Analysis and development of finite volume schemes asymptotically preserving in the low Mach number limit for the Euler and Navier-Stokes equations

Paola Allegrini

THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Paola ALLEGRINI

Le 20 septembre 2023

Analyse et développement de schémas Volumes Finis asymptotiquement préservants dans la limite bas-Mach pour les équations d'Euler et de Navier-Stokes

---

Ecole doctorale : **EDMITT - Ecole Doctorale Mathématiques, Informatique et Télécommunications de Toulouse**

Spécialité : **Mathématiques et Applications**

Unité de recherche :
IMT : Institut de Mathématiques de Toulouse

Thèse dirigée par
Marie-Helene VIGNAL

Jury

M. Christophe CHALONS, Rapporteur
M. Angelo IOLLO, Rapporteur
M. Vincent PERRIER, Examinateur
Mme Marie-Hélène VIGNAL, Directrice de thèse
M. Raphaël LOUBERE, Président

Remerciements

Premièrement, je tiens à remercier ma directrice de thèse, Marie-Hélène Vignal, qui m'a encadrée durant ces quatre années.
Merci Marie-Hélène pour m'avoir offert l'opportunité de réaliser cette thèse. Tu as toujours pris le temps de m'écouter, de me guider et de répondre à mes questions. Tu as su me mettre en confiance pour présenter mon travail et m'encourager dans les moments difficiles.
Je remercie Christophe Chalons et Angelo Iollo d'avoir consacré leur temps à la lecture et à l'évaluation de mon travail.
Je remercie Raphaël Loubère et Vincent Perrier qui ont accepté de participer au jury.
Je remercie l'ensemble des membres de l'Institut de Mathématiques de Toulouse.
Je remercie les doctorants et post-doctorants, anciens, visiteurs et actuels, avec qui j'ai passé la majeure partie de mon temps à Toulouse. En particulier, Corentin et Perla, avec qui j'ai partagé le bureau 202. Perla, nous sommes restées un an de plus ensemble, merci pour ton écoute et tes encouragements.
Je souhaite également exprimer ma reconnaissance envers les personnes rencontrées à Toulouse qui m'ont apporté leur soutien tout au long de cette aventure.
Un grand merci s'adresse également à mes amis qui sont venus me voir dans les moments les plus difficiles et les plus joyeux de cette période. En particulier Priya, tu as toujours répondu présente lorsque j'en avais besoin.
Enfin, je tiens à exprimer toute ma gratitude envers ma famille, qui a toujours été à mes côtés. Alberto, tu en fais désormais partie, merci pour ton soutien.

Contents

List of Figures

List of Tables

Chapitre 1

Chapitre Introductif

Contents

1.1 Contexte général

L'étude et la simulation des écoulements fluides revêtent une grande importance dans de nombreux domaines, tels que l'aéronautique, l'ingénierie des fluides et la météorologie. Cependant, lorsqu'il s'agit de simuler des écoulements à faibles nombres de Mach, c'est-à-dire des écoulements à des vitesses bien inférieures à la vitesse du son, des problèmes numériques spécifiques se posent.
En effet, lorsque le fluide atteint de faibles nombres de Mach, les ondes acoustiques se propagent très rapidement par rapport au fluide. Cela peut se produire sur seulement une partie du domaine et peut varier en fonction du temps, introduisant ainsi des contraintes numériques importantes. Les schémas numériques explicites classiquement utilisés perdent alors en précision et en stabilité. Ils perdent aussi en consistance dans la limite des faibles nombres de Mach et nécessitent souvent des pas de temps extrêmement petits pour suivre ces ondes acoustiques rapides, ce qui entraîne des coûts de calcul prohibitifs.
Dans cette perspective, cette thèse se concentre sur le développement et l'étude d'un schéma numérique préservant de bonnes propriétés de stabilité et de consistance lorsque le fluide évolue vers des régimes à faibles nombres de Mach. Nous

commencerons par nous intéresser aux équations d'Euler complet, qui permettent la modélisation de fluides compressibles et non visqueux. Nous étendrons ensuite notre travail au cas du modèle de Navier-Stokes, prenant en compte les effets visqueux du fluide.

1.2 Les équations d'Euler complet

Les équations d'Euler complet permettent de décrire l'écoulement de fluides compressibles et non visqueux en milieu continu. Elles permettent de modéliser et d'analyser le comportement des fluides en mouvement, qu'il s'agisse de l'écoulement de l'air autour d'un avion, de l'eau dans un canal ou même du sang dans les vaisseaux sanguins. Les équations sont obtenues à partir des lois de conservation de la masse, de la quantité de mouvement et de l'énergie du fluide considéré :

$$\partial_t \rho + \nabla_x \cdot q = 0, \tag{1.1a}$$

$$\partial_t q + \nabla_x \cdot \left(\frac{q \otimes q}{\rho} \right) + \nabla_x p = 0, \tag{1.1b}$$

$$\partial_t E + \nabla_x \cdot \left((E + p) \, \frac{q}{\rho} \right) = 0, \tag{1.1c}$$

où $\rho(t,x) > 0$ est la densité du fluide, $q(t,x) = \rho(t,x)u(t,x)$ sa quantité de mouvement, $u(t,x)$ le vecteur de vitesse de taille d (où d est la dimension de l'espace), $E(t,x)$ l'énergie totale et $p(t,x)$ la pression du fluide avec $t \in \mathbb{R}$ et $x \in \mathbb{R}^d$ les variables de temps et espace. Nous fermons le système (nous sommes avec quatre variables et trois équations) avec une loi d'état reliant la pression aux autres variables ρ, q et E. Dans ce travail, nous considérons celle des gazs parfaits :

$$E = \frac{p}{\gamma - 1} + \frac{1}{2} \frac{|q|^2}{\rho}, \tag{1.1d}$$

où $\gamma = \dfrac{c_p}{c_v}$ est le rapport des chaleurs spécifiques c_p et c_v à pression et volume constant, respectivement.
En fonction du phénomène physique que l'on cherche à étudier, la vitesse du fluide peut prendre de très grandes ou de très petites vitesses. Dans le cas où sa vitesse est très petite comparée à la vitesse du son dans le fluide, c, celui-ci devient faiblement compressible et même incompressible. Le rapport entre ces vitesses est donné par le nombre de Mach

$$M = \frac{|u|}{c},$$

où $|u|$ est la norme du vecteur vitesse. A la limite, lorsque M tend vers 0, le fluide devient incompressible, et dans ce cas, le modèle d'Euler incompressible peut être utilisé pour décrire l'écoulement. Cependant, il existe des situations dans lesquelles le fluide peut être légèrement compressible, ou des situations dans lesquelles le fluide présente des ondes de pression très rapides dans certaines régions du domaine et pas

dans d'autres, ou même des ondes rapides qui ne sont présentes que pendant de courtes durées par rapport à l'échelle d'observation. La valeur du nombre de Mach varie en fonction de l'espace et du temps, et seul le modèle d'Euler compressible est valable pour toutes les valeurs du nombre de Mach. Ces situations se produisent notamment dans le contexte des écoulements géophysiques et environnementaux. Il s'agit donc d'un problème multi-échelles et la simulation numérique de ces écoulements devient très complexe.
Dans ce travail, je m'intéresse à l'analyse et au développement de méthodes numériques robustes et efficaces dans tous les régimes, allant du Mach lorsque le fluide passe d'un régime compressible à incompressible et vice-versa. Pour mettre en évidence cette transition, nous travaillons sur le système adimensionné en fonction du nombre de Mach que je présente dans le paragraphe ci-dessous.

1.3 Adimensionnenement des équations d'Euler complet

Pour adimensionner le modèle d'Euler compressible nous introduisons les quantités suivantes :

$$\begin{aligned} \tilde{\rho} = \rho/\rho_0, \qquad \tilde{u} = u/u_0, \qquad \tilde{p} = p/p_0, \qquad \tilde{E} = E/p_0, \\ \tilde{x} = x/x_0, \qquad \tilde{t} = t/t_0, \end{aligned} \tag{1.2}$$

où ρ_0, p_0, x_0, t_0 et $u_0 = x_0/t_0$ sont les ordres de grandeur des valeurs prises par le fluide pour les situations considérées. Alors

$$\frac{\partial \rho}{\partial t}(x,t) = \frac{\partial}{\partial t}\left(\rho_0 \tilde{\rho}(\tilde{x},\tilde{t})\right) = \rho_0 \frac{\partial \tilde{t}}{\partial t}\frac{\partial \tilde{\rho}}{\partial \tilde{t}}(\tilde{x},\tilde{t}) = \frac{\rho_0}{t_0}\frac{\partial \tilde{\rho}}{\partial \tilde{t}}(\tilde{x},\tilde{t}) = \frac{\rho_0 u_0}{x_0}\frac{\partial \tilde{\rho}}{\partial \tilde{t}}(\tilde{x},\tilde{t}),$$

où nous avons utilisé $u_0/x_0 = 1/t_0$. En faisant de même pour les autres termes du système (1.1), nous obtenons :

$$\frac{\rho_0 u_0}{x_0}\partial_{\tilde{t}}\tilde{\rho} + \frac{\rho_0 u_0}{x_0}\nabla_{\tilde{x}} \cdot \tilde{q} = 0, \tag{1.3a}$$

$$\frac{\rho_0 u_0^2}{x_0}\partial_{\tilde{t}}\tilde{q} + \frac{\rho_0 u_0^2}{x_0}\nabla_{\tilde{x}} \cdot \left(\frac{\tilde{q}\otimes\tilde{q}}{\tilde{\rho}}\right) + \frac{p_0}{x_0}\nabla_{\tilde{x}}\tilde{p} = 0, \tag{1.3b}$$

$$\frac{p_0 u_0}{x_0}\partial_{\tilde{t}}\tilde{E} + \frac{p_0\rho_0 u_0}{\rho_0 x_0}\nabla_{\tilde{x}} \cdot \left((\tilde{E}+\tilde{p})\frac{\tilde{q}}{\tilde{\rho}}\right) = 0, \tag{1.3c}$$

$$p_0\tilde{p} = (\gamma-1)\left(p_0\tilde{E} - \frac{\rho_0^2 u_0^2}{2\rho_0}\frac{\tilde{q}^2}{\tilde{\rho}}\right). \tag{1.3d}$$

Nous divisons respectivement par $\frac{\rho_0 u_0}{x_0}$, $\frac{\rho_0 u_0^2}{x_0}$, $\frac{p_0 u_0}{x_0}$ et p_0 les équations sur la densité (1.3a), la quantité de mouvement (1.3b), l'énergie (1.3c) et l'équation d'état (1.3d). En définissant le paramètre

$$\varepsilon = \frac{\rho_0 u_0^2}{p_0},$$

et en omettant les tildes, on obtient le système d'Euler adimensionné suivant

$$\partial_t \rho + \nabla \cdot q = 0, \tag{1.4a}$$
$$\partial_t q + \nabla \cdot \left(\frac{q \otimes q}{\rho} \right) + \frac{1}{\varepsilon} \nabla p = 0, \tag{1.4b}$$
$$\partial_t E + \nabla \cdot \left((E + p) \, \frac{q}{\rho} \right) = 0, \tag{1.4c}$$
$$E = \frac{p}{\gamma - 1} + \frac{\varepsilon}{2} \frac{|q|^2}{\rho}. \tag{1.4d}$$

Le paramètre $\varepsilon > 0$ est relié au nombre de Mach M_0 à travers la relation

$$M_0 = \frac{u_0}{c_0} = \sqrt{\frac{\varepsilon}{\gamma}} \quad \text{ou encore} \quad \varepsilon = \gamma M_0^2,$$

où la vitesse du son est définie par $c_0 = \sqrt{\gamma p_0 / \rho_0}$.
Pour construire un schéma AP (Asymptotiquement Préservant) dans la limite bas-Mach, il faut, dans un premier temps identifier le modèle limite. Je rappelle dans le paragraphe ci-dessous les grandes lignes du passage à la limite dans le modèle d'Euler.

1.4 Limite bas-Mach pour les équations d'Euler

A la limite, c'est-à-dire lorsque ε (le nombre de Mach) tend vers 0, on obtient le modèle d'Euler incompressible (voir Section 2.1 pour les détails du passage à la limite). Formellement, la conservation de la quantité de mouvement, nous donne

$$\nabla p = 0,$$

qui donne avec la loi d'état une énergie constante en espace. Avec des conditions aux limites bien choisies, l'équation d'énergie nous donne

$$\nabla \cdot u = 0,$$

et donc l'incompressibilité.
Ainsi le modèle limite $\varepsilon \to 0$ est le modèle d'Euler incompressible

$$\partial_t \rho + \nabla \cdot q = 0, \tag{1.5a}$$
$$\partial_t q + \nabla \cdot \left(\frac{q \otimes q}{\rho} \right) + \nabla p^1 = 0, \tag{1.5b}$$
$$\nabla \cdot u = 0, \tag{1.5c}$$

où l'on a supposé que p^1, la correction d'ordre ε de la pression ($p^\varepsilon = p + \varepsilon p^1$ avec p^ε la pression dans le système (1.4)), existe.
On voit donc que pour des valeurs de ε élevées, l'écoulement est gouverné par des effets compressibles, tandis que dans la limite des faibles valeurs de ε, les équations compressibles convergent vers le régime incompressible.

1.5 Problématiques liées à la limite bas-Mach

Cette transition est particulièrement difficile à capter numériquement parce que les schémas classiquement utilisés pour discrétiser le système d'Euler compressible (1.4), sont explicites en temps et de type Godunov en espace. En effet, les schémas de Godunov sont des schémas conservatifs de type volumes finis bien adaptés pour des systèmes hyperbolique de lois de conservations comme le modèle d'Euler compressible qui peut se réécrire sous la forme

$$\partial_t W + \nabla \cdot F(W) = 0,$$

où $W = (\rho,\, q,\, E)$ est le vecteur des variables conservatives et

$$F(W) = \begin{pmatrix} q \\ \dfrac{q \otimes q}{\rho} + p\, I_d \\ (E+p)\, \dfrac{q}{\rho} \end{pmatrix},$$

est le flux associé.
Ces schémas sont capables de bien capturer des solutions discontinues comme des ondes de choc qui sont des solutions du modèle compressible. De plus, le choix d'une discrétisation explicite en temps est classiquement choisie puisque qu'elle est simple à implémenter et la condition sur le pas de temps n'est pas restrictive dans le régime compressible. D'autre part, le système est extrêmement non linéaire donc une discrétisation implicite est très couteuse. Cependant, si ces schémas permettent une bonne description des régimes pour des nombres de Mach d'ordre 1, ils sont inutilisables pour des faibles nombres de Mach. En effet, ils ne sont pas consistants dans la limite bas-Mach, c'est-à-dire que le système discrétisé ne tend pas à la limite $\varepsilon \to 0$ vers une approximation du modèle incompressible. Par ailleurs, pour que ces schémas soient stables, le pas de temps doit résoudre l'échelle des ondes de pression qui est de l'ordre de $\sqrt{\varepsilon}$. Par conséquent, le pas de temps tend vers 0 lorsque ε tend vers 0.

1.5.1 Consistance des schémas de type Godunov dans la limite bas-Mach

Commençons, par donner la semi-discrétisation du système (1.4) avec un schéma explicite :

$$\frac{\rho^{n+1} - \rho^n}{\Delta t} + \nabla \cdot q^n = 0, \tag{1.6a}$$

$$\frac{q^{n+1} - q^n}{\Delta t} + \nabla \cdot \left(\frac{q^n \otimes q^n}{\rho^n} \right) + \frac{1}{\varepsilon} \nabla p^n = 0, \tag{1.6b}$$

$$\frac{E^{n+1} - E^n}{\Delta t} + \nabla \cdot \left((E^n + p^n)\, \frac{q^n}{\rho^n} \right) = 0, \tag{1.6c}$$

$$E^{n+1} = \frac{p^{n+1}}{\gamma - 1} + \varepsilon\, \frac{\rho^{n+1}\, |u^{n+1}|^2}{2}. \tag{1.6d}$$

Lorsque nous passons formellement, comme dans le cas continu, à la limite dans la semi-discretization, l'équation sur la quantité de mouvement nous donne $\nabla p^n = 0$ lors du calcul de la solution W^{n+1}. Cela impose uniquement une contrainte sur la condition initiale mais n'impose pas une discrétisation du modèle limite, soit : $\forall n \geq 0$, $\nabla p^{n+1} = 0$ et $\nabla \cdot u^{n+1} = 0$.
L'étude détaillée des problèmes de consistance à la limite des schémas explicites classiques de type Godunov a été effectuée pour la première fois par Guillard et Viozat [43]. En menant une analyse asymptotique sur le nombre de Mach pour un schéma upwind, les auteurs montrent que cette discrétisation autorise pour des faibles nombres de Mach, des fluctuations sur la pression de l'ordre du nombre de Mach (M), soit de l'ordre de $\sqrt{\varepsilon}$, tandis que dans le cas continu les fluctuations ne sont que de l'ordre de M^2, c'est-à-dire ε. On ne retrouve donc pas le correct ordre de magnitude pour la pression et des résultats numériques montrent que la solution obtenue est parfois très loin de la solution incompressible. En modifiant la viscosité numérique du flux dans le schéma de Roe à l'aide d'un préconditionneur, il est possible de retrouver le bon ordre de magnitude. Il existe de nombreux travaux proposant des schémas de Godunov modifiés basés sur des méthodes de préconditionneurs pour résoudre ce problème de précision [79, 18, 19, 15, 55, 56]. Dans d'autres travaux, le comportement des schémas de type Godunov est étudié dans le régime faiblement compressible, voir [27] ou [42].

1.5.2 Stabilité des schémas classiques explicites

La stabilité des schémas explicites est assurée sous une condition C.F.L. reliée aux valeurs propres de la matrice jacobienne associée au flux F. En dimension $d = 1$, les valeurs propres sont données par

$$\lambda_1 = u - \frac{c}{\sqrt{\varepsilon}}, \quad \lambda_2 = u, \quad \lambda_3 = u + \frac{c}{\sqrt{\varepsilon}},$$

où $c^2 = \gamma p/\rho$. Ainsi, pour les pas de temps et d'espace Δt et Δx, la condition de C.F.L. est donnée par

$$\Delta t \leq \frac{\Delta x}{|u| + \frac{c}{\sqrt{\varepsilon}}}.$$

A la limite $\varepsilon \to 0$, la vitesse de fluide est négligeable devant celle des ondes acoustiques et on a

$$\Delta t \leq \Delta x \sqrt{\varepsilon}.$$

Cela implique que pour des faibles valeurs du nombre de Mach, Δt est très petit et le coût de calcul devient extrêmement élevé.

1.5.3 Travaux proposés dans la littérature

Nous avons évoqué des méthodes de préconditionneurs basées sur la modification des schémas de type Godunov utilisés pour le modèle compressible. Par ailleurs, il existe

des méthodes basées sur la résolution d'une équation sur la pression permettant de controler ses variations. On trouve leur origine dans les méthodes développées pour le modèle incompressible où une équation elliptique sur la pression apparait naturellement dans la reformulation de ce dernier (voir Section 2.1.2). La perte de consistance à faible nombre de Mach est corrigée à l'aide de méthodes de splitting avec correction de pression. Nous pouvons citer les travaux pionners de Harlow et Amsden [45, 4] avec un schéma de différences finies semi-implicite mais aussi des schémas volumes finis [48]. De nombreux travaux sont proposés [8, 52, 73, 68], on y trouve notamment des discrétisations sur maillages colocalisés ou décalés et des formulations en variables physiques (ρ, u et p). Il est important de noter que ces méthodes sont souvent adaptées pour des fluides faiblement compressibles. En effet, dans ces régimes la reformulation est équivalente puisque les solutions sont régulières mais pour de plus grandes valeurs du nombre de Mach, où des ondes de choc peuvent se développer, elle peut entrainer une perte de la structure conservative des équations.
De plus, l'ensemble des techniques énoncées propose des solutions pour résoudre la consistance des schémas existants à faible nombre de Mach mais n'enlève pas la contrainte de type C.F.L. extrêmement restrictive. Elles demandent toutes de résoudre l'échelle des ondes acoustiques (rapides dans ce régime) pour préserver la stabilité du schéma. Notre travail consiste à développer un schéma qui soit à la fois consistant à la limite et stable sous une condition C.F.L. non contrainte par le nombre de Mach, avec un coût de calcul abordable. Ce type de schéma est connu sous le nom de schémas asymptotiquement préservants (AP).

1.6 Schémas asymptotiquement préservants

Cette classe de schémas fut introduite par Shi Jin [47] dans le contexte des équations cinétiques et de leurs limite fluide ou diffusive. Ces schémas sont stables uniformément par rapport au paramètre ε, il est donc possible d'utiliser un maillage indépendant des petites échelles. De plus, ils sont asymptotiquement consistants : lorsque le paramètre $\varepsilon \to 0$, ils redonnent une discrétisation du modèle limite qui lui est indépendant de ε (Figure 1.1).
Une communauté scientifique importante se consacre à l'étude et au développement de schémas AP pour différents modèles et différentes limites. Parmi ces travaux, nous pouvons citer [25] pour le modèle de Navier-Stokes et sa limite bas-Mach basé sur une méthode de Gauge, [24] pour le modèle d'Euler-Lorentz et sa limite fluide de dérive, basé sur une reformulation du modèle.
De nombreux schémas AP proposent une discrétisation semi-implicite où une partie des termes est traitée de manière explicite tandis que l'autre partie est traitée implicitement. Dans ce contexte, et pour des modèles similaires à celui considéré ici, nous citons par exemple les travaux de [64] pour un modèle bi-fluide, [26] pour le modèle d'Euler isentropique en introduisant un paramètre permettant de choisir le degrés d'implicitation du flux de pression. Il est ensuite étendu au cas d'Euler complet

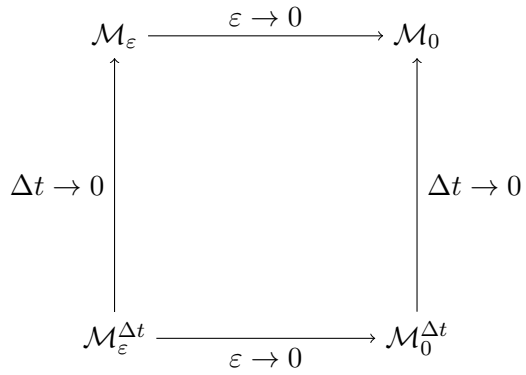


FIGURE 1.1 – Illustration du principe des schémas AP : $\mathcal{M}_\varepsilon$ correspond au modèle d'Euler compressible et $\mathcal{M}_\varepsilon^{\Delta t}$ sa discrétisation avec un schéma AP. Lorsque $\varepsilon \to 0$, $\mathcal{M}_\varepsilon^{\Delta t}$ donne une discrétisation $\mathcal{M}_0^{\Delta t}$ stable indépendamment de ε et consistante avec le modèle d'Euler incompressible M_0.

et de Navier-Stokes [21]. Dans [65], le même splitting en temps que celui proposé dans [21] est utilisé avec une décomposition de la pression proposée par R. Klein [51] dans le flux d'énergie. Dans [67], une adaptation des algorithmes de correction de pression sur maillages décalés est proposée pour les équations de Saint-Venant et d'Euler complet puis étendue aux équations de Navier-Stokes [32].

D'autres méthodes proposent un splitting basé sur la séparation entre les ondes matérielles ou de transport (lentes) et acoustiques (rapides) du modèle compressible. Dans les travaux de [17] pour Euler-friction puis pour Euler complet avec une équation d'état générale [16], le schéma semi-implicite développé est résolu avec une méthode de Lagrange-projection. Le système est divisé en un sous-système acoustique et un sous-système matériel. Le premier est traité de manière implicite, à l'aide d'une technique de relaxation pour faciliter le traitement des termes non linéaires, puis dans une seconde étape le système matériel est résolu explicitement.

Dans cette thèse, nous nous intéressons au développement d'un schéma AP basé sur une discrétisation IMEX (Implicite-Explicite) [80, 66] où le flux est séparé en une partie qui sera traitée de manière totalement explicite et une autre de manière totalement implicite. Dans ce cadre, nous faisons référence aux travaux de [30] pour le modèle d'Euler isentropique où un schéma d'ordre 1 est proposé avec un solveur de Rusanov [69, 39] utilisé en espace. Une analyse de stabilité sur le système linéarisé leur permet également de définir la viscosité numérique nécessaire pour obtenir un schéma qui soit L^2 ou L^∞ stable. Dans [29], un schéma assurant la propriété TVD (total variational diminishing) et plus précis que le schéma d'ordre 1 est construit en couplant les discrétisations d'ordre 1 et 2. Ce nouveau schéma est ensuite utilisé dans une procédure MOOD (Multi-dimensional Optimal Order Detection) [20, 33] permettant de réduire les oscillations du schéma d'ordre 2. Dans [10], une discrétisation sur maillages décalés est présentée pour le modèle d'Euler isentropique puis étendue au cas d'Euler complet avec le même splitting en temps proposé dans [30].

Dans [6], deux schémas semi-implicites pour les équations d'Euler complet sont comparés, l'un avec un splitting sur la pression similaire à celui proposé dans [26] et l'autre avec un splitting de flux basé sur la séparation entre les ondes matérielles et acoustiques. Dans [37] un schéma volumes finis sur maillages décalés d'ordre 1 pour le modèle d'Euler complet. Le splitting IMEX utilisé pour la discrétisation en temps est inspiré de [38]. Cette discrétisation mène à la résolution d'une équation non linéaire sur la pression à l'aide d'une méthode itérative. Ce schéma peut traiter différentes lois d'états et traite aussi le cas des équations de Navier-Stokes avec un traitement explicite des termes diffusifs. Ces idées sont reprises dans [76] pour développer des méthodes de Galerkine discontinues d'ordre élevé sur des maillages décalés également. Enfin elle sont aussi utilisées dans [11] pour construire un schéma volumes finis sur maillages colocalisés. Le même splitting en temps [38] est utilisé et en s'appuyant sur les résultats de [30] pour la construction d'un schéma L^2 stable, un solveur de Rusanov est utilisé sur la partie explicite et un solveur centré sur la partie implicite. De plus, le caractère bien posé du système non linéaire sur la pression à résoudre est prouvé puis une extension du schéma à l'ordre 2 en temps [80, 66] et en espace est proposée. Un schéma volumes finis colocalisés avec un splitting similaire est également proposé dans [14] qui a l'avantage de résoudre une équation linéaire sur la pression pour la loi d'état des gazs parfaits. Ce schéma est également étendu à l'ordre 3 et aux équations de Navier-Stokes.

1.7 Le cas des équations de Navier-Stokes

1.7.1 Problématique

On s'intéresse également au cas des équations de Navier-Stokes qui nous permettent de prendre en compte les effets visqueux et de conduction du fluide. Elles sont constituées d'une partie non visqueuse, correspondant aux équations d'Euler, et d'une partie diffusive correspondant au tenseur des contraintes visqueuses et au flux de température.
Pour étudier la limite à faibles nombres de Mach, nous travaillons aussi sur le système de Navier-Stokes adimensionné. Lorsque $M \to 0$, le modèle tend vers un modèle limite. Il est intéressant de noter que lorsque les coefficients de viscosité et de conductivité sont nuls, nous retrouvons le modèle d'Euler incompressible. De plus, lorsque uniquement les effets conductifs sont négligeables, le modèle limite est celui de Navier-Stokes incompressible avec la contrainte d'incompressibilité $\nabla \cdot u = 0$. Nous rencontrons les mêmes problématiques que dans le cas d'Euler complet : lorsque $M \to 0$, les schémas classiques explicites ne sont pas stables et consistants avec le modèle limite.
Généralement, les schémas AP proposés pour les équations d'Euler complet sont directement étendus au cas de Navier-Stokes. Les termes convectifs sont traités de manière identique et pour les termes diffusifs plusieurs choix sont possibles. Un traitement explicite est le plus simple mais impose une contrainte parabolique sur le pas de temps [41]. Nous avons

$$\Delta t \leq \frac{\Delta x^2}{2\max\left(\frac{4\mu}{3\rho}, \frac{\lambda}{c_v\rho}\right)},$$

où μ et λ sont respectivement les coefficients de viscosité et de conductivité du fluide et c_v est la chaleur spécifique du fluide à volume constant. Cette contrainte peut devenir très restrictive dans les régimes fortement visqueux. Une autre solution est d'impliciter ces termes mais cela ajoute des termes non linéaires sur l'équation de pression qu'il faut traiter. De plus, puisque les équations sur la quantité de mouvement sont couplées à travers le tenseur des contraintes visqueuses cela demande la résolution d'un système linéaire qui grandit avec la dimension en espace. Parmi les travaux dans le contexte des schémas AP basés sur des méthodes IMEX avec un splitting en temps similaire à celui que nous utilisons, nous avons par example [76] avec des méthodes de Galerkine discontinues et les termes diffusifs explicites, [12] qui étend la discrétisation de [11] avec un traitement implicite des termes additionnels. Pour cela un double algorithme de Picard est utilisé pour résoudre l'équation de pression : une boucle pour approcher les termes visqueux qui sont ensuite injectés dans la seconde pour résoudre l'équation de pression.
Afin de formellement passer à limite dans les équations de Navier-Stokes, nous présentons ci-dessous l'adimenssionnement du système. La limite sera ensuite détaillée dans la Section 3.1.1.

1.7.2 Adimensionnement des équations de Navier-Stokes

Nous présentons brièvement l'adimensionnenment des équations de Navier-Stokes qui sont les suivantes

$$\partial_t \rho + \nabla_x \cdot q = 0, \tag{1.7a}$$
$$\partial_t q + \nabla_x \cdot \left(\frac{q \otimes q}{\rho}\right) + \nabla_x p = \nabla_x \cdot \sigma, \tag{1.7b}$$
$$\partial_t E + \nabla_x \cdot \left((E+p)\,\frac{q}{\rho}\right) = \nabla_x \cdot \left(\sigma \frac{q}{\rho}\right) + \nabla_x \cdot (\lambda \nabla T)\,, \tag{1.7c}$$

où la pression p et la température T du fluide sont données pour un gaz parfait par

$$E = \frac{p}{\gamma - 1} + \frac{\varepsilon}{2}\frac{|q|^2}{\rho}, \quad T = \frac{p}{R\rho}, \tag{1.7d}$$

avec $R = c_p - c_v$ et $\gamma = c_p/c_v$ le rapport des chaleurs spécifiques à pression et volume constant. Le tenseur des contraintes visqueuses σ est défini par

$$\sigma = \mu\left(\nabla u + (\nabla u)^T\right) - \frac{2}{3}\mu\,(\nabla \cdot u)\,I. \tag{1.8}$$

Les paramètre μ et λ sont respectivement les coefficients de viscosité et de conductivité du fluide.

Pour adimensionner le modèle de Navier-Stokes compressible, nous introduisons les quantités suivantes :

$$\begin{gathered}\tilde{\rho} = \rho/\rho_0, \qquad \tilde{u} = u/u_0, \qquad \tilde{p} = p/p_0, \qquad \tilde{E} = E/p_0, \quad \tilde{T} = T/T_0, \\ \tilde{x} = x/x_0, \qquad \tilde{t} = t/t_0,\end{gathered} \tag{1.9}$$

où ρ_0, p_0, x_0, t_0, $u_0 = x_0/t_0$ et $T_0 = p_0/\rho_0$ sont les ordres de grandeur des valeurs prises par le fluide pour les situations considérées. Alors

$$\frac{\partial \rho}{\partial x}(x,t) = \frac{\partial}{\partial x}\left(\rho_0 \tilde{\rho}(\tilde{x},\tilde{t})\right) = \rho_0 \frac{\partial \tilde{x}}{\partial x}\frac{\partial \tilde{\rho}}{\partial \tilde{x}}(\tilde{x},\tilde{t}) = \frac{\rho_0}{x_0}\frac{\partial \tilde{\rho}}{\partial \tilde{x}}(\tilde{x},\tilde{t}) = \frac{\rho_0}{x_0}\frac{\partial \tilde{\rho}}{\partial \tilde{t}}(\tilde{x},\tilde{t}),$$

$$\frac{\partial^2 T}{\partial x^2}(x,t) = \frac{T_0}{x_0^2}\frac{\partial^2 \tilde{T}}{\partial \tilde{x}^2}(\tilde{x},\tilde{t}) = \frac{p_0}{\rho_0 x_0^2}\frac{\partial^2 \tilde{T}}{\partial \tilde{x}^2}(\tilde{x},\tilde{t}).$$

Nous faisons de même pour les autres termes du système (1.7) et divisons respectivement par $\frac{\rho_0 u_0}{x_0}$, $\frac{\rho_0 u_0^2}{x_0}$, $\frac{p_0 u_0}{x_0}$ et p_0 les équations sur la densité, la quantité de mouvement, l'énergie et l'équation d'état. En définissant, comme pour les équations d'Euler, $\varepsilon = \rho_0 u_0^2/p_0 = \gamma M_0^2$ nous obtenons :

$$\partial_t \tilde{\rho} + \nabla_{\tilde{x}} \cdot \tilde{q} = 0, \tag{1.10a}$$

$$\partial_t \tilde{q} + \nabla_{\tilde{x}} \cdot \left(\frac{\tilde{q}\otimes\tilde{q}}{\tilde{\rho}}\right) + \frac{1}{\varepsilon}\nabla_{\tilde{x}}\tilde{p} = \frac{\mu}{\rho_0 u_0 x_0}\left(\Delta_{\tilde{x}}\tilde{u} + \nabla_{\tilde{x}}\left(\nabla_{\tilde{x}}\cdot\tilde{u}\right)\right), \tag{1.10b}$$

$$\begin{aligned}\partial_t \tilde{E} + \nabla_{\tilde{x}} \cdot ((\tilde{E}+\tilde{p})\frac{\tilde{q}}{\tilde{\rho}}) = \frac{\varepsilon\mu}{\rho_0 u_0 x_0}\nabla_{\tilde{x}} \cdot \left(\left(\nabla_{\tilde{x}}\tilde{u} + (\nabla_{\tilde{x}}\tilde{u})^T\right)\tilde{u} - \frac{2}{3}\left(\nabla_{\tilde{x}}\cdot\tilde{u}\right)\tilde{u}\right) \\ + \frac{\lambda}{\rho_0 u_0 x_0}\Delta_{\tilde{x}}\tilde{T},\end{aligned} \tag{1.10c}$$

$$\tilde{p} = (\gamma-1)\left(\tilde{E} - \frac{\varepsilon}{2}\frac{\tilde{q}^2}{\tilde{\rho}}\right). \tag{1.10d}$$

Enfin, en définissant les paramètres :

$$\tilde{\mu} = \frac{\mu}{\rho_0 u_0 x_0}, \quad \tilde{\lambda} = \frac{\lambda}{\rho_0 u_0 x_0},$$

qui sont respectivement les coefficients de viscosité et de conductivité adimensionnnés, et en omettant les tildes nous obtenons le système de Navier-Stokes adimensionné :

$$\partial_t \rho + \nabla \cdot q = 0, \tag{1.11a}$$

$$\partial_t q + \nabla \cdot \left(\frac{q\otimes q}{\rho}\right) + \frac{1}{\varepsilon}\nabla p = \nabla\cdot\sigma, \tag{1.11b}$$

$$\partial_t E + \nabla\cdot\left((E+p)\,\frac{q}{\rho}\right) = \varepsilon\nabla\cdot\left(\sigma\frac{q}{\rho}\right) + \nabla\cdot\left(\lambda\nabla T\right), \tag{1.11c}$$

$$E = \frac{p}{\gamma-1} + \frac{\varepsilon}{2}\frac{|q|^2}{\rho}. \tag{1.11d}$$

1.8 Synthèse des travaux et organisation du manuscrit

Dans cette thèse nous développons et analysons un schéma volumes finis asymptotiquement préservant dans la limite des faibles nombre de Mach pour les équations d'Euler et de Navier-Stokes.
Dans le Chapitre 2, nous nous intéressons aux équations d'Euler complet. Dans l'introduction 2.1, nous détaillons la limite formelle des équations d'Euler compressible vers les équations d'Euler incompressible, la reformulation de ce dernier et l'intérêt des schémas AP pour l'obtention d'un schéma robuste et efficace dans tous les régimes liés au nombre de Mach. Puis, dans la Section 2.2, nous effectuons une analyse sur les splitting de flux existants pour la construction d'un schéma AP basé sur une discrétisation IMEX en temps. A partir de cette étude sur la stabilité et la consistante asymptotique ainsi que sur la préservation de solutions stationnaires, nous montrons que le splitting choisi est le plus adapté à notre problème. Dans la Section 2.3, nous présentons la discrétisation de notre schéma AP d'ordre 1. La semi-discrétisation en temps proposée est basée sur une linéarisation de l'équation de pression initialement résolue à l'aide d'un algorithme de Picard dans [11]. La consistance asymptotique et la préservation des états stationnaires est prouvée. Nous présentons ensuite la discrétisation en espace choisie à partir des résultats démontrés dans [30]. Elle repose sur la discrétisation des flux explicites et implicites avec un solveur de Rusanov pour un schéma L^∞ stable ou une discrétisation centrée sur uniquement la partie implicite pour un schéma L^2 stable. De plus, à travers d'une analyse de Fourier sur le système linéarisé, nous montrons la stabilité de notre de schéma. Nous concluons cette section avec des résultats numériques où nous comparons notre schéma au schéma non linéaire [11] et nous montrons son bon comportement. Dans la Section 2.4, nous proposons une extension d'ordre 2 avec une approche de type IMEX Runge-Kutta en temps et une méthode MUSCL en espace. A partir des travaux [29] pour le modèle d'Euler isentropique, nous construisons un schéma TVD qui à travers une procédure de type MOOD permet de réduire les oscillations propres aux schémas IMEX d'ordre 2. Nous terminons ce chapitre avec des résultats numériques comparant les schémas AP d'ordre 1, d'ordre 2 et MOOD qui mettent en valeur dans le schéma MOOD la montée en précision du schéma d'ordre 1 et la réduction des oscillations du schéma d'ordre 2.
Dans le Chapitre 3, nous traitons le cas des équations de Navier-Stokes qui permettent de prendre en compte les effets visqueux et conductifs du fluide. Nous commençons par donner des détails sur le passage à la limite bas-Mach et les difficultés numériques associées aux termes diffusifs. Dans la Section 3.2, nous présentons la discrétisation du schéma AP d'ordre 1. Il repose sur le même splitting en temps que dans le Chapitre 2 pour la partie non visqueuse et en un traitement implicite des termes diffusifs afin de s'affranchir de la contrainte C.F.L. restrictive pour des régimes hautement visqueux. En suivant la même stratégie que pour Euler complet, nous proposons une approximation des termes visqueux dans l'équation sur la pression résultant en un système linéaire simple à résoudre. Le schéma proposé est ensuite étendu à l'ordre 2 et l'on prouve son bon comportement sur un ensemble de

cas tests numériques en deux dimensions faisant intervenir des fluides non visqueux, visqueux, en présence de conduction dans des régimes compressibles et à faibles nombres de Mach.

Chapter 2

Second order all Mach number IMEX scheme for the full Euler equations

The content of this chapter is the subject of an article written in collaboration with Marie-Hélène Vignal and submitted for publication [3].

Contents

2.1 Introduction

We consider the modeling of a compressible fluid described by the compressible full Euler equations and we are interested in numerical methods valid in all Mach regimes. Let $\Omega \subset \mathbb{R}^d$ ($d = 1$, 2 or 3) be an open bounded domain, the full Euler equations in rescaled variables are given by

$$\partial_t \rho + \nabla \cdot q = 0, \tag{2.1a}$$

$$\partial_t q + \nabla \cdot \left(\frac{q \otimes q}{\rho} \right) + \frac{1}{\varepsilon} \nabla p = 0, \tag{2.1b}$$

$$\partial_t E + \nabla \cdot \left((E + p) \, \frac{q}{\rho} \right) = 0, \tag{2.1c}$$

with $\rho(t,x) > 0$ is the density of the fluid, $q(t,x) = \rho(t,x)u(t,x)$ its momentum, $u(t,x)$ its velocity field, $E(t,x)$ its total energy and $p(t,x)$ its pressure and where $x \in \Omega$ and $t \in \mathbb{R}^+$ are the space and time variables. The pressure is given by an equation of state, here that of perfect gases:

$$E = \frac{p}{\gamma - 1} + \frac{\varepsilon}{2} \frac{|q|^2}{\rho}, \tag{2.1d}$$

with $\gamma > 1$ the given ratio of specific heats.
The rescaled parameter ε is related to the Mach number

$$M^2 = \frac{u_0^2}{c_0^2} = \frac{\varepsilon}{\gamma},$$

with $c_0^2 = \gamma \; p_0/\rho_0$, u_0, p_0 and ρ_0 being the typical values of the velocity, pressure and density in the fluid. The previous system can be rewritten in compact form as

$$\partial_t W(x,t) + \nabla \cdot F(W(x,t)) = 0, \tag{2.2}$$

where $W = (\rho, q, E)$ is the vector of conservative variables and

$$F(W) = \begin{pmatrix} q \\ \rho u \otimes u + \dfrac{1}{\varepsilon} \, p \, Id_{\mathbb{R}^3} \\ (E + p) \, u \end{pmatrix},$$

the flux.
In low Mach number regimes, the typical sound speed in the fluid, c_0, is very large compared to the typical speed of the fluid itself, u_0, and so ε is very small. It is well known that in such situations (see [78]), if an explicit scheme is used, the time step must satisfy a severe C.F.L. (Courant-Friedrichs-Levy) stability condition. Indeed, for $d = 1$, the Jacobian matrix associated to F is given by

$$DF(W) = \begin{pmatrix} 0 & 1 & 0 \\ \dfrac{1}{2}(\gamma - 3)u^2 & (3 - \gamma)u & \dfrac{\gamma - 1}{\varepsilon} \\ \dfrac{-\gamma p u}{(\gamma - 1)\rho} + \dfrac{1}{2}(\gamma - 2)\varepsilon u^3 & \dfrac{\gamma p}{(\gamma - 1)\rho} + \dfrac{\varepsilon}{2}(3 - 2\gamma)u^2 & \gamma u \end{pmatrix},$$

and its eigenvalues are

$$\lambda_1 = u - c/\sqrt{\varepsilon}, \quad \lambda_2 = u, \quad \lambda_3 = u + c/\sqrt{\varepsilon},$$

with $c^2 = \gamma\, p/\rho$ and u the fluid velocity. And, the C.F.L. condition, ensuring the stability of explicit schemes, for the time and space steps Δt and Δx, is given by

$$\Delta t \leq \frac{\Delta x}{\max(|u \pm c/\sqrt{\varepsilon}|)}. \tag{2.3}$$

Then, for a given space step Δx, the time step Δt is of order $\sqrt{\varepsilon}$ and tends to 0 with ε. Furthermore, even if this constraint is satisfied, it is also well known (see [43], [42] or [27]) that explicit schemes suffer from a consistency problem in the limit $\varepsilon \to 0$. They are not capable to capture the right asymptotic regime.
A possible way to bypass these limitations is to use in the regions where ε is sufficiently small, the incompressible Euler equations obtained as the low Mach number limit of the compressible Euler equations (2.1). Here, we prefer to use the compressible model and an asymptotic preserving scheme as explained in Section 2.1.3. But, first, let us recall the formal low Mach number limit in the next section.

