Thèse de doctorat de l'Université Pierre et Marie Curie
Organisme d'accueil : Commissariat à l'Energie Atomique
et aux énergies alternatives

Présentée et soutenue publiquement le 9 décembre 2014

pour l'obtention du grade de
Docteur de L'Université Pierre et Marie Curie
Spécialité : Mathématiques Appliquées

par
Mathieu Girardin

Méthodes numériques tout-régime et préservant l'asymptotique de type Lagrange-Projection.

Application aux écoulements diphasiques en régime bas Mach

Après avis des rapporteurs
M. Francis Filbet
M. Thierry Gallouët

Devant le jury composé de

M.	Christophe Chalons	Directeur de Thèse
M.	Frédéric Coquel	Examinateur
M.	Stéphane Dellacherie	Examinateur
M.	Francis Filbet	Rapporteur
M.	Thierry Gallouët	Rapporteur
M^{me}	Edwige Godlewski	Présidente du jury
M.	Samuel Kokh	Co-Encadrant
M.	Quang Huy Tran	Examinateur

à Mado,

Remerciements

Je tiens à remercier toutes les personnes qui ont contribué à l'aboutissement de ce projet de recherche et m'excuse si les quelques lignes qui vont suivre ne sont pas à la hauteur de l'aide précieuse qui m'a été fournie tout au long de ma thèse.

Mes premiers remerciements vont à Christophe Chalons et Samuel Kokh qui ont encadré cette thèse. Ils ont toujours été disponibles et n'ont pas hésité à donner de leur temps pour mener à bien ce projet. Leurs connaissances en analyse numérique, en calcul scientifique et leur sens de la pédagogie ont grandement contribué au niveau scientifique et à la clarté de ce manuscrit. Néanmoins, si je ne devais les remercier que pour une seule chose, ce serait pour leurs qualités humaines sans lesquelles cette collaboration n'aurait pas été aussi agréable. Encore merci pour tout.

J'aimerais remercier Edwige Richebois, Didier Jamet et Danielle Gallo qui ont assuré, au cours de ma thèse, la direction du LMEC (anciennement LETR) du CEA Saclay, ainsi que les conseillers scientifiques Jacques Segré et Alain Forestier avec qui j'ai pu avoir des discussions très enrichissantes, et de manière plus générale l'ensemble des membres du LMEC, du LJLL et du LRC Manon. Un grand merci en particulier à Pierre-Arnaud Raviart, Stéphane Dellacherie, Nicolas Seguin, Yohan Penel et Jonathan Jung pour les échanges constructifs que nous avons pu avoir sur mes travaux de thèse.

Je souhaite remercier vivement Francis Filbet et Thierry Gallouët d'avoir accepter de rédiger un rapport sur ce travail, ainsi que Christophe Chalons, Frédéric Coquel, Stéphane Dellacherie, Francis Filbet, Thierry Gallouët, Edwige Godlewski, Samuel Kokh et Quang Huy Tran d'avoir accepté de participer à mon jury. C'est pour moi un honneur de les associer à mon travail.

Bien qu'ils n'aient pas participé au contenu scientifique de cette thèse, je tiens à saluer ma famille et mes amis qui ont aidé à maintenir le moral des troupes pendant cette formidable aventure qu'est la thèse.

Enfin, je ne sais comment remercier Marie pour son rôle déterminant durant cette période. J'ai toujours pu compter sur son soutien indéfectible, en particulier lorsque tout partait à vau-l'eau.

Résumé

Les écoulements diphasiques dans les circuits primaire et secondaire des centrales de type réacteur à eau pressurisée (REP) appartiennent à des régimes très variés allant du faible nombre de Mach (en régime nominal) jusqu'aux ondes de chocs (pour certains cas accidentels). Ces différents régimes d'écoulements peuvent également apparaître simultanément ou successivement. Calculer des solutions approchées précises de ces écoulements peut s'avérer délicat dans certains régimes. Par exemple, les schémas numériques classiques de type volumes finis sont trop diffusifs dans le régime des faibles nombre de Mach et requièrent alors d'utiliser une discrétisation extrêmement fine pour calculer de bonnes solutions approchées.

On s'intéresse dans le cadre de cette thèse à la conception et à l'étude de méthodes numériques robustes et stables à grand pas de temps, capables de calculer des solutions approchées précises quel que soit le régime d'écoulement, y compris sur maillage grossier. Un des points importants de cette thèse est la stratégie en trois étapes qui permet de construire de tels schémas :

1. Utiliser un *splitting* d'opérateur pour séparer la résolution approchée des phénomènes rapides de celles des phénomènes lents. On utilisera notamment une approche Lagrange-Projection qui permettra de décomposer naturellement les ondes acoustiques et les ondes matières. On construit ainsi des schémas semi-implicites stables à grand pas de temps.
2. Utiliser un solveur de relaxation de type Suliciu pour traiter les ondes acoustiques. Ce solveur est robuste et permet de gérer les non-linéarités issues de la loi de pression. On obtient ainsi un schéma implicite peu coûteux en temps de calcul, qui nécessite seulement la résolution d'un système linéaire.
3. Introduire une modification des flux numériques à partir de l'analyse du comportement de l'erreur de troncature du schéma en fonction du nombre de Mach. Cette stratégie anti-diffusive permet en particulier d'améliorer la précision du schéma dans le régime des faibles nombres de Mach.

Deux approches sont utilisées pour analyser la capacité du schéma numérique à gérer plusieurs régimes d'écoulement. La première approche est celle des schémas *asymptotic preserving* que l'on a utilisée pour traiter le système de la dynamique des gaz avec termes sources raides. Une seconde approche basée sur la notion de schéma tout-régime a ensuite été utilisée pour le système de la dynamique des gaz à bas nombre de Mach ainsi que pour les systèmes diphasiques homogénéisés HRM et HEM à bas nombre de Mach.

Des propriétés garantissant la stabilité et la robustesse des schémas ont également été obtenues. On a en particulier étudié l'obtention d'inégalités d'entropie discrètes. Finalement, l'implémentation de ces méthodes a permis de mener des expériences numériques en 1D et 2D sur maillage non structuré, qui confirment le gain en précision et en temps de calcul des schémas *asymptotic preserving* et tout-régime ainsi construits par rapport à des schémas numériques classiques.

Abstract

Two-phase flows in Pressurized Water Reactors (PWR) belong to a wide range of Mach number flows, going from low Mach number value (in nominal configuration) to shock waves (for some accidental configurations). Computing accurate approximate solutions of those flows may be challenging from a numerical point of view as classical finite volume methods are too diffusive in the low Mach regime, and so require to use very fine discretizations.

In this thesis, we are interested in designing and studying some robust numerical schemes that are stable for large time steps and accurate even on coarse meshes for a wide range of flow regimes. An important feature is the three-step strategy to construct those schemes :

1. Use an operator splitting to solve fast and slow phenomena separately. In particular, we use a Lagrange-Projection strategy to decouple acoustic and material waves. This will help to design mixed implicit-explicit schemes that are stable for large time steps.
2. Use a Suliciu type relaxation scheme to solve acoustic waves. This solver is robust and deals with pressure non-linearities. The resulting implicit scheme is cheap as it only requires to solve a linear problem.
3. Introduce a modification of numerical fluxes based on the behavior of the truncation error with respect to the Mach number. This anti-diffusive strategy improves the accuracy of the numerical scheme in the low Mach regime.

Two approaches have been used to assess the ability of our numerical schemes to deal with a wide range of flow regimes. The first approach, based on the asymptotic preserving property, has been used for the gas dynamics equations with stiff source terms. The second approach, based on the all-regime property, has been used for the gas dynamics equations and the homogeneous two-phase flows models HRM and HEM in the low Mach regime.

We also obtained some robustness and stability properties for our numerical schemes. In particular, some discrete entropy inequalities are shown. Numerical evidences, in 1D and in 2D on unstructured meshes, assess the gain in term of accuracy and CPU time of those asymptotic preserving and all-regime numerical schemes in comparison with classical finite volume methods.

Table des matières

Introduction générale

0.1 Contexte général

La modélisation et la simulation des écoulements dans les circuits primaire et secondaire des centrales de type réacteur à eau pressurisée (REP) a suscité de nombreux travaux de recherche sur les écoulements diphasiques au sein du CEA, d'EDF et de l'IRSN notamment. Plusieurs codes de calcul de thermohydraulique utilisés dans l'industrie nucléaire pour les études de sûreté s'appuient sur des modèles d'écoulements à un ou plusieurs constituants compressibles. C'est le cas par exemple des codes CATHARE, NEPTUNE, et FLICA (France), ainsi que de RELAP (USA).

Les études de sûreté demandent d'examiner des écoulements appartenant à des régimes très variés, allant du très faible nombre de Mach (comme les écoulements dans un cœur en régime nominal) jusqu'aux ondes de chocs (pour certains cas d'accidents tels que la perte de réfrigérant primaire, les phénomènes de crise d'ébullition, ou le renoyage des cœurs). L'étude d'écoulements présentant plusieurs régimes simultanément ou successivement doit également être considérée.

Gérer ces différents régimes et les transitions entre eux peut s'avérer délicat d'un point de vue numérique. Il se trouve en effet que les schémas classiques de type volume fini, souvent utilisé dans ce contexte, sont trop diffusifs dans le régime des faibles nombres de Mach et requièrent d'utiliser une discrétisation extrêmement fine pour calculer de bonnes solutions approchées. L'utilisation d'une discrétisation aussi fine en temps et en espace pose des problèmes de coût en temps de calcul et de stockage de données rendant ces schémas difficilement utilisables pour des applications industrielles.

De nouvelles méthodes numériques permettant d'effectuer des simulations dans le régime bas Mach ont été proposées dans la littérature depuis les années 1980. Néanmoins, la mise au point et l'étude de méthodes numériques robustes, rapides et permettant de gérer avec précision plusieurs régimes d'écoulements est encore aujourd'hui un défi et un domaine actif de recherche, et concerne également d'autres secteurs industriels comme l'aéronautique ou l'industrie pétrolière.

Dans le cadre de cette thèse, nous développons et étudions des techniques permettant de construire et de caractériser des schémas numériques tout-régime, c'est à dire des schémas précis et stables quel que soit le régime d'écoulement considéré.

0.2 Formulation mathématique

Dans le cadre de cette thèse on s'intéressera à des systèmes d'équations aux dérivées partielles modélisant la dynamique d'un ou plusieurs fluides, notés $\mathbf{S}^\epsilon$, qui dépendent d'un paramètre ϵ représentant par exemple le nombre de Mach. On considèrera des schémas numériques $\mathbf{S}^\epsilon_{\mathbf{h},\mathbf{\Delta t}}$ où h est le pas de maillage et Δt le pas de temps, qui permettent de calculer des solutions approchées de $\mathbf{S}^\epsilon$. Pour les systèmes que l'on va considérer, on observera à l'aide de simulations que les méthodes numériques usuelles nécessitent de choisir $\Delta t = O(\epsilon h)$ pour des raisons de stabilité et $h = O(\epsilon)$ pour des raisons de précision alors que les variations caractéristiques en temps et en espace de la solution sont d'ordre 1 par rapport à ϵ. Ces contraintes sur la discrétisation en temps et en espace sont donc coûteuses en terme de temps de calcul et de stockage de données pour des faibles valeurs de ϵ. On va détailler ici quelques méthodes proposées dans la littérature pour construire des méthodes qui s'affranchissent de ces contraintes sur la finesse de la discrétisation à utiliser pour calculer des solutions approchées précises de $\mathbf{S}^\epsilon$.

1. **Construction d'un schéma numérique pour le régime $\epsilon \ll 1$ à l'aide du système limite.** Une première méthode consiste à étudier la limite $\mathbf{S^0}$ du système $\mathbf{S}^\epsilon$ quand ϵ tend vers zéro. La nature du système $\mathbf{S^0}$ peut être différente de celle du système $\mathbf{S}^\epsilon$. On construit alors un schéma numérique $\mathbf{S^0_{h,\Delta t}}$ stable et consistant avec $\mathbf{S^0}$. Ce schéma permet de calculer de bonnes approximations des solutions du système $\mathbf{S^0}$ pour un pas d'espace et un pas de temps donnés, indépendants du paramètre ϵ que l'on a fait tendre vers 0. Ce sont également de bonnes approximations des solutions de $\mathbf{S}^\epsilon$ tant que $\epsilon \ll 1$.

2. **Couplage d'un schéma pour le régime $\epsilon \ll 1$ avec un schéma pour le régime ϵ d'ordre 1.** Un inconvénient de la méthode précédente est que le schéma $\mathbf{S^0_{h,\Delta t}}$ est consistant avec $\mathbf{S^0}$ et non pas avec $\mathbf{S}^\epsilon$. C'est pourquoi ce schéma numérique n'est pas en mesure de calculer de bonnes approximations des solutions de $\mathbf{S}^\epsilon$ lorsque ϵ est d'ordre 1. Une façon de palier à ce problème est d'utiliser le schéma $\mathbf{S^0_{h,\Delta t}}$ consistant avec $\mathbf{S^0}$ dans les régions où $\epsilon \ll 1$ et un schéma $\mathbf{S^\epsilon_{h,\Delta t}}$ consistant avec $\mathbf{S}^\epsilon$ dans les régions où ϵ est d'ordre 1. La construction d'une telle stratégie numérique pose néanmoins les questions suivantes : Quel indicateur choisir pour délimiter les régions où utiliser $\mathbf{S^0_{h,\Delta t}}$ des régions où utiliser $\mathbf{S^\epsilon_{h,\Delta t}}$? Comment coupler les schémas $\mathbf{S^0_{h,\Delta t}}$ et $\mathbf{S^\epsilon_{h,\Delta t}}$ consistants respectivement avec les systèmes $\mathbf{S^0}$ et $\mathbf{S}^\epsilon$ qui sont de natures potentiellement différentes et peuvent ne pas avoir le même nombre d'équations ?

3. **Construction de schémas *asymptotic preserving*.** On considère un schéma $\mathbf{S^\epsilon_{h,\Delta t}}$ consistant avec $\mathbf{S}^\epsilon$ et on définit au moins formellement $\mathbf{S^0_{h,\Delta t}} = \lim_{\epsilon \to 0} \mathbf{S^\epsilon_{h,\Delta t}}$. Un schéma est dit *asymptotic preserving* ou préservant l'asymptotique si le schéma $\mathbf{S^0_{h,\Delta t}}$ est consistant avec le système limite $\mathbf{S^0}$. Un tel schéma ne nécessite pas de couplage entre deux schémas numériques distincts car il effectue naturellement la transition entre consistance avec $\mathbf{S}^\epsilon$ pour $\epsilon > 0$ et consistance avec $\mathbf{S^0}$ lorsque ϵ tend vers zéro. Un point délicat est de savoir comment construire de tels schémas numériques. En effet, pour les systèmes considérés au cours de cette thèse, les méthodes numériques classiques ne préservent pas l'asymptotique. Par ailleurs, cette propriété se concentre sur la capacité du schéma numérique à reproduire le comportement du système continu quand on passe à la limite ϵ tend vers zéro, ce qui soulève la question du comportement des schémas préservant l'asymptotique dans les régimes intermédiaires.

4. **Construction de schémas tout-régime.** Un schéma numérique $\mathbf{S^\epsilon_{h,\Delta t}}$ est dit tout-régime s'il permet de calculer des solutions approchées précises avec un pas de temps et un pas de maillage indépendants de ϵ. Pour construire un tel schéma, on s'intéresse au comportement vis-à-vis du paramètre ϵ des propriétés de stabilité et de consistance du schéma $\mathbf{S^\epsilon_{h,\Delta t}}$ étudié comme discrétisation

du système $\mathbf{S}^\epsilon$. Ainsi, ces études ne nécessitent pas de faire tendre ϵ vers 0 pour essayer de retrouver le comportement du système limite $\mathbf{S}^\mathbf{0}$, mais sont plutôt basées sur des propriétés uniformes par rapport au paramètre ϵ, qui garantissent le bon comportement du schéma numérique quel que soit le régime considéré.

On s'intéresse dans le cadre de cette thèse à la conception et à l'étude de méthodes numériques robustes et stables à grand pas de temps, capables de calculer des solutions approchées précises quelque soit le régime d'écoulement, y compris sur maillage grossier. Un des points importants de cette thèse est la stratégie en trois étapes qui permet de construire de tels schémas :

1. Utiliser un *splitting* d'opérateur pour séparer la résolution approchée des phénomènes rapides de celles des phénomènes lents. On utilisera notamment une approche Lagrange-Projection qui permettra de décomposer naturellement les ondes acoustiques et les ondes matières. On construit ainsi des schémas semi-implicites stables à grand pas de temps.
2. Utiliser un solveur de relaxation de type Suliciu pour traiter les ondes acoustiques. Ce solveur est robuste et permet de gérer les non-linéarités issues de la loi de pression. On obtient ainsi un schéma implicite peu coûteux en temps de calcul, qui nécessite seulement la résolution d'un système linéaire.
3. Introduire une modification des flux numériques à partir de l'analyse du comportement de l'erreur de troncature du schéma en fonction du nombre de Mach. Cette stratégie anti-diffusive permet en particulier d'améliorer la précision du schéma dans le régime des faibles nombres de Mach.

Deux approches sont utilisées pour analyser la capacité du schéma numérique à gérer plusieurs régimes d'écoulement. La première approche est celle des schémas *asymptotic preserving* que l'on a utilisée pour traiter le système de la dynamique des gaz avec termes sources raides. Une seconde approche basée sur la notion de schéma tout-régime a ensuite été utilisée pour le système de la dynamique des gaz à bas nombre de Mach ainsi que pour les systèmes diphasiques homogénéisés HRM et HEM à bas nombre de Mach.

Des propriétés garantissant la stabilité et la robustesse des schémas ont également été obtenues. On a en particulier étudié l'obtention d'inégalités d'entropie discrètes. Finalement, l'implémentation de ces méthodes à permis de mener des expériences numériques en 1D et 2D sur maillage non structuré, qui viennent confirmer le gain en précision et en temps de calcul des schémas *asymptotic preserving* et tout-régime ainsi construits par rapport à des schémas numériques classiques.

Ce manuscrit de thèse se présente comme suit. Dans les chapitres 1 et 2, on considère la construction et l'étude de schémas *asymptotic preserving* pour le système de la dynamique des gaz avec termes sources raides. On introduira notamment la décomposition Lagrange-Projection mentionnée ci-dessus. Ensuite on propose des schémas tout-régime, où le nombre de Mach M jouera le rôle du paramètre ϵ, pour le système de la dynamique des gaz dans le chapitre 3 et les systèmes diphasiques homogénéisés HRM et HEM dans le chapitre 4 basés sur la même décomposition Lagrange-Projection. Finalement, le chapitre 5 présente certains aspects de programmation qui ont permis d'obtenir les résultats numériques présentés dans les chapitres précédents. Dans les cinq sections ci-dessous, on présente les principaux résultats obtenus au cours de cette thèse, qui sont présentés plus en détails au cours des cinq chapitres de ce manuscrit.

0.3 Chapitre 1 : Schémas préservant l'asymptotique et stables à grand pas de temps pour le système de la dynamique des gaz avec termes sources raides

Dans ce premier chapitre, on considère le système hyperbolique de la dynamique des gaz avec termes sources raides en une dimension d'espace

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p) = \rho(g - \frac{\alpha}{\epsilon} u), \\ \partial_t(\rho E) + \partial_x((\rho E + p)u) = \rho u(g - \frac{\alpha}{\epsilon} u). \end{cases} \tag{1}$$

Le paramètre ϵ influe sur le coefficient de friction dans les termes sources. En considérant le comportement en temps long $t' = \epsilon t$ et le développement asymptotique pour la vitesse $u = u^0 + \epsilon u^1 + \mathcal{O}(\epsilon^2)$. On obtient en faisant tendre ϵ vers zéro dans (1) le système parabolique limite

$$\begin{cases} \partial_{t'} \rho + \partial_x(\rho u^1) = 0, \\ \partial_x p = \rho(g - \alpha u^1), \\ \partial_{t'}(\rho e) + \partial_x((\rho e + p)u^1) = \rho u^1(g - \alpha u^1). \end{cases} \tag{2}$$

Plusieurs schémas numériques *asymptotic preserving* pour ce système ont été proposés dans la littérature. Le but de notre étude est de proposer un schéma numérique pour le système (1) qui réponde à deux objectifs. Tout d'abord, le schéma doit être *asymptotic preserving* pour éviter les contraintes sur la discrétisation lorsque $\epsilon \ll 1$. De plus, dans les applications qui nous intéressent, l'amplitude des ondes (rapides) acoustiques est faible mais ce sont ces ondes qui pilotent le choix du pas de temps. On proposera donc une méthode stable à grand pas de temps où la condition de stabilité CFL fera intervenir uniquement la vitesse des ondes (lentes) matières.

Pour atteindre ce second objectif, on introduit un *splitting* d'opérateurs de type Lagrange-Projection qui permet naturellement de résoudre séparément les phénomènes rapides, à savoir les ondes acoustiques et les termes sources, et les phénomènes lents, à savoir les ondes matières. On construit alors un schéma semi-implicite en traitant implicitement les phénomènes rapides pour éviter les contraintes sur le choix du pas de temps et explicitement les phénomènes lents afin de rester précis.

Il reste à choisir des schémas numériques pour ces deux étapes qui permettent d'obtenir la propriété *asymptotic preserving*. Pour la première étape, on propose un schéma prenant en compte simultanément les termes sources et les termes acoustiques grâce à la notion de consistance au sens intégral avec terme source. On utilise un solveur de Riemann approché, basé sur une approche de relaxation en pression qui permet de gérer à moindre coût les non-linéarités issues de la loi d'état du fluide. On obtient ainsi un schéma numérique implicite qui nécessite seulement la résolution d'un système linéaire penta-diagonal et est donc peu coûteux en temps de calcul. L'étape de transport est ensuite résolue avec un schéma décentré amont explicite qui permet de rester précis pour la résolution des ondes matières. C'est cette étape qui détermine le pas de temps du schéma complet qui est donc dirigé par la vitesse des ondes (lentes) matières.

On prouve certaines propriétés de ce schéma numérique. Il est en particulier conservatif et préserve la positivité de la densité. Le schéma est *asymptotic preserving*. On dispose également d'une inégalité d'entropie discrète pour une version explicite du schéma. Des résultats numériques viennent confirmer le bon comportement du schéma et montrent en particulier le gain en précision et en temps de calcul par rapport à un schéma qui n'est pas *asymptotic preserving* quand ϵ est faible.

0.4 Chapitre 2 : Schémas de splitting d'opérateur préservant l'asymptotique pour le système de la dynamique des gaz avec termes sources raides

On considère à nouveau le système de la dynamique des gaz avec termes sources raides (1) et sa limite quand ϵ tend vers zéro (2). Dans le chapitre 1, on a construit un schéma *asymptotic preserving* grâce à la théorie de la consistance au sens intégral avec termes sources qui traitait simultanément les termes sources et les termes convectifs. La propriété *asymptotic preserving* était alors montrée par le calcul mais ne jouait pas de rôle explicite dans la construction du schéma.

On va s'intéresser ici à un autre processus de construction de schéma *asymptotic preserving*, où l'étude de l'erreur de troncature du schéma va permettre de modifier un schéma numérique pour le rendre *asymptotic preserving*. Cette méthode de construction est plus facilement transposable Ã d'autres systÃĺmes et d'autres asymptotiques que celle du chapitre 1, on l'utilisera en particulier dans les chapitres 3 et 4 pour traiter l'asymptotique des faibles nombres de Mach.

On utilise à nouveau un *splitting* d'opérateur mais on sépare cette fois-ci le système (1) en trois sous-systèmes contenant respectivement les termes associés aux trois vitesses du système, à savoir les ondes acoustiques, les termes sources et les ondes de transport. La résolution des termes sources est ainsi effectuée dans une étape à part que l'on résout implicitement afin d'éviter la forte contrainte sur le pas de temps lié à ϵ. Les étapes acoustiques et de transport sont quant à elles traitées de manière explicite.

On utilise alors des schémas numériques pour résoudre de manière approchée chacune de ces étapes. L'analyse de l'erreur de troncature de ces différents schémas quand ϵ tend vers zéro montre l'utilité d'introduire un nouveau degré de liberté $\theta(\epsilon)$ dans l'étape acoustique afin d'obtenir des flux numériques consistants avec les flux du système limite (2) quand ϵ tend vers zéro.

Le schéma ainsi construit est conservatif et préserve la positivité de la densité. Par ailleurs, pourvu que $\theta(\epsilon) = O(\epsilon)$ le schéma ainsi construit est bien *asymptotic preserving*. Des résultats numériques montrent le bon comportement du schéma lorsque $\epsilon \ll 1$, pourvu que la condition sur la modification $\theta(\epsilon)$ soit vérifiée. Une généralisation de cette méthode à deux dimensions d'espace et aux maillages non structurés permet d'obtenir des résultats numériques similaires.

Finalement, des études annexes viennent compléter ces analyses en apportant un éclairage sur l'implémentation des conditions aux limites, le comportement du schéma en régime intermédiaire et l'extension à l'ordre 2 en espace.

0.5 Chapitre 3 : Schémas Lagrange-Projection tout-régime pour le système de la dynamique des gaz sur maillage non structuré

Dans le chapitre 3, on considère le système de la dynamique des gaz en deux dimensions d'espace

$$\left\{\begin{array}{rcl} \partial_t \rho + \nabla \cdot (\rho \mathbf{u}) & = & 0, \\ \partial_t (\rho \mathbf{u}) + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \nabla p & = & 0, \\ \partial_t (\rho E) + \nabla \cdot [(\rho E + p)\mathbf{u}] & = & 0. \end{array}\right. \tag{3}$$

On introduit des grandeurs caractéristiques pour adimensionner le système (3) et on obtient

$$\left\{\begin{array}{rcl} \partial_t \rho + \nabla \cdot (\rho \mathbf{u}) & = & 0, \\ \partial_t (\rho \mathbf{u}) + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \frac{1}{M^2} \nabla p & = & 0, \\ \partial_t (\rho E) + \nabla \cdot [(\rho E + p)\mathbf{u}] & = & 0, \end{array}\right. \tag{4}$$

où le nombre de Mach $M = \frac{u_0}{c_0}$, u_0 et c_0 sont respectivement une vitesse du fluide et une vitesse du son caractéristiques. Le nombre de Mach M joue ici le rôle du paramètre ϵ dont dépend le système. Dans les chapitres 1 et 2, le paramètre ϵ apparait dans les termes sources tandis qu'ici il intervient dans les termes de flux. Sous certaines conditions, un système limite quand M tend vers zéro peut être obtenu (voir l'annexe 3.D pour plus de détails). Néanmoins, on privilégie ici l'approche des schémas tout-régime qui ne nécessitent pas d'utiliser le système limite. En effet, on étudie le comportement par rapport à M de la condition de stabilité CFL et de l'erreur de troncature du schéma comme discrétisation du système (4).

On considère tout d'abord un *splitting* d'opérateurs Lagrange-Projection qui permet de découpler la résolution des ondes acoustiques de celles des ondes matières. Dans le régime des faibles nombres de Mach, c'est l'étape acoustique qui contient les phénomènes rapides que l'on souhaite traiter implicitement pour éviter les contraintes sur le pas de temps liées à la vitesse du son dans le fluide. On utilise un solveur de relaxation en pression pour l'étape acoustique afin d'obtenir un schéma implicite peu coûteux en temps de calcul. L'étape de transport est ensuite effectuée à l'aide d'un schéma décentré amont explicite.

L'analyse de l'erreur de troncature du schéma suggère d'introduire une modification $\theta(M)$ afin de contrôler la diffusion numérique introduite par la discrétisation du gradient de pression qui peut venir polluer la solution approchée dans le régime des faibles nombres de Mach ($M \ll 1$). On propose ensuite un solveur de Riemann approché permettant de retrouver les flux numériques du schéma modifié. On utilise ce résultat afin d'étudier les propriétés de stabilité de ce schéma numérique anti-diffusif, on obtient en particulier une inégalité d'entropie discrète.

Le schéma numérique est écrit en deux dimensions d'espace et sur maillage non structuré. Il permet de calculer des solutions approchées précises dans le régime des faibles nombres de Mach pourvu que $\theta(M) = O(M)$. Il est de plus conservatif et présente de bonnes propriétés de stabilité. On montre par exemple que le schéma préserve la positivité de la densité et on exhibe une inégalité d'entropie discrète pour la version explicite du schéma. Des tests numériques en régime bas Mach et en régime compressible montrent le gain en précision et en temps de calcul ainsi que la robustesse du schéma tout-régime ainsi construit.

Des études annexes viennent apporter un éclairage sur l'adimensionnement du système de la dynamique des gaz, un résultat de stabilité L^2, l'influence de la forme du maillage sur le comportement du schéma à bas nombre de Mach et la comparaison avec un schéma doté d'une correction bas Mach écrit directement en coordonnée Eulérienne (par opposition au formalisme Lagrange-Projection proposé ici).

0.6 Chapitre 4 : Schémas Lagrange-Projection tout-régime pour les modèles diphasiques homogénéisés HEM et HRM sur maillage non structuré

Dans le chapitre 4, on s'intéresse au système diphasique homogénéisé HRM

$$\left\{\begin{array}{rcl} \partial_t(\rho Y) + \nabla\,.\,(\rho Y \mathbf{u}) & = & \lambda_0 \rho \left(Y^*(\rho,e) - Y\right), \\ \partial_t \rho + \nabla\,.\,(\rho \mathbf{u}) & = & 0, \\ \partial_t(\rho \mathbf{u}) + \nabla\,.\,(\rho \mathbf{u} \otimes \mathbf{u}) + \nabla p & = & 0, \\ \partial_t(\rho E) + \nabla\,.\,\left[(\rho E + p)\,\mathbf{u}\right] & = & 0, \end{array}\right. \tag{5}$$

ainsi qu'au système diphasique homogénéisé HEM que l'on obtient formellement pour $\lambda_0 = +\infty$. Comme précédemment, on introduit des grandeurs caractéristiques pour adimensionner ce système et on obtient

$$\left\{\begin{array}{rcl} \partial_t(\rho Y) + \nabla\,.\,(\rho Y \mathbf{u}) & = & \lambda_0 T \rho \left(Y^*(\rho,e) - Y\right), \\ \partial_t \rho + \nabla\,.\,(\rho \mathbf{u}) & = & 0, \\ \partial_t(\rho \mathbf{u}) + \nabla\,.\,(\rho \mathbf{u} \otimes \mathbf{u}) + \frac{1}{M^2}\nabla p & = & 0, \\ \partial_t(\rho E) + \nabla\,.\,\left[(\rho E + p)\,\mathbf{u}\right] & = & 0. \end{array}\right. \tag{6}$$

Le nombre de Mach M associé au mélange joue ici le rôle du paramètre ϵ.

On propose pour les systèmes HRM et HEM une méthode numérique inspirée de celle présentée au cours du chapitre 3 pour le système de la dynamique des gaz. La différence entre les systèmes (5) et (3) est la nouvelle variable Y qui correspond à la fraction de masse d'un des deux constituants du mélange. Cette nouvelle variable, en plus d'être transportée à la vitesse du mélange, fait apparaître un terme source qui correspond aux transitions de phases. Par ailleurs, la loi de pression de mélange dans (5) est une fonction de la fraction de masse, ainsi que de la densité et de l'énergie interne du mélange.

On considère un *splitting* d'opérateurs Lagrange-Projection-Source où on sépare le système (5) en trois sous-systèmes contenant respectivement les termes acoustiques, les termes de transport et les termes sources. Il est à noter que la vitesse de relaxation vers l'équilibre thermodynamique λ_0 peut s'avérer grande devant la vitesse des ondes matières. C'est pourquoi on traitera implicitement les phénomènes de transition de phase. Comme précédemment, on traite implicitement l'étape acoustique, à l'aide d'un schéma de relaxation en pression, et explicitement l'étape de transport, afin d'éviter la forte contrainte sur le pas de temps liée à la vitesse du son du mélange dans le régime des faibles nombres de Mach.

L'analyse de l'erreur de troncature du schéma suggère, comme dans le chapitre 3, de modifier celui-ci à l'aide d'un nouveau degré de liberté $\theta(M)$ pour le rendre plus précis dans le régime des faibles nombres de Mach. On obtient, pourvu que $\theta(M) = O(M)$, de bonnes propriétés de stabilité et de consistance du schéma modifié, quel que soit le régime considéré. Cela confirme le caractère tout-régime du schéma modifié en deux dimensions d'espace et sur maillage non structuré ainsi construit. Les simulations numériques confirment le bon comportement du schéma que ce soit en régime bas Mach ou en régime compressible.

Des études annexes viennent compléter cette étude en proposant un solveur de Riemann approché permettant de retrouver les flux du schéma modifié, on utilise alors ce résultat pour prouver une inégalité d'entropie discrète pour une version explicite du schéma.

0.7 Chapitre 5 : Implémentation

Le chapitre 5 présente certains aspects des codes développés au cours de cette thèse. On présente notamment la bibliothèque *YAFiVoC* qui a été développée et utilisée pour implémenter les schémas numériques des chapitres 2, 3 et 4. On s'intéresse ensuite plus précisément à quelques éléments de l'implémentation du schéma semi-implicite sur maillage non structuré introduit dans le chapitre 3.

0.8 Publications

Les travaux présentés dans ce manuscrit ont fait l'objet de publications.

1. Les travaux du chapitre 1 ont fait l'objet d'une publication dont les références sont : C. Chalons, M. Girardin and S. Kokh, *Large time step and asymptotic preserving numerical schemes for the gas dynamics equations with source terms*, SIAM J. Sci. Comput., 35(6) : a2874–a2902, (2013).
2. Les travaux du chapitre 2 ont fait l'objet d'une publication dont les références sont : C. Chalons, M. Girardin and S. Kokh, *Operator-splitting-based asymptotic preserving scheme for the gas dynamics equations with stiff source terms*, AIMS on Applied Mathematics, Proceedings of the 2012 International Conference on Hyperbolic Problems : Theory, Numerics, Applications, 8 : 607–614 , (2014).
3. Les travaux du chapitre 3 ont fait l'objet d'un article soumis à la revue Communications in Computational Physics : C. Chalons, M. Girardin and S. Kokh, *An all-regime Lagrange-Projection like scheme for the gas dynamics equations on unstructured meshes.*
4. Les travaux du chapitre 4 ont fait l'objet d'un article soumis : C. Chalons, M. Girardin and S. Kokh, *An all-regime Lagrange-Projection like scheme for 2D homogeneous models for two-phase flows on unstructured meshes.*

Chapitre 1

Schémas préservant l'asymptotique et stables à grand pas de temps pour le système de la dynamique des gaz avec termes sources raides

Ce chapitre a fait l'objet d'une publication dont les références sont : C. Chalons, M. Girardin and S. Kokh, *Large time step and asymptotic preserving numerical schemes for the gas dynamics equations with source terms*, SIAM J. Sci. Comput., 35(6) : a2874–a2902, (2013).

Large time step and asymptotic preserving numerical schemes for the gas dynamics equations with source terms

Abstract

We propose a *large time step* and *asymptotic preserving* scheme for the gas dynamics equations with external forces and friction terms. By asymptotic preserving, we mean that the numerical scheme is able to reproduce at the discrete level the parabolic-type asymptotic behaviour satisfied by the continuous equations. By large time-step, we mean that the scheme is stable under a CFL stability condition driven by the (slow) material waves, and not by the (fast) acoustic waves as it is customary in Godunov-type schemes. Numerical evidence are proposed and show a gain of several orders of magnitude in both accuracy and efficiency.

1.1 Introduction

Motivation. In this paper, we consider the system of gas dynamics in Eulerian coordinates with external body forces and friction terms in one space dimension. Our first motivation is the simulation of the flow in the core of a nuclear reactor whose geometry is composed of many channels. From a practical point of view, this very complex geometry is not resolved but is very often taken into account by means of a subgrid model in order to save computational cost. More precisely, the core is modelled as a porous medium in many industrial codes and the friction coefficient is associated with the medium porosity and is used to model the wall-friction influence of the channels upon the fluid. We refer for instance the reader to [2], [54], [53] and the references therein.

First feature and difficulty. In such applications, the flow is *subsonic* as the fluid velocity is low and stationary or nearly stationary flow profiles are of particular interest. These profiles express the balance between gravity, friction and a pressure drop between the inlet and the outlet of the core. From a numerical point of view, we are thus naturally interested in the development of numerical schemes able to reach such stationary solutions as quickly as possible, that is to say using large time steps in the frame of a time-marching strategy which considers a stationary solution as limit of unsteady processes when time t goes to ∞. As acoustic waves are not predominant in our industrial processes unlike transport waves, we first more precisely require our method to enable the use of large time steps in order to avoid classic Courant-Friedrichs-Lewy (CFL) restriction based on the (fast) acoustic waves of the model. Second, we want our method to accurately approximate (slow) waves that account for material transport as this one is actually predominant. Classic means to fulfil our first requirement for avoiding CFL based on the acoustic waves, consists in deriving an implicit in time discretization. Unfortunately and apart from the question of boundary conditions which deserves a particular attention, this usually induces more numerical diffusion, including for the approximation of the material waves. In order to meet both first and second needs, we will propose a mixed implicit-explicit strategy : the terms responsible for the acoustic waves receive a time implicit treatment while the ones responsible for the transport waves are treated by an explicit update. This task is achieved by means of a Lagrange-Projection [32] algorithm as in Coquel *et al.* [19]. This approach provides a natural decoupling of the acoustic waves and the material waves. On the other hand, an approximation based on a now well-known relaxation strategy introduced by Suliciu [55] and Jin and Xin [45], see also for instance [20, 18, 17, 8], will provide us with a simple mean to circumvent the nonlinearities involved with the equation of state of the fluid.

Even though the fluid velocity can be very small in practice for certain flow configurations, we do not

consider in this paper the nearly incompressible regime of low Mach numbers. It is indeed well-known that when the Mach number is small, standard shock-capturing methods that work correctly when the Mach number is of order one are not accurate and need a great attention. We refer for instance the reader to Dellacherie *et al.* [25, 26] (and the references therein) for recent contributions on this topic, and to Haack, Jin and Liu [38] for a new numerical method for the isentropic Euler and Navier-Stokes equations that is valid for all Mach numbers and whose ideas are close to the ones proposed here. In the present paper, we focus on the subsonic regime with *moderate* difference between the speeds of the flow and the acoustic waves and postpone to a future work the extension of our strategy to the case of several space dimensions and the low Mach regime.

Second feature and difficulty. Another feature encountered in our target applications is the use of coarse or very coarse spacial discretizations in order to save computational cost. Therefore, we need to build a numerical scheme that is as accurate as possible for a given coarse space discretization and large time steps. The accuracy of the classic splitting strategy with pointwise (implicit) evaluation of the source term turns out to be severely affected in this context and does not provide satisfactory results. For such goal, we propose to develop a so-called asymptotic-preserving scheme introduced in the pioneer work of Jin [40] and Jin and Levermore [43]. When one considers the solutions of our system of gas dynamics in Eulerian coordinates with external body forces and friction terms in the asymptotic regime obtained for both long time and large friction coefficients α, the solutions of the system are expected to behave like the solutions of a typical parabolic system, see for instance [39, 46, 21, 49, 50]... Therefore, we aim at deriving a scheme that preserves this property for the discrete approximation of the solution. Such property is referred to as an asymptotic preserving (AP) property. Observe that while the magnitude of the friction coefficient α used in our target applications may be considered 0.5 up to 1.0, which is by no mean a large value, the large friction coefficient limit $\alpha \to \infty$ is of real interest as it can be considered as a model worst-case scenario for testing the accuracy of the method in the presence of friction source term and coarse meshes.

Since its introduction in Jin [40] and Jin and Levermore [43], the notion of AP numerical schemes has been investigated and implemented in the past years in a wide range of context. In collisional kinetic theory with applications to plasmas, semiconductors, rarefied gas dynamics, radiative transfer (to mention only a few of them), let us quote for instance (see also the references therein) Klar [48], Jin [41], Jin, Pareschi and Toscani [44], Naldi and Pareschi [51], Gosse and Toscani [36], Buet *et al.* [12, 11], Berthon *et al.* [3, 5] Carillo, Goudon and Lafitte and Vecil [14, 15], Degond *et al.* [24], [23], Filbet and Jin [29], Crouseilles and Lemou [22], Dimarco and Pareschi [28], Després, Buet and Franck [27]. For problems similar to our model and related methods, we can quote for instance Bouchut, Ounaissa and Perthame [9], Berthon and Turpault [5] and Chalons *et al.* [16]. Without being exhaustive, we also refer the reader to [4, 1] and [27] for the recent development of asymptotic-preserving finite volume schemes on unstructured meshes.

Let us briefly discuss the methods proposed in [9, 5, 16] as they are certainly the closest to the one proposed here, at least in its explicit version. In [9], the authors extend the so-called USI (Upwinding Sources at Interfaces) approach initiated by Cargo and Le Roux [13], Greenberg and Le Roux [37], Gosse and Le Roux [34] (see also Gosse [33], Perthame and Simeoni [52], Jin [42], Katsaounis, Perthame and Simeoni [47], R. Botchorishvili, B.Perthame and A.Vasseur [7]...) whose principle is to upwind the sources at interfaces, to the Euler equations with high friction in the barotropic case and without gravity. The approach then uses a classic finite volume scheme together with the upwinding of source terms involving the reconstruction of interface variables while preserving Darcy steady states. In order to prove the AP property, a restrictive hypothesis for the basic scheme which is not valid for all schemes is assumed. In [5], the authors modify the well-known HLL Approximate Riemann Solver (ARS) for the associated homogeneous hyperbolic system by introducing a free parameter into the source term in order to obtain the AP property. The resulting numerical procedure is robust as the source term discretization preserves

the physical admissible states and can be applied to several models of physical interest. In [16], the authors derive an explicit asymptotic-preserving scheme for the same flow model as the one considered here, the scheme being also well-balanced when the model is written in Lagrangian coordinates. The method is a Godunov-type method based on the definition of a relevant ARS. For defining such an ARS, the effects of both gravity and friction source terms are incorporated into the solver thanks to the concept of simple Approximate Riemann Solver and consistency with the integral form introduced by Gallice [30, 31]. This powerful method allows to account for both source terms and convective fluxes at the same time. In particular, the source terms are taken into account at interfaces. This method therefore also falls into the general framework of USI schemes.
In this paper, we extend this approach within a mixed implicit-explicit framework for designing an AP scheme which is stable for large time steps in order to meet both requirements discussed in the previous paragraphs.

Outline of the paper. The outline of the paper is as follows. In the next section, we give the model under consideration and its parabolic-type asymptotic limit. In section 1.3, we first briefly recall the Lagrange-Projection decomposition and the pressure relaxation strategy. We then recall the concepts of simple Approximate Riemann Solver and consistency in the integral sense in section 1.4. The last part of this section gives the explicit in time numerical scheme and section 1.5 focuses on the Lagrangian system. At the end of this section, an important discussion is proposed on the definition of the time step and the interfacial velocities in the asymptotic regime. In particular, we explain why the proposed strategy is well suited to preserve the asymptotic limit, unlike the classic splitting strategy, and why an implicit treatment is needed for the time step not to be zero in the asymptotic regime. At last, the proposed implicit in time scheme for the Lagrangian system is given in section 1.6 and the overall mixed implicit-explicit scheme is described in section 1.7. Finally, section 1.8 gives the main properties of the mixed implicit-explicit scheme, including a discussion on the scheme obtained in the asymptotic limit, and the last section provides some numerical illustrations.

1.2 Governing equations and asymptotic behaviour

The gas dynamics equations with gravity and friction terms in Eulerian coordinates are given by

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p) = \rho(g - \alpha u), \\ \partial_t(\rho E) + \partial_x((\rho E + p)u) = \rho u(g - \alpha u), \end{cases} \tag{1.1}$$

where ρ, u and E denote the density, the velocity and the total energy of the fluid, g the gravitational acceleration and α the friction parameter. The pressure law $p = p(\rho, e)$ is assumed to be a given function of the density ρ and the internal energy e defined by $e = E - \frac{u^2}{2}$, satisfying the usual Weyl assumptions [56]. Under these assumptions and when the source terms are omitted, (1.1) is shown to be strictly hyperbolic over the phase space Ω given by $\Omega = \{(\rho, \rho u, \rho E)^T \in \mathbb{R}^3,\ \rho > 0,\ e > 0\}$, with eigenvalues given by $\lambda_1 = u - c < \lambda_2 = u < \lambda_3 = u + c$, where $c = [(\partial p/\partial \rho)_\epsilon + (p/\rho^2)(\partial p/\partial \epsilon)_\rho]^{1/2}$ is the sound speed. Moreover, the characteristic fields associated with λ_1 and λ_3 are genuinely non linear while the characteristic field associated with λ_2 is linearly degenerate.

Let us also recall that λ_1 and λ_3 give rise to the so-called acoustic waves, while λ_2 is associated to the transport phenomenon. We refer for instance the reader to [32] for more details.

Let $s = s(\rho, e)$ be the strictly convex mathematical specific entropy which satisfies $\partial_e s(\rho, e) < 0$ and

$$-Tds = de + pd\left(\frac{1}{\rho}\right), \tag{1.2}$$

where $T > 0$ is the temperature. We obtain for a smooth solution of (1.1) the following equation $\partial_t(\rho s) + \partial_x(\rho s u) = 0$. We are particularly interested in studying the long time behavior of (1.1) when the friction parameter goes to infinity. Let us assume that (1.1) is in dimensionless form and let ϵ be a dimensionless and small positive parameter. We model this flow regime by replacing α with α/ϵ with a slight abuse of notation and by performing the change of variable $t' = \epsilon t$ in the system (1.1). We obtain the system

$$\epsilon\partial_{t'}\rho + \partial_x(\rho u) = 0, \quad \epsilon\partial_{t'}(\rho u) + \partial_x(\rho u^2 + p) = \rho\left(g - \frac{\alpha}{\epsilon}u\right), \quad \epsilon\partial_{t'}(\rho E) + \partial_x((\rho E + p)u) = \rho u\left(g - \frac{\alpha}{\epsilon}u\right).$$

Let us assume that the velocity u admits an asymptotic expansion in powers of ϵ of the following form

$$u = u^0 + \epsilon u^1 + \mathcal{O}(\epsilon^2). \tag{1.3}$$

Multiplying the second equation by ϵ and letting ϵ go to 0 first gives $u^0 = 0$. Then inserting $u = \epsilon u^1 + \mathcal{O}(\epsilon^2)$ in the first equation, dividing by ϵ and letting ϵ go to 0 gives $\partial_{t'}\rho + \partial_x\rho u^1 = 0$. If we now insert $u = \epsilon u^1 + \mathcal{O}(\epsilon^2)$ in the second equation and let ϵ go to 0, we get $\partial_x p = \rho g - \rho\alpha u^1$. At last, inserting $u = \epsilon u^1 + \mathcal{O}(\epsilon^2)$ in the third equation, dividing by ϵ and letting ϵ go to 0 gives $\partial_{t'}(\rho e) + \partial_x(\rho e u^1 + p u^1) = \rho u^1(g - \alpha u^1)$. The long time behaviour of the solutions of (1.1) for large friction coefficients is then given by the following system of partial differential equations

$$\begin{cases} \partial_{t'}\rho + \partial_x(\rho u^1) = 0, \\ \partial_x p = \rho(g - \alpha u^1), \\ \partial_{t'}(\rho e) + \partial_x((\rho e + p)u^1) = \rho u^1(g - \alpha u^1), \end{cases} \tag{1.4}$$

where we note that in comparison to (1.1) the flow speed u has been replaced by its first order corrector u^1 in *both* the mass flux of the first equation and the friction term of the second equation. This observation will play a crucial role in the forthcoming developments. In addition, we note that this model has to be understood as the diffusive or parabolic limit of the hyperbolic model (1.1) since the second derivative of the pressure p naturally appears in the first equation of the limit system using its second equation : $\partial_{t'}\rho + \partial_x\left(\frac{\rho g - \partial_x p}{\alpha}\right) = 0$. The formal derivation of (1.4) proposed here can be given a rigorous meaning from the analysis point of view. Such a problem has indeed been studied by many authors and at least in the simplified situation of the barotropic case, many contributions are concerned with the existence and convergence of the solutions of the gaz dynamics equations to the Darcy's law in (1.4). To mention only a few of them, let us quote for instance [39, 46, 21, 49, 50]. Note that these works essentially differ from the underlying assumptions (Lagrangian or Eulerian coordinates, smooth or possibly discontinuous solutions, one or several space dimensions, linear or non linear pressure laws...) and techniques. See also the references therein and [6] for further results.

From a numerical point of view and as already said in the previous section, one of our objectives is to preserve this asymptotic behaviour at the discrete level. In other words, we aim at proposing a consistent numerical scheme for (1.1) leading to a consistent numerical scheme for (1.4) when ϵ goes to zero and up to the expected changes of variables. Before proceeding and to conclude this section, let us formally rephrase this property in terms of limits with respect to the small parameter ϵ and to the time and space steps Δt, Δx used in the numerical approximation. Let us denote M^ϵ the initial model (1.1), M^0 the limit model (1.4), $S^\epsilon_{\Delta t,\Delta x}$ a consistent numerical scheme for (1.1) and $S^0_{\Delta t,\Delta x}$ its asymptotic limit. Recall that consistency of $S^\epsilon_{\Delta t,\Delta x}$ means that $\lim\limits_{\Delta t,\Delta x\to 0} S^\epsilon_{\Delta t,\Delta x} = M^\epsilon$, for all $\epsilon > 0$. By definition and with a little abuse in the notations, $S^\epsilon_{\Delta t,\Delta x}$ is said to be asymptotic preserving if $S^0_{\Delta t,\Delta x}$ is consistent with M^0, that is $\lim\limits_{\Delta t,\Delta x\to 0} S^0_{\Delta t,\Delta x} = M^0$, or equivalently $\lim\limits_{\Delta t,\Delta x\to 0}\lim\limits_{\epsilon\to 0} S^\epsilon_{\Delta t,\Delta x} = \lim\limits_{\epsilon\to 0} M^\epsilon$. In other words, the asymptotic preserving property is formally equivalent to the following order of limits interchange

property : $\lim_{\Delta t,\Delta x\to 0}\lim_{\epsilon\to 0} S^{\epsilon}_{\Delta t,\Delta x} = \lim_{\epsilon\to 0}\lim_{\Delta t,\Delta x\to 0} S^{\epsilon}_{\Delta t,\Delta x}$. From a practical point of view this equality formally means that for large friction coefficients, or equivalently for small values of ϵ, an asymptotic preserving scheme is expected to give good numerical results even for reasonable mesh sizes (with respect to ϵ). This will be observed in the last section devoted to the numerical experiments.

1.3 Lagrange-Projection approach and relaxation procedure

In this section we briefly recall the so-called Lagrange-Projection strategy applied to (1.1) and propose a relaxation procedure for approximating the solutions of the underlying Lagrangian system. As motivated in the introduction, these are two key ingredients of the method we propose, together with the notion of consistency in the integral sense that will be recalled in the next section. Let us begin with the Lagrange-Projection decomposition.

1.3.1 Lagrange-Projection decomposition

We describe here a procedure that allows to approximate the evolution of the system (1.1) over a time interval $[t_0, t_0+\Delta t]$. The guideline of the method consists in decoupling the terms responsible for the acoustic waves and the transport waves. By using the chain rule for the space derivatives we split up the operators of system (1.1) and obtain two subsystems. The first subsystem describes the transport process and reads

$$\left\{\begin{array}{l} \partial_t\rho + u\partial_x\rho = 0, \\ \partial_t(\rho u) + u\partial_x(\rho u) = 0, \\ \partial_t(\rho E) + u\partial_x(\rho E) = 0. \end{array}\right. \tag{1.5}$$

The second subsystem accounts for acoustic, gravity and friction effects, namely

$$\partial_t\rho + \rho\partial_x u = 0, \quad \partial_t(\rho u) + \rho u\partial_x u + \partial_x p = \rho(g-\alpha u), \quad \partial_t(\rho E) + \rho E\partial_x u + \partial_x(pu) = \rho u(g-\alpha u).$$

If we note $\tau = \frac{1}{\rho}$ the specific volume, the above system also reads

$$\left\{\begin{array}{l} \partial_t\tau - \tau\partial_x u = 0, \\ \partial_t u + \tau\partial_x p = g - \alpha u, \\ \partial_t E + \tau\partial_x(pu) = u(g-\alpha u). \end{array}\right. \tag{1.6}$$

Then for $t\in[t_0, t_0+\Delta t]$, we propose to approximate $\tau(x,t)\partial_x\cdot$ by $\tau(x,t_0)\partial_x\cdot$ in (1.6). If one introduces the mass variable m defined by $\mathrm{d}m = \tau(x,t_0)^{-1}\mathrm{d}x$, we obtain

$$\left\{\begin{array}{l} \partial_t\tau - \partial_m u = 0, \\ \partial_t u + \partial_m p = g - \alpha u, \\ \partial_t E + \partial_m(pu) = u(g-\alpha u), \end{array}\right. \tag{1.7}$$

that will be referred to as the Lagrangian system. Let us note that system (1.7) is consistent with the usual form of the gas dynamics equations in Lagrangian coordinates with friction and gravity terms. We refer for instance the reader to [32] for more details. It is worth noticing that (1.7) is easily shown to be hyperbolic over the phase space Ω^{Lag} given by $\Omega^{\mathrm{Lag}} = \{(\tau,u,E)^T\in\mathbb{R}^3, \tau>0, e>0\}$, with eigenvalues given by $\lambda_1^{\mathrm{Lag}} = -\rho c < \lambda_2^{\mathrm{Lag}} = 0 < \lambda_3^{\mathrm{Lag}} = \rho c$, where c still denotes the sound speed. Here again, the extreme characteristic fields associated with λ_1^{Lag} and λ_3^{Lag} are genuinely non linear while the intermediate characteristic field associated with λ_2^{Lag} is linearly degenerate. Importantly, we note that the sound speed only appears in the characteristic speeds of this Lagrangian model. The flow speed u is

no longer present but is on the other hand the unique characteristic velocity of the first subsystem (1.5). System (1.7) (respectively (1.5)) then only contains the so-called acoustic (resp. material or transport) waves. From a numerical point of view, the scheme associated with this decomposition simply consists of an usual two-step splitting strategy where (1.7) is solved in the first step and (1.5) in the second one. Both steps will be solved over the same time interval. Recall that the idea will be here to propose a time implicit treatment of the Lagrangian system (1.7) to avoid a too restrictive CFL condition involving the sound speed c and an explicit treatment of the transport part (1.5) to keep accuracy on the material waves. A key ingredient to get a low cost implicit scheme for the Lagrangian system will rely on a relevant pressure relaxation approximation described in the next section.

1.3.2 Relaxation approximation

We propose in this section a relaxation approximation of the Lagragian system (1.7). The main objective is to overcome the non linearities that make difficult the resolution of this system. From a numerical point of view, this strategy will be used to design a low cost time implicit treatment. We first consider the model neglecting the source terms and then extend the approach to the full model.

Relaxation approximation of the homogeneous model

The design principle of the so-called pressure relaxation methods is to introduce a larger system than the original one but easier to solve. More precisely, the objective is to discard the non linearities induced by the pressure law $p = p(\rho, e)$. Such a strategy is now well known in the literature and we refer for instance the reader to [55, 45, 20, 18, 17, 8] and the references therein. To do so, we introduce a new variable Π that can be seen as a linearization of the pressure p and that is considered as a new unknown. In particular, it evolves according to its own partial differential equation. More precisely, we propose the following relaxation system for (1.7) when the source terms are omitted :

$$\begin{cases} \partial_t \tau - \partial_m u = 0, \\ \partial_t u + \partial_m \Pi = 0, \\ \partial_t \Pi + a^2 \partial_m u = \lambda(p - \Pi), \\ \partial_t E + \partial_m (\Pi u) = 0, \end{cases} \tag{1.8}$$

where a is a constant to be precised and λ the relaxation parameter. At least formally, we observe that in the asymptotic regime $\lambda \to +\infty$ we have $\Pi \to p$ and we recover the initial system (1.7) without the source terms. In order to prevent this relaxation procedure from forming instabilities, it is now well established that a must be chosen sufficiently large and according to the subcharacteristic condition

$$a > \max(\rho c), \tag{1.9}$$

for all the states under consideration (see for instance [17] for a rigorous proof). In addition, it can be easily proved that (1.8) with $\lambda = 0$ is strictly hyperbolic with three eigenvalues given by $-a$, 0 and a which are nothing but approximations of the exact eigenvalues $-\rho c$, 0 and ρc for system (1.7). Then and in particular, the subcharacteristic condition means that information propagates faster in the relaxation model. More importantly, the characteristic fields associated with these new eigenvalues are shown to be linearly degenerate. This property allows to solve analytically the Riemann problem associated with (1.8) when $\lambda = 0$, that is when considering an initial data made of two constant states separated by an initial discontinuity. This property justifies by itself the introduction of the proposed relaxation model and its simplicity. If we go further into the details and as it is customary, the exact Riemann solutions are self-similar and made of three contact discontinuities propagating with velocities $-a$, a and 0 and

separating two intermediate states.

From a numerical point of view, the numerical strategy for approximating the solutions of (1.7) using (1.8) consists in first solving (1.8) with $\lambda = 0$, that is

$$\left\{\begin{array}{l} \partial_t \tau - \partial_m u = 0, \\ \partial_t u + \partial_m \Pi = 0, \\ \partial_t \Pi + a^2 \partial_m u = 0, \\ \partial_t E + \partial_m (\Pi u) = 0, \end{array}\right. \tag{1.10}$$

and then to take into account the source term $\partial_t \tau = 0, \quad \partial_t u = 0, \quad \partial_t \Pi = \lambda(p - \Pi), \quad \partial_t E = 0$, in the asymptotic regime $\lambda \to +\infty$. Which amounts to set $\Pi = p(\rho, E)$, before solving again (1.10) as the time goes on. The new variable Π is said to be at equilibrium. Importantly, note that λ does not appear explicitly in the scheme, but its value is always implicitly equal to ∞ which is expressed by the relation $\Pi = p(\rho, E)$. In particular, no confusion can be made in regards to the order of the limits in λ and ϵ when the AP property will be considered.

To conclude this section, notice that the self-similar Riemann solutions associated with (1.10) being explicitly known, it is natural to use an *exact* Godunov scheme to numerically solve (1.10). See for instance the references above for more details.

Relaxation approximation of the Lagrangian system with source terms

Let a be a real parameter chosen in agreement with the subcharacteristic condition (1.9). In order to approximate the solution of (1.7), we propose to supplement (1.10) with the friction and gravity terms so that we have to solve

$$\left\{\begin{array}{l} \partial_t \tau - \partial_m u = 0, \\ \partial_t u + \partial_m \Pi = g - \alpha u, \\ \partial_t \Pi + a^2 \partial_m u = 0, \\ \partial_t E + \partial_m (\Pi u) = u(g - \alpha u). \end{array}\right. \tag{1.11}$$

Let us underline that the solutions of the Riemann problem associated with (1.11) are neither self-similar nor explicitly known anymore. This makes the use of the exact Godunov method quite a complex task. In what follows, we decide nevertheless to approximate the non self-similar Riemann solutions to (1.11) by self-similar approximate Riemann solutions and to use an approximate Godunov-type method for solving (1.11). In order to guarantee the consistency of the proposed self-similar approximate Riemann solutions to (1.11) with the exact ones, we will impose a generalized notion of consistency in the integral sense due to Gallice [30, 31] and adapted to systems with source terms. This is recalled in the next section.

To conclude this section, let us observe that (1.11) can be given the following equivalent form

$$\left\{\begin{array}{l} \partial_t \tau - \partial_m u = 0, \\ \partial_t \overrightarrow{w} + a \partial_m \overrightarrow{w} = a(g - \alpha u), \\ \partial_t \overleftarrow{w} - a \partial_m \overleftarrow{w} = -a(g - \alpha u), \\ \partial_t E + \partial_m (\Pi u) = u(g - \alpha u), \end{array}\right. \tag{1.12}$$

where the new variables $\overrightarrow{w}$ and $\overleftarrow{w}$ are defined by $\overrightarrow{w} = \Pi + au, \quad \overleftarrow{w} = \Pi - au$. These quantities are nothing but the strong Riemann invariants associated with the characteristic speeds $\pm a$ of the relaxation system (1.11) when the source terms are omitted. The closure relations for (1.12) are naturally given by $u = \frac{\overrightarrow{w} - \overleftarrow{w}}{2a}, \quad \Pi = \frac{\overrightarrow{w} + \overleftarrow{w}}{2}$. This new formulation will be used hereafter to define the proposed implicit in time numerical strategy.

1.4 Consistency in the integral sense and explicit in time Godunov-type scheme

We briefly recall in this section the notion of consistency in the integral sense of a self-similar approximate Riemann solver for a given set of hyperbolic equations with source terms that we write in the condensed form

$$\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = \mathbf{S}(\mathbf{U}), \tag{1.13}$$

supplemented with the validity of an entropy inequality

$$\partial_t \eta + \partial_x q \leq 0, \tag{1.14}$$

where (η, q) is a strictly convex entropy-entropy flux pair. We also derive the corresponding explicit in time Godunov-type scheme for approximating the solutions to (1.13) and refer to [30, 31] for the details.

Solving the Riemann problem amounts to find the solution to (1.13) with the following piecewise constant initial data $\mathbf{U}(x, t=0) = \mathbf{U}_L \quad \text{if} \quad x < 0, \qquad \mathbf{U}_R \quad \text{if} \quad x > 0$, for any given $\mathbf{U}_L$ and $\mathbf{U}_R$ in the phase space. Unlike the homogeneous case corresponding to the choice $\mathbf{S}(\mathbf{U}) = 0$, the exact Riemann solution that we denote $\mathbf{U}(x, t; \mathbf{U}_L, \mathbf{U}_R)$ is not self-similar. Notice however that an approximate Riemann solver $\mathbf{W}(\frac{x}{t}; \mathbf{U}_L, \mathbf{U}_R)$ may be self-similar as in the homogeneous case provided that some consistency relations are imposed. More precisely, let us consider a simple approximate Riemann solver $\mathbf{W}(\frac{x}{t}; \mathbf{U}_L, \mathbf{U}_R)$ made of $l+1$ intermediate states $\mathbf{U}_k$ separated by discontinuities propagating with velocities λ_k, namely

$$\mathbf{W}\left(\frac{x}{t}; \mathbf{U}_L, \mathbf{U}_R\right) = \begin{cases} \mathbf{U}_1 = \mathbf{U}_L, & \frac{x}{t} < \lambda_1, \\ \vdots & \\ \mathbf{U}_k, & \lambda_{k-1} < \frac{x}{t} < \lambda_k, \\ \vdots & \\ \mathbf{U}_{l+1} = \mathbf{U}_R, & \frac{x}{t} > \lambda_l. \end{cases} \tag{1.15}$$

From Gallice [30, 31], if $\Delta x = \frac{1}{2}(\Delta x_L + \Delta x_R)$ with $\Delta x_L > 0$, $\Delta x_R > 0$ and $\Delta t > 0$ are respectively space and time steps that verify the CFL condition

$$\max_{1 \leq k \leq l} |\lambda_k| \frac{\Delta t}{\min(\Delta x_L, \Delta x_R)} \leq \frac{1}{2}, \tag{1.16}$$

the approximate Riemann solver is said to be consistent with the integral form of (1.13) over the interval $[\frac{-\Delta x_L}{2}, \frac{\Delta x_R}{2}]$ if the integral of (1.15) approximates correctly the integral of the exact solution in the sense that there exists a function $\widetilde{\mathbf{S}}$ such that

$$\mathbf{F}(\mathbf{U}_R) - \mathbf{F}(\mathbf{U}_L) - \Delta x \widetilde{\mathbf{S}}(\Delta x, \Delta t; \mathbf{U}_L, \mathbf{U}_R) = \sum_{k=1}^{l} \lambda_k (\mathbf{U}_{k+1} - \mathbf{U}_k), \tag{1.17}$$

where $\widetilde{\mathbf{S}}(\Delta x, \Delta t; \mathbf{U}_L, \mathbf{U}_R)$ is consistent with the source terms $\mathbf{S}(\mathbf{U})$ in the sense that

$$\lim_{\substack{\mathbf{U}_L, \mathbf{U}_R \to \mathbf{U} \\ \Delta t, \Delta x \to 0}} \widetilde{\mathbf{S}}(\Delta x, \Delta t; \mathbf{U}_L, \mathbf{U}_R) = \mathbf{S}(\mathbf{U}).$$

Hereafter and using very classic notations, $(\Delta x_j)_{j \in \mathbb{Z}}$ and Δt represent the constant time and variable space steps of the mesh under consideration for defining the approximate solutions. More precisely and in order to define the Godunov-type scheme associated with this approximate Riemann solver, we define

the mesh interfaces $x_{j+1/2} = x_{j-1/2} + \Delta x_j$ for $j \in \mathbb{Z}$, and the intermediate times $t^n = n\Delta t$ for $n \in \mathbb{N}$. Note that Δx in (1.17) then plays the role of $\Delta x_{j+1/2} = \frac{1}{2}(\Delta x_j + \Delta x_{j+1})$. In the sequel, $\mathbf{U}_j^n$ denotes the approximate value of $\mathbf{U}$ at time t^n and on the cell $[x_{j-1/2}, x_{j+1/2})$. For $n = 0$ and $j \in \mathbb{Z}$, we set $\mathbf{U}_j^0 = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{U}_0(x)\,\mathrm{d}x$ where $\mathbf{U}_0(x)$ is the initial condition. Then, the explicit in time Godunov-type scheme reads

$$\left\{\begin{array}{l} \mathbf{U}_j^{n+1} = \mathbf{U}_j^n - \dfrac{\Delta t}{\Delta x_j}(\mathbf{F}_{j+\frac{1}{2}}^n - \mathbf{F}_{j-\frac{1}{2}}^n) + \dfrac{\Delta t}{2}(\dfrac{\Delta x_{j+1/2}}{\Delta x_j}\mathbf{S}_{j+\frac{1}{2}}^n + \dfrac{\Delta x_{j-1/2}}{\Delta x_j}\mathbf{S}_{j-\frac{1}{2}}^n), \\ \\ \mathbf{F}_{j+\frac{1}{2}}^n = \mathbf{F}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n), \\ \\ \mathbf{S}_{j+\frac{1}{2}}^n = \widetilde{\mathbf{S}}(\Delta x_{j+1/2}, \Delta t; \mathbf{U}_j^n, \mathbf{U}_{j+1}^n), \end{array}\right. \tag{1.18}$$

with $\mathbf{F}(\mathbf{U}_L, \mathbf{U}_R) = \frac{1}{2}\left\{\mathbf{F}(\mathbf{U}_L) + \mathbf{F}(\mathbf{U}_R) - \sum_{k=1}^{l} |\lambda_k|(\mathbf{U}_{k+1} - \mathbf{U}_k)\right\}$. As far as the consistency with the entropy inequality (1.14) is concerned, the simple approximate Riemann solver is said to be consistent with the integral form of (1.14) if and only if there exists a function $\widetilde{\sigma}$ such that under the CFL condition (1.16) we have

$$q(\mathbf{U}_R) - q(\mathbf{U}_L) - \Delta x\, \widetilde{\sigma}(\Delta x, \Delta t; \mathbf{U}_L, \mathbf{U}_R) \leq \sum_{k=1}^{l} \lambda_k\big(\eta(\mathbf{U}_{k+1}) - \eta(\mathbf{U}_k)\big), \tag{1.19}$$

with $\lim\limits_{\substack{\mathbf{U}_L, \mathbf{U}_R \to \mathbf{U} \\ \Delta t, \Delta x \to 0}} \widetilde{\sigma}(\Delta x, \Delta t; \mathbf{U}_L, \mathbf{U}_R) = 0$. Then, the numerical scheme defined by (1.18) satisfies the following discrete entropy inequality

$$\left\{\begin{array}{l} \eta(\mathbf{U}_j^{n+1}) \leq \eta(\mathbf{U}_j^n) - \dfrac{\Delta t}{\Delta x_j}(q_{j+\frac{1}{2}}^n - q_{j-\frac{1}{2}}^n) + \dfrac{\Delta t}{2}(\dfrac{\Delta x_{j-1/2}}{\Delta x_j}\sigma_{j-\frac{1}{2}}^n + \dfrac{\Delta x_{j+1/2}}{\Delta x_j}\sigma_{j+\frac{1}{2}}^n), \\ \\ q_{j+\frac{1}{2}}^n = \widetilde{q}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n), \\ \\ \sigma_{j+\frac{1}{2}}^n = \widetilde{\sigma}(\Delta x_{j+1/2}, \Delta t; \mathbf{U}_j^n, \mathbf{U}_{j+1}^n), \end{array}\right. \tag{1.20}$$

with

$$\widetilde{q}(\mathbf{U}_L, \mathbf{U}_R) = \frac{1}{2}\left\{q(\mathbf{U}_L) + q(\mathbf{U}_R) - \sum_{k=1}^{l} |\lambda_k|\big(S(\mathbf{U}_{k+1}) - S(\mathbf{U}_k)\big)\right\}. \tag{1.21}$$

The CFL condition associated with this explicit in time Godunov-type scheme naturally reads $\max\limits_{1\leq k\leq l} |\lambda_k(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n)| \dfrac{\Delta t}{\min(\Delta x_j, \Delta x_{j+1})} \leq \dfrac{1}{2}$, for all j. Again, we refer to [30, 31, 16] for more details. To conclude this section, let us observe that the numerical flux $\mathbf{F}(\mathbf{U}_L, \mathbf{U}_R)$ and the entropy numerical flux $\widetilde{q}(\mathbf{U}_L, \mathbf{U}_R)$ are clearly consistent in the classic senses $\mathbf{F}(\mathbf{U}, \mathbf{U}) = \mathbf{F}(\mathbf{U})$ and $\widetilde{q}(\mathbf{U}, \mathbf{U}) = q(\mathbf{U})$ provided that the intermediate states of the approximate Riemann solver are such that $\mathbf{U}_k = \mathbf{U}$ for all $k = 1, ..., l$ as soon as $\mathbf{U} := \mathbf{U}_L = \mathbf{U}_R$.

