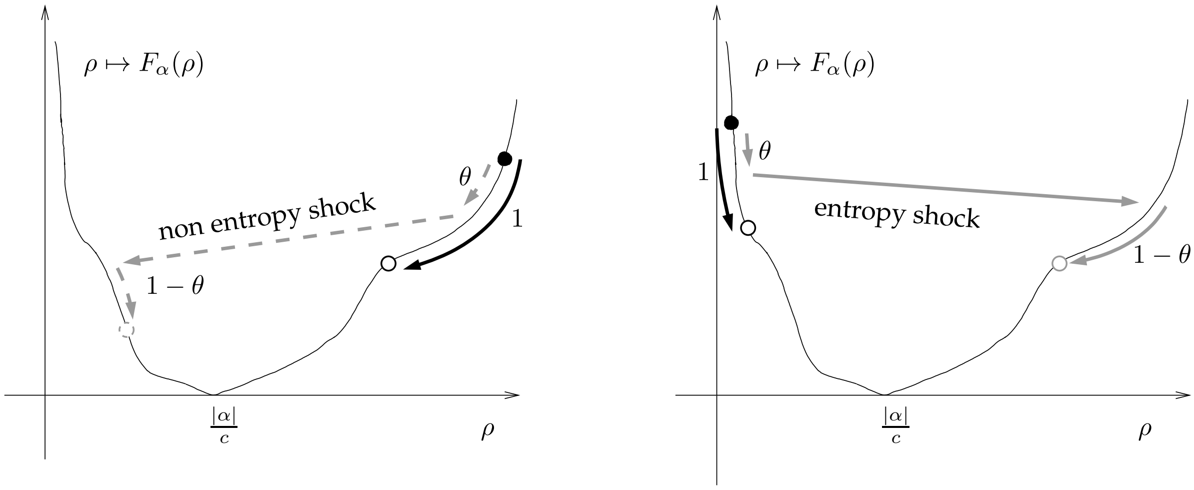


Figure 4.1. *How to reach the density ρ_+ (white circles) from the density ρ_- (black circles) when α is positive. On the left is the supersonic case $c\rho_- < \alpha$; on the right is the subsonic case $\alpha \leq c\rho_-$.*

Proof. Suppose that α is positive. We already emphasized in the proof of Theorem 4.3 that if the velocity at the entry of the particle is subsonic (i.e. $c\rho_- \geq \alpha$) then there is no discontinuity in the solution and the solution remains subsonic: $c\rho_- \geq \alpha$. In that case, $F_\alpha(\rho_-) - F_\alpha(\rho_+) = 1$, and as F_α increases on $(\alpha/c, +\infty)$ we obtain that $\rho_- \geq \rho_+$.

□

Remark 6. When the drag force depends only on α,

$$F_\alpha(\rho) = \frac{1}{|D(\rho,\alpha)|}\left(\frac{\alpha^2}{\rho} + c^2\rho\right)$$

for some real C, and F_α has the remarkable property of being compatible with shocks at speed v in the Euler equations:

$$\forall \alpha \neq 0,\, \forall \rho > 0,\; F_\alpha\left(\frac{\alpha^2}{c^2\rho}\right) = F_\alpha(\rho).$$

It follows that a shock corresponds to a horizontal jump on the graph of F_α, which is not the case in general. A consequence is that the right state (ρ_+, q_+) is the same whatever the value of θ is. The contrast between the two situations is depicted on Figure 4.2 below. In that case, the germ $\mathcal{G}_D(v)$ can be described more concisely: the second point of Definition 4.1 becomes

$$\left(\frac{\alpha^2}{\rho_-} + c^2\rho_-\right) - \left(\frac{\alpha^2}{\rho_+} + c^2\rho_+\right) = D(\rho,\alpha), \tag{4.12}$$

and we still have the two inequalities of Corollary 4.4

Remark 7. System (4.8) may not have any solution continuous on the whole interval $[-\varepsilon/2, \varepsilon/2]$ if ρ_- is too close from α/c when $\alpha > 0$. In the case of the linear drag force $D(\rho,\alpha) = \lambda\alpha$, where $\lambda \geq 0$ is a friction coefficient, an explicit computation shows that when α is positive there is no solution ρ_- belongs to

$$\left[\frac{\alpha}{c} + \frac{\lambda\alpha - \alpha\sqrt{4c\lambda+\lambda^2}}{2c^2}, \frac{\alpha}{c} + \frac{\lambda\alpha + \alpha\sqrt{4c\lambda+\lambda^2}}{2c^2}\right].$$

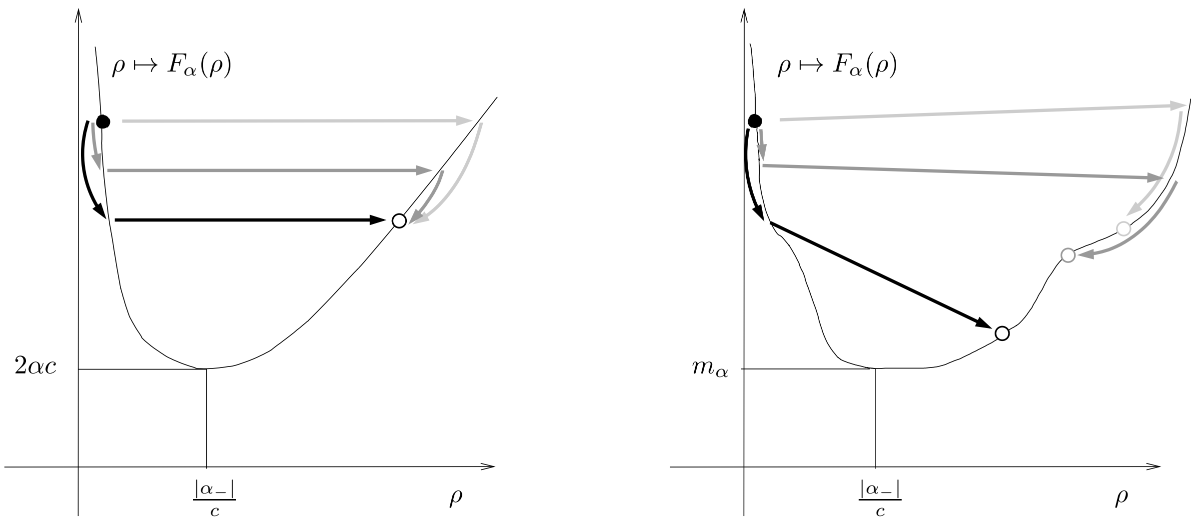


Figure 4.2. *Densities ρ_+ (white dots) accessible from ρ_- with $\theta = 0$ (light grey) $0 < \theta < 1$ (medium grey) and $\theta = 1$ (black). On the left, the drag force depends only on α, while on the right, it also depends on ρ.*

4.2.3 Definition of the solution

Let us now reformulate the ODE (4.1c). The source term in (4.1b) is the exact opposite of the left hand side in (4.1c), so the total impulsion is formally conserved through time:

$$\frac{d}{dt}\left[\int_{\mathbb{R}} \rho(t,x)u(t,x)dx + mh'(t)\right] = 0. \tag{4.13}$$

We can use this additional property of the model to give a precise definition of the ODE (4.1c).

Proposition 4.5. *Let (ρ, u) be a solution of* (4.1a)-(4.1b) *such that for almost every $t > 0$, the traces around the particle exist and are such that*

$$((\rho_-(t), q_-(t)), (\rho_+(t), q_+(t))) \in \mathcal{G}_D(h'(t)).$$

Then, it satisfies the conservation of total impulsion (4.13) *if and only if for almost every $t > 0$,*

$$mh''(t) = c^2(\rho_-(t) - \rho_+(t))\left(1 - \frac{(u_-(t) - h'(t))(u_+(t) - h'(t))}{c^2}\right), \tag{4.14}$$

where the subscripts $\pm$ indicates the left and right traces around the particle: $\rho_\pm(t) = \rho(t, h(t)_\pm)$.

Proof. If $((\rho_-(t), q_-(t)), (\rho_+(t), q_+(t)))$ belongs to $\mathcal{G}_D(h'(t))$, then $q_-(t) - h'(t)\rho_-(t)$ is equal to $q_+(t) -$

$h'(t)\rho_+(t)$. As usual, we denote by $\alpha(t)$ this quantity, and replace $q = \rho u$ by $\alpha + h'\rho$. We have

$$\begin{aligned}
mh''(t) &= -\partial_t \int_{\mathbb{R}} \rho(t,x)u(t,x)dx \\
&= -\partial_t \int_{-\infty}^{h(t)} \rho(t,x)u(t,x)dx - \partial_t \int_{h(t)}^{+\infty} \rho(t,x)u(t,x)dx \\
&= h'(\rho_+u_+ - \rho_-u_-) + (\rho_-u_-^2 + c^2\rho_-) - (\rho_+u_+^2 + c^2\rho_+) \\
&= h'^2(\rho_+ - \rho_-) + \rho_-\left(\frac{\alpha^2}{\rho_-^2} + 2h'\frac{\alpha}{\rho_-} + h'^2\right) - \rho_+\left(\frac{\alpha^2}{\rho_+^2} + 2h'\frac{\alpha}{\rho_+} + h'^2\right) + c^2(\rho_- - \rho_+) \\
&= h'^2(\rho_+ - \rho_-) + \alpha^2\left(\frac{1}{\rho_-} - \frac{1}{\rho_+}\right) + h'^2(\rho_- - \rho_+) + c^2(\rho_- - \rho_+) \\
&= (\rho_- - \rho_+)\left(c^2 - \frac{\alpha^2}{\rho_-\rho_+}\right) \\
&= c^2(\rho_- - \rho_+)\left(1 - \frac{(u_- - h')(u_+ - h')}{c^2}\right).
\end{aligned}$$

□

Remark 8. The ODE (4.14) does not seem to depend on the drag force: this dependence is hidden in the assumption

$$((\rho_-(t), q_-(t)), (\rho_+(t), q_+(t))) \in \mathcal{G}_D(h'(t))$$

because the germ does depend on the drag force.

Remark 9. When the drag force depends only on α, which is conserved through the particle, the initial ODE (4.1c) makes sense. In that case, it is not difficult to use the more concise description of the germ (4.12) to prove that (4.1c) and (4.14) are equivalent.

Thanks to the previous reformulation of the ODE and on the entropy inequality (4.5), we define the entropy solutions of the coupled problem (4.1):

Definition 4.2. *Assume that $(\rho_0, q_0) \in L^\infty(\mathbb{R})^2$ and $v_0 \in \mathbb{R}$. A triplet $(\rho, q, h) \in L^\infty(\mathbb{R}_+ \times \mathbb{R}) \times L^\infty(\mathbb{R}_+ \times \mathbb{R}) \times W^{2,\infty}_{loc}(\mathbb{R}_+)$ is called an entropy solution of the problem* (4.1) *if:*

- *(ρ, q) is a weak solution of the isothermal Euler equations on the sets $\{(t,x) \in \mathbb{R}_+^* \times \mathbb{R} : x > h(t)\}$ and $\{(t,x) \in \mathbb{R}_+^* \times \mathbb{R} : x < h(t)\}$.*
- *For any entropy-entropy flux pair (E, G) such that E is convex and $\partial_q E(\rho, q) = J\left(\frac{q}{\rho}\right)$, for any non-negative test function $\Phi \in \mathcal{C}_0^\infty(\mathbb{R}_+ \times \mathbb{R})$, we have*

$$\begin{aligned}\iint_{\mathbb{R}_+\times\mathbb{R}} E(U)\partial_t\Phi\, dt\, dx + \iint_{\mathbb{R}_+\times\mathbb{R}} G(U)\partial_x\Phi\, dt\, dx + \int_{\mathbb{R}_+} mP(h'(t))\partial_t(\Phi(t,h(t)))dt \\ + \int_{\mathbb{R}} \Phi(0,\cdot)E(U_0)\, dx + mP(v_0)\Phi(0,v_0) \quad \geq 0\end{aligned} \tag{4.15}$$

  where P is a given antiderivative of J;
- *For almost every $t > 0$, the traces around the particle exist and belong to the germ at speed $h'(t)$:*

$$((\rho(t,h(t)_-), q(t,h(t)_-)), (\rho(t,h(t)_+), q(t,h(t)_+))) \in \mathcal{G}_D(h'(t));$$

- *For almost every $t > 0$, the particle is driven by the ODE:*

$$mh''(t) = c^2(\rho_-(t) - \rho_+(t))\left(1 - \frac{(u_-(t) - h'(t))(u_+(t) - h'(t))}{c^2}\right).$$

Remark 10. The entropy inequality (4.15) implies that the solution is an entropy solution of the Euler equations on the sets $\{x < h(t)\}$ and $\{x > h(t)\}$. Moreover, when the test function Φ tends to a Dirac measure at $(t, h(t))$ for which the traces exist, (4.15) yields that

$$h'(t)(E(U_-) - E(U_+)) - (G(U_-) - G(U_+)) + mh''(t)P'(h(t)) \leq 0 \tag{4.16}$$

where we denote by $U_\pm$ the left and right traces around the particle:

$$U_- = (\rho(t, h(t)_-), q(t, h(t)_-)).$$

In other words, the total energy is dissipated through the particle. This property is consistent with the Definition 4.1 of the germ and the ODE (4.14). Indeed, if we introduce α and replace h'' by its expression in (4.14), Equation (4.16) becomes

$$\alpha \left[\alpha^2 \frac{\rho_-^2 - \rho_+^2}{2\rho_-^2\rho_+^2} + c^2(\ln(\rho_+) - \ln(\rho_-))\right] \leq 0.$$

This holds true if (U_-^0, U_+^0) belongs to $\mathcal{G}_D(h'(t^0))$. This can be checked by treating sperately the subsonic case and the supersonic case. In the latter case, we have to use that $\rho_+ \leq \frac{\alpha^2}{c^2\rho_-}$ which is easily deduced from Definition 4.1.

Remark 11. In the case of the scalar conservation law, it is not necessary to require the existence of the traces. Indeed, as the solution is a classic solution on $\{x < h'\}$, strong traces exist (see [Pan07] and [Vas01]). However, such a result is much harder to obtain in the system case.

4.3 Riemann problem for a particle with a constant fixed velocity

In this section we focus on the uncoupled problem where the particle has a constant speed equal to some given v in $\mathbb{R}$. Moreover, we consider a class of very specific initial datum, which consists in piecewise constant functions for the density ρ and for the momentum $q = \rho u$, with a single discontinuity falling exactly on the initial position of the particle. The problem under study in this section is the Riemann problem:

$$\begin{cases} \partial_t \rho + \partial_x q & = 0, \\ \partial_t q + \partial_x \left(\frac{q^2}{\rho} + c^2\rho\right) & = -D(\rho, \alpha)\delta_{vt}, \\ \rho(0, x) & = \rho_L \mathbf{1}_{x<0} + \rho_R \mathbf{1}_{x>0}, \\ q(0, x) & = q_L \mathbf{1}_{x<0} + q_R \mathbf{1}_{x>0}, \end{cases} \tag{4.17}$$

where (ρ_L, q_L) and (ρ_R, q_R) belong to $\mathbb{R}_+^* \times \mathbb{R}$. We recall once for all that α denotes the quantity $q - v\rho$. As the particle's trajectory is a straight line, we look for self-similar solutions of (**??**), i.e. solutions that only depend on x/t. This section is devoted to the proof of the following theorem.

Theorem 4.6. *Consider a drag force*

$$\begin{array}{rccc} D: & \mathbb{R}_+^* \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ & (\rho, \alpha) & \mapsto & D(\rho, \alpha) \end{array}$$

having the same sign as α, vanishing in $\alpha = 0$ and $\mathcal{C}^1$. Suppose that D is an increasing function of α and that $|D|$ is a decreasing function of ρ. Then for all states (ρ_L, q_L) and (ρ_R, q_R) in $\mathbb{R}_+^* \times \mathbb{R}$ and for every particle velocity v in $\mathbb{R}$, the Riemann problem (**??**) *has a unique self-similar solution.*

Example 2. All frictions of the form

$$\Gamma(\rho, u, h') = \rho^n |u - h'|^{m-1}(u - h') = \rho^{n-m}|\alpha|^{m-1}\alpha$$

with $n \geq 0$, $m \geq 1$ and $m \geq n$, verify the two conditions of Theorem 4.6.

Let (ρ, q) be a self-similar solution of (**??**) and denote by (ρ_-, q_-) (respectively (ρ_+, q_+)) the traces of the density and the momentum on the left (respectively on the right) of the particle, i.e. on the line $(t, (vt)_-)$ (respectively on the line $(t, (vt)_+)$). Then

$$(\rho|_L, q|_L) = \begin{cases} (\rho, q) & \text{on } x < vt \\ (\rho_-, q_-) & \text{on } x \geq vt \end{cases} \quad \text{and } (\rho|_R, q|_R) = \begin{cases} (\rho_+, q_+) & \text{on } x \leq vt \\ (\rho, q) & \text{on } x > vt \end{cases}$$

are self-similar solution of the *classical* isothermal Euler equations (without source term) for the initial datum

$$(\rho(0, x), q(0, x)) = (\rho_L \mathbf{1}_{x<0} + \rho_- \mathbf{1}_{x>0}, q_L \mathbf{1}_{x<0} + q_- \mathbf{1}_{x>0})$$

and

$$(\rho(0, x), q(0, x)) = (\rho_+ \mathbf{1}_{x<0} + \rho_R \mathbf{1}_{x>0}, q_+ \mathbf{1}_{x<0} + q_R \mathbf{1}_{x>0}).$$

Therefore, on $\{(t, x), x < vt\}$, a solution of (**??**), is just the restriction on this set of a Riemann problem for the Euler equations. The same holds true on $\{(t, x), x > vt\}$, for a different Riemann problem.

In Section 4.3.1 we describe the set $U_-(\rho_L, q_L, v)$ of all the values that a Riemann solution for the classical Euler equations between (ρ_L, q_L) and a state $(\underline{\rho}, \underline{q}) \in \mathbb{R}_+ \times \mathbb{R}$ can take on the line $x = vt$ and the set $U_+(\rho_R, q_R, v)$ of all the values that a Riemann solution for the classical Euler equations between $(\bar{\rho}, \bar{q}) \in \mathbb{R}_+ \times \mathbb{R}$ and (ρ_R, q_R) can take on the line $x = vt$. The traces (ρ_-, q_-) and (ρ_+, q_+) around the particle should be respectively chosen in those sets. The existence and uniqueness to the Riemann problem (**??**) boils down to prove that there is a unique way to pick (ρ_-, q_-) in $U_-(\rho_L, q_L, v)$ and (ρ_+, q_+) in $U_+(\rho_R, q_R, v)$ such that $((\rho_-, q_-), (\rho_+, q_+))$ belongs to the germ $\mathcal{G}_D(v)$. In Section 4.3.2, we prove that if $u_L - v \leq c$ and $u_R - v \geq -c$, any potential traces $(\rho_\pm, q_\pm)$ around the particle inherits the property $|u_\pm - v| \leq c$ and conclude in that case. It will be referred to as the subsonic case. The other case, referred to as the supersonic case and studied in Section 4.3.3, is more complicated because different types of solutions arise.

Remark 12. In all the sequel, the terms *subsonic* and *supersonic* are used *in the framework of the particle*. For example we say that a solution is subsonic if the difference between the velocity of the particle and the fluid's velocity on both side of the particle is smaller than the speed of sound c.

4.3.1 Accessible states around the particle

Let us start with some very classical results on the isothermal Euler equations without source term that will be useful to determine the solution of (**??**).

Lemma 4.7. *The isothermal Euler equations without source term*

$$\begin{cases} \partial_t \rho + \partial_x q & = 0, \\ \partial_t q + \partial_x (\frac{q^2}{\rho} + c^2 \rho) & = 0, \end{cases}$$

is a strictly hyperbolic system. The eigenvalues of its Jacobian matrix are

$$\lambda_i(\rho, q) = \frac{q}{\rho} + (-1)^i c$$

and the corresponding normalized eigenvectors are

$$r_i(\rho, q) = \begin{pmatrix} \frac{(-1)^i \rho}{c} \\ \frac{(-1)^i q}{c} + \rho \end{pmatrix}.$$

They define two genuinely nonlinear fields. The i-th rarefaction waves express

$$\rho(s) = \begin{cases} \rho_L & \text{if } s \leq s_L = \frac{q_L}{\rho_L} + (-1)^i c, \\ \rho_L e^{\frac{(-1)^i}{c}(s - s_L)} & \text{if } s_L \leq s \leq s_R, \\ \rho_R = \rho_L e^{\frac{(-1)^i}{c}(s_R - s_L)} & \text{if } s_R \leq s, \end{cases}$$

and

$$q(s)=\begin{cases} q_L & \text{if } s\le s_L=\frac{q_L}{\rho_L}+(-1)^i c,\\ [q_L+\rho_L(s-s_L)]\,e^{\frac{(-1)^i}{c}(s-s_L)} & \text{if } s_L\le s\le s_R,\\ q_R=[q_L+\rho_L(s_R-s_L)]\,e^{\frac{(-1)^i}{c}(s_R-s_L)} & \text{if } s_R\le s.\end{cases}$$

The speed of the i-shock is

$$\sigma_i=\frac{q_L}{\rho_L}+(-1)^i c\sqrt{\frac{\rho_R}{\rho_L}}$$

and shocks are entropy satisfying if and only if $u_L\ge u_R$.

We recall below a well-known result on the structure of the Riemann solution.

Proposition 4.8. *For all (ρ_a,q_a) and (ρ_b,q_b) in $\mathbb{R}_+^*\times\mathbb{R}$, there exists a unique self-similar entropy solution of the Riemann problem*

$$\begin{cases} \partial_t\rho+\partial_x q & =0,\\ \partial_t q+\partial_x\left(\frac{q^2}{\rho}+c^2\rho\right) & =0,\\ \rho(0,x) & =\rho_a\mathbf{1}_{x<0}+\rho_b\mathbf{1}_{x>0},\\ q(0,x) & =q_a\mathbf{1}_{x<0}+q_b\mathbf{1}_{x>0}.\end{cases}$$

It consists of the succession of a 1-wave (rarefaction or shock) linking (ρ_L,q_L) to an intermediate state (ρ_*,q_*) followed by a 2-wave (rarefaction or shock) linking (ρ_*,q_*) to (ρ_R,q_R). We denote by

$$W(x/t;(\rho_a,q_a),(\rho_b,q_b))$$

this unique solution.

Proof. The proof of those two results are classical and can be found for example in [GR96]. □

Those tools allow us to describe the set of the accessible states from (ρ_L,q_L) on the left of a particle moving at speed v

$$U_-(\rho_L,q_L,v)=\left\{W(v_-;(\rho_L,q_L),(\underline{\rho},\underline{q})),\ (\underline{\rho},\underline{q})\in\mathbb{R}_+^*\times\mathbb{R}\right\}.$$

and the set of the accessible states from (ρ_R,q_R) on the right of a particle moving at speed v

$$U_+(\rho_R,q_R,v)=\left\{W(v_+;(\bar\rho,\bar q),(\rho_R,q_R)),\ (\bar\rho,\bar q)\in\mathbb{R}_+^*\times\mathbb{R}\right\}$$

in which the left and right traces around the particle must be chosen. According to Definition 4.1, the quantity $q-v\rho$ must be conserved through the particle, so it is more convenient to reason with the variables $(\rho,q-v\rho)$ rather than with the initial unknowns (ρ,q). We introduce the sets

$$V_-(\rho_L,\alpha_L,v)=\{(\rho_-,q_--v\rho_-)\ :\ (\rho_-,q_-)\in U_-(\rho_L,\alpha_L+v\rho_L,v)\}\tag{4.18}$$

and

$$V_+(\rho_R,\alpha_R,v)=\{(\rho_+,q_+-v\rho_+)\ :\ (\rho_+,q_+)\in U_+(\rho_R,\alpha_R+v\rho_R,v)\}.\tag{4.19}$$

Proving that there exists a unique solution to the Riemann problem (**??**) is equivalent to prove that

$$\mathcal{G}_D(v)\cap(U_-(\rho_L,q_L,v)\times U_+(\rho_R,q_R,v))$$

consists in a unique pair of states $((\rho_-,q_-),(\rho_+,q_+))$. This approach follows [DL88] and [Dub01], in which the links between partial Riemann problems and boundary conditions are investigated.

We now give a precise description of the sets $V_-(\rho_L,\alpha_L,v)$ and $V_+(\rho_R,\alpha_R,v)$.

Lemma 4.9. *Let* $(\rho_L, \alpha_L) \in \mathbb{R}_+^* \times \mathbb{R}$ *and* $v \in \mathbb{R}$. *Then*

$$V_-(\rho_L, \alpha_L, v) = \{(\rho_L, \alpha_L)\} \cup \Gamma_-^{sub} \cup \Omega_-^{sup},$$

where Γ_-^{sub} *is the graph of a decreasing function* $f_-^{sub} : [\rho_{L,ex}, +\infty] \to \mathbb{R}$ *for some* $\rho_{L,ex} > 0$, *included in* $\{(\rho, \alpha) : -c\rho \leq \alpha \leq c\rho\}$, *and* Ω_-^{sup} *is the strict hypograph of a decreasing function* $f_-^{sup} : (0, +\infty) \to \mathbb{R}$, *included in* $\{(\rho, \alpha) : \alpha < -c\rho\}$.

Lemma 4.10. *Let* $(\rho_R, \alpha_R) \in \mathbb{R}_+^* \times \mathbb{R}$ *and* $v \in \mathbb{R}$. *Then*

$$V_+(\rho_R, \alpha_R, v) = \{(\rho_R, \alpha_R)\} \cup \Gamma_+^{sub} \cup \Omega_+^{sup}$$

where Γ_+^{sub} *is the graph of a increasing function* $f_+^{sub} : [\rho_{R,ex}, +\infty] \to \mathbb{R}$ *for some* $\rho_{R,ex} > 0$, *included in* $\{(\rho, \alpha) : -c\rho \leq \alpha \leq c\rho\}$, *and* Ω_+^{sup} *is the strict epigraph of a increasing function* $f_+^{sup} : (0, +\infty) \to \mathbb{R}$, *included in* $\{(\rho, \alpha) : \alpha > c\rho\}$.

Those sets are respectively depicted on the left and on the right of Figure 4.3. The subscript ex refers to the extremity of Γ_-^{sub} and Γ_+^{sub}.

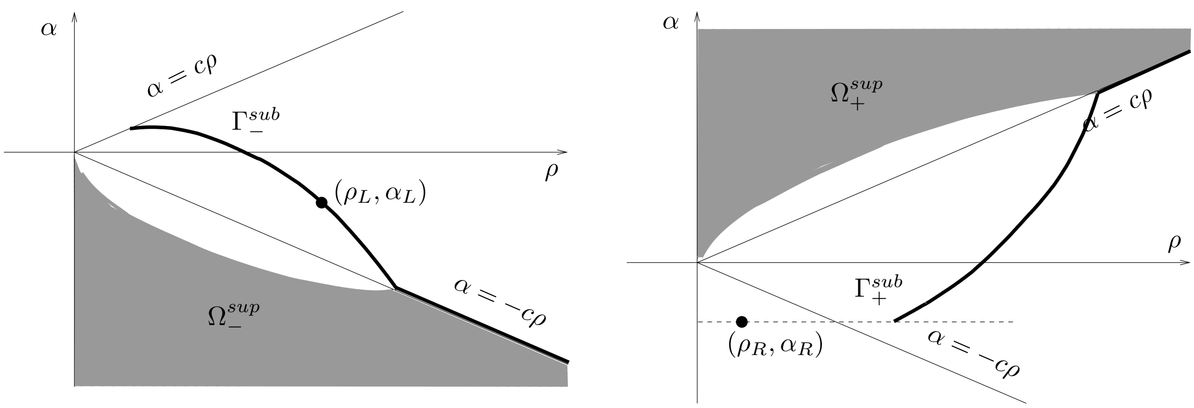


Figure 4.3. *On the left, the set* $V_-(\rho_L, \alpha_L, v)$ *when* $-c\rho_L < \alpha_L < c\rho_L$. *On the right, the set* $V_+(\rho_R, \alpha_R, v)$ *when* $\alpha_R < c\rho_R$.

Remark 13. By definition of α we have

$$\{(\rho, \alpha) : -c\rho \leq \alpha \leq c\rho\} = \{(\rho, u) - c \leq u - v \leq c\},$$

so Γ_-^{sub} and Γ_+^{sub} contain only subsonic states (in the framework of the particle), while Ω_-^{sup} and Ω_+^{sup} contain only supersonic states (in the framework of the particle).

Proof (Lemma 4.9). We fix the particle velocity v and a left state (ρ_L, α_L). The element of (ρ, α) of the set $V_-(\rho_L, \alpha_L, v)$ is the value on the line $x = vt$ of a Riemann problem with (ρ_L, α_L) on its left. Thus it is linked with (ρ_L, α_L) by a succession of two waves, both traveling slower than v. Let us first exhibit all the states (ρ_*, q_*) that can be linked to (ρ_L, q_L) with a 1-wave traveling at speed smaller than v. According to Lemma (4.7), with a 1-rarefaction wave we can reach all the states (ρ, q) in the set

$$\left\{ \left(\rho_L e^{-\frac{(s-(u_L-c))}{c}}, [q_L + \rho_L(s - (u_L - c))]\, e^{-\frac{(s-(u_L-c))}{c}} \right),\ u_L - c \leq s \leq v \right\}.$$

This set is empty if $u_L - c > v$. Parametrized by ρ it rewrites

$$\left\{ \left(\rho, \left[u_L - c \ln\left(\frac{\rho}{\rho_L}\right) \right] \rho \right), \rho_L e^{-\frac{v-(u_L-c)}{c}} \leq \rho \leq \rho_L \right\}.$$

Let us denote in this case $\rho_{L,ex} = \rho_L e^{-\frac{v-(u_L-c)}{c}}$. The states (ρ_*, q_*) accessible through a 1-shock traveling slower than v satisfy

$$\begin{cases} \frac{q_L}{\rho_L} - c\sqrt{\frac{\rho_*}{\rho_L}} < v \quad \text{and} \quad \rho_* > \rho_L, \\ q_* = q_L + \left(\frac{q_L}{\rho_L} - c\sqrt{\frac{\rho_*}{\rho_L}}\right)(\rho_* - \rho_L). \end{cases}$$

We easily pass to the (ρ, α) variable: with a 1-wave traveling slower than v, we can reach all the states (ρ, α) with $\alpha = f_-(\rho)$, where

$$f_-(\rho) = \begin{cases} \left[\frac{\alpha_L}{\rho_L} - c\ln\left(\frac{\rho}{\rho_L}\right)\right]\rho & \text{if } \rho_{L,ex} \leq \rho \leq \rho_L, \\ \alpha_L + \left(\frac{\alpha_L}{\rho_L} - c\sqrt{\frac{\rho}{\rho_L}}\right)(\rho - \rho_L) & \text{if } \max(\rho_{L,ex}, \rho_L) < \rho, \end{cases}$$

and

$$\rho_{L,ex} = \begin{cases} \rho_L e^{-\frac{v-s_L}{c}} & \text{if } s_L = u_L - c \leq v, \\ \left(\frac{u_L - v}{c}\right)^2 \rho_L & \text{if } u_L - c > v. \end{cases}$$

If $u_L - v < c$, this graph regroups all the 1-shocks and the 1-rarefaction waves leading to a density higher than $\rho_{L,ex} = \rho_L e^{-\frac{v-(u_L-c)}{c}}$, while if $u_L - v \geq c$, this graphs contains only the 1-shocks leading to a density higher than $\rho_{L,ex} = \rho_L \left(\dfrac{u_L - v}{c}\right)^2 = \dfrac{\alpha_L}{c^2 \rho_L}$. We check that f_- is concave and decreasing. Moreover, if $u_L - c \leq v$, $f_-(\rho_{L,ex}) = c\rho_{L,ex}$ and $f_-'(\rho_{L,ex}) = 0$, while if $u_L - c > v$, $f_-(\rho_{L,ex}) = \alpha_L$. In particular, $f_-(\rho) < c\rho$ for any $\rho > \rho_{L,ex}$ as shown on Figure (4.4). Let us now stop at any state (ρ_*, q_*)

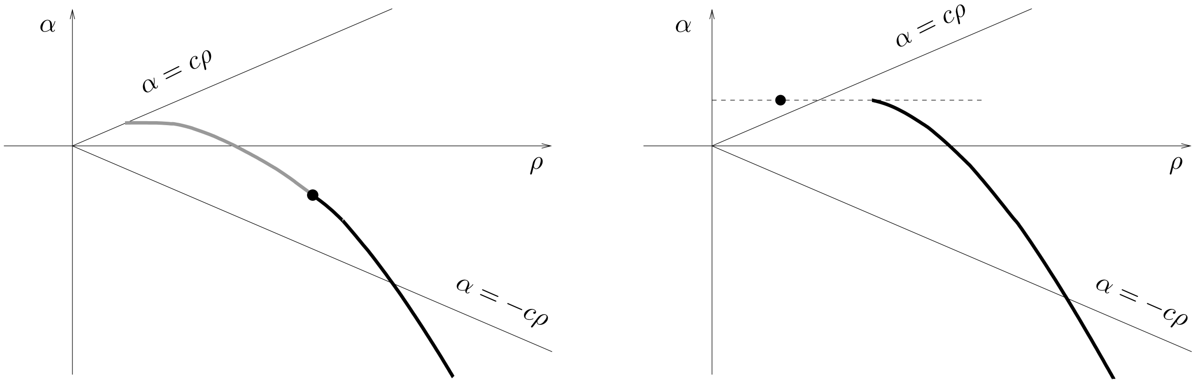


Figure 4.4. *Accessible states via a* 1*-wave starting from* (ρ_L, q_L)*(black dot) in the* (ρ, α)*-plane. Left, the case* $u_L - v \leq c$ *and right, the case* $u_L - v > c$*. In gray are the* 1*-rarefaction waves and in black are the* 1*-shocks.*

belonging to the graph of f_-, and continue with a 2-wave traveling at speed smaller than v. The set of all the states (ρ_-, q_-) that can be joined from (ρ_*, q_*) with a 2-wave traveling at speed smaller than v is

$$\left\{\left(\rho_* e^{\frac{(s-(u_*+c))}{c}}, [q_* + \rho_L(s - s_*)]\, e^{\frac{(s-(u_*+c))}{c}}\right),\ u_* + c \leq s \leq v\right\}.$$

This set is empty if $u_* + c > v$, and can be parametrized by ρ by

$$\left\{\left(\rho, \left[u_* + c\ln\left(\frac{\rho}{\rho_*}\right)\right]\rho\right), \rho_* \leq \rho \leq \rho_* e^{\frac{v-(u_*+c)}{c}}\right\}.$$

With a 2-shock slower than v, we can reach all the states (ρ_-, q_-) such that

$$\begin{cases} \frac{q_*}{\rho_*} + c\sqrt{\frac{\rho_-}{\rho_*}} < v \quad \text{and} \quad \rho_- < \rho_*, \\ q_- = q_* + \left(\frac{q_*}{\rho_*} + c\sqrt{\frac{\rho_-}{\rho_*}}\right)(\rho_- - \rho_*). \end{cases}$$

Therefore the 2-waves traveling at speed smaller than v, starting from (ρ_*, q_*), are:

- If $u_* - v \leq -c$, all the 2-shocks and the 2-rarefaction waves leading to a density smaller than $\rho_* e^{\frac{v-(u_*+c)}{c}}$;
- If $-c < u_* - v < 0$, only the 2-shock leading to a density smaller than
$$\rho_{*,ex} = \rho_* \left(\frac{v-u_*}{c}\right)^2 = \frac{\alpha_*^2}{c^2 \rho_*};$$
- If $u_* - v \geq 0$, there are no such 2-waves.

The Figure (4.5) resumes the first two cases. We now prove that, as emphasized in Figure 4.5, all the

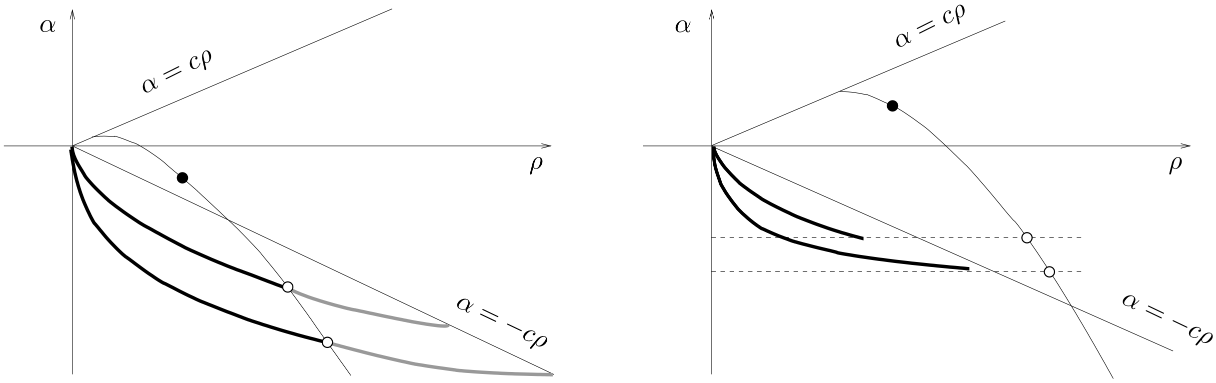


Figure 4.5. *Accessible states from (ρ_L, q_L) (black dot) via a 1-wave stopped in (ρ_*, q_*) (white dots) followed by a 2-wave in the (ρ, α)-plane. Left, the case where $u_* - v \leq -c$ and right, the case $-c < u_{L*} - v < 0$. In gray are the 2-rarefaction waves and in black are the 2-shocks.*

states reached from (ρ_*, q_*) through a 2-wave traveling at speed smaller than v are such that $\alpha \leq -c\rho$. This is easy to check for the 2-rarefaction waves and for the 2-shock when $u_* - v \leq -c$. When $-c < u_* - v < 0$, the reachable densities are smaller than $\frac{\alpha_*^2}{c^2\rho_*} < \rho_*$. We want to prove that
$$\alpha = \rho \left(\frac{\alpha_*}{\rho_*} + c\sqrt{\frac{\rho}{\rho_*}} - c\sqrt{\frac{\rho_*}{\rho}}\right) \leq -c\rho.$$
This is true because the function $\rho \mapsto \frac{\alpha_*}{\rho_*} + c\sqrt{\frac{\rho}{\rho_*}} - c\sqrt{\frac{\rho_*}{\rho}}$ is increasing, and is equal to $\frac{-c^2\rho_*}{|\alpha_*|}$, which is smaller than $-c$.

To conclude, the set of the states reached from (ρ_L, α_L) through a 1-wave slower than the particle is the graph of f_-. The set of the stated reached from (ρ_L, α_L) through a 1-wave followed by a 2-wave, both of them traveling at speed smaller than v, is a family of curves, entirely included in $\{(\rho, \alpha), \alpha \leq -c\rho\}$. Those curves are portions of Lax curves, which fill the half plane $\mathbb{R}_+^* \times \mathbb{R}$ and do not cross each other. For a fixes $\alpha < 0$ such that $f_-(\rho) \geq -c\rho$, it is not possible to reach densities higher than the density obtain thanks to a shock at speed v, which is $\frac{f_-(\rho)^2}{c^2\rho}$. We can separate those states in two categories, as depicted on the left of Figure 4.3:

- the subsonic ones (i.e the ones such that $-c\rho \leq \alpha \leq c\rho$), which constitute the graph of the function
$$f_-^{sub}(\rho) = \begin{cases} f_-(\rho) & \text{if } \rho_{L,ex}\rho \leq \rho_{-c} \\ -c\rho & \text{if } \rho_{-c} \leq \rho \end{cases}$$
where ρ_{-c} is the only density such that $f_-(\rho_{-c}) = -c\rho$. We regroup those states in Γ_-^{sub};

- the supersonic ones, and more precisely the states such that $\alpha < -c\rho$, which formed the hypograph of the function
$$f_-^{sup}(\rho) = \begin{cases} f_-\left(\frac{f_-(\rho)^2}{c^2\rho}\right) & \text{if } 0 < \rho < \rho_{-c} \\ -c\rho & \text{if } \rho_{-c} \leq \rho. \end{cases}$$
We regroup those states in Ω_-^{sup}. □

Proof (Lemma 4.10). We do not prove this lemma which is exactly similar to Lemma 4.9. The curve of the states accessible by a 2-wave traveling faster than v and ending in (ρ_R, q_R), can be parametrized by
$$f_+(\rho) = \begin{cases} \left[\frac{\alpha_R}{\rho_R} + c\ln\left(\frac{\rho}{\rho_R}\right)\right]\rho & \text{if } \rho_{R,ex} \leq \rho \leq \rho_R, \\ \alpha_R + \left(\frac{\alpha_R}{\rho_R} + c\sqrt{\frac{\rho}{\rho_R}}\right)(\rho - \rho_R) & \text{if } \max(\rho_{R,ex}, \rho_R) < \rho, \end{cases}$$
where
$$\rho_{R,ex} = \begin{cases} \rho_R e^{\frac{v-(u_R+c)}{c}} & \text{if } u_R + c \geq v, \\ \left(\frac{v-u_R}{c}\right)^2 \rho_R & \text{if } u_R + c < v. \end{cases}$$
If $u_R - v \geq -c$ it is possible to follow all the 2-shocks and some 2-rarefaction waves. In this case $f_+(\rho_{R,ex})$ belongs to the line $\alpha = -c\rho$ and $f_+'(\rho_{R,ex}) = 0$. On the contrary if $u_R - v < -c$, we can only follow 2-shocks, and the state $(\rho_{R,ex}, \alpha_{R,ex})$ with the highest density we can reach verifies
$$\alpha_{R,ex} = \alpha_R \quad \text{and} \quad \rho_{R,ex} = \frac{\alpha_R^2}{c^2\rho_R}.$$
We easily check that that f_+ is convex, increasing and crosses the line $\alpha = c\rho$ for a unique density that we denote by ρ_c.

After similar computations for the 1-wave ending on a state $(\rho_*, q_* = f_+(\rho_*))$ we obtain
$$f_+^{sub}(\rho) = \begin{cases} f_+(\rho) & \text{if } \rho_{R,ex} \leq \rho \leq \rho_c, \\ c\rho & \text{if } \rho_c \leq \rho, \end{cases}$$
and
$$f_+^{sup} = \begin{cases} f_+\left(\frac{f_+(\rho)^2}{c^2\rho}\right) & \text{if } 0 < \rho < \rho_c, \\ c\rho & \text{if } \rho_c \leq \rho. \end{cases}$$
□

4.3.2 Resolution of the Riemann problem in the subsonic case

We are now in position to solve the Riemann problem (**??**) for a particle moving at speed v. In the sequel, we denote by (ρ_+, q_+) the trace on the left of the particle, i.e. on the line $x = (vt)_-$ and by (ρ_-, q_-) the trace on its right, i.e. on the line $x = (vt)_+$. In this section, we treat a special case implying that (ρ_L, α_L) does not belong to Ω_+^{sup} and (ρ_R, α_R) does not belong to Ω_-^{sup}. A consequence is that the left trace belongs to Γ_-^{sub} and the right trace belong to Γ_+^{sub}.

Lemma 4.11. *If (ρ_L, q_L) and (ρ_R, q_R) verify*
$$u_L - v \leq c \quad \textit{and} \quad -c \leq u_R - v$$
then (ρ_-, q_-) and (ρ_+, q_+) are necessary subsonic, i.e,
$$-c \leq u_- - v \leq c \quad \textit{and} \quad -c \leq u_+ - v \leq c.$$

Proof. This is a straightforward consequence of the form of $V_-(\rho_L, \alpha_L, v)$ and $V_+(\rho_R, \alpha_R, v)$ exhibited in Lemmas 4.9 and 4.10 and resumed on Figure 4.3. Definition 4.1 implies that $q_- - v\rho_- = q_+ - v\rho_+$. We denote by α this quantity. If $\alpha = 0$ there is nothing to prove. Suppose $\alpha > 0$. Lemma 4.9 shows that $V_-(\rho_L, \alpha_L, v)$ contains only states such that $u - v \leq c$. As a consequence, we necessary have that $c\rho_- \geq \alpha$. The velocity at the entry of the particle is subsonic, thus it is also subsonic at its exit: $\rho_+ \geq \frac{\alpha}{c}$ (see Corollary 4.4). If $\alpha < 0$, Lemma 4.9 shows that $V_+(\rho_R, \alpha_R, v)$ contains only states verifying $u - v \geq -c$. Using Corollary 4.4 again, we obtain that both ρ_+ and ρ_- are larger than $\frac{|\alpha|}{c}$. ☐

We are now looking for (ρ_-, α_-) in Γ_-^{sub} and (ρ_+, α_+) in Γ_+^{sub}. The quantity $\alpha = q - v\rho$ being conserved through the particle, it is more convenient to parametrize the accessible sets Γ_-^{sub} and Γ_+^{sub} by α rather than by ρ. For this purpose, we introduce g_-^{sub} and g_+^{sub}, the inverses of f_-^{sub} and f_+^{sub}. They exist because these two functions are strictly monotone. We now have

$$\Gamma_-^{sub} = \{(\rho, \alpha), \rho = g_-^{sub}(\alpha),\ \alpha \leq f_-^{sub}(\rho_{L,ex})\}$$

and

$$\Gamma_+^{sub} = \{(\rho, \alpha), \rho = g_+^{sub}(\alpha),\ \alpha \geq f_+^{sub}(\rho_{R,ex})\}.$$

Lemma 4.12. *If D has the same sign as α an is a nondecreasing function of α, the function*

$$\Delta(\alpha) = F_\alpha(g_-^{sub}(\alpha)) - F_\alpha(g_+^{sub}(\alpha))$$

is strictly decreasing on any interval included in $(0, +\infty)$ where $g_-^{sub} \geq g_+^{sub}$ and on any interval included in $(-\infty, 0)$ where $g_-^{sub} \leq g_+^{sub}$;

Proof. We compute the derivative of Δ:

$$\begin{aligned}\left[F_\alpha(g_-^{sub}(\alpha)) - F_\alpha(g_+^{sub}(\alpha))\right]' =& F_\alpha'(g_-^{sub}(\alpha))(g_-^{sub})'(\alpha) + \left[\frac{\partial}{\partial\alpha}F_\alpha\right](g_-^{sub}(\alpha)) \\ &- F_\alpha'(g_+^{sub}(\alpha))(g_+^{sub})'(\alpha) - \left[\frac{\partial}{\partial\alpha}F_\alpha\right](g_+^{sub}(\alpha))\end{aligned}$$

On the one hand, $\alpha \mapsto g_-^{sub}(\alpha)$ decreases, $\alpha \mapsto g_+^{sub}(\alpha)$ increases. We recall that for all α in there interval of definition, $cg_\pm^{sub}(\alpha) \geq |\alpha|$. If $c\rho$ is larger than $|\alpha|$,

$$F_\alpha'(\rho) = \frac{1}{|D(\rho,\alpha)|}\left(c^2 - \frac{\alpha^2}{\rho^2}\right) \geq 0.$$

and it follows that

$$F_\alpha'(g_-^{sub}(\alpha))(g_-^{sub})'(\alpha) - F_\alpha'(g_+^{sub}(\alpha))(g_+^{sub})'(\alpha) \leq 0.$$

On the other hand, by definition of F_α, we have

$$\begin{aligned}\left[\frac{\partial}{\partial\alpha}F_\alpha\right](\rho_2) - \left[\frac{\partial}{\partial\alpha}F_\alpha\right](\rho_2) &= \int_{\rho_1}^{\rho_2} \frac{\partial}{\partial\alpha}\left[\frac{1}{|D(r,\alpha)|}\left(c^2 - \frac{\alpha^2}{r^2}\right)\right] dr \\ &= \int_{\rho_1}^{\rho_2}\left[\frac{-\operatorname{sign}(D(r,\alpha))\partial_\alpha D(r,\alpha)}{|D(r,\alpha)|^2}\left(c^2 - \frac{\alpha^2}{r^2}\right) - \frac{2\alpha}{r^2|D(r,\alpha)|}\right] dr.\end{aligned}$$

If $c\rho_1$ and $c\rho_2$ are both larger than α, the term $c^2 - \frac{\alpha^2}{r^2}$ is nonnegative. Moreover $D(r,\alpha)$ and α have the same sign and D is an non-decreasing function of α. Thus the integral has the opposite sign as $\alpha(\rho_2 - \rho_1)$. In particular, on any interval where α is positive and $g_-^{sub} \geq g_+^{sub}$ and on any interval where α is negative and $g_-^{sub} \leq g_+^{sub}$,

$$\left[\frac{\partial}{\partial\alpha}F_\alpha\right](g_-^{sub}(\alpha)) - \left[\frac{\partial}{\partial\alpha}F_\alpha\right](g_+^{sub}(\alpha)) \leq 0,$$

which proves the lemma. ☐

We are now in position to state the result in the subsonic case.

Proposition 4.13. *Let $v \in \mathbb{R}$, $(\rho_L, \alpha_L) \in \mathbb{R}_+^* \times \mathbb{R}$ and $(\rho_R, \alpha_R) \in \mathbb{R}_+^* \times \mathbb{R}$ such that*

$$-c \leq u_L - v \leq c \quad \text{and} \quad -c \leq u_R - v \leq c.$$

Then the set

$$\mathcal{G}_D(v) \cap (V_-(\rho_L, \alpha_L, v) \times V_+(\rho_R, \alpha_R, v))$$

is reduced to a unique element. Moreover, both traces belong to the subsonic triangle

$$\{(\rho, \alpha), -c\rho \leq \alpha \leq c\rho\} = \{(\rho, u), -c \leq u - v \leq c\}.$$

In other words, the Riemann problem (**??**) *admits a unique solution, which is entirely subsonic.*

Proof. Step 1: Properties of the traces.

Suppose that

$$((\rho_-, \alpha_-), (\rho_+, \alpha_+)) \in \mathcal{G}_D(v) \cap (V_-(\rho_L, \alpha_L, v) \times V_+(\rho_R, \alpha_R, v))$$

then $\alpha_- = \alpha_+$. We denote by α this quantity. In Lemma 4.11 we proved that

$$-c \leq u_- - v \leq c \quad \text{and} \quad -c \leq u_+ - v \leq c.$$

Thus we can use Corollary 4.4 to obtain that

- If $\alpha = 0$ then $\rho_- = \rho_+$ and $q_- = q_+$;
- If $\alpha > 0$, then $\rho_- > \rho_+$;
- If $\alpha < 0$, then $\rho_- < \rho_+$.

Step 2: It exists a unique α_0 such that $g_-^{sub}(\alpha_0) = g_+^{sub}(\alpha_0)$.

As $c\rho_L \geq \alpha$, the upper extremity of g_-^{sub} is the point $(\rho_{L,ex}, c\rho_{L,ex})$ (see the left of Figure 4.4, $\rho_{L,ex}$ has been defined in Lemma 4.9). Similarly as $c\rho_R \geq |\alpha|$, the lower extremity of g_+^{sub} is the point $(\rho_{R,ex}, -c\rho_{R,ex})$. If $(\rho_{L,ex}, c\rho_{L,ex})$ belongs to Γ_+^{sub}, as depicted on the left of Figure 4.6, we directly obtain the existence of $\alpha_0 = c\rho_{L,ex}$. Similarly if $(\rho_{R,ex}, -c\rho_{R,ex})$ belong to Γ_+^{sub} we have $\alpha_0 = -c\rho_{R,ex}$. If we are not on one of these cases, as in the right of Figure 4.6, we have

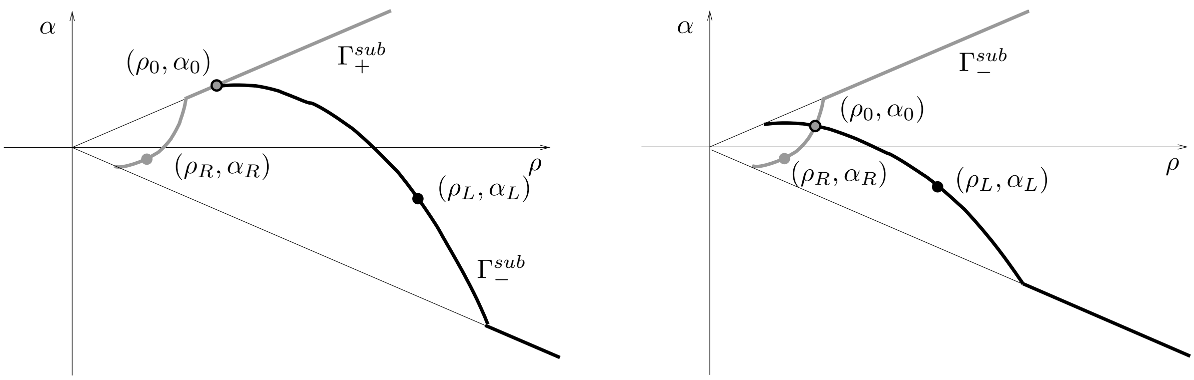


Figure 4.6. *In black, the graph of g_-^{sub}, in grey the graph of g_+^{sub}. The intersection point $(\rho_0 = g_\pm^{sub}(\alpha_0), \alpha_0)$ is either on the sonic line $\alpha = c\rho$ or somewhere inside the triangle.*

$$g_-^{sub}(c\rho_{L,ex}) < g_+^{sub}(c\rho_{L,ex}) \quad \text{and} \quad g_+^{sub}(-c\rho_{R,ex}) < g_-^{sub}(-c\rho_{R,ex}),$$

and α_0 exists continuity of $g_-^{sub} - g_+^{sub}$.

Step 3: Conclusion. Let us suppose that $\alpha_0 > 0$. According to the first step, the solution lies in the set $0 \le \alpha \le \alpha_0$ (we check on Figure 4.6 that the relative positions of $g_-^{sub}(\alpha)$ and $g_+^{sub}(\alpha)$ are only correct in that zone). Moreover $\Delta(\alpha_0) = 0$ and

$$\Delta(\alpha) = \int_{g_+^{sub}(\alpha)}^{g_-^{sub}(\alpha)} \frac{1}{|D(\alpha,\rho)|}\left(c^2 - \frac{\alpha^2}{r^2}\right) dr \underset{\alpha\to 0_+}{\longrightarrow} +\infty,$$

because $D(\rho, 0) = 0$ and $g_+^{sub}(0) \neq g_-^{sub}(0)$. By Lemma 4.12, the function Δ is monotonous, thus there exists a unique α in $[0, \alpha_0]$ such that $g_+^{sub}(\alpha) - g_-^{sub}(0) = 1$. Therefore, $(\rho_-, \alpha_-) = (g_-^{sub}(\alpha), \alpha)$ and $(\rho_+, \alpha_+) = (g_+^{sub}(\alpha), \alpha)$ verify Definition 4.1 and give a solution. □

4.3.3 Resolution of the Riemann problem in the supersonic case

In this Section we prove Theorem 4.6 when $\alpha_L > c\rho_L$ or $\alpha_R < -c\rho_R$. Without loss of generality, let us assume that $\alpha_L > c\rho_L$; the case $\alpha_R < -c\rho_L$ may be treated in a symmetrical way. Lemma 4.4 does not hold anymore. We must study in detail the case where (ρ_L, α_L) belongs to $V_+(\rho_R, \alpha_R, v)$, which was excluded in the subsonic case. For this purpose, we introduce some notation, summarized on Figure 4.7, which also recall the notation introduced in Lemmas 4.9 and 4.10. In the sequel we denote by:

- for any subscript i and for any point (ρ_i, α_i), we denote by $(\tilde{\rho}_i := \frac{\alpha_i^2}{c^2\rho_i}, \alpha_i)$ the state reached with a shock at speed v. Remark that $\tilde{\tilde{\rho}}_i = \rho_i$;
- $\rho_{L,ex}$ the extremity of the curve g_-^{sub}. Lemma 4.9 shows that when $u_L - v > c$, $\rho_{L,ex} = \tilde{\rho_L}$, and that
$$\forall \alpha < 0,\ (g_-^{sup}(\alpha), \alpha) = (\widetilde{g_-^{sub}(\alpha)}, \alpha) \ \text{ and } \ \forall \alpha > 0,\ (g_+^{sup}(\alpha), \alpha) = (\widetilde{g_+^{sub}(\alpha)}, \alpha).$$
- $\rho_E = g_+^{sup}(\alpha_L)$ the intersection of the line $\alpha = \alpha_L$ with Γ_+^{sup}. Note that $\tilde{\rho_E} = g_+^{sub}(\alpha_L)$.
- $(\rho_{R,ex}, \alpha_{R,ex})$ the extremity of the curve g_+^{sub}. We recall that, by Lemma 4.10, if $\alpha_R < -c\rho_R$, $\alpha_{R,ex} = \alpha_R$ and $\rho_{R,ex} = \frac{\alpha_R^2}{c^2\rho_R}$; and that if $\alpha_R \ge -c\rho_R$, $\rho_{R,ex} \le \rho_R$ and $\alpha_{R,ex} = -c\rho_{R,ex}$.
- $\rho_F = g_-^{sup}(\alpha_{R,ex})$ the intersection of the line $\alpha = \alpha_{R,ex}$ with Γ_-^{sup}. Note that $\tilde{\rho_F} = g_+^{sub}(\alpha_{R,ex})$.