2.1.1 Low Mach number limit for the full Euler equations

The rigorous low Mach number limit of the compressible Euler system has been widely studied in the last years [50, 49, 70, 5, 57, 60, 59, 1]. Results in the case of non-isentropic Euler equations with general initial data can be found in [60] in the space $\Omega = \mathbb{R}^d$, in [1] for an exterior domain and in a bounded toroidal domain in [59]. Here, we briefly recall the formal limit. We denote by $(\rho^\varepsilon, q^\varepsilon, E^\varepsilon, p^\varepsilon)$ the solution of (2.1) with general initial conditions and with the impermeability boundary condition

$$u^\varepsilon \cdot \nu = 0, \quad \text{on } \partial\Omega,$$

where ν is the unit normal to $\partial\Omega$ outward to Ω. Performing an asymptotic expansion such that:

$$\begin{aligned}
\rho^\varepsilon &= \rho^0 + \varepsilon\rho^1,\\
q^\varepsilon &= q^0 + \varepsilon q^1,\\
p^\varepsilon &= p^0 + \varepsilon p^1,\\
E^\varepsilon &= E^0 + \varepsilon E^1,
\end{aligned}$$

inserting the following expansions into the compressible Euler equations (2.1) and collecting the different order terms we obtain for all $x \in \Omega$ and $t > 0$,

$$\varepsilon^{-1}: \quad \nabla p^0 = 0, \tag{2.4a}$$

$$\varepsilon^{0}: \quad \partial_t \rho^0 + \nabla \cdot q^0 = 0, \tag{2.4b}$$

$$\partial_t q^0 + \nabla \cdot \left(\frac{q^0 \otimes q^0}{\rho^0} \right) + \nabla p^1 = 0, \tag{2.4c}$$

$$\partial_t E^0 + \nabla \cdot \left((E^0 + p^0)\, \frac{q^0}{\rho^0} \right) = 0, \tag{2.4d}$$

$$E^0 = \frac{p^0}{\gamma - 1}. \tag{2.4e}$$

Note that, $p^1(x,t) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}(p^\varepsilon(x,t) - p^0)$ is the order one correction of the pressure.
Since $\nabla p^0(x,t) = 0$, $p^0(x,t) = p^0(t)$ then for all $x \in \Omega$ and $t > 0$

$$E^0(x,t) = E^0(t) = \frac{p^0(t)}{\gamma - 1}.$$

Integrating now (2.4d) on $[0,T] \times \Omega$ we obtain,

$$\int_\Omega \big(E^0(t) - E^0(x,0)\big)\, dx + \int_0^t \left(\int_{\partial\Omega} (E^0(t) + p^0(t)\,)\, u^0(x,t) \cdot \nu(x)\, d\sigma(x) \right) dt = 0. \tag{2.5}$$

Using the boundary condition (**??**), we obtain for all $t > 0$,

$$|\Omega|\, E^0(t) - \int_\Omega E^0(x,0) dx = 0.$$

And so,

$$E^0(t) = E^0 = \frac{1}{|\Omega|} \int_\Omega E^0(x,0)\, dx.$$

Thereafter, having constant energy and pressure, we recover from (2.4d) the incompressibility constraint

$$\nabla \cdot u^0(x,t) = 0, \tag{2.6}$$

for all $x \in \Omega$ and $t > 0$. Finally, the **"incompressible limit system"** [1] writes

$$\partial_t \rho^0 + \nabla \cdot q^0 = 0, \tag{2.7a}$$

$$\partial_t q^0 + \nabla \cdot \left(\frac{q^0 \otimes q^0}{\rho^0} \right) + \nabla p^1 = 0, \tag{2.7b}$$

$$\nabla \cdot u^0 = 0, \tag{2.7c}$$

$$E^0 = \frac{p^0}{\gamma - 1} = \frac{1}{|\Omega|} \int_\Omega E^0(x,0)\, dx. \tag{2.7d}$$

Let us note that p^1 is given implicitly by the incompressibility constraint (2.7c). We will see in the next section how an explicit equation can be recovered.

2.1.2 Reformulation of the incompressible limit model

In this section, we present the reformulation of the limit model and its equivalence with the original incompressible model.

Lemma 2.1.1. *(Formal)*

1. *The incompressible model* $\mathbf{M_0}$ *given by* (2.7) *can be reformulated into* $\mathbf{RM_0}$ *replacing the incompressibility constraint* (2.7c) *by an explicit equation for* p^1.

$$\mathbf{M_0} \implies \mathbf{RM_0}$$

where the reformulated incompressible system noted $\mathbf{RM_0}$ *reads*

$$\partial_t \rho^0 + \nabla \cdot q^0 = 0, \tag{2.8a}$$

$$\partial_t q^0 + \nabla \cdot \left(\frac{q^0 \otimes q^0}{\rho^0} \right) + \nabla p^1 = 0, \tag{2.8b}$$

$$-\nabla \cdot \left(\frac{1}{\rho^0} \nabla p^1 \right) = \nabla \cdot \Big((u^0 \cdot \nabla) u^0 \Big), \tag{2.8c}$$

$$E^0 = \frac{p^0}{\gamma - 1} = \frac{1}{|\Omega|} \int_\Omega E^0(x, 0)\, dx. \tag{2.8d}$$

2. *The reformulated incompressible Euler system* $\mathbf{RM_0}$ *is equivalent to* $\mathbf{M_0}$ *if and only if the initial condition is well prepared to the incompressible regime. More precisely*

$$\mathbf{RM_0} \implies \mathbf{M_0}$$

if and only if

$$\nabla \cdot u(\,\cdot\,, 0) = 0.$$

Proof. Let us start by the proof of *1. To recover an explicit equation for the first-order pressure correction* p^1*, first we write the velocity equation:*

$$\rho^0\, \partial_t u^0 + \rho^0\, (u^0 \cdot \nabla) u^0 + \nabla p^1 = 0,$$

obtained by inserting the density equation into the momentum equation. Dividing this equation by ρ^0*,*

$$\partial_t u^0 + (u^0 \cdot \nabla) u^0 + \frac{1}{\rho^0} \nabla p^1 = 0,$$

taking the divergence

$$\nabla \cdot \partial_t u^0 + \nabla \cdot \Big((u^0 \cdot \nabla) u^0 \Big) + \nabla \cdot \left(\frac{1}{\rho^0} \nabla p^1 \right) = 0, \tag{2.9}$$

and using the incompressibility constraint $\nabla \cdot u^0 = 0$*, the derivative in time* $\nabla \cdot \partial_t u^0$ *vanishes and we obtain an elliptic equation for* p^1*:*

$$-\nabla \cdot \left(\frac{1}{\rho^0} \nabla p^1 \right) = \nabla \cdot \Big((u^0 \cdot \nabla) u^0 \Big). \tag{2.10}$$

Thus, changing (2.7c) *into* (2.10)*, we obtain* $\mathbf{RM_0}$*.*

Remark 1. In the literature, the reformulated model can be found written in physical variables with the density equation reduced to a transport equation and replacing (2.8b) by the velocity equation

$$\rho^0\, \partial_t u^0 + \rho^0\, (u^0 \cdot \nabla) u^0 + \nabla p^1 = 0.$$

Let us now prove 2. From the momentum equation (2.8b), we can still recover (2.9). Then, using (2.8c) we obtain for all $x \in \Omega$ and $t > 0$

$$\partial_t \left(\nabla \cdot u^0 \right) = 0,$$

and thus

$$\nabla \cdot u^0(\cdot, t) = \nabla \cdot u^0(\cdot, 0).$$

This shows that $\mathbf{RM_0} \implies \mathbf{M_0}$ *if and only if* $\nabla \cdot u^0(\cdot, 0) = 0$. □

In the next section, we address the numerical challenges that arise when dealing both with the compressible Euler model and its incompressible limit model.

2.1.3 Principle of asymptotic preserving schemes

First let us clarify that in the rescaled Euler system (2.1), our parameter ε is constant and is related to the reference Mach number. However, the physical Mach number can vary both in space and time. In practice, the schemes can be written on the non rescaled equations (1.1), but in order to study the low Mach number limit, it is convenient to work on the rescaled equations. In this section, in order to explain the numerical difficulties associated to the physical Mach number, we can equivalently consider in the rescaled system our parameter ε to vary in space and time.

The limit model (2.7) does no longer depend on ε and so is no more constrained by the small values of ε. But, it can be used only where ε is sufficiently small. Where ε takes on order one or intermediate values, the compressible Euler equations (2.1) must be used. Then, two models must be used which leads to other difficulties like the detection of the interface between the two models, the reconnection at the interface...
For example, in Figure 2.1, we represent ε as a function of the space variable. Around 0.75 on the x-axis, the interface between the compressible ($\varepsilon = 10^0$) and incompressible ($\varepsilon = 10^{-5}$) regime is very sharp and so it is clear where to switch models. On the other side, between 0 and 0.4, the interface is more diffused and therefore it becomes difficult to know where we can switch models. For the intermediate values of ε the limit model is not valid yet, so it cannot be used and so we must put the interface for sufficiently small values of ε. But, the only valid model for intermediate values or order one values of ε is the compressible model and an explicit discretization starts to become rather costly due to the restrictive C.F.L. (2.3) even for intermediate values of ε. In addition, as already mentioned, it can suffer from consistency issues.
Another possible solution consists in using only one model, the compressible Euler equations (2.1), valid everywhere and at every time. But, an asymptotic preserving scheme, free of the constraints related to the Mach number ε, must be used. Such schemes have been developed in the literature, see [25, 24, 26, 44, 75, 64, 9, 30, 10, 29] for the isentropic Euler system and [63, 21, 65, 37, 16, 30, 11] for the full Euler system. They permit to avoid the time step limitations, the schemes are said to

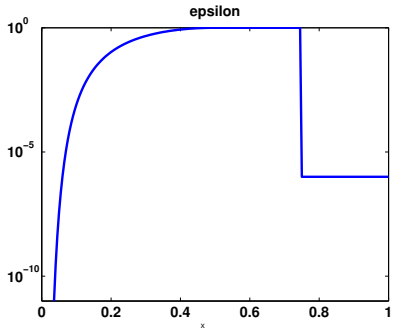


Figure 2.1 – Parameter ε as a function of the space domain. At $x = 0.75$, the interface between the compressible and limit model is sharp while between $x = 0$ and $x = 0.4$, we observe a diffused interface.

be asymptotically stable. And, they lead to consistent approximations of the limit incompressible model when the Mach number goes to zero, this corresponds to the asymptotic consistency property.
In practice, a possible strategy to obtain asymptotic preserving finite volume schemes consists in using IMEX (implicit-explicit) methods ([80], [66]). For such schemes, the C.F.L. restriction is only related to the part treated explicitly. In our case, the flux of the Euler system (2.1) is split into two parts

$$F = F_e + F_i.$$

The first one, F_e, will be treated explicitly while the other one, F_i, will be treated implicitly. This flux splitting must be well chosen in order to obtain both asymptotic stability and asymptotic consistency and such that the computational cost of the scheme is not too high especially when ε is of order 1, in compressible regimes. It is important to conserve the properties of the classical explicit schemes like the cost of the scheme and the preservation of stationary states. The goal is then to construct an asymptotic preserving scheme impliciting as few terms as possible.
If in the case of the isentropic Euler equations, the flux splitting is now well known, for the full Euler equations, it is not so clear since we can find different flux splittings in the literature leading to different asymptotic preserving schemes, see [21, 37, 16, 30, 10, 11]. Here, in Section 2.2, based on the analysis of the asymptotic preserving properties (stability and consistency), we select the flux splitting for constructing an IMEX AP (asymptotic preserving) scheme and we compare it to the different flux splittings proposed in the literature.
We propose in Section 2.3, a new linear asymptotic preserving scheme based on the non linear scheme proposed in [11]. We prove the asymptotic consistency as well as its preservation of contact discontinuities.
Furthermore, we numerically show that an upwinding on the implicit numerical fluxes is necessary for ensuring the low oscillatory property. Then, using a Fourier linear stability analysis, we prove that this new AP scheme is linearly L^2 stable under a C.F.L. condition independent of the Mach number ε. Additionally, this

analysis emphasizes that the upwinding on the implicit part improves the stability. Numerical simulations presented in Section 2.3.4 show the good behavior of the scheme in the non linear case. Like in [29] for the isentropic case, we propose in Section 2.4 a second order extension based on the ARS-IMEX scheme ([80]) and we use in Section 2.4.5 a MOOD process (see [33], [34]) in order to preserve the low oscillatory properties of the order one scheme to the second order schemes.

2.2 Analysis of the flux splitting

2.2.1 Choice of the flux splitting

In order to build an IMEX finite volume AP scheme in low Mach number regimes, we must decide which terms should be treated implicitly and so, we must set a flux decomposition $F = F_i + F_e$. If an order 1 scheme is considered, the parts F_e and F_i will be respectively explicitly and implicitly discretized leading to the following semi-discretization

$$\frac{W^{n+1} - W^n}{\Delta t} + \nabla \cdot F_e(W^n) + \nabla \cdot F_i(W^{n+1}) = 0. \tag{2.11}$$

We can find different flux splittings in the literature leading to different asymptotic preserving schemes, see [21, 16, 37, 30, 10, 77, 11]. Let us first note that a fully implicit scheme ($F_e = 0$, $F_i = F$) is of course asymptotic preserving but its cost, especially in compressible regimes, is too high due to the nonlinearity of the system. Moreover, the introduction of implicit terms increases the viscosity of the scheme and thus decreases its accuracy. Finally, an implicit treatment of the whole flux is not necessary to build an AP scheme. We therefore wish to treat implicitly as few terms as possible while guaranteeing the AP character of the scheme.
Furthermore, in the case of an explicit scheme, $F_e = F$ and $F_i = 0$, the semi-discretization reads

$$\frac{\rho^{n+1} - \rho^n}{\Delta t} + \nabla \cdot q^n = 0, \tag{2.12a}$$

$$\frac{q^{n+1} - q^n}{\Delta t} + \nabla \cdot \left(\frac{q^n \otimes q^n}{\rho^n} \right) + \frac{1}{\varepsilon} \nabla p^n = 0, \tag{2.12b}$$

$$\frac{E^{n+1} - E^n}{\Delta t} + \nabla \cdot \left((E^n + p^n) \frac{q^n}{\rho^n} \right) = 0, \tag{2.12c}$$

$$E^{n+1} = \frac{p^{n+1}}{\gamma - 1} + \varepsilon \frac{\rho^{n+1} |u^{n+1}|^2}{2}, \tag{2.12d}$$

and we have seen that it is not asymptotically stable. Moreover, the scheme is not asymptotically consistent either, meaning that we do not obtain a constant pressure and a divergence free velocity when ε tends to 0. Indeed, following the formal asymptotic limit in the continuous case, we can perform the formal asymptotic limit of the explicit semi-discretization. Then, multiplying (2.12b) by ε and passing to the limit gives only the constraint $\nabla p^n = 0$ on W^n when W^{n+1} is calculated.

It does not impose $\nabla p^{n+1} = 0$ and so an implicit discretization of the pressure gradient term in the momentum equation is necessary to obtain the asymptotic consistency. But, this is not sufficient since we must recover the incompressibility constraint $\nabla \cdot u^{n+1} = 0$ from the energy equation like in the continuous case (see Section 2.1.1). This means that a part of the energy flux term $(E + p)\, u$ must be treated implicitly.

To well choose the implicit part of the energy flux, we need to better understand the transition from the compressible Euler model (2.1) to the incompressible model (2.7). For this purpose, we complete and study in Sections 2.2.1.1 and 2.2.1.2 the following diagram (Figure 2.2).

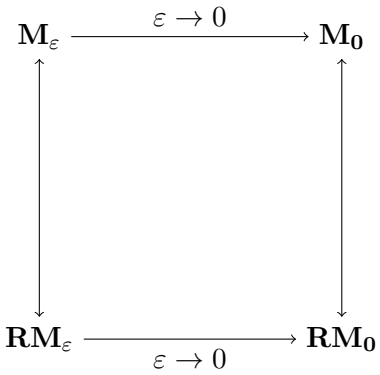


Figure 2.2 – $\mathbf{M}_\varepsilon$ is the compressible model given by (2.1), $\mathbf{M_0}$ is the incompressible limit model given by (2.7), $\mathbf{RM_0}$ is the reformulated incompressible model given by (2.8) and $\mathbf{RM}_\varepsilon$ the reformulated compressible model to be determined.

Until now, we have studied the upper part of the diagram. We first study the reformulation of the compressible model and its equivalence with the original model (see Section 2.2.1.1). Then, we look at the limit $\varepsilon \to 0$ of the resulting system and compare it to the reformulated incompressible model (see Section 2.2.1.2).

2.2.1.1 Reformulation of the full Euler system : pressure wave equation

As done for the incompressible model (see Section 2.1.2), we can reformulate the compressible model to obtain an explicit equation for the pressure.

Lemma 2.2.1. *(Formal)*

1. *The compressible Euler model* $\mathbf{M}_\varepsilon$ *given by* (2.1) *can be reformulated replacing the energy equation* (2.1c) *by a nonlinear wave equation for* p*.*

$$\mathbf{M}_\varepsilon \implies \mathbf{RM}_\varepsilon$$

where the reformulated compressible system $\mathbf{RM}_\varepsilon$ *reads*

$$\partial_t \rho + \nabla \cdot q = 0, \tag{2.13a}$$

$$\partial_t q + \nabla \cdot \left(\frac{q \otimes q}{\rho} \right) + \frac{1}{\varepsilon} \nabla p = 0, \tag{2.13b}$$

$$\partial_t \left(\frac{\partial_t p}{\gamma\, p} + \frac{u \cdot \nabla p}{\gamma\, p} \right) - \frac{1}{\varepsilon} \nabla \cdot \left(\frac{1}{\rho}\, \nabla p \right) = \nabla \cdot \Big((u \cdot \nabla)\, u \Big), \tag{2.13c}$$

$$E = \frac{p}{\gamma - 1} + \frac{\varepsilon}{2} \frac{|q|^2}{\rho}. \tag{2.13d}$$

2. *The reformulated compressible Euler system* $\mathbf{RM}_\varepsilon$ *is conditionally equivalent to the compressible Euler system* $\mathbf{M}_\varepsilon$*. More precisely*

$$\mathbf{RM}_\varepsilon \implies \mathbf{M}_\varepsilon$$

if and only if

$$\partial_t p(\,\cdot\,, 0) = -u(\,\cdot\,, 0) \cdot \nabla p(\,\cdot\,, 0) - \gamma p(\,\cdot\,, 0)\, \nabla \cdot u(\,\cdot\,, 0).$$

Proof. Let us start by the proof of *1. We begin from* (2.1)

$$\begin{aligned}
&\partial_t \rho + \nabla \cdot q = 0, \\
&\partial_t q + \nabla \cdot \left(\frac{q \otimes q}{\rho} \right) + \frac{1}{\varepsilon} \nabla p = 0, \\
&\partial_t E + \nabla \cdot \left((E + p)\, \frac{q}{\rho} \right) = 0.
\end{aligned}$$

The velocity equation is obtained from the momentum and mass conservation:

$$\rho\, \partial_t u + \rho\, (u \cdot \nabla)\, u + \frac{1}{\varepsilon}\, \nabla p = 0.$$

In the case of the incompressible system, we used the incompressibility constraint $\nabla \cdot u = 0$ *which comes from the limit of the energy equation and is no longer true in the compressible regime. But, we can still use the energy equation. Using the relations*

$$\begin{aligned}
E &= k_\varepsilon + \frac{p}{\gamma - 1}, \\
E + p &= k_\varepsilon + h,
\end{aligned}$$

where $k_\varepsilon = \varepsilon\, \rho\, \frac{|u|^2}{2}$ *is the kinetic energy,* $\frac{p}{\gamma - 1}$ *is the internal energy and* $h = \frac{\gamma p}{\gamma - 1}$ *is the enthalpy. The energy equation gives*

$$\begin{aligned}
\partial_t E + \nabla \cdot ((E + p)\, u) &= \partial_t k_\varepsilon + \partial_t \frac{p}{\gamma - 1} + \nabla \cdot ((k_\varepsilon + h)\, u) \\
&= \partial_t k_\varepsilon + \partial_t \frac{p}{\gamma - 1} + \nabla \cdot (k_\varepsilon\, u) + \nabla \cdot \left(\frac{\gamma p}{\gamma - 1} u \right) = 0.
\end{aligned}$$

Noting that $\frac{\gamma p}{\gamma-1}u = p\,u + \frac{1}{\gamma-1}\,p\,u$ *and* $\nabla\cdot(p\,u) = u\cdot\nabla p \;+\; p\nabla\cdot u$, *we get*

$$\partial_t k_\varepsilon + \partial_t\frac{p}{\gamma-1} + \nabla\cdot(k_\varepsilon\,u) + u\cdot\nabla p \;+ p\nabla\cdot u + \frac{1}{\gamma-1}\nabla\cdot(p\,u) = 0.$$

Then, using the mass and velocity equations, we eliminate the kinetic part since

$$\partial_t k_\varepsilon + \nabla\cdot(k_\varepsilon\,u) + u\cdot\nabla p \;= \varepsilon u\cdot\left(\rho\,\partial_t u + \rho\,(u\cdot\nabla)\,u + \frac{1}{\varepsilon}\,\nabla p\right) + \;\frac{\varepsilon}{2}|u|^2\,(\partial_t\rho + \nabla\cdot(\rho u)) = 0.$$

And so, having

$$\partial_t\frac{p}{\gamma-1} + p\nabla\cdot u + \frac{1}{\gamma-1}\nabla\cdot(p\,u) = 0,$$

the energy equation yields

$$\partial_t\frac{p}{\gamma-1} + \frac{u\cdot\nabla p}{\gamma-1} + \frac{\gamma p}{\gamma-1}\,\nabla\cdot u = 0. \tag{2.14}$$

Now, dividing the velocity equation $\rho\,\partial_t u + \rho\,(u\cdot\nabla)\,u + \frac{1}{\varepsilon}\,\nabla p = 0$ *by* ρ, *and taking the divergence of the resulting equation we obtain an expression for* $\nabla\cdot\partial_t u$:

$$\nabla\cdot\partial_t u + \;\nabla\cdot\Big((u\cdot\nabla)u\Big) + \frac{1}{\varepsilon}\,\nabla\cdot\left(\frac{1}{\rho}\,\nabla p\right) = 0. \tag{2.15}$$

Then, dividing the internal energy equation (2.14) *by the enthalpy* $\frac{\gamma p}{\gamma-1}$ *and taking the time derivative, the resulting equation is*

$$\partial_t\left(\frac{\partial_t p}{\gamma\,p} + \frac{u\cdot\nabla p}{\gamma\,p}\right) + \nabla\cdot\partial_t u = 0.$$

We insert (2.15) *into the previous equation and we obtain*

$$\partial_t\left(\frac{\partial_t p}{\gamma\,p} + \frac{u\cdot\nabla p}{\gamma\,p}\right) - \frac{1}{\varepsilon}\nabla\cdot\left(\frac{1}{\rho}\,\nabla p\right) = \nabla\cdot\Big((u\cdot\nabla)\,u\Big), \tag{2.16}$$

which is the nonlinear pressure wave equation in the fluid.
Thus, changing the energy equation (2.1c) *into* (2.16) *we recover* $\mathbf{RM}_\varepsilon$.

Let us prove 2. The mass and momentum equations yield an expression for $\partial_t\nabla\cdot u$ *given by* (2.15). *Then,* (2.13c) *yields*

$$\partial_t\left(\frac{\partial_t p}{\gamma\,p} + \frac{u\cdot\nabla p}{\gamma\,p} + \nabla\cdot u\right) = 0.$$

Integrating the above equation on time gives

$$\frac{\partial_t p(\cdot\,,t)}{\gamma p(\cdot\,,t)} + \frac{u(\cdot\,,t)\cdot\nabla p(\cdot\,,t)}{\gamma p(\cdot\,,t)} + \;\nabla\cdot u(\cdot\,,t) = \frac{\partial_t p(\cdot\,,0)}{\gamma p(\cdot\,,0)} + \frac{u(\cdot\,,0)\cdot\nabla p(\cdot\,,0)}{\gamma p(\cdot\,,0)} + \;\nabla\cdot u(\cdot\,,0). \tag{2.17}$$

Then, (2.17) holds the internal energy equation (2.14) if and only if it is satisfied at time $t = 0$. That is, if and only if

$$\partial_t p(\,\cdot\,,0) = -u(\,\cdot\,,0)\cdot\nabla p(\,\cdot\,,0) - \gamma p(\,\cdot\,,0)\,\nabla\cdot u(\,\cdot\,,0). \tag{2.18}$$

And, thus the energy equation (2.1c) is recovered if and only if (2.18) is satisfied. □

Next, let us investigate the limit $\varepsilon \to 0$ of $\mathbf{RM}_\varepsilon$ and compare it to the reformulated incompressible model (see 2.1.2).

2.2.1.2 Limit of the reformulated compressible model and reformulated incompressible model

We prove the following Lemma.

Lemma 2.2.2. *(Formal)*
The limit of the reformulated compressible model $\mathbf{RM}_\varepsilon$ given by (2.13) is not unconditionally equal to the reformulated incompressible model $\mathbf{RM_0}$ given by (2.8). More precisely

$$\lim_{\varepsilon\to 0} \mathbf{RM}_\varepsilon = \mathbf{RM_0}$$

if and only if

$$\partial_t p(\,\cdot\,,0) = 0.$$

Proof. Let us first study the formal limit of $\mathbf{RM}_\varepsilon$. We denote by $(\rho^\varepsilon, q^\varepsilon, E^\varepsilon, p^\varepsilon)$ the solution of (2.13) with general initial conditions and the impermeability boundary condition $u^\varepsilon\cdot\nu = 0$ on $\partial\Omega$. Inserting the asymptotic expansion of the previous solution gives:

$$\varepsilon^{-1}: \quad \nabla p^0 = 0, \tag{2.19a}$$

$$\nabla\cdot\left(\frac{1}{\rho^0}\nabla p^0\right) = 0, \tag{2.19b}$$

$$\varepsilon^{0}: \quad \partial_t\rho^0 + \nabla\cdot q^0 = 0, \tag{2.19c}$$

$$\partial_t q^0 + \nabla\cdot\left(\frac{q^0\otimes q^0}{\rho^0}\right) + \nabla p^1 = 0, \tag{2.19d}$$

$$\frac{d}{dt}\left(\frac{p^{0\prime}(t)}{\gamma\, p^0(t)}\right) - \nabla\cdot\left(\frac{1}{\rho^0}\nabla p^1\right) = \nabla\cdot\Big((u^0\cdot\nabla)\,u^0\Big), \tag{2.19e}$$

$$E^0(t) = \frac{p^0(t)}{\gamma-1}. \tag{2.19f}$$

Using (2.19c) into (2.19d) and taking the divergence yields

$$\nabla\cdot\partial_t u^0 + \nabla\cdot\Big((u^0\cdot\nabla)u^0\Big) = -\nabla\cdot\left(\frac{1}{\rho^0}\nabla p^1\right), \tag{2.20}$$

Inserting the result into the limit pressure equation (2.19e) gives

$$\frac{\partial}{\partial_t}\left(\frac{p^{0'}(t)}{\gamma\, p^0(t)} - \nabla\cdot u^0(x,t)\right) = 0. \tag{2.21}$$

Integrating (2.21) on Ω and using the boundary condition $u^0\cdot\nu = 0$ on $\partial\Omega$, we obtain for all $t>0$

$$\frac{\partial}{\partial t}\left(\frac{|\Omega| p^{0'}(t)}{\gamma\, p^0(t)} - \int_{\partial\Omega} u^0(x,t)\cdot\nu(x)\, d\sigma(x)\right) = 0.$$

And so,

$$|\Omega|\frac{\partial}{\partial_t}\left(\frac{p^{0'}(t)}{\gamma\, p^0(t)}\right) = 0. \tag{2.22}$$

Then, with (2.19e) we recover the reformulated pressure equation

$$-\nabla\cdot\left(\frac{1}{\rho^0}\nabla p^1\right) = \nabla\cdot\Big((u^0\cdot\nabla)\, u^0\Big).$$

Now, that we know the limit of $\mathbf{RM}_\varepsilon$ we can prove under which condition it is equal to the reformulated incompressible model. In $\mathbf{RM_0}$, the state equation reads

$$E^0 = \frac{p^0}{\gamma-1} = \frac{1}{|\Omega|}\int_\Omega E^0(x,0)\, dx,$$

whereas $\lim_{\varepsilon\to 0}\mathbf{RM}_\varepsilon$ only gives a constant pressure and energy in space but not in time. We have for all $t>0$

$$E^0(t) = \frac{p^0(t)}{\gamma-1}.$$

Furthermore, integrating (2.21) in time and space gives for $t>0$

$$\big(p^0(t)\big)' = p^0(t)\frac{1}{|\Omega|}\int_\Omega \frac{\partial_t p^0(x,0)}{p^0(x,0)} dx,$$

And so for all $t\geq 0$, $\partial_t p^0(\,\cdot\,,t) = 0$ if and only if $\partial_t p^0(\,\cdot\,,0) = 0$. In this case, we recover for all $t>0$

$$E^0 = \frac{p^0}{\gamma-1} = \frac{1}{|\Omega|}\int_\Omega \frac{p^0(x,0)}{\gamma-1}\, dx = \frac{1}{|\Omega|}\int_\Omega E^0(x,0)\, dx.$$

□

To conclude on the remarks, the reformulation of the compressible model can be useful to better understand the transition to the low Mach number limit but the resulting system is not unconditionally equivalent. This also highlights the importance of constructing a scheme based on the conservative variables.

From a numerical point of view, looking at the pressure wave equation of the reformulated Euler system (2.13)

$$\partial_t \left(\frac{\partial_t p}{\gamma\, p} + \frac{u \cdot \nabla p}{\gamma\, p} \right) - \frac{1}{\varepsilon} \nabla \cdot \left(\frac{1}{\rho} \nabla p \right) = \nabla \cdot \Big((u \cdot \nabla)\, u \Big), \tag{2.23}$$

we see that if the second order term in space

$$\frac{1}{\varepsilon} \nabla \cdot \left(\frac{1}{\rho} \nabla p \right),$$

is explicit, we recover the classical constrained C.F.L. which imposes Δt of the order of $\sqrt{\varepsilon}\, \Delta x$ for ensuring the stability of the scheme.

Then, for constructing a scheme, for the Euler system (2.1), uniformly stable with respect to ε, the time discretization must lead to an implicit discretization of the above term and thus, we need an implicit discretization of the term

$$h\, \nabla \cdot u = \frac{\gamma\, p}{\gamma - 1} \nabla \cdot u,$$

in equation (2.14). And, since we must work with the conservative variables, especially when ε is of order 1 to compute the entropic solution, we need an implicit discretization of the term $\nabla \cdot (h\, u)$ in the energy flux in the Euler equations.

2.2.1.3 Chosen flux splitting : asymptotic preserving properties

Following the previous analysis, we conclude that a possible choice for the flux splitting is given by

$$F_e(W) = \begin{pmatrix} \rho u \\ \rho u \otimes u \\ k_\varepsilon(W)\, u \end{pmatrix}, \qquad F_i(W) = \begin{pmatrix} 0 \\ \frac{1}{\varepsilon}\, p\, Id_{\mathbb{R}^3} \\ h(W)\, u \end{pmatrix}, \tag{2.24}$$

where $k_\varepsilon(W) = \varepsilon\, \rho\, |u|^2/2$ and $h(W) = \frac{\gamma p}{\gamma - 1} = \gamma\, (E - k_\varepsilon(W))$.

Note that this flux splitting was first introduced in [38] but not in the context of AP schemes in the low Mach number limit. The authors consider this flux splitting for building schemes which ensure the recognition of contact discontinuities and shear waves. The same splitting is used in [37, 76, 11, 14] for constructing an AP scheme in the low Mach number limit for the full Euler equations with a finite volume discretization on staggered grids in [37], on collocated grids in [11, 14] and with a staggered DG (discontinuous Galerkine) method in [76].

The semi-discretization writes

$$\frac{\rho^{n+1}-\rho^n}{\Delta t}+\nabla\cdot q^n=0, \tag{2.25a}$$

$$\frac{q^{n+1}-q^n}{\Delta t}+\nabla\cdot\left(\frac{q^n\otimes q^n}{\rho^n}\right)+\frac{1}{\varepsilon}\nabla p^{n+1}=0, \tag{2.25b}$$

$$\frac{E^{n+1}-E^n}{\Delta t}+\nabla\cdot\left(k_\varepsilon(W^n)\,\frac{q^n}{\rho^n}\right)+\nabla\cdot\left(h(W^{n+1})\,\frac{q^{n+1}}{\rho^{n+1}}\right)=0, \tag{2.25c}$$

$$E^{n+1}=\frac{p^{n+1}}{\gamma-1}+\varepsilon\,\frac{\rho^{n+1}\,|u^{n+1}|^2}{2}. \tag{2.25d}$$

Let us study its properties. We prove the following result

Lemma 2.2.3. *Assuming impermeability boundary conditions $u\cdot\nu=0$ on $\partial\Omega$, the semi-discretization (2.25) satisfies the following properties:*

(i)- Necessary condition for the asymptotic stability: In one dimension, each of the Jacobian matrices DF_e and DF_i have real eigenvalues. Those of DF_e are bounded uniformly when ε tends to 0.

(ii)- Asymptotic consistency: Formally passing to the limit ε tends to zero into the semi-discretization in time, we recover an approximation of the incompressible Euler equations (2.7). In particular, we have $E^{n+1}=\frac{p^{n+1}}{\gamma-1}=\frac{1}{|\Omega|}\int_\Omega E^0(x)\,dx$, and so for all $n\geq 0$,

$$\nabla p^{n+1}=\nabla E^{n+1}=0,$$

and for all $n\geq 1$,

$$\nabla\cdot u^{n+1}=0.$$

(iii)- Preservation of contact discontinuities:

$$\left\{\begin{array}{l} Du^n(x)=0,\\ \nabla p^n(x)=0.\end{array}\right.\Rightarrow \textit{There exists a solution } W^{n+1} \textit{ such that } \left\{\begin{array}{l} Du^{n+1}(x)=0,\\ \nabla p^{n+1}(x)=0,\end{array}\right.$$

where $Du(x)=(\partial_{x_j}u_i(x))_{1\leq i,j\leq d}$ denotes the Jacobian matrix of $u=(u_1,\cdots,u_d)\in\mathbb{R}^d$.

Before proving this result, let us first make two important remarks.

1. Property (i) is related to the asymptotic stability of the scheme while Property (ii) is related to the asymptotic consistency and Property (iii) to the preservation of stationary incompressible solutions.

   Indeed, in the case of the isentropic Euler system, it has been shown in [30] with a one dimensional linear stability analysis, that for ensuring the asymptotic stability, the discretization must satisfy the following essential properties: the upwinding of the explicit numerical fluxes must be related only on

the eigenvalues of the explicit matrix DF_e, the Jacobian matrix of the explicit part of the flux. Then, the implicit part of the flux, F_i, can be discretized with a centered solver. In this case, the scheme will be only L^2 stable and some oscillations may appear during the simulations but they are damped for long times.

These oscillations can be removed, or at least reduced, using an upwind solver for the implicit part, this upwinding can be related on the eigenvalues of the implicit matrix DF_i, the Jacobian matrix of the part of the flux which is implicitly discretized. These eigenvalues are proportional to $1/\sqrt{\varepsilon}$ and so the added numerical viscosity can be important in the low Mach number regime. Nevertheless, the C.F.L. stability condition of such IMEX schemes is only related to the eigenvalues of the explicit matrix DF_e.

Following this result, for ensuring the asymptotic stability of the scheme, it is necessary that each of the Jacobian matrices (DF_e and DF_i) have real eigenvalues, and those of DF_e must be uniformly bounded when ε tends to 0. Furthermore, let us note that the hyperbolicty of each "sub-matrix", DF_e and DF_i, is not required neither in the linear stability analysis conducted in Section 2.3.3, neither in the numerical scheme. Then, we assume that they could have multiple real eigenvalues and not be diagonalizable.

Furthermore, Property (ii) shows that passing to the limit ε tends to zero into (2.11), (2.24), we recover an approximation of the incompressible Euler equations (2.7).

Finally, it is important to note that a fluid with constant velocity and pressure is a "stationary solution" of the compressible and incompressible Euler equations. Then, Property $iii)$ shows that the scheme preserves these stationary incompressible solutions and so contact discontinuities.

2. Note that the implicit treatment of the full "enthalpy term" in the energy equation is necessary for the asymptotic consistency: if we change the flux splitting by

$$F_e(W) = \begin{pmatrix} \rho\, u \\ \rho u \otimes u \\ k_\varepsilon(W)\, u + \beta\, h(W)\, u \end{pmatrix}, \quad F_i(W) = \begin{pmatrix} 0 \\ \dfrac{1}{\varepsilon}\, p\, Id_{\mathbb{R}^3} \\ (1-\beta)\, h(W)\, u \end{pmatrix},$$

with $\beta \in [0, 1[$, then Property (ii) is lost. Indeed, formally passing to the limit ε tends to 0 into the momentum equation, we obtain like for $\beta = 0$, $\nabla p^{n+1} = \nabla E^{n+1} = 0$ for all $n \geq 0$. Then, integrating on Ω the energy equation using the boundary condition $u^n \cdot \nu = 0$ on $\partial\Omega$ for all $n \geq 0$ (where ν is the unit normal to $\partial\Omega$ outward to Ω) gives:

$$|\Omega| E^{n+1} = \int_\Omega E^n(x)dx + \int_{\partial\Omega} (k_\varepsilon^n(x)\, \beta h^n(x)) u^n(x) \cdot \nu(x)\, d\sigma(x)$$
$$+ \int_{\partial\Omega} (1-\beta) h^{n+1}(x) u^{n+1}(x) \cdot \nu(x)\, d\sigma(x) = \int_\Omega E^n(x)dx = \int_\Omega E^0(x)dx.$$

Then, for all $n \geq 1$, $h^{n+1} = h^n = \gamma \int_\Omega E^0(x)dx$. But now, the energy equation gives

$$E^{n+1} = E^n - \Delta t\, \beta\, h^n \nabla \cdot u^n - \Delta t\, (1-\beta)\, h^{n+1} \nabla \cdot u^{n+1},$$

and implies, for all $n \geq 1$,

$$\nabla \cdot u^{n+1} = \frac{\beta}{\beta - 1} \nabla \cdot u^n.$$

The incompressibility constraint is recovered only if $\beta = 0$ or if the initial velocity is well-prepared to the low Mach number regimes, that is, close to a divergence free velocity. And so, the energy flux splitting corresponding to $\beta = 0$ is the best one for ensuring the asymptotic consistency with general initial data. To the best of our knowledge, this is the first time that it is explained why this is the better choice for building a finite volumes IMEX AP scheme in the low Mach number limit.

Proof. Concerning assertion (i). In one dimension, the Jacobian matrices $DF_e(W)$ and $DF_i(W)$ associated to respectively the explicit and implicit fluxes are given by

$$DF_e(W) = \begin{pmatrix} 0 & 1 & 0 \\ -u^2 & 2u & 0 \\ -\varepsilon u^3 & \dfrac{3}{2}\varepsilon u^2 & 0 \end{pmatrix},$$

$$DF_i(W) = \begin{pmatrix} 0 & 0 & 0 \\ \frac{\gamma-1}{2}u^2 & -(\gamma-1)u & \frac{\gamma-1}{\varepsilon} \\ -\frac{\gamma E}{\rho}u + \gamma\varepsilon\, u^3 & \frac{\gamma E}{\rho} - \frac{3\varepsilon\gamma}{2}u^2 & \gamma u \end{pmatrix}.$$

A simple calculation shows that the eigenvalues of DF_e are given by 0 and u of multiplicity 2 while those of DF_i are given by 0 and $u/2 \pm \sqrt{u^2/4 + c^2/\varepsilon}$ where we recall that $c^2 = \gamma\, p/\rho$.

It has been already proven in [11], that this spliting satisfies (ii) with initial conditions which are not necessarily well prepared to the low Mach number regime (that is, initial conditions with a pressure and a velocity that are not necessarily close to a constant and a divergence free vector field, respectively). Let us recall the proof. Formally passing to the limit ε tends to 0 into the momentum equation multiplied by ε, we obtain $\nabla p^{n+1} = 0$ and thanks to the limit equation of state, we obtain $\nabla E^{n+1} = 0$. Integrating, the limit of the energy equation and using the boundary condition ($u^{n+1} \cdot \nu = u^n \cdot \nu = 0$ on $\partial\Omega$ where ν is the unit normal to $\partial\Omega$ outward to Ω) yields for all $n \geq 0$,

$$|\Omega| E^{n+1} = \int_\Omega E^n(x)dx + \int_{\partial\Omega} k_\varepsilon^n(x)\, u^n(x) \cdot \nu(x)\, d\sigma(x) + \int_{\partial\Omega} h^{n+1}(x) u^{n+1}(x) \cdot \nu(x)\, d\sigma(x) = \int_\Omega E^n(x)dx = \int_\Omega E^0(x)dx.$$

We recover that for all $n \geq 0$,

$$E^{n+1} = \frac{1}{|\Omega|} \int_\Omega E^0(x)\, dx.$$

Then, $E^{n+1} = E^n$ for all $n \geq 1$, and so using another time the limit energy equation: $E^{n+1} = E^n - \Delta t\, h^{n+1} \nabla \cdot u^{n+1}$, we recover for all $n \geq 1$, the incompressibility constraint

$$\nabla \cdot u^{n+1}(x) = 0,$$

for all $x \in \Omega$. Moreover, if the initial data are well prepared, that is, if $\nabla E(\,\cdot\,, 0) = 0$ and so $\nabla E^0 = 0$, then, $E^1 = E^0$ and we also recover

$$\nabla \cdot u^1(x) = 0,$$

for all $x \in \Omega$. Otherwise, if the initial data are not well prepared, that is, if $\nabla E(\,\cdot\,, 0) \neq 0$, then we have $E^1 = \dfrac{1}{|\Omega|} \displaystyle\int_\Omega E^0(x)\, dx$ and

$$\nabla \cdot u^1(x) = -\frac{E^0(x) - E^1}{\Delta t\, \gamma E^1},$$

for all $x \in \Omega$.
Let us now prove (iii). If $u^n(x) = \overline{u} \in \mathbb{R}^d$ and $p^n(x) = \overline{p} > 0$ are constant, then the mass equation gives

$$\rho^{n+1} = \rho^n - \Delta t\, \overline{u} \cdot \nabla \rho^n.$$

And, thanks to the momentum equation

$$\rho^{n+1}\, u^{n+1} = (\rho^n - \Delta t\, \overline{u} \cdot \nabla \rho^n)\, \overline{u} - \frac{\Delta t}{\varepsilon} \nabla p^{n+1} = \rho^{n+1}\, \overline{u} - \frac{\Delta t}{\varepsilon} \nabla p^{n+1}.$$

And so, there exists a solution such that $u^{n+1} = \overline{u}$ and $p^{n+1} = \overline{p}$, since the previous equation yields $\rho^{n+1}\, \overline{u} = \rho^{n+1}\, \overline{u}$ and the energy equation yields

$$\begin{aligned} E^{n+1} &= E^n - \Delta t\, \overline{u} \cdot \nabla k_\varepsilon^n \\ &= \frac{\overline{p}}{(\gamma - 1)} + \varepsilon\, \frac{\rho^n\, |\overline{u}|^2}{2} - \Delta t\, \overline{u} \cdot \nabla \left(\varepsilon\, \frac{\rho^n\, |\overline{u}|^2}{2} \right) \\ &= \frac{\overline{p}}{(\gamma - 1)} + \frac{\varepsilon}{2}\, |\overline{u}|^2\, (\rho^n - \Delta t\, \overline{u} \cdot \nabla \rho^n)\,. \end{aligned}$$

And so, we recover the state equation $E^{n+1} = \frac{\overline{p}}{(\gamma-1)} + \varepsilon\, \frac{\rho^{n+1}\, |\overline{u}|^2}{2}$. □

2.2.2 State of the Art of all Mach number IMEX finite volume schemes

Let us look at the properties $(i) - (iii)$ of Lemma 2.2.3 for some flux splittings found in the literature.