1.5 Application to the Lagrangian system and explicit in time Godunov-type scheme

We suppose again that a is a parameter that complies with the subcharacteristic constraint (1.9). We consider a step Δm of the space variable expressed through the mass variable and a time step Δt. The objective of this section is to define a consistent simple approximate Riemann solver for (1.12). We have

in this case

$$\mathbf{U} = \begin{pmatrix} \tau \\ \overrightarrow{w} \\ \overleftarrow{w} \\ E \end{pmatrix}, \ \mathbf{F}(\mathbf{U}) = \begin{pmatrix} -u \\ a\overrightarrow{w} \\ -a\overleftarrow{w} \\ \Pi u \end{pmatrix}, \ \mathbf{S}(\mathbf{U}) = \begin{pmatrix} 0 \\ a(g-\alpha u) \\ -a(g-\alpha u) \\ u(g-\alpha u) \end{pmatrix}, \ \mathbf{F}(\mathbf{U}_L, \mathbf{U}_R) = \begin{pmatrix} F^{\tau}(\mathbf{U}_L, \mathbf{U}_R) \\ F^{\overrightarrow{w}}(\mathbf{U}_L, \mathbf{U}_R) \\ F^{\overleftarrow{w}}(\mathbf{U}_L, \mathbf{U}_R) \\ F^{E}(\mathbf{U}_L, \mathbf{U}_R) \end{pmatrix}.$$

Note in particular that introducing the notation $\mathbf{F}(\mathbf{U}) = (F^{\tau}, F^{\overrightarrow{w}}, F^{\overleftarrow{w}}, F^{E})^T$, the energy flux satisfies the relation

$$F^{E} = -\frac{F^{\overrightarrow{w}} + F^{\overleftarrow{w}}}{2a} F^{\tau}. \tag{1.22}$$

This relation will be used in the calculations below.

In order to mimic the self-similar solution to (1.12) when the source terms are omitted, we propose to consider a simple approximate Riemann solver made of 3 waves, namely a stationary wave and two waves propagating with velocities $\pm a$:

$$\mathbf{W}\left(\frac{m}{t}; \mathbf{U}_L, \mathbf{U}_R\right) = \begin{cases} \mathbf{U}_L, & \frac{m}{t} < -a, \\ \mathbf{U}_L^*, & -a < \frac{m}{t} < 0, \\ \mathbf{U}_R^*, & 0 < \frac{m}{t} < a, \\ \mathbf{U}_R, & \frac{m}{t} > a. \end{cases}$$

Following [16], we define $\tilde{\mathbf{S}}$ as follows $\tilde{\mathbf{S}}(\Delta m, \Delta t; \mathbf{U}_L, \mathbf{U}_R) = \begin{pmatrix} 0 \\ a(g-\alpha\tilde{u}) \\ -a(g-\alpha\tilde{u}) \\ \tilde{u}(g-\alpha\tilde{u}) \end{pmatrix}$, where $\tilde{u}$ represents a consistent approximation of the velocity u, in the sense that $\lim_{\substack{\mathbf{U}_L, \mathbf{U}_R \to \mathbf{U} \\ \Delta t, \Delta x \to 0}} \tilde{u}(\Delta m, \Delta t; \mathbf{U}_L, \mathbf{U}_R) = u$. The definition of $\tilde{u}$ will be specified later on.

We now turn to the definition of the intermediate states $\mathbf{U}_L^*$ and $\mathbf{U}_R^*$. Each state containing four components, eight relations are expected. As motivated in the previous section, we first impose the consistency relations (1.17) which gives here

$$\begin{cases} (u_L - u_R) = -a(\tau_L^* - \tau_L) + a(\tau_R - \tau_R^*), \\ a(\overrightarrow{w}_R - \overrightarrow{w}_L) - \Delta m\, a(g-\alpha\tilde{u}) = -a(\overrightarrow{w}_L^* - \overrightarrow{w}_L) + a(\overrightarrow{w}_R - \overrightarrow{w}_R^*), \\ -a(\overleftarrow{w}_R - \overleftarrow{w}_L) + \Delta m\, a(g-\alpha\tilde{u}) = -a(\overleftarrow{w}_L^* - \overleftarrow{w}_L) + a(\overleftarrow{w}_R - \overleftarrow{w}_R^*), \\ (\Pi_R u_R - \Pi_L u_L) - \Delta m\, \tilde{u}(g-\alpha\tilde{u}) = -a(E_L^* - E_L) + a(E_R - E_R^*). \end{cases} \tag{1.23}$$

Then, we make the natural choice of imposing the Rankine Hugoniot relations associated with the mass conservation across each wave of the approximate Riemann solver. We get

$$\begin{cases} u_L - a\tau_L = u_L^* - a\tau_L^*, \\ u_R + a\tau_R = u_R^* + a\tau_R^*, \\ u_L^* = u_R^*. \end{cases} \tag{1.24}$$

Note that (1.24) provides only two independent relations since the first equation of (1.23) is a linear combination of the three equations in (1.24). Then, two equations are still missing. In the sequel we note $u^* = u_L^* = u_R^*$.

In order to account for the source terms, we propose to impose a generalized jump condition across the stationary wave and associated with the momentum equation in (1.11). This amounts to take into account

the source term at the interface of the initial condition. More precisely, we impose $\Pi_R^* - \Pi_L^* = \Delta m(g - \alpha\tilde{u})$. So that only one equation is now missing. At last and in regards to the energy equation, we then propose to mimic the relation (1.22) at the discrete level by imposing the following relation on the numerical flux of the Godunov-type method : $F^E(\mathbf{U}_L, \mathbf{U}_R) = -F^\tau(\mathbf{U}_L, \mathbf{U}_R) \times \frac{F^{\overrightarrow{w}}(\mathbf{U}_L, \mathbf{U}_R) + F^{\overleftarrow{w}}(\mathbf{U}_L, \mathbf{U}_R)}{2a}$. It remains to define $\tilde{u}$ and following [16] we set $\tilde{u} = u^*$. This choice aims at proposing the same approximation of the velocity u in the mass flux at the interface (that is u^*) as in the friction term (that is $\tilde{u}$). As mentioned in section 1.2, this is also true at the continuous level for the parabolic system (1.4). It turns out to be essential in order to obtain the asymptotic preserving property.

At this point, we can now define the intermediate states $\mathbf{U}_L^*$ and $\mathbf{U}_R^*$. We have after easy calculations

$$\left\{\begin{array}{l} u^* = \dfrac{1}{2a + \alpha\Delta m}\big(a(u_R + u_L) - (\Pi_R - \Pi_L) + g\Delta m\big), \\ \tau_L^* = \tau_L + \dfrac{u^* - u_L}{a}, \qquad \tau_R^* = \tau_R + \dfrac{u_R - u^*}{a}, \\ \Pi_R^* = \dfrac{\Pi_R + \Pi_L}{2} - a\dfrac{u_R - u_L}{2} + \dfrac{(g - \alpha u^*)\Delta m}{2}, \qquad \Pi_L^* = \Pi_R^* - (g - \alpha u^*)\Delta m, \\ E_L^* = E_L + \dfrac{1}{a}\Big(p_L u_L - u^*(p^* - \dfrac{\Delta m}{2}\big(g - \alpha u^*)\big)\Big), \\ E_R^* = E_R - \frac{1}{a}\Big(p_R u_R - u^*(p^* + \frac{\Delta m}{2}\big(g - \alpha u^*)\big)\Big), \end{array}\right. \tag{1.25}$$

where we have set $p^* = \frac{\Pi_R + \Pi_L}{2} - a\frac{u_R - u_L}{2} = \frac{\overrightarrow{w}_L + \overleftarrow{w}_R}{2}$. We find the numerical flux of the Godunov-type scheme

$$\left\{\begin{array}{l} F^\tau(\mathbf{U}_L, \mathbf{U}_R) = -u^*, \qquad F^{\overrightarrow{w}}(\mathbf{U}_L, \mathbf{U}_R) = a\overrightarrow{w}_L + \dfrac{a(g - \alpha u^*)\Delta m}{2} = a^2 u^* + a p^*, \\ F^{\overleftarrow{w}}(\mathbf{U}_L, \mathbf{U}_R) = -a\overleftarrow{w}_R + \dfrac{a(g - \alpha u^*)\Delta m}{2} = a^2 u^* - a p^*, \qquad F^E(\mathbf{U}_L, \mathbf{U}_R) = p^* u^*. \end{array}\right.$$

We now use the above flux definition to derive an explicit in time Godunov-type scheme for (1.12). We consider the approximate variable $\mathbf{U}_j^n$ known for $j \in \mathbb{Z}$ and we set $\Delta m_j = \rho_j^n \Delta x$, $\Delta m_{j+1/2} = \frac{\Delta m_j + \Delta m_{j+1}}{2}$. The superscript $^{\text{Lag}}$ will denote the updated values after the approximation of (1.12). Following (1.18), we are led to the following scheme

$$\left\{\begin{array}{l} \tau_j^{\text{Lag}} = \tau_j^n + \dfrac{\Delta t}{\Delta m_j}(u^*_{j+\frac12} - u^*_{j-\frac12}), \\ \overrightarrow{w}_j^{\text{Lag}} = \overrightarrow{w}_j^n - a\dfrac{\Delta t}{\Delta m_j}(\overrightarrow{w}_j^n - \overrightarrow{w}_{j-1}^n) + \Delta t\, a\, \dfrac{\Delta m_{j-1/2}}{\Delta m_j}(g - \alpha u^*_{j-\frac12}), \\ \overleftarrow{w}_j^{\text{Lag}} = \overleftarrow{w}_j^n + a\dfrac{\Delta t}{\Delta m_j}(\overleftarrow{w}_{j+1}^n - \overleftarrow{w}_j^n) - \Delta t\, a\, \dfrac{\Delta m_{j+1/2}}{\Delta m_j}(g - \alpha u^*_{j+\frac12}), \\ E_j^{\text{Lag}} = E_j^n - \dfrac{\Delta t}{\Delta m_j}((up)^*_{j+\frac12} - (up)^*_{j-\frac12}) + g\dfrac{\Delta t}{\Delta m_j}\dfrac{\Delta m_{j+1/2} u^*_{j+\frac12} + \Delta m_{j-1/2} u^*_{j-\frac12}}{2} \\ \qquad - \alpha\dfrac{\Delta t}{\Delta m_j}\dfrac{\Delta m_{j+1/2}(u^*_{j+\frac12})^2 + \Delta m_{j-1/2}(u^*_{j-\frac12})^2}{2}, \end{array}\right. \tag{1.26}$$

where

$$u^*_{j+\frac12} = \frac{1}{2a + \alpha\Delta m_{j+1/2}}(\overrightarrow{w}_j^n - \overleftarrow{w}_{j+1}^n + g\Delta m_{j+1/2}), \quad p^*_{j+\frac12} = \frac{\overrightarrow{w}_j^n + \overleftarrow{w}_{j+1}^n}{2}. \tag{1.27}$$

Notice that the source terms clearly appear to receive an upwind treatment in (1.26). This scheme is easily shown to be stable under the CFL condition given by

$$\max_{j\in\mathbb{Z}} \frac{\Delta t}{\Delta m_j} a \le \frac12, \tag{1.28}$$

for all j. Let us note that as a satisfies the subcharacteristic condition (1.9) this CFL condition naturally involves the sound speed c.

1.5.1 Remarks on Δt and $u^*_{j+\frac{1}{2}}$.

About Δt.

If we perform the scale change

$$\Delta t \to \frac{\Delta t}{\epsilon}, \quad \alpha \to \frac{\alpha}{\epsilon}, \tag{1.29}$$

in the scheme (1.26), we obtain the CFL condition $\max_{j\in\mathbb{Z}} \frac{\Delta t}{\Delta m_j} a \leq \frac{\epsilon}{2}$, which is too restrictive in the asymptotic limit as $\lim_{\epsilon\to 0} \Delta t = 0$. One way to overcome this difficulty is to treat implicitly the centered part of the source term as in [35, 10]. This gives an hyperbolic CFL condition that converges towards a parabolic CFL condition associated to the limit system in the asymptotic limit $2\Delta t < \epsilon \frac{\Delta m}{a} + \frac{\alpha(\Delta x)^2}{a^2}$.

As motivated before, here we will make the choice to get rid of (1.28) by treating implicitly both the fluxes and the whole source terms to construct a numerical scheme that has no CFL condition either in classic or asymptotic regime. A rigorous proof of the non-linear stability of the scheme is an open problem. This scheme is presented in the next section.

About $u^*_{j+\frac{1}{2}}$.

Let us underline here the importance to choose a specific covolume flux $u^*_{j+\frac{1}{2}}$ to obtain the asymptotic preserving property. Assuming asymptotic expansions $u^n_j = u^{n,(0)}_j + \mathcal{O}(\epsilon)$, similar to (1.3) at the continuous level, the key property to obtain the asymptotic preserving property will be seen in proof 1.8.1 to be

$$\text{If } \; u^{n,(0)}_j = 0 \;\; \forall j, \quad \text{then} \quad \frac{u^*_{j+\frac{1}{2}}}{\epsilon} = \frac{0}{\epsilon} + v^*_{j+\frac{1}{2}} + \mathcal{O}(\epsilon), \tag{1.30}$$

where $v^*_{j+\frac{1}{2}}$ is consistent with $u^1 = \frac{1}{\alpha}\left(g - \frac{1}{\rho}\partial_x p\right)$.
Indeed, considering the scale change (1.29), the first equation of (1.26) gives,

$$\tau^{\text{Lag}}_j = \tau^n_j + \frac{\Delta t}{\Delta m_j}\left(\frac{u^*_{j+\frac{1}{2}}}{\epsilon} - \frac{u^*_{j-\frac{1}{2}}}{\epsilon}\right) = \tau^n_j + \frac{\Delta t}{\Delta m_j}\left(v^*_{j+\frac{1}{2}} - v^*_{j-\frac{1}{2}}\right) + \mathcal{O}(\epsilon),$$

which is consistent with the first equation of (1.38).

If one uses a more natural definition of the interfacial velocity $u^*_{j+\frac{1}{2}}$, which corresponds to the interfacial velocity associated with the relaxation scheme applied to (1.8), instead of (1.27) if we set

$$u^{*,classic}_{j+\frac{1}{2}} = \frac{\overrightarrow{w}^n_j - \overleftarrow{w}^n_{j+1}}{2a} = \frac{1}{2}\left(u^n_{j+1} + u^n_j\right) - \frac{1}{2a}\left(\Pi^n_{j+1} - \Pi^n_j\right),$$

then we have

$$\frac{u^{*,classic}_{j+\frac{1}{2}}}{\epsilon} = \frac{1}{2\epsilon}\left(u^{n,(0)}_{j+1} + u^{n,(0)}_j\right) + \frac{1}{2}\left(u^{n,(1)}_{j+1} + u^{n,(1)}_j\right) - \frac{1}{2a\epsilon}\left(\Pi^n_{j+1} - \Pi^n_j\right) + \mathcal{O}(\epsilon) = v^*_{j+\frac{1}{2}} + \mathcal{O}(\frac{\Delta x}{\epsilon}) + \mathcal{O}(\epsilon)$$

Thus this classic flux does not verify the property (1.30) because of the term $\frac{1}{2a\epsilon}\left(\Pi^n_{j+1} - \Pi^n_j\right) = \mathcal{O}(\frac{\Delta x}{\epsilon})$. Therefore the numerical diffusion becomes of order $\mathcal{O}(\frac{\Delta x}{\epsilon})$ and we cannot recover the good asymptotic behaviour of the covolume flux with the more natural choice of $u^{*,classic}_{j+\frac{1}{2}}$.

While the specific choice $u^*_{j+\frac{1}{2}}$ in (1.27) gives

$$\frac{u^*_{j+\frac{1}{2}}}{\epsilon} = \frac{1}{\alpha\Delta m_{j+1/2}} \frac{1}{\left(1+\frac{2a\epsilon}{\alpha\Delta m_{j+1/2}}\right)} \left(a(u^n_{j+1}+u^n_j) - (\Pi^n_{j+1}-\Pi^n_j) + g\Delta m_{j+1/2}\right)$$
$$= \frac{1}{\alpha\Delta m_{j+1/2}} \left(a(u^{n,(0)}_{j+1}+u^{n,(0)}_j) - (\Pi^n_{j+1}-\Pi^n_j) + g\Delta m_{j+1/2}\right) + \mathcal{O}\left(\frac{\epsilon}{(\Delta x)^2}\right)$$
$$= v^*_{j+\frac{1}{2}} + \mathcal{O}\left(\frac{\epsilon}{(\Delta x)^2}\right)$$

which verifies (1.30). The consistency with u^1 arises from the specific form of non centered terms in $u^*_{j+\frac{1}{2}}$ which coincides with the equation verified by u^1. This will be the key ingredient to prove the asymptotic preserving property in Lagrangian coordinates. Asymptotic expansion as ϵ goes to zero have been performed for a given Δx so that we have in particular $\epsilon \ll \Delta x$ and $\mathcal{O}\left(\frac{\epsilon}{(\Delta x)^2}\right) = \mathcal{O}(\epsilon)$. Of course, if we set $\Delta x = 0$ for a given ϵ in (1.27) it is clear that $u^*_{j+\frac{1}{2}}$ is consistent with u.

1.6 Implicit in time Godunov-type scheme for the Lagrangian system

We follow here a standard approach for deriving an implicit in time Godunov-type scheme from (1.26) by simply replacing the terms evaluated at time t^n with the terms noted with the superscript $^{\rm Lag}$. We get

$$\begin{cases} \overrightarrow{w}^{\rm Lag}_j = \overrightarrow{w}^n_j - a\dfrac{\Delta t}{\Delta m_j}(\overrightarrow{w}^{\rm Lag}_j - \overrightarrow{w}^{\rm Lag}_{j-1}) + \Delta t\, a\, \dfrac{\Delta m_{j-1/2}}{\Delta m_j}(g - \alpha u^*_{j-\frac{1}{2}}), \\ \overleftarrow{w}^{\rm Lag}_j = \overleftarrow{w}^n_j + a\dfrac{\Delta t}{\Delta m_j}(\overleftarrow{w}^{\rm Lag}_{j+1} - \overleftarrow{w}^{\rm Lag}_j) - \Delta t\, a\, \dfrac{\Delta m_{j+1/2}}{\Delta m_j}(g - \alpha u^*_{j+\frac{1}{2}}), \end{cases} \tag{1.31}$$

with

$$u^*_{j+\frac{1}{2}} = \frac{1}{2a+\alpha\Delta m_{j+1/2}}(\overrightarrow{w}^{\rm Lag}_j - \overleftarrow{w}^{\rm Lag}_{j+1} + g\Delta m_{j+1/2}), \tag{1.32}$$

and

$$\begin{cases} \tau^{\rm Lag}_j = \tau^n_j + \dfrac{\Delta t}{\Delta m_j}(u^*_{j+\frac{1}{2}} - u^*_{j-\frac{1}{2}}), \\ E^{\rm Lag}_j = E^n_j - \dfrac{\Delta t}{\Delta m_j}((up)^*_{j+\frac{1}{2}} - (up)^*_{j-\frac{1}{2}}) + g\dfrac{\Delta t}{\Delta m_j}\dfrac{\Delta m_{j+1/2}u^*_{j+\frac{1}{2}} + \Delta m_{j-1/2}u^*_{j-\frac{1}{2}}}{2} \\ \qquad -\alpha\dfrac{\Delta t}{\Delta m_j}\dfrac{\Delta m_{j+1/2}(u^*_{j+\frac{1}{2}})^2 + \Delta m_{j-1/2}(u^*_{j-\frac{1}{2}})^2}{2}, \end{cases} \tag{1.33}$$

with $p^*_{j+\frac{1}{2}} = \dfrac{\overrightarrow{w}^{\rm Lag}_j + \overleftarrow{w}^{\rm Lag}_{j+1}}{2}$.

It is important to notice that (1.31) is independent of (1.33). More precisely, once (1.31) is solved, the update values (1.33) for τ and E follow *explicitly*. As far as $\overrightarrow{w}$ and $\overleftarrow{w}$ are concerned, the update formulas (1.31) are coupled and require the resolution of a *linear* system. The corresponding matrix is shown to be pentadiagonal and *strictly diagonally dominant*. Therefore, it is invertible and (1.31) can be solved for any $\Delta t > 0$. The proposed implicit in time numerical scheme for solving the Lagrangian system is then actually *cheap* thanks to the relaxation strategy.

1.7 Implicit-explicit in time Godunov-type scheme for the Eulerian system (1.1)

In order to complete the definition of our numerical scheme, it remains to define the second step of the operator splitting associated with the Lagrange-Projection decomposition, which corresponds to the Eulerian projection (1.5). Following [32], we consider a very classic upwind and explicit in time numerical scheme given by

$$X_j^{n+1} = X_j^{\text{Lag}} + \frac{\Delta t}{\Delta x}\Big((u^*_{j-\frac{1}{2}})^+ X_{j-1}^{\text{Lag}} + \Big[(u^*_{j+\frac{1}{2}})^- - (u^*_{j-\frac{1}{2}})^+\Big] X_j^{\text{Lag}} - (u^*_{j+\frac{1}{2}})^- X_{j+1}^{\text{Lag}}\Big), \tag{1.34}$$

where $X \in \{\rho, \rho u, \rho E\}$ and $u^+ = \frac{u+|u|}{2}, \quad u^- = \frac{u-|u|}{2}$, for all u. The update formula (1.34) is shown to be stable under the CFL condition given by

$$\frac{\Delta t}{\Delta x}\Big((u^*_{j-\frac{1}{2}})^+ - (u^*_{j+\frac{1}{2}})^-\Big) < 1, \tag{1.35}$$

see again [32]. Note that this CFL condition involves only the flow speed and is not based on the acoustic waves.

Note that, if (1.30) holds and (1.29) is performed, the CFL condition (1.35) becomes as ϵ goes to zero

$$\frac{\Delta t}{\Delta x}\Big((v^*_{j-\frac{1}{2}})^+ - (v^*_{j+\frac{1}{2}})^-\Big) < 1, \tag{1.36}$$

which will allow a strictly positive time step in the asymptotic regime.

For the sake of clarity, let us briefly recall the different steps of the overall method called LP-IMEX, suppose that at the instant n we know $(\rho_j^n, (\rho u)_j^n, (\rho E)_j^n)$ for $j \in \mathbb{Z}$, we perform the following steps :

(i) compute $(\tau_j^n, \overrightarrow{w}_j^n, \overleftarrow{w}_j^n, E_j^n)$ at equilibrium by evaluating $\Pi_j^n = p(\rho_j^n, e_j^n)$,

(ii) compute $(\tau_j^{\text{Lag}}, \overrightarrow{w}_j^{\text{Lag}}, \overleftarrow{w}_j^{\text{Lag}}, E_j^{\text{Lag}})$ thanks to the implicit scheme defined by (1.31)-(1.33),

(iii) evaluate $(\rho_j^{\text{Lag}}, (\rho u)_j^{\text{Lag}}, (\rho E)_j^{\text{Lag}})$ thanks to $(\tau_j^{\text{Lag}}, \overrightarrow{w}_j^{\text{Lag}}, \overleftarrow{w}_j^{\text{Lag}}, E_j^{\text{Lag}})$,

(iv) compute $(\rho_j^{n+1}, (\rho u)_j^{n+1}, (\rho E)_j^{n+1})$ thanks to the explicit scheme defined by (1.34) and (1.32).

In the numerical experiments and for the sake of comparison, we will also consider the following explicit-explicit numerical scheme that will be referred to as the LP-EXEX scheme. Suppose that at the instant n we know $(\rho_j^n, (\rho u)_j^n, (\rho E)_j^n)$ for $j \in \mathbb{Z}$, we perform the following steps :

(i) compute $(\tau_j^n, \overrightarrow{w}_j^n, \overleftarrow{w}_j^n, E_j^n)$ at equilibrium by evaluating $\Pi_j^n = p(\rho_j^n, e_j^n)$,

(ii) compute $(\tau_j^{\text{Lag}}, \overrightarrow{w}_j^{\text{Lag}}, \overleftarrow{w}_j^{\text{Lag}}, E_j^{\text{Lag}})$ thanks to the explicit scheme defined by (1.26),

(iii) evaluate $(\rho_j^{\text{Lag}}, (\rho u)_j^{\text{Lag}}, (\rho E)_j^{\text{Lag}})$ thanks to $(\tau_j^{\text{Lag}}, \overrightarrow{w}_j^{\text{Lag}}, \overleftarrow{w}_j^{\text{Lag}}, E_j^{\text{Lag}})$,

(iv) compute $(\rho_j^{n+1}, (\rho u)_j^{n+1}, (\rho E)_j^{n+1})$ thanks to the explicit scheme defined by (1.34) and (1.27).

1.8 Main properties

We give in this section the main properties of the proposed numerical scheme.

Theorem 1. *Under the CFL condition (1.35), the implicit-explicit in time numerical scheme LP-IMEX is well defined and satisfies the following stability properties :*

(i) it is a conservative scheme for the density ρ. It is also a conservative scheme for ρu and ρE when the source terms are omitted,

(ii) the density ρ_j^n is positive for all j and $n > 0$ provided that ρ_j^0 is positive for all j,

(iii) it is asymptotic preserving.

In addition, under the CFL condition (1.28), the explicit-explicit overall numerical scheme LP-EXEX

(iv) satisfies an entropy inequality.

This result is worth a few comments. We first note that the overall numerical scheme is stable under the CFL condition (1.35) which only involves the flow speed and not the acoustic waves. It is then less restrictive than the classic CFL restrictions of the usual explicit Godunov-type numerical schemes. The CFL condition (1.35) involves however the solution computed at the end of the Lagrangian step and then is not determined explicitly. Finding an explicit formula for the time step restriction that guarantees (1.35) and then the stability of the scheme is an open question at the moment.

Properties *(i)* and *(ii)* are obtained from standard manipulations[32, 19]. We first prove *(iii)*, then comment on the limit system we obtain, and at last we prove *(iv)*.

1.8.1 Proof of *(iii)*

This section aims at proving that the proposed LP-IMEX numerical scheme is asymptotic preserving. More precisely and similarly to the continuous analysis proposed in section 1.2, we aim at proving that if we perform the scale change

$$\Delta t \to \frac{\Delta t}{\epsilon}, \quad \alpha \to \frac{\alpha}{\epsilon}, \tag{1.37}$$

which corresponds to the study of the large time behaviour for a large friction coefficient, then we get a consistent approximation of (1.4) when ϵ goes to zero. We propose a two-part proof : we first prove the property in Lagrangian coordinates and then consider the Eulerian framework.

Asymptotic preserving property in Lagrangian coordinates

In Lagrangian coordinates, the limit system (1.4) reads

$$\begin{cases} \partial_t \tau - \partial_m u^1 = 0, \\ \partial_m p = g - \alpha u^1, \\ \partial_t e + \partial_m p u^1 = u^1 (g - \alpha u^1). \end{cases} \tag{1.38}$$

Let us perform the scale change (1.37) in (1.31)-(1.33). We get

$$\begin{cases} \overrightarrow{w}_j^{\text{Lag}} = \overrightarrow{w}_j^n - a\dfrac{\Delta t}{\epsilon \Delta m_j}(\overrightarrow{w}_j^{\text{Lag}} - \overrightarrow{w}_{j-1}^{\text{Lag}}) + \dfrac{\Delta t a}{\epsilon}\dfrac{\Delta m_{j-1/2}}{\Delta m_j}(g - \dfrac{\alpha}{\epsilon}u^*_{j-\frac{1}{2}}), \\ \overleftarrow{w}_j^{\text{Lag}} = \overleftarrow{w}_j^n + a\dfrac{\Delta t}{\epsilon \Delta m_j}(\overleftarrow{w}_{j+1}^{\text{Lag}} - \overleftarrow{w}_j^{\text{Lag}}) - \dfrac{\Delta t a}{\epsilon}\dfrac{\Delta m_{j+1/2}}{\Delta m_j}(g - \dfrac{\alpha}{\epsilon}u^*_{j+\frac{1}{2}}), \\ \tau_j^{\text{Lag}} = \tau_j^n + \dfrac{\Delta t}{\epsilon \Delta m_j}(u^*_{j+\frac{1}{2}} - u^*_{j-\frac{1}{2}}), \\ E_j^{\text{Lag}} = E_j^n - \dfrac{\Delta t}{\epsilon \Delta m_j}\Big((up)^*_{j+\frac{1}{2}} - (up)^*_{j-\frac{1}{2}}\Big) + \dfrac{g\Delta t}{\epsilon \Delta m_j}\dfrac{\Delta m_{j+1/2}u^*_{j+\frac{1}{2}} + \Delta m_{j-1/2}u^*_{j-\frac{1}{2}}}{2} \\ \qquad - \dfrac{\alpha \Delta t}{\epsilon^2 \Delta m_j}\dfrac{\Delta m_{j+1/2}(u^*_{j+\frac{1}{2}})^2 + \Delta m_{j-1/2}(u^*_{j-\frac{1}{2}})^2}{2}, \end{cases}$$

with $u^*_{j+\frac{1}{2}} = \frac{\epsilon}{2a\epsilon + \alpha \Delta m_{j+1/2}}(\overrightarrow{w}_j^{\text{Lag}} - \overleftarrow{w}_{j+1}^{\text{Lag}} + g\Delta m_{j+1/2})$ and $p^*_{j+\frac{1}{2}} = \frac{\overrightarrow{w}_j^{\text{Lag}} + \overleftarrow{w}_{j+1}^{\text{Lag}}}{2}$.

This system may be recast in variables (u, Π, τ, E)

$$
\begin{cases}
u_j^{Lag} = u_j^n - \dfrac{\Delta t}{\epsilon \Delta m_j}\left(p^*_{j+\frac{1}{2}} - p^*_{j-\frac{1}{2}}\right) + \dfrac{g\Delta t}{\epsilon \Delta m_j}\dfrac{\Delta m_{j+1/2} + \Delta m_{j-1/2}}{2} \\
\qquad - \dfrac{\alpha \Delta t}{\epsilon \Delta m_j}\dfrac{\Delta m_{j+1/2} u^*_{j+\frac{1}{2}} + \Delta m_{j-1/2} u^*_{j-\frac{1}{2}}}{2}, \\
\Pi_j^{Lag} = \Pi_j^n - a^2 \dfrac{\Delta t}{\epsilon \Delta m_j}\left(u^*_{j+\frac{1}{2}} - u^*_{j-\frac{1}{2}}\right), \\
\tau_j^{\mathrm{Lag}} = \tau_j^n + \dfrac{\Delta t}{\epsilon \Delta m_j}\left(u^*_{j+\frac{1}{2}} - u^*_{j-\frac{1}{2}}\right), \\
E_j^{\mathrm{Lag}} = E_j^n - \dfrac{\Delta t}{\epsilon \Delta m_j}\left((up)^*_{j+\frac{1}{2}} - (up)^*_{j-\frac{1}{2}}\right) + \dfrac{g\Delta t}{\epsilon \Delta m_j}\dfrac{\Delta m_{j+1/2} u^*_{j+\frac{1}{2}} + \Delta m_{j-1/2} u^*_{j-\frac{1}{2}}}{2} \\
\qquad - \dfrac{\alpha \Delta t}{\epsilon^2 \Delta m_j}\dfrac{\Delta m_{j+1/2} (u^*_{j+\frac{1}{2}})^2 + \Delta m_{j-1/2} (u^*_{j-\frac{1}{2}})^2}{2},
\end{cases}
\tag{1.39}
$$

with $p^*_{j+\frac{1}{2}} = \frac{1}{2}\left(\Pi^{\mathrm{Lag}}_{j+1} + \Pi^{\mathrm{Lag}}_j\right) - \frac{a}{2}\left(u^{\mathrm{Lag}}_{j+1} - u^{\mathrm{Lag}}_j\right)$,
and $u^*_{j+\frac{1}{2}} = \frac{\epsilon}{\alpha \Delta m_{j+1/2}} \frac{1}{(1+\frac{2a\epsilon}{\alpha \Delta m_{j+1/2}})}\left(a(u^{\mathrm{Lag}}_{j+1} + u^{\mathrm{Lag}}_j) - (\Pi^{\mathrm{Lag}}_{j+1} - \Pi^{\mathrm{Lag}}_j) + g\Delta m_{j+1/2}\right)$.

Asymptotic development as ϵ goes to zero are performed for a given Δx so that we have in particular $\epsilon \ll \Delta x$ and $\mathcal{O}\left(\frac{\epsilon}{\Delta x}\right) = \mathcal{O}(\epsilon)$. Of course, if we set $\Delta x = 0$ for a given ϵ in (1.27) it is clear that $u^*_{j+\frac{1}{2}}$ is consistent with u. Besides, the CFL condition (1.36) gives $\mathcal{O}(\frac{\Delta t}{\Delta x}) = \mathcal{O}(1)$.

Let us assume that the following discrete asymptotic development similar to (1.3) $u_j^{\mathrm{Lag}} = u_j^{\mathrm{Lag},(0)} + \mathcal{O}(\epsilon)$, then

$$
u^*_{j+\frac{1}{2}} = \frac{\epsilon}{\alpha \Delta m_{j+1/2}}\left(a(u^{\mathrm{Lag},(0)}_{j+1} + u^{\mathrm{Lag},(0)}_j) - (\Pi^{\mathrm{Lag}}_{j+1} - \Pi^{\mathrm{Lag}}_j) + g\Delta m_{j+1/2}\right) + \mathcal{O}\left(\epsilon^2\right). \tag{1.40}
$$

Multiplying the first equation of (1.39) by ϵ, we obtain $\frac{4a\Delta t}{\Delta m_j} u_j^{\mathrm{Lag},(0)} = \mathcal{O}(\epsilon)$, that implies

$$
u_j^{\mathrm{Lag},(0)} = 0 \tag{1.41}
$$

Now that we have (1.41), property (1.30) is easily proved with (1.40) and we obtain

$$
\frac{u^*_{j+\frac{1}{2}}}{\epsilon} = v^*_{j+\frac{1}{2}} + \mathcal{O}\left(\epsilon\right), \qquad \text{where} \qquad v^*_{j+\frac{1}{2}} = \frac{1}{\alpha \Delta m_{j+1/2}}\left(g\Delta m_{j+1/2} - (\Pi^{\mathrm{Lag}}_{j+1} - \Pi^{\mathrm{Lag}}_j)\right),
$$

is consistent with u^1 thanks to the second equation of (1.38).

We also have from (1.41) $p^*_{j+\frac{1}{2}} = P^*_{j+1/2} + \mathcal{O}(\epsilon)$, where $P^*_{j+1/2} = \frac{\Pi^{\mathrm{Lag}}_j + \Pi^{\mathrm{Lag}}_{j+1}}{2}$.

By reinjecting those expressions of $u^*_{j+\frac{1}{2}}$ and $p^*_{j+\frac{1}{2}}$ in the last three equations of (1.39), we obtain finally

$$
\left\{
\begin{array}{l}
\tau_j^{\mathrm{Lag}} = \tau_j^n + \dfrac{\Delta t}{\Delta m_j}\left(v^*_{j+\frac12} - v^*_{j-\frac12}\right) + \mathcal{O}(\epsilon), \\
v^*_{j+\frac12} = \dfrac{1}{\alpha}\left(g - \dfrac{\Pi_{j+1}^{\mathrm{Lag}} - \Pi_j^{\mathrm{Lag}}}{\Delta m_{j+\frac12}}\right), \\
P^*_{j+\frac12} = \dfrac{\Pi_j^{\mathrm{Lag}} + \Pi_{j+1}^{\mathrm{Lag}}}{2}, \\
\Pi_j^{\mathrm{Lag}} = \Pi_j^n - a^2\dfrac{\Delta t}{\Delta m_j}\left(v^*_{j+\frac12} - v^*_{j-\frac12}\right) + \mathcal{O}(\epsilon), \\
e_j^{\mathrm{Lag}} = e_j^n - \dfrac{\Delta t}{\Delta m_j}\left((vP)^*_{j+\frac12} - (vP)^*_{j-\frac12}\right) + g\dfrac{\Delta t}{\Delta m_j}\dfrac{\Delta m_{j+1/2}v^*_{j+\frac12} + \Delta m_{j-1/2}v^*_{j-\frac12}}{2} \\
\qquad\qquad - \alpha\dfrac{\Delta t}{\Delta m_j}\dfrac{\Delta m_{j+1/2}(v^*_{j+\frac12})^2 + \Delta m_{j-1/2}(v^*_{j-\frac12})^2}{2} + \mathcal{O}(\epsilon),
\end{array}
\right. \tag{1.42}
$$

which is consistent with (1.38) when ϵ tends to 0. Let us remark that the scheme for the variable Π is consistent with the limit pressure equation $\partial_t \Pi + a^2 \partial_m u^1 = 0$.

This limit scheme is implicit, as was the scheme (1.31)-(1.33). We assume that the scheme is stable with no CFL condition, a rigorous proof of this assertion is an open problem.

Asymptotic Preserving property in Eulerian coordinates

It remains to prove that after the Eulerian projection, the overall scheme is consistent with (1.4). In (1.34) we perform the scale change (1.29). We have for $X \in \{\rho, \rho u, \rho E\}$:

$$
X_j^{n+1} = X_j^{\mathrm{Lag}} + \frac{\Delta t}{\Delta x}\left((v^*_{j-\frac12})^+ X_{j-1}^{\mathrm{Lag}} + \left((v^*_{j+\frac12})^- - (v^*_{j-\frac12})^+\right) X_j^{\mathrm{Lag}} - (v^*_{j+\frac12})^- X_{j+1}^{\mathrm{Lag}}\right) + \mathcal{O}(\epsilon). \tag{1.43}
$$

Since we have $u_j^{\mathrm{Lag},(0)} = 0$, then $(\rho u)_j^{\mathrm{Lag},(0)} = 0$ for all $j \in \mathbb{Z}$ and considering $X = \rho u$ in the previous equality gives $(\rho u)_j^{n+1,(0)} = 0$ and then

$$
u_j^{n+1,(0)} = 0. \tag{1.44}
$$

Let us remark that the first equation in (1.42) reads also (just multiply by $\rho_j^n \rho_j^{\mathrm{Lag}}$)

$$
\rho_j^n = \rho_j^{\mathrm{Lag}}\left(1 + \frac{\Delta t}{\Delta x}(v^*_{j+\frac12} - v^*_{j-\frac12})\right) + \mathcal{O}(\epsilon).
$$

Then, for $X \in \{\rho, \rho E\}$ the Eulerian projection (1.43) may be recast into

$$
\begin{aligned}
(X)_j^{n+1} &= (X)_j^{\mathrm{Lag}}\left(1 + \frac{\Delta t}{\Delta x}(v^*_{j+\frac12} - v^*_{j-\frac12})\right) - \mathscr{L}^a((X)^{\mathrm{Lag}}, v^*) + \mathcal{O}(\epsilon) \\
&= \rho_j^n \frac{X_j^{\mathrm{Lag}}}{\rho_j^{\mathrm{Lag}}} - \mathscr{L}^a((X)^{\mathrm{Lag}}, v^*) + \mathcal{O}(\epsilon),
\end{aligned} \tag{1.45}
$$

where the advection operator $\mathscr{L}^a((X)^{\mathrm{Lag}}, v^*)$ is consistent with $\partial_x(Xv)$ and is defined by

$$
\mathscr{L}^a((X)^{\mathrm{Lag}}, v^*) = \frac{\Delta t}{\Delta x}\Big\{\left[(X)_j^{\mathrm{Lag}}(v^*_{j+\frac12})^+ + (X)_{j+1}^{\mathrm{Lag}}(v^*_{j+\frac12})^-\right] - \left[(X)_j^{\mathrm{Lag}}(v^*_{j-\frac12})^- + (X)_{j-1}^{\mathrm{Lag}}(v^*_{j-\frac12})^+\right]\Big\}.
$$

Taking $X = \rho$ in (1.45) gives with (1.42)

$$\begin{cases} \rho_j^{n+1} = \rho_j^n - \mathscr{L}^a(\rho^{\mathrm{Lag}}, v^*) + \mathcal{O}(\epsilon), \\ v^*_{j+\frac{1}{2}} = \dfrac{1}{\alpha}\left(g - 2\dfrac{\Pi_{j+1}^{\mathrm{Lag}} - \Pi_j^{\mathrm{Lag}}}{(\rho_j^n + \rho_{j+1}^n)\Delta x}\right), \\ P^*_{j+\frac{1}{2}} = \dfrac{\Pi_j^{\mathrm{Lag}} + \Pi_{j+1}^{\mathrm{Lag}}}{2} \\ \Pi_j^{\mathrm{Lag}} = \Pi_j^n - a^2\dfrac{\Delta t}{\Delta m_j}(v^*_{j+\frac{1}{2}} - v^*_{j-\frac{1}{2}}) + \mathcal{O}(\epsilon). \end{cases} \tag{1.46}$$

Then taking $X = \rho E$ in (1.45), that is $X = \rho e + \mathcal{O}(\epsilon)$ thanks to (1.41) and (1.44), we inject the last equation of (1.42)

$$\begin{cases} (\rho e)_j^{n+1} = (\rho e)_j^n - \dfrac{\Delta t}{\Delta x}\left((vP)^*_{j+\frac{1}{2}} - (vP)^*_{j-\frac{1}{2}}\right) \\ \qquad\qquad + \rho_j^n \Delta t\left(g\dfrac{v^*_{j+\frac{1}{2}} + v^*_{j-\frac{1}{2}}}{2} - \alpha\dfrac{(v^*_{j+\frac{1}{2}})^2 + (v^*_{j-\frac{1}{2}})^2}{2}\right) - \mathscr{L}^a((\rho e)^{\mathrm{Lag}}, v^*) + \mathcal{O}(\epsilon). \end{cases} \tag{1.47}$$

This limit scheme (1.46)-(1.47) is clearly consistent with (1.4) when ϵ goes to zero. Hence, the overall scheme is asymptotic preserving and the proof is completed. We recall that the limit scheme CFL condition reads (1.36) and that Π is consistent with the limit pressure equation $\partial_t \Pi + a^2 \partial_m u^1 = 0$. The non-linear stability of the limit scheme (1.42)-(1.46)-(1.47) is still an open problem.

1.8.2 Comments on the limit system (1.42)-(1.46)-(1.47)

We consider here the parabolic problem

$$\partial_t \rho + \partial_x(\rho u^1) = 0, \qquad u^1 = \frac{1}{\alpha}\left(g - \frac{1}{\rho}\partial_x p\right)$$

which corresponds to (1.4) in the barotropic case. An approach to discretize this convective diffusive system, is to split it into two sub system

$\partial_t \tau - \partial_m u^1 = 0, \qquad u^1 = \frac{1}{\alpha}(g - \partial_m p), \quad$ and $\quad \partial_t \rho + u^1 \partial_x \rho = 0, \qquad u^1 = \frac{1}{\alpha}\left(g - \frac{1}{\rho}\partial_x p\right).$

Following this splitting of operator, a two step implicit-explicit numerical strategy consists in solving the first system with an implicit scheme and then solving the second system with an explicit scheme. In the implicit step we may use a linearisation of the pressure given by its own evolution law to obtain a linear scheme.

Since the first step scheme is implicit, we assume that it has no CFL restriction. The second step scheme is explicit and stable under the CFL condition $\frac{\Delta t}{\Delta x}\left((u^1_{j-\frac{1}{2}})^+ - (u^1_{j+\frac{1}{2}})^-\right) < 1$, Thus the overall scheme for the parabolic problem is stable under this hyperbolic CFL condition in $\mathcal{O}(\frac{\Delta t}{\Delta x})$.

The limit scheme (1.42)-(1.46)-(1.47) corresponds to such a splitting of operator discretization of the parabolic limit system (1.4). It is thus natural to recover an hyperbolic CFL condition.

1.8.3 Proof of *(iv)*

We now prove that the explicit scheme LP-EXEX satisfies a discrete entropy inequality. As above, we first prove such an inequality in Lagrangian coordinates considering the scheme (1.26) that may be

recast into

$$\left\{\begin{array}{l} \tau_j^{\mathrm{Lag}} = \tau_j^n + \dfrac{\Delta t}{\Delta m_j}(u^*_{j+\frac12} - u^*_{j-\frac12}), \\ u_j^{\mathrm{Lag}} = u_j^n - \dfrac{\Delta t}{\Delta m_j}(p^*_{j+\frac12} - p^*_{j-\frac12}) + \dfrac{\Delta t}{\Delta m_j}\Delta m_{j-1/2}\left(g - \alpha u_{j-\frac12}\right), \\ \qquad\qquad + \dfrac{\Delta t}{\Delta m_j}\Delta m_{j+1/2}\left(g - \alpha u_{j+\frac12}\right) \\ E_j^{\mathrm{Lag}} = E_j^n - \dfrac{\Delta t}{\Delta m_j}\left((up)^*_{j+\frac12} - (up)^*_{j-\frac12}\right) + g\dfrac{\Delta t}{\Delta m_j}\dfrac{\Delta m_{j+1/2}u^*_{j+\frac12} + \Delta m_{j-1/2}u^*_{j-\frac12}}{2} \\ \qquad\qquad - \alpha\dfrac{\Delta t}{\Delta m_j}\dfrac{\Delta m_{j+1/2}(u^*_{j+\frac12})^2 + \Delta m_{j-1/2}(u^*_{j-\frac12})^2}{2}, \end{array}\right. \tag{1.48}$$

with

$$\left\{\begin{array}{l} u^*_{j+\frac12} = \dfrac{1}{2a + \alpha\Delta m_{j+1/2}}\Big(a(u^n_{j+1} + u^n_j) - \big(p(\rho^n_{j+1}, e^n_{j+1}) - p(\rho^n_j, e^n_j)\big) + g\Delta m_{j+1/2}\Big), \\ p^*_{j+\frac12} = \dfrac{p(\rho^n_{j+1}, e^n_{j+1}) + p(\rho^n_j, e^n_j)}{2} - \dfrac{a(u^n_{j+1} - u^n_j)}{2}. \end{array}\right.$$

Then we take the Eulerian projection (1.34) into account to obtain a discrete entropy inequality for the overall scheme in Eulerian coordinates.
Let $s = s(\rho, e)$ be the strictly convex mathematical entropy. The entropy inequality associated with (1.1) writes

$$\partial_t(\rho s) + \partial_x(\rho u s) \leq 0.$$

In Lagrangian coordinates, (1.7) is associated with the following entropy inequality

$$\partial_t s \leq 0. \tag{1.49}$$

In the sequel and with a little abuse in the notations, we will consider the pressure as a function of the density and the entropy $p = p(\rho, s)$.

Entropy inequality in Lagrangian coordinates

The scheme (1.48) is a Godunov-type scheme. We prove here that the associated approximate Riemann solver is consistent with the entropy inequality (1.49). Let us check that

$$0 \leq -a(s^*_L - s_L) + a(s_R - s^*_R), \tag{1.50}$$

where $s^*_L = s(\rho^*_L, e^*_L)$ and $s^*_R = s(\rho^*_R, e^*_R)$ so that (1.19) is verified with $\eta = s$, $q = 0$ and $\tilde\sigma(\Delta m, \Delta t; \mathbf{U}_L, \mathbf{U}_R) = 0$. Let us first prove the following result.

Proposition 1. *If $a > 0$ is such that*

$$\left\{\begin{array}{l} \rho^*_L > 0, \qquad \rho^*_R > 0, \\ \rho^2\partial_\rho p(\rho, s_L) \leq a^2, \quad \forall \rho \in I(\rho_L, \rho^*_L), \\ \rho^2\partial_\rho p(\rho, s_R) \leq a^2, \quad \forall \rho \in I(\rho_R, \rho^*_R), \end{array}\right.$$

then we have

$$e^*_L \geq e(\rho^*_L, s_L), \qquad e^*_R \geq e(\rho^*_R, s_R).$$

Proof. Recall that we have

$$\begin{cases} E_L^* = E_L + \dfrac{1}{a}\Big(p_L u_L - u^*[p^* - \dfrac{\Delta m}{2}(g - \alpha u^*)]\Big), \\ E_R^* = E_R + \dfrac{1}{a}\Big(u^*[p^* + \dfrac{\Delta m}{2}(g - \alpha u^*)] - p_R u_R\Big), \end{cases}$$

and $\Pi_R^* = p^* + \frac{\Delta m}{2}(g - \alpha u^*)$, $\quad \Pi_L^* = p^* - \frac{\Delta m}{2}(g - \alpha u^*)$, which gives

$$\begin{cases} E_L^* = E_L - \dfrac{1}{a}(\Pi_L^* u^* - \Pi_L u_L), \\ E_R^* = E_R + \dfrac{1}{a}(\Pi_R^* u^* - \Pi_R u_R), \end{cases}$$

so that $e_R^* = E_R^* - \frac{1}{2}u^{*2} = e_R + \frac{1}{2}(u_R^2 - u^{*2}) + \frac{1}{a}(\Pi_R^* u^* - \Pi_R u_R)$. Using the definitions of u^* and Π_R^* given in (1.25), straightforward calculations then lead to $e_R^* = e_R + \frac{1}{2a^2}(\Pi_R^{*2} - \Pi_R^2)$ and $\Pi_R^* = \Pi_R + a(u^* - u_R)$. The latter equality, together with the last two equalities in (1.24), gives $\frac{\Pi_R^*}{a^2} = \frac{1}{\rho_R} + \frac{p_R}{a^2} - \frac{1}{\rho_R^*}$.
It is then not difficult to check that

$$\begin{cases} e_R^* - e(\rho_R^*, s_R) = \dfrac{1}{2a^2}(p(\rho_R^*, s_R) - \Pi_R^*)^2 + \phi(\rho_R), \\ \phi(\rho) = e(\rho, s_R) - \dfrac{p(\rho, s_R)^2}{2a^2} - e(\rho_R^*, s_R) + \dfrac{p(\rho_R^*, s_R)^2}{2a^2} + p(\rho_R^*, s_R)\left(\dfrac{1}{\rho} + \dfrac{p(\rho, s_R)}{a^2} - \dfrac{1}{\rho_R^*} - \dfrac{p(\rho_R^*, s_R)}{a^2}\right), \end{cases}$$

with, using the well-known relation (1.2), $\phi'(\rho) = (p(\rho, s_R) - p(\rho_R^*, s_R))(\frac{1}{\rho^2} - \frac{1}{a^2}\partial_\rho p(\rho, s_R))$, $\quad \phi(\rho_R^*) = 0$. Therefore $\phi(\rho) \geq 0$ for all $\rho \in I(\rho_R, \rho_R^*)$ since $\partial_\rho p \geq 0$ under the assumptions of Weyl and by (1.9). We have thus proved in particular that

$$e_R^* \geq e(\rho_R^*, s_R).$$

The proof of the second inequality follows the same idea and the proof of the proposition is then completed.

□

In order to get (1.50), we then note that we have $\partial_e s(\rho, e) < 0$ where $s = s(\rho, e(\rho, s))$. So that $s_R - s_R^* = s(\rho_R^*, e(\rho_R^*, s_R)) - s(\rho_R^*, e_R^*) = \partial_e s(\rho_R^*, \overline{e}_L^*)(e(\rho_R^*, s_R) - e_R^*) \geq 0$,

$$s_R - s_R^* \geq 0.$$

We get in the same way $s_L - s_L^* \geq 0$ which gives (1.50) since $a > 0$. Then, under the CFL condition (1.28), the scheme in Lagrangian coordinates satisfies the discrete entropy inequality (1.20) which reads

$$s_j^{\text{Lag}} \leq s_j^n - \frac{\Delta t}{\Delta m_j}(q_{j+\frac{1}{2}}^n - q_{j-\frac{1}{2}}^n), \tag{1.51}$$

where $q_{j+\frac{1}{2}}^n$ naturally follows from (1.21).

Entropy inequality in Eulerian coordinates

Let us now prove an entropy inequality for the whole Lagrange-Projection scheme. We first recall that the function $\rho s = \rho s(\rho, \rho u, \rho E)$ is strictly convex by assumption, see [32]. The Eulerian projection (1.34) being a convex combination for each $X = \rho, \rho u, \rho E$, we thus have

$$(\rho s)_j^{n+1} \leq (\rho s)_j^{\text{Lag}} + \frac{\Delta t}{\Delta x}\left((u_{j-\frac{1}{2}}^*)^+(\rho s)_{j-1}^{\text{Lag}} + \left((u_{j+\frac{1}{2}}^*)^- - (u_{j-\frac{1}{2}}^*)^+\right)(\rho s)_j^{\text{Lag}} - (u_{j+\frac{1}{2}}^*)^-(\rho s)_{j+1}^{\text{Lag}}\right).$$

Then using the first equation of (1.48) in (1.51), together with the definition of Δm_j, we have

$$(\rho s)_j^{\mathrm{Lag}} \leq (\rho s)_j^n - \frac{\Delta t}{\Delta x}(q_{j+\frac{1}{2}}^n - q_{j-\frac{1}{2}}^n) - \frac{\Delta t}{\Delta x}(\rho s)_j^{\mathrm{Lag}}(u_{j+\frac{1}{2}}^* - u_{j-\frac{1}{2}}^*).$$

Then using these two inequalities together with the simple relation $u = u^+ + u^-$, we get

$$(\rho s)_j^{n+1} \leq (\rho s)_j^n - \frac{\Delta t}{\Delta x}(g_{j+\frac{1}{2}}^n - g_{j-\frac{1}{2}}^n),$$

where $g_{j+\frac{1}{2}}^n = (u_{j+\frac{1}{2}}^*)^+(\rho s)_j^{\mathrm{Lag}} + (u_{j+\frac{1}{2}}^*)^-(\rho s)_{j+1}^{\mathrm{Lag}} + q_{j+\frac{1}{2}}^n$.
Which is nothing but the expected discrete entropy inequality for the whole Lagrange-Projection scheme (note indeed that $g_{j+\frac{1}{2}}^n$ is clearly consistent with the entropy flux $\rho s u$).

1.9 Numerical results

For the sake of stressing the importance of the source term discretization, we propose to also consider in the following a simpler approximation strategy that will be referred to as the LP-EXEX SP scheme. In this scheme, the source terms will be treated by means of a separate operator splitting. Suppose that for some instant n we know $(\rho_j^n, (\rho u)_j^n, (\rho E)_j^n)$ for $j \in \mathbb{Z}$, the LP-EXEX SP numerical scheme reads :

(i) compute $(\tau_j^n, \overrightarrow{w}_j^n, \overleftarrow{w}_j^n, E_j^n)$ at equilibrium by evaluating $\Pi_j^n = p(\rho_j^n, e_j^n)$,

(ii) compute $(\tau_j^{Lag}, \overrightarrow{w}_j^{Lag}, \overleftarrow{w}_j^{Lag}, E_j^{Lag})$ thanks to the use the explicit scheme defined by (1.26) with $\alpha = 0$ and $g = 0$, so that the gravity and friction terms are not taken into account yet

(iii) evaluate $(\rho_j^{Lag}, (\rho u)_j^{Lag}, (\rho E)_j^{Lag})$ thanks to
$(\tau_j^{Lag}, \overrightarrow{w}_j^{Lag}, \overleftarrow{w}_j^{Lag}, E_j^{Lag})$,

(iv) compute $(\rho_j^{n+1,\sharp}, (\rho u)_j^{n+1,\sharp}, (\rho E)_j^{n+1,\sharp})$ thanks to the explicit scheme defined by (1.34) and (1.27),

(v) account for the gravity and friction terms by integrating the system of ordinary differential equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\rho\\ \rho u\\ \rho E\end{bmatrix} = \begin{bmatrix}0\\ \rho(g-\alpha u)\\ \rho u(g-\alpha u)\end{bmatrix}.$$

The update in term of the variable ρ, u, and e then reads

$$\rho_j^{n+1} = \rho_j^{n+1,\sharp}, \quad u_j^{n+1} = u_j^{n+1,\sharp}e^{-\alpha\Delta t} + \frac{g}{\alpha}(1 - e^{-\alpha\Delta t}), \quad e_j^{n+1} = e_j^{n+1,\sharp}.$$

(vi) evaluate $(\rho_j^{n+1}, (\rho u)_j^{n+1}, (\rho E)_j^{n+1})$.

We propose to test both LP-EXEX SP and LP-IMEX scheme against a test case that has been proposed in [16]. In the sequel, we shall consider that the fluid is equipped with a perfect gas equation of state $p = (\gamma - 1)\rho e$ and we set the gravity acceleration, the friction coefficient and the specific heat ratio to the following values $g = 9.81\,\mathrm{m}\cdot\mathrm{s}^{-2}$, $\alpha = 10^6\,\mathrm{s}^{-1}$, $\gamma = 1.4$.

The initial condition is defined by $\begin{cases}(\rho, u, p) = (1.0, 0, 10000.0), & \text{if } x \in [0, 0.35] \cap [0.65, 1],\\ (\rho, u, p) = (2.0, 0, 26390.2), & \text{if } x \in [0.35, 0.65].\end{cases}$
At the boundaries, we impose periodic boundary conditions thanks to a fictious cell at each end of the domain. In the sequel, both LP-EXEX and LP-EXEX SP computations will be performed with a time step defined by $\Delta t = \frac{\min(\rho_j^n)\Delta x}{2a}$, in order to agree with the classic acoustic CFL (1.28). The choice of the time step for the LP-IMEX scheme will be specified case by case. For each test, we will compute a reference solution thanks to the LP-EXEX scheme over a 10 000-cell grid. If we refer to this solution thanks to the superscript $^{\mathrm{ref}}$ and if Y denote a fluid variable, for the sake of comparison we shall consi-

der in the sequel the L^1 relative error with respect to the reference solution at the instant t defined by $\mathrm{err}(Y,t) = \frac{\|Y(\cdot,t)-Y^{\mathrm{ref}}(\cdot,t)\|_{L^1([0,1])}}{\|Y^{\mathrm{ref}}(\cdot,t)\|_{L^1([0,1])}}$.

1.9.1 Test case 1 : sensitivity with respect to the space step for large friction

We run our numerical tests with the LP-EXEX SP scheme using a spatial discretization over 100 cells, 1000 cells and 10 000 cells. This leads to time step values Δt that are respectively of magnitude $\frac{10}{\alpha}$, $\frac{1}{\alpha}$ and $\frac{1}{10\alpha}$. In figure 1.1, we display the result obtained at $t = 0.01\,\mathrm{s}$ and we can see that there is a large amount of numerical diffusion due to the discretization of the source term with large values of α. It is necessary to choose small values of Δt relatively to $\frac{1}{\alpha}$ in order to preserve the accuracy of the solution.

We now consider the same test performed with the LP-IMEX scheme. We now choose Δt in agreement with the CFL condition (1.35) by setting

$$\Delta t = \min\left(\frac{\Delta x}{2\max\left(u_j^n\right)}, \frac{1}{\alpha}\right), \tag{1.52}$$

so that we always have $\Delta t \leq \frac{1}{\alpha}$. The results obtained with LP-IMEX scheme at instant $t = 0.01\,\mathrm{s}$ are presented in figure 1.2 for discretization grids of 100 cells and 1000 cells. It is clear that the approximate solution is much more accurate than the one computed with the LP-EXEX SP scheme, even for coarse mesh. Let us note that for the relatively large space steps that have been used here, choice (1.52) imposes $\Delta t = \frac{1}{\alpha}$. Table 1.1 displays the relative error obtained with both the LP-EXEX SP and the LP-IMEX scheme. It shows that the asymptotic preserving property significantly lessen the numerical diffusion and improve accuracy by several orders of magnitude.

TABLE 1.1 – Comparison of the relative errors between the approximated solutions obtained with both LP-EXEX SP and LP-IMEX schemes. The space domain is discretized with a 1 000-cell space discretization and $\Delta t = \frac{1}{\alpha}$ for both schemes.

numerical scheme	$\mathrm{err}(\rho, t = 0.01)$	$\mathrm{err}(u, t = 0.01)$	$\mathrm{err}(P, t = 0.01)$
LP-EXEX SP	1.686931×10^{-2}	6.858335×10^{-1}	2.539820×10^{-2}
LP-IMEX	3.959560×10^{-4}	1.195630×10^{-2}	5.635518×10^{-4}

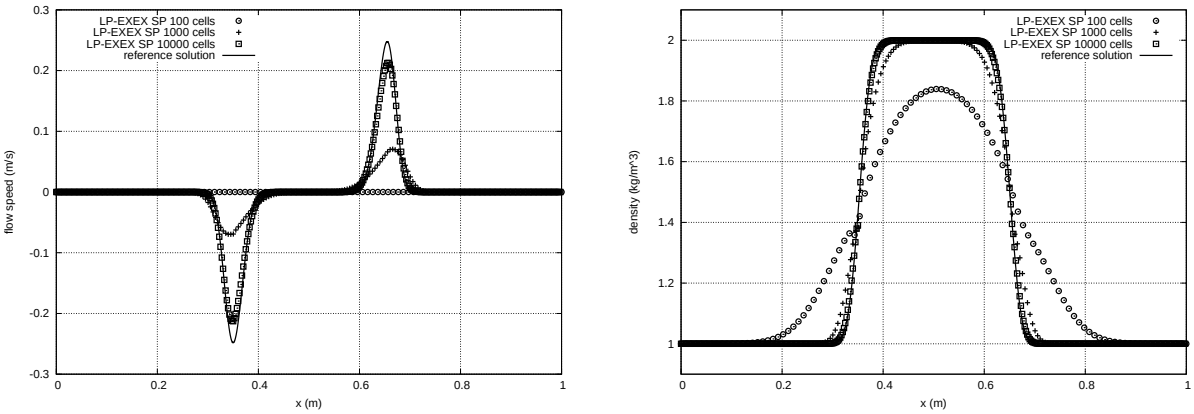


FIGURE 1.1 – Profile at time $t = 0.01\,\mathrm{s}$ of the velocity (left) and the density (right) obtained for a 100-cell, 1000-cell and 10 000-cell grid with the LP-EXEX SP scheme and the reference solution (LP-EXEX scheme with 10 000-cell mesh).