We first exhibit a link between the position of (ρ_L, α_L) and the position of (ρ_R, α_R).

Lemma 4.14. *Let $(\rho_L, \alpha_L) \in \mathbb{R}_+^* \times \mathbb{R}$ and $(\rho_R, \alpha_R) \in \mathbb{R}_+^* \times \mathbb{R}$ such that $\alpha_L > c\rho_L$ and $\alpha_R < -c\rho_R$. It is not possible to have*
$$(\rho_L, \alpha_L) \in \Omega_+^{sup} \quad \text{and} \quad (\rho_R, \alpha_R) \in \Omega_-^{sup}.$$

Proof. The hypothesis $\alpha_R < -c\rho_R$ implies that $\rho_{R,ex} = \tilde{\rho_R}$. Suppose that (ρ_L, α_L) belongs to Ω_+^{sup}. Then $\rho_L < \rho_E$, and we have $\tilde{\rho_E} < \tilde{\rho_L}$. The monotonicity of g_+^{sub} and g_-^{sub} gives $\tilde{\rho_R} = \rho_{R,ex} < \tilde{\rho_F}$. Thus $\rho_R > \rho_F$, which means that (ρ_R, α_R) does not belong to Ω_-^{sup}. □

It allows us to exclude the case $(\rho_R, \alpha_R) \in \Omega_-^{sup}$ of our study. Indeed, if $(\rho_R, \alpha_R) \in \Omega_-^{sup}$, then $\alpha_R < -c\rho$ and (ρ_L, α_L) does not belong to Ω_-^{sup}, and we treat that case by symmetry. We now state the result in the supersonic case.

Proposition 4.15. *Let $(\rho_L, \alpha_L) \in \mathbb{R}_+^* \times \mathbb{R}$ and $(\rho_R, \alpha_R) \in \mathbb{R}_+^* \times \mathbb{R}$ such that $\alpha_L > c\rho_L$ and $(\rho_R, \alpha_R) \notin \Omega_-^{sup}$ Then the set*
$$\mathcal{G}_D(\lambda) \cap (V_-(\rho_L, \alpha_L, v) \times V_+(\rho_R, \alpha_R, v))$$
is reduced to a unique element $((\rho_-, \alpha_-), (\rho_+, \alpha_+))$.

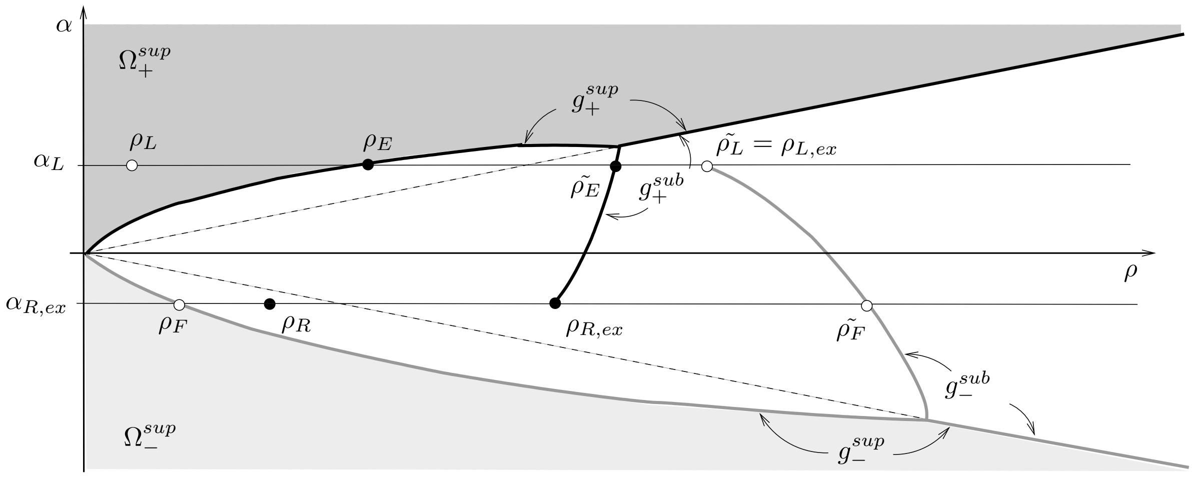


Figure 4.7. *Notation for the supersonic case $\alpha_L > c\rho_L$. $V_+(v, \rho_R, \alpha_R)$ is the union of the open set Ω_+^{sub} above the graph of g_+^{sup} and of the graph of g_+^{sub}; $V_-(v, \rho_L, \alpha_L)$ is the union of the open set Ω_-^{sub} below the graph of g_-^{sup} and of the graph of g_-^{sub}.*

Proof. We prove the result for a drag force only depend on α, in which case we simply denote it by $D(\alpha)$. Because of the compatibility between the germ $\mathcal{G}_D(v)$ and the shocks at speed v, the germ is smaller and easier to describe. Instead of Definition (4.1), we can use the three relations of Corollary 4.4 and (4.12). Even though it simplifies the proof, all the key ingredients are present in that case. The proof of Theorem 4.6 in the general case is given in Appendix 4.5.

The proof relies on the relative positions of ρ_L and ρ_E, and on the fact that the transformation $(\rho, \alpha) \mapsto \left(\tilde{\rho} = \frac{\alpha^2}{c^2\rho}, \alpha\right)$ reverses the positions of points around the point $(\alpha/c, \alpha)$, as depicted on Figure 4.7.

Case 1*: $\rho_L \le \rho_E$, or equivalently $(\rho_L, \alpha_L) \in \Omega_+^{sup}$.*

In that case, we can chose $(\rho_-, \alpha_-) = (\rho_L, \alpha_L)$ if and only if

$$\left(\frac{\alpha_L^2}{\rho_L} + c^2\rho_L\right) - \left(\frac{\alpha_L^2}{\rho_E} + c^2\rho_E\right) > D(\alpha_L).$$

Assume that this inequality is fulfilled. Then, as $f_{\alpha_L} : \rho \mapsto \frac{\alpha_L^2}{\rho} + c^2\rho$ decreases on $(0, \frac{\alpha_L}{c})$, there exists a unique $\rho_+ \in (\rho_L, \rho_E)$ such that

$$\left(\frac{\alpha_L^2}{\rho_L} + c^2\rho_L\right) - \left(\frac{\alpha_L^2}{\rho_+} + c^2\rho_+\right) = D(\alpha_L).$$

Therefore we obtain the solution

$$((\rho_L, \alpha_L), (\rho_+, \alpha_L)) \in \mathcal{G}_D(v) \cap (V_-(\rho_L, \alpha_L, v) \times V_+(\rho_R, \alpha_R, v)).$$

In that case we have

$$\left(\frac{\alpha_L^2}{\tilde{\rho_L}} + c^2\tilde{\rho_L}\right) - \left(\frac{\alpha_L^2}{\tilde{\rho_E}} + c^2\tilde{\rho_E}\right) > D(\alpha_L),$$

so $\Delta(\alpha_L) > 1$. The function $g_-^{sub} - g_+^{sub}$ decreases, so it remain positive on $(0, \alpha_L)$. Therefore, Lemma 4.11 shows that there is no other solution. Suppose now that

$$\left(\frac{\alpha_L^2}{\rho_L} + c^2\rho_L\right) - \left(\frac{\alpha_L^2}{\rho_E} + c^2\rho_E\right) < D(\alpha_L).$$

Thus we have that $\Delta(\alpha_L) < 1$ and $\Delta(0) = +\infty$ and we conclude for the existence and the uniqueness as in Proposition 4.13.

Case 2: $\rho_L > \rho_E$*, or equivalently* $(\rho_L, \alpha_L) \notin \Omega_+^{sup}$.

It implies that $\tilde{\rho_E} > \tilde{\rho_L}$. Moreover, (ρ_R, α_R) does not belong to Ω_-^{sup}. The fact that $\rho_F < \rho_R$ implies that $\tilde{\rho_F} \geq \rho_{R,ex}$ and that $\Delta(\alpha_L) < 0$. We conclude exactly as in the subsonic case, see the proof of Proposition 4.13. $\square$

4.3.4 Asymptotics

Depending on the intensity of the drag force D, the model (4.1) exhibit a whole range of behavior, from the lack of particle to the presence of a solid wall.

Proposition 4.16. *Suppose that the drag force writes* $D(\alpha) = \lambda D_0(\alpha)$*, with* D_0 *a fixed drag force. When* λ *tends to infinity, the solution of* (**??**) *tends to the solution of the Riemann problem with the same initial data for the Euler equation with a solid wall along* $x = vt$*. When* λ *vanishes, the solution of* (**??**) *tends to the solution of the Riemann problem with the same initial data for the Euler equation without particle.*

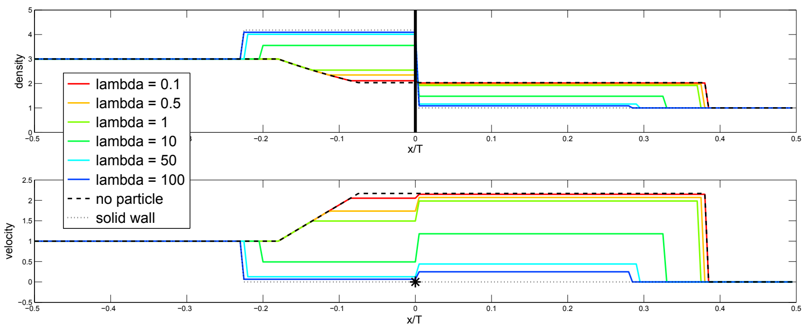


Figure 4.8. *Exact solutions of the same Riemann problem for different friction coefficients* λ*.*

Proof. Let us begin by the case $\lambda \longrightarrow +\infty$. Then in the supersonic case, it becomes impossible to chose $(\rho_-, \alpha_-) = (\rho_L, \alpha_L)$ because the inequality

$$\left(\frac{\alpha_L^2}{\rho_L} + c^2\rho_L\right) - \left(\frac{\alpha_L^2}{\rho_E} + c^2\rho_E\right) > \lambda D_0(\alpha_L)$$

will always fail. This is illustrated by Figure 4.9, where the drag force is $D(\alpha) = \lambda\alpha$: this property holds for $\lambda = 1$, while it is lost for $\lambda = 20$. In the subsonic case, as g_-^{sub} and g_+^{sub} take value in a bounded interval, for all non null α, $\Delta(\alpha)$ tends to 0 as λ tends to infinity. Therefore, (ρ_-, α_-) tends to $(g_-^{sub}(0), 0)$ while (ρ_+, α_+) tends to $(g_+^{sub}(0), 0)$. This is exactly the solution for the Riemann problem with a solid wall. When λ tends to zero, remark that in the subsonic case (or the supersonic case when (ρ_L, α_L) is not in Ω_+^{sup}) the solution of $\Delta(\alpha) = 0$ tends to the crossing point α_0. In the supersonic case when $(\rho_L, \alpha_L) \in \Omega_+^{sup}$, the inequality

$$\left(\frac{\alpha_L^2}{\rho_L} + c^2\rho_L\right) - \left(\frac{\alpha_L^2}{\rho_E} + c^2\rho_E\right) > \lambda D_0(\alpha_L)$$

is verified for small enough λ, and (ρ_-, α_-) is (ρ_L, α_L) while (ρ_+, α_+) tends to (ρ_L, α_L). It corresponds to the value of the solution of the Euler equation without particle on the line $x = vt$. Those two asymptotics behaviors are depicted, in the subsonic case, on Figure 4.8. $\square$

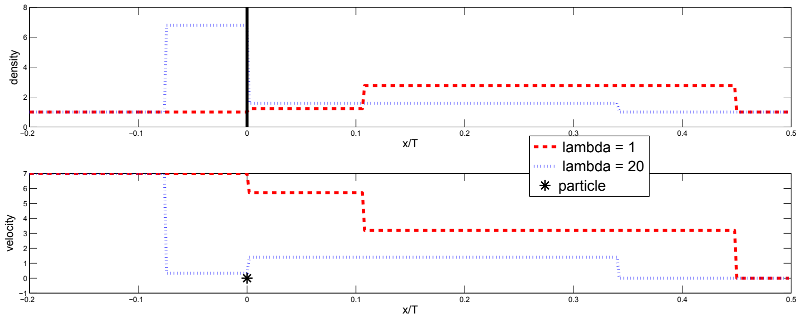


Figure 4.9. *Two types of solutions when $\alpha_L > c\rho_L$, for the drag force $D(\alpha) = \lambda\alpha$. For small enough λ there is no wave on the left of the particle. As λ becomes greater, we recover the subsonic solution.*

4.4 Existence of up to three solutions when $|D|$ is not a decreasing function of ρ

When $|D|$ is not a decreasing function of ρ, Condition (4.22) may not hold and we can lose uniqueness and obtain up to three solutions. This is not surprising: the choice $D(\rho, u - v) = \rho$ is, up to a change of variable and pressure law, similar to the problem of the shallow water with a discontinuous topography, where these three solutions arise. In a more general framework, hyperbolic systems with resonant source term like ours have been investigate in [IT92] and [GL04], where the possible coexistence of three solutions is proved.

Suppose that this condition is reversed in

$$\forall\alpha,\ \forall\rho_1 < \rho_2 \leq \frac{|\alpha|}{c},\quad F_\alpha(\rho_1) - F_\alpha(\rho_2) \geq F_\alpha(\tilde{\rho_1}) - F_\alpha(\tilde{\rho_2}). \tag{4.20}$$

Fix $\alpha_L > 0$ and $\rho_L < \frac{\alpha_L}{c}$. For $0 \leq \theta_1 \leq \theta_2 \leq 1$, consider ρ_1 and ρ_2 such that

$$F_{\alpha_L}(\rho_L) - F_{\alpha_L}(\rho_i) = \theta_i,$$

and $\bar{\rho}_1$ and $\bar{\rho}_2$ such that

$$F_{\alpha_L}(\tilde{\rho}_i) - F_{\alpha_L}(\bar{\rho}_i) = 1 - \theta_i.$$

We recall that $\tilde{\rho}_i = \frac{\alpha_L^2}{c^2\rho_i}$. Then we have

$$\begin{aligned}
F_{\alpha_L}(\bar{\rho}_1) - F_{\alpha_L}(\bar{\rho}_2) &= (F_{\alpha_L}(\bar{\rho}_1) - F_{\alpha_L}(\tilde{\rho}_1)) + (F_{\alpha_L}(\tilde{\rho}_1) - F_{\alpha_L}(\tilde{\rho}_2)) + (F_{\alpha_L}(\tilde{\rho}_2) - F_{\alpha_L}(\bar{\rho}_2)) \\
&\leq -(1-\theta_1) + (F_{\alpha_L}(\rho_1) - F_{\alpha_L}(\rho_2)) + (1-\theta_2) \quad \text{with (4.20)} \\
&\leq 0.
\end{aligned}$$

It makes possible the coexistence of three facts that exclude each other under Hypothesis (4.22).

- First, there exists $(\rho_0, \alpha_L) \in \Omega_+^{sup}$ such that:

  $$F_{\alpha_L}(\rho_L) - F_{\alpha_L}(\rho_0) = 1.$$

  Thus $(\rho_-, \alpha_-) = (\rho_L, \alpha_L)$ and $(\rho_+, \alpha_+) = (\rho_0, \alpha_L)$ gives a solution that has no wave on the left of the particle, and two supersonic waves on its right, as depicted in Figure 4.10;

- Second,

  $$F_{\alpha_L}(\tilde{\rho_L}) - F_{\alpha_L}(\tilde{\rho_E}) \leq 1.$$

  If D is still an increasing function of α, we can apply the proof of Proposition 4.13 to obtain the existence of a pair of subsonic traces. This solution has a 1-wave on the left of the particle, and a 2-wave on its right, as depicted in Figure 4.11;

- Finally, the state reached on the right of the particle by jumping immediately ($\xi^0 = -\varepsilon/2$ or $\theta = 0$ in Theorem 4.3) is smaller than $\tilde{\rho_E}$, while the state reached by jumping in the end ($\xi^0 = \varepsilon/2$ or $\theta = 1$) is larger than $\tilde{\rho_E}$. Then $(\rho_-, \alpha_-) = (\rho_I, \alpha_L)$ and $(\rho_+, \alpha_+) = (\tilde{\rho_E}, \alpha_L)$ are admissible traces around the particle. The corresponding solution has no wave on its left, and just a 2-wave on its right, as depicted in Figure 4.12.

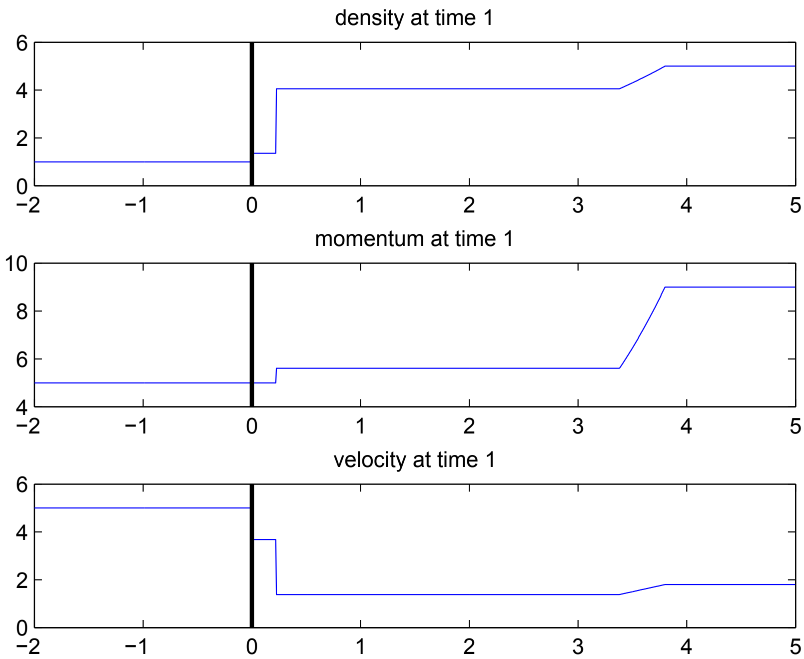


Figure 4.10. *Solution with two supersonic traces.*

Figures 4.10, 4.12 and 4.11 represent the three exact solutions at time $T = 1$ that are obtained for the Riemann problem

$$\begin{cases} \partial_t \rho + \partial_x q = 0, \\ \partial_t q + \partial_x \left(\frac{q^2}{\rho} + 4\rho \right) = -0.9\rho^2 u \delta_0, \\ \rho(0,x) = 1\,\mathbf{1}_{x<0} + 5\,\mathbf{1}_{x>0}, \\ q(0,x) = 5\,\mathbf{1}_{x<0} + 9\,\mathbf{1}_{x>0}. \end{cases} \tag{4.21}$$

In that case we have $F_\alpha(\rho) = \frac{\alpha}{2\rho^2} + \frac{c^2}{\alpha}\log(\rho)$. The traces have been numerically computed thanks to the previous analysis. Another way to see that we can lose uniqueness, and obtain up to three solutions, is depicted on Figure 4.13. Following [GL04], we introduce a merged 1-wave, which regroups all the state (ρ_I, α_I) that can be reached from (ρ_L, α_L) through three successive steps:

- From (ρ_L, α_L) we reach a state (ρ_-, α_-) on the left of the particle, by following a 1-wave and a 2-wave, both traveling at a speed smaller than the particle's velocity v:

$$(\rho_-, \alpha_-) \in V_-(\rho_L, \alpha_L, v);$$

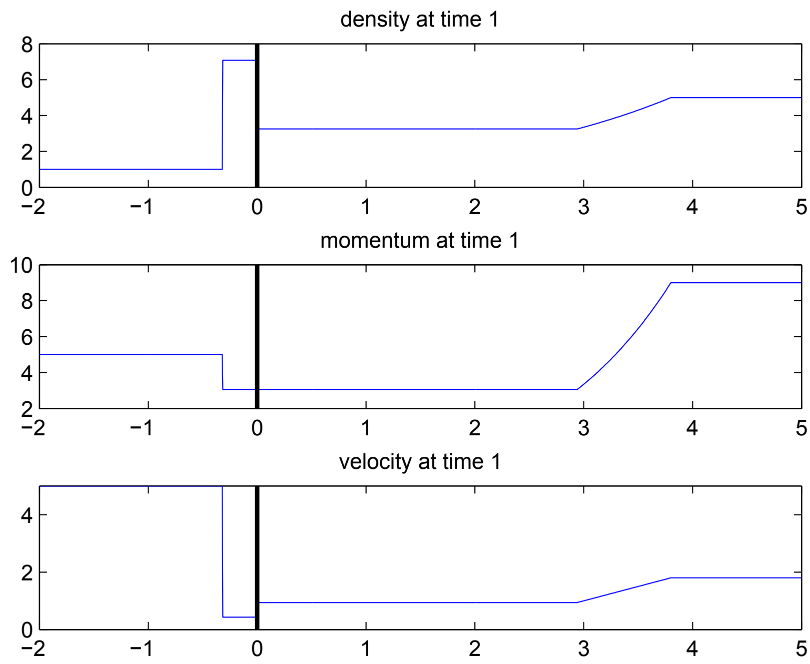


Figure 4.11. *Solution with two subsonic traces.*

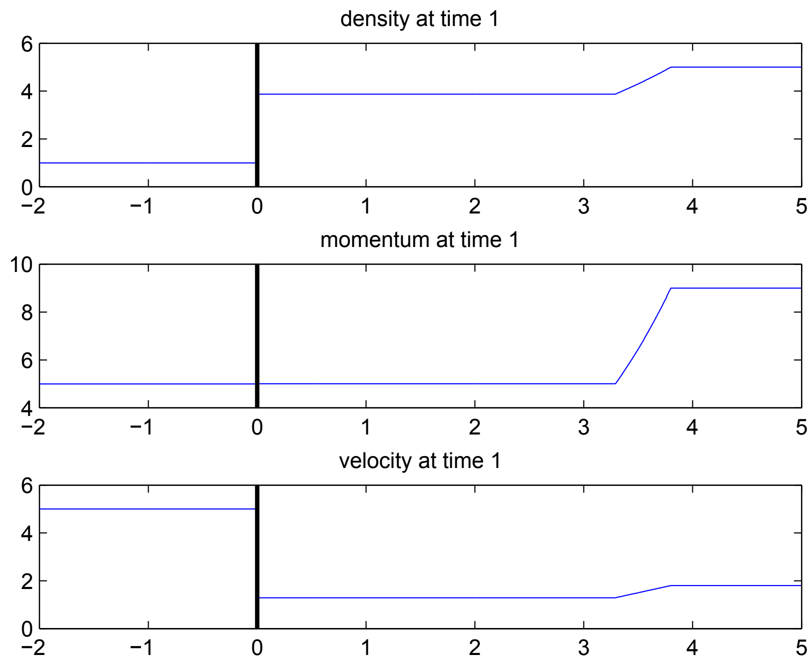


Figure 4.12. *Solution with one supersonic trace (on the left) and one subsonic trace (on the right).*

- From (ρ_-, α_-) we reach, through the particle, a state (ρ_+, α_+):

$$((\rho_-, \alpha_-), (\rho_+, \alpha_+)) \in \mathcal{G}_D(v);$$

- From (ρ_+, α_+) we reach (ρ_I, α_I) with a 1-wave traveling faster than v.

The solutions of the Riemann problem (**??**) are the intersections between this merged 1-wave and the usual curve of 2-waves arriving in (ρ_R, α_R). In the case of a supersonic left state the merged 1-wave contains three different types of state:

- Those obtained by taking $(\rho_-, \alpha_-) = (\rho_L, \alpha_L)$, then by decreasing continuously the quantity F_{α_L} of 1 inside the particle. In that case (ρ_+, α_+) is supersonic, and we can carry on with any 1-rarefaction wave and some 1-shocks to reach (ρ_I, α_I). This is part 1 of the black curve on Figures 4.13 and 4.14.

- Those obtained by taking $(\rho_-, \alpha_-) = (\rho_L, \alpha_L)$, then by decreasing continuously the quantity F_{α_L} of θ for some $\theta \in [0, 1]$, making a shock inside the particle, and finally continuously decreasing of $(1 - \theta)$ along F_{α_L}. In that case (ρ_+, q_+) is subsonic and there exists no 1-wave faster than v starting from (ρ_+, q_+). This is part 2 of the black curve on Figures 4.13 and 4.14.

- Those obtained by starting from (ρ_L, α_L) with a 1-shock slower than the particle. There exists no such 1-rarefaction wave and we reach a subsonic state (ρ_-, α_-), which lies on the dashed gray line on Figure 4.13. The state (ρ_+, α_+) is necessarily obtained by decreasing continuously of 1 along the graph of F_{α_-}. Therefore, (ρ_+, α_+) is subsonic and there exists no 1-wave faster than v starting from (ρ_+, q_+). This is part 3 of the black curve on Figures 4.13 and 4.14.

As we can see on Figures 4.13 and 4.14, the shape of the merged 1-wave depends on the relative positions of the densities $\rho_{0,+}$ and $\rho_{1,+}$. The Hypothesis (4.22) ensures that $\rho_{0,+} \leq \rho_{1,+}$, and the merged 1-wave curve can be parametrized by ρ as in Figure 4.14. If this hypothesis does not hold, it becomes possible that $\rho_{0,+} > \rho_{1,+}$, in which case the merged 1-wave curve has a Z-shape and can intersect the 2-waves curve up to three times as in Figure 4.13.

Remark 14. When D depends only on α, $\rho_{0,+} = \rho_{1,+}$ and the segment disappears.

4.5 Proof of Proposition 4.15 when D also depends on the density

When D also depends on ρ, the germ is much larger than when it depends only of α. For (ρ_-, α) is fixed, with $\alpha > 0$ and $c\rho_- < \alpha$, it contains one supersonic state (ρ_0, α), and a whole set of subsonic states (ρ, α) with ρ taking values in some interval $[\rho_{0+}, \rho_{1+}]$. When D depends only on α, this interval reduces to a single point $\rho_{0+} = \rho_{1+} = \tilde{\rho_0}$. In that case we only had to worry about the relative positions of ρ_0, ρ_E and ρ_L on the one hand, and of $\tilde{\rho_0}$, $\tilde{\rho_E}$ and $\tilde{\rho_L}$ on the other hand (see Figure 4.7 for the notation). Those relative positions were easy to deduce from one another. In the general case, we have to study the relative positions of ρ_0, ρ_E and ρ_L on the one hand, and of the whole interval $[\rho_{0+}, \rho_{1+}]$ with $\tilde{\rho_E}$ and $\tilde{\rho_L}$ on the other hand. The following property insures that those relative are linked to each other nicely.

Proposition 4.17. *if $|D|$ is a decreasing function of ρ, then for every $\alpha \in \mathbb{R}$, for all states (ρ_1, α) and (ρ_2, α) in $\mathbb{R}_+^* \times \mathbb{R}^*$ with $\rho_1 < \rho_2 \leq \frac{|\alpha|}{c}$ we have*

$$F_\alpha(\rho_1) - F_\alpha(\rho_2) \leq F_\alpha\left(\frac{\alpha^2}{c^2 \rho_1}\right) - F_\alpha\left(\frac{\alpha^2}{c^2 \rho_2}\right). \tag{4.22}$$

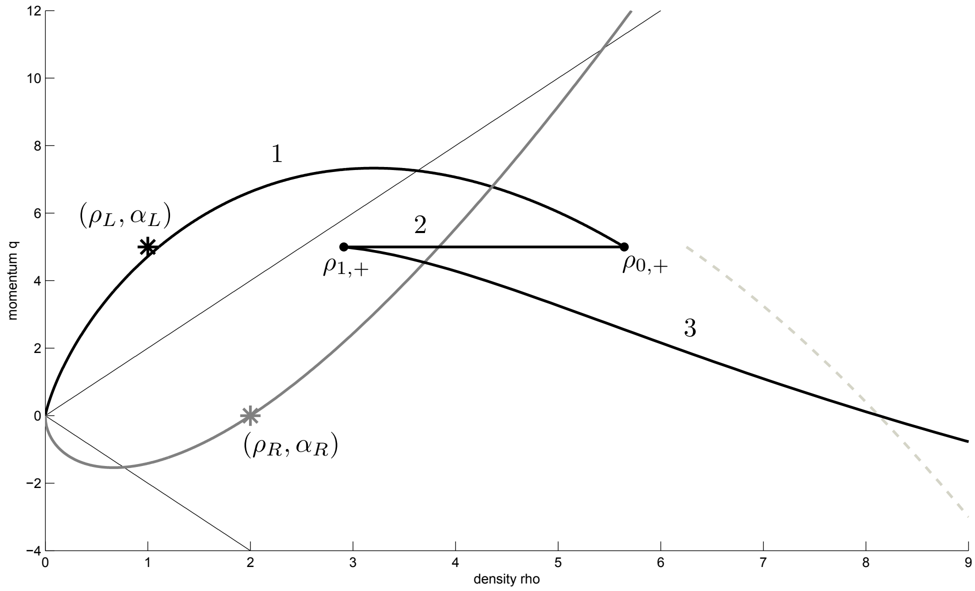


Figure 4.13. *Example of non uniqueness with the friction $D(\rho, q, h') = \rho^2(u - h') = \rho\alpha$. The gray line is the usual curve of 2-wave arriving in (ρ_R, α_R) (gray star). The black line is the merged 1-curve from (ρ_L, α_L) (black star).*

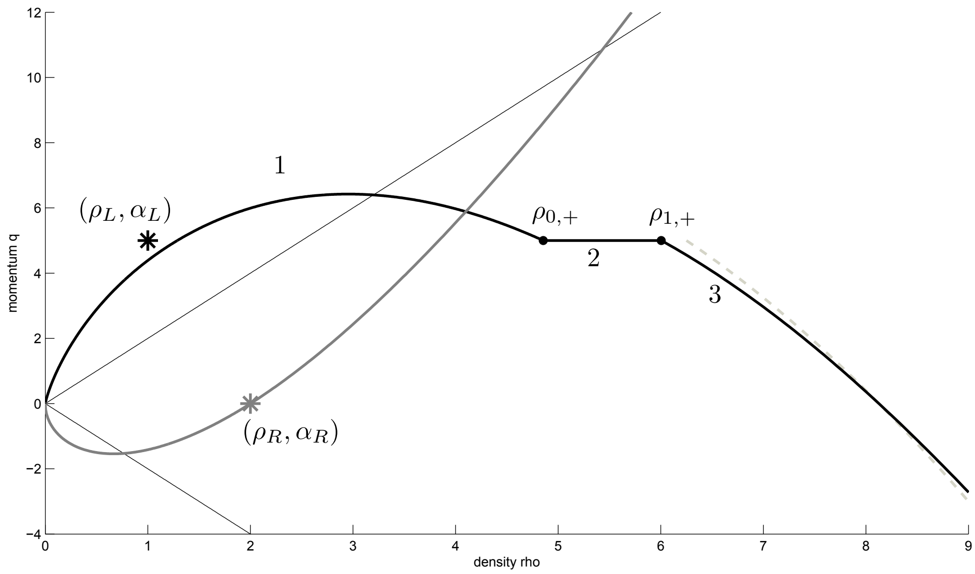


Figure 4.14. *Example of uniqueness with the friction $\Gamma(\rho, q, h') = u - h' = \frac{\alpha}{\rho}$. The gray line is the usual curve of 2-wave arriving in (ρ_R, α_R) (gray star). The black line is the merged 1-curve from (ρ_L, α_L) (black star).*

Proof. To prove (4.22) it is sufficient to prove that the function

$$\rho \mapsto F_\alpha(\rho) - F_\alpha\left(\frac{\alpha^2}{c^2\rho}\right)$$

increases on $(0, \frac{|\alpha|}{c})$. We compute its derivative:

$$\begin{aligned}
\frac{\partial}{\partial\rho}\left[F_\alpha(\rho) - F_\alpha\left(\frac{\alpha^2}{c^2\rho}\right)\right] &= F'_\alpha(\rho) + \frac{\alpha^2}{c^2\rho^2}F'_\alpha\left(\frac{\alpha^2}{c^2\rho}\right)\\
&= \frac{1}{|D|(\rho,\alpha)}\left(c^2 - \frac{\alpha^2}{\rho^2}\right) + \frac{\frac{\alpha^2}{c^2\rho^2}}{|D|(\frac{\alpha^2}{c^2\rho},\alpha)}\left(c^2 - \frac{\alpha^2}{\frac{\alpha^4}{c^4\rho^2}}\right)\\
&= \left[\frac{1}{|D|(\rho,\alpha)} - \frac{1}{|D|(\frac{\alpha^2}{c^2\rho},\alpha)}\right]\left(c^2 - \frac{\alpha^2}{\rho^2}\right).
\end{aligned}$$

On $(0, \frac{|\alpha|}{c})$, this quantity has same sign as

$$|D|(\rho,\alpha) - |D|\left(\frac{\alpha^2}{c^2\rho},\alpha\right),$$

which is positive if $|D|$ is a decreasing function of ρ, because $\rho \leq \frac{\alpha^2}{c^2\rho}$ on $(0, \frac{|\alpha|}{c})$. □

In the sequel we use the notation of Section 4.3.3 summarized on Figure 4.7. We recall in particular that for any (ρ,α), the state denoting by $(\tilde{\rho},\alpha)$ is reached with a shock at speed v: $\tilde{\rho} = \frac{\alpha^2}{c^2\rho}$. Let us first describe the form of the germ imposed by Hypothesis (4.22).

Lemma 4.18. *Suppose that the drag force D verifies the Hypothesis* (4.22)*. Fix (ρ_-, q_-) such that $\alpha_- = q_- - v\rho_-$ is positive and that $\rho_- \leq \frac{\alpha_-}{c}$. Then:*

- *If it exists $\rho_0 \leq \frac{\alpha_-}{c}$ such that $F_{\alpha_-}(\rho_-) - F_{\alpha_-}(\rho_0) = 1$, there exists a unique density $\rho_{1,+}$ greater than $\frac{\alpha_-}{c}$, such that $F_{\alpha_-}(\tilde{\rho_-}) - F_{\alpha_-}(\rho_{1,+}) = 1$. Then solutions of* (4.8) *in $\xi = \varepsilon/2$ can take the values (ρ_+, q_+) with*

  $$\rho_+ \in [\rho_{0,+}, \rho_{1,+}] \quad \textit{and} \quad q_+ = \alpha_- + v\rho_+,$$

  where we denote by $\rho_{0,+} = \tilde{\rho}_0$;

- *If ρ_0 does not exist but $\rho_{1,+}$ does, the solutions of* (4.8) *in $\xi = \varepsilon/2$ can take the values (ρ_+, q_+) with*

  $$\rho_+ \in \left[\frac{\alpha_-}{c}, \rho_{1,+}\right] \quad \textit{and} \quad q_+ = \alpha_- + v\rho_+,$$

  and we denote by $\rho_{0,+} = \frac{\alpha_-}{c}$;

- *If neither ρ_0 nor $\rho_{1,+}$ exist, system* (4.8) *does not admit any solution on the whole interval $(-\varepsilon/2, \varepsilon/2)$.*

Those three cases are illustrated in Figure 4.15 below.

Proof (Lemma 4.18). We first suppose that, as in top of Figure 4.16,

$$F_{\alpha_L}(\rho_L) - F_{\alpha_L}(\rho_E) \geq 1.$$

Then, there exists $\rho_0 \in [\rho_L, \rho_E]$ such that

$$F_{\alpha_L}(\rho_L) - F_{\alpha_L}(\rho_0) = 1,$$

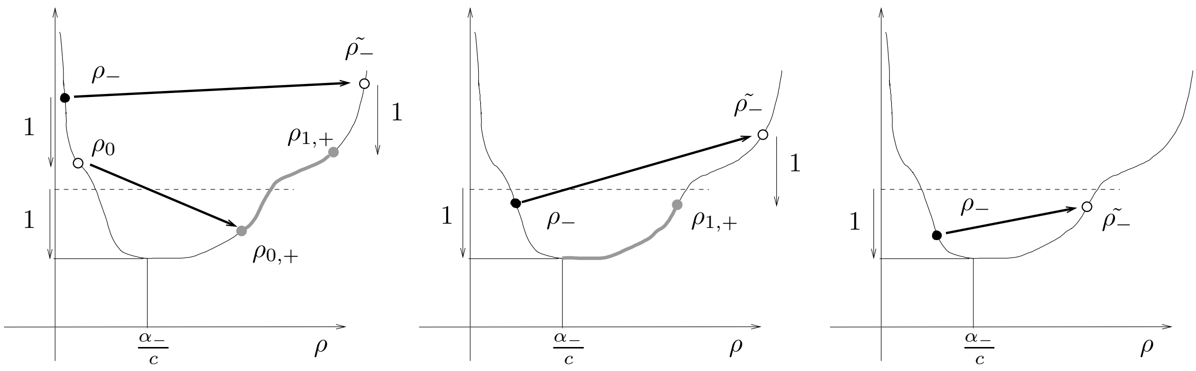


Figure 4.15. *The three cases of Lemma 4.18 from left to right. The bold arrows are entropy shocks at speed v.*

and the state (ρ_0, α_L) belongs to Ω_+^{sup} and provides an admissible state on the right of the particle. Let us prove that it is the unique solution. As $\rho_0 < \rho_E$, Hypothesis (4.22) yields $\rho_{0,+} > \tilde{\rho_E}$. We also proved in Lemma 4.18 that $\rho_{1,+} \geq \rho_{0,+}$. Therefore, the interval $[\rho_{0,+}, \rho_{1,+}]$ does not intersect Γ_+^{sub}. Eventually, Hypothesis (4.22) gives that $F_\alpha(\tilde{\rho_L}) - F_\alpha(\tilde{\rho_E}) \geq 1$. Thus we cannot choose any subsonic traces by Lemma 4.12. We now suppose that

$$F_{\alpha_L}(\rho_L) - F_{\alpha_L}(\rho_E) < 1.$$

If $F_\alpha(\tilde{\rho_L}) - F_\alpha(\tilde{\rho_E}) < 1$ (Figure 4.16, in the middle), which rewrites $\Delta(\alpha_L) < 0$, Lemma 4.12 implies that there exists a unique solution inside the subsonic triangle $|\alpha| < c$. Moreover, Hypothesis (4.22) yield $F_\alpha(\rho_L) - F_\alpha(\rho_E) < 1$, and $[\rho_{0,+}, \rho_{1,+}]$ is included in $[\frac{\alpha_L}{c}, \tilde{\rho_E}]$, so there is no solution with (ρ_L, α_L) for the left trace. Eventually, if $F_\alpha(\tilde{\rho_L}) - F_\alpha(\tilde{\rho_E}) > 1$ (at the bottom of Figure 4.16), there is no solution in the subsonic triangle. But in that case, $\rho_{1,+}$ exists and is greater than $\tilde{\rho_E}$, while $\rho_{0,+}$ is smaller than $\tilde{\rho_E}$ (and might be equal to $\frac{\alpha_L}{c}$). Therefore we can take (ρ_L, α_L) for the left trace and $(\tilde{\rho_E}, \alpha_L)$ for the right trace. □

4.6 A few words on the fully coupled system

We now concentrate on the fully coupled system (4.1), where the particle is free to move under the influence of the fluid. The solution has been defined in Section 4.2, Definition 4.2. We recall that the movement of the particle is governed by the ODE (4.14). In this section, we prove that the acceleration of the particle has the expected sign, i.e. that the particle accelerates if it travels slower than the fluid. We also prove that the total energy decreases through time.

A natural extension of Theorem 4.6 would be to prove existence and uniqueness to the Riemann problem for the fully coupled system:

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x\left(\rho u^2 + c^2 \rho\right) = -D(\rho, u, h'(t))\delta_{h(t)}(x), \\ m h''(t) = D(\rho, u, h'(t))(t, h(t)), \\ h(0) = h^{ini}, \; h'(0) = v^{ini} \\ \rho(0, x) = \rho_L \mathbf{1}_{x < h^{ini}} + \rho_R \mathbf{1}_{x > h^{ini}}, \\ q(0, x) = q_L \mathbf{1}_{x < h^{ini}} + q_R \mathbf{1}_{x > h^{ini}}. \end{cases} \tag{4.23}$$

However, such a result seems hard to obtain. Indeed, when the particle is free to move under the influence of the fluid, its trajectory is not a straight line anymore. Therefore, we cannot search a self-similar solution anymore. In the simpler case where the fluid is governed by the Burgers' equation, a solution can be explicitly constructed. This construction relies on the fact that either the state on the left of the particle, either the state on its right, remains constant on a small time interval. However, this is not the

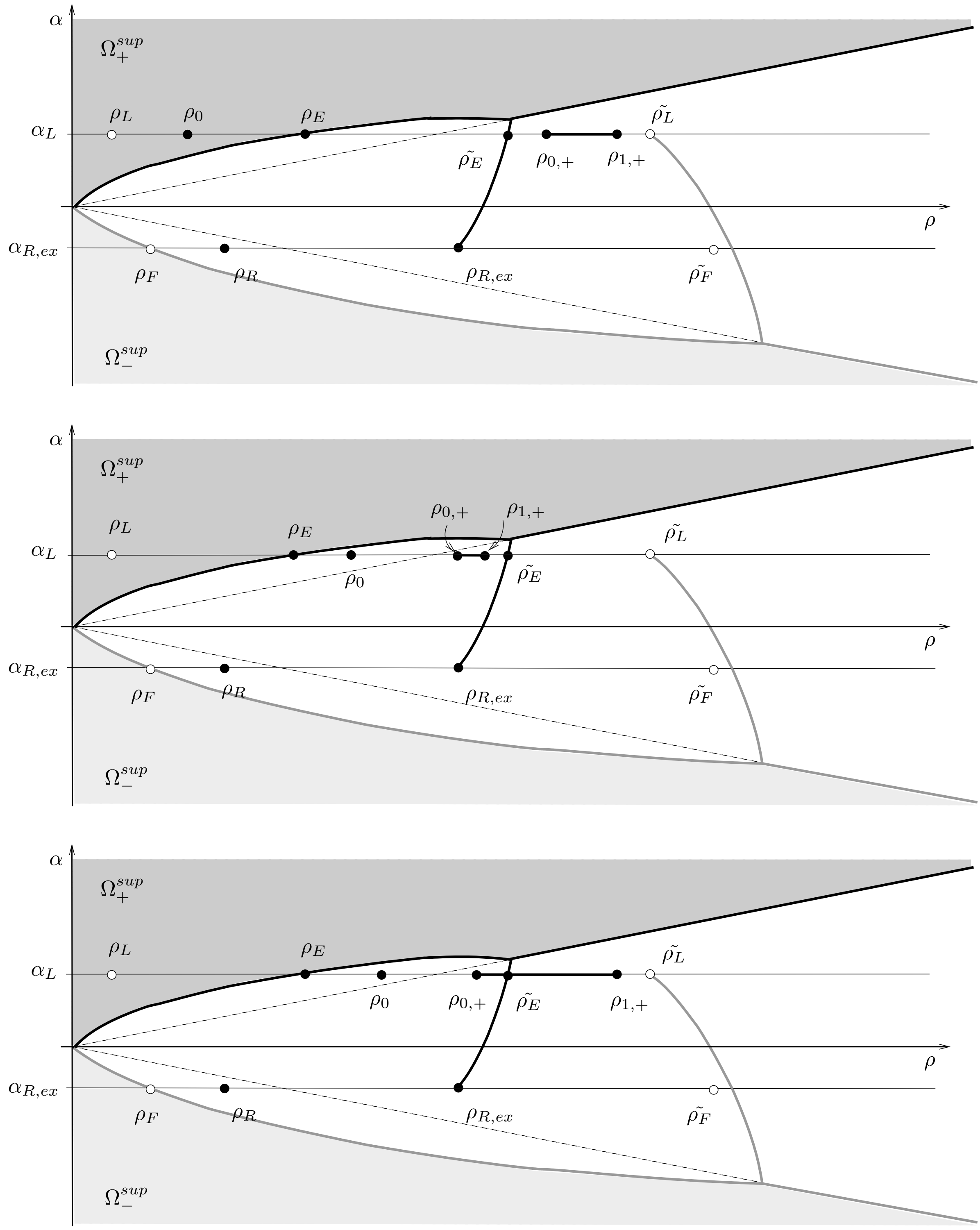


Figure 4.16. *Relative positions of* $\rho_{0,+}$ *and* $\rho_{1,+}$.

case when the fluid follows the Euler equations: the traces around the particle change at all time. This is why we believe that solving the Riemann problem for the fully coupled System (4.23) is as difficult as solving the Cauchy problem, say for initial data having total bounded variation. In [BCG14], R. Borsche, R. Colombo and M. Garavello were able to prove a local in time existence result when the initial data is small, subsonic, and when the initial traces around the particle belong to the germ. They use a wave front tracking method in the framework of PDE-ODE coupling, see [BCG14], [BCG12], [CG10] and references therein. Without the assumption of smallness, we were able to prove a local in time existence result for (4.23).

4.6.1 Sign of the particle's acceleration

We are now interested in the fully coupled system (4.1) (which is (4.23) with general initial data).The first line (4.1a) implies that the total mass

$$\partial_t \left(\int_{\mathbb{R}} \rho(t,x)dx + m \right) = 0$$

is conserved. As usual, we denote by (ρ_-, u_-) and (ρ_+, u_+) the left and right traces of the fluid's density and velocity around the particle. If (ρ, u, h) is a solution of (4.1), the conservation of total mass yields

$$\begin{aligned} 0 = \partial_t \int_{\mathbb{R}} \rho(t,x)dx &= \partial_t \int_{-\infty}^{h(t)} \rho(t,x)dx + \partial_t \int_{h(t)}^{+\infty} \rho(t,x)dx \\ &= \int_{-\infty}^{h(t)} \partial_t \rho(t,x)dx + \int_{h(t)}^{+\infty} \partial_t \rho(t,x)dx + h'(\rho_- - \rho_+) \\ &= -\int_{-\infty}^{h(t)} \partial_x(\rho u)(t,x)dx - \int_{h(t)}^{+\infty} \partial_x(\rho u)(t,x)dx + h'(\rho_- - \rho_+) \\ &= \rho_+ u_+ - \rho_- u_- + h'(\rho_- - \rho_+). \end{aligned}$$

Hence, the quantity $\rho(u - h')$ is conserved through the particle, and denoted as before by α. We recover the first point of Definition 4.1. In particular, as explained in the remark below, the phrase "the particle travels slower than the fluid" makes sense.

Remark 15. We always are in one of the following cases;

1. Either $u_+ < h'(t)$ and $u_- < h'(t)$: the fluid is slower than the particle (α<0).
2. Either $u_+ = h'(t)$ and $u_- = h'(t)$: the fluid and the particle have the same velocity (α=0).
3. Either $u_+ > h'(t)$ and $u_- > h'(t)$: the fluid is faster than the particle (α>0).

Combining formally the two last lines (4.1b) and (4.1c), we find that the total impulsion

$$\partial_t \left(\int_{\mathbb{R}} q(t,x)dx + mh'(t) \right) = 0$$

is also conserved. In Proposition 4.13, we used that fact to rewrite the ordinary differential equation driving the particle on the form

$$mh''(t) = c^2(\rho_- - \rho_+) \left(1 - \frac{(u_- - h')(u_+ - h')}{c^2} \right).$$

Unlike Expression (4.1c), this ODE makes sense if ρ and u are discontinuous along the particle's trajectory h, but it is not clear whether the particle's acceleration h'' has a consistent sign with Remark 15 or not.

Proposition 4.19. *For any drag force D having the same sign as α, the acceleration of the particle h'' and α have the same sign.*

Proof. Definition 4.1 yields that

- if $\alpha > 0$, and $u^- - h'(t) > c$, then $\rho_+ \in \left[\rho_-, \tilde{\rho_-} = \dfrac{\alpha^2}{c^2\rho_-}\right]$;
- if $\alpha > 0$, and $u^- - h'(t) < c$, then $\rho_+ \in \left[\dfrac{\alpha}{c}, \rho_-\right]$;
- if $\alpha = 0$, then $\rho_- = \rho_+$ and $u_- = u_+$.
- if $\alpha < 0$ and $u^+ - h'(t) < -c$, then $\rho_- \in \left[\rho_+, \tilde{\rho_+} = \dfrac{\alpha^2}{c^2\rho_+}\right]$;
- if $\alpha < 0$ and $u^+ - h'(t) < -c$, then$\rho_- \in \left[\frac{|\alpha|}{c}, \rho_+\right]$.

The case where α null is trivial. By symmetry, we only treat the case where α is stricly positive. If $u_- - h' > c$, we have that

$$u_+ - h' = \frac{\alpha}{\rho_+} \in \left[\frac{c^2\rho_-}{\alpha}, \frac{\alpha}{\rho_-} = u_- - h'\right]$$

and it follows that

$$(u_- - h')(u_+ - h') > (u_- - h')\frac{c^2\rho_-}{\alpha} = c^2.$$

Moreover, ρ_+ is bigger than ρ_-, which yields the positivity of h''. Eventually, if $u_- - h' < c$, we necessarily have $u_+ - h' < c$ and $\rho_+ < \rho_-$, and h'' is also non-negative. ☐

4.6.2 Decrease of the total energy

We now turn to the decrease of the total energy

$$\mathcal{E}(t) = \int_{\mathbb{R}} E(\rho, q)(t, x)dx + \frac{m}{2}\left(h'(t)\right)^2,$$

where $E = \frac{1}{2}\rho u^2 + c^2\rho\ln(\rho)$ denotes the energy of the fluid, as previously. We denote by $G = u(E + c^2\rho)$ the energy flux. The pair (E, G) is a flux- entropy flux pair for the barotropic Euler equation. We denote by $f^\rho(\rho, q) = q$ and $f^q(\rho, q) = \frac{q^2}{\rho} + c^2\rho$ the two component of the flux F associated to this equation. A formal computation yields

$$\begin{aligned}
\partial_t \int_{\mathbb{R}} E(\rho, q)(t, x)dx &= \int_{\mathbb{R}} \partial_t E(\rho, q)(t, x)dx \\
&= \int_{\mathbb{R}} \nabla_{(\rho,q)} E \cdot \begin{pmatrix} \partial_t \rho \\ \partial_t q \end{pmatrix} dx \\
&= -\int_{\mathbb{R}} \nabla_{(\rho,q)} E \cdot \left[\begin{pmatrix} \partial_x f^\rho(\rho, q) \\ \partial_x f^q(\rho, q) \end{pmatrix} + \begin{pmatrix} 0 \\ D(\rho, u, h')\delta_{h(t)}(x) \end{pmatrix}\right] dx \\
&= -\int_{\mathbb{R}} \nabla_{(\rho,q)} E \cdot D_{(\rho,q)} F \cdot \left[\begin{pmatrix} \partial_x \rho \\ \partial_x q \end{pmatrix} + \begin{pmatrix} 0 \\ D(\rho, u, h')\delta_{h(t)}(x) \end{pmatrix}\right] dx \\
&= -\int_{\mathbb{R}} \nabla_{(\rho,q)} G \cdot \begin{pmatrix} \partial_x \rho \\ \partial_x q \end{pmatrix} + \nabla_{(\rho,q)} E \cdot \begin{pmatrix} 0 \\ D(\rho, u, h')\delta_{h(t)}(x) \end{pmatrix} dx \\
&= -\int_{\mathbb{R}} \partial_x G(\rho, q)dx - u(t, h(t))D(\rho, u, h')(t, h(t)) \\
&= -mu(t, h(t))h''(t),
\end{aligned}$$

and therefore

$$\partial_t \mathcal{E}(t) = -mh''(t)(u(t,h(t)) - h'(t)).$$

In particular, as h'' and $u - h'$ have the same sign (formally with (4.1c), and rigorously with Proposition 4.19), the total energy is decreasing. The following proposition states this result in a more rigorous framework, and completes Remark 10.

Proposition 4.20. *If the particle is driven by the ODE* (4.14) *recall below,*

$$mh''(t) = c^2(\rho_- - \rho_+)\left(1 - \frac{(u_- - h')(u_+ - h')}{c^2}\right),$$

then the total energy $\mathcal{E}$ decreases.

Proof. We denote by (E_-, G_-) and (E_+, G_+) the left and right traces of the energy and energy flux of the fluid around the particle. Cutting the integral on $\{x < h\}$ and $\{x > h\}$, we obtain

$$\begin{aligned}
\mathcal{E}' &= h'(E_- - E_+) + (G_+ - G_-) + h'(\rho_- - \rho_+)\left(c^2 - \frac{\alpha^2}{\rho_-\rho_+}\right)\\
&= (h' - u_-)E_- - (h' - U_+)E_+ + c^2(\rho_+ u_+ - \rho_- u_-) + h'(\rho_- - \rho_+)\left(c^2 - \frac{\alpha^2}{\rho_-\rho_+}\right)\\
&= \alpha\left(\frac{E_+}{\rho_+} - \frac{E_-}{\rho_-}\right) - h'(\rho_- - \rho_+)\frac{\alpha^2}{\rho_-\rho_+} \quad \text{because} \quad \rho_+ u_+ - \rho_- u_- = h'(\rho_+ - \rho_-)\\
&= \alpha\left[\frac{u_+^2}{2} - \frac{u_-^2}{2} + c^2(\ln(\rho_+) - \ln(\rho_-))\right] - h'(\rho_- - \rho_+)\frac{\alpha^2}{\rho_-\rho_+}\\
&= \alpha\left[\frac{1}{2}\left(\frac{\alpha^2}{\rho_+^2} + 2h'\frac{\alpha}{\rho_+} + h'^2\right) - \frac{1}{2}\left(\frac{\alpha^2}{\rho_-^2} + 2h'\frac{\alpha}{\rho_-} + h'^2\right) + c^2(\ln(\rho_+) - \ln(\rho_-))\right] - h'(\rho_- - \rho_+)\frac{\alpha^2}{\rho_-\rho_+}\\
&= \alpha\left[\alpha^2\frac{(\rho_-^2 - \rho_+^2)}{2\rho_-^2\rho_+^2} + \alpha\frac{h'(\rho_- - \rho_+)}{\rho_-\rho_+} + c^2(\ln(\rho_+) - \ln(\rho_-))\right] - h'(\rho_- - \rho_+)\frac{\alpha^2}{\rho_-\rho_+}\\
&= \alpha\left[\alpha^2\frac{(\rho_-^2 - \rho_+^2)}{2\rho_-^2\rho_+^2} + c^2(\ln(\rho_+) - \ln(\rho_-))\right]
\end{aligned}$$

Let us analyze the sign of this last quantity. By symmetry, we can suppose that α is strictly positive. We first treat the case where $u_- - h'$ is larger than the speed of sound c, or equivalently where $\rho_- < \frac{\alpha}{c}$. In that case, ρ_+ belongs to the interval $\left[\rho_-, \frac{\alpha^2}{c^2\rho_-}\right]$. We introduce the function

$$s(\rho) = \alpha^2\frac{(\rho_-^2 - \rho^2)}{2\rho_-^2\rho^2} + c^2(\ln(\rho) - \ln(\rho_-))$$

Its derivative is

$$s'(\rho) = -\frac{\alpha^2}{\rho^3} + \frac{c^2}{\rho} = \frac{1}{\rho}\left(c^2 - \frac{\alpha^2}{\rho^2}\right),$$

therefore s decreases then increases, reaching its minimum in $\frac{\alpha}{c}$. Moreover, $s(\rho_-)$ is null and

$$s\left(\frac{\alpha^2}{c^2\rho_-}\right) = \frac{\rho_-^2 c^4}{2\alpha^2} - \frac{\alpha^2}{2\rho_-^2} + c^2\ln\left(\frac{\alpha^2}{c^2\rho_-^2}\right).$$

With $X = \frac{\alpha^2}{c^2\rho_-^2}$, which is larger than 1, $s\left(\frac{\alpha^2}{c^2\rho_-}\right)$ rewrites

$$s\left(\frac{\alpha^2}{c^2\rho_-}\right) = g(X) = c^2\left(\frac{1}{2X} - \frac{X}{2} + \ln(X)\right).$$

We check that $g(1)=0$ and that

$$g'(X)=\frac{c^2}{X}-\frac{c^2}{2}-\frac{c^2}{2X^2}=-\frac{c^2}{2}\left(\frac{1}{X}-1\right)^2<0.$$

Therefore, s is negative on $\left[\rho_-, \frac{\alpha^2}{c^2\rho_-}\right]$, which concludes the proof in that case. Let us now treat the case where $u_- - h' \leq c$. We have that $u_+ - h' \leq c$ and $\frac{\alpha}{c} \leq \rho_+ < \rho_-$. In particular, $\frac{\alpha^2}{\rho_+^2} \leq c^2$ and we obtain:

$$\begin{aligned} s(\rho_+) &= \alpha^2\frac{(\rho_-^2-\rho_+^2)}{2\rho_-^2\rho_+^2}+c^2(\ln(\rho_+)-\ln(\rho_-)) \\ &\leq \frac{c^2}{2}\left(1-\frac{\rho_+^2}{\rho_-^2}\right)+c^2\ln\left(\frac{\rho_+}{\rho_-}\right). \end{aligned}$$

Introducing $X=\frac{\rho_+}{\rho_-}<1$ allows us to write $s(\rho_+)=c^2g(X)$, with

$$g(X)=\frac{1}{2}(1-X^2)+\ln(X).$$

We have that $g(1)=0$ and we check that g increases on $(0,1)$. Therefore, $s(\rho_+)$ is negative, which concludes the proof. ☐

4.6.3 Coupled Riemann problem in the subsonic non-characteristic case

In the sequel we prove the local in time existence of a solution to the following Riemann (4.23). We consider the particular case where the drag force is linear:

$$D(\rho,u,h')=\lambda\rho(u-h')$$

and where the initial data are subsonic: $|u_L-v^{ini}|<c$ and $|u_R-v^{ini}|<c$. Moreover, we suppose that $((\rho_L,u_L),(\rho_R,u_R))$ belongs to the germ for a particle at speed v^{ini}. As usual, we denote by $q=\rho u$ the momentum of the fluid and by α the quantity $q-v\rho$, which is conserved through the particle.