2.2.2.1 First family of splittings

It is not possible to be exhaustive, as there are far too many works on the subject, so we limit ourselves to flux splittings close to the one considered in this work. In the pioneering work [21], the authors proposed the following semi-discretization:

$$\frac{\rho^{n+1}-\rho^n}{\Delta t} + \nabla\cdot q^n = 0, \tag{2.26a}$$

$$\frac{q^{n+1}-q^n}{\Delta t} + \nabla\cdot\left(\frac{q^n\otimes q^n}{\rho^n}\right) + \alpha\,\nabla p^n + (1/\varepsilon-\alpha)\,\nabla p^{n+1} = 0, \tag{2.26b}$$

$$\frac{E^{n+1}-E^n}{\Delta t} + \nabla\cdot\left((E^n+p^n)\,\frac{q^{n+1}}{\rho^n}\right) = 0, \tag{2.26c}$$

$$E^{n+1} = \frac{p^{n+1}}{\gamma-1} + \varepsilon\,\frac{\rho^n\,|u^n|^2}{2}, \tag{2.26d}$$

and where $\alpha \geq 0$ must be well chosen for ensuring the asymptotic stability and avoid oscillations. In many test cases, the authors choose $\alpha = 0$ but $\alpha = 10$ is also considered. The previous semi-discretization can be rewritten as

$$\frac{W^{n+1}-W^n}{\Delta t} + \nabla\cdot F(\tilde{W}^{n,n+1}) = 0,$$

with $F(\tilde{W}^{n,n+1})$ given by

$$\left(q^n,\, \rho^n u^n\otimes u^n + \alpha\, p^n\, Id_{\mathbb{R}^3} + \left(\frac{1}{\varepsilon}-\alpha\right)\, p^{n+1}\, Id_{\mathbb{R}^3}\,,\, \frac{E^n+p^n}{\rho^n}\, q^{n+1}\right), \tag{2.27}$$

and $\nabla\cdot F(\tilde{W}^{n,n+1})$ cannot be written in a conservative flux splitting form $\nabla\cdot F_e(W^n) + \nabla\cdot F_i(W^{n+1})$. And, even for $\alpha = 0$, it is a multi-step method since p^n depends on E^n and (ρ^{n-1}, q^{n-1}). Hence, it cannot either be written (in one dimension) into the non-conservative following form $A_e\,\partial_x W^n + A_i\,\partial_x W^{n+1}$. Then, (i) is meaningless and the semi-discretization is outside the scope of our study. Note that Property (ii) is satisfied up to $\mathcal{O}(\Delta t)$ but not (iii) in general. Let us briefly recall the proofs.

Proof. For the proof of (ii), passing to the limit into the momentum and state equations we have $\nabla p^{n+1} = \nabla E^{n+1} = 0$ for all $n \geq 0$. Then integrating the energy equation on Ω and using the impermeability boundary condition gives $E^{n+1} = 1/|\Omega|\int_\Omega E^0(x)dx$ for all $n \geq 1$. Noting that $\rho^{n+1} = \rho^n + \mathcal{O}(\Delta t)$, the energy equation gives

$$\nabla\cdot\left(\frac{\rho^{n+1}}{\rho^n}u^{n+1}\right) = \nabla\cdot\left(\frac{\rho^n+\mathcal{O}(\Delta t)}{\rho^n}u^{n+1}\right) = 0,$$

and so the incompressiblity constraint $\nabla\cdot u^{n+1} = 0$ is recovered up to an error of $\mathcal{O}(\Delta t)$ for all $n \geq 1$.
For the proof of (iii), if $u^n(x) = \overline{u} \in \mathbb{R}^d$ and $p^n(x) = \overline{p} > 0$ are constant, the mass equation gives $\rho^{n+1} = \rho^n - \Delta t\,\overline{u}\cdot\nabla\rho^n$. And thanks to the momentum equation

$\rho^{n+1}\,u^{n+1} = \rho^{n+1}\,\overline{u} - \Delta t(1/\varepsilon - \alpha)\nabla p^{n+1}$. Therefore, there exists a solution such that $u^{n+1} = \overline{u}$ and $p^{n+1} = \overline{p}$, since the previous equation yields $\rho^{n+1}\,\overline{u} = \rho^{n+1}\,\overline{u}$. Then, using $\rho^{n+1} = \rho^n + \mathcal{O}(\Delta t)$, the energy equation yields

$$\begin{aligned} E^{n+1} &= E^n - \Delta t\,\overline{u}\nabla\cdot\left((E^n+p^n)\frac{\rho^{n+1}}{\rho^n}\right) \\ &= \frac{\overline{p}}{(\gamma-1)} + \frac{\varepsilon}{2}\,\rho^{n-1}|\overline{u}|^2 - \Delta t\,\overline{u}\cdot\nabla\left(\left(\frac{\gamma}{\gamma-1}\overline{p} + \frac{\varepsilon}{2}\,\rho^{n-1}|\overline{u}|^2\right)\frac{\rho^n+\mathcal{O}(\Delta t)}{\rho^n}\right) \\ &= \frac{\overline{p}}{(\gamma-1)} + \frac{\varepsilon}{2}\,|\overline{u}|^2\left(\rho^{n-1} - \Delta t\,\overline{u}\cdot\nabla\rho^{n-1}\right) + \mathcal{O}(\Delta t^2). \end{aligned}$$

And so, using the mass transport equation $\rho^n = \rho^{n-1} - \Delta t\,\overline{u}\cdot\nabla\rho^{n-1}$, we recover the state equation

$$E^{n+1} = \frac{\overline{p}}{(\gamma-1)} + \frac{\varepsilon}{2}\,\rho^n\,|\overline{u}|^2,$$

only up to $\mathcal{O}(\Delta t^2)$. The contact discontinuities are therefore not exactly preserved. □

In [30], a modified flux splitting is proposed to obtain a one-step method. And so, the system can be rewritten in a non-conservative form as follows (in one dimension):

$$\frac{W^{n+1}-W^n}{\Delta t} + A_e(W^n, W^{n+1})\,\partial_x W^n + A_e(W^{n+1}, W^n)\;\partial_x W^{n+1} = 0. \tag{2.28}$$

In this work, a first flux splitting was proposed having the advantage that it lead to a conservative flux splitting. The main advantages of having a conservative form are that the theoretical results obtained for Implicit-Explicit (IMEX) methods can be applied. Unfortunately, both its conservative and non-conservative versions, do not satisfy Property (i), in particular the eigenvalues associated to the explicit part are not always real. We will also prove that Property (ii) is satisfied but not Property (iii). Then, they are able to modify the original flux splitting to satisfy Property (i). This last flux splitting (2.32) is the one used in [30] and the analysis is given below (see Lemma 2.2.5).

First let us present and study the properties of the first non-hyperbolic splitting and its conservative version. The original splitting consists in setting $\alpha = 0$ into (2.26) and changing into the energy equation (2.26c), $\nabla\cdot\left((E^n+p^n)q^{n+1}/\rho^n\right)$ by $\nabla\cdot\left((E^n+\tilde{p}^n)q^{n+1}/\rho^n\right)$ where $\tilde{p}^n = (\gamma-1)\,(E^n - \varepsilon\,\rho^n\,|u^n|^2/2)$, then the energy equation is changed into

$$\frac{E^{n+1}-E^n}{\Delta t} + \nabla\cdot\left(\left(\gamma E - \frac{(\gamma-1)\varepsilon}{2}\frac{|q|^2}{\rho}\right)^n\frac{q^{n+1}}{\rho^n}\right) = 0. \tag{2.29}$$

In [30], this semi-discretization is modified into a conservative one. This conservative semi-discretization is given by

$$\frac{W^{n+1}-W^n}{\Delta t} + \nabla\cdot F_e(W^n) + \nabla\cdot F_i(W^{n+1}) = 0, \tag{2.30a}$$

with

$$F_e(W) = \begin{pmatrix} q \\ \dfrac{q\otimes q}{\rho} - \dfrac{\gamma-1}{2}\left(\dfrac{|q|^2}{\rho}\right) Id_{\mathbb{R}^3} \\ -\dfrac{(\gamma-1)\varepsilon}{2}\dfrac{|q|^2 q}{\rho^2} \end{pmatrix}, \quad F_i(W) = \begin{pmatrix} 0 \\ \dfrac{\gamma-1}{\varepsilon}E \\ \dfrac{\gamma E q}{\rho} \end{pmatrix}. \tag{2.30b}$$

It is obtained expressing p^{n+1} in the momentum equation in terms of W^n and E^{n+1} using the state equation $E^{n+1} = p^{n+1}/(\gamma-1) + \varepsilon\,\rho^n\,|u^n|^2/2$.

Lemma 2.2.4. *Assuming impermeability boundary conditions $u\cdot\nu = 0$ on $\partial\Omega$, the semi-discretization (2.30) does not satisfy Properties (i) and (iii) of Lemma 2.2.3 but does satisfy Property (ii).*

Proof. Let us prove *(i). The Jacobian matrices DF_e and DF_i associated respectively to F_e and F_i are*

$$DF_e(W^n) = \begin{pmatrix} 0 & 1 & 0 \\ \dfrac{\gamma-3}{2}(u^n)^2 & (3-\gamma)u^n & 0 \\ (\gamma-1)(u^n)^3 & -(\gamma-1)\varepsilon\dfrac{3(u^n)^2}{2}, & 0 \end{pmatrix},$$

$$DFi(W^{n+1}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \dfrac{\gamma-1}{\varepsilon} \\ -\dfrac{\gamma E^{n+1}}{(\rho^n+1)^2} & 0 & \dfrac{\gamma}{\rho^{n+1}} \end{pmatrix}. \tag{2.31}$$

Then, the eigenvalues of DF_e are

$$(3-\gamma)\,\frac{|u|}{2} \pm \sqrt{(\gamma-3)\,(\gamma-1)}\frac{|u|}{2}, \qquad 0.$$

Those of DF_i are 0 with multiplicity 2 and γ/ρ. Since $\gamma > 1$, the explicit eigenvalues are complex when $\gamma < 3$ and so does not satisfy Property (i).

Remark 2. Let us note that, the non-conservative form (where the energy is given by (2.29)) does not satisfy Property (i) either. The eigenvalues of the explicit matrix $A_e(W^n, W^{n+1})$ are $(3-\gamma)\,u^n/2 \pm \sqrt{(\gamma-3)\,(\gamma-1)}|u^n|/2$ and $\gamma\,q^{n+1}/\rho^n$. Those of the implicit matrix $A_i(W^n, W^{n+1})$ are $\pm\sqrt{(\gamma-1)\,(E^n+\tilde{p}^n)/(\rho^{n+1}\,\varepsilon)}$ and 0.

For the proof of (ii), passing to the limit in the momentum equation multiplied by ε and the state equation gives $\nabla E^{n+1} = \nabla p^{n+1} = 0$ for all $n \geq 0$. Then, integrating on Ω the limit energy equation and using the boundary condition $u^n\cdot\nu = 0$ on $\partial\Omega$ gives for all $n \geq 1$,

$$E^{n+1} = E^n = \frac{1}{|\Omega|}\int_\Omega E^0(x)dx.$$

Using again the limit energy equation, we recover the incompressibility constraint

$$\nabla\cdot u^{n+1} = 0,$$

for all $n \geq 1$.
For the proof of (iii), if $u^n(x) = \overline{u} \in \mathbb{R}^d$ *and* $p^n(x) = \overline{p} > 0$ *are constant, then the mass equation gives* $\rho^{n+1} = \rho^n - \Delta t\, \overline{u} \cdot \nabla \rho^n$. *Now, using the state equation, the momentum equation* $\rho^{n+1}\, u^{n+1} = (\rho^n - \Delta t\, \overline{u} \cdot \nabla \rho^n)\, \overline{u} - \Delta t/\varepsilon \nabla p^{n+1} = \rho^{n+1}\, \overline{u} - \Delta t/\varepsilon \nabla p^{n+1}$. *So* $u^{n+1} = \overline{u}$ *and* $p^{n+1} = \overline{p}$ *is a solution for the momentum equation. Now, the energy equation yields*

$$\begin{aligned} E^{n+1} &= E^n + \Delta t\, (\gamma - 1) \overline{u} \cdot \nabla \left(\frac{\varepsilon}{2}\, \rho^n\, |\overline{u}|^2 \right) - \Delta t \left(\frac{\gamma E^{n+1} q^{n+1}}{\rho^{n+1}} \right) \\ &= \frac{\overline{p}}{(\gamma - 1)} + \frac{\varepsilon}{2}\, \rho^{n-1} |\overline{u}|^2 - \Delta t\, (\gamma - 1) \overline{u} \cdot \nabla \overline{p} + \Delta t \overline{u} \cdot \nabla E^{n+1} \\ &= \frac{\overline{p}}{(\gamma - 1)} + \frac{\varepsilon}{2}\, |\overline{u}|^2 \left(\rho^{n-1} - \Delta t \overline{u} \cdot \nabla \rho^n \right) \\ &= \frac{\overline{p}}{(\gamma - 1)} + \frac{\varepsilon}{2}\, |\overline{u}|^2 \rho^n \ + \mathcal{O}(\Delta t^2), \end{aligned}$$

where we used the mass equation $\rho^n = \rho^{n-1} - \Delta t \overline{u} \cdot \nabla \rho^{n-1}$ *twice. Thus, the state* $E^{n+1} = \frac{\overline{p}}{(\gamma - 1)} + \frac{\varepsilon}{2}\, \rho^n\, |\overline{u}|^2$ *is not exactly satisfied.*

□

To recover real eigenvalues for the explicit part, a non conservative flux splitting is proposed in [30]. It consists in replacing $\nabla \cdot \left((E^n + p^n) q^{n+1} / \rho^n \right)$ by $\nabla \cdot \left((E^n + \tilde{p}^n) q^{n+1} / \rho^n \right)$ but this time, an implicit discretization of the mass equation is considered. The semi-discretization is given by

$$\frac{\rho^{n+1} - \rho^n}{\Delta t} + \nabla \cdot q^{n+1} = 0, \tag{2.32a}$$

$$\frac{q^{n+1} - q^n}{\Delta t} + \nabla \cdot \left(\frac{q^n \otimes q^n}{\rho^n} \right) + \frac{1}{\varepsilon}\, \nabla p^{n+1} = 0, \tag{2.32b}$$

$$\frac{E^{n+1} - E^n}{\Delta t} + \nabla \cdot \left((E^n + \tilde{p}^n)\, \frac{q^{n+1}}{\rho^n} \right) = 0, \tag{2.32c}$$

$$E^{n+1} = \frac{p^{n+1}}{\gamma - 1} + \varepsilon\, \frac{\rho^n\, |u^n|^2}{2}, \tag{2.32d}$$

where we recall that $\tilde{p}^n = (\gamma - 1)\, (E^n - \varepsilon\, \rho^n\, |u^n|^2 / 2)$. The density is treated explicitly into the energy flux in order to have an uncoupled scheme. Note that this flux splitting is also used in [10].
The flux splitting can be rewritten in one dimension, into the form $A_e\, \partial_x W^n + A_i\, \partial_x W^{n+1}$ (2.28) where

$$A_e(W^n, W^{n+1}) = \begin{pmatrix} 0 & 0 & 0 \\ \dfrac{\gamma - 3}{2} (u^n)^2 & (3 - \gamma) u^n & 0 \\ \left(-\dfrac{\gamma E^n}{(\rho^n)^2} + (\gamma - 1) \varepsilon \dfrac{(u^n)^2}{\rho^n} \right) q^{n+1} & -(\gamma - 1) \varepsilon u^n \dfrac{q^{n+1}}{\rho^n} & \gamma \dfrac{q^{n+1}}{\rho^n} \end{pmatrix},$$

$$A_i(W^n, W^{n+1}) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & \dfrac{\gamma-1}{\varepsilon} \\ 0 & \dfrac{E^n + \tilde{p}^n}{\rho^n} & 0 \end{pmatrix}. \tag{2.33}$$

Lemma 2.2.5. *Assuming impermeability boundary conditions $u \cdot \nu = 0$ on $\partial\Omega$, the semi-discretization (2.32) satisfies the necessary condition for asymptotic stability: Property (i) of Lemma 2.2.3 but does not satisfy Properties (ii) and (iii), respectively the asymptotic consistency and the preservation of contact discontinuities.*

Proof. The eigenvalues of $A_e(W^n, W^{n+1})$ are given by 0, $\gamma\, q^{n+1}/\rho^n$, $(3-\gamma)u^n$ and those of $A_i(W^n, W^{n+1})$ are

$$\pm\sqrt{\frac{\gamma-1}{\varepsilon}\,\frac{E^n + \tilde{p}^n}{\rho^n}}, \quad 0.$$

Then, Property (i) is satisfied.
Formally passing to the limit into (2.32) yields $\nabla p^{n+1} = \nabla E^{n+1} = 0$ and $E^{n+1} = \frac{p^{n+1}}{\gamma-1} = 1/|\Omega| \int_\Omega E^0(x)\, dx$ for all $n \geq 0$. But, now, the energy equation gives the incompressibility constraint up to an error of order Δt, since the energy equation gives for all $n \geq 1$: $\nabla \cdot (q^{n+1}/\rho^n) = \nabla \cdot (u^{n+1}\, \rho^{n+1}/\rho^n) = 0$ and so, using the mass equation gives $\nabla \cdot \left(\rho^{n+1}/\rho^n u^{n+1}\right) = \nabla \cdot u^{n+1} - \Delta t \nabla \cdot \left(\nabla \cdot q^{n+1}/\rho^n u^{n+1}\right) = 0$. And so, the asymptotic consistency is obtained up to an error of order Δt and (ii) is not exactly satisfied.
Now, let us prove that (iii) is not satisfied in general. If $u^n(x) = \overline{u} \in \mathbb{R}^d$ and $p^n(x) = \overline{p} > 0$ are constant, then we still have $\rho^{n+1}\, u^{n+1} = (\rho^n - \Delta t\, \overline{u} \cdot \nabla \rho^n)\, \overline{u} - \Delta t/\varepsilon \nabla p^{n+1}$. And, if $u^{n+1} = \overline{u}$ and $p^{n+1} = \overline{p}$, we obtain $\rho^{n+1}\, \overline{u} = (\rho^n - \Delta t\, \overline{u} \cdot \nabla \rho^n)\, \overline{u}$ but thanks to the mass equation,

$$\rho^{n+1} = \rho^n - \Delta t\, \nabla \cdot (\rho^{n+1}\, u^{n+1}) = \rho^n - \Delta t\, \overline{u} \cdot \nabla \rho^{n+1} \neq \rho^n - \Delta t\, \overline{u} \cdot \nabla \rho^n,$$

in general. We only get

$$\rho^{n+1}\, \overline{u} = (\rho^n - \Delta t\, \overline{u} \cdot \nabla (\rho^n + \mathcal{O}(\Delta t))\,)\, \overline{u} = \rho^{n+1}\, \overline{u} + \mathcal{O}(\Delta t^2).$$

Then, there does not always exist a solution such that $u^{n+1} = \overline{u}$ and $p^{n+1} = \overline{p}$. □

Finally, let us note that with this family of flux splittings, an important advantage is that the resulting schemes can be solved in an uncoupled way. This is very important because it allows the construction of schemes with a not to high computational cost. This is especially true when dealing with multi-dimensional systems, it avoids solving large linear systems. The uncoupled form of the scheme is obtained with a reformulation of the energy equation like in the continuous case (see Section 2.2.1.1). For the semi-discretization (2.32), we obtain inserting the expression of q^{n+1} given by (2.32b) into the energy equation (2.32c):

$$E^{n+1} = E^n - \Delta t\, \nabla \cdot \left(\frac{E^n + \tilde{p}^n}{\rho^n} \left(q^n - \Delta t \nabla \cdot \left(\frac{q^n \otimes q^n}{\rho^n} \right) - \Delta t \frac{1}{\varepsilon} \nabla p^{n+1} \right) \right) \tag{2.34}$$

Then, replacing p^{n+1} by the state equation $p^{n+1} = (\gamma - 1)(E^{n+1} - \varepsilon\,\rho^n\,|u^n|^2)$, we obtain a linear equation on the unknown E^{n+1}. With this information the momentum can be updated with (2.32b) (where the pressure p^{n+1} is expressed with the state equation) and finally the density is given by (2.32a).

2.2.2.2 Second family of splittings

Let us remark that all considered flux splittings do not perfectly separate pressure and fluid waves. Indeed, the eigenvalues of the implicit part depend on the fluid velocity in each case. It is in fact possible to find such a flux splitting. Let us consider the flux splitting inspired by the operator splitting strategy like proposed in [16]. Then, the semi-discretization consists in an explicit treatment for the transport terms with the fluid velocity:

$$\frac{\rho^{n+1} - \rho^n}{\Delta t} + \nabla \cdot q^n = 0, \tag{2.35a}$$

$$\frac{q^{n+1} - q^n}{\Delta t} + \nabla \cdot \left(\frac{q^n \otimes q^n}{\rho^n}\right) + \frac{1}{\varepsilon}\,\nabla p^{n+1} = 0, \tag{2.35b}$$

$$\frac{E^{n+1} - E^n}{\Delta t} + \nabla \cdot \left(E^n\,\frac{q^n}{\rho^n}\right) + \nabla \cdot \left(p^{n+1}\,\frac{q^{n+1}}{\rho^{n+1}}\right) = 0, \tag{2.35c}$$

$$E^{n+1} = \frac{p^{n+1}}{\gamma - 1} + \varepsilon\,\frac{\rho^{n+1}\,|u^{n+1}|^2}{2}. \tag{2.35d}$$

Lemma 2.2.6. *Assuming impermeability boundary conditions $u \cdot \nu = 0$ on $\partial\Omega$, the semi-discretization (2.35) satisfies the necessary condition for asymptotic stability and the preservation of contact discontinuities: Properties (i) and (iii) of Lemma 2.2.3 but satisfies the asymptotic consistency, Property (ii), only if the initial velocity is well-prepared to the low Mach number regime, more precisely if and only if $\lim_{\varepsilon\to 0} u(x,0) = u_0(x)$ with $\nabla \cdot u_0(x) = 0$.*

Proof. In one dimension, the Jacobian matrices associated to F_e and F_i are given by

$$DF_e(W) = \begin{pmatrix} 0 & 1 & 0 \\ -u^2 & 2u & 0 \\ -\dfrac{u}{\rho}E & \dfrac{E}{\rho} & u \end{pmatrix},$$

$$DF_i(W) = \begin{pmatrix} 0 & 0 & 0 \\ \dfrac{\gamma-1}{2}u^2 & -(\gamma-1)u & \dfrac{\gamma-1}{\varepsilon} \\ +\dfrac{\varepsilon c^2 u}{\gamma} - \dfrac{(\gamma-1)(\gamma-2)}{2\gamma}\varepsilon u^2 & -\dfrac{\varepsilon c^2}{\gamma} + \varepsilon(\gamma-1)u^2 & -(\gamma-1)u \end{pmatrix},$$

where the eigenvalue of DF_e is u of multiplicity 3 and those of the implicit matrix DF_i are 0 and $\pm\sqrt{(\gamma-1)\,c^2/(\gamma\,\varepsilon)}$. Then, Property (i) is satisfied and we can see that this flux splitting perfectly separates the transport and the pressure waves

since the eigenvalues of the implicit matrix do no more depend on u and those of the explicit matrix do not depend on c.
Let us look at the asymptotic consistency. Formally passing to the limit ε tends to 0 into the momentum equation multiplied by ε, we obtain $\nabla p^{n+1} = 0$ and thanks to the limit equation of state, we obtain $\nabla E^{n+1} = 0$. Integrating, the limit of the energy equation and using the boundary condition ($u^{n+1} \cdot \nu = u^n \cdot \nu = 0$ on $\partial\Omega$ where ν is the unit normal to $\partial\Omega$ outward to Ω), we recover that for all $n \geq 0$, $E^{n+1} = 1/|\Omega| \int_\Omega E^0(x)\, dx =< E_0 >$. And, using another time the limit energy equation we obtain

$$< E_0 > \nabla \cdot u^n(x) + (\gamma - 1) < E_0 > \nabla \cdot u^{n+1}(x) = 0,$$

for all $n \geq 1$ and $x \in \Omega$. And so, we recover Property (ii) if and only if $\nabla \cdot u^0(x) = 0$ that is if and only if the initial velocity is well-prepared to the low Mach number regime.
The flux splitting also satisfies Property (iii). Indeed, if $u^n(x) = \overline{u} \in \mathbb{R}^d$ and $p^n(x) = \overline{p} > 0$ are constant, then the mass equation gives $\rho^{n+1} = \rho^n - \Delta t\, \overline{u} \cdot \nabla \rho^n$. Now, thanks to the momentum equation $\rho^{n+1}\, u^{n+1} = (\rho^n - \Delta t\, \overline{u} \cdot \nabla \rho^n)\, \overline{u} - \Delta t/\varepsilon \nabla p^{n+1} = \rho^{n+1}\, \overline{u} - \Delta t/\varepsilon \nabla p^{n+1}$. And so, there exists a solution such that $u^{n+1} = \overline{u}$ and $p^{n+1} = \overline{p}$, since the previous equation yields $\rho^{n+1}\, \overline{u} = \rho^{n+1}\, \overline{u}$ and the energy equation yields

$$\begin{aligned} E^{n+1} &= E^n - \Delta t\, \overline{u} \cdot \nabla E^n \\ &= \frac{\overline{p}}{\gamma - 1} + \frac{\varepsilon}{2}\, \rho^n |\overline{u}|^2 - \Delta t\, \overline{u} \cdot \nabla (\frac{\varepsilon}{2}\, \rho^n |\overline{u}|^2\,) \\ &= \frac{\overline{p}}{\gamma - 1} + \frac{\varepsilon}{2}\, \rho^{n+1} |\overline{u}|^2. \end{aligned}$$

And so, the state equation is recovered. □

2.2.2.3 Flux splitting considered

Let us conclude this section with a review of existing low Mach number IMEX schemes using the flux splitting (2.24). We recall the semi-discretization:

$$\frac{\rho^{n+1} - \rho^n}{\Delta t} + \nabla \cdot q^n = 0, \tag{2.36a}$$

$$\frac{q^{n+1} - q^n}{\Delta t} + \nabla \cdot \left(\frac{q^n \otimes q^n}{\rho^n} \right) + \frac{1}{\varepsilon}\, \nabla p^{n+1} = 0, \tag{2.36b}$$

$$\frac{E^{n+1} - E^n}{\Delta t} + \nabla \cdot \left(k_\varepsilon(W^n)\, \frac{q^n}{\rho^n} \right) + \nabla \cdot \left(h(W^{n+1})\, \frac{q^{n+1}}{\rho^{n+1}} \right) = 0, \tag{2.36c}$$

$$E^{n+1} = \frac{p^{n+1}}{\gamma - 1} + \varepsilon\, \frac{\rho^{n+1}\, |u^{n+1}|^2}{2}. \tag{2.36d}$$

The first article in which (2.24) has been used for building an all Mach number scheme is [37], the scheme is based on staggered grids. In [11], an all Mach number

scheme on collocated grids is presented. The scheme consists in discretizing in space the semi-discretization (2.36). The mass equation can be advanced since the implicit flux is zero on it, then a nonlinear system on the momentum and energy has to be solved with a Picard algorithm. The resolution of this nonlinear system can be prepared to the low Mach number regimes by reformulating it: inserting the expression of q^{n+1} into the energy equation and using the state equation, yields

$$E^{n+1} = \frac{p^{n+1}}{\gamma-1} + k_\varepsilon^{n+1} = E^n - \Delta t \nabla \cdot \left(k_\varepsilon^n \frac{q^n}{\rho^n} \right) - \Delta t \nabla \cdot \left(\frac{h^{n+1}}{\rho^{n+1}} \left(q^n - \Delta t \nabla \cdot \left(\frac{q^n \otimes q^n}{\rho^n} \right) + \frac{1}{\varepsilon} \nabla p^{n+1} \right) \right),$$

Then, one recovers the discretization of the pressure wave equation into the fluid and the resulting scheme reads

$$q^{n+1,exp} = q^n - \Delta t \nabla \cdot (\rho^n u^n \otimes u^n), \tag{2.37a}$$

$$E^{n+1,exp} = E^n - \Delta t \, \nabla \cdot (k_\varepsilon^n \, u^n), \tag{2.37b}$$

$$\rho^{n+1} = \rho^n - \Delta t \nabla \cdot q^n, \tag{2.37c}$$

$$\frac{\varepsilon}{\gamma-1} p^{n+1} - \Delta t^2 \nabla \cdot \left(\frac{h^{n+1}}{\rho^{n+1}} \nabla p^{n+1} \right) = \varepsilon E^{n+1,exp} - \varepsilon k_\varepsilon^{n+1} - \varepsilon \Delta t \nabla \cdot \left(\frac{h^{n+1}}{\rho^{n+1}} q^{n+1,exp} \right), \tag{2.37d}$$

$$q^{n+1} = q^{n+1,exp} - \Delta t \frac{1}{\varepsilon} \nabla p^{n+1}, \tag{2.37e}$$

$$E^{n+1} = E^{n+1,exp} - \Delta t \nabla \cdot \left(h^{n+1} \frac{q^{n+1}}{\rho^{n+1}} \right). \tag{2.37f}$$

$$\tag{2.37g}$$

where

$$k_\varepsilon^{n+1} = k_\varepsilon(W^{n+1}) = \varepsilon \, \rho^{n+1} \frac{|u^{n+1}|^2}{2}, \quad h^{n+1} = \frac{\gamma \, p^{n+1}}{\gamma - 1}.$$

Note that, (2.37d) yields to an elliptic equation for determining p^{n+1}, but it is still coupled with the momentum equation. In [11], this reformulation is done on the fully discretized equations (in time and space) leading to a five points discretization of the second order operator on the pressure into (2.37d). To avoid checkerboard effects a non standard discretization of the enthalpy $\gamma \, p/((\gamma-1) \, \rho$ is introduced and then a Picard algorithm is used to solve this nonlinear system in terms of p and q. An unlimited in time order 2 discretization is also proposed and multidimensional simulations (2D and 3D) are performed. Note that, in [14], such a discretization is also proposed and extended to the Navier-Stokes system but the proposed linearization yields an all speed scheme such that the asymptotic consistency is obtained up to an order Δt term.

Here, we propose and study an all Mach number IMEX finite volume scheme also based on the flux splitting (2.24) but we prefer to discretize the previous reformulated semi-discrete equation instead of reformulating the full discrete one, thus, eliminating the checkerboard effect problems. Moreover, we propose a linearization of this reformulated equation in order to avoid the Picard algorithm for which the convergence is not always guaranteed during the simulations. The resolution is therefore uncoupled, since ρ is calculated first, then p, then q can be calculated and finally E is updated. We prove the asymptotic consistency on the semi-discretization as well as its preservation of contact discontinuities. Furthermore, we perform a linear stability analysis showing that our scheme is linearly L^2 stable and we propose a second order in space and time scheme. Then, we use a MOOD process to reduce the oscillations that are common for second order schemes.

2.3 Our new Order 1 AP scheme

2.3.1 A linear semi-discretization

Our linear first order all Mach number IMEX semi-discretization consists in replacing in (2.37d), the terms h^{n+1} and k_ε^{n+1} by their values calculated with the explicit convected part of the conservative variables. It is given by

$$q^{n+1,exp} = q^n - \Delta t \nabla \cdot (\rho^n u^n \otimes u^n), \tag{2.38a}$$

$$E^{n+1,exp} = E^n - \Delta t \, \nabla \cdot (k_\varepsilon^n \, u^n), \tag{2.38b}$$

$$\rho^{n+1} = \rho^n - \Delta t \nabla \cdot q^n, \tag{2.38c}$$

$$\frac{\varepsilon \, p^{n+1}}{\gamma - 1} - \Delta t^2 \nabla \cdot \left(\frac{h^{n+1,exp}}{\rho^{n+1}} \nabla p^{n+1} \right) = \varepsilon E^{n+1,exp} - \varepsilon k_\varepsilon^{n+1,exp} - \varepsilon \, \Delta t \nabla \cdot \left(\frac{h^{n+1,exp}}{\rho^{n+1}} q^{n+1,exp} \right), \tag{2.38d}$$

$$q^{n+1} = q^{n+1,exp} - \Delta t \frac{1}{\varepsilon} \nabla p^{n+1}, \tag{2.38e}$$

$$E^{n+1} = E^{n+1,exp} - \Delta t \nabla \cdot \left(\frac{\gamma \, p^{n+1}}{(\gamma - 1) \, \rho^{n+1}} q^{n+1} \right), \tag{2.38f}$$

where $k_\varepsilon^l = k_\varepsilon(W^l) = \varepsilon \, |q^l|^2/(2\rho^l)$, for $l \in \{n, "n+1, exp", n+1\}$ and the enthalpy in the pressure equation is defined by

$$h^{n+1,exp} = h(W^{n+1,exp}) = \gamma \, (E^{n+1,exp} - k_\varepsilon(W^{n+1,exp})).$$

Note that the scheme is linear and uncoupled since ρ^{n+1}, p^{n+1} , q^{n+1} and E^{n+1} can be computed sequentially. Moreover, note that like in [11], we do not use the state equation for computing E^{n+1} but the conservative equation (2.38f). Indeed, using the equation of state (2.1d) leads to inconsistent results: in some shock test cases the intermediate state is not well calculated, see Figure 2.3. We do not obtain the entropic solution. The same problem was noticed for the nonlinear scheme proposed in [11].

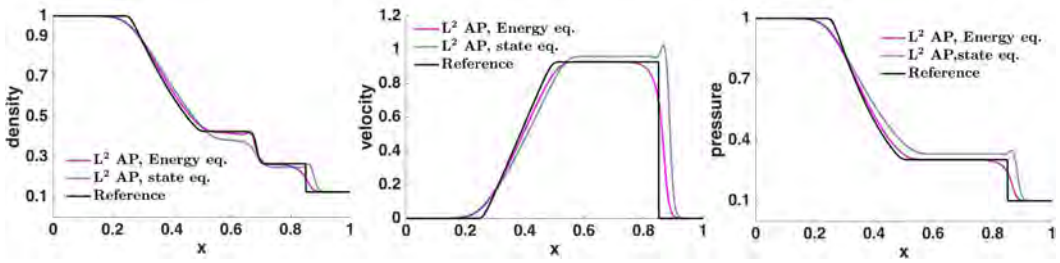


Figure 2.3 – State equation versus energy equation : Solution of the Sod shock tube problem (see Section 2.3.4.1) at $t_{final} = 0.2$ for 200 cells. Results for the Order 1 L^2 AP scheme using the equation of energy (2.38f) for updating the energy (pink lines) and using the equation of state (2.1d) for updating the energy (blue dotted lines).

Let us prove that this semi-discretization is asymptotically consistent and preserves the contact discontinuities:

Lemma 2.3.1. *The semi-discretization* (2.38) *satisfies Property* (*iii*) *of Lemma 2.2.3 and so preserves the contact discontinuities.*

Furthermore, assuming impermeability boundary conditions $u \cdot \nu = 0$ *on* $\partial\Omega$*, the semi-discretization* (2.38) *gives* $p_0^{n+1} = (\gamma - 1) < E_0^0 >$ *and* $< E_0^{n+1} >=< E_0^0 >$ *for all* $n \geq 0$*. Additionally, if the initial energy is well-prepared to the low Mach number regime, more precisely if* $\lim_{\varepsilon \to 0} E(x, 0) = \overline{E}_0$ *with* $\overline{E}_0$ *constant, the semi-discretization* (2.38) *is asymptotically consistent. The formal low Mach number limit of the system gives* $p^{n+1} = (\gamma - 1)\, E^{n+1} = (\gamma - 1)\overline{E}_0$ *and* $\nabla \cdot u^{n+1} = 0$ *for all* $n \geq 0$*.*

Proof. If $u^n(x) = \overline{u} \in \mathbb{R}^d$ and $p^n(x) = \overline{p} > 0$ are constant with $E^n(x) = \overline{p}/(\gamma - 1) + \varepsilon\, \rho^n\, |\overline{u}|^2/2$, then the mass equation gives $\rho^{n+1} = \rho^n - \Delta t\, \overline{u} \cdot \nabla \rho^n$. And we have

$$
\begin{aligned}
q^{n+1,exp} &= \rho^{n+1,exp}\, u^{n+1,exp} = \rho^n\, \overline{u} - \Delta t \overline{u} \cdot \nabla \rho^n\, \overline{u} = \rho^{n+1}\, \overline{u}, \\
k_\varepsilon^{n+1,exp} &= \frac{\varepsilon}{2}\, \rho^{n+1,exp}\, |u^{n+1,exp}|^2 = \frac{\varepsilon}{2}\, \rho^{n+1}\, |\overline{u}|^2, \\
E^{n+1,exp} &= E^n - \Delta t\, \frac{\varepsilon}{2}\, |\overline{u}|^2\, \overline{u} \cdot \nabla \rho^n \\
&= \frac{\overline{p}}{\gamma - 1} + \frac{\varepsilon}{2}\, \rho^n\, |\overline{u}|^2 - \Delta t\, \frac{\varepsilon}{2}\, |\overline{u}|^2\, \overline{u} \cdot \nabla \rho^n \\
&= \frac{\overline{p}}{\gamma - 1} + \frac{\varepsilon}{2}\, \rho^{n+1}\, |\overline{u}|^2, \\
h^{n+1,exp} &= \frac{\gamma \overline{p}}{\gamma - 1}.
\end{aligned}
$$

Now, the pressure equation and energy equations lead to

$$\begin{aligned}\frac{p^{n+1}}{\gamma-1} &= E^{n+1,exp} - k_\varepsilon^{n+1,exp} - \Delta t\nabla\cdot\left(\frac{\gamma\overline{p}}{\gamma-1}\frac{q^{n+1,exp}}{\rho^{n+1}}\right)\\ &= \frac{\overline{p}}{\gamma-1} - \Delta t\nabla\cdot\left(\frac{\gamma\overline{p}}{\gamma-1}\,\overline{u}\right)\\ &= \frac{\overline{p}}{\gamma-1},\\ E^{n+1} &= \frac{\overline{p}}{\gamma-1} + \frac{\varepsilon}{2}\,\rho^{n+1}\,|\overline{u}|^2 - \Delta t\nabla\cdot\left(\gamma\,\overline{p}\,u^{n+1}\right)\\ &= \frac{\overline{p}}{\gamma-1} + \frac{\varepsilon}{2}\,\rho^{n+1}\,|\overline{u}|^2.\end{aligned}$$

And so, there exists a solution W^{n+1} such that $u^{n+1} = \overline{u}$ and $p^{n+1} = \overline{p}$. The semi-discretization (2.38) preserves the contact discontinuities.
Let us prove the asymptotic consistency. We perform an asymptotic expansion, assuming that all the quantities $f^l = f_0^l + \varepsilon\, f_1^l$ for $l = n$, "$n+1, exp$" or $n+1$, then we obtain:

$$\varepsilon^{-1}: \quad \nabla\cdot\left(\frac{h_0^{n+1,exp}}{\rho_0^{n+1}}\nabla p_0^{n+1}\right) = 0, \tag{2.39a}$$

$$\nabla p_0^{n+1} = 0, \tag{2.39b}$$

$$\varepsilon^{0}: \quad q_0^{n+1,exp} = q_0^n - \Delta t\nabla\cdot(\rho_0^n u_0^n\otimes u_0^n), \tag{2.39c}$$

$$E_0^{n+1,exp} = E_0^n, \tag{2.39d}$$

$$\rho_0^{n+1} = \rho_0^n - \Delta t\nabla\cdot q_0^n \tag{2.39e}$$

$$\frac{p_0^{n+1}}{\gamma-1} = E_0^n - \Delta t\nabla\cdot\left(\frac{\gamma\,E_0^n}{\rho_0^{n+1}}q_0^{n+1}\right), \tag{2.39f}$$

$$q_0^{n+1} = q_0^{n+1,exp} - \Delta t\nabla p_1^{n+1}, \tag{2.39g}$$

$$E_0^{n+1} = E_0^n - \Delta t\frac{\gamma p_0^{n+1}}{\gamma-1}\nabla\cdot u_0^{n+1}. \tag{2.39h}$$

where $h_0^{n+1,exp} = \gamma\, E_0^n$.
Integrating the pressure equation (2.39f) on Ω and using the impermeability boundary condition $u_0^{n+1}\cdot\nu = 0$ on $\partial\Omega$, we get

$$|\Omega|\,\frac{p_0^{n+1}}{\gamma-1} = \int_\Omega E_0^n(x)\,dx - \Delta t\int_{\partial\Omega}\gamma\,E_0^n(x)u_0^{n+1}(x)\cdot\nu(x)d\sigma(x) = \int_\Omega E_0^n(x)\,dx.$$

And so, $p_0^{n+1} = (\gamma-1) < E_0^n >$ with $< E_0^n >= \frac{1}{|\Omega|}\int_\Omega E_0^n(x)\,dx$. Now, integrating the energy equation on Ω gives $< E_0^{n+1} >=< E_0^n >$. By induction, we have for all $n \geq 0$

$$p_0^{n+1} = (\gamma-1) < E_0^0 >, \quad < E_0^{n+1} >=< E_0^0 > .$$

Furthermore, assuming the initial energy is well-prepared i.e. $E_0^0 = \overline{E}_0$ is constant, then then $p_0^{n+1} = (\gamma-1)\,E_0^0 = (\gamma-1)\,\overline{E}_0$. In particular for $n = 0$, the pressure

equation writes

$$\frac{p_0^1}{\gamma-1} = E_0^0 = E_0^0 - \Delta t \gamma\, E_0^0 \nabla \cdot u_0^1,$$

and so $\nabla \cdot u_0^1 = 0$. On the other side, the energy equation gives for $n = 0$,

$$E_0^1 = E_0^0 - \Delta t \nabla \cdot \left(\frac{\gamma\, p_0^1}{(\gamma-1)} u_0^1 \right) = E_0^0 = \frac{p_0^1}{\gamma-1}.$$

By induction on the property "$\frac{p_0^{n+1}}{\gamma-1} = E_0^n = \overline{E}_0$", we obtain $\frac{p_0^{n+1}}{\gamma-1} = E_0^{n+1} = \overline{E}_0$ and $\nabla \cdot u_0^{n+1} = 0$ for all $n \geq 0$. □

Remark 3. For non well-prepared initial conditions, it is possible to recover the asymptotic consistency modifying the semi-discretization changing in (2.38f) the term $\gamma\, p^{n+1}/(\gamma-1)$ in the flux by $h^{n+1,exp}$. Let us briefly prove it. Performing an asymptotic expansion on the modified energy equation gives

$$E^{n+1} = E_0^n - \Delta t \nabla \cdot \left(\frac{\gamma\, E_0^n}{\rho_0^{n+1}} q_0^{n+1} \right). \tag{2.40}$$

And so by identification with the pressure equation (2.39f):

$$\frac{p_0^{n+1}}{\gamma-1} = E_0^n - \Delta t \nabla \cdot \left(\frac{\gamma\, E_0^n}{\rho_0^{n+1}} q_0^{n+1} \right),$$

we have $E_0^{n+1} = \dfrac{p_0^{n+1}}{\gamma-1}$. Then, E_0^{n+1} is constant and integrating (2.40) on Ω gives $E_0^{n+1} = E_0^n = \dfrac{1}{|\Omega|} \displaystyle\int_\Omega E_0^0(x)dx$ for all $n \geq 1$. Therefore, the incompressibility constraint is retrieved for all $n \geq 1$ without assuming well-prepared initial conditions.