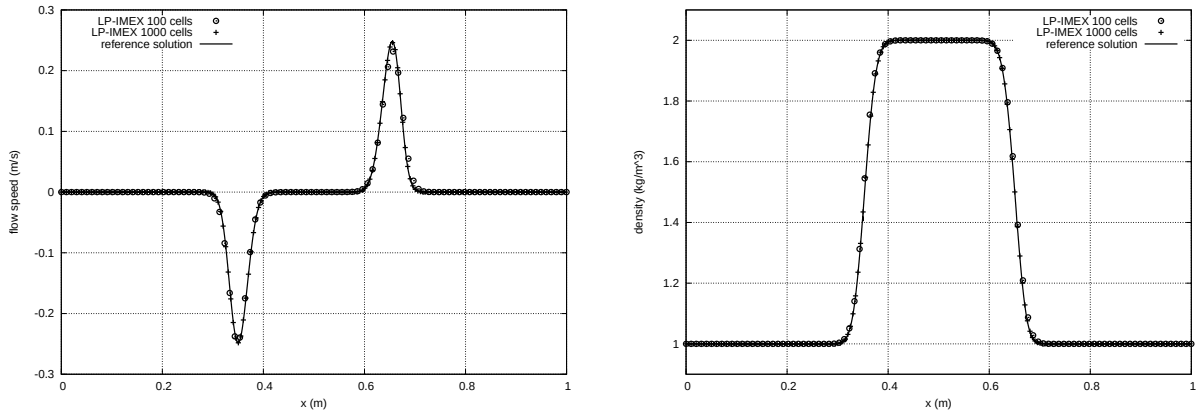


FIGURE 1.2 – Profile at time $t = 0.01\,\mathrm{s}$ of the velocity (left) and the density (right) obtained for a 100-cell and 1000-cell grid with the LP-IMEX scheme and the reference solution (LP-EXEX scheme with 10 000-cell mesh).

1.9.2 Test case 2 : sensitivity with respect to the time step

We are now interested in testing the LP-IMEX scheme in situations where Δt is much bigger than $\frac{1}{\alpha}$. In order to proceed, we relax the previous time step choice (1.52) by suppressing the control of with respect to $\frac{1}{\alpha}$. We simply choose to define Δt in agreement with the CFL condition (1.35) based on the material velocity by setting

$$\Delta t = \min\left(\frac{\Delta x}{2\max\left(u_j^n\right)}\right). \tag{1.53}$$

The test is performed with a 1000-cell mesh. The graph of the approximate solution at $t = 0.01\,\mathrm{s}$ are displayed in figure 1.3. For this grid choice, we obtain that the magnitude of Δt is $\frac{1000}{\alpha}$. Let us underline that these time steps are 1000 times larger than the time steps used in section 1.9.1. It appears that even for such relatively large time step and space step the approximate solution remains very accurate. This results is the consequence of the good behaviour of the numerical scheme for large friction coefficients. We compare in table 1.2 the relative errors of the fluid variables for this choice of time step values. We can verify here that the LP-IMEX scheme enables the use of time steps that are much larger than the acoustic based time step and also much larger than $\frac{1}{\alpha}$ while preserving accurate simulation results.

TABLE 1.2 – Comparison of the relative L^1-errors obtained with the LP-IMEX scheme for a 1 000-cell space discretization and two different Δt values.

numerical scheme	Δt	$\mathrm{err}(\rho, t = 0.01)$	$\mathrm{err}(u, t = 0.01)$	$\mathrm{err}(P, t = 0.01)$
LP-IMEX	$\frac{1}{\alpha}$	3.959560×10^{-4}	1.195630×10^{-2}	5.635518×10^{-4}
LP-IMEX	$\frac{1000}{\alpha}$	2.607495×10^{-3}	1.099137×10^{-1}	3.288768×10^{-3}

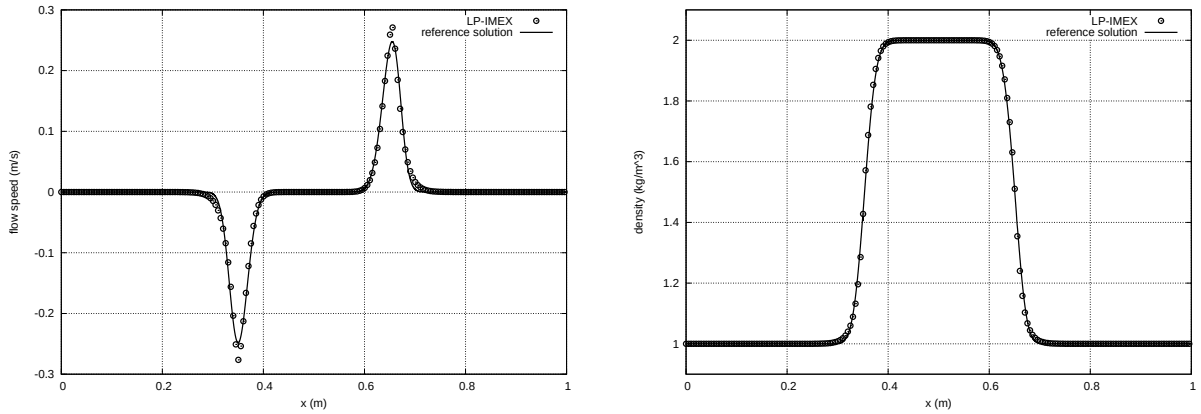


FIGURE 1.3 – Profile at time $t = 0.01\,\mathrm{s}$ of the velocity (left) and the density (right) obtained for a 1000-cell grid with the LP-IMEX scheme and the reference solution (LP-EXEX scheme with 10 000-cell mesh).

1.9.3 Test case 3 : sensitivity with respect to the friction parameter α

Previous tests use a single value of the friction parameter equal to $10^6 s^{-1}$. Here, we run tests with the LP-IMEX scheme using a spatial discretization over 1000 cells and friction parameter values $10^5 s^{-1}$, $10^6 s^{-1}$ and $10^7 s^{-1}$. We can observe on figure 1.4, figure 1.3 and figure 1.5, that numerical results are close to their respective reference solution for each value of α. Besides, numerical simulations took respectively 37, 14 and 5 iterations to reach time 0.01s for the friction parameter $10^5 s^{-1}$, $10^6 s^{-1}$ and $10^7 s^{-1}$. This confirms the good behavior of the LP-IMEX scheme CFL condition (1.35) for large friction parameter. Indeed, the larger the friction parameter the smaller the velocity and so the bigger the time step.

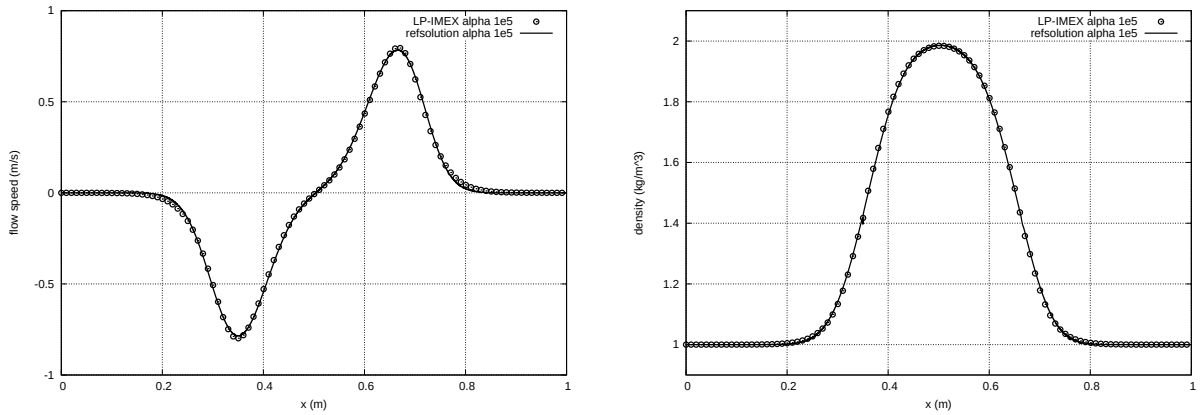


FIGURE 1.4 – Profile at time $t = 0.01\,\mathrm{s}$ of the velocity (left) and the density (right) obtained for a 1000-cell grid and $10^5 s^{-1}$ friction parameter with the LP-IMEX scheme and the reference solution (LP-EXEX scheme with 10 000-cell mesh).

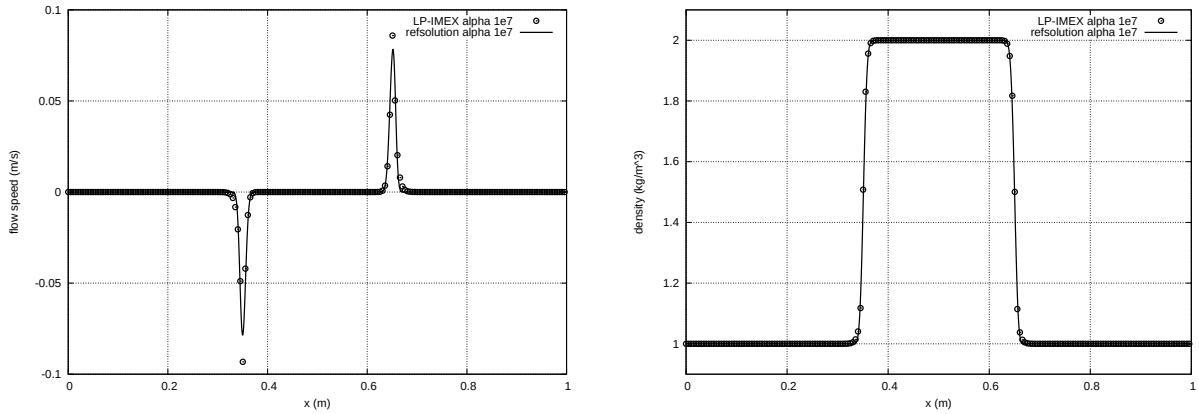


FIGURE 1.5 – Profile at time $t = 0.01\,\mathrm{s}$ of the velocity (left) and the density (right) obtained for a 1000-cell grid and $10^7 s^{-1}$ friction parameter with the LP-IMEX scheme and the reference solution (LP-EXEX scheme with 10 000-cell mesh).

Conclusion

We designed a large time step and asymptotic preserving scheme for the gas dynamics equations with gravity and linear friction. The stability is proved under a time step CFL restriction based on the velocity u only and not on the sound speed c. The scheme gives very good results and combines accuracy thanks to the asymptotic preserving property and efficiency thanks to the large time step stability condition. Future developments include an extension to several space dimensions and high-order accuracy, to low-Mach number flows and to more complicated systems of partial differential equations used in the modeling of two phase flows.

Acknowledgement. *This present work has been partially achieved within the framework of the research project LATSAP. This project was funded by the CEA/DEN/DANS/DM2S for the 2011 edition of the CEMRACS.*

Bibliographie

[1] R. Abgrall, D. Aregba, C. Berthon, M. Castro and C. Parés, Special issue : Numerical approximations of hyperbolic systems with source terms and applications, J. Sci. Comput. 48, (2011).

[2] G. Basque, L. Delapierre, D. Grand and P. Mercier, *BACCHUS; A numerical approach to two-phase flow in a rod bundle*, Nuclear Engineering and Design, 82(2-3) : 191–204, (1984).

[3] C. Berthon, P. Charrier and B. Dubroca, *An HLLC scheme to solve the M1 Model of radiative transfer in two space dimensions*, J. Sci. Comput., 31(3) : 347–389, (2007).

[4] C. Berthon, P.G. LeFloch and R. Turpault, *Late-time relaxation limits of nonlinear hyperbolic systems. A general framework*, Math. of Comput., 82 : 831–860, (2013).

[5] C. Berthon and R. Turpault, *Asymptotic preserving HLL schemes.* Numer, Methods Partial Differential Equations, 27(6) : 1396–1422, (2011).

[6] S. Bianchini, B. Hanouzet and R. Natalini, *Asymptotic behavior of smooth solutions for partially dissipative hyperbolic systems with a convex entropy*, Communications on Pure and Applied Mathematics, 60(11) : 1559–1622, (2007).

[7] R. Botchorishvili, B.Perthame and A.Vasseur, *Equilibrium schemes for scalar conservation laws with stiff sources*, Math. of Comp., 72 : 131–157, (2003).

[8] F. Bouchut, *A reduced stability condition for nonlinear relaxation to conservative laws*, J. Hyp. Diff. Eq., 1(1) : 149–170, (2004).

[9] F. Bouchut, H. Ounaissa and B. Perthame, *Upwinding of the source term at interfaces for Euler equations with high friction*, Comput. Math. Appl., 53 : 361–375, (2007).

[10] A.-C. Boulanger, C. Cancès, H. Mathis, K. Saleh and N. Seguin, *OS-AMOAL : Optimized simulations by adapted models using asymptotic limits*, ESAIM : Proceedings, 38 : 183–201, (2012).

[11] C. Buet and S. Cordier, *An asymptotic preserving scheme for hydrodynamics radiative transfer models : numerics for radiative transfer*, Numer. Math., 108 : 199-221, (2007).

[12] C. Buet and B. Després, *Asymptotic preserving and positive schemes for radiation hydrodynamics*, J. Comput. Phys., 215(2) : 717–740, (2006).

[13] P. Cargo and A.-Y. Le Roux, *Un schéma équilibre adapté au modèle d'atmosphère avec termes de gravité.*, C. R. Acad. Sci. Paris, Série I 318 : 73–76, (1994).

[14] J.-A. Carillo, T. Goudon and P. Lafitte, *Simulation of fluid and particles flows : Asymptotic preserving schemes for bubbling and flowing regimes*, Journal of Computational Physics, 227(16) : 7929–7951, (2008).

[15] J.-A. Carrillo, Th. Goudon, P. Lafitte and F. Vecil, *Numerical schemes of diffusion asymptotics and moment closures for kinetic equations*, J. Sci. Comput, 36(1) : 113–149, (2008).

[16] C. Chalons, F. Coquel, E. Godlewski, P-A Raviart and N. Seguin, *Godunov-type schemes for hyperbolic systems with parameter dependent source. The case of Euler system with friction*, Math. Models Methods Appl. Sci., 20(11) : 2109–2166, (2010).

[17] C. Chalons and J.F. Coulombel, *Relaxation approximation of the Euler equations*, Journal of Mathematical Analysis and Applications, 348(2) : 872–893, (2008).

[18] C. Chalons and F. Coquel, *Navier-stokes equations with several independant pressure laws and explicit predictor-corrector schemes*, Numerisch Math, 101(3) : 451–478, (2005).

[19] F. Coquel, Q.L. Nguyen, M. Postel, and Q.H. Tran, *Entropy-satisfying relaxation method with large time-steps for Euler IBVPs*, Math. Comp., 79 : 1493–1533, (2010).

[20] F. Coquel, E. Godlewski, B. Perthame, A. In and P. Rascle, Some new Godunov and relaxation methods for two-phase flow problems, In Godunov methods, Springer, 179–188, (2001).

[21] J.-F. Coulombel and T. Goudon, *The strong relaxation limit of the multidimensional isothermal Euler equations*, Trans. Amer. Math. Soc., 359(2) : 637–648, (2007).

[22] N. Crouseilles and M. Lemou, *An asymptotic preserving scheme based on a micro-macro decomposition for collisional Vlasov equations : diffusion and high-field scaling limits*, KRM, 4 : 441–477, (2011).

[23] P. Degond, F. Deluzet, M.-H. Vignal, L. Navoret and A-B Sun, *Asymptotic-Preserving Particle-In-Cell method for the Vlasov-Poisson system near quasineutrality*, Journal of Computational Physics, 229(16) 5630–5652, (2010).

[24] P. Degond, F. Deluzet, M.-H. Vignal, A. Sangam, *An Asymptotic Preserving Scheme for the Euler equations in a strong magnetic field*, Journal of Computational Physics, 228(10) : 3540-3558, (2009).

[25] S. Dellacherie, *Analysis of Godunov type schemes applied to the compressible Euler system at low Mach number*, J. Comp. Phys., 229(4) : 978–1016, (2010).

[26] S. Dellacherie, P. Omnes and F. Rieper, *The influence of cell geometry on the Godunov scheme applied to the linear wave equation*, J. Comp. Phys., 229(14) : 5315–5338, (2010).

[27] B. Després, C. Buet and E. Franck, *Design of asymptotic preserving finite volume schemes for the hyperbolic heat equation on unstructured meshes*, Numerische Math., 122(2) : 227–278, (2012).

[28] G. Dimarco and L. Pareschi, *Exponential methods for kinetic equations*, SIAM J. Num. Anal, 49 : 2057–2077, (2011).

[29] F. Filbet and S. Jin, *A class of asymptotic preserving schemes for kinetic equations and related problems with stiff sources*, J. Comp. Phys., 229(20) : 7625–7648, (2010).

[30] G. Gallice, *Solveurs simples positifs et entropiques pour les systèmes hyperboliques avec terme source*, C. R. Math. Acad. Sci. Paris, 334(8) : 713–716, (2002).

[31] G. Gallice, *Positive and entropy stable Godunov-type schemes for gas dynamics and MHD equations in Lagrangian or Eulerian coordinates*, Numer. Math., 94(4) : 673–713, (2003).

[32] E. Godlewski and P.-A. Raviart. Numerical approximation of hyperbolic systems of conservation laws. Applied Mathematical Sciences, Springer-Verlag, New York, (1996).

[33] L. Gosse, *A priori error estimate for a well-balanced scheme designed for inhomogeneous scalar conservation laws*, C. R. Acad. Sci. Paris, 327(5) : 467-472, (1998).

[34] L. Gosse and A.-Y. Le Roux, *A well-balanced scheme designed for inhomogeneous scalar conservation laws*, C. R. Acad. Sci. Paris, 323(5) : 543-546, (1996).

[35] L. Gosse and G. Toscani, *Space Localization and Well-Balanced Schemes for Discrete Kinetic Models in Diffusive Regimes*, SIAM J. Numer. Anal., 41(2) : 641–658, (2003).

[36] L. Gosse and G. Toscani, *Asymptotic-preserving and well-balanced schemes for radiative transfer and the Rosseland approximation*, Numer. Math., 98(2) 223–250, (2004).

[37] J.-M. Greenberg and A.-Y. Le Roux, *A well-balanced scheme for the numerical processing of source terms in hyperbolic equations*, SIAM, J. of Num. Anal., 33 : 1–16, (1996).

[38] J. Haack, S. Jin and J.-G. Liu, *An all-speed asymptotic-preserving method for the isentropic Euler and Navier-Stokes equation*, Commun. Comp. Phys., 12 : 955–980, (2012).

[39] L. Hsiao, T.-P. Liu, *Convergence to nonlinear diffusion waves for solutions of a system of hyperbolic conservation laws with damping*, Comm. Math. Phys., 143(3) : 599–605, (1992).

[40] S. Jin, *Runge-Kutta methods for hyperbolic conservation laws with stiff relaxation terms*, J. Comput. Phys., 122 : 51–67, (1995).

[41] S. Jin, *Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations*, SIAM J. Sci. Comput., 21 : 441–454, (1999).

[42] S. Jin, *A steady-state capturing method for hyperbolic systems with geometrical source terms*, M2AN, 35(4) : 631–645, (2001).

[43] S. Jin and C.D. Levermore, *Numerical Schemes for Hyperbolic Conservation Laws with Stiff Relaxation Terms*, J. Computational Physics, 126 : 449–467, (1996).

[44] S.Jin, L. Pareschi and G. Toscani, *Uniformly accurate diffusive relaxation schemes for multiscale transport equations*, SIAM J. Numerical Analysis, 38(13) : 913–936, (2000).

[45] S. Jin and Z. P. Xin. *The relaxation schemes for systems of conservation laws in arbitrary space dimension*. Comm. Pure Appl. Math., 48(3) : 235–276, (1995).

[46] S. Junca and M. Rascle. *Strong relaxation of the isothermal Euler system to the heat equation*. Z. Angew. Math. Phys., 53(2) : 239–264, (2002).

[47] Th. Katsaounis, B. Perthame and C. Simeoni, *Upwinding Sources at Interfaces in Conservation Laws*, Applied Mathematics Letters, 17 : 309–316, (2004).

[48] A. Klar, *An asymptotic-induced scheme for nonstationary transport equations in the diffusive limit*, SIAM J. Numer. Anal., 35(3) : 1073–1094, (1998).

[49] C. Lin and J.-F. Coulombel, *The strong relaxation limit of the multidimensional Euler equations*, Nonlinear Differential Equations and Applications, 20 : 447–461, (2012).

[50] P. Marcati and A. Milani, *The one-dimensional Darcys law as the limit of a compressible Euler flow*, J. Differential Equations, 84(1) : 129–147, (1990).

[51] G. Naldi and L. Pareschi, *Numerical schemes for hyperbolic systems of conservation laws with stiff diffusive relaxation*, SIAM J. Numer. Anal, 37(4) : 1246–1270, (2000).

[52] B. Perthame and C. Simeoni, *A kinetic scheme for the Saint-Venant system with a source term*, Calcolo, 38(4) : 201–231, (2001).

[53] M. F. Robbe and F. Bliard, *A porosity method to describe the influence of internal structures on a fluid flow in case of fast dynamics problems*. Nuclear Engineering and Design, 215 : 217–242, (2002).

[54] W. Sha, B.T Chao and S.L Soo, *Porous-media formulation for multiphase flow with heat transfer*, Nuclear Engineering and Design, 82(2-3) : 93–106, (1984).

[55] I. Suliciu, *On the thermodynamics of fluids with relaxation and phase transitions. Fluids with relaxation*, Int. J. Engag. Sci., 36 : 921–947, (1998).

[56] H. Weyl, *Shock waves in arbitrary fluids*, Comm. Pure Appl. Math, 2 : 103–122, (1949).

Chapitre 2

Schémas de splitting d'opérateur préservant l'asymptotique pour le système de la dynamique des gaz avec termes sources raides

Ce chapitre a fait l'objet d'une publication dont les références sont : C. Chalons, M. Girardin and S. Kokh, *Operator-splitting-based asymptotic preserving scheme for the gas dynamics equations with stiff source terms*, AIMS on Applied Mathematics, Proceedings of the 2012 International Conference on Hyperbolic Problems : Theory, Numerics, Applications, 8 : 607–614 , (2014).

Dans le chapitre 1, on a proposé un schéma numérique basé sur la consistance au sens intégral. Le calcul a montré que ce schéma est *asymptotic preserving*. Ce processus de construction est spécifique à ce système et cette asymptotique. Dans ce chapitre, on va s'intéresser à une méthode de construction de schémas préservant l'asymptotique plus générale qui pourra être utilisée pour étudier d'autres systèmes et d'autres régimes asymptotiques, comme par exemple le régime des faibles nombres de Mach.

Les annexes viennent compléter l'article *Operator-splitting-based asymptotic preserving scheme for the gas dynamics equations with stiff source terms*, en apportant un éclairage particulier sur le traitement des conditions aux limites dans la section 2.A, la précision des schémas dans les régimes intermédiaires dans la section 2.B et l'extension à l'ordre 2 en espace dans la section 2.C.

Operator-splitting-based asymptotic preserving scheme for the gas dynamics equations with stiff source terms

Abstract

We propose a numerical scheme for the gas dynamics equations with external forces and friction terms that is able to accurately approximate the flow in the large friction regime with a coarse spatial discretization. The key idea is to use a classic convection/source operator splitting and then to reduce the numerical diffusion involved with the approximation of the convective terms when the friction effects are dominant. The overall resulting scheme satisfies a discrete entropy inequality under a condition on the numerical diffusion reduction. Numerical tests are proposed in 1D and 2D that show a gain of accuracy.

2.1 Introduction

We are interested in the simulation of subsonic compressible flows where the driving phenomena are stiff source terms and material transport. Such a flow configuration may be encountered in several industrial processes like the flows within the core of a nuclear reactor. We consider the gas dynamics equations with external forces and friction and propose a formal asymptotic analysis for large friction. In this context, it is well known that classic operator splitting techniques fail to produce accurate approximate solutions when the small scales are not resolved, that is to say when coarse meshes are used. For a solution whose long-time behaviour is governed by the solution of a parabolic system, asymptotic preserving numerical schemes may be used to transpose this feature to the discrete setting without using very fine meshes (see e.g. [1, 4, 5, 8, 14, 15, 16] and the references therein). The present work is dedicated to the *short time* behaviour of the solution and relies on a careful asymptotic analysis of the classic operator splitting technique. In particular, we focus on the numerical diffusion created by the upwind part of the numerical flux function associated with the treatment of the convective part. This numerical diffusion is indeed what prevents the scheme from computing accurate solutions in the asymptotic limit of large friction coefficients. Based on this analysis, we propose a numerical diffusion reduction technique : the resulting scheme is an operator splitting algorithm with a good asymptotic behaviour. It satisfies a discrete entropy inequality under a condition on the correction. The proposed diffusion reduction technique is similar to the one used for the low Mach limit in [9].

2.2 Governing equations and large friction asymptotic behaviour

Governing equations. The gas dynamics equations with gravity and friction terms in Eulerian coordinates are given by

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p) = \rho(g - \alpha u), \\ \partial_t(\rho E) + \partial_x((\rho E + p)u) = \rho u(g - \alpha u), \end{cases} \tag{2.1}$$

where ρ, u and E denote the density, the velocity and the total energy of the fluid, g the gravitational acceleration and α a friction parameter. The pressure law $p = p(\rho, e)$ is assumed to be a given function of the density ρ and the internal energy $e = E - \frac{u^2}{2}$ that satisfies the usual Weyl assumptions [21]. The sound speed c is given by $c = [\left(\frac{\partial p}{\partial \rho}\right)_e + \left(\frac{p}{\rho^2}\right)\left(\frac{\partial p}{\partial e}\right)_\rho]^{1/2}$.

Large friction asymptotic behaviour. We are interested in the behaviour of (2.1) when the friction parameter α goes to infinity. We model this flow regime by simply replacing α with $\frac{\alpha}{\epsilon}$ in (2.1), where

$\epsilon > 0$ denotes a small parameter. The asymptotic regime is obtained when $\epsilon \to 0$.

Let us then assume that the velocity u admits an asymptotic expansion in powers of ϵ of the form $u = u^0 + \epsilon u^1 + \mathcal{O}(\epsilon^2)$. Multiplying the second equation of (2.1) by ϵ and letting ϵ go to 0 gives $u^0 = 0$. Then system (2.1) reads

$$\begin{cases} \partial_t \rho + \epsilon \partial_x(\rho u^1) = \mathcal{O}(\epsilon^2), \\ \partial_x p = \rho(g - \alpha u^1) + \mathcal{O}(\epsilon), \\ \partial_t(\rho E) + \epsilon \partial_x((\rho E + p)u^1) = \epsilon \rho u^1 (g - \alpha u^1) + \mathcal{O}(\epsilon^2). \end{cases} \tag{2.2}$$

Note that compared to (2.1), the fluxes now involve the first order corrector u^1 (or equivalently the small scale contribution ϵu^1) instead of u. From a numerical point of view, our objective is to propose numerical fluxes that are able to capture those small scales in the limit $\epsilon \to 0$.

Lagrange-Source-Projection operator splitting. We first propose an operator splitting between the terms accounting for the transport waves, the acoustic waves and the source terms. More precisely and using the chain rule for the space derivatives we split up the system (2.1) into the following three subsystems. The first subsystem describes the transport process and reads

$$\begin{cases} \partial_t \rho + u \partial_x \rho = 0, \\ \partial_t(\rho u) + u \partial_x(\rho u) = 0, \\ \partial_t(\rho E) + u \partial_x(\rho E) = 0. \end{cases} \tag{2.3}$$

The second subsystem governs the acoustic phenomena, namely

$$\begin{cases} \partial_t \rho + \rho \partial_x u = 0, \\ \partial_t(\rho u) + \rho u \partial_x u + \partial_x p = 0, \\ \partial_t(\rho E) + \rho E \partial_x u + \partial_x(pu) = 0, \end{cases}$$

Or equivalently with $\tau = \frac{1}{\rho}$ the specific volume and the mass variable m such that $\tau \partial_x = \partial_m$

$$\begin{cases} \partial_t \tau - \partial_m u = 0, \\ \partial_t u + \partial_m p = 0, \\ \partial_t E + \partial_m(pu) = 0. \end{cases} \tag{2.4}$$

The third subsystem accounts for gravity and friction effects and reads

$$\begin{cases} \partial_t \tau = 0, \\ \partial_t u = g - \alpha u, \\ \partial_t E = u(g - \alpha u). \end{cases} \tag{2.5}$$

Let us mention that the acoustic/transport splitting provides a simple setting for our analysis, however it does not play a crucial role in the forthcoming developments and it could be replaced by a classic Eulerian finite volume discretization.

2.3 Naive operator splitting numerical scheme

We begin with a natural discretization of (2.1) based on the previous Lagrange/Source/Projection decomposition. Let Δx and Δt represent the constant time and space steps. We set $x_{j+\frac{1}{2}} = x_{j-\frac{1}{2}} + \Delta x$ and $t^n = t^{n-1} + \Delta t$. In the sequel, X_j^n denotes the approximate value of X at time t^n within

the cell $[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}})$. We propose a three-step numerical strategy based on the splitting of operator (2.4),(2.5),(2.3).

Regarding the Lagrange step (2.4) and using clear notations, we propose the following update formulas

$$\left\{\begin{array}{l} \tau_j^{Lag} = \tau_j^n + \frac{\Delta t}{\rho_j^n \Delta x}\left(u^*_{j+\frac{1}{2}} - u^*_{j-\frac{1}{2}}\right), \\ u_j^{Lag} = u_j^n - \frac{\Delta t}{\rho_j^n \Delta x}\left(p^*_{j+\frac{1}{2}} - p^*_{j-\frac{1}{2}}\right), \\ E_j^{Lag} = E_j^n - \frac{\Delta t}{\rho_j^n \Delta x}\left((pu)^*_{j+\frac{1}{2}} - (pu)^*_{j-\frac{1}{2}}\right), \end{array}\right. \tag{2.6}$$

where the numerical fluxes are defined by

$$\left\{\begin{array}{l} u^*_{j+\frac{1}{2}} = \frac{1}{2}\left(u^n_{j+1} + u^n_j\right) - \frac{1}{2a}\left(\Pi^n_{j+1} - \Pi^n_j\right), \\ p^*_{j+\frac{1}{2}} = \frac{1}{2}\left(\Pi^n_{j+1} + \Pi^n_j\right) - \frac{a}{2}\left(u^n_{j+1} - u^n_j\right), \\ \Pi^n_j = p(\rho^n_j, e^n_j), \end{array}\right. \tag{2.7}$$

and associated with a classic pressure relaxation process (see [3, 6, 7, 17, 18]). In order to ensure stability, the parameter a must be chosen sufficiently large according to the so-called subcharacteristic condition $a > \rho c$. This scheme is shown to be nonlinearly stable under the CFL condition

$$2a\Delta t \leq \rho_j^n \Delta x. \tag{2.8}$$

For the system (2.5), we propose a point-wise implicit evaluation

$$\left\{\begin{array}{l} \overline{\tau}_j = \tau_j^{Lag}, \\ \overline{u}_j = u_j^{Lag} + g\Delta t - \alpha \Delta t \overline{u}_j, \\ \overline{E}_j = E_j^{Lag} + g\Delta t \overline{u}_j - \alpha \Delta t (\overline{u}_j)^2. \end{array}\right. \tag{2.9}$$

This *implicit* treatment is particularly important in the large friction regime since an explicit scheme would require $\Delta t = \mathcal{O}(\frac{1}{\alpha})$, leading to $\Delta t = 0$ in the limit $\alpha \to +\infty$.

Regarding the transport step (2.4), we consider a standard upwind and time-explicit numerical scheme given by

$$X_j^{n+1} = \overline{X}_j - \frac{\Delta t}{\Delta x}\Big((u^*_{j-\frac{1}{2}})^+ \left[\overline{X}_j - \overline{X}_{j-1}\right] + (u^*_{j+\frac{1}{2}})^- \left[\overline{X}_{j+1} - \overline{X}_j\right]\Big), \tag{2.10}$$

where $X \in \{\rho, \rho u, \rho E\}$ and $u^\pm = \frac{u \pm |u|}{2}$ (see for instance [13]). The update formula (2.10) is shown to be stable under the CFL condition

$$\Delta t\Big((u^*_{j-\frac{1}{2}})^+ - (u^*_{j+\frac{1}{2}})^-\Big) < \Delta x. \tag{2.11}$$

For the sake of clarity, let us briefly recall the different steps of the overall method that shall be referred to as LSP-IMEX. Assume that $(\rho_j^n, (\rho u)_j^n, (\rho E)_j^n)$ is known, $(\rho_j^{n+1}, (\rho u)_j^{n+1}, (\rho E)_j^{n+1})$ is computed by the following three steps :

(i) compute $(\tau_j^{Lag}, u_j^{Lag}, E_j^{Lag})$ from $(\rho_j^n, (\rho u)_j^n, (\rho E)_j^n)$ with (2.6)-(2.7),

(ii) compute $(\overline{\tau}_j, \overline{u}_j, \overline{E}_j)$ from $(\tau_j^{Lag}, u_j^{Lag}, E_j^{Lag})$ with (2.9),

(iii) compute $(\rho_j^{n+1}, (\rho u)_j^{n+1}, (\rho E)_j^{n+1})$ from $(\overline{\tau}_j, \overline{u}_j, \overline{E}_j)$ with (2.10).

Asymptotic analysis. Let us study the asymptotic behaviour of this scheme in the large friction limit. For the sake of simplicity, we focus only on the first equation governing the evolution of τ. Let us

first observe that for the continuous setting in Lagrangian coordinates, the first equation of (2.2) reads

$$\partial_t \tau + \epsilon \partial_m u^1 = \mathcal{O}(\epsilon^2). \tag{2.12}$$

We now perform a similar asymptotic analysis in the discrete setting by replacing α with $\frac{\alpha}{\epsilon}$ and assuming expansions of the form $u_j^n = u_j^{n,(0)} + \epsilon u_j^{n,(1)} + \mathcal{O}(\epsilon^2)$. Injecting these expansions in the first equation of (2.7) gives

$$u^*_{j+\frac{1}{2}} = \frac{1}{2}\left(u_{j+1}^{n,(0)} + u_j^{n,(0)}\right) - \frac{\Delta x}{2a}\left(\frac{\Pi_{j+1}^n - \Pi_j^n}{\Delta x}\right) + \frac{\epsilon}{2}\left(u_{j+1}^{n,(1)} + u_j^{n,(1)}\right) + \mathcal{O}(\epsilon^2).$$

Then, if $u_j^{n,(0)} = 0$ the first equation of (2.6) reads

$$\tau_j^{Lag} = \tau_j^n + \frac{\epsilon \Delta t}{\rho_j^n \Delta x}\left(\frac{1}{2}\left(u_{j+1}^{n,(1)} + u_j^{n,(1)}\right) - \frac{1}{2}\left(u_j^{n,(1)} + u_{j-1}^{n,(1)}\right)\right)$$

$$-\frac{\Delta t \Delta x}{\rho_j^n}\left(\frac{\Pi_{j+1}^n - 2\Pi_j^n + \Pi_{j-1}^n}{2a(\Delta x)^2}\right) + \mathcal{O}(\epsilon^2).$$

The above equation is clearly not consistent with (2.12) because of the third term which is of order 1 with respect to ϵ. This term is dominant compared to the effects of order ϵ we are interested in. Nonetheless, it is important to note that this problem no longer occurs if $\Delta x = \mathcal{O}(\epsilon)$. In other words, such a scheme requires a fine mesh to get accurate numerical results in the regime $\epsilon \to 0$.

Moreover, the second equation of (2.9) leads to

$$\overline{u}_j = \frac{u_j^{Lag} + g\Delta t}{1 + \alpha \epsilon^{-1} \Delta t},$$

which implies that $\overline{u}_j^{(0)} = 0$. Then (2.10) gives $u_j^{n+1,(0)} = 0$.

It is important to emphasize that the bad asymptotic behaviour of the scheme comes from the upwind part of the numerical fluxes and not from the source terms discretization based on an operator splitting technique. The non-centred part of the flux $u^*_{j+\frac{1}{2}}$ may be interpreted as numerical diffusion that becomes predominant in the asymptotic regime $\epsilon \to 0$ and prevents the scheme from capturing the right interface velocity of order ϵ. As briefly discussed above, a possible way to circumvent this difficulty consists in using a grid size of order ϵ to reduce the importance of the non-centred term in $u^*_{j+\frac{1}{2}}$. Of course, this cannot be envisaged in practice. In the next section, we propose a correction of $u^*_{j+\frac{1}{2}}$ involving the friction parameter α to obtain a scheme that provides good numerical results for large friction test cases with coarse meshes. The proposed correction can be easily understood as a numerical diffusion reduction technique.

2.4 Suitable operator splitting numerical scheme and numerical diffusion reduction technique.

Numerical diffusion reduction technique. We propose a correction of the previous scheme to obtain the expected large friction asymptotic behaviour. We replace the definition of the relaxation fluxes (2.7) by

$$\begin{cases} u^*_{j+\frac{1}{2}} = \frac{1}{2}\left(u_{j+1}^n + u_j^n\right) - \frac{\theta_{j+\frac{1}{2}}}{2a}\left(\Pi_{j+1}^n - \Pi_j^n\right), \\ p^*_{j+\frac{1}{2}} = \frac{1}{2}\left(\Pi_{j+1}^n + \Pi_j^n\right) - \frac{a}{2}\left(u_{j+1}^n - u_j^n\right), \quad \Pi_j^n = p(\rho_j^n, e_j^n), \end{cases} \tag{2.13}$$

and still use (2.6)-(2.9) and (2.10). This scheme will be referred to as LSP-IMEX COR. We recover the LSP-IMEX scheme by choosing $\theta_{j+\frac{1}{2}} = 1$. In order to obtain a good asymptotic behaviour, $\theta_{j+\frac{1}{2}} = \mathcal{O}(\epsilon)$ is expected. Since the parameter $\theta_{j+\frac{1}{2}}$ allows to reduce the importance of the non-centred terms, it may be interpreted as a numerical diffusion reduction technique. As those non-centred terms are crucial for the sake of stability, a condition on this correction to obtain a discrete entropy inequality is expected and given in Theorem 2.

Asymptotic analysis We replace α by $\frac{\alpha}{\epsilon}$ and assume the same asymptotic expansions as before that we inject in the first equation of (2.13) to obtain

$$u^*_{j+\frac{1}{2}} = \frac{1}{2}\left(u^{n,(0)}_{j+1} + u^{n,(0)}_j\right) + \frac{\epsilon}{2}\left(u^{n,(1)}_{j+1} + u^{n,(1)}_j\right) - \frac{\theta_{j+\frac{1}{2}}\Delta x}{2a}\left(\frac{\Pi^n_{j+1} - \Pi^n_j}{\Delta x}\right) + \mathcal{O}(\epsilon^2)$$

If $u^{n,(0)}_j = 0$ and $\theta_{j+\frac{1}{2}} = \epsilon\tilde{\theta}_{j+\frac{1}{2}} = \mathcal{O}(\epsilon)$ then the first equation of (2.6) reads

$$\tau^{Lag}_j = \tau^n_j + \frac{\epsilon\Delta t}{\rho^n_j \Delta x}\left(\frac{1}{2}\left(u^{n,(1)}_{j+1} + u^{n,(1)}_j\right) - \frac{1}{2}\left(u^{n,(1)}_j + u^{n,(1)}_{j-1}\right)\right)$$

$$-\frac{\epsilon\Delta t}{\rho^n_j}\left(\tilde{\theta}_{j+\frac{1}{2}}\frac{(\Pi^n_{j+1} - \Pi^n_j)}{\Delta x} - \tilde{\theta}_{j-\frac{1}{2}}\frac{(\Pi^n_j - \Pi^n_{j-1})}{\Delta x}\right) + \mathcal{O}(\epsilon^2),$$

that is consistent with (2.12) as the third term is of order $\mathcal{O}(\epsilon)$ and consistent with 0. A straight forward analysis, similar to the one just performed for the density in Lagrangian coordinates, proves that the overall numerical scheme is consistent in the asymptotic limit with (2.2).

Main properties. We now give the main properties of the LSP-IMEX COR scheme.

Theorem 2. *Under the CFL condition (2.8) and (2.11) the numerical scheme LSP-IMEX COR is well-defined and satisfies the following properties*

(i) it is a conservative scheme for the density ρ. It is also a conservative scheme for ρu and ρE when the source terms are omitted.

(ii) $\rho^0_j > 0$ for all j implies that $\rho^n_j > 0$ for all j and $n > 0$.

(iii) it preserves the large friction asymptotic behaviour in the sense that the scheme is consistent with (2.2) in the asymptotic limit $\epsilon \to 0$.

(iv) in addition, under the following condition on the correction $\theta_{j+\frac{1}{2}}$

$$\frac{1}{2}(\overline{u}_j - u^{Lag}_j)^2 + \frac{a\Delta t}{\rho^n_j \Delta x}\left(\frac{1}{2a^2}(p(\rho^{R,*}_{j-\frac{1}{2}}, s^n_j) - p^*_{j-\frac{1}{2}})^2 - \frac{\gamma_{j-\frac{1}{2}}}{a}(p(\rho^{R,*}_{j-\frac{1}{2}}, s^n_j) - p^*_{j-\frac{1}{2}})\right)$$

$$+\frac{a\Delta t}{\rho^n_j \Delta x}\left(\frac{1}{2a^2}(p(\rho^{L,*}_{j+\frac{1}{2}}, s^n_j) - p^*_{j+\frac{1}{2}})^2 + \frac{\gamma_{j+\frac{1}{2}}}{a}(p(\rho^{L,*}_{j+\frac{1}{2}}, s^n_j) - p^*_{j+\frac{1}{2}})\right) \geq 0,$$

where $\gamma_{j+\frac{1}{2}} = \frac{1-\theta_{j+\frac{1}{2}}}{2a}\left(\Pi^n_{j+1} - \Pi^n_j\right)$, $\rho^{R,*}_{j-\frac{1}{2}} = \frac{\rho^n_j}{1+\rho^n_j a^{-1}\left(u^n_j - (u^*_{j-\frac{1}{2}} + \gamma_{j-\frac{1}{2}})\right)}$,

and $\rho^{L,*}_{j+\frac{1}{2}} = \frac{\rho^n_j}{1-\rho^n_j a^{-1}\left(u^n_j - (u^*_{j+\frac{1}{2}} + \gamma_{j+\frac{1}{2}})\right)}$, *it verifies a discrete entropy inequality*

$$(\rho s)^{n+1}_j \leq (\rho s)^n_j - \frac{\Delta t}{\Delta x}\left(G_{j+\frac{1}{2}} - G_{j-\frac{1}{2}}\right)$$

where $s^n_j = s(\rho^n_j, e^n_j)$, s denotes the mathematical entropy associated to the Euler system and $G_{j+\frac{1}{2}} = (u^*_{j+\frac{1}{2}})^+(\rho s)^{Lag}_j + (u^*_{j+\frac{1}{2}})^-(\rho s)^{Lag}_{j+1}$.

Note that, this condition in *(iv)* is of course true for $\theta_{j+\frac{1}{2}} = 1$ so that the LSP-IMEX scheme verifies this discrete entropy inequality.

2.5 Numerical results

We propose to test both LSP-IMEX and LSP-IMEX COR scheme against a test case that has been proposed in [5] and [8]. The friction coefficient is given by $\alpha = 10^6\, s^{-1}$. Both LSP-IMEX and LSP-IMEX COR computations are performed with a time step defined by $2a\Delta t = \min(\rho_j^n)\Delta x$ which complies with (2.8) and (2.11).

We run our numerical test using a spatial discretization over 100 cells, 1000 cells and 10 000 cells for LSP-IMEX scheme. In figure 2.1, we can see that there is a large amount of numerical diffusion for coarse meshes and so the numerical solution is not accurate.

Then we set $\theta_{j+\frac{1}{2}} = \min\left(\frac{2a}{\alpha(\rho_{j+1}^n+\rho_j^n)\Delta x}, 1\right)$. This choice is nondimensional and verifies $\theta_{j+\frac{1}{2}} = \mathcal{O}(\epsilon)$ so that the correct asymptotic behaviour is expected when $\epsilon \to 0$. Moreover, we recover $\theta_{j+\frac{1}{2}} = 1$ when $\alpha \to 0$ or $\Delta x \to 0$. In other words, the correction is not activated when the friction parameter is not large enough or when we use a fine discretization. We run the same test for LSP-IMEX COR scheme. In figure 2.2, we observe that the approximate solutions are much more accurate than the one computed without the correction, especially for a coarse mesh. Let us note $J = \max(j \in \mathbb{N}/j\Delta x \leq 1)$, by plotting $n \mapsto \sum_j \left((\rho s)_j^{n+1} - (\rho s)_j^n\right) + \left((\rho u s)_J^n - (\rho u s)_0^n\right)$, we observe in figure 2.3 that the average entropy inequality is verified.

We derive a multi-dimensional version of the LSP-IMEX COR scheme by using the 1D fluxes of the scheme and the rotational invariance of the equations. We consider the 2D domain $[0.1]^2$, discretized over a 5000-triangle grid. The friction parameter is set to $\alpha = 10^6 s^{-1}$ and the initial condition reads

$$\begin{cases} (\rho, u, p) = (1.0, 0, 10000.0), & \text{if } (x-0.5)^2 + (y-0.5)^2 \geq 0.3, \\ (\rho, u, p) = (2.0, 0, 26390.2), & \text{if } (x-0.5)^2 + (y-0.5)^2 < 0.3. \end{cases}$$

In figure 2.4, we observe that LSP-IMEX scheme leads to a large amount of numerical diffusion while LSP-IMEX COR scheme with $\theta = \frac{1}{\alpha\Delta t}$ gives much more accurate numerical results.

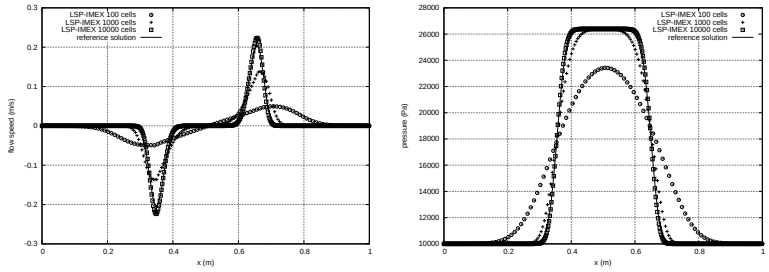


FIGURE 2.1 – Profile at time $t = 0.01s$ of the velocity (left) and the pressure (right) obtained for a 100-cell, 1000-cell and 10 000-cell grid with the LSP-IMEX scheme and the reference solution (LSP-IMEX COR scheme with 10 000-cell mesh).

Acknowledgement. *The authors are supported by the LRC MANON (CEA/DEN/DANS/DM2S and UPMC-CNRS/LJLL).*

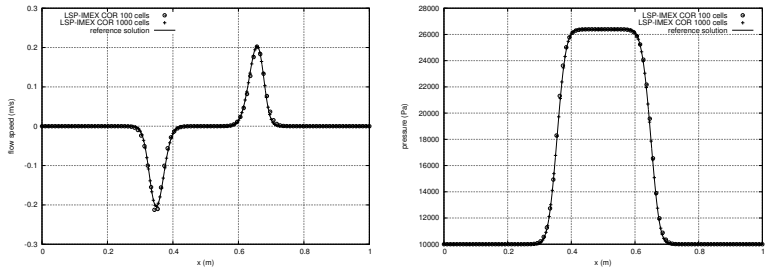


FIGURE 2.2 – Profile at time $t = 0.01s$ of the velocity (left) and the pressure (right) obtained for a 100-cell and 1000-cell grid with the LSP-IMEX COR scheme and the reference solution (LSP-IMEX COR scheme with 10 000-cell mesh).

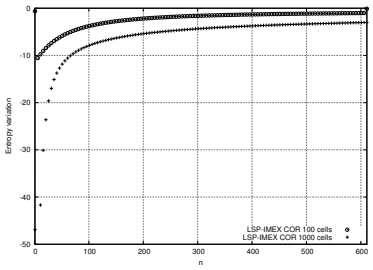


FIGURE 2.3 – Variation of mathematical entropy averaged over the domain for each time step obtained for a 100-cell and 1000-cell grid with LSP-IMEX COR scheme.

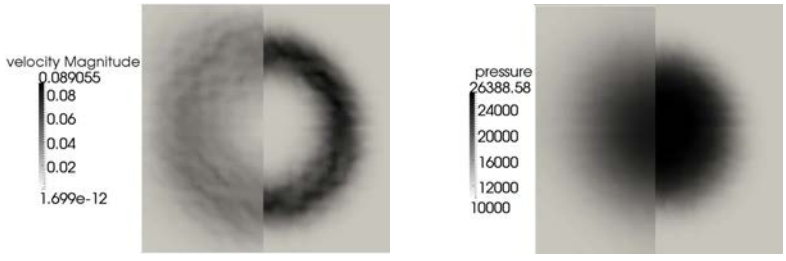


FIGURE 2.4 – Profile at time $t = 0.01s$ of the velocity (left) and the pressure (right) obtained for a 5000-triangle mesh with the LSP-IMEX scheme (left) and the LSP-IMEX COR scheme (right).

Annexes

Ces annexes viennent compléter l'article *Operator-splitting-based asymptotic preserving scheme for the gas dynamics equations with stiff source terms*. L'annexe 2.A précise la manière d'imposer les conditions aux limites pour le schéma LSP-IMEX, ainsi que pour les schémas basés sur le *splitting* d'opérateur Lagrange-Projection proposés dans les chapitres suivants. L'annexe 2.B introduit la notion de régime intermédiaire qui permet de distinguer les schémas préservant l'asymptotique des schémas dits tout-régime. Finalement, un premier pas vers l'extension à des schémas préservant l'asymptotique d'ordre élevé est proposé dans l'annexe 2.C, avec un schéma d'ordre 2 en espace.

2.A Conditions aux limites

Le schéma de splitting d'opérateur Lagrange/Source/Projection (2.6)-(2.9)-(2.10) est une discrétisation en trois étapes du système de la dynamique des gaz avec gravité et friction (2.1). Ce schéma a été écrit et étudié pour un domaine infini $x \in \mathbb{R}$ et donc une infinité de cellules $j \in \mathrm{Z}$. En pratique, le domaine d'étude est fini $x \in [x_0, x_L]$ où x_0 et x_L sont deux réels, et on a un nombre N fini de cellules $1 \leq j \leq N$. Lors de la modélisation du problème, des conditions à imposer aux bords du domaine, en $x = x_0$ et $x = x_L$ viennent s'ajouter au système d'équations aux dérivées partielles (2.1). Il faut alors discrétiser ces conditions aux limites pour pouvoir mettre à jour la solution approchée dans les cellules $j = 1$ et $j = N$. Sans perte de généralité, on considère dans la suite les conditions aux limites en x_0.

Une méthode pour traiter les conditions aux limites des systèmes hyperboliques est le formalisme des demi-problèmes de Riemann [10, 11, 12]. Pour mettre à jour la solution dans la cellule $j = 1$, on calcule en x_0 un flux de frontière à l'aide du schéma de Godunov. Pour calculer ce flux, on résout un problème de Riemann en x_0 entre :

- un état droit entièrement connu car situé dans le domaine physique $x > x_0$;
- un état gauche admissible, choisi de manière à imposer les conditions aux limites en $x = x_0$. Cet état n'est que partiellement connu.

La résolution analytique du problème de Riemann pour des états droit et gauche quelconques permet alors de déterminer le nombre et les quantités de l'état gauche qui doivent être fixées afin de connaître la solution dans le domaine physique $x > x_0$ et donc le flux de frontière. Pour avoir un problème bien posé, il est nécessaire que les conditions aux limites que l'on souhaite imposer permettent de fixer ces quantités de l'état gauche.

Nous allons utiliser ce formalisme pour le schéma de l'étape Lagrangienne (2.6) puis pour le schéma de l'étape de projection (2.10). Nous présenterons ensuite les relations permettant de définir les flux de frontières pour quelques choix classiques de conditions aux limites pour le système de la dynamique des gaz.

Remarque. Le schéma pour les termes sources (2.9) a un stencil de taille un, il n'est donc pas nécessaire d'imposer de conditions aux limites pour effectuer cette étape.

Conditions aux limites pour l'étape Lagrangienne (2.6)

Le système de relaxation homogène en coordonnées Lagrangienne s'écrit

$$\begin{cases} \partial_t \tau - \partial_m u = 0, \\ \partial_t u + \partial_m \Pi = 0, \\ \partial_t \Pi + a^2 \partial_m u = 0, \\ \partial_t E + \partial_m (pu) = 0. \end{cases} \tag{2.14}$$

La résolution du problème de Riemann pour le système de relaxation homogène (2.14) étant plus aisé que pour le système de la dynamique des gaz en coordonnées Lagrangienne (2.4), on préférera utiliser ce système pour imposer les conditions aux limites. Le système (2.14) est hyperbolique, les trois champs caractéristiques $-a$, 0 et a sont linéairement dégénérés et on dispose de quatre invariants de Riemann forts :

- $\overleftarrow{w} = \Pi - au$ pour le champ $-a$;
- $I = \Pi + a^2\tau$ et $S = E - \frac{u^2}{2} - \frac{\Pi^2}{2a^2}$ pour le champ 0 ;
- $\overrightarrow{w} = \Pi + au$ pour le champ a.

Ces invariants de Riemann forts permettent de résoudre analytiquement le problème de Riemann. On utilisera le jeu de variables $\mathbf{V} = (I, \overrightarrow{w}, \overleftarrow{w}, S)^T$ pour imposer les conditions aux limites. On considère le problème de Riemann au temps t^n et au niveau de l'interface $1/2$, c'est à dire en x_0. La solution est autosimilaire, composée de trois ondes de vitesse $-a$, 0 et a :

$$\mathbf{V}\left(\frac{m}{t}, \mathbf{V}_L, \mathbf{V}_R\right) = \begin{cases} \mathbf{V}_L, & \frac{m}{t} < -a, \\ \mathbf{V}_L^*, & -a < \frac{m}{t} < 0, \\ \mathbf{V}_R^*, & 0 < \frac{m}{t} < a, \\ \mathbf{V}_R, & a < \frac{m}{t}. \end{cases}$$

Comme $\overleftarrow{w}$ est un 1-invariant de Riemann fort, il est constant à la traversée des 2-ondes et 3-ondes, $\overleftarrow{w}_L^* = \overleftarrow{w}_R^* = \overleftarrow{w}_R$. I et S sont des 2-invariants de Riemann fort, ils sont donc constants à la traversée des 1-ondes et des 3-ondes, $I_L^* = I_L$, $I_R^* = I_R$, $S_L^* = S_L$ et $S_R^* = S_R$. Finalement $\overrightarrow{w}$ est un 3-invariant de Riemann fort, il est donc constant à la traversée des 1-ondes et des 2-ondes : $\overrightarrow{w}_L^* = \overrightarrow{w}_R^* = \overrightarrow{w}_L$. Les états intermédiaires sont donc donnés par

$$\mathbf{V}_L^* = (I_L, \overleftarrow{w}_R, \overrightarrow{w}_L, S_L)^T, \quad \mathbf{V}_R^* = (I_R, \overleftarrow{w}_R, \overrightarrow{w}_L, S_R)^T,$$

et on a ainsi résolu de manière analytique le problème de Riemann. On peut alors définir une vitesse et une pression à l'interface par

$$\begin{cases} u^* = u_L^* = u_R^* = \frac{\overrightarrow{w}_L - \overleftarrow{w}_R}{2a}, \\ \Pi^* = \Pi_L^* = \Pi_R^* = \frac{\overrightarrow{w}_L + \overleftarrow{w}_R}{2}. \end{cases}$$

Ces deux grandeurs permettent d'évaluer le flux $(u^*, \Pi^*, a^2 u^*, \Pi^* u^*)$ au niveau de l'interface et sont donc les grandeurs que l'on doit calculer pour mettre à jour le schéma Lagrangien (2.6).

Pour connaître la solution dans le domaine physique $\frac{m}{t} > 0$, il suffit donc de connaître les valeurs de $\mathbf{V}_R$ et $\overrightarrow{w}_L$. Les valeurs de I_L, $\overleftarrow{w}_L$ et S_L ne sont pas utilisées pour calculer la solution dans cette partie du domaine et donc ne sont pas nécessaires pour évaluer le flux d'interface en x_0. Ainsi, lors de l'étape Lagrangienne, une seule caractéristique est entrante au domaine et il suffit donc d'imposer une quantité au bord permettant de fixer la valeur de $\overrightarrow{w}$.

Conditions aux limites pour l'étape de projection (2.10)

Pour l'étape de projection, le flux du schéma décentré amont (2.10) à l'interface 1/2 peut être vu comme le flux du schéma de Godunov du système

$$\partial_t \mathbf{U} + u^*_{1/2} \partial_x \mathbf{U} = 0,$$

où $u^*_{1/2}$ est la vitesse d'interface calculée lors de l'étape Lagrangienne et $\mathbf{U} = (\rho, \rho u, \rho E)^T$. Ce système est hyperbolique et l'unique champ caractéristique $u^*_{1/2}$ est linéairement dégénéré.

On considère le problème de Riemann au temps t^n au niveau de l'interface 1/2. La solution est autosimilaire, composée d'une seule onde de vitesse $u^*_{1/2}$:

$$\mathbf{U}\left(\frac{x}{t}, \mathbf{U}_L, \mathbf{U}_R\right) = \begin{cases} \mathbf{U}_L, & \frac{x}{t} < u^*_{1/2}, \\ \mathbf{U}_R, & u^*_{1/2} < \frac{x}{t}. \end{cases}$$

Pour connaître la solution dans le domaine physique $\frac{x}{t} > 0$, il y a deux cas de figure :

- si $u^*_{1/2} \leq 0$, il suffit de connaître $\mathbf{U}_R$. Il n'est donc pas nécessaire d'imposer des conditions aux limites pour l'étape de projection car l'unique caractéristique est sortante au domaine.
- si $0 < u^*_{1/2}$, il faut connaître $\mathbf{U}_R$ et $\mathbf{U}_L$. Il est nécessaire d'imposer les quatre grandeurs de $\mathbf{U}_L$ pour calculer le flux d'interface en x_0. L'unique champ de multiplicité quatre est entrant au domaine.

Conditions aux limites pour le schéma Lagrange-Projection (2.6)-(2.10)

On a vu précédemment les quantités qu'il était nécessaire de connaître lors de l'étape Lagrangienne et de l'étape de projection pour pouvoir calculer le flux d'interface de ces deux étapes. Nous allons utiliser ces résultats pour imposer des conditions aux limites usuelles pour le système de la dynamique des gaz.

1. Condition à la limite de type Neumann

La condition à la limite de type Neumann s'écrit au niveau continu pour $\varphi \in \{\rho, \rho u, \rho E\}$

$$\partial_x \varphi = 0, \text{ en } x = x_0$$

- Étape Lagrangienne : on impose $\partial_x u = 0$ et $\partial_x p = 0$ en x_0, on a donc aussi $\partial_x \overrightarrow{w} = 0$. Une discrétisation naturelle de cette condition est :

$$\overrightarrow{w}_L = \overrightarrow{w}_R.$$

- Étape de projection : on impose $\partial_x \mathbf{U} = 0$ en x_0, une discrétisation naturelle de cette condition est :

$$\mathbf{U}_L = \mathbf{U}_R.$$

2. Conditions à la limite de type paroi

La condition à la limite de type paroi s'écrit au niveau continu

$$u = 0, \text{ en } x = x_0$$

.

- Étape Lagrangienne : une discrétisation naturelle de cette condition est d'imposer $u^*_{1/2} = 0$. Or comme $u^*_{1/2} = \frac{\overrightarrow{w}_L - \overleftarrow{w}_R}{2a}$, on est amené à imposer la condition

$$\overrightarrow{w}_L = \overleftarrow{w}_R.$$

- Étape de projection : comme $u^*_{1/2} = 0$, il n'est pas nécessaire d'imposer de quantité lors de cette étape.

3. Conditions à la limite de sortie subsonique (pression imposée)

Dans le cas d'une sortie subsonique ($u^*_{1/2} \leq 0$), on souhaite imposer une pression en sortie

$$p = p_{out}, \quad \text{en } x = x_0.$$

- Étape Lagrangienne : une discrétisation naturelle de cette condition est d'imposer $p^*_{1/2} = p_{out}$. Comme $p^*_{1/2} = \frac{\overrightarrow{w}_L + \overleftarrow{w}_R}{2}$, on est amené à imposer la condition

$$\overrightarrow{w}_L = 2p_{out} - \overleftarrow{w}_R.$$

- Étape de projection : comme $u^*_{1/2} \leq 0$, il n'est pas nécessaire d'imposer de quantités lors de cette étape.

4. Condition à la limite d'entrée subsonique (débit-enthalpie imposés)

Dans le cas d'une entrée subsonique ($u^*_{1/2} > 0$), on souhaite imposer un débit et une enthalpie

$$\begin{cases} \rho u = Q_{in}, & \text{en } x = x_0, \\ h = h_{in}, & \text{en } x = x_0. \end{cases}$$

On a par définition des variables et des invariants de Riemann de l'étape Lagrangienne :

$$\begin{cases} \rho u = a\left(\frac{\overrightarrow{w} - \overleftarrow{w}}{2}\right)\left(I - \left(\frac{\overrightarrow{w} + \overleftarrow{w}}{2}\right)\right)^{-1}, \\ h = E - \frac{u^2}{2} + \Pi\tau = S + \frac{I}{2a^2}\left(\frac{\overrightarrow{w} + \overleftarrow{w}}{2}\right) - \frac{1}{2a^2}\left(\frac{\overrightarrow{w} + \overleftarrow{w}}{2}\right)^2. \end{cases}$$

- Étape Lagrangienne : on utilise une technique d'annulation des ondes pour la variable I, on a alors la relation $I_L = I_R$. Cela est dû au fait que I est un invariant de Riemann fort pour le second champ et est donc lié au degré de multiplicité supplémentaire introduit par le processus de relaxation. On impose alors le débit $\rho^*_L u^*_L = Q_{in}$ et l'enthalpie $h^*_L = h_{in}$, on obtient les relations

$$\begin{cases} (a + Q_{in})\,\overrightarrow{w}_L = 2Q_{in}I_L + (a - Q_{in})\,\overleftarrow{w}_R, \\ S_L = h_{in} - \frac{I_L}{2a^2}\left(\frac{\overrightarrow{w}_L + \overleftarrow{w}_R}{2}\right) + \frac{1}{2a^2}\left(\frac{\overrightarrow{w}_L + \overleftarrow{w}_R}{2}\right)^2. \end{cases}$$

  La première relation est la seule nécessaire pour calculer le flux d'interface lors de l'étape Lagrangienne. La suivante permet de définir complètement l'état $\mathbf{V}^*_L$ dont on va se servir lors de l'étape de projection.
- Étape de projection : le champ est entrant ($u^*_{1/2} > 0$), il faut déterminer un état $\mathbf{U}_L$ à advecter

dans le domaine. On choisit ici de fixer

$$\mathbf{U}_L = \mathbf{U}(\mathbf{V}_L^*)$$

où $\mathbf{V}_L^*$ a été déterminé lors de l'étape Lagrangienne.

Généralisation pour les schémas implicites et multi-dimensionnels

On s'est intéressé précédemment aux conditions aux limites d'un schéma Lagrange-Projection explicite en une dimension d'espace. Dans les chapitres 3 et 4, on va construire des schémas où l'étape Lagrangienne sera traitée de façon implicite. On prendra alors des évaluations implicites des relations écrites ci-dessus pour l'étape Lagrangienne. Par ailleurs les schémas seront écrits en deux dimensions d'espace et sur maillage non structuré, on utilise l'invariance par rotation des systèmes d'équations pour se ramener à un système quasi 1D le long de la normale au domaine. On utilisera alors les relations précédentes pour calculer les flux de frontières.

2.B Régime intermédiaire et précision pour les schémas asymptotic preserving

On a étudié précédemment le comportement du schéma (2.6)-(2.13)-(2.9)-(2.10) dans deux régimes particuliers, d'une part le régime classique ϵ d'ordre 1, et d'autre part le régime asymptotique $\epsilon \to 0$. On va s'intéresser ici au régime intermédiaire situé entre le régime classique et le régime asymptotique.

Approche à Δx fixé

On considère un pas d'espace Δx fixé et on fait varier la valeur de ϵ. Lorsque ϵ est d'ordre 1, c'est à dire en régime classique, le schéma LSP-IMEX obtenu pour $\theta = 1$ donne de bons résultats numériques. Néanmoins, au vu de l'analyse de l'erreur de troncature menée dans la section 2.4, le choix $\theta = O(\epsilon)$ est nécessaire pour avoir un bon comportement en régime asymptotique $\epsilon \to 0$. L'idée est alors de choisir la valeur de θ en fonction du régime dans lequel on se trouve afin d'obtenir de bons résultats numériques quel que soit le régime de ϵ considéré. Un tel schéma est dit tout-régime.

Pour construire un tel schéma, un choix naturel est de prendre la plus petite valeur de θ vérifiant un certain critère de stabilité. En effet, tant que le schéma reste stable, plus la valeur de θ est faible plus la diffusion numérique est faible et on s'attend donc à calculer des solutions approchées plus précises. Néanmoins, compte tenu de la nature non linéaire du système d'équations et de l'importance du caractère entropique des solutions dans certains régimes, il est difficile d'obtenir un tel critère en pratique. On est alors amené à définir la valeur de θ de manière empirique, on a fait le choix par exemple pour le schéma LSP-IMEX COR de fixer $\theta^{COR} = \min(\epsilon \frac{a}{\alpha\rho\Delta x}, 1)$. Ce choix permet d'effectuer la transition entre $\theta = O(\epsilon)$ en régime asymptotique et $\theta = 1$ en régime classique. Des résultats numériques obtenus avec ce schéma sont présentés dans la section 2.5.