Proposition 4.21. *Assume that at some time t^0, the particle has speed v^0 and the left and right traces (ρ_-^0,u_-^0) and (ρ_+^0,u_+^0) belong to the germ at speed v^0, and that $|u_\pm^0-c|<0$. Assume that the same holds at time $t^1>t^0$ and denote with an exponent 1 the different quantities at this time. Moreover, suppose that (ρ_-^1,q_-^1) is linked to (ρ_-^0,q_-^0) by a 1-wave, and that (ρ_+^1,q_+^1) is linked to (ρ_+^0,q_+^0) by a 2-wave. The variations $\Delta\rho_-=\rho_-^1-\rho_-^0$, $\Delta\rho_+$, $\Delta\alpha$ and Δv verify at the first order the linear system:*

$$\begin{pmatrix} 1 & c-\frac{\alpha^0}{\rho_-^0} & 0 & \rho_-^0 \\ 1 & 0 & -\frac{\alpha^0}{\rho_+^0}-c & \rho_+^0 \\ \frac{2\alpha^0}{\rho_-^0}-\frac{2\alpha^0}{\rho_+^0}-\lambda & c^2-\frac{(\alpha^0)^2}{(\rho_-^0)^2} & \frac{(\alpha^0)^2}{(\rho_+^0)^2}-c^2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix}\Delta\alpha \\ \Delta\rho_- \\ \Delta\rho_+ \\ \Delta v\end{pmatrix}=\begin{pmatrix}0\\0\\0\\\frac{\lambda}{m}\alpha^0\Delta t\end{pmatrix} \tag{4.24}$$

Proof. According to Remark 9 it is not necessary to reformulate the ODE driving the particle and we can use the simple expression

$$mh''(t)=\lambda\alpha(t).$$

A first order Taylor expansion of this relation gives

$$m\frac{\Delta v}{\Delta t}=\lambda\alpha^0.$$

We assumed that (ρ_-^1,q_-^1) lies on the 1-wave curve passing through (ρ_-^0,q_-^0). The slope of this curve parametrizing the momentum q as a function of ρ, is $u-c$. Therefore we can write

$$\begin{aligned}\alpha^1 &= q_-^1 - v^1\rho_-^1\\ &= q_-^0 + (u_-^0 - c)\Delta\rho_- + O(\Delta\rho_-) - (v^0 + \Delta v + O(\Delta v))(\rho_-^0 + \Delta\rho_- + O(\Delta\rho_-))\\ &= \alpha^0 + \left(\frac{\alpha^0}{\rho_-^0} - c\right)\Delta\rho_- - \rho_-^0\Delta v + O(\Delta v) + O(\Delta\rho_-).\end{aligned}$$

Similarly, we assumed that (ρ_+^1, q_+^1) lies on the 2-wave curve passing through (ρ_+^0, q_+^0):

$$\begin{aligned}\alpha^1 &= q_+^1 - v^1\rho_+^1\\ &= q_+^0 + (u_+^0 + c)\Delta\rho_+ + O(\Delta\rho_+) - (v^0 + \Delta v + O(\Delta v))(\rho_+^0 + \Delta\rho_+ + O(\Delta\rho_+))\\ &= \alpha^0 + \left(\frac{\alpha^0}{\rho_+^0} + c\right)\Delta\rho_+ - \rho_+^0\Delta v + O(\Delta v) + O(\Delta\rho_+).\end{aligned}$$

Eventually, a first order Taylor expansion of the relation (4.12) on the charge yields

$$\begin{aligned}\left(\frac{(\alpha^1)^2}{\rho_-^1} + c^2\rho_-^1\right) - \left(\frac{(\alpha^1)^2}{\rho_+^1} + c^2\rho_+^1\right) - \lambda\alpha^1 &= \left(\frac{(\alpha^0)^2}{\rho_-^0} + c^2\rho_-^0\right) - \left(\frac{(\alpha^0)^2}{\rho_+^0} + c^2\rho_+^0\right) - \lambda\alpha^0\\ &\quad + \left(c^2 - \frac{(\alpha^0)^2}{(\rho_-^0)^2}\right)\Delta\rho_- + \left(\frac{(\alpha^0)^2}{(\rho_+^0)^2} - c^2\right)\Delta\rho_+\\ &\quad + \left(\frac{2\alpha^0}{\rho_-^0} - \frac{2\alpha^0}{\rho_+^0} - \lambda\right)\Delta\alpha\\ &\quad + O(\Delta\alpha) + O(\Delta\rho_-) + O(\Delta\rho_+).\end{aligned}$$

We suppose that both pairs of traces belong to the germ associated to their particle's velocities. Therefore the two terms on the first line are null and we obtain the last line of system (4.24). $\square$

Proposition 4.22. *If the hypothesis of Proposition 4.21 hold for every time t^0 and t^1 in the interval $(0,T)$, then on this interval ρ_-, ρ_+, α and $v = h'$ verify the following system of ODEs:*

$$\begin{cases}\alpha'(t) &= -\frac{\lambda\alpha(t)}{m}\dfrac{\rho_-(t)\left(c+\frac{\alpha(t)}{\rho_-(t)}\right)+\rho_+(t)\left(c-\frac{\alpha(t)}{\rho_+(t)}\right)}{2c+\lambda-\alpha(t)\left(\frac{1}{\rho_-(t)}-\frac{1}{\rho_+(t)}\right)},\\ \rho_-'(t) &= -\frac{\lambda\alpha(t)}{cm}\dfrac{\left(c+\frac{\alpha(t)}{\rho_-(t)}\right)\left[\lambda\rho_-(t)+\alpha(t)\left(\frac{\rho_-(t)}{\rho_+(t)}-1\right)\right]}{\left(c-\frac{\alpha(t)}{\rho_-(t)}\right)\left[2c+\lambda-\alpha(t)\left(\frac{1}{\rho_-(t)}-\frac{1}{\rho_+(t)}\right)\right]},\\ \rho_+'(t) &= \frac{\lambda\alpha(t)}{cm}\dfrac{\left(c-\frac{\alpha(t)}{\rho_+(t)}\right)\left[\rho_+(t)\lambda+\alpha(t)\left(1-\frac{\rho_+(t)}{\rho_-(t)}\right)\right]}{\left(c+\frac{\alpha(t)}{\rho_+(t)}\right)\left[2c+\lambda-\alpha(t)\left(\frac{1}{\rho_-(t)}-\frac{1}{\rho_+(t)}\right)\right]},\\ v_-'(t) &= \frac{\lambda\alpha(t)}{m}\end{cases}\tag{4.25}$$

It follows that if $\alpha > 0$, α and ρ_- decrease and h' and ρ_+ increase, while if $\alpha < 0$, α and ρ_- increase and h' and ρ_+ decrease. Moreover we have that:

$$\begin{cases}u_+'(t) &= \frac{\lambda c^2}{m}(\rho_-(t)-\rho_+(t))\dfrac{c\rho_+(t)-\alpha(t)}{(\alpha(t)+c\rho_+(t))\left[2c+\lambda-\alpha(t)\left(\frac{1}{\rho_-(t)}-\frac{1}{\rho_+(t)}\right)\right]},\\ u_-'(t) &= \frac{\lambda c^2}{m}(\rho_-(t)-\rho_+(t))\dfrac{c\rho_-(t)+\alpha(t)}{(c\rho_-(t)-\alpha(t))\left[2c+\lambda-\alpha(t)\left(\frac{1}{\rho_-(t)}-\frac{1}{\rho_+(t)}\right)\right]},\end{cases}\tag{4.26}$$

and u_+ and u_- increase if $\alpha > 0$ and decrease if $\alpha < 0$.

Proof. Dividing by $t^1 - t^0$ and letting t^1 tends to t^0, Proposition 4.21 yields

$$\begin{pmatrix} 1 & c - \frac{\alpha}{\rho_-} & 0 & \rho_- \\ 1 & 0 & -\frac{\alpha}{\rho_+} - c & \rho_+ \\ \frac{2\alpha}{\rho_-} - \frac{2\alpha}{\rho_+} - \lambda & c^2 - \frac{\alpha^2}{(\rho_-)^2} & \frac{\alpha^2}{(\rho_+)^2} - c^2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha' \\ \rho'_- \\ \rho'_+ \\ h'' \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{\lambda}{m}\alpha \end{pmatrix}. \tag{4.27}$$

We denote by M this matrix, and $\tilde{M}$ the comatrix of M. To solve this system we need to compute $\tilde{M}_{41}$, $\tilde{M}_{42}$, $\tilde{M}_{43}$ and $\tilde{M}_{44}$ which is also the determinant of M. We first show that this determinant is not zero:

$$\begin{aligned} \begin{vmatrix} 1 & c - \frac{\alpha}{\rho_-} & 0 \\ 1 & 0 & -\frac{\alpha}{\rho_+} - c \\ \frac{2\alpha}{\rho_-} - \frac{2\alpha}{\rho_+} - \lambda & c^2 - \frac{\alpha^2}{(\rho_-)^2} & \frac{\alpha^2}{(\rho_+)^2} - c^2 \end{vmatrix} &= \left(\frac{\alpha}{\rho_+} + c\right)\left[\left(c^2 - \frac{\alpha^2}{(\rho_-)^2}\right) + \left(\frac{\alpha}{\rho_-} - c\right)\left(\frac{2\alpha}{\rho_-} - \frac{2\alpha}{\rho_+} - \lambda\right)\right] \\ &\quad + \left(c^2 - \frac{\alpha^2}{(\rho_+)^2}\right)\left(\frac{\alpha}{\rho_-} - c\right) \\ &= \left(c + \frac{\alpha}{\rho_+}\right)\left(c - \frac{\alpha}{\rho_-}\right)\left[\left(c + \frac{\alpha}{\rho_-}\right) - \left(\frac{2\alpha}{\rho_-} - \frac{2\alpha}{\rho_+} - \lambda\right) + \left(c - \frac{\alpha}{\rho_+}\right)\right] \\ &= \left(c + \frac{\alpha}{\rho_+}\right)\left(c - \frac{\alpha}{\rho_-}\right)\left[2c + \lambda + \alpha\left(\frac{1}{\rho_+} - \frac{1}{\rho_-}\right)\right]. \end{aligned}$$

As (ρ_-, α) and (ρ_+, α) belongs to the triangle $-c\rho < \alpha < c\rho$ we have that

$$c - \frac{\alpha}{\rho_\pm} > 0 \quad \text{and} \quad c + \frac{\alpha}{\rho_\pm} < 0.$$

Moreover $\rho_- - \rho_+$ has the same sign than α, so $\alpha\left(\frac{1}{\rho_+} - \frac{1}{\rho_-}\right)$ is always positive. Let us now compute $-\tilde{M}_{41}$:

$$\begin{aligned} \begin{vmatrix} c - \frac{\alpha}{\rho_-} & 0 & \rho_- \\ 0 & -\frac{\alpha}{\rho_+} - c & \rho_+ \\ c^2 - \frac{\alpha^2}{(\rho_-)^2} & \frac{\alpha^2}{(\rho_+)^2} - c^2 & 0 \end{vmatrix} &= \rho_- \left(c + \frac{\alpha}{\rho_+}\right)\left(c^2 - \frac{\alpha^2}{(\rho_-)^2}\right) + \rho_+\left(c - \frac{\alpha}{\rho_-}\right)\left(c^2 - \frac{\alpha^2}{(\rho_+)^2}\right) \\ &= \left(c + \frac{\alpha}{\rho_+}\right)\left(c - \frac{\alpha}{\rho_-}\right)\left[\rho_-\left(c + \frac{\alpha}{\rho_-}\right) + \rho_+\left(c - \frac{\alpha}{\rho_+}\right)\right] \\ &> 0\,. \end{aligned}$$

The coefficient $\tilde{M}_{42}$ writes

$$\begin{aligned} \begin{vmatrix} 1 & 0 & \rho_- \\ 1 & -\frac{\alpha}{\rho_+} - c & \rho_+ \\ \frac{2\alpha}{\rho_-} - \frac{2\alpha}{\rho_+} - \lambda & \frac{\alpha^2}{(\rho_+)^2} - c^2 & 0 \end{vmatrix} &= \rho_-\left(\frac{\alpha^2}{(\rho_+)^2} - c^2 + \left(\frac{\alpha}{\rho_+} + c\right)\left(\frac{2\alpha}{\rho_-} - \frac{2\alpha}{\rho_+} - \lambda\right)\right) - \rho_+\left(\frac{\alpha^2}{(\rho_+)^2} - c^2\right) \\ &= \left(\frac{\alpha}{\rho_+} + c\right)\left[\rho_-\left(\frac{\alpha}{\rho_+} - c\right) + \rho_-\left(\frac{2\alpha}{\rho_-} - \frac{2\alpha}{\rho_+} - \lambda\right) - \rho_+\left(\frac{\alpha}{\rho_+} - c\right)\right] \\ &= \left(\frac{\alpha}{\rho_+} + c\right)\left[c(\rho_+ - \rho_-) + \alpha\left(1 - \frac{\rho_-}{\rho_+}\right) - \lambda\rho_-\right]. \end{aligned}$$

When the drag force depend only of α, (4.12) yields that

$$\alpha^2\left(\frac{1}{\rho_-} - \frac{1}{\rho_+}\right) + c^2(\rho_- - \rho_+) = \lambda\alpha$$

and

$$c(\rho_+ - \rho_-) = \frac{\alpha^2}{c}\left(\frac{1}{\rho_-} - \frac{1}{\rho_+}\right) - \frac{\lambda}{c}\alpha.$$

It follows that

$$\tilde{M}_{42} = -\frac{1}{c}\left(\frac{\alpha}{\rho_+} + c\right)\left(\frac{\alpha}{\rho_-} + c\right)\rho_-\left[\lambda + \alpha\left(\frac{1}{\rho_+} - \frac{1}{\rho_-}\right)\right],$$

which is negative. Finally we compute $-\tilde{M}_{43}$:

$$\begin{aligned}
\begin{vmatrix} 1 & c - \frac{\alpha}{\rho_-} & \rho_- \\ 1 & 0 & \rho_+ \\ \frac{2\alpha}{\rho_-} - \frac{2\alpha}{\rho_+} - \lambda & c^2 - \frac{\alpha^2}{(\rho_-)^2} & 0 \end{vmatrix} &= \rho_-\left(c^2 - \frac{\alpha^2}{(\rho_-)^2}\right) - \rho_+\left[\left(c^2 - \frac{\alpha^2}{(\rho_-)^2}\right) - \left(c - \frac{\alpha}{\rho_-}\right)\left(\frac{2\alpha}{\rho_-} - \frac{2\alpha}{\rho_+} - \lambda\right)\right] \\
&= \left(c - \frac{\alpha}{\rho_-}\right)\left[\rho_-\left(c + \frac{\alpha}{\rho_-}\right) - \rho_+\left(c + \frac{\alpha}{\rho_-}\right) + \rho_+\left(\frac{2\alpha}{\rho_-} - \frac{2\alpha}{\rho_+} - \lambda\right)\right] \\
&= \left(c - \frac{\alpha}{\rho_-}\right)\left[c(\rho_- - \rho_+) + \alpha\left(\frac{\rho_+}{\rho_-} - 1\right) - \lambda\rho_+\right] \\
&= -\frac{1}{c}\left(c - \frac{\alpha}{\rho_-}\right)\left(c - \frac{\alpha}{\rho_+}\right)\rho_+\left[\alpha\left(\frac{1}{\rho_+} - \frac{1}{\rho_-}\right) - \lambda\right] \\
&\leq 0\,.
\end{aligned}$$

Considering that

$$\begin{pmatrix} \alpha \\ \rho_- \\ \rho_+ \\ h'' \end{pmatrix} = \frac{\lambda\alpha}{m\det(M)}\begin{pmatrix} \tilde{M}_{41} \\ \tilde{M}_{42} \\ \tilde{M}_{43} \\ \det(M) \end{pmatrix}$$

we obtain first part of the result. For the differential equations on the velocities, we denote by B the quantity

$$B = \left[2c + \lambda - \alpha\left(\frac{1}{\rho_-} - \frac{1}{\rho_+}\right)\right] = 2c + \frac{c^2}{\alpha}(\rho_- - \rho_+),$$

and we compute the derivatives of u_- and u_+:

$$\begin{aligned}
u'_+ &= \frac{\alpha'}{\rho_+} - \frac{\alpha\rho'_+}{\rho_+^2} + v' \\
&= \frac{\lambda\alpha}{\rho_+ m}\left(-\frac{c(\rho_- + \rho_+)}{B} - \frac{\alpha}{c}\frac{(c\rho_+ - \alpha)\left[\lambda + \alpha\left(\frac{1}{\rho_+} - \frac{1}{\rho_-}\right)\right]}{(c\rho_+ + \alpha)\,B} + \rho_+\right) \\
&= \frac{\lambda\alpha}{\rho_+ mB}\left(\rho_+\left[2c + \frac{c^2}{\alpha}(\rho_- - \rho_+)\right] - c(\rho_+ + \rho_-) - \frac{\alpha}{c}\frac{(c\rho_+ - \alpha)\,\frac{c^2}{\alpha}(\rho_- - \rho_+)}{(c\rho_+ + \alpha)}\right) \\
&= \frac{c\lambda\alpha}{\rho_+ mB}(\rho_+ - \rho_-)\left[1 - \frac{\rho_+ c}{\alpha} + \frac{c\rho_+ - \alpha}{c\rho_+ + \alpha}\right] \\
&= \frac{c^2\lambda\alpha}{mB}(\rho_+ - \rho_-)\frac{\alpha - c\rho_+}{\alpha + c\rho_+},
\end{aligned}$$

and

$$
\begin{aligned}
u'_- &= \frac{\alpha'}{\rho_-} - \frac{\alpha\rho'_-}{\rho_-^2} + v' \\
&= \frac{\lambda\alpha}{\rho_- m}\left(-\frac{c(\rho_- + \rho_+)}{B} + \frac{\alpha}{c}\frac{(c\rho_- + \alpha)\left[\lambda + \alpha\left(\frac{1}{\rho_+} - \frac{1}{\rho_-}\right)\right]}{(c\rho_- - \alpha)\,B} + \rho_-\right) \\
&= \frac{\lambda\alpha}{\rho_- mB}\left(\rho_-\left[2c + \frac{c^2}{\alpha}(\rho_- - \rho_+)\right] - c(\rho_+ + \rho_-) + \frac{\alpha}{c}\frac{(c\rho_- + \alpha)\,\frac{c^2}{\alpha}(\rho_- - \rho_+)}{(c\rho_- - \alpha)}\right) \\
&= \frac{c\lambda\alpha}{\rho_- mB}(\rho_- - \rho_+)\left[1 + \frac{\rho_- c}{\alpha} + \frac{c\rho_- + \alpha}{c\rho_- - \alpha}\right] \\
&= \frac{c^2\lambda\alpha}{mB}(\rho_- - \rho_+)\frac{c\rho_- + \alpha}{c\rho_- - \alpha}
\end{aligned}
$$

□

Proposition 4.23. *Let $(\alpha, \rho_-, \rho_+, h)$ be a solution of System* (4.25) *supplied with initial conditions*

$$\rho_-(0) = \rho_-^0, \rho_+(0) = \rho_+^0, \alpha(0) = \alpha^0, h(0) = h^0 \;\text{ and } h'(0) = v^0,$$

such that $|u_\pm^0 - v^0| < c$. Then the solution is global in time, α always keeps the same sign and tends to 0, ρ_+ and ρ_- tends to a common density ρ_∞, and the traces remain subsonic for all time: $|u_\pm - h'| < c$.

Proof. System (4.25) verifies the hypothesis of the Cauchy-Lipschitz theorem. Every constant quadruplet $(0, \rho_-^0, \rho_+^0, h^0)$ is a solution of this system. It follows that if $\alpha(0)$ is not zero, it remains non null. If $\alpha^0 > 0$, then the form of the germ impose that ρ_- is bigger than ρ_+. As ρ_- decreases and ρ_- increases, these densities stay in $[\rho_+^0, \rho_-^0]$. Moreover, h' increases and the decrease of the total energy prevents it from blowing up. Therefore h' remains bounded, hence α remains bounded as well. Eventually, all the quantities are monotonous and bounded, so the solution is global. To prove that the solution is subsonic, it suffices to notice than when $\alpha > 0$,

$$0 < u_- - h' < u_+ - h' = \frac{\alpha}{\rho_+} \le \frac{\alpha(0)}{\rho_+(0)} < c.$$

□

The system of ODE (4.25) drives the particle's velocity and the traces around the particle. This analysis is valid if two successive left traces are linked by a 1-wave (and similarly, if two successive right traces are linked by a 2-wave). This is possible if the waves emitted by the particle do not interact immediately with each other or with the particle. The following Theorem states that this is true for short time if we start from a subsonic element of the germ. To avoid the deal with vertical characteristic in the proof, we suppose that $u_L \neq c$ and $u_R \neq -c$. Up to a change of reference frame, this is not a restriction.

Theorem 4.24. *Suppose that initial data (ρ_L, u_L) and (ρ_R, u_R) and the initial particle's velocity v^{ini} of* (4.23) *are such that $|u_L - v^{ini}| < c$, $|u_R - v^{ini}| < c$ and*

$$((\rho_L, u_L), (\rho_R, u_R)) \in \mathcal{G}_D(v^{ini}).$$

Then, there exists a local in time solution to the Riemann problem (4.23). *On its interval of definition $(0, T)$, the left and right traces of the fluid around the particle and the particle's velocity verify* (4.25). *Moreover, the solution is constant along the lines*

$$(t - t^0)(u_+(t^0) + c) = x - h(t^0) \;\text{ for } t^0 \in (0, T) \;\text{ and }\; x > h(t^0)$$

and

$$(t - t^0)(u_-(t^0) - c) = x - h(t^0) \;\text{ for } t^0 \in (0, T) \;\text{ and }\; x < h(t^0).$$

Proof. Let $(\rho_-, \rho_+, \alpha, h)$ be the solution of System (4.25) with initial data

$$\rho_-(0) = \rho_L,\ \rho_+(0) = \rho_R,\ \alpha(0) = \alpha_L = \alpha_R,\ h(0) = h^{ini} \text{ and } h'(0) = v^{ini}.$$

By symmetry, it is sufficient to treat the case $\alpha(0) > 0$. Let us consider the sets of half lines

$$\mathcal{S}_+ = \{(x,t) : x = (t - t^0)(u_+(t^0) + c) + h(t^0),\ t \geq t^0, x \geq h(t^0)\}$$

and

$$\mathcal{S}_- = \{(x,t) : x = (t - t^0)(u_-(t^0) - c) + h(t^0),\ t \geq t^0, x \leq h(t^0)\}.$$

Some lines of those sets, for different starting point t^0, are drew on Figure 4.17.

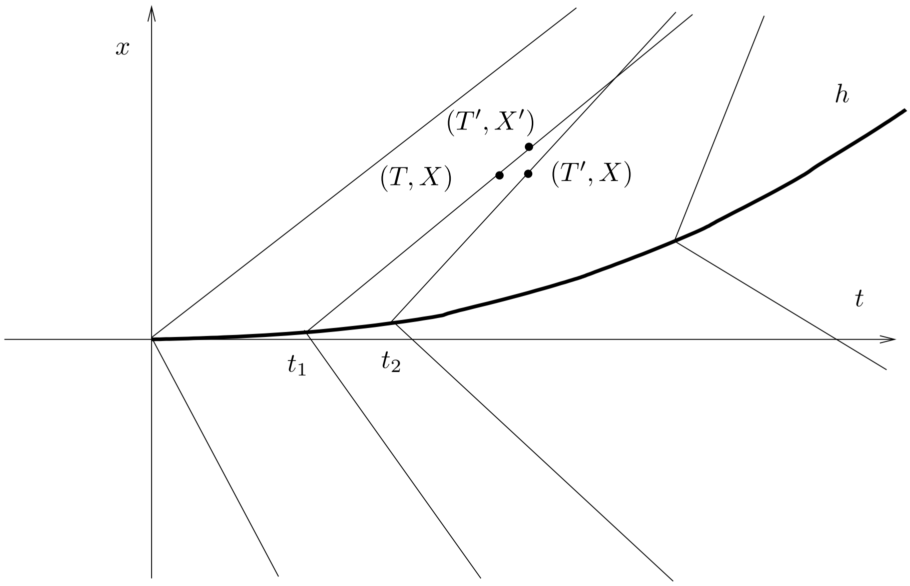


Figure 4.17. *Characteristic lines of the solution to the fully coupled Riemann problem.*

by Propositions 4.22 and 4.23, h is convex and $u_- - c$ is increasing and smaller than h'. Therefore, the half lines in $\mathcal{S}_-$ do not intersect. Every point (T, X) such that $X < h(T)$ and $X > (u_L - c)T + h^{ini}$ belongs to a unique half line of $\mathcal{S}_-$. We denote by $\Phi(T, X)$ the unique time t such that

$$X = (u_-(t) - c)(T - t) + h(t).$$

For the same reasons, two half-line in $\mathcal{S}_+$ always have an intersection point above h. Consider two times $t^A < t^B$ and denote by (T_{AB}, X_{AB}) the coordinates of the intersection between the two half lines starting from t^A and t^B. In the sequel we denote by p the slope of the lines: $p = u_+ + c$. We have that

$$T_{AB} = \frac{t^B p(t^B) - t^A p(t^A)}{p(t^B) - p(t^A)} - \frac{h(t^B) - h(t^A)}{p(t^B) - p(t^A)}.$$

Passing through the limit $t^B \longrightarrow t^A$ we obtain that T_{AB} tends to $\frac{p-h'}{p'}(t^A)$. This quantity is stricly positive around $t = 0$, indeed

$$\left(\frac{p - h'}{p'}\right)(0) = \frac{(u_R + c - v^0)}{u'_+(0)} > 0.$$

Thus there exists a time $T_f > 0$ such that all (T, X) in the set

$$\{(T, X), 0 < T < T_f, h(T) < X < (u_R + c)T + h^{ini}\}$$

belongs to a unique half line of $\mathcal{S}_+$. We also denote by $\Phi(T, X)$ the unique time t such that

$$X = (u_+(t) + c)(T - t) + h(t).$$

We extend Φ under the line $x = (u_L - c)t + h^{ini}$ and above the line $x = (u_R + c)t + h^{ini}$ by 0. Let us prove that on $(0, T_f)$ the function

$$(T, X) \mapsto \rho_+(\Phi(T, X)) \quad \text{and } (T, X) \mapsto u_+(\Phi(T, X))$$

are solution of the Euler equations on $\{x < h(t)\}$ and $\{x > h(t)\}$. Consider three points (X, T), (X', T) and (X, T') as depicted on Figure (4.17). By definition,

$$\Phi(T, X) = \Phi(T', X') = t_1 \quad \text{and} \quad \Phi(T, X') = t_2.$$

The points are placed in such a way that

$$\frac{\Phi(T', X') - \Phi(T', X)}{T' - T} + p(t_1)\frac{\Phi(T, X) - \Phi(T', X)}{X - X'} = \frac{t_1 - t_2}{T' - T}\left(1 - p(t_1)\frac{T' - T}{X' - X}\right) = 0$$

which yields when t_2 tends to t_1 to the following transport equation for Φ:

$$p(\Phi)\partial_T \Phi + \partial_X \Phi = 0.$$

We now prove the total mass conservation:

$$\begin{aligned}
\partial_T(\rho_+(\Phi)) + \partial_X(\rho_+ u_+(\Phi)) &= (\partial_T \Phi)\rho_+'(\Phi) + (\partial_X \Phi)(\rho_+ u_+)'(\Phi) \\
&= (\partial_X \Phi)\left[\rho_+' u_+ + u_+' \rho_+ - \rho_+'(u_+ + c)\right](\Phi) \\
&= (\partial_X \Phi)\left[u_+' \rho_+ - c\rho_+'\right](\Phi)
\end{aligned}$$

We use the second line of (4.27) to conclude:

$$\begin{aligned}
u_+' \rho_+ - c\rho_+' &= \rho_+\left(\frac{\alpha}{\rho_+} + h'\right)' - c\rho_+' \\
&= \rho_+\left(\frac{\alpha' \rho_+ - \rho_+' \alpha}{\rho_+^2} + h''\right) - c\rho_+' \\
&= \alpha' - \rho_+'\left(\frac{\alpha}{\rho_+} + c\right) + \rho_+ h'' \\
&= 0.
\end{aligned}$$

The conservation of impulsion follows from the same argument:

$$\begin{aligned}
&\partial_T(\rho_+ u_+(\Phi)) + \partial_X(\rho_+ u_+^2(\Phi) + c^2 \rho_+(\Phi)) \\
&\quad = \partial_X(\Phi)\left[\rho_+' u_+^2 + 2\rho_+ u_+ u_+' - (u_+ + c)(u_+' \rho_+ + \rho_+' u_+) + c^2 \rho_+'\right](\Phi) \\
&\quad = \partial_X(\Phi)\left[\rho_+' u_+^2 + 2\rho_+ u_+ u_+' - \rho_+ u_+ u_+' - c\rho_+ u_+' - \rho_+' u_+^2 - c\rho_+' u_+ + c^2 \rho_+'\right](\Phi) \\
&\quad = \partial_X(\Phi)\left[\rho_+ u_+ u_+' - c\rho_+ u_+' - c\rho_+' u_+ + c^2 \rho_+'\right](\Phi) \\
&\quad = \partial_X(\Phi)\left[u_+(\rho_+ u_+' - c\rho_+') + c(c\rho_+' - \rho_+ u_+')\right](\Phi) \\
&\quad = 0.
\end{aligned}$$

To conclude, it remains to notice that by construction of the solution, the traces belong to the germ for all time. $\square$

5 Schémas numériques pour le couplage entre les équations d'Euler et une particule

Dans ce chapitre, nous présentons quelques schémas numériques approchant le modèle de couplage fluide-particule défini au cours du chapitre précédent. Les deux difficultés principales sont, d'une part, de trouver des solveurs qui ne nécessitent pas la résolution complète du problème de Riemann avec particule et, d'autre part, de traiter correctement le déplacement de la particule lorsque celle-ci est mobile. Dans une première partie, on présente un schéma basé sur un splitting entre le terme source et la partie fluide. Il ne converge pas vers la bonne solution, ce qui illustre la nécessité de reformuler le système comme un problème d'interface. Lorsque la mesure de Dirac est régularisée, c'est-à-dire lorsqu'on approche le problème (4.8) dans lequel la particule est épaissie, on montre que la perte d'unicité du théorème 4.6 se pose aussi au niveau numérique. Dans une seconde partie, on présente deux schémas basés sur la résolution exacte du problème de Riemann (**??**), l'un de type Godunov et l'autre de type Glimm. Ce dernier schéma permet également de gérer le mouvement d'une particule sans changer le maillage à chaque itération en temps. Nous présentons différentes simulations numériques obtenues avec ces schémas, deux d'entre elles étant tirées de l'acte de conférence [Agu14b]. Dans la dernière partie, on présente quelques pistes pour se passer du solveur exact de Riemann : un schéma dit d'entraînement qui, comme dans la section 2.4.2, consiste à deviner les traces autour de la particule de manière à conserver une partie du germe, et un schéma de relaxation modifié pour prendre en compte la particule, inspiré de la partie 2.4.3.

Table des matières

5.1 Schéma de splitting

5.1.1 Échec du schéma de splitting naïf

On adopte les notations classiques pour les schémas de volumes finis. Comme auparavant on note $q = \rho u$ la quantité de mouvement, et $U_j^n = (\rho_j^n, q_j^n)$ une approximation des variables conservées à la n-ième itération en temps sur la j-ième maille. On note g le flux numérique ; on ne considérera que des flux à deux points. Un premier exemple montre l'échec de l'algorithme de splitting naïf entre les équations fluide et le terme source singulier d'interaction entre le fluide et la particule. Comme dans la partie 2.4.1, qui illustrait cet échec dans le cas du couplage Burgers-particule, on suppose que la particule est immobile : $v = 0$. On suppose que la particule se situe initialement au milieu de la maille0. Le schéma à trois points

$$\begin{cases} U_j^{n+1/2} & = U_j^n - \frac{t^{n+1}-t^n}{\Delta x}(g(U_j^n, U_{j+1}^n) - g(U_{j-1}^n, U_j^n)), \\ U_j^{n+1} & = U_j^{n+1/2} \text{ si } j \neq 0, \\ U_0^{n+1} & = U_0^{n+1/2} - \frac{t^{n+1}-t^n}{\Delta x} \begin{pmatrix} 0 \\ D(\rho_0^{n+1/2}, \rho_{j_0}^{n+1/2} u_0^{n+1/2}) \end{pmatrix}, \end{cases}$$

ne converge pas vers la solution, même dans le cas le plus simple où $D(\rho, \alpha) = \lambda\alpha$ et où la donnée initiale est un problème de Riemann dont les états gauche et droit sont initialement dans le germe. Cela peut être vu sur la figure 5.1. Cette incapacité à capturer correctement un phénomène de petit échelle

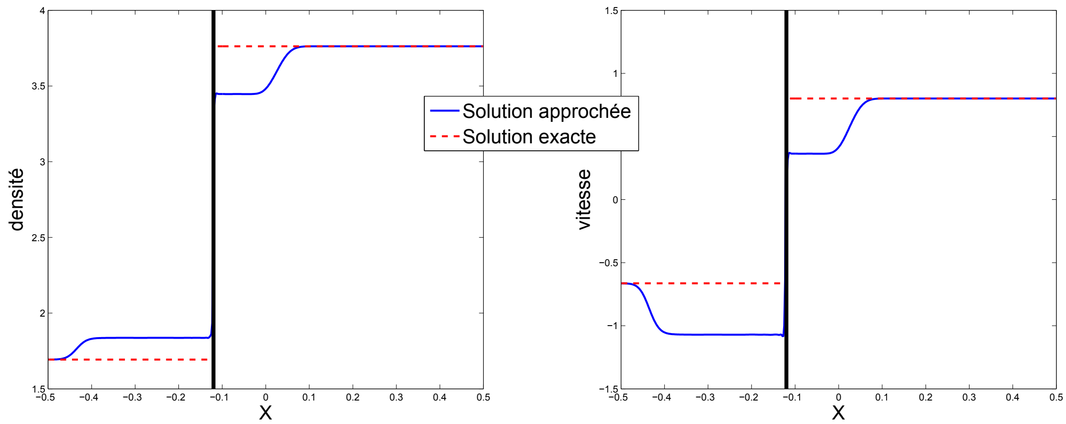


FIGURE 5.1. *Solution au $T = 0.04$ donnée par le splitting naïf entre fluide et particule.*

rappelle les difficultés rencontrées pour approcher les chocs non classiques (voir par exemple [HL00]) ou les systèmes non conservatifs (voir par exemple [DMLM95], [Par06], [Ber99] ou [CC06]). Cela souligne le caractère mal posé de l'équation (4.1), et la nécessité de reformuler ce système comme un problème d'interface.

5.1.2 Épaississement de la particule et perte d'unicité

On peut aussi explorer numériquement la perte d'unicité dans le théorème 4.17. Cette perte d'unicité est un phénomène connu dans le cadre des systèmes hyperboliques résonants, conservatifs ou non (voir [IT92] et [GL04]), et bien étudié sur des systèmes particuliers, comme les équations de Saint Venant à fond discontinues (voir les articles [CLS04] et [LT03]) et les équations de la dynamique des gaz dans une tuyère (voir [LT07]). Le premier exemple correspond, à un changement de loi de pression près, au frottement $D(\rho, \alpha) = \rho$. Aucune des deux hypothèses du théorème 4.6 pour ce frottement. Dans ce cas, le problème de Riemann (**??**) avec $\rho_L = 0.7$, $\rho_R = 5$, $q_L = 5$, $q_R = 9$, $c = 2$ et $\lambda = 1.5$

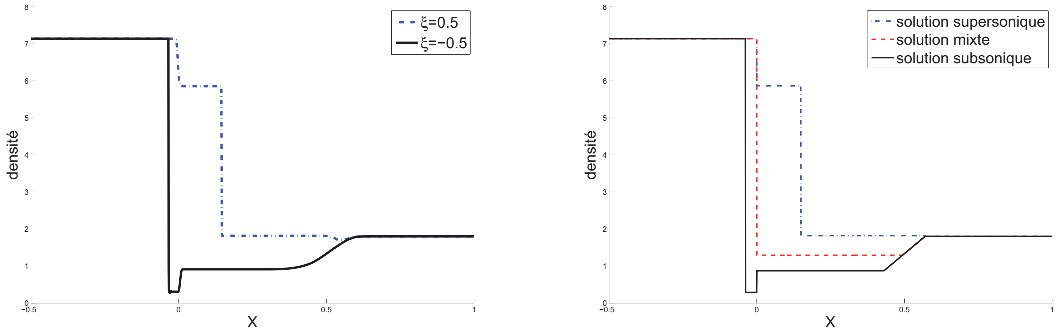


FIGURE 5.2. *À gauche : solutions au temps* $T = 0.15$ *fournies par le schéma de Godunov pour différentes régularisations de la mesure de Dirac. À droite : les trois solutions du problème de Riemann.*

admets trois solutions, tracées sur la droite de la figure 5.2. Comme dans [BCL13], la coexistence de ces solutions subsite au niveau numérique. On peut voir à gauche de la figure 5.2 deux solutions sélectionnées par le schéma de Godunov quand on remplace la mesure de Dirac par

$$x \mapsto \exp((x/\eta - \xi)^2)/(\eta\sqrt{2}),$$

avec $\eta = 0.005$ et $\xi = -0.5$ ou $\xi = 0.5$. On a utilisé un splitting entre la partie fluide et le terme source régularisé, et on a pris un maillage assez fin (5 000 mailles) pour lequel la particule épaissie contient une soixantaine de mailles. Les solutions supersonique et subsonique sont obtenues pour de larges plages du paramètre ξ, avec une transition très rapide entre les deux, passant par la solution mixte.

5.2 Schémas basés sur la résolution exacte du problème de Riemann

5.2.1 Cas de la particule immobile

Dans le cas où la particule est immobile, on dérive facilement des schémas basés sur la résolution exacte du problème de Riemann, explicitement construite dans la preuve du théorème 4.6. On place la particule sur l'interface $1/2$ du maillage et on utilise une approche de mailles fantôme (voir [FAMO99] et [AK01]). Le schéma s'écrit

$$\left\{ \begin{array}{lll} U_j^{n+1} & = U_j^n - \frac{\Delta t}{\Delta x}(g(U_j^n, U_{j+1}^n) - g(U_{j-1}^n, U_j^n)) & \text{pour } j \notin \{0, 1\}, \\ U_0^{n+1} & = U_0^n - \frac{\Delta t}{\Delta x}(g(U_0^n, U_{\text{part},-}^n) - g(U_{-1}^n, U_0^n)) & \text{pour } j = 0, \\ U_1^{n+1} & = U_1^n - \frac{\Delta t}{\Delta x}(g(U_1^n, U_2^n) - g(U_{\text{part},+}^n, U_1^n)) & \text{pour } j = 1. \end{array} \right. \tag{5.1}$$

Dans cette section on prend $U_{\text{part},-}^n = U_{\text{exa},-}^n$ et $U_{\text{part},+}^n = U_{\text{exa},+}^n$, où $U_{\text{exa},-}^n$ et $U_{\text{exa},+}^n$ regroupent les valeurs prises par la densité et la quantité de mouvement du fluide sur les lignes $x = 0^-$ et $x = 0^+$ de la solution de (**??**), avec

$$\rho_L = \rho_0^n, \ u_L = \frac{q_0^n}{\rho_0^n}, \ \rho_R = \rho_1^n \ \text{ et } \ u_R = \frac{q_1^n}{\rho_1^n}.$$

Remarque 16. Lorsque g est le flux de Godunov, U_0^{n+1} et U_1^{n+1} sont obtenues comme les moyennes sur les mailles 0 et 1 de la solution exacte *avec particule* de **??**. En d'autres termes, c'est le schéma original de Godunov adapté au cas du couplage fluide-particule.

Remarque 17. Si la donnée initiale est un problème de Riemann avec $(\rho_L, q_L), (\rho_R, q_R)$ qui appartient au germe $\mathcal{G}_D(0)$, alors

$$U^n_{\text{part},-} = U_L \quad \text{et} \quad U^n_{\text{part},+} = U_R,$$

et par conséquent

$$\forall n \geq 0,\ \forall j \leq 0,\ U^n_j = (\rho_L, q_L) \ \text{ et } \ \forall j \geq 1,\ U^n_j = (\rho_R, q_R).$$

En d'autres termes, le schéma conservent tous les équilibres associés au germe.

5.2.2 Cas d'une particule mobile

On autorise désormais la particule à bouger sous l'influence du fluide. Il est nécessaire de replacer la particule sur une interface du maillage à la fin de chaque itération en temps. Pour ce faire, il est possible d'utiliser un maillage suivant la particule, comme dans le chapitre 3. Toutefois pour des applications plus complexes, avec plusieurs particules par exemple, on décide d'utiliser un maillage fixe et de replacer aléatoirement la particule avec une méthode à la Glimm. À chaque itération, un réel x_r est tiré uniformément dans $[0, \Delta x]$. La valeur du fluide dans la maille j à l'itération $n+1$ est actualisé comme la valeur au point x_r et au temps $t^{n+1} - t^n$ de la solution de (4.1) avec

$$U^0(x) = U^n_{j-1}\mathbf{1}_{x<0} + U^n_j \mathbf{1}_{0<x<\Delta x} + U^n_{j+1}\mathbf{1}_{\Delta x<x}.$$

Notons (ρ, u) la solution exacte de ce problème, où la particule est prise en compte si $j = 0$ ou $j = 1$. Sous la condition CFL classique $\Delta t < \frac{\Delta x}{2(c+\max_x |u|)}$, la solution consiste en la juxtaposition des solutions de ces deux problèmes de Riemann. La position de la particule est actualisée en accord avec le choix de x_r. Si la particule est située sur l'interface $j^n_0 + 1/2$ au temps n, et a vitesse v^n, on la place :

1. sur l'interface $j^n_0 + 3/2$ si $v^n > 0$ et $x_r < v^n \Delta t$, dans ce cas $j^{n+1}_0 = j^n_0 + 1$;
2. sur l'interface $j^n_0 - 1/2$ si $v^n < 0$ et $x_r > \Delta x + v^n \Delta t$, dans ce cas $j^{n+1}_0 = j^n_0 - 1$;
3. sur l'interface $j^n_0 + 1/2$ dans les autres cas, on a alors $j^{n+1}_0 = j^n_0$.

Finalement, la vitesse de la particule est actualisée avec (4.14), en utilisant les traces numériques.

5.3 Simulations numériques

5.3.1 Tuyau d'orgue bouché

On peut utiliser le schéma (5.1) pour simuler un tuyau d'orgue bouché. Le tuyau est initialement rempli d'un fluide au repos, de densité $5\,kg/m$, et on prend $c = 1\,m/s$. Pour $t \geq 0$, on impose un débit constant de $3\,kg/s$ à l'extrémité gauche de l'orgue. L'air peut s'échapper librement à la sortie droite du tuyau, qui est bloqué en son milieu par un tampon poreux qu'on modélise en utilisant le terme force $D(\rho, \alpha) = \alpha$. Au temps $0.041\,s$, le choc crée par la condition au bord de gauche rencontre la particule. La solution du problème de Riemann correspondant contient un choc de chaque côté de la particule. Grosso modo, la plus grosse partie de l'air est bloquée par la particule, ce qui cause une augmentation de sa densité et une diminution de sa vitesse. Comme le tampon est poreux, une petite partie de l'air parvient à traverser l'obstacle, et aura une vitesse élevée à la sortie de la particule puisque la quantité de mouvement est constante de part et d'autre de l'obstacle. Un choc se crée à gauche de la particule, et interagit avec l'extrémité de gauche au temps $0.114\,s$, ce qui crée un nouveau choc qui rencontre la particule au temps $0.153\,s$. En temps long, la quantité de mouvement du fluide est constante à travers tout le tube, avec une densité élevée et une faible vitesse à gauche de la particule, et une densité faible et une vitesse élevée à sa droite. L'allure de la solution après les deux premières interactions entre la particule et le fluide est donnée sur la figure 5.3. Cette simulation illustre la convergence du schéma (5.1). On a utilisé un flux de Godunov, mais les résultats sont similaires avec un flux de Rusanov.

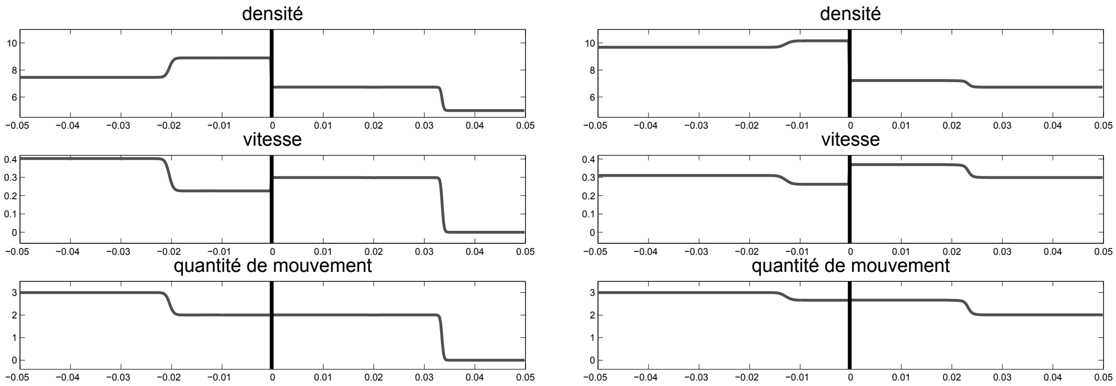


FIGURE 5.3. *Deux interactions successives entre un choc et une particule dans un tuyau d'orgue bouché.*

5.3.2 Fluide ralenti par un grillage de plus en plus épais

Dans cette simulation, la vitesse du son est égale à $1\,m/s$ et le fluide a initialement une densité prise constante égale à $1\,kg/m$ et une vitesse constante égale à $2\,m/s$. En $x = 0$ se situe un grillage modélisé par le terme de frottement

$$D(\rho, q) = \lambda(t)q$$

avec

$$\lambda(t) = \begin{cases} 0 & \text{si } t < 0.01, \\ (t - 0.01)/1.2 & \text{si } t \geq 0.01. \end{cases}$$

Ainsi, le paramètre de friction λ est de plus en plus élevé, ce qui traduit le fait que le grillage est de plus en plus épais. On utilise le schéma (5.1) avec g le flux de Godunov. L'intervalle d'espace $[-0.2, 0.2]$ est discrétisé avec $10\,000$ mailles. On impose un débit de $2\,kg.m^{-1}.s^{-1}$ à gauche et une sortie libre à droite. Tant que le coefficient reste petit, l'écoulement est supersonique et le fluide est ralenti à l'arrière de la particule. Au temps $t = 0.39$ environ (ce qui correspond à $\lambda = 0.63$ environ), le vitesse à la sortie de la particule est égale à celle du son. À partir de cet instant, l'écoulement devient subsonique et le fluide est freiné à l'avant de la particule, où il se forme un bouchon comme dans le cas du tuyau d'orgue bouché.

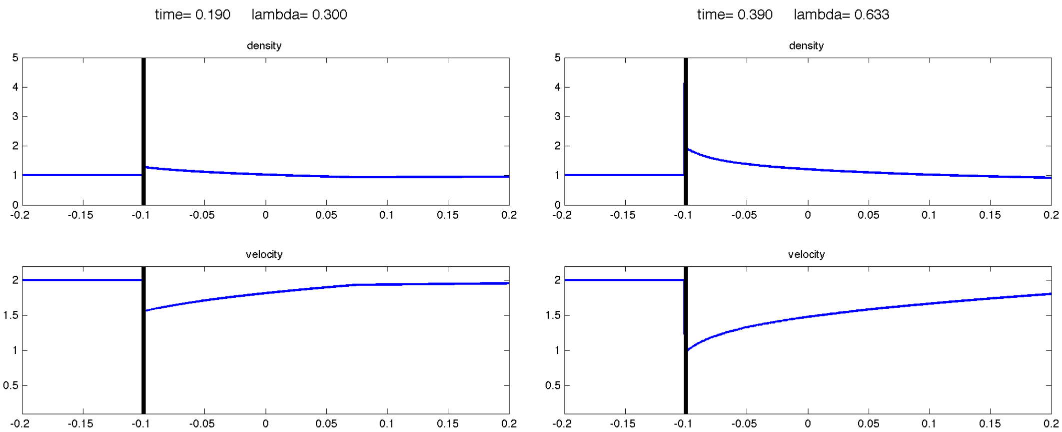


FIGURE 5.4. *Allure supersonique de la solution tant que le coefficient de frottement est petit.*

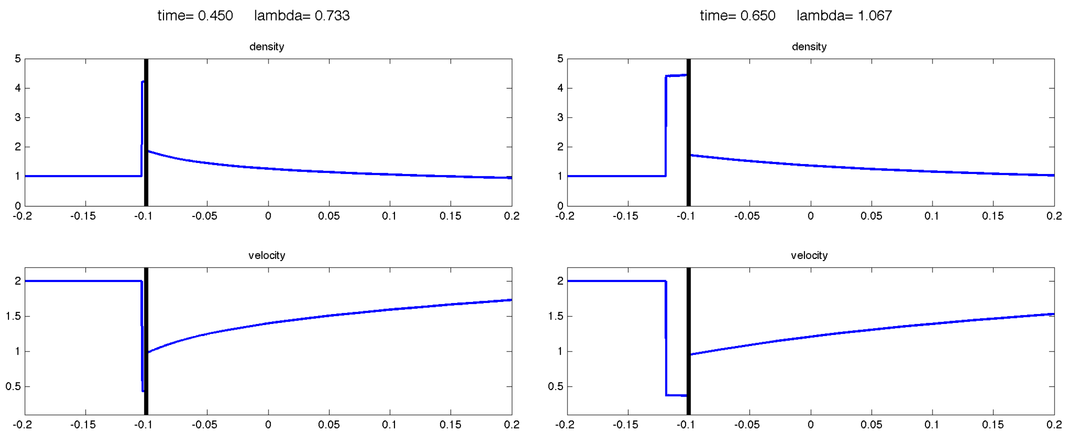


FIGURE 5.5. *Allure subsonique de la solution quand le coefficient de friction est plus grand.*

5.3.3 Bille tombant dans un tube

La simulation numérique suivante est inspirée par celle présentée dans [BCG14]. Elle est réalisée avec le schéma de Glimm présenté dans la partie 5.2.2. Une bille tombe dans un cylindre vertical rempli d'un fluide compressible non visqueux, initialement au repos et de densité constante $1/225\,kg/m$. Le fluide comme la particule sont soumis à des forces de pesanteur et de friction. Le système complet s'écrit :

$$\begin{cases} \partial_t\rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + c^2\rho) = -\lambda(u-h')\delta_{h(t)}(x) - \rho g - \nu_F(\rho,u), \\ mh''(t) = \lambda(u(t,h(t)) - h'(t)) - mg - m\nu_S(h'(t)). \end{cases}$$

Comme dans [BCG14], on prend $\nu_S(h') = 10^{-2}h'$, $\nu_F(\rho,u) = 10^{-8}\rho u|u|$, $c = 15\,m/s$, $m = 0.004\,kg$ et $g = 9.81\,m/s^2$. On choisit $\lambda = 5\,m^2 \cdot kg/s$. Le premier terme de l'EDO sue la particule doit être

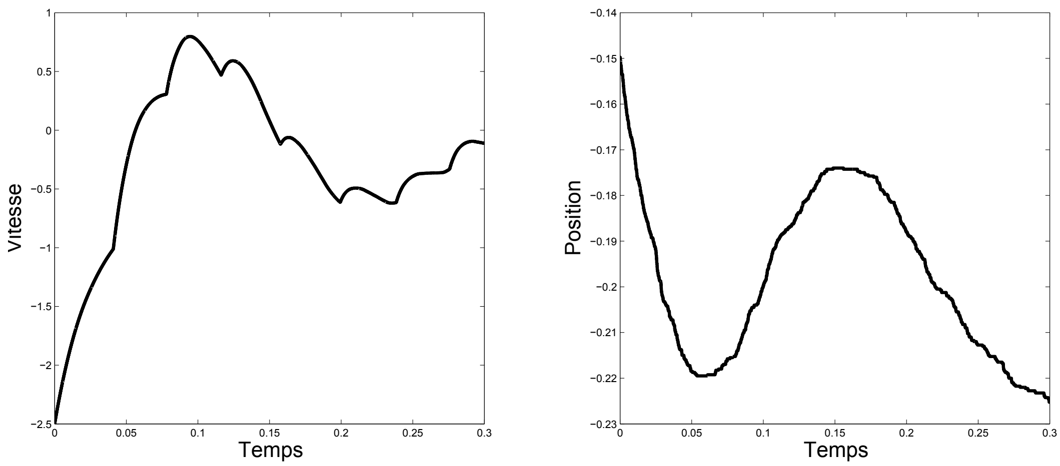


FIGURE 5.6. *À gauche : vitesse de la particule au cours du temps. Chaque discontinuité dans son accélération est causée par un choc la rencontrant. À droite : position de la particule au cours du temps.*

comprise comme dans l'équation (4.14). Au début, la bille chute vers le cas, ce qui crée un choc dans le fluide qui se situe en dessous. Le choc se réfléchit sur la paroi du bas du tube, pour frapper la particule. L'accélération de la particule est discontinue à ce temps et la chute de la particule est ralentie, comme on peut le voir sur la partie gauche de la figure 5.6. Quand le choc interagit avec la particule, il est scindé en deux comme dans l'exemple précédent. La majorité du fluide reste bloquée en dessous de la particule, mais une petite partie parvient à « passer sur les côtés » et s'échappera par le haut du tube. Ainsi, un choc d'amplitude de plus en plus faible effectue des aller-retours entre le fond du tube et la particule (voir la figure 5.7). D'autre part, la particule étant très légère, elle est très sensible à la vitesse du fluide. Hors celle ci est positive (dirigée vers le haut) après les premières fois que le choc a rebondi au fond du tube. Par conséquent, la bille remonte un instant, puis la gravité redevient prédominante. Les résultats obtenus sont qualitativement les mêmes que ceux présentés dans [BCG14]. Toutefois, la correspondance n'est pas parfaite, puisque la modélisation est différente. En particulier dans [BCG14], la friction entre la bille et le fluide est traduite par le terme source

$$\nu_I = 5\left(h' - \frac{u_- + u_+}{2}\right)^2,$$

alors qu'elle est décrite par les conditions d'interface de la définition 4.1 dans cette thèse.

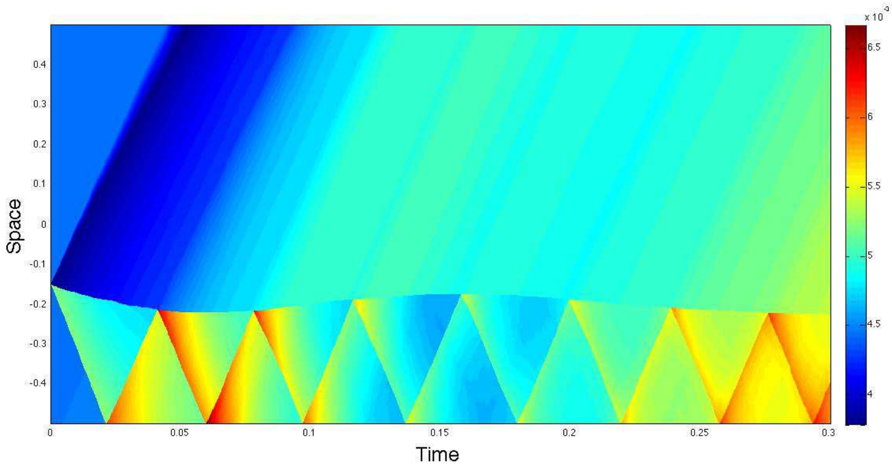


FIGURE 5.7. *Densité du fluide au cours du temps. Le choc effectuant des aller-retours entre le fond du tube et la particule est de force de plus en plus faible.*

5.4 Se passer du solveur de Riemann exact

Les schémas présentés ci-dessus reposent sur la résolution exacte du problème de Riemann (**??**). Sans surprise, cette résolution est coûteuse. Dans cette section, nous présentons deux tentatives pour obtenir un schéma plus rapide. Les schémas correspondant convergent dans le cas du couplage Burgers-particule (voir la partie 2.4), mais ils échouent tous les deux dans le cas du couplage Euler-particule. Ils sont présentés dans le cas où la particule est immobile, et nous nous restreignons au cas du frottement le plus simple, $D(\rho, \alpha) = \lambda\alpha$ avec $\lambda \geq 0$.