However, this version of the semi-discretization leads to a less diffusive scheme in the L^2 stable version of the scheme (see next section) and so oscillations appear in some test cases (see Figure 2.4). These oscillations are diffused for large times and disappear in the L^∞ version of the scheme. They are the trace of the L^∞ instability of the L^2 stable version of the scheme which remains L^2 stable. We will see in the following that all L^∞ versions of the schemes, require an upwinding on the implicit flux proportional to $1/\sqrt{\varepsilon}$ and therefore does not allow to have an asymptotic accuracy. The L^2 versions, although showing oscillations, can be interesting for some test cases. This is why, even if the asymptotic consistency is obtained only for well prepared initial data, we choose to use the semi-discretisation (2.38).

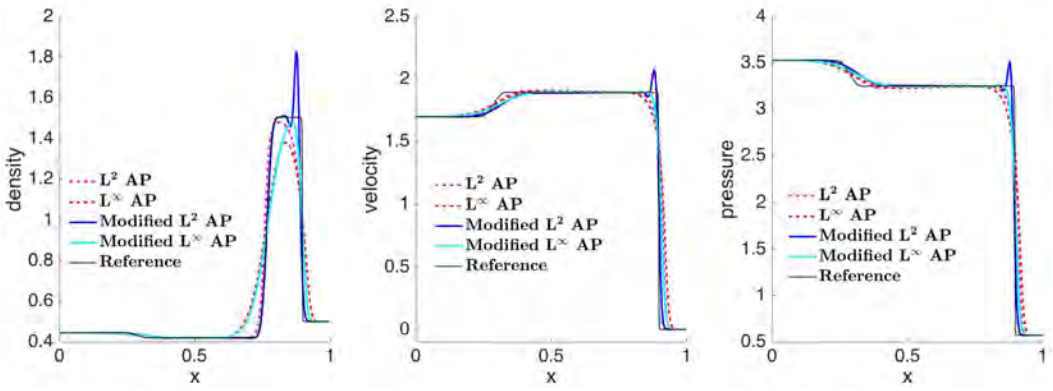


Figure 2.4 – Lax problem (see Section 2.3.4.1) for 200 cells. Results for the order 1 L^2 AP and L^∞ AP schemes associated to the semi-discretization (2.38) (pink and red curves) and the modified one (blue and cyan curves).

2.3.2 The order 1 schemes

In [29], it has been shown that a centered discretization for the implicit flux terms is sufficient to ensure an L^2 AP scheme if the explicit flux terms are discretized with an upwind discretization like Roe type solvers. The resulting scheme gives consistent and stable results but can present oscillations which are the signature of the non L^∞ stability. Since the scheme is L^2 stable, these oscillations do not propagate in the domain, remain localized and do not increase over time. This problem can be cured introducing an upwinding in the implicit discrete flux leading to the so-called L^∞ AP scheme.

2.3.2.1 L^2 stable discretization in one dimension

First, let us present the L^2 stable full discretizations in space and time for the previous semi-discretization (2.38). It is based on the modified Lax-Friedrichs solver for the explicit flux F_e and a centered solver for the implicit flux F_i.

We consider a uniform discretization in space and time for clarity, with $\Delta x > 0$ and $\Delta t > 0$ the space and time steps. The fully L^2 stable discrete version of (2.38) in

one dimension reads

$$W_j^{n+1,exp} = W_j^n - \Delta t \, \frac{(\mathcal{F}_e)_{j+1/2}^n - (\mathcal{F}_e)_{j-1/2}^n}{\Delta x}, \tag{2.41a}$$

$$\rho_j^{n+1} = \rho_j^n - \Delta t \frac{(\mathcal{F}_{e\rho})_{j+1/2}^n - (\mathcal{F}_{e\rho})_{j-1/2}^n}{\Delta x}, \tag{2.41b}$$

$$\begin{aligned} \frac{\varepsilon}{\gamma-1} p_j^{n+1} - \frac{\Delta t^2}{\Delta x}\left(\left(\frac{h^{n+1,exp}}{\rho^{n+1}}\right)_{j+1/2} \frac{p_{j+1}^{n+1} - p_j^{n+1}}{\Delta x} - \left(\frac{h^{n+1,exp}}{\rho^{n+1}}\right)_{j-1/2} \frac{p_j^{n+1} - p_{j-1}^{n+1}}{\Delta x} \right) \\ = \varepsilon E_j^{n+1,exp} - \varepsilon k_j^{n+1,exp} - \varepsilon \frac{\Delta t}{\Delta x}\left(\frac{h_{j+1}^{n+1,exp}}{\rho_{j+1}^{n+1}} q_{j+1}^{n+1,exp} - \frac{h_{j-1}^{n+1,exp}}{\rho_{j-1}^{n+1}} q_{j-1}^{n+1,exp} \right), \end{aligned} \tag{2.41c}$$

$$q_j^{n+1} = q_j^{n+1,exp} - \Delta t \, \frac{p_{j+1}^{n+1} - p_{j-1}^{n+1}}{2\varepsilon \Delta x}, \tag{2.41d}$$

$$E_j^{n+1} = E_j^{n+1,exp} - \frac{\Delta t}{2\Delta x}\left(\left(\frac{\gamma \, p^{n+1}}{(\gamma-1)\,\rho^{n+1}} q^{n+1} \right)_{j+1} - \left(\frac{\gamma \;\, p^{n+1}}{(\gamma-1)\,\rho^{n+1}} q^{n+1} \right)_{j-1} \right), \tag{2.41e}$$

where the explicit numerical flux $(\mathcal{F}_e)^n = ((\mathcal{F}_{e\rho})^n, (\mathcal{F}_{eq})^n, (\mathcal{F}_{eE})^n)$ is given by

$$(\mathcal{F}_e)_{j+1/2}^n = \frac{F_e(W_{j+1}^n) + F_e(W_j^n)}{2} - (\mathcal{D}_e)_{j+1/2}^n (W_{j+1}^n - W_j^n), \tag{2.41f}$$

where F_e is given by (2.24) and with the explicit viscosity coefficient,

$$(\mathcal{D}_e)_{j+1/2}^n = \frac{1}{2} \max\left(|u_{j+1}^n|, |u_j^n|\right),$$

taken as half of the maximum explicit eigenvalues of DF_e.
In (2.41c), h/ρ at the interfaces is computed as the arithmetic average:

$$\left(\frac{h^{n+1,exp}}{\rho^{n+1}}\right)_{j+1/2} = \frac{1}{2}\left(\frac{h_j^{n+1,exp}}{\rho_j^{n+1}} + \frac{h_{j+1}^{n+1,exp}}{\rho_{j+1}^{n+1}} \right).$$

In the following this scheme is called the Order 1 L^2 AP scheme.

2.3.2.2 L^∞ stable discretization in one dimension

Now, we present the scheme with an upwinding on the implicit fluxes. We first compute the L^2 stable solution $W_j^{n+1,L2}$ given by (2.41) and we add numerical dissipation as done for the explicit numerical flux $(\mathcal{F}_e)_{j+1/2}^n$, thus leading to a modified scheme for the density, momentum and energy equations.

$$W_j^{n+1} = \begin{pmatrix} \rho_j^{n+1,L2} \\ q_j^{n+1,L2} \\ E_j^{n+1,L2} \end{pmatrix} + \frac{\Delta t}{\Delta x}\left((\mathcal{D}_i)_{j+1/2}^n (W_{j+1}^{n+1} - W_j^{n+1}) - (\mathcal{D}_i)_{j-1/2}^n (W_j^{n+1} - W_{j-1}^{n+1}) \right), \tag{2.42}$$

where $(\mathcal{D}_i)^n_{j+1/2}$ is the implicit viscosity coefficient, taken as half of the maximum implicit eigenvalue

$$(\mathcal{D}_i)^n_{j+1/2} := \frac{1}{2}\max\left(|\lambda_i(W^n_{j+1})|, |\lambda_i(W^n_j)|\right),$$

where

$$|\lambda_i(W)| = \frac{|u|}{2} + \sqrt{\frac{u^2}{4} + \frac{c^2}{\varepsilon}}.$$

Let us note that, to avoid the computational effort of solving a nonlinear system, the numerical viscosity is applied after the calculation of the L^2 stable solution and the implicit viscosity $(\mathcal{D}_i)$ is chosen to be taken at time n and not $n+1$. In the following, this scheme is called the Order 1 L^∞ stable scheme.
Moreover, it is important to note that the upwinding on ρ must be applied after the calculation of the pressure, otherwise the scheme does not preserve exactly contact discontinuities (see Figure 2.5).

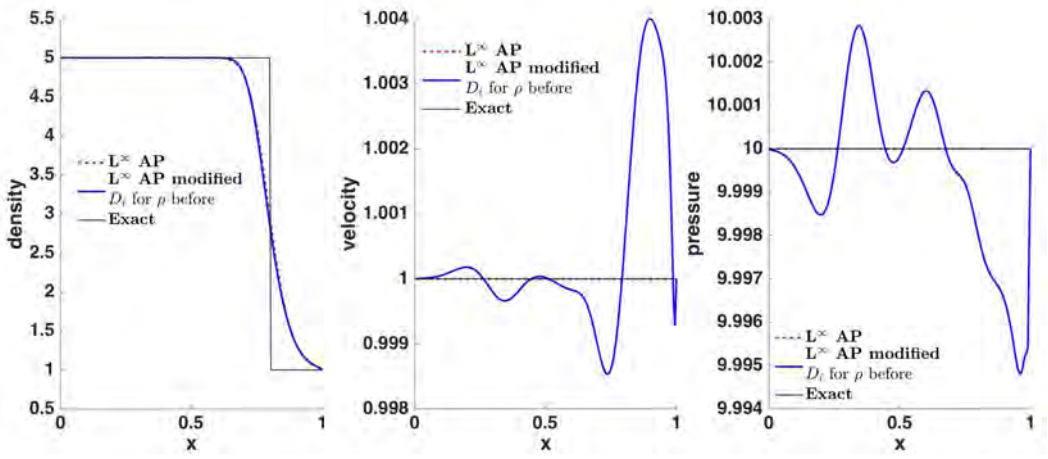


Figure 2.5 – Contact problem (see Section 2.3.4.1) for 200 cells. Results for the Order 1 L^∞ AP schemes associated to the semi-discretization (2.38) (red curve) and the modified one where the upwinding on the variable ρ is applied juster after calculing ρ^{n+1,L^2} with (2.41b) .

2.3.2.3 Modified nonlinear Order 1 AP scheme ([11])

In numerical simulations, we compare our scheme to the scheme proposed in [11] slightly modified. Indeed, in [11] the reformulated pressure equation is calculated on the fully discretized scheme giving a discretization of the elliptical term in the pressure equation spread over 5 cells. Here, we prefer to discretize the reformulated pressure equation (2.37d) and obtain a discretization of the elliptic term spread over 3 cells. This modification allows for a less diffusive scheme. Furthermore, in [11], the implicit part of the energy flux $\gamma/(\gamma-1)\,p/\rho\;q$ is discretized as the product of the centered approximations of $\gamma/(\gamma-1)\,p/\rho$ and q, here we prefer to use the centered

approximation of the quantity. This modification improves the convergence of the Picard algorithm. Then, the resulting nonlinear scheme, called NL Order 1 L^2 AP scheme consists in the following

$$\rho_j^{n+1} = \rho_j^n - \Delta t \frac{(\mathcal{F}_{e\rho})_{j+1/2}^n - (\mathcal{F}_{e\rho})_{j-1/2}^n}{\Delta x}, \tag{2.43a}$$

$$\begin{aligned}\frac{\varepsilon}{\gamma-1} p_j^{n+1} - \frac{\Delta t^2}{\Delta x}\left(\left(\frac{h^{n+1}}{\rho^{n+1}}\right)_{j+1/2} \frac{p_{j+1}^{n+1} - p_j^{n+1}}{\Delta x} - \left(\frac{h^{n+1}}{\rho^{n+1}}\right)_{j-1/2} \frac{p_j^{n+1} - p_{j-1}^{n+1}}{\Delta x}\right) \\ = \varepsilon\left(E_j^{n+1,exp} - k_j^{n+1}\right) - \varepsilon \frac{\Delta t}{\Delta x}\left(\left(\frac{h^{n+1}}{\rho^{n+1}} q^{n+1,exp}\right)_{j+1/2} \right. \\ \left. - \left(\frac{h^{n+1}}{\rho^{n+1}} q^{n+1,exp}\right)_{j-1/2}\right),\end{aligned} \tag{2.43b}$$

$$q_j^{n+1} = q_j^{n+1,exp} - \Delta t \, \frac{p_{j+1}^{n+1} - p_{j-1}^{n+1}}{2\varepsilon \Delta x}, \tag{2.43c}$$

$$E_j^{n+1} = E_j^{n+1,exp} - \frac{\Delta t}{2\Delta x}\left(\left(\frac{\gamma\, p^{n+1}}{(\gamma-1)\,\rho^{n+1}} q^{n+1}\right)_{j+1} - \left(\frac{\gamma\, p^{n+1}}{(\gamma-1)\,\rho^{n+1}} q^{n+1}\right)_{j-1}\right). \tag{2.43d}$$

We add the upwinding on the implicit part to the NL Order 1 L^2 AP scheme following the same process described in the previous section. We obtain the NL Order 1 L^∞ AP scheme.

2.3.3 One dimensional linear Fourier stability analysis

In this section, we consider $d = 1$. We linearize the Euler system (2.1) around a constant solution $\overline{W} = (\overline{\rho}, \overline{q}, \overline{E})$ such that $\overline{\rho} > 0$ and $\overline{p} = \overline{E} - \varepsilon/2\,\overline{q}^2/\overline{\rho} > 0$. We denote by $\overline{u} = \overline{q}/\overline{\rho}$ and $\overline{c}^2 = \gamma\,\overline{p}/\overline{\rho}$. The linearized system is given by

$$\partial_t W + A\,\partial_x W = 0, \tag{2.44}$$

where $A = DF(\overline{W}) = DF_e(\overline{W}) + DF_i(\overline{W})$ with

$$\begin{aligned} DF_e(\overline{W}) &= \begin{pmatrix} 0 & 1 & 0 \\ -\overline{u}^2 & 2\,\overline{u} & 0 \\ -\varepsilon\,\overline{u}^3 & \frac{3}{2}\,\varepsilon\,\overline{u}^2 & 0 \end{pmatrix}, \\ DF_i(\overline{W}) &= \begin{pmatrix} 0 & 0 & 0 \\ \frac{\gamma-1}{2}\overline{u}^2 & (1-\gamma)u & \frac{\gamma-1}{\varepsilon} \\ -\frac{\overline{c}^2}{(\gamma-1)} + \gamma\varepsilon\frac{\overline{u}^3}{2} & \frac{\overline{c}^2}{\gamma-1} - \gamma\,\varepsilon\overline{u}^2 & \gamma\overline{u} \end{pmatrix}. \end{aligned} \tag{2.45}$$

The eigenvalues of A are

$$\overline{u} - \frac{\overline{c}}{\sqrt{\varepsilon}}, \qquad \overline{u}, \qquad \overline{u} + \frac{\overline{c}}{\sqrt{\varepsilon}}.$$

We denote by P, the matrix of the eigenvectors of A, then $P^{-1}AP = D$ where D is the diagonal matrix with the eigenvalues of A on the diagonal. The matrices P and P^{-1} are defined by

$$P = \begin{pmatrix} 1 & 1 & 1 \\ u - \dfrac{c}{\sqrt{\varepsilon}} & u & u + \dfrac{c}{\sqrt{\varepsilon}} \\ \frac{M\,c^2}{2}(M-2) + \dfrac{c^2}{\gamma-1} & \frac{M\,c^2}{2} & \frac{M\,c^2}{2}(M+2) + \dfrac{c^2}{\gamma-1} \end{pmatrix},$$

$$P^{-1} = \begin{pmatrix} \frac{M}{4}\left((\gamma-1)M+2\right) & -\dfrac{\sqrt{\varepsilon}}{2c}\left((\gamma-1)M+1\right) & \dfrac{\gamma-1}{2c^2} \\ (1-\gamma)\frac{M}{2}+1 & (\gamma-1)M^2 & -\dfrac{\gamma-1}{2c^2} \\ \frac{M}{4}\left((\gamma-1)M-2\right) & -\dfrac{\sqrt{\varepsilon}}{2c}\left((\gamma-1)M-1\right) & \dfrac{\gamma-1}{2c^2} \end{pmatrix},$$

with $M = \sqrt{\varepsilon}u/c$. We denote by V the coordinates of W in the eigenvectors basis, then $W = PV$ and $\partial_t V + D\,\partial_x V = 0$. Since D is diagonal,

$$\frac{1}{2}\partial_x(DV \cdot V) = \frac{1}{2}\left((D\partial_x V)\cdot V + (D^t\,\partial_x V)\cdot V\right) = (D\partial_x V)\cdot V.$$

Then, taking the scalar product of this equation with V, we obtain

$$\partial_t V \cdot V + (D\,\partial_x V)\cdot V = \partial_t \|V\|_2^2(x,t) + \partial_x(DV\cdot V) = 0,$$

where $\|\cdot\|_2$ is the Euclidean norm of $\mathbb{R}^3$. Integrating on the space domain and assuming periodic boundary conditions, we have

$$\partial_t \int_0^1 \|V\|_2^2(x,t) + (DV\cdot V)(1,t) - (DV\cdot V)(0,t) = \partial_t \|V\|_{L^2(]0,1[)}(t) = 0.$$

We recover for all $t > 0$,

$$\|V\|_{L^2(]0,1[)}(t) = \|V\|_{L^2(]0,1[)}(0).$$

Let us note that this result holds only in the eigenvector basis and is not true for W since the matrix A is not symmetric.
Following this proof, using Fourier analysis, proving the decrease of the L^2 norm of the discretized vector V, we prove the L^2 stability of our Order 1 L^2 AP (2.41) and Order 1 L^∞ AP (2.42) schemes. We conclude this section with the same analysis for the nonlinear AP scheme NL Order 1 L^2 AP scheme (2.43).
To prove the L^2 stability of the scheme, we prove the decrease of the L^2 norm of the discretized vector V.

2.3.3.1 Linearization of our semi-discretization

We begin, linearizing the semi-discretized system (2.38) around $\overline{W} = (\overline{\rho}, \overline{q}, \overline{E})$ a constant solution. Then, we set $W^k = \overline{W} + \epsilon\,\check{W}^k$ for $k = n+1, n$, "$n+1, exp$" and $p^{n+1} = \overline{p} + \epsilon\,\check{p}^{n+1}$. Using a Taylor expansion we get

$$\overline{W} + \epsilon\,\check{W}^{n+1,exp} = \overline{W} + \epsilon\,\check{W}^{n} - \Delta t\,\partial_x\Big(F_e(\overline{W}) + \epsilon\,DF_e(\overline{W})\,\check{W}^{n} + o(\epsilon)\Big).$$

Neglecting all terms of order greater than ϵ, an omitting the omitting the "checks" we obtain

$$W^{n+1,exp} = W^{n} - \Delta t\,DF_e(\overline{W})\,\partial_x W^{n},$$

and so

$$\begin{aligned}
\rho^{n+1,exp} &= \rho^{n} - \Delta t\,\partial_x q^{n},\\
q^{n+1,exp} &= q^{n} + \Delta t\,\overline{u}^2\partial_x\rho^{n} - 2\,\Delta t\,\overline{u}\,\partial_x q^{n},\\
E^{n+1,exp} &= E^{n} + \Delta t\,\varepsilon\,\overline{u}^3\,\partial_x\rho^{n} - \Delta t\,\frac{3\,\varepsilon}{2}\,\overline{u}^2\,\partial_x q^{n}.
\end{aligned} \tag{2.46}$$

Now the pressure equation gives

$$\begin{aligned}
\frac{\varepsilon}{\gamma-1}(\overline{p} + \epsilon\,\check{p}^{n+1}) - \epsilon\,\Delta t^2\partial_x\,\Big(\Big(H(\overline{W}) + o(\epsilon)\Big)\partial_x\check{p}^{n+1}\Big) =\\
\varepsilon\,\Big(\overline{E} + \epsilon\,\check{E}^{n+1,exp} - k(\overline{W}) - \epsilon\,Dk(\overline{W})\,\check{W}^{n+1,exp} + o(\epsilon)\Big)\\
-\varepsilon\,\Delta t\,\partial_x\,\Big(\Big(H(\overline{W}) + \epsilon\,DH(\overline{W})\,\check{W}^{n+1,exp}\Big)\,(\overline{q} + \epsilon\,\check{q}^{n+1,exp})\,\Big),
\end{aligned}$$

where

$$H(\overline{W}) = \gamma\,\frac{\overline{E}}{\overline{\rho}} - \frac{\gamma\,\varepsilon}{2}\,\frac{\overline{q}^2}{\overline{\rho}^2} = \frac{1}{\gamma-1}\,\frac{\gamma\,\overline{p}}{\overline{\rho}} = \frac{\overline{c}^2}{\gamma-1},$$

$$\begin{aligned}
DH(\overline{W}) = \Big(-\gamma\,\frac{\overline{E}}{\overline{\rho}^2} + \gamma\,\varepsilon\,\frac{\overline{u}^2}{\overline{\rho}}, -\gamma\,\varepsilon\,\frac{\overline{u}}{\overline{\rho}}, \frac{\gamma}{\overline{\rho}}\Big) &= \Big(-\frac{\gamma\,\overline{p}}{(\gamma-1)\,\overline{\rho}^2} + \gamma\,\varepsilon\,\frac{\overline{u}^2}{2\,\overline{\rho}}, -\gamma\,\varepsilon\,\frac{\overline{u}}{\overline{\rho}}, \frac{\gamma}{\overline{\rho}}\Big)\\
&= \Big(-\frac{\overline{c}^2}{(\gamma-1)\,\overline{\rho}} + \gamma\,\varepsilon\,\frac{\overline{u}^2}{2\,\overline{\rho}}, -\gamma\,\varepsilon\,\frac{\overline{u}}{\overline{\rho}}, \frac{\gamma}{\overline{\rho}}\Big).
\end{aligned}$$

$$Dk(\overline{W}) = (-\varepsilon/2\,\overline{u}^2, \varepsilon\,\overline{u}, 0),$$

Neglecting the error terms and omitting the "checks", we obtain

$$\begin{aligned}
\frac{\varepsilon}{\gamma-1}p^{n+1} - \frac{\overline{c}^2}{\gamma-1}\Delta t^2\partial^2_{xx}p^{n+1} &= \varepsilon\,\Big(E^{n+1,exp} + \frac{\varepsilon\,\overline{u}^2}{2}\,\rho^{n+1} - \varepsilon\,\overline{u}\,q^{n+1,exp}\Big)\\
&\quad - \varepsilon\,\Delta t\,\Big(\frac{\overline{c}^2}{\gamma-1}\,\partial_x q^{n+1,exp} + \overline{q}\,DH(\overline{W})\,\partial_x W^{n+1,exp}\Big)\\
&= \varepsilon\,\Big(E^{n+1,exp} + \frac{\varepsilon\,\overline{u}^2}{2}\,\rho^{n+1} - \varepsilon\,\overline{u}\,q^{n+1,exp}\Big) - \varepsilon\,\Delta t\,\frac{\overline{c}^2}{\gamma-1}\,\partial_x q^{n+1,exp}\\
&\quad - \varepsilon\,\Delta t\,\Big(-\frac{c^2\,\overline{u}}{\gamma-1} + \gamma\,\varepsilon\,\frac{\overline{u}^3}{2}\Big)\,\partial_x\rho^{n+1} + \varepsilon\,\Delta t\,\gamma\,\varepsilon\,\overline{u}^2\,\partial_x q^{n+1,exp} - \gamma\,\overline{u}\,\partial_x E^{n+1,exp},
\end{aligned}$$

And so,

$$\begin{aligned}
p^{n+1} - \frac{\overline{c}^2}{\varepsilon}\Delta t^2\partial^2_{xx}p^{n+1} &= (\gamma-1)\,\Big(E^{n+1,exp} + \frac{\varepsilon\,\overline{u}^2}{2}\,\rho^{n+1} - \varepsilon\,\overline{u}\,q^{n+1,exp}\Big)\\
&\quad - \Delta t\,\Big(-\overline{c}^2\,\overline{u} + \gamma\,(\gamma-1)\,\frac{\overline{u}^3}{2}\Big)\,\partial_x\rho^{n+1} - \Delta t\,\left(\overline{c}^2 - \gamma\,(\gamma-1)\,\varepsilon\,\overline{u}^2\right)\,\partial_x q^{n+1,exp}\\
&\quad - \Delta t\gamma\,(\gamma-1)\,\overline{u}\,\partial_x E^{n+1,exp},
\end{aligned}$$

Finally, for the energy equation, we obtain

$$E^{n+1} = E^{n+1,exp} + \Delta t \, \frac{\overline{c}^2}{(\gamma-1)\,\overline{\rho}} \, \partial_x \rho^{n+1} - \Delta t \, \frac{\overline{c}^2}{(\gamma-1)} \, \partial_x q^{n+1} - \Delta t \, \frac{\gamma\,\overline{u}}{\gamma-1} \, \partial_x p^{n+1}.$$

The resulting one dimensional linearized semi-discretized system reads:

$$W^{n+1,exp} = W^n - \Delta t \, DF_e(\overline{W}) \, \partial_x W^n, \tag{2.47a}$$

$$\rho^{n+1} = \rho^{n+1,exp}, \tag{2.47b}$$

$$\begin{aligned} p^{n+1} - \Delta t^2 \, \frac{c^2}{\varepsilon} \, \partial^2_{xx} p^{n+1} &= (\gamma-1)\left(E^{n+1,exp} + \varepsilon\,\overline{u}^2/2\,\rho^{n+1} - \varepsilon\,\overline{u}\,q^{n+1,exp}\right) \\ &\quad - \Delta t\left[\left(-c^2\,\overline{u} + \gamma\,(\gamma-1)/2\,\varepsilon\,\overline{u}^3\right)\,\partial_x\rho^{n+1} + \left(c^2 - \gamma\,(\gamma-1)\,\varepsilon\,\overline{u}^2\right)\,\partial_x q^{n+1,exp}\right. \\ &\qquad \left. + \gamma\,(\gamma-1)\,\overline{u}\,\partial_x E^{n+1,exp}\right], \end{aligned} \tag{2.47c}$$

$$q^{n+1} = q^{n+1,exp} - \frac{\Delta t}{\varepsilon} \, \partial_x p^{n+1}, \tag{2.47d}$$

$$E^{n+1} = E^{n+1,exp} + \Delta t \, \frac{c^2\,\overline{u}}{\gamma-1} \, \partial_x \rho^{n+1} - \Delta t \, \frac{c^2}{\gamma-1} \, \partial_x q^{n+1} - \Delta t \, \frac{\gamma\,\overline{u}}{\gamma-1} \, \partial_x p^{n+1}. \tag{2.47e}$$

2.3.3.2 L^2 Stability of our Order 1 L^2 AP scheme

Discretizing the linearized semi-discretized system (2.47) with the Order 1 L^2 AP scheme, we obtain:

$$W_j^{n+1,exp} = W_j^n - \frac{\Delta t}{\Delta x}\left(DF_e(\overline{W})\,\frac{W_{j+1}^n - W_{j-1}^n}{2} - \frac{\overline{u}}{2}\,(W_{j+1}^n - 2\,W_j^n + W_{j-1}^n)\right), \tag{2.48a}$$

$$\rho_j^{n+1} = \rho_j^{n+1,exp}, \tag{2.48b}$$

$$\begin{aligned} &p_j^{n+1} - \frac{c^2\,\Delta t^2}{\varepsilon\,\Delta x^2}(p_{j+1}^{n+1} - 2\,p_j^{n+1} + p_{j-1}^{n+1}) = (\gamma-1)\,E_j^{n+1,exp} \\ &+ (\gamma-1)\,\varepsilon\left(\overline{u}^2/2\,\rho_j^{n+1} - \varepsilon\,\overline{u}\,q_j^{n+1,exp}\right) - \frac{\Delta t}{\Delta x}\,\left(-c^2\,\overline{u} + \gamma\,(\gamma-1)/2\,\varepsilon\,\overline{u}^3\right)\,\frac{\rho_{j+1}^{n+1} - \rho_{j-1}^{n+1}}{2} \\ &- \frac{\Delta t}{\Delta x}\left(c^2 - \gamma\,(\gamma-1)\,\varepsilon\,\overline{u}^2\right)\,\frac{q_{j+1}^{n+1,exp} - q_{j-1}^{n+1,exp}}{2} - \frac{\Delta t}{\Delta x}\,\gamma\,(\gamma-1)\,\overline{u}\,\frac{E_{j+1}^{n+1,exp} - E_{j-1}^{n+1,exp}}{2}, \end{aligned} \tag{2.48c}$$

$$q_j^{n+1} = q_j^{n+1,exp} - \frac{\Delta t}{\varepsilon\,\Delta x}\,\frac{p_{j+1}^{n+1} - p_{j-1}^{n+1}}{2}, \tag{2.48d}$$

$$E_j^{n+1} = E_j^{n+1,exp} + \frac{\Delta t}{(\gamma-1)\,\Delta x}\left(\overline{c}^2\,\overline{u}\frac{\rho_{j+1}^{n+1} - \rho_{j-1}^{n+1}}{2} - \overline{c}^2\,\frac{q_{j+1}^{n+1} - q_{j-1}^{n+1}}{2} - \gamma\,\overline{u}\,\frac{p_{j+1}^{n+1} - p_{j-1}^{n+1}}{2}\right). \tag{2.48e}$$

where F_e is given by (2.24). We assume periodic boundary conditions. We prove the following result

Lemma 2.3.2 (L^2 Stability of our Order 1 L^2 AP scheme)**.** *Let $W_0 \in L^2(]0,1[)$ and $\Delta x > 0$ and $\Delta t > 0$ the space and time steps satisfying the C.F.L. condition*

$$\gamma\,|\overline{u}|\,\Delta t = \Delta x.$$

We denote by $W_j^0 = 1/\Delta x \int_{(j-1)\,\Delta x}^{j\,\Delta x} W_0(x)\,dx$, by (W_j^n) the solution of (2.48) and $P\,V^n(x) = W^n(x) = W_j^n = P\,V_j^n$ if $x \in](j-1)\,\Delta x, j\,\Delta x[$ where P is the matrix of the eigenvectors of $DF(\overline{W})$. We set $M_\varepsilon = \sqrt{\varepsilon}\,\overline{u}/\overline{c}$.
Then, for all $M_\varepsilon \in]0,25[$ and all $\gamma \in [1,10]$, there exists $C > 0$ depending on γ, $\overline{u}$, $\overline{c}$ and ε such that for all $n \geq 0$

$$\|V^n\|_{L^2(]0,1[)} \leq C\,\|V^0\|_{L^2(]0,1[)}.$$

Proof. We assume $\overline{u} > 0$ for clarity. The same proof can be done for $\overline{u} < 0$.
For $l = n+1$, $n+1, exp$ or n, we define on $[0,1[$ $W^l(x) = W_j^l$ if $x \in [(j-1)\,\Delta x, j\,\Delta x[$ for $j = 1, \cdots, L = 1/\Delta x$. Then,

$$W^l(x) = \sum_{k\in\mathbb{Z}} \hat{W}^l(k) e^{2\,i\pi\,k\,x} \quad \text{with } \hat{W}^l(k) = \int_0^1 W^l(x)\,e^{-2\,i\pi\,k\,x}\,dx,$$

and $W^l = \sum_{k\in\mathbb{Z}} \hat{W}^l(k)$. Then, (2.48a) reads

$$W^{n+1,exp}(x) = W^n(x) - \alpha\,DF_e\,\frac{W^n(x+\Delta x) - W^n(x-\Delta x)}{2} + \alpha\,\frac{\overline{u}}{2}\,(W^n(x+\Delta x) - 2\,W^n(x) + W^n(x-\Delta x)),$$

where $\alpha = \frac{\Delta t}{\Delta x}$ and DF_e is given by (2.45).
Taking the Fourier transform of the resulting equation, we get

$$\hat{W}^{n+1,exp}(k) = B_w^e(k)\hat{W}^n(k),$$

where

$$B_w^e(k) = (1 - \alpha\,\overline{u}\,(1-\cos\varphi))\,I_3 - i\,\alpha\,\sin\varphi\,DF_e,$$

with $\varphi = 2\pi\,k\,\Delta x$.
Now, we can move into the eigenvector basis. Denoting by $V^k = P^{-1}\,W^k$ for $k = n$, "$n+1, exp$" or '$n+1$ yields $\hat{V}^k = P^{-1}\,\hat{W}^k$. Multiplying $\hat{W}^{n+1,exp}(k)$ by P^{-1}, we obtain

$$P^{-1}\hat{W}^{n+1,exp}(k) = \hat{V}^{n+1,exp}(k) = P^{-1}\,B_w^e(k)\,P\,\hat{V}^n(k) = B^e(k)\,\hat{V}^n(k),$$

where

$$B^e(k) = P^{-1}\,B_w^e(k)\,P = (1 - \alpha\,\overline{u}\,(1-\cos\varphi))\,I_3 - i\,\alpha\,\sin\varphi\,A_e,$$

with

$$A_e = P^{-1} DF_e P = \begin{pmatrix} \frac{\overline{u}}{2} & 0 & -\frac{\overline{u}}{2} \\ \overline{u} - \frac{\overline{c}}{\sqrt{\varepsilon}} & \overline{u} & \overline{u} + \frac{\overline{c}}{\sqrt{\varepsilon}} \\ -\frac{\overline{u}}{2} & 0 & \frac{\overline{u}}{2} \end{pmatrix}. \tag{2.49}$$

The Eigenvalues of $B^e(k)$ are given by

$$\begin{aligned} \lambda_1 &= 1 - \alpha\,\overline{u}\,(1-\cos\varphi), \\ \lambda_2 &= \lambda_3 = 1 - \alpha\,\overline{u}\,(1-\cos\varphi) - i\,\alpha\,\overline{u}\,\sin\varphi. \end{aligned}$$

We have,

$$\begin{aligned} |\lambda_1|^2 &= 1 - \alpha\,\overline{u}\,(1-\cos\varphi)\,(2 - \alpha\,\overline{u}(1-\cos\varphi)), \\ |\lambda_{2,3}|^2 &= 1 - 2\,\alpha\,\overline{u}\,(1-\cos\varphi)\,(1-\alpha\,\overline{u}). \end{aligned}$$

In order to have the spectral radius of $B^e(k)$, noted $r(B^e(k))$, lower than one, we need

$$\alpha\,\overline{u}\,(1-\cos\varphi)\,(2 - \alpha\,\overline{u}(1-\cos\varphi)) \geq 0 \;\Leftrightarrow\; (2 - \alpha\,\overline{u}(1-\cos\varphi)) \geq 0,$$

and,

$$2\,\alpha\,\overline{u}\,(1-\cos\varphi)\,(1-\alpha\,\overline{u}) \geq 0 \;\Leftrightarrow\; (1-\alpha\,\overline{u}) \geq 0,$$

where $0 \leq (1-\cos\varphi) \leq 2$.
Then, $r(B^e(k))$ is lower than 1 under the C.F.L. condition $\overline{u}\alpha \leq 1$, that is

$$\overline{u}\,\Delta t \leq \Delta x.$$

Note that $\hat{\rho}^{n+1} = \hat{\rho}^{n+1,exp}$. Now, the Fourier transforms of Eqs. (2.48c), (2.48d) and (2.48e) give

$$\begin{aligned} &(1 + 2\,\alpha^2\,\frac{\overline{c}^2}{\varepsilon}\,(1-\cos\varphi))\,\hat{p}^{n+1}(k) \\ &\quad = \left((\gamma-1)\,\frac{\varepsilon\,\overline{u}^2}{2} - i\,\alpha\,\overline{u}\sin\varphi\left(-\overline{c}^2 + \frac{\gamma\,(\gamma-1)\varepsilon\,\overline{u}^2}{2}\right)\right)\hat{\rho}^{n+1,exp}(k) \\ &\qquad\quad - \left((\gamma-1)\,\varepsilon\,\overline{u} + i\,\alpha\,\sin\varphi\,\left(\overline{c}^2 - \gamma\,(\gamma-1)\,\varepsilon\,\overline{u}^2\right)\right)\hat{q}^{n+1,exp}(k) \\ &\qquad\quad + (\gamma-1)\,(1 - i\,\alpha\,\overline{u}\,\sin\varphi\,\gamma)\,\hat{E}^{n+1,exp}(k), \\ &\hat{q}^{n+1}(k) = \hat{q}^{n+1,exp}(k) - \frac{\alpha}{\varepsilon}\,i\sin\varphi\,\hat{p}^{n+1}(k), \\ &\hat{E}^{n+1}(k) = \hat{E}^{n+1,exp}(k) + \alpha\,i\sin\varphi\,\left(\frac{c^2\,\overline{u}}{\gamma-1}\,\hat{\rho}^{n+1} - \frac{\overline{c}^2}{\gamma-1}\,\hat{q}^{n+1} - \frac{\gamma\,\overline{u}}{\gamma-1}\,\hat{p}^{n+1}(k)\right), \\ &\quad = \hat{E}^{n+1,exp}(k) + \alpha\,i\sin\varphi\,\frac{c^2\,\overline{u}}{\gamma-1}\,\hat{\rho}^{n+1} \\ &\qquad\quad - \alpha i\sin\varphi\,\left(\frac{c^2}{\gamma-1}\,q^{n+1,exp}(k) - \left(-\frac{\overline{c}^2\alpha\,i\sin\varphi}{\varepsilon\,(\gamma-1)} + \frac{\gamma\,\overline{u}}{\gamma-1}\right)\,\hat{p}^{n+1}(k)\right). \end{aligned}$$

We note that $\hat{p}^{n+1}(k)$ given by the first equation only depends on the variables $\hat{\rho}^{n+1,exp}(k)$, $\hat{q}^{n+1,exp}(k)$ and $\hat{E}^{n+1,exp}(k)$. So inserting the expression of $\hat{p}^{n+1}(k)$ into the equations for $\hat{q}^{n+1}(k)$ and $\hat{E}^{n+1}(k)$, yields

$$\hat{W}^{n+1}(k) = \begin{pmatrix} \hat{\rho}^{n+1}(k) \\ \hat{q}^{n+1}(k) \\ \hat{E}^{n+1}(k) \end{pmatrix} = B_w^i(k) \begin{pmatrix} \hat{\rho}^{n+1,exp}(k) \\ \hat{q}^{n+1,exp}(k) \\ \hat{E}^{n+1,exp}(k) \end{pmatrix} = B_w^i(k)\, B_w^e(k)\, \hat{W}^n(k). \quad (2.50)$$

where $B_w^i(k)$ is of the form $B_w^i(k) = \begin{pmatrix} 1 & 0 & 0 \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{pmatrix}$.
By multipliying (2.50) by P^{-1}, we move into the eigenvector basis and have

$$\begin{aligned} \hat{V}^{n+1}(k) &= P^{-1} B_w^i(k) \hat{W}^{n+1,exp}(k) \\ &= P^{-1} B_w^i(k) P\, \hat{V}^{n+1,exp}(k) \\ &= B^i(k)\, B^e(k)\, \hat{V}^n(k), \end{aligned} \quad (2.51)$$

where we used $\hat{V}^{n+1,exp}(k) = B^e(k)\, \hat{V}^n(k)$ and we noted $B^i(k) = P^{-1} B_w^i(k) P$.
Let us note that, we also have $B^i(k)B^e(k) = P^{-1}B_w^i(k)P\, P^{-1}B_w^e(k)P = P^{-1}B_w^i(k)\, B_w^e(k)P$.
Thanks to the software Maple (we refer to Appendix A for all the calculations obtained with the software), we see that the resulting matrix $B^i(k)B^e(k)$ is of the form

$$B^i(k)B^e(k) = \begin{pmatrix} d_{11} & 0 & d_{13} \\ d_{21} & 1 + \alpha\overline{u}(\cos\varphi - 1) - i\,\alpha\overline{u}\,\sin\varphi & d_{23} \\ d_{31} & 0 & d_{33} \end{pmatrix}.$$

Therefore, the characteristic polynomial of $B^i(k)B^e(k)$ reads

$$det(B^i(k)\, B^e(k) - \lambda\, I_3) = (1 + \alpha\overline{u}(\cos\varphi - 1) - i\,\alpha\overline{u}\,\sin\varphi\ - \lambda) \begin{vmatrix} d_{11} - \lambda & d_{13} \\ d_{31} & d_{33} - \lambda \end{vmatrix}.$$

One eigenvalue is directly given by

$$\lambda_1 = 1 - \alpha\overline{u}\,(1 - \cos\varphi + i\sin\varphi)\,,$$

and we rewrite

$$det(B^i(k)\, B^e(k) - \lambda\, I_3) = (1 - \alpha\overline{u}\,(1 - \cos\varphi + i\sin\varphi) - \lambda)\ Q^i(\lambda),$$

where $Q^i(\lambda)$ is a second order polynomial given by

$$Q^i(\lambda) = \lambda^2 + a_i\,\lambda + b_i,$$

with a_i and b_i complex coefficients. We refer to equation (51) in Appendix A for there explicit expressions obtained with the software Maple. There expression is given in terms of four variables:

$$cfl = \alpha\,\overline{u}, \quad M = M_\varepsilon = \sqrt{\varepsilon}\frac{\overline{u}}{\overline{c}}, \quad \gamma, \quad \varphi = 2\pi k\Delta x.$$

Then, setting $cfl = 1/\gamma$, such that,

$$\gamma\,\alpha\,\overline{u} = 1,$$

the coefficients a_i and b_i are expressed in terms of M_ε, γ and φ. We denote by

$$\lambda_2(M_\varepsilon, \varphi, \gamma), \quad \lambda_3(M_\varepsilon, \varphi, \gamma),$$

the roots of Q^i. Hence, the eigenvalues of $B^i(k)\,B^e(k)$ are $\lambda_1 = 1 - \alpha\overline{u}(1 - \cos\varphi + i\,\sin\varphi)$, λ_2 and λ_3. We always have $|\lambda_1| \leq 1$ since, as shown for the eigenvalues of A_e, $\alpha\overline{u} \leq 1$ implies $|\lambda_1| \leq 1$ and here we have $\alpha\overline{u} = 1/\gamma \leq 1$.
We plot on Figure 2.6, the maximum modulus of the roots of Q^i as a function of $M_\varepsilon \in]0, 25]$ and $\gamma \in [1, 10]$

$$f(M_\varepsilon, \gamma) := \max(\max_{\varphi\in[0,2\,\pi[}|\lambda_2(M_\varepsilon, \varphi, \gamma)|, \max_{\varphi\in[0,2\,\pi[}|\lambda_3(M_\varepsilon, \varphi, \gamma)|).$$

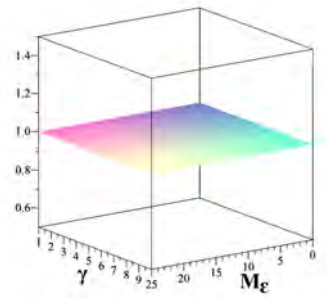


Figure 2.6 – Stability of our Order 1 L^2 AP scheme (Lemma 2.3.2): Maximum modulus of the roots of Q^i.