D'autres choix sont possibles pour θ. On considère ici en particulier

$$\theta^{CUTOFF} = \begin{cases} 1, & \text{si } \epsilon > \overline{\epsilon}, \\ \epsilon, & \text{si } \epsilon \leq \overline{\epsilon}, \end{cases}$$

où on se donne une valeur $\overline{\epsilon} > 0$. Ce choix coïncide avec le schéma LSP-IMEX, c'est à dire $\theta = 1$, pour ϵ supérieur à $\overline{\epsilon}$, mais est *asymptotic preserving* car on a bien $\theta^{CUTOFF} = O(\epsilon)$ lorsque $\epsilon \to 0$. Ainsi, si

on choisit une valeur $\bar{\epsilon} \ll 1$, le schéma obtenu avec θ^{CUTOFF} est bien *asymptotic preserving* mais calcule des solutions approchées aussi mauvaises que le schéma LSP-IMEX pour des cas tests où $\epsilon > \bar{\epsilon}$.

Ainsi, la propriété *asymptotic preserving* n'est pas suffisante pour garantir le bon comportement d'un schéma numérique en régime intermédiaire, qui correspond ici à une valeur de ϵ telle que $0 < \epsilon \ll 1$. Il est donc nécessaire de chercher d'autres propriétés permettant de construire des schémas tout-régime. Par exemple, dans les chapitres 3 et 4, on regardera des propriétés d'uniformité de l'erreur de troncature et de la condition CFL par rapport au petit paramètre ϵ qui sera joué par le nombre de Mach M.

Approche à ϵ fixé

On considère une valeur de ϵ fixée et on fait varier le pas d'espace Δx. Le schéma (2.6)-(2.13)-(2.9)-(2.10) est consistant quel que soit le choix de θ d'ordre 1 par rapport à Δx. On s'attend donc à obtenir de bons résultats numériques à condition que l'on raffine suffisamment. Néanmoins, le schéma LSP-IMEX, obtenu pour $\theta = 1$, requiert d'utiliser un pas d'espace Δx de l'ordre de ϵ pour calculer des solutions approchées suffisamment précises, tandis que le schéma LSP-IMEX COR, obtenu pour $\theta = \theta^{COR}$, permet de calculer de bonnes solutions approchées même avec un maillage grossier, c'est à dire un pas d'espace Δx d'ordre 1 par rapport à ϵ.

On trace des courbes de convergence en norme L^1 pour étudier l'influence du choix de la modification θ sur la précision des solutions approchées pour divers maillages. On considère le cas test de la section 2.5 et on calcule une solution de référence sur un maillage fin de 2.10^5 mailles. avec une méthode d'ordre 2 en espace LSP-IMEX-OD2X qui sera présentée dans la section 2.C. On peut alors calculer pour une grandeur Y l'erreur relative L^1 au temps final $t_f = 0.01s$:

$$\mathrm{err}(Y) = \frac{\|Y(\cdot, t_f) - Y^{\mathrm{ref}}(\cdot, t_f)\|_{L^1([0,1])}}{\|Y^{\mathrm{ref}}(\cdot, t_f)\|_{L^1([0,1])}}.$$

On calcule cette erreur pour les variables u et p, les schémas LSP-IMEX, LSP-IMEX COR et LSP-IMEX COR2 qui correspondent respectivement aux choix $\theta = 1$, $\theta = \theta^{COR}$ et $\theta = \frac{\epsilon}{\alpha}$, et des maillages allant de 100 mailles à 100000 mailles.

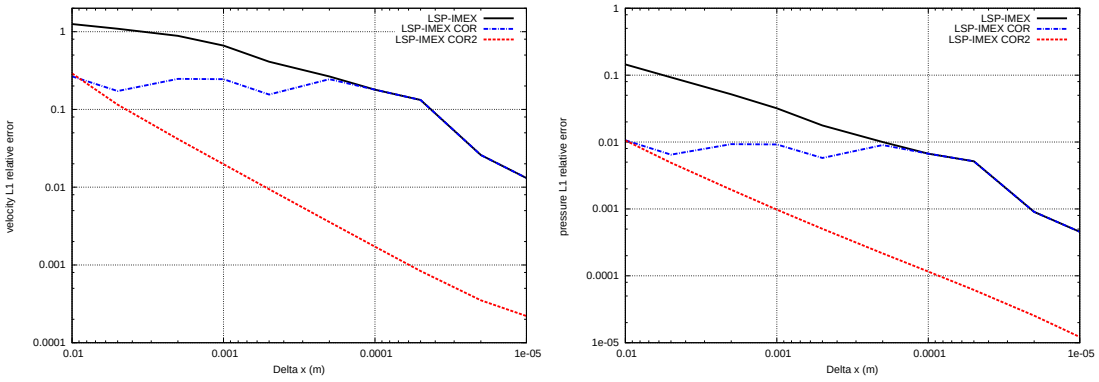


FIGURE 2.B.1 – Courbes de convergence en norme L^1 pour la vitesse (gauche) et la pression (droite) des schémas LSP-IMEX, LSP-IMEX COR et LSP-IMEX COR2. Le cas test considéré est celui de la section 2.5.

On observe sur la figure 2.B.1 que le schéma LSP-IMEX ne calcule pas de bonnes approximations quand $\epsilon \ll \Delta x$, contrairement au schéma LSP-IMEX COR qui est bien plus précis pour ces maillages grossiers. Compte tenu du choix de θ^{COR}, ces deux schémas coïncident quand $\Delta x \ll \epsilon$. Le comportement du schéma LSP-IMEX COR en régime intermédiaire, c'est à dire pour $\Delta x \simeq \epsilon$, n'est pas satisfaisant car

on observe un palier et on ne gagne donc pas en précision lorsque l'on raffine en maillage. Ceci est dû à la transition entre $\theta = O(\epsilon)$ quand $\epsilon \ll \Delta x$ et $\theta = 1$ quand $\Delta x \ll \epsilon$. Cette observation motive le choix $\theta = \frac{\epsilon}{\alpha}$ qui correspond à une volonté de corriger le schéma de manière uniforme par rapport à ϵ. On observe en effet sur la figure 2.B.1 que le schéma LSP-IMEX COR2 permet d'obtenir de meilleurs résultats que le schéma LSP-IMEX COR, quel que soit le régime et en particulier pour le régime intermédiaire $\Delta x \simeq \epsilon$.

2.C Extension à l'ordre 2 en espace

On construit ici une extension du schéma (2.6)-(2.13)-(2.9)-(2.10) à l'ordre 2 en espace. Ceci permet d'améliorer la précision des approximations des solutions régulières. Pour obtenir un tel schéma on utilise l'approche MUSCL classique, voir [20, 19] ainsi que [2] pour une application à un schéma Lagrange-Projection. Cette approche consiste à effectuer une reconstruction des variables à l'aide des fonctions affines par morceaux et non pas constantes par morceaux comme c'est le cas pour des méthodes d'ordre 1. Ainsi, pour une variable w, à partir de la connaissance des valeurs des w_i au centre des mailles, on construit dans chaque maille la fonction $\tilde{w}_i(x) = w_i + \sigma_i x$, où σ_i correspond à la pente dans la maille i et est évaluée à partir des valeurs des w_i, w_{i-1} et w_{i+1}. En effet, on fixe ici

$$\sigma_i = minmod\left(\frac{w_i - w_{i-1}}{\Delta x}, \frac{w_{i+1} - w_i}{\Delta x}\right),$$

où la fonction *minmod* est un limiteur de pente qui évite les oscillations lorsque la solution n'est plus régulière,

$$minmod(a,b) = \begin{cases} a, & \text{si } |a| \leq |b| \text{ et } a.b \geq 0, \\ b, & \text{si } |b| \leq |a| \text{ et } a.b \geq 0, \\ 0, & \text{si } a.b \leq 0. \end{cases}$$

Cette reconstruction permet de définir des valeurs au niveau des interfaces de la maille i :

- $w_i^L = \tilde{w}_i\left(-\frac{\Delta x}{2}\right)$ au niveau de l'interface $i - \frac{1}{2}$;
- $w_i^R = \tilde{w}_i\left(\frac{\Delta x}{2}\right)$ au niveau de l'interface $i + \frac{1}{2}$.

On utilise alors ces valeurs plutôt que les valeurs aux centres des mailles pour exprimer les flux numériques et on obtient ainsi une méthode d'ordre 2 en espace. On précise maintenant la définition des schémas numériques d'ordre 2 pour les différentes étapes du schéma Lagrange-Source-Projection.

L'étape Lagrangienne (2.6)-(2.7) devient

$$\begin{cases} \tau_j^{Lag} = \tau_j^n + \frac{\Delta t}{\rho_j^n \Delta x}\left(u_{j+\frac{1}{2}}^{*,2} - u_{j-\frac{1}{2}}^{*,2}\right), \\ u_j^{Lag} = u_j^n - \frac{\Delta t}{\rho_j^n \Delta x}\left(p_{j+\frac{1}{2}}^{*,2} - p_{j-\frac{1}{2}}^{*,2}\right), \\ E_j^{Lag} = E_j^n - \frac{\Delta t}{\rho_j^n \Delta x}\left((pu)_{j+\frac{1}{2}}^{*,2} - (pu)_{j-\frac{1}{2}}^{*,2}\right), \end{cases} \tag{2.15}$$

où les flux numériques sont donnés par

$$\begin{cases} u_{j+\frac{1}{2}}^{*,2} = \frac{1}{2}\left(u_{j+1}^{n,L} + u_j^{n,R}\right) - \frac{\theta_{j+\frac{1}{2}}}{2a}\left(\Pi_{j+1}^{n,L} - \Pi_j^{n,R}\right), \\ p_{j+\frac{1}{2}}^{*,2} = \frac{1}{2}\left(\Pi_{j+1}^{n,L} + \Pi_j^{n,R}\right) - \frac{a}{2}\left(u_{j+1}^{n,L} - u_j^{n,R}\right), \\ \Pi_j^n = p(\rho_j^n, e_j^n). \end{cases} \tag{2.16}$$

On remarque que pour calculer les flux numériques de l'étape Lagrangienne, il suffit d'effectuer une reconstruction affine par morceaux pour les variables u et Π.

Pour l'étape des termes sources, on garde le schéma (2.9)

$$\left\{\begin{array}{l} \overline{\tau}_j = \tau_j^{Lag}, \\ \overline{u}_j = u_j^{Lag} + g\Delta t - \alpha\Delta t\overline{u}_j, \\ \overline{E}_j = E_j^{Lag} + g\Delta t\overline{u}_j - \alpha\Delta t(\overline{u}_j)^2. \end{array}\right. \tag{2.17}$$

En effet, comme les termes sources sont centrés, il n'est pas nécessaire d'utiliser les valeurs reconstruites aux interfaces.

L'étape de projection (2.10) devient pour $X \in \{\rho, \rho u, \rho E\}$

$$X_j^{n+1} = \overline{X}_j - \frac{\Delta t}{\Delta x}\Big((u_{j-\frac{1}{2}}^{*,2})^+ \left[\overline{X}_j^L - \overline{X}_{j-1}^R\right] + (u_{j+\frac{1}{2}}^{*,2})^- \left[\overline{X}_{j+1}^L - \overline{X}_j^R\right]\Big). \tag{2.18}$$

Par analogie avec les schémas d'ordre 1, on définit les schémas LSP-IMEX-OD2X et LSP-IMEX-OD2X COR à partir de (2.15)-(2.16)-(2.17)-(2.18) pour des valeurs respectives de $\theta_{j+\frac{1}{2}} = 1$ et $\theta_{j+\frac{1}{2}} = \min\left(\frac{2a}{\alpha(\rho_{j+1}^n+\rho_j^n)\Delta x}, 1\right)$. On considère le cas test de la section 2.5. Sur la figure 2.C.1, on observe que le schéma LSP-IMEX OD2X est très diffusif sur maillage grossier et ne permet donc pas de calculer des solutions approchées précises. Pour le schéma LSP-IMEX-OD2X COR, on observe figure 2.C.2 qu'il permet de calculer des solutions approchées précises même sur maillage grossier. Par ailleurs, si on compare les résultats de la figure 2.C.2 à ceux de la figure 2.1, on observe que, bien que très diffusés, les résultats obtenus avec LSP-IMEX-OD2X sont meilleurs que ceux obtenus avec la méthode d'ordre 1 LSP-IMEX. Ceci est confirmé pour les schémas LSP-IMEX-OD2X COR et LSP-IMEX COR, à l'aide des courbes de convergence en norme L^1 figure 2.C.3.

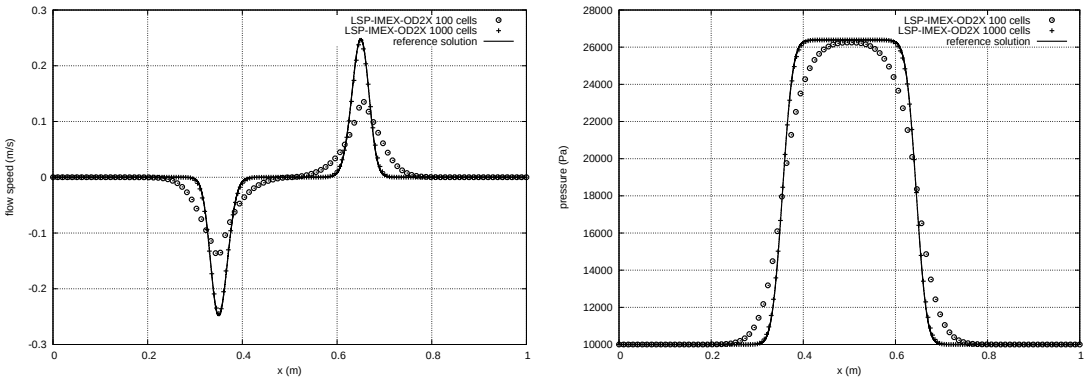


FIGURE 2.C.1 – Profil au temps $t = 0.01s$ de la vitesse (gauche) et la pression (droite) obtenus avec un maillage de 100 mailles et 1000 mailles pour le schéma LSP-IMEX-OD2X et une solution de référence (LSP-IMEX-OD2X avec un maillage de 2.10^5 mailles).

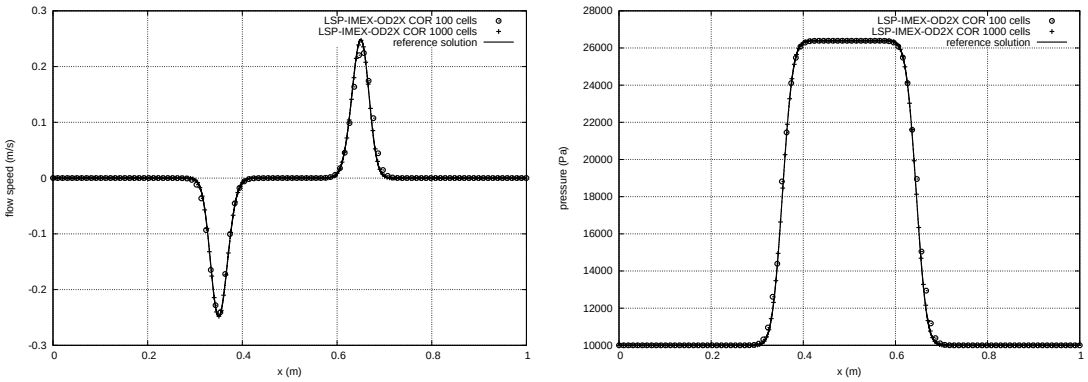


FIGURE 2.C.2 – Profil au temps $t = 0.01s$ de la vitesse (gauche) et la pression (droite) obtenus avec un maillage de 100 mailles et 1000 mailles pour le schéma LSP-IMEX-OD2X COR et une solution de référence (LSP-IMEX-OD2X avec un maillage de 2.10^5 mailles).

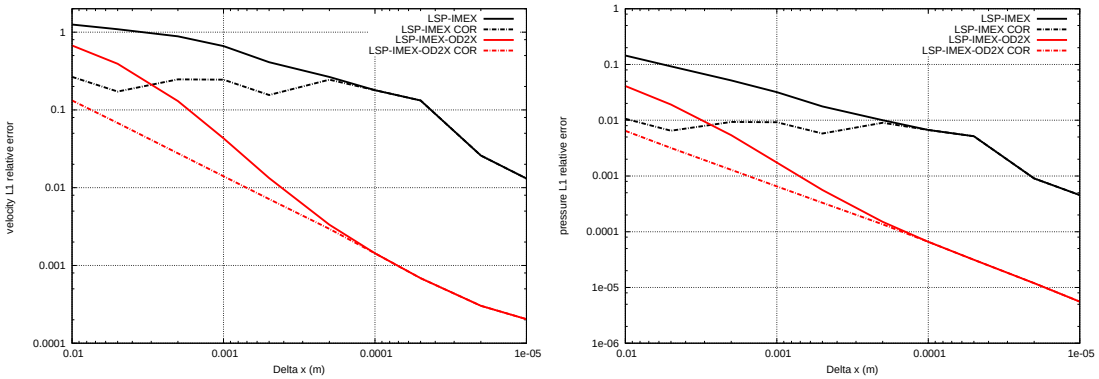


FIGURE 2.C.3 – Courbes de convergence en norme L^1 pour la vitesse (gauche) et la pression (droite) des schémas LSP-IMEX, LSP-IMEX COR, LSP-IMEX-OD2X et LSP-IMEX-OD2X COR. Le cas test considéré est celui de la section 2.5.

Bibliographie

[1] C. Berthon and R. Turpault, *Asymptotic preserving HLL schemes.* Numer, Methods Partial Differential Equations, 27(6) : 1396–1422, (2011).

[2] M. Billaud Friess, B. Boutin, F. Caetano, G. Faccanoni, S. Kokh, F. Lagoutière and L. Navoret, *A second order anti-diffusive Lagrange-remap scheme for two-component flows*, ESAIM : Proceedings, 32 : 149–162, (2011).

[3] F. Bouchut, *A reduced stability condition for nonlinear relaxation to conservative laws*, J. Hyp. Diff. Eq, 1(1) : 149-170, (2004).

[4] F. Bouchut, H. Ounaissa and B. Perthame, *Upwinding of the source term at interfaces for Euler equations with high friction*, Comput. Math. Appl., 53 : 361–375, (2007).

[5] C. Chalons, F. Coquel, E. Godlewski, P-A Raviart and N. Seguin, *Godunov-type schemes for hyperbolic systems with parameter dependent source. The case of Euler system with friction*, Math. Models Methods Appl. Sci, 20(11) : 2109–2166, (2010).

[6] C. Chalons and J.-F. Coulombel, *Relaxation approximation of the Euler equations*, Journal of Mathematical Analysis and Applications, 348(2) : 872–893, (2008).

[7] C. Chalons and F. Coquel, *Navier-stokes equations with several independant pressure laws and explicit predictor-corrector schemes*, Numerisch Math, 101(3) : 451–478, (2005).

[8] C. Chalons, M. Girardin and S. Kokh, *Large time step and asymptotic preserving numerical schemes for the gas dynamics equations with source terms*, SIAM J. Sci. Comput., 35(6) : a2874–a2902, (2013).

[9] S. Dellacherie, *Analysis of Godunov type schemes applied to the compressible Euler system at low Mach number*, J. Comput Phys., 229(4) : 978–1016, (2012).

[10] B. Després and F. Dubois, Systèmes hyperboliques de lois de conservation, application à la dynamique des gaz. Éditions de l'École Polytechnique. 169–177, (2005).

[11] F. Dubois and P. Le Floch, *Boundary conditions for nonlinear hyperbolic systems of conservation laws*, J. Diff. Eq., 71 : 93–122, (1988).

[12] F. Dubois, *Partial Riemann problems, boundary conditions and gas dynamics*, Absorbing Boundaries and Layers, Domain Decomposition Methods, Nova Sci. Publ., 16–77, (2001).

[13] E. Godlewski and P.-A Raviart, Numerical Approximation Of Hyperbolic Systems Of Conservation Laws, Springer-Verlag, New York, (1996).

[14] L. Gosse and G. Toscani, *An asymptotic-preserving well-balanced schemes for the hyperbolic heat equations*, C.R. Acad. Sci. Paris, Ser. I 334 : 337–342, (2002).

[15] S. Jin, *Runge-Kutta methods for hyperbolic conservations law with stiff relaxation terms*, J. Comput. Phys., 122 : 51–67, (1995).

[16] S. Jin and C.D. Levermore, *Numerical schemes for hyperbolic conservation laws with stiff relaxation terms*, J. Comput. Phys., 126 : 449–467, (1996).

[17] S. Jin and Z. P. Xin, *The relaxation schemes for systems of conservation laws in arbitrary space dimension*, Comm. Pure Appl. Math., 28 : 235–276, (1995).

[18] I. Suliciu, *On the thermodynamics of fluids with relaxation and phase transitions. Fluids with relaxation*, Int. J. Engag. Sci., 36 : 921–947, (1998).

[19] E.F. Toro, Riemann solvers and numerical methods for fluid dynamics, Springer-Verlag, Berlin, second edition, (1999).

[20] B. van Leer, *Towards the ultimate conservative difference scheme V. A second order sequel to Godunov's method*, J. Comput. Phys., 32 : 101–136, (1979).

[21] H. Weyl, Shock waves in arbitrary fluids, Comm. Pure Appl. Math., 2 : 103–122, (1949).

Chapitre 3

Schémas Lagrange-Projection tout-régime pour le système de la dynamique des gaz sur maillage non structuré

Ce chapitre a fait l'objet d'un article soumis à la revue Communications in Computational Physics : C. Chalons, M. Girardin and S. Kokh, *An all-regime Lagrange-Projection like scheme for the gas dynamics equations on unstructured meshes.*

Dans les chapitres 1 et 2, on a considéré le régime asymptotique du système de la dynamique des gaz avec termes sources raides. Dans ce chapitre, on s'intéresse au système de la dynamique des gaz homogène dans le régime des faibles nombres de Mach.

Les annexes viennent compléter l'article *An all-regime Lagrange-Projection like scheme for the gas dynamics equations on unstructured meshes*, en apportant un éclairage particulier sur l'obtention du système adimensionné de la dynamique des gaz et de sa limite en régime bas Mach dans la section 3.D, un résultat de stabilité L^2 dans le cas barotrope linéarisé dans la section 3.E, l'influence de la forme du maillage sur les résultats numériques à bas nombre de Mach dans la section 3.F et la comparaison à un schéma en coordonnée Eulérienne avec correction bas Mach dans la section 3.G.

An all-regime Lagrange-Projection like scheme for the gas dynamics equations on unstructured meshes

Abstract

We propose an *all-regime* Lagrange-Projection like numerical scheme for the gas dynamics equations. By *all-regime*, we mean that the numerical scheme is able to compute accurate approximate solutions with an under-resolved discretization, i.e. a mesh size and time step much bigger than the Mach number M. The key idea is to decouple acoustic and transport phenomenon and then alter the numerical flux in the acoustic approximation to obtain a uniform truncation error in term of M. This modified scheme is conservative and endowed with good stability properties with respect to the positivity of the density and the internal energy. A discrete entropy inequality under a condition on the modification is obtained thanks to a reinterpretation of the modified scheme in the Harten Lax and van Leer formalism. A natural extension to multi-dimensional problems discretized over unstructured mesh is proposed. Then a simple and efficient semi implicit scheme is also proposed. The resulting scheme is stable under a CFL condition driven by the (slow) material waves and not by the (fast) acoustic waves and so verifies the *all-regime* property. Numerical evidences are proposed and show the ability of the scheme to deal with tests where the flow regime may vary from low to high Mach values.

3.1 Introduction

In this paper, we consider the system of gas dynamics in two space dimension in situations when the flow regime may vary in terms of Mach number M across the computational domain. We propose a collocated Finite Volume method that addresses two important issues.

The first issue concerns the lack of accuracy in the low Mach regime of Godunov-type schemes. While these methods performs well at capturing shocks, they may generate spurious numerical diffusion when they are used for simulating low Mach flows over relatively coarse mesh, i.e. mesh size much bigger the Mach number. Improvements of Godunov-type schemes more generally of collocated methods have been proposed by many authors like [30, 17, 22, 5, 9, 7, 29, 19, 25, 24, 15, 11, 8, 20]. The analysis of these authors may rely on different arguments like the analysis of the viscosity matrix [30], an asymptotic expansion in terms of Mach number [17], a detailed study in [11] that seek for invariance properties of the numerical scheme transposing the framework of Schochet [26] to the discrete setting, and also an analysis based on the so-called Asymptotic Preserving property [21] in [20]. Nevertheless the resulting cure usually boils down to reduce the numerical diffusion in the momentum equation for low Mach number values.

The second problem we address deals with subsonic flow when the fluid velocity is slow and the acoustic waves are not driving phenomenons. In this case, the Courant-Friedrichs-Lewy (CFL) condition on the time step for explicit Godunov-type methods that involves the (fast) acoustic wave velocity may lead to very small time steps choices and thus costly computations. It seems natural to seek for numerical schemes that enable the use of a large time steps that are not constrained by the sound velocity. This question has been examined by several authors like [22, 7, 8, 9, 20] (see also [4, 6]) who derived mixed implicit-explicit strategies that allows to choose the time step independently of the Mach Number.

Numerical schemes that can tackle both issues, namely : accuracy for mesh sizes that do not depend on the Mach number and also stability for time steps that are not constrained by the Mach value are usually referred to as *all-regime*, like the methods proposed by [22, 7, 8, 9, 20].

In the present work, we first propose an operator splitting strategy that allows to decouple the acoustic and the transport phenomenons. The approximation algorithm is split into two steps : an acoustic step

and a transport step. For one-dimensional problems, this strategy is equivalent to an explicit Lagrange-Projection [16, 13] method, however the present splitting does not involve any moving Lagrangian mesh and can be naturally expressed for multi-dimensional problems. Following simple lines inspired by [11, 10] we investigate the dependence of the truncation error with respect to the Mach number. Let us mention that our study does not involve a Taylor expansion in the vicinity of the zero-Mach limit, nor a near-divergence free condition for the velocity field. Although this analysis is by no mean a thorough explanation of the low Mach regime behavior of our solver, it is enough to suggest simple means to obtain a truncation error with a uniform dependence on the Mach number for $M < 1$. The cure simply relies on modifying the pressure terms in the flux of the acoustic operator that is coherent with the correction proposed by [11, 10, 19, 25, 15].

Although this modified scheme is based on a modified flux definition, one can shows that it can also be rephrased as a simple approximate Riemann solver in the sense of Harten, Lax and van Leer [18] that is consistent with the integral form of the gas dynamics equation. This scheme is endowed with good stability properties under a CFL condition that involves the Mach number as the time step is still constrained by the sound velocity.

We propose to circumvent this time-step restriction by implementing a mixed implicit-explicit method following the ideas developped by [6] for one-dimensional problems using a genuine Lagrange-Projection framework. This idea was also used in [4] and consists in using an implicit update for the acoustic step and an explicit march in time for the transport step. This enables stability under a CFL condition that only involves the (slow) material waves without the (fast) acoustic waves. Finally, let us mention that the overall procedure is a conservative discretization that relies on a Suliciu relaxation approach [28] that allows to cope with compressible fluids equipped with very general Equation of State (EOS).

The paper is structured as follows : we first present the operator splitting considering only one-dimensional problems. Then we study the behavior of the scheme in the low Mach regime. This allows to lead to an explicit corrected scheme for the sole acoustic step that preserves the accuracy of the scheme at low Mach. Interestingly, we show that this flux-based corrected method may be expressed thanks to an approximate Riemann solver for the acoustic step. Next and thanks to this property, we investigate the ability of the corrected scheme to satisfy to a discrete entropy inequality. Afterwards, we present the extension of the operator splitting method to unstructured meshes either with a semi-implicit or full-explicit march in time. Finally we present numerical results involving low Mach and multi-regime flows.

3.2 Governing equations

We are interested in the two-dimensional gas dynamics equations

$$\begin{cases} \partial_t \rho + \nabla \cdot (\rho \mathbf{u}) = 0, & \text{(3.1a)} \\ \partial_t (\rho \mathbf{u}) + \nabla \cdot (\rho \mathbf{u} \otimes \mathbf{u}) + \nabla p = 0, & \text{(3.1b)} \\ \partial_t (\rho E) + \nabla \cdot [(\rho E + p)\mathbf{u}] = 0, & \text{(3.1c)} \end{cases}$$

where ρ, $\mathbf{u} = (u_1, u_2)^t$, E denote respectively the density, the velocity vector and the total energy of the fluid. Let $e = E - \frac{|\mathbf{u}|^2}{2}$ be the specific internal energy of the fluid and s its specific entropy. We note $\tau = 1/\rho$ and we suppose given an Equation of State (EOS) through the mapping $(\tau, s) \mapsto e^{\text{EOS}}$ which

satisfies to the usual Weyl assumptions [31]

$$\begin{array}{lll} \partial_\tau e^{\mathrm{EOS}} < 0, & \partial_s e^{\mathrm{EOS}} > 0, & \partial_{\tau\tau} e^{\mathrm{EOS}} > 0 \\ \partial_{ss} e^{\mathrm{EOS}} > 0, & \partial_{\tau\tau} e^{\mathrm{EOS}} \partial_{ss} e^{\mathrm{EOS}} > (\partial_{\tau s} e^{\mathrm{EOS}})^2, & \partial_{\tau\tau\tau} e^{\mathrm{EOS}} < 0. \end{array} \tag{3.2}$$

The entropy $s = s^{\mathrm{EOS}}(\tau, e)$ verifies $e = e^{\mathrm{EOS}}(\tau, s)$ thanks to (3.2) and we can define the pressure $p = -\partial_\tau e^{\mathrm{EOS}}$ and the sound velocity $c = \tau\sqrt{\partial_{\tau\tau} e^{\mathrm{EOS}}}$. The above assumptions imply that $(\tau, s) \mapsto e^{\mathrm{EOS}}$ and $(\tau, e) \mapsto -s^{\mathrm{EOS}}$ are strictly convex. Using a slight abuse of notation, we shall also consider p as a function of (τ, e) and note $p = p^{\mathrm{EOS}}(\tau, e)$.

3.3 Acoustic/transport operator splitting strategy for the one-dimensional problem

In this section we will consider for the sake of simplicity one-dimensional problems and propose a two-step approximation strategy based on an operator splitting. The aim of this splitting is to decouple acoustic and transport phenomena. Using this guideline we will propose an explicit numerical solver. We shall propose two simple extensions of this method to two-dimensional problems discretized over unstructured grids using either an explicit or a semi-implicit time update in section 3.5.4.

Before going any further, we introduce classical notations for the one-dimensional setting : let $\Delta t > 0$ and $\Delta x > 0$ be respectively the time and space steps. We define the Eulerian mesh interfaces $x_{j+1/2} = j\Delta x$ for $j \in \mathbb{Z}$, and the intermediate times $t^n = n\Delta t$ for $n \in \mathbb{N}$. If b is a fluid parameter, in the sequel, we will note b_j^n (resp. b_j^{n+1}) the approximate value b respectively within the j^{th} cell $[x_{j-1/2}, x_{j+1/2})$ at instant $t = t^n$ (resp. $t = t^{n+1}$).

For one-dimensional problems, (3.1) supplemented with a passive scalar variable v (that will account for the transverse velocity in two-dimensional problems) reads

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, & \text{(3.3a)} \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p) = 0, & \text{(3.3b)} \\ \partial_t(\rho v) + \partial_x(\rho u v) = 0, & \text{(3.3c)} \\ \partial_t(\rho E) + \partial_x[(\rho E + p)u] = 0. & \text{(3.3d)} \end{cases}$$

Our discretization strategy of (3.3) consists in approximating successively the solutions of the following systems (3.4) and (3.5) where

$$\begin{cases} \partial_t \rho + \rho\partial_x u = 0, & \text{(3.4a)} \\ \partial_t(\rho u) + \rho u\partial_x u + \partial_x p = 0, & \text{(3.4b)} \\ \partial_t(\rho v) + \rho v\partial_x u = 0, & \text{(3.4c)} \\ \partial_t(\rho E) + \rho E\partial_x u + \partial_x(pu) = 0, & \text{(3.4d)} \end{cases}$$

and

$$\begin{cases} \partial_t \rho + u\partial_x \rho = 0, & \text{(3.5a)} \\ \partial_t(\rho u) + u\partial_x(\rho u) = 0, & \text{(3.5b)} \\ \partial_t(\rho v) + u\partial_x(\rho v) = 0, & \text{(3.5c)} \\ \partial_t(\rho E) + u\partial_x(\rho E) = 0. & \text{(3.5d)} \end{cases}$$

In the sequel, system (3.4) and (3.5) will be respectively referred to as the acoustic system and the

transport system.

Given a fluid state $(\rho, \rho u, \rho v, \rho E)_j^n$, $j \in \mathbb{Z}$ at instant t^n, this splitting algorithm can be decomposed as follows.

1. Update the fluid state $(\rho, \rho u, \rho v, \rho E)_j^n$ to the value $(\rho, \rho u, \rho v, \rho E)_j^{n+1-}$ by approximating the solution of (3.4) ;
2. Update the fluid state $(\rho, \rho u, \rho v, \rho E)_j^{n+1-}$ to the value $(\rho, \rho u, \rho v, \rho E)_j^{n+1}$ by approximating the solution of (3.5).

3.3.1 Properties and approximation of the one-dimensional acoustic system

First, we notice that the acoustic system (3.4) reads equivalently

$$\partial_t \tau - \tau \partial_x u = 0, \qquad \partial_t u + \tau \partial_x p = 0, \qquad \partial_t v = 0, \qquad \partial_t E + \tau \partial_x (pu) = 0. \tag{3.6}$$

The acoustic system (3.6) is a quasilinear system that can be simply checked to be strictly hyperbolic. Indeed, the Jacobian of the system (3.6) has three eigenvalues $(\lambda_1, \lambda_2, \lambda_3) = (-c, 0, +c)$. The waves associated with λ_1 and λ_3 are genuinely nonlinear waves while the wave of velocity $\lambda_2 = 0$ is a stationnary contact discontinuity.

In order to derive an update process from $(\rho, \rho u, \rho v, \rho E)_j^n$ to $(\rho, \rho u, \rho v, \rho E)_j^{n+1-}$, we will perform several approximations. We notice that for a smooth solution (3.6) we also have $\partial_t p + \tau(\rho c)^2 \partial_x u = 0$ and we thus choose to perform a Suliciu-type approximation of (3.6) for $t \in [t^n, t^n + \Delta t)$ by introducing a surrogate pressure Π and considering the relaxed system

$$\begin{cases} \partial_t \tau - \tau \partial_x u = 0, & \text{(3.7a)} \\ \partial_t u + \tau \partial_x \Pi = 0, & \text{(3.7b)} \\ \partial_t v = 0, & \text{(3.7c)} \\ \partial_t E + \tau \partial_x (\Pi u) = 0, & \text{(3.7d)} \\ \partial_t \Pi + \tau a^2 \partial_x u = \nu(\Pi - p), & \text{(3.7e)} \end{cases}$$

where $a > 0$ is a parameter whose choice will be specified later. In the regime $\nu \to +\infty$ we formally recover (3.6). In our numerical solver context, we classically mimic the $\nu \to +\infty$ regime enforcing at each time step $\Pi_j^n = p^{\mathrm{EOS}}(\tau_j^n, e_j^n)$ and then solving (3.7) with $\nu = 0$.

At last, for $t \in [t^n, t^n + \Delta t)$ we choose to approximate $\tau(x,t)\partial_x$ by $\tau(x, t^n)\partial_x$ in (3.7). If one introduces the mass variable m defined by $\mathrm{d}m = \rho(x, t^n)\mathrm{d}x$ our approximation of (3.6) (up to an abuse of notation) can be expressed in the following fully conservative form

$$\partial_t \mathbf{W} + \partial_m \mathbf{F}(\mathbf{W}) = 0, \tag{3.8}$$

where $\mathbf{W} = (\tau, u, v, E, \Pi)^T$ and $\mathbf{F}(\mathbf{W}) = (-u, \Pi, 0, \Pi u, a^2 u)^T$. Let us remark that (3.8) is consistent with a Suliciu relaxation of the gas dynamics equation written in Lagrangian coordinates using a mass variable formulation. The solution of the Riemann problem associated with (3.8) can be derived explicitly (see section 3.C). This allows to write an exact Godunov solver for (3.8) that turns out to be an approximate Riemann solver for (3.6) following the Harten-Lax-van Leer formalism (see section 3.B and [18, 1]). It provides us with the update formula

$$\begin{cases} \mathbf{W}_j^{n+1-} = \mathbf{W}_j^n - \dfrac{\Delta t}{\Delta x}\left(\mathbf{F}_{j+1/2} - \mathbf{F}_{j-1/2}\right), & (3.9a) \\ \mathbf{F}_{j+1/2} = \mathbf{F}(\mathbf{W}_j^n, \mathbf{W}_{j+1}^n), & (3.9b) \\ \mathbf{F}(\mathbf{W}_L, \mathbf{W}_R) = (-u^*, \Pi^*, 0, \Pi^* u^*, a^2 u^*)^T, & (3.9c) \end{cases}$$

where

$$\begin{cases} u^* = \dfrac{(u_R + u_L)}{2} - \dfrac{1}{2a}(\Pi_R - \Pi_L), & (3.10a) \\ \Pi^* = \dfrac{(\Pi_R + \Pi_L)}{2} - \dfrac{a}{2}(u_R - u_L). & (3.10b) \end{cases}$$

The update of the conservative variables is obtained by setting $\rho_j^{n+1-} = 1/\tau_j^{n+1-}$, $(\rho u)_j^{n+1-} = \rho_j^{n+1-} \times u_j^{n+1-}$, $(\rho v)_j^{n+1-} = \rho_j^{n+1-} \times v_j^{n+1-}$ and $(\rho E)_j^{n+1-} = \rho_j^{n+1-} \times E_j^{n+1-}$. This can be summed up by the following update formulas

$$\begin{cases} L_j \rho_j^{n+1-} = \rho_j^n, & (3.11a) \\ L_j (\rho u)_j^{n+1-} = (\rho u)_j^n - \dfrac{\Delta t}{\Delta x}(\Pi_{j+1/2}^* - \Pi_{j-1/2}^*), & (3.11b) \\ L_j (\rho v)_j^{n+1-} = (\rho v)_j^n, & (3.11c) \\ L_j (\rho E)_j^{n+1-} = (\rho E)_j^n - \dfrac{\Delta t}{\Delta x}(\Pi_{j+1/2}^* u_{j+1/2}^* - \Pi_{j-1/2}^* u_{j-1/2}^*), & (3.11d) \\ L_j = 1 + \dfrac{\Delta t}{\Delta x}(u_{j+1/2}^* - u_{j-1/2}^*). & (3.11e) \end{cases}$$

Let us remark that (3.9) also proposes an update relation for Π. However in this case Π is just a disposable intermediate value whose role only consists in providing a formula for the interface pressure terms and the udpated value Π_k^{n+1-} will be discarded. Indeed in this explicit scheme, Π is updated after each time step by the equilibrium formula $\Pi_j^n = p^{\mathrm{EOS}}(\tau_j^n, e_j^n)$. However, this will no longer be the case for *semi-implicit strategy* as we will see in section 3.5.4.

Let us finally note that the relaxation scheme (3.9) is equivalent to the acoustic scheme [12]. In order to avoid numerical instabilities, the parameter a must complies with the subcharacteristic condition

$$a > \max \rho c, \tag{3.12}$$

for all possible values of ρc when considering a solution of the equilibrium system (3.6). In practice we will choose a value a_{LR} for each interface by setting

$$a_{LR} = K \max(\rho_L^n c_L^n, \rho_R^n c_R^n), \tag{3.13}$$

where $K \geq 1$, $LR = j + 1/2$, $L = j$ and $R = j + 1$. We refer the reader to [1, 3, 2, 13] and the reference therein for more details.

3.3.2 Properties and approximation of the one-dimensional transport system

The transport system equation discretization is quite simple. Indeed, system (3.5) is a quasi-linear hyperbolic system that only involves the transport of the conservative variables with the velocity u. We choose to approximate the solution of (3.5) thanks to a standard upwind Finite-Volume approximation for $\varphi \in \{\rho, \rho u, \rho v, \rho E\}$

$$\varphi_j^{n+1} = \varphi_j^{n+1-} - \frac{\Delta t}{\Delta x}\left(u_{j+1/2}^* \varphi_{j+1/2}^{n+1-} - u_{j-1/2}^* \varphi_{j-1/2}^{n+1-}\right) + \frac{\Delta t}{\Delta x}\varphi_j^{n+1-}\left(u_{j+1/2}^* - u_{j-1/2}^*\right), \tag{3.14}$$

where

$$\varphi_{j+1/2}^{n+1-} = \begin{cases} \varphi_j^{n+1-}, \text{ if } u_{j+1/2}^* \geq 0, \\ \varphi_{j+1}^{n+1-}, \text{ if } u_{j+1/2}^* < 0. \end{cases}$$

Let us finally remark that (3.14) can be recast into

$$\varphi_j^{n+1} = \varphi_j^{n+1-} L_j + \frac{\Delta t}{\Delta x} \left(u_{j+1/2}^* \varphi_{j+1/2}^{n+1-} - u_{j-1/2}^* \varphi_{j-1/2}^{n+1-} \right). \tag{3.15}$$

3.3.3 Properties of the operator splitting scheme

We present here a few properties of the operator splitting scheme defined by (3.9) and (3.14). Let us first remark that this algorithm performs the same update as a classical Lagrange-Remap (or equivalently Lagrange-Projection) algorithm for one-dimensional problems (see appendix 3.A) although the design of our algorithm does not involve a moving mesh for following the variables in a Lagrangian reference frame. *This feature will be the key element of the multi-dimensional extension of the present scheme.* It is also interesting to mention that the operator splitting strategy also provided a mean of treating the waves of the gas dynamics system (3.3) separately : the acoustic step only involves acoustic waves while freezing the transport waves. The transport step only deals with the contact discontinuity of the material transport. Let us mention that a similar operator splitting was used in [14].

The overall update from variable at instant t^n to variables at instant t^{n+1} is fully conservative with respect to ρ, ρu, ρv and ρE. Indeed, we have

$$\begin{cases} \rho_j^{n+1} = \rho_j^n + \dfrac{\Delta t}{\Delta x} \left(u_{j+1/2}^* \rho_{j+1/2}^{n+1-} - u_{j-1/2}^* \rho_{j-1/2}^{n+1-} \right), & \text{(3.16a)} \\ (\rho u)_j^{n+1} = (\rho u)_j^n + \dfrac{\Delta t}{\Delta x} \left(u_{j+1/2}^* (\rho u)_{j+1/2}^{n+1-} + \Pi_{j+1/2}^* - u_{j-1/2}^* (\rho u)_{j-1/2}^{n+1-} - \Pi_{j-1/2}^* \right), & \text{(3.16b)} \\ (\rho v)_j^{n+1} = (\rho v)_j^n + \dfrac{\Delta t}{\Delta x} \left(u_{j+1/2}^* (\rho v)_{j+1/2}^{n+1-} - u_{j-1/2}^* (\rho v)_{j-1/2}^{n+1-} - \right), & \text{(3.16c)} \\ (\rho E)_j^{n+1} = (\rho E)_j^n + \dfrac{\Delta t}{\Delta x} \left(u_{j+1/2}^* (\rho E)_{j+1/2}^{n+1-} + \Pi_{j+1/2}^* u_{j+1/2}^* \right) \\ \qquad\qquad - \dfrac{\Delta t}{\Delta x} \left(u_{j-1/2}^* (\rho E)_{j-1/2}^{n+1-} + \Pi_{j-1/2}^* u_{j-1/2}^* \right). & \text{(3.16d)} \end{cases}$$

The scheme (3.9)-(3.10) for the acoustic step is stable under the Courant-Friedrichs-Lewy (CFL) condition

$$\frac{\Delta t}{\Delta x} \max_{j \in \mathbb{Z}} \left(\max(\tau_j^n, \tau_{j+1}^n) a_{j+1/2} \right) \leq \frac{1}{2}. \tag{3.17}$$

If one notes $b^{\pm} = \frac{b \pm |b|}{2}$, then a classical result states that the CFL condition associated with the transport scheme (3.14) reads

$$\Delta t \max_{j \in \mathbb{Z}} \left((u_{j-\frac{1}{2}}^*)^+ - (u_{j+\frac{1}{2}}^*)^- \right) < \Delta x. \tag{3.18}$$

Entropy-related stability properties of the scheme will be examined in section 3.5.3.

One can also remark that both the acoustic steps and the transport steps are achieved thanks to genuine Godunov solvers applied to simplified subsystems.

3.4 Behavior of the scheme with respect to the Mach regime

We are now interested in the behavior of the numerical scheme with respect to the variations of the Mach regime. In order to characterize the Mach regime of the flow, we consider a classical rescaling of

the equations (3.3) : let us introduce the following non-dimensional quantities :

$$\tilde{x} = \frac{x}{L}, \quad \tilde{t} = \frac{t}{T}, \quad \tilde{\rho} = \frac{\rho}{\rho_0}, \quad \tilde{u} = \frac{u}{u_0}, \quad \tilde{v} = \frac{v}{v_0}, \quad \tilde{e} = \frac{e}{e_0}, \quad \tilde{p} = \frac{p}{p_0}, \quad \tilde{c} = \frac{c}{c_0}. \tag{3.19}$$

The parameters L, T, $u_0 = v_0 = \frac{L}{T}$, ρ_0, $e_0 = p_0\rho_0$, p_0 and $c_0 = \sqrt{\frac{p_0}{\rho_0}}$ denote respectively a characteristic length, time, velocity, density, internal energy, pressure and sound speed. If $M = \frac{u_0}{c_0}$ is the so-called Mach-number then system (3.3) reads

$$\begin{cases} \partial_{\tilde{t}}\tilde{\rho} + \partial_{\tilde{x}}(\tilde{\rho}\tilde{u}) = 0, & \text{(3.20a)} \\ \partial_{\tilde{t}}(\tilde{\rho}\tilde{u}) + \partial_{\tilde{x}}(\tilde{\rho}\tilde{u}^2) + \dfrac{1}{M^2}\partial_{\tilde{x}}\tilde{p} = 0, & \text{(3.20b)} \\ \partial_{\tilde{t}}(\tilde{\rho}\tilde{v}) + \partial_{\tilde{x}}(\tilde{\rho}\tilde{u}\tilde{v}) = 0, & \text{(3.20c)} \\ \partial_{\tilde{t}}(\tilde{\rho}\tilde{E}) + \partial_{\tilde{x}}[(\tilde{\rho}\tilde{E} + \tilde{p})\tilde{u}] = 0, & \text{(3.20d)} \end{cases}$$

where $\tilde{E} = \tilde{e} + \frac{1}{2}M^2\tilde{u}^2$. For a given small value of the Mach number, we distinguish two cases :

- the term $\partial_{\tilde{x}}\tilde{p}$ remains of magnitude $O(M^2)$. Then the variations of $\tilde{\rho}\tilde{u}$ are of order 1 which implies that all the tilde variables will remain of order 1. We shall refer this case as the low Mach regime ;
- the term $\partial_{\tilde{x}}\tilde{p}$ does not remain of magnitude $O(M^2)$. Then the variations of $\tilde{\rho}\tilde{u}$ will reach a magnitude $O(1/M)$ or $O(1/M^2)$. These large magnitude variations of the momentum will induce a growth of the Mach number and thus a change of Mach regime.

Before going any further, let us underline that in the present approach we do not intend to study the behavior of the rescaled system (3.3) in the limit regime $M \to 0$. This delicate question has been widely investigated over the past years and is still a rich field of research [17, 11, 20]. We focus here on a simpler task that consists in examining the consistency of a rescaled approximate solution provided by the splitting operator algorithm with the solution of (3.20) in the low Mach regime. The framework we will place ourselves in does not require sophisticated hypotheses and may deal with the evaluation of a local behavior of the solution (a few neighbouring cells in the discrete setting). More precisely, if one considers smooth solutions of (3.20) and considers the truncation error of the rescaled numerical scheme in the sense of Finite Difference, how does it depends on M in the low Mach regime ?

Introducing the rescaling defined earlier into (3.10) we get

$$\tilde{u}^*_{j+1/2} = \frac{1}{2}(\tilde{u}^n_j + \tilde{u}^n_{j+1}) - \frac{1}{2\tilde{a}_{j+1/2}M}(\tilde{\Pi}^n_{j+1} - \tilde{\Pi}^n_j), \quad \tilde{\Pi}^*_{j+1/2} = \frac{1}{2}(\tilde{\Pi}^n_j + \tilde{\Pi}^n_{j+1}) - \frac{\tilde{a}_{j+1/2}M}{2}(\tilde{u}^n_{j+1} - \tilde{u}^n_j),$$

for (3.11) we have

$$\begin{cases} \tilde{L}_j\tilde{\rho}^{n+1-}_j = \tilde{\rho}^n_j, & \text{(3.21a)} \\ \tilde{L}_j(\tilde{\rho}\tilde{u})^{n+1-}_j = (\tilde{\rho}\tilde{u})^n_j - \dfrac{\Delta\tilde{t}}{M^2\Delta\tilde{x}}(\tilde{\Pi}^*_{j+1/2} - \tilde{\Pi}^*_{j-1/2}), & \text{(3.21b)} \\ \tilde{L}_j(\tilde{\rho}\tilde{v})^{n+1-}_j = (\tilde{\rho}\tilde{v})^n_j, & \text{(3.21c)} \\ \tilde{L}_j(\tilde{\rho}\tilde{E})^{n+1-}_j = (\tilde{\rho}\tilde{E})^n_j - \dfrac{\Delta\tilde{t}}{\Delta\tilde{x}}(\tilde{\Pi}^*_{j+1/2}\tilde{u}^*_{j+1/2} - \tilde{\Pi}^*_{j-1/2}\tilde{u}^*_{j-1/2}), & \text{(3.21d)} \\ \tilde{L}_j = L_j = 1 + \dfrac{\Delta\tilde{t}}{\Delta\tilde{x}}(\tilde{u}^*_{j+1/2} - \tilde{u}^*_{j-1/2}), & \text{(3.21e)} \end{cases}$$

and finally if $\tilde{\varphi} \in \{\tilde{\rho}, \tilde{\rho}\tilde{u}, \tilde{\rho}\tilde{v}, \tilde{\rho}\tilde{E}\}$ the rescaling of (3.15), reads

$$\frac{1}{\Delta\tilde{t}}(\tilde{\varphi}^{n+1}_j - \tilde{L}_j\tilde{\varphi}^{n+1-}_j) + \frac{1}{\Delta\tilde{x}}\left(\tilde{\varphi}^{n+1-}_{j+1/2}\tilde{u}^*_{j+1/2} - \tilde{u}^*_{j-1/2}\tilde{\varphi}^{n+1-}_{j-1/2}\right) = 0. \tag{3.22}$$

Note that the CFL restriction of the acoustic step reads now

$$\frac{\Delta\tilde{t}}{\Delta\tilde{x}}\max(\tilde{\tau}_j^n,\tilde{\tau}_{j+1}^n)\tilde{a}_{j+1/2}^n \leq \frac{M}{2}, \tag{3.23}$$

while the CFL restriction associated with the transport step is

$$\left((\tilde{u}_{j-1/2}^*)^+ - (\tilde{u}_{j+1/2}^*)^-\right)\frac{\Delta\tilde{t}}{\Delta\tilde{x}} \leq 1. \tag{3.24}$$

In order to evaluate the truncation error (in the Finite Difference sense) in the low Mach regime, we use the classical tool of equivalent equations. Let $(\tilde{x},\tilde{t}) \mapsto \tilde{b}$ be a parameter of (rescaled) functions that describe a smooth flow. With a classical slight abuse of notation, we consider that $\tilde{\varphi}(x_j,t^n) = \tilde{\varphi}_j^n$ so that we can substitute these functions into the discrete update formula when $\tilde{\varphi} \in \{\tilde{\rho},\tilde{u},\tilde{v},\tilde{E},\tilde{\Pi}\}$. We suppose that we are in low Mach regime, namely $\partial_{\tilde{x}}\tilde{p} = O(M^2)$. This hypothesis yields that $\tilde{\Pi}_{j+1}^n = \tilde{\Pi}_j^n + O(M^2\Delta\tilde{x})$ for the discrete unknowns. We have the following result.

Proposition 2. *In the low Mach regime, the rescaled discretization of the acoustic step is consistent with*

$$\partial_{\tilde{t}}\tilde{\tau} - \tilde{\tau}\partial_{\tilde{x}}\tilde{u} = O(\Delta\tilde{t}) + O(M\Delta\tilde{x}), \qquad \partial_{\tilde{t}}\tilde{u} + \frac{\tilde{\tau}}{M^2}\partial_{\tilde{x}}\tilde{p} = O(\Delta\tilde{t}) + O\left(\frac{\Delta\tilde{x}}{M}\right),$$
$$\partial_{\tilde{t}}\tilde{v} = O(\Delta\tilde{t}), \qquad \partial_{\tilde{t}}\tilde{E} + \tilde{\tau}\partial_{\tilde{x}}(\tilde{p}\tilde{u}) = O(\Delta\tilde{t}) + O(M\Delta\tilde{x}).$$

The rescaled discretization of the transport step is consistent with

$$\partial_{\tilde{t}}\tilde{\varphi} + \tilde{u}\partial_{\tilde{x}}\tilde{\varphi} = O(\Delta\tilde{t}) + O(\Delta\tilde{x}) + O(M\Delta\tilde{x}),$$

and the equivalent equation verified by the rescaled scheme reads

$$\begin{cases} \partial_{\tilde{t}}\tilde{\rho} + \partial_{\tilde{x}}(\tilde{\rho}\tilde{u}) = O(\Delta\tilde{t}) + O(\Delta\tilde{x}) + O(M\Delta\tilde{x}), & \text{(3.25a)} \\ \partial_{\tilde{t}}(\tilde{\rho}\tilde{u}) + \partial_{\tilde{x}}(\tilde{\rho}\tilde{u}^2) + \dfrac{1}{M^2}\partial_{\tilde{x}}\tilde{p} = O(\Delta\tilde{t}) + O(\Delta\tilde{x}) + O(M\Delta\tilde{x}) + O\left(\dfrac{\Delta\tilde{x}}{M}\right), & \text{(3.25b)} \\ \partial_{\tilde{t}}(\tilde{\rho}\tilde{v}) + \partial_{\tilde{x}}(\tilde{\rho}\tilde{u}\tilde{v}) = O(\Delta\tilde{t}) + O(\Delta\tilde{x}) + O(M\Delta\tilde{x}), & \text{(3.25c)} \\ \partial_{\tilde{t}}(\tilde{\rho}\tilde{E}) + \partial_{\tilde{x}}[(\tilde{\rho}\tilde{E} + \tilde{p})\tilde{u}] = O(\Delta\tilde{t}) + O(\Delta\tilde{x}) + O(M\Delta\tilde{x}). & \text{(3.25d)} \end{cases}$$

Proof. There exists three smooth functions A, B and C of magnitude 1 with respect to M such that

$$\tilde{u}_{j+1/2}^* = \frac{\tilde{u}_{j+1}^n + \tilde{u}_j^n}{2} + M\Delta\tilde{x}A(x_{j+1/2},t^n) + O(M\Delta\tilde{x}^2),$$
$$\tilde{\Pi}_{j+1/2}^* = \frac{\tilde{\Pi}_{j+1}^n + \tilde{\Pi}_j^n}{2} + M\Delta\tilde{x}B(x_{i+1/2},t^n) + O(M\Delta\tilde{x}^2),$$
$$\tilde{\Pi}_{j+1/2}^*\tilde{u}_{j+1/2}^* = \frac{(\tilde{u}_{j+1}^n + \tilde{u}_j^n)(\tilde{\Pi}_{j+1}^n + \tilde{\Pi}_j^n)}{4} + M\Delta\tilde{x}C(x_{i+1/2},t^n) + O(M\Delta\tilde{x}^2).$$

Injecting the above relation into (3.21) we get

$$
\begin{cases}
\tilde{L}_j = 1 + \Delta\tilde{t}\dfrac{\tilde{u}^n_{j+1} - \tilde{u}^n_{j-1}}{2\Delta\tilde{x}} + O(M\Delta\tilde{x}\Delta\tilde{t}). & (3.26a)\\
\tilde{L}_j\tilde{\rho}^{n+1-}_j = \tilde{\rho}^n_j, & (3.26b)\\
\tilde{L}_j(\tilde{\rho}\tilde{u})^{n+1-}_j = (\tilde{\rho}\tilde{u})^n_j - \dfrac{\Delta\tilde{t}}{M^2}\dfrac{\tilde{\Pi}^n_{j+1} - \tilde{\Pi}^n_{j-1}}{2\Delta\tilde{x}} + O\left(\dfrac{\Delta\tilde{x}\Delta\tilde{t}}{M}\right), & (3.26c)\\
\tilde{L}_j(\tilde{\rho}_j\tilde{v})^{n+1-} = (\tilde{\rho}\tilde{v})^n_j, & (3.26d)\\
L_j(\tilde{\rho}\tilde{E})^{n+1-}_j = (\tilde{\rho}\tilde{E})^n_j - \Delta\tilde{t}\left(\dfrac{(\tilde{u}^n_{j+1} + \tilde{u}^n_j)(\tilde{\Pi}^n_{j+1} + \tilde{\Pi}^n_j)}{4\Delta\tilde{x}} - \dfrac{(\tilde{u}^n_{j-1} + \tilde{u}^n_j)(\tilde{\Pi}^n_{j-1} + \tilde{\Pi}^n_j)}{4\Delta\tilde{x}}\right)\\
\qquad + O(M\Delta\tilde{x}\Delta\tilde{t}), & (3.26e)
\end{cases}
$$

This yields

$$
\begin{cases}
\tilde{L}_j\tilde{\rho}^{n+1-}_j = \tilde{\rho}^n_j, & (3.27a)\\
\tilde{L}_j(\tilde{\rho}\tilde{u})^{n+1-}_j = (\tilde{\rho}\tilde{u})^n_j - \dfrac{\Delta\tilde{t}}{M^2}\partial_{\tilde{x}}\tilde{p} + O\left(\dfrac{\Delta\tilde{x}\Delta\tilde{t}}{M}\right) + O\left(\Delta\tilde{x}^2\Delta\tilde{t}\right), & (3.27b)\\
\tilde{L}_j(\tilde{\rho}_j\tilde{v})^{n+1-} = (\tilde{\rho}\tilde{v})^n_j, & (3.27c)\\
L_j(\tilde{\rho}\tilde{E})^{n+1-}_j = (\tilde{\rho}\tilde{E})^n_j - \Delta\tilde{t}\partial_{\tilde{x}}(\tilde{p}\tilde{u}) + O(M\Delta\tilde{x}\Delta\tilde{t}) + O(\Delta\tilde{x}^2\Delta\tilde{t}), & (3.27d)\\
\tilde{L}_j = 1 + \Delta\tilde{t}\partial_{\tilde{x}}\tilde{u} + O(M\Delta\tilde{x}\Delta\tilde{t}) + O(\Delta\tilde{x}^2\Delta\tilde{t}). & (3.27e)
\end{cases}
$$

Remark 1. *For smooth solutions in the low Mach regime, we have $\partial_{\tilde{x}\tilde{x}}\tilde{p} = O(M^2)$. We used this relation to obtain the term $O\left(\Delta\tilde{x}^2\Delta\tilde{t}\right)$ in* (3.27b).

Let us remark that (3.27) is indeed consistent at order 1 with respect to Δx with

$$
\partial_{\tilde{t}}\tilde{\tau} - \tilde{\tau}\partial_{\tilde{x}}\tilde{u} = O(\Delta\tilde{t}) + O(M\Delta\tilde{x}), \qquad \partial_{\tilde{t}}\tilde{u} + \frac{\tilde{\tau}}{M^2}\partial_{\tilde{x}}\tilde{p} = O(\Delta\tilde{t}) + O\left(\frac{\Delta\tilde{x}}{M}\right),
$$
$$
\partial_{\tilde{t}}\tilde{v} = O(\Delta\tilde{t}), \qquad \partial_{\tilde{t}}\tilde{E} + \tilde{\tau}\partial_{\tilde{x}}(\tilde{p}\tilde{u}) = O(\Delta\tilde{t}) + O(M\Delta\tilde{x}).
$$

Now we turn to the transport step. Accounting for the low Mach hypothesis, (3.22) becomes

$$
\frac{1}{\Delta\tilde{t}}(\tilde{\varphi}^{n+1}_j - \tilde{L}_j\tilde{\varphi}^{n+1-}_j) + \frac{1}{2\Delta\tilde{x}}\left(\tilde{\varphi}^{n+1-}_{j+1/2}(\tilde{u}^n_{j+1} + \tilde{u}^n_j) - \tilde{\varphi}^{n+1-}_{j-1/2}(\tilde{u}^n_j + \tilde{u}^n_{j-1})\right) = O(M\Delta\tilde{x}),
$$

hence

$$
\frac{1}{\Delta\tilde{t}}(\tilde{\varphi}^{n+1}_j - \tilde{L}_j\tilde{\varphi}^{n+1-}_j) + \partial_{\tilde{x}}(\tilde{\varphi}\tilde{u}) = O(\Delta\tilde{x}) + O(M\Delta\tilde{x}), \tag{3.28}
$$

which is consistent with $\partial_{\tilde{t}}\tilde{\varphi} + \tilde{u}\partial_{\tilde{x}}\tilde{\varphi} = O(\Delta\tilde{t}) + O(\Delta\tilde{x}) + O(M\Delta\tilde{x})$. Finally, using (3.27) into (3.28) we finally obtain the desired result. □

Remark 2. *It is important to note that the analysis we proposed in this section cannot be considered as an exhaustive explanation for the behavior of the numerical scheme in the Low Mach regime. It just merely provides magnitude estimate of the truncation error. Considering the same lines with additional hypotheses : $\tilde{\rho}$, $\tilde{u}$, $\tilde{v}$, $\tilde{E}$ are solution of the rescaled gas dynamics equations in the low Mach regime with well-prepared conditions [11], then one can show that the $O(\Delta\tilde{x}/M)$ term in* (3.25b) *does vanish [11] for one-dimensional problems set over the whole real line. The analysis is delicate and depends on many hypotheses : for two-dimensional problems same results can be obtained for discretization over a triangular mesh with periodic boundary conditions. However, this no longer works for two-dimensional Cartesian meshes where the classical Godunov-type solvers perform poorly with periodic boundary conditions. More general boundary conditions require a specific study for each case [11, 10].*

3.5 Low Mach correction

The equivalent equation (3.25) satisfied by the rescaled scheme is clearly not satisfactory because of the term $O(\frac{\Delta\tilde{x}}{M})$ which behaves badly when $M \ll \Delta\tilde{x}$. This suggests to modify the scheme accordingly.

3.5.1 Correction of the low Mach behavior : a simple flux modification

In the light of the previous asymptotic analysis, we propose to leave the projection step unchanged and rather focus on the acoustic step of the scheme. In the acoustic step, we suggest to simply replace $\Pi^*_{j+1/2}$ by

$$\Pi^{*,\theta}_{j+1/2} = \frac{1}{2}(\Pi^n_j + \Pi^n_{j+1}) - \theta_{j+1/2}\frac{a_{j+1/2}}{2}(u^n_{j+1} - u^n_j). \tag{3.29}$$

The associated dimensionless flux reads

$$\tilde{\Pi}^{*,\theta}_{j+1/2} = \frac{1}{2}(\tilde{\Pi}^n_j + \tilde{\Pi}^n_{j+1}) - \theta_{j+1/2}\frac{\tilde{a}_{j+1/2}M}{2}(\tilde{u}^n_{j+1} - \tilde{u}^n_j). \tag{3.30}$$

This yields the following modified scheme for the acoustic step.

$$\begin{cases} \mathbf{W}^{n+1-}_j = \mathbf{W}^n_j - \dfrac{\Delta t}{\Delta x}\left(\mathbf{F}_{j+1/2} - \mathbf{F}_{j-1/2}\right), & \text{(3.31a)} \\ \mathbf{F}_{j+1/2} = \mathbf{F}^\theta(\mathbf{W}^n_j, \mathbf{W}^n_{j+1}), & \text{(3.31b)} \\ \mathbf{F}^\theta(\mathbf{W}_L, \mathbf{W}_R) = (-u^*, \Pi^{*,\theta}, 0, \Pi^{*,\theta}u^*, a^2u^*)^T. & \text{(3.31c)} \end{cases}$$

Let us underline that this modification solely alters the non-centered terms of the pressure flux. In other words this does not modify the ultimate consistency of $\Pi^{*,\theta}_{j+1/2}$ with the pressure value, it does impact the numerical dissipation involved with the discretization of the pressure terms. This approach complies with several previous works that have been investigating the approximation of the low Mach regime like [19, 25, 15]. While such modification is usually delicate with regards to the stability of the numerical scheme, we will nevertheless see that the resulting modified numerical scheme is still endowed with stability properties (see section 3.5.3).

In the sequel, in order to perform an equivalent equation analysis with the modified pressure flux, we consider a smooth function $x \mapsto \theta$ such that $\theta_{j+1/2} = \theta(x_{j+1/2})$. We have the following consistency properties for the numerical scheme with the modified pressure flux $\Pi^{*,\theta}_{j+1/2}$.