5.4.1 Schéma d'entrainement

Dans un premier temps, essayons de déterminer $U^n_{part,-}$ et $U^n_{part,+}$ dans (5.1) sans résoudre le problème de Riemann avec particule entre U^n_0 et U^n_1. Comme dans [AS12] et dans la partie 2.4.2, on suppose que

U_0^n est l'état à gauche de la particule, et on cherche $U^n_{part,+}$ de sorte que

$$(U_0^n, U^n_{part,+}) \in \mathcal{G}_D(0).$$

Ainsi d'après la remarque 6, on doit avoir $q^n_{part,+} = q_0^n$ et

$$\left(\frac{(q_0^n)^2}{\rho_0^n} + c^2\rho_0^n\right) - \left(\frac{(q_0^n)^2}{\rho^n_{part,+}} + c^2\rho^n_{part,+}\right) = \lambda q_0^n \tag{5.2}$$

L'équation (5.2) a toujours deux solutions si q_0^n est négatif. Si q_0^n est strictement positif, elle a deux solutions si $\frac{(q_0^n)^2}{\rho_0^n} + c^2\rho_0^n > 2q_0^n c + \lambda q_0^n$, une seule en cas d'égalité et aucune si $\frac{(q_0^n)^2}{\rho_0^n} + c^2\rho_0^n < 2q_0^n c + \lambda q_0^n$ (voir les remarques 6 et 7). On sélectionne la solution qui correspond à une solution régulière dans le théorème 4.3, autrement dit celle pour laquelle la vitesse est subsonique partout, ou supersonique partout. Dans le cas où il n'y a pas de solution, on fait décroitre la charge autant que possible. Ce choix est dépeint sur la figure 5.8 ci dessous. Tous calculs faits, en posant

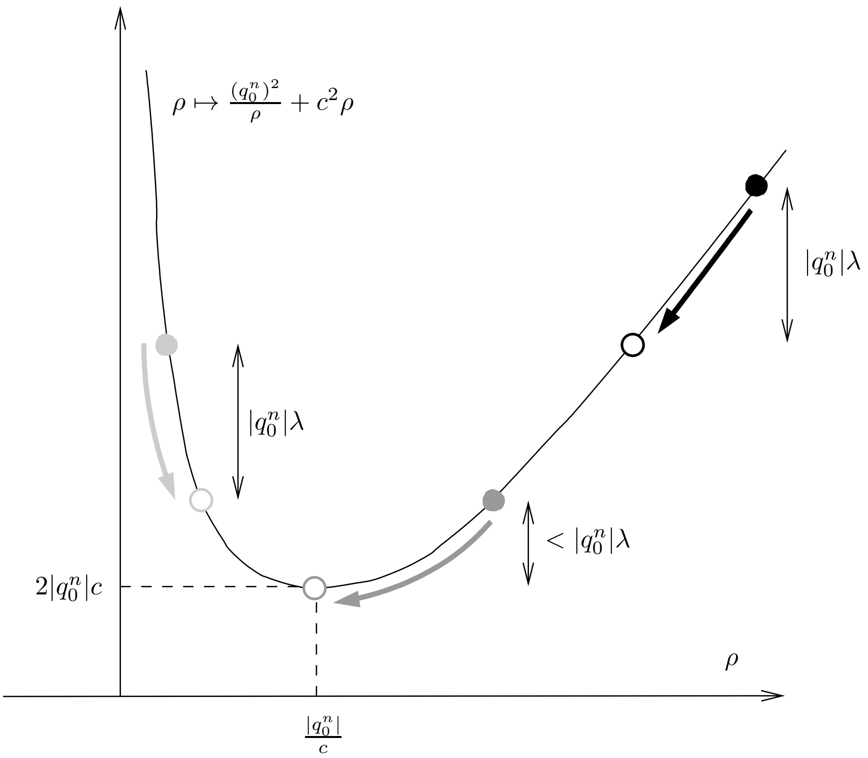


FIGURE 5.8. *Choix de $\rho^n_{part,+}$ (cercles creux) selon ρ_0^n (cercles pleins).*

$$\beta(U_0^n) = \lambda q_0^n - \left(\frac{(q_0^n)^2}{\rho_0^n} + c^2\rho_0^n\right) \quad \text{et} \quad \Delta(U_0^n) = \beta^2(U_0^n) - 4(q_0^n)^2c^2,$$

cela mène au choix

$$\rho^n_{part,+} = \begin{cases} \dfrac{-\beta(U^n_0) + \sqrt{\Delta(U^n_0)}}{2c^2} & \text{si } \Delta(U^n_0) > 0 \text{ et } c\rho^n_0 > |q^n_0|, \\ \dfrac{-\beta(U^n_0) - \sqrt{\Delta(U^n_0)}}{2c^2} & \text{si } \Delta(U^n_0) > 0 \text{ et } c\rho^n_0 \leq |q^n_0|, \\ \dfrac{|q^n_0|}{c} & \text{si } \Delta(U^n_0) \leq 0. \end{cases} \tag{5.3}$$

On peut raisonner de même pour déterminer $U^n_{part,-}$ à partir de U^n_1. On obtient $q^n_{part,-} = q^n_1$ et

$$\rho^n_{part,-} = \begin{cases} \dfrac{-\beta(U^n_1) + \sqrt{\Delta(U^n_1)}}{2c^2} & \text{si } \Delta(U^n_1) > 0 \text{ et } c\rho^n_1 > |q^n_1|, \\ \dfrac{-\beta(U^n_1) - \sqrt{\Delta(U^n_1)}}{2c^2} & \text{si } \Delta(U^n_1) > 0 \text{ et } c\rho^n_1 \leq |q^n_1|, \\ \dfrac{|q^n_1|}{c} & \text{si } \Delta(U^n_1) \leq 0, \end{cases} \tag{5.4}$$

où

$$\beta(U^n_1) = -\left(\lambda q^n_1 + \frac{(q^n_1)^2}{\rho^n_1} + c^2\rho^n_1\right) \quad \text{et} \quad \Delta(U^n_1) = \beta(U^n_1)^2 - 4(q^n_1)^2c^2.$$

Ce schéma ne converge pas vers la bonne solution. Ceci est illustré sur la figure 5.9. La donnée initiale est le problème de Riemann

$$\begin{cases} \rho^0(x) = \mathbf{1}_{x<0} + 2 * \mathbf{1}_{x\geq 0}, \\ q^0(x) = \mathbf{1}_{x<0} + 2 * \mathbf{1}_{x\geq 0}, \end{cases}$$

et on prend $c = 1$. D'après la proposition 4.13, les traces de part et d'autre de la solution sont subsoniques. Le schéma utilisé est (5.1) avec $U^n_{part,+}$ défini par (5.3) et $U^n_{part,-}$ défini par (5.4). Le flux numérique g est celui de Godunov. Le pas de temps est fixé de sorte que

$$\forall n,\ t^{n+1} - t^n \leq \frac{0.45\Delta x}{max_j(|U^n_j|, |U^n_{part,-}|, |U^n_{part,+}|) + c}.$$

Le schéma converge pour $\lambda = 1$ mais pas pour $\lambda = 2$. Le temps final est 0.2. Lorsque la donnée initiale

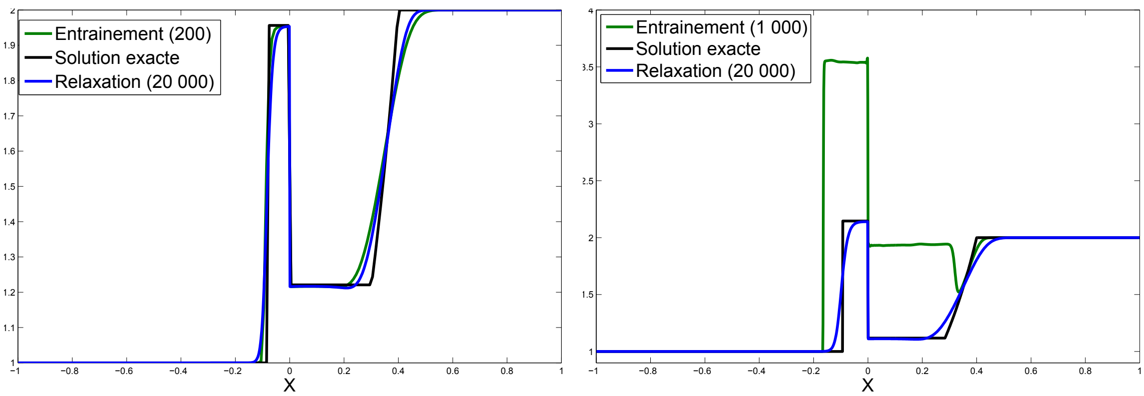


FIGURE 5.9. *Le schéma d'entrainement converge pour $\lambda = 1$ mais pas pour $\lambda = 2$. Le schéma de relaxation converge dans les deux cas. Le nombre de maille est indiqué entre parenthèse dans la légende.*

est

$$\begin{cases} \rho^0(x) = 2 * \mathbf{1}_{x<0} + 0.5 * \mathbf{1}_{x\geq 0}, \\ q^0(x) = 2 * \mathbf{1}_{x<0}, \end{cases}$$

et quand λ est assez petit, la solution est résonante, c'est à dire qu'une détente démarre juste à droite de la particule. Dans ce cas, aussi petit que soit λ, le schéma ne parvient pas à capturer la bonne trace à gauche de la particule, et le choc qui se détache de ce côté de la particule n'apparaît pas dans la solution numérique. Ceci est illustré pour $\lambda = 0.2$ par la figure 5.10. L'intervalle d'espace contient $5\,000$ mailles, le temps final est 0.4.

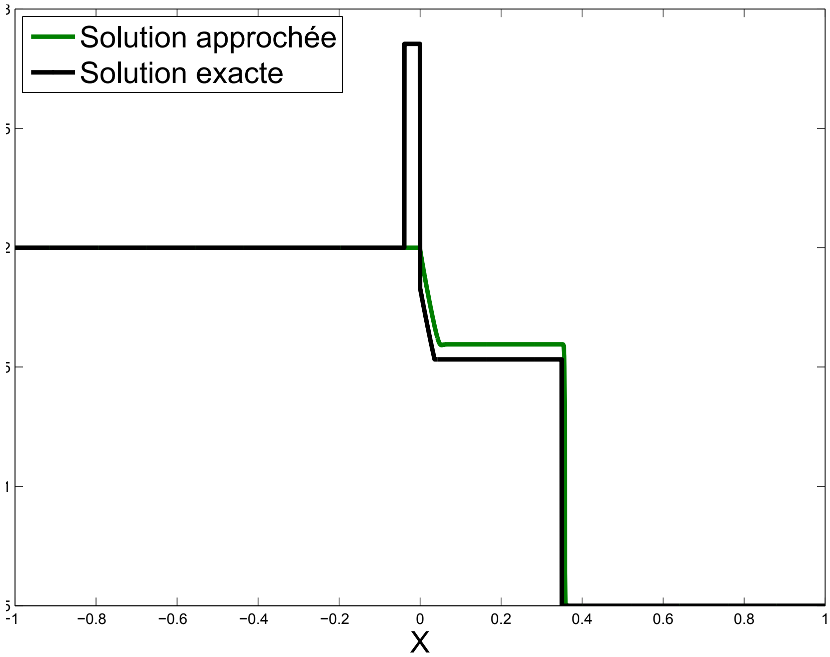


FIGURE 5.10. *Le schéma d'entrainement ne converge pas dans le cas résonant, même si λ est petit.*

5.4.2 Schéma de relaxation

On s'intéresse maintenant au schéma de relaxation basé sur le problème de Riemann pour le système augmenté suivant :

$$\begin{cases} \partial_t \rho + \partial_x m = 0; \\ \partial_t m + A^2 \partial_x \rho = \mathcal{M}\delta_0 - \frac{1}{\varepsilon}(m - q) \\ \partial_t q + \partial_x \eta = -\lambda q \delta_0 \\ \partial_t \eta + A^2 \partial_x q = -\frac{1}{\varepsilon}\left(\left(\eta - \frac{q^2}{\rho} + c^2 \rho\right)\right) \\ \rho^0(x) = \rho_L \mathbf{1}_{x<0} + \rho_R \mathbf{x \geq 0} \\ q^0(x) = q_L \mathbf{1}_{x<0} + q_R \mathbf{x \geq 0} \\ m^0(x) = m_L \mathbf{1}_{x<0} + m_R \mathbf{x \geq 0} \\ \eta^0(x) = \rho_L \mathbf{1}_{x<0} + \eta_R \mathbf{x \geq 0} \end{cases} \tag{5.5}$$

Comme dans la section 2.4.3, le terme source $\mathcal{M}\delta_0$ autorise la variable ρ à être discontinue à travers la particule. La solution contient trois discontinuités, dont les vitesses sont $-A$, 0 et A. On note avec un indice $-$ les états à gauche de la particule et avec un indice $+$ ceux à sa droite. La première, troisième et quatrième lignes de (5.5) donnent

$$\begin{cases} m_- = m_+ := m, \\ q_- = q_+ := q, \\ \eta_- - \eta_+ = \lambda q. \end{cases}$$

La seconde ligne de (5.5) fournit, selon le choix du poids $\mathcal{M}$, une quatrième relation, éventuellement non linéaire, liant les états de part et d'autre de la particule. Les relations de Rankine–Hugoniot à travers les ondes à vitesses $-A$ et A donnent

$$\begin{cases} -A(\rho_L - \rho_-) & = m_L - m, \\ A(\rho_R - \rho_+) & = m_R - m, \\ -A(q_L - q) & = \eta_L - \eta_-, \\ A(q_R - q) & = \eta_R - \eta_+. \end{cases}$$

Trois équations ne portent que sur q, η_- et η_+, ce qui permet de trouver

$$q = \frac{A(q_L + q_R) + (\eta_L - \eta_R)}{2A + \lambda}.$$

On peut alors exprimer toutes les inconnues en fonction de q et m :

$$\begin{cases} \eta_- & = \eta_L + A(q_L - q), \\ \eta_+ & = \eta_L - A(q_R - q), \\ \rho_-(m) & = \rho_L + \frac{m_L - m}{A}, \\ \rho_+(m) & = \rho_R - \frac{m_R - m}{A}. \end{cases}$$

La valeur de m est déterminée par une dernière relation qui correspond à la seconde ligne de (5.5) qu'on avait laissé de côté. Il semble souhaitable que :

- Si $\lambda = 0$, on retrouve le schéma de relaxation pour Euler sans particule, soit

$$m := m_0 = \frac{m_L + m_R}{2} + \frac{A}{2}(\rho_L - \rho_R). \tag{5.6}$$

- Si $((\rho_L, q_L), (\rho_R, q_R))$ appartient au germe $\mathcal{G}_D(0)$, $m = q_L = q_R$ et donc $\rho_- = \rho_L$ et $\rho_+ = \rho_R$.

On a choisi d'imposer

$$\left(\frac{m^2}{\rho_-(m)} + c^2\rho_-(m)\right) - \left(\frac{m^2}{\rho_+(m)} + c^2\rho_+(m)\right) = \lambda m. \tag{5.7}$$

Cette équation a au moins une solution. En effet la fonction

$$\Phi(m) = \left(\frac{m^2}{\rho_-(m)} + c^2\rho_-(m)\right) - \left(\frac{m^2}{\rho_+(m)} + c^2\rho_+(m)\right) - \lambda m$$

tend vers $+\infty$ lorsque $\rho_-(m)$ tends vers 0 et vers $+\infty$ lorsque $\rho_+(m)$ tends vers 0. Le problème est qu'elle peut avoir plusieurs solutions, quel que soit le choix de A assez grand. En effet, $\Phi(0) = c^2(\rho_-(0) - \rho_+(0))$ et $\Phi(m_0) = -\lambda m_0$ peuvent être de signes opposés. On résout (5.7) par une méthode itérative en prenant m_0 défini par (5.6) comme point de départ, pour assurer le premier point. Comme on peut le voir sur la figure 5.9, ce schéma converge sur le premier cas test. Il ne converge pas non plus dans le cas résonant. Cela est illustré sur la figure 5.11. La donnée initiale est

$$\begin{cases} \rho^0(x) = 5 * \mathbf{1}_{x<0} + 0.5 * \mathbf{1}_{x\geq 0}, \\ q^0(x) = 3 * \mathbf{1}_{x<0} \end{cases}$$

avec $\lambda = 1$. L'intervalle d'espace $[-1, 1]$ est discrétisé avec 20 000 mailles, et le temps final est $T = 0.2$.

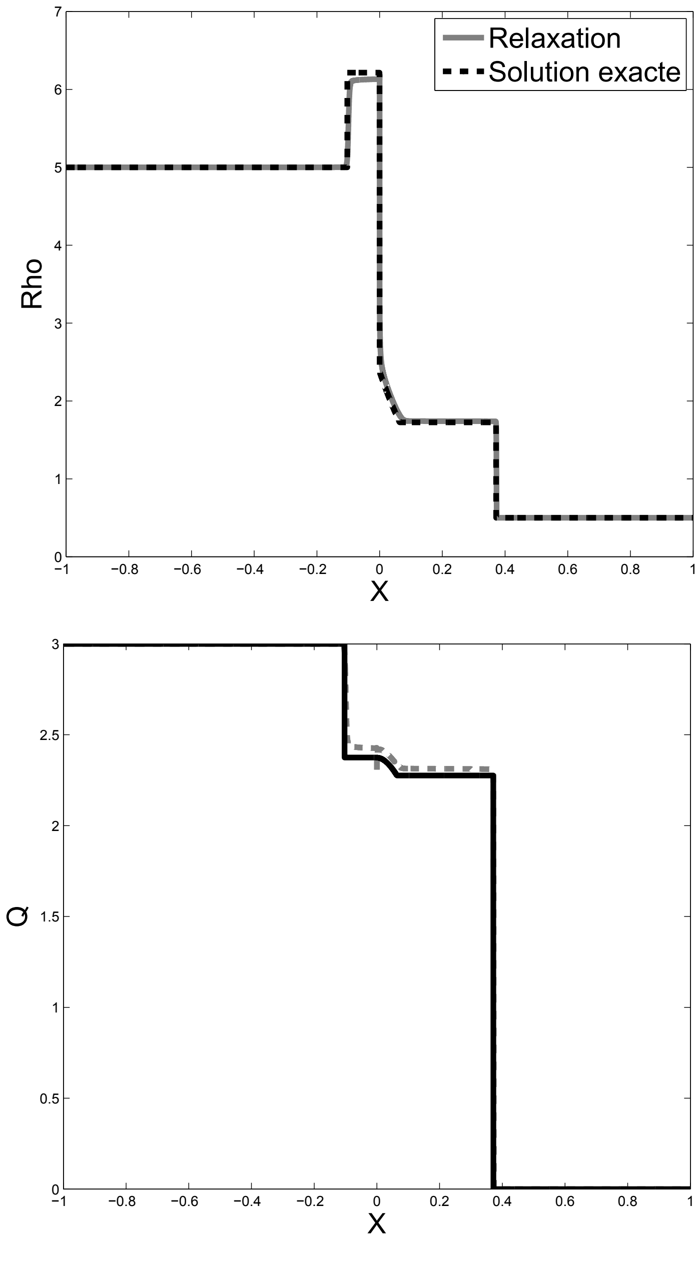


FIGURE 5.11. *Le schéma de relaxation ne converge pas dans le cas résonant.*

III

Schémas de reconstruction discontinue

6 Discontinuous reconstruction schemes

Ce chapitre est consacré à la dérivation de schémas de volumes finis capables d'approcher exactement les chocs. Cela permet de surmonter plusieurs difficultés liées à la diffusion numérique introduite par la phase de projection sur le maillage, qui peut avoir des conséquences majeures dans certains cas. Contrairement à la plupart des schémas de volumes finis disponibles, les schémas présentés dans ce chapitre sont capables :

- d'approcher correctement les chocs qui se déplacent lentement par rapport à la vitesse du son, sans produire d'oscillations parasites dans la quantité de mouvement (dans la dynamique des gaz avec ou sans énergie) ;
- d'approcher correctement les réflexions de choc sur une paroi, sans créer d'élévation non physique de la température près du mur (dans la dynamique des gaz avec énergie) ;
- d'approcher correctement les chocs non classiques, qui sont des chocs qui ne vérifient pas les critères usuels de sélection de chocs et qui apparaissent comme limite de solutions de systèmes diffusif-dispersif.

Ces schémas sont dits *de reconstruction discontinue* car ils reposent sur la reconstruction de chocs à l'intérieur de certaines mailles à chaque itérations en temps. Dans la première partie, nous présentons un critère sur l'opérateur de reconstruction qui assure que le schéma est exact lorsque la donnée initiale est une choc. Dans les trois parties suivantes, qui correspondent peu ou proue à la prépublication [Agu14a], cette stratégie est appliquée à l'équation de Burgers, aux équations d'Euler isothermes et à la dynamique des gaz avec énergie. Ces schémas reposent sur la résolution exacte du problème de Riemann. Dans la partie 6.6, qui correspond au travail [AC14] en collaboration avec Christophe Chalons, nous couplons la stratégie de reconstruction discontinue avec l'utilisation d'un solveur de Riemann approché. Enfin, dans la dernière partie, nous proposons un schéma de reconstruction discontinue qui permet d'approcher les chocs non classiques d'un modèle d'élastodynamique.

Contents

6.1 Introduction

When approaching the solutions of hyperbolic systems, finite volume schemes usually introduce a small amount of numerical dissipation. The main effect of this numerical viscosity is to smear out the discontinuities over several cells. Let us recall from where comes this numerical diffusion. Consider the Cauchy problem for the hyperbolic system

$$\begin{cases} \partial_t U + \partial_x f(U) = 0. \\ U(t=0,x) = U^0(x), \end{cases} \tag{6.1}$$

At the n-th iteration in time, the finite volume scheme give a piecewise constant approximation of the solution

$$U^n(x) = \sum_{j\in\mathbb{Z}} U_j^n \mathbf{1}_{[x_{j-1/2}^n, x_{j+1/2}^n]}(x),$$

where the vector U regroups the n unknowns ($n \in \mathbb{N}^*$) and is defined on a open convex set Ω of $\mathbb{R}^n$. Let us denote by $(t,x) \mapsto U_{exa}^n(t,x)$ the exact solution of the Cauchy problem (6.1) with initial data U^n. The celebrated Godunov scheme consists in updated U as the average of U_{exa}^n at time t^n on each cell:

$$U_j^{n+1} = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} U_{exa}^n(t^{n+1}-t^n, x)dx.$$

This procedure is described on the left of Figure 6.1. If the time step verifies the Courant–Friedrichs–Lewy condition

$$t^{n+1} - t^n \leq \frac{\Delta x}{2V_{\text{waves}}^n}. \tag{6.2}$$

$U_{exa}^n(t^{n+1}-t^n,\cdot)$ is the juxtaposition of the solutions of each Riemann problems between two adjacents cells. Here, W_{waves}^n denotes the maximum of the waves speeds appearing in those Riemann problems.

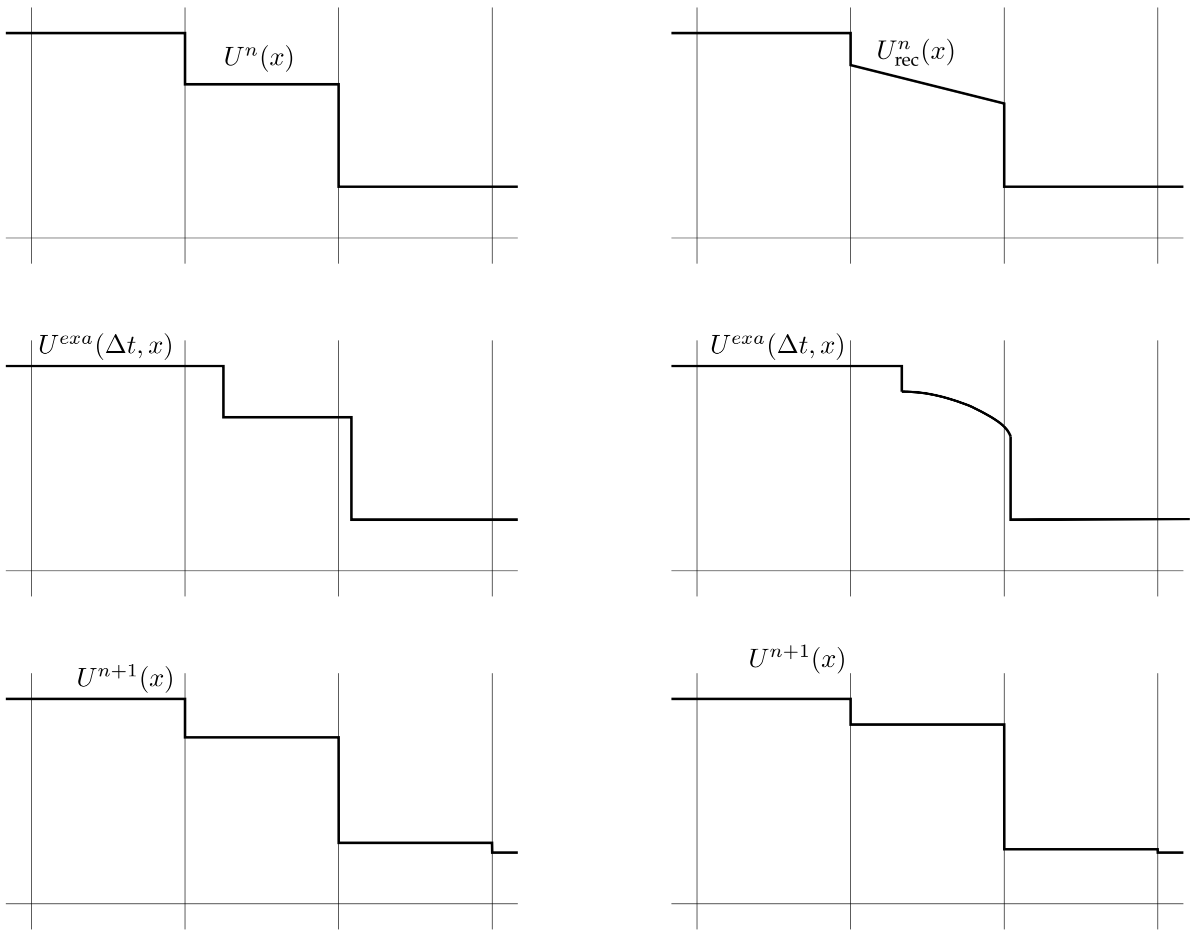


Figure 6.1. *Left: principle of the Godunov scheme. The solution at time t^{n+1} is the average on the mesh of the exact solution at time Δt of the Cauchy problem* (6.1) *with the piecewise initial data U^n. Right: higher order schemes are based on the same idea, but use a reconstruction U^n_{rec} that contains more details than U^n for the initial data.*

Moreover, using the integral Formulation (6.1), we obtain than under the condition (6.2), the above formula is equivalent to

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x}(f(W(0;U_j^n, U_{j+1}^n)) - f(W(0;U_j^n, U_{j+1}^n))),$$

where $\xi \mapsto W(\xi; U_L, U_R)$ denotes the value on the line $\xi = x/t$ of the selfsimilar solution to the Riemann problem between U_L and U_R. The staggered Lax–Friedrichs scheme can be described in a similar way than the Godunov scheme, but the exact solution is averaged the on staggered cells

$$[x_{j-1/2}^{n+1}, x_{j+1/2}^{n+1}] = [x_{j-1/2}^n - \Delta x/2, x_{j+1/2}^n - \Delta x/2].$$

This scheme has the same order of the Godunov scheme but it is more diffusive. In the example of Figure 6.1, in which all the waves travel to the right, the staggered Lax–Friedrichs scheme creates two new intermediate values instead of one for the Godunov scheme. The advantage of the Lax–Friedrichs approach is that under the CFL condition (6.2), computing the flux in $x = x_{j-1/2} + \Delta x/2$ is a trivial matter since the waves have no time to reach this interface.

It is clear that information is lost in the averaging of the exact solution over cells. Many schemes, included ours, are based on the idea of reversing the averaging process to reconstruct solutions containing more details inside the cell $[x_{j-1/2}^n, x_{j+1/2}^n]$. In the MUSCL scheme of Van Leer [vL97] and the

Nessyahu–Tadmor scheme [NT90], which extend to the order 2 the Godunov scheme and the staggered Lax–Friedrichs scheme respectively, the piecewise constant function U^n is reconstructed as a piecewise linear function

$$U^n_{\text{rec}}(x) = \sum_{j\in\mathbb{Z}} U^n_j + V^n_j(x - x^n_j)\mathbf{1}_{[x^n_{j-1/2},x^n_{j+1/2}]}(x) \tag{6.3}$$

The vector of slopes V^n_j in cell j depends on U^n_{j-1}, U^n_j and U^n_{j+1}, and is chosen so that the scheme, for scalar conservation laws, is total variation diminishing. It is of course possible to use higher order polynomial reconstructions, which yields to higher order schemes. However, this reconstruction procedure is accurate in smooth regions only. Indeed, we easily imagine that approximating a function with a piecewise linear function instead of a piecewise constant function is an improvement if the function is smooth, but not if the function is discontinuous.

The discontinuous reconstruction scheme takes the opposite direction. It is not designed to be precise in smooth region, but to approach exactly isolated shock. To this purpose, instead of a piecewise regular reconstruction like (6.3), the reconstructed initial data is itself piecewise constant, with additional discontinuities inside the cells:

$$U^n_{\text{rec}}(x) = \sum_{j\in\mathbb{Z}} \left(U^n_{j,L} \times \mathbf{1}_{x<x^n_{j-1/2}+d^n_j}(x) + U^n_{j,R} \times \mathbf{1}_{x\geq x^n_{j-1/2}+d^n_j}(x)\right) \mathbf{1}_{[x^n_{j-1/2},x^n_{j+1/2}]}(x).$$

Here, d^n_j is a vector (that will be defined later) and $\times$ denotes the component by component product. This reconstruction is only of order 1 in smooth regions, but it is of "infinite order" when the initial data is a shock. Indeed, the reconstruction parameters $U^n_{j,L}$, $U^n_{j,R}$ and d^n_j are chosen such that if U^n is the exact average on the mesh of an isolated shock, then U^n_{rec} coincides that shock. Moreover, U^{n+1} will be the average on the mesh of this same shock evolved during the time $t^{n+1} - t^n$. This situation is depicted on Figure 6.2.

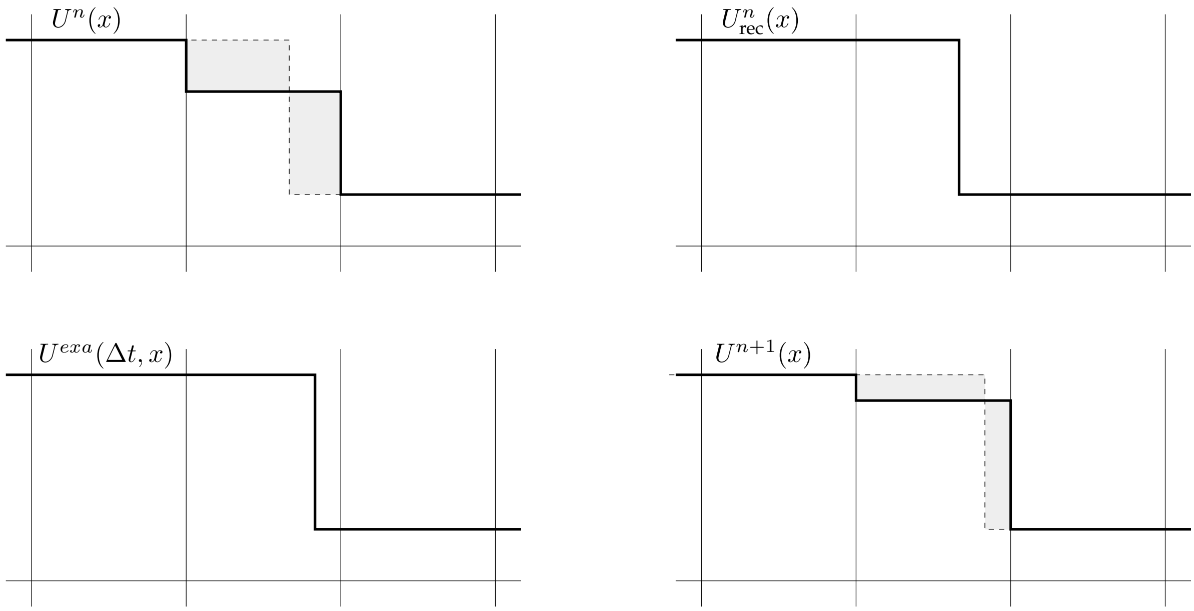


Figure 6.2. *In the discontinuous reconstruction scheme, the initial data is reconstructed as a piecewise constant data, with discontinuities lying inside the cells. In the case where the solution is an isolated shock (dotted line) it is able to reconstruct it perfectly. The gray-colored rectangles are of equal areas.*

This type of scheme was introduced in [DL01] in terms of uphill decentering under downhill constraints (the two settings are equivalent for the linear advection equation). The interpretation in terms of discontinuous reconstruction appears in [BCLL08]. Other versions of discontinuous reconstruction can be found in [ADVLCL08], where a reservoir counter is used to mimic the non-diffusive behavior of a scheme with an optimal Courant number, and in [LWM08], where the reconstruction is driven by the entropy. In [Har84], Harten proposed an ENO scheme based on a reconstruction which can have a discontinuity inside the cell. This scheme is able to approximate exactly contact discontinuities.

This strategy has been applied to the linear advection equation in [DL01], to nonclassical solutions of scalar conservation laws in [BCLL08], to gas dynamics in [DL01], [ADVLCL08] and [LWM08], to the Lifshitz–Slyozov equation in [GLT13], to two-components fluid in [DL07] and [BFBC$^+$11], ... A general framework in the case of scalar conservation laws with increasing fluxes is investigated in [LagXX], [Lag08], [Bou04a] and [Lag99], where general discussions on TVD bounds and discrete entropy inequalities can be found.

The scheme relies on two main ingredients: the detection and reconstruction of discontinuities inside each cells of the mesh, and the use of a moving mesh, just like in the staggered Lax–Friedrichs scheme. This time, it allows to avoid the resolution of multiple shock interactions. The discontinuous reconstruction scheme has the following properties:

- the scheme is exact whenever the initial datum is a pure shock, in the sense that the approximated solution is the exact solution averaged over the cells of the mesh;
- in general, the shocks are very narrow;
- no spurious oscillations in the momentum appears behind slowly moving shocks in the Euler equations;
- the scheme is able to capture nonclassical shocks, which are shocks that verify an entropy inequality but not the Liu criterion [Liu74].

The last two points are due to the fact that the discontinuous reconstruction scheme does not introduce any numerical dissipation near shocks. As we said, the main effect of numerical viscosity is to smear out shocks on several cells. But in some cases, it has more spectacular effects. In the context of gas dynamics, Quirk gave in [Qui94] a list of situations in which usual finite volume schemes are not able to capture the solution in a satisfactory way. For example the Lax–Friedrichs and Godunov schemes create an artificial spike in the momentum, followed by spurious oscillations, when computing a slowly moving shock (i.e. a shock that travels slowly compared to the speed of sound). The situation can be even worse when a second order scheme is used. Several fixes have been proposed, based on the idea of adding a sufficient amount of dissipation to the scheme, either "by hand" or by interlacing two different schemes, see for example [DM96]. Let us sketch an explanation of this problem, proposed in the article [JL96].

In this paper, Jin and Liu study the traveling waves (i.e. the solution only depending on $\frac{x-st}{\varepsilon}$, where s is a free parameter) of the system

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = \varepsilon \partial_{xx}\rho, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + c^2\rho) = \varepsilon\partial_{xx}(\rho u). \end{cases} \tag{6.4}$$

They prove than when the density is monotonous, the momentum is not monotonous if s is small. The diffusion matrix of finite volume scheme is more complicated than the one appearing in (6.4), but has an impact on the ρ equation as well. Thus, when the speed of the shock is small compared to the speed of the sound c, a spike appears in the momentum of viscous shock profiles of (6.4). However, the solutions we are interesting in are morally limits of the Navier–Stokes equations

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0 \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p(\rho)) = \varepsilon\partial_{xx} u \end{cases} \tag{6.5}$$

for which no momentum spike appears in this system when ε tends to 0. It explains why spurious oscillations appear when approaching a slowly moving shock with a scheme whose numerical diffusion looks like (6.4). To eliminate this phenomenon, the numerical diffusion must mimic the physical

Navier–Stokes viscosity (6.5), which seems to be possible with the discontinuous reconstruction strategy.

The second example concerns the numerical capture of shocks appearing as limit of diffusive-dispersive model. For example, let us look at the solutions of

$$\partial_t u + \partial_x u^3 = 0, \tag{6.6}$$

that are limits, as $\varepsilon > 0$ tends to 0, of the modified Burgers–Korteweg de Vries equation

$$\partial_t u_\varepsilon + \partial_x u_\varepsilon^3 = \varepsilon\beta\partial_{xx} u_\varepsilon + \varepsilon^2\gamma\partial_{xxx} u_\varepsilon, \tag{6.7}$$

where β and γ are positive parameters. In [JMS95], the authors prove the existence of families of traveling waves of (6.7) which tend, as $\varepsilon > 0$ tends to 0, toward a undercompressive shock of (6.6). By under compressive shock, we mean that the characteristics are not impinging on the shock wave, or in other words that the shock does not verify the Oleinik's condition [Ole57].

In Section 6.8 we focus on the system arising in nonlinear elastodynamic:

$$\begin{cases} \partial_t w - \partial_x v = 0 \\ \partial_t v - \partial_x \sigma(w) = 0 \end{cases} \tag{6.8}$$

We are interesting to the limit solution as ε tends to 0 of the system

$$\begin{cases} \partial_t w - \partial_x v = 0, \\ \partial_t v - \partial_x \sigma(w) = \varepsilon\partial_{xx} v - \alpha\varepsilon^2\partial_{xxx} w, \end{cases} \tag{6.9}$$

where $\sigma(w) = w^3 + mw$ and ε, α and m are positive. In [HL98] Hayes and Le Floch give numerical evidence that here again, shocks violating the Liu criterion [Liu74] (which extend the Lax criterion [Lax57], which extend to systems the Oleinik's criterion) appear.

Once again, the numerical diffusion competes with the physical one, and it that case it destroys the equilibrium between the diffusion and the dispersion, making the diffusion predominant. The disastrous consequence is that usual finite volume schemes, such as Lax–Friedrichs or Godunov (even if the Riemann solver takes into account the small scales of (6.9), in a sense that will be precise in Section 6.8), are unable to capture the shocks that appear as limit of the diffusive-dispersive Systems like (6.7) or (6.9). It can be seen easily on Equation (6.7): the flux is increasing, thus the Godunov flux is a simple left decentering flux, and the scheme converges toward the classical solution.

This chapter is organized as follow. In the following section, we present the discontinuous reconstruction strategy, which has three steps:

- from U^n, build the piecewise constant reconstruction

$$U^n_{\text{rec}}(x) = \sum_{j\in\mathbb{Z}} \left(U^n_{j,L} \times \mathbf{1}_{x<x^n_{j-1/2}+d^n_j} + U^n_{j,R} \times \mathbf{1}_{x<x^n_{j-1/2}+d^n_j} \right) \mathbf{1}_{[x^n_{j-1/2}, x^n_{j+1/2}]};$$

- compute (an approximation of) the exact solution at time $t^{n+1} - t^n$;
- average this exact solution on the cells $[x^{n+1}_{j-1/2}, x^{n+1}_{j+1/2}]$ to obtain $(U^{n+1}_j)_{j\in\mathbb{Z}}$.

Those steps are described in Section 6.2. The use of a moving mesh permits us to regroup the last two steps by computing the exact fluxes on the edges $x = x^n_{j-1/2} + V^n_{\text{mesh}} t$, where

$$V^n_{\text{mesh}} = \frac{x^{n+1}_j - x^n_j}{t^{n+1} - t^n}.$$

When V^n_{mesh} is large enough and under a CFL condition looking like (6.2), the flux is piecewise constant, and thus is easily computed. In Sections 6.3 and 6.4, the scheme is applied to the Burgers equations and to the barotropic Euler equations. We present in details the reconstruction procedure, and present various numerical simulations. Then in Section 6.5, we present prospective results in two space dimensions and for the Euler equations with energy. In both cases, it is necessary to handle contact discontinuities, which yields new difficulties. However the results are encouraging. In particular, the scheme approximates correctly shock reflections on a wall. Section 6.6 explains how to couple the discontinuous reconstruction strategy with the use of an approximate Riemann solver. The approximate Riemann solver must be exact on isolated shocks, and we use the one proposed in [CC13]. This a joined work with Christophe Chalons. Eventually, the last Section 6.8 focuses on System (6.8). The difficulty is to impose, at the numerical level, the so-called kinetic relation, which keeps an accurate trace on the small scales balance in the augmented System (6.9).

6.2 Construction of the scheme

We adopt classical notation for finite volume schemes. The scheme writes

$$U_j^{n+1} = U_j^n - \frac{t^{n+1} - t^n}{\Delta x}(f_{j+1/2}^n - f_{j-1/2}^n). \tag{6.10}$$

Our aim is to describe how the numerical flux $f_{j+1/2}^n$ is computed as a function of U_{j-1}^n, U_j^n, U_{j+1}^n and U_{j+2}^n. We use a moving mesh: every cells of the mesh have the same sign Δx, but their centers moves from time to time:

$$x_{j+1/2}^{n+1} = x_{j+1/2}^n + V_{\text{mesh}}^n(t^{n+1} - t^n).$$

The real V_{mesh}^n is called the speed of the mesh.

6.2.1 Detection and reconstruction of interesting discontinuities

Our aim is to build a finite volume scheme which is exact for a particular class of initial datum, called discontinuity of interest. In the case of the Euler equations, this class is the set of all entropy satisfying shocks (and contacts discontinuities when there is an equation on the energy). In the case of nonlinear elastodynamic, we will focus on the smaller class of nonclassical shocks. Consider an initial data on the form

$$U^0(x) = U_L(x)\mathbf{1}_{x\leq 0} + U_R(x)\mathbf{1}_{x>0},$$

which correspond to a shock or a contact: the left and right states are linked by the Rankine–Hugoniot relations:

$$\exists s(U_L, U_R) \in \mathbb{R} : \ s(U_L, U_R)(U_L - U_R) = f(U_L) - f(U_R). \tag{6.11}$$

$s(U_L, U_R)$ is the speed of the shock (if $U_L = U_R$, we set arbitrarily $s(U_L, U_L) = \lambda_1(U_L)$). Suppose that the shock is initially in $x_0^0 = 0$: the initial sampling is

$$U_j^0 = \begin{cases} U_L & \text{if } j < 0, \\ \frac{U_L + U_R}{2} & \text{if } j = 0, \\ U_R & \text{if } j > 0. \end{cases}$$

The middle value does not correspond to any pointwise value of the initial data, and this is the one we would like to reconstruct. To this purpose, we need a detection criterion which is able to recognize that the states $U_{-1}^0 = U_L$ and $U_1^0 = U_R$ are linked by a discontinuity of interest.

Definition 6.1. *A* reconstruction operator *is a function*

$$\mathcal{R} : \begin{array}{ccc} \Omega \times \Omega \times \Omega & \to & \Omega \times \Omega \\ (U_L, U_M, U_R) & \mapsto & (U_{M,L}, U_{M,R}) \end{array}$$

such that for all (U_L, U_R) linked by a discontinuity of interest, for all α in $(0,1)$,

$$\forall U_M, \ \mathcal{R}(U_L, \alpha U_L + (1-\alpha)U_R, U_R) = (U_L, U_R)$$

and in general, $(U_{M,L}, U_{M,R})$ *are linked by an entropy satisfying shock or a contact discontinuity. The states* $U_{M,L}$ *and* $U_{M,R}$ *are called the left and right reconstructed states (in cell* j *at time* t^n*). When*

$$\mathcal{R}(U_L, U_M, U_R) = (U_M, U_M),$$

we say that no reconstruction is performed.

The next step consists in replacing the average value U_j^n by a piecewise constant function between the left and right reconstructed states $U_{j,L}^n$ and $U_{j,R}^n$. The discontinuity is placed by conservation of the variables of U. We define the vector of distances to the left interface d_j^n by

$$\Delta x U_j^n = d_j^n \times U_{j,L} + (\Delta x - d_j^n) \times U_{j,R}. \tag{6.12}$$

The operator $\times$ denotes the component by component product of two vectors having the same size. For the components such that $U_{j,L}^n = U_{j,R}^n$, we set arbitrarily the corresponding component of d_j^n to 0.

Definition 6.2. *The* reconstructed solution at time t^n *is the piecewise constant function*

$$U_{\text{rec}}^n(x) = \sum_{j\in\mathbb{Z}} \left(U_{j,L}^n \times \mathbf{1}_{x<x_{j-1/2}^n + d_j^n} + U_{j,R}^n \times \mathbf{1}_{x<x_{j-1/2}^n + d_j^n} \right) \mathbf{1}_{[x_{j-1/2}^n, x_{j+1/2}^n]} \tag{6.13}$$

where $(U_{j,L}^n, U_{j,R}^n) = \mathcal{R}(U_{j-1}^n, U_j^n, U_{j+1}^n)$ *and* d_j^n *is defined by* (6.12).

Choice (6.12) permits to reconstruct perfectly isolated shock.

Proposition 6.1. *Suppose that* U_{j-1}^n *and* U_{j+1}^n *are linked by a discontinuity of interest, and that there exists* $\alpha \in [0,1]$ *such that* $U_j^n = \alpha U_{j-1}^n + (1-\alpha) U_{j+1}^n$*. Then* $U_{j,L}^n = U_{j-1}^n$*,* $U_{j,R}^n = U_{j+1}^n$ *and all the component of* d_j^n *are equals to* $\alpha \Delta x$*.*

Proof. The proof is straightforward: by definition of the detection criterion, $U_{j,L}^n = U_{j-1}^n$ and $U_{j,R}^n = U_{j+1}^n$, and the definition of d_j^n yields immediately that all its components are equal to $\alpha\Delta x$. □

When U_{j-1}^n and U_{j+1}^n are not linked by a shock of interest, the components of d_j^n are different and it is possible that one of them (or more) does not belong to $[0, \Delta x]$. In that case, the reconstruction of the corresponding variable is not conservative:

$$U_j^n \neq \int_{x_{j-1/2}}^{x_{j+1/2}} U_{j,L}^n \times \mathbf{1}_{x<x_{j-1/2}^n + d_j^n} + U_{j,R}^n \mathbf{1}_{x>x_{j-1/2}^n + d_j^n} \, dx.$$

Authorizing one component of d_j^n to be outside of the interval $[0, \Delta x]$ is the key point to approach correctly slowly moving shocks (see Section 6.4.2). Even though the reconstruction procedure is not conservative, the finite volume scheme will be conservative as it will have the form (6.10).

The choice of the reconstruction operator $\mathcal{R}$ is crucial. On the one hand, if

$$\mathcal{R}(U_L, U_M, U_R) = (U_M, U_M)$$

too often, U_{rec}^n is very close to U^n and the scheme behaves like the Lax–Friedrichs scheme. Once again, this will be illustrated by the slowly moving shocks test case. On the other hand, if reconstructions are not performed carefully, i.e. if

$$\mathcal{R}(U_L, U_M, U_R) \neq (U_M, U_M)$$

too often, U_n^{rec} is too far from U^n and the scheme may roughly speaking give anything. In Theorem 6.3 below, we give constraints on $\mathcal{R}$ which ensure that the scheme is exact on isolated shocks.

6.2.2 Advection of the reconstructed discontinuities and computation of the flux

We now want to compute $(U_j^{n+1})_{j\in\mathbb{Z}}$ as the average on the mesh of the exact solution of

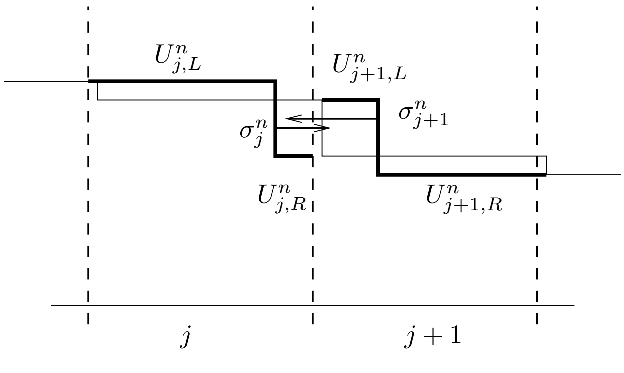
$$\begin{cases} \partial_t U + \partial_x f(U) = 0, \\ U^0(x) = U^n_{\text{rec}}(x). \end{cases}$$

However, when the component of d_j^n are not all equals, this is a difficult task that involves the resolution of multiple wave interactions. We make a rough approximation of this solution by considering that the exact solution of

$$\begin{cases} \partial_t U + \partial_x f(U), \\ U^0(x) = U^n_{j,L} \times \mathbf{1}_{x<x^n_{j-1/2}+d^n_j} + U^n_{j,R} \times \mathbf{1}_{x<x^n_{j-1/2}+d^n_j}, \end{cases}$$

is $U(t,x) = U^0(x - s(U^n_{j,L}, U^n_{j,R})t)$. In other words, we consider that the discontinuities in each variables of U travels independently from the others, but that they all have the same speed $s(U^n_{j,L}, U^n_{j,R})$. It recalls the ambiguity in stationary shock position underlined in [Zai12]. This uncertainty is somehow dispatched between the components of U by considering different positions of the shock for each components.

To illustrate the importance of using a moving mesh, consider the case of a single conservation law in which two shocks are reconstructed in two adjacent cells. Suppose that the reconstructed shock in cell j has a positive speed, that the one in cell $j+1$ has a negative speed, and that they are both close to $x^n_{j+1/2}$, as depicted on the left of Figure 6.3. Then the two shocks can interact during the time step, and the waves emitted by their interaction can cross the interface $x = x^n_{j+1/2}$. A reconstructed shock can also interact with the waves created by the discontinuity between $U^n_{j,R}$ and $U^n_{j+1,L}$ located in $x^n_{j+1/2}$. In any case, it becomes impossible to compute the flux across this interface without resolving the waves interactions. Moreover, imposing a time step $t^{n+1} - t^n$ smaller than the time when the waves interact only postpones the problem to the next time iteration. A solution is to fix a mesh speed V^n_{mesh} larger



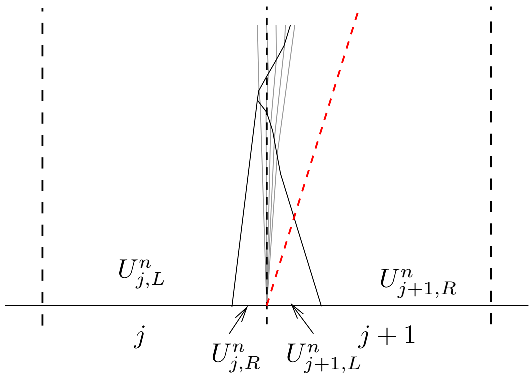


Figure 6.3. *Interactions of two reconstructed shocks with each other and with a rarefaction fan. On the left, the bold lines represent the reconstruction and the arrows the speed of the shocks. On the right, the waves created by the reconstructed solution are drew in the (x,t) plane (black corresponds to discontinuity, grey to smooth area). Computing the flux along $x = x_{j+1/2}$ (black dotted line) is very difficult, but compute it along $x = x_{j+1/2} + V^n_{\text{mesh}}$ (red dotted line) only requires to know when the right shock crosses the interface.*

than the maximum of the waves speed

$$|V^n_{\text{mesh}}| > V^n_{\text{waves}}, \tag{6.14}$$

where V^n_{waves} is the maximum of the waves speed appearing in the Riemann problems between $U^n_{j,R}$ and $U^n_{j+1,L}$ and of the reconstructed shocks, and to compute the flux on the bended interfaces $x =$

$x^n_{j+1/2} + V^n_{\text{mesh}} t$. If (6.14) and the CFL condition

$$t^{n+1} - t^n < \frac{\Delta x}{V^n_{\text{waves}} + |V^n_{\text{mesh}}|} \tag{6.15}$$

hold, this is an easy task. Indeed:

- Condition (6.14) ensures that the waves emitted at $x^n_{j+1/2}$ do not cross the interface $x = x^n_{j+1/2} + V^n_{\text{mesh}} t$;
- Condition (6.15) ensures that they do not cross the interfaces $x = x^n_{j-1/2} + V^n_{\text{mesh}} t$ and $x = x^n_{j+3/2} + V^n_{\text{mesh}} t$ either;
- similarly, the shock reconstructed in cell j may cross the interface $x = x^n_{j-1/2} + V^n_{\text{mesh}} t$ if V^n_{mesh} is positive (the interface $x = x^n_{j+1/2} + V^n_{\text{mesh}} t$ if V^n_{mesh} is negative) and just this one;
- Eventually, if two waves interact, the resulting interaction will not have time to catch up any interface.

This is illustrated in Figure 6.3, on the right. The last point holds true for scalar conservation law, because the maximum of the waves speed after the interaction is smaller than the maximum of the waves speed before it. In system of conservation laws, it is possible that an interaction of waves emits a wave traveling faster than the initial ones. Thus is it wise to stay away from the equality in Equations (6.14) and (6.15), at least for the first iterations in time.

The *Courant number* λ links the time step proportionally to the time step:

$$\text{for all } n, \quad t^{n+1} - t^n = \frac{\lambda \Delta x}{V^n_{\text{waves}}}.$$

In the sequel the mesh speed is taken such that for all n,

$$V^n_{\text{mesh}} = \frac{(-1)^{n+1} \Delta x}{2(t^{n+1} - t^n)}. \tag{6.16}$$

If λ is smaller than $1/2$, Conditions (6.14) and (6.15) hold, and more precisely

$$t^{n+1} - t^n = \frac{(\lambda + 1/2)\Delta x}{|V^n_{\text{mesh}}| + V^n_{\text{waves}}}.$$

Choice (6.16) is not optimal, and we obtain a lower numerical diffusion (in smooth regions) when it is taken closer to V^n_{mesh}. We are now in position to compute the fluxes.

- For odd iteration in time, the mesh speed is positive and the only wave that can cross the left interface $x = x^n_{j-1/2} + V^n_{\text{mesh}}$ during the time step is the shock reconstructed in the j-th cell. This shock is placed at different positions for each conserved variables. Thus we compute the *vector of crossing times* T^n_j, which verifies

  $$d^{2n-1}_j + s(U^{2n-1}_{j,L}, U^{2n-1}_{j,R}) T^{2n-1}_j = V^{2n-1}_{\text{mesh}} T^{2n-1}_j.$$

  Thus T^{2n-1}_j is given by

  $$T^{2n-1}_j = \frac{d^{2n-1}_j}{V^{2n-1}_{\text{mesh}} - s(U^{2n-1}_{j,L}, U^{2n-1}_{j,R})}. \tag{6.17}$$

  The flux passing through the left interface is piecewise constant. Indeed for $t < t^{2n} - t^{2n-1}$ we have

  $$U(t, x^{2n-1}_{j-1/2} + V^{2n-1}_{\text{mesh}} t) = \begin{cases} U^{2n-1}_{n,L} & \text{before the crossing time,} \\ U^{2n-1}_{n,R} & \text{after the crossing time,} \end{cases}$$

as illustrated on Figure 6.3. Considering that the discontinuity may not cross the interface during the time step (if $T_j^{2n-1} > t^{n+1} - t^n$) and that T_j^{2n-1} has negative components where d_j^{2n-1} has some, we obtain the numerical flux:

$$\begin{aligned} f_{j-1/2}^{2n-1} = & (f(U_{j,L}^{2n-1}) - V_{\text{mesh}}^{2n-1} U_{j,L}^{2n-1}) \times \max\left(0, \min\left(\frac{T_j^{2n-1}}{t^{2n} - t^{2n-1}}, 1\right)\right) \\ & + (f(U_{j,R}^{2n-1}) - V_{\text{mesh}}^{2n-1} U_{j,R}^{2n-1}) \times \left[1 - \max\left(0, \min\left(\frac{T_j^{2n-1}}{t^{2n} - t^{2n-1}}, 1\right)\right)\right]. \end{aligned} \tag{6.18}$$

- For even iteration in time, V_{mesh}^{2n} is negative and verifies (6.14). The reconstructed shock in cell j crosses the right interface $x = x_{j+1/2}^{2n} + V_{\text{mesh}}^{2n} t$. Thus the vector of crossing times is

$$T_j^{2n} = \frac{\Delta x - d_j^{2n}}{s(U_{j,L}^{2n}, U_{j,R}^{2n}) - V_{\text{mesh}}^{2n}}, \tag{6.19}$$

and the flux is

$$\begin{aligned} f_{j+1/2}^{2n} = & (f(U_{j,R}^{2n}) - V_{\text{mesh}}^{2n} U_{j,R}^{2n-1}) \times \max\left(0, \min\left(\frac{T_j^{2n}}{t^{2n+1} - t^{2n}}, 1\right)\right) \\ & + (f(U_{j,L}^{2n}) - V_{\text{mesh}}^{2n} U_{j,L}^{2n-1}) \times \left[1 - \max\left(0, \min\left(\frac{T_j^{2n}}{t^{2n+1} - t^{2n}}, 1\right)\right)\right]. \end{aligned} \tag{6.20}$$

Remark 18. On a component where d_j^n is negative, the quantity

$$\max\left(0, \min\left(\frac{T_j^{2n}}{t^{2n+1} - t^{2n}}, 1\right)\right)$$

is equal to 0, while on a component where d_j^n is larger than Δx, the CFL condition (6.15) yields that it is equal to 1. Thus, the fluxes (6.18) and (6.20) are the same if they are computed with $\max(0, \min(\Delta x, d_j^n))$ instead of d_j^n in (6.17) and (6.19). This corresponds to nonconservative reconstructions, where U_j^n is replaced by another constant value in cell j ($U_{j,R}^n$ and $U_{j,L}^n$ respectively). However, the scheme is conservative in that case, since it has the form (6.10).