As we can see on the figure, the maximum value of $|\lambda_2|$ and $|\lambda_3|$ is one. This proves that, under the C.F.L. condition $\gamma\overline{u}\Delta t \leq \Delta x$, for all $M_\varepsilon \in]0, 25]$, all $\varphi \in [0, 2\,\pi]$ and all $\gamma \in [1, 10]$, the spectral radius of $B_i(k)\,B_e(k)$, denoted by $r^i(B_i(k)\,B_e(k))$, is lower than 1.
Since there exists at least one norm matrix $\|\cdot\|$ in $\mathbb{R}^n$ (depending on $B_i(k)\,B_e(k)$), such that $\|B_i(k)\,B_e(k)\| \leq r(B_i(k)\,B_e(k))$, we obtain

$$\|\hat{V}^{n+1}(k)\| \leq \|B_i(k)\,B_e(k)\|\|\hat{V}^n(k)\| \leq \|\hat{V}^n(k)\| \leq \cdots \leq \|\hat{V}^0(k)\|.$$

But, all norms are equivalent in finite dimension then, there exists $C_1 > 0$ and $C_2 > 0$ depending on $\overline{u}$, $\overline{c}$ and ε such that for all $n \geq 0$,

$$C_1\|\hat{V}^n(k)\|_2 \leq \|\hat{V}^n(k)\| \leq C_2\,\|\hat{V}^n(k)\|_2.$$

where $\|\cdot\|_2$ is the Euclidean norm. And so,

$$C_1\|\hat{V}^n(k)\|_2 \leq \|\hat{V}^n(k)\| \leq \|\hat{V}^0(k)\| \leq C_2\,\|\hat{V}^0(k)\|_2,$$

Furthermore, using Plancherel's theorem, we have that

$$\|V^n\|^2_{L^2(]0,1[)} = \sum_{k\in\mathbb{Z}} |\hat{V}^n(k)|^2 = \sum_{k\in\mathbb{Z}} \|\hat{V}^n(k)\|_2^2.$$

Therefore, noting $C = C_2/C_1 > 0$ we prove that for all $n > 0$

$$\|V^n\|_{L^2(]0,1[)} \leq C\|V^0\|_{L^2(]0,1[)}.$$

□

Remark 4. Note that, this C.F.L. seems to be optimal, indeed for larger values of the C.F.L, i.e $|\overline{u}|\,\Delta t/\Delta x = C'$ with $C' > 1/\gamma$, there exists γ such that the spectral radius is bigger than 1. In Figure 2.7 we show that for $C' = 1$, the maximum modulus of the roots of Q^i are greater than 1 from approximately $\gamma = 2.5$. However, for $\gamma < 2.5$, it is sufficient to assume $|\overline{u}|\,\Delta t/\Delta x = 1$.

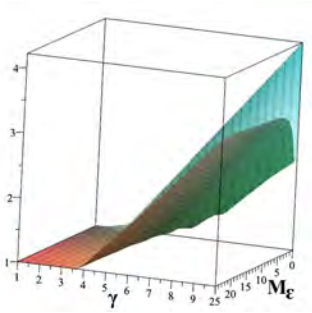


Figure 2.7 – Maximum modulus of the roots of Q^i for $C' = 1 > 1/\gamma$, i.e., under the C.F.L. condition $|\overline{u}|\,\Delta t = \Delta x$ instead of $|\overline{u}|\,\Delta t = \Delta x/\gamma$.

Remark 5. Also note that, for higher values of γ, this condition is not sufficient. Indeed, in Figure 2.8), we plotted under the C.F.L. condition $|\overline{u}|\,\Delta t = \Delta x/\gamma$, the maximum modulus of the roots of Q^i for all $M_\varepsilon \in]0, 25]$, all $\varphi \in [0, 2\,\pi]$ and all $\gamma \in [1, 15]$. The maximum modulus of the roots of Q^i are greater than 1 from approximately $\gamma = 11$.

2.3.3.3 L^2 Stability of our Order 1 L^∞ AP scheme

Let us now perform the stability analysis of the Order 1 L^∞ AP scheme. The Order 1 L^∞ AP scheme on the linearized system consists in the following

$$W_j^{n+1} - \frac{|\lambda_i|\,\Delta t}{2\,\Delta x}\left(W_{j+1}^{n+1} - 2\,W_j^{n+1} + W_{j-1}^{n+1}\right) = W_j^{n+1,L2}, \tag{2.52}$$

where $W_j^{n+1,L2}$ is given by the Order 1 L^2 AP scheme (2.48) and where λ_i is the maximum implicit eigenvalue associated to the implicit flux F_i

$$|\lambda_i| = \frac{|\overline{u}|}{2} + \sqrt{\frac{\overline{u}^2}{4} + \frac{\overline{c}^2}{\varepsilon}} = |\overline{u}|\left(\frac{1}{2} + \sqrt{\frac{1}{4} + \frac{1}{M_\varepsilon^2}}\right),$$

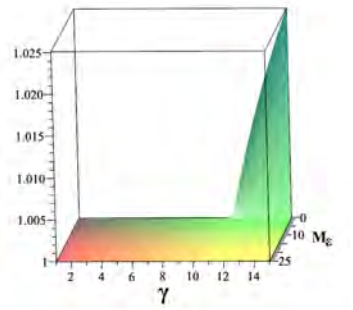


Figure 2.8 – Maximum modulus of the roots of Q^i for higher values of γ under the C.F.L. condition $|\overline{u}|\,\Delta t = \Delta x/\gamma$.

with $M_\varepsilon = \sqrt{\varepsilon}\overline{u}/\overline{c}$.
We assume periodic boundary conditions and prove the following result:

Lemma 2.3.3 (L^2 Stability of our Order 1 L^∞ AP scheme)**.** *Let $W_0 \in L^2(]0,1[)$ and $\Delta x > 0$ and $\Delta t > 0$ the space and time steps satisfying the C.F.L. condition*

$$|\overline{u}|\,\Delta t = \Delta x.$$

We denote by $W_j^0 = 1/\Delta x \int_{(j-1)\,\Delta x}^{j\,\Delta x} W_0(x)\,dx$, by (W_j^n) the solution of (2.52) and $P\,V^n(x) = W^n(x) = W_j^n = P\,V_j^n$ if $x \in](j-1)\,\Delta x, j\,\Delta x[$ where P is the matrix of the eigenvectors of $DF(\overline{W})$. We set $M_\varepsilon = \sqrt{\varepsilon}\,\overline{u}/\overline{c}$.
Then, for all $M_\varepsilon \in]0,25[$ and all $\gamma \in [1,5]$, there exists $C > 0$ depending on γ, $\overline{u}$, $\overline{c}$ and ε such that for all $n \geq 0$

$$\|V^n\|_{L^2(]0,1[)} \leq C\,\|V^0\|_{L^2(]0,1[)}.$$

Proof. We assume $\overline{u} > 0$ for clarity. The same proof can be done for $\overline{u} < 0$. We proceed like in the proof of the previous Lemma. We take the Fourier transform of (2.52). Noting $\varphi = 2\pi k \Delta x$, we obtain

$$\begin{aligned}(1 + |\lambda_i|\,\alpha)\hat{W}_j^{n+1}(k) - \frac{|\lambda_i|}{2}\,\alpha\,e^{i\varphi}\,\hat{W}_j^{n+1}(k) - \frac{|\lambda_i|}{2}\,\alpha\,e^{-i\varphi}\,\hat{W}_j^{n+1}(k)\\ = (1 + |\lambda_i|\,(1-\cos\phi)\,)\,\hat{W}_j^{n+1}(k)\\ = \left(1 + \alpha\,|\overline{u}|\left(1/2 + \sqrt{1/4 + 1/M_\varepsilon^2}\right)\,\alpha\,(1-\cos\phi)\right)\,\hat{W}_j^{n+1}(k) = \hat{W}_j^{n+1,L^2}(k).\end{aligned}$$

By multiplying the resulting equation by P^{-1}, we pass into the eigenvector basis of DF and using the coordinates of W (2.51), we obtain

$$P^{-1}\,\beta\,P\,\hat{V}^{n+1}(k) = \beta\,\hat{V}^{n+1}(k) = B^i(k)\,B^e(k)\,\hat{V}^n(k),$$

where $\beta = \left(1 + \alpha\,\overline{u}\,\left(1/2 + \sqrt{1/4 + 1/M_\varepsilon^2}\right)\,(1-\cos\varphi)\right)$.

Then, using the results of the proof of the previous Lemma, we obtain

$$\begin{aligned} det\left(\frac{1}{\beta}\,B^i(k)\,B^e(k)-\lambda\,I_3\right) &= \frac{1}{\beta^3}\,det\left(B^i(k)\,B^e(k)-\beta\,\lambda\,I_3\right)\\ &= \frac{1}{\beta^3}\left(1-\alpha\overline{u}\,(1-\cos\varphi+i\sin\varphi)-\beta\lambda\right)\,Q^i(\beta\lambda)\\ &= \left(\frac{1-\alpha\overline{u}\,(1-\cos\varphi+i\sin\varphi)}{\beta}-\lambda\right)\,\overline{Q}^i(\lambda), \end{aligned}$$

where $\overline{Q}^i(\lambda)=\frac{Q^i(\beta\,\lambda)}{\beta^2}$.
We set

$$\overline{u}\,\Delta t=\Delta x,$$

and we denote by $\overline{\lambda}_2(M_\varepsilon,\varphi,\gamma)$ and $\overline{\lambda}_3(M_\varepsilon,\varphi,\gamma)$ the roots of $\overline{Q}^i$. As for the previous proof, we plot on Figure 2.9 the maximum modulus of the roots of $\overline{Q}^i$ that is

$$\overline{f}(M_\varepsilon,\gamma):=\max(\max_{\varphi\in[0,2\,\pi[}|\overline{\lambda}_2(M_\varepsilon,\varphi,\gamma)|,\max_{\varphi\in[0,2\,\pi[}|\overline{\lambda}_3(M_\varepsilon,\varphi,\gamma)|),$$

as a function of $M_\varepsilon\in]0,25]$ and $\gamma\in[1,5]$.

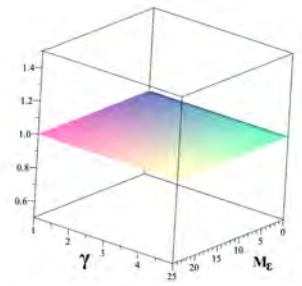


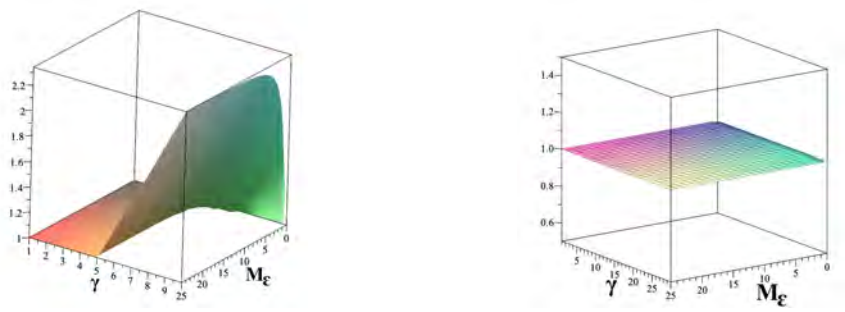
Figure 2.9 – Stability of our Order 1 L^∞ AP scheme (Lemma 2.3.3): Maximum modulus of the roots of $\overline{Q}^i$.

This proves that under the C.F.L. condition $\overline{u}\Delta t\leq\Delta x$, for all $M_\varepsilon\in]0,25]$, all $\varphi\in[0,2\,\pi]$ and all $\gamma\in[1,5]$, the spectral radius of $1/\beta\,B_i(k)\,B_e(k)$ is lower than 1. We conclude like in the proof of Lemma 2.3.2. □

Remark 6. Note that for higher values of γ, this condition is not sufficient (see Figure 2.10). Nevertheless, stability can be recovered under the C.F.L. condition $1.7\gamma\overline{u}\Delta t\leq\Delta x$, for all $M_\varepsilon\in]0,25]$, all $\varphi\in[0,2\,\pi]$ and all $\gamma\in[1,25]$.

2.3.3.4 L^2 Stability of the NL Order 1 L^2 AP scheme ([11])

We conclude with the linear stability analysis of the nonlinear scheme, we prove that the NL Order 1 L^2 AP scheme is L^2 stable. The scheme on the linearized system



(a) C.F.L. condition $|\overline{u}|\,\Delta t = \Delta x$ (b) C.F.L. condition $1.7\gamma|\overline{u}|\,\Delta t = \Delta x$

Figure 2.10 – Maximum modulus of the roots of $\overline{Q}^i$ for higher values of γ.

with periodic boundary conditions, is given by

$$\frac{W_j^{n+1}-W_j^n}{\Delta t}+DF_e(\overline{W})\,\frac{W_{j+1}^n-W_{j-1}^n}{2\,\Delta x}-\frac{|\overline{u}|}{2\,\Delta x}\left(W_{j+1}^n-2\,W_j^n+W_{j-1}^n\right)+DF_i(\overline{W})\,\frac{W_{j+1}^{n+1}-W_{j-1}^{n+1}}{2\,\Delta x}=0, \quad (2.53)$$

for $j=1,\cdots L=1/\Delta x$ with $W_0^n=W_L^n$ and $W_{L+1}^n=W_1^n$.

Lemma 2.3.4 (Stability of the NL Order 1 L^2 AP scheme)**.** *Let $W_0\in L^2(]0,1[)$ and $\Delta x>0$ and $\Delta t>0$ the space and time steps satisfying the C.F.L. condition*

$$|\overline{u}|\,\Delta t=\Delta x.$$

We denote by $W_j^0=1/\Delta x\int_{(j-1)\,\Delta x}^{j\,\Delta x}W_0(x)\,dx$ and by (W_j^n) the solution of (2.53) and $P\,V^n(x)=W^n(x)=W_j^n=P\,V_j^n$ if $x\in](j-1)\,\Delta x,j\,\Delta x[$ where P is the matrix of the eigenvectors of $DF(\overline{W})$. We set $M_\varepsilon=\sqrt{\varepsilon}\,\overline{u}/\overline{c}$.
Then, for all $M_\varepsilon\in]0,25]$, there exists $C>0$ depending on $\overline{u}$, $\overline{c}$ and ε such that for all $n\geq 0$

$$\|V^n\|_{L^2(]0,1[)}\leq C\,\|V^0\|_{L^2(]0,1[)}.$$

Proof. We assume $\overline{u}>0$ for clarity. The same proof can be done for $\overline{u}<0$.
We proceed like in the proof of the previous lemmas, we take the Fourier transform of (2.53) and move into the eigenvector basis of DF by multiplying the resulting equation by P^{-1}. We obtain

$$\hat{V}^{n+1}(k)=\overline{B}_i^{-1}(k)\hat{V}^{n+1,exp}(k)=\overline{B}_i^{-1}(k)\,B_e(k)\,\hat{V}^n(k),$$

where

$$\overline{B}_i(k)=I_3+\alpha\,i\sin\varphi\,A_i,$$
$$B_e(k)=\big(1-\alpha\,\overline{u}\,(1-\cos\varphi)\big)Id-\alpha\,i\sin\varphi\,A_e,$$

with $\varphi = 2\pi\, k\, \Delta x$, $A_e = P^{-1} DF_e P$ given by (2.49) and

$$A_i = P^{-1} DF_i P = \begin{pmatrix} \frac{\overline{u}}{2} - \frac{\overline{c}}{\sqrt{\varepsilon}} & 0 & \frac{\overline{u}}{2} \\ -\overline{u} + \frac{\overline{c}}{\sqrt{\varepsilon}} & 0 & -\overline{u} - \frac{\overline{c}}{\sqrt{\varepsilon}} \\ \frac{\overline{u}}{2} & 0 & \frac{\overline{u}}{2} + \frac{\overline{c}}{\sqrt{\varepsilon}} \end{pmatrix}. \tag{2.54}$$

The matrix $B_e(k)$ is the same as in the proof of Lemma 2.3.2, its eigenvalues are

$$\begin{aligned} &\lambda_{e1} = 1 - \alpha\, \overline{u}\, (1 - \cos\varphi), \\ &\lambda_{e2} = \lambda_{e3} = 1 - \alpha\, \overline{u}\, (1 - \cos\varphi) - i\, \alpha\, \overline{u}\, sin\varphi. \end{aligned}$$

And $0 \leq \alpha\, \overline{u} \leq 1 \implies |\lambda_{ei}| \leq 1$ for $i = 1, 2, 3$.
Those of $\overline{B}_i(k)$ are

$$\begin{aligned} \lambda_{i1} &= 1 + i \left(\sin\varphi\, \alpha\, \overline{u}/2 - \alpha\, |\sin\varphi| \sqrt{\overline{c}^2/\sqrt{\varepsilon} + \overline{u}^2/4} \right), \\ \lambda_{i2} &= 1, \\ \lambda_{i3} &= 1 + i \left(\sin\varphi\, \alpha\, \overline{u}/2 + \alpha\, |\sin\varphi| \sqrt{\overline{c}^2/\sqrt{\varepsilon} + \overline{u}^2/4} \right). \end{aligned}$$

Now we set

$$\alpha\, \overline{u} = 1,$$

and a quite long calculus using the software Maple (see Appendix B for more details on the calculations) yields

$$det(\overline{B}_i^{-1}(k)\, B_e(k) - \lambda\, I_3) = (\cos\varphi - i\, \sin\varphi - \lambda)\, P_2(\lambda),$$

where

$$P_2(\lambda) = \lambda^2 + c_i\, \lambda + d_i,$$

and setting $M_\varepsilon = \sqrt{\varepsilon}\, \overline{u}/\overline{c}$ the coefficients c_i and d_i are given by

$$c_i = -M_\varepsilon^2\, \frac{2\, \cos\varphi + \sin^2\varphi - i\, \sin\varphi\, (1 - \cos\varphi)}{M_\varepsilon^2 + \sin^2\varphi + i \sin\varphi\, M_\varepsilon^2},$$

$$\begin{aligned} d_i = &\frac{M_\varepsilon^4\, (\sin^2\varphi\, \cos\varphi + \cos^2\varphi - i\, \sin\varphi\, \cos\varphi\, (1 - \cos\varphi))}{(M_\varepsilon^2 + \sin^2\varphi + i \sin\varphi\, M_\varepsilon^2)^2} \\ &+ \frac{M_\varepsilon^2\, (\sin^2\varphi\, \cos^2\varphi - i\, \sin^3\varphi\, \cos\varphi)}{(M_\varepsilon^2 + \sin^2\varphi + i \sin\varphi\, M_\varepsilon^2)^2}. \end{aligned}$$

We denote $\lambda_1 = \cos\varphi - i\, \sin\varphi$ and by $\lambda_2(M_\varepsilon, \varphi)$ and $\lambda_3(M_\varepsilon, \varphi)$ the roots of P_2. We have $|\lambda_1| = 1$ and we plot on Figure 2.11, on the left: the maximum modulus $|\lambda_2(M_\varepsilon, \varphi)|$ and $|\lambda_3(M_\varepsilon, \varphi)|$, that is

$$\max(|\lambda_2(M_\varepsilon, \varphi)|, |\lambda_3(M_\varepsilon, \varphi)|)$$

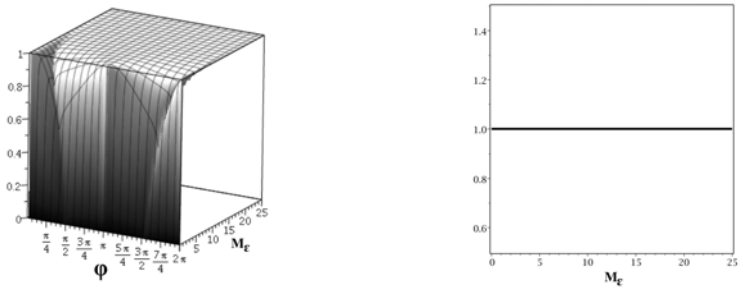


Figure 2.11 – Stability of the NL Order 1 L^2 AP scheme (Lemma 2.3.4): Maximum modulus of the roots of P_2.

as a function of $\varphi \in [0, 2\,\pi]$ and $M_\varepsilon \in]0, 25]$. On the right: the maximum modulus of the roots of P_2 as a function of $M_\varepsilon \in]0, 25]$

$$g(M_\varepsilon) := \max(\max_{\varphi\in[0,2\,\pi[} |\lambda_2(M_\varepsilon, \varphi)|, \max_{\varphi\in[0,2\,\pi[} |\lambda_3(M_\varepsilon, \varphi)|)$$

.

This proves that for all $M_\varepsilon \in]0, 25]$ and all $\varphi \in [0, 2\,\pi]$, the spectral radius of $\overline{B}_i^{-1}(k)\, B_e(k)$, denoted by $r(\overline{B}_i^{-1}(k)\, B_e(k))$, is lower than 1. We conclude like in the proof of Lemma 2.3.2.

Note that when $\sin\varphi = 0$, P_2 reduces in $P_2(\lambda) = \lambda^2 - 2\,\cos\varphi + \cos^2\varphi$ and

$$det(\overline{B}_i^{-1}(k)\, B_e(k) - \lambda\, I_3) = (\cos\varphi - \lambda)\, \left(\lambda^2 - 2\,\cos\varphi + \cos^2\varphi\right).$$

And we have $\lambda_1 = \lambda_2 = \lambda_3 = \cos\varphi$. Moreover, $\overline{B}_i(k) = \Big(I_3 + \alpha\, 2\, i \sin\varphi\, A_i\Big) = Id$ and so

$$det(\overline{B}_i^{-1}(k)\, B_e(k) - \lambda\, I_3) = det(B_e(k) - \lambda\, Id).$$

Therefore, the eigenvalues are those of $B_e(k)$ which when $\alpha\overline{u} = 1$ and $\sin\varphi = 0$, are well given by $\lambda_{e1} = \lambda_{e2} = \lambda_{e3} = \cos\varphi$. □

2.3.4 Numerical results for order 1 schemes

In this part, we present several numerical test cases which show the good behavior of our new order 1 linear AP scheme (2.41), (2.42). We compare it to the nonlinear AP scheme (2.43) inspired by [11]. If not mentioned, the reference solution is computed using an order one explicit scheme with the Rusanov solver on a refined grid ($N = 3000$). For all test cases the space domain is set to $\Omega = [0, 1]$ and we choose $\gamma = 1.4$. If not mentioned $\varepsilon = 1$.

2.3.4.1 Classical Riemann problems: The Sod, Lax and Contact problems

The initial data of the classical Riemann problems is given by

$$(\rho, u, p)(0, x) = \begin{cases} w_L = (\rho_L, u_L, p_L) & \text{if } x < x_d, \\ w_R = (\rho_R, u_R, p_R) & \text{otherwise,} \end{cases}$$

where the initial left and right states values, w_L and w_R respectively, are summarized in Table 2.1. For these test cases Dirichlet conditions are imposed at the boundaries.

Name	t_{final}	x_d		ρ	u	p
Sod	0.2	0.5	w_L	1	0	1
			w_R	0.125	0	0.1
Lax	0.14	0.5	w_L	0.445	1.698	3.528
			w_R	0.5	0	0.571
Contact	0.3	0.5	w_L	5	1	10
			w_R	1	1	10

Table 2.1 – Initial data for the Sod and Contact problems

The Sod problem, represents a benchmarks in gas dynamics [78]. Its solution consists of a left-moving rarefaction fan, an intermediate contact discontinuity and a right-moving shock wave, see Figure 2.12. The Mach number, $M^2 = u^2\,\rho/(\gamma\,p)$, ranges from 0 to 1. Since at $x = 0$ or $x = 1$, $M = 0$ and at $x = 0.5$, $M \approx 0.9$. We can see that our new simpler scheme gives similar results to the NL AP scheme. The linearization of the nonlinear system does not alter the results.

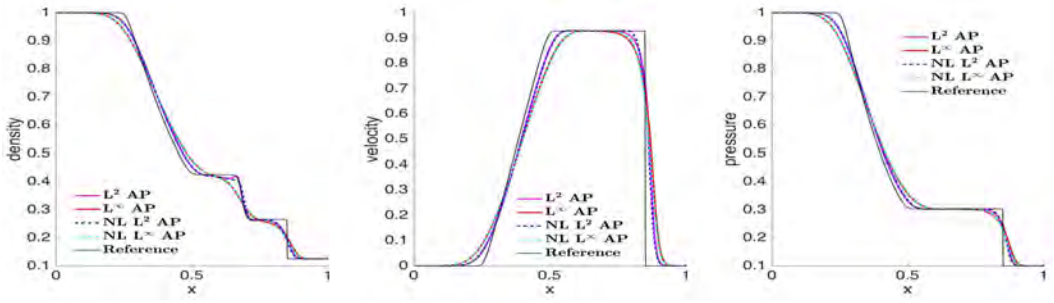


Figure 2.12 – Sod problem for 200 cells. Comparison of our order 1 AP scheme (solid lines) with the NL AP scheme (dashed lines) for both L^2 and L^∞ discretizations.

The solution of the Lax problem also consists of a rarefaction wave, an intermediate contact discontinuity and a shock wave. This test case is more complex compared to the previous one due to the discontinuity in the initial velocity which is not present in the Sod problem. The results are again very similar with a slightly more diffusive behavior for our scheme (see the density profile).
The Contact problem is constituted of a right-moving contact discontinuity on the density with constant velocity and pressure see Figure 2.14. The exact solution is given by:

$$\rho(x,t) = \rho(x-t,0), \quad u(x,t) = u(x,0) \quad \text{and} \quad p(x,t) = p(x,0).$$

It is a particular incompressible solution and is used to test numerically the preservation of the contact discontinuities which is the property (iii) of Lemma 2.2.3. We

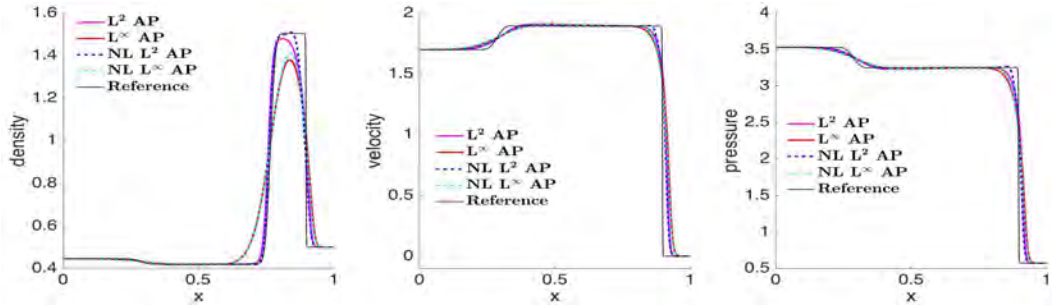


Figure 2.13 – Lax problem for 200 cells. Comparison of our order 1 AP scheme (solid lines) with the NL AP scheme (dashed lines) for both L^2 and L^∞ discretizations.

can see that all schemes preserved the contact discontinuity showing again that the exact resolution of the nonlinear problem is not necessary.

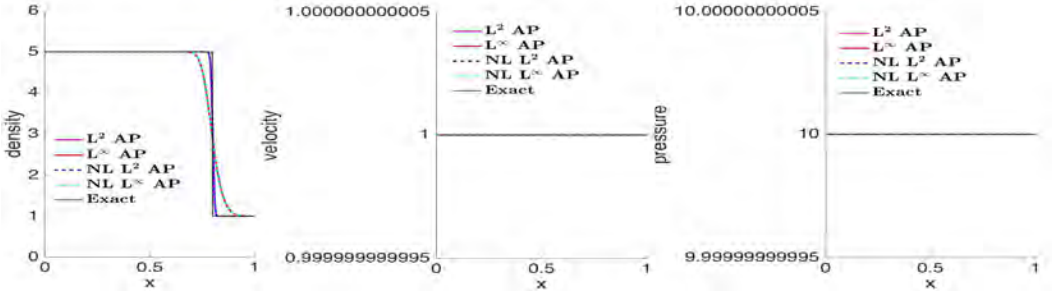


Figure 2.14 – Contact problem for 500 cells. Comparison of our order 1 AP scheme (solid lines) with the NL AP scheme (dashed lines) for both L^2 and L^∞ discretizations.

Note that the implicit upwinding, necessary to obtain the L^∞ stability property, introduces numerical diffusion as it is expected but does not appear to be necessary in these test cases. We will see in the next test case that without this upwinding, non-physical oscillations may appear.

2.3.4.2 Several interacting Riemann problems

The initial data is given by

$$\rho(0,x)=1, \qquad u(0,x)=\begin{cases} 1-\dfrac{\varepsilon}{2} & \text{if } x\in[0,0.2[,\\ 1 & \text{if } x\in[0.2,0.3],\\ 1+\dfrac{\varepsilon}{2} & \text{if } x\in]0.3,0.7[,\\ 1 & \text{if } x\in[0.7,0.8],\\ 1-\dfrac{\varepsilon}{2} & \text{if } x\in]0.8,1], \end{cases} \qquad p(0,x)=1, \quad (2.55)$$

The system is supplemented with periodic boundary conditions. The results are given for different values of the Mach number: $\varepsilon = 1$, $\varepsilon = 10^{-1}$ and $\varepsilon = 10^{-2}$ respectively in Figures 2.15, 2.16 and 2.17. In each corresponding regime we observe oscillations on the density profile when using the L^2 discretization for both our linear AP schemes and also for the NL AP schemes. When using the L^∞ discretizations, these oscillations are significantly reduced. Note that we have also applied this upwinding to the NL AP scheme, the results are similar to those for our new linear scheme. This illustrates the need to add the upwinding on the L^2 discretization. Therefore, for the rest of the paper we will keep only the L^∞ AP scheme called Order 1 AP scheme.

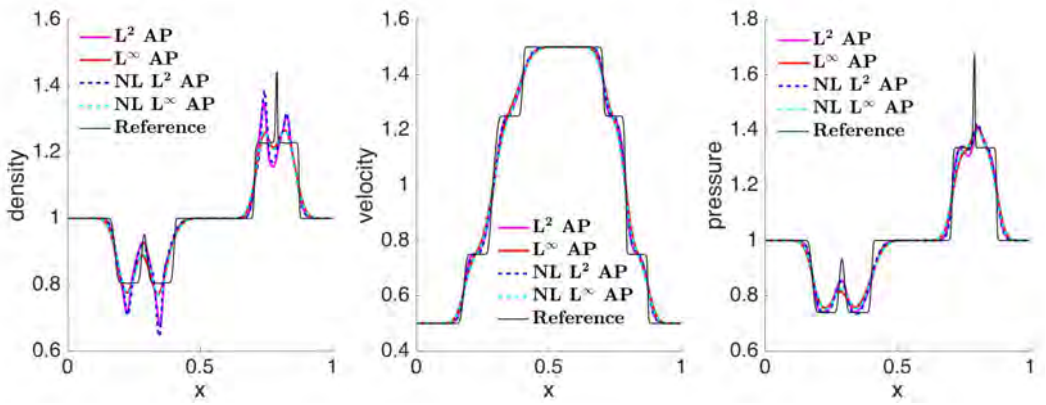


Figure 2.15 – Several interacting Riemann problems experiment. Comparison of our order 1 L^∞ AP scheme against our order 1 L^2 AP scheme and the nonlinear (NL) scheme. $\varepsilon = 1$, $t_{final} = 0.04$ with 200 cells.

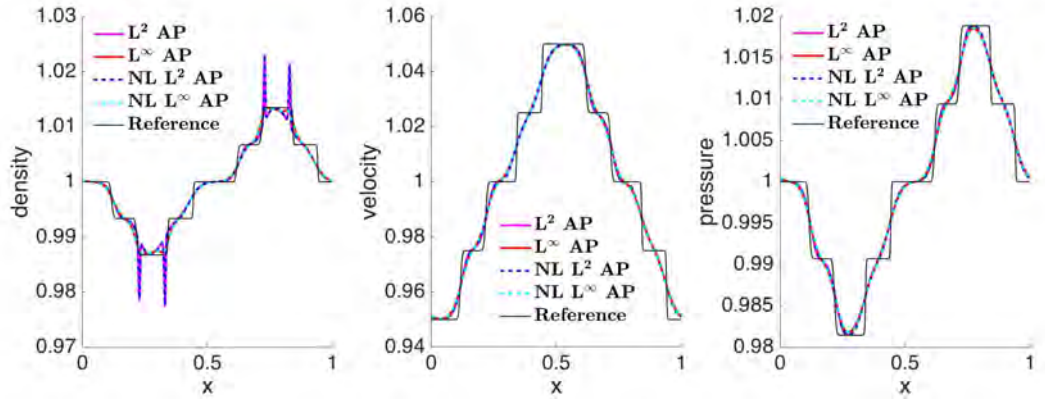


Figure 2.16 – Several interacting Riemann problems experiment. Comparison of our order 1 L^∞ AP scheme against our order 1 L^2 AP scheme and the nonlinear (NL) scheme. Results for $\varepsilon = 10^{-1}$, $t_{final} = 0.03$ with 500 cells.

In Figure 2.19, in order to show the asymptotic stability of our scheme, we compare the time steps of our L^∞ AP scheme against the classical explicit one. On the left

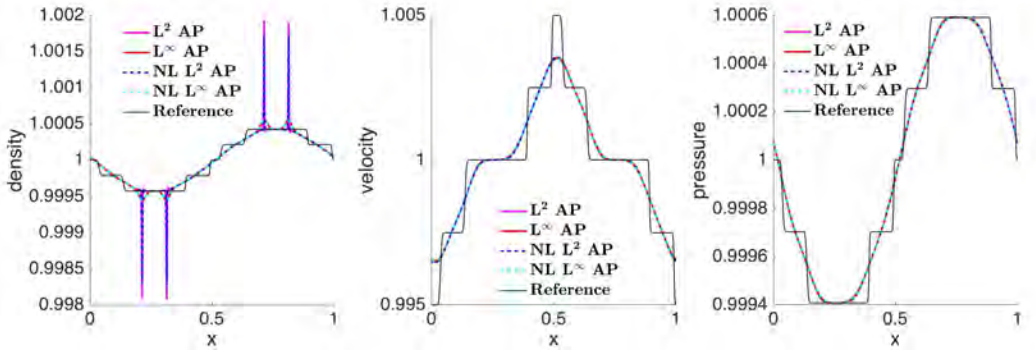


Figure 2.17 – Several interacting Riemann problems experiment. Comparison of our order 1 L^∞ AP scheme against our order 1 L^2 AP scheme and the nonlinear (NL) scheme. Results for $\varepsilon = 10^{-2}$, $t_{final} = 0.015$ with 1000 cells.

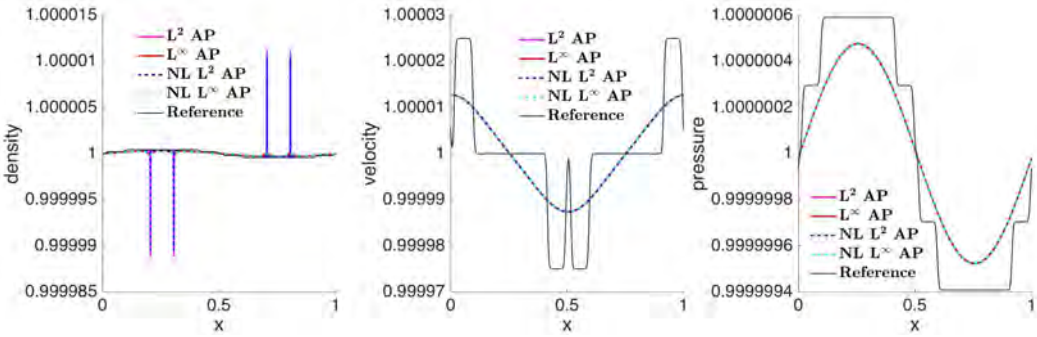
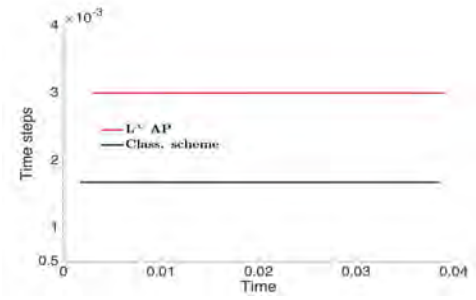
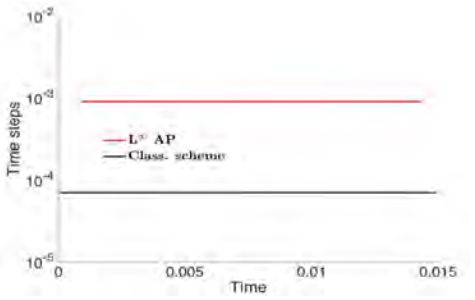


Figure 2.18 – Several interacting Riemann problems experiment. Comparison of our order 1 L^∞ AP scheme against our order 1 L^2 AP scheme and the nonlinear (NL) scheme. Results for $\varepsilon = 10^{-4}$, $t_{final} = 0.006$ with 3500 cells.

picture for $\varepsilon = 1$, we can see that the time steps have the same order of magnitude. On the contrary, on the right picture for $\varepsilon = 10^{-2}$, we observe that the time steps of the AP scheme are around $1/\sqrt{\varepsilon} = 10$ times bigger than the ones required by an explicit method. This shows that the AP scheme can employ time steps which are independent of ε while time steps of explicit schemes remain close to the Mach number value.



(a) $\varepsilon = 1$, final time $t_{final} = 0.04$ for 200 cells



(b) $\varepsilon = 10^{-2}$, final time $t_{final} = 0.015$ for 1000 cells

Figure 2.19 – Several interacting Riemann problems experiment. Comparison of the time step sizes Δt as a function of time between the classical and the AP scheme for different values of ε.

2.4 Low oscillating order 2 AP scheme

In this section, we extend our new linear AP scheme to second order accuracy in time and space. Like in [29, 11], this extension is based on an Implict-Explicit (IMEX) Runge-Kutta approach [80, 66, 31, 9]. In particular, we make use of the second order Ascher, Ruuth and Spiteri [80] scheme denoted in the sequel by ARS(2,2,2) which have been shown in [61] to be the better choice for order 2 discretizations. We recall that implicit methods of order higher than one for hyperbolic problems cannot be TVD nor L^∞ stable for unconstrained time steps [40] and this situation does not change when IMEX methods are employed [29]. To bypass this limitation, we use the same approach as in [29] for the isentropic Euler equations. The idea consists in blending together first and second order implicit time-space discretizations giving rise to an accurate but order 1 AP scheme which guarantees the preservation of the L^∞ stability and TVD property and, to use the order 2 scheme as often as possible by setting up a MOOD (Multidimensional Optimal Order Detection) method [20]. We detail in the next sections the different steps of the scheme.

2.4.1 Order 2 AP semi-discretization in time

First, let us present the order 2 semi-discretization in time. The Butcher tableaux of the ARS(2,2,2) discretization are given in Table 2.2 where $\beta = 1 - \sqrt{2}/2$ and $\alpha = 1 - 1/(2\beta)$. Remarking that $\alpha = \beta - 1$ and $1 - \alpha = 2 - \beta$, the following order

Explicit

0	0	0	0
β	β	0	0
1	α	1 - α	0
	α	1 - α	0

Implicit

0	0	0	0
β	0	β	0
1	0	1 - β	β
	0	1 - β	β

Table 2.2 – Butcher tableaux for the ARS(2,2,2) time discretization. Left panel: explicit tableau. Right panel: implicit tableau.

2 in time scheme is obtained:

$$\begin{aligned}
&\frac{W^\star - W^n}{\Delta t} + \beta\,\nabla\cdot F_e(W^n) + \beta\,\nabla\cdot F_i(W^\star) = 0,\\
&\frac{W^{n+1} - W^n}{\Delta t} + \nabla\cdot F(W^\star) + (1-\beta)\nabla\cdot\left(F_e(W^\star) - F_e(W^n)\right)\\
&\qquad\qquad + \beta\nabla\cdot\left(F_i(W^{n+1}) - F_i(W^\star)\right) = 0.
\end{aligned}$$

Following the reformulation and linearization used for the order 1 AP scheme, we

obtain

$$q^{\star,exp} = q^n - \beta\Delta t\nabla\cdot(\rho^n u^n\otimes u^n), \tag{2.56a}$$
$$E^{\star,exp} = E^n - \beta\Delta t\,\nabla\cdot(k_\varepsilon^n\, u^n), \tag{2.56b}$$
$$\rho^\star = \rho^n - \beta\Delta t\nabla\cdot q^n, \tag{2.56c}$$
$$\frac{\varepsilon}{\gamma-1}p^\star - \beta^2\Delta t^2\nabla\cdot\left(\frac{h^{\star,exp}}{\rho^\star}\nabla p^\star\right) = \varepsilon\left(E^{\star,exp}-k^{\star,exp}\right) - \varepsilon\beta\Delta t\nabla\cdot\left(\frac{h^{\star,exp}}{\rho^\star}q^{\star,exp}\right), \tag{2.56d}$$
$$q^\star = q^{\star,exp} - \beta\Delta t\frac{1}{\varepsilon}\nabla p^\star, \tag{2.56e}$$
$$E^\star = E^{\star,exp} - \beta\Delta t\nabla\cdot\left(\frac{\gamma p^\star}{(\gamma-1)\rho^\star}q^\star\right), \tag{2.56f}$$

where $h^{\star,exp} = \gamma\left(E^{\star,exp} - k_\varepsilon(W^{\star,exp})\right.$.

$$W^{n+1,exp} = \begin{pmatrix}\rho^{n+1,exp}\\ q^{n+1,exp}\\ E^{n+1,exp}\end{pmatrix} = W^n - \Delta t[(\beta-1)\nabla\cdot F_e(W^n) + (2-\beta)\nabla\cdot F_e(W^\star) + (1-\beta)\nabla\cdot F_i(W^\star)], \tag{2.57a}$$
$$\rho^{n+1} = \rho^{n+1,exp} \tag{2.57b}$$
$$\frac{\varepsilon}{\gamma-1}p^{n+1} - \beta^2\Delta t^2\nabla\cdot\left(\frac{h^{n+1,exp}}{\rho^{n+1}}\nabla p^{n+1}\right) = \varepsilon\left(E^{n+1,exp}-k^{n+1,exp}\right) - \varepsilon\beta\Delta t\nabla\cdot\left(\frac{h^{n+1,exp}}{\rho^{n+1}}q^{n+1,exp}\right), \tag{2.57c}$$
$$q^{n+1} = q^{n+1,exp} - \frac{\beta\Delta t}{\varepsilon}\nabla p^{n+1}, \tag{2.57d}$$
$$E^{n+1} = E^{n+1,exp} - \beta\Delta t\nabla\cdot\left(\frac{\gamma p^{n+1}}{(\gamma-1)\rho^{n+1}}q^{n+1}\right), \tag{2.57e}$$

with $h^{n+1,exp} = \gamma\left(E^{n+1,exp} - k_\varepsilon(W^{n+1,exp})\right)$.