Proposition 3. *In the low Mach regime, the rescaled discretization (3.31) of the acoustic step is consistent with*

$$\partial_{\tilde{t}}\tilde{\tau} - \tilde{\tau}\partial_{\tilde{x}}\tilde{u} = O(\Delta\tilde{t}) + O(M\Delta\tilde{x}), \qquad \partial_{\tilde{t}}\tilde{u} + \frac{\tilde{\tau}}{M^2}\partial_{\tilde{x}}\tilde{p} = O(\Delta\tilde{t}) + O\left(\frac{\theta\Delta\tilde{x}}{M}\right),$$
$$\partial_{\tilde{t}}\tilde{v} = O(\Delta\tilde{t}), \qquad \partial_{\tilde{t}}\tilde{E} + \tilde{\tau}\partial_{\tilde{x}}(\tilde{p}\tilde{u}) = O(\Delta\tilde{t}) + O(M\Delta\tilde{x}) + O(M\theta\Delta\tilde{x}).$$

The rescaled discretization of the transport step is consistent with

$$\partial_{\tilde{t}}\tilde{\varphi} + \tilde{u}\partial_{\tilde{x}}\tilde{\varphi} = +O(\Delta\tilde{t}) + O(\Delta\tilde{x}) + O(M\Delta\tilde{x}),$$

and the equivalent equation verified by the rescaled scheme reads

$$\begin{cases} \partial_{\tilde{t}}\tilde{\rho} + \partial_{\tilde{x}}(\tilde{\rho}\tilde{u}) = O(\Delta\tilde{t}) + O(\Delta\tilde{x}) + O(M\Delta\tilde{x}), & (3.32a) \\ \partial_{\tilde{t}}(\tilde{\rho}\tilde{u}) + \partial_{\tilde{x}}(\tilde{\rho}\tilde{u}^2)\dfrac{1}{M^2} + \partial_{\tilde{x}}\tilde{p} = O(\Delta\tilde{t}) + O(\Delta\tilde{x}) + O\left(\dfrac{\theta\Delta\tilde{x}}{M}\right), & (3.32b) \\ \partial_{\tilde{t}}(\tilde{\rho}\tilde{v}) + \partial_{\tilde{x}}(\tilde{\rho}\tilde{u}\tilde{v}) = O(\Delta\tilde{t}) + O(\Delta\tilde{x}) + O(M\Delta\tilde{x}). & (3.32c) \\ \partial_{\tilde{t}}(\tilde{\rho}\tilde{E}) + \partial_{\tilde{x}}[(\tilde{\rho}\tilde{E} + \tilde{p})\tilde{u}] = O(\Delta\tilde{t}) + O(\Delta\tilde{x}) + O(M\Delta\tilde{x}) + O(M\theta\Delta\tilde{x}). & (3.32d) \end{cases}$$

As a consequence, provided that we impose the asymptotic behavior $\theta_{j+1/2} = O(M)$, the truncation error is uniform with respect to M.

Proof. Following similar lines as in the proof of proposition 2 and using the same notations, there exists four smooth functions A, B, C and D of magnitude 1 with respect to M such that

$$\begin{aligned} \tilde{u}^*_{j+1/2} &= \frac{\tilde{u}^n_{j+1} + \tilde{u}^n_j}{2} + M\Delta\tilde{x}A(x_{j+1/2}, t^n) + O(M\Delta\tilde{x}^2), \\ \tilde{\Pi}^{*,\theta}_{j+1/2} &= \frac{\tilde{p}^n_{j+1} + \tilde{p}^n_j}{2} + \theta_{j+1/2}M\Delta\tilde{x}B(x_{i+1/2}, t^n) + O(M\Delta\tilde{x}^2), \\ \tilde{\Pi}^{*,\theta}_{j+1/2}\tilde{u}^*_{j+1/2} &= \frac{(\tilde{u}^n_{j+1} + \tilde{u}^n_j)(\tilde{p}^n_{j+1} + \tilde{p}^n_j)}{4} + M\Delta\tilde{x}C(x_{i+1/2}, t^n) + M\theta_{j+1/2}\Delta\tilde{x}D(x_{i+1/2}, t^n) + O(M\Delta\tilde{x}^2). \end{aligned}$$

The rest of the analysis follows the same line as the proof of proposition 2. Using (3.21) we get

$$\begin{cases} \tilde{L}_j\tilde{\rho}^{n+1-}_j = \tilde{\rho}^n_j, & (3.33a) \\ \tilde{L}_j(\tilde{\rho}\tilde{u})^{n+1-}_j = (\tilde{\rho}\tilde{u})^n_j - \dfrac{\Delta\tilde{t}}{M^2}\partial_{\tilde{x}}\tilde{p} + O\left(\dfrac{\theta\Delta\tilde{x}\Delta\tilde{t}}{M}\right) + O\left(\Delta\tilde{x}^2\Delta\tilde{t}\right), & (3.33b) \\ \tilde{L}_j(\tilde{\rho}_j\tilde{v})^{n+1-} = (\tilde{\rho}\tilde{v})^n_j, & (3.33c) \\ L_j(\tilde{\rho}\tilde{E})^{n+1-}_j = (\tilde{\rho}\tilde{E})^n_j - \Delta\tilde{t}\partial_{\tilde{x}}(\tilde{p}\tilde{u}) + O(M\Delta\tilde{x}\Delta\tilde{t}) + O(M\theta\Delta\tilde{x}\Delta\tilde{t}) + O(\Delta\tilde{x}^2\Delta\tilde{t}), & (3.33d) \\ \tilde{L}_j = 1 + \Delta\tilde{t}\partial_{\tilde{x}}\tilde{u} + O(M\Delta\tilde{x}\Delta\tilde{t}) + O(\Delta\tilde{x}^2\Delta\tilde{t}). & (3.33e) \end{cases}$$

and (3.22) yields again

$$\frac{1}{\Delta\tilde{t}}(\tilde{\varphi}^{n+1}_j - \tilde{L}_j\tilde{\varphi}^{n+1-}_j) + \partial_{\tilde{x}}(\tilde{\varphi}\tilde{u}) = O(\Delta\tilde{x}) + O(M\Delta\tilde{x}). \tag{3.34}$$

Relations (3.33) and (3.34) provides the desired results. □

Remark 3. *In the light of the truncation error that appears in* (3.25)*, one can see that it is not necessary to involve a correction for the energy flux term in* (3.31c)*. It would be possible to consider a numerical scheme with the definition* (3.10a) *for the velocity at the interface, the modified pressure* (3.29) *for interface pressure terms and $\Pi^* u^*$ for the energy flux.*

3.5.2 Approximate Riemann solver for the modified acoustic scheme

The modified numerical scheme (3.31) for the acoustic step belongs to the category of flux-based solver. Indeed, this solver relies on an update formula (3.31a) that involves the modified flux (3.31c). We will prove in this section that this modified flux solver can also be obtained thanks to an approximate Riemann solver in the sense of Harten, Lax and van Leer [18, 1], see also Annex B for a quick refresh on this, that is consistent with the integral form of (3.8). This formalism is useful to establish stability properties. We have the following proposition.

Proposition 4. *There exists a simple approximate Riemann solver that is an approximation of the Riemann problem associated with the relaxed acoustic problem* (3.8) *and whose associated flux matches the flux of the modified acoustic solver. More precisely, there exists a self-similar function*

$$\mathbf{W}_{RP}^{\theta}\left(\frac{m}{t};\mathbf{W}_L,\mathbf{W}_R\right)=(\tau,u,v,E,\Pi)\left(\frac{m}{t};\mathbf{W}_L,\mathbf{W}_R\right)=\begin{cases}\mathbf{W}_L, & \text{if} \quad m/t<-a,\\ \mathbf{W}_L^{*,\theta}, & \text{if} \quad -a\le m/t<0,\\ \mathbf{W}_R^{*,\theta}, & \text{if} \quad 0\le m/t<+a,\\ \mathbf{W}_R, & \text{if} \quad +a\le m/t.\end{cases} \tag{3.35}$$

such that

$$\begin{aligned}\mathbf{F}^{\theta}(\mathbf{W}_R,\mathbf{W}_L)&=\mathbf{F}(\mathbf{W}_L)-\int_{-\infty}^{0}[\mathbf{W}_{RP}^{\theta}(\xi;\mathbf{W}_L,\mathbf{W}_R)-\mathbf{W}_L]\,\mathrm{d}\xi\\ &=\mathbf{F}(\mathbf{W}_R)+\int_{0}^{+\infty}[\mathbf{W}_{RP}^{\theta}(\xi;\mathbf{W}_L,\mathbf{W}_R)-\mathbf{W}_R]\,\mathrm{d}\xi\\ &=\frac{1}{2}\left(\mathbf{F}(\mathbf{W}_L)+\mathbf{F}(\mathbf{W}_R)\right)-\frac{a}{2}\left(\mathbf{W}_L^{*,\theta}-\mathbf{W}_L\right)-\frac{a}{2}\left(\mathbf{W}_R-\mathbf{W}_R^{*,\theta}\right).\end{aligned} \tag{3.36}$$

The states $\mathbf{W}_L^{*,\theta}=(\tau_L^{*,\theta},u_L^{*,\theta},v_L^{*,\theta},\Pi_L^{*,\theta})^T$ *and* $\mathbf{W}_R^{*,\theta}=(\tau_R^{*,\theta},u_R^{*,\theta},v_R^{*,\theta},\Pi_R^{*,\theta})^T$ *are given by*

$$\tau_L^{*,\theta}=\tau_L+\frac{1}{a}(u^*-u_L), \qquad \tau_R^{*,\theta}=\tau_R+\frac{1}{a}(u_R-u^*), \tag{3.37a}$$

$$u_L^{*,\theta}=u^*+\frac{1}{2}(\theta-1)(u_R-u_L), \qquad u_R^{*,\theta}=u^*+\frac{1}{2}(1-\theta)(u_R-u_L), \tag{3.37b}$$

$$v_L^{*,\theta}=v_L, \qquad v_R^{*,\theta}=v_R, \tag{3.37c}$$

$$E_L^{*,\theta}=E_L+\frac{1}{a}(\Pi_L u_L-\Pi^{*,\theta}u^*), \qquad E_R^{*,\theta}=E_R+\frac{1}{a}(\Pi^{*,\theta}u^*-\Pi_R u_R) \tag{3.37d}$$

$$\Pi_L^{*,\theta}=\Pi^*, \qquad \Pi_R^{*,\theta}=\Pi^*. \tag{3.37e}$$

Proof. Suppose that $\mathbf{W}_{\mathrm{RP}}^{\theta}$ is consistent with the integral form of the relaxed acoustic problem (3.8) then for a given $\mathbf{W}_L$ and $\mathbf{W}_R$ we have

$$\mathbf{F}(\mathbf{W}_R)-\mathbf{F}(\mathbf{W}_L)=-a(\mathbf{W}_L^{*,\theta}-\mathbf{W}_L)+a(\mathbf{W}_R-\mathbf{W}_R^{*,\theta}),$$

which reads

$$\mathbf{W}_R^{*,\theta}+\mathbf{W}_L^{*,\theta}=\mathbf{W}_R+\mathbf{W}_L-\frac{1}{a}(\mathbf{F}(\mathbf{W}_R)-\mathbf{F}(\mathbf{W}_L)). \tag{3.38}$$

If the resulting flux of this approximate Riemann solver is $F^{\theta}(\mathbf{W}_L,\mathbf{W}_R)$ then (3.36) is verified and yields

$$2F^{\theta}(\mathbf{W}_L,\mathbf{W}_R)=\mathbf{F}(\mathbf{W}_R)+\mathbf{F}(\mathbf{W}_L)-a(\mathbf{W}_L^{*,\theta}-\mathbf{W}_L)-a(\mathbf{W}_R-\mathbf{W}_R^{*,\theta})$$

or equivalently

$$\mathbf{W}_R^{*,\theta}-\mathbf{W}_L^{*,\theta}=\mathbf{W}_R-\mathbf{W}_L+\frac{1}{a}\left(2F^{\theta}(\mathbf{W}_L,\mathbf{W}_R)-\mathbf{F}(\mathbf{W}_L)-\mathbf{F}(\mathbf{W}_R)\right). \tag{3.39}$$

Both (3.38) and (3.39) provide

$$\mathbf{W}_L^{*,\theta}=\mathbf{W}_L-\frac{1}{a}(F^{\theta}(\mathbf{W}_L,\mathbf{W}_R)-\mathbf{F}(\mathbf{W}_L)), \quad \mathbf{W}_R^{*,\theta}=\mathbf{W}_R+\frac{1}{a}(F^{\theta}(\mathbf{W}_L,\mathbf{W}_R)-\mathbf{F}(\mathbf{W}_R)).$$

This yields the desired results. □

Using this approximate Riemann solver, we can deduce that the modified acoustic solver (3.31) is stable under the same CFL conditions (3.17) that does not depend on the modification θ. Moreover, when $\theta = 1$ the self-similar function $\mathbf{W}_{RP}^{\theta}$ defined in proposition 4 degenerates to the exact solution of the Riemann problem associated with relaxed acoustic system (3.8).

Finally, if one takes into account the equilibrium projection step of the relaxation strategy into the approximate Riemann solver of proposition 4, we have $\Pi_L = p^{\mathrm{EOS}}(\tau_L, e_L)$, and $\Pi_R = p^{\mathrm{EOS}}(\tau_R, e_R)$. Under this assumption, it is easy to check that the first coordinates (τ, u, v, E) of the self similar function $\mathbf{W}_{\mathrm{RP}}^{\theta}$ are consistent with the integral form of the acoustic system (3.6).

3.5.3 Properties of the modified operator splitting scheme

We start this section by examining the ability of the modified operator splitting scheme to satisfies a discrete entropy inequality. In the sequel, $I(b, b') \subset \mathbb{R}$ will denote the interval whose bounds are $b \in \mathbb{R}$ and $b' \in \mathbb{R}$. We consider the following slightly more restrictive subcharacteristic condition

$$\begin{aligned} &\tau_L^* > 0, \quad -\partial_\tau p^{\mathrm{EOS}}(\tau, s_L) \leq a^2, \quad \forall \tau \in I(\tau_L, \tau_L^*), \\ &\tau_R^* > 0, \quad -\partial_\tau p^{\mathrm{EOS}}(\tau, s_R) \leq a^2, \quad \forall \tau \in I(\tau_R, \tau_R^*), \end{aligned} \tag{3.40}$$

and we start with the two following technical results. we also refer the reader to Annex B for a quick refresh on this topic.

Lemma 1. *Consider the solution of Riemann problem for the relaxed acoustic system* (3.8)*. Suppose that* (3.40) *is verified. Let* $s_k = s^{\mathrm{EOS}}(\tau_k, e_k)$*,* $k = L, R$*, we have*

$$e_k^* - e^{\mathrm{EOS}}(\tau_k^*, s_k) - \frac{\left(p^{\mathrm{EOS}}(\tau_k^*, s_k) - \Pi^*\right)^2}{2a^2} \geq 0. \tag{3.41}$$

Proof. We consider the case $k = R$ and set for $\tau \in I(\tau_R, \tau_R^*)$

$$\begin{aligned} \phi(\tau) = e^{\mathrm{EOS}}(\tau, s_R) - \frac{p^{\mathrm{EOS}}(\tau, s_R)^2}{2a^2} - e^{\mathrm{EOS}}(\tau_R^*, s_R) + \frac{p^{\mathrm{EOS}}(\tau_R^*, s_R)^2}{2a^2} \\ + p^{\mathrm{EOS}}(\tau_R^*, s_R)\left(\tau + \frac{p^{\mathrm{EOS}}(\tau, s_R)}{a^2} - \tau_R^* - \frac{p^{\mathrm{EOS}}(\tau_R^*, s_R)}{a^2}\right). \end{aligned}$$

We have $\phi'(\tau) = \left(p^{\mathrm{EOS}}(\tau, s_R) - p^{\mathrm{EOS}}(\tau_R^*, s_R)\right)\left(1 - \rho^2 c^2(\tau, s_R)/a^2\right)$. If $\tau_R > \tau > \tau_R^*$ (resp. $\tau_R < \tau < \tau_R^*$) the Weyl assumptions (3.2) provides $p^{\mathrm{EOS}}(\tau, s_R) - p^{\mathrm{EOS}}(\tau_R^*, s_R) < 0$ (resp. $p^{\mathrm{EOS}}(\tau, s_R) - p^{\mathrm{EOS}}(\tau_R^*, s_R) > 0$) and together with hypothesis (3.40) this yields $\phi'(\tau) \geq 0$ (resp. $\phi'(\tau) \leq 0$). As $\phi(\tau_R^*) = 0$ we obtain that $\phi(\tau_R) > \phi(\tau_R^*) = 0$ for $\tau \in I(\tau_R, \tau_R^*)$. Using the Riemann invariant jump relation $(e_R^* - \frac{\Pi^*}{2a^2}) = (e_R - \frac{\Pi_R}{2a^2})$, one obtains $0 < \phi(\tau_R) = e_R^* - e^{\mathrm{EOS}}(\tau_R^*, s_R) - \frac{1}{2a^2}(p^{\mathrm{EOS}}(\tau_R^*, s_R) - \Pi^*)^2$. The same lines applies for the case $k = L$. □

Lemma 2. *Let* $\theta \in \mathbb{R}$*, and* $e_k^{*,\theta} = E_k^{*,\theta} - (u_k^{*,\theta})^2/2$ *for* $k = L, R$ *then we have*

$$e_k^{*,\theta} - e^{\mathrm{EOS}}(\tau_k^{*,\theta}, s_k) - \frac{1}{2a^2}\left(p^{\mathrm{EOS}}(\tau_k^{*,\theta}, s_k) - \Pi^*\right)^2 + \frac{(1-\theta)^2(u_R - u_L)^2}{8} \geq 0, \quad k = L, R. \tag{3.42}$$

Proof. One has $u_R^{*,\theta} = u^* + (1-\theta)(u_R - u_L)/2$ and $\Pi^{*,\theta} = \Pi^* + (1-\theta)a(u_R - u_L)/2$ and together with (3.37) one obtains $e_R^{*,\theta} = e_R^* - (1-\theta)^2(u_R - u_L)^2/8$. Injecting this relation into (3.41) and noticing that $\tau_R^{*,\theta} = \tau_R^*$ provides the desired result for $k = R$. The case $k = L$ is obtained with the same lines. □

It is now clear that the inequalities

$$-\frac{1}{2a^2}\left(p^{\mathrm{EOS}}(\tau_k^{*,\theta},s_R)-\Pi^*\right)^2+\frac{(1-\theta)^2(u_R-u_L)^2}{8}\leq 0,\quad k=L,R \tag{3.43}$$

can help us equip the modified numerical scheme with a discrete entropy inequality.

Proposition 5. *Let $s_k^{*,\theta}=s^{\mathrm{EOS}}(\tau_k^{*,\theta},e_k^{*,\theta})$ for $k=L,R$. If assumption* (3.43) *is verified, we have*

$$0\leq -a(s_L^{*,\theta}-s_L)+a(s_R-s_R^{*,\theta}). \tag{3.44}$$

Inequality (3.44) *implies that the modified scheme* (3.31) *for the acoustic step is consistent with the integral form of the entropy inequality*

$$\partial_t s(\tau,e)\leq 0. \tag{3.45}$$

Moreover, the explicit modified scheme (3.31) *is equipped with a discrete entropy inequality. Indeed there exists a numerical flux function $q_{j+1/2}^n=q(\mathbf{W}_j^n,\mathbf{W}_{j+1}^n)$ that is consistent with 0 when Δt and Δx tend to 0 such that*

$$s(\tau_j^{n+1-},e_j^{n+1-})-s(\tau_j^n,e_j^n)+\tau_j^n\frac{\Delta t}{\Delta x}(q_{j+1/2}^n-q_{j-1/2}^n)\leq 0. \tag{3.46}$$

Proof. Let $k=L,R$, under hypothesis (3.43), we have that $e_k^{*,\theta}\geq e^{\mathrm{EOS}}(\tau_k^{*,\theta},s_k)$. According to (3.2) $\epsilon\mapsto s^{\mathrm{EOS}}(\tau_k^{*,\theta},\epsilon)$ is increasing, thus $s^{\mathrm{EOS}}(\tau_k^{*,\theta},e_k^{*,\theta})=s_k^{*,\theta}\geq s^{\mathrm{EOS}}(\tau_k^{*,\theta},e^{\mathrm{EOS}}(\tau_k^{*,\theta},s_k))=s_k$. Inequality (3.44) follows trivially. Relation (3.44) expresses the consistency with the integral form of (3.45) and it provides the entropy inequality (3.46) (see [1] and Annex B). □

We can now state the following entropic property for the full modified operator splitting explicit scheme composed by (3.31) and (3.14).

Proposition 6. *If the assumptions* (3.43), (3.17) *and* (3.18) *are verified, then the explicit scheme defined by* (3.31) *and* (3.14) *verifies the following discrete entropy inequality*

$$\rho_j^{n+1}s(\tau_j^{n+1},e_j^{n+1})-\rho_j^n s(\tau_j^n,e_j^n)+\frac{\Delta t}{\Delta x}\left(g_{j+1/2}^n-g_{j-1/2}^n\right)\leq 0, \tag{3.47}$$

where the numerical entropy flux is defined by

$$g_{j+1/2}^n=(u_{j+1/2}^*)^+\rho_j^{n+1-}s(\tau_j^{n+1-},e_j^{n+1-})+(u_{j+1/2}^*)^-\rho_j^{n+1-}s(\tau_{j+1}^{n+1-},e_{j+1}^{n+1-})+q_{j+1/2}^n. \tag{3.48}$$

Proof. Let $\phi\in(\rho,\rho u,\rho v,\rho E)$, under the CFL assumption (3.18) the transport scheme (3.14) expresses ϕ_j^{n+1} as a convex combination of ϕ_i^{n+1-}, $i=j-1,j,j+1$, indeed one has

$$\phi_j^{n+1}=\frac{\Delta t}{\Delta x}(u_{j-1/2}^*)^+\phi_{j-1}^{n+1-}+\left(1-\frac{\Delta t}{\Delta x}((u_{j+1/2}^*)^--(u_{j-1/2}^*)^+)\right)\phi_j^{n+1-}+\frac{\Delta t}{\Delta x}(u_{j+1/2}^*)^-\phi_{j+1}^{n+1-}.$$

As the mapping $(\rho,\rho u,\rho v,\rho E)\mapsto -(\rho s)(\tau,e)$ is a strictly convex function (see for example [16]) we obtain that

$$\begin{aligned}-(\rho s)(\tau_j^{n+1},e_j^{n+1})\leq &-\frac{\Delta t}{\Delta x}(u_{j-1/2}^*)^+(\rho s)(\tau_{j-1}^{n+1-},e_{j-1}^{n+1-})-\frac{\Delta t}{\Delta x}(u_{j+1/2}^*)^-(\rho s)(\tau_{j-1}^{n+1-},e_{j+1}^{n+1-})\\&-\left(1-\frac{\Delta t}{\Delta x}((u_{j+1/2}^*)^--(u_{j-1/2}^*)^+)\right)(\rho s)(\tau_j^{n+1-},e_j^{n+1-}).\end{aligned}$$

Using relation (3.46) one obtains (3.47). □

We now sum up the main properties of the modified operator splitting scheme.

Theorem 3. *Suppose that* (3.17), (3.18) (3.12) *are satisfied, the explicit scheme defined by* (3.31) *and* (3.14) *verifies*

1. *the scheme is conservative with respect to the density ρ, the momentum ρu and total energy ρE,*
2. *the density ρ_j^n is positive for all j and $n > 0$ provided that ρ_j^0 is positive for all j,*
3. *if $\theta = \mathcal{O}(M)$, then the truncation error of the numerical scheme is uniform with respect to $M < 1$,*
4. *if* (3.43) *is verified then the numerical scheme is equipped with a discrete entropy inequality,*
5. *if* (3.43) *is verified then $e_j^n > 0$ for all $j \in \mathbb{Z}$ and all $n \in \mathbb{N}$.*

It is clear from (3.32b) that the choice $\theta = O(M)$ is natural for the modified scheme to have an equivalent equation which is satisfactory when $M \ll \Delta\tilde{x}$ (uniform consistency w.r.t. M). At this stage $\theta = O(M)$ is not made precise, see section 3.6 below. Let us now discuss the new condition which is related to the correction θ.

Behavior of condition (3.43) **in the low-Mach regime for a perfect gas equation of state**

We have just seen that the scheme is entropic provided that (3.43) is satisfied. In this section, we study the compatibility in the low Mach regime between the condition (3.43) that is required to obtain a discrete entropy inequality and the condition $\theta = O(M)$ that is required to have uniform consistency with respect to M (see section 3.5). If $|u_R - u_L| = 0$, any value of $\theta \in \mathbb{R}$ verifies condition (3.43), we can then assume that $|u_R - u_L| > 0$. We consider the case of a Perfect Gas EOS defined by $p^{\mathrm{EOS}}(\rho, e) = (\gamma - 1)\rho e$, where γ is the specific heat ratio. First, let us recall that $\tau_k^{*,\theta} = \tau_k^*$ and $\Pi_k = p^{\mathrm{EOS}}(\tau_k, s_k)$, $k = R, L$. For $k = R$, relation (3.43) reads

$$|1 - \theta| \leq \frac{2}{a} \frac{|p^{\mathrm{EOS}}(\tau_R^*, s_R) - \Pi^*|}{|u_R - u_L|}. \tag{3.49}$$

Let us remark that the right hand side of this inequality does not depend on θ. The Perfect Gas assumption provides that $p^{\mathrm{EOS}}(\tau_R^*, s_R) = \Pi_R \, (\tau_R/\tau_R^*)^\gamma$, therefore thanks to the definition of Π^* we get

$$p^{\mathrm{EOS}}(\tau_R^*, s_R) - \Pi^* = \Pi_R \, (\tau_R/\tau_R^*)^\gamma - \frac{\Pi_L + \Pi_R}{2} + \frac{a}{2}(u_R - u_L). \tag{3.50}$$

The definition of $\tau_R^{*,\theta} = \tau_R^*$ by (3.37) using the dimensionless parameters defined by (3.19) gives $\tilde{\tau}_R^* = \tilde{\tau}_R + (\tilde{\Pi}_R - \tilde{\Pi}_L)/(2\tilde{a}^2) + M(\tilde{u}_R - \tilde{u}_L)/(2\tilde{a})$. If one now supposes that the flow is locally in the low Mach regime, then we have $\partial_{\tilde{x}}\tilde{\Pi} = O(M^2)$, therefore $\Pi_R - \Pi_L = O(M^2\Delta\tilde{x})$. Thus we obtain

$$\frac{\tilde{\tau}_R^*}{\tilde{\tau}_R} = 1 + M\frac{\tilde{u}_R - \tilde{u}_L}{2\tilde{a}\tilde{\tau}_R} + O(M^2\Delta\tilde{x}).$$

Injecting the above relation into (3.50), we obtain

$$\frac{p^{\mathrm{EOS}}(\tau_R^*, s_R) - \Pi^*}{p_0} = -\frac{M}{2}\left[1 - \frac{\gamma\tilde{\Pi}_R}{\tilde{a}^2\tilde{\tau}_R}\right]\tilde{a}(\tilde{u}_R - \tilde{u}_L) + O(M^2\Delta\tilde{x}).$$

Using the fact that $\gamma p^{\mathrm{EOS}}(\tau_R, s_R) = \gamma\Pi_R = \rho_R(c_R)^2$ for a Perfect Gas in the previous relation allows to recast (3.49) into

$$|1 - \theta| \leq \left|1 - \left(\frac{\tilde{\rho}_R\tilde{c}_R}{\tilde{a}}\right)^2 + O\left(\frac{M\Delta\tilde{x}}{|\tilde{u}_R - \tilde{u}_L|}\right)\right|. \tag{3.51}$$

Let us recall that by definition : $\tilde{a} = K\max(\tilde{\rho}_R\tilde{c}_R, \tilde{\rho}_L\tilde{c}_L)$ with $K \geq 1$. Suppose without loss of

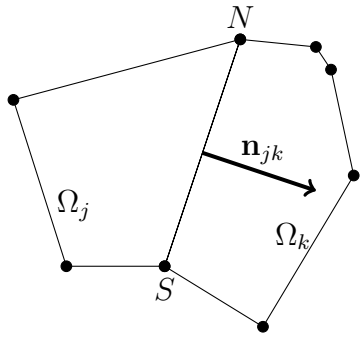


FIGURE 3.1 – the face $\Gamma_{jk} = \overline{\Omega_j} \cap \overline{\Omega_k}$ defined the segment (NS) has a unit normal vector $\mathbf{n}_{jk}$ oriented from Ω_j to Ω_k.

generality that $\tilde{\rho}_R \tilde{c}_R = \max(\tilde{\rho}_R \tilde{c}_R, \tilde{\rho}_L \tilde{c}_L)$ then $(\tilde{\rho}_R \tilde{c}_R)^2/\tilde{a}^2 = 1/K$ and the condition (3.51) becomes

$$|1-\theta| \leq \left| 1 - \left(\frac{1}{K}\right)^2 + O\left(\frac{M\Delta\tilde{x}}{|\tilde{u}_R - \tilde{u}_L|}\right) \right|.$$

Behavior when $M \to 0$. When $M \to 0$ the above inequality yields that $\theta \geq (1/K)^2$ if one wants to enforce uniform consistency with respect to M by setting $\theta = O(M)$. This leads to a contradiction. As a conclusion, a correction scheme with $\theta = O(M)$ does not provide an entropic scheme in the asymptotic limit $M \to 0$. On the contrary, $\theta = 1$ which correspond to the classic unmodified scheme is still entropic. Nevertheless, it is reasonable to consider that in the limit $M \to 0$, the solution of the gas dynamics equation is smooth and therefore the consistency with an entropy criterion is a less critical matter.

3.5.4 Extension to several space dimensions with unstructured grids

Without loss of generality, we suppose that $\Omega \subset \mathbb{R}^2$ is a polygonal domain that is covered by a set of N polygonal cells $(\Omega_j)_{1\leq j\leq N}$. Let Γ be a face of a cell Ω_j, $1 \leq j \leq N$. If $\Gamma \subset \partial\Omega$, we suppose that there exists a single $k > N$ that will help to index ghost values for boundary conditions and we shall note $\Gamma = \Gamma_{jk}$. If $\Gamma \cap \partial\Omega = \emptyset$, we suppose that the mesh is admissible in the sense that there exists a single $1 \leq k \leq N$ such that $\Gamma = \overline{\Omega_j} \cap \overline{\Omega_k}$. Moreover, for $1 \leq j \leq N$ and $1 \leq k \leq N$ we suppose that $\overline{\Omega_i} \cap \overline{\Omega_j}$ can either be empty, a vertex or a single face of the mesh. If Γ_{jk} be the face of a cell Ω_j then $\mathbf{n}_{jk}$ will denote the unit vector normal to Γ_{jk} pointing out of Ω_j. We define $N(j)$ the set of indices k such that Γ_{jk} is a face of Ω_j. Let $E = \{(j,k) \mid 1 \leq j,k \leq N,\ k \in N(j)\}$ and $E^{\text{ext}} = \{(j,k) \mid 1 \leq j \leq N,\ k \in N(j), \Gamma_{jk} \subset \partial\Omega\}$. In sequel $x = (x_1, x_2) \in \mathbb{R}^2$ will denote the space variable.

We will now present a natural extension of our discretization strategy for the case of multi-dimensional problems with unstructured grids. Within this framework, the classical Lagrange-Remap algorithm involves tracking a genuine multi-dimensional moving mesh. This task is a very delicate matter as the mesh may be dramatically distorted during the simulation. We will here present a much simpler approach that relies on the alternative guideline proposed in section 3.3. A similar approach was used to derive an explicit scheme for two-component interface problems in [14].

Consider the operator splitting of (3.1) into the following systems

$$\begin{cases} \partial_t \rho + \rho \mathrm{div}(\mathbf{u}) = 0, & (3.52a) \\ \partial_t(\rho\mathbf{u}) + \rho\mathbf{u}\mathrm{div}(\mathbf{u}) + \nabla p = 0, & (3.52b) \\ \partial_t(\rho E) + \rho E \mathrm{div}(\mathbf{u}) + \mathrm{div}(P\mathbf{u}) = 0, & (3.52c) \end{cases}$$

and

$$\begin{cases} \partial_t \rho + (\mathbf{u}\cdot\nabla)\rho = 0, & (3.53a) \\ \partial_t(\rho\mathbf{u}) + (\mathbf{u}\cdot\nabla)\rho\mathbf{u} = 0, & (3.53b) \\ \partial_t(\rho E) + (\mathbf{u}\cdot\nabla)\rho E = 0. & (3.53c) \end{cases}$$

Before going any further, let us note that we obtain similar properties as for the systems (3.4) and (3.5). Indeed system (3.52) is a quasilinear hyperbolic system that involves the two nonlinear acoustic waves of velocity $\pm c$ and two null velocity contact discontinuities waves. System (3.52) only involves acoustic phenomena while freezing the material transport, while (3.53) is pure multi-dimensional transport system at the material velocity $\mathbf{u}$.

We adopt the same strategy as in section 3.3 : given a fluid state $(\rho, \rho\mathbf{u}, \rho E)_j^n$,
– update the fluid state to the value $(\rho, \rho\mathbf{u}, \rho E)_j^{n+1-}$ by approximating the solution of (3.52),
– update the fluid state to the value $(\rho, \rho\mathbf{u}, \rho E)_j^{n+1}$ by approximating the solution of (3.53).

Approximation of the acoustic system (3.52)
System (3.52) can be expressed

$$\partial_t \tau - \tau(x,t)\mathrm{div}(\mathbf{u}) = 0, \quad \partial_t \mathbf{u} + \tau(x,t)\nabla p = 0, \quad E_t + \tau(x,t)\mathrm{div}(p\mathbf{u}) = 0.$$

Using the same lines as in section 3.3 we consider a Suliciu-type relaxation approximation

$$\begin{aligned} &\partial_t \tau - \tau(x,t)\mathrm{div}(\mathbf{u}) = 0, && \partial_t \mathbf{u} + \tau(x,t)\nabla\Pi = 0, \\ &E_t + \tau(x,t)\mathrm{div}(p\mathbf{u}) = 0, && \Pi_t + \tau(x,t)a^2\mathrm{div}(\mathbf{u}) = \nu(p-\Pi), \end{aligned}$$

in the regime $\nu \to +\infty$. Once again, for $t \in [t^n, t^n + \Delta t)$, this task is achieved by setting $\Pi(x,t^n) = p(x,t^n)$ and then solving the relaxation system for $\nu = 0$. We approximate again $\tau(x,t)\partial_{x_r}$ by $\tau(x,t^n)\partial_{x_r}$ for $r = 1, 2$ when $t \in [t^n, t^n + \Delta t)$. In the regime $\lambda = 0$ our approximation of (3.52) becomes

$$\begin{aligned} &\partial_t \tau - \tau(x,t^n)\mathrm{div}(\mathbf{u}) = 0, && \partial_t \mathbf{u} + \tau(x,t^n)\nabla\Pi &&= 0, \\ &E_t + \tau(x,t^n)\mathrm{div}(p\mathbf{u}) = 0, && \Pi_t + \tau(x,t^n)a^2\mathrm{div}(\mathbf{u}) &&= 0. \end{aligned} \tag{3.54}$$

If b is a flow parameter and b_j^n is an approximation of $\frac{1}{|\Omega_j|}\int_{\Omega_j} b(x,t^n)\,\mathrm{d}x$, we solve (3.54) thanks to the following classical Finite-Volume method

$$\begin{cases} \mathbf{u}_j^{n+1-} = \mathbf{u}_j^n - \tau_j^n \Delta t \sum_{k\in N(j)} \sigma_{jk}\, \Pi_{jk}^{*,\theta}\, \mathbf{n}_{jk}, & (3.55a) \\ \Pi_j^{n+1-} = \Pi_j^n - \tau_j^n \Delta t \sum_{k\in N(j)} \sigma_{jk}\, (a_{jk})^2 u_{jk}^*, & (3.55b) \\ \tau_j^{n+1-} = \tau_j^n + \tau_j^n \Delta t \sum_{k\in N(j)} \sigma_{jk}\, u_{jk}^*, & (3.55c) \\ E_j^{n+1-} = E_j^n - \tau_j^n \Delta t \sum_{k\in N(j)} \sigma_{jk}\, \Pi_{jk}^{*,\theta}\, u_{jk}^*, & (3.55d) \end{cases}$$

where $\sigma_{jk} = |\Gamma_{jk}|/|\Omega_j|$.

The three scalar quantities a_{jk}, $\Pi_{jk}^{*,\theta}$ and u_{jk}^* that respectively represent an average sound velocity, a pressure and the normal velocity at the face Γ_{jk}. In order to define these quantities, we classicaly take advantage of the fact that (3.54) is rotational invariant. This allows to associate in the referential of each face Γ_{jk} a Suliciu-type relaxation approximation of a one-dimensional Riemann problem in the frame of

the face. Noting $\sharp \in \{n, n+1-\}$, this leads us to set

$$a_{jk} \geq \max[(\rho c)_j^n, (\rho c)_k^n], \tag{3.56a}$$

$$u_{jk}^* = \frac{1}{2}\mathbf{n}_{jk}^T(\mathbf{u}_j^\sharp + \mathbf{u}_k^\sharp) - \frac{1}{2a_{jk}}(\Pi_k^\sharp - \Pi_j^\sharp), \tag{3.56b}$$

$$\Pi_{jk}^{*,\theta} = \frac{1}{2}(\Pi_j^\sharp + \Pi_k^\sharp) - \frac{a_{jk}\theta_{jk}}{2}\mathbf{n}_{jk}^T(\mathbf{u}_k^\sharp - \mathbf{u}_j^\sharp). \tag{3.56c}$$

When $\sharp = n$ the solver is explicit and when $\sharp = n+1-$, the solver is implicit.

Approximation of the transport system (3.53)
In order to approximate the solution of (3.53), we simply use an upwind Finite-Volume scheme. Let $\varphi \in \{\rho, \rho u_1, \rho u_2, \rho E\}$, we set

$$\varphi_j^{n+1} = \varphi_j^{n+1-} - \Delta t \sum_{k\in N(j)} \sigma_{jk}\, u_{jk}^* \varphi_{jk}^{n+1-} + \Delta t \varphi_j^{n+1-} \sum_{k\in N(j)} \sigma_{jk}\, u_{jk}^*, \tag{3.57}$$

where φ_{jk}^{n+1-} is defined by the upwind choice with respect to the sign of u_{jk}^*, namely

$$\varphi_{jk}^{n+1-} = \begin{cases} \varphi_j^{n+1-}, & \text{if } u_{jk}^* > 0, \\ \varphi_k^{n+1-}, & \text{if } u_{jk}^* \leq 0. \end{cases}$$

Proposition 7. *The overall numerical scheme composed by the discretization steps* (3.55a)-(3.55d) *and* (3.57) *is conservative with respect to the variable ρ, $\rho\mathbf{u}$ and ρE, for both the implicit solver and the explicit solver. The update of these variables from t^n to t^{n+1} reads*

$$\rho_j^{n+1} - \rho_j^n + \Delta t \sum_{k\in N(j)} \sigma_{jk}\, \rho_{jk}^{n+1-} u_{jk}^* = 0, \tag{3.58a}$$

$$(\rho\mathbf{u})_j^{n+1} - (\rho\mathbf{u})_j^n + \Delta t \sum_{k\in N(j)} \sigma_{jk} \left((\rho\mathbf{u})_{jk}^{n+1-} u_{jk}^* + \Pi_{jk}^{*,\theta}\mathbf{n}_{jk} \right) = 0, \tag{3.58b}$$

$$(\rho E)_j^{n+1} - (\rho E)_j^n + \Delta t \sum_{k\in N(j)} \sigma_{jk} \left((\rho E)_{jk}^{n+1-} + \Pi_{jk}^{*,\theta} \right) u_{jk}^* = 0. \tag{3.58c}$$

The semi-implicit solver obtained for $\sharp = n+1-$ can be decomposed along the following steps : the acoustic step first involves solving the linear system (3.55a)-(3.55b) for computing the acoustic velocity $\mathbf{u}_j^{n+1-}$ and pressure term Π_j^{n+1-}. The acoustic step is completed by the update of τ_j^{n+1-} and E_j^{n+1-} thanks to the explicit procedures (3.55c) and (3.55d). The last stage of the semi-implicit solver is achieved thanks to the explicit transport scheme (3.57).

We want now to investigate further the implicit system involved with the semi-implicit method for the specific case of wall-boundary conditions that we implement by imposing ghost values Π_k^{n+1-} and $\mathbf{n}_{jk}^T\mathbf{u}_k^{n+1-}$ for a boundary face $\Gamma_{jk} \subset \partial\Omega$, where $1 \leq j \leq N$ and $k \in N(j)$, $k > N$ with

$$\Pi_k^{n+1-} = \Pi_j^{n+1-}, \quad \mathbf{n}_{jk}^T\mathbf{u}_k^{n+1-} = -\mathbf{n}_{jk}^T\mathbf{u}_j^{n+1-}. \tag{3.59}$$

We have the following proposition.

Proposition 8. *We consider the case of the semi-implicit solver with implementation of wall boundary conditions* (3.59) *and a uniform choice of a, i.e. $a_{jk} = a$ for all $1 \leq j \leq N$ and $k \in N(j)$. If $\tau_j^n > 0$ for all $1 \leq j \leq N$, then the linear system* (3.55a)-(3.55b) *always possesses a single solution for any $\Delta t > 0$ and $\theta_{jk} > 0$.*

Proof. For the sake of readability, we shall note here $\mathbf{u}_j^{n+1-} = \mathbf{u}_j$ and $\Pi_j^{n+1-} = \Pi_j$. The finite-dimension linear system (3.55a)-(3.55b) reads

$$\begin{cases} |\Omega_j|\mathbf{u}_j + \tau_j^n \Delta t \sum\limits_{k\in N(j)} |\Gamma_{jk}| \left[\dfrac{1}{2}(\Pi_j + \Pi_k) - \dfrac{a\theta_{jk}}{2}\mathbf{n}_{jk}^T(\mathbf{u}_k - \mathbf{u}_j)\right]\mathbf{n}_{jk} = |\Omega_j|\mathbf{u}_j^n, & \text{(3.60a)} \\ |\Omega_j|\Pi_j + \tau_j^n \Delta t \sum\limits_{k\in N(j)} |\Gamma_{jk}| a^2 \left[\dfrac{1}{2}\mathbf{n}_{jk}^T(\mathbf{u}_j + \mathbf{u}_k) - \dfrac{1}{2a}(\Pi_k - \Pi_j)\right] = |\Omega_j|\Pi_j^n. & \text{(3.60b)} \end{cases}$$

This system admits a unique solution if and only if $\mathbf{u}_j = 0$, $\Pi_j = 0$, $1 \le j \le N$ is the only solution of the particular case obtained for $\mathbf{u}_j^n = 0$, $\Pi_j^n = 0$, $1 \le j \le N$. Thus, let us now suppose that the right members of (3.60) are null, we proceed using an energy estimate type proof. Let us multiply (3.60a) by $\dfrac{2\mathbf{u}_j^T}{\tau_j^n \Delta t}$ and sum over j, we obtain

$$\begin{aligned} 0 = \sum_{j=1}^N \frac{2|\Omega_j||\mathbf{u}_j|^2}{\tau_j^n \Delta t} + \sum_{j=1}^N \sum_{k\in N(j)} |\Gamma_{jk}|(\Pi_j + \Pi_k)(\mathbf{u}_j^T \mathbf{n}_{jk}) \\ - \sum_{j=1}^N \sum_{k\in N(j)} |\Gamma_{jk}| a\theta_{jk}(\mathbf{u}_j^T \mathbf{n}_{jk})(\mathbf{u}_k - \mathbf{u}_j)^T \mathbf{n}_{jk}. \end{aligned} \tag{3.61}$$

Accounting for the fact that $\sum\limits_{k\in N(j)} |\Gamma_{jk}|\mathbf{n}_{jk} = \mathbf{0}$, the second term of (3.61) verifies

$$\sum_{j=1}^N \sum_{k\in N(j)} |\Gamma_{jk}|(\Pi_j + \Pi_k)(\mathbf{u}_j^T \mathbf{n}_{jk}) = \sum_{j=1}^N \sum_{k\in N(j)} |\Gamma_{jk}|\Pi_k \mathbf{u}_j^T \mathbf{n}_{jk}.$$

Using boundary conditions (3.59), the third term of (3.61) reads

$$\begin{aligned} &\sum_{j=1}^N \sum_{k\in N(j)} |\Gamma_{jk}| a\theta_{jk}(\mathbf{u}_j^T \mathbf{n}_{jk})(\mathbf{u}_k - \mathbf{u}_j)^T \mathbf{n}_{jk} \\ &\quad = \sum_{(j,k)\in E} |\Gamma_{jk}| a\theta_{jk}\Big[(\mathbf{u}_j^T \mathbf{n}_{jk})(\mathbf{u}_k - \mathbf{u}_j)^T \mathbf{n}_{jk} + (\mathbf{u}_k^T \mathbf{n}_{kj})(\mathbf{u}_j - \mathbf{u}_k)^T \mathbf{n}_{kj}\Big] \\ &\qquad + \sum_{(j,k)\in E^{\text{ext}}} |\Gamma_{jk}| a\theta_{jk}(\mathbf{u}_j^T \mathbf{n}_{jk})(\mathbf{u}_k - \mathbf{u}_j)^T \mathbf{n}_{jk} \\ &\quad = -\sum_{(j,k)\in E} |\Gamma_{jk}| a\theta_{jk}\Big[(\mathbf{u}_k - \mathbf{u}_j)^T \mathbf{n}_{jk}\Big]^2 - 2\sum_{(j,k)\in E^{\text{ext}}} |\Gamma_{jk}| a\theta_{jk}[(\mathbf{u}_j^T \mathbf{n}_{jk})]^2. \end{aligned}$$

Finally we see that (3.61) is equivalent to

$$\begin{aligned} 0 = \sum_{j=1}^N \frac{2|\Omega_j||\mathbf{u}_j|^2}{\tau_j^n \Delta t} + \sum_{j=1}^N \sum_{k\in N(j)} |\Gamma_{jk}|\Pi_k \mathbf{u}_j^T \mathbf{n}_{jk} + \sum_{(j,k)\in E} |\Gamma_{jk}| a\theta_{jk}\Big[(\mathbf{u}_k - \mathbf{u}_j)^T \mathbf{n}_{jk}\Big]^2 \\ + 2\sum_{(j,k)\in E^{\text{ext}}} |\Gamma_{jk}| a\theta_{jk}[(\mathbf{u}_j^T \mathbf{n}_{jk})]^2. \end{aligned} \tag{3.62}$$

Let us turn to the pressure equation (3.60b), we multiply by $\dfrac{2\Pi_j}{\tau_j^n a^2 \Delta t}$ and sum over all $1 \le j \le N$, this

yields

$$0 = \sum_{j=1}^{N} \frac{2|\Omega_j|\Pi_j^2}{\tau_j^n a^2 \Delta t} + \sum_{j=1}^{N} \sum_{k \in N(j)} |\Gamma_{jk}| \mathbf{n}_{jk}^T (\mathbf{u}_j + \mathbf{u}_k) \Pi_j - \sum_{j=1}^{N} \sum_{k \in N(j)} \frac{1}{a} |\Gamma_{jk}| (\Pi_k - \Pi_j) \Pi_j. \tag{3.63}$$

Using once again $\sum_{k \in N(j)} |\Gamma_{jk}| \mathbf{n}_{jk} = \mathbf{0}$, we have for the second term of (3.63) that

$$\sum_{j=1}^{N} \sum_{k \in N(j)} |\Gamma_{jk}| \mathbf{n}_{jk}^T (\mathbf{u}_j + \mathbf{u}_k) \Pi_j = \sum_{j=1}^{N} \sum_{k \in N(j)} |\Gamma_{jk}| \mathbf{n}_{jk}^T \mathbf{u}_k \Pi_j.$$

Accounting for (3.59), the third term of (3.63) verifies

$$\begin{aligned} \sum_{j=1}^{N} \sum_{k \in N(j)} \frac{|\Gamma_{jk}|}{a} (\Pi_k - \Pi_j) \Pi_j &= \frac{1}{a} \sum_{(j,k) \in E} |\Gamma_{jk}| \Big[(\Pi_k - \Pi_j) \Pi_j - (\Pi_j - \Pi_k) \Pi_k \Big] \\ &\quad + \frac{1}{a} \sum_{(j,k) \in E^{\text{ext}}} |\Gamma_{jk}| (\Pi_k - \Pi_j) \Pi_j \\ &= -\frac{1}{a} \sum_{(j,k) \in E} |\Gamma_{jk}| (\Pi_k - \Pi_j)^2. \end{aligned}$$

Then, we see that (3.63) also reads

$$0 = \sum_{j=1}^{N} \frac{2|\Omega_j|\Pi_j^2}{\tau_j^n a^2 \Delta t} + \sum_{j=1}^{N} \sum_{k \in N(j)} |\Gamma_{jk}| \mathbf{n}_{jk}^T \mathbf{u}_k \Pi_j + \frac{1}{a} \sum_{(j,k) \in E} |\Gamma_{jk}| (\Pi_k - \Pi_j)^2. \tag{3.64}$$

We now remark that

$$\begin{aligned} \sum_{j=1}^{N} \sum_{k \in N(j)} |\Gamma_{jk}| \mathbf{n}_{jk}^T \mathbf{u}_k \Pi_j &= \sum_{(j,k) \in E} |\Gamma_{jk}| (\mathbf{n}_{jk}^T \mathbf{u}_k \Pi_j + \mathbf{n}_{kj}^T \mathbf{u}_j \Pi_k) + \sum_{(j,k) \in E^{\text{ext}}} |\Gamma_{jk}| \mathbf{n}_{jk}^T \mathbf{u}_k \Pi_j \\ &= \sum_{(j,k) \in E} |\Gamma_{jk}| \mathbf{n}_{jk}^T (\mathbf{u}_k \Pi_j - \mathbf{u}_j \Pi_k) - \sum_{(j,k) \in E^{\text{ext}}} |\Gamma_{jk}| \mathbf{n}_{jk}^T \mathbf{u}_j \Pi_j, \end{aligned}$$

and also that

$$\begin{aligned} \sum_{j=1}^{N} \sum_{k \in N(j)} |\Gamma_{jk}| \Pi_k \mathbf{u}_j^T \mathbf{n}_{jk} &= \sum_{(j,k) \in E} |\Gamma_{jk}| (\Pi_k \mathbf{u}_j^T \mathbf{n}_{jk} + \Pi_j \mathbf{u}_k^T \mathbf{n}_{kj}) + \sum_{(j,k) \in E^{\text{ext}}} |\Gamma_{jk}| \Pi_k \mathbf{u}_j^T \mathbf{n}_{jk} \\ &= \sum_{(j,k) \in E} |\Gamma_{jk}| \mathbf{n}_{jk}^T (\Pi_k \mathbf{u}_j - \Pi_j \mathbf{u}_k) + \sum_{(j,k) \in E^{\text{ext}}} |\Gamma_{jk}| \Pi_j \mathbf{n}_{jk}^T \mathbf{u}_j. \end{aligned}$$

Therefore

$$\sum_{j=1}^{N} \sum_{k \in N(j)} |\Gamma_{jk}| \Pi_j \mathbf{u}_k^T \mathbf{n}_{jk} + \sum_{j=1}^{N} \sum_{k \in N(j)} |\Gamma_{jk}| \Pi_k \mathbf{u}_j^T \mathbf{n}_{jk} = 0.$$

Thus, summing (3.62) and (3.64), we obtain

$$\begin{aligned} 0 = &\sum_{j=1}^{N} \frac{2|\Omega_j|}{\tau_j^n \Delta t} \left(|\mathbf{u}_j|^2 + \frac{\Pi_j^2}{a^2} \right) + \sum_{(j,k) \in E} |\Gamma_{jk}| \left\{ a \theta_{jk} \Big[(\mathbf{u}_k - \mathbf{u}_j)^T \mathbf{n}_{jk} \Big]^2 + \frac{(\Pi_k - \Pi_j)^2}{a} \right\} \\ &+ 2 \sum_{(j,k) \in E^{\text{ext}}} |\Gamma_{jk}| \left\{ a \theta_{jk} [(\mathbf{u}_j^T \mathbf{n}_{jk})]^2 \right\}. \end{aligned}$$

This implies that $|\mathbf{u}_j| = \Pi_j = 0$ for all $1 \leq j \leq N$.

□

Remark 4. *It is possible to derive a similar proof for the case of periodic boundary conditions.*

We now examine the stability of the multi-dimensional operator splitting strategy (3.55), (3.56) and (3.57). The acoustic step (3.55) in the explicit cases $\sharp = n$ is stable under the CFL condition

$$\Delta t \max_{1\leq j\leq N} \left[\tau_j^n \left(\max_{k\in N(j)} \sigma_{jk} a_{jk}\right)\right] \leq \frac{1}{2}. \tag{3.65}$$

For both the explicit scheme $\sharp = n$ and semi-implicit scheme $\sharp = n+1-$, the transport step (3.57) is stable under the CFL condition

$$\Delta t \max_{1\leq j\leq N} \left(\sum_{k\in N(j)} \left|\sigma_{jk}(\mathbf{n}_{jk}^T \mathbf{u}_{jk}^{*,\theta})\right|\right) \leq 1. \tag{3.66}$$

When one uses the semi-implicit scheme $\sharp = n+1-$, the condition (3.66) becomes implicit as the computation of $\mathbf{u}_{jk}^{*,\theta}$ depends on a given Δt. In our simulations with the semi-implicit scheme, we chose to compute Δt thanks to the CFL condition (3.66) with the value $\mathbf{u}_{jk}^{*,\theta}$ given by the fully explicit scheme $\sharp = n$. It is then possible to check *a posteriori* that this Δt value matches (3.66).

We gather thereafter the properties of the explicit and semi-implicit multi-dimensional schemes.

Theorem 4. *Suppose that* (3.65), (3.66) *and* (3.12) *are satisfied. The explicit scheme defined by* (3.55) *and* (3.57) *with* $\sharp = n$ *verifies*

1. *the scheme is conservative with respect to the density* ρ*, the momentum* ρu *and total energy* ρE*,*
2. *the density* ρ_j^n *is positive for all* j *and* $n > 0$ *provided that* ρ_j^0 *is positive for all* j*,*
3. *if* $\theta = O(M)$*, then the truncation error of the numerical scheme is uniform with respect to* $M < 1$*,*
4. *if* (3.43) *is verified then the numerical scheme is equipped with a discrete entropy inequality,*
5. *if* (3.43) *is verified then* $e_j^n > 0$ *for all* $j \in \mathbb{Z}$ *and all* $n \in \mathbb{N}$*.*

Theorem 5. *Suppose that* (3.66) *and* (3.12) *are satisfied. The semi-implicit scheme defined by* (3.55) *and* (3.57) *with* $\sharp = n+1-$ *verifies*

1. *the scheme is conservative with respect to the density* ρ*, the momentum* ρu *and total energy* ρE*,*
2. *the density* ρ_j^n *is positive for all* j *and* $n > 0$ *provided that* ρ_j^0 *is positive for all* j*,*
3. *if* $\theta = O(M)$*, then the truncation error of the numerical scheme is uniform with respect to* $M < 1$*.*

Let us note that the implicit treatment of the acoustic step leads to a CFL restriction (3.66) based only on (slow) material waves.

3.6 Numerical results

In this section, we present numerical results computed thanks to the general operator splitting strategy (3.55), (3.56) and (3.57) with the following schemes :

- EX($\theta = 1$) : the explicit operator splitting scheme obtained for $\theta_{jk} = 1$ and $\sharp = n$,
- EX($\theta = O(M)$) : the explicit modified operator splitting scheme obtained with the low Mach correction $\theta_{jk} = \min\left(|u_{jk}^*|/\max(c_j^n, c_k^n), 1\right)$ and $\sharp = n$,
- EX($\theta = 0$) : the explicit modified operator splitting scheme with centered pressure gradient $\theta_{jk} = 0$ and $\sharp = n$,

– IMEX($\theta = 1$) : the semi-implicit operator splitting scheme with $\theta_{jk} = 1$ and $\sharp = n + 1-$,
– IMEX($\theta = O(M)$) : the modified semi-implicit operator splitting scheme with $\sharp = n + 1-$ and a low Mach correction θ_{jk} defined as in the case of EX($\theta = O(M)$),
– IMEX($\theta = 0$) : the modified semi-implicit operator splitting scheme with a centered pressure gradient $\theta_{ij} = 0$ and $\sharp = n + 1-$.

Remark 5. *The choice of the modification* $\theta_{jk} = \min\left(\frac{|u^*_{jk}|}{\max(c^n_j, c^n_k)}, 1\right)$ *corresponds to a low Mach correction. Indeed, this choice is non-dimensional, in* $(0, 1)$*, such that* $\theta = O(M)$ *in the low Mach regime and* $\theta = 1$ *for large Mach numbers. In this latter case, we then recover the classical scheme without modification.*

In the sequel, we shall consider that the fluid follows a perfect gas equation of state $p = (\gamma - 1)\rho e$ with a specific heat ratio $\gamma = 1.4$. We will test schemes on both low Mach and order 1 Mach number test cases.

3.6.1 Low Mach number examples

In this section we will consider low Mach tests and try to examine two questions : the accuracy gain for simulations on coarse grid in the low Mach regime thanks to the proposed correction, then the benefit of using a semi-implicit strategy in term of CPU time.

Vortex in a Box

We consider a test performed in [5]. The computational domain is $\Omega = [0, 1]^2$ with an initial condition given by

$$\rho_0(x_1, x_2) = 1 - \frac{1}{2}\tanh\left(x_2 - \frac{1}{2}\right), \qquad u_0(x_1, x_2) = 2\sin^2(\pi x_1)\sin(\pi x_2)\cos(\pi x_2),$$
$$p_0(x_1, x_2) = 1000, \qquad v_0(x_1, x_2) = -2\sin(\pi x_1)\cos(\pi x_1)\sin^2(\pi x_2).$$

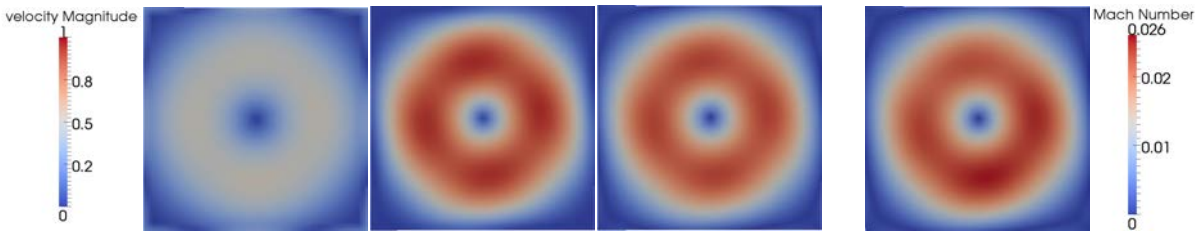
No-slip boundary conditions are imposed on the domain boundaries. The Mach number for the resulting flows is of order 0.026, so that we are in the low Mach regime. Results are displayed in table 3.1 and figures 3.1 and 3.2.

We first use the schemes EX($\theta = 1$) with a 400×400-cell and a 50×50-cell mesh. As expected the scheme performs poorly on the coarse mesh and the gain of accuracy is obvious when one refines the mesh : a mesh size of order M is required, but it comes at a much higher price in terms of CPU time as we can see on table 3.1. The EX($\theta = O(M)$) scheme gives good results even with the coarse 50×50-cell grid. With the low Mach correction scheme the connection between the accuracy of the solution and the mesh size does not seem to be constrained by M. Therefore, for a given target accuracy on a relatively coarse mesh, this numerical scheme is also much cheaper in term of CPU time.

Let us now turn to the semi-implicit strategies where the time step was chosen in agreement with the material CFL condition (3.66). While the IMEX($\theta = 1$) scheme is not CPU intensive on a coarse mesh the results are very altered by the numerical diffusion. The IMEX($\theta = O(M)$) scheme performs fast and allows to recover numerical results that are as good as EX($\theta = O(M)$). As with the EX($\theta = M$) scheme the accuracy seems much less constrained by the Mach number when it comes to choosing the mesh size. As we can see in table 3.1, the IMEX($\theta = O(M)$) scheme is 3.34 times faster than the EX($\theta = O(M)$).

TABLE 3.1 – Vortex in a box test case. Comparison of the number of iterations and CPU time of EX($\theta = 1$), EX($\theta = O(M)$), IMEX($\theta = 1$) and IMEX($\theta = O(M)$) schemes to obtain solutions of figure 3.1 and figure 3.2.

Numerical scheme Mesh	EX($\theta = 1$) 400×400	EX($\theta = 1$) 50×50	EX($\theta = O(M)$) 50×50	IMEX($\theta = 1$) 50×50	IMEX($\theta = O(M)$) 50×50
Number of iterations	18 457	2 306	2 305	43	56
CPU time (s)	9 263.04 (2h 34 min)	17.14	19.3	3.75	5.77



(a) (b) (c) (d)

FIGURE 3.1 – Vortex in a box test case with a Cartesian mesh. Profile at time $t = 0.125$ s of the velocity magnitude for (a) EX($\theta = 1$), (b) EX($\theta = O(M)$) with a 50×50-cell Cartesian mesh, (c) velocity magnitude obtained with EX($\theta = 1$) using a 400×400 Cartesian mesh and (d) Mach number obtained with EX($\theta = 1$) using a 400×400 Cartesian mesh.

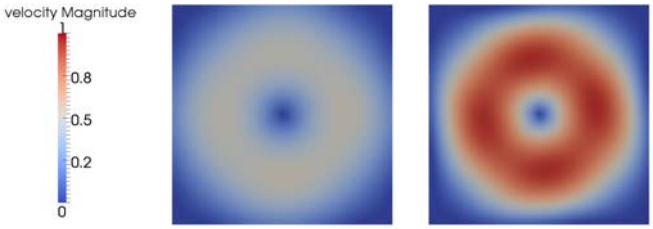


FIGURE 3.2 – Vortex in a box test case with a Cartesian mesh. Profile at time $t = 0.125$ s of the velocity magnitude for the IMEX($\theta = 1$) scheme (left) and the IMEX($\theta = O(M)$) scheme (right) on a 50×50-cell Cartesian mesh.

TABLE 3.2 – Backward facing step test case. Comparison of the number of iterations and CPU time of EX($\theta = 1$), EX($\theta = O(M)$), IMEX($\theta = 1$) and IMEX($\theta = O(M)$) schemes to obtain solutions of figure 3.3 and figure 3.5.

Numerical scheme	EX($\theta = 1$)	EX($\theta = O(M)$)	IMEX($\theta = 1$)	IMEX($\theta = O(M)$)
Number of time steps	5 258 803	5 258 784	4 384	4 970
CPU time (s)	62 805.6 (17h 27min)	69 764.7 (19h 22min)	418.37 (6min 58s)	500.83 (8min 20s)

Backward facing step

We consider now the case of an inviscid flow passing a backward facing step as derived from [7]. The computational domain is $\Omega = [0, 18] \times [0, 2] \setminus (0, 4) \times (0, 1)$. The initial condition is given by

$$\rho_0(x_1, x_2) = 10, \qquad u_0(x_1, x_2) = 1, \qquad p_0(x_1, x_2) = 10^5, \qquad v_0(x_1, x_2) = 0.$$

We impose an inlet boundary condition at $\{0\} \times [1, 2]$ and an outlet boundary condition at $\{12\} \times [0, 2]$. Wall boundary conditions are set on other boundaries. This configuration leads to a low Mach flow with the order of magnitude $10^{-3} \leq M \leq 10^{-2}$. All tests are performed with a 220×20 Cartesian space grid.

Figure 3.3 and 3.5 display the flow profile at $t = 50\,\mathrm{s}$, we observe that EX($\theta = 1$) and IMEX($\theta = 1$) schemes do not capture the vortex of the fluid in the low Mach velocity region. On the contrary, thanks to the low Mach correction EX($\theta = O(M)$) and IMEX($\theta = O(M)$) schemes are both able to capture this vortex with a coarse Cartesian Mesh. In term of CPU cost, measure are presented in table 3.2. We observe that the IMEX($\theta = O(M)$) scheme is 139.29 times faster than the EX($\theta = O(M)$) scheme thanks to the use of material velocity CFL condition (3.66), due to the implicit treatment of the acoustic step.

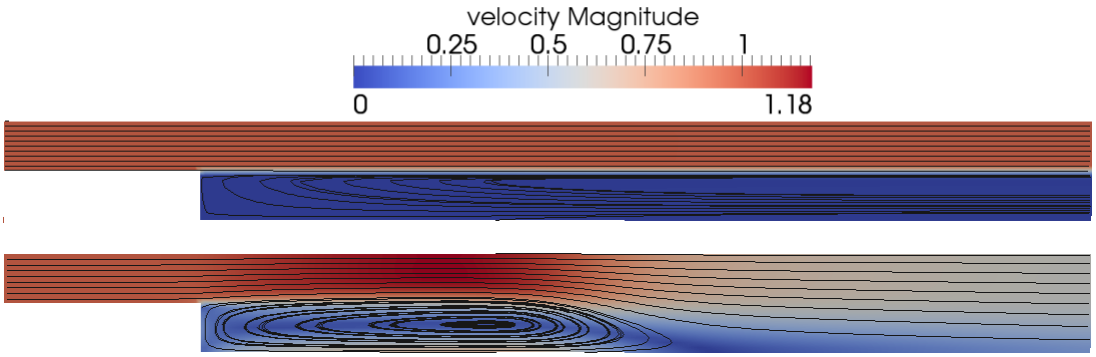


FIGURE 3.3 – Backward facing step test case. Profile at $t = 50\,\mathrm{s}$ of the velocity magnitude and stream lines for the EX($\theta = 1$) scheme (top) and the EX($\theta = O(M)$) scheme (bottom) on a 220×20-cell Cartesian mesh.

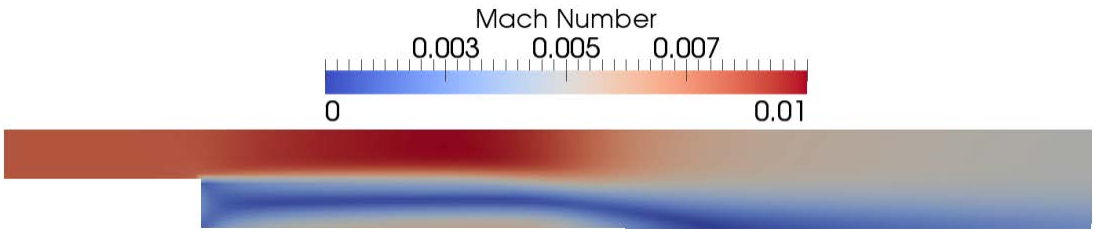


FIGURE 3.4 – Backward facing step test case. Mapping at $t = 50\,\mathrm{s}$ of the Mach number values for EX($\theta = O(M)$) scheme with a 220×20-cell Cartesian mesh.