Remark 19. When no reconstruction is performed, we have $U_{j,L}^n = U_{j,R}^n = U_j^n$ and the fluxes (6.18) and (6.20) coincide with the staggered Lax–Friedrichs fluxes:

$$\begin{cases} f_{j-1/2}^{2n-1} & = f(U_j^{2n-1}) - V_{\text{mesh}}^{2n-1} U_j^{2n-1}, \\ f_{j+1/2}^{2n} & = f(U_j^{2n}) - V_{\text{mesh}}^{2n} U_j^{2n}. \end{cases}$$

Remark 20. On a component where $U_{j,L}^n = U_{j,R}^n$, Fluxes (6.18) and (6.20) do not depend on d_j^n and $s(U_{j,L}^n, U_{j,R}^n)$. Thus the arbitrary values we gave to those quantities is not important.

Remark 21. In practice for all the numerical simulation, we take

$$\forall n \in \mathbb{N}, |V_{\text{mesh}}^n| = V_{\text{waves}}^n = \frac{\Delta x}{2(t^n - t^{n-1})}.$$

6.2.3 Properties of the scheme

The first proposition, which is trivial to prove, states that the scheme is consistent.

Proposition 6.2. *If the reconstruction operator $\mathcal{R}$ is such that*

$$\forall U,\ \mathcal{R}(U,U,U) = (U,U)$$

then the flux is consistent: if $U_j^n = U$ for all $j \in \{j_0-2, j_0-1, j_0, j_0+1, j_0+2\}$, then

$$f^n_{j_0-1/2} = f^n_{j_0+1/2} = f(U) - V^n_{\text{mesh}} U.$$

The next theorem explores the link between the reconstruction operator $\mathcal{R}$ and the approximation of discontinuities of interest.

Theorem 6.3. *Let U_L and U_R be two states linked by a discontinuity of interest. If the reconstruction operator $\mathcal{R}$ verifies*

$$\forall \alpha \in [0,1],\ \mathcal{R}(U_L, \alpha U_L + (1-\alpha) U_R, U_R) = (U_L, U_R)$$

$$\forall \alpha \in (0,1),\ \mathcal{R}(U_L, U_L, \alpha U_L + (1-\alpha) U_R) = (U_L, U_L)$$

and

$$\forall \alpha \in (0,1),\ \mathcal{R}(\alpha U_L + (1-\alpha) U_R, U_R, U_R) = (U_R, U_R).$$

Then, the discontinuous reconstruction scheme is exact when the initial data is

$$U^0(x) = U_L \mathbf{1}_{x<0} + U_R \mathbf{1}_{x \geq 0}.$$

In other words,

$$\forall n \geq 0, \forall j \in \mathbb{Z},\ U_j^n = \frac{1}{\Delta x} \int_{x^n_{j-1/2}}^{x^n_{j+1/2}} U_L \mathbf{1}_{x < s(U_L,U_R)t} + U_R \mathbf{1}_{x \geq s(U_L,U_R)t},$$

where $s(U_L, U_R)$ is the speed of the discontinuity, defined by (6.11).

Proof. We proceed by induction. For the sake of simplicity we take $V^n_{\text{mesh}} = 0$ for all n and we suppose that $s(U_L, U_R)$ is positive. It follows that the reconstructed shock crosses the right interface of the cell where it lies, so the flux is given by (6.18).

The property holds at time $n = 0$ by definition of the initial sampling. Suppose that is holds true at time n, and denote by j_0^n the cell where the shock lies and by δ^n its distance to the left interface:

$$s(U_L,U_R)t^n \in [x_{j_0^n-1/2}, x_{j_0^n-1/2}), \quad s(U_L,U_R)t^n - x_{j_0^n-1/2} = \delta^n$$

and

$$\begin{cases} U_j^n = U_L \text{ if } j < j_0, \\ U_{j_0}^n = \frac{\delta^n}{\Delta x} U_L + \frac{\Delta x - \delta^n}{\Delta x} U_R, \\ U_j^n = U_R \text{ if } j > j_0/ \end{cases}$$

We want to prove that

$$U^n_{\text{rec}}(x) = U_L \mathbf{1}_{x < s(U_L,U_R)t} + U_R \mathbf{1}_{x \geq s(U_L,U_R)t}.$$

Let us first treat the special case where the shock is at an interface: $\delta^n = 0$. By Definitions 6.1 and 6.12; we have

$$\mathcal{R}(U_{j_0^n-2}, U_{j_0^n-1}, U_{j_0^n}) = (U_L, U_R) \text{ and } d^n_{j_0^n-1} = (\Delta x, \cdots, \Delta x)$$

and

$$\mathcal{R}(U_{j_0^n-1}, U_{j_0^n}, U_{j_0^n+1}) = (U_L, U_R) \text{ and } d^n_{j_0^n} = (0, \cdots, 0).$$

Thus the numerical flux (6.18) coincides with the Lax–Friedrichs flux everywhere, which gives the correct solution at the next time iteration.

If δ^n is strictly positive, the first hypothesis of the theorem ensures that the correct reconstruction is performed in the j_0^n-th cell:

$$\mathcal{R}(U_{j_0^n-1}, U_{j_0}, U_{j_0^n+1}) = (U_L, U_R) \text{ and } d_{j_0^n}^n = (\delta^n, \cdots, \delta^n)$$

The last two conditions imply that no reconstruction is performed in cells $j_0^n - 1$ and $j_0^n + 1$, and thus $f_{j_0^n-1/2}^n = f(U_L)$ and $f_{j_0^n+3/2}^n = f(U_R)$.

Let us now focus on the computation of the flux $f_{j_0^n+1/2}^n$. Under the CFL condition (6.15), only two cases arise: either the shock is still in cell j_0^n at time t^{n+1}, or it crosses the right interface to end up in cell j_0^n+1. We set accordingly $j_0^{n+1} = j_0^n$ and $j_0^{n+1} = j_0^n+1$. In the first case, $\delta^n + s(U_L, U_R)(t^{n+1} - t^n) < \Delta x$ and at time t^{n+1}, the shock is at a distance $\delta^{n+1} = \delta^n + s(U_L, U_R)(t^{n+1} - t^n)$ from $x_{j_0^{n+1}-1/2}$. All the component of d_j^n are equals to δ^n, thus all the crossing times are equals to

$$T_{j_0^n} = \frac{\Delta x - \delta^n}{s(U_j, U_R)},$$

which is larger than Δt. It yields that $f_{j_0^n+1/2}^n$ is a simple right decentering. Hence we have

$$\begin{cases} U_j^{n+1} & = U_L \qquad \text{if } j < j_0^{n+1} \\ U_{j_0^{n+1}}^{n+1} & = U_j^n - \frac{t^{n+1}-t^n}{\Delta x}(f(U_R) - f(U_L)) \\ & = \frac{\delta^n}{\Delta x} U_L + \frac{\Delta x - \delta^n}{\Delta x} U_R - \frac{t^{n+1}-t^n}{\Delta x} s(U_R, U_L)(U_R - U_L) \\ & = \frac{\delta^{n+1}}{\Delta x} U_L + \frac{\Delta x - \delta^{n+1}}{\Delta x} U_R \\ U_j^{n+1} & = U_R \qquad \text{if } j > j_0^{n+1} \end{cases}$$

and the result hold true at time $n+1$.

In the second case, $\delta^n + s(U_L, U_R)(t^{n+1} - t^n) \geq \Delta x$ and at time t^{n+1}, the shock is at a distance $\delta^{n+1} = \delta^n + s(U_L, U_R)(t^{n+1} - t^n) - \Delta x$ from $x_{j_0^{n+1}-1/2}$. A straightforward computation gives

$$\begin{cases} U_{j+1}^n & = U_L \qquad \text{if } j < j_0^{n+1} \\ U_{j_0^{n+1}-1}^{n+1} & = U_j^n - \frac{t^{n+1}-t^n}{\Delta x}\left(\frac{T_{j_0^n}}{t^{n+1}-t^n} f(U_R) + \left(1 - \frac{T_{j_0^n}}{t^{n+1}-t^n}\right) f(U_L) - f(U_L)\right) \\ & = U_j^n - \frac{T_{j_0^n}}{\Delta x}(f(U_R) - f(U_L)) \\ & = \frac{\delta^n}{\Delta x} U_L - \frac{\Delta x - \delta^n}{s(U_L, U_R)} \frac{s(U_L, U_R)(U_R - U_L)}{\Delta x} \\ & = U_L \\ U_{j_0^{n+1}}^{n+1} & = U_R - \frac{t^{n+1}-t^n}{\Delta x}\left(f(U_R) - \left(\frac{T_{j_0^n}}{t^{n+1}-t^n} f(U_R) + \left(1 - \frac{T_{j_0^n}}{t^{n+1}-t^n}\right) f(U_L)\right)\right) \\ & = U_R - \frac{t^{n+1}-t^n-T_{j_0^n}}{\Delta x}(f(U_R) - f(U_L)) \\ & = U_R - \frac{t^{n+1}-t^n-\frac{\Delta x-\delta^n}{s(U_L,U_R)}}{\Delta x} s(U_L, U_R)(U_R - U_L) \\ & = U_R - \frac{\delta^n + (t^{n+1}-t^n)s(U_L,U_R) - \Delta x}{\Delta x}(U_R - U_L) \\ & = \frac{\delta^{n+1}}{\Delta x} U_L + \frac{\Delta x - \delta^{n+1}}{\Delta x} U_R \\ U_j^{n+1} & = U_R \qquad \text{if } j > j_0^{n+1}. \end{cases}$$

which is exactly the average of the exact solution at time t^{n+1}.

□

6.3 The reconstruction scheme for the Burgers equation

The Burgers equation is an example of scalar conservation law with stricly convex flux. The Cauchy problem writes

$$\begin{cases} \partial_t u + \partial_x \frac{u^2}{2} = 0, \\ u(t=0, x) = u^0(x). \end{cases} \tag{6.21}$$

There exists a unique solution that also verifies the following entropy inequality:

$$\partial_t \frac{u^2}{2} + \partial_x \frac{u^3}{3} \leq 0.$$

This case is particularly simple because all states u_L and u_R with u_L larger than u_R are linked by an entropy satisfying shock. To define the reconstruction operator $\mathcal{R}$, we first introduce an operator $\mathcal{CR}$ which gives candidates to the reconstructed states. Then, we allow some reconstructions to be rejected. We only want to reconstruct entropy shocks, so we set

$$\mathcal{CR}(u_L, u_M, u_R) = (\bar{u}_{M,L}, \bar{u}_{M,R}) := \begin{cases} (u_L, u_R) & \text{if } u_L > u_R, \\ (u_M, u_M) & \text{otherwise .} \end{cases}$$

We accept the reconstruction if d_j^n, which is

$$d_j^n = \Delta x \frac{u_j^n - \bar{u}_{j,R}^n}{\bar{u}_{j,L}^n - \bar{u}_{j,R}^n},$$

by (6.12), belongs to $(0, \Delta x)$:

$$\mathcal{R}(u_{j-1}^n, u_j^n, u_{j+1}^n) = \begin{cases} \mathcal{CR}(u_{j-1}^n, u_j^n, u_{j+1}^n) & \text{if } d_j^n \in (0, \Delta x) \\ (u_j^n, u_j^n) & \text{otherwise .} \end{cases}$$

We perform a numerical test that exhibits the non diffusive behavior of the reconstruction scheme. The initial data is

$$u^0(x) = \begin{cases} 3 & \text{if } x \leq -3, \\ 3 - (x+3) & \text{if } -3 \leq x \leq -1, \\ 1 & \text{if } -1 \leq x. \end{cases}$$

The exact solution is given by

$$u(t,x) = \begin{cases} 3 * \mathbf{1}_{x \leq -3+3t} + \left(3 - \frac{x-(3+3t)}{(1-t)}\right) \mathbf{1}_{-3+3t < x < -1+t} + \mathbf{1}_{-1+t \leq x} & \text{if } t < 1, \\ 3 * \mathbf{1}_{x \leq 2(t-1)} + \mathbf{1}_{2(t-1) < x} & \text{if } t \geq 1. \end{cases}$$

We compare the Godunov scheme and the reconstruction scheme on the interval $[-4, 2]$ discretized with 100 cells. The Courant number is $\lambda = 0.4$. As depicted on Figure 6.4, the shock is perfectly reconstructed at time $T = 1$ with the reconstruction scheme, while it is smeared on about 10 cells by the Godunov scheme. After time $T = 1$, the shock is exactly advected by the reconstruction scheme, while the Godunov scheme keeps diffusing it. The apparition of steps inside the isentropic compression is very similar to those which appears in the smooth regions in [DL01] and [Lag08]. It does not affect the order of convergence of the scheme. Entropy corrections have been proposed in [Lag99], [LagXX] and [Bou04a] and no such steps appear in [LWM08].

6.4 The reconstruction scheme for the barotropic Euler equations

This section is devoted to the generalization of the reconstruction scheme to a particular system of conservation laws, the isothermal Euler equations. It describes the evolution of an inviscid compressible gas, having density ρ and velocity u, when the pressure law is $p = p(\rho) = c^2\rho$. Here, c is the speed of sound. This pressure law avoids us to deal with vacuum, but the scheme extend straightforwardly to pressure laws $p(\rho) = a\rho^\gamma$ with $a > 0$ and $\gamma > 1$ when there is no vacuum. The Cauchy problem for this system writes

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + c^2 \rho) = 0, \\ \rho(t=0, x) = \rho^0(x), \\ u(t=0, x) = u^0(x). \end{cases}$$

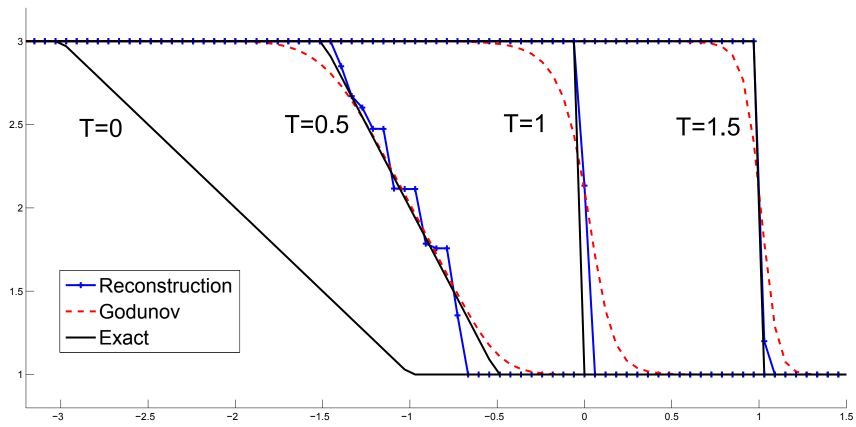


Figure 6.4. *Simulation of an isentropic compression with the Godunov scheme and the reconstruction scheme.*

Introducing the momentum $q = \rho u$, it rewrites as a system of two conservation laws

$$\begin{cases} \partial_t \rho + \partial_x q = 0, \\ \partial_t q + \partial_x \left(\frac{q^2}{\rho} + c^2 \rho \right) = 0, \\ \rho(t = 0, x) = \rho^0(x), \\ q(t = 0, x) = q^0(x) = \rho^0(x) u^0(x). \end{cases} \tag{6.22}$$

In the sequel, $U = (\rho, q)$ and

$$\Omega = \{(\rho, q) \in \mathbb{R}_+^* \times \mathbb{R}\}.$$

6.4.1 How to detect shock and when to accept the reconstruction

System (6.22) is strictly hyperbolic, the eigenvalues of the Jacobian matrix being $u - c$ and $u + c$. The corresponding fields are genuinely nonlinear. This system has been widely studied. We are here interesting in the structure of the Riemann problem

$$\begin{cases} \rho^0(x) & = \rho_L \mathbf{1}_{x<0} + \rho_R \mathbf{1}_{x>0}, \\ q^0(x) & = q_L \mathbf{1}_{x<0} + q_R \mathbf{1}_{x>0}. \end{cases}$$

Its solution consists of the succession of two simple waves, rarefactions or shocks, separated by an intermediate state. As a consequence, unless we start with very specific initial data, there is no reason for the states $(\rho_{j-1}^n, q_{j-1}^n)$ and $(\rho_{j+1}^n, q_{j+1}^n)$ to be linked by an entropy satisfying shock, and the Riemann problem between these two states is more likely to develop the full pattern of two waves. How to decide for a reconstruction procedure in this framework? We try to reconstruct the state (ρ_j^n, q_j^n) as the average of *one of the waves* appearing in the Riemann problem between its neighbor states $(\rho_{j-1}^n, q_{j-1}^n)$ and $(\rho_{j+1}^n, q_{j+1}^n)$. This wave must be an entropy satisfying shock and must somehow be the dominant wave in the Riemann problem. We use the following lemma to decide wether we try to reconstruct and which wave is chosen.

Lemma 6.4. *If $u_L > u_R$ and $\rho_L < \rho_R$, then the Riemann problem between (ρ_L, q_L) and (ρ_R, q_R) contains a 1-shock . If $u_L > u_R$ and $\rho_L > \rho_R$, it contains a 2-shock.*

Remark 22. In the previous lemma, the Riemann problem may contain a 1-shock and a 2-shock.

Proof. Let us denote by (ρ_I, u_I) the intermediate state of this Riemann problem. As the velocity increases through rarefaction waves, the Riemann problem contains at least one shock. Moreover, the density decreases through 1-rarefaction waves and 2-shocks, and increases through 1-shocks and 2-rarefaction waves. Hence, if $\rho_L < \rho_R$, the Riemann problem contains a 1-shock, while if $\rho_L > \rho_R$, it contains a 2-shock. $\square$

Lemma 6.4 gives us a criterion to define an operator $\mathcal{CR}$ that provides candidates reconstructed states

$$(\bar{U}^n_{j,L}, \bar{U}^n_{j,R}) = \mathcal{CR}(U^n_{j-1}, U^n_j, U^n_{j+1}).$$

The reconstruction operator $\mathcal{R}$ is defined from $\mathcal{CR}$ buy canceling some reconstructions.

We denote by $(\rho^n_{j,*}, u^n_{j,*})$ the intermediate state in the Riemann problem between $(\rho^n_{j-1}, q^n_{j-1})$ and $(\rho^n_{j+1}, q^n_{j+1})$. The left and right candidates reconstructed states $(\bar{\rho}^n_{j,L}, \bar{q}^n_{j,L})$ and $(\bar{\rho}^n_{j,R}, \bar{q}^n_{j,R})$ are

$$(\bar{\rho}^n_{j,L}, \bar{q}^n_{j,L}) = \begin{cases} (\rho^n_{j-1}, q^n_{j-1}) & \text{if } u^n_{j-1} > u^n_{j+1} \text{ and } \rho^n_{j-1} < \rho^n_{j+1}; \\ (\rho^n_{j,*}, q^n_{j,*}) & \text{if } u^n_{j-1} > u^n_{j+1} \text{ and } \rho^n_{j-1} > \rho^n_{j+1}; \\ (\rho^n_j, q^n_j) & \text{otherwise;} \end{cases} \tag{6.23}$$

and

$$(\bar{\rho}^n_{j,R}, \bar{q}^n_{j,R}) = \begin{cases} (\rho^n_{j,*}, q^n_{j,*}) & \text{if } u^n_{j-1} > u^n_{j+1} \text{ and } \rho^n_{j-1} > \rho^n_{j+1}; \\ (\rho^n_{j+1}, q^n_{j+1}) & \text{if } u^n_{j-1} > u^n_{j+1} \text{ and } \rho^n_{j-1} < \rho^n_{j+1}; \\ (\rho^n_j, q^n_j) & \text{otherwise.} \end{cases} \tag{6.24}$$

It corresponds respectively to a 1-shock reconstruction, a 2-shock reconstruction and no reconstruction.

Remark 23. If $(\rho^n_{j-1}, q^n_{j-1})$ and $(\rho^n_{j+1}, q^n_{j+1})$ are linked by a shock, we have

$$(\bar{\rho}^n_{j,L}, \bar{q}^n_{j,L}) = (\rho^n_{j-1}, u^n_{j-1}) \text{ and } (\bar{\rho}^n_{j,R}, \bar{q}^n_{j,R}) = (\rho^n_{j+1}, q^n_{j+1}),$$

as required in Definition 6.1. This is crucial to approach exactly shocks, see Theorem 6.3.

The Rankine–Hugoniot relation (6.11) gives an explicit formula for the speed $s(U^n_{j-1}, U^n_{j+1})$ associated to the shock:

$$s(U^n_{j-1}, U^n_{j+1}) = \begin{cases} u^n_{j-1} - c\sqrt{\frac{\rho^n_{j,*}}{\rho^n_{j-1}}} & \text{if } u^n_{j-1} > u^n_{j+1} \text{ and } \rho^n_{j-1} > \rho^n_{j+1}; \\ u^n_{j+1} + c\sqrt{\frac{\rho^n_{j,*}}{\rho^n_{j+1}}} & \text{if } u^n_{j-1} > u^n_{j+1} \text{ and } \rho^n_{j-1} < \rho^n_{j+1}; \\ u^n_j - c\sqrt{\frac{\rho^n_j}{\rho^n_j}} & \text{otherwise;} \end{cases} \tag{6.25}$$

Following (6.12), the vector $d^n_j = (d^{n,\rho}_j, d^{n,q}_j)$ of distances to the left interface is given by

$$d^{n,\rho}_j = \Delta x \frac{\bar{\rho}^n_{j,R} - \bar{\rho}^n_j}{\bar{\rho}^n_{j,R} - \bar{\rho}^n_{j,L}} \text{ and } d^{n,q}_j = \Delta x \frac{\bar{q}^n_{j,R} - \bar{q}^n_j}{\bar{q}^n_{j,R} - \bar{q}^n_{j,L}}. \tag{6.26}$$

and we set them arbitrarily to 0 its components that are not defined.

We now define $\mathcal{R}$. The most natural choice is to accept the reconstruction when $0 < d^{n,\rho}_j < \Delta x$ and $0 < d^{n,q}_j < \Delta x$, i.e. when the reconstructed shock lies inside the cell for both reconstruction in ρ and q. This choice will be referred to as the *fully conservative reconstruction*. It corresponds to the reconstruction operator

$$\mathcal{R}_{FC}(U^n_{j-1}, U^n_j, U^n_{j+1}) = \begin{cases} \mathcal{CR}(U^n_{j-1}, U^n_j, U^n_{j+1}) \text{ if } d^{n,\rho}_j \in (0, \Delta x) \text{ and } d^{n,q}_j \in (0, \Delta x), \\ (U^n_j, U^n_j) \text{ otherwise .} \end{cases}$$

We also consider the case where the reconstruction is accepted whenever $0 < d_j^{n,\rho} < \Delta x$. No condition is required on $d_{j,n}^q$, which means that the reconstruction on the momentum can be nonconservative. This choice will be referred to as the *half conservative reconstruction*, defined by

$$\mathcal{R}_{HC}(U_{j-1}^n, U_j^n, U_{j+1}^n) = \begin{cases} \mathcal{CR}(U_{j-1}^n, U_j^n, U_{j+1}^n) \text{ if } d_j^{n,\rho} \in (0, \Delta x), \\ (U_j^n, U_j^n) \text{ otherwise }. \end{cases}$$

The corresponding reconstruction operator is richer than $\mathcal{R}_{FC}$. We recall that authorizing nonconservative reconstruction does not prevent the scheme from being conservative. We now prove that those schemes are exact on isolated shocks.

Proposition 6.5. *The schemes corresponding to the conditions $\mathcal{R}_{FC}$ or $\mathcal{R}_{HC}$ are consistent and exact on isolated shocks.*

Proof. The consistency follow immediately from Proposition 6.2. To prove that the schemes are exact when the initial data is a shock, we prove that the three hypothesis of Theorem 6.3 hold for any shock. Fix a real α in $(0, 1)$ and denote by

$$(\rho_M, q_M) = \alpha(\rho_L, q_L) + (1-\alpha)(\rho_R, q_R).$$

Denote by $u_M = \frac{q_M}{\rho_M}$ the corresponding velocity. We prove the result for a 1-shock, thus we suppose $\rho_L < \rho_R$. Let us first check that $u_R \geq u_M \geq u_L$. This ensures that Lemma 6.4 detects a 1-shock:

$$\mathcal{R}(U_L, U_M, U_R) = (U_L, U_R).$$

We have

$$\begin{aligned} u_M - u_L &= \frac{\alpha q_L + (1-\alpha)q_R}{\alpha\rho_L + (1-\alpha)\rho_R} - \frac{q_L}{\rho_L} \\ &= \frac{\rho_L(\alpha q_L + (1-\alpha)q_R) - q_L(\alpha\rho_L + (1-\alpha)\rho_R)}{\rho_L(\alpha\rho_L + (1-\alpha)\rho_R)} \\ &= \frac{(1-\alpha)(\rho_L q_R - \rho_R q_L)}{\rho_L(\alpha\rho_L + (1-\alpha)\rho_R)} \\ &= \frac{(1-\alpha)\rho_L\rho_R(u_R - u_L)}{\rho_L(\alpha\rho_L + (1-\alpha)\rho_R)} \\ &\leq 0, \end{aligned}$$

and we obtain similarly that u_M is larger than u_R. This proves the first point of Theorem 6.3. However for the same reason, Lemma 6.4 detects a 1-shock between U_L and U_M and between U_M and U_R:

$$\mathcal{CR}(U_L, U_L, U_M) \neq (U_L, U_L)$$

and

$$\mathcal{CR}(U_M, U_R, U_R) \neq (U_R, U_R)$$

Between U_L and U_M, the left candidate reconstructed state is (ρ_L, q_L). Thus the density cannot be reconstructed in a conservative way and $\mathcal{R}_{HC}(U_L, U_L, U_M) = (U_L, U_L)$. This proves the second hypothesis of Theorem 6.3.

We now prove the last hypothesis of this theorem. We denote by U_* the intermediate state appearing in the Riemann problem between U_M and U_R. We know that the solution contains a 1-shock by Lemma 6.4. If U_* and U_R are linked by a 2-rarefaction wave, both the left and the right candidates reconstructed densities are larger than ρ_R. Once again, we cannot reconstruct ρ in a conservative fashion. We now prove that U_* and U_R can not be linked by a 2-shock. If it was possible, U_* would belong to the light gray region in Figure 6.5. Indeed, we would have $u_* > u_R$ and $\rho_* > \rho_R$. On the other hand, (ρ_*, q_*) would also lie on the 1-Lax curve of (ρ_M, q_M). This state lies on the chord joining (ρ_L, q_L) and

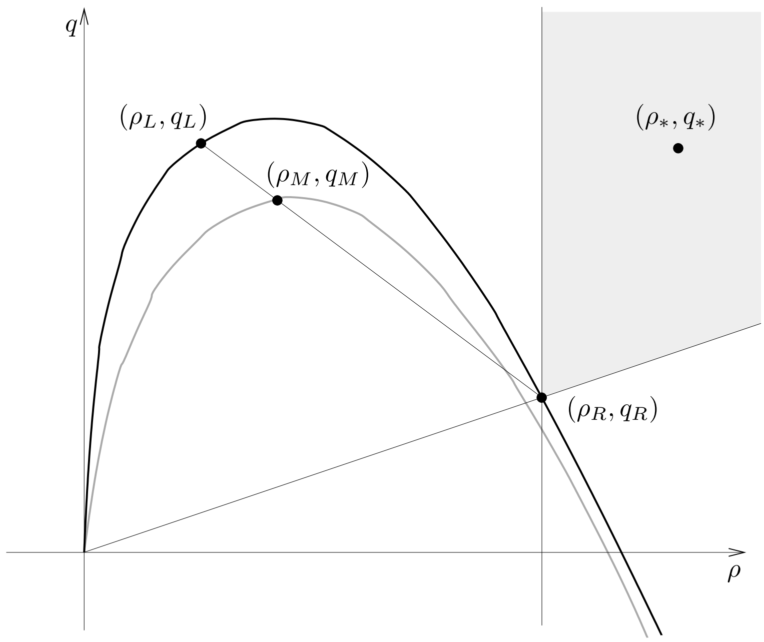


Figure 6.5. *The Riemann problem between cells (ρ_M, q_M) and (ρ_R, q_R) cannot consist of a 1-shock and a 2-shock, because (ρ_*, q_*) should belong at the same time to the gray region and to the gray curve.*

(ρ_R, q_R), and hence below the 1-wave curve of (ρ_M, q_M) (the dark gray curve in the Figure 6.5) is below the 1-Lax curve of (ρ_L, q_L) (the black curve in the Figure 6.5). This last curve does not intersect the light gray area, because it is concave and its slope in ρ_R is smaller than u_R, and we obtain a contradiction. □

6.4.2 Slowly moving shocks

Many finite volume schemes fails to approach correctly shocks that move slowly compared to the speed of sound. Typically, a spike in the momentum appears in the first iterations in time, and is then, by conservation of the momentum, counterbalance by a hollow. The spike oscillates through time in an almost periodic manner, where the period corresponds to the time that the shock needs to cross an entire cell. Therefore, even though the oscillations are diffused by the scheme, it is a constant source of error that is blamed for slow convergence to the steady state (see [Noh87] and [Men94]). Moreover, higher order schemes tend to better preserve the spurious oscillations than first order schemes. This problem was first reported by Colella and Woodward in [WC84]. In [Rob90] and [AR97], numerous numerical tests and comparison between schemes are performed. Unlike all schemes tested in those papers, the half and fully conservative reconstruction schemes are exact not only on steady shocks, but on all shocks (see Proposition 6.5). This can be checked on the top of Figure 6.6, where the spurious oscillations created by the Godunov, Lax–Friedrichs, MUSCL ([vL97]) and Nessyahu–Tadmor ([NT90]) schemes can be seen. These schemes were briefly described in the introduction of the chapter. On the bottom of Figure 6.6, the initial data is a Riemann problem whose solution consists of a 1-slowly moving shock and a 2-shock of smaller strength. The half conservative reconstruction scheme approaches the slowly moving shock correctly, but spurious oscillations appear with the fully conservative reconstruction scheme, as depicted on the bottom of Figure 6.6. Those simulations are done using a mesh with 200 cells and a Courant number of $\lambda = 0.45$. The final time is $T = 0.5$. The initial data on top of Figure 6.6 is a 1-shock moving at speed 0.1: $\rho_L = 2$, $\rho_R = 10$, $u_L = 0.1 + c * \sqrt{\frac{\rho_R}{\rho_L}}$ and $u_R = 0.1 + c * \sqrt{\frac{\rho_L}{\rho_R}}$,

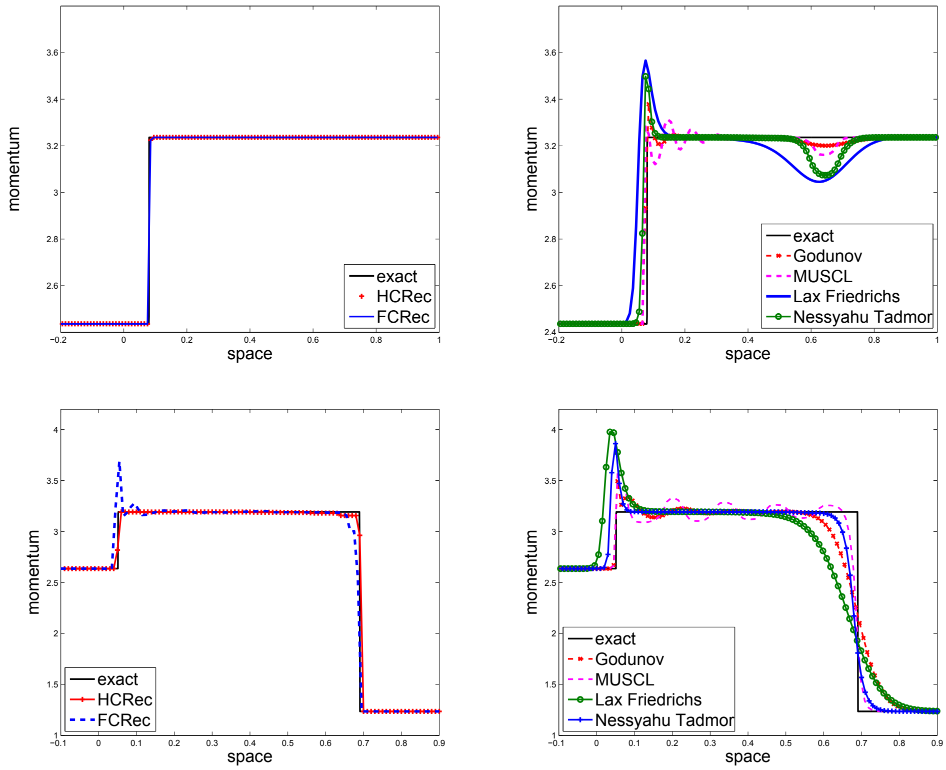


Figure 6.6. *Comparison of the half conservative (HCRec) and the fully conservative (FCRec) reconstruction schemes with classical schemes on a unique slowly moving shock (top) and on a Riemann problem consisting of a slowly moving* 1*-shock followed by a* 2*-shock (bottom).*

which yields $q_L \approx 2.4361$ and $q_R \approx 3.2361$. The speed of sound is $c = 0.5$. The shock crosses a cell in approximately 6.25 iterations. For the second simulation, the momentum is modified: $q_L \approx 2.6361$ and $q_R \approx 1.2361$. When the initial data is a Riemann problem containing only shocks, even if one of them is a slowly moving one, the reconstruction scheme is of order 2. This is illustrated on the left of Figure 6.7. On the right of this figure, we can see that the error is entirely due to the first iterations in time. Once the two shocks are separated, the shocks are better approximated, and no diffusion appears. The initial data is the same than in the bottom of Figure 6.6.

An explanation of the superiority of the half conservative reconstruction scheme over the other schemes (including the fully conservative reconstruction scheme) might be linked to the study of slowly moving shocks by Jin and Liu in [JL96]. In this paper, they prove with a traveling wave analysis that a momentum spike appears in the viscous shock profile for the Euler equations with a linear viscosity

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = \varepsilon \partial_{xx}\rho, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + c^2\rho) = \varepsilon \partial_{xx} q, \end{cases}$$

when the density has a monotonous profile. With the viscosity of Navier–Stokes, the momentum remains monotonous . We believe that the half conservative reconstruction scheme is the only one that has a numerical viscosity looking like the physical Navier–Stokes' viscosity. Indeed, as the reconstruction in the density is performed whenever it is possible (while there is an additional constraint on the momentum in the fully conservative reconstruction), the numerical viscosity is likely to be zero on the mass conservation law as often as possible. Another argument is that if the momentum spike appears

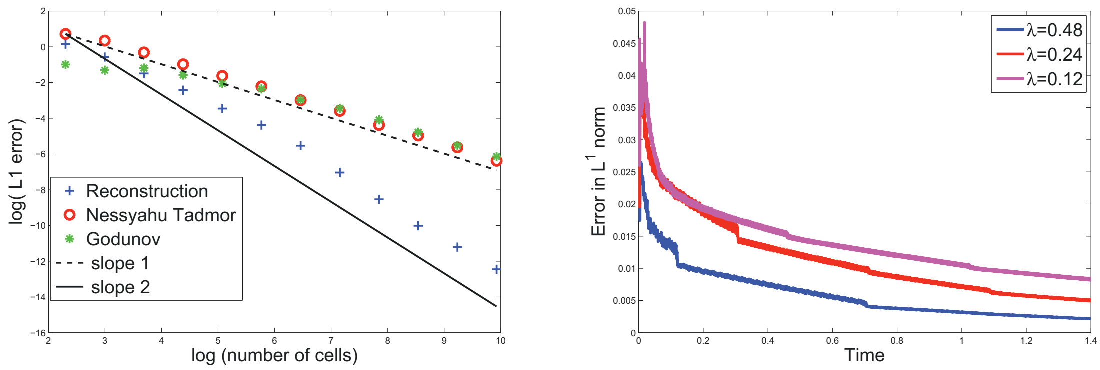


Figure 6.7. *Left: order of the half conservative reconstruction on the Riemann problem only containing shocks. Right: L^1 error through time for different Courant number conditions, for a mesh of* 500 *cells.*

in the fully conservative reconstruction scheme, it blocks the reconstruction ($\mathcal{R}_{FC}$ becomes false) and the scheme will later behave like the Lax–Friedrichs scheme, while the half conservative reconstruction scheme $\mathcal{R}_{HC}$ is more flexible, and continues to reconstruct.

6.4.3 Use of other schemes in the rarefaction waves

When no shocks are detected, or when the reconstruction proposed by $\mathcal{CR}$ is cancelled, the flux degenerates toward the Lax–Friedrichs flux. As a consequence, it is very diffusive inside the rarefaction waves. A simple cure is to the replace, whenever it is chosen, the Lax–Friedrichs flux by a more accurate one. The fluxes write (for the half conservative scheme):

$$f^n_{j+1/2} = \begin{cases} f^{n,REC}_{j+1/2} & \text{if a reconstruction if performed in cell } j+\frac{1}{2}+\frac{1}{2}\,\text{sign}(V^n_{\text{mesh}}) \\ f^{n,NT}_{j+1/2} & \text{otherwise,} \end{cases}$$

where $f^{n,REC}_{j+1/2}$ is the flux for the half conservative reconstruction scheme and $f^{n,NT}_{j+1/2}$ is another flux on a staggered grid. In our implementation we use the simplest version of the Nessyahu–Tadmor scheme [NT90]. Its principle has been briefly described in the introduction. There is no Riemann problems to solve, so it almost does not impact the computation time. As expected, this coupled scheme behaves like the Nessyahu–Tadmor scheme in the rarefaction wave, and like the reconstruction scheme on the shock. This is illustrated by the Figure 6.8 below, where we approximate the solution of the Riemann problem

$$\begin{cases} \rho^0(x) & = 2 * \mathbf{1}_{x<0} + 5 * \mathbf{1}_{x\geq 0} \\ q^0(x) & = 3 * \mathbf{1}_{x<0} + 2 * \mathbf{1}_{x\geq 0} \end{cases}$$

at time $T = 0.1$. The Courant constant is $\lambda = 0.1$, the speed of sound is $c = 3$ and the mesh has 100 cells. On Figure 6.9, we see that we recover the same order of convergence than with the Nessyahu–Tadmor scheme, with lower L^1-error as the discontinuous reconstruction scheme approximates the shock more precisely.

6.5 Handling contacts

We now turn to more prospective results concerning two dimensional simulations and extensions to the Euler equations with energy. In both cases we have to deal with contact discontinuities.

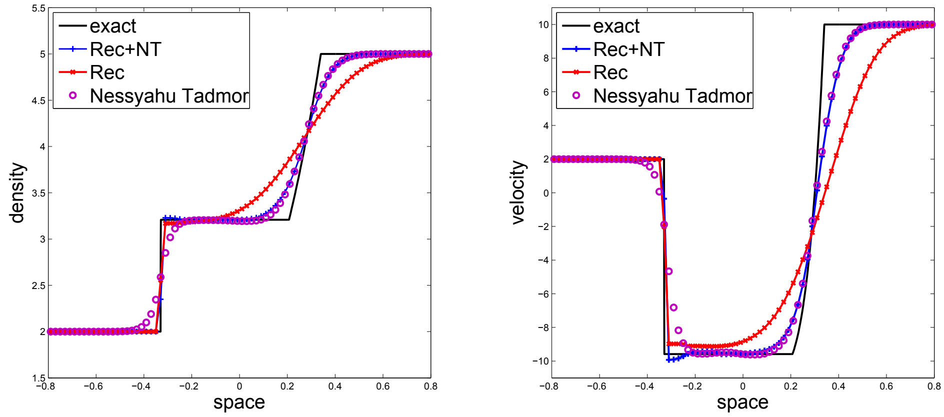


Figure 6.8. *Comparison of the reconstruction scheme (Rec), the Nessyahu–Tadmor scheme and the coupled scheme (Rec+NT) on a Riemann problem consisting of a* 1*-shock and a* 2*-rarefaction wave.*

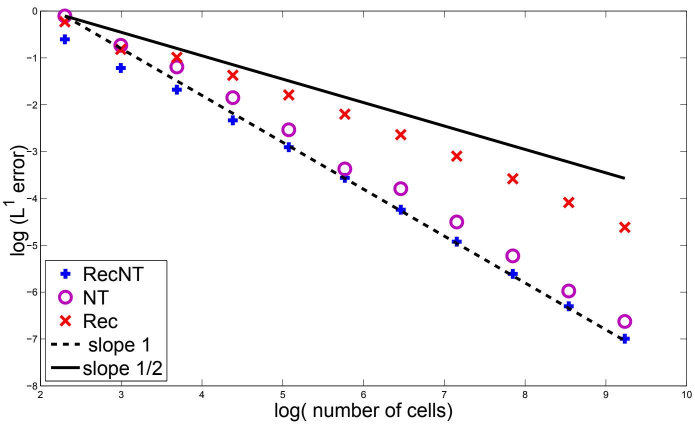


Figure 6.9. *Order of convergence in* L^1*-norm.*

6.5.1 Two dimensional simulations for the barotropic Euler equations

The discontinuous reconstruction scheme described in the previous section extends without any modifications to the barotropic Euler equations when no vacuum appears. In that case the pressure law is $p(\rho) = a\rho^\gamma$, where $a > 0$ and $\gamma > 1$. The case $\gamma = 2$ corresponds to the shallow water system, which describe the height of water h in a canal of small depth. The components of the horizontal velocity of the flow in the x and y direction are denoted by u and v, and g is the gravitational constant. When the canal is flat, the system writes

$$\begin{cases} \partial_t h + \partial_x(hu) + \partial_y(hv) = 0, \\ \partial_t(hu) + \partial_x(hu^2 + gh^2) + \partial_y(huv) = 0, \\ \partial_t(hv) + \partial_x(huv) + \partial_y(hv^2 + gh^2) = 0, \\ (h, u, v)(t = 0, x) = (h^0, u^0, v^0)(x, y). \end{cases} \tag{6.27}$$

We use a cartesian grid and a directional splitting to simulate the solution of (6.27). The solution at time Δt of (6.27) is approximated by the solution at time Δt of

$$\begin{cases} \partial_t h + \partial_y(hv) = 0, \\ \partial_t(hu) + \partial_y(huv) = 0, \\ \partial_t(hv) + \partial_y(hv^2 + gh^2) = 0, \\ (h, u, v)(t = 0, x, y) = (h^{1/2}, u^{1/2}, v^{1/2})(x, y), \end{cases}$$

where $(\rho^{1/2}, u^{1/2}, v^{1/2})$ is the solution at time Δt of

$$\begin{cases} \partial_t h + \partial_x(hu) = 0, \\ \partial_t(hu) + \partial_x(hu^2 + gh^2) = 0, \\ \partial_t(hv) + \partial_x(huv) = 0, \\ (h, u, v)(t = 0, x) = (h^0, u^0, v^0)(x, y). \end{cases} \tag{6.28}$$

Those systems are 1-dimensional, thus we can use the techniques developed in the previous section to approximate numerically their solutions. The two first lines of (6.28) correspond to the 1-dimensional shallow water equations. Thus the solution of the Riemann problems for (h, u) consists in the succession of two simple waves separated by a constant state (h_*, u_*), which are shocks or rarefaction waves. The tangent velocity v is constant through those waves. The third equation corresponds to a linearly degenerate fields. The associated wave is a contact discontinuity has speed u^*.

We would like to propose a reconstruction operator $\mathcal{R}$ which is exact on isolated shock and isolated contact. The difficulty is that h is constant through the contact. Thus if we try to reconstruct along the contact, we have $\bar{h}^n_{j,L} = \bar{h}^n_{j,R}$ and it is almost always impossible to reconstruct the density in a conservative manner. In the sequel we use the reconstruction scheme only for the first two equations and another scheme for the equation in v (here the Nessayahu-Tadmor flux). It would be interesting to find a way to use also the reconstruction strategy in the contact. For example in [DL01], an anti diffusive remapping from Lagrangian to Eulerian coordinates is used to capture contacts sharply.

We now present the numerical results of a 2-dimensional simulation of three columns of water falling into a rectangular basin. We simulate the shallow water equations on the square $[-1, 1] \times [-1, 1]$, with the pressure law $p(h) = h^2$, reflective boundary conditions and initial data

$$\begin{cases} h^0(x, y) = 3 + \mathbf{1}_{(x-0.5)^2+(y-0.5)^2<0.15^2} + \mathbf{1}_{(x+0.5)^2+(y+0.5)^2<0.15^2} + 2 * \mathbf{1}_{x^2+y^2<0.2^2}, \\ u^0(x, y) = 0, \\ v^0(x, y) = 0. \end{cases}$$

We chose this initial data because of the number of details appearing after the water reflects a few time on the boundaries. On Figures 6.10 and 6.11, the Courant number is set to 0.36 and we compare the reconstruction scheme and the Nessyahu–Tadmor scheme. The results are displayed for different times and for meshes with, from left to right, 128×128 cells, 256×256 cells and 512×512 cells. On the top row are the results obtained with the reconstruction scheme coupled with the Nessyahu–Tadmor

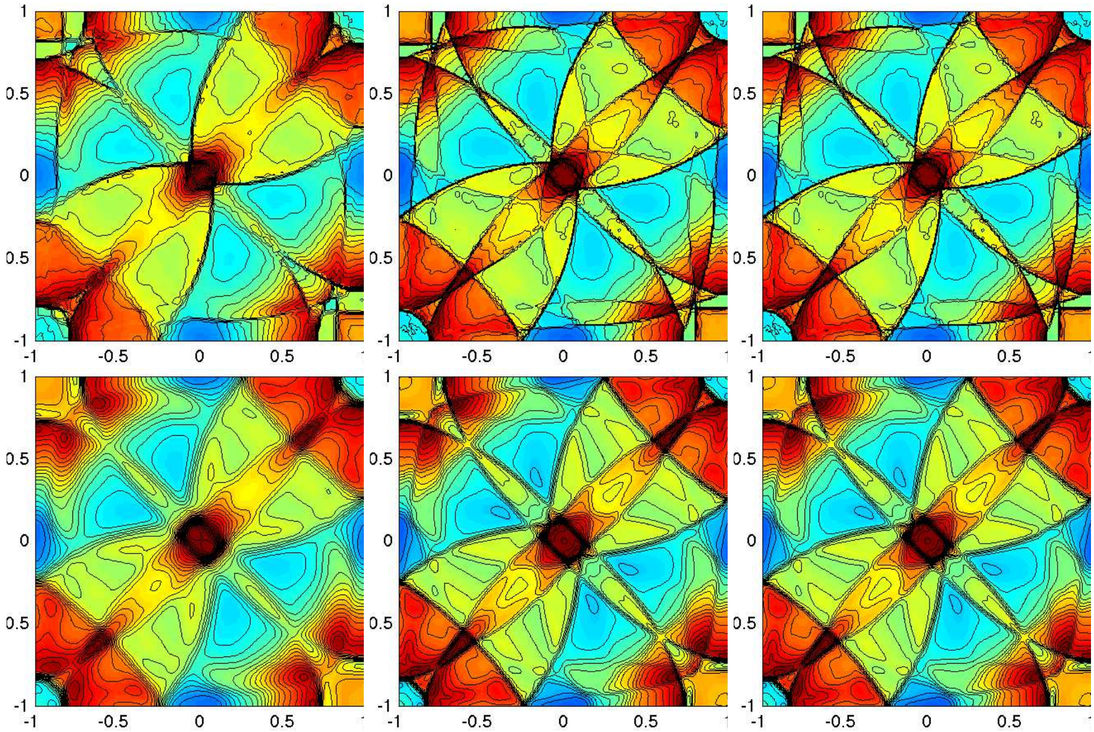


Figure 6.10. *Droplet falling in a square basin: height of water at time* 1.035.

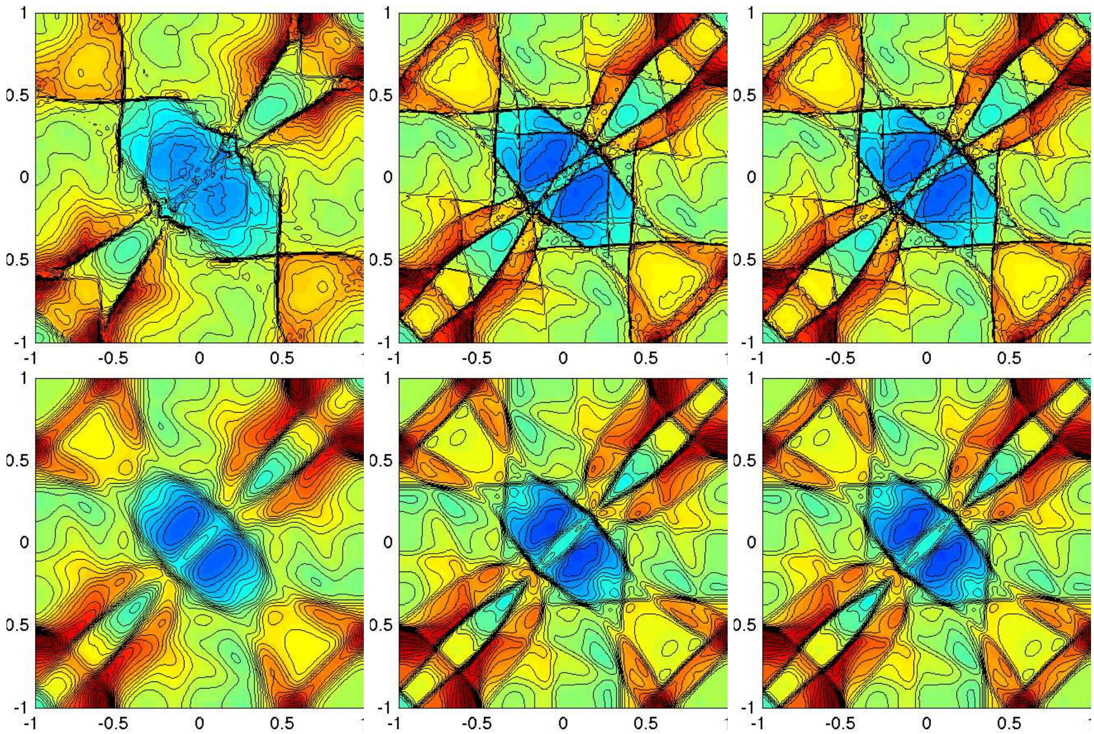


Figure 6.11. *Droplet falling in a square basin: height of water at time* 2.085.

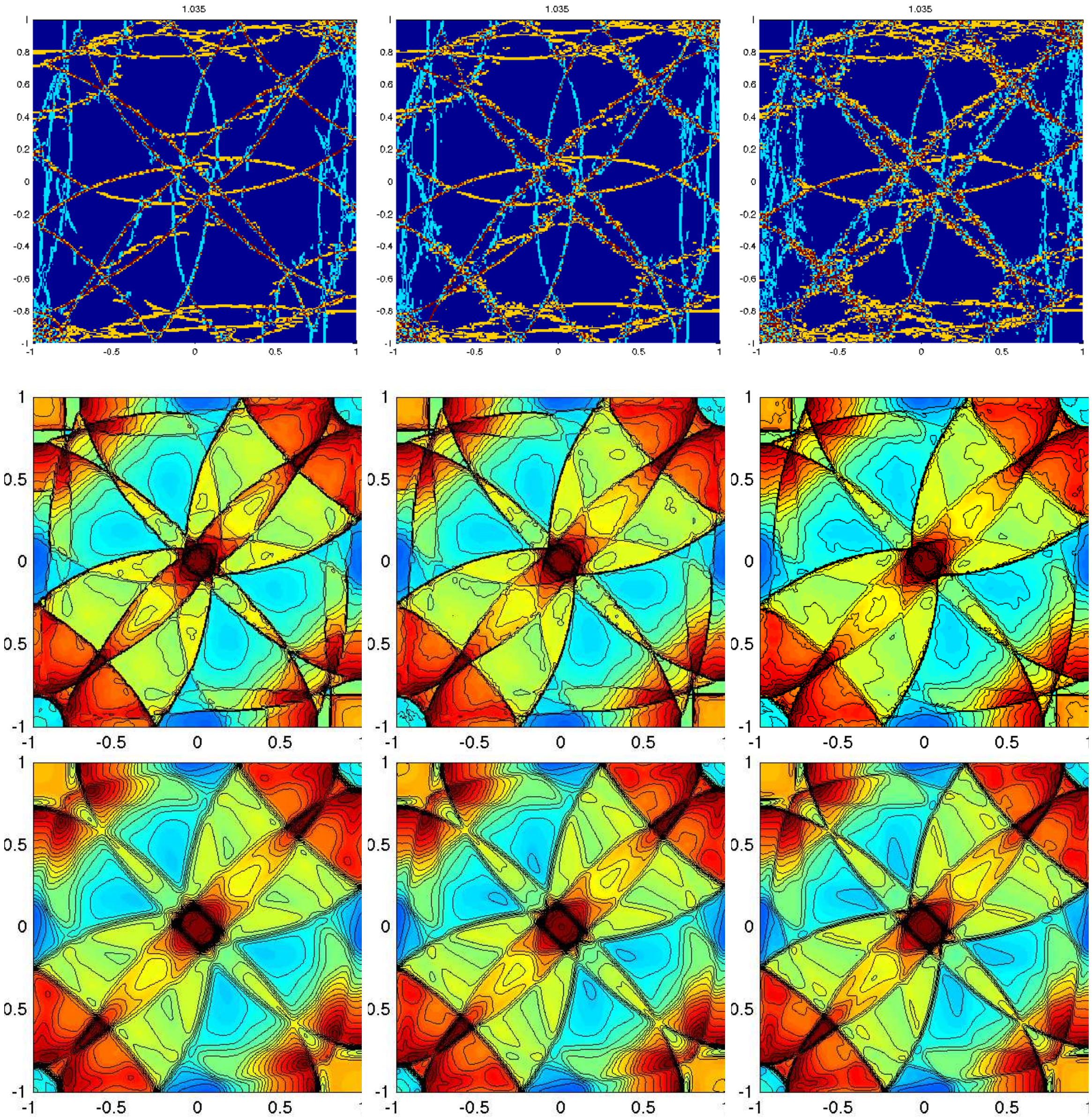


Figure 6.12. *Height of water at time* 1.035 *obtained on a mesh of* 256×256 *cells for different Courant numbers: from left to right,* λ *is* 0.24, 0.36 *and* 0.48. *Second line: with the reconstruction scheme coupled with the Nessyahu–Tadmor scheme, third line: with the Nessyahu–Tadmor scheme. On the first line, one can see in which cells a reconstruction is performed (blue: none, yellow: in the* y*-direction, cyan: in the* x*-direction, red: in both direction).*

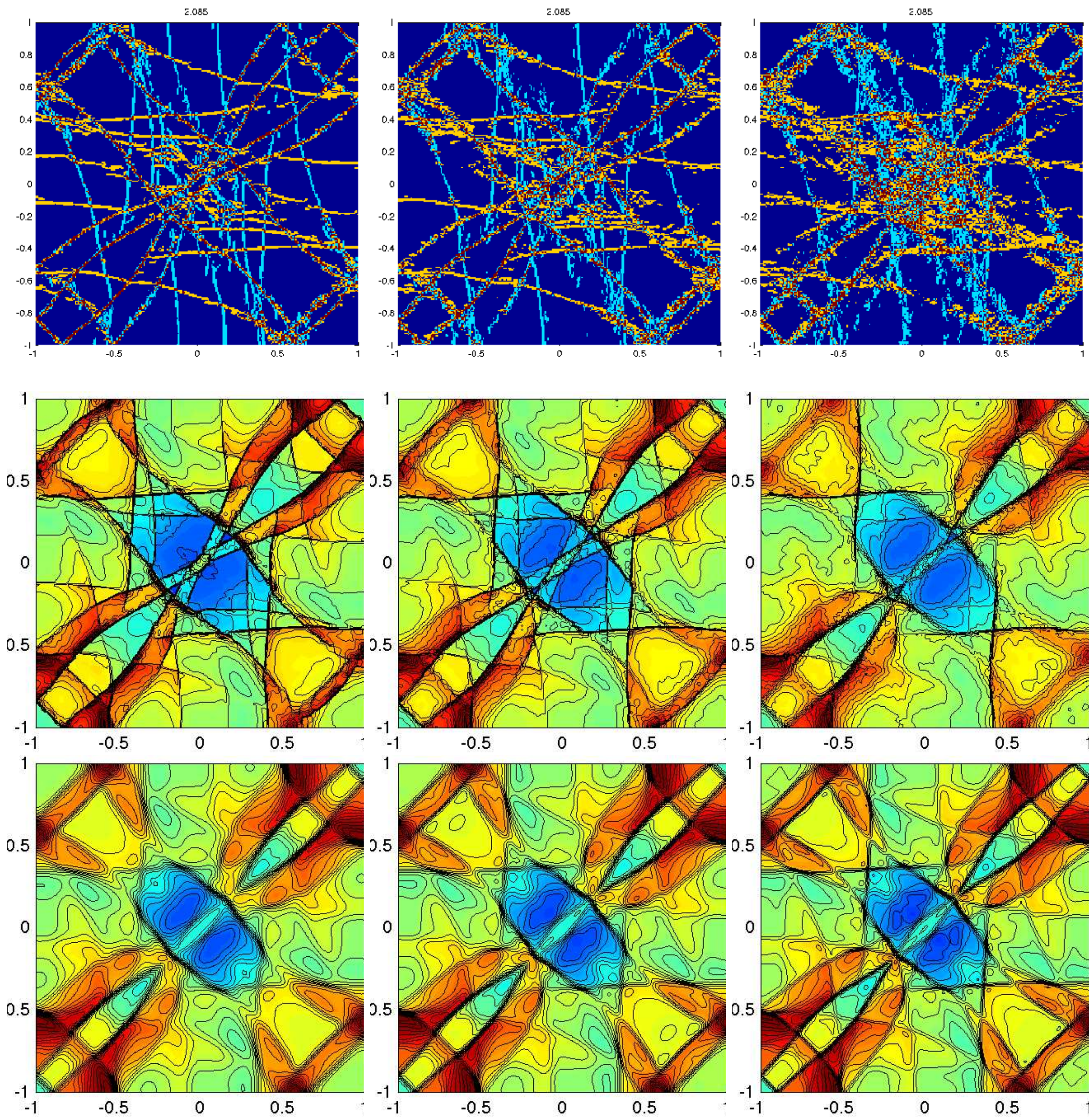


Figure 6.13. *Density at time* 2.085 *obtained on a mesh of* 256×256 *cells for different Courant numbers: from left to right,* λ *is* 0.24, 0.36 *and* 0.48.