2.4.2 Order 2 space discretization in one dimension

In order to extend the space accuracy to second order, we use classically the MUSCL technique [54] and so a piecewise linear reconstruction of W_j^n given by

$$\widehat{W}_j^n(x) = W_j^n + \sigma_j^n(x-x_j)$$

where α_j^n is a limited slope and is computed for each component using a minmod limiter:

$$\alpha_j^n = minmod\left(\frac{W_{j+1}^n - W_j^n}{\Delta x}, \frac{W_j^n - W_{j-1}^n}{\Delta x}\right)$$

where the limiter is defined as

$$minmod(a,b) = \frac{1}{2}(sign(a)+sign(b))\min(|a|,|b|) = \begin{cases} a & if\ |a|<|b|, ab>0, \\ b & if\ |b|<|a|, ab>0, \\ 0 & otherwise. \end{cases}$$

This piecewise linear reconstruction is used for defining the numerical flux at the interfaces using the notations introduced for the order 1 AP scheme

$$(\mathcal{F}_e)^n_{j+1/2} := \frac{F_e(W^n_{j+1,-}) + F_e(W^n_{j,+})}{2} - (\mathcal{D}_e)^n_{j+1/2}(W^n_{j+1,-} - W^n_{j,+}), \tag{2.58}$$

where $(\mathcal{D}_e)^n_{j+1/2} = \frac{1}{2}\max\left(|u^n_{j,+}|, |u^n_{j+1,-}|\right)$ and where

$$W^n_{j,\pm} = \widehat{W}^n_j\left(x_j \pm \frac{\Delta x}{2}\right) = W^n_j \pm \frac{\Delta x}{2}\,\sigma^n_j.$$

In the momentum and energy equations the implicit flux F_i is discretized with a centered solver which ensures second order accuracy in space so no reconstruction is needed. Moreover, we note that the discretization of the flux operator $\nabla\cdot\left(\frac{h}{\rho}\nabla p\right)_j$ is also second order accurate. Indeed, its discreatization in one dimension reads:

$$\frac{1}{\Delta x}\left(\frac{1}{2}(\phi_{i+1}+\phi_i)\frac{\psi_{i+1}-\psi_i}{\Delta x} - \frac{1}{2}(\phi_i+\phi_{i-1})\frac{\psi_i-\psi_{i-1}}{\Delta x}\right), \tag{2.59}$$

where we set $\phi = \frac{h}{\rho}$, $\psi = p$ and $f_k = f(x_k)$ for $f = \phi, \psi$ and $k \in \{i, i+1, i-1\}$. Using a Taylor expansion in terms of $f(x_i)$ for $f(x_{i+1})$ and $f(x_{i-1})$, we have:

$$f(x_{i+1}) = f(x_i) + \Delta x\partial_x f(x_i) + \frac{\Delta x^2}{2}\partial^2_{xx} f(x_i) + \frac{\Delta x^3}{6}\partial^3_{xxx} f(x_i) + \mathcal{O}(\Delta x^4), \tag{2.60}$$

$$f(x_{i-1}) = f(x_i) - \Delta x\partial_x f(x_i) + \frac{\Delta x^2}{2}\partial^2_{xx} f(x_i) - \frac{\Delta x^3}{6}\partial^3_{xxx} f(x_i) + \mathcal{O}(\Delta x^4). \tag{2.61}$$

And so,

$$\begin{aligned}
\frac{\psi_{i+1}-\psi_i}{\Delta x} &= \partial_x\psi_i + \frac{\Delta x}{2}\partial^2_{xx}\psi_i + \frac{\Delta x^2}{6}\partial^3_{xxx}\psi_i + \mathcal{O}(\Delta x^3), \\
\frac{\psi_i-\psi_{i-1}}{\Delta x} &= \partial_x\psi_i - \frac{\Delta x}{2}\partial^2_{xx}\psi_i + \frac{\Delta x^2}{6}\partial^3_{xxx}\psi_i + \mathcal{O}(\Delta x^3), \\
\frac{1}{2}(\phi_{i+1}+\phi_i) &= \phi_i + \frac{\Delta x}{2}\partial_x\phi_i + \frac{\Delta x^2}{4}\partial^2_{xx}\phi_i + \mathcal{O}(\Delta x^3), \\
\frac{1}{2}(\phi_i+\phi_{i-1}) &= \phi_i - \frac{\Delta x}{2}\partial_x\phi_i + \frac{\Delta x^2}{4}\partial^2_{xx}\phi_i + \mathcal{O}(\Delta x^3).
\end{aligned}$$

Then, injecting those expressions into (2.59), gives

$$\begin{aligned}
D(\partial_x(\phi\partial_x\psi))_j &= \frac{1}{\Delta x}\left(\frac{1}{2}(\phi_{i+1}+\phi_i)\frac{\psi_{i+1}-\psi_i}{\Delta x} - \frac{1}{2}(\phi_i+\phi_{i-1})\frac{\psi_i-\psi_{i-1}}{\Delta x}\right) \\
&= \frac{1}{\Delta x}\left(\Delta x\partial_x\psi_i\partial_x\phi_i + \Delta x\partial^2_{xx}\psi_i\phi_i + \mathcal{O}(\Delta x^3)\right) \\
&= \partial_x(\phi\partial_x\psi)(x_j) + \mathcal{O}(\Delta x^2).
\end{aligned}$$

The space accuracy with an upwind discretization of the implicit flux could also be increased. In that case, it would require a linear reconstruction of the fluxes

$$\frac{1}{\varepsilon}\, p\, I_d \quad \text{and} \quad \frac{\gamma p}{(\gamma-1)\rho} q$$

in the momentum and energy equations of (2.56)-(2.57). In our discretization, the pressures $p^\star$ and p^{n+1} are not defined as a function of the conservative variables (not defined with the state equation) since they are given respectively by (2.56d) when computing $W^\star$ and by (2.57c) for W^{n+1}. Therefore, the definition of the implicit numerical flux by using a piecewise linear reconstruction on the conservative variables (ρ, q, E) as done for the explicit part in (2.58) is not possible here. We also started to investigate a linear reconstruction on the primitive variables to compute the implicit fluxes $\frac{1}{\varepsilon}\, p\, I_d$ and $\frac{\gamma p\, q}{(\gamma-1)\rho} = \frac{\gamma p\, u}{\gamma-1}$ but did not have time to go further in that direction.
In this work, we propose to use a linear reconstruction only when adding numerical dissipation on the conservative variables. Moreover, numerical tests intend to show that adding implicit diffusion only at the end of the second step is sufficient (see Figure 2.20). As done for the Order 1 L^∞ stable scheme, we compute the L^2 stable solution $W_j^{n+1,L2}$ with (2.56)-(2.57), (2.58) and then add numerical dissipation on the conservative variables:

$$W_j^{n+1} = \begin{pmatrix} \rho_j^{n+1,L2} \\ q_j^{n+1,L2} \\ E_j^{n+1,L2} \end{pmatrix} + \frac{\beta\Delta t}{\Delta x}\left((\mathcal{D}_i)_{j+1/2}^n(\tilde{W}_{j+1,-}^{n+1} - \tilde{W}_{j,+}^{n+1})\right) - \frac{\beta\Delta t}{\Delta x}\left((\mathcal{D}_i)_{j-1/2}^n(\tilde{W}_{j,-}^{n+1} - \tilde{W}_{j-1,+}^{n+1})\right), \quad (2.62)$$

where the implicit viscosity coefficient

$$(\mathcal{D}_i)_{j+1/2}^n = \frac{1}{2}\max\left(|\lambda_i(W_{j+1,-}^n)|, |\lambda_i(W_{j,+}^n)|\right), \quad (2.63)$$

with $|\lambda_i(W)| = \frac{|u|}{2} + \sqrt{\frac{u^2}{4} + \frac{c^2}{\varepsilon}}$ and where like in [29]

$$\tilde{W}_{j,\pm}^{n+1} = W_j^{n+1} \pm \frac{\Delta x}{2}\sigma_j^n.$$

Let us also remark that in this case, we impose explicit slopes for $\tilde{W}_{j,\pm}^{n+1}$ in order to avoid the nonlinearity arising from an implicit reconstruction of W_j^{n+1}.

2.4.3 The accurate TVD AP scheme

It is well known that a second order discretization in space introduces oscillations which can be eliminated by using limiters such as the minmod limiter. The same problem occurs with time discretization. Indeed, using this second order discretization in time with an order 1 discretization in space leads to numerical oscillations.

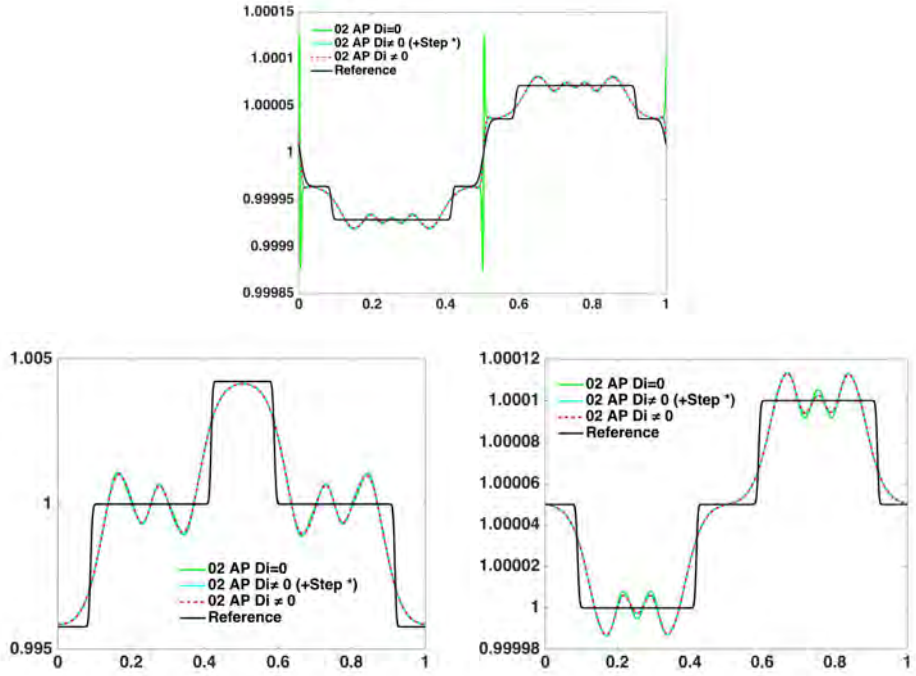


Figure 2.20 – Approximation of the physical variables ρ (top), u (bottom left) and p (bottom right) for a shock tube test case when $\varepsilon = 10^{-4}$ (see Section 2.4.5.1 for its description). Comparison of the Order 2 AP schemes against a reference solution for various choices on the implicit upwinding. Results for "**02 AP Di=0**": no implicit upwinding added (green curve), for "**02 AP Di≠0 (+Step*)**": implicit upwinding added at the end of the first step, i.e, after computing $W^\star$ and at the end of second step (cyan curve) and for "**02 AP Di≠0**": implicit upwinding added only at the end of the second step (red dashed curve). The cyan and red curves overlap intending to show that adding numerical viscosity only at the end of the second step (red dashed curve) is sufficient.

On Figure 2.21, we can see that the Order 2 AP scheme gives more accurate results than the Order 1 AP scheme but we can also remark that when the Mach number decreases oscillations appear. In [29], it has been proved for the following linear advection equation

$$\partial_t w + c_e \partial_x w + \frac{c_i}{\sqrt{\varepsilon}} \partial_x w = 0, \tag{2.64}$$

where $c_e > 0$ and $c_i > 0$, that these oscillations are the result of the loss of the L^∞ stability and TVD (total variation diminishing) properties of the second order semi-discretisation. Let us recall that a scheme is L^∞ stable when for all $n \geq 0$,

$$\|w^{n+1}\|_\infty \leq \|w^n\|_\infty = \max_{j \in \{1 \dots N\}} |w_j^n|,$$

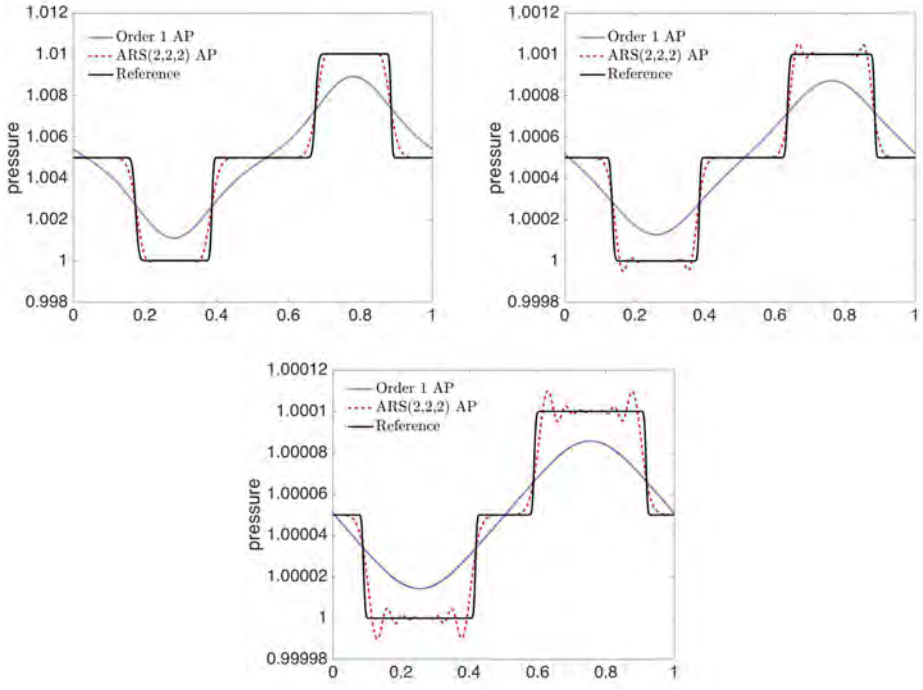


Figure 2.21 – Approximations of the pressure for a shock tube test case (see Section 2.4.5.1 for its description) for different Mach numbers. Comparison of the first-order AP scheme (blue dotted line) and of the second-order in time AP scheme (red dashed line) against a reference solution (black solid line) for different values of ε: for $\varepsilon = 1$ (top left), for $\varepsilon = 10^{-2}$ (top right) and for $\varepsilon = 10^{-4}$ (bottom). An order 1 space discretization is used for the AP schemes.

and TVD when

$$TV(w^{n+1}) \leq TV(w^n) = \sum_{j=1}^{N} |w_{j+1}^n - w_j^n|,$$

where $w_j^n = w(x_j, t^n)$. In [29], the equation (2.64) was discretized with an upwind scheme in space and similarly to the Euler equations, the slow scale was discretized explicitly and the fast scale (where we note a dependency in $1/\sqrt{\varepsilon}$) implicitly. The full discretization of the order one scheme used reads

$$w_j^{n+1} = w_j^n - \frac{\Delta t}{\Delta x} c_e (w_j^n - w_{j-1}^n) - \frac{\Delta t}{\Delta x} \frac{c_i}{\sqrt{\varepsilon}} (w_j^{n+1} - w_{j-1}^{n+1}). \tag{2.65}$$

For the second-order time discretization, an ARS(2,2,2) scheme was used. Unfortunately, for the IMEX second order discretization in time the L^∞ stability and TVD properties can be recovered only if the time steps are of the order of that of the explicit semi-discretization and so constrained by the Mach number. It has been shown in [40] that there does not exist TVD implicit Runge-Kutta schemes with unconstrained time steps of order higher than one for an hyperbolic equation and this situation does not change when IMEX methods are employed [29].

To tackle this problem and obtain a L^∞ stable and TVD numerical semi-discretization more accurate than the Order 1 AP scheme, we introduce a convex combination between the order 1 and order 2 schemes, as proposed in [29] for the isentropic Euler equations:

$$W^{n+1} = (1-\theta)W^{n+1,01} + \theta W^{n+1,02}, \tag{2.66}$$

where $W^{n+1,01}$ is given by the Order 1 AP scheme (2.38), $W^{n+1,02}$ by the second-order AP one (2.56)-(2.57) and $\theta \in [0,1]$. We aim to choose the largest possible value of θ to be as accurate as possible while ensuring the TVD property. We set θ to $\frac{\beta}{1-\beta} = \sqrt{2}-1 \approx 0.4142$ since in [29], it has been proved that this value is the largest possible value to ensure the TVD property in the case of the linear advection equation (2.64). It was shown for $w^{n+1,01}$ computed with (2.65) and an upwind scheme with the ARS(2,2,2) second-order time-discretization for computing $w^{n+1,02}$. Finally, we obtain the following scheme referred as the TVD AP scheme

$$\frac{W^\star - W^n}{\Delta t} + \beta \nabla \cdot F_e(W^n) + \beta \nabla \cdot F_i(W^\star) = 0, \tag{2.67a}$$

$$\begin{aligned}\frac{W^{n+1} - W^n}{\Delta t} &+ (1-\theta+\theta(\beta-1))\nabla \cdot F_e(W^n) + \theta(2-\beta)\nabla \cdot F_e(W^\star) \\ &+ \theta(1-\beta)\nabla \cdot F_i(W^\star) + (1-\theta+\theta\beta)\nabla \cdot F_i(W^{n+1}) = 0.\end{aligned} \tag{2.67b}$$

The space discretization is the same as for the Order 2 AP scheme where the explicit flux F_e is given by (2.58) and the implicit flux is discretized with a centered solver. The upwinding on the implicit part is also added at the end of the second step. As done for the Order 1 L^∞ stable scheme, we compute the L^2 stable solution $W_j^{n+1,L2}$ with (2.67), (2.58) and then add numerical dissipation on the conservative variables:

$$\begin{aligned}W_j^{n+1} = \begin{pmatrix} \rho_j^{n+1,L2} \\ q_j^{n+1,L2} \\ E_j^{n+1,L2} \end{pmatrix} &+ (1-\theta+\theta\beta)\frac{\Delta t}{\Delta x}\left((\mathcal{D}_i)^n_{j+1/2}(W^{n+1}_{j+1} - W^{n+1}_j)\right) \\ &- (1-\theta+\theta\beta)\frac{\Delta t}{\Delta x}\left((\mathcal{D}_i)^n_{j-1/2}(W^{n+1}_{j,} - W^{n+1}_{j-1})\right),\end{aligned} \tag{2.68}$$

where $W_j^{n+1,L2}$ is computed with (2.67), (2.58) and

$$(\mathcal{D}_i)^n_{j+1/2} = \frac{1}{2}\max\left(|\lambda_i(W^n_{j+1})|, |\lambda_i(W^n_j)|\right).$$

In the following, this scheme is referred as the TVD AP scheme and defined by (2.67) in time and (2.58), (2.68) in space. Note that this scheme is more accurate than the Order 1 AP scheme but is still of order 1. We will see in Section 2.4.5, how we can bypass this limitation.

Remark 7. Our first choice was to perform a reconstruction on the conservative variables when adding the implicit upwinding with (2.68) as done for the Order 2

AP scheme with (2.62) (see Section 2.4.2 for more details). But, numerical tests showed that with this choice, spurious oscillations around shocks are not always eliminated. On Figure 2.22, we compare the results given on the density for the shock tube problem (see Section 2.4.5.1) with various values of the Mach number when: dissipation is added with reconstruction ("**TVD AP Di≠0 (W+/-)**", blue curve) and when dissipation is added with (2.68) ("**TVD AP Di≠0**", red dashed curve). We see that we obtain more accurate results with the reconstruction for $\varepsilon = 1$ and 10^{-2} but for $\varepsilon = 10^{-4}$ oscillations are present with this choice. Therefore, in order to ensure the TVD property we choose to add the upwinding with (2.68).

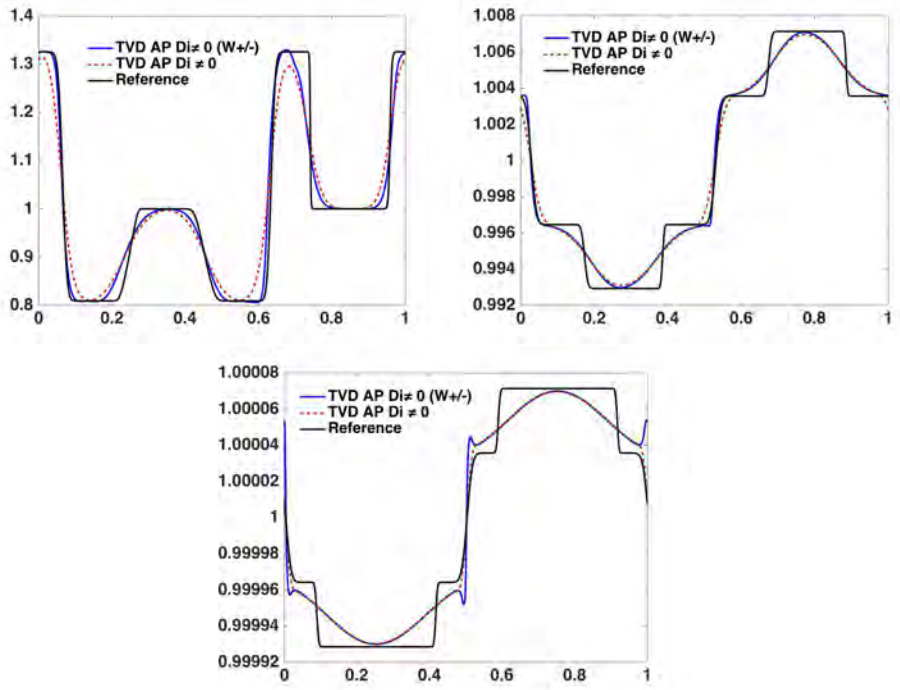


Figure 2.22 – Approximations of the density for a shock tube test case (see Section 2.4.5.1 for its description) for different values of ε: for $\varepsilon = 1$ (top left), for $\varepsilon = 10^{-2}$ (top right) and for $\varepsilon = 10^{-4}$ (bottom). Comparison of the TVD AP schemes against a reference solution when: the upwinding is added at the end of the second step with a linear reconstruction on the conservative variables as done for the Order 2 AP scheme (blue curve) and when the upwinding is added at the end of the second step with (2.68) (red dashed curve). The results are more accurate with the reconstruction procedure but for $\varepsilon = 10^{-4}$ spurious oscillations around shocks are not completely eliminated and thus, the TVD proprety is lost in this case.

2.4.4 Numerical results

In this part, we present several numerical test cases which show the behavior of our AP schemes. We resume here, for eaech scheme, the corresponding discretization

used:

— The **Order 1 AP scheme** is given by the full discretization (2.41), (2.42).
— The **Order 2 AP scheme** is given by the semi-discretization (2.56), (2.57) and the space discretization (2.58), (2.62).
— The **TVD AP scheme** is given by the semi-discretization (2.67) and the space discretization (2.58), (2.68).

In each presented test case we compare the Order 1 AP scheme (blue dotted line), the TVD AP scheme (red line) and the Order 2 AP scheme (green line) against a reference solution that is, if not mentioned, computed using an order one explicit scheme with the Rusanov solver on a refined grid ($N = 3000$). For all test cases the space domain is set to $\Omega = [0, 1]$ and we choose $\gamma = 1.4$. If not mentioned $\varepsilon = 1$.

On Figure 2.23, we see that with the constructed TVD AP scheme we are able as expected to obtain a more accurate first order scheme that ensures the TVD (Total variation diminishing) property. The scheme is less accurate than the Order 2 AP scheme but does not show any oscillation even as the Mach number decreases.

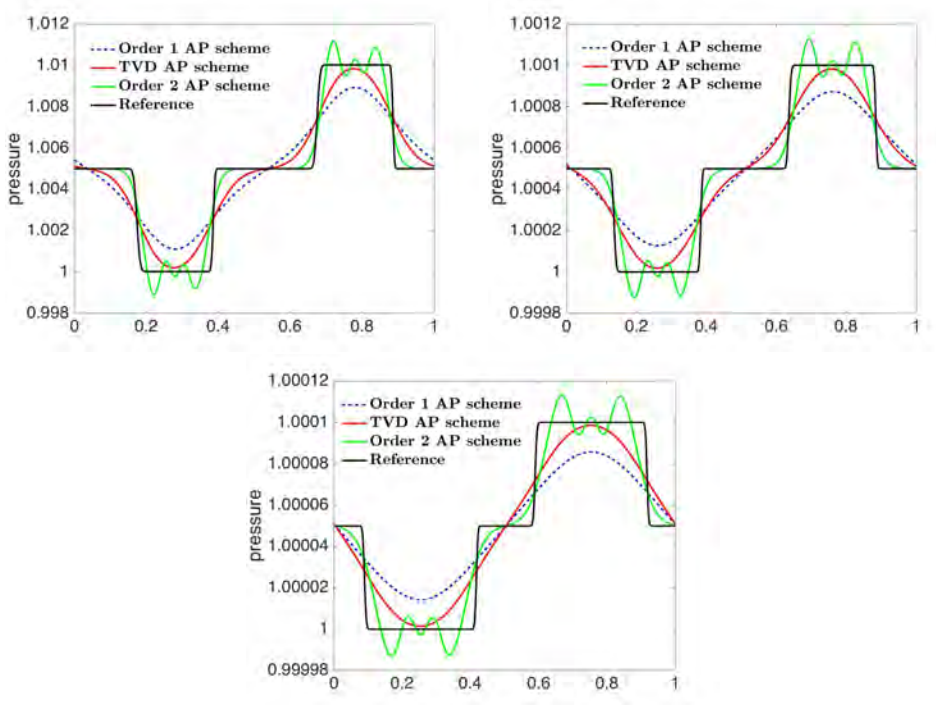


Figure 2.23 – Approximations of the pressure for a shock tube test case (see Section 2.4.5.1 for its description) for different Mach numbers: Comparison of the first-order AP scheme (blue dotted line), the second-order AP scheme (green line) and of the TVD AP scheme (red line) against a reference solution (black solid line) for different values of ε: for $\varepsilon = 1$ (top left), for $\varepsilon = 10^{-2}$ (top right) and for $\varepsilon = 10^{-4}$ (bottom).

Classical Riemann problems: The Sod, Lax and Contact (stiff) problems

Here we validate the behavior of the TVD and Order 2 AP schemes for the Riemann problems introduced in Section 2.3.4.1. The initial data of the classical Riemann problems is given by

$$(\rho, u, p)(0, x) = \begin{cases} w_L = (\rho_L, u_L, p_L) & \text{if } x < x_d, \\ w_R = (\rho_R, u_R, p_R) & \text{otherwise.} \end{cases}$$

The initial left and right states values, w_L and w_R respectively, are summarized in Table 2.3. For all tests Dirichlet conditions are imposed at the boundaries.
The results are shown for the Lax and Sod problem and for a stiff contact problem. On the stiff contact problem, the difference between the left and right initial states for the density is of order 10^5.

Name	t_{final}	x_d		ρ	u	p
Sod	0.2	0.5	w_L	1	0	1
			w_R	0.125	0	0.1
Lax	0.14	0.5	w_L	0.445	1.698	3.528
			w_R	0.5	0	0.571
Stiff contact	0.5	0.25	w_L	1000	1	10^5
			w_R	0.01	1	10^5

Table 2.3 – Initial data for the Sod, Lax and Contact (stiff) problems

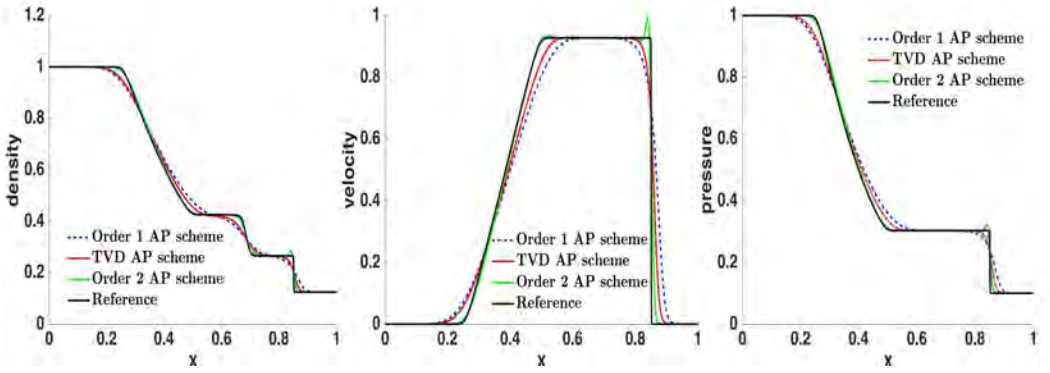


Figure 2.24 – Sod problem for 200 cells: Comparison of the first-order AP scheme (blue line), the second-order AP scheme (green line) and the TVD AP scheme (red line) against the reference solution (black solid line).

The 2D isentropic vortex : Numerical convergence

The isentropic vortex problem was initially introduced by [46] to test the accuracy of numerical methods since the analytical solution is regular and known. It corre-

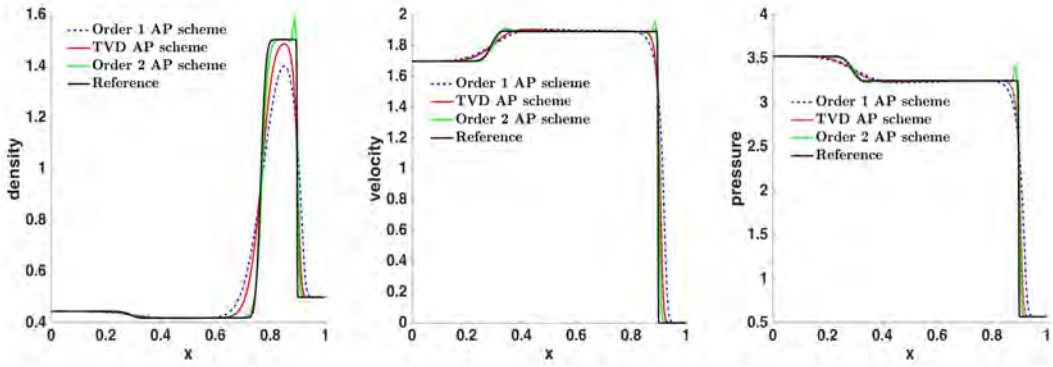


Figure 2.25 – Lax problem for 200 cells: Comparison of the first-order AP scheme (blue line), the second-order AP scheme (green line) and the TVD AP scheme (red line) against the reference solution (black solid line).

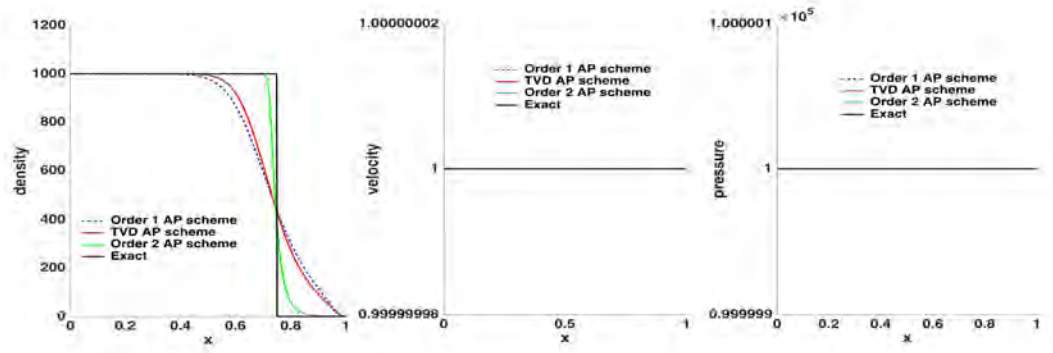


Figure 2.26 – Contact problem (stiff) for 500 cells: Comparison of the first-order AP scheme (blue line), the second-order AP scheme (green line) and the TVD AP scheme (red line) against the reference solution (black solid line).

sponds to a flow characterized by $(\rho_\infty, u_\infty, v_\infty, p_\infty) = (1, 1, 1, 1)$ to which we add an isentropic vortex given by perturbations in (u, v) and the temperature $T = p/\rho$ but no perturbation in the entropy $S = p/\rho^\gamma$:

$$(\delta u, \delta v) = \frac{d}{2\pi} e^{\frac{1-r^2}{2}}(-y, x), \quad \delta T = -\frac{(\gamma-1)d^2}{8\gamma\pi^2} e^{1-r^2}, \quad \delta S = 0,$$

where $r = \sqrt{x^2 + y^2}$ and the vortex strength $d = 5$. The initial data is given by

$$(\rho, u, v, p)(0, x, y) = (\rho_\infty + \delta\rho, u_\infty + \delta u, v_\infty + \delta v, p_\infty + \delta p), \tag{2.69}$$

where the perturbations for the density and pressure read

$$\delta\rho = (1 + \delta T)^{1/(\gamma-1)} - 1 \quad \text{and} \quad \delta p = (1 + \delta T)^{\gamma/(\gamma-1)} - 1.$$

The domain is set to $\Omega = [-5, 5]^2$ and periodic boundary conditions are used. The exact solution of this problem with the above initial data is the initial vortex

convected with the mean velocity, i.e.,

$$W_{ex}(t,x,y) = W_0(x - u_\infty t, y - v_\infty t).$$

For details on the space discretization of the AP schemes in dimension two, see Section 3.2.4.
To assess the numerical order of accuracy, we compute the relative L^2 errors on the density for several uniform meshes:

$$e_{L2} = \frac{||\rho^n - \rho^n_{ex}||_{L^2}}{||\rho^n_{ex}||_{L^2}} = \frac{\sqrt{\sum_{i,j} |\rho^n_{i,j} - \rho_{ex}(t^n, x_i, y_j)|^2}}{\sqrt{\sum_{i,j} |\rho_{ex}(t^n, x_i, y_j)|^2}}.$$

The L^2 errors are computed at $t_{final} = 1$ and shown in logarithmic scale as a function of the number of cells on Figure 2.28 on the right. We get the right orders for each of the schemes. Note that the TVD AP scheme is of order 1 but with an error always lower than the first order schemes, which confirms that the accuracy has been increased.

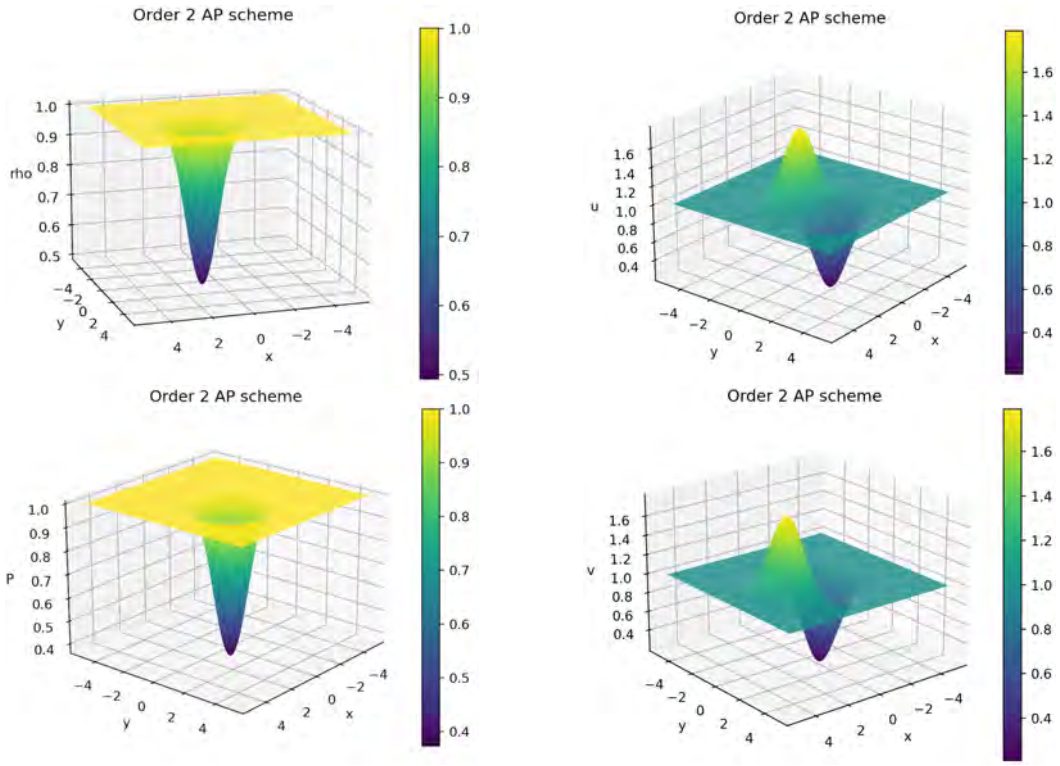


Figure 2.27 – Isentropic vortex (Section 2.4.4): Left panels: Density (top) and pressure (bottom) profiles. Right panels: velocity in x direction (top) and velocity in y direction (bottom). Surface plots for the unlimited Order 2 AP scheme with 128×128 grid points.

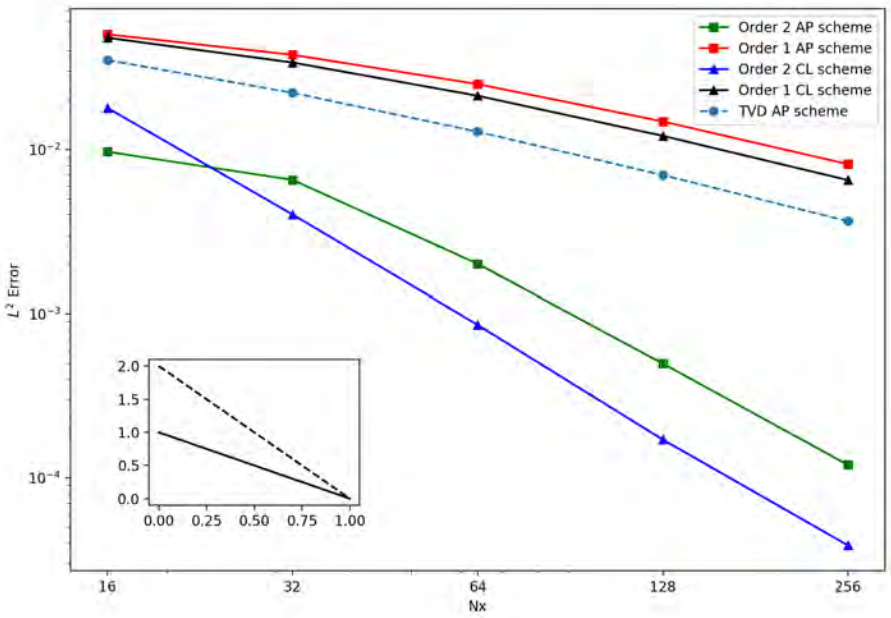


Figure 2.28 – Isentropic vortex (Section 2.4.4): Logscale of the L^2 norm of the density error at time $t_{final} = 1$ for the Order 1 and Order 2 unlimited AP (squares) and explicit schemes (triangles) and for the TVD AP scheme (dots) as a function of the number of cells.

2.4.5 Mood procedure

Since we observed that in many situations the full second order AP scheme can be employed without formation of spurious oscillations and since the TVD AP scheme is only first-order accurate we aim in constructing an optimized AP scheme using the MOOD technique [33], [34] (Multidimensional Optimal Order Detection).
It consists in using at each time step the second-order oscillatory discretization (2.56)-(2.57) whenever possible, i.e., when no oscillations appear. Instead, if the numerical solution presents oscillations, we discard it and we replace it by the limited TVD AP scheme (2.67) the most accurate order 1 scheme. The procedure can be summarized by the following algorithm:

Algorithm 1.
1. Compute a candidate solution $W^{n+1,O2}$ using the second order AP scheme (2.56)-(2.57), (2.58);
2. Detect in each cells if this candidate presents spurious oscillations applying the local detection criteria (see bellow);
3. If in a cell j, spurious oscillations are detected then compute the solution with the TVD AP scheme (2.67), (2.58), (2.68), $W^{n+1} = W^{n+1,TVD}$. Otherwise, if for all cells the local detection criteria is satisfied, set $W^{n+1} = W^{n+1,O2}$.

We refer to this algorithm as the AP-MOOD scheme for the Euler equations.

Detection criterion : Discrete maximum principle on u and p

The detection of these oscillations is a difficult problem. Many different detection criteria can be found in the literature such as the positivity of the density, a maximum principle on the physical quantities, on the conservative variables... Note that the maximum principle on physical or conservative quantities is not a property satisfied at the continuous level.
Here we prefer to choose a criterion based on a property verified by the continuous problem. It is known that at a continuous level and for a Riemann problem at least u or p satisfy the maximum principle. We propose to use this property for detecting spurious oscillations. We introduce a local detection criterion which relies on testing whether both u and p break the maximum principle at the same time. For each cell j, we calculate the following bounds:

$$\begin{aligned} (u_{min})_j^n &= \min(u_{j-1}^n, u_j^n, u_{j+1}^n), & (u_{max})_j^n &= \max(u_{j-1}^n, u_j^n, u_{j+1}^n), \\ (p_{min})_j^n &= \min(p_{j-1}^n, p_j^n, p_{j+1}^n), & (p_{max})_j^n &= \max(p_{j-1}^n, p_j^n, p_{j+1}^n). \end{aligned}$$

We detect a loss of the maximum principle on u and p simultaneously on a given cell $j \in 1, \dots, N$ if for $f = p$ and u:

$$f_j^{n+1} < m_{f,j}^n - tol \qquad \text{or} \qquad f_j^{n+1} > M_{f,j}^n + tol, \tag{2.70}$$

where $m_{f,j}^n = (f_{min})_j^n$, $M_{f,j}^n = (f_{max})_j^n$ and

$$tol = \mu_{tol} \frac{\max_j(f_j^0)}{\min_j(f_j^0)} (\max_j(f_j^0) - \min_j(f_j^0)).$$

The tolerance parameter μ_{tol} allows us to relax the discrete maximum principle. It must be chosen such as not to activate the procedure too much in order to be as close as possible to the second-order solution, but so as to activate it enough to significantly reduce the oscillations. The best choice for μ_{tol} is case-dependent and therefore its value is different for each test case. With this choice, we observe that oscillations are reduced and the accuracy of this AP-MOOD scheme is better than the one of the TVD AP scheme, see figures in Section 2.4.5.1.

Other considered detection criteria

- ***Global criterion***

We have also considered a global detection criterion as proposed in [29] for the isentropic case, this means that we detect a loss of the maximum principle on u and p simultaneously if for $f = p$ and u:

$$f^{n+1} < m_f^n - tol \text{ or } f_j^{n+1} > M_f^n + tol, \tag{2.71}$$

where

$$m_f^n = \min_j(f_j^n) \text{ and } M_f^n = \max_j(f_j^n).$$

We have replaced $(f_{min})_j^n$ and $(f_{max})_j^n$ by $(f_{min})^n$ and $(f_{max})^n$. This method is less costly and the tolerance μ_{tol} is less test case sensitive but in some cases, oscillations are not detected even diminishing μ_{tol}. We can observe it for the Shock tube problem (see Section 2.4.5.1 for the test case description). In Figure 2.29, we compare for the velocity profile the results given using our local detection criterion (red line on the left), and using the global criterion (magenta line on the right). For each value of ε, we use the same number of cells and final times as in Section 2.4.5.1. We observe that the oscillations on the intermediate state are not well-captured when using the global criterion. Actually, they are at most detected only for the first two iterations even though we have set $\mu_{tol} = 0$ (i.e. the maximum principle on u and p is not relaxed). On the other hand, with the local criterion we have fixed $\mu_{tol} = 1.4 \times 10^{-1}$ and we obtain much better results (red line on the left figures).

- ***Local correction procedure***

We also considered a local correction that is, we have changed Step 3 of Algorithm 1 by

3. If in a cell j, spurious oscillations are detected, we set $W_j^{n+1} = W_j^{n+1,TVD}$ where $W_j^{n+1,TVD}$ is computed with the TVD AP scheme (2.67), (2.58), (2.68), otherwise set $W_j^{n+1} = W_j^{n+1,O2}$.