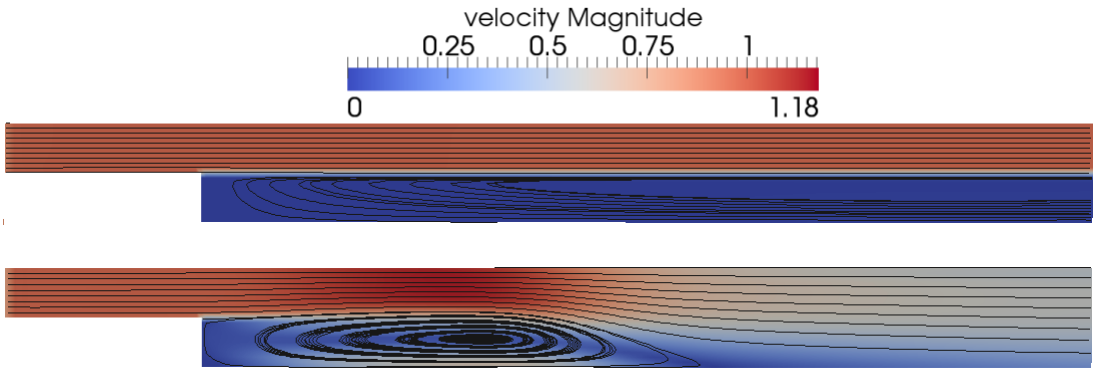


FIGURE 3.5 – Backward facing step test case. Mapping at $t = 50\,\mathrm{s}$ of the velocity magnitude and stream lines obtained with the IMEX($\theta = 1$) scheme (top) and the IMEX($\theta = O(M)$) scheme (bottom) using a 220×20-cell Cartesian mesh.

3.6.2 Compressible flow examples

In this section, we assess the ability of our operator splitting scheme to handle cases where the flow may not remain uniformly in the same Mach regime over the whole computational domain Ω. We will see that even with a centred pressure discretization (which corresponds to the choice $\theta = 0$), the solution remains stable but may be less precise in area where the Mach number is of order 1. The semi-implicit scheme becomes slower than the explicit scheme when the Mach number is of order 1 as the benefit from using a material CFL (3.66) condition instead of an acoustic CFL (3.65) becomes less beneficial but requires solving a linear system.

1D Sod shock tube

We consider a variant of the classical Sod shock tube [27], that consists in solving the one-dimensional Riemann problem over $\Omega = [0, 1]$ defined by the initial conditions $(\rho, u, P) = (1.0, 0.0, 10^5)$ for $x < 0.5$ and $(\rho, u, P) = (0.1, 0.0, 10^4)$ for $x > 0.5$.

We impose Neumann boundary conditions during the test. The domain is discretized over a 1000-cell grid. This resulting Mach number verifies $0 < M < 0.95$, so that we have both low Mach and order 1 Mach values. We plot the solution at $t = 3.1 \times 10^{-4}\,\mathrm{s}$.

Figure 3.6 displays the results obtained with EX(θ) and IMEX(θ) for $\theta = 1$ and $\theta = 0$. We use as reference solution an approximation computed with EX($\theta = 1$) using a 10 000-cell mesh. All schemes show a good agreement with the reference solution. The schemes EX($\theta = 0$) and IMEX($\theta = 0$) schemes are slightly less diffused than the EX($\theta = 1$) and IMEX($\theta = 1$) schemes results. Let us underline that despite part of the solutions clearly do not belong to the low Mach regime since $M \simeq 0.95$, the schemes EX($\theta = 0$) and IMEX($\theta = 0$) are stable and provide good numerical results while involving a centered pressure discretization with $\theta_{ij} = 0$.

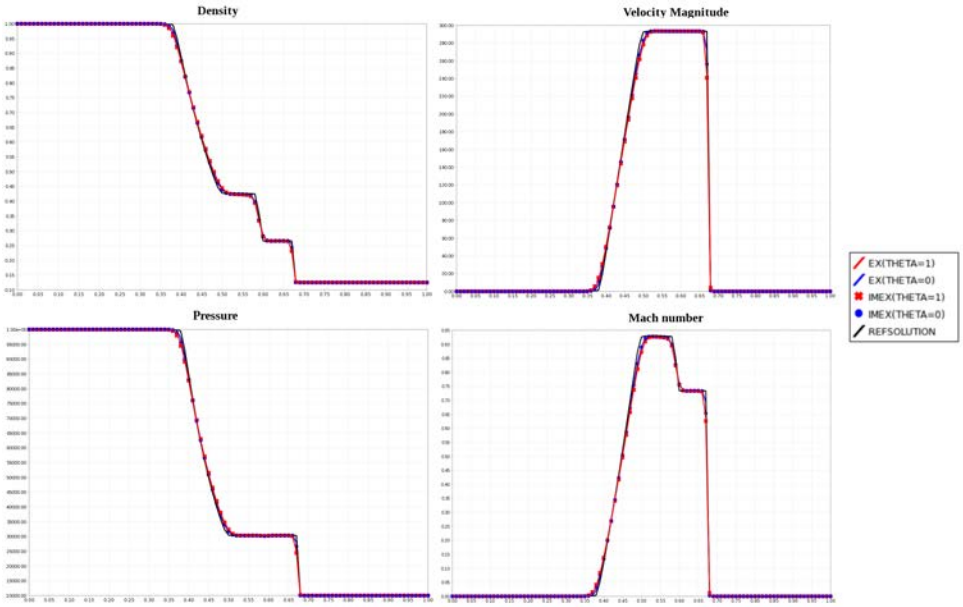


FIGURE 3.6 – 1D Sod shock tube test case. Profile at $t = 3.1 \times 10^{-4}$ s of the density (top left), velocity magnitude (top right), pressure (bottom left) and Mach number (bottom right) for the EX($\theta = 1$), EX($\theta = 0$), IMEX($\theta = 1$), IMEX($\theta = 0$) using a 1000-cell grid, together with reference solution.

2D-Riemann problem

We consider a 2D Riemann problem that consists of 4 shock waves [23]. We consider the domain $\Omega = [0, 1]^2$. The initial condition is

$$(\rho, u_1, u_2, P)(x_1, x_2, t = 0) = \begin{cases} (0.1380,\ 1.206,\ 1.206,\ 0.029), & \text{for } x_1 < 0.5, \quad x_2 < 0.5 \\ (0.5323,\ 0.000,\ 1.206,\ 0.300), & \text{for } x_1 > 0.5, \quad x_2 < 0.5 \\ (0.5323,\ 1.206,\ 0.000,\ 0.300), & \text{for } x_1 < 0.5, \quad x_2 > 0.5 \\ (1.5000,\ 0.000,\ 0.000,\ 1.500), & \text{for } x_1 > 0.5, \quad x_2 > 0.5 \end{cases}$$

We impose Neumann boundary conditions. This configuration leads to a Mach number that ranges from 10^{-5} to 3.15, i.e. according to the regions of the computation domain, the flow belongs to the low Mach regime or the order 1 Mach regime. We consider as a reference solution the approximation obtained with EX($\theta = 1$) for a 200×200-cell Cartesian mesh. Figures 3.7, 3.8, 3.9 and 3.10 display the result at $t = 0.4$ s.

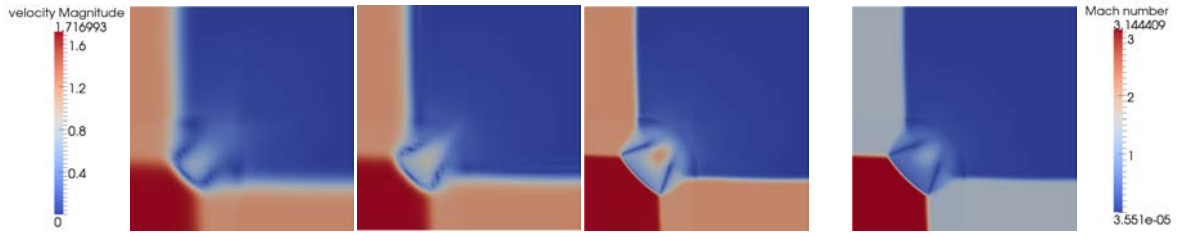
We observe in figure 3.7 and figure 3.8 that EX($\theta = 0$) and IMEX($\theta = 0$) schemes are stable for this test case with both low Mach and order 1 Mach number values regions. Both figures show that the wave pattern at the center of the domain shape is better captured with coarse meshes when one uses the corrected schemes ($\theta = 0$). A 1D cut along the axis $y = x$ as depicted in figure 3.9, also corroborates this observation : the approximation obtained with EX($\theta = 0$) and IMEX($\theta = 0$) schemes are closer to the 200×200-cell reference solution thanks to the numerical diffusion reduction. Nonetheless, we observe on a 1D cut along the $x = 0.75$ axis in figure 3.10 a spurious overshot for both density and pressure located at the shock front with EX($\theta = 0$) and IMEX($\theta = 0$). This suggests that a small value of θ allows to improve the precision of the scheme by reducing the numerical diffusion but it may cause overshoots if the value of θ becomes too small relatively to the local behavior of flow. In all our numerical experiments the scheme seems to remain stable for any value of $\theta \in (0, 1)$. Let us note that even if the pressure gradient is given a centred treatment ($\theta = 0$), the transport step introduce some numerical diffusion (independent

TABLE 3.3 – 2D Riemann problem test case. Comparison of the number of time steps and CPU time necessary for reaching $t = 0.4\,\mathrm{s}$ with a 50×50-cell Cartesian grid with EX($\theta = 1$), EX($\theta = 0$), IMEX($\theta = 1$) and IMEX($\theta = 0$).

Numerical scheme	EX($\theta = 1$)	EX($\theta = 0$)	IMEX($\theta = 1$)	IMEX($\theta = 0$)
Number of iterations	323	343	216	218
CPU time (s)	2.59	2.79	10.28	10.33

of M) that stabilize the scheme see (3.32).

In table 3.3 we observe that the choice of θ does not impact the number of time steps and CPU time. For this case, while the number of time steps is slightly reduced by about 30%, the semi-implicit schemes are much slower due to the time required for solving the linear system involved with the schemes.



(a) (b) (c) (d)

FIGURE 3.7 – 2D Riemann problem with a Cartesian mesh. Profile at $t = 0.4\,\mathrm{s}$ of the velocity magnitude for (a) EX($\theta = 1$), (b) EX($\theta = 0$) with a 50×50-cell mesh, (c) velocity magnitude and (d) Mach number with EX($\theta = 1$) using a 200×200 mesh.

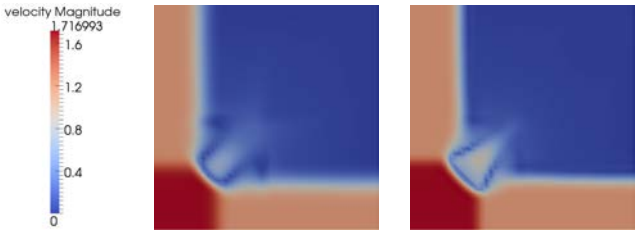


FIGURE 3.8 – 2D Riemann problem test case. Profile at $t = 0.4\,\mathrm{s}$ of the velocity magnitude for IMEX($\theta = 1$) (left) and IMEX($\theta = 0$) (right) on a 50×50-cell Cartesian mesh.

As a partial conclusion of this section, we can observe that for tests that strongly involve the compressibility of the fluid both semi-implicit and explicit schemes seem to be very robust, independantly of the choice of θ within $[0, 1]$. However, if the low Mach correction is too important, *i.e.* the value of θ is too close to 0 we witnessed a deterioration of the numerical approximation with the appearance of overshoots in the vicinity of shock fronts.

Then some numerical criterion may be constructed with good properties, $\theta_{ij} = \min\left(\frac{|u^*_{ij}|}{\max(c^n_i, c^n_j)}, 1\right)$ for instance.

We also observed that the benefit in terms of CPU time of the semi-implicit scheme vanishes when the Mach number becomes of order 1.

The implementation of the criterion on θ to recover a discrete entropy inequality does not allow to recover a good low Mach behaviour as it was expected from its low Mach analysis. Finding from a theoretical point of view a criterion on θ that allows to recover a good low Mach behaviour and avoid

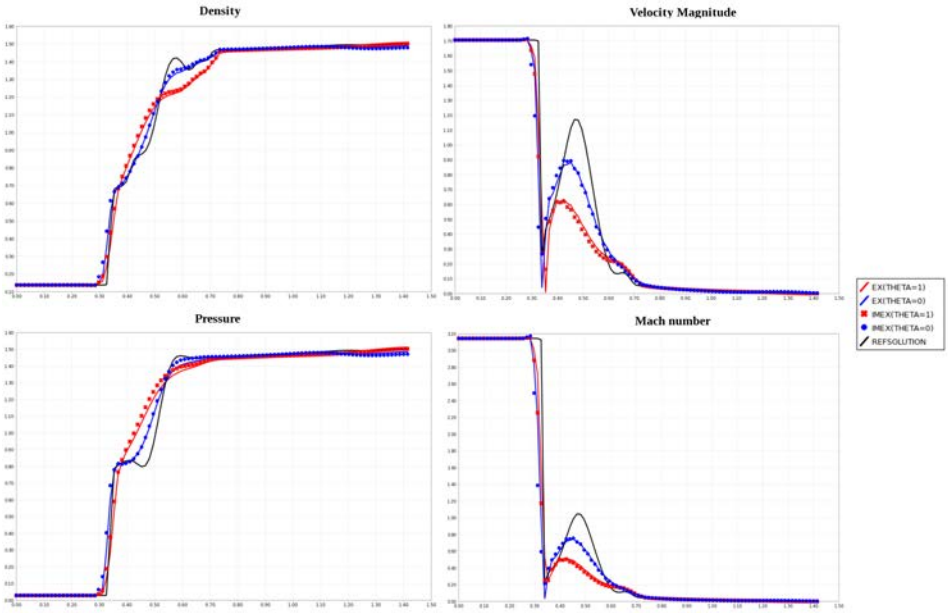


FIGURE 3.9 – 2D Riemann problem test case. Cut profile along $y = x$ at $t = 0.4\,\mathrm{s}$ of the density, velocity magnitude, pressure and Mach number for EX($\theta = 1$), EX($\theta = 0$), IMEX($\theta = 1$) and IMEX($\theta = 0$) using a 50×50 mesh together with the 200×200-cell reference solution.

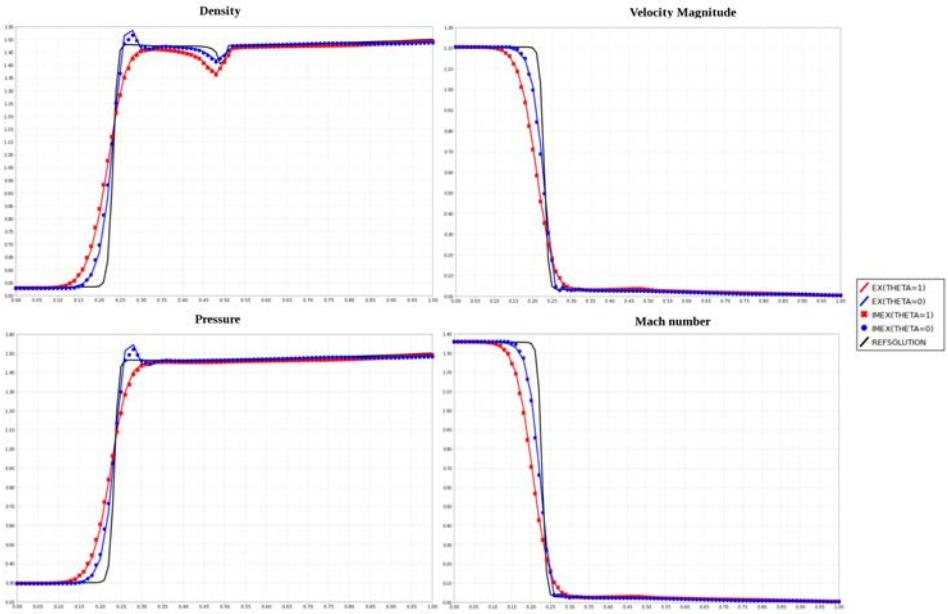


FIGURE 3.10 – 2D Riemann problem test case. Cut profile along $x = 0.75$ at $t = 0.4\,\mathrm{s}$ of the density, velocity magnitude, pressure and Mach number for EX($\theta = 1$), EX($\theta = 0$), IMEX($\theta = 1$) and IMEX($\theta = 0$) using a 50×50 mesh together with the 200×200-cell reference solution.

spurious phenomenon that may occur if θ is too small for a given configuration is still an open problem.

Remark 6. *The robustness of the scheme with respect to the modification $\theta \geq 0$ seems to be linked to the Lagrange-Projection decomposition approach. Indeed, numerical evidences not presented here show that a modified relaxation scheme written in Eulerian coordinates is unstable outside of the low-Mach regime for value of θ that are too small.*

3.7 Conclusion

We proposed a conservative operator splitting based Lagrange-Projection like numerical strategy for approximating the gas dynamics that decouples acoustic and transport phenomenons. The operator splitting scheme is positive for the density, the internal energy and entropic under classical CFL conditions. For one-dimensional problem, this procedure is equivalent to a Lagrange-Projection discretization. We presented an analysis of the way the truncation error depends on the Mach number for one-dimensional problems. In the low Mach regime, the truncation error of the scheme showed to be non-uniform with respect to the Mach number M. This allowed us to modify the operator splitting scheme in order to recover a uniform truncation error in term of M by altering the numerical flux in the acoustic approximation. We showed that this modification can be obtained thanks to a simple approximate Riemann solver that is consistent with the integral form of the PDEs. This modified operator splitting scheme is conservative and endowed with good stability properties with respect to the positivity of the density, the internal energy under classical acoustic CFL conditions that depend on M. The resulting scheme allows to deal with tests where the flow regime may vary from low to high Mach values.

We showed that this splitting strategy has a natural extension to multi-dimensional problems discretized over unstructured meshes. A simple and efficient semi-implicit scheme that is stable under CFL conditions based on the material velocity is also proposed and leads to an *all-regime* numerical scheme, following the ideas paved by [6] for one-dimensional problems.

Future developments include extensions to high-order methods and approximation of other systems for the simulation of multi-material flows.

Acknowledgement

The authors would like to thank Pierre-Arnaud Raviart for sharing his personal notes and for the many rich discussions that brought an invaluable contribution to this work.

Annexes

3.A Classical Lagrange-Projection for one-dimensional gas dynamics

In this section we briefly recall the classical Lagrange-Projection (or Lagrange-Remap) procedure for deriving a Finite Volume discretization within a one-dimensional framework. For a detailed description we refer the reader to [16, 13]. Let $(X,t)\mathbb{R}\times[t^n,t^n+\Delta t)\mapsto\chi$ be the mapping defined by

$$\partial_t\chi = u(\chi(X,t),t), \qquad \chi(X,t=t^n)=X.$$

The pair (X,t) is usually referred to as the Lagrangian system of coordinates : a particle of fluid at the position X at instant $t=t^n$ will be located at $x=\chi(X,t)$, $t\in[t^n,t^n+\Delta t]$. If $(x,t)\mapsto b$ is a mapping that provides an Eulerian representation of a parameter b, one defines a Lagrangian representation of b as the function $(X,t)\mapsto b^{\text{Lag}}$ by setting $b^{\text{Lag}}(X,t)=b(\chi(X,t),t)$. The system (3.3) is equivalent to

$$\begin{cases} \partial_t\mathbf{V}^{\text{Lag}}(X,t)+\tau^{\text{Lag}}(X,t^n)\partial_X\mathbf{F}^{\text{Lag}}(\mathbf{V}^{\text{Lag}})(X,t)=0, \\ \mathbf{V}^{\text{Lag}}=(\tau^{\text{Lag}},u^{\text{Lag}},v^{\text{Lag}},E^{\text{Lag}})^T, \quad \mathbf{F}^{\text{Lag}}(\mathbf{V}^{\text{Lag}})=(-u^{\text{Lag}},p^{\text{Lag}},0,p^{\text{Lag}}u^{\text{Lag}})^T. \end{cases} \tag{3.67}$$

It is common to introduce a mass coordinate m defined by $\mathrm{d}m=\rho(X,t^n)\mathrm{d}X$ in order to obtain the equivalent conservation laws (with a slight abuse of notation)

$$\partial_t\mathbf{V}^{\text{Lag}}(m,t)+\partial_m\mathbf{F}^{\text{Lag}}(\mathbf{V}^{\text{Lag}})(m,t)=0. \tag{3.68}$$

Straightforward calculations show that (3.68) (which is nothing but (3.6)) is hyperbolic over the phase space $\Omega^{\text{Lag}}=\{(\tau^{\text{Lag}},u^{\text{Lag}},v^{\text{Lag}},E^{\text{Lag}})^T\in\mathbb{R}^4,\ \tau^{\text{Lag}}>0,\ e^{\text{Lag}}>0\}$, with eigenvalues given by $\lambda_1^{\text{Lag}}=-\rho c<\lambda_2^{\text{Lag}}=0<\lambda_3^{\text{Lag}}=\rho c$, where c still denotes the Eulerian sound speed. Here again, the extreme characteristic fields associated with λ_1^{Lag} and λ_3^{Lag} are genuinely non linear while the intermediate characteristic field associated with λ_2^{Lag} is linearly degenerate. It is important to note that the material transport phenomenons are frozen in system (3.68) which explains why the characteristics speeds of the system only involve the sound velocity c.

Before going any further, we introduce classical notations : let $\Delta t>0$ and $\Delta x>0$ be respectively the time and space steps. We define the Eulerian mesh interfaces $x_{j+1/2}=j\Delta x$ for $j\in\mathbb{Z}$, and the intermediate times $t^n=n\Delta t$ for $n\in\mathbb{N}$. If b is a fluid parameter, in the sequel, we will note b_j^n (resp. b_j^{n+1}) the approximate value b respectively within the j^{th} Eulerian cell $[x_{j-1/2},x_{j+1/2})$ at instant $t=t^n$ (resp. $t=t^{n+1}$). We need to introduce a moving Lagrangian mesh (with respect to the Eulerian mesh) whose cell j at instant t^n is $[x_{j-1/2},x_{j+1/2})$ and at instant $t=t^{n+1}$ is $[x^*_{j-1/2},x^*_{j+1/2})$. The value of the parameter b at instant t^n (resp. $t=t^{n+1}$) in the Lagrangian cell j is noted b_j^{Lag} (resp. b_j^{n+1-}). Given a fluid state $(\rho,\rho u,\rho v,\rho E)_j^n$, $j\in\mathbb{Z}$ at instant t^n, the Lagrange-Projection strategy proposes the following update procedure.

1. Build the discrete Lagrangian fluid state at instant t^n by setting $(\mathbf{V}^{\text{Lag}})_j=(\tau_j^n,u_j^n,v_j^n,E_j^n)$;
2. Update the Lagrangian fluid state into the value $(\mathbf{V}^{\text{Lag}})_j^{n+1-}=(\tau_j^{n+1-},u_j^{n+1-},v_j^{n+1-},E_j^{n+1-})$ by approximating the solution of (3.68) ;
3. Build the updated value $(\rho,\rho u,\rho v,\rho E)_j^{n+1}$ by remapping the Lagrangian state $(\mathbf{V}^{\text{Lag}})_j^{n+1-}$ onto the Eulerian mesh.

The Lagrangian step ($t^n\to t^{n+1-}$)

We propose to approximate the solution of (3.68) using the acoustic scheme [12, 13]. This leads to

$$\begin{cases} \tau_j^{n+1-} = \tau_j^n + \dfrac{\Delta t}{\Delta x}\tau_j^n(u^*_{j+1/2} - u^*_{j-1/2}), & \text{(3.69a)} \\ u_j^{n+1-} = u_j^n - \dfrac{\Delta t}{\Delta x}\tau_j^n(p^*_{j+1/2} - p^*_{j-1/2}), & \text{(3.69b)} \\ v_j^{n+1-} = v_j^n, & \text{(3.69c)} \\ E_j^{n+1-} = E_j^n - \dfrac{\Delta t}{\Delta x}\tau_j^n\big((pu)^*_{j+1/2} - (pu)^*_{j-1/2}\big), & \text{(3.69d)} \end{cases}$$

where the interfaces terms are defined by

$$\begin{aligned} u^*_{j+1/2} &= \frac{(u_j^n + u_{j+1}^n)}{2} + \frac{1}{2a^n_{j+1/2}}(p_j^n - p_{j+1}^n), \quad & p^*_{j+1/2} &= \frac{(p_j^n + p_{j+1}^n)}{2} + \frac{a^n_{j+1/2}}{2}(u_j^n - u_{j+1}^n), \\ (pu)^*_{j+1/2} &= p^*_{j+1/2}u^*_{j+1/2}, & a^n_{j+1/2} &= \max((\rho c)_j^n, (\rho c)_{j+1}^n). \end{aligned} \tag{3.70}$$

The acoustic scheme (3.69) with (3.70) provides the same update of the flow variable as the scheme (3.9) with (3.10). Let us mention that a direct proof of stability for the acoustic scheme is available in [12] under the CFL criterion (3.17).

The projection (or remapping) step ($t^{n+1-} \to t^{n+1}$)
The aim of this step is to project the solution obtained at the end of the Lagrangian step onto the Eulerian cells $[x_{j-1/2}, x_{j+1/2})$. If one notes $\mathbb{1}_{[x^*_{j-1/2}, x^*_{j+1/2})}$ the characteristic function of $[x^*_{j-1/2}, x^*_{j+1/2})$, a standard way to achieve to goal consists in : first, approximating the position of the Lagrangian mesh interfaces at instant t^{n+1} by setting $x^*_{j+1/2} = x_{j+1/2} + u^*_{j+1/2}\Delta t$; second reaveraging the conservative variable unknowns over the Eulerian mesh by setting [16]

$$\varphi_j^{n+1} = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \left[\sum_{j\in\mathbb{Z}} \varphi_j^{n+1-}\mathbb{1}_{[x^*_{j-1/2}, x^*_{j+1/2})}(x)\right] \mathrm{d}x, \quad \text{where } \varphi \in \{\rho, \rho u, \rho v, \rho E\}. \tag{3.71}$$

Noting $\Delta x_j^* = x^*_{j+1/2} - x^*_{j-1/2}$ and $\varepsilon(j,n) = -\operatorname{sign}(u^*_{j+1/2})1/2$ one obtains the update formula

$$\begin{aligned} \varphi_j^{n+1} &= \frac{1}{\Delta x}\left\{\Delta x_j^*\varphi_j^{n+1-} - \Delta t\left(u^*_{j+1/2}\varphi^{n+1-}_{j+1/2+\varepsilon(j,n)} - u^*_{j-1/2}\varphi^{n+1-}_{j-1/2+\varepsilon(j-1,n)}\right)\right\} \\ &= \frac{\Delta t}{\Delta x}(u^*_{j+1/2} - u^*_{j-1/2})\varphi_j^{n+1-} - \frac{\Delta t}{\Delta x}\left(u^*_{j+1/2}\varphi^{n+1-}_{j+1/2+\varepsilon(j,n)} - u^*_{j-1/2}\varphi^{n+1-}_{j-1/2+\varepsilon(j-1,n)}\right) \end{aligned}. \tag{3.72}$$

The update formula (3.72) matches the classic upwind scheme. Consequently this is the same numerical scheme as (3.14).

3.B Approximate Riemann solvers : Harten Lax and van Leer formalism

We briefly recall the Harten, Lax and van Leer formalism associated with the numerical approximation of the solutions $(x,t) \in \mathbb{R}\times[0,+\infty) \mapsto \mathbf{U} \in \mathbb{R}^m$ of the general hyperbolic system of conservation laws

$$\partial_t \mathbf{U} + \partial_x \mathbf{G}(\mathbf{U}) = 0, \qquad x \in \mathbb{R},\ t > 0, \tag{3.73}$$

by means of the so-called approximate Riemann solvers and Godunov-type methods, where $\mathbf{G} : \mathbb{R}^m \to \mathbb{R}^m$ is a smooth function. System (3.73) is supplemented with the validity of an entropy inequality

$$\partial_t \eta(\mathbf{U}) + \partial_x q(\mathbf{U}) \leq 0, \tag{3.74}$$

where $\mathbf{U} \mapsto (\eta, q)$ is a strictly convex entropy-entropy flux pair (see [16]).

Solving the Riemann problem amounts to find the solution of (3.73) with the following piecewise constant initial data

$$\mathbf{U}(x, t = 0) = \begin{cases} \mathbf{U}_L, & \text{if} \quad x < 0, \\ \mathbf{U}_R, & \text{if} \quad x > 0, \end{cases}$$

for any given $\mathbf{U}_L$ and $\mathbf{U}_R$ in the phase space. It is well-known that the exact Riemann solution $\mathbf{U}(x/t; \mathbf{U}_L, \mathbf{U}_R)$ is self-similar, *i.e.* depends only on the ration x/t. In order to approximate this solution, we consider a (self-similar) simple approximate Riemann solver $\mathbf{U}_{\mathrm{RP}}(\frac{x}{t}; \mathbf{U}_L, \mathbf{U}_R)$ made of $l+1$ intermediate states $\mathbf{U}_k$ separated by discontinuities propagating with velocities $\lambda_1 \leq \cdots \leq \lambda_l$, namely

$$\mathbf{U}_{\mathrm{RP}}\left(\frac{x}{t}; \mathbf{U}_L, \mathbf{U}_R\right) = \begin{cases} \mathbf{U}_1 = \mathbf{U}_L, & \text{if} \quad x/t < \lambda_1, \\ \vdots & \\ \mathbf{U}_k, & \text{if} \quad \lambda_{k-1} < x/t < \lambda_k, \\ \vdots & \\ \mathbf{U}_{l+1} = \mathbf{U}_R, & \text{if} \quad x/t > \lambda_l. \end{cases} \tag{3.75}$$

From [18, 1], if $\Delta x = \frac{1}{2}(\Delta x_L + \Delta x_R)$ with $\Delta x_L > 0$, $\Delta x_R > 0$ and $\Delta t > 0$ are respectively space and time steps that verify the CFL condition

$$\max_{1 \leq k \leq l} |\lambda_k| \frac{\Delta t}{\min(\Delta x_L, \Delta x_R)} \leq \frac{1}{2}, \tag{3.76}$$

such an approximate Riemann solver is said to be consistent with the integral form of (3.73) over the interval $[\frac{-\Delta x_L}{2}, \frac{\Delta x_R}{2}] \times [0, \Delta t]$ if $\iint_{[\frac{-\Delta x_L}{2}, \frac{\Delta x_R}{2}] \times [0, \Delta t]} [\partial_t \mathbf{U}_{\mathrm{RP}} + \partial_x \mathbf{G}(\mathbf{U}_{\mathrm{RP}})] \, \mathrm{d}x\mathrm{d}t = 0$, in other words if

$$\mathbf{G}(\mathbf{U}_R) - \mathbf{G}(\mathbf{U}_L) = \sum_{k=1}^{l} \lambda_k (\mathbf{U}_{k+1} - \mathbf{U}_k). \tag{3.77}$$

Regarding the consistency with the entropy inequality (3.74), the simple approximate Riemann solver is said to be consistent with the integral form of (3.74) if and only if under the CFL condition (3.76)

we have

$$q(\mathbf{U}_R) - q(\mathbf{U}_L) \leq \sum_{k=1}^{l} \lambda_k \big(\eta(\mathbf{U}_{k+1}) - \eta(\mathbf{U}_k)\big). \tag{3.78}$$

Hereafter and using classic notations, $(\Delta x_j)_{j \in \mathbb{Z}}$ and Δt represent the space steps and constant time step of the mesh under consideration to define the approximate solutions. More precisely,

we define the mesh interfaces $x_{j+1/2} = x_{j-1/2} + \Delta x_j$ for $j \in \mathbb{Z}$, the intermediate times $t^n = n\Delta t$ for $n \in \mathbb{N}$,

and we note $\mathbf{U}_j^n$ the approximate value of $\mathbf{U}$ at time t^n and on the cell $[x_{j-1/2}, x_{j+1/2})$.

For $n = 0$ and $j \in \mathbb{Z}$, we set $\mathbf{U}_j^0 = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{U}_0(x) \, \mathrm{d}x$ where $\mathbf{U}_0(x)$ is the initial condition. Then,

the explicit in time Godunov-type scheme reads

$$\begin{cases} \mathbf{U}_j^{n+1} = \mathbf{U}_j^n - \dfrac{\Delta t}{\Delta x_j}(\mathbf{G}_{j+\frac{1}{2}}^n - \mathbf{G}_{j-\frac{1}{2}}^n), & (3.79a) \\ \mathbf{G}_{j+\frac{1}{2}}^n = \mathbf{G}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n), & (3.79b) \end{cases}$$

with

$$\mathbf{G}(\mathbf{U}_L, \mathbf{U}_R) = \frac{1}{2}\left[\mathbf{G}(\mathbf{U}_L) + \mathbf{G}(\mathbf{U}_R) - \sum_{k=1}^{l} |\lambda_k|(\mathbf{U}_{k+1} - \mathbf{U}_k)\right]. \tag{3.80}$$

Moreover, if the simple approximate Riemann solver is consistent with the entropy inequality (3.74), then the numerical scheme defined by (3.79) satisfies the following discrete entropy inequality

$$\begin{cases} \eta(\mathbf{U}_j^{n+1}) \leq \eta(\mathbf{U}_j^n) - \dfrac{\Delta t}{\Delta x_j}(q_{j+\frac{1}{2}}^n - q_{j-\frac{1}{2}}^n), \\ q_{j+\frac{1}{2}}^n = \widetilde{q}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n), \end{cases}$$

with

$$\widetilde{q}(\mathbf{U}_L, \mathbf{U}_R) = \frac{1}{2}\left[q(\mathbf{U}_L) + q(\mathbf{U}_R) - \sum_{k=1}^{l} |\lambda_k|\big(S(\mathbf{U}_{k+1}) - S(\mathbf{U}_k)\big)\right]. \tag{3.82}$$

The CFL condition associated with this (explicit in time) Godunov-type scheme reads

$$\max_{1\leq k\leq l} \mid \lambda_k(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n) \mid \frac{\Delta t}{\min(\Delta x_j, \Delta x_{j+1})} \leq \frac{1}{2},$$

for all j. Again, we refer to [18, 1]

for more details. To conclude this paragraph, let us observe that the numerical flux $\mathbf{G}(\mathbf{U}_L, \mathbf{U}_R)$ and the entropy numerical flux $\widetilde{q}(\mathbf{U}_L, \mathbf{U}_R)$ are clearly consistent in the classical sense, namely $\mathbf{G}(\mathbf{U}, \mathbf{U}) = \mathbf{G}(\mathbf{U})$ and $\widetilde{q}(\mathbf{U}, \mathbf{U}) = q(\mathbf{U})$ provided that the intermediate states of the approximate Riemann solver are such that $\mathbf{U}_k = \mathbf{U}$ for all $k = 1, ..., l$ as soon as $\mathbf{U}_L = \mathbf{U}_R = \mathbf{U}$.

3.C Riemann problem for the relaxation approximation of the acoustic system

We consider the Suliciu relaxation approximation of the Lagrangian gas dynamics equations expressed using a mass coordinate. The system reads

$$\begin{cases} \partial_t \tau - \partial_m u = 0, & (3.83a) \\ \partial_t u + \partial_m \Pi = 0, & (3.83b) \\ \partial_t v = 0, & (3.83c) \\ \partial_t E + \partial_m(\Pi u) = 0, & (3.83d) \\ \partial_t \Pi + a^2 \partial_m u = \lambda(p - \Pi), & (3.83e) \end{cases}$$

where a is a constant that verifies the subcharacteristic condition $a > \max(\rho c)$ in order to prevent instabilities (see for instance [3] for a rigorous proof). It is easy to prove that the convective part of (3.83) is strictly hyperbolic with three eigenvalues given by $-a$, 0 and a which correspond to linearizations of the exact eigenvalues $-\rho c$, 0 and ρc for system (3.67). Interestingly, the characteristic fields are linearly degenerate, which allows to solve analytically the Riemann problem associated with (3.83) with $\lambda = 0$.

More precisely, the exact Riemann solution

$$\overline{\mathbf{W}}(\frac{m}{t};\overline{\mathbf{U}}_L,\overline{\mathbf{U}}_R)=(\tau,u,v,E,\Pi)^T(\frac{m}{t};\overline{\mathbf{U}}_L,\overline{\mathbf{U}}_R)$$

associated with given left state $\overline{\mathbf{U}}_L=(\tau,u,v,E,\Pi)_L^T$ and right state $\overline{\mathbf{U}}_R=(\tau,u,v,E,\Pi)_R^T$, is made of three contact discontinuities propagating with velocities $-a$, a and 0 and separating two intermediate states $\overline{\mathbf{U}}_L^*$ and $\overline{\mathbf{U}}_R^*$, namely

$$\overline{\mathbf{W}}\Big(\frac{m}{t};\overline{\mathbf{U}}_L,\overline{\mathbf{U}}_R\Big)=\begin{cases}\overline{\mathbf{U}}_L, & \text{if } \frac{m}{t}<-a,\\ \overline{\mathbf{U}}_L^*, & \text{if } -a<\frac{m}{t}<0,\\ \overline{\mathbf{U}}_R^*, & \text{if } 0<\frac{m}{t}<a,\\ \overline{\mathbf{U}}_R, & \text{if } \frac{m}{t}>a.\end{cases}\tag{3.84}$$

The intermediate states are easily recovered from the following formulas

$$\begin{cases} u^*=u_L^*=u_R^*=\dfrac{u_R+u_L}{2}-\dfrac{\Pi_R-\Pi_L}{2a}, & \Pi_L^*=\Pi_R^*=\Pi^*=\dfrac{\Pi_R+\Pi_L}{2}-a\dfrac{u_R-u_L}{2}, & \text{(3.85a)}\\ v_L^*=v_L, & v_R^*=v_R, & \text{(3.85b)}\\ \tau_L^*=\tau_L+\dfrac{u^*-u_L}{a}, & \tau_R^*=\tau_R+\dfrac{u_R-u^*}{a}, & \text{(3.85c)}\\ E_L^*=E_L+\dfrac{1}{a}\Big(p_Lu_L-u^*\Pi^*\Big), & E_R^*=E_R-\dfrac{1}{a}\Big(p_Ru_R-u^*\Pi^*\Big). & \text{(3.85d)}\end{cases}$$

Then, setting $\mathbf{U}_L=(\tau,u,v,E)^T$, the classical scheme can be understood in the Harten, Lax and van Leer formalism by considering the following approximate Riemann solver $\mathbf{W}\left(\frac{m}{t};\mathbf{U}_L,\mathbf{U}_R\right)$ obtained by simply extracting the first four components from $\overline{\mathbf{W}}\left(\frac{m}{t};\overline{\mathbf{U}}_L,\overline{\mathbf{U}}_R\right)$, in which we take Π at equilibrium, namely

$$\Pi_L=p_L,\quad \Pi_R=p_R.$$

More precisely, we have

$$\mathbf{W}\Big(\frac{m}{t};\mathbf{U}_L,\mathbf{U}_R\Big)=\begin{cases}\mathbf{U}_L, & \frac{m}{t}<-a,\\ \mathbf{U}_L^*, & -a<\frac{m}{t}<0,\\ \mathbf{U}_R^*, & 0<\frac{m}{t}<a,\\ \mathbf{U}_R, & \frac{m}{t}>a,\end{cases}\tag{3.86}$$

where the intermediate states are given by (3.85), together with $\Pi_L=p_L$ and $\Pi_R=p_R$.

3.D Adimensionnement et système limite

On s'intéresse ici à l'obtention du système adimensionné (3.20), puis on considère sa limite lorsque le nombre de Mach M tend vers zéro. On injecte les expressions des grandeurs adimensionnées définies par

(3.19) avec $u_0 = v_0$ dans le système de la dynamique des gaz (3.1) pour obtenir

$$\begin{cases} \left(\dfrac{L}{u_0 T}\right) \partial_{\tilde{t}} \tilde{\rho} + \tilde{\nabla} \cdot (\tilde{\rho} \tilde{\mathbf{u}}) = 0, & \text{(3.87a)} \\ \left(\dfrac{L}{u_0 T}\right) \partial_{\tilde{t}} (\tilde{\rho} \tilde{\mathbf{u}}) + \tilde{\nabla} \cdot (\tilde{\rho} \tilde{\mathbf{u}} \otimes \tilde{\mathbf{u}}) + \left(\dfrac{p_0}{\rho_0 (u_0)^2}\right) \tilde{\nabla} \tilde{p} = 0, & \text{(3.87b)} \\ \left(\dfrac{L}{u_0 T}\right) \partial_{\tilde{t}} (\tilde{\rho} \tilde{e}) + \tilde{\nabla} \cdot (\tilde{\rho} \tilde{e} \tilde{\mathbf{u}}) + \left(\dfrac{p_0}{\rho_0 e_0}\right) \tilde{\nabla} \cdot (\tilde{p} \tilde{\mathbf{u}}) \\ + \left(\dfrac{L u_0}{e_0 T}\right) \partial_{\tilde{t}} \left(\tilde{\rho} \dfrac{|\tilde{\mathbf{u}}|^2}{2}\right) + \left(\dfrac{(u_0)^2}{e_0}\right) \tilde{\nabla} \cdot \left(\tilde{\rho} \dfrac{|\tilde{\mathbf{u}}|^2}{2} \tilde{\mathbf{u}}\right) = 0. & \text{(3.87c)} \end{cases}$$

On considère alors que les grandeurs caractéristiques spatiale et temporelle sont associées au phénomène de transport, c'est à dire $u_0 = \frac{L}{T}$. Par ailleurs, on définit une vitesse caractéristique du son $c_0 = \sqrt{\frac{p_0}{\rho_0}}$. On a alors $\frac{p_0}{\rho_0 (u_0)^2} = \frac{1}{M^2}$, où $M = \frac{u_0}{c_0}$ est le nombre de Mach. Le système (3.87) devient

$$\begin{cases} \partial_{\tilde{t}} \tilde{\rho} + \tilde{\nabla} \cdot (\tilde{\rho} \tilde{\mathbf{u}}) = 0, & \text{(3.88a)} \\ \partial_{\tilde{t}} (\tilde{\rho} \tilde{\mathbf{u}}) + \tilde{\nabla} \cdot (\tilde{\rho} \tilde{\mathbf{u}} \otimes \tilde{\mathbf{u}}) + \dfrac{1}{M^2} \tilde{\nabla} \tilde{p} = 0, & \text{(3.88b)} \\ \partial_{\tilde{t}} (\tilde{\rho} \tilde{e}) + \tilde{\nabla} \cdot (\tilde{\rho} \tilde{e} \tilde{\mathbf{u}}) + \left(\dfrac{p_0}{\rho_0 e_0}\right) \tilde{\nabla} \cdot (\tilde{p} \tilde{\mathbf{u}}) \\ + \left(\dfrac{p_0}{\rho_0 e_0}\right) M^2 \left[\partial_{\tilde{t}} \left(\tilde{\rho} \dfrac{|\tilde{\mathbf{u}}|^2}{2}\right) + \tilde{\nabla} \cdot \left(\tilde{\rho} \dfrac{|\tilde{\mathbf{u}}|^2}{2} \tilde{\mathbf{u}}\right)\right] = 0. & \text{(3.88c)} \end{cases}$$

On définit le rapport $K = \frac{p_0}{\rho_0 e_0}$ et on obtient finalement le système adimensionné

$$\begin{cases} \partial_{\tilde{t}} \tilde{\rho} + \tilde{\nabla} \cdot (\tilde{\rho} \tilde{\mathbf{u}}) = 0, & \text{(3.89a)} \\ \partial_{\tilde{t}} (\tilde{\rho} \tilde{\mathbf{u}}) + \tilde{\nabla} \cdot (\tilde{\rho} \tilde{\mathbf{u}} \otimes \tilde{\mathbf{u}}) + \dfrac{1}{M^2} \tilde{\nabla} \tilde{p} = 0, & \text{(3.89b)} \\ \partial_{\tilde{t}} (\tilde{\rho} \tilde{e}) + \tilde{\nabla} \cdot (\tilde{\rho} \tilde{e} \tilde{\mathbf{u}}) + K \tilde{\nabla} \cdot (\tilde{p} \tilde{\mathbf{u}}) \\ + K M^2 \left[\partial_{\tilde{t}} \left(\tilde{\rho} \dfrac{|\tilde{\mathbf{u}}|^2}{2}\right) + \tilde{\nabla} \cdot \left(\tilde{\rho} \dfrac{|\tilde{\mathbf{u}}|^2}{2} \tilde{\mathbf{u}}\right)\right] = 0. & \text{(3.89c)} \end{cases}$$

Remarque 1. *L'introduction des grandeurs caractéristiques p_0, ρ_0 et e_0 pour la pression, la densité et l'énergie interne est liée au comportement thermodynamique du gaz considéré. En effet, on a $p_0 \simeq p(\rho_0, e_0)$ et donc $K \simeq \frac{p(\rho_0, e_0)}{\rho_0 e_0}$. Pour une loi d'état de type gaz parfait, il est naturel de considérer $p_0 = \rho_0 e_0$ et donc $K = 1$.*

On s'intéresse maintenant à la limite à bas nombre de Mach du système (3.89). Des manipulations classiques montrent que les solutions régulières de (3.89) sont équivalentes aux solutions régulières du système suivant

$$\begin{cases} \partial_{\tilde{t}} \tilde{\rho} + \tilde{\nabla} \cdot (\tilde{\rho} \tilde{\mathbf{u}}) = 0, & \text{(3.90a)} \\ \partial_{\tilde{t}} \tilde{\mathbf{u}} + \left(\tilde{\mathbf{u}} \cdot \tilde{\nabla}\right) \tilde{\mathbf{u}} + \dfrac{1}{M^2 \tilde{\rho}} \tilde{\nabla} \tilde{p} = 0, & \text{(3.90b)} \\ \partial_{\tilde{t}} \tilde{p} + K \tilde{\mathbf{u}} \cdot \tilde{\nabla} \tilde{p} + \tilde{\rho} \tilde{c}^2 \tilde{\nabla} \cdot \tilde{\mathbf{u}} = 0. & \text{(3.90c)} \end{cases}$$

On suppose que K est d'ordre 1 par rapport au nombre de Mach M et que l'on a les développements

asymptotiques

$$\begin{cases} \tilde{\rho} = \tilde{\rho}_0 & + M\tilde{\rho}_1 & + M^2\tilde{\rho}_2 & + ... \\ \tilde{\mathbf{u}} = \tilde{\mathbf{u}}_0 & + M\tilde{\mathbf{u}}_1 & + M^2\tilde{\mathbf{u}}_2 & + ... \\ \tilde{p} = \tilde{p}_0 & + M\tilde{p}_1 & + M^2\tilde{p}_2 & + ... \\ \tilde{c} = \tilde{c}_0 & + M\tilde{c}_1 & + M^2\tilde{c}_2 & + ... \end{cases}$$

En faisant tendre M vers 0 dans (3.90b), on obtient à l'ordre -2 et -1 en M

$$\tilde{\nabla}\tilde{p}_0 = 0, \quad \tilde{\nabla}\tilde{p}_1 = 0. \tag{3.91}$$

Le système (3.90) donne ensuite à l'ordre 0 en M

$$\begin{cases} \partial_{\tilde{t}}\tilde{\rho}_0 + \tilde{\nabla}\cdot(\tilde{\rho}_0\tilde{\mathbf{u}}_0) = 0, & \text{(3.92a)} \\ \partial_{\tilde{t}}\tilde{\mathbf{u}}_0 + \left(\tilde{\mathbf{u}}_0\cdot\tilde{\nabla}\right)\tilde{\mathbf{u}}_0 + \dfrac{1}{\tilde{\rho}_0}\tilde{\nabla}\tilde{p}_2 = 0, & \text{(3.92b)} \\ \partial_{\tilde{t}}\tilde{p}_0 + \tilde{\rho}_0\tilde{c}_0^2\tilde{\nabla}\cdot\tilde{\mathbf{u}}_0 = 0. & \text{(3.92c)} \end{cases}$$

Ce système n'est pas fermé à cause de l'inconnue supplémentaire $\tilde{p}_2$. Afin de définir le système limite de (3.90), on suppose que l'écoulement est étudié dans un domaine fini $\mathcal{D}$ pour la variable spatiale. On intègre sur $\mathcal{D}$ l'équation (3.92c)

$$\partial_{\tilde{t}}\tilde{p}_0\int_{\mathcal{D}}\frac{1}{\tilde{\rho}_0\tilde{c}_0^2}d\mathbf{x} + \int_{\mathcal{D}}\tilde{\nabla}\cdot\tilde{\mathbf{u}}_0 d\mathbf{x} = 0.$$

On considère de plus des conditions aux limites périodiques ou de glissement ($\tilde{\mathbf{u}}\cdot\mathbf{n} = 0$, où $\mathbf{n}$ est la normale unitaire sortante au domaine), on obtient $\partial_{\tilde{t}}\tilde{p}_0 = 0$, ainsi $\tilde{p}_0$ est une constante en temps et en espace. D'après (3.92c), on a directement $\tilde{\nabla}\cdot\tilde{\mathbf{u}}_0 = 0$. Le comportement asymptotique quand M tend vers 0 est donc dirigé par le système suivant

$$\begin{cases} \partial_{\tilde{t}}\tilde{\rho}_0 + \tilde{\nabla}\cdot(\tilde{\rho}_0\tilde{\mathbf{u}}_0) = 0, & \text{(3.93a)} \\ \partial_{\tilde{t}}\tilde{\mathbf{u}}_0 + \left(\tilde{\mathbf{u}}_0\cdot\tilde{\nabla}\right)\tilde{\mathbf{u}}_0 + \dfrac{1}{\tilde{\rho}_0}\tilde{\nabla}\tilde{p}_2 = 0, & \text{(3.93b)} \\ \tilde{\nabla}\cdot\tilde{\mathbf{u}}_0 = 0. & \text{(3.93c)} \end{cases}$$

L'inconnue supplémentaire $\tilde{p}_2$ correspond au multiplicateur de Lagrange associé à la contrainte d'incompressibilité (3.93c). En effet, le terme de transport de (3.93b) ne préserve pas l'espace des champs de vitesse à divergence nulle. Le système (3.93) permet alors d'étudier le caractère *asymptotic preserving* des schémas numériques, l'obtention d'un bon équivalent discret de la condition d'incompressibilité (3.93c) jouant un rôle primordial.

Remarque 2. (importante) Le système limite (3.93) **requiert des arguments globaux pour être valide, à cause de l'intégration sur le domaine $\mathcal{D}$ (il faut alors utiliser les conditions aux limites pour fermer le système). En particulier, ce cadre ne permet pas de considérer le cas d'un nombre de Mach variant des petites aux grandes valeurs en fonction de la région du domaine $\mathcal{D}$ dans laquelle on se trouve. C'est pourquoi le système** (3.93) **n'a pas été évoqué dans le cadre de l'article *An all-regime Lagrange-Projection like scheme for the gas dynamics equations on unstructured meshes* où le comportement en fonction du nombre de Mach est étudié à l'aide du système adimensionné** (3.20)**. Plus précisément, on étudie l'ordre de grandeur par rapport au nombre de Mach M de la condition de stabilité CFL et de l'erreur de troncature du schéma numérique.**

3.E Un résultat de stabilité L^2

On présente ici un résultat de stabilité L^2 d'un schéma modifié, en suivant l'approche proposé précédemment dans ce chapitre, pour le p-système barotrope. On considère le p-système barotrope en une dimension d'espace

$$\begin{cases} \partial_t \tau - \partial_m u = 0, & \text{(3.94a)} \\ \partial_t u + \partial_m p(\tau) = 0, & \text{(3.94b)} \end{cases}$$

où la loi de pression est linéarisée $p(\tau) = p_0 + a(\tau_0 - \tau)$, et où p_0, τ_0 et $a > 0$ sont respectivement une pression, un covolume et le carré d'une vitesse du son Lagrangienne de référence.

Le système (3.94) se réécrit

$$\begin{cases} \partial_t \tau - \partial_m u = 0, & \text{(3.95a)} \\ \partial_t u - a\partial_m \tau = 0. & \text{(3.95b)} \end{cases}$$

On obtient en multipliant (3.95a) par $a\tau$ et (3.95b) par u, puis en sommant

$$\partial_t \left(\frac{a\tau^2}{2} + \frac{u^2}{2} \right) - a\partial_m \left(\tau u \right) = 0.$$

On considère des conditions aux limites périodiques et on intègre cette égalité sur le domaine $[m_0, m_L]$ pour obtenir la conservation d'énergie suivante

$$\frac{d}{dt} \left(\int_{m_0}^{m_L} \left(\frac{a\tau^2}{2} + \frac{u^2}{2} \right) dm \right) = 0. \tag{3.96}$$

On essaye de montrer un équivalent de cette égalité d'énergie pour le schéma numérique semi-discret suivant

$$\begin{cases} \dfrac{d\tau_j}{dt} = \dfrac{1}{\Delta m} \left(u^*_{j+\frac{1}{2}} - u^*_{j-\frac{1}{2}} \right), & \text{(3.97a)} \\ \dfrac{du_j}{dt} = -\dfrac{1}{\Delta m} \left(p^{*,\theta}_{j+\frac{1}{2}} - p^{*,\theta}_{j-\frac{1}{2}} \right), & \text{(3.97b)} \end{cases}$$

où

$$\begin{cases} u^*_{j+\frac{1}{2}} = \dfrac{u_{j+1} + u_j}{2} - \dfrac{\Pi_{j+1} - \Pi_j}{2a}, & \text{(3.98a)} \\ p^{*,\theta}_{j+\frac{1}{2}} = \dfrac{\Pi_{j+1} + \Pi_j}{2} - a\theta \left(\dfrac{u_{j+1} - u_j}{2} \right), & \text{(3.98b)} \\ \Pi_j = p(\tau_j) = p_0 + a(\tau_0 - \tau_j). & \text{(3.98c)} \end{cases}$$

La constante θ permet de réduire la diffusion numérique du schéma en régime bas Mach à l'image de ce qui a été proposé précédemment dans ce chapitre.

On injecte (3.98) dans (3.97) pour obtenir

$$\begin{cases} \dfrac{d\tau_j}{dt} = \dfrac{1}{\Delta m} \left(\dfrac{u_{j+1} - u_{j-1}}{2} + \dfrac{\tau_{j+1} - 2\tau_j + \tau_{j-1}}{2} \right), & \text{(3.99a)} \\ \dfrac{du_j}{dt} = \dfrac{a}{\Delta m} \left(\dfrac{\tau_{j+1} - \tau_{j-1}}{2} + \theta \left(\dfrac{u_{j+1} - 2u_j + u_{j-1}}{2} \right) \right). & \text{(3.99b)} \end{cases}$$

On multiplie (3.99a) par $a\tau_j$ et (3.99b) par u_j puis on somme pour obtenir

$$\frac{d}{dt}\left(\frac{a(\tau_j)^2}{2}+\frac{(u_j)^2}{2}\right)=\frac{a\tau_j}{\Delta m}\left(\frac{u_{j+1}-u_{j-1}}{2}\right)+\frac{au_j}{\Delta m}\left(\frac{\tau_{j+1}-\tau_{j-1}}{2}\right)+\frac{a\tau_j}{\Delta m}\left(\frac{\tau_{j+1}-2\tau_j+\tau_{j-1}}{2}\right)$$
$$+\theta\frac{au_j}{\Delta m}\left(\frac{u_{j+1}-2u_j+u_{j-1}}{2}\right).$$

On somme cette égalité sur les cellules du domaine $1\leq j\leq N$ pour obtenir

$$\frac{d}{dt}\left[\sum_{j=1}^{N}\left(\frac{a(\tau_j)^2}{2}+\frac{(u_j)^2}{2}\right)\right]=\sum_{j=1}^{N}\frac{a\tau_j}{\Delta m}\left(\frac{u_{j+1}-u_{j-1}}{2}\right)+\sum_{j=1}^{N}\frac{au_j}{\Delta m}\left(\frac{\tau_{j+1}-\tau_{j-1}}{2}\right)$$
$$+\sum_{j=1}^{N}\frac{a\tau_j}{\Delta m}\left(\frac{\tau_{j+1}-2\tau_j+\tau_{j-1}}{2}\right)+\sum_{j=1}^{N}\theta\frac{au_j}{\Delta m}\left(\frac{u_{j+1}-2u_j+u_{j-1}}{2}\right). \quad (3.100)$$

En considérant des conditions aux limites périodiques, $u_0=u_N$, $\tau_0=\tau_N$, $u_{N+1}=u_1$ et $\tau_{N+1}=\tau_1$, on a l'égalité suivante

$$\sum_{j=1}^{N}\frac{a\tau_j}{\Delta m}\left(\frac{u_{j+1}-u_{j-1}}{2}\right)+\sum_{j=1}^{N}\frac{au_j}{\Delta m}\left(\frac{\tau_{j+1}-\tau_{j-1}}{2}\right)=0.$$

Par ailleurs, on a

$$\sum_{j=1}^{N}\tau_j\left(\tau_{j+1}-2\tau_j+\tau_{j-1}\right)=\sum_{j=1}^{N}\tau_j\left(\tau_{j+1}-\tau_j\right)-\sum_{j=1}^{N}\tau_j\left(\tau_j-\tau_{j-1}\right)=-\sum_{j=1}^{N}\left(\tau_{j+1}-\tau_j\right)^2.$$

De même, on montre que

$$\sum_{j=1}^{N}u_j\left(u_{j+1}-2u_j+u_{j-1}\right)=-\sum_{j=1}^{N}\left(u_{j+1}-u_j\right)^2.$$

En injectant ces trois égalités dans (3.100), on a finalement

$$\frac{d}{dt}\left[\sum_{j=1}^{N}\left(\frac{a(\tau_j)^2}{2}+\frac{(u_j)^2}{2}\right)\right]=-\frac{a}{\Delta m}\sum_{j=1}^{N}\left(\tau_{j+1}-\tau_j\right)^2-\frac{a\theta}{\Delta m}\sum_{j=1}^{N}\left(u_{j+1}-u_j\right)^2 \quad (3.101)$$

Ainsi pour toute valeur positive de la modification $\theta\geq 0$, on a l'inégalité d'énergie suivante

$$\frac{d}{dt}\left[\sum_{j=1}^{N}\left(\frac{a(\tau_j)^2}{2}+\frac{(u_j)^2}{2}\right)\right]\leq 0.$$

Ce résultat de stabilité du schéma (3.97)-(3.98) pour toute valeur de $\theta\geq 0$ est indépendante du régime du nombre de Mach considéré. On peut en particulier choisir $\theta=O(M)$ en régime bas Mach pour obtenir un bon comportement du schéma numérique dans ce régime. Le choix $\theta=0$ est optimal au sens où il permet de réduire le plus possible la diffusion numérique tout en restant stable. On obtient ainsi un schéma numérique anti-diffusif.

L'obtention d'un résultat de stabilité similaire pour le schéma Lagrangien (3.55)-(3.56) est un problème ouvert. Les résultats numériques obtenus dans la section 3.6.2 montrent la stabilité des schémas EX($\theta=0$) et IMEX($\theta=0$) pour les cas tests considérés.

3.F Influence de la forme des cellules du maillage sur les résultats numériques à bas nombre de Mach

Les schémas EX(θ) et IMEX(θ) ont été écrits pour des maillages non structurés, on étudie ici l'influence de la forme du maillage sur le comportement de ces schémas numériques à bas nombre de Mach. On considère en particulier deux types de maillage, les maillages cartésiens et les maillages triangulaires. En se plaçant dans le cas test du vortex dans une boîte, on a vu dans la section 3.6.1 les résultats obtenus sur un maillage cartésien de 2500 cellules. On s'intéresse ici aux résultats obtenus sur un maillage triangulaire de 2450 cellules.

On observe sur la figure 3.F.1 que les résultats obtenus sur un maillage triangulaire avec les schémas EX($\theta = 1$) et EX($\theta = O(M)$) sont de bonnes approximations. On obtient sur la figure 3.F.2 des résultats similaires pour les schémas IMEX($\theta = 1$) et IMEX($\theta = O(M)$). En comparant les résultats des figures 3.F.1 et 3.F.2 obtenus sur des maillages triangulaires à ceux des figures 3.1 et 3.2 obtenus sur des maillages cartésiens, on observe que les schémas EX($\theta = 1$) et IMEX($\theta = 1$) sont nettement plus précis sur maillage triangulaire que sur maillage cartésien à nombre de degré de liberté fixé.

On considère une solution de référence calculée avec le schéma EX($\theta = 1$) sur un maillage triangulaire fin de 40000 cellules. On trace la coupe 1D en $y = 0.5$ des solutions approchées obtenues précédemment. On observe sur la figure 3.F.3 que les schémas EX($\theta = O(M)$) et IMEX($\theta = O(M)$) sont moins diffusifs que les schémas EX($\theta = 1$) et IMEX($\theta = 1$), que ce soit sur maillage cartésien ou triangulaire. Néanmoins, le gain de précision est plus important dans le cas des maillages cartésiens car la solution obtenue pour $\theta = 1$ est très diffusée.

On a donc observé que :

- à bas nombre de Mach, les résultats sont corrects sur maillage triangulaire même sans correction,
- néanmoins la correction permet d'améliorer la précision des solutions approchées,
- sur maillage cartésien il est primordial d'utiliser une correction à bas nombre de Mach pour éviter la forte diffusion numérique dans ce régime.

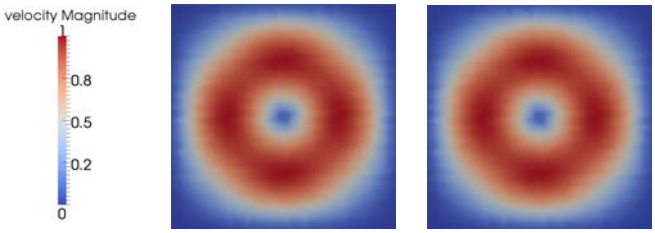


FIGURE 3.F.1 – Cas test du vortex dans une boîte. Profil au temps $t = 0.125\,\mathrm{s}$ de la norme du vecteur vitesse pour les schémas EX($\theta = 1$) (gauche) et EX($\theta = O(M)$) (droite) pour un maillage triangulaire de 2450 cellules.

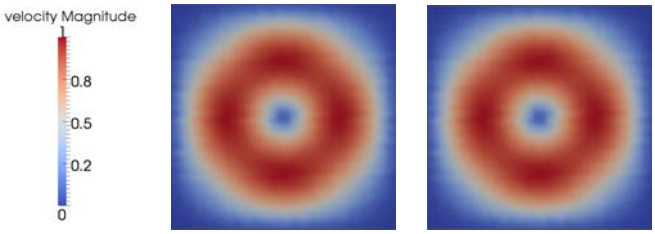


FIGURE 3.F.2 – Cas test du vortex dans une boîte. Profil au temps $t = 0.125\,\mathrm{s}$ de la norme du vecteur vitesse pour les schémas IMEX($\theta = 1$) (gauche) et IMEX($\theta = O(M)$) (droite) pour un maillage triangulaire de 2450 cellules.

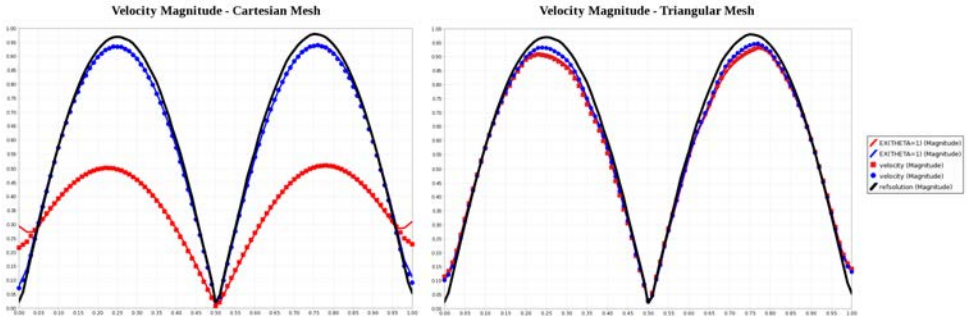


FIGURE 3.F.3 – Cas test du vortex dans une boîte. Profil le long de l'axe $y = 0.5$ au temps $t = 0.125\,\mathrm{s}$ de la norme du vecteur vitesse pour les schémas EX($\theta = 1$), EX($\theta = O(M)$), IMEX($\theta = 1$) et IMEX($\theta = O(M)$) pour un maillage cartésien de 2500 cellules (gauche) et un maillage triangulaire de 2450 cellules (droite), la solution de référence est calculée à l'aide du schéma EX($\theta = 1$) sur un maillage triangulaire de 40000 cellules.