	$\lambda = 0.24$				$\lambda = 0.36$				$\lambda = 0.48$			
	RNT	Rec	NT	LF	RNT	Rec	NT	LF	RNT	Rec	NT	LF
128×128	502	394	242	190	470	392	248	190	458	382	253	210
256×256	1485	1047	497	329	1163	814	452	376	1069	849	448	332
512×512	9165	7260	2622	1893	6671	5396	2030	1499	5706	4750	1801	1364

Table 6.1. *CPU times at times 3 (in seconds), for the discontinuous reconstruction (Rec), the Lax–Friedrichs scheme (LF), the Nessyahu–Tadmor scheme (NT) and the discontinuous reconstruction scheme coupled with the Nessyahu–Tadmor scheme (RNT).*

scheme, and on the bottom row those obtained with the Nessyahu–Tadmor scheme. The CPU times are given for different Courant numbers and a final time of 3 in Table 6.1.

On Figure 6.12 and 6.13, the mesh is the same everywhere (256×256) and we compare the behavior of the schemes with respect to the Courant number. We took, from left to right, $\lambda = 0.24$, $\lambda = 0.36$ and $\lambda = 0.48$. The Nessyahu–Tadmor scheme behaves as expected and gives sharper resolution when the Courant number is closer to 0.5. The reconstruction scheme is surprisingly much better at low Courant numbers. A lot of details captured for $\lambda = 0.24$ that are completely missed with $\lambda = 0.48$. For this last Courant number, we can check that the reconstruction scheme is not competitive with the Nessyahu–Tadmor scheme, especially if we take into account the computational times given in Table 6.1.

The phenomenon does not appear in $1D$ or when we compare the "basic" reconstruction scheme (not coupled with the Nessyahu–Tadmor scheme) with the Lax–Friedrichs scheme. A possible explanation is that some reconstruction are performed in zones of contacts, but as we only know how to reconstruct shocks, it gives a more diffusive flux than the Nessyahu–Tadmor scheme. This hypothesis is supported by Figures 6.12 and 6.13, top lines, which show in which cells a reconstruction is performed. It appears clearly that too much reconstructions are performed at high Courant number, whereas reconstructions are sharply concentrated around discontinuities at low Courant number.

6.5.2 Extension to the full Euler equations

Let us now consider the full gas dynamics equation for an ideal gas:

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p) = 0, \\ \partial_t E + \partial_x(u(E+p)) = 0, \end{cases}$$

where ρ is the density, u is the fluid's velocity and E is the total energy per unit volume. We take the pressure law

$$p = p(\rho, e) = e(\gamma - 1)\rho,$$

where $e = \frac{E}{\rho} - \frac{1}{2}u^2$ is the specific internal energy and γ is the ratio of specific heats. The complete solution of the Riemann problem can be found, for example, in [Tor09]. The novelty here is that the 2-wave is a contact discontinuity. The density and the velocity of the fluid are constant through this contact discontinuity. We denote by $U^n_{j,*L/R} = (\rho_{j,*L/R}, u^n_{j,*}, p^n_{j,*})$ the left or right density, velocity and pressure appearing around the contact discontinuity in the Riemann problem between $U^n_{j-1} = (\rho^n_{j-1}, u^n_{j-1}, p^n_{j-1})$ and $U^n_{j+1} = (\rho^n_{j+1}, u^n_{j+1}, p^n_{j+1})$. We recall that λ denotes the Courant number. The candidate reconstruction operator $\mathcal{RC}$

$$\mathcal{RC}(U^n_{j-1}, U^n_j, U^n_{j+1}) = (\bar{U}^n_{j,L}, \bar{U}^n_{j,R})$$

is built as follow.

- If $u^n_{j-1} \geq u^n_{j+1}$, $\rho^n_{j-1} \leq \rho^n_{j+1}$, $p^n_{j-1} \leq p^n_{j+1}$ and if

  $$|\rho^n_{j-1} - \rho^n_{j,*L}| > \lambda \max(|\rho^n_{j,*L} - \rho^n_{j,*R}|, |\rho^n_{j,*R} - \rho^n_{j+1}|),$$

  we set the desired reconstructed states to be

  $$\bar{U}^n_{j,L} = U^n_{j-1} \quad \text{and} \quad \bar{U}^n_{j,R} = U^n_{j,*L}.$$

  In other words, we will try to reconstruct a 1-shock.

- If $u^n_{j-1} \geq u^n_{j+1}$, $\rho^n_{j-1} \geq \rho^n_{j+1}$, $p^n_{j-1} \geq p^n_{j+1}$, and if

  $$|\rho^n_{j,*R} - \rho^n_{j+1}| > \lambda \max(|\rho^n_{j,*L} - \rho^n_{j,*R}|, |\rho^n_{j-1} - \rho^n_{j,*L}|),$$

  we set the desired reconstructed states to be

  $$\bar{U}^n_{j,L} = U^n_{j,*R} \quad \text{and} \quad \bar{U}^n_{j,R} = U^n_{j+1}.$$

  In other words, we will try to reconstruct a 3-shock.

- If

  $$|\rho^n_{j,*L} - \rho^n_{j,*R}| > \lambda \max(|\rho^n_{j-1} - \rho^n_{j,*L}|, |\rho^n_{j,*R} - \rho^n_{j+1}|),$$

  we set

  $$\bar{U}^n_{j,L} = U^n_{j,*L} \quad \text{and} \quad \bar{U}^n_{j,R} = U^n_{j,*R}.$$

  and try to reconstruct as a 2-contact discontinuity.

- Otherwise, we do not perform any reconstruction and set

  $$\bar{U}^n_{j,L} = U^n_j \quad \text{and} \quad \bar{U}^n_{j,R} = U^n_j.$$

The reconstruction operator $\mathcal{R}(U^n_{j-1}, U^n_j, U^n_{j+1})$ is equal to $\mathcal{RC}(U^n_{j-1}, U^n_j, U^n_{j+1})$, if and only if the following conditions are all fulfilled:

- Both d_j^ρ and d_j^E are between 0 and Δx;
- Both triplets $(\rho_{j-1}^n, \rho_j^n, \rho_{j+1}^n)$ and $(u_{j-1}^n, \frac{1}{\Delta x}\int_{x_{j-1/2}^n}^{x_{j+1/2}^n} u_{\mathrm{rec}}^n(x)dx, u_{j+1}^n)$ are monotonous;
- Eventually, e_{rec}^n remains positive on the cell $[x_{j-1/2}^n, x_{j+1/2}^n]$.

Otherwise, we set $\mathcal{R}(U_{j-1}^n, U_j^n, U_{j+1}^n) = (U_j^n, U_j^n)$. Here, u_{rec}^n and e_{rec}^n are the piecewise constant reconstructed velocity and internal energy at time t^n. The reconstruction scheme appears to be much more complicated in this case. This is due to the lack of criterion to detect a dominant contact discontinuity. Indeed, Proposition 6.4 (even with an additional test on the pressure) is still valid for this system as p and u remains constant through contact discontinuities. But it is not entirely satisfactory to decide which wave prevails in the Riemann problem. Indeed, a lot of Riemann problems with a 1-shock and a 2-contact discontinuity verify the inequalities

$$u_L^n \geq u_R^n,\ \rho_L^n \leq \rho_R^n \quad \text{and} \quad p_L^n \geq p_R^n$$

even though the shock is small and the contact is strong. This is why we added additional constraints like

$$|\rho_{j-1} - \rho_j^{*,L}| > \lambda \max(|\rho_j^{*,L} - \rho_j^{*,R}|, |\rho_j^{*,R} - \rho_{j+1}|),$$

which is a criterion on the force of the wave. Note that with this criterion, some shocks may not be detected, while without it, some contact discontinuities will be detected as shocks. We observe numerically that using the Courant number tunes the scheme pretty well. This very same difficulty to detect contact discontinuities led us to add more constraints to accept a reconstruction. The first one is similar to the one for the isothermal case. We do not impose a conservation constraint on the momentum to keep the good behavior near slowly moving shocks, see Section 6.4.2 and Test 4 below. The density and velocity are monotonous along the profile of a viscous shock, which justifies the second constraint. Eventually, the last constraint forces the pressure to remain positive.

We perform various numerical tests to compare the reconstruction scheme with other classical schemes. On all these figures, `Rec` indicates the half conservative reconstruction scheme and `Rec+NT` indicates the half conservation reconstruction scheme coupled with the Nessyahu–Tadmor scheme. These schemes are compared with the Godunov and Rusanov schemes, and with the simplest version of the Nessyahu–Tadmor scheme [NT90] and the simplest version of the MUSCL scheme [vL97]

Test 1: A Riemann problem with three discontinuities

On Figure 6.14, we compare numerous schemes on a Riemann problem with

$$\begin{cases} \rho_L = 5.99924\, u_L = 19.5975 \quad \text{and} \quad p_L = 460.894, \\ \rho_R = 5.99242\, u_R = -6.19633 \quad \text{and} \quad p_R = 46.0950. \end{cases}$$

The solution consists of three discontinuities moving to the right (cf the book [Tor09]). We took a Courant number of 0.4 and discretized the interval in 400 cells. The shocks are very well computed, an we observe an improvement on the contact discontinuity. This improvement is often much better when the reconstruction scheme is coupled with a higher order scheme.

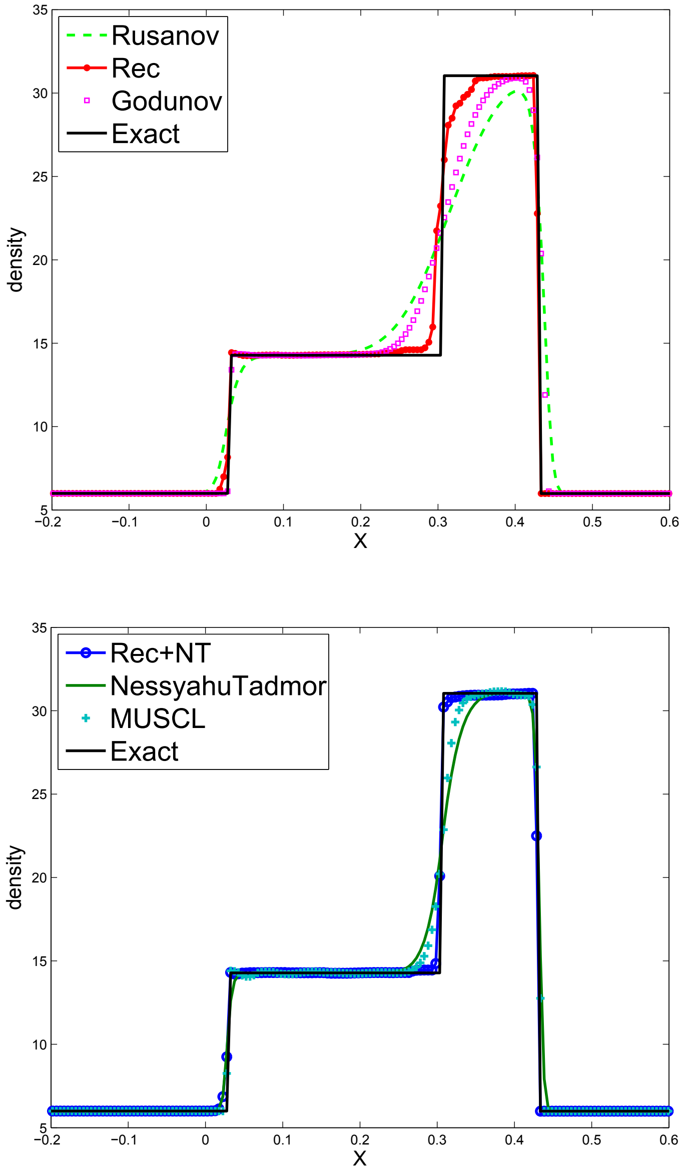


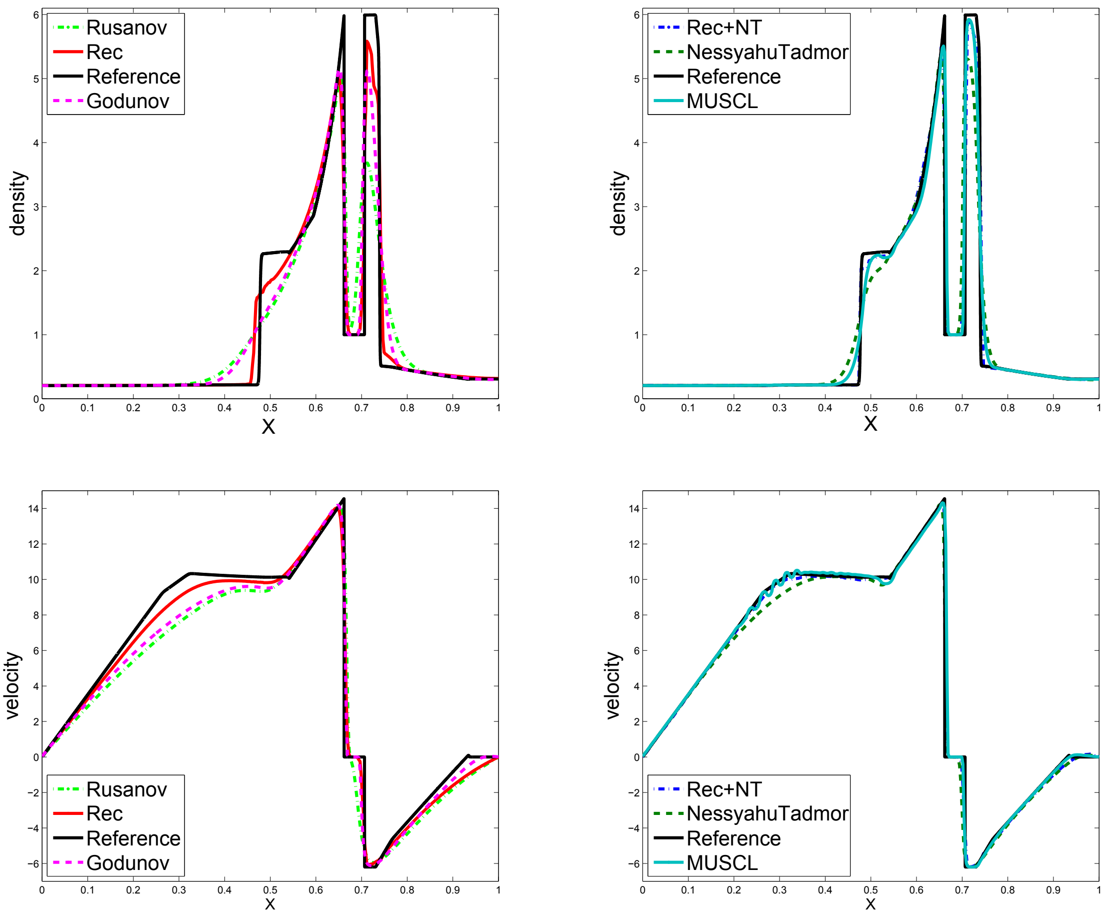
Figure 6.14. *Test* 1. *Density at time* 0.035.

Test 2: Blast Wave

We compare those same schemes in the Colella and Woodward blast wave test case, introduced in [WC84]. The initial data is

$$\begin{cases} \rho^0(x) = 1, \\ u^0(x) = 0, \\ p^0(x) = 1\,000 * \mathbf{1}_{x<0.1} + 0.01 * \mathbf{1}_{0.1<x<0.9} + 100 * \mathbf{1}_{x\geq 0.9}\,. \end{cases}$$

The solution is computed on the interval $[0, 1]$, with reflective boundary conditions at the two extremities of the interval. The reference solution is obtained by with the Nessyahu–Tadmor scheme and $30\,000$ cells. A more accurate reference solution can be found in [WC84]. On Figure 6.15, we used a mesh containing 400 cells, a Courant number set to 0.45 and the final time is $T = 0.026$. We plot the density, velocity and internal energy. The nondissipative feature of the reconstruction scheme is particularly obvious on this last plot. On Figure 6.16, the final time is $T = 0.038$. We took a Courant number of 0.48 and 400, $1\,200$ and $2\,000$ cells. Observe that even with few points, the discontinuities are sharply captured, even though no reconstruction is performed in between, where we recover the behavior of the Lax–Friedrichs or the Nessyahu–Tadmor schemes.



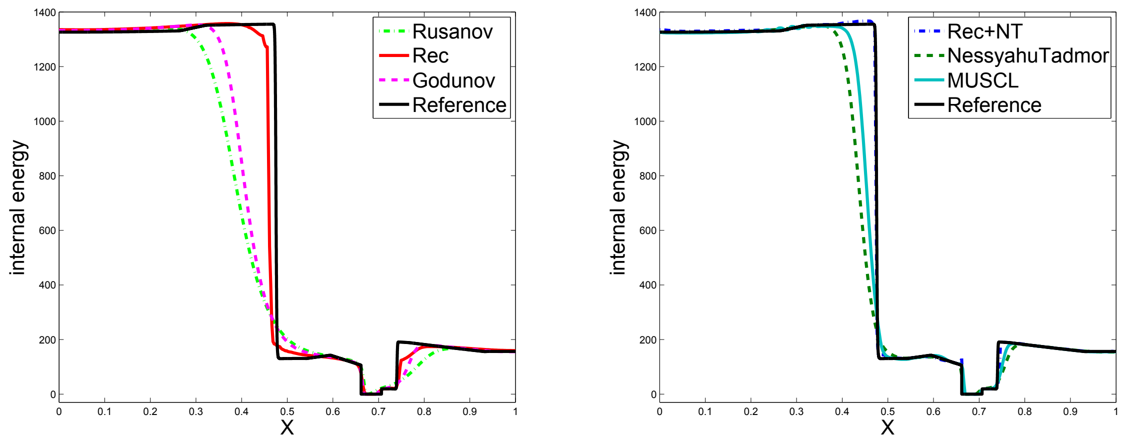


Figure 6.15. *Test* 2*. From top to bottom, density, velocity and internal energy at time* 0.026 *in the blast wave test case, with a mesh of* 400 *cells.*

Test 3**: Interaction of a shock with a sine wave**

The Figure 6.17 presents the results at time 1.8 of an entropy satisfying shock interacting with a sine wave, which mimics turbulence. It is presented in more details in [LW03]. The space interval $[-5,5]$ is discretized with 400 cells, the Courant number is 0.45 and the initial data is:

$$\begin{cases} \rho^0(x) = 3.897143 * \mathbf{1}_{x<-4} + (1+0.2\sin(5x))\mathbf{1}_{x\geq -4}, \\ u^0(x) = 2.629369 * \mathbf{1}_{x<-4}, \\ p^0(x) = 10.33333 * \mathbf{1}_{x<-4} + \mathbf{1}_{x\geq -4}. \end{cases}$$

The reference solution is, once again, the result given by the Nessyahu–Tadmor scheme with $30\,000$ cells and a Courant number of 0.48. The frequency of the oscillations are sharply captured, even though their amplitude is underestimated by the Lax–Friedrichs scheme.

Test 4**: A Riemann problem with a slowly moving shock**

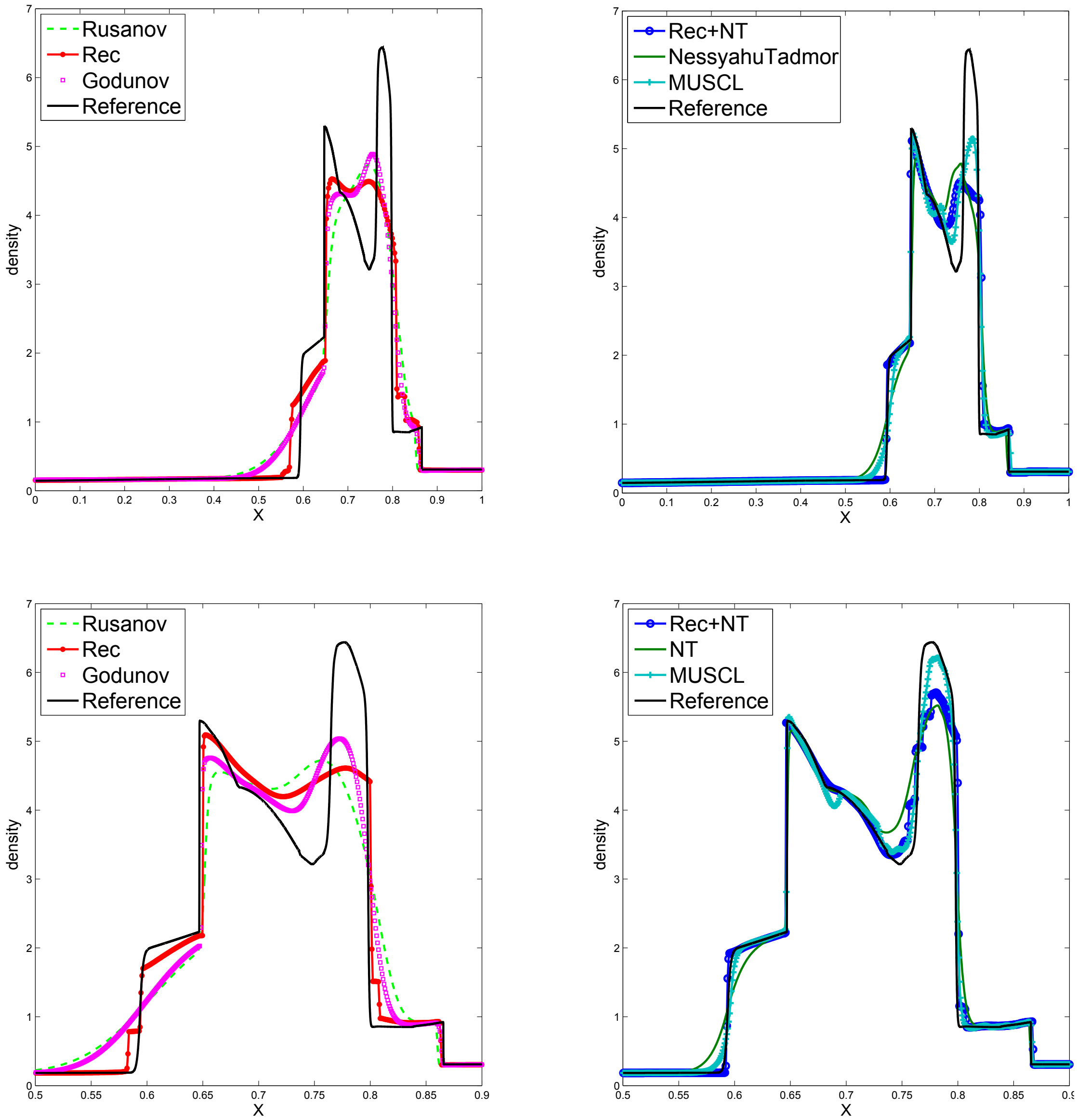
On Figure 6.18, the test case consists in a Riemann problem containing a slowly moving 3-shock. We plot the momentum for a mesh of 800 cells and a Courant number of 0.3. More precisely, we have

$$\rho_L = 3.86,\ u_L = -0.81,\ p_L = 10.33 \quad \text{and } \rho_R = 1.02,\ u_R = -3.44,\ p_R = 1.02.$$

The solution contains three shock, and the 3-shock is slowly moving. This test case is a slight perturbation of the pure slowly moving shock introduced by [Qui94], in which $\rho_R = p_R = 1$, in which case the 3-shock is isolated. The aim is to test if the good behavior on shocks predicted by Theorem 6.3 stands under perturbations. It does: no spurious oscillations appears.

Test 5**: Wall heating phenomenon**

Another nice feature of the reconstruction scheme is that is seems to drastically diminish the wall heating phenomenon reported in [Noh87]. It occurs when a shock reflects on a solid wall, and takes the form of a hollow in a density, or a spike in the temperature, near the wall. We test our scheme on two cases considered by Donat and Marquinat in [DM96], who proposed a cure by interlacing two



Rusanov
Rec
Godunov
Reference
density
X
Rec+NT
NessyahuTadmor
MUSCL
Reference
density
X
Rusanov
Rec
Godunov
Reference
density
X
Rec+NT
NT
MUSCL
Reference
density
X

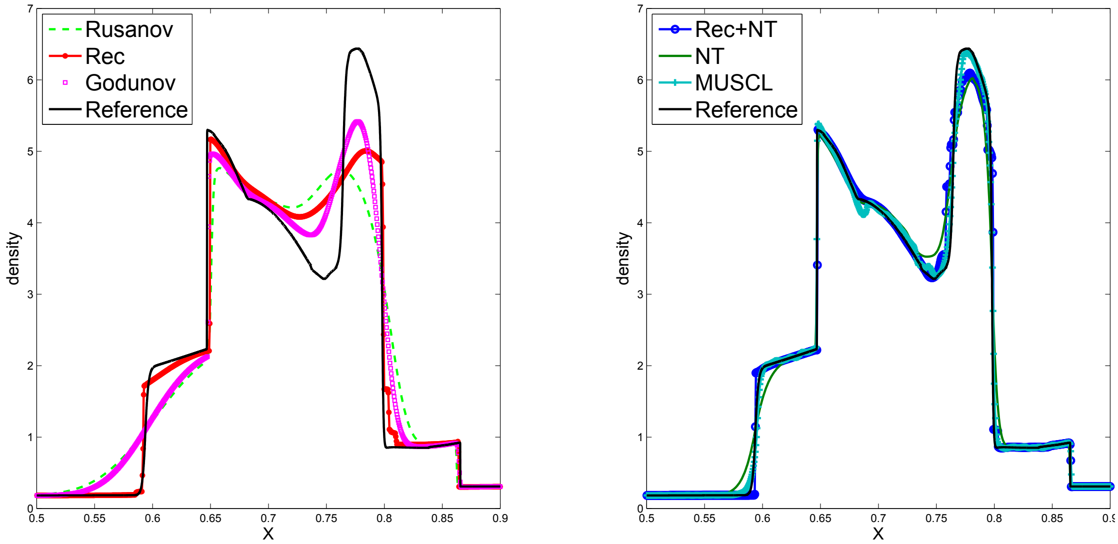


Figure 6.16. *Test* 2*. Density at time* 0.038 *in the blast wave test case, with a mesh of* 400 *cells (top),* 1 200 *cells (middle) and* 2 000 *cells (bottom).*

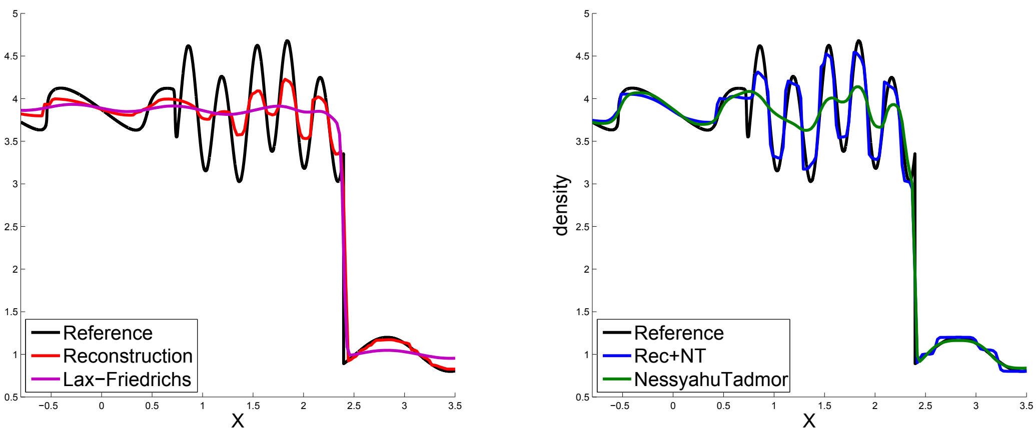


Figure 6.17. *Test* 3*. Density at time* 1.8 *for the shock entropy wave interaction (zoom around the high frequency oscillations).*

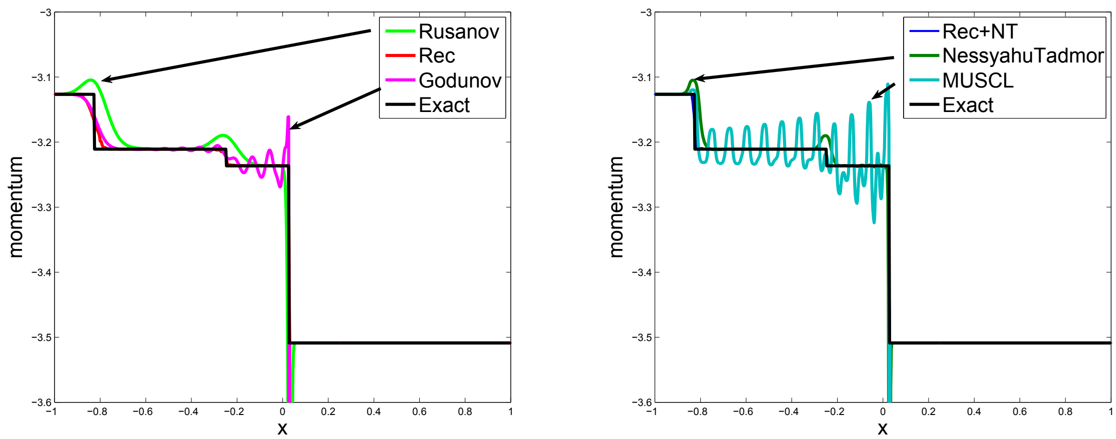


Figure 6.18. *Test* 4*. Momentum at time* 0.3 *for a Riemann problem with a slowly moving shock. The numerical solution given by the reconstruction schemes are very close to the exact solution.*

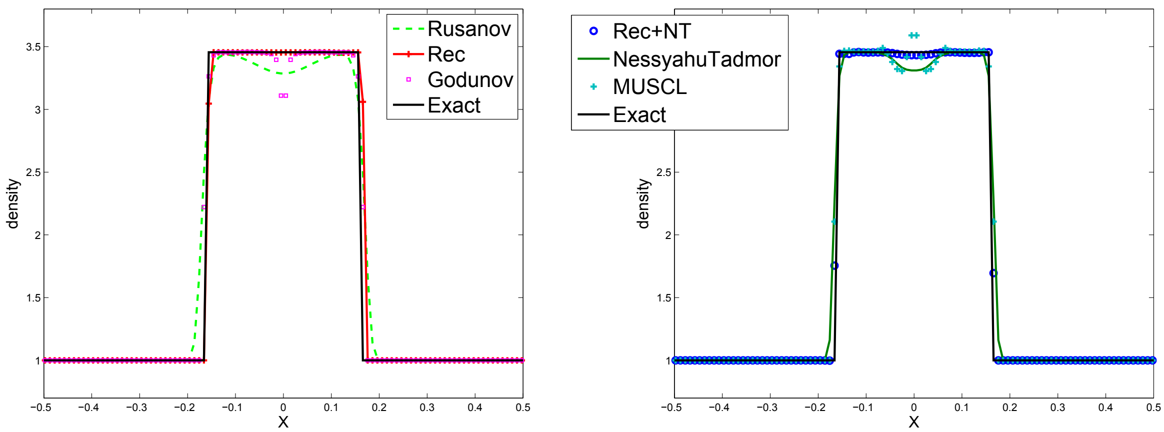


Figure 6.19. *Test* 5*. Density time* 0.1 *for a Riemann problem with two symmetric shocks.*

schemes. For those two tests only, $\gamma = 5/3$. The first case is a Riemann problem developing two symmetric shocks. The initial data is

$$\rho_L = 1,\ u_L = 4,\ p_L = 1 \quad \text{and}\ \rho_R = 1,\ u_R = -4,\ p_R = 1\,.$$

The mesh has 200 cells and the Courant number of 0.4. The results, shown on Figure 6.19, show that the wall heating phenomenon is drastically diminished with the reconstruction schemes. The second test is the reflection of a gas of density 1, pressure 0.001 and velocity 1 on a solid wall on its right. On Figure 6.20 is a zoom around the wall, and we can clearly see the wall heating phenomenon and the resulting spurious oscillations arising with the Godunov and MUSCL schemes, and the good behavior of the reconstruction schemes. We took a Courant number of 0.45 and 1 000 cells.

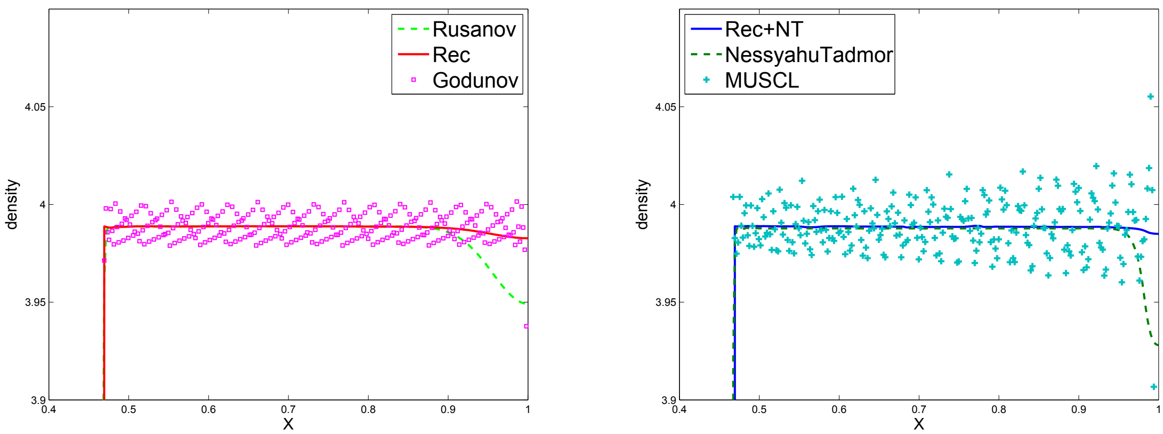


Figure 6.20. *Test* 5*. Density at time* 1.6 *of a fluid reflecting on a solid wall.*

6.6 Use of an approximate Riemann solver: the example of the p-system

So far, the reconstruction operators $\mathcal{R}$ we presented used the exact solution of the Riemann problem. We would like to avoid the exact resolution of Riemann problem for two reasons. First, it is numerically costly. Second, the explicit solution is not known for every hyperbolic system. Thus in this section we present an operator $\mathcal{R}$ that is based on a approximate Riemann solver. The key point is that according to Definition 6.1, $\mathcal{R}$ must be exact on discontinuity of interest. It yields that the approximate Riemann solver must be exact on those shocks as well. Unfortunately, this is not the case of the approximate Riemann solvers developed so far.

In the sequel, we focus on the barotropic gas dynamics equations in Lagrangian coordinates :

$$\begin{cases} \partial_t \tau - \partial_x u = 0, \\ \partial_t u + \partial_x p(\tau) = 0, \\ \tau(0,x) = \tau^0(x),\ u(0,x) = u^0(x), \end{cases} \tag{6.29}$$

where the pressure law p is a convex decreasing function. This system is well-known to be strictly hyperbolic, with eigenvalues $\pm\sqrt{-p'(\tau)}$ and two genuinely nonlinear characteristic fields. The phase space Ω of this system is

$$\Omega = \{(\tau,u) \in \mathbb{R}^2, \tau > 0\ \}.$$

As usual, we supplement (6.29) with the validity of the so-called entropy inequality

$$\partial_t \mathcal{U}(\tau,u) + \partial_x \mathcal{W}(\tau,u) \leq 0, \tag{6.30}$$

where $(\mathcal{U},\mathcal{W})$ is the entropy-entropy flux pair given by

$$\mathcal{U}(\tau,u) = \frac{u^2}{2} - \int^{\tau} p(y)\,dy \quad \text{and} \quad \mathcal{W}(\tau,u) = p(\tau)u,$$

and we consider entropy weak solutions of (6.29)-(6.30), see [GR96] for example. We focus particularly on the case where the initial data is a Riemann initial data

$$\begin{cases} \tau^0(x) = \tau_L \mathbf{1}_{x<0} + \tau_R \mathbf{1}_{x>0}, \\ u^0(x) = u_L \mathbf{1}_{x<0} + u_R \mathbf{1}_{x>0}, \end{cases} \tag{6.31}$$

where the left and right states (τ_L,u_L) and (τ_R,u_R) verify the Rankine–Hugoniot relations (see [GR96] for example)

$$u_L - u_R = -s(\tau_L - \tau_R) \quad \text{and} \quad p(\tau_L) - p(\tau_R) = s(u_L - u_R), \tag{6.32}$$

for a given speed of propagation s, and the triple (τ_L,u_L), (τ_R,u_R), s satisfies (6.30) in the weak sense, i.e. that

$$-s\,(\mathcal{U}(\tau_R,u_R) - \mathcal{U}(\tau_L,u_L)) + (\mathcal{W}(\tau_R,u_R) - \mathcal{W}(\tau_L,u_L)) \leq 0.$$

In that case the initial condition under consideration is such that the initial discontinuity will move at velocity s to form an admissible entropy weak solution of (6.29)-(6.30), see [GR96]. We say that we have an isolated shock wave.

Let us illustrate the fact that usual approximate Riemann solvers (abbreviated ARS in the sequel) are not exact for isolated shock waves. We consider here the Suliciu relaxation procedure [Sul98] (see also [BdL09]), which is also the starting point of the solver proposed in [CC13] and used in the present section. The idea is to approach the solutions of (6.29) by the solutions of a larger but simpler system, namely

$$\begin{cases} \partial_t \tau - \partial_x u = 0, \\ \partial_t u + \partial_x \pi = 0, \\ \partial_t \mathcal{T} = \frac{1}{\varepsilon}(\tau - \mathcal{T}), \end{cases} \tag{6.33}$$

where π is a new unknown related to the pressure p via the expansion

$$\pi = \pi(\mathcal{T}) = p(\mathcal{T}) + a^2(\mathcal{T} - \tau).$$

When ε tends to 0, we recover asymptotically the p-system we are interested in, and for stability reason a must be chosen larger than the maximum speed of the wave appearing in the initial system (6.29), see the next section and [Bou04b], [CLL94] and [JX95] for more details. This is the so-called sucharacteristic stability condition.

System (6.33) is actually simpler to solve than the original p-system (6.29) since the Riemann problem associated to (6.33), namely

$$\begin{cases} \partial_t \tau - \partial_x u = 0, \\ \partial_t u + \partial_x \pi = 0, \\ \partial_t \mathcal{T} = 0, \\ (\tau, u, \mathcal{T})(t=0,x) = (\tau_L, u_L, \mathcal{T}_L)\mathbf{1}_{x<0} + (\tau_R, u_R, \mathcal{T}_R)\mathbf{1}_{x\geq 0}, \\ \mathcal{T}_L = \tau_L \text{ and } \mathcal{T}_R = \tau_R, \end{cases} \tag{6.34}$$

where the initial data is taken at equilibrium (i.e. $\mathcal{T} = \tau$), can be explicitly solved. The characteristic fields are easily shown to be linearly degenerate, and the solution of (6.34) contains three contact discontinuities, propagating with velocities $-a$, 0, and a. We denote by $(\tau_{L,*}, u_{L,*}, \mathcal{T}_{L,*})$ the state on the left of the stationary wave, and by $(\tau_{R,*}, u_{R,*}, \mathcal{T}_{R,*})$ the state on its right (see Figure 6.21). Similarly we denote by $p_L := \pi(\mathcal{T}_L) = p(\tau_L)$ and $p_R := \pi(\mathcal{T}_R) = p(\tau_R)$.

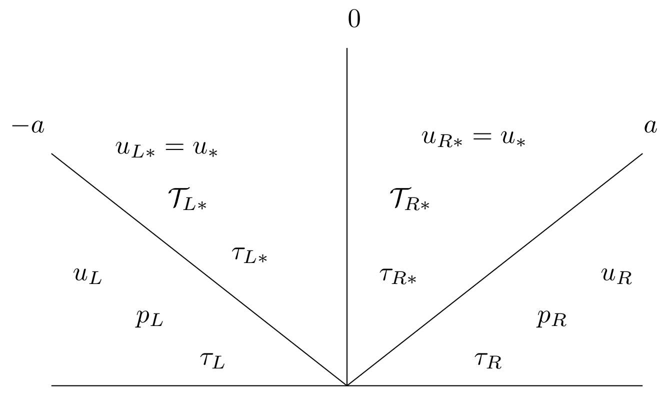


Figure 6.21. *The approximate Riemann solver of Suliciu.*

The intermediate states can be obtained using the Rankine–Hugoniot relations across each contact discontinuities, and are given by $\mathcal{T}_{L*} = \mathcal{T}_L$, $\mathcal{T}_{R*} = \mathcal{T}_R$, and

$$\begin{cases} u_{L*} = u_{R*} := u_* = \frac{u_L + u_R}{2} - \frac{p_R - p_L}{2a}, \\ \pi_{L*} = \pi_{R*} := \pi_* = \frac{p_L + p_R}{2} - \frac{a}{2}(u_R - u_L), \\ \tau_{L*} = \tau_L + \frac{u_* - u_L}{2a}, \\ \tau_{R*} = \tau_R - \frac{u_* - u_R}{2a}. \end{cases} \tag{6.35}$$

We refer once again to [GR96] for details. Importantly, since a is chosen larger than any speed of propagation of (6.29) for stability reasons, there is no hope to capture exactly isolated shocks with any Godunov-type method using this approximate Riemann solver. The idea of [CC13] was to modify the approximate Riemann solver given by (6.34) by introducing a new wave propagating with a velocity

σ in order for the new ARS to be exact on isolated shock waves. The new wave pattern is depicted on Figure 6.22. Attached to this new wave is a parameter θ, which makes the link between the case of an isolated shock ($\theta = 1$, the other three waves are trivial) and the classical solver ($\theta = 0$, the new wave is trivial and the solver coincides with (6.35)). Moreover, this parameter can be chosen in such a way that the new ARS is entropy satisfying and exact on isolated shocks. The parameter θ can be seen as a detection parameter for shocks. Thus the shock detection step of Lemma 6.4 is already enclosed in the ARS of [CC13].

In the next section, we describe the approximate of [CC13]. Then we propose a reconstruction operator based on this ARS, and we prove that the corresponding discontinuous reconstruction scheme is exact when the initial data is an isolated shock. The last sections are devoted to numerical simulations.

6.6.1 The approximate Riemann solver

In this section we briefly describe the approximate Riemann solver introduced in [CC13] to solve (6.29)-(6.30)-(6.31) and refer to this paper for details. The proposed approximate solution is the exact solution of the following system with Riemann initial data at equilibrium:

$$\begin{cases} \partial_t \tau - \partial_x u = 0, \\ \partial_t u + \partial_x \pi(\mathcal{T}) = 0, \\ \partial_t \mathcal{T} = \mathcal{M}(\theta)\delta_{x=\sigma t}, \\ (\tau, u, \mathcal{T})(t=0,x) = (\tau_L, u_L, \mathcal{T}_L)\mathbf{1}_{x<0} + (\tau_R, u_R, \mathcal{T}_R)\mathbf{1}_{x>0}, \\ \mathcal{T}_L = \tau_L \text{ and } \mathcal{T}_R = \tau_R. \end{cases} \tag{6.36}$$

System (6.36) is a modified version of (6.33), where the measure valued right hand side $\mathcal{M}(\theta)\delta_{x=\sigma t}$ allows the solution to jump along the line $x = \sigma t$. The parameter σ is chosen such that if the initial datum is an isolated shock, σ is the exact speed of this shock. The Rankine–Hugoniot relations (6.32) yield

$$s^2 = -\frac{p_L - p_R}{\tau_L - \tau_R}.$$

In [CC13] the authors then propose to set

$$\sigma = \operatorname{sign}(\tau_R - \tau_L)\sqrt{-\frac{p_L - p_R}{\tau_L - \tau_R}}, \tag{6.37}$$

the role of the sign function being to distinguish in between entropy shocks of the first (resp. the second) eigenvalue $-\sqrt{-p'}$ (resp. $\sqrt{-p'}$), for which the shock has a negative velocity and $\tau_R < \tau_L$ (resp. a positive velocity and $\tau_R > \tau_L$). The parameter θ will be chosen so that the approximate Riemann solver is entropy satisfying. The solution of System (6.36) has four waves: the three usual waves having speeds $-a$, 0 and a, plus a wave at speed σ driven by the source term. According to [CC13], this σ-wave has to be understood as an approximation of the shock wave with the largest amplitude in the exact Riemann solution of (6.29)-(6.30)-(6.31). This approximation turns out to be exact in the case of an isolated shock wave. The structure of the new ARS is depicted on Figure 6.22.

To define the three intermediate states (and hence the nine unknowns), consistency relations and Rankine–Hugoniot relations are imposed by the authors in [CC13]. We note in particular that u and τ are constant through the stationary wave while the quantity denoted by

$$\mathcal{I} = \pi(\mathcal{T}) + a^2\tau = p(\mathcal{T}) + a^2\mathcal{T} \tag{6.38}$$

is constant through the $-a$- and a-waves. We denote with a minus subscript (resp. a plus subscript) their values on the left (resp. on the right) of the σ-wave. The intermediate states are found by solving

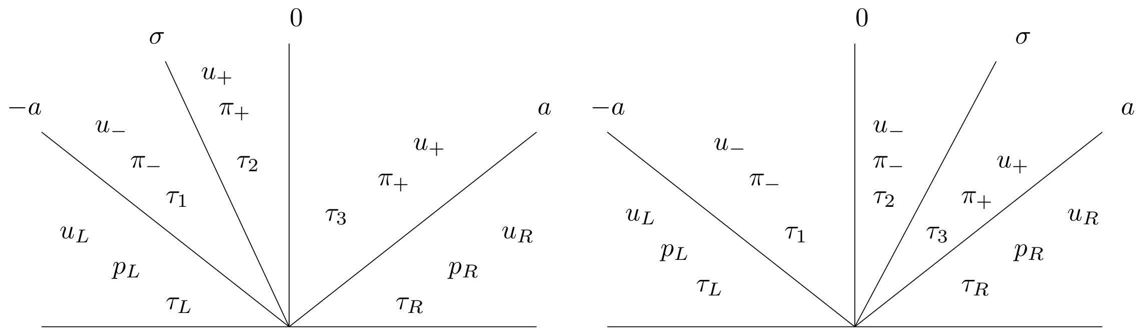


Figure 6.22. *Structure of the Riemann solution depending on the sign of σ.*

a *linear* system (in τ, u and $\mathcal{I}$) and are given , when $\sigma > 0$, by

$$\begin{cases} u_- &= u_* + \frac{\sigma\theta}{2a}(a-\sigma)(\tau_R - \tau_L), \\ u_+ &= u_* - \frac{\sigma\theta}{2a}(a+\sigma)(\tau_R - \tau_L), \\ \tau_1 &= \tau_L + \frac{1}{a}(u_- - u_L), \\ \tau_3 &= \tau_R + \frac{1}{a}(u_R - u_+), \\ \tau_2 &= \tau_3 - \theta(\tau_R - \tau_L), \end{cases} \quad \text{and} \quad \begin{cases} \pi_- &= \pi_* - \frac{\sigma\theta}{2}(a-\sigma)(\tau_R - \tau_L), \\ \pi_+ &= \pi_* - \frac{\sigma\theta}{2}(a+\sigma)(\tau_R - \tau_L), \\ \mathcal{I}_1 &= \mathcal{I}_L, \\ \mathcal{I}_2 &= (1-\theta)\mathcal{I}_R + \theta\mathcal{I}_L, \\ \mathcal{I}_3 &= \mathcal{I}_R. \end{cases} \tag{6.39}$$

and when $\sigma < 0$, by

$$\begin{cases} u_- &= u_* + \frac{\sigma\theta}{2a}(a-\sigma)(\tau_R - \tau_L), \\ u_+ &= u_* - \frac{\sigma\theta}{2a}(a+\sigma)(\tau_R - \tau_L), \\ \tau_1 &= \tau_L + \frac{1}{a}(u_- - u_L), \\ \tau_3 &= \tau_R + \frac{1}{a}(u_R - u_+), \\ \tau_2 &= \tau_1 + \theta(\tau_R - \tau_L), \end{cases} \quad \text{and} \quad \begin{cases} \pi_- &= \pi_* - \frac{\sigma\theta}{2}(a-\sigma)(\tau_R - \tau_L), \\ \pi_+ &= \pi_* - \frac{\sigma\theta}{2}(a+\sigma)(\tau_R - \tau_L), \\ \mathcal{I}_1 &= \mathcal{I}_L, \\ \mathcal{I}_2 &= (1-\theta)\mathcal{I}_L + \theta\mathcal{I}_R \\ \mathcal{I}_3 &= \mathcal{I}_R. \end{cases} \tag{6.40}$$

The states denoted with a star subscript $*$ correspond to intermediate states in the Suliciu ARS. Their expressions have been given in (6.35).

The key point of [CC13] is to pick θ is such a way that the solver satisfies a discrete entropy inequality and gives the exact solution when the initial data is an isolated shock propagating with speed s (in which case $\sigma = s$ thanks to the choice (6.37)). Under the classical subcharacteristic condition

$$|a| > \max_{\tau \in \{\tau_L, \tau_{L*}, \tau_{R*}, \tau_R\}} \sqrt{-p'(\tau)}, \tag{6.41}$$

it can be achieved with the choice

$$\begin{aligned} \sigma\theta(\mathcal{I}_R - \mathcal{I}_L) = \max\Big(0, \min\Big(&\sigma(\mathcal{I}_R - \mathcal{I}_L), -2a(a^2-\sigma^2)\mathcal{A}(v_L, v_R), \\ &\frac{a^2|\sigma|(a+|\sigma|)}{a+|\sigma|/2}(\tau_{R*} - \tau_{L*})\,\mathrm{sign}(\sigma)\Big)\Big), \end{aligned} \tag{6.42}$$

where

$$\mathcal{A}(v_L, v_R) = \frac{\int^{\tau_R} p(s)\,ds + \frac{p_R}{2a^2} - \int^{\tau_L} p(s)\,ds - \frac{p_L}{2a^2}}{\mathcal{I}_R - \mathcal{I}_L} - \frac{\pi_*}{a^2}.$$

We recall below the main result of [CC13].

Theorem 6.6. *Under the subcharacteristic condition* (6.41)*, the approximate Riemann solver defined by* (6.37)*,* (6.39)-(6.40) *and* (6.42) *is conservative and entropy satisfying, preserves the phase space Ω, is Lipschitz continuous with respect to the initial Riemann data, and is exact on isolated shocks (which, again, means that if the initial condition is such that the exact solution is an isolated entropy shock wave, the proposed approximate solution coincides with the exact one).*

6.6.2 Detection of shock and when to accept the reconstruction

The candidate reconstructed states are chosen thanks to the approximate Riemann solver we just presented. We take

$$\mathcal{CR}(U_L, U_M, U_R) = (U_-, U_+),$$

where U_- and U_+ are the states on the left and of the right of the σ-wave in (6.36) (see also Figure 6.22). This criterion verifies Proposition 6.1 because the approximate solver is exact on isolated shocks. The reconstruction is accepted if both the specific volume τ and the velocity u can be reconstruct in a conservative manner:

$$\mathcal{R}(U_{j-1}^n, U_j^n, U_{j+1}^n) = \begin{cases} \mathcal{RC}(U_{j-1}^n, U_j^n, U_{j+1}^n) \text{ if } d_j^{n,\rho} \in (0,\Delta x) \text{ and } d_j^{n,\rho} \in (0,\Delta x) \in (0,\Delta x) \\ (U_j^n, U_j^n) \text{ otherwise.} \end{cases}$$

We set the speed of the reconstructed shock $s(U_{j,L}^n, U_{j,R}^n)$ according to (6.37).

Proposition 6.7. *The reconstruction operator $\mathcal{R}$ defined above verifies the hypothesis of Theorem* (6.3)*. Thus the associated reconstruction scheme is exact on isolated shocks.*

Proof. Let U_L and U_R be linked by a shock. Theorem 6.6 insures that for all α in $[0,1]$,

$$\mathcal{R}(U_L, \alpha U_L + (1-\alpha)U_R, U_R) = (U_L, U_R),$$

and that the speed associated to the reconstructed shock coincides with the speed of the shock.