In this case, the results are not satisfactory, with sometimes higher oscillations appearing in neighboring cells. We can observe it on Figure 2.30 where we obtain good results for $\varepsilon = 10^{-2}$ (top right) and $\varepsilon = 10^{-3}$ (top left) but not for $\varepsilon = 10^{-4}$ (bottom).

2.4.5.1 Results for the shock tube problem

We consider a Riemann problem with the following initial data:

$$\rho(0,x) = 1, \quad u(0,x) = 1, \quad p(0,x) = \begin{cases} 1+\varepsilon & \text{if } x < 0.5, \\ 1 & \text{otherwise.} \end{cases} \tag{2.72}$$

Periodic boundary conditions are prescribed. In Figures 2.31, we compare the results given by the four schemes in several regimes corresponding to different values of the Mach number: $\varepsilon = 1$, $\varepsilon = 10^{-2}$, $\varepsilon = 10^{-3}$ and $\varepsilon = 10^{-4}$. We observe that the Order 1 AP scheme is very diffusive while the Order 2 scheme better approximates the solution but presents oscillations increasing as the Mach number decreases. The TVD AP scheme shows almost no oscillation as the Mach number decreases. In red, we show the results of our AP-MOOD scheme which reduces significantly the oscillations produced by the Order 2 AP scheme and gives results as accurate as the Order 2 AP scheme.

2.4.5.2 Results for the Several interacting Riemann problems test case

We propose to also test the developed AP-MOOD scheme for the severe test case "Several interacting Riemann problems" introduced in Section (2.3.4.2). We remind the initial data below

$$\rho(0,x)=1, \qquad u(0,x)=\begin{cases} 1-\dfrac{\varepsilon}{2} & \text{if } x\in[0,0.2[,\\ 1 & \text{if } x\in[0.2,0.3],\\ 1+\dfrac{\varepsilon}{2} & \text{if } x\in]0.3,0.7[,\\ 1 & \text{if } x\in[0.7,0.8],\\ 1-\dfrac{\varepsilon}{2} & \text{if } x\in]0.8,1], \end{cases} \qquad p(0,x)=1. \quad (2.73)$$

In Figure 2.32, we show the physical variable profiles for decreasing values of ε: $\varepsilon=10^{-2}$, $\varepsilon=10^{-3}$ and $\varepsilon=10^{-4}$. The tolerance in the detection criterion (2.70) is set to $\mu_{tol}=5\times10^{-2}$. With this choice, we observe that the oscillations on the velocity and pressure are well captured by the AP-MOOD procedure.

2.5 Conclusion

In this chapter, we have developed and studied a new linear AP IMEX scheme for the compressible Euler system in the low Mach number limit. We have shown that the chosen flux splitting has good properties for all regimes from compressible to incompressible. We have proved that our resulting AP scheme is asymptotically consistent, it degenerates into a consistent discretization of the incompressible system when the Mach number is sufficiently small. We have performed a Fourier stability analysis showing that our scheme is linearly L^2 stable under a C.F.L. condition independent of the Mach number. Furthermore, we have constructed an accurate TVD first order scheme and using a MOOD process, we preserve the low oscillatory properties of the order one scheme to the second order scheme. One dimensional and two dimensional numerical experiments supported the proposed analysis. In the future, we aim in focusing on local domain decomposition techniques using this AP scheme, the classical scheme for the compressible Euler equations and the classical scheme for the incompressible Euler equations.

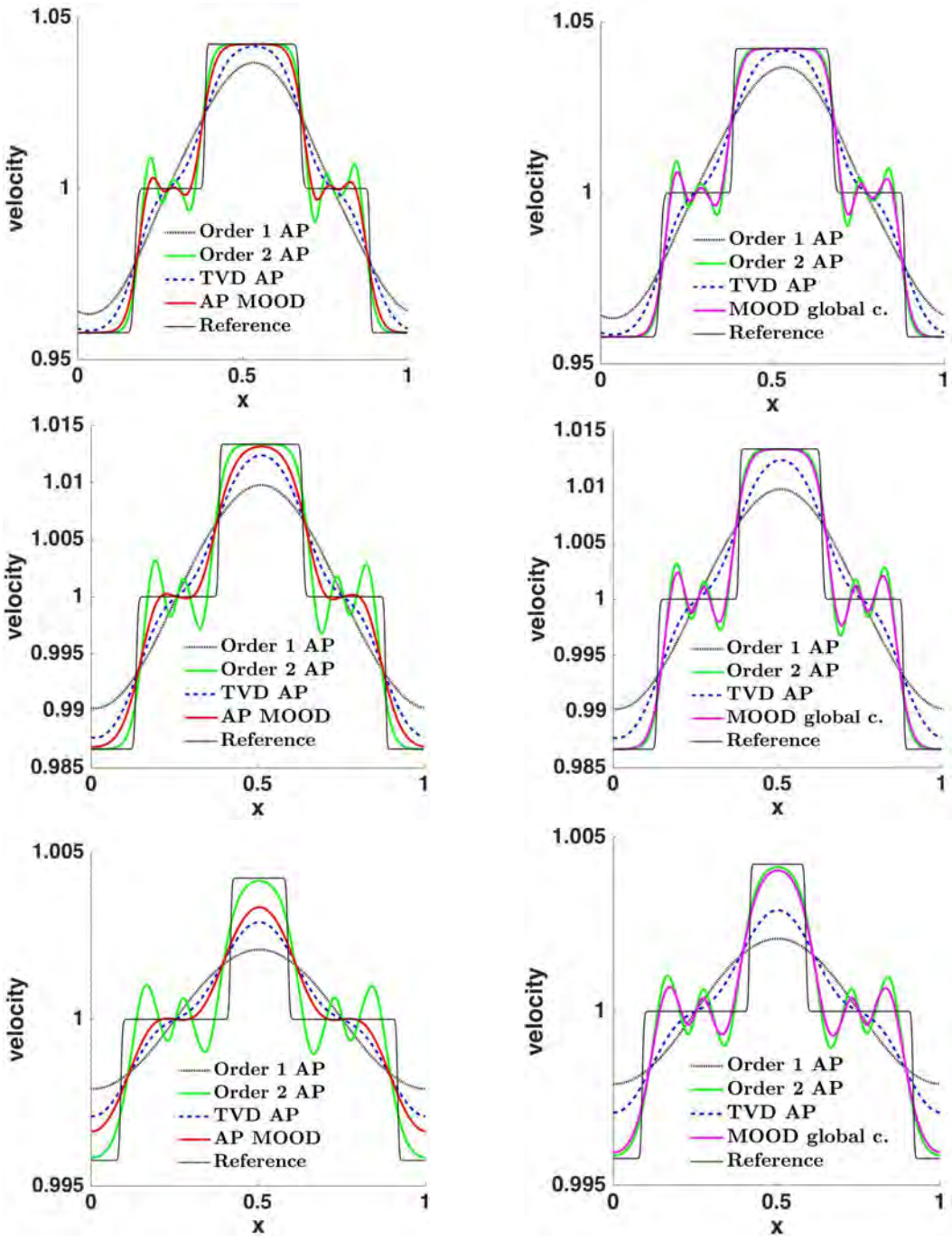


Figure 2.29 – Comparison of the local detection criterion (2.70) (red line on the left) and the global one (2.71) (magenta line on the right) against the reference solution (black solid line). Velocity profile for the shock tube problem (2.72) where $\varepsilon = 10^{-2}$ (top), $\varepsilon = 10^{-3}$ (middle) and $\varepsilon = 10^{-4}$ (bottom).

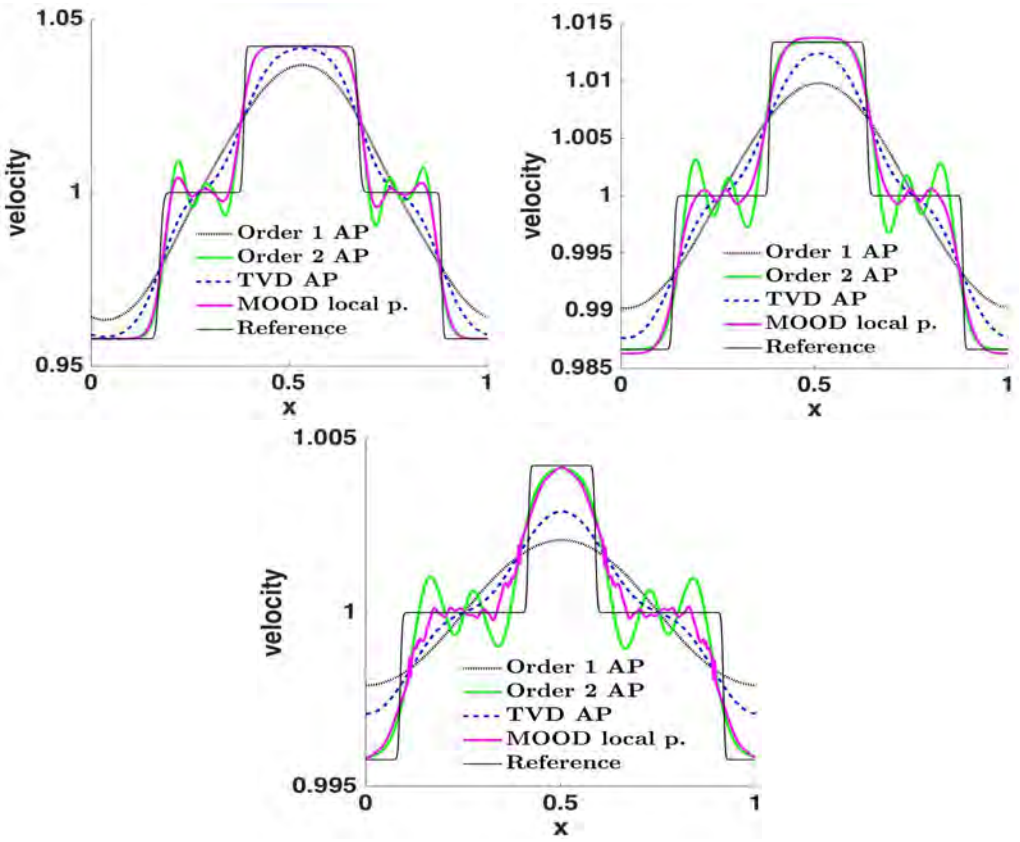
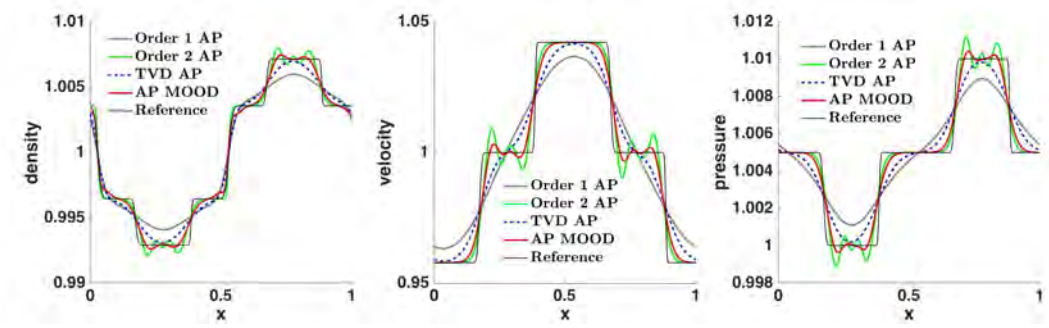
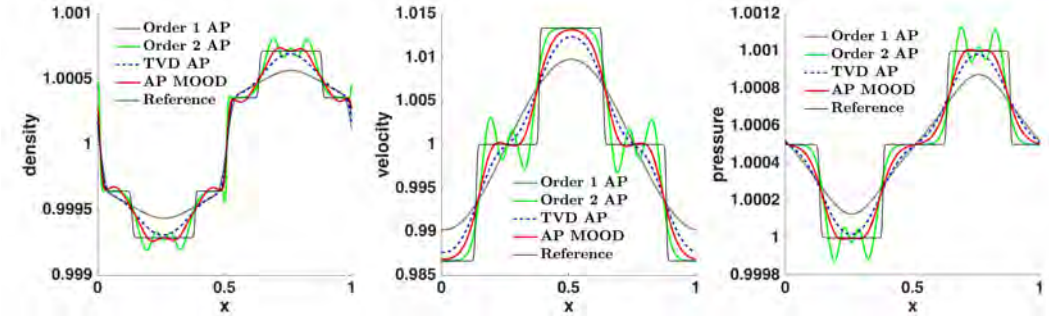


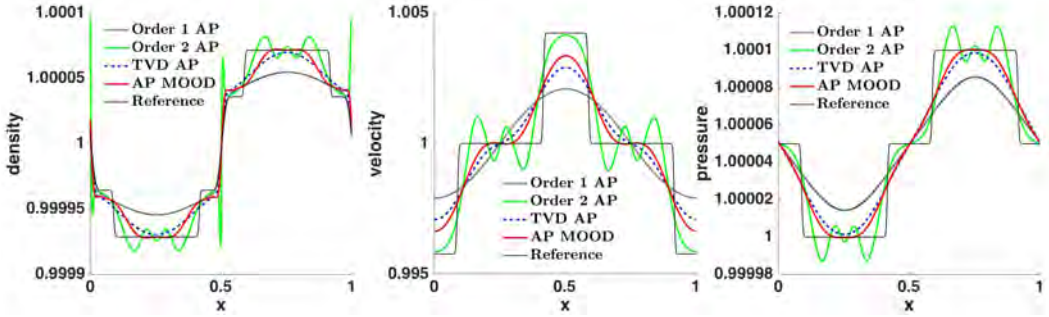
Figure 2.30 – Results when applying a local procedure: replacing Step 3 of Algorithm 1 (magenta line). Velocity profile for the shock tube problem (2.72) where $\varepsilon = 10^{-2}$ (top left), $\varepsilon = 10^{-3}$ (top right) and $\varepsilon = 10^{-4}$ (bottom).



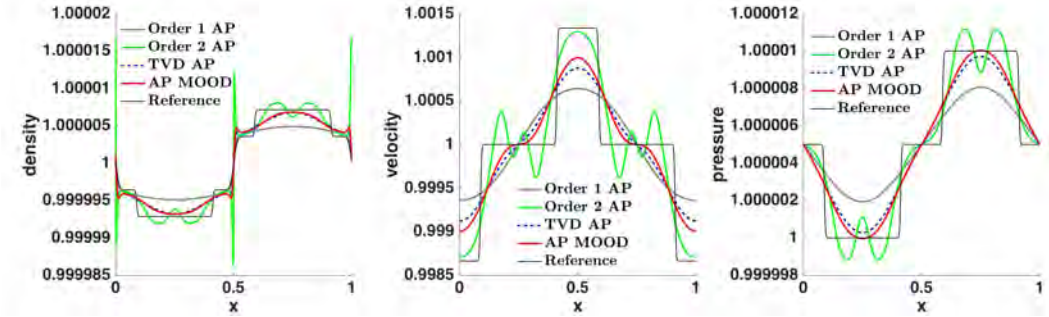
(a) Shock tube problem for $\varepsilon = 10^{-2}$ at $t_{final} = 0.03$ for 500 cells.



(b) Shock tube problem for $\varepsilon = 10^{-3}$ at $t_{final} = 0.01$ for 1000 cells.

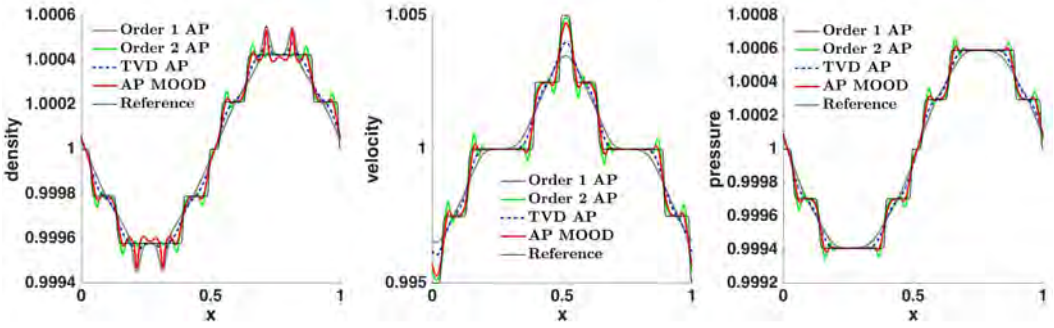


(c) Shock tube problem for $\varepsilon = 10^{-4}$ at $t_{final} = 0.0035$ for 2000 cells.

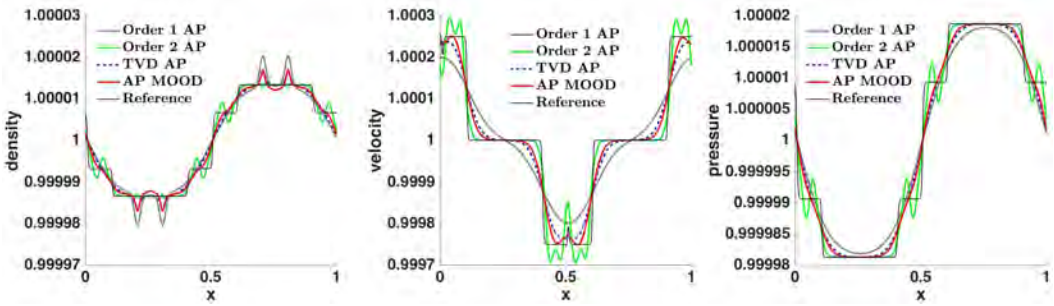


(d) Shock tube problem for $\varepsilon = 10^{-5}$ at $t_{final} = 0.0011$ for 5000 cells.

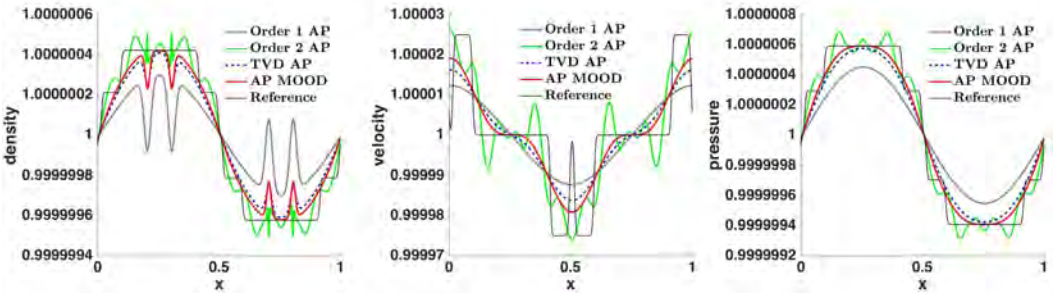
Figure 2.31 – Shock tube problem (2.72): Comparison of the first-order AP scheme (black line), the second-order AP scheme (green line), the TVD AP scheme (blue line) and of the AP MOOD scheme fixing the tolerance $\mu_{tol} = 1.4 \times 10^{-1}$ (red line) against the reference solution (black solid line).



(a) Results for $\varepsilon = 10^{-2}$ at $t_{final} = 0.015$ for 1000 cells.



(b) Results for $\varepsilon = 10^{-3}$ at $t_{final} = 0.01$ for 2000 cells.



(c) Results for $\varepsilon = 10^{-4}$ at $t_{final} = 0.006$ for 3500 cells.

Figure 2.32 – Several interacting Riemann problems (2.73): Comparison of the first-order AP scheme (black line), the second-order AP scheme (green line), the TVD AP scheme (blue line) and of the AP-MOOD scheme fixing the tolerance $\mu_{tol} = 5 \times 10^{-2}$ (red line) against the reference solution (black solid line).

Chapter 3

Extension to the Navier-Stokes equations and two dimensional numerical tests

The content of this chapter is the subject of an article in preparation in collaboration with Marie-Hélène Vignal.

Contents

3.1 Introduction

We consider the modeling of a compressible fluid described by the Navier-Stokes equations and we are interested in numerical methods valid in all Mach regimes.

Let $\Omega \subset \mathbb{R}^d$ ($d = 1$, 2 or 3) be an open bounded domain, the Navier-Stokes equations in rescaled variables are given by

$$\partial_t \rho + \nabla \cdot q = 0, \tag{3.1a}$$

$$\partial_t q + \nabla \cdot \left(\frac{q \otimes q}{\rho} \right) + \frac{1}{\varepsilon} \nabla p = \nabla \cdot \sigma, \tag{3.1b}$$

$$\partial_t E + \nabla \cdot \left((E + p) \, \frac{q}{\rho} \right) = \varepsilon \nabla \cdot \left(\sigma \frac{q}{\rho} \right) + \nabla \cdot (\lambda \nabla T) \,, \tag{3.1c}$$

with $\rho > 0$ the density of the fluid, $q = \rho U$ its momentum, U its velocity field, E its total energy, p its pressure given by an equation of state, here that of perfect gases:

$$E = \frac{p}{\gamma - 1} + \frac{\varepsilon}{2} \frac{|q|^2}{\rho}, \tag{3.1d}$$

with $\gamma = c_p / c_v > 1$ the ratio of specific heats.
The rescaled parameter ε is related to the Mach number

$$M^2 = \frac{U_0^2}{c_0^2} = \frac{\varepsilon}{\gamma},$$

with $c_0^2 = \gamma \, p_0 / \rho_0$. U_0, p_0 and ρ_0 being the typical values of the velocities, pressure and density in the fluid.
On the right hand side of the system, the diffusion terms are reported. We have in the momentum equation the viscous stress tensor related to the derivatives of the velocity field U and given by

$$\sigma = \mu \left(\nabla U + (\nabla U)^T \right) - \frac{2}{3} \mu \, (\nabla \cdot U) \, I, \tag{3.2}$$

with I the identity matrix and μ the dynamical viscosity coefficient. While in the energy equation there are the viscous work $\sigma q / \rho$ and the conductive heat flux $\lambda \nabla T$ where T is the fluid temperature and λ is the given thermal conductivity coefficient. Finally, for a perfect gas the temperature is given by the relation

$$p = R \rho T, \tag{3.3}$$

with $R = c_p - c_v$ the specific gas constant.

Remark 8. We have different relations between the parameters of the fluid mainly $\gamma = c_p / c_v$ with c_p the specific heat at constant pressure and c_v the specific heat at constant volume. Moreover, since $R = c_p - c_v$ with $c_p > c_v > 0$, it gives the relations $c_v = R/(\gamma - 1)$ and $c_p = \gamma R/(\gamma - 1)$. At last, sometimes the fluid can also be caracterized by the Prandtl number defined as the ratio of momentum diffusivity to thermal diffusivity, $Pr = \mu c_p / \lambda = \mu \gamma c_v / \lambda$.

The previous system can be rewritten in compact form as

$$\partial_t W + \nabla \cdot F(W) = \nabla \cdot G(W, \nabla W), \tag{3.4}$$

where $W = (\rho, q, E)$ is the vector of conserved variables,

$$F(W) = \begin{pmatrix} q \\ \rho U \otimes U + \dfrac{1}{\varepsilon}\, p\, Id^3_{\mathbb{R}} \\ (E+p)\, U \end{pmatrix},$$

is the inviscid flux for the Euler equations and

$$G(W, \nabla W) = \begin{pmatrix} 0 \\ \sigma \\ \varepsilon\, \sigma\, U + \lambda \nabla T \end{pmatrix},$$

is the diffusion flux.
In low Mach number regimes, the typical sound speed in the fluid, c_0, is very large compared to the typical speed of the fluid itself, U_0, and so ε is very small. In such situations, if an explicit scheme is used, the time step must satisfy a severe C.F.L. (Courant-Friedrichs-Levy) stability condition due to the acoustic waves. Indeed, for $d = 1$, the eigenvalues of the Jacobian matrix, $DF(W)$, are given by

$$\lambda_1 = u - \frac{c}{\sqrt{\varepsilon}}, \quad \lambda_2 = u, \quad \lambda_3 = u + \frac{c}{\sqrt{\varepsilon}},$$

with u the fluid velocity and $c^2 = \gamma\, p/\rho$.
For $d = 1$, the diffusion flux may be rewritten with the help of the diffusion matrix $D(W)$ as $G(W, \partial_x W) = D(W)\partial_x W$ (see [41]). Using the relation $\frac{\lambda(\gamma-1)}{R} = \frac{\lambda}{c_v}$, the diffusion matrix reads

$$D(W) = \begin{pmatrix} 0 & 0 & 0 \\ -\dfrac{4\mu u}{3\rho} & \dfrac{4}{3\rho}\mu & 0 \\ -\dfrac{4\mu\varepsilon u^2}{3\rho} + \dfrac{\lambda}{c_v}\left(\dfrac{u^2}{\rho} - \dfrac{E}{\rho^2}\right) & \dfrac{\varepsilon u}{\rho}\left(\dfrac{4\mu}{3} - \dfrac{\lambda}{c_v}\right) & \dfrac{\lambda}{c_v \rho} \end{pmatrix},$$

and its eigenvalues are

$$\lambda_{\nu 1} = 0, \quad \lambda_{\nu 2} = \frac{4\mu}{3\rho}, \quad \lambda_{\nu 3} = \frac{\lambda}{c_v \rho}.$$

Then, the C.F.L. condition ensuring the stability of explicit schemes, for the time and space steps Δt and Δx is given by

$$\Delta t \leq \frac{1}{\dfrac{\max\left(|u \pm \dfrac{c}{\sqrt{\varepsilon}}|\right)}{\Delta x} + 2\dfrac{\max\left(\dfrac{4\mu}{3\rho}, \dfrac{\lambda}{c_v \rho}\right)}{\Delta x^2}}. \tag{3.5}$$

Then, for a given space step Δx, the time step Δt is of order $\sqrt{\varepsilon}$ and tends to 0 with ε. This constraint is still present even when the diffusive part is implicit. Furthermore, even if this constraint is satisfied, it is well known (see [43], [42] or [27]) that explicit schemes suffer from a consistency problem in the limit $\varepsilon \to 0$. Indeed, they are not capable to capture the right asymptotic regime. A possible way to bypass these limitations is to use when ε is sufficiently small, the limit model obtained as the low Mach number limit of the compressible Navier-Stokes equations (3.1). Let us recall the formal low Mach number limit in the next section.

3.1.1 Low Mach number limit of the Navier-Stokes equations

The rigorous low Mach number limit of the isentropic Navier-Stokes equations has been well investigated in the last years [23, 28, 58, 71]. The non-isentropic case is more complicated and results in the case of the Euler equations can be found in [60, 1, 59]. For the rigorous low Mach number limit for the non-isentropic compressible Navier-Stokes equations one can refer to [2] for the whole space domain and [35] on a bounded domain. Here, we recall how to recover formally the limit model. We denote by $(\rho^\varepsilon, q^\varepsilon, E^\varepsilon, p^\varepsilon, T^\varepsilon)$ the solution of (3.1) and assume that it converges towards $(\rho^0, q^0, E^0, p^0, T^0)$ when ε tends to 0. In order to pass to the limit we consider the slip boundary condition

$$U^\varepsilon \cdot n = 0,$$

and the Neumann boundary condition

$$\frac{\partial T^\varepsilon}{\partial n} = 0,$$

on $\partial\Omega$ where n is the unit normal to $\partial\Omega$ outward to Ω. Inserting the asymptotic expansion of $(\rho^\varepsilon, q^\varepsilon, E^\varepsilon, p^\varepsilon, T^\varepsilon)$, we get:

$$\varepsilon^{-1}: \quad \nabla p^0 = 0, \tag{3.6a}$$

$$\varepsilon^{0}: \quad \partial_t \rho^0 + \nabla \cdot q^0 = 0, \tag{3.6b}$$

$$\partial_t q^0 + \nabla \cdot \left(\frac{q^0 \otimes q^0}{\rho^0} \right) + \nabla p^1 = \nabla \cdot \sigma^0, \tag{3.6c}$$

$$\partial_t E^0 + \nabla \cdot \left((E^0 + p^0) \, \frac{q^0}{\rho^0} \right) = \nabla \cdot \left(\lambda \nabla T^0 \right), \tag{3.6d}$$

$$E^0 = \frac{p^0}{\gamma - 1}, \tag{3.6e}$$

where $p^1(x,t) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}(p^\varepsilon(x,t) - p^0)$ is the order one correction of the pressure and

$$\sigma^0 = \mu \left(\nabla U^0 + (\nabla U^0)^T \right) - \frac{2}{3} \mu \left(\nabla \cdot U^0 \right) I,$$

with $U^0 = q^0/\rho^0$. Since $\nabla p^0(x,t) = 0$, $p^0(x,t) = p^0(t)$ then for all $x \in \Omega$ and $t > 0$

$$E^0(x,t) = E^0(t) = \frac{p^0(t)}{\gamma - 1}.$$

Integrating now (3.6d) on $[0,T] \times \Omega$ we obtain,

$$\int_\Omega \left(E^0(t) - E^0(x,0)\right) dx \; + \; \int_0^t \left(\int_{\partial\Omega} (E^0(t) + p^0(t)\,)\, u^0(x,t) \cdot \nu(x)\, d\sigma(x) \right) dt \\ - \lambda \int_0^t \left(\int_{\partial\Omega} \frac{\partial T^0(x,t)}{\partial n(x)}\, d\sigma(x) \right) dt = 0. \quad (3.7)$$

Using the boundary conditions we obtain for all $t > 0$,

$$|\Omega|\, E^0(t) - \int_\Omega E^0(x,0) dx = 0.$$

And so,

$$E^0(t) = E^0 = \frac{1}{|\Omega|} \int_\Omega E^0(x,0)\, dx.$$

Thereafter, having constant energy and pressure and $T^0 = \dfrac{p^0}{R\rho^0}$, we obtain from the energy equation (3.6d):

$$\left(E^0 + p^0\right) \nabla \cdot U^0 = \frac{\gamma\, p^0}{\gamma - 1} \nabla \cdot U^0 = p^0 \nabla \cdot \left(\lambda \nabla \left(\frac{1}{R\rho^0} \right) \right),$$

and so,

$$\gamma \nabla \cdot\; U^0 = (\gamma - 1)\, \nabla \cdot \left(\frac{\lambda}{R} \nabla \left(\frac{1}{\rho^0} \right) \right).$$

Finally, the low Mach number limit system [2] reads

$$\partial_t \rho^0 + \nabla \cdot q^0 = 0, \quad (3.8a)$$

$$\partial_t q^0 + \nabla \cdot \left(\frac{q^0 \otimes q^0}{\rho^0} \right) + \nabla p^1 = \nabla \cdot \sigma^0, \quad (3.8b)$$

$$\gamma \nabla \cdot U^0 = (\gamma - 1) \nabla \cdot \left(\frac{\lambda}{R} \nabla \left(\frac{1}{\rho^0} \right) \right), \quad (3.8c)$$

$$E^0 = \frac{p^0}{\gamma - 1} = \frac{1}{|\Omega|} \int_\Omega E^0(x,0)\, dx. \quad (3.8d)$$

Let us remark that unlike for the full Euler equations, we do not have the incompressibility constraint $\nabla \cdot U^0 = 0$. This is due to the combined effects of large temperature variations and thermal conduction. Note that neglecting the heat conduction effects, i.e., setting $\lambda = 0$, we do obtain the Navier-Stokes incompressible model. Moreover, neglecting also the viscous forces, i.e., setting $\mu = 0$, the low Mach number limit for the compressible Euler equations is recovered (see Section 2.1.1).

3.1.2 Asymptotic preserving schemes

First, let us clarify that the rescaled Navier-Stokes system (3.1) is used for the analysis. But, in practice, simulations are performed on the non rescaled Navier-Stokes system (1.7). In the rescaled system, our parameter ε is constant. In practice, it is like if we considered ε not constant and varying in space and time.
The limit model (3.8) does no longer depend on the Mach number and so is no more constrained by the small values of ε. But, it can be used only where ε is sufficiently small. Where ε takes on order one or intermediate values, the Navier-Stokes equations (3.1) must be used. Then, two models must be used which leads to other difficulties like the detection of the interface between the two models, the reconnection at the interface... (see Section 2.1.3 for more details). A solution can be to construct an asymptotic preserving scheme which is consistent with the limit and free of the constraints related to the Mach number ε. Such schemes have been developed in the literature, see [25, 24, 26, 44, 75, 64, 9, 30, 10, 29] for the isentropic Euler system, [63, 21, 65, 37, 16, 30, 11] for the full Euler system or [25, 21, 32, 37, 76, 14, 12] for the full Navier-Stokes system. They permit to avoid the time step limitations, the schemes are said to be asymptotically stable. And, they lead to consistent approximations of the limit incompressible model when the low Mach number goes to zero, this corresponds to the asymptotic consistency property. In practice, one strategy among others for constructing asymptotic preserving schemes is using IMEX methods [80, 66]. The flux F is split into two parts,

$$F = F_e + F_i.$$

The first part F_e will be treated explicitly while the second one, F_i, implicitly. The choice of the decomposition $F = F_e + F_i$ must be well chosen in order to obtain asymptotic stability, asymptotic consistency and with a reasonable computational cost. In Chapter 2, we defined, for the full Euler equations, criteria in order to choose correctly the right flux decomposition respecting the above properties. For the Navier-Stokes equations, we propose to use the same decomposition of F for the hyperbolic part, the one introduced by E.F. Toro and M.E. Vázquez-Cendón in [38] (see Chapter 2). With this choice, in dimension 1 the eigenvalues of the Jacobian matrix related to F_e are given by $\lambda_1 = \lambda_2 = u$ the fluid velocity and $\lambda_3 = 0$. Moreover, in the case of viscous flows, the time step must obey a quadratic restriction proportional to $\Delta x^2 / \max(\mu, \lambda)$, and this condition on the maximum time step can become rather severe for large values of μ or λ. To also overcome those restrictions, we propose to treat implicitly the diffusion flux. Since G contains only linear terms we consider that this choice of treatment adds a reasonable computational cost and allows us to use larger time steps in strongly viscous regimes as well. Then, for the chosen flux decomposition, the C.F.L. condition of such IMEX scheme is given by

$$\Delta t \leq \frac{\Delta x}{|u|}.$$

In the next section, we present the proposed order 1 IMEX scheme based on a finite volume discretization in space. Following the same strategy as for the full Euler

equations, the system is reformulated to solve a nonlinear equation on the pressure for which a linearization is proposed. This choice allows to avoid the use of an iterative method for solving it and the resolution is therefore uncoupled. Then, the asymptotic consistency is proved on the semi-discretization. Furthermore, we propose a second-order extension using an IMEX Runge-Kutta approach in time. Then, we show the good behavior of the order 2 scheme on a variety of two dimensional numerical tests involving non-viscous fluids, viscous fluids, fluids in presence of heat conduction. They take into account fluids in the compressible and low Mach number regimes.

3.2 Order 1 AP scheme

3.2.1 Semi-discretization in time

3.2.1.1 Order 1 discretization

For clarity, we consider a uniform discretization in time denoting by Δt the time step. Using for F the flux decomposition proposed in [38] and treating the diffusion flux implicitly we rewrite the system in a conservative system form

$$\frac{W^{n+1}-W^n}{\Delta t}+\nabla\cdot F_e(W^n)+\nabla\cdot F_i(W^{n+1})=\nabla\cdot G\left(W^{n+1},\nabla W^{n+1}\right). \tag{3.9a}$$

where the explicit and implicit inviscid fluxes are given by

$$F_e(W)=\begin{pmatrix}\rho U\\ \rho U\otimes U\\ k_\varepsilon U\end{pmatrix},\quad F_i(W)=\begin{pmatrix}0\\ \dfrac{p}{\varepsilon}I_d\\ h\,U\end{pmatrix},$$

with $k_\varepsilon=k_\varepsilon(W)=\varepsilon\rho|U|^2/2$ the kinetic energy and $h=h(W)=\gamma(E-k_\varepsilon)$ the specific enthalpy and where the implicit diffusion flux is given by

$$G(W,\nabla W)=\begin{pmatrix}0\\ \sigma\\ \varepsilon\,\sigma\,U+\lambda\nabla T\end{pmatrix}=\begin{pmatrix}0\\ \sigma\\ \varepsilon\,\sigma\,U+\dfrac{\lambda}{R}\nabla\left(\dfrac{p}{\rho}\right)\end{pmatrix}.$$

Then, we have

$$\rho^{n+1}=\rho^n-\Delta t\nabla\cdot(\rho^nU^n), \tag{3.10a}$$

$$q^{n+1}=q^n-\Delta t\nabla\cdot(\rho^nU^n\otimes U^n)-\Delta t\frac{1}{\varepsilon}\nabla p^{n+1}+\Delta t\nabla\cdot\sigma^{n+1}, \tag{3.10b}$$

$$E^{n+1}=E^n-\Delta t\nabla\cdot(k_\varepsilon^nU^n)-\Delta t\nabla\cdot\left(h^{n+1}U^{n+1}\right)+\varepsilon\Delta t\nabla\cdot\left(\sigma^{n+1}U^{n+1}\right)+\frac{\Delta t\lambda}{R}\Delta\left(\frac{p^{n+1}}{\rho^{n+1}}\right), \tag{3.10c}$$

$$E^{n+1}=\frac{p^{n+1}}{\gamma-1}+k_\varepsilon^{n+1}, \tag{3.10d}$$

where $k_\varepsilon^{n+1} = k_\varepsilon(W^{n+1})$ and $h^{n+1} = h(W^{n+1})$.
Inserting the momentum equation (3.10b) into the implicit inviscid flux for the energy equation (3.10c), expressing E^{n+1} with the state equation we have

$$p^{n+1} + k_\varepsilon^{n+1} = E^{n+1,exp} - \Delta t \nabla \cdot \left(\frac{h^{n+1}}{\rho^{n+1}} \left(q^{n+1,exp} - \frac{\Delta t}{\varepsilon} \nabla p^{n+1} + \Delta t \nabla \cdot \sigma^{n+1} \right) \right) + \varepsilon \Delta t \nabla \cdot \left(\sigma^{n+1} \frac{q^{n+1}}{\rho^{n+1}} \right) + \frac{\Delta t \lambda}{R} \Delta \left(\frac{p^{n+1}}{\rho^{n+1}} \right), \quad (3.11)$$

where $W^{n+1,exp}$ is the explicit convected part of the conservative variables,

$$W^{n+1,exp} = \begin{pmatrix} \rho^{n+1,exp} \\ q^{n+1,exp} \\ E^{n+1,exp} \end{pmatrix} = W^n - \Delta t \nabla \cdot F_e(W^n) = \begin{pmatrix} \rho^n - \Delta t \nabla \cdot (\rho^n U^n) \\ q^n - \Delta t \nabla \cdot (\rho^n U^n \otimes U^n) \\ E^n - \Delta t \nabla \cdot (k_\varepsilon^n U^n) \end{pmatrix}. \quad (3.12)$$

Then, by passing the terms p^{n+1} to the left hand side and multiplying by ε, one obtains the discretization of an elliptic equation for determining the unknown pressure p^{n+1}:

$$\frac{\varepsilon}{\gamma - 1} p^{n+1} - \Delta t^2 \nabla \cdot \left(\frac{h^{n+1}}{\rho^{n+1}} \nabla p^{n+1} \right) - \frac{\varepsilon \Delta t \lambda}{R} \Delta \left(\frac{p^{n+1}}{\rho^{n+1}} \right) = -\varepsilon k_\varepsilon^{n+1} + \varepsilon E^{n+1,exp} - \varepsilon \Delta t \nabla \cdot \left(\frac{h^{n+1}}{\rho^{n+1}} \left(q^{n+1,exp} + \Delta t \nabla \cdot \sigma^{n+1} \right) \right) + \varepsilon^2 \Delta t \nabla \cdot \left(\sigma^{n+1} \frac{q^{n+1}}{\rho^{n+1}} \right). \quad (3.13)$$

The equation on the unknown pressure is highly nonlinear, it is completely coupled with the momentum equation. Within the framework of finite volume schemes on collocated grids, the scheme proposed in [11], for the full Euler case, i.e., for $\lambda = \mu = 0$, consists in solving (3.13) with the use of a Picard algorithm. In [14], a linearization of the kinetic term k_ε is proposed for the state equation and the enthalpy h is treated explicitly resulting on a linear version for the pressure equation (3.13). The algorithm is also extended to the Navier-Stokes model discretizing the diffusive terms explicitly. In [12], following the strategy proposed in [11], it is extended to the Navier-Stokes equations considering as here, the diffusion flux term implicit. The nonlinear pressure equation (3.13) is solved using two Picard iterates. A first loop is used to update the viscous terms and then in a second loop the new viscous values are used to solve the pressure equation. Here, we propose a linearization of the reformulated pressure equation (3.13) in order to simplify the scheme and reduce the computational cost. Following the same strategy as for the full Euler equations (see Section 2.3) for the inviscid part, the quantities h^{n+1} and k_ε^{n+1} are approximated in the pressure equation by their explicit convected values $k_\varepsilon^{n+1,exp} = k_\varepsilon(W^{n+1,exp})$ and $h^{n+1,exp} = h(W^{n+1,exp})$. We propose to also approximate in the

pressure equation the implicit viscous terms by their explicit convected values setting

$$\sigma^{n+1} = \sigma(W^{n+1,exp}), \quad \left(\sigma\frac{q}{\rho}\right)^{n+1} = \left(\sigma(W^{n+1,exp})\frac{q^{n+1,exp}}{\rho^{n+1}}\right). \tag{3.14}$$

We will see that the C.F.L. condition is not impacted by this strategy. Doing so, we obtain an elliptic linear equation for p^{n+1}. After expressing in the energy equation the temperature with the state equation $T = p/(R\rho)$, the resulting semi-discretization reads

$$q^{n+1,exp} = q^n - \Delta t\nabla\cdot(\rho^n U^n\otimes U^n), \tag{3.15a}$$
$$E^{n+1,exp} = E^n - \Delta t\,\nabla\cdot(k_\varepsilon^n\,U^n), \tag{3.15b}$$
$$\rho^{n+1} = \rho^n - \Delta t\nabla\cdot(\rho^n U^n), \tag{3.15c}$$
$$\begin{aligned}\frac{\varepsilon}{\gamma-1}p^{n+1} - \Delta t^2\nabla\cdot\left(\frac{h^{n+1,exp}}{\rho^{n+1}}\nabla p^{n+1}\right) - \frac{\varepsilon\Delta t\lambda}{R}\Delta\left(\frac{p^{n+1}}{\rho^{n+1}}\right) &= -\varepsilon k_\varepsilon^{n+1,exp}\\ + \varepsilon E^{n+1,exp} - \varepsilon\Delta t\nabla\cdot\left(\frac{h^{n+1,exp}}{\rho^{n+1}}\left(q^{n+1,exp}+\Delta t\nabla\cdot\sigma^{n+1,exp}\right)\right)&\\ + \varepsilon^2\Delta t\nabla\cdot\left(\sigma^{n+1,exp}\frac{q^{n+1,exp}}{\rho^{n+1}}\right),&\end{aligned} \tag{3.15d}$$
$$q^{n+1} - \Delta t\nabla\cdot\sigma^{n+1} = q^{n+1,exp} - \Delta t\frac{1}{\varepsilon}\nabla p^{n+1}, \tag{3.15e}$$
$$\begin{aligned}E^{n+1} = E^{n+1,exp} - \Delta t\nabla\cdot\left(\frac{\gamma p^{n+1}}{(\gamma-1)\rho^{n+1}}q^{n+1}\right) + \varepsilon\Delta t\nabla\cdot\left(\sigma^{n+1}\frac{q^{n+1}}{\rho^{n+1}}\right)&\\ + \frac{\Delta t\lambda}{R}\Delta\left(\frac{p^{n+1}}{\rho^{n+1}}\right).&\end{aligned} \tag{3.15f}$$

The scheme consists in computing sequentially ρ^{n+1} with (3.15c) that is given explicitly, p^{n+1} with (3.15d) solving a linear system, q^{n+1} with (3.15e) solving again a linear system. Then σ^{n+1} is calculated to update E^{n+1} explicitly with (3.15f). Note that to compute q^{n+1} we need to solve a linear system since the viscous forces are taken implicitly. Moreover, the momentum equations are coupled through σ^{n+1}, so they are not solved independently. This implicit treatment allows us to have a condition on the time step Δt that is not constrained.