3.G Un schéma en coordonnée Eulérienne avec correction bas Mach

On considère dans cette section un schéma en coordonnée Eulérienne avec correction bas Mach pour le système de la dynamique des gaz 3.1. Ce schéma a été proposé par P.-A. Raviart et s'écrit

$$\begin{cases} \rho_i^{n+1} = \rho_i^n - \Delta t \sum_{j\in N(i)} \frac{\sigma_{ij}}{2} \left[u_{ij}^* \left(\rho_{ij,R} + \rho_{ij,L} \right) - |u_{ij}^*| \left(\rho_{ij,R} - \rho_{ij,L} \right) \right], & (3.102a) \\ (\rho \mathbf{u})_i^{n+1} = (\rho \mathbf{u})_i^n - \Delta t \sum_{j\in N(i)} \frac{\sigma_{ij}}{2} \left[u_{ij}^* \left((\rho \mathbf{u})_{ij,R} + (\rho \mathbf{u})_{ij,L} \right) - |u_{ij}^*| \left((\rho \mathbf{u})_{ij,R} - (\rho \mathbf{u})_{ij,L} \right) \right] \\ \qquad - \Delta t \sum_{j\in N(i)} \sigma_{ij} p_{ij}^* \mathbf{n}_{ij}, & (3.102b) \\ (\rho E)_i^{n+1} = (\rho E)_i^n - \Delta t \sum_{j\in N(i)} \frac{\sigma_{ij}}{2} \left[u_{ij}^* \left((\rho E)_{ij,R} + (\rho E)_{ij,L} \right) - |u_{ij}^*| \left((\rho E)_{ij,R} - (\rho E)_{ij,L} \right) \right] \\ \qquad - \Delta t \sum_{j\in N(i)} \sigma_{ij} p_{ij}^* u_{ij}^*, & (3.102c) \end{cases}$$

avec

$$\begin{cases} a_{ij} \geq \max[(\rho c)_i^n, (\rho c)_j^n], \quad p_i^n = p(\rho_i^n, e_i^n), & (3.103a) \\ u_{ij}^* = \frac{1}{2} \mathbf{n}_{ij}^T (\mathbf{u}_i^n + \mathbf{u}_j^n) - \frac{1}{2a_{ij}} (p_j^n - p_i^n), & (3.103b) \\ p_{ij}^* = \frac{1}{2} (p_i^n + p_j^n) - \theta_{ij} \frac{a_{ij}}{2} \mathbf{n}_{ij}^T (\mathbf{u}_j^n - \mathbf{u}_i^n), & (3.103c) \\ \rho_{ij,L} = \rho_i^n \left(1 - \frac{\rho_i^n}{2(a_{ij})^2} \left(p_j^n - p_i^n \right) - \theta_{ij} \frac{\rho_i^n}{2a_{ij}} \mathbf{n}_{ij}^T \left(\mathbf{u}_j^n - \mathbf{u}_i^n \right) \right)^{-1}, & (3.103d) \\ \rho_{ij,R} = \rho_j^n \left(1 + \frac{\rho_j^n}{2(a_{ij})^2} \left(p_j^n - p_i^n \right) - \theta_{ij} \frac{\rho_j^n}{2a_{ij}} \mathbf{n}_{ij}^T \left(\mathbf{u}_j^n - \mathbf{u}_i^n \right) \right)^{-1}, & (3.103e) \\ \mathbf{u}_{ij,L} = u_{ij}^* \mathbf{n}_{ij} + \left(\mathbf{t}_{ij}^T \mathbf{u}_i^n \right) \mathbf{t}_{ij}, \quad \mathbf{u}_{ij,R} = u_{ij}^* \mathbf{n}_{ij} + \left(\mathbf{t}_{ij}^T \mathbf{u}_j^n \right) \mathbf{t}_{ij}, & (3.103f) \\ E_{ij,L} = E_i^n - \frac{1}{a_{ij}} \left(p_{ij}^* u_{ij}^* - p_i^n \left(\mathbf{n}_{ij}^T \mathbf{u}_i^n \right) \right), \quad E_{ij,R} = E_j^n + \frac{1}{a_{ij}} \left(p_{ij}^* u_{ij}^* - p_j^n \left(\mathbf{n}_{ij}^T \mathbf{u}_j^n \right) \right), & (3.103g) \end{cases}$$

où le vecteur $\mathbf{t}_{ij}$ est unitaire et tangent au côté Γ_{ij}. On note RELAX(θ) le schéma numérique composé de (3.102)-(3.103) et on considère la condition CFL suivante

$$\Delta t \max_{1\leq i\leq N} \left[\max_{j\in N(i)} \sigma_{ij} \left(|\mathbf{n}_{ij}^T \mathbf{u}_i^n| + c_i^n \right) \right] \leq \frac{1}{2}. \tag{3.104}$$

Cette condition sur le pas de temps Δt pour le schéma explicite (3.102)-(3.103) est dirigée par la vitesse du son et donc par la vitesse des ondes (rapides) acoustiques. La construction d'une version implicite ou semi-implicite de ce schéma afin d'obtenir une condition CFL basée sur la vitesse des ondes (lentes) matières peut s'avérer délicate en raison de la non linéarité du schéma (3.102)-(3.103). On comprend ainsi l'utilité du *splitting* d'opérateurs introduit dans la section 3.5.4 et qui permet de se ramener aux coordonnées Lagrangiennes pour construire le schéma semi-implicite LSP-IMEX(θ).

La modification θ_{ij} correspond à une correction du schéma numérique qui permet d'obtenir un bon comportement en régime bas Mach. En effet, des calculs non présentés dans ce manuscrit montrent que le schéma RELAX(θ) est *asymptotic preserving* si $\theta = O(M)$. Par ailleurs, on retrouve les flux du schéma de relaxation classique en coordonnée Eulérienne pour $\theta_{ij} = 1$.

On s'intéresse aux résultats obtenus, pour le cas test du vortex dans une boîte, avec les schémas

RELAX($\theta = 1$) et RELAX($\theta = O(M)$) qui correspondent respectivement aux choix $\theta_{ij} = 1$ et $\theta_{ij} = \min\left(|u^*_{ij}|/\max(c^n_i, c^n_j), 1\right)$. On observe sur la figure 3.G.1 que le schéma RELAX($\theta = O(M)$) est beaucoup moins diffusif que le schéma RELAX($\theta = 1$). Ainsi, la modification $\theta = O(M)$ permet bien d'améliorer nettement la précision du schéma dans le régime des faibles nombres de Mach.

Néanmoins, pour des valeurs de θ trop faibles en régime intermédiaire, le schéma RELAX(θ) peut faire apparaître des oscillations ou bien encore être instable. Pour le cas test du tube à choc de Sod, on observe que le schéma RELAX(θ) est instable pour $\theta = O(M)$ ainsi que pour $\theta = \theta_0$ avec $\theta_0 \leq 0.52$. De plus, pour $\theta = \theta_0$ avec $0.52 \leq \theta_0 \leq 0.63$, les résultats obtenus avec le schéma RELAX(θ) présentent des oscillations. En effet, on observe sur la figure 3.G.2 que les résultats obtenus avec le schéma RELAX($\theta = 1$) sont bons, tandis que ceux obtenus avec le schéma RELAX($\theta = 0.53$) présentent des oscillations.

Étudier la stabilité du schéma RELAX(θ) pour $\theta \neq 1$ est un problème ouvert qui permettrait d'obtenir un critère sur le choix de θ garantissant de bons résultats numériques quel que soit le régime du nombre de Mach considéré. Il est à noter que les schémas numériques LSP-EX(θ) et LSP-IMEX(θ) semblent être plus robustes vis-à-vis du choix de la modification θ que le schéma RELAX(θ), grâce à la diffusion numérique de l'étape de transport qui ne dépend pas de θ.

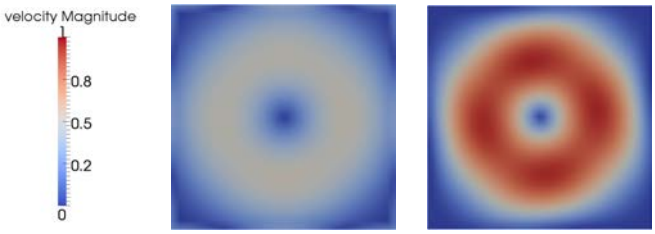


FIGURE 3.G.1 – Cas test du vortex dans une boîte. Profil au temps $t = 0.125\,\mathrm{s}$ de la norme du vecteur vitesse pour les schémas RELAX($\theta = 1$) (gauche) et RELAX($\theta = O(M)$) (droite) pour un maillage cartésien de 2500 cellules.

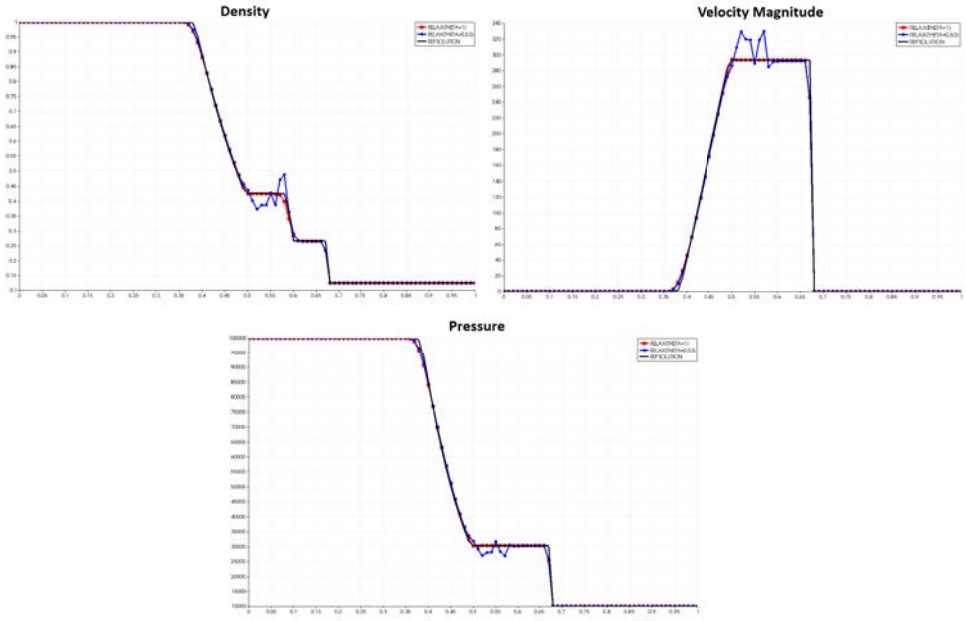


FIGURE 3.G.2 – Cas test du tube à choc de Sod. Profil au temps $t = 3.1 \times 10^{-4}\,\mathrm{s}$ de la densité (en haut gauche), de la norme du vecteur vitesse (en haut droit) et de la pression (en bas) pour les schémas RELAX($\theta = 1$) et RELAX($\theta = 0.53$) pour un maillage de 1000 cellules, ainsi qu'une solution de référence.

Bibliographie

[1] F. Bouchut, Nonlinear stability of finite volume methods for hyperbolic conservation laws, and well-balanced schemes for sources, Frontiers in Mathematics series, Birkhäuser, (2004).

[2] C. Chalons and F. Coquel, *Navier-stokes equations with several independant pressure laws and explicit predictor-corrector schemes*, Numerisch Math, 101(3) : 451–478, (2005).

[3] C. Chalons and J.-F. Coulombel, *Relaxation approximation of the Euler equations*,J. Math. Anal. Appl., 348(2) : 872–893, (2008).

[4] C. Chalons, M. Girardin, and S. Kokh, *Large time step and asymptotic preserving numerical schemes for the gas dynamics equations with source terms*, SIAM J. Sci. Comput., 35(6) : a2874–a2902, (2013).

[5] P. Colella and K. Pao, *A projection method for low speed flows*, J. Comp. Phys., 149(2) : 245–269, (1999).

[6] F. Coquel, Q. L. Nguyen, M. Postel, and Q. H. Tran, *Entropy-satisfying relaxation method with large time-steps for Euler IBVPs*, Math. Comput., 79(271) : 1493–1533, (2010).

[7] F. Cordier, P. Degond and A. Kumbaro, *An Asymptotic-Preserving all-speed scheme for the Euler and NavierÛStokes equations*, J. Comp. Phys., 231(17) : 5685–5704, (2012).

[8] P. Degond and M. Tang, *All speed method for the Euler equation in the low mach number limit*, Communications in Computational Physics, 10 : 1–31, (2011).

[9] P. Degond, S. Jin and J.-G. Liu, *Mach-number uniform asymptotic-preserving gauge schemes for compressible flows*, Bull. Inst. Math., Acad. Sin. (N.S.), 2(4) : 851–892, (2007).

[10] S. Dellacherie, P. Omnes and P.A. Raviart, *Construction of modified Godunov type schemes accurate at any Mach number for the compressible Euler system*, submitted, (2013).

[11] S. Dellacherie, *Analysis of Godunov type schemes applied to the compressible euler system at low Mach number*, J. Comp. Phys., 229(4) : 978–1016, (2010).

[12] B. Després, *Inégalité entropique pour un solveur conservatif du système de la dynamique des gaz en coordonnées de lagrange*, C. R. Acad. Sci. Paris, Série I, 324 : 1301–1306, (1997).

[13] B. Després, Lois de conservations Eulériennes, Lagrangiennes et méthodes numériques, volume 68 of Mathématiques et applications, SMAI, Springer, (2010).

[14] B. Després, E. Labourasse, F. Lagoutière, and I. Marmajou, *An antidissipative transport scheme on unstructured meshes for multicomponent flows*, Int. J. Finite. Vol. Meth., 7 : 30–65, (2010).

[15] F. Dauvergne, J.-M. Ghidaglia, F. Pascal, and J.-M. Rovarch, *Renormalization of the numerical diffusion for an upwind finite volume method. application to the simulation of Kelvin-Helmholtz instability*, Finite volumes for complex applications. V. Proceedings of the 5th International Symposium, Aussois, June 2008, R. Eymard and J.-M. Hérard editors, 321–328, (2008).

[16] E. Godlewski and P.-A. Raviart, Numerical approximation of hyperbolic systems of conservation laws, Springer, (1996).

[17] H. Guillard and C. Viozat, *On the behavior of upwind schemes in the low Mach limit*, Comp. & Fluid, 28 : 63–86, (1999).

[18] A. Harten, P.D. Lax, and B. Van Leer, *On upstream differencing and godunov-type schemes for hyperbolic conservation laws*, SIAM Review, 25 : 35–61, (1983).

[19] I. Toumi, A. Kumbaro, and H. Paillere, *Approximate Riemann solvers and flux vector splitting schemes for two-phase flow*, In VKI LS 1999-03, Computational Fluid Dynamics, (1999).

[20] J.-G. Liu J. Haack, S. Jin, *An all-speed asymptotic-preserving method for the isentropic Euler and Navier-Stokes equations*, Commun. Comp. Phys., 12 : 955–980, (2012).

[21] S. Jin, *Runge-Kutta methods for hyperbolic conservation laws with stiff relaxation terms*, J. Comp. Phys., 122(1) : 51–67, (1995).

[22] R. Klein, *Semi-implicit extension of a Godunov-type scheme based on low Mach number asymptotics I : One-dimensional flow*, J. Comp. Phys., 121(2) : 213–237, (1995).

[23] P. D. Lax and X.-D. Liu, *Solution of two-dimensional Riemann problems of gas dynamics by positive schemes*, SIAM J. Sci. Comput., 19(2) : 319–340, (1998).

[24] M.-S. Liou, *A sequel to AUSM, part ii : AUSM+-up for all speeds*, J. Comp. Phys., 214(1) : 137–170, (2006).

[25] H. Paillère, C. Viozat, A. Kumbaro, and I. Toumi, *Comparison of low mach number models for natural convection problems*, Heat and Mass Transfer, 36(6) : 567–573, (2000).

[26] S. Schochet, *Fast singular limits of hyperbolic PDEs*, J. Differ. Equations, 114(2) : 476–512, (1994).

[27] G. A. Sod, Numerical methods in fluid dynamics, Initial and initial-boundary value problems, Cambridge : Cambridge University Press, (1985).

[28] I. Suliciu, *On the thermodynamics of rate-type fluids and phase transitions. i. rate-type fluids*, International Journal of Engineering Science, 36(9) : 921–947, (1998).

[29] B. Thornber, A. Mosedale, D. Drikakis, D. Youngs, and R.J.R. Williams, *An improved reconstruction method for compressible flows with low Mach number features*, J. Comp. Phys., 227(10) : 4873–4894, (2008).

[30] E. Turkel, *Preconditioned methods for solving the incompressible and low speed compressible equations*, J. Comp. Phys., 72(2) : 277–298, (1987).

[31] H. Weyl, *Shock waves in arbitrary fluids*, Commun. Pure Appl. Math., 2 : 103–122, (1949).

Chapitre 4

Schémas Lagrange-Projection tout-régime pour les modèles diphasiques homogénéisés HEM et HRM sur maillage non structuré

Ce chapitre a fait l'objet d'un article soumis : C. Chalons, M. Girardin and S. Kokh, *An all-regime Lagrange-Projection like scheme for 2D homogeneous models for two-phase flows on unstructured meshes.*

Dans les chapitres précédents, on a considéré des systèmes modélisant l'écoulement d'un seul constituant. Dans ce chapitre, on s'intéresse aux modèles diphasiques homogénéisés HEM et HRM qui permettent de modéliser un mélange de deux constituants. On va généraliser pour ces modèles diphasiques la méthodes de construction de schéma numérique et les propriétés obtenues dans le cadre monophasique du chapitre 3.

Les annexes viennent compléter l'article *An all-regime Lagrange-Projection like scheme for 2D homogeneous models for two-phase flows on unstructured meshes*, en apportant un éclairage particulier sur l'obtention d'un *solver* de Riemann approché pour le schéma acoustique modifié dans la section (4.A) et la preuve d'une inégalité d'entropie discrète multi-d sur maillage non structuré dans la section (4.B). La démonstration de cette inégalité d'entropie discrète peut être facilement adaptée au schéma proposé dans le chapitre 3 pour le système de la dynamique des gaz, généralisant ainsi le résultat obtenu en une dimension d'espace au cours de ce chapitre.

An all-regime Lagrange-Projection like scheme for 2D homogeneous models for two-phase flows on unstructured meshes

Abstract

We propose an *all regime* Lagrange-Projection like numerical scheme for 2D homogeneous models for two-phase flows. By *all regime*, we mean that the numerical scheme is able to compute accurate approximate solutions with an under-resolved discretization, i.e. a mesh size and time step much bigger than the Mach number M of the mixture. The key idea is to decouple acoustic, transport and phase transition phenomenon using a Lagrange-Projection decomposition in order to treat implicitly (fast) acoustic and phase transition phenomenon and explicitly the (slow) transport phenomena. Then, extending a strategy developed in the case of the usual gas dynamics equations, we alter the numerical flux in the acoustic approximation to obtain an uniform truncation error in term of M. This modified scheme is conservative and endowed with good stability properties with respect to the positivity of the density and preserving the mass fraction within the interval $(0, 1)$. Numerical evidences are proposed and show the ability of the scheme to deal with tests where the flow regime may vary from low to high Mach values.

4.1 Introduction

We are interested in the simulation of two-phase flows in situations where the flow regime may vary in terms of Mach number M across the computational domain. Among the numerous models that describe two-phase flows we consider here an Homogeneous Relaxation Model frequently referred to as HRM and its related Homogeneous Equilibrium Model HEM, see [4, 16, 10, 1] and references therein. We propose a collocated Finite Volume method that addresses two important issues.

The first issue concerns the lack of accuracy in the low-Mach regime of Godunov-type schemes when using an under-resolved mesh. This problem has been widely investigated in the case of the gas dynamics equations, see [23, 18, 13, 15, 14, 12, 21, 9, 3, 8]. The analysis of these authors may rely on different arguments like the analysis of the viscosity matrix [23], an asymptotic expansion in terms of Mach number [18], a detailed study in [15] that seek for invariance properties of the numerical scheme transposing the framework of Schochet to the discrete setting, and also an analysis based on the so-called Asymptotic Preserving property in [21]. Nevertheless the resulting cure usually boils down to reduce the numerical diffusion in the momentum equation for low Mach number values. Some works have been devoted to the extension of those strategies to two-phase flows [10, 2, 17, 20].

The second problem we address is the CFL restriction on the time step for explicit Finite Volume methods that involve the (fast) acoustic and phase transition phenomena. It seems natural to seek for numerical schemes that enable the use of large time steps constrained only by the (slow) material phenomena, see [13, 11, 21, 7, 8].

Numerical schemes that can tackle both issues, namely : accuracy for mesh sizes that do not depend on the Mach number and also stability for time steps that are not constrained by the Mach value are usually referred to as *all regime*, like the methods proposed in [13, 21, 8].

In the present work, we propose an extension of the method proposed in [8] for the gas dynamics equations to the case of homogeneous models for two-phase flows. An operator splitting strategy allows to decouple the acoustic, transport and phase transition phenomena. The approximation algorithm is split into three steps : an acoustic step, a transport step and a phase transition step. A mixed implicit-explicit method is obtained by using implicit updates for the acoustic and phase transition steps, and an

explicit update for the transport step. Then, a modification of fluxes for the acoustic step allow to recover a truncation error that is uniform with respect to the Mach number. The resulting scheme allows to cope with unstructured meshes and compressible flows equipped with very general Equation of State (EOS). Finally, let us mention that the overall procedure is shown to be a conservative discretization (except for the mass fraction due to phase transition) and endowed with good stability properties with respect to the positivity of the density and ensuring that the mass fraction remains in the interval $(0, 1)$. We also prove the validity of a discrete entropy inequality in 2D and for general meshes.

4.2 Governing equations and low-Mach number regime

Governing equations. We are interested in the two-dimensional homogeneous relaxation model (HRM)

$$\left\{\begin{array}{rcl} \partial_t(\rho Y) + \nabla\cdot(\rho Y\mathbf{u}) &=& \lambda_0\rho\left(Y^*(\rho,e)-Y\right), \\ \partial_t\rho + \nabla\cdot(\rho\mathbf{u}) &=& 0, \\ \partial_t(\rho\mathbf{u}) + \nabla\cdot(\rho\mathbf{u}\otimes\mathbf{u}) + \nabla p &=& 0, \\ \partial_t(\rho E) + \nabla\cdot\left[(\rho E + p)\,\mathbf{u}\right] &=& 0, \end{array}\right. \tag{4.1}$$

where Y, ρ, $\mathbf{u} = (u, v)^t$, E denote the mass fraction, the density, the velocity vector and the total energy of the mixture. The pressure $p = p(\rho, e, Y)$ is assumed to be a given function of the density ρ, the internal energy $e = E - \frac{|\mathbf{u}|^2}{2}$ of the mixture and the mass fraction Y. The mass fraction at thermodynamic equilibrium $Y^*(\rho, e)$ is a given function of the density and the internal energy of the mixture. For HRM, the thermodynamic equilibrium $Y = Y^*(\rho, e)$ is not instantaneously achieved but is reached at speed $\lambda_0 > 0$. We refer for instance the reader to [4, 16, 10, 1] and the references therein.

Remark 7. *We note that in the limit $\lambda_0 \to \infty$, HRM converges at least formally toward the homogeneous equilibrium model (HEM) given by*

$$\left\{\begin{array}{rcl} Y &=& Y^*(\rho,e), \\ \partial_t\rho + \nabla\cdot(\rho\mathbf{u}) &=& 0, \\ \partial_t(\rho\mathbf{u}) + \nabla\cdot(\rho\mathbf{u}\otimes\mathbf{u}) + \nabla p^{HEM} &=& 0, \\ \partial_t(\rho E) + \nabla\cdot\left[\left(\rho E + p^{HEM}\right)\mathbf{u}\right] &=& 0, \end{array}\right. \tag{4.2}$$

where $p^{HEM}(\rho, e) = p(\rho, e, Y^*(\rho, e))$.

Dimensionless governing equations. We are now interested in the behavior of the HRM system with respect to the variation of the Mach regime. In order to characterize the Mach regime of the flow, we consider a rescaling of the equations. Let us introduce the following non-dimensional quantities :

$$\tilde{x} = \frac{x}{L}, \quad \tilde{t} = \frac{t}{T}, \quad \tilde{\rho} = \frac{\rho}{\rho_0}, \quad \tilde{u} = \frac{u}{u_0}, \quad \tilde{v} = \frac{v}{v_0}, \quad \tilde{e} = \frac{e}{e_0}, \quad \tilde{p} = \frac{p}{p_0}, \quad \tilde{c} = \frac{c}{c_0}$$

The parameters L, T, $u_0 = v_0 = \frac{L}{T}$, ρ_0, $e_0 = \rho_0 p_0$ and $c_0 = \sqrt{\frac{p_0}{\rho_0}}$ denote respectively a characteristic length, time, velocity, density, internal energy, pressure, and sound speed of the mixture. If $M = \frac{u_0}{c_0}$ is the so-called Mach-number then system (4.1) reads

$$\left\{\begin{array}{rcl} \partial_{\tilde{t}}(\tilde{\rho}Y) + \tilde{\nabla}\cdot(\tilde{\rho}Y\tilde{\mathbf{u}}) - \lambda_0 T\tilde{\rho}\left(Y^*(\rho_0\tilde{\rho}, e_0\tilde{e}) - Y\right) &=& 0, \\ \partial_{\tilde{t}}\tilde{\rho} + \tilde{\nabla}\cdot(\tilde{\rho}\tilde{\mathbf{u}}) &=& 0, \\ \partial_{\tilde{t}}(\tilde{\rho}\tilde{\mathbf{u}}) + \tilde{\nabla}\cdot(\tilde{\rho}\tilde{\mathbf{u}}\otimes\tilde{\mathbf{u}}) + \frac{1}{M^2}\tilde{\nabla}\tilde{p} &=& 0, \\ \partial_{\tilde{t}}(\tilde{\rho}\tilde{e}) + \tilde{\nabla}\cdot[(\tilde{\rho}\tilde{e} + \tilde{p})\tilde{\mathbf{u}}] + \frac{M^2}{2}\left[\partial_{\tilde{t}}(\tilde{\rho}|\tilde{\mathbf{u}}|^2) + \tilde{\nabla}\cdot(\tilde{\rho}|\tilde{\mathbf{u}}|^2\tilde{\mathbf{u}})\right] &=& 0, \end{array}\right. \tag{4.3}$$

This system motivates the following definition.

Definition 6. *In the following, the flow is said to be in the low-Mach regime if and only if the Mach number $M \ll 1$ and $\tilde{\nabla}\tilde{p} = \mathcal{O}(M^2)$.*

Remark 8. *If the term $\tilde{\nabla}\tilde{p}$ does not remain of magnitude $\mathcal{O}(M^2)$ then the variation of $\tilde{\rho}\tilde{u}$ will reach a magnitude $\mathcal{O}\left(\frac{1}{M}\right)$ or $\mathcal{O}\left(\frac{1}{M^2}\right)$. These large magnitude variations of the momentum will induce a growth of the Mach number and thus change the Mach regime.*

Remark 9. *The source term in the mass fraction equation may be stiff if the relaxation toward thermodynamic equilibrium is much faster than the convective part of the system ($\lambda_0 T \gg 1$).*

4.3 Acoustic-transport-phase transition operator splitting strategy

In this section, we propose a three-step approximation strategy based on an operator splitting for approximating the solutions of (4.1). The aim of this splitting is to decouple acoustic, transport and phase transition phenomena. Using the chain rule for the space derivatives, we split system (4.1) into the following three subsystems. The first subsystem describe the transport process and reads

$$\left\{\begin{array}{lll} \partial_t(\rho Y) & + \mathbf{u}\cdot\nabla(\rho Y) & = 0, \\ \partial_t\rho & + \mathbf{u}\cdot\nabla(\rho) & = 0, \\ \partial_t(\rho\mathbf{u}) & + \mathbf{u}\cdot\nabla(\rho\mathbf{u}) & = 0, \\ \partial_t(\rho E) & + \mathbf{u}\cdot\nabla\big(\rho E\big) & = 0. \end{array}\right. \tag{4.4}$$

The second subsystem governs the acoustic phenomena, namely

$$\begin{array}{rr} \partial_t(\rho Y) + (\rho Y)\nabla\cdot\mathbf{u} = 0, & \partial_t\rho + \rho\nabla\cdot\mathbf{u} = 0, \\ \partial_t(\rho\mathbf{u}) + (\rho\mathbf{u})\nabla\cdot\mathbf{u} + \nabla p = 0, & \partial_t(\rho E) + (\rho E)\nabla\cdot\mathbf{u} + \nabla\cdot[p\mathbf{u}] = 0, \end{array}$$

Or equivalently with $\tau = \frac{1}{\rho}$ the specific volume

$$\left\{\begin{array}{rcl} \partial_t Y & = & 0, \\ \partial_t\tau - \tau\nabla\cdot\mathbf{u} & = & 0, \\ \partial_t\mathbf{u} + \tau\nabla p & = & 0, \\ \partial_t E + \tau\nabla\cdot(p\mathbf{u}) & = & 0. \end{array}\right. \tag{4.5}$$

This system is nothing but the gas dynamics equations in Lagrangian coordinates, so that the proposed transport-acoustic decomposition is nothing but the natural (and physically relevant) Lagrange-Projection strategy. This is an original approach for treating low Mach regimes that was first proposed in [8].
The third subsystem accounts for mass transfer between phases and reads

$$\left\{\begin{array}{rcl} \partial_t(\rho Y) & = & \lambda_0\rho\left(Y^*(\rho,e) - Y\right), \\ \partial_t\rho & = & 0, \\ \partial_t(\rho\mathbf{u}) & = & 0, \\ \partial_t(\rho E) & = & 0, \end{array}\right.$$

or equivalently

$$\left\{\begin{array}{rcl} \partial_t Y &=& \lambda_0 \left(Y^*(\rho,e) - Y\right), \\ \partial_t \rho &=& 0, \\ \partial_t(\rho \mathbf{u}) &=& 0, \\ \partial_t(\rho E) &=& 0. \end{array}\right. \tag{4.6}$$

Let us mention that this transport/acoustic/phase transition splitting separates physical phenomena that happen at speed $u_0/c_0/\lambda_0$ that may differ from several order of magnitude. From a numerical point of view, such a decomposition is very helpful to design large time step implicit-explicit strategy with CFL restriction based only on the slow phenomenon [11, 7, 8].

4.4 Numerical scheme

Let us suppose that the domain $\Omega \subset \mathbb{R}^2$ is discretized by N cells Ω_i. Let Γ_{ij} be the common edge of two neighbouring cells Ω_i and Ω_j and $\mathbf{n}_{ij}$ be the unit vector normal to Γ_{ij} pointing from Ω_i to Ω_j. We define $N(i)$ the set of indices $1 \leq j \leq N$ such that Ω_i and Ω_j have a common face. Let $\Delta t > 0$ be the time step, we define the intermediate times $t^n = n\Delta t$ for $n \in \mathbb{N}$. If b is a fluid parameter, in the sequel, we will note b_i^n (resp. b_i^{n+1}) the approximate value of b within the cell Ω_i at instant $t = t^n$ (resp. $t = t^{n+1}$).

Given a fluid state $(Y, \rho, \rho u, \rho v, \rho E)_i^n$, $1 \leq i \leq N$ at instant t^n, this splitting algorithm may be decomposed as follows

1. Acoustic step : Update the fluid state $(Y, \rho, \rho u, \rho v, \rho E)_i^n$ to the value $(Y, \rho, \rho u, \rho v, \rho E)_i^{n+1-}$ by approximating the solution of (4.5) ;
2. Transport step : Update the fluid state $(Y, \rho, \rho u, \rho v, \rho E)_i^{n+1-}$ to the value $(\overline{Y}, \overline{\rho}, \overline{\rho u}, \overline{\rho v}, \overline{\rho E})_i$ by approximating the solution of (4.4) ;
3. Phase transition step : Update the fluid state $(\overline{Y}, \overline{\rho}, \overline{\rho u}, \overline{\rho v}, \overline{\rho E})_i$ to the value $(Y, \rho, \rho u, \rho v, \rho E)_i^{n+1}$ by approximating the solution of (4.6).

Let us enter the details of each step.

Acoustic step (Lagrange step). Regarding the acoustic step (4.5) we propose the following update formulas

$$\left\{\begin{array}{rcl} \mathbf{u}_i^{n+1-} &=& \mathbf{u}_i^n - \tau_i^n \Delta t \sum\limits_{j \in N(i)} \sigma_{ij} p_{ij}^{*,\theta} \mathbf{n}_{ij}, \\ \Pi_i^{n+1-} &=& \Pi_i^n - \tau_i^n \Delta t \sum\limits_{j \in N(i)} \sigma_{ij} (a_{ij})^2 u_{ij}^*, \\ Y_i^{n+1-} &=& Y_i^n, \\ \tau_i^{n+1-} &=& \tau_i^n + \tau_i^n \Delta t \sum\limits_{j \in N(i)} \sigma_{ij} u_{ij}^*, \\ E_i^{n+1-} &=& E_i^n - \tau_i^n \Delta t \sum\limits_{j \in N(i)} \sigma_{ij} p_{ij}^{*,\theta} u_{ij}^*, \end{array}\right. \tag{4.7}$$

where $\sigma_{ij} = |\Gamma_{ij}|/|\Omega_i|$ and Π is an unknown associated with the so-called Suliciu relaxation approximation and given at time t^n by $\Pi_i^n = p(\rho_i^n, e_i^n, Y_i^n)$, see [22, 5, 6].
The three scalar quantities a_{ij}, $p_{ij}^{*,\theta}$, and u_{ij}^* represent respectively an average sound velocity, a pressure and normal velocity at the face Γ_{ij} and are given by

$$\left\{\begin{array}{rcl} a_{ij} &=& \max\left((\rho c)_i^n, (\rho c)_j^n\right), \\ u_{ij}^* &=& \frac{1}{2}\mathbf{n}_{ij}^T\left(\mathbf{u}_i^n + \mathbf{u}_j^n\right) - \frac{1}{2a_{ij}}\left(\Pi_j^n - \Pi_i^n\right), \\ p_{ij}^{*,\theta} &=& \frac{1}{2}\left(\Pi_i^n + \Pi_j^n\right) - \frac{a_{ij}\theta_{ij}}{2}\mathbf{n}_{ij}^T\left(\mathbf{u}_j^n - \mathbf{u}_i^n\right). \end{array}\right. \tag{4.8}$$

Remark that the modification of classical fluxes thanks to θ_{ij} will allow to avoid spurious numerical diffusion in the low Mach regime, as proposed recently in [8] such a modification is the key point to make the scheme accurate in the low Mach regime. The classical Suliciu relaxation fluxes correspond to the choice $\theta_{ij} = 1$.
At this stage, the CFL restriction of this *explicit* scheme is based on the (fast) acoustic waves and reads

$$\Delta t \max_{1\leq j\leq N}\left[\tau_j^n\left(\max_{i\in N(j)}\sigma_{ij}a_{ij}\right)\right] \leq \frac{1}{2}. \tag{4.9}$$

To obtain a time step definition based only on the slow waves, following ideas developed in [11, 7, 8], we propose to use an implicit scheme for the acoustic step. We use (4.7) with a new definition of the pressure and normal velocity at the interface Γ_{ij} given by

$$\left\{\begin{array}{lcl} a_{ij} & = & \max\left((\rho c)_i^n,(\rho c)_j^n\right), \\ u_{ij}^* & = & \frac{1}{2}\mathbf{n}_{ij}^T\left(\mathbf{u}_i^{n+1-}+\mathbf{u}_j^{n+1-}\right)-\frac{1}{2a_{ij}}\left(\Pi_j^{n+1-}-\Pi_i^{n+1-}\right), \\ p_{ij}^{*,\theta} & = & \frac{1}{2}\left(\Pi_i^{n+1-}+\Pi_j^{n+1-}\right)-\frac{a_{ij}\theta_{ij}}{2}\mathbf{n}_{ij}^T\left(\mathbf{u}_j^{n+1-}-\mathbf{u}_i^{n+1-}\right). \end{array}\right. \tag{4.10}$$

Thanks to the Suliciu-type relaxation strategy, scheme (4.7)-(4.10) is valid for any pressure law and only requires to solve a linear problem with respect to variables $\mathbf{u}$ and Π. Then other update formulas for variables Y, τ and E are evaluated explicitly, while the scheme is actually implicit.

Transport step (Projection step). In order to approximate the solutions of (4.4), we simply use an upwind Finite-Volume scheme : Let $\varphi \in \{\rho Y, \rho, \rho\mathbf{u}, \rho E\}$, we set

$$\overline{\varphi}_i = \varphi_i^{n+1-} - \Delta t\left[\sum_{j\in N(i)}(\sigma_{ij}u_{ij}^*\varphi_{ij}^{n+1-})\right] + \Delta t\varphi_i^{n+1-}\left[\sum_{j\in N(i)}(\sigma_{ij}u_{ij}^*)\right], \tag{4.11}$$

where φ_{ij}^{n+1-} is defined by the upwind choice with respect to the sign of u_{ij}^*, namely

$$\varphi_{ij}^{n+1-} = \left\{\begin{array}{ll} \varphi_i^{n+1-}, & \text{if } u_{ij}^* > 0, \\ \varphi_j^{n+1-}, & \text{if } u_{ij}^* \leq 0. \end{array}\right. \tag{4.12}$$

The CFL restriction of this explicit scheme is based on the (slow) material waves and reads

$$\Delta t \max_{1\leq j\leq N}\left[\sum_{i\in N(j)}(\sigma_{ij}|u_{ij}^*|)\right] \leq 1. \tag{4.13}$$

Phase transition step (Source terms step). To approximate system (4.6), we propose a pointwise implicit evaluation

$$\left\{\begin{array}{lcl} Y_i^{n+1} & = & \overline{Y}_i + \lambda_0\Delta t\left(Y^*(\overline{\rho}_i,\overline{e}_i)-Y_i^{n+1}\right), \\ \varphi_i^{n+1} & = & \overline{\varphi}_i, \quad \varphi\in\{\rho,\rho\mathbf{u},\rho E\}. \end{array}\right. \tag{4.14}$$

The implicit treatment is particularly important for large values of λ_0 to avoid a CFL restriction based on the (fast) phase transition phenomenon.

Remark 10. *In the limit $\lambda_0 \to +\infty$, this step may be replaced by the projection on the thermodynamic equilibrium*

$$\left\{\begin{array}{lcl} Y_i^{n+1} & = & Y^*(\overline{\rho}_i,\overline{e}_i), \\ \varphi_i^{n+1} & = & \overline{\varphi}_i, \quad \varphi\in\{\rho,\rho\mathbf{u},\rho E\}. \end{array}\right. \tag{4.15}$$

to obtain a numerical scheme for the HEM system (4.2).

Overall numerical scheme. The overall numerical scheme composed by the discretization (4.7)-(4.11)-(4.14) is conservative with respect to the variables ρ, $\rho\mathbf{u}$, ρE for both the implicit solver (4.10) and the explicit solver (4.8). The update from t^n to t^{n+1} reads after easy calculations

$$\left\{\begin{array}{rcll}(\rho Y)_i^{n+1} &=& (\rho Y)_i^n &- \Delta t \displaystyle\sum_{j\in N(i)} \sigma_{ij}(\rho Y)_{ij}^{n+1-} u_{ij}^* + \lambda_0 \Delta t \rho_i^{n+1}\left(Y^*(\overline{\rho}_i,\overline{e}_i) - Y_i^{n+1}\right),\\ \rho_i^{n+1} &=& \rho_i^n &- \Delta t \displaystyle\sum_{j\in N(i)} \sigma_{ij}\rho_{ij}^{n+1-} u_{ij}^*,\\ (\rho\mathbf{u})_i^{n+1} &=& (\rho\mathbf{u})_i^n &- \Delta t \displaystyle\sum_{j\in N(i)} \sigma_{ij}\left((\rho\mathbf{u})_{ij}^{n+1-} u_{ij}^* + p_{ij}^{*,\theta}\mathbf{n}_{ij}\right),\\ (\rho E)_i^{n+1} &=& (\rho E)_i^n &- \Delta t \displaystyle\sum_{j\in N(i)} \sigma_{ij}\left((\rho E)_i^{n+1-} + p_{ij}^{*,\theta}\right) u_{ij}^*.\end{array}\right. \tag{4.16}$$

For the sake of clarity, let us briefly recall the different steps of the method that shall be referred to as LPS-IMEX(θ). Assume that $(\rho Y,\rho,\rho\mathbf{u},\rho E)_j^n$ is known, $(\rho Y,\rho,\rho\mathbf{u},\rho E)_j^{n+1}$ is computed by the following three steps :

(i) compute $(\rho Y,\rho,\rho\mathbf{u},\rho E)_j^{n+1-}$ from $(\rho Y,\rho,\rho\mathbf{u},\rho E)_j^n$ with (4.7)-(4.10),
(ii) compute $(\overline{\rho Y},\overline{\rho},\overline{\rho\mathbf{u}},\overline{\rho E})_j$ from $(\rho Y,\rho,\rho\mathbf{u},\rho E)_j^{n+1-}$ with (4.11)-(4.12),
(iii) compute $(\rho Y,\rho,\rho\mathbf{u},\rho E)_j^{n+1}$ from $(\overline{\rho Y},\overline{\rho},\overline{\rho\mathbf{u}},\overline{\rho E})_j$ with (4.14) for HRM or with (4.15) for HEM.

We also define the method that shall be referred to as LPS-EX(θ). Assume that $(\rho Y,\rho,\rho\mathbf{u},\rho E)_j^n$ is known, $(\rho Y,\rho,\rho\mathbf{u},\rho E)_j^{n+1}$ is computed by the following three steps :

(i) compute $(\rho Y,\rho,\rho\mathbf{u},\rho E)_j^{n+1-}$ from $(\rho Y,\rho,\rho\mathbf{u},\rho E)_j^n$ with (4.7)-(4.8),
(ii) compute $(\overline{\rho Y},\overline{\rho},\overline{\rho\mathbf{u}},\overline{\rho E})_j$ from $(\rho Y,\rho,\rho\mathbf{u},\rho E)_j^{n+1-}$ with (4.11)-(4.12),
(iii) compute $(\rho Y,\rho,\rho\mathbf{u},\rho E)_j^{n+1}$ from $(\overline{\rho Y},\overline{\rho},\overline{\rho\mathbf{u}},\overline{\rho E})_j$ with (4.14) for HRM or with (4.15) for HEM.

The difference between those two methods is that the Lagrange step is implicit for the LPS-IMEX(θ) scheme and explicit for the LPS-EX(θ) scheme. The source terms step is treated implicitly and the transport step explicitly in both schemes.

4.5 Main properties

We now give the main properties of the LPS-EX(θ) and LPS-IMEX(θ) schemes.

Theorem 7. *Under the acoustic CFL condition (4.9) and the material CFL condition (4.13), the LPS-EX(θ) scheme is well-defined and satisfies the following properties*

(i) it is a conservative scheme for ρ, $\rho\mathbf{u}$ and ρE. It is also a conservative scheme for ρY when there is no mass transfer between phases ($\lambda_0 = 0$).

(ii) the density ρ_i^n is positive for all i and $n > 0$ provided that ρ_i^0 is positive for all i.

(iii) $Y_i^n \in [0,1]$ for all i and $n > 0$ provided that $Y_i^0 \in [0,1]$ for all i and $\overline{e}_i > 0$ for all i and $n \geq 0$.

(iv) if $\theta = \mathcal{O}(M)$, then the truncation error of the numerical scheme is uniform with respect to $M < 1$.

Theorem 8. *Under the material CFL condition (4.13), the LPS-IMEX(θ) scheme is well-defined and satisfy the following properties*

(i) it is a conservative scheme for ρ, $\rho\mathbf{u}$ and ρE. It is also a conservative scheme for ρY when there is no mass transfer between phases ($\lambda_0 = 0$).

(ii) the density ρ_i^n is positive for all i and $n > 0$ provided that ρ_i^0 is positive for all i.

(iii) $Y_i^n \in [0,1]$ for all i and $n > 0$ provided that $Y_i^0 \in [0,1]$ for all i and $\overline{e}_i > 0$ for all i and $n \geq 0$.

(iv) if $\theta = \mathcal{O}(M)$, then the truncation error of the numerical scheme is uniform with respect to $M < 1$.
(v) it is stable in the uniform sense with respect to the Mach number M.

Remark 11. *For the LPS-EX(θ) scheme, we may prove the positivity of the internal energy and a discrete entropy inequality under a condition on the modification θ, see appendix 4.B for more details. Under this condition, we have in particular $\overline{e}_i > 0$ for all i and $n \geq 0$.*

Proof of property *(i)* is easily obtained from (4.16) and is thus left to the reader (see also [8]).

Proof of properties *(ii)* **and** *(iii)*. Let us consider that $Y_i^n \in [0,1]$, $\rho_i^n > 0$ and $\overline{e}_i > 0$ for all i, we are going to show that $Y_i^{n+1} \in [0,1]$ and $\rho_i^{n+1} > 0$ for all i :

- Acoustic step : the mass fraction is unchanged in this step $Y_i^{n+1-} = Y_i^n$. Thus, we have $Y_i^{n+1-} \in [0,1]$ for all i.
  The density is given by
  $$\rho_i^{n+1-} = \rho_i^n \left(1 + \Delta t \sum_{j \in N(i)} \sigma_{ij} u_{ij}^*\right)^{-1},$$
  so that we have $\rho_i^{n+1-} > 0$ for all i thanks to the CFL condition (4.13).
- Transport step : the upwind choice (4.12) is such that
  $$u_{ij}^* \varphi_{ij}^{n+1-} = (u_{ij}^*)^+ \varphi_i^{n+1-} + (u_{ij}^*)^- \varphi_j^{n+1-},$$
  where $u^+ = \frac{u+|u|}{2}$ and $u^- = \frac{u-|u|}{2}$. Injecting those expressions in the transport step (4.11) for the density and the mass fraction holds
  $$\begin{cases} \overline{\rho}_i = \left(1 + \Delta t \sum_{j \in N(i)} \sigma_{ij} (u_{ij}^*)^-\right) \rho_i^{n+1-} - \Delta t \sum_{j \in N(i)} \sigma_{ij} \left(u_{ij}^*\right)^- \rho_j^{n+1-}, \\ \overline{Y}_i = \left(\frac{\rho_i^{n+1-}}{\overline{\rho}_i}\right) \left(1 + \Delta t \sum_{j \in N(i)} \sigma_{ij} (u_{ij}^*)^-\right) Y_i^{n+1-} - \Delta t \sum_{j \in N(i)} \sigma_{ij} \left(u_{ij}^*\right)^- \left(\frac{\rho_j^{n+1-}}{\overline{\rho}_i}\right) Y_j^{n+1-}. \end{cases}$$
  As $(u_{ij}^*)^- \leq 0$, under the CFL condition (4.13), $\overline{\rho}_i$ (resp. $\overline{Y}_i$) is a convex combination of ρ_i^{n+1-} (resp. Y_i^{n+1-}) and ρ_j^{n+1-} (resp. Y_j^{n+1-}) for $j \in N(i)$. Thus, we have $\overline{\rho}_i > 0$ and $\overline{Y}_i \in [0,1]$ for all i.
- Phase transition step : the density is unchanged in this step $\rho_i^{n+1} = \overline{\rho}_i$, so that $\rho_i^{n+1} > 0$ for all i.
  The mass fraction update writes for HRM
  $$Y_i^{n+1} = \left(\frac{1}{1 + \lambda_0 \Delta t}\right) \overline{Y}_i + \left(\frac{\lambda_0 \Delta t}{1 + \lambda_0 \Delta t}\right) Y^*(\overline{\rho}_i, \overline{e}_i),$$
  and for HEM
  $$Y_i^{n+1} = Y^*(\overline{\rho}_i, \overline{e}_i),$$
  In both cases, Y_i^{n+1} may be seen as a convex combination of $Y^*(\overline{\rho}_i, \overline{e}_i)$ and $\overline{Y}_i$. We assumed that $\overline{e}_i > 0$ and proved that $\overline{\rho}_i > 0$ for all i, so that $Y^*(\overline{\rho}_i, \overline{e}_i) \in [0,1]$ by definition of the function Y^*. Thus $Y_i^{n+1} \in [0,1]$ for all i.

This conclude the proof. □

Behavior with respect to the Mach regime. In order to prove *(iv)* and *(v)*, we are now interested in the behavior of the numerical scheme with respect to the Mach regime. Namely, we study the dependence with respect to the Mach number M of both the CFL stability condition and the truncation error.
Introducing the rescaling and tilde variables defined earlier into (4.8) we get

$$\left\{\begin{array}{lcl} \tilde{u}_{ij}^{*} & = & \frac{1}{2}\mathbf{n}_{ij}^{T}\left(\tilde{\mathbf{u}}_{i}^{n}+\tilde{\mathbf{u}}_{j}^{n}\right)-\frac{1}{M}\frac{1}{2\tilde{a}_{ij}}\left(\tilde{\Pi}_{j}^{n}-\tilde{\Pi}_{i}^{n}\right), \\ \tilde{p}_{ij}^{*,\theta} & = & \frac{1}{2}\left(\tilde{\Pi}_{i}^{n}+\tilde{\Pi}_{j}^{n}\right)-M\frac{\tilde{a}_{ij}\theta_{ij}}{2}\mathbf{n}_{ij}^{T}\left(\tilde{\mathbf{u}}_{j}^{n}-\tilde{\mathbf{u}}_{i}^{n}\right). \end{array}\right. \tag{4.17}$$

For (4.10) we get

$$\left\{\begin{array}{lcl} \tilde{u}_{ij}^{*} & = & \frac{1}{2}\mathbf{n}_{ij}^{T}\left(\tilde{\mathbf{u}}_{i}^{n+1-}+\tilde{\mathbf{u}}_{j}^{n+1-}\right)-\frac{1}{M}\frac{1}{2\tilde{a}_{ij}}\left(\tilde{\Pi}_{j}^{n+1-}-\tilde{\Pi}_{i}^{n+1-}\right), \\ \tilde{p}_{ij}^{*,\theta} & = & \frac{1}{2}\left(\tilde{\Pi}_{i}^{n+1-}+\tilde{\Pi}_{j}^{n+1-}\right)-M\frac{\tilde{a}_{ij}\theta_{ij}}{2}\mathbf{n}_{ij}^{T}\left(\tilde{\mathbf{u}}_{j}^{n+1-}-\tilde{\mathbf{u}}_{i}^{n+1-}\right). \end{array}\right. \tag{4.18}$$

The rescaling of the acoustic step (4.7) reads

$$\left\{\begin{array}{lcl} \tilde{\mathbf{u}}_{i}^{n+1-} & = & \tilde{\mathbf{u}}_{i}^{n}-\dfrac{1}{M^{2}}\tilde{\tau}_{i}^{n}\Delta\tilde{t}\displaystyle\sum_{j\in N(i)}\tilde{\sigma_{ij}}\tilde{p}_{ij}^{*,\theta}\mathbf{n}_{ij}, \\ \tilde{\Pi}_{i}^{n+1-} & = & \tilde{\Pi}_{i}^{n}-\tilde{\tau}_{i}^{n}\Delta\tilde{t}\displaystyle\sum_{j\in N(i)}\tilde{\sigma}_{ij}(\tilde{a}_{ij})^{2}\tilde{u}_{ij}^{*}, \\ Y_{i}^{n+1-} & = & Y_{i}^{n}, \\ \tilde{\tau}_{i}^{n+1-} & = & \tilde{\tau}_{i}^{n}+\tilde{\tau}_{i}^{n}\Delta\tilde{t}\displaystyle\sum_{j\in N(i)}\tilde{\sigma}_{ij}\tilde{u}_{ij}^{*}, \\ \tilde{E}_{i}^{n+1-} & = & \tilde{E}_{i}^{n}-\tilde{\tau}_{i}^{n}\Delta\tilde{t}\displaystyle\sum_{j\in N(i)}\tilde{\sigma}_{ij}\tilde{p}_{ij}^{*,\theta}\tilde{u}_{ij}^{*}, \end{array}\right. \tag{4.19}$$

where $\tilde{\sigma_{ij}}=\frac{\sigma_{ij}}{L}$ and $\tilde{a}_{ij}=\frac{a_{ij}}{\rho_0 c_0}$. Note that the CFL restriction of the explicit acoustic step reads now

$$\Delta\tilde{t}\max_{1\leq j\leq N}\left[\tilde{\tau}_{j}^{n}\left(\max_{i\in N(j)}\tilde{\sigma}_{ij}\tilde{a}_{ij}\right)\right]\leq\frac{M}{2}. \tag{4.20}$$

The rescaling of the transport step (4.11) reads for $\tilde{\varphi}\in\{\tilde{\rho}Y,\tilde{\rho},\tilde{\rho}\tilde{\mathbf{u}},\tilde{\rho}\tilde{E}\}$

$$\overline{\tilde{\varphi}}_{i}=\tilde{\varphi}_{i}^{n+1-}-\Delta\tilde{t}\left[\sum_{j\in N(i)}\tilde{\sigma_{ij}}\tilde{u}_{ij}^{*}\tilde{\varphi}_{ij}^{n+1-}\right]+\Delta\tilde{t}\tilde{\varphi}_{i}^{n+1-}\left[\sum_{j\in N(i)}\tilde{\sigma_{ij}}\tilde{u}_{ij}^{*}\right]. \tag{4.21}$$

The CFL restriction associated with the transport step is

$$\Delta\tilde{t}\max_{1\leq j\leq N}\left[\sum_{i\in N(i)}(\tilde{\sigma_{ij}}|\tilde{u}_{ij}^{*}|)\right]\leq 1. \tag{4.22}$$

Finally the phase transition step (4.14) becomes

$$\left\{\begin{array}{lcl} Y_{i}^{n+1} & = & \overline{Y}_{i}+\lambda_{0}T\Delta\tilde{t}\left(Y^{*}(\overline{\rho}_{i},\overline{e}_{i})-Y_{i}^{n+1}\right), \\ \tilde{\varphi}_{i}^{n+1} & = & \overline{\tilde{\varphi}}_{i},\quad \tilde{\varphi}\in\{\tilde{\rho},\tilde{\rho}\tilde{\mathbf{u}},\tilde{\rho}\tilde{E}\}. \end{array}\right. \tag{4.23}$$

Proof of property *(v)*. We define $\tilde{h}=h/L$ where h is the mesh size. The acoustic CFL restriction (4.20) is very restrictive in low Mach regime as $\Delta\tilde{t}=\mathcal{O}\left(M\tilde{h}\right)$, while the transport CFL restriction (4.22) is uniform with respect to the Mach number $\Delta\tilde{t}=\mathcal{O}\left(\tilde{h}\right)$. Thus the LPS-IMEX($\theta$) scheme is stable in the uniform sense with respect to the Mach number M, while the LPS-EX(θ) scheme is not. □

Proof of property *(iv)*. In order to evaluate the truncation error in the low Mach regime, we use the classical tool of equivalent equation. With a slight abuse of notation, we consider $\tilde{\varphi}(\mathbf{x}_i,t^n)=\tilde{\varphi}_i^n$ so that we can substitute these functions in discrete formulas.
We assume that we are in low Mach regime, namely $M\ll 1$ and $\tilde{\nabla}\tilde{p}=\mathcal{O}(M^2)$. This hypothesis yields

that for $j \in N(i)$, we have $\tilde{\Pi}_j^n = \tilde{\Pi}_i^n + \mathcal{O}(M^2\tilde{h})$ and $\tilde{\Pi}_j^{n+1-} = \tilde{\Pi}_i^{n+1-} + \mathcal{O}(M^2\tilde{h})$ for the discrete unknowns. The rescaled discretization of the acoustic step (4.19) is consistent with

$$\left\{\begin{array}{rcl} \partial_{\tilde{t}} Y & = & \mathcal{O}\left(\Delta\tilde{t}\right), \\ \partial_{\tilde{t}}\tilde{\tau} - \tilde{\tau}\tilde{\nabla}\cdot\tilde{\mathbf{u}} & = & \mathcal{O}\left(\Delta\tilde{t}\right) + \mathcal{O}\left(M\tilde{h}\right), \\ \partial_{\tilde{t}}\tilde{\mathbf{u}} + \frac{\tilde{\tau}}{M^2}\tilde{\nabla}\tilde{p}(\tilde{\rho},\tilde{e},Y) & = & \mathcal{O}\left(\Delta\tilde{t}\right) + \mathcal{O}\left(\frac{\theta}{M}\tilde{h}\right), \\ \partial_{\tilde{t}}\tilde{E} + \tilde{\tau}\tilde{\nabla}\cdot\left(\tilde{p}(\tilde{\rho},\tilde{e},Y)\tilde{\mathbf{u}}\right) & = & \mathcal{O}\left(\Delta\tilde{t}\right) + \mathcal{O}\left(M\tilde{h}\right) + \mathcal{O}\left(\theta M\tilde{h}\right). \end{array}\right.$$

for both the implicit solver (4.18) and the explicit solver (4.17).
The rescaled discretization of the transport step (4.22) is consistent with

$$\partial_{\tilde{t}}\tilde{\varphi} + \tilde{\mathbf{u}}\cdot\tilde{\nabla}\tilde{\varphi} = \mathcal{O}\left(\Delta\tilde{t}\right) + \mathcal{O}\left(\tilde{h}\right) + \mathcal{O}\left(M\tilde{h}\right) \quad for\ \tilde{\varphi} \in \{\tilde{\rho}Y, \tilde{\rho}, \tilde{\rho}\tilde{\mathbf{u}}, \tilde{\rho}\tilde{E}\}$$

The rescaled discretization of the phase transition step (4.23) is consistent with

$$\left\{\begin{array}{rcl} \partial_{\tilde{t}}(\tilde{\rho}Y) & = & \lambda_0 T\left(\tilde{\rho}Y^*(\tilde{\rho},\tilde{e}) - \tilde{\rho}Y\right) + \mathcal{O}\left(\Delta\tilde{t}\right), \\ \partial_{\tilde{t}}\tilde{\varphi} & = & \mathcal{O}\left(\Delta\tilde{t}\right) \quad \text{for } \tilde{\varphi} \in \{\tilde{\rho}, \tilde{\rho}\tilde{\mathbf{u}}, \tilde{\rho}\tilde{E}\}. \end{array}\right.$$

So that the equivalent equation verified by the overall rescaled scheme reads

$$\left\{\begin{array}{rcl} \partial_t(\rho Y) + \nabla\cdot(\rho Y\mathbf{u}) & = & \lambda_0\rho\left(Y^*(\rho,e) - Y\right) + \mathcal{O}\left(\Delta\tilde{t}\right) + \mathcal{O}\left(\tilde{h}\right) + \mathcal{O}\left(M\tilde{h}\right), \\ \partial_t\rho + \nabla\cdot(\rho\mathbf{u}) & = & \mathcal{O}\left(\Delta\tilde{t}\right) + \mathcal{O}\left(\tilde{h}\right) + \mathcal{O}\left(M\tilde{h}\right), \\ \partial_t(\rho\mathbf{u}) + \nabla\cdot(\rho\mathbf{u}\otimes\mathbf{u}) + \nabla p & = & \mathcal{O}\left(\Delta\tilde{t}\right) + \mathcal{O}\left(\frac{\theta\tilde{h}}{M}\right) + \mathcal{O}\left(\tilde{h}\right) + \mathcal{O}\left(M\tilde{h}\right), \\ \partial_t(\rho E) + \nabla\cdot\left[(\rho E + p)\,\mathbf{u}\right] & = & \mathcal{O}\left(\Delta\tilde{t}\right) + \mathcal{O}\left(\tilde{h}\right) + \mathcal{O}\left(M\tilde{h}\right) + \mathcal{O}\left(\theta M\tilde{h}\right). \end{array}\right.$$

As a consequence, provided that we impose the asymptotic behavior $\theta = \mathcal{O}(M)$, the truncation error of scheme (4.19)-(4.21)-(4.23) is uniform with respect to M for both the implicit solver (4.18) and the explicit solver (4.17). This concludes the proof of property *(iv)*. □

Let us note that the classical Suliciu relaxation fluxes obtained for $\theta = 1$ do not have a truncation error that is uniform with respect to the Mach number.

4.6 Numerical results

We propose to test both LPS-IMEX(θ) and LPS-EX(θ) scheme against low Mach number test cases and order 1 Mach number test cases. LPS-EX(θ) computations are performed with a time step satisfying both (4.9) and (4.13), while LPS-IMEX(θ) computations are performed with a time step defined by an explicit evaluation of (4.13) (explicit means here that u^* defined by (4.8) is used to evaluate Δt). We consider a mixture of two perfect gas with different adiabatic coefficients $\gamma_1 > \gamma_2 > 1$. The pressure, sound speed and mass fraction of the mixture are given by

$$\left\{\begin{array}{l} p(\rho,e,Y) = (\gamma_{mix}(Y)-1)\rho e, \quad c^2(\rho,e,Y) = \gamma_{mix}(Y)\frac{p(\rho,e,Y)}{\rho}, \\ Y^*(\rho,e) = \left\{\begin{array}{ll} 1 & \text{if } \rho < \rho_1^*, \\ \left(\frac{\rho_1^*}{\rho}\right)\left(\frac{\rho-\rho_2^*}{\rho_1^*-\rho_2^*}\right) & \text{if } \rho_1^* \le \rho \le \rho_2^*, \\ 0 & \text{if } \rho_2^* < \rho, \end{array}\right. \end{array}\right.$$

where

$$\gamma_{mix}(Y) = Y\gamma_1 + (1-Y)\gamma_2, \quad \rho_1^* = \frac{1}{\exp(1)}\left(\frac{\gamma_2-1}{\gamma_1-1}\right)^{\frac{\gamma_2}{\gamma_2-\gamma_1}}, \quad \rho_2^* = \frac{1}{\exp(1)}\left(\frac{\gamma_2-1}{\gamma_1-1}\right)^{\frac{\gamma_1}{\gamma_2-\gamma_1}}.$$

We refer for instance the reader to [1] and the references therein. We assume that $\lambda_0 \to \infty$, so that the thermodynamic equilibrium is instantaneously achieved.

4.6.1 Low Mach number examples

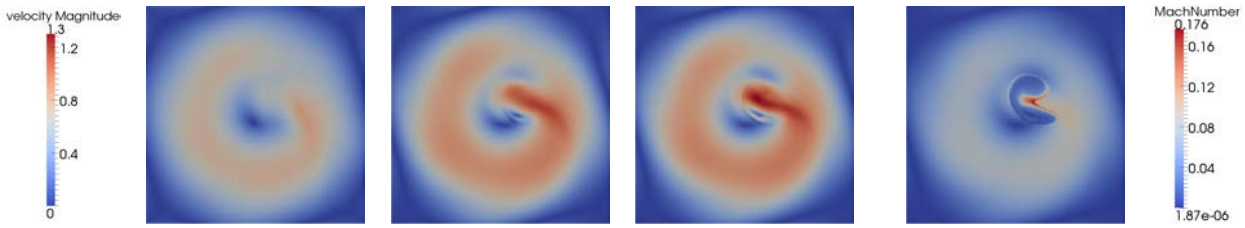
We consider low Mach test cases and try to examine two questions : the accuracy gain for simulations on coarse grid in the low Mach regime, then the benefit of using a semi-implicit strategy in term of CPU time.

Bubble in a vortex test case. The computational domain is $\Omega = [0,1]^2$. The adiabatic coefficients are $\gamma_1 = 2$, $\gamma_2 = 1.4$, which gives $\rho_1^* \simeq 3.1205576$, $\rho_2^* \simeq 7.801394$. The initial condition is given by

$$\begin{cases} p(x,y,t=0) = 1000, \\ \rho(x,y,t=0) = \begin{cases} 1 & \text{if } (x-0.5)^2+(y-0.25)^2 \le 0.01, \\ 10 & \text{if } (x-0.5)^2+(y-0.25)^2 > 0.01, \end{cases} \\ Y(x,y,t=0) = \begin{cases} 1 & \text{if } (x-0.5)^2+(y-0.25)^2 \le 0.01, \\ 0 & \text{if } (x-0.5)^2+(y-0.25)^2 > 0.01, \end{cases} \\ u(x,y,t=0) = 2\sin^2(\pi x)\sin(\pi y)\cos(\pi y), \\ v(x,y,t=0) = -2\sin(\pi x)\cos(\pi x)\sin^2(\pi y). \end{cases}$$

We impose no-slip boundary conditions. The Mach number for the resulting flows is of order 10^{-4} in phase 1 ($Y = 1$) and 10^{-3} in phase 2 ($Y = 0$) so that pure phases are in the low Mach regime. Nevertheless, since the sound speed of the mixture is smaller than the sound speed of pure phase, we observed a Mach number that goes up to 10^{-1} in the mixture. We plot the solution at time $t = 0.5s$.

Figure (4.1)-(4.3) displays the results obtained with LPS-EX(θ) scheme for $\theta_{ij} = 1$ and $\theta_{ij} = M_{ij}^n$, where $M_{ij}^n = \frac{|u_{ij}^*|}{max(c_i,c_j)}$ is an evaluation of the Mach number on each interface at time t^n. We use as a reference solution an approximation computed with LPS-EX($\theta = 1$) using a 1.6×10^5-cell triangular mesh. The choice $\theta_{ij} = M_{ij}$ leads to approximation that are much more accurate than $\theta_{ij} = 1$. On figures (4.2)-(4.3), we obtain similar results for the LPS-IMEX(θ) scheme. The LPS-IMEX(θ) and LPS-EX(θ) schemes for $\theta_{ij} = M_{ij}^n$ require respectively 2479s=41min19s and 16465s=4h34min25s of CPU time, so that using an implicit solver for the acoustic step is 6.6 times faster.



(a) (b) (c) (d)

FIGURE 4.1 – Bubble in a vortex test case. Profile at time t = 0.5 s of the velocity magnitude for (a) LSP-EX($\theta = 1$), (b) LSP-EX($\theta = \mathcal{O}(M)$) with a 200×200-cell Cartesian mesh, (c) velocity magnitude obtained with the reference solution and (d) Mach number obtained with the reference solution.

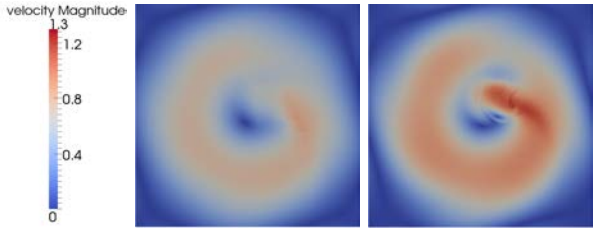


FIGURE 4.2 – Bubble in a vortex test case. Profile at time t = 0.5 s of the velocity magnitude for the LSP-IMEX($\theta = 1$) scheme (left) and the LSP-IMEX($\theta = \mathcal{O}(M)$) scheme (right) on a 200 × 200-cell Cartesian mesh. To be compared with (a), (b) and (c) on figure 4.1.