By construction, the approximate Riemann solver of [CC13] is exact on $(\mathcal{R}_{j_0^n})$. Therefore, the shock is correctly reconstructed in cell j_0^n, it has the correct speed

$$\sigma := \frac{u_L - u_R}{\tau_R - \tau_L}.$$

Now, consider the Riemann problem between the left state (τ_L, u_L) and $(\tilde\tau_R, \tilde u_R) := (\alpha\tau_R + (1-\alpha)\tau_L, \alpha u_R + (1-\alpha)u_L)$, with $\alpha \in (0,1)$. We denote $\tilde\sigma$ the speed of the σ-wave in this Riemann problem, which is

$$\tilde\sigma := \sqrt{\frac{p_L - \tilde p_R}{\tilde\tau_R - \tau_L}},$$

by (6.37). As $\tau_+ \geq \tau_-$, if τ_- is larger than τ_L, it is impossible to reconstruct τ in a conservative manner, which insures

$$\mathcal{R}(U_L, U_L, \tilde U_R) = (U_L, U_L).$$

The difference $\tau_- - \tau_L$ expresses as a function of θ:

$$\begin{aligned}
\tau_-(\theta) - \tau_L &= \tau_+(\theta) - \theta(\tilde\tau_R - \tau_L) - \tau_L \\
&= \tilde\tau_R + \frac{1}{a}(\tilde u_R - u_+(\theta)) - \theta(\tilde\tau_R - \tau_L) - \tau_L \\
&= (1-\theta)(\tilde\tau_R - \tau_L) + \frac{1}{a}\left(\tilde u_R - u_* + \frac{\tilde\sigma\theta}{2a}(a+\tilde\sigma)(\tilde\tau_R - \tau_L)\right) \\
&= (\tilde\tau_R - \tau_L) + \frac{\tilde u_R - u_*}{a} - \theta(\tilde\tau_R - \tau_L)\left(1 - \frac{\tilde\sigma}{2a^2}(a+\tilde\sigma)\right).
\end{aligned}$$

This function is decreasing with respect to θ. Indeed, $\tilde\tau_R$ il larger than τ_L, and as a is larger than $\tilde\sigma$,

$$\frac{\tilde\sigma(a+\tilde\sigma)}{2a^2} \leq 1.$$

Let us prove that the quantity

$$Q := \tilde u_R - u_* + \frac{\tilde\sigma}{2a}(a+\tilde\sigma)(\tilde\tau_R - \tau_L)$$

is positive. Replacing u_* by its value, we obtain

$$Q = \frac{\tilde{u}_R - u_L}{2} + \frac{\tilde{p}_R - p_L}{2a} + \frac{\tilde{\sigma}}{2a}(a + \tilde{\sigma})(\tilde{\tau}_R - \tau_L)$$

Now, we divide by $\tilde{\tau}_R - \tau_L$ and use the definition of $\tilde{\sigma}$ to write

$$\frac{2Q}{\tilde{\tau}_R - \tau_L} = -\sigma - \frac{\tilde{\sigma}^2}{a} + \frac{\tilde{\sigma}}{a}(a + \tilde{\sigma}) = \tilde{\sigma} - \sigma$$

We conclude by remarking that as the pressure is convex,

$$\tilde{p}_R = p(\alpha\tau_R + (1-\alpha)\tau_L) \leq \alpha p_R + (1-\alpha)p_L$$

and thus

$$\tilde{\sigma} = \sqrt{\frac{p_L - \tilde{p}_R}{\tilde{\tau}_R - \tau_L}} \geq \sqrt{\frac{p_L - p_R}{\tau_R - \tau_L}} = \sigma.$$

It is easier to prove that

$$\mathcal{R}(\tilde{U}_L, U_R, U_R) = (U_R, U_R),$$

where $\tilde{U}_L := \alpha U_L + (1-\alpha)U_R$. We have

$$\begin{aligned} u_+ - u_R &\geq u_* - \frac{\tilde{\sigma}}{2a}(a + \tilde{\sigma})(\tau_R - \tilde{\tau}_L) - u_R \\ &= \frac{\tilde{u}_L - u_R}{2} + \frac{p_R - \tilde{p}_L}{2a} - \frac{\tilde{\sigma}}{2a}(a + \tilde{\sigma})(\tau_R - \tilde{\tau}_L) \\ &= \frac{\tau_R - \tilde{\tau}_L}{2}\left(\sigma - \frac{\tilde{\sigma}^2}{a} - \frac{\tilde{\sigma}}{a}(a + \tilde{\sigma})\right) \\ &= \frac{\tau_R - \tilde{\tau}_L}{2}\left(\sigma - \tilde{\sigma}\right). \end{aligned}$$

It follows that u_+ is larger than u_R:

$$\tilde{\sigma} = \sqrt{\frac{\tilde{p}_L - p_R}{\tau_R - \tilde{\tau}_L}} \leq \sqrt{\frac{p_L - p_R}{\tau_R - \tau_L}} = \sigma.$$

Thus, $\tau_+ = \tau_R + \frac{1}{a}(u_R - u_+)$ is smaller than τ_L, and as τ_- is smaller than τ_+, it is once again impossible to reconstruct τ in a conservative manner. ☐

6.6.3 On a fixed grid

In this section we present a discontinuous reconstruction scheme on a fixed grid. In this section only we take $x_{j+1/2}^{n+1} = x_{j+1/2}^{n}$ and thus $V_{\text{mesh}}^{n} = 0$. There is at least three advantages in using a fixed grid:

- the Courant number can be taken in $(0, 1)$, instead of in $(0, 0.5)$ with a staggered grid. Thus we can use time step twice as large;
- at equal Courant number, the scheme on a fixed grid is less diffusive than with a staggered grid;
- dealing with boundary conditions is easier.

The problem is to deal with interactions between reconstructed shocks near an interface. The basic idea is to cancel some reconstructions when it happens. Typically in the example of Figure 6.23, where two reconstructed shocks cross the blue interface within the time step, we decide to cancel those two reconstructions and to use a classical numerical flux. This will be sufficient to obtain the main property of the scheme, namely to be exact when the initial data is an isolated shock.

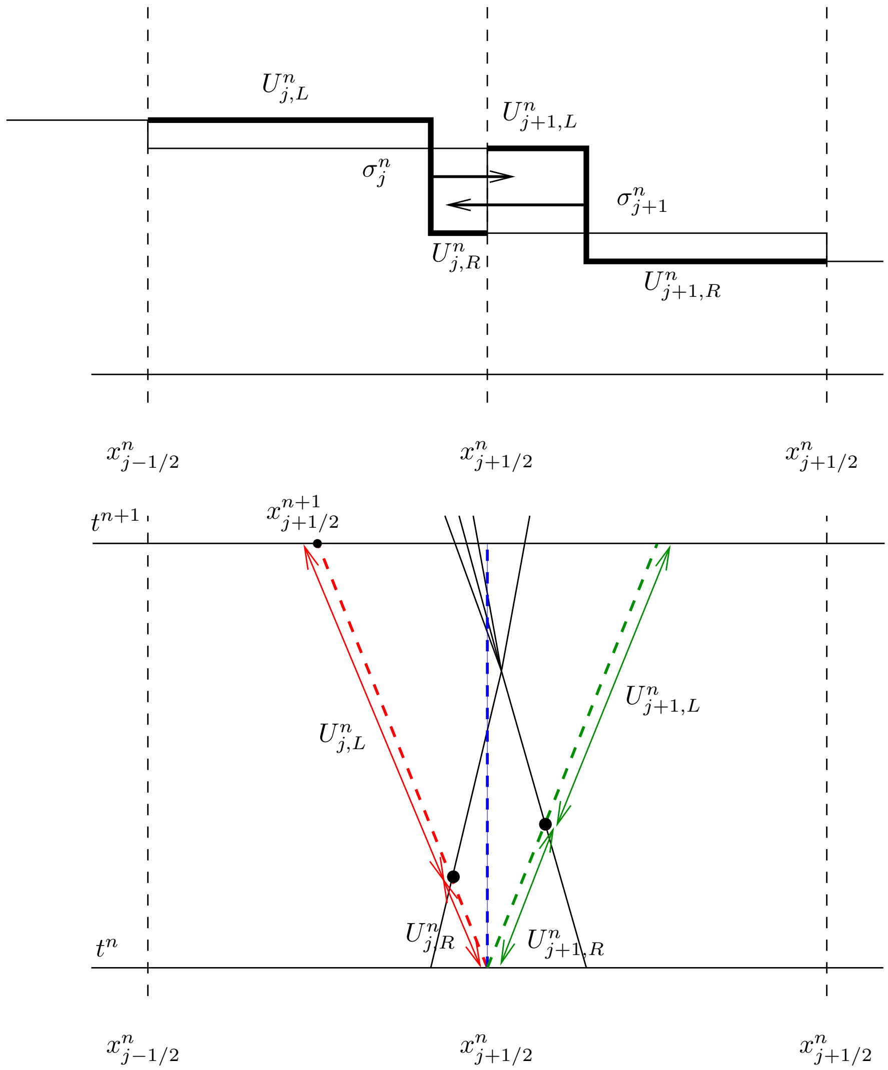


Figure 6.23. *Top: Reconstruction of a* 1*-shock in the* j*-th cell* ($\sigma^n_j > 0$) *and of a* 2*-shock in the* $j+1$*-th cell* ($\sigma^n_{j+1} < 0$)*. Bottom: in black, positions of the reconstructed shocks through times, and the waves emitted when they interact. It is easy to compute the flux through the red and green interfaces, but not through the blue one.*

Introducing discussion

If a shock is reconstructed in the j-th cell and has a positive speed ($\sigma_j^n > 0$), it is tempting to use the flux Formula (6.20) which gives the flux passing through the $j + 1/2$ interface if the initial data were

$$\begin{cases} \tau^0(x) = \tau_{j,L}^n \times \mathbf{1}_{x<x_{j-1/2}^n + d_j^{n,\tau}} + \tau_{j,R}^n \times \mathbf{1}_{x>x_{j-1/2}^n + d_j^{n,\tau}}, \\ u^0(x) = u_{j,L}^n \times \mathbf{1}_{x<x_{j-1/2}^n + d_j^{n,u}} + u_{j,R}^n \times \mathbf{1}_{x>x_{j-1/2}^n + d_j^{n,u}}, \end{cases}$$

(see what happens through the red interface of the bottom of Figure 6.23).

However, if a shock is reconstructed in the $j + 1$ cell and has a negative speed ($\sigma_{j+1}^n < 0$), it is similarly tempting to use the flux Formula (6.18), which gives the flux passing though the $j + 1/2$ interface if the initial data were

$$\begin{cases} \tau^0(x) = \tau_{j+1,L}^n \times \mathbf{1}_{x<x_{j+1/2}^n + d_{j+1}^{n,\tau}} + \tau_{j+1,R}^n \times \mathbf{1}_{x>x_{j+1/2}^n + d_{j+1}^{n,\tau}}, \\ u^0(x) = u_{j+1,L}^n \times \mathbf{1}_{x<x_{j+1/2}^n + d_{j+1}^{n,u}} + u_{j+1,R}^n \times \mathbf{1}_{x>x_{j+1/2}^n + d_{j+1}^{n,u}}, \end{cases}$$

(see what happens through the green interface of the bottom of Figure 6.23).

In the variable τ, the two shocks interact at the time $T_{j+1/2}^{n,\text{inter},\tau}$ such that

$$d_j^{n,\tau} + \sigma_j^n T_{j+1/2}^{n,\text{inter},\tau} = \Delta x + d_{j+1}^{n,\tau} + \sigma_{j+1}^n T_{j+1/2}^{n,\text{inter},\tau},$$

and we obtain a different interaction time in the variable u. We evaluate the time at which the two shocks interact by

$$T_{j+1/2}^{n,\text{inter}} = \min(T_{j+1/2}^{n,\text{inter},\tau}, T_{j+1/2}^{n,\text{inter},u}) = \min\left(\frac{d_{j+1}^n + \Delta x - d_j^n}{s_j^n - s_{j+1}^n}\right),$$

where the second minimum is taken on the two components of the vectors d_j^n and d_{j+1}^n. The interaction can also occur when when σ_j^n and σ_{j+1}^n have the same sign and $\sigma_j^n > \sigma_{j+1}^n$. If this time is smaller than Δt^n, the shocks interact within the time step and the waves created by the resulting interaction are likely to meet the $j + 1/2$-th interface (which was not the case under Condition (6.14), see once again Figure (6.23)). In that case, we decide to not to take into account the reconstructions in cells j and $j + 1$, and to simply use the Godunov-type flux (associated with (6.36)) instead of the reconstruction flux (6.20) or (6.18).

Definition of the fluxes

Let us now be more precise and introduce some notation. We denote by $\mathcal{F}^{\text{GOD}}(U_L, U_R)$ the Godunov-type scheme associated to the ARS (6.36):

$$\mathcal{F}^{\text{GOD}}(U_L, U_R) = \begin{cases} (-u_-, \pi_-) & \text{if } \sigma > 0, \\ (-u_+, \pi_+) & \text{if } \sigma \leq 0, \end{cases}$$

(see Figure 6.22). We also denote by $\mathcal{F}_{j-1/2}^{n,\leftarrow}$ the flux given by (6.18) and by $\mathcal{F}_{j+1/2}^{n,\rightarrow}$ the flux given by (6.20). Moreover, we use the convention $\sigma_j^n = 0$ if no reconstruction is performed in cell j, and we extend the definition of the interaction time by

$$T_{j+1/2}^{n,\text{inter}} = \begin{cases} \min\left(\frac{d_{j+1}^n + \Delta x - d_j^n}{\sigma_j^n - \sigma_{j+1}^n}\right) & \text{if } \sigma_j^n > \sigma_{j+1}^n, \text{ and } \sigma_{j+1}^n \sigma_j^n \neq 0 \\ +\infty & \text{otherwise.} \end{cases} \tag{6.43}$$

We propose the following fluxes:

$$\mathcal{F}^n_{j+1/2} = \begin{cases} \mathcal{F}^{n,\rightarrow}_{j+1/2} & \text{if } \sigma^n_j > 0,\ T^{n,\text{inter}}_{j+1/2} > \Delta t \text{ and } \sigma^n_{j+1} \geq 0, \\ \mathcal{F}^{n,\leftarrow}_{j+1/2} & \text{if } \sigma^n_{j+1} < 0,\ T^{n,\text{inter}}_{j+1/2} > \Delta t \text{ and } \sigma^n_j \leq 0, \\ \mathcal{F}^{\text{GOD}}(U^n_j, U^n_{j+1}) & \text{if } (\sigma^n_j > 0 \text{ and } \sigma^n_{j+1} < 0) \text{ or } T^{n,\text{inter}}_{j+1/2} \leq \Delta t, \\ \mathcal{F}^{\text{GOD}}(U^n_{j,R}, U^n_{j+1,L}) & \text{otherwise.} \end{cases} \tag{6.44}$$

In particular for simplicity, we cancel the reconstructions in cells j and $j+1$ if $\sigma^n_j > 0$ and $\sigma^n_{j+1} < 0$. We now justify the choice of flux $F^{n,\rightarrow}_{j+1/2}$ if $\sigma^n_j > 0$ and $\sigma^n_{j+1} = 0$, which corresponds to the case where a shock with positive speed is reconstructed in the j-th cell but no shock is reconstructed in the $j+1$-th cell. In that case the $-a^n_j$, 0 and σ^n_j-waves in the solution of (6.36) with the left state $U^n_{j,R}$ and the right state U^n_{j+1} are trivial, and only the a^n_j-wave is nontrivial. Thus, as $a^n_j > 0$ and as depicted on Figure 6.24, no wave interaction occur at interface $j+1/2$. The flux through the interface $j+1/2$ is easily computed and given by $\mathcal{F}^{n,\rightarrow}_{j+1/2}$.

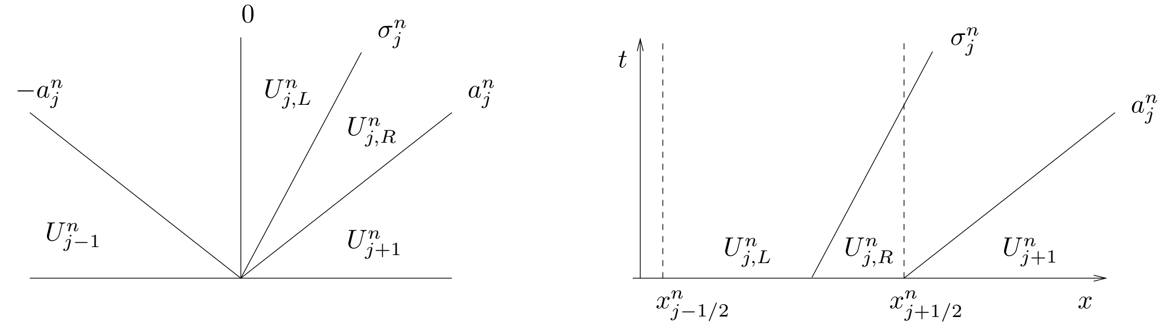


Figure 6.24. *Left: the structure of the ARS used to determine $U^n_{j,L}$ and $U^n_{j,R}$. Right: the waves emitted in the reconstructed solution: there is only a wave at speed a^n_j between $U^n_{j,R}$ and U^n_{j+1}, which is faster than the reconstructed shock. Thus the flux through the interface $j+1/2$ is easily computed and given by $\mathcal{F}^{n,\rightarrow}_{j+1/2}$.*

6.6.4 One dimensional simulations

To begin with we test the discontinuous reconstruction scheme on a staggered grid. The space interval $[-1,1]$ is discretized with 100 cells and the pressure law is $p(\tau) = \tau^{-2}$. We take

$$\Delta t^n = 0.45 \frac{\Delta x}{V^n_{\text{waves}}} \text{ and } V^n_{\text{mesh}} = (-1)^n \frac{\Delta x}{2\Delta t^n}.$$

Thus (6.2) and (6.14) hold. In each case, we compare the discontinuous reconstruction scheme and the Godunov-type scheme, both based on the approximate Riemann solver of [CC13]. Note that the "Godunov-type scheme" is computed on a fixed grid and its fluxes are given by

$$\mathcal{F}^n_{j+1/2} = \begin{cases} (-u^n_{j+1/2,-}, \pi^n_{j+1/2,-}) & \text{if } \sigma > 0, \\ (-u^n_{j+1/2,+}, \pi^n_{j+1/2,+}) & \text{if } \sigma \leq 0. \end{cases}$$

Here,

$$(u^n_{j+1/2,-}, u^n_{j+1/2,+}, \pi^n_{j+1/2,-}, \pi^n_{j+1/2,+})$$

is defined by (6.39)-(6.40), with left and right states

$$(\tau_L, u_L, \mathcal{T}_L) = (\tau^n_j, u^n_j, \tau^n_j) \text{ and } (\tau_R, u_R, \mathcal{T}_R) = (\tau^n_{j+1}, u^n_{j+1}, \tau^n_{j+1}).$$

The first three test cases are taken from [CC13], the fourth one is a variant of the fast shock of Arora and Roe [AR97]. On this test cases, the results given by the proposed schemes on staggered and fixed grids are very close to each other (the scheme on a fixed grid being less diffusive in some cases, but with almost no impact in those cases). This is why we only show the results given by the scheme on a staggered grid.

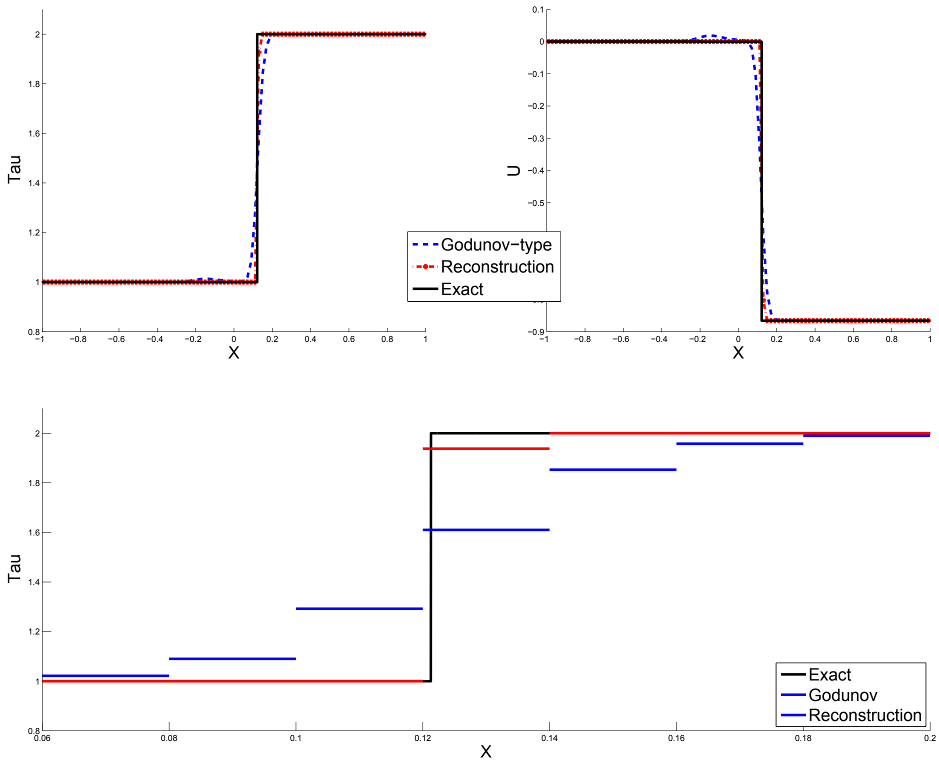


Figure 6.25. *Test* 1 *: specific volume (left) and velocity (right) at time* $t = 0.15$.

Test 1: Isolated shock

The first Riemann problem is

$$\tau_L = 1,\ u_L = 0,\ \tau_R = 0 \ \text{ and } \ u_R = -\sqrt{3}/2.$$

It corresponds to an isolated shock. Theorem 6.3 is illustrated on Figure 6.25: the shock is perfectly advected by the reconstruction scheme, while it is diffused by the Godunov-type scheme. On the bottom of this figure, we can see that there is only one intermediate value in the shock profile given by the reconstruction scheme, which corresponds to the average of the exact solution on the cell.

Test 2: Rarefaction and shock

The second Riemann problem corresponds to

$$\tau_L = 0.3,\ u_L = 0,\ \tau_R = 0.6 \ \text{ and } \ u_R = 0.$$

It contains a 1-rarefaction wave and a 2-shock. On Figure 6.26, we can see that the shock is sharply captured by the reconstruction scheme.

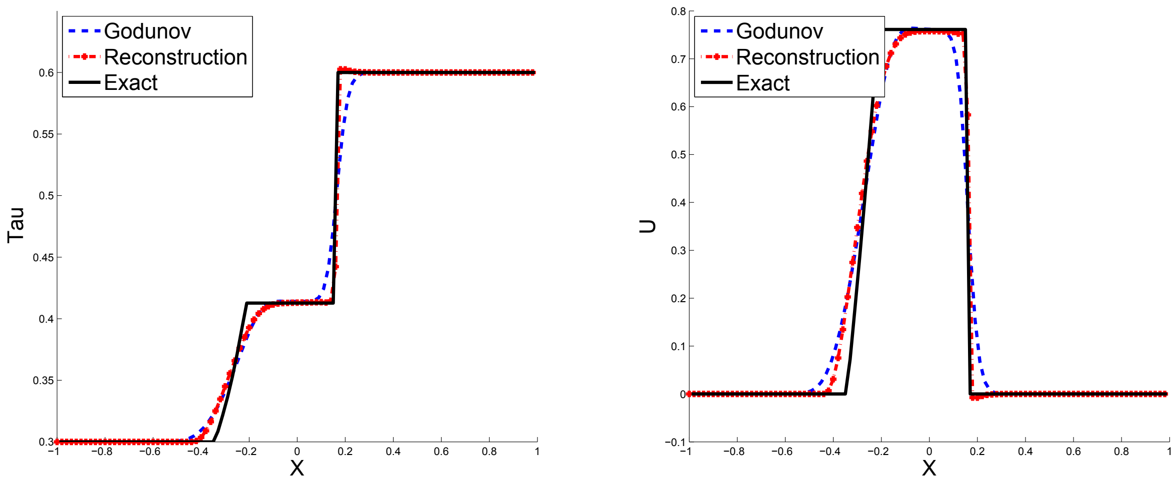


Figure 6.26. *Test 2 : specific volume (left) and velocity (right) at time $t = 0.04$.*

Test 3: Shock and shock

Here the Riemann problem corresponds to

$$\tau_L = 0.5,\ u_L = 2,\ \tau_R = 0.6 \text{ and } u_R = 0.$$

It contains a 1-shock and a 2-shock. Figure 6.27 clearly demonstrates that they are both sharply captured, with only one point of numerical diffusion.

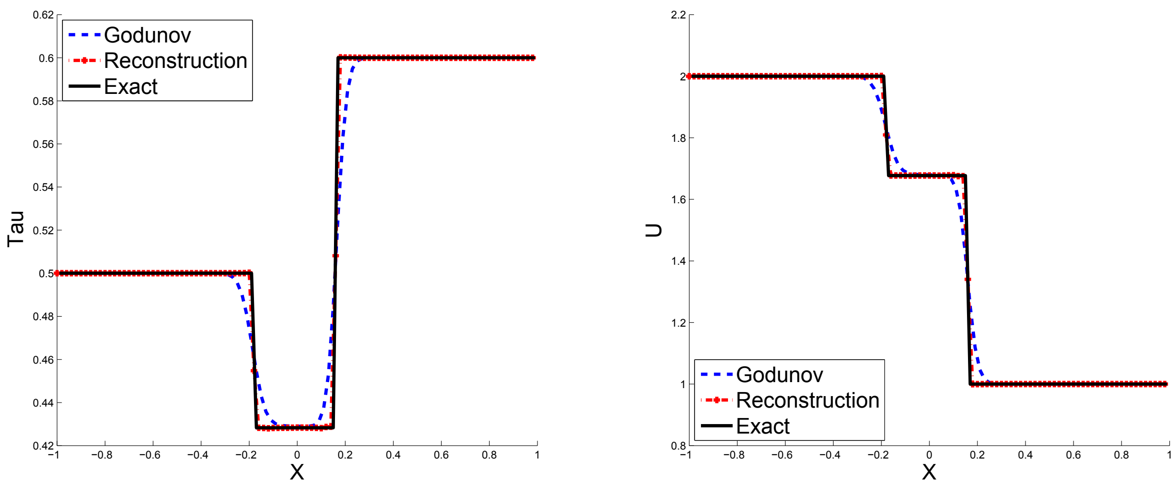


Figure 6.27. *Test 3 : specific volume (left) and velocity (right) at time $t = 0.05$.*

Test 4: shock and strong shock

For this test case we have

$$\tau_L = 1,\ u_L = 1,\ \tau_R = 2 \text{ and } u_R = -8,$$

which is a slight modification of the isolated fast shock considered in [AR97]. When computing a strong shock (also called fast shock), spurious oscillations appear in the velocity when using usual Godunov-type methods. They also appear clearly in the characteristic variable $u \pm \sqrt{8/\tau}$. Those two quantities are plotted on Figure 6.28. Theorem 6.3 states that the scheme is exact on isolated strong shocks. We can see on Figure 6.28 that the ability of the reconstruction scheme to approach correctly fast shocks is not lost when we introduce a perturbation (here, a second wave in the Riemann problem).

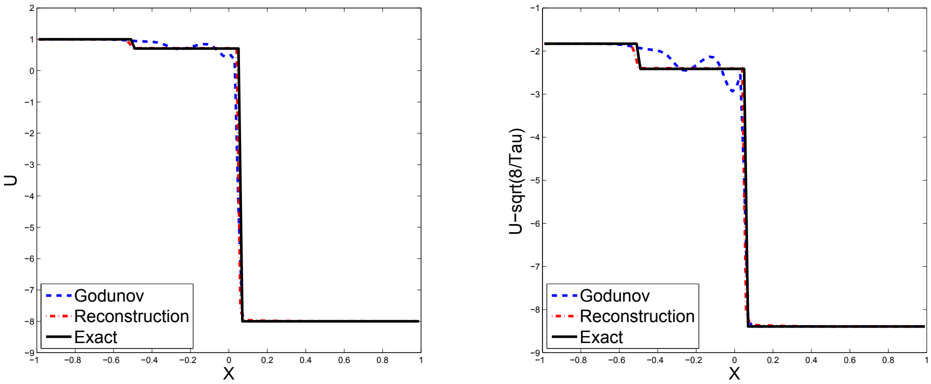


Figure 6.28. *Test* 4 *: velocity (left) and* $u - \sqrt{8/\tau}$ *(right) at time* $t = 0.3$.

Test 5**: Isentropic compression**

This test case consists in reversing the time starting from a developed 1-rarefaction wave. More precisely, let us denote

$$(t, x) \mapsto U_{\text{rar}}(x/t; U_L, U_R) = (\tau_{\text{rar}}(x/t; U_L, U_R), u_{\text{rar}}(x/t; U_L, U_R))$$

the self-similar rarefaction wave associated with the Riemann initial condition $U_L \mathbf{1}_{x<0} + U_R \mathbf{1}_{x\geq 0}$, with $U_L = (\tau_L, u_L)$, $U_R = (\tau_L, u_L)$ and

$$\tau_L = 0.5,\ u_L = 10,\ \tau_R = 5 \ \text{ and } \ u_R = u_L - \sqrt{2}(\tau_R^{-1/2} - \tau_L^{-1/2}).$$

For this test case, the initial data is

$$\tau^0(x) = \tau_{rar}\left(\frac{-x}{T_{\text{rar}}}\right) \quad u^0(x) = u_{rar}\left(\frac{-x}{T_{\text{rar}}}\right)$$

with $T_{\text{rar}} = 0.2$. This initial condition is plotted on Figure 6.29.

For times $t < T_{\text{rar}}$, the exact solution is $U(t, x) = U_{rar}\left(\frac{-x}{T_{\text{rar}}-t}\right)$, and for times $t \geq T_{\text{rar}}$, the solution coincides with the solution of the Riemann problem associated with left state U_R and right state U_L (at time $t - T_{\text{rar}}$), which contains a 1-shock and a 2-rarefaction. In particular at time $t = T_{\text{rar}}$ the exact solution is nothing but a discontinuity between U_R and U_L and located at $x = 0$.

We compare the reconstruction scheme on a staggered grid (abbreviated `RecSG` in the legends) and on a fixed grid (abbreviated `RecFG`) at five different times. This is the only test case where the results are significantly different for the two schemes. On Figures 6.30 and 6.31 we plot the solutions for $t = 0.14$, $t = 0.17$, $t = 0.2$, $t = 0.23$ and $t = 0.26$. We observe a spike near the discontinuity when the time t is close to T_{rar}. It seems that the reconstruction scheme starts to reconstruct a little before $t = T_{\text{rar}}$

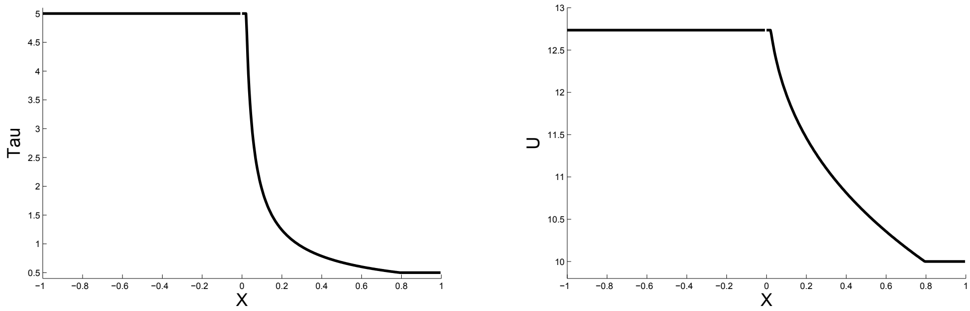


Figure 6.29. *Initial condition for test* 5.

the expected shock that will appear in the exact solution for times $t > T_{\mathrm{rar}}$, and that will be eventually sharply computed by the proposed scheme. The maximal height of the spike is at $t = T_{\mathrm{rar}}$, and it diminishes rapidly afterwards, the shock being correctly approximated after T_{rar}. The spike does not prevent the scheme from converging in L^1, and its size is independent from the Courant number. On this test case, we observe that the results given on a fixed grid are much better than the ones given on a staggered grid. On the other hand, we note that the Godunov-type scheme does not create any spike, but introduces numerical diffusion on the shock wave for times $t > T_{\mathrm{rar}}$, as it is expected.

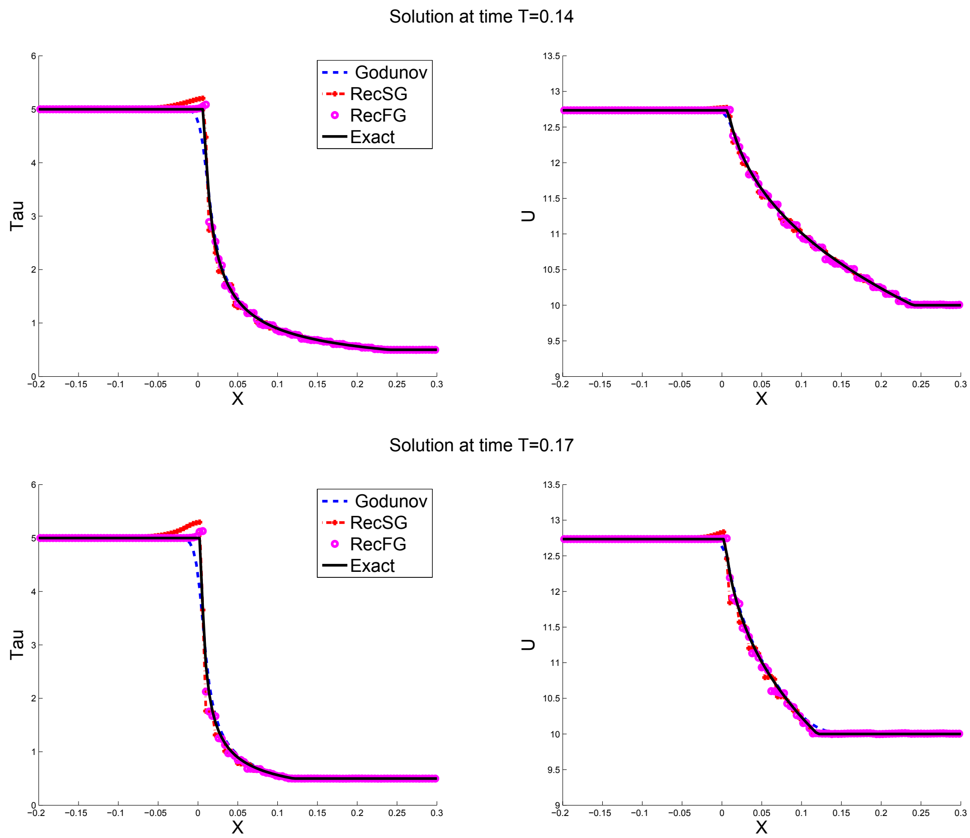


Figure 6.30. *Solution of test* 5 *with, from top to bottom,* $t = 0.14$ *and* $t = 0.17$.

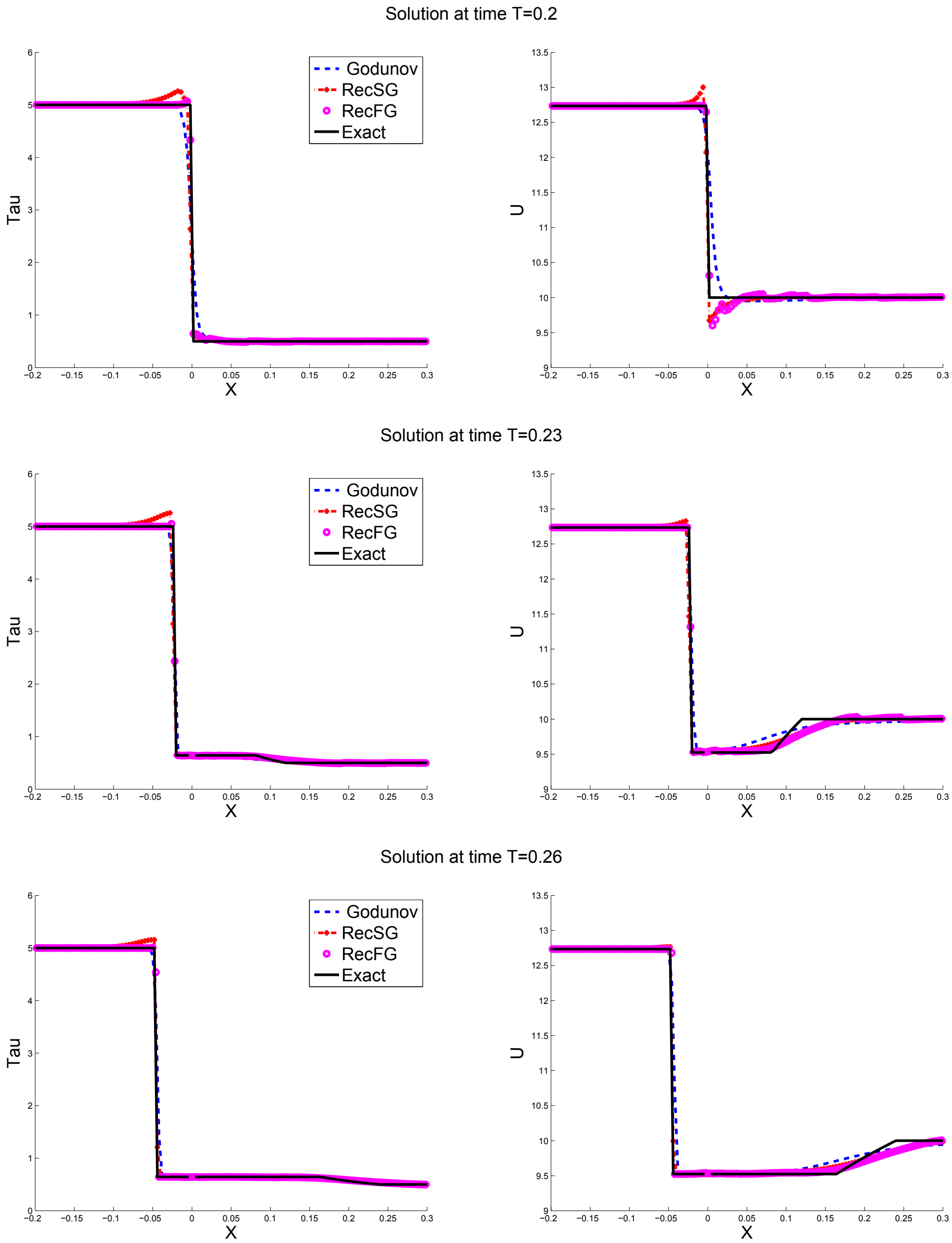


Figure 6.31. *Solution of test* 5 *with, from top to bottom,* $t = 0.2$, $t = 0.22$ *and* $t = 0.26$.

On Figure 6.32, we plot the evolution of the entropy during the time interval $[0, 4]$ for this test case. More precisely, we plot the evolution of the quantity

$$\int \mathcal{U}(\tau^n_{\text{rec}}(x), u^n_{\text{rec}}(x))dx - \int \mathcal{U}(\tau^{n-1}_{\text{rec}}(x), u^{n-1}_{\text{rec}}(x))dx$$
$$+ (t^n - t^{n-1})((\mathcal{W}(\tau_R, u_R) - V^{n-1}_{\text{mesh}}\mathcal{U}(\tau_R, u_R)) - (\mathcal{W}(\tau_L, u_L) - V^{n-1}_{\text{mesh}}\mathcal{U}(\tau_L, u_L)))$$

where

$$\tau^n_{\text{rec}}(x) = \sum_j (\tau^n_{j,L}\mathbf{1}_{x<x^n_{j-1/2}+d^{n,\tau}_j} + \tau^n_{j,R}\mathbf{1}_{x\geq x^n_{j-1/2}+d^{n,\tau}_j})\mathbf{1}_{x\in[x^n_{j-1/2},x^n_{j+1/2}]}$$

and

$$u^n_{\text{rec}}(x) = \sum_j (u^n_{j,L} \mathbf{1}_{x < x^n_{j-1/2} + d^{n,u}_j} + u^n_{j,R} \mathbf{1}_{x \geq x^n_{j-1/2} + d^{n,u}_j}) \mathbf{1}_{x \in [x^n_{j-1/2}, x^n_{j+1/2}]}$$

are the reconstructed solution at time t^n (we recall than on a fixed grid, $V^{n-1}_{\text{mesh}} = 0$ for all n).

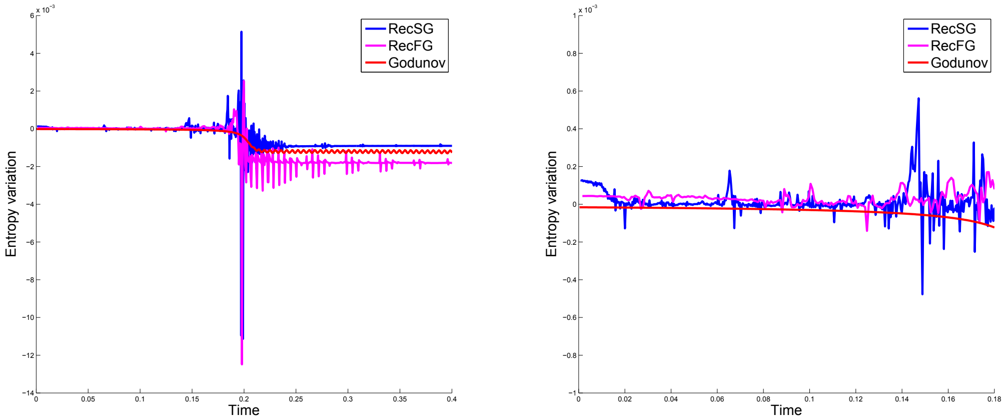


Figure 6.32. *Evolution of the entropy through time in test 5, on the time interval* $[0, 4]$ *(top), and a zoom before time* T_{rar} *(bottom).*

On Figure 6.32, we can see that this quantity oscillates around 0 and often takes positive values until the time $T_{\text{rar}} = 0.2$. It indicates that the scheme is not entropy satisfying in the strict sense, but only in a "weak" sense to be defined. Indeed, the scheme seems to converge toward the correct solution for the L^1-norm. The order of convergence is plotted on Figure 6.33.

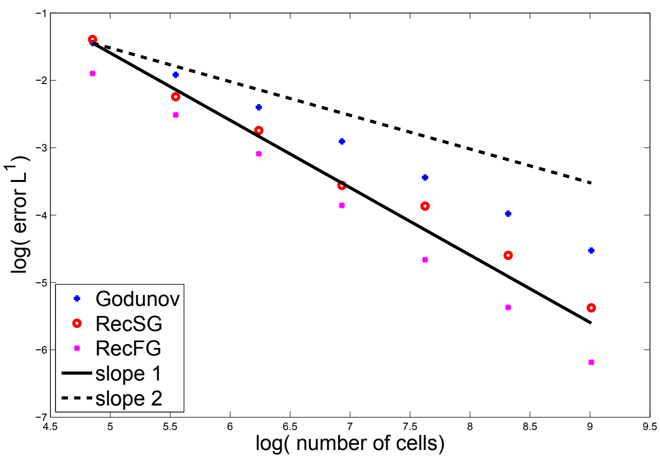


Figure 6.33. L^1*-error at time* T_{rar} *with the Godunov-type scheme and with the discontinuous reconstruction schemes.*

Note also that even though the scheme is not entropy satisfying, it contains entropy information as the underlying approximate Riemann solver is entropy satisfying. It is thus likely that the problem might come from the discontinuous reconstruction strategy. More precisely, we believe that some reconstructions are not entropy satisfying and should be cancelled using a stronger constraint than $d_j^{n,\tau} \in [0, \Delta x]$ and $d_j^{n,u} \in [0, \Delta x]$. For example in this case, we expect the reconstructed solution $U^n_{\text{rec}} = (\tau^n_{\text{rec}}, u^n_{\text{rec}})$ to be monotonous for times $t < T_{\text{rar}}$, but as depicted on Figure 6.34, this is not the case at all in the course of the first iteration ($n = 0$). This contrasts with the case of a single conservation law with a convex flux, illustrated by the Burgers equation on Figure 6.4.

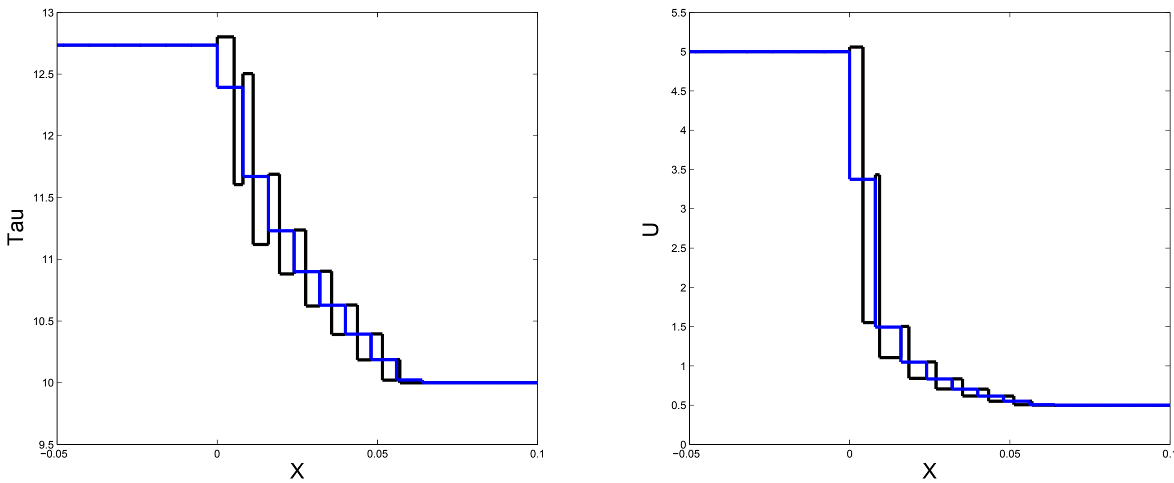


Figure 6.34. *In black, the reconstructed solution U^0_{rec}, i.e. the solution reconstructed from the initial data before any iteration in time is done. We take $T_{\text{rar}} = 0.015$ and 500 cells. In blue, the piecewise constant function corresponding to $(U^0_j)_{j\in\mathbb{Z}}$*

In the scalar case, entropy satisfying versions of the anti-diffusive scheme of [DL01] (which can be reinterpreted in terms of discontinuous reconstruction) have been proposed in [LagXX] and [Bou04a], but remain to be adapted to the present setting. It is a challenging and open problem at the stage of the present work.

6.7 Two dimensional simulations

In two space dimensions, the p-system writes

$$\begin{cases} \partial_t \tau - \partial_x u - \partial_y v = 0, \\ \partial_t u + \partial_x p(\tau) = 0, \\ \partial_t v + \partial_y p(\tau) = 0, \\ \tau(0,x) = \tau^0(x),\ u(0,x) = u^0(x),\ v(0,x) = v^0(x). \end{cases} \tag{6.45}$$

The velocity of the fluid is (u, v). We use a directional splitting to approach the solution of (6.45) and we present below the simulation of a column falling in a basinand reflecting on its sides. We take $p(\tau) = \tau^{-2}$, which correspond to the pressure law for the shallow-water equation written in Lagrangian coordinates (in that case τ represents the inverse of the height of water). The initial data is

$$\begin{cases} \tau^0(x,y) = \mathbf{1}_{\sqrt{x^2+y^2} \geq 0.5} + 2 * \mathbf{1}_{\sqrt{x^2+y^2} < 0.5}, \\ u^0(x,y) = v^0(x,y) = 0. \end{cases}$$

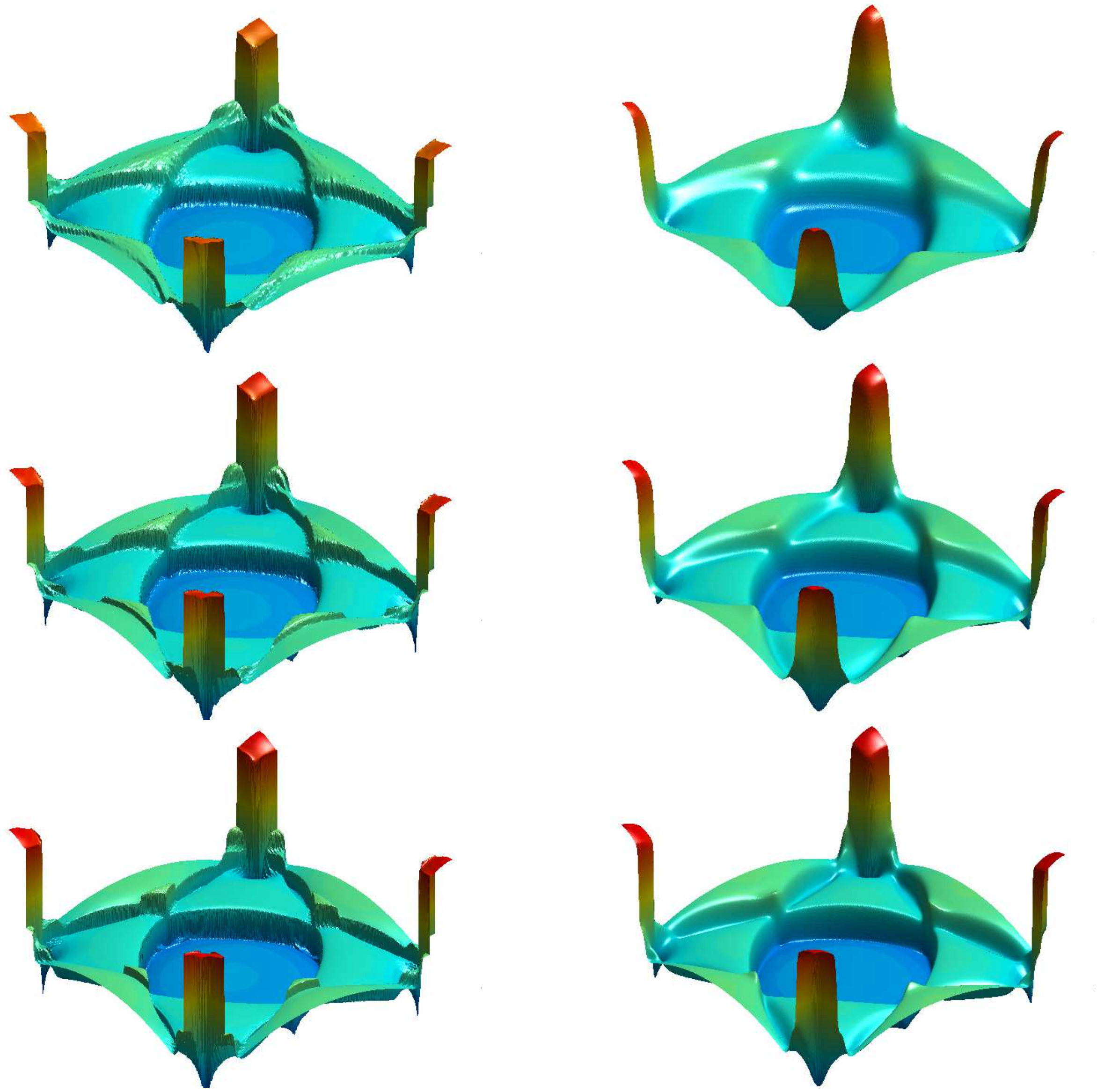


Figure 6.35. *Comparison of τ for the discontinuous reconstruction scheme (left) and the Godunov-type scheme (right) at time $t = 3.05$ for meshes with (from top to bottom):* 200×200, 400×400 *and* 800×800 *cells*

The basin is the square $[-1,1]^2$ and we take reflective boundary conditions on its sides. We compare the solution given by the discontinuous reconstruction scheme on a fixed grid and the Godunov-type scheme at different times.

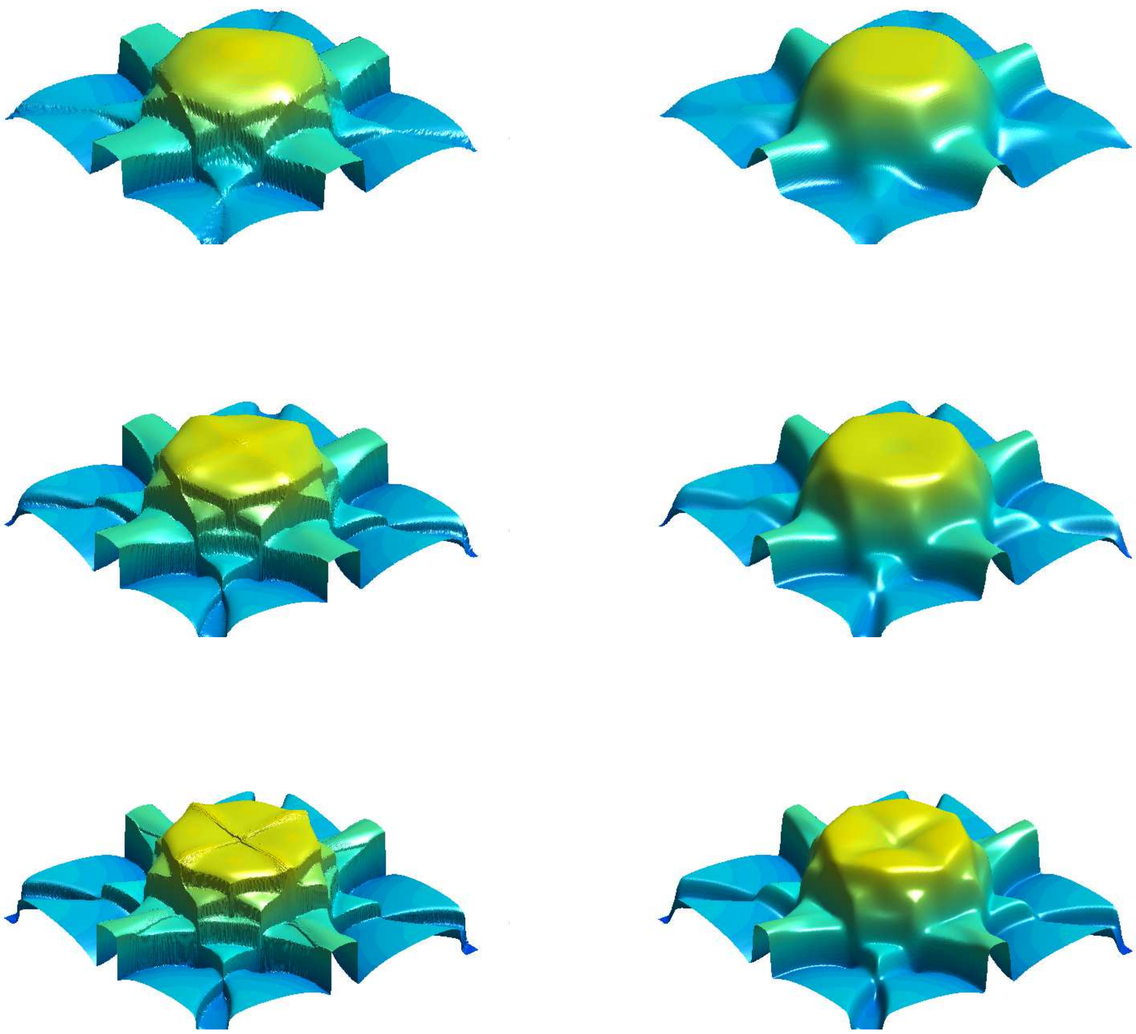


Figure 6.36. *Comparison of τ for the discontinuous reconstruction scheme (left) and the Godunov-type scheme (right) at time $t = 3.82$ for meshes with (from top to bottom):* 200×200, 400×400 *and* 800×800 *cells.*

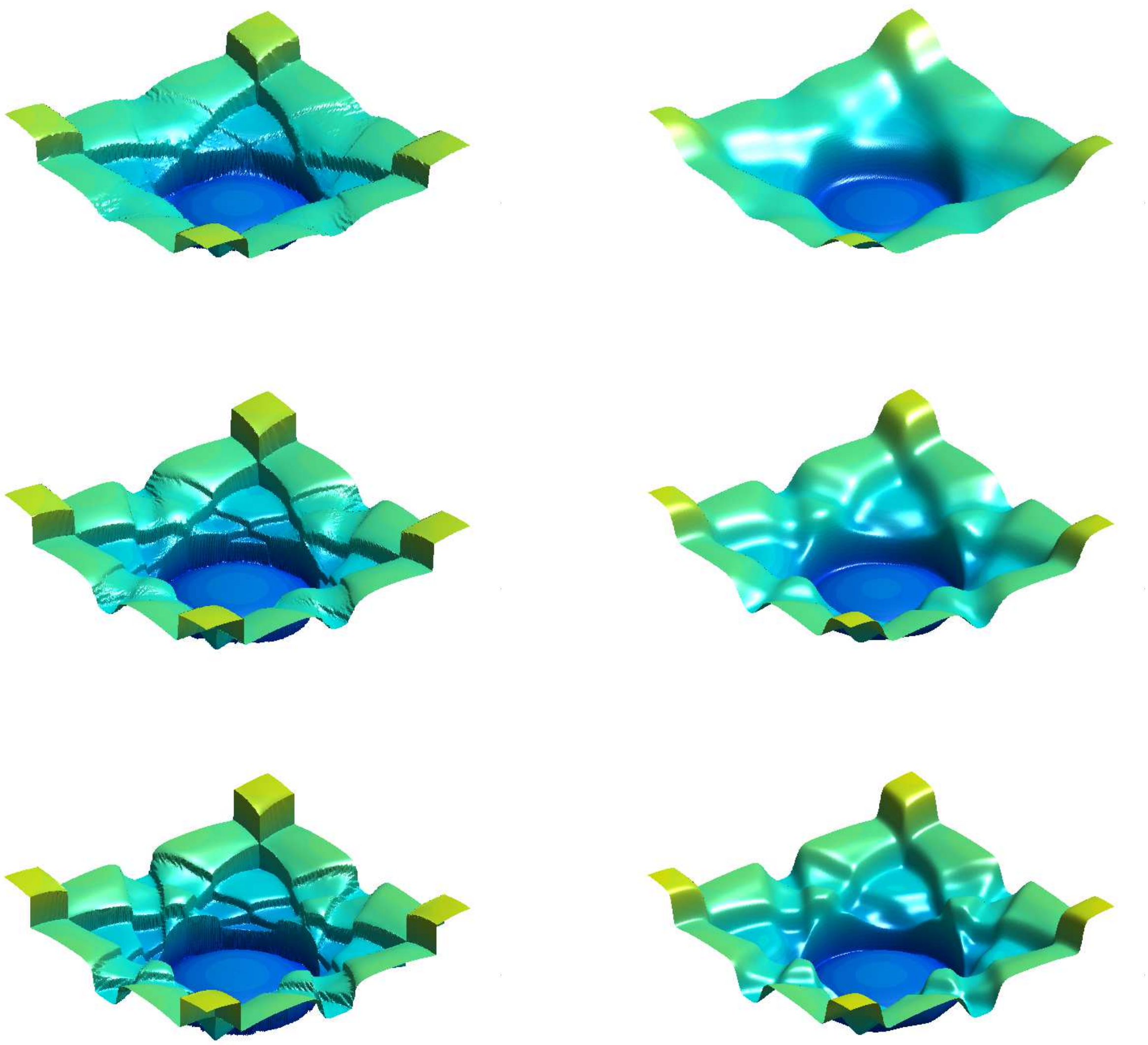


Figure 6.37. *Comparison of τ for the discontinuous reconstruction scheme (left) and the Godunov-type scheme (right) at time $t = 4.58$ for meshes with (from top to bottom):* 200×200, 400×400 *and* 800×800 *cells.*

6.8 The reconstruction scheme for the nonlinear elastodynamic

We now apply the discontinuous reconstruction to the system of conservation laws

$$\begin{cases} \partial_t v - \partial_x \sigma(w) = 0, \\ \partial_t w - \partial_x v = 0, \\ v(t=0,x) = v^0(x),\ w(t=0,x) = w^0(x), \end{cases} \tag{6.46}$$

where the stress σ is twice differentiable, and verifies

$$w\sigma''(w) > 0,\ \sigma'(w) > 0 \ \text{and} \ \lim_{|w|\to+\infty} \sigma'(w) = +\infty. \tag{6.47}$$

Such a system describes phase transitions between two different atomic lattices in a solid. The only difference between this system and the gas dynamics in Lagrangian coordinates described by (6.29) is that $-\sigma$, which plays the role of the pressure law p, is now convex-concave and not convex. A major consequence is that the equivalence between a single strictly convex entropy inequality and classical criteria to select shocks (such as Oleinik's criterion [Ole57] for scalar conservation law, or Lax [Lax57] and Liu [Liu74] criteria for systems) is lost. There exists shocks that dissipate entropy but do not verify those criteria. They are called *nonclassical shocks*, and these are the discontinuities of interest we want to approach with a discontinuous reconstruction scheme.

6.8.1 Nonclassical shocks and kinetic relations

Consider the Cauchy problem for a system of conservation law

$$\begin{cases} \partial_t U + \partial_x F(U) = 0, \\ U(t=0,x) = U^0(x), \end{cases} \tag{6.48}$$

where U is the vector regrouping the n unknowns belonging to a open set Ω of $\mathbb{R}^n$, and $F = \mathbb{R}^n \to \mathbb{R}^n$ is the flux function. In the case of nonlinear elasticity (6.46), $n = 2$, $U = (v, w)$ and $F(U) = (-\sigma(w), -v)$. When the Jacobian matrix of F is diagonalizable with separated eigenvalues $\lambda_1(U) < \cdots < \lambda_n(U)$, System (6.48) is said to be strictly hyperbolic. We denote by $r_i(U)$ a right eigenvector associated to the eigenvalue $\lambda_i(U)$. In the case of nonlinear elasticity (6.46), the eigenvalues are $\lambda_i(w) = (-1)^i\sqrt{\sigma'(w)}$ and the eigenvectors are $r_i(w) = ((-1)^{i+1}\sqrt{\sigma'(w)}, 1)^{\,t}$.

Solutions of hyperbolic systems are in general not smooth, even when the initial data is smooth. Discontinuities appear in finite time and it is necessary to consider weak solutions. Doing so, uniqueness is lost, and an additional criterion has to be imposed to select a unique solution. For example, the vector of unknown U must verify, in addition to the initial system, a so called entropy inequality

$$\partial_t \mathcal{U}(U) + \partial_x \mathcal{W}(U) \le 0. \tag{6.49}$$

The entropy $\mathcal{U}$ and the entropy flux $\mathcal{W}$ usually come from considerations on the system under study, where small physical terms of diffusion and dispersion are not neglected. Such an augmented system has additional properties, for example the total energy is conserved, and the inequality (6.49) keeps a trace of those properties when the small scales vanishe. In general, from the augmented system only one entropy inequality is deduced, thus the solution of (6.48) are not expected to verify other entropy inequalities than (6.49). In the sequel we suppose that the entropy $\mathcal{U}$ is strictly convex.

Example 3. In the case of the nonlinear elastodynamic, the augmented system is

$$\begin{cases} \partial_t v^\varepsilon - \partial_x \sigma(w^\varepsilon) = \varepsilon \partial_{xx} v + \alpha \varepsilon^2 \partial_{xxxx} w, \\ \partial_t w - \partial_x v = 0, \end{cases} \tag{6.50}$$

with ε and α positive. It yields the strictly convex entropy inequality (6.49) with

$$\mathcal{U}(v,w) = \frac{v^2}{2} + \int_0^w \sigma(z)dz \ \text{and} \ \mathcal{W}(v,w) = -v\sigma(w). \tag{6.51}$$

A crucial question is the one of criterion to decide wether a shock is admissible or not. Suppose that U_L and U_R are linked by a shock, and denote by s its velocity. The shocks is entropy satisfying, i.e. verifies (6.49) , if and only if

$$s(\mathcal{U}(U_L) - \mathcal{U}(U_R)) - (\mathcal{W}(U_L) - \mathcal{W}(U_R)) \leq 0 \tag{6.52}$$

When the system is genuinely nonlinear, i.e. when

$$\nabla\lambda_i(U) \cdot r_i(U) \neq 0 \, \forall i \in \{1, \cdots, n\}, \; \forall U \in \Omega,$$

and when the entropy $\mathcal{U}$ is stricly convex, (6.52) is equivalent to the Lax criterion [Lax57]:

$$\exists i \in \{1, \cdots, n\} : \; \lambda_i(U_R) < s < \lambda_i(U_L), \; s < \lambda_{i+1}(U_R) \; \text{ and } \; s > \lambda_{i-1}(U_L) \tag{6.53}$$

In [Liu74], Liu generalized this criterion to the case where $\nabla\lambda_i \cdot r_i$ may vanish. Its criterion is equivalent to (6.53) and (6.52) when the system is genuinely nonlinear. He proved that under this criterion, there exists a unique selfsimilar solution to the Riemann problem (6.48–6.49). However, for system that are not genuinely nonlinear like (6.46), the Liu criterion is stronger than (6.52). An admissible shock in the sense Liu verifies (6.52) for *any* strictly convex entropy $\mathcal{U}$, but some shocks verify a single strictly convex entropy inequality without satisfying the Liu criterion. These shocks are called nonclassical.

When System (6.48) is not genuinely nonlinear, the addition of the single entropy inequality (6.49) is not enough to regain uniqueness of the solution. In some sense, this inequality does not contain enough information on the small scale effects in the augmented system. The characterization of shocks that arise as limits of (6.50) is explored, for systems like (6.46) but with a pressure law that decreases on an interval, in [Sle89] and [AK91] for example. In general, a possible characterization is

$$U_L = \Phi^\flat(U_R),$$

where $\Phi^\flat$ is called a *kinetic relation*. We refer to the monograph [LeF02] for a review about kinetic relations and their links with diffusive-dispersive models. The literature on the subject is so wide that we only cite a few key references on System (6.46). The interested reader will find a more comprehensive bibliography on this topic in [LeF02].

The numerical approximation of nonclassical shocks is a challenging tasks. Roughly speaking, usual finite volume schemes introduce a numerical viscosity which destroys the equilibrium between the diffusion and the capillarity effects in (6.50). The diffusion becomes dominant and schemes converge toward the classical solution, even though they are based on nonclassical Riemann solver.

Two types of schemes have been proposed. On the one hand, it is possible to use high order schemes consistent with the augmented system (see for example [HL97], [LMR02], [LR00] and [KR10]). As outlined in [HL97], it is necessary to discretize the flux if the hyperbolic part with a high enough order, once again to keep the exact balance between the small scales. On the other hand, some schemes relies on the tracking of the nonclassical shocks (see for example [CL03], [CG08], [BCLL08], [Per11]). In that case the kinetic relations (6.56) are taken into account the scheme, for example through the use of an exact nonclassical Riemann solver. The discontinuous reconstruction strategy can be applied to construct a fully conservative scheme that approximates nonclassical shocks exactly. It extends to systems the scheme presented in [BCLL08].

6.8.2 The Riemann problem for the nonlinear elasticity model

In this section we briefly recall the results of Thanh and LeFloch [LT01] about the Riemann problem for System (6.46). We adopt the notation of this paper.

If the left state (v_L, w_L) and the right state (v_R, w_R) are linked by a 1-shock, its speed is equal to $-s(w_L, w_R)$, where

$$s(w_L, w_R) = \sqrt{\frac{\sigma(w_L) - \sigma(w_R)}{w_L - w_R}}, \tag{6.54}$$

and v_R is given by

$$v_R := \mathcal{H}_1(w_R, v_L, w_L) = v_L - s(w_L, w_R)(w_R - w_L). \tag{6.55}$$

The 2-shocks have the speed $s(w_L, w_R)$ defined in (6.54) and v_L can be parametrized as a function $\mathcal{H}^2$ of w_L, v_R and w_R. The 1-shock verifies the entropy inequality (6.51) if and only if

$$w_R^2 \leq w_L w_R \quad \text{or} \quad w_R w_L \leq \Phi^\flat_\infty(w_L) w_R,$$

where $\Phi^\flat_\infty(w_L)$ is the real having the opposite sign than w_L for which the dissipation of entropy (6.52) vanishes. There exists a real $\Phi^{-\sharp}(w_L)$ such that

$$w_L \Phi^{-\sharp}(w_L) \leq w_L \Phi^\flat_\infty(w_L),$$

and such that if w_R in $|\Phi^\flat_\infty(w_L), \Phi^{-\sharp}(w_L)|$, the shock between (v_L, w_L) and (v_R, w_R) is not admissible in the sense of Liu. If nothing is added to (6.51), some Riemann problems admit an infinity of solutions (see [LT01], Section 3). Uniqueness can be obtained by imposing a *kinetic relation* that strongly constraint the nonclassical shocks. It states that

$$\begin{cases} w_L = \Phi^{\flat,1}(w_R) & \text{for a 1-nonclassical shock,} \\ w_R = \Phi^{\flat,2}(w_L) & \text{for a 2-nonclassical shock,} \end{cases} \tag{6.56}$$

where $\Phi^{\flat,1}$ et $\Phi^{\flat,2}$ both verify

$$w\Phi^\flat_\infty(w) \leq w\Phi^{\flat,i}(w) \leq w\Phi^\sharp(w).$$

We recall the main result of [LT01].

Theorem 6.8. *Thanh, LeFloch If the kinetic function $\Phi^{\flat,1}$ and $\Phi^{\flat,2}$ are monotone decreasing, the Riemann problem for System (6.46-6.51-6.56) has a unique selfsimilar solution.*

The solution is explicitly constructed. This exact Riemann solver is the foundation of the scheme presented in the next section.

6.8.3 The scheme

In the sequel we call *phase transition* a shock in which w changes sign. It can be either nonclassical or classical.

Lemma 6.9. *If the states (v_L, w_L) and (v_R, w_R) are linked by a 1-phase transition, then*

$$w_L w_R < 0 \ \textit{ and } (w_L - w_R)(v_L - v_R) > 0. \tag{6.57}$$

If they are linked by a 2-phase transition, then

$$w_L w_R < 0 \ \textit{ and } (w_L - w_R)(v_L - v_R) < 0.$$

Proof. It is a straightforward consequence of the Rankine–Hugoniot relations. In the case of nonlinear elastodynamic (6.46), they write

$$\begin{cases} s(v_L - v_R) = \sigma(w_R) - \sigma(w_L); \\ s(w_L - w_R) = v_R - v_L. \end{cases}$$

The second line yields the result as the speed shock s is negative for the 1-shocks and positive for the 2-shocks. ☐

Similarly, we obtain the following characterization of classical shocks in which w does not change sign.

Lemma 6.10. *If the states (v_L, w_L) and (v_R, w_R) are linked by a 1-shock in which w does not change sign, then either*

$$w_R < w_L \leq 0 \text{ and } v_R < v_L,$$

or

$$0 \leq w_L < w_R \text{ and } v_R > v_L.$$

If they are linked by a 2-shock in which w does not change sign, then either

$$w_L < w_R \leq 0 \text{ and } v_R < v_L,$$

or

$$0 \leq w_R < w_L \text{ and } v_R > v_L.$$

Remark 24. The fact that (v_L, w_L) and (v_R, w_R) verify (6.57) does not imply that there is a 1-transition phase in the Riemann problem between (v_L, w_L) and (v_R, w_R). This is for example the case of the Riemann problem

$$v_L = -1,\ w_L = 1,\ v_R = -2 \text{ and } w_R = -0.1$$

which verifies (6.57) but whose solution contains a 1-rarefaction wave and a 2-classical shock, as depicted on Figure 6.38 below.

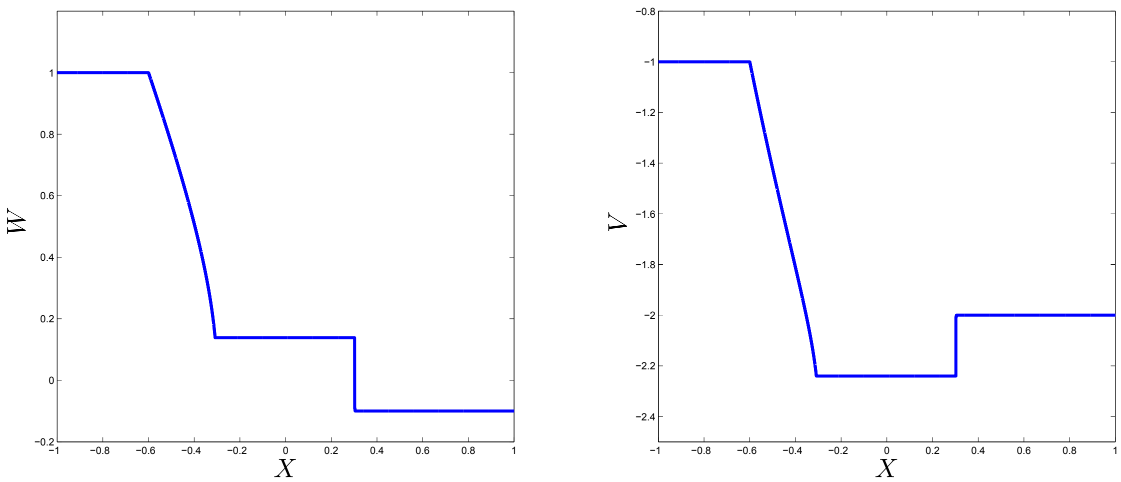


Figure 6.38. *Example of a Riemann problem that verifies* (6.57) *but that does not contain a* 1*-transition phase.*

Consider three states $U_L = (w_L, v_L)$, $U_M = (w_M, v_M)$ and $U_R = (w_R, v_R)$, and denote by $U_{M,\star}$ the intersection point between the 1-nonclassical and the 2-nonclassical Lax curves in the Riemann problem between U_L and U_R. Lemma (6.9) yields the following definition of the candidate reconstruction operator

$$(\bar{U}_{M,L}, \bar{U}_{M,L}) = \mathcal{CR}(U_L, U_M, U_R).$$

- If $w_L w_R < 0$, $(w_L - w_R)(v_L - v_R) > 0$, and $w_L w_{L,\star} < 0$, the solution contains a 1-phase transition.
  - If the 1-phase transition is a nonclassical shock, it is preceded by a classical wave and we set

$$\begin{cases} (\bar{v}_{M,L}, \bar{w}_{M,L}) & = (\mathcal{H}_1(\Phi^{\flat,1}(w_{M,\star}), v_{M,\star}, w_{M,\star}), \Phi^{\flat,1}(w_{M,\star})), \\ (\bar{v}_{M,R}, \bar{w}_{M,R}) & = (v_{M,\star}, w_{M,\star}). \end{cases}$$

- If the 1-phase transition is a classical shock, it is the only contribution in the 1-wave and we set
$$\begin{cases}(\bar v_{M,L}, \bar w_{M,L}) & = (v_L, w_L),\\ (\bar v_{M,R}, \bar w_{M,R}) & = (v_{M,\star}, w_{M,\star}).\end{cases}$$

- If $w_L w_R < 0$, $(w_L - w_R)(v_L - v_R) < 0$, and $w_R w_{M,\star} < 0$, the solution contains a 2-phase transition.
  - If the 2-phase transition is a nonclassical shock, it is followed by a classical wave and we set
$$\begin{cases}(\bar v_{M,L}, \bar w_{M,L}) & = (v_{M,\star}, w_{M,\star}),\\ (\bar v_{M,R}, \bar w_{M,R}) & = (\mathcal{H}_2(\Phi^{\flat,2}(w_{M,\star}), v_{M,\star}, w_{M,\star}), \Phi^{\flat,2}(w_{M,\star})).\end{cases}$$
  - If the 2-phase transition is a classical shock, it is the only contribution in the 2-wave and we set
$$\begin{cases}(\bar v_{M,L}, \bar w_{M,L}) & = (v_{M,\star}, w_{M,\star}),\\ (\bar v_{M,R}, \bar w_{M,R}) & = (v_R, w_R).\end{cases}$$

- In the other cases, we do not detect any relevant phase transition and we set
$$\begin{cases}(\bar v_{M,L}, \bar w_{M,L}) & = (v_M, w_M),\\ (\bar v_{M,R}, \bar w_{M,R}) & = (v_M, w_M).\end{cases}$$

With Lemma (6.10) we can extend "for free" the candidate detection criterion $\mathcal{CR}$ to reconstruct also shocks in which w does not change sign. Those shocks are classical.

- If either ($w_R < w_L \le 0$ and $v_R < v_L$) or ($0 \le w_L < w_R$ and $v_R > v_L$), and if $w_L w_{M,\star} > 0$, the 1-wave is just a 1-shock and we set
$$\begin{cases}(\bar v_{M,L}, \bar w_{M,L}) & = (v_L, w_L),\\ (\bar v_{M,R}, \bar w_{M,R}) & = (v_{M,\star}, w_{M,\star}).\end{cases}$$

- If either ($w_L < w_R \le 0$ and $v_R < v_L$) or ($0 \le w_R < w_L$ and $v_R > v_L$), and if $w_R w_{M,\star} > 0$, the 2-wave is just a 2-shock and we set
$$\begin{cases}(\bar v_{M,L}, \bar w_{M,L}) & = (v_{M,\star}, w_{M,\star}),\\ (\bar v_{M,R}, \bar w_{M,R}) & = (v_R, w_R).\end{cases}$$

We accept the reconstruction if both v and w can be reconstructed in a conservative manner:
$$\mathcal{R}(U_{j-1}^n, U_j^n, U_{j+1}^n) = \begin{cases}\mathcal{CR}(U_{j-1}^n, U_j^n, U_{j+1}^n) \text{ if } d_j^{n,v}, d_j^{n,w} \in [0, \Delta x],\\ (U_j^n, U_j^n) \text{ otherwise.}\end{cases}$$

Remark 25. The reconstructed operator $\mathcal{R}$ we just presented relies on the exact resolution of the Riemann problem. In [CCER12], the authors construct an approximate Riemann solver which is exact on isolated nonclassical shocks in the hyperbolic-elliptic case where p decreases on an interval. This type of solvers could be used to avoid the exact resolution of Riemann problems, in the spirit of Section 6.6.

6.8.4 Exact approximation of isolated nonclassical shocks

Theorem 6.11. *The discontinuous reconstruction scheme applied to the elastodynamic System* (6.46) *is exact when the initial data is an isolated nonclassical shock.*

Proof. We verify the three assumptions of Theorem 6.3. The proof repeatedly uses the monotonicity of w and v through shocks given by Lemmas 6.9 and 6.10. We prove the result for a 1-nonclassical shock such that $w_L < 0 < w_R$, the other cases being exactly similar. Let U_L and U_R be two states linked by a 1-nonclassical shock: $w_L = \Phi^{\flat,1}(w_R)$ by (6.56), and $v_L = \mathcal{H}^1(\Phi^{\flat,1}(w_R), v_R, w_R)$ by (6.55). Moreover by Lemma 6.9 gives that v_L is smaller than v_R. Let α be a real in $[0,1]$ and $U_M = \alpha U_L + (1-\alpha)U_R$. A 1-transition phase is detected between U_L and U_R and we have

$$\mathcal{R}(U_L, U_M, U_R) = (U_L, U_R).$$

Suppose now that α belongs to $(0,1)$. Our aims is to prove the last two hypothesis of Theorem 6.3.

- If w_M is positive, we have $w_L < 0 < w_M < w_R$ and $v_L < v_M < v_R$. It yields the following detections.
  - A 1-classical shock is detected between U_M and U_R. The left candidate reconstructed state is U_M, and the right candidate reconstructed state is the intermediate state in the Riemann problem between U_M and U_R, we denote it by U_*:

    $$\mathcal{CR}(U_M, U_R, U_R) = (U_M, U_*).$$

    To reconstruct w in a conservative manner, it is necessary to have $w_* > w_R$. It yields that U_* and U_R are linked by a 2-classical shock, and thus that $v_R > v_*$. Thus it is not possible to reconstruct v in a conservative manner and $\mathcal{R}(U_M, U_R, U_R) = (U_R, U_R)$.
  - a 1-phase transition is detected between U_L and U_M. Let us denote by U_* the intermediate state appearing in the solution of Riemann problem between U_L and U_M. If U_L and U_* are linked by a 1-classical shock, we will have

    $$\mathcal{CR}(U_L, U_L, U_M) = (U_L, U_*)$$

    and consequently, $d_L^U = (\Delta x, \Delta x)$. Thus we have $\mathcal{R}(U_L, U_L, U_M) = (U_L, U_L)$. We now suppose that U_L and U_* are linked by 1-classical shock followed by a 1-nonclassical shock:

    $$\mathcal{CR}(U_L, U_L, U_M) = (\bar{U}_{L,L}, U_*),$$

    with $\bar{w}_{L,L} = \Phi^{\flat,1}(w_*)$ and $\bar{v}_{L,L} = \mathcal{H}^1(\bar{w}_{L,L}, v_*, w_*)$. It is not possible to reconstruct w in a conservative manner if $w_L \leq \bar{w}_{L,L}$. Suppose by contradiction that $w_L > \bar{w}_{L,L}$: U_L and $\bar{U}_{L,L}$ are linked by a 1-classical shock. The allure of w is depicted on Figure 6.39, which also recall the notation of Definition 6.1. As the kinetic function $\phi^{\flat,1}$ decreases,

    $$\bar{w}_{L,R} = \phi^{-\flat,1}(\bar{w}_{L,L}) > \phi^{-\flat,1}(w_L) = w_R$$

    and there is a classical shock between U_* and U_R. The states U_* and U_R are both on the 1-wave curve. Theorem 4.5 of [LT01] states that along this curve, v is an increasing function of w. Therefore $\bar{v}_{L,R}$ is larger than v_R. On the other hand, the Rankine–Hugoniot relations applied to the 2-shock yields that $\bar{v}_{L,R}$ is smaller than v_M, which contradict the fact that $v_M < v_R$.
- If w_M is negative, we have $w_L < w_M \leq 0 < w_R$ and $v_L < v_M < v_R$. This case is much simpler.
  - Nothing is detected between U_L and U_M:

    $$\mathcal{CR}(U_L, U_L, U_M) = \mathcal{C}(U_L, U_L, U_M) = (U_L, U_L).$$

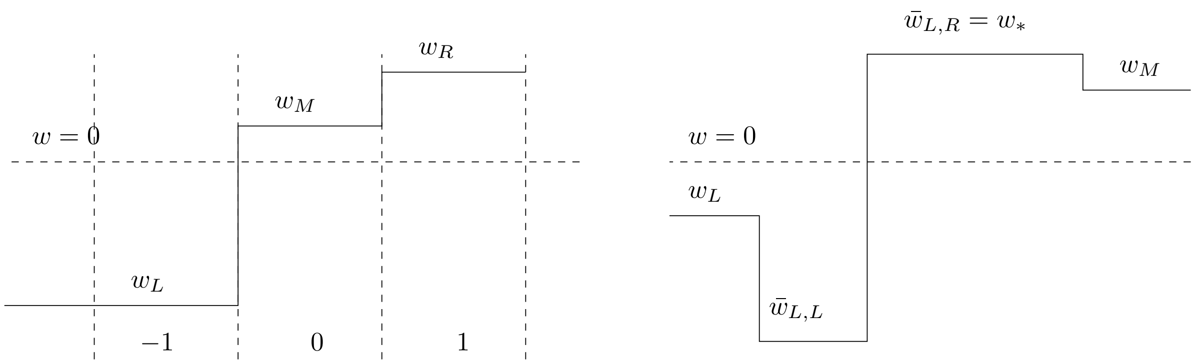


Figure 6.39. *Structure of w when a 1-phase transition is detected in cell -1. In that case, the Riemann solution (on the right) contains three shocks.*

– A 1-phase transition is detected between U_M and U_R. We denote by U_* the intermediate state in the Riemann problem. We have

$$\mathcal{CR}(U_M, U_R, U_R) = (\bar{U}_{R,L}, U_*).$$

It is not possible to reconstruct w in a conservative manner if $w_* \leq w_R$. On the other hand, if $w_* > w_R$, there is a 2-classical shock between U_* and U_R, thus $v_* \leq v_R$. Moreover through the 1-phase transition, $\bar{v}_{L,L} \leq v_\star$. Thus it is impossible to reconstruct v in a conservative manner.

□

Remark 26. It is crucial to require that both v and w are reconstructed in a conservative manner. For example, the Riemann problem

$$w_L = -6,\ v_L = -10,\ w_R = 9 \text{ and } v_R = 110$$

corresponds to an isolated 1-nonclassical shock for the stress $\sigma(w) = w^3 + w$ and the kinetic function $\phi^{\flat,1}(w) = -\frac{2}{3}w$, but $\mathcal{R}(U_L, U_L, U_M) \neq (U_L, U_L)$ with $\alpha = 0.2$ is true if we only impose that the reconstruction on v is conservative, while $\mathcal{R}(U_M, U_R, U_R) \neq (u_R, U_R)$ with $\alpha = 0.6$ is true if we only impose that the reconstruction on w is conservative.

6.8.5 Numerical Simulations

For all the numerical simulations presented below, $m = 1$ and the kinetic functions are $\Phi^{\flat,1}(w) = \Phi^{\flat,2}(w) = -\frac{2}{3}w$.

Isolated nonclassical shock

This test case illustrates Theorem 6.11. The initial data is:

$$\begin{cases} v^0(x) = -10 * \mathbf{1}_{x<0} + 100 * \mathbf{1}_{x>0} \\ w^0(x) = -6 * \mathbf{1}_{x<0} + 9 * \mathbf{1}_{x>0} \end{cases}$$

On the top if Figure 6.40, we plot the exact nonclassical solution and the solution given by the reconstruction scheme. As expected they are exactly the same. On the bottom of Figure 6.40, we plot the solution given by the Godunov scheme based on the exact *nonclassical* Riemann solver. It appears that is does not capture the nonclassical solution but the classical one, which in that case is a rarefaction

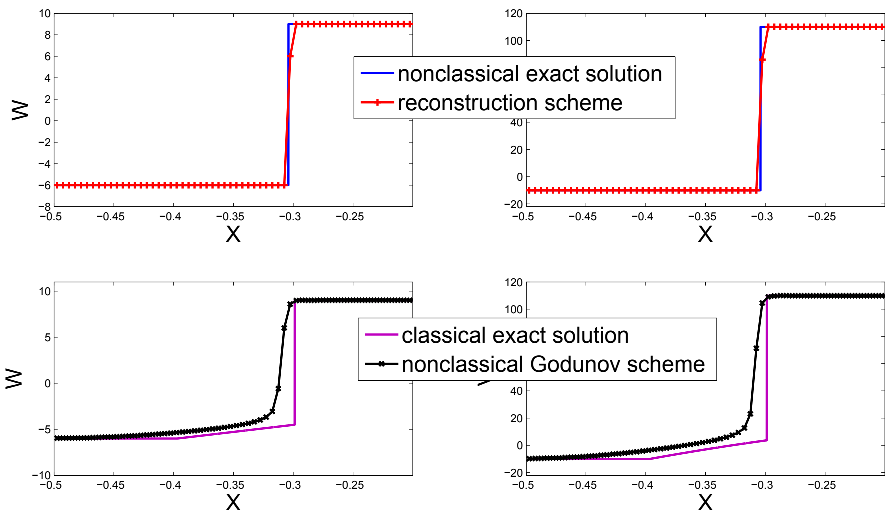


Figure 6.40. *Usual schemes are unable to capture nonclassical shocks. On the contrary, the reconstruction scheme captures isolated nonclassical shocks exactly.*

followed by a shock. This is a general phenomenon: usual finite volume scheme are not able to capture nonclassical solutions. Thus in the sequel, we only have the Glimm scheme to compare our scheme with. The Courant number is set to 0.45, the final time is $T = 0.038$ and the mesh has 200 cells.

A Riemann problem with two nonclassical shocks

The initial data is now

$$\begin{cases} v^0(x) = 1 * \mathbf{1}_{x<0} + 2 * \mathbf{1}_{x>0}, \\ w^0(x) = 6 * \mathbf{1}_{x<0} - 10 * \mathbf{1}_{x>0}. \end{cases}$$

The exact solution consists in a 1-classical shock, a 1-nonclassical shock, a 2-nonclassical shock and a 2-rarefaction wave. On Figure 6.41, we plot the exact solution and the numerical solution given by the reconstruction scheme. We compare the reconstruction scheme with the detection of the phase transitions only (with Lemma 6.9) and the reconstruction scheme with the detection of phase transition and the classical shocks (with Lemma 6.10). Both of them capture very sharply the wave transitions; the 1-classical shock is much more diffused if it is not reconstructed, and both scheme behaves in the same way in the rarefaction wave.

Perturbation of a classical shock

This test case is taken from [CL03]. The initial data is

$$\begin{cases} v^0(x) = 1 * \mathbf{1}_{x<0} - 11 * \mathbf{1}_{x>0}, \\ w^0(x) = (1+\varepsilon) * \mathbf{1}_{x<0} - 3 * \mathbf{1}_{x>0}. \end{cases}$$

with $\varepsilon \in \{0, 0.05, 0.1\}$ When $\varepsilon = 0$, is corresponds to a classical shock. When $\varepsilon > 0$, the classical shock is split in a classical shock followed by a nonclassical shock. Their speeds are close to each other. We plot the solutions at time $T = 0.4$ on Figure 6.42. The solutions are very well approximated when $\varepsilon > 0$. When $\varepsilon = 0$, a spike appears. This spike corresponds to the state linked to (v_R, w_R) by a nonclassical shock. This shock is immediately caught up by the classical shock behind (v_L, w_R) and the spike, so

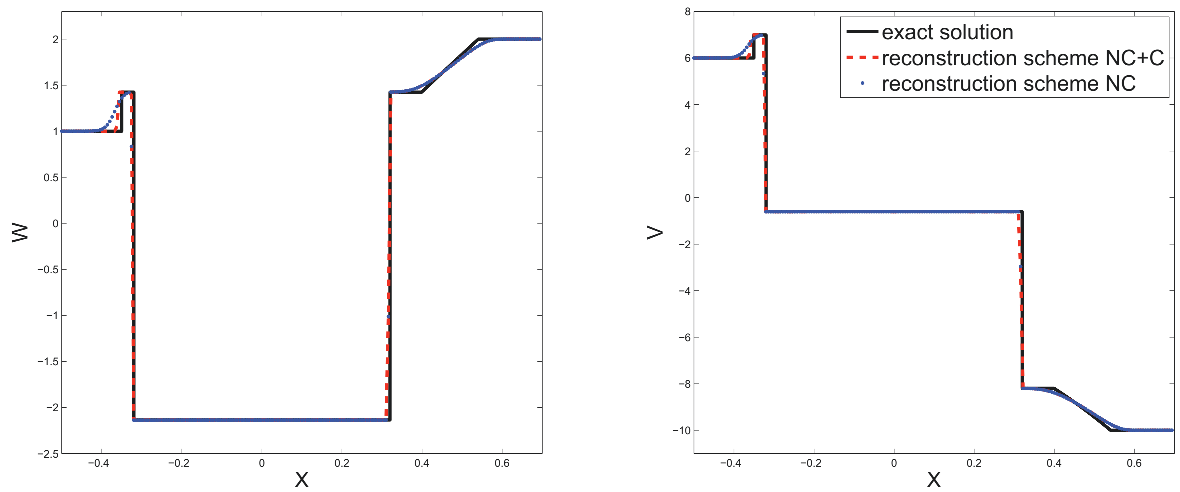


Figure 6.41. *A Riemann problem with four waves at time* 0.15*. The Courant number is* 0.45*, the mesh has* 200 *cells.*

the spike is only 1 cell wide. Il illustrates that the reconstruction scheme is not able to approach exactly classical shocks when w changes sign. We could add other conditions in the detection criterion $\mathcal{R}$ to

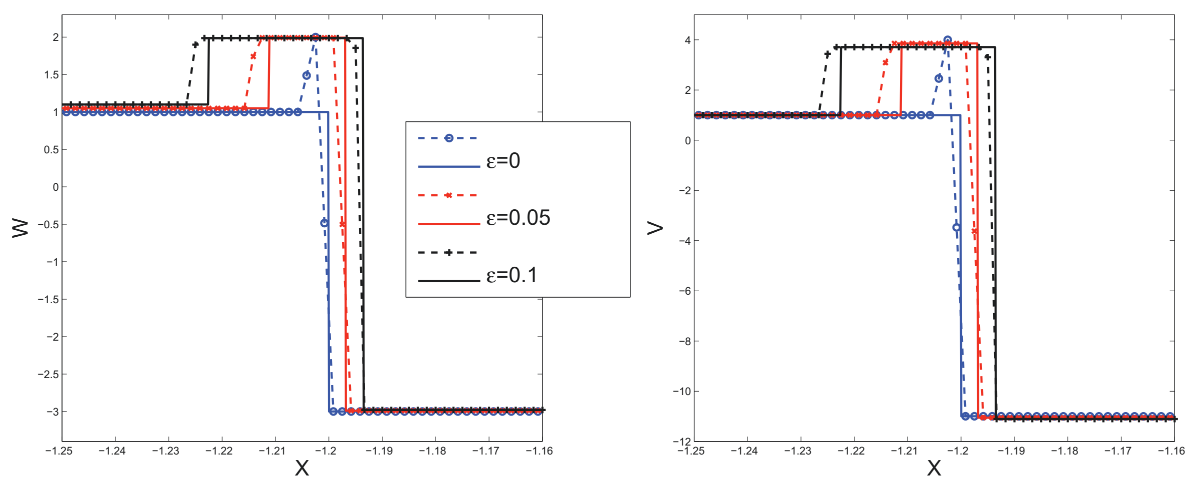


Figure 6.42. *Perturbations of an isolated classical shock. Dashed lines: numerical solutions; plain lines: exact solutions.*

make the scheme exact on isolated classical shock where w changes sign. A possibility is to ask for monotonicity on three adjacent cells, for example to cancel the reconstruction if

$$U_M \notin [\min(U_L, U_R), \max(U_L, U_R)].$$

However, this criterion is not robust toward perturbation and the spike reappears if we slightly change the initial data.

Liste des figures

Liste des figures

Références

[AC14] Nina Aguillon and Christophe Chalons. Non diffusive schemes based on approximate Riemann solvers for Lagrangian gas dynamics. *In preparation*, 2014.

[ADVLCL08] François Alouges, Florian De Vuyst, Gérard Le Coq, and Emmanuel Lorin. The reservoir technique: a way to make Godunov-type schemes zero or very low diffuse. Application to Colella-Glaz solver. *Eur. J. Mech. B Fluids*, 27(6):643–664, 2008.

[Agu14a] Nina Aguillon. A non-dissipative reconstruction scheme for the compressible Euler equations. *Preprint, arXiv:1403.7497*, 2014.

[Agu14b] Nina Aguillon. Numerical simulations of a fluid-particle coupling. In *Finite Volumes for Complex Applications VII-Elliptic, Parabolic and Hyperbolic Problems*, pages 759–767. Springer Proceedings in Mathematics & Statistics, 2014.

[Agu14c] Nina Aguillon. Riemann problem for a particle-fluid coupling. *Math. Models Methods Appl. Sci.*, 2014.

[AK91] Rohan Abeyaratne and James K. Knowles. Kinetic relations and the propagation of phase boundaries in solids. *Arch. Rational Mech. Anal.*, 114(2):119–154, 1991.

[AK01] Rémi Abgrall and Smadar Karni. Computations of compressible multifluids. *J. Comput. Phys.*, 169(2):594–623, 2001.

[AKR10] Boris Andreianov, Kenneth H. Karlsen, and Nils H. Risebro. On vanishing viscosity approximation of conservation laws with discontinuous flux. *Netw. Heterog. Media*, 5(3):617–633, 2010.

[AKR11] Boris Andreianov, Kenneth Hvistendahl Karlsen, and Nils Henrik Risebro. A theory of L^1-dissipative solvers for scalar conservation laws with discontinuous flux. *Arch. Ration. Mech. Anal.*, 201, 2011.

[ALS14] Nina Aguillon, Frédéric Lagoutière, and Nicolas Seguin. Convergence of finite volumes schemes for the coupling between the inviscid Burgers equation and a particle. *in preparation*, 2014.

[ALST10] Boris Andreianov, Frédéric Lagoutière, Nicolas Seguin, and Takéo Takahashi. Small solids in an inviscid fluid. *Netw. Heterog. Media*, 5(3):385–404, 2010.

[ALST13] Boris Andreianov, Frédéric Lagoutière, Nicolas Seguin, and Takéo Takahashi. Well-posedness for a one-dimensional fluid-particle interaction model. *To appear in SIAM J. Appl. Math.*, 5, 2013.

[AR97] Mohit Arora and Philip L. Roe. On postshock oscillations due to shock capturing schemes in unsteady flows. *J. Comput. Phys.*, 130(1):25–40, 1997.

[AS12] Boris Andreianov and Nicolas Seguin. Analysis of a Burgers equation with singular resonant source term and convergence of well-balanced schemes. *Discrete Contin. Dyn. Syst.*, 32(6):1939–1964, 2012.

[BCG10] Raul Borsche, Rinaldo M. Colombo, and Mauro Garavello. On the coupling of systems of hyperbolic conservation laws with ordinary differential equations. *Nonlinearity*, 23(11):2749–2770, 2010.

[BCG12] Raul Borsche, Rinaldo M. Colombo, and Mauro Garavello. Mixed systems: ODEs - balance laws. *J. Differential Equations*, 252(3):2311–2338, 2012.

[BCG14] Raul Borsche, Rinaldo M. Colombo, and Mauro Garavello. On the interactions between a solid body and a compressible inviscid fluid. *To appear in Interfaces Free Bound.*, 2014.

[BCL13] Benjamin Boutin, Frédéric Coquel, and Philippe G. LeFloch. Coupling techniques for nonlinear hyperbolic equations. III. The well-balanced approximation of thick interfaces. *SIAM J. Numer. Anal.*, 51, 2013.

[BCLL08] Benjamin Boutin, Christophe Chalons, Frédéric Lagoutière, and Philippe G. LeFloch. Convergent and conservative schemes for nonclassical solutions based on kinetic relations. I. *Interfaces Free Bound.*, 10(3):399–421, 2008.

[BD06] Céline Baranger and Laurent Desvillettes. Coupling Euler and Vlasov equations in the context of sprays: the local-in-time, classical solutions. *J. Hyperbolic Differ. Equ.*, 3(1):1–26, 2006.

[BdL09] François Bouchut and Tomás Morales de Luna. Semi-discrete entropy satisfying approximate Riemann solvers. The case of the Suliciu relaxation approximation. *J. Sci. Comput.*, 41(3):483–509, 2009.

[Ber99] Christophe Berthon. *Contribution à l'analyse numérique des équations de Navier–Stokes compressibles à deux entropies spécifiques. Applications à la turbulence compressible*. PhD thesis, Université Paris VI, 1999.

[BFBC+11] Marie Billaud Friess, Benjamin Boutin, Filipa Caetano, Gloria Faccanoni, Samuel Kokh, Frédéric Lagoutière, and Laurent Navoret. A second order anti-diffusive Lagrange-remap scheme for two-component flows. In *CEMRACS'10 research achievements: numerical modeling of fusion*, volume 32 of *ESAIM Proc.*, pages 149–162. EDP Sci., Les Ulis, 2011.

[BLN79] Claude Bardos, Alain-Yves LeRoux, and Jean-Claude Nédélec. First order quasilinear equations with boundary conditions. *Comm. Partial Differential Equations*, 4(9):1017–1034, 1979.

[Bou04a] François Bouchut. An antidiffusive entropy scheme for monotone scalar conservation laws. *J. Sci. Comput.*, 21(1):1–30, 2004.

[Bou04b] François Bouchut. A reduced stability condition for nonlinear relaxation to conservation laws. *J. Hyperbolic Differ. Equ.*, 1(1):149–170, 2004.

[BST12] Muriel Boulakia, Erica Schwindt, and Takéo Takahashi. Existence of strong solutions for the motion of an elastic structure in an incompressible viscous fluid. *Interfaces Free Bound.*, 14(3):273–306, 2012.

[CC06] Christophe Chalons and Frédéric Coquel. Euler equations with several independent pressure laws and entropy satisfying explicit projection schemes. *Math. Models Methods Appl. Sci.*, 16(9):1469–1504, 2006.

[CC13] Christophe Chalons and Frédéric Coquel. Modified suliciu relaxation system and exact resolution of isolated shock waves. *Math. Models Methods Appl. Sci.*, 2013.

[CCER12] Christophe Chalons, Frédéric Coquel, Patrick Engel, and Christian Rohde. Fast relaxation solvers for hyperbolic-elliptic phase transition problems. *SIAM J. Sci. Comput.*, 34(3):A1753–A1776, 2012.

[CDEG05] Antonin Chambolle, Benoît Desjardins, Maria J. Esteban, and Céline Grandmont. Existence of weak solutions for the unsteady interaction of a viscous fluid with an elastic plate. *J. Math. Fluid Mech.*, 7(3):368–404, 2005.

[CG08] Christophe Chalons and Paola Goatin. Godunov scheme and sampling technique for computing phase transitions in traffic flow modeling. *Interfaces Free Bound.*, 10(2):197–221, 2008.

[CG10] Rinaldo M. Colombo and Graziano Guerra. On general balance laws with boundary. *J. Differential Equations*, 248(5):1017–1043, 2010.

[CL03] Christophe Chalons and Philippe G. LeFloch. Computing undercompressive waves with the random choice scheme. Nonclassical shock waves. *Interfaces Free Bound.*, 5(2):129–158, 2003.

[CLL94] Gui Qiang Chen, C. David Levermore, and Tai-Ping Liu. Hyperbolic conservation laws with stiff relaxation terms and entropy. *Comm. Pure Appl. Math.*, 47(6):787–830, 1994.

[CLS04] Ashwin Chinnayya, Alain-Yves LeRoux, and Nicolas Seguin. A well-balanced numerical scheme for the approximation of the shallow-water equations with topography: the resonance phenomenon. *Int. J. Finite Vol.*, 1(1):33, 2004.

[CT80] Michael G. Crandall and Luc Tartar. Some relations between nonexpansive and order preserving mappings. *Proc. Amer. Math. Soc.*, 78(3):385–390, 1980.

[DE99] Benoît Desjardins and Maria J Esteban. Existence of weak solutions for the motion of rigid bodies in a viscous fluid. *Arch. Ration. Mech. Anal.*, 146(1):59–71, 1999.

[DE00] Benoît Desjardins and Maria J. Esteban. On weak solutions for fluid-rigid structure interaction: compressible and incompressible models. *Comm. Partial Differential Equations*, 25(7-8):1399–1413, 2000.

[DL88] François Dubois and Philippe G. LeFloch. Boundary conditions for nonlinear hyperbolic systems of conservation laws. *J. Differential Equations*, 71(1):93–122, 1988.

[DL01] Bruno Després and Frédéric Lagoutière. Contact discontinuity capturing schemes for linear advection and compressible gas dynamics. *J. Sci. Comput.*, 16(4):479–524 (2002), 2001.

[DL07] Bruno Després and Frédéric Lagoutière. Numerical resolution of a two-component compressible fluid model with interfaces. *Prog. Comput. Fluid Dyn.*, 7(6):295–310, 2007.

[DM96] Rosa Donat and Antonio Marquina. Capturing shock reflections: an improved flux formula. *J. Comput. Phys.*, 125(1):42–58, 1996.

[DMLM95] Gianni Dal Maso, Philippe G. Lefloch, and François Murat. Definition and weak stability of nonconservative products. *J. Math. Pures Appl. (9)*, 74(6):483–548, 1995.

[Dom02] Komla Domelevo. Well-posedness of a kinetic model of dispersed two-phase flow with point-particles and stability of travelling waves. *Discrete Contin. Dyn. Syst. Ser. B*, 2(4):591–607, 2002.

[Dub01] François Dubois. Partial Riemann problem, boundary conditions, and gas dynamics. In *Absorbing boundaries and layers, domain decomposition methods*, pages 16–77. Nova Sci. Publ., Huntington, NY, 2001.

[Eva10] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[FAMO99] Ronald P. Fedkiw, Tariq Aslam, Barry Merriman, and Stanley Osher. A non-oscillatory Eulerian approach to interfaces in multimaterial flows (the ghost fluid method). *J. Comput. Phys.*, 152(2):457–492, 1999.

[Fei03a] Eduard Feireisl. On the motion of rigid bodies in a viscous compressible fluid. *Arch. Ration. Mech. Anal.*, 167(4):281–308, 2003.

[Fei03b] Eduard Feireisl. On the motion of rigid bodies in a viscous incompressible fluid. *J. Evol. Equ.*, 3(3):419–441, 2003. Dedicated to Philippe Bénilan.

[GL96] Joshua M Greenberg and Alain-Yves LeRoux. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM J. Numer. Anal.*, 33(1):1–16, 1996.

[GL04] Paola Goatin and Philippe G. LeFloch. The Riemann problem for a class of resonant hyperbolic systems of balance laws. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 21(6):881–902, 2004.

[GLS00] Max D. Gunzburger, Hyung-Chun Lee, and Gregory A. Seregin. Global existence of weak solutions for viscous incompressible flows around a moving rigid body in three dimensions. *J. Math. Fluid Mech.*, 2(3):219–266, 2000.

[GLT13] Thierry Goudon, Frédéric Lagoutière, and Léon Matar Tine. Simulations of the Lifshitz-Slyozov equations: the role of coagulation terms in the asymptotic behavior. *Math. Models Methods Appl. Sci.*, 23(7):1177–1215, 2013.

[GR96] Edwige Godlewski and Pierre-Arnaud Raviart. *Numerical approximation of hyperbolic systems of conservation laws*, volume 118 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1996.

[Har84] Ami Harten. On a class of high resolution total-variation-stable finite-difference schemes. *SIAM J. Numer. Anal.*, 21(1):1–23, 1984. With an appendix by Peter D. Lax.

[Har89] Ami Harten. ENO schemes with subcell resolution. *J. Comput. Phys.*, 83(1):148–184, 1989.

[Hil07a] Matthieu Hillairet. Do Navier-Stokes equations enable to predict contact between immersed solid particles? In *Analysis and simulation of fluid dynamics*, Adv. Math. Fluid Mech., pages 109–127. Birkhäuser, Basel, 2007.

[Hil07b] Matthieu Hillairet. Lack of collision between solid bodies in a 2D incompressible viscous flow. *Comm. Partial Differential Equations*, 32(7-9):1345–1371, 2007.

[HL97] Brian T. Hayes and Philippe G. LeFloch. Non-classical shocks and kinetic relations: scalar conservation laws. *Arch. Rational Mech. Anal.*, 139(1):1–56, 1997.

[HL98] Brian T. Hayes and Philippe G. Lefloch. Nonclassical shocks and kinetic relations: finite difference schemes. *SIAM J. Numer. Anal.*, 35(6):2169–2194 (electronic), 1998.

[HL00] Brian T. Hayes and Philippe G. Lefloch. Nonclassical shocks and kinetic relations: strictly hyperbolic systems. *SIAM J. Math. Anal.*, 31(5), 2000.

[HT09] Matthieu Hillairet and Takéo Takahashi. Collisions in three-dimensional fluid structure interaction problems. *SIAM J. Math. Anal.*, 40(6):2451–2477, 2009.

[IT92] Eli Isaacson and Blake Temple. Nonlinear resonance in systems of conservation laws. *SIAM J. Appl. Math.*, 52(5):1260–1278, 1992.

[JL96] Shi Jin and Jian-Guo Liu. The effects of numerical viscosities. I. Slowly moving shocks. *J. Comput. Phys.*, 126(2):373–389, 1996.

[JMS95] Doug Jacobs, Bill McKinney, and Michael Shearer. Travelling wave solutions of the modified Korteweg-de Vries-Burgers equation. *J. Differential Equations*, 116(2):448–467, 1995.

[JX95] Shi Jin and Zhou Ping Xin. The relaxation schemes for systems of conservation laws in arbitrary space dimensions. *Comm. Pure Appl. Math.*, 48(3):235–276, 1995.

[KR10] Frederike Kissling and Christian Rohde. The computation of nonclassical shock waves with a heterogeneous multiscale method. *Netw. Heterog. Media*, 5(3):661–674, 2010.

[Kru70] Stanislav Nikolaevich Kruzhkov. First-order quasilinear equations in several independent variables. *Math. USSR Sb.*, 10:217–243, 1970.

[Lag99] Frédéric Lagoutiere. Numerical resolution of scalar convex equations: explicit stability, entropy and convergence conditions. In *CEMRACS 1999 (Orsay)*, volume 10 of *ESAIM Proc.*, pages 183–199 (electronic). Soc. Math. Appl. Indust., Paris, 1999.

[Lag08] Frédéric Lagoutière. Stability of reconstruction schemes for scalar hyperbolic conservation laws. *Commun. Math. Sci.*, 6(1):57–70, 2008.

[LagXX] Frédéric Lagoutière. Non-dissipative entropy satisfying discontinuous reconstruction schemes for hyperbolic conservation laws. *Preprint*, XX.

[Lax57] Peter D. Lax. Hyperbolic systems of conservation laws. II. *Comm. Pure Appl. Math.*, 10:537–566, 1957.

[LeF02] Philippe G. LeFloch. *Hyperbolic systems of conservation laws*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2002. The theory of classical and nonclassical shock waves.

[LeR77] Alain-Yves LeRoux. A numerical conception of entropy for quasi-linear equations. *Math. Comp.*, 31(140):848–872, 1977.

[Liu74] Tai Ping Liu. The Riemann problem for general 2×2 conservation laws. *Trans. Amer. Math. Soc.*, 199:89–112, 1974.

[LMR02] Philippe. G. Lefloch, Jean-Marc Mercier, and Christian Rohde. Fully discrete, entropy conservative schemes of arbitrary order. *SIAM J. Numer. Anal.*, 40(5):1968–1992 (electronic), 2002.

[LR00] Philippe G. Lefloch and Christian Rohde. High-order schemes, entropy inequalities, and nonclassical shocks. *SIAM J. Numer. Anal.*, 37(6):2023–2060 (electronic), 2000.

[LST08] Frédéric Lagoutière, Nicolas Seguin, and Takéo Takahashi. A simple 1D model of inviscid fluid-solid interaction. *J. Differential Equations*, 245(11):3503–3544, 2008.

[LT01] Philippe G. LeFloch and Mai Duc Thanh. Nonclassical Riemann solvers and kinetic relations. I. A nonconvex hyperbolic model of phase transitions. *Z. angew. Math. Phys.*, 52, 2001.

[LT03] Philippe G. Lefloch and Mai Duc Thanh. The Riemann problem for fluid flows in a nozzle with discontinuous cross-section. *Commun. Math. Sci.*, 1(4):763–797, 2003.

[LT07] Philippe G. LeFloch and Mai Duc Thanh. The Riemann problem for the shallow water equations with discontinuous topography. *Commun. Math. Sci.*, 5(4):865–885, 2007.

[LW03] Richard Liska and Burton Wendroff. Comparison of several difference schemes on 1D and 2D test problems for the Euler equations. *SIAM J. Sci. Comput.*, 25(3):995–1017 (electronic), 2003.

[LWM08] Hongxia Li, Zhigang Wang, and De-kang Mao. Numerically neither dissipative nor compressive scheme for linear advection equation and its application to the Euler system. *J. Sci. Comput.*, 36(3):285–331, 2008.

[Men94] Ralph Menikoff. Errors when shock waves interact due to numerical shock width. *SIAM J. Sci. Comput.*, 15(5):1227–1242, 1994.

[Noh87] William F Noh. Errors for calculations of strong shocks using an artificial viscosity and an artificial heat flux. *J. Comp. Phys.,*, 72(1):78–120, 1987.

[NT90] Haim Nessyahu and Eitan Tadmor. Nonoscillatory central differencing for hyperbolic conservation laws. *J. Comput. Phys.*, 87(2):408–463, 1990.

[Ole57] Olga Arsen'evna Oleĭnik. On the uniqueness of the generalized solution of the Cauchy problem for a non-linear system of equations occurring in mechanics. *Uspehi Mat. Nauk (N.S.)*, 12(6(78)):169–176, 1957.

[Pan07] Eugeny Yur'evich Panov. Existence of strong traces for quasi-solutions of multidimensional conservation laws. *J. Hyperbolic Differ. Equ.*, 4(4):729–770, 2007.

[Par06] Carlos Parés. Numerical methods for nonconservative hyperbolic systems: a theoretical framework. *SIAM J. Numer. Anal.*, 44, 2006.

[Per11] Vincent Perrier. A conservative method for the simulation of the isothermal Euler system with the van-der-Waals equation of state. *J. Sci. Comput.*, 48(1-3):296–303, 2011.

[Qui94] James J. Quirk. A contribution to the great Riemann solver debate. *Internat. J. Numer. Methods Fluids*, 18(6):555–574, 1994.

[Rob90] Thomas W. Roberts. The behavior of flux difference splitting schemes near slowly moving shock waves. *J. Comput. Phys.*, 90(1):141–160, 1990.

[Sle89] Marshall Slemrod. A limiting "viscosity" approach to the Riemann problem for materials exhibiting change of phase. *Arch. Rational Mech. Anal.*, 105(4):327–365, 1989.

[SMST02] Jorge Alonso San Martín, Victor Starovoitov, and Marius Tucsnak. Global weak solutions for the two-dimensional motion of several rigid bodies in an incompressible viscous fluid. *Arch. Ration. Mech. Anal.*, 161(2):113–147, 2002.

[Sta04] Victor N. Starovoitov. Behavior of a rigid body in an incompressible viscous fluid near a boundary. In *Free boundary problems (Trento, 2002)*, volume 147 of *Internat. Ser. Numer. Math.*, pages 313–327. Birkhäuser, Basel, 2004.

[Sul98] Ion Suliciu. On the thermodynamics of rate-type fluids and phase transitions. I. Rate-type fluids. *Internat. J. Engrg. Sci.*, 36(9):921–947, 1998.

[SV03] Nicolas Seguin and Julien Vovelle. Analysis and approximation of a scalar conservation law with a flux function with discontinuous coefficients. *Math. Models Methods Appl. Sci.*, 13(2):221–257, 2003.

[Tor09] Eleuterio F. Toro. *Riemann solvers and numerical methods for fluid dynamics*. Springer-Verlag, Berlin, third edition, 2009. A practical introduction.

[Vas01] Alexis Vasseur. Strong traces for solutions of multidimensional scalar conservation laws. *Arch. Ration. Mech. Anal.*, 160(3):181–193, 2001.

[vL97] Bram van Leer. Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method. *J. Comput. Phys.*, 135(2):227–248, 1997.

[Vov02] Julien Vovelle. Convergence of finite volume monotone schemes for scalar conservation laws on bounded domains. *Numer. Math.*, 90(3):563–596, 2002.

[VZ03] Juan Luis Vázquez and Enrique Zuazua. Large time behavior for a simplified 1D model of fluid-solid interaction. *Comm. Partial Differential Equations*, 28(9-10):1705–1738, 2003.

[VZ06] Juan Luis Vázquez and Enrique Zuazua. Lack of collision in a simplified 1D model for fluid-solid interaction. *Math. Models Methods Appl. Sci.*, 16(5):637–678, 2006.

[WC84] Paul Woodward and Phillip Colella. The numerical simulation of two-dimensional fluid flow with strong shocks. *J. Comput. Phys.*, 54(1):115–173, 1984.

[Zai12] Daniel Wei-Ming Zaide. *Numerical Shockwave Anomalies*. ProQuest LLC, Ann Arbor, MI, 2012. Thesis (Ph.D.)–University of Michigan.