3.2.1.2 Asymptotic consistency

Lemma 3.2.1. *Assuming the boundary conditions $U.n = 0$ and $\partial T/\partial n = 0$ on $\partial\Omega$, the semi-discretization (3.15) gives $p_0^{n+1} = (\gamma-1) < E_0^0 > + \mathcal{O}(\Delta t^2)$ and $< E_0^{n+1} >=< E_0^0 >$ for all $n \geq 0$. Furthermore, if the initial energy is well-prepared to the low Mach number regime, more precisely if $\lim_{\varepsilon\to 0} E(x,0) = \overline{E}_0$ with $\overline{E}_0$ constant, the semi-discretization (3.15) is asymptotically consistent up to*

order $\mathcal{O}(\Delta t)$. The formal low Mach number limit of the system gives for all $n \geq 0$

$$\begin{aligned}
&\nabla p_0^{n+1} = 0,\\
&E_0^{n+1} = \frac{p_0^{n+1}}{\gamma - 1} + \mathcal{O}(\Delta t^2) = \overline{E}_0 + \mathcal{O}(\Delta t^2),\\
&\gamma \nabla \cdot U_0^{n+1} = (\gamma - 1)\frac{\lambda}{R}\Delta\left(\frac{1}{\rho_0^{n+1}}\right) + \mathcal{O}(\Delta t).
\end{aligned}$$

Where for all functions f, $f_0 = \lim_{\varepsilon \to 0} f$.

Proof. Let us prove the asymptotic consistency. Reformulating the pressure equation (3.15d), we obtain

$$\begin{aligned}
\frac{\varepsilon}{\gamma - 1} p^{n+1} = &-\varepsilon k_\varepsilon^{n+1,exp} + \varepsilon E^{n+1,exp} + \frac{\varepsilon \Delta t \lambda}{R}\Delta\left(\frac{p^{n+1}}{\rho^{n+1}}\right)\\
&- \varepsilon \Delta t \nabla \cdot \left(\frac{h^{n+1,exp}}{\rho^{n+1}}\left(q^{n+1} + \Delta t \nabla \cdot \left(\sigma^{n+1,exp} - \sigma^{n+1}\right)\right)\right)\\
&+ \varepsilon^2 \Delta t \nabla \cdot \left(\sigma^{n+1,exp}\frac{q^{n+1,exp}}{\rho^{n+1}}\right). \quad (3.16)
\end{aligned}$$

Then, we perform an asymptotic expansion, assuming that all quantities $f^l = f_0^l + \varepsilon f_1^l$ with $l = n$, $(n+1, exp)$, $n+1$ we obtain:

$$\begin{aligned}
\varepsilon^{-1}: \quad & \nabla \cdot \left(\frac{h_0^{n+1,exp}}{\rho_0^{n+1}}\nabla p_0^{n+1}\right) = 0, && (3.17a)\\
& \nabla p_0^{n+1} = 0, && (3.17b)\\
\varepsilon^{0}: \quad & q_0^{n+1,exp} = q_0^n - \Delta t \nabla \cdot (\rho_0^n U_0^n \otimes U_0^n), && (3.17c)\\
& E_0^{n+1,exp} = E_0^n, && (3.17d)\\
& \rho_0^{n+1} = \rho_0^n - \Delta t \nabla \cdot (\rho_0^n U_0^n) && (3.17e)\\
& \frac{p_0^{n+1}}{\gamma - 1} = E_0^n + \frac{\Delta t \lambda p_0^{n+1}}{R}\Delta\left(\frac{1}{\rho_0^{n+1}}\right) - \Delta t \nabla \cdot \left(\gamma\, E_0^n\, U_0^{n+1}\right) && (3.17f)\\
& \qquad - \Delta t^2 \nabla \cdot \left(\frac{\gamma\, E_0^n}{\rho_0^{n+1}}\nabla \cdot \left(\sigma_0^{n+1,exp} - \sigma_0^{n+1}\right)\right), && (3.17g)\\
& q_0^{n+1} = q_0^{n+1,exp} - \Delta t \nabla p_1^{n+1} + \Delta t \nabla \cdot \sigma_0^{n+1}, && (3.17h)\\
& E_0^{n+1} = E_0^n - \Delta t \frac{\gamma p_0^{n+1}}{\gamma - 1}\nabla \cdot U_0^{n+1} + \frac{\Delta t \lambda p_0^{n+1}}{R}\Delta\left(\frac{1}{\rho_0^{n+1}}\right). && (3.17i)
\end{aligned}$$

where $h_0^{n+1,exp} = \lim_{\varepsilon \to 0} \gamma \left(E^{n+1,exp} - k_\varepsilon^{n+1,exp}\right) = \gamma\, E_0^n$.

Integrating the pressure equation (3.17g) on Ω we obtain:

$$|\Omega|\frac{p_0^{n+1}}{\gamma-1} = \int_\Omega E_0^n(x)dx + \Delta t\lambda \int_{\partial\Omega} \frac{\partial T_0^{n+1}(x)}{\partial n(x)}\, d\sigma(x) - \Delta t \int_{\partial\Omega} \gamma E_0^n(x) U_0^{n+1}(x)\cdot n(x)\, d\sigma(x) - \Delta t^2 \int_{\partial\Omega} \frac{\gamma E_0^n(x)}{\rho_0^{n+1}(x)} \left(\nabla\cdot\left(\sigma_0^{n+1,exp} - \sigma_0^{n+1}\right)\right)\cdot n(x)\, d\sigma(x),$$

where we recall that $T_0^{n+1} = p_0^{n+1}/(R\rho_0^{n+1})$.
And so, using the boundary conditions $U.n = 0$ and $\dfrac{\partial T}{\partial n} = 0$, we get:

$$\frac{p_0^{n+1}}{\gamma-1} = \frac{1}{|\Omega|}\int_\Omega E_0^n(x)dx - \frac{\Delta t^2}{|\Omega|}\int_{\partial\Omega} \frac{\gamma E_0^n(x)}{\rho_0^{n+1}(x)} \left(\nabla\cdot\left(\sigma_0^{n+1,exp} - \sigma_0^{n+1}\right)\right)\cdot n(x)\, d\sigma(x).$$

Moreover, integrating the energy equation (3.17i) on Ω gives

$$< E_0^{n+1} >=< E_0^n >,$$

with $< E_0^n >= 1/|\Omega| \int_\Omega E_0^n(x)\, dx$ for all $n \geq 0$.
By induction, we have for all $n \geq 0$,

$$\frac{p_0^{n+1}}{\gamma-1} =< E_0^0 > + C_n\Delta t^2, \quad < E_0^{n+1} >=< E_0^0 >,$$

where C_n is a constant arising from the approximation of the viscous terms in the pressure equation. For all $n \geq 0$, it is given by

$$C_n = -\frac{\gamma}{|\Omega|}\int_{\partial\Omega} \frac{E_0^n(x)}{\rho_0^{n+1}(x)} \left(\nabla\cdot\left(\sigma_0^{n+1,exp} - \sigma_0^{n+1}\right)\right)\cdot n(x)\, d\sigma(x).$$

Furthermore, assuming the initial energy well-prepared, i.e., $E_0^0 = \overline{E}_0$ is constant, then,

$$\frac{p_0^{n+1}}{\gamma-1} = E_0^0 + C_n\Delta t^2 = \overline{E}_0 + C_n\Delta t^2.$$

Therefore, for $n = 0$ the energy equation (3.17i) gives:

$$\begin{aligned} E_0^1 &= E_0^0 - \Delta t\frac{\gamma p_0^1}{\gamma-1}\nabla\cdot U_0^1 + \frac{\Delta t\lambda p_0^1}{R}\Delta\left(\frac{1}{\rho_0^1}\right) \\ &= \overline{E}_0 - \Delta t\gamma\overline{E}_0\nabla\cdot U_0^1 - \Delta t^3\,\gamma\, C_1\nabla\cdot U_0^1 + \frac{\Delta t\lambda p_0^1}{R}\Delta\left(\frac{1}{\rho_0^1}\right). \end{aligned}$$

And since the pressure equation (3.17g) reads for $n = 0$:

$$\frac{p_0^1}{\gamma-1} = \overline{E}_0 + \frac{\Delta t\lambda p_0^1}{R}\Delta\left(\frac{1}{\rho_0^1}\right) - \Delta t\gamma\,\overline{E}_0\nabla\cdot U_0^1 - \Delta t^2\gamma\,\overline{E}_0\nabla\cdot\left(\frac{1}{\rho_0^1}\nabla\cdot\left(\sigma_0^{1,exp} - \sigma_0^1\right)\right).$$

We get,

$$\begin{aligned} E_0^1 &= \frac{p_0^1}{\gamma-1} + \Delta t^2 \gamma\, \overline{E}_0 \nabla \cdot \left(\frac{1}{\rho_0^1} \nabla \cdot \left(\sigma_0^{1,exp} - \sigma_0^1 \right) \right) - \Delta t^3\, \gamma\, C_1 \nabla \cdot U_0^1 \\ &= \frac{p_0^1}{\gamma-1} + C_1' \Delta t^2 \\ &= \overline{E}_0 + (C_1 + C_1') \Delta t^2, \end{aligned}$$

where $C_1' = \gamma\, \overline{E}_0 \nabla \cdot \left(\frac{1}{\rho_0^1} \nabla \cdot \left(\sigma_0^{1,exp} - \sigma_0^1 \right) \right) - \Delta t \gamma\, C_1 \nabla \cdot U_0^1$. Let us note that if $\mu = 0$, then $C_1 = C_1' = 0$ and so, $E_0^1 = p_0^1/(\gamma - 1)$.
By induction, for all $n \geq 0$,

$$E_0^{n+1} = \frac{p_0^{n+1}}{\gamma-1} + \mathcal{O}(\Delta t^2) = \overline{E}_0 + \mathcal{O}(\Delta t^2).$$

Thereafter, the energy equation (3.17i) yields

$$\overline{E}_0 + \mathcal{O}(\Delta t^2) = \overline{E}_0 + \mathcal{O}(\Delta t^2) - \Delta t \frac{\gamma p_0^{n+1}}{\gamma-1} \nabla \cdot U_0^{n+1} + \frac{\Delta t \lambda p_0^{n+1}}{R} \Delta \left(\frac{1}{\rho_0^{n+1}} \right),$$

and thus, we recover up to an order Δt term, a discretization of the divergence equation for the velocity field in the limit model (3.8):

$$\gamma \nabla \cdot U_0^{n+1} = (\gamma - 1) \nabla \cdot \left(\frac{\lambda}{R} \nabla \left(\frac{1}{\rho_0^{n+1}} \right) \right) + \mathcal{O}(\Delta t).$$

This ends the proof and shows that our scheme is asymptotitcally consistent with the limit model up to $\mathcal{O}(\Delta t)$ assuming $E_0^0 = \overline{E}_0$. □

Remark 9. Let us note that considering an inviscid fluid, i.e., setting $\mu = 0$, we recover exactly a discretization of the limit model: for all $n \geq 0$,

$$\begin{aligned} &\frac{p_0^{n+1}}{\gamma-1} = E_0^{n+1} = \overline{E}_0, \\ &\gamma \nabla \cdot U_0^{n+1} = (\gamma - 1) \nabla \cdot \left(\frac{\lambda}{R} \nabla \left(\frac{1}{\rho_0^{n+1}} \right) \right). \end{aligned}$$

3.2.2 Full discretization in one dimension

The discretization of the space domain follows the usual finite volume framework. The explicit flux F_e is discretized with a Rusanov-type solver (3.24) and the implicit fluxes F_i and G are discretized with a centered solver. The resulting scheme gives consistent and stable results but can present oscillations which are the signature of the non L^∞ stability. In Section 3.2.2.2, based on the results obtained for the Euler equations (see Section 2.3.4), we also propose a discretization where we add an upwinding on the implicit flux F_i in order to eliminate the oscillations.
For simplicity, we start by presenting the space discretizations in one dimension. Then we will present them in two dimensions omitting some details when the generalization is easily extended from the discretization of the one-dimensional setting.

3.2.2.1 Order 1 L^2 discretization

Let us first recall the semi-discretization (3.15) in one dimension:

$$W^{n+1,exp} = \begin{pmatrix} \rho^{n+1,exp} \\ q^{n+1,exp} \\ E^{n+1,exp} \end{pmatrix} = \begin{pmatrix} \rho^n - \Delta t \partial_x \left(\rho^n u^n\right) \\ q^n - \Delta t \partial_x \left(\rho^n (u^n)^2\right) \\ E^n - \Delta t\, \partial_x (k_\varepsilon^n\, u^n) \end{pmatrix}, \tag{3.18}$$

$$\rho^{n+1} = \rho^{n+1,exp}, \tag{3.19}$$

$$\begin{aligned} \frac{\varepsilon}{\gamma-1} p^{n+1} - \Delta t^2 \partial_x \left(\frac{h^{n+1,exp}}{\rho^{n+1}} \partial_x p^{n+1} \right) - \frac{\varepsilon \Delta t \lambda}{R} \partial_{xx}^2 \left(\frac{p^{n+1}}{\rho^{n+1}} \right) = -\varepsilon k_\varepsilon^{n+1,exp} \\ + \varepsilon E^{n+1,exp} - \varepsilon \Delta t \partial_x \left(\frac{h^{n+1,exp}}{\rho^{n+1}} \left(q^{n+1,exp} + \Delta t \partial_x \sigma^{n+1,exp} \right) \right) \\ + \varepsilon^2 \Delta t \partial_x \left(\sigma^{n+1,exp} \frac{q^{n+1,exp}}{\rho^{n+1}} \right), \end{aligned} \tag{3.20}$$

$$q^{n+1} - \Delta t \partial_x \sigma^{n+1} = q^{n+1,exp} - \Delta t \frac{1}{\varepsilon} \partial_x p^{n+1}, \tag{3.21}$$

$$\begin{aligned} E^{n+1} \ = E^{n+1,exp} - \Delta t \partial_x \left(\frac{\gamma p^{n+1}}{(\gamma-1)\rho^{n+1}} q^{n+1} \right) + \varepsilon \Delta t \partial_x \left(\sigma^{n+1} \frac{q^{n+1}}{\rho^{n+1}} \right) \\ + \frac{\Delta t \lambda}{R} \partial_{xx}^2 \left(\frac{p^{n+1}}{\rho^{n+1}} \right). \end{aligned} \tag{3.22}$$

Now, we consider a uniform discretization in space and denote by $\Delta x > 0$ the space step. The fully L^2 stable discrete version reads

$$W_j^{n+1,exp} = W_j^n - \Delta t \frac{(\mathcal{F}_e)_{j+1/2}^n - (\mathcal{F}_e)_{j-1/2}^n}{\Delta x}, \tag{3.23a}$$

$$\rho_j^{n+1} = \rho_j^{n+1,exp}, \tag{3.23b}$$

$$\begin{aligned} &\frac{\varepsilon}{\gamma-1} p_j^{n+1} - \frac{\Delta t^2}{\Delta x} \left(\left(\frac{h^{n+1,exp}}{\rho^{n+1}} \right)_{j+1/2} \frac{p_{j+1}^{n+1} - p_j^{n+1}}{\Delta x} - \left(\frac{h^{n+1,exp}}{\rho^{n+1}} \right)_{j-1/2} \frac{p_j^{n+1} - p_{j-1}^{n+1}}{\Delta x} \right) \\ &- \frac{\varepsilon \Delta t \lambda}{R} \left(\partial_{xx}^2 \left(\frac{p^{n+1}}{\rho^{n+1}} \right) \right)_j = \varepsilon \left(E_j^{n+1,exp} + \varepsilon \Delta t \frac{(\sigma u)_{j+1/2}^{n+1,exp} - (\sigma u)_{j-1/2}^{n+1,exp}}{\Delta x} - k_j^{n+1,exp} \right) \\ &\quad - \varepsilon \frac{\Delta t}{2\Delta x} \left(\frac{h_{j+1}^{n+1,exp}}{\rho_{j+1}^{n+1}} \left(q_{j+1}^{n+1,exp} + \Delta t \frac{\sigma_{j+3/2}^{n+1,exp} - \sigma_{j+1/2}^{n+1,exp}}{\Delta x} \right) \right. \\ &\quad \left. - \frac{h_{j-1}^{n+1,exp}}{\rho_{j-1}^{n+1}} \left(q_{j-1}^{n+1,exp} + \Delta t \frac{\sigma_{j-1/2}^{n+1,exp} - \sigma_{j-3/2}^{n+1,exp}}{\Delta x} \right) \right), \end{aligned} \tag{3.23c}$$

$$q_j^{n+1} - \Delta t \frac{\sigma_{j+1/2}^{n+1} - \sigma_{j-1/2}^{n+1}}{\Delta x} = q_j^{n+1,exp} - \frac{\Delta t}{\varepsilon} \frac{p_{j+1}^{n+1} - p_{j-1}^{n+1}}{2\Delta x}, \tag{3.23d}$$

$$E_j^{n+1} = E_j^{n+1,exp} + \varepsilon\Delta t \frac{(\sigma u)_{j+1/2}^{n+1} - (\sigma u)_{j-1/2}^{n+1}}{\Delta x} - \frac{\Delta t}{2\Delta x}\left(\left(\frac{\gamma p^{n+1}}{(\gamma-1)\rho^{n+1}} q^{n+1}\right)_{j+1} - \left(\frac{\gamma p^{n+1}}{(\gamma-1)\rho^{n+1}}\right)_{j-1}\right) + \frac{\Delta t\lambda}{R}\left(\partial_{xx}^2\left(\frac{p^{n+1}}{\rho^{n+1}}\right)\right)_j. \tag{3.23e}$$

The explicit numerical flux $(\mathcal{F}_e)^n = ((\mathcal{F}_{e\rho})^n, (\mathcal{F}_{eq})^n, (\mathcal{F}_{eE})^n)$ is the Rusanov solver and given by

$$(\mathcal{F}_e)_{j+1/2}^n := \frac{F_e(W_{j+1}^n) + F_e(W_j^n)}{2} - (\mathcal{D}_e)_{j+1/2}^n (W_{j+1}^n - W_j^n), \tag{3.24}$$

with $(\mathcal{D}_e)_{j+1/2}^n$ the explicit viscosity coefficient, taken as half of the maximum explicit eigenvalue of the Jacobian matrix associated to $(\mathcal{F}_e^x)^n$:

$$(\mathcal{D}_e)_{j+1/2}^n = \frac{1}{2}\max\left(|u_{j+1}^n|, |u_j^n|\right).$$

Concerning the second order derivatives for the pressure (3.23c) and energy (3.23e) equations, we set

$$\left(\partial_{xx}^2\left(\frac{p^{n+1}}{\rho^{n+1}}\right)\right)_j = \frac{\left(\frac{p^{n+1}}{\rho^{n+1}}\right)_{j+1} - 2\left(\frac{p^{n+1}}{\rho^{n+1}}\right)_j + \left(\frac{p^{n+1}}{\rho^{n+1}}\right)_{j-1}}{\Delta x^2}.$$

In (3.23c), $\left(\frac{h}{\rho}\right)_{j+1/2}$ is computed as the arithmetic average:

$$\left(\frac{h^{n+1,exp}}{\rho^{n+1}}\right)_{j+1/2} = \frac{1}{2}\left(\frac{h_j^{n+1,exp}}{\rho_j^{n+1}} + \frac{h_{j+1}^{n+1,exp}}{\rho_{j+1}^{n+1}}\right).$$

We are left with computing the viscous terms $\sigma_{j+1/2}$ and $(\sigma u)_{j+1/2}$. Let us note that in one dimension the viscous stress tensor is simply given by $\sigma = \frac{4}{3}\mu\partial_x u$ and discretized by

$$\sigma_{j+1/2} = \frac{4\mu}{3}(\partial_x u)_{j+1/2} = \frac{4\mu}{3}\frac{u_{j+1} - u_j}{\Delta x}, \qquad (\sigma u)_{j+1/2} = \sigma_{j+1/2} u_{j+1/2},$$

where $u_{j+1/2} = \frac{1}{2}(u_j + u_{j+1})$. The discretization is now complete and we refer to this scheme as the Order 1 L^2 AP scheme.

3.2.2.2 Order 1 implicit upwinding

As mentionned before, choosing a centered discretization for the implicit inviscid flux F_i leads to an L^2 stable scheme (see Section 2.3.4 for more details). In some cases, we may want to add some stabilization to reduce appearing spurious oscillations. Adding numerical dissipation on the implicit inviscid flux, we are able to construct an order 1 L^∞ stable scheme [29]. For that, we compute the L^2 stable solution

$W_j^{n+1,L2}$ given by (3.23) and we add an upwinding as done for the explicit numerical flux $(\mathcal{F}_e)_{j+1/2}^n$, thus leading to a modified scheme for the density, momentum and energy equations.

$$W_j^{n+1} = \begin{pmatrix} \rho_j^{n+1,L2} \\ q_j^{n+1,L2} \\ E_j^{n+1,L2} \end{pmatrix} + \frac{\Delta t}{\Delta x}\left((\mathcal{D}_i)_{j+1/2}^n(W_{j+1}^{n+1} - W_j^{n+1}) - (\mathcal{D}_i)_{j-1/2}^n(W_j^{n+1} - W_{j-1}^{n+1})\right). \tag{3.25a}$$

where $(\mathcal{D}_i)_{j+1/2}^n$ is the implicit viscosity coefficient, taken as half of the maximum implicit eigenvalue

$$(\mathcal{D}_i)_{j+1/2}^n := \frac{1}{2}\max\left(|\lambda_i(W_{j+1}^n)|, |\lambda_i(W_j^n)|\right), \tag{3.25b}$$

where

$$|\lambda_i(W)| = \frac{|u|}{2} + \sqrt{\frac{u^2}{4} + \frac{c^2}{\varepsilon}}.$$

As already mentioned in Chapter 2, the upwinding must be applied after the computation of all conservative variables. Moreover, it is important to note that the viscosity coefficient $(\mathcal{D}_i)$ depends on the scaling parameter ε. It is inversely proportional to the Mach number and may lead to excessive diffusion in the low Mach number regime, see Section 3.4.4 for illustration. Therefore, the proposed stabilization technique is used only if needed, depending on the problem to solve. We refer to the modified scheme (3.23)-(3.25) as the Order 1 L^∞ AP scheme. In Figure 3.1, we compare the density profiles for a 2D Riemman problem (see Section 3.4.1 for its description) computed with the Order 1 L^2 scheme (left) against the Order 1 L^∞ scheme (right). As expected, with an upwinding on the implicit flux we are able to eliminate the oscillations appearing when a centered discretization for the implicit part is chosen. However, the solution is also more diffused and it shows the need to extend the schemes to a higher order of accuracy.

3.2.3 Asymptotic stability : C.F.L. condition on the time step

For the full Euler equations, when we set $\mu = \lambda = 0$, we conducted a one-dimensional linear Fourier stability analysis of the presented scheme (see Section 2.3). It proved on the linearized system that our Order 1 L^2 AP scheme (3.23), (3.24) is stable under the C.F.L. condition $\gamma\Delta t = |u|\Delta x$. Furthermore, when adding the implicit upwinding on the implicit inviscid part (see Section 3.25), our Order 1 L^∞ AP scheme (3.23)-(3.25) is stable under the C.F.L. condition $\Delta t = |u|\Delta x$. Then, both discretizations are asymptotically stable. The time step restriction does not depend on the Mach number and is only related to the fluid velocity.
Moreover, we show numerically that the situation does not change in case of viscous flows. In Figure 3.2, we compare the time step sizes Δt between different schemes for the first problem of Stokes (see Section 3.4.5 for its description) for which $\varepsilon = 10^{-6}$. On the left figure, we compare the time step sizes between the explicit scheme and

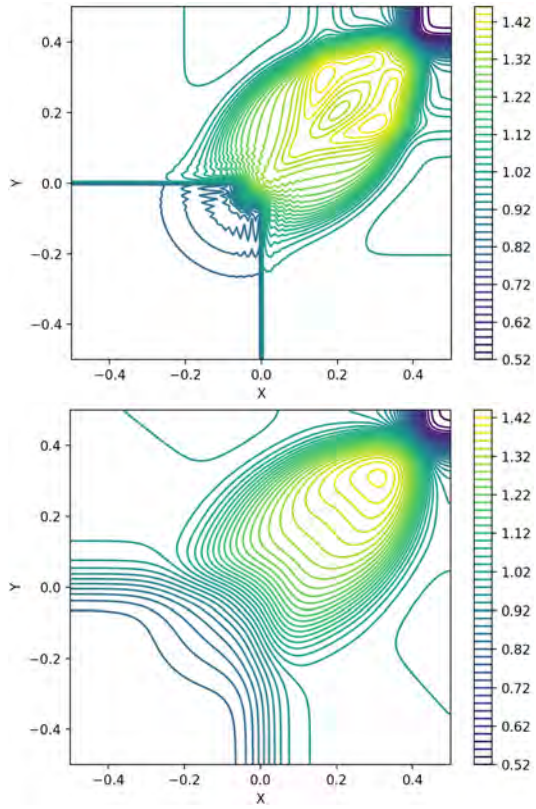


Figure 3.1 – 2D Riemann problem (see Section 3.4.1 for its description): Density isolines with the Order 1 L^2 AP scheme (top) and the Order 1 L^∞ AP scheme (bottom).

the Order 1 AP schemes. We observe that the time steps of the Order 1 L^2 AP scheme with an explicit discretization of the viscous terms are around $1/\sqrt{\varepsilon} = 10^3$ times bigger than the ones required by the fully explicit scheme. With an implicit discretization of the viscous terms, the time steps sizes of our Order 1 L^2 and L^∞ AP schemes are even bigger. This shows that the AP schemes can employ time step sizes independently of the Mach number regime. On the right figure, we observe the advantages of an implicit discretization of the viscous terms in the case of highly viscous flows.

3.2.4 L^2 discretization in two dimensions

Let us consider a uniform space discretization where we denote by Δx and Δy the space steps respectively in x and y direction. A cell is labeled by the indices i, j, i and j for respectively the x and y directions. We denote by $W_{i,j} = (\rho_{i,j}, q_{i,j}, E_{i,j})$ the conservative variables. The momentum is defined by $q_{i,j} = (\rho U)_{i,j} = (\rho u, \rho v)_{i,j}$ where u and v are the velocities in the x and y direction respectively. Then, the

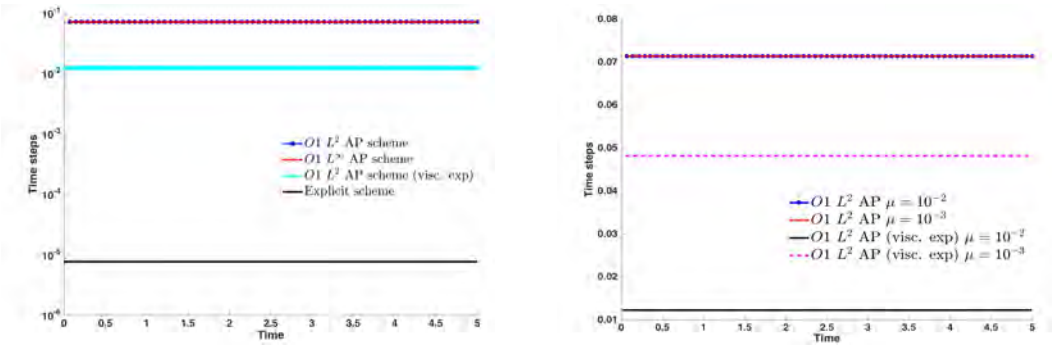


Figure 3.2 – Time step sizes Δt as a function of time for the first problem of Stokes (see Section sec:stokes for its description): Left panel: Comparison of the Order 1 AP schemes against the explicit scheme for $\varepsilon = 10^{-6}$ and $\mu = 10^{-2}$. Right panel: Comparison of the Order 1 L^2 AP scheme against the Order 1 L^2 AP scheme with an explicit discretization of the viscous terms for $\varepsilon = 10^{-6}$ and with $\mu = 10^{-2}$ and $\mu = 10^{-3}$.

fully discrete scheme reads

$$W_{i,j}^{n+1,exp} = \begin{pmatrix} \rho_{i,j}^{n+1,exp} \\ q_{i,j}^{n+1,exp} \\ E_{i,j}^{n+1,exp} \end{pmatrix} = W_{i,j}^{n} - \Delta t \left(\nabla \cdot F_e(W^n)\right)_{i,j}, \tag{3.26a}$$

$$\rho_{i,j}^{n+1} = \rho_{i,j}^{n+1,exp}, \tag{3.26b}$$

$$\begin{aligned} \frac{\varepsilon}{\gamma-1} p_{i,j}^{n+1} - \Delta t^2 \nabla \cdot \left(\frac{h^{n+1,exp}}{\rho^{n+1}} \nabla p^{n+1} \right)_{i,j} - \frac{\varepsilon \Delta t \lambda}{R} \Delta \left(\frac{p^{n+1}}{\rho^{n+1}} \right)_{i,j} &= \varepsilon E_{i,j}^{n+1,exp} \\ - \varepsilon k_{i,j}^{n+1,exp} + \varepsilon^2 \Delta t \nabla \cdot \left(\sigma^{n+1,exp} \frac{q^{n+1,exp}}{\rho^{n+1}} \right)_{i,j} \\ - \varepsilon \Delta t \nabla \cdot \left(\frac{h^{n+1,exp}}{\rho^{n+1}} \left(q^{n+1,exp} + \Delta t \nabla \cdot \sigma^{n+1,exp} \right) \right)_{i,j}, \end{aligned} \tag{3.26c}$$

$$q_{i,j}^{n+1} - \Delta t \left(\nabla \cdot \sigma^{n+1}\right)_{i,j} = q_{i,j}^{n+1,exp} - \frac{\Delta t}{2\varepsilon} \begin{pmatrix} \dfrac{p_{i+1,j}^{n+1} - p_{i-1,j}^{n+1}}{\Delta x} \\ \dfrac{p_{i,j+1}^{n+1} - p_{i,j-1}^{n+1}}{\Delta y} \end{pmatrix}, \tag{3.26d}$$

$$\begin{aligned} E_{i,j}^{n+1} = E_{i,j}^{n+1,exp} - \Delta t \nabla \cdot \left(\frac{\gamma p^{n+1}}{(\gamma-1)\rho^{n+1}} q^{n+1} \right)_{i,j} + \varepsilon \Delta t \nabla \cdot \left(\sigma^{n+1} \frac{q^{n+1}}{\rho^{n+1}} \right)_{i,j} \\ + \frac{\Delta t \lambda}{R} \Delta \left(\frac{p^{n+1}}{\rho^{n+1}} \right)_{i,j}. \end{aligned} \tag{3.26e}$$

The discretization of the explicit fluxes and of the viscous terms is discussed bellow.

3.2.4.1 Discretization of $W_{i,j}^{n+1,exp}$, the explicit convected part of the conservative variables

Let us first consider the discretization of the explicit convected part of the conservative variables. The discretization reads

$$W_{i,j}^{n+1,exp} = W_{i,j}^{n} - \Delta t \left(\frac{(\mathcal{F}_e^x)_{i+1/2,j}^n - (\mathcal{F}_e^x)_{i-1/2,j}^n}{\Delta x} + \frac{(\mathcal{F}_e^y)_{i,j+1/2}^n - (\mathcal{F}_e^y)_{i,j-1/2}^n}{\Delta y} \right). \tag{3.27}$$

For the explicit numerical fluxes $(\mathcal{F}_e^x)^n = \left((\rho u)^n, (\rho u^2)^n, (\rho uv)^n, (k_\varepsilon u)^n\right)$ we consider a Rusanov solver. The fluxes are given by

$$(\mathcal{F}_e^x)_{i+1/2,j}^n := \frac{\mathcal{F}_e^x(W_{i+1,j}^n) + \mathcal{F}_e^x(W_{i+1,j}^n)}{2} - (\mathcal{D}_e^x)_{i+1/2,j}^n (W_{i+1,j}^n - W_{i,j}^n), \tag{3.28}$$

where $(\mathcal{D}_e^x)_{i+1/2,j}^n$ the explicit viscosity coefficient, is taken as half of the maximum explicit eigenvalue of the Jacobian matrix associated to $(\mathcal{F}_e^x)^n$:

$$(\mathcal{D}_e^x)_{i+1/2,j}^n = \frac{1}{2} \max\left(|u_{i+1,j}^n|, |u_{i,j}^n|\right).$$

Likewise, in the y direction the explicit fluxes $(\mathcal{F}_e^y)^n = \left((\rho v)^n, (\rho uv)^n, (\rho v^2)^n, (k_\varepsilon v)^n\right)$ are given by

$$(\mathcal{F}_e^y)_{i,j+1/2}^n := \frac{\mathcal{F}_e^y(W_{i,j+1}^n) + \mathcal{F}_e^y(W_{i,j+1}^n)}{2} - (\mathcal{D}_e^y)_{i,j+1/2}^n (W_{i,j+1}^n - W_{i,j}^n), \tag{3.29}$$

where $(\mathcal{D}_e^y)_{i,j+1/2}^n$ the explicit viscosity coefficient, is taken as half of the maximum explicit eigenvalue of the Jacobian matrix associated to $(\mathcal{F}_e^y)^n$:

$$(\mathcal{D}_e^y)_{i,j+1/2}^n = \frac{1}{2} \max\left(|v_{i,j+1}^n|, |v_{i,j}^n|\right).$$

Let us note that the upwinding depends only on the fluid velocity.
In the next section, we present the discretization of the viscous terms.

3.2.4.2 Discretization of the viscous stress tensor and viscous forces

Let us remember that the viscous stress tensor is defined by

$$\sigma = \mu\left(\nabla U + (\nabla U)^T\right) - \frac{2}{3}\mu\left(\nabla \cdot U\right) I,$$

where $\nabla U = \begin{pmatrix} \partial_x u & \partial_y u \\ \partial_x v & \partial_y v \end{pmatrix}$ and $(\nabla \cdot U)\, I = \begin{pmatrix} \partial_x u + \partial_y v & 0 \\ 0 & \partial_x u + \partial_y v \end{pmatrix}$. And so,

$$\sigma = \frac{\mu}{3} \begin{pmatrix} 4\partial_x u - 2\partial_y v & 3\partial_y u + 3\partial_x v \\ 3\partial_y u + 3\partial_x v & 4\partial_y v - 2\partial_x u \end{pmatrix}. \tag{3.30}$$

Moreover, the term $\nabla \cdot \sigma$ in the pressure equation (3.26c) and the momentum equations (3.26d) is given by:

$$\nabla \cdot \sigma = \mu \Delta U + \frac{\mu}{3} \nabla (\nabla \cdot U) = \frac{\mu}{3} \begin{pmatrix} 4\partial_{xx}^2 u + 3\partial_{yy}^2 u + \partial_{xy}^2 v \\ 3\partial_{xx}^2 v + 4\partial_{yy}^2 v + \partial_{yx}^2 u \end{pmatrix}. \tag{3.31}$$

In order to compute the viscous term $(\nabla \cdot \sigma)_{i,j}$, we need the velocity derivatives at the interfaces. They are discretized with a centered solver by

$$\nabla U_{i+1/2,j} = \begin{pmatrix} \dfrac{u_{i+1,j} - u_{i,j}}{\Delta x} & \dfrac{u_{i+1,j+1} - u_{i+1,j-1} + u_{i,j+1} - u_{i,j-1}}{4\Delta y} \\ \dfrac{v_{i+1,j} - v_{i,j}}{\Delta x} & \dfrac{v_{i+1,j+1} - v_{i+1,j-1} + v_{i,j+1} - v_{i,j-1}}{4\Delta y} \end{pmatrix},$$

$$\text{and,} \quad \nabla U_{i,j+1/2} = \begin{pmatrix} \dfrac{u_{i+1,j+1} - u_{i-1,j+1} + u_{i+1,j} - u_{i-1,j}}{4\Delta x} & \dfrac{u_{i,j+1} - u_{i,j}}{\Delta y} \\ \dfrac{v_{i+1,j+1} - v_{i-1,j+1} + v_{i+1,j} - v_{i-1,j}}{4\Delta x} & \dfrac{v_{i,j+1} - v_{i,j}}{\Delta y} \end{pmatrix}, \tag{3.32}$$

where the discretization of $\partial_y f_{i+1/2,j}$ for $f = u, v$ is obtained as follows

$$\begin{aligned} (\partial_y f)_{i+1/2,j} &= \frac{1}{2} \left((\partial_y f)_{i+1,j} + (\partial_y f)_{i,j} \right) \\ &= \frac{1}{2} \left(\frac{u_{i+1,j+1} - u_{i+1,j-1}}{2\Delta y} + \frac{u_{i,j+1} - u_{i,j-1}}{2\Delta y} \right) \\ &= \frac{u_{i+1,j+1} - u_{i+1,j-1} + u_{i,j+1} - u_{i,j-1}}{4\Delta y}. \end{aligned}$$

Thus, the discretization of the second order derivatives $\partial_{xx}^2 f$ and $\partial_{yy}^2 f$ with (3.32) is given by:

$$\left(\partial_{xx}^2 f\right)_{i,j} = \frac{(\partial_x f)_{i+1/2,j} - (\partial_x f)_{i-1/2,j}}{\Delta x} = \frac{f_{i+1,j} - 2f_{i,j} + f_{i+1,j}}{\Delta x^2},$$

and

$$\left(\partial_{yy}^2 f\right)_{i,j} = \frac{(\partial_y f)_{i,j+1/2} - (\partial_x f)_{i,j-1/2}}{\Delta y} = \frac{f_{i,j+1} - 2f_{i,j} + f_{i,j+1}}{\Delta y^2}.$$

Furthermore, the discretization of the crossed derivatives $\left(\partial_{yx}^2 f\right)_{i,j} = \left(\partial_{xy}^2 f\right)_{i,j}$ reads:

$$\begin{aligned} \left(\partial_{yx}^2 f\right)_{i,j} &= \frac{(\partial_x f)_{i,j+1/2} - (\partial_x f)_{i,j-1/2}}{\Delta y} \\ &= \frac{1}{2\Delta y} \left(\frac{u_{i+1,j+1} - u_{i-1,j+1}}{2\Delta x} + \frac{u_{i+1,j-1} - u_{i-1,j-1}}{2\Delta x} \right). \end{aligned}$$

Then, the fully discretized momentum equations (3.26d) read

$$\begin{aligned}(\rho u)_{i,j}^{n+1}+\mu\Delta t\Big(\ 2(\frac{4}{3\Delta x^2}+\frac{1}{\Delta y^2})u_{ij}^{n+1}-\frac{4}{3\Delta x^2}(u_{i+1,j}^{n+1}+u_{i-1,j}^{n+1})-\frac{1}{\Delta y^2}(u_{i,j+1}^{n+1}+u_{i,j-1}^{n+1})\\ -\frac{1}{12\Delta x\Delta y}(v_{i+1,j+1}^{n+1}+v_{i-1,j-1}^{n+1}-v_{i-1,j+1}^{n+1}-v_{i+1,j-1}^{n+1})\Big)\\ =(\rho u)_{ij}^{n+1,exp}-\frac{\Delta t}{\varepsilon}\frac{p_{i+1,j}^{n+1}-p_{i-1,j}^{n+1}}{2\Delta x},\end{aligned}$$

$$\begin{aligned}(\rho v)_{i,j}^{n+1}+\mu\Delta t\Big(\ 2(\frac{1}{\Delta x^2}+\frac{4}{3\Delta y^2})v_{ij}^{n+1}-\frac{1}{\Delta x^2}(v_{i+1,j}^{n+1}+v_{i-1,j}^{n+1})-\frac{4}{3\Delta y^2}(v_{i,j+1}^{n+1}+v_{i,j-1}^{n+1})\\ -\frac{1}{12\Delta x\Delta y}(u_{i+1,j+1}^{n+1}+u_{i-1,j-1}^{n+1}-u_{i-1,j+1}^{n+1}-u_{i+1,j-1}^{n+1})\Big)\\ =(\rho v)_{ij}^{n+1,exp}-\frac{\Delta t}{\varepsilon}\frac{p_{i,j+1}^{n+1}-p_{i,j-1}^{n+1}}{2\Delta y}.\end{aligned} \tag{3.33}$$

Additionally, to compute the term $\nabla\cdot\left(\sigma\frac{q}{\rho}\right)_{i,j}$ in the pressure equation (3.26c) and in the energy equation (3.26e), the discretization of $\sigma_{i+1/2,j}$ and $\sigma_{i,j+1/2}$ is given by (3.32) and the velocities at the interfaces are defined as the arithmetic average with

$$U_{i+1/2,j}=\frac{1}{2}\left(U_{i+1,j}+U_{i,j}\right),\quad U_{i,j+1/2}=\frac{1}{2}\left(U_{i,j+1}+U_{i,j}\right). \tag{3.34}$$

We will skip the full discretization of the remaining terms since it can be easily derived from the one-dimensional case. We present in the next section, the order 2 extension of the proposed schemes.

3.3 Order 2 AP scheme

3.3.1 Semi-discretization in time

Let us first present the order 2 semi-discretization in time. Using the ARS(2,2,2) [80], the following order 2 semi-discretization in time scheme is obtained

$$\frac{W^\star-W^n}{\Delta t}+\beta\,\nabla\cdot F_e(W^n)+\beta\,\nabla\cdot F_i(W^\star)=\beta\nabla\cdot G(W^\star,\nabla W^\star), \tag{3.35a}$$

$$\begin{aligned}&\frac{W^{n+1}-W^n}{\Delta t}+(\beta-1)\,\nabla\cdot F_e(W^n)+(2-\beta)\,\nabla\cdot F_e(W^\star) \qquad (3.35b)\\ &\quad+(1-\beta)\,\nabla\cdot F_i(W^\star)+\beta\,\nabla\cdot F_i(W^{n+1})\\ &\qquad=(1-\beta)\,\nabla\cdot G(W^\star,\nabla W^\star)+\beta\nabla\cdot G(W^{n+1},\nabla W^{n+1}), \qquad (3.35c)\end{aligned}$$

with $\beta=1-1/\sqrt{2}$.
Following the reformulation used for the order 1 AP scheme for each step, we have

Step 1:

$$W^{\star,exp} = W^n - \beta\Delta t\nabla\cdot F_e(W^n), \tag{3.36a}$$

$$\rho^\star = \rho^{\star,exp}, \tag{3.36b}$$

$$\begin{aligned}\frac{\varepsilon}{\gamma-1}p^\star - \beta^2\Delta t^2\nabla\cdot\left(\frac{h^{\star,exp}}{\rho^\star}\nabla p^\star\right) - \frac{\varepsilon\beta\Delta t\lambda}{R}\Delta\left(\frac{p^\star}{\rho^\star}\right) &= \varepsilon\left(E^{\star,exp} - k^{\star,exp}\right)\\ - \varepsilon\beta\Delta t\nabla\cdot\left(\frac{h^{\star,exp}}{\rho^\star}\left(q^{\star,exp} + \beta\Delta t\nabla\cdot\sigma^{\star,exp}\right)\right) &+ \varepsilon^2\beta\Delta t\nabla\cdot