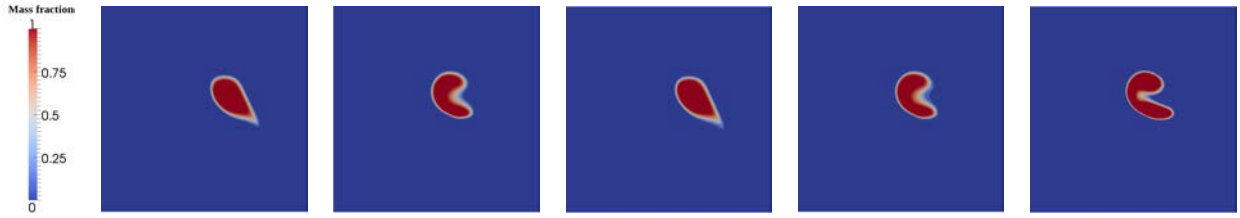


FIGURE 4.3 – Bubble in a vortex test case. Profile at time t = 0.5 s of the mass fraction for (a) LSP-EX($\theta = 1$), (b) LSP-EX($\theta = \mathcal{O}(M)$), (c) LSP-IMEX($\theta = 1$) , (d) LSP-IMEX($\theta = \mathcal{O}(M)$) with a 200 × 200-cell Cartesian mesh and (e) mass fraction of the reference solution.

Subsonic flow in a channel with bump. We consider now the case of a subsonic flow passing a channel with a 20% sinusoidal bump, see figure 4.4.

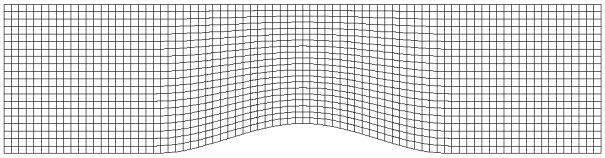


FIGURE 4.4 – 80 × 20 quadrangular mesh of a 20% sinusoidal bump.

The adiabatic coefficients are $\gamma_1 = 2$, $\gamma_2 = 1.4$, which gives $\rho_1^* \simeq 3.1205576$, $\rho_2^* \simeq 7.801394$. The initial condition is given by

$$(\rho, Y, p, u, v)(x, y, t = 0) = (7.81, 0, 3124, 0, 0).$$

We impose an inlet boundary condition at $0 \times [0, 1] : (h, u, v) = (1400, 0.2, 0)$ and an outlet boundary condition at $4 \times [0, 1] : p = 3124$. Wall boundary conditions are set on other boundaries. There is no occurence of phase 1 for this configuration as $\rho > \rho_2^*$ in the whole computational domain. Besides, the Mach number is of order 10^{-2} so that we are in the low mach regime. All tests are performed on a 80×20 quadrangular mesh.

Figure 4.5,4.6 and 4.7 display the flow profile at $t = 2s$, we observe that LPS-EX(θ) and LPS-IMEX(θ) schemes are more diffusive for $\theta = 1$ than for $\theta_{ij} = M_{ij}^n$, where $M_{ij}^n = \frac{|u_{ij}^*|}{max(c_i, c_j)}$ is an evaluation of the Mach number on each interface at time t^n. In term of CPU time, for $\theta_{ij} = M_{ij}^n$, we observe that the LPS-IMEX(θ) scheme is 79.2 times faster than the LPS-EX(θ) scheme thanks to the use of a material velocity CFL condition 4.13.

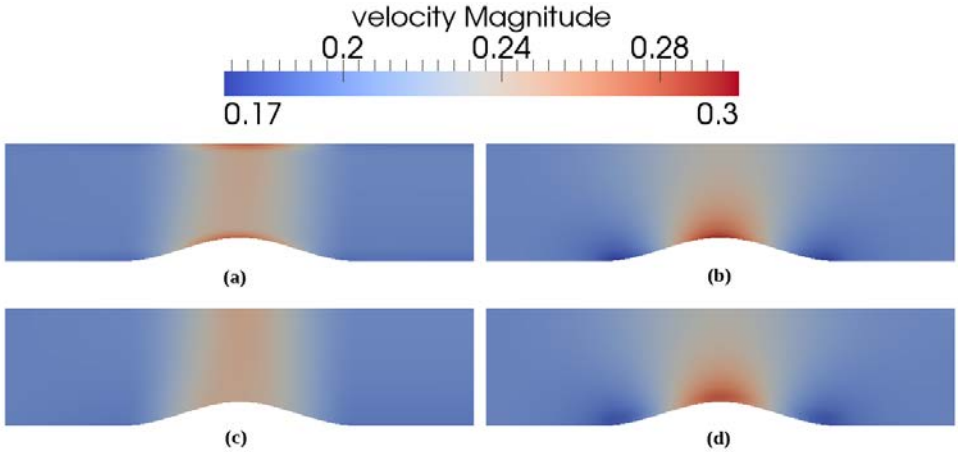


FIGURE 4.5 – Subsonic flow in a channel with bump. Profile at t = 2s of the velocity magnitude for (a) LPS-EX($\theta = 1$), (b) LPS-EX($\theta = \mathcal{O}(M)$), (c) LPS-IMEX($\theta = 1$), (d) LPS-IMEX($\theta = \mathcal{O}(M)$) using a 80×20 quadrangular mesh.

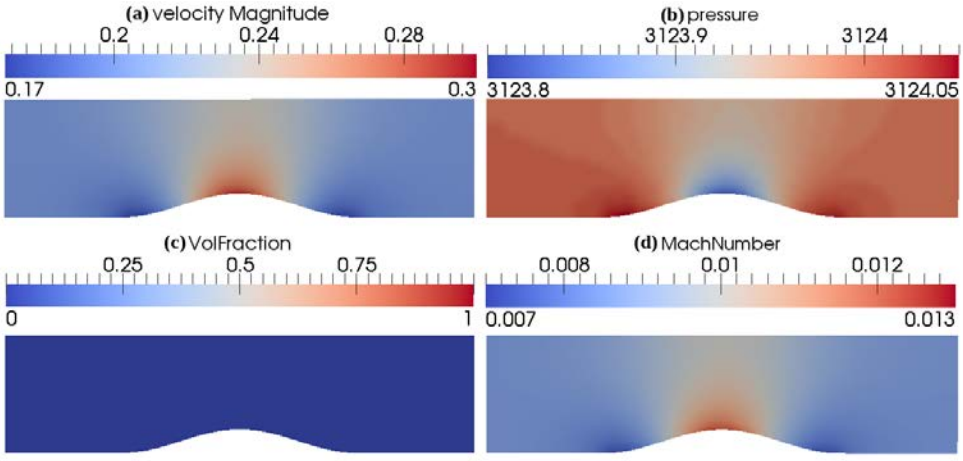


FIGURE 4.6 – Subsonic flow in a channel with bump. Profile at t = 2s of the (a) velocity magnitude, (b) pressure, (c) mass fraction and (d) Mach number for the LPS-EX($\theta = \mathcal{O}(M)$) using a 80×20 quadrangular mesh.

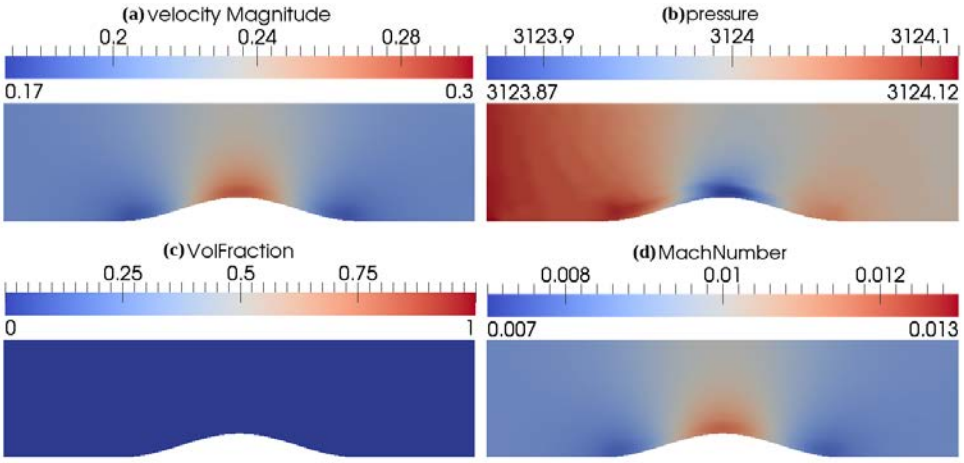


FIGURE 4.7 – Subsonic flow in a channel with bump. Profile at t = 2s of the (a) velocity magnitude, (b) pressure, (c) mass fraction and (d) Mach number for the LPS-IMEX($\theta = \mathcal{O}(M)$) using a 80×20 quadrangular mesh.

4.6.2 Compressible flow examples

In this section, we assess the ability of our operator splitting scheme to handle cases where the flow is not in the low Mach regime over the whole computational domain. This is an important issue since the modification introduced by θ modifies the numerical diffusion of the scheme, which is relevant in the low Mach number regime, but might give rise to instabilities when the Mach number is of order 1 or larger. *We will see that even with a centred pressure discretization, ie. for $\theta = 0$, the solution remains stable.*

Two-rarefaction Riemann problem with appearance of phase 1. The computational domain is $\Omega = [0, 1]$. The adiabatic coefficients are $\gamma_1 = 1.6$, $\gamma_2 = 1.4$, which yields $\rho_1^* \simeq 6.2855651$, $\rho_2^* \simeq 9.4283477$. The initial condition is given by

$$(\rho, Y, p, u)(x, t = 0) = \begin{cases} (10, 0, 1, -2), & \text{for } x < 0.5, \\ (10, 0, 1, 1), & \text{for } x > 0.5. \end{cases}$$

We impose Neumann boundary conditions and plot the solution at time $t = 0.1s$. The initial condition is in phase 2 but an intermediate zone with $\rho < \rho_1^*$ appears for $t > 0$.

Figure (4.8) displays the results obtained with LPS-EX(θ) and LPS-IMEX(θ) schemes for $\theta_{ij} = 1$ and $\theta_{ij} = 0$. We use as a reference solution an approximation computed with LPS-EX($\theta = 1$) using a 10 000-cell mesh. All schemes show a good agreement with the reference solution. The LPS-EX($\theta = 0$) and LPS-IMEX($\theta = 0$) schemes are slightly less diffusive than the LPS-EX($\theta = 1$) and LPS-IMEX($\theta = 1$) schemes. Let us underline that despite part of the solutions clearly do not belong to the low Mach regime since $M \approx 5.4$ in the left part of the domain and $M \approx 2.7$ in the right part of the domain, the schemes LPS-EX($\theta = 0$) and LPS-IMEX($\theta = 0$) are stable and provide good numerical results while involving a centred pressure discretization with $\theta_{ij} = 0$.

Remark 12. *The numerical diffusion of the mass fraction generates a mixture zone where the sound speed is much smaller than in pure phases. This leads to an overshot for the Mach number value to be observed on figure 4.8. One may use an anti-diffusive scheme for the mass fraction during the transport step to reduce this phenomenon.*

Transonic flow in a channel with bump. We consider now the case of a transonic flow passing a channel with a 20% sinusoidal bump, see figure 4.4. The adiabatic coefficients are $\gamma_1 = 2$, $\gamma_2 = 1.4$, which

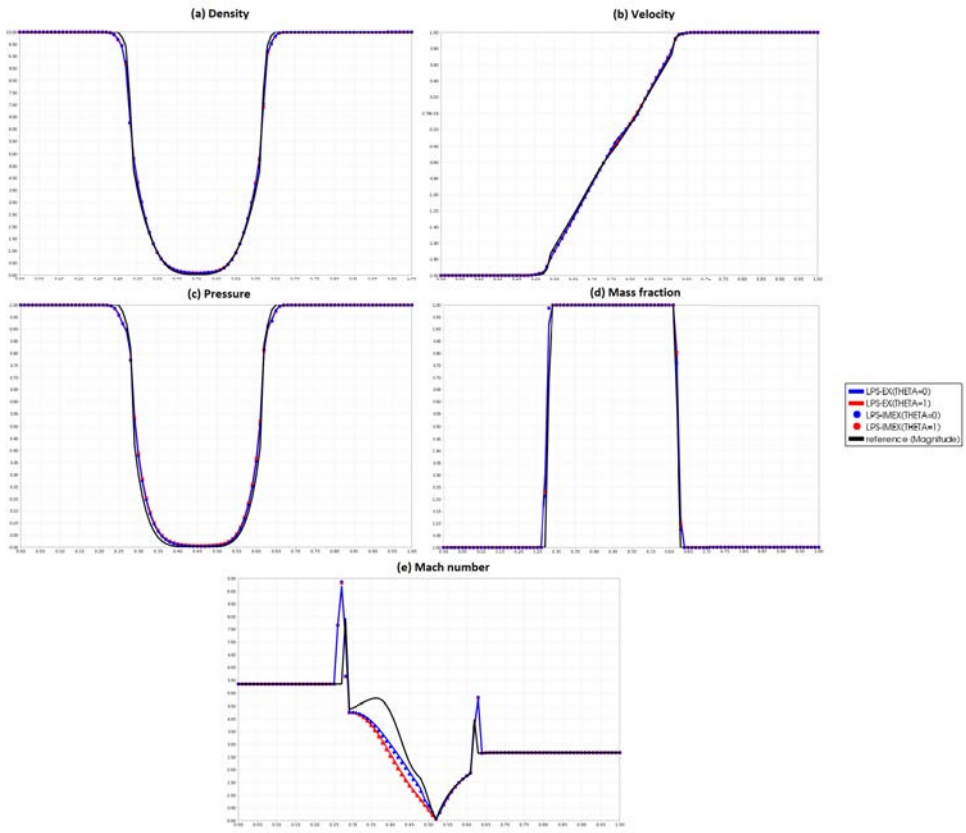


FIGURE 4.8 – Two-rarefaction Riemann problem with appearance of phase 1. Profile at t = 0.1 s of the (a) density, (b) velocity magnitude, (c) pressure, (d) mass fraction and (e) Mach number for the LPS-EX($\theta = 1$), LPS-EX($\theta = 0$), LPS-IMEX($\theta = 1$), LPS-IMEX($\theta = 0$) using a 1000-cell grid, together with reference solution.

gives $\rho_1^* \simeq 3.1205576$, $\rho_2^* \simeq 7.801394$. The initial condition is given by

$$(\rho, Y, p, u, v)(x, y, t = 0) = (7.81, 0, 3124, 0, 0).$$

We impose an inlet boundary condition at $0 \times [0,1] : (h, u, v) = (1400, 12, 0)$ and an outlet boundary condition at $4 \times [0,1] : p = 3124$. Wall boundary conditions are set on other boundaries. This configuration leads to a transonic flow as M reaches 1.12. For this configuration, there is occurence of phase 1 near the bump. All tests are performed on a 80×20 quadrangular mesh and we plot solution at time $t = 2s$.

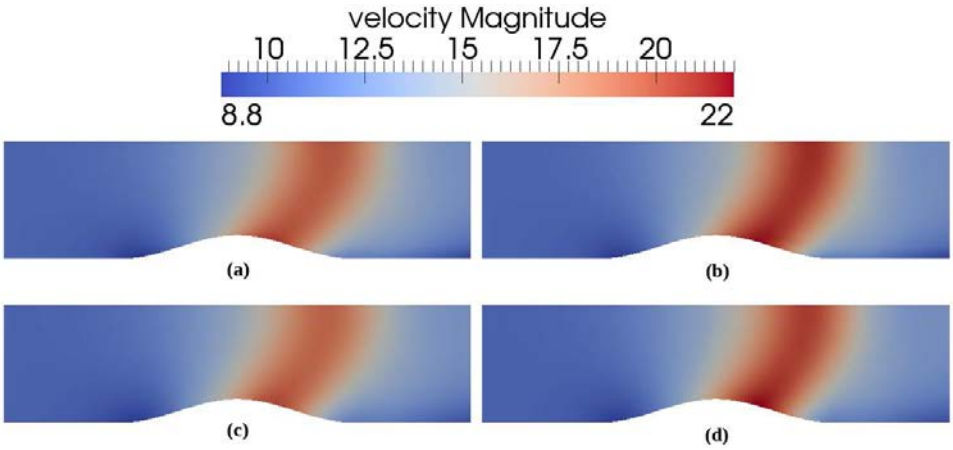


FIGURE 4.9 – Transonic flow in a channel with bump. Profile at t = 2s of the velocity magnitude for (a) LPS-EX($\theta = 1$), (b) LPS-EX($\theta = 0$), (c) LPS-IMEX($\theta = 1$), (d) LPS-IMEX($\theta = 0$) using a 80×20 quadrangular mesh.

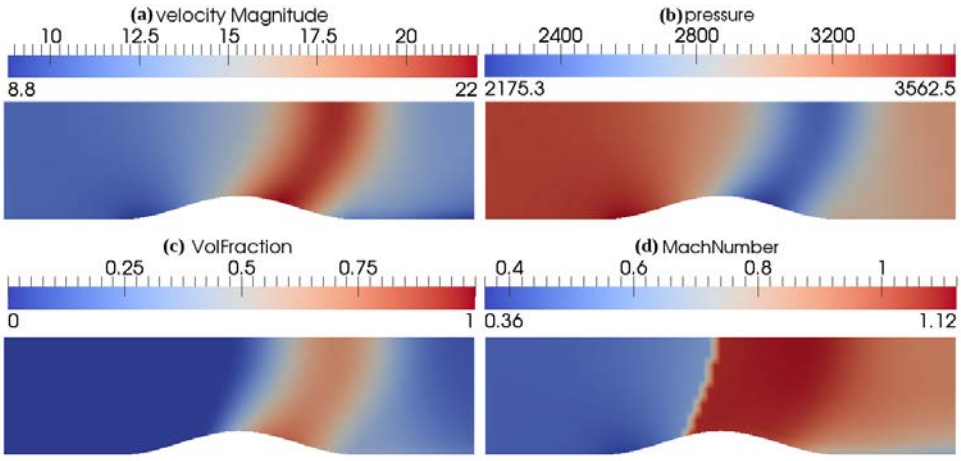


FIGURE 4.10 – Transonic flow in a channel with bump. Profile at t = 2s of the (a) velocity magnitude, (b) pressure, (c) mass fraction and (d) Mach number for LPS-EX($\theta = 0$) using a 80×20 quadrangular mesh.

Figure 4.9,4.10 and 4.11 display the results obtained with LPS-EX(θ) and LPS-IMEX(θ) schemes for $\theta_{ij} = 1$ and $\theta_{ij} = 0$. All schemes are able to compute accurate approximate solutions. The results obtained with $\theta_{ij} = 0$ are slightly less diffused than the results obtained with $\theta_{ij} = 1$.

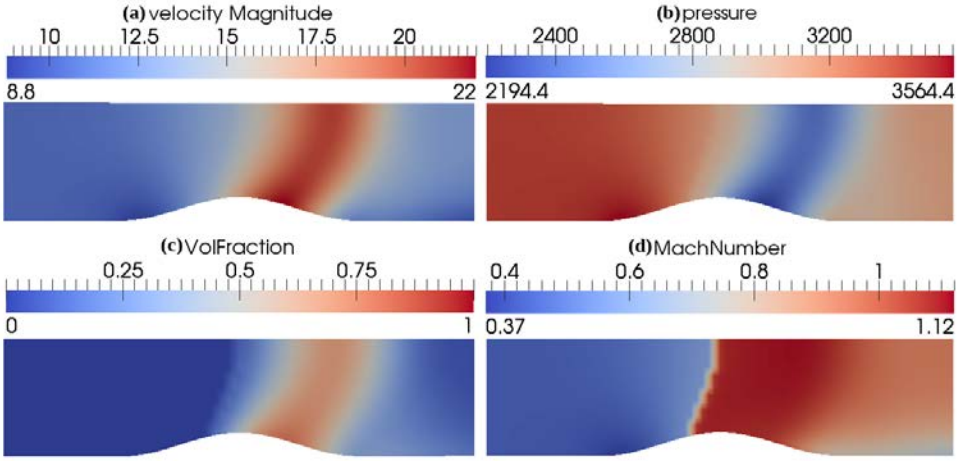


FIGURE 4.11 – Transonic flow in a channel with bump. Profile at t = 2s of the (a) velocity magnitude, (b) pressure, (c) mass fraction and (d) Mach number for LPS-IMEX($\theta = 0$) using a 80×20 quadrangular mesh.

Annexes

4.A Solver de Riemann approché pour le schéma acoustique modifié

Le schéma numérique (4.7)-(4.8) est une méthode de type volume fini. En effet, pour mettre à jour la solution à l'aide de (4.7), on calcule sur les faces du maillage des flux modifiés avec (4.8). Ces flux peuvent être obtenus à partir des flux d'un système quasi-1D grâce à l'invariance par rotation du système (4.5). On va montrer dans cette section que ces flux modifiés quasi-1D peuvent être obtenus à partir d'un *solver* de Riemann approché, au sens de Harten, Lax et van Leer [19], qui est consistant au sens intégral avec le système

$$\partial_t \mathbf{W} + \partial_m \mathbf{F}(\mathbf{W}) = 0, \tag{4.24}$$

où $\mathbf{W} = (Y, \tau, u, v, E, \Pi)^T$, $\mathbf{F}(\mathbf{W}) = (0, -u, \Pi, 0, \Pi u, a^2 u)^T$, et la variable de masse m est définie par $\mathrm{d}m = \rho(x, t^n)\mathrm{d}x$.

On considère le schéma numérique quasi-1D

$$\begin{cases} \mathbf{W}_j^{n+1-} = \mathbf{W}_j^n - \dfrac{\Delta t}{\Delta m}\left(\mathbf{F}_{j+1/2} - \mathbf{F}_{j-1/2}\right), & \text{(4.25a)} \\ \mathbf{F}_{j+1/2} = \mathbf{F}^\theta(\mathbf{W}_j^n, \mathbf{W}_{j+1}^n), & \text{(4.25b)} \\ \mathbf{F}^\theta(\mathbf{W}_L, \mathbf{W}_R) = (0, -u^*, \Pi^{*,\theta}, 0, \Pi^{*,\theta} u^*, a^2 u^*)^T, & \text{(4.25c)} \end{cases}$$

où la vitesse et la pression d'interface sont données par

$$\begin{cases} u^* = \dfrac{(u_R + u_L)}{2} - \dfrac{1}{2a}(\Pi_R - \Pi_L), & \text{(4.26a)} \\ \Pi^{*,\theta} = \dfrac{(\Pi_R + \Pi_L)}{2} - \theta\dfrac{a}{2}(u_R - u_L). & \text{(4.26b)} \end{cases}$$

Ce schéma quasi-1D permet bien de retrouver les flux du schéma (4.7)-(4.8) en prenant pour u (resp. v) la vitesse normale (resp. la vitesse tangentielle) à la face où on estime le flux. De plus on a la propriété suivante :

Proposition 9. *Il existe un solver de Riemann approché simple pour le système* (4.24) *dont le flux numérique associé au sens de Harten, Lax, Van Leer correspond à celui du solver acoustique modifié* (4.25)-(4.26). *Plus précisément il existe une fonction auto-similaire*

$$\mathbf{W}^{\theta}_{RP}\left(\frac{m}{t};\mathbf{W}_L,\mathbf{W}_R\right)=(Y,\tau,u,v,E,\Pi)\left(\frac{m}{t};\mathbf{W}_L,\mathbf{W}_R\right)=\begin{cases}\mathbf{W}_L, & \text{if} \quad m/t<-a,\\ \mathbf{W}_L^{*,\theta}, & \text{if} \quad -a\le m/t<0,\\ \mathbf{W}_R^{*,\theta}, & \text{if} \quad 0\le m/t<+a,\\ \mathbf{W}_R, & \text{if} \quad +a\le m/t.\end{cases} \tag{4.27}$$

telle que

$$\begin{aligned}\mathbf{F}^{\theta}(\mathbf{W}_R,\mathbf{W}_L)&=\mathbf{F}(\mathbf{W}_L)-\int_{-\infty}^{0}[\mathbf{W}^{\theta}_{RP}(\xi;\mathbf{W}_L,\mathbf{W}_R)-\mathbf{W}_L]\,\mathrm{d}\xi\\ &=\mathbf{F}(\mathbf{W}_R)+\int_{0}^{+\infty}[\mathbf{W}^{\theta}_{RP}(\xi;\mathbf{W}_L,\mathbf{W}_R)-\mathbf{W}_R]\,\mathrm{d}\xi\\ &=\frac{1}{2}\left(\mathbf{F}(\mathbf{W}_L)+\mathbf{F}(\mathbf{W}_R)\right)-\frac{a}{2}\left(\mathbf{W}_L^{*,\theta}-\mathbf{W}_L\right)-\frac{a}{2}\left(\mathbf{W}_R-\mathbf{W}_R^{*,\theta}\right).\end{aligned} \tag{4.28}$$

Les états $\mathbf{W}_L^{*,\theta}=(Y_L^{*,\theta},\tau_L^{*,\theta},u_L^{*,\theta},v_L^{*,\theta},E_L^{*,\theta},\Pi_L^{*,\theta})^T$ *et* $\mathbf{W}_R^{*,\theta}=(Y_R^{*,\theta},\tau_R^{*,\theta},u_R^{*,\theta},v_R^{*,\theta},E_R^{*,\theta},\Pi_R^{*,\theta})^T$ *sont donnés par*

$$Y_L^{*,\theta}=Y_L, \qquad Y_R^{*,\theta}=Y_R, \tag{4.29a}$$

$$\tau_L^{*,\theta}=\tau_L+\frac{1}{a}(u^*-u_L), \qquad \tau_R^{*,\theta}=\tau_R+\frac{1}{a}(u_R-u^*), \tag{4.29b}$$

$$u_L^{*,\theta}=u^*+\frac{1}{2}(\theta-1)(u_R-u_L), \qquad u_R^{*,\theta}=u^*+\frac{1}{2}(1-\theta)(u_R-u_L), \tag{4.29c}$$

$$v_L^{*,\theta}=v_L, \qquad v_R^{*,\theta}=v_R, \tag{4.29d}$$

$$E_L^{*,\theta}=E_L+\frac{1}{a}(\Pi_L u_L-\Pi^{*,\theta}u^*), \qquad E_R^{*,\theta}=E_R+\frac{1}{a}(\Pi^{*,\theta}u^*-\Pi_R u_R) \tag{4.29e}$$

$$\Pi_L^{*,\theta}=\Pi^{*,\theta=1}, \qquad \Pi_R^{*,\theta}=\Pi^{*,\theta=1}. \tag{4.29f}$$

On définit alors les grandeurs $\Pi^*=\Pi_R^{*,\theta}=\Pi_L^{*,\theta}$, $(Y_L^*,\tau_L^*,v_L^*)=(Y_L^{*,\theta},\tau_L^{*,\theta},v_L^{*,\theta})$ *et* $(Y_R^*,\tau_R^*,v_R^*)=(Y_R^{*,\theta},\tau_R^{*,\theta},v_R^{*,\theta})$ *qui sont indépendantes de* θ.

Démonstration. La consistance au sens intégral de $\mathbf{W}^{\theta}_{\mathrm{RP}}$ avec le système (4.24), pour des états $\mathbf{W}_L$ et $\mathbf{W}_R$ donnés, s'écrit : $\mathbf{F}(\mathbf{W}_R)-\mathbf{F}(\mathbf{W}_L)=-a(\mathbf{W}_L^{*,\theta}-\mathbf{W}_L)+a(\mathbf{W}_R-\mathbf{W}_R^{*,\theta})$, ce qui donne

$$\mathbf{W}_R^{*,\theta}+\mathbf{W}_L^{*,\theta}=\mathbf{W}_R+\mathbf{W}_L-\frac{1}{a}(\mathbf{F}(\mathbf{W}_R)-\mathbf{F}(\mathbf{W}_L)). \tag{4.30}$$

Si le flux de ce solver de Riemann approché est $F^{\theta}(\mathbf{W}_L,\mathbf{W}_R)$, alors la formule (4.28) est vérifiée et on a

$$2F^{\theta}(\mathbf{W}_L,\mathbf{W}_R)=\mathbf{F}(\mathbf{W}_R)+\mathbf{F}(\mathbf{W}_L)-a(\mathbf{W}_L^{*,\theta}-\mathbf{W}_L)-a(\mathbf{W}_R-\mathbf{W}_R^{*,\theta})$$

ou de manière équivalente

$$\mathbf{W}_R^{*,\theta}-\mathbf{W}_L^{*,\theta}=\mathbf{W}_R-\mathbf{W}_L+\frac{1}{a}\left(2F^{\theta}(\mathbf{W}_L,\mathbf{W}_R)-\mathbf{F}(\mathbf{W}_L)-\mathbf{F}(\mathbf{W}_R)\right). \tag{4.31}$$

On utilise (4.30) et (4.31) pour obtenir

$$\mathbf{W}_L^{*,\theta}=\mathbf{W}_L-\frac{1}{a}(F^{\theta}(\mathbf{W}_L,\mathbf{W}_R)-\mathbf{F}(\mathbf{W}_L)), \quad \mathbf{W}_R^{*,\theta}=\mathbf{W}_R+\frac{1}{a}(F^{\theta}(\mathbf{W}_L,\mathbf{W}_R)-\mathbf{F}(\mathbf{W}_R)).$$

Les états $\mathbf{W}_L^{*,\theta}$ et $\mathbf{W}_R^{*,\theta}$ ainsi définis en fonction de $\mathbf{W}_L$ et $\mathbf{W}_R$ correspondent à (4.29) et vérifient bien (4.30) et (4.31), fournissant ainsi le résultat recherché. □

En utilisant ce *solver* de Riemann approché, on montre que le schéma acoustique modifié (4.25) est stable sous la condition CFL

$$2a\Delta t \leq \Delta m,$$

qui ne dépend pas de la modification θ. De plus, pour $\theta = 1$ la fonction auto-similaire $\mathbf{W}_{RP}^{\theta}$ définie dans la proposition 9 dégénère vers la solution exacte du problème de Riemann associé au système (4.24).

Par ailleurs, si on prend en compte la projection sur l'état d'équilibre de la stratégie de relaxation, on a $\Pi_L = p(\tau_L, e_L, Y_L)$, et $\Pi_R = p(\tau_R, e_R, Y_R)$. On peut alors montrer que les premières coordonnées (Y, τ, u, v, E) de la fonction auto-similaire $\mathbf{W}_{\mathrm{RP}}^{\theta}$ sont consistantes au sens intégral avec le système

$$\partial_t \mathbf{V} + \partial_m \mathbf{F}(\mathbf{V}) = 0, \tag{4.32}$$

où $\mathbf{V} = (Y, \tau, u, v, E)^T$ et $\mathbf{F}(\mathbf{V}) = (0, -u, p, 0, pu)^T$.
Finalement, le schéma numérique (4.7)-(4.8) peut se réécrire à l'aide des états intermédiaires du *solver* de Riemann approché de la proposition 9 sous la forme

$$\mathbf{W}_i^{n+1-} = \left(1 - \tau_i^n \Delta t \sum_{j \in N(i)} \sigma_{ij} a_{ij}\right) \mathbf{W}_i^n + \tau_i^n \Delta t \sum_{j \in N(i)} \sigma_{ij} a_{ij} \mathbf{W}_{ij}^{n,*,\theta}, \tag{4.33}$$

où $\mathbf{W}_{ij}^{n,*,\theta} = \mathbf{W}_L^{*,\theta}$ est donné par (4.29) pour $\mathbf{W}_L = \mathbf{W}_i^n$ et $\mathbf{W}_R = \mathbf{W}_j^n$, avec $u = \mathbf{n}_{ij}^T \mathbf{u}$ et $v = ||\mathbf{u} - (\mathbf{n}_{ij}^T \mathbf{u})\mathbf{n}_{ij}||$. Cette écriture est utile pour étudier les propriétés de stabilité du schéma acoustique. On montre en particulier une inégalité d'entropie discrète dans la section suivante.

4.B Inégalité d'entropie discrète

Dans cette section, on montre une inégalité d'entropie discrète pour le schéma LPS-EX(θ). On considère le système HRM (4.1). On note $\tau = 1/\rho$ et s l'entropie spécifique. On suppose donnée une équation d'état de mélange $(\tau, s, Y) \mapsto e^{\mathrm{EOS}}$ qui satisfait

$$\partial_\tau e^{\mathrm{EOS}} < 0, \quad \partial_s e^{\mathrm{EOS}} > 0, \quad \partial_{\tau\tau} e^{\mathrm{EOS}} > 0. \tag{4.34}$$

L'entropie de mélange $s = s^{\mathrm{EOS}}(\tau, e, Y)$ vérifie $e = e^{\mathrm{EOS}}(\tau, s, Y)$, on peut définir la pression $p = -\partial_\tau e^{\mathrm{EOS}}$ et la vitesse du son de mélange $c = \tau\sqrt{\partial_{\tau\tau} e^{\mathrm{EOS}}}$. On considère aussi p comme une fonction de (τ, e, Y) que l'on note $p = p^{\mathrm{EOS}}(\tau, e, Y)$. On suppose de plus que

$$(\tau, e, Y) \mapsto -s^{\mathrm{EOS}} \text{ est convexe.} \tag{4.35}$$

Finalement, comme $Y^*(\rho, e)$ correspond à l'équilibre thermodynamique, on suppose que la fonction $Y \mapsto s^{\mathrm{EOS}}(\tau, e, Y)$ est maximale en $Y = Y^*(\rho, e)$.

Dans la suite, on note $I(b, b') \subset \mathbb{R}$ l'intervalle entre $b \in \mathbb{R}$ et $b' \in \mathbb{R}$. On considère la condition sous-caractéristique

$$\begin{aligned} &\tau_L^* > 0, \quad -\partial_\tau p^{\mathrm{EOS}}(\tau, s_L, Y_L) \leq a^2, \quad \forall \tau \in I(\tau_L, \tau_L^*), \\ &\tau_R^* > 0, \quad -\partial_\tau p^{\mathrm{EOS}}(\tau, s_R, Y_R) \leq a^2, \quad \forall \tau \in I(\tau_R, \tau_R^*), \end{aligned} \tag{4.36}$$

et on commence par prouver deux résultats techniques.

Lemme 1. *On considère la solution du problème de Riemann* (4.24) *donnée par* (4.29) *pour* $\theta = 1$. *On*

suppose que (4.36) *est vérifiée. Soit* $s_k = s^{\mathrm{EOS}}(\tau_k, e_k, Y_k)$, $k = L, R$ *et* $e_k^* = E_k^{*,\theta=1} - (u_k^{*,\theta=1})^2/2 - (v_k^*)^2/2$, *on a*

$$e_k^* - e^{\mathrm{EOS}}(\tau_k^*, s_k, Y_k) - \frac{\left(p^{\mathrm{EOS}}(\tau_k^*, s_k, Y_k) - \Pi^*\right)^2}{2a^2} \geq 0. \tag{4.37}$$

Démonstration. On considère le cas $k = R$ et on définit pour $\tau \in I(\tau_R, \tau_R^*)$

$$\begin{aligned}\phi(\tau) = e^{\mathrm{EOS}}(\tau, s_R, Y_R) - \frac{p^{\mathrm{EOS}}(\tau, s_R, Y_R)^2}{2a^2} - e^{\mathrm{EOS}}(\tau_R^*, s_R, Y_R) + \frac{p^{\mathrm{EOS}}(\tau_R^*, s_R, Y_R)^2}{2a^2} \\ + p^{\mathrm{EOS}}(\tau_R^*, s_R, Y_R)\left(\tau + \frac{p^{\mathrm{EOS}}(\tau, s_R, Y_R)}{a^2} - \tau_R^* - \frac{p^{\mathrm{EOS}}(\tau_R^*, s_R, Y_R)}{a^2}\right).\end{aligned}$$

On a $\phi'(\tau) = \left(p^{\mathrm{EOS}}(\tau, s_R, Y_R) - p^{\mathrm{EOS}}(\tau_R^*, s_R, Y_R)\right)\left(1 - \rho^2 c^2(\tau, s_R, Y_R)/a^2\right)$. Si $\tau_R > \tau > \tau_R^*$ (resp. $\tau_R < \tau < \tau_R^*$), les conditions sur l'équation d'état (4.34) donnent $p^{\mathrm{EOS}}(\tau, s_R, Y_R) - p^{\mathrm{EOS}}(\tau_R^*, s_R, Y_R) < 0$ (resp. $p^{\mathrm{EOS}}(\tau, s_R, Y_R) - p^{\mathrm{EOS}}(\tau_R^*, s_R, Y_R) > 0$) en utilisant aussi la condition sous-caractéristique (4.36) on a $\phi'(\tau) \geq 0$ (resp. $\phi'(\tau) \leq 0$). Comme $\phi(\tau_R^*) = 0$, on obtient $\phi(\tau_R) > \phi(\tau_R^*) = 0$ pour $\tau \in I(\tau_R, \tau_R^*)$. En utilisant la relation de saut $(e_R^* - \frac{\Pi^*}{2a^2}) = (e_R - \frac{\Pi_R}{2a^2})$, on a $0 < \phi(\tau_R) = e_R^* - e^{\mathrm{EOS}}(\tau_R^*, s_R, Y_R) - \frac{1}{2a^2}(p^{\mathrm{EOS}}(\tau_R^*, s_R, Y_R) - \Pi^*)^2$. Le même raisonnement s'applique pour $k = L$. □

Lemme 2. *Soient* $\theta \in \mathbb{R}$ *et* $e_k^{*,\theta} = E_k^{*,\theta} - (u_k^{*,\theta})^2/2 - (v_k^*)^2/2$ *pour* $k = L, R$, *on a*

$$e_k^{*,\theta} - e^{\mathrm{EOS}}(\tau_k, s_k, Y_k) - \frac{1}{2a^2}\left(p^{\mathrm{EOS}}(\tau_k, s_k, Y_k) - \Pi^*\right)^2 + \frac{(1-\theta)^2(u_R - u_L)^2}{8} \geq 0, \quad k = L, R. \tag{4.38}$$

Démonstration. On a $u_R^{*,\theta} = u^* + (1-\theta)(u_R - u_L)/2$, $v_R^* = v_R$ et $\Pi^{*,\theta} = \Pi^* + (1-\theta)a(u_R - u_L)/2$. En utilisant (4.29) on obtient $e_R^{*,\theta} = e_R^* - (1-\theta)^2(u_R - u_L)^2/8$. On injecte cette relation dans (4.37) pour obtenir le résultat voulu. □

Il est maintenant clair que l'inégalité

$$-\frac{1}{2a^2}\left(p^{\mathrm{EOS}}(\tau_k^*, s_k, Y_k) - \Pi^*\right)^2 + \frac{(1-\theta)^2(u_R - u_L)^2}{8} \leq 0, \quad k = L, R \tag{4.39}$$

peut aider à obtenir une inégalité d'entropie discrète pour le schéma numérique modifié.

Proposition 10. *Soit* $s_k^{*,\theta} = s^{\mathrm{EOS}}(\tau_k^{*,\theta}, e_k^{*,\theta}, Y_k^{*,\theta})$ *pour* $k = L, R$. *Si la condition* (4.39) *est vérifiée, on a*

$$s_k^{*,\theta} \geq s_k. \tag{4.40}$$

L'inégalité (4.40) *implique que le schéma modifié* (4.33) *pour l'étape acoustique (qui peut aussi s'écrire* (4.7)-(4.8)*) vérifie l'inégalité d'entropie discrète suivante*

$$s^{\mathrm{EOS}}(\tau_j^{n+1-}, e_j^{n+1-}, Y_j^{n+1-}) \geq s^{\mathrm{EOS}}(\tau_j^n, e_j^n, Y_j^n). \tag{4.41}$$

Démonstration. Soit $k = L, R$, sous l'hypothèse (4.39), on a $e_k^{*,\theta} \geq e^{\mathrm{EOS}}(\tau_k^*, s_k, Y_k)$. De plus, on a $Y_k^{*,\theta} = Y_k$, $\tau_k^{*,\theta} = \tau_k^*$ et donc $e^{\mathrm{EOS}}(\tau_k^*, s_k, Y_k) = e^{\mathrm{EOS}}(\tau_k^{*,\theta}, s_k, Y_k^{*,\theta})$. D'après les conditions sur la loi d'état (4.34) $\epsilon \mapsto s^{\mathrm{EOS}}(\tau_k^{*,\theta}, \epsilon, Y_k^{*,\theta})$ est croissante, donc $s^{\mathrm{EOS}}(\tau_k^{*,\theta}, e_k^{*,\theta}, Y_k^{*,\theta}) \geq s^{\mathrm{EOS}}(\tau_k^{*,\theta}, e^{\mathrm{EOS}}(\tau_k^{*,\theta}, s_k, Y_k^{*,\theta}), Y_k^{*,\theta}) = s_k$ et on a bien l'inégalité (4.40).
Sous la condition de stabilité CFL (4.9), le schéma (4.33) pour l'étape acoustique est une combinaison convexe. De plus, comme les fonctions $u \mapsto -\frac{u^2}{2}$ et $v \mapsto -\frac{v^2}{2}$ sont concaves, on a

$$e_i^{n+1-} \geq \tilde{e}_i^{n+1-},$$

où $e_{ij}^{n,*,\theta} = E_{ij}^{n,*,\theta} - (u_{ij}^{n,*,\theta})^2/2 - (v_{ij}^{n,*,\theta})^2/2$ et

$$\tilde{e}_i^{n+1-} = \left(1 - \tau_i^n \Delta t \sum_{j\in N(i)} \sigma_{ij} a_{ij}\right) e_i^n + \tau_i^n \Delta t \sum_{j\in N(i)} \sigma_{ij} a_{ij} e_{ij}^{n,*,\theta}.$$

Comme $\epsilon \mapsto s^{\mathrm{EOS}}(\tau_j^{n+1-}, \epsilon, Y_j^{n+1-})$ est croissante, on a alors

$$s^{\mathrm{EOS}}(\tau_j^{n+1-}, e_j^{n+1-}, Y_j^{n+1-}) \geq s^{\mathrm{EOS}}(\tau_j^{n+1-}, \tilde{e}_j^{n+1-}, Y_j^{n+1-}). \tag{4.42}$$

La définition de $\tilde{e}_j^{n+1-}$ et le schéma (4.33) pour les variables τ et Y sont des combinaisons convexes, on utilise le fait que la fonction $(\tau, s, Y) \mapsto s^{\mathrm{EOS}}$ est concave pour obtenir

$$s^{\mathrm{EOS}}(\tau_j^{n+1-}, \tilde{e}_j^{n+1-}, Y_j^{n+1-}) \geq \left(1 - \tau_i^n \Delta t \sum_{j\in N(i)} \sigma_{ij} a_{ij}\right) s_i^n + \tau_i^n \Delta t \sum_{j\in N(i)} \sigma_{ij} a_{ij} s_{ij}^{n,*,\theta},$$

où $s_{ij}^{n,*,\theta} = s^{\mathrm{EOS}}(\tau_{ij}^{n,*,\theta}, e_{ij}^{n,*,\theta}, Y_{ij}^{n,*,\theta})$. On injecte l'inégalité (4.40) pour obtenir

$$s^{\mathrm{EOS}}(\tau_j^{n+1-}, \tilde{e}_j^{n+1-}, Y_j^{n+1-}) \geq s^{\mathrm{EOS}}(\tau_j^n, e_j^n, Y_j^n). \tag{4.43}$$

On utilise alors (4.42) et (4.43) pour obtenir l'inégalité d'entropie discrète voulue (4.41). □

On peut maintenant proposer une inégalité d'entropie discrète pour le schéma LPS-EX(θ).

Proposition 11. *Si les conditions* (4.39), (4.9) *et* (4.13) *sont vérifiées, alors le schéma LPS-EX(θ) constitué des étapes* (4.7)-(4.11)-(4.14) *vérifie l'inégalité d'entropie discrète suivante*

$$\begin{aligned}\rho_i^{n+1} s^{\mathrm{EOS}}(\tau_i^{n+1}, e_i^{n+1}, Y_i^{n+1}) \geq \rho_i^n s^{\mathrm{EOS}}(\tau_i^n, e_i^n, Y_i^n) - \Delta t \sum_{j\in N(i)} \sigma_{ij} (u_{ij}^*)^+ \rho_i^{n+1-} s^{\mathrm{EOS}}(\tau_i^{n+1-}, e_i^{n+1-}, Y_i^{n+1-}) \\ - \Delta t \sum_{j\in N(i)} \sigma_{ij} \left(u_{ij}^*\right)^- \rho_j^{n+1-} s^{\mathrm{EOS}}(\tau_j^{n+1-}, e_j^{n+1-}, Y_j^{n+1-}).\end{aligned} \tag{4.44}$$

Démonstration. D'après la proposition 10, on a directement pour l'étape acoustique (4.7)

$$s^{\mathrm{EOS}}(\tau_j^{n+1-}, e_j^{n+1-}, Y_j^{n+1-}) \geq s^{\mathrm{EOS}}(\tau_j^n, e_j^n, Y_j^n). \tag{4.45}$$

L'étape de projection (4.11) peut se réécrire pour $\varphi \in \{\rho Y, \rho, \rho \mathbf{u}, \rho E\}$ sous la forme

$$\overline{\varphi}_i = \left(1 + \Delta t \sum_{j\in N(i)} \sigma_{ij} (u_{ij}^*)^-\right) \varphi_i^{n+1-} - \Delta t \sum_{j\in N(i)} \sigma_{ij} \left(u_{ij}^*\right)^- \varphi_j^{n+1-},$$

qui est une combinaison convexe sous la condition CFL (4.13). Comme $(\rho Y, \rho, \rho u, \rho v, \rho E) \mapsto (\rho s^{\mathrm{EOS}})(\tau, e, Y)$ est concave, on a alors directement

$$\begin{aligned}\overline{\rho}_i s^{\mathrm{EOS}}(\overline{\tau}_i, \overline{e}_i, \overline{Y}_i) \geq \left(1 + \Delta t \sum_{j\in N(i)} \sigma_{ij} (u_{ij}^*)^-\right) \rho_i^{n+1-} s^{\mathrm{EOS}}(\tau_i^{n+1-}, e_i^{n+1-}, Y_i^{n+1-}) \\ - \Delta t \sum_{j\in N(i)} \sigma_{ij} \left(u_{ij}^*\right)^- \rho_j^{n+1-} s^{\mathrm{EOS}}(\tau_j^{n+1-}, e_j^{n+1-}, Y_j^{n+1-}).\end{aligned} \tag{4.46}$$

Pour l'étape de transition de phase (4.14), on a la combinaison convexe suivante

$$Y_i^{n+1} = \frac{1}{1+\lambda_0\Delta t}\overline{Y}_i + \frac{\lambda_0\Delta t}{1+\lambda_0\Delta t}Y^*(\overline{\rho}_i, \overline{e}_i).$$

Comme $(\tau, e, Y) \mapsto s^{\mathrm{EOS}}(\tau, e, Y)$ est concave, on a donc

$$s^{\mathrm{EOS}}(\tau_i^{n+1}, e_i^{n+1}, Y_i^{n+1}) \geq \frac{1}{1+\lambda_0\Delta t}s^{\mathrm{EOS}}(\tau_i^{n+1}, e_i^{n+1}, \overline{Y}_i) + \frac{\lambda_0\Delta t}{1+\lambda_0\Delta t}s^{\mathrm{EOS}}(\tau_i^{n+1}, e_i^{n+1}, Y^*(\overline{\rho}_i, \overline{e}_i)).$$

On injecte dans cette équation $\tau_i^{n+1} = \overline{\tau}_i$, $e_i^{n+1} = \overline{e}_i$ et le fait que la fonction $Y \mapsto s^{\mathrm{EOS}}(\tau, e, Y)$ est maximale en $Y = Y^*(\rho, e)$ pour obtenir

$$\rho_i^{n+1} s^{\mathrm{EOS}}(\tau_i^{n+1}, e_i^{n+1}, Y_i^{n+1}) \geq \overline{\rho}_i s^{\mathrm{EOS}}(\overline{\tau}_i, \overline{e}_i, \overline{Y}_i). \tag{4.47}$$

On combine (4.45), (4.46), (4.47) et le schéma acoustique pour τ (4.7) afin d'obtenir l'inégalité d'entropie discrète pour le schéma complet (4.44). □

On a montré dans la proposition 11 une inégalité d'entropie discrète pour le schéma LPS-EX(θ) sous la condition (4.39) sur la modification θ. L'étude des propriétés de stabilité L^2 des schémas LPS-EX(θ) et LPS-IMEX(θ) pour toute valeur de $\theta \geq 0$ est un problème ouvert.

Bibliographie

[1] A. Ambroso, C. Chalons, F. Coquel, E. Godlewski, F. Lagoutiere and P.-A. Raviart, *The coupling of homogeneous models for two-phase flows*, Int. Journal for Finite Volume, 4(1) : 39–54, (2007).

[2] M. Bernard, S. Dellacherie, G. Faccanoni, B. Grec and Y. Penel, *Study of a low Mach nuclear core model for two-phase flows with phase transition I : stiffened gas law*, ESAIM : Mathematical Modelling and Numerical Analysis, 48(6) : 1639–1679, (2014).

[3] S. Dellacherie, P. Omnes and P.-A. Raviart, *Construction of modified Godunov type schemes accurate at any Mach number for the compressible Euler system*, submitted, hal-00776629, (2013).

[4] Z. Bilicki and J. Kestin, *Physical aspects of the relaxation model in two phase flow*, Proc. R. Soc. Lond., A428 : 379–397, (1990).

[5] F. Bouchut, Nonlinear stability of finite volume methods for hyperbolic conservation laws, and well-balanced schemes for sources, Frontiers in Mathematics series. Birkhäuser, 2004.

[6] C. Chalons and J.-F. Coulombel, *Relaxation approximation of the Euler equations*, Journal of Mathematical Analysis and Applications, 348(2) : 872–893, (2008).

[7] C. Chalons, M. Girardin, and S. Kokh, *Large time step and asymptotic preserving numerical schemes for the gas dynamics equations with source terms*, SIAM J. Sci. Comput., 35(6) : a2874–a2902, (2013).

[8] C. Chalons, M. Girardin and S. Kokh, *An all-regime Lagrange-Projection like scheme for the gas dynamics equations on unstructured meshes*, submitted, hal-01007622, (2014).

[9] D. Chauveheid, J.-P. Braeunig and J.-M. Ghidaglia, *A totally Eulerian Finite Volume solver for multi-material fluid flows III : the low Mach number case*, to appear in European Journal of Mechanics-B/Fluids, (2013).

[10] S. Clerc, *Numerical simulation of the homogeneous equilibrium model for two-phase flows*, J. Comput. Phys., 161(1) : 354–375, (2000).

[11] F. Coquel, Q.-L. Nguyen, M. Postel, and Q.H. Tran, *Entropy-satisfying relaxation method with large time-steps for Euler IBVPs*, Math. Comput., 79(271) : 1493–1533, (2010).

[12] F. Cordier, P. Degond and A. Kumbaro, *An Asymptotic-Preserving all-speed scheme for the Euler and Navier-Stokes equations*, J. Comput. Phys., 231(17) : 5685–5704, (2012).

[13] P. Degond, S. Jin, and J.-G. Liu, *Mach-number uniform asymptotic-preserving gauge schemes for compressible flows*, Bull. Inst. Math., Acad. Sin. (N.S.), 2(4) : 851–892, (2007).

[14] P. Degond and M. Tang, *All speed method for the Euler equation in the low Mach number limit*, Commun. Comp. Phys., 10 : 1–31, (2011).

[15] S. Dellacherie, *Analysis of Godunov type schemes applied to the compressible Euler system at low Mach number*, J. Comput Phys., 229(4) : 978–1016, (2010).

[16] P. Downar-Zapolski, Z. Bilicki, L. Bolle and J. Franco, *The non equilibrium model for one-dimensional flashing liquid flow*, Int. J. multiphase Flow, 22(3) : 473–483, (1996).

[17] N. Grenier, J.-P. Vila and P. Villedieu, *An accurate low-Mach scheme for a compressible two-fluid model applied to free-surface flows*, J. Comput. Phys., 252 : 1–19, (2013).

[18] H. Guillard and C. Viozat, *On the behavior of upwind schemes in the low Mach limit*, Comput. Fluid., 28 : 63–86, (1999).

[19] A. Harten, P.-D. Lax, and B. Van Leer, *On upstream differencing and godunov-type schemes for hyperbolic conservation laws*, SIAM Review, 25 : 35–61, (1983).

[20] S. LeMartelot, B. Nkonga and R. Saurel, *Liquid and liquid-gas flows at all speeds*, J. Comput. Phys., 255 : 53–82, (2013).

[21] J.-G. Liu J. Haack and S. Jin, *An all-speed asymptotic-preserving method for the isentropic Euler and Navier-Stokes equations*, Commun. Comp. Phys., 12 : 955–980, (2012).

[22] I. Suliciu, *On the thermodynamics of fluids with relaxation and phase transitions. Fluids with relaxation*, Int. J. Engag. Sci., 36 : 921–947, (1998).

[23] E. Turkel, *Preconditioned methods for solving the incompressible and low speed compressible equations*, J. Comp. Phys., 72(2) : 277–298, (1987).

Chapitre 5

Implémentation

Dans ce chapitre, nous allons nous intéresser à l'implémentation des schémas numériques présentés dans les chapitres précédents. Cette implémentation a été effectuée dans le cadre du code *YAFiVoC*. Le langage de programmation de ce code que nous avons développé avec S. Kokh est le *C* et la compilation s'effectue à l'aide de *CMake*. Nous rappelons que les schémas présentés dans les chapitres précédents sont explicites et/ou implicites en temps, et définis en 1D ou en 2D sur maillages non-structurés. L'objectif de ce court chapitre est de donner un aperçu des différents aspects du code utilisé principalement lors de cette thèse et notamment des structures de données, ainsi que de l'implémentation de quelques méthodes.

5.1 YAFiVoC

YAFiVoC (Yet Another Finite Volume Code) est une bibliothèque dédiée à l'implémentation de méthodes de type volumes finis en 2D et sur maillages non-structurés. Cette bibliothèque permet d'effectuer les opérations de base relatives à ces méthodes :

- lecture de fichier de configuration,
- lecture de maillage,
- lecture de données décrivant la condition initiale,
- sauvegarde des données,
- système de stockage des données via des tableaux dont les éléments sont des vecteurs d'inconnues,
- interface de programmation (API, *Application Programming Interface*) pour accéder aux éléments du maillage, de la géométrie, aux éléments décrivant les conditions aux limites, l'accès aux inconnues associés aux faces et aux arêtes, syntaxe non-ambigue pour les boucles,
- un système de description des lois d'états via des pointeurs de fonction dont l'affectation est réalisé dynamiquement par un système de *plugins*,
- un système de description via pointeur de fonction de la mise à jour du temps t^n au temps t^{n+1} des variables dont l'affectation est réalisé dynamiquement par un système de *plugins*.

Le fichier de configuration est basé sur le format libconfig de la bibliothèque libconfig projet open source. Les formats des fichiers contenant le maillage et les données d'entrée et de sortie sont spécifiés dans le fichier de configuration utilisés par *YAFiVoC* lors de l'exécution. Dans le cadre de cette thèse on a utilisé les formats suivants :

- VTK (*legacy unstructured mesh* en ASCII) pour la donnée initiale et la sauvegarde de données à un temps ultérieur,
- format du maillage 2D Triangle pour les maillages triangulaires,
- format interne à *YAFiVoC* (inspiré du format 2D Triangle) pour les maillages de quadrangles. Il n'y a pas véritablement de logiciel de maillage sous-jacent à proprement parlé, mais des modules python permettant de créer des maillages de quadrangle au format correspondant.

Un aspect important de *YAFiVoC* est sa modularité. En effet, l'utilisation de pointeurs de fonctions permet de séparer les fichiers contenant le code associé aux fonctionnalités mentionnées précédemment des fichiers contenant le code associé à un *solver* ou une loi d'état spécifique.

Pour implémenter un nouveau schéma numérique, il suffit de créer à partir d'un *template* de base les fichiers qui précisent la définition de la fonction permettant de mettre à jour la solution entre les temps t^n et t^{n+1}.

De même, pour implémenter une nouvelle loi d'état, on crée à partir d'un *template* de base les fichiers qui précisent la définition des fonctions thermodynamiques (pression, vitesse du son, enthalpie, etc.).

Une fois les fichiers de notre nouveau *solver* ou de notre nouvelle loi d'état créés, il faut les ajouter dans le fichier **CMakeLists.txt** utilisé par *CMake* lors de la compilation.

Une fois la compilation effectuée, on crée dans un même *directory* (dit *directory* d'exécution)

- un fichier de configuration,
- un fichier qui décrit l'état initial du fluide,
- des fichiers qui décrivent le maillage.

On exécute alors le binaire principal *YAFiVoC.exe* en lui donnant comme argument le *directory* d'exécution. Le fichier de configuration permet de fixer certains paramètres de la simulation tels que

- **tMax**, le temps final de la simulation,

- **nbSaveMax**, le nombre de solutions intermédiaires que l'on souhaite sauvegarder,
- **maillage**, le type de maillage et la liste des fichiers du maillage,
- **solver**, le choix du *solver* utilisé pour calculer le pas de temps et mettre à jour la solution,
- **eos**, le choix de l'équation d'état utilisée pour calculer la thermodynamique.

On utilise finalement *ParaView* pour afficher les données sauvegardées au format VTK.

5.2 Structures de données

La structure de données principale de *YAFiVoC* est **problem_t**. Elle contient des pointeurs vers toutes les autres données et en particulier vers :

- les arguments passés à l'exécutable,
- le temps actuel, le pas de temps et le temps final,
- la solution actuelle,
- le maillage,
- les données du *solver* de type **solverData_t** (la définition de ce type est spécifique à chaque *solver* afin de répondre aux besoins des différentes méthodes numériques).

Une autre structure importante est **mesh_t**. Comme son nom l'indique, ce type correspond au maillage et contient :

- une liste des noeuds de type **node_t**,
- une liste des côtés de type **edge_t**,
- une liste des cellules de type **cell_t**,
- des listes qui contiennent les côtés situés sur le bord du domaine classés par région.

Ces structures de données permettent d'accéder aux informations usuelles sur la connectivité des éléments du maillage.

La structure **varArray_t** permet de créer des tableaux de données associés à un nombre d'éléments $nbElts$ (par exemple le nombre de cellules ou le nombre de faces du maillage) et à un certain nombre de variables $nbVar$.

5.3 Implémentation de la méthode IMEX(θ)

On présente ici, dans un soucis d'illustration, l'implémentation du schéma numérique IMEX(θ) introduit au cours du chapitre 3. Pour ajouter IMEX(θ) à la liste des *solvers* disponibles dans *YAFiVoC*, on crée les fichiers suivants :

- **LagProjIMEXSolver.h** qui contient la définition du type **solverData_t** et la déclaration des fonctions spécifiques à ce *solver*,
- **DefineFluxes.c** qui contient la définition des fonctions spécifiques à ce *solver*,
- **LagProjIMEXSolver.c** qui contient la définition de la fonction **MySolver** qui est appelée dans la boucle principale en temps pour passer la solution du temps t^n au temps t^{n+1}. Cette fonction appelle les fonctions définies dans **DefineFluxes.c** et utilise la donnée de type **solverData_t**,
- **Initialize_FinalizeSolver.c** qui contient les constructeurs et destructeurs associés au type **solverData_t**. C'est dans ce constructeur que l'on peut lire les arguments passés dans le document de configuration **config.cfg** à l'aide du **SimpleConfigReader** de *YAFiVoC*,
- **CMakeLists.txt** qui est nécessaire pour compiler le code avec *CMake*. Il faut ensuite ajouter le dossier contenant tout ces fichiers à la liste des dossiers à parcourir dans le **CMakeLists.txt**

principal.

Pour le *solver* IMEX(θ), la structure de donnée **SolverData_t** contient :

- **aRelax**, le coefficient de relaxation a_{jk} sur chaque face du maillage,
- **theta**, la modification θ_{jk} sur chaque face du maillage,
- **UStar**, la valeur de la vitesse d'interface u^*_{jk} sur chaque face du maillage,
- **PStar**, la valeur de la pression d'interface $\Pi^{*,\theta}_{jk}$ sur chaque face du maillage,
- **upwindVal**, la valeur décentrée amont φ^{n+1-}_{jk} sur chaque face du maillage pour $\varphi \in \{\rho, \rho\mathbf{u}, \rho E\}$,
- **A**,**x** et **b**, qui sont respectivement la matrice, la solution et le membre de droite du système linéaire A x=b que l'on doit résoudre lors de l'étape Lagrangienne. On utilise la bibliothèque *PETSc*, et plus précisément les types **Mat** et **Vec** pour le stockage et l'assemblage des matrices et des vecteurs,
- **coefCFL**, coefficient utilisé pour calculer le pas de temps à partir de la condition de stabilité CFL,
- **coeftheta**, coefficient utilisé pour choisir la valeur de la modification θ_{jk},
- **coefBC**, coefficient utilisé pour savoir quelles conditions aux limites doivent être imposées.

Les valeurs de **coefCFL**, **coeftheta** et **coefBC** sont lues dans le fichier de configuration **config.cfg**.

Dans **DefineFluxes.c**, on définit les fonctions suivantes :

- **void DefineRelaxationParameter(problem_t* pb)**
  Cette fonction calcule **aRelax**, le coefficient de relaxation sur chaque face avec (3.56a).

- **void DefineModificationParameter(problem_t* pb)**
  Cette fonction calcule **theta**, la modification sur chaque face en fonction du choix précisé dans le fichier de configuration grâce à **coeftheta**.

- **void ComputeTimeStep(problem_t* pb)**
  Cette fonction calcule le pas de temps **pb→dt** à partir d'une évaluation explicite de la condition CFL (3.66).

- **void AssembleMatrixAndRHS(problem_t* pb)**
  Cette fonction assemble la matrice **A** et le membre de droite **b** du système linéaire (3.60).

- **void SolveLinearSystem(problem_t* pb)**
  Cette fonction calcule **x** la solution approchée du système linéaire $\mathbf{Ax} = \mathbf{b}$. On utilise pour cela la bibliothèque *PETSc*, et plus précisément le type **KSP**. La méthode de Krylov retenue est BICG-Stab avec un préconditionneur de Jacobi. On initialise cette méthode itérative à l'aide de la solution au temps t^n.

- **void ComputeUStarAndPStar(problem_t* pb)**
  Cette fonction calcule **UStar** et **PStar** sur chaque face à l'aide de la solution du problème linéaire **x** et des formules (3.56b)-(3.56c).

- **void UpdateLagrangian(problem_t* pb)**
  Cette fonction calcule la solution à la fin de l'étape Lagrangienne à l'aide de (3.55).

- **void ComputeUpwindVal(problem_t* pb)**
  Cette fonction calcule **upwindVal**, les valeurs décentrées amont sur chaque face pour l'étape de projection (3.57).

- **void UpdateProjection(problem_t* pb)**
  Cette fonction calcule la solution à la fin de l'étape de projection à l'aide de (3.57).

Finalement, la fonction **MySolver** qui permet de passer la solution du temps t^n au temps t^{n+1} à l'aide du schéma IMEX(θ) s'écrit simplement

```
void MySolver(problem_t* pb)
{
//Calcul du pas de temps et des paramètres
DefineRelaxationParameter(pb) ;
DefineModificationParameter(pb) ;
ComputeTimeStep(pb) ;

//Étape Lagrangienne
AssembleMatrixAndRHS(pb) ;
SolveLinearSystem(pb) ;
ComputeUStarAndPStar(pb) ;
UpdateLagrangian(pb) ;

//Étape de projection
ComputeUpwindVal(pb) ;
UpdateProjection(pb) ;
}
```

5.4 Implémentation pour les modèles diphasiques

La bibliothèque *YAFiVoC* a initialement été écrite pour le système d'Euler, c'est pourquoi la solution est stockée sur un **varArray_t** de taille fixée **nbVarEuler**. Les indices des différentes grandeurs (densité, vitesses, énergie, pression, ...) sont stockés dans des constantes globales et on dispose de fonctions d'accès et d'écritures spécifiques à ces tableaux de données. Par ailleurs, les pointeurs de fonctions thermodynamiques sont eux aussi associés au cas des équations d'Euler.

L'implémentation des schémas du chapitre 4 pour les modèles homogénéisés diphasiques HRM/HEM a nécessité de modifier en profondeur la bibliothèque YAFiVoC. Il a fallu en particulier changer le nombre de variables et modifier le nombre d'arguments des pointeurs de fonctions associées à la thermodynamique du code. Ces modifications sont inhérentes au changement de système d'équations. Néanmoins, une fois ces modifications de *YAFiVoC* effectuées, on peut aisément implémenter de nouveaux *solvers* et de nouvelles lois d'état sans avoir à modifier à nouveau le code de *YAFiVoC*.

Perspectives

On détaille ici trois pistes de réflexions qui constituent le prolongement naturel des travaux menés au cours de cette thèse.

Une première piste est la montée en ordre en temps et en espace de schémas *asymptotic preserving* ou tout-régime. De premiers éléments sur la montée à l'ordre deux en espace ont été donnés dans l'annexe 2.C. Néanmoins, on a constaté lorsqu'on s'est intéressé aux régimes intermédiaires dans l'annexe 2.B, que l'ordre de convergence d'un schéma pouvait être différent sur maillage grossier $h \gg \epsilon$ et sur maillage fin $h \ll \epsilon$. De plus, l'ordre de convergence d'un schéma modifié sur maillage grossier $h \gg \epsilon$ dépend fortement du choix de la modification $\theta(\epsilon)$.

Une seconde piste est l'obtention d'un critère sur le choix de la modification $\theta(\epsilon)$, garantissant certaines propriétés de stabilité (L^2, TVD, ...), moins restrictif que le critère garantissant l'inégalité d'entropie discrète obtenu au cours de cette thèse. On pourrait alors choisir une valeur optimale du paramètre θ qui pilote la stratégie anti-diffusive, en prennant la plus petite valeur vérifiant ce critère de stabilité.

La troisième piste est l'extension des méthodes construites au cours de cette thèse à des systèmes diphasiques plus complexes comme par exemple le système de Baer-Nunziato. Ce type de système soulève de nouvelles questions liées aux termes non conservatifs ainsi qu'à la nécessité d'utiliser deux paramètres, à savoir le nombre de Mach dans chaque phase, au lieu d'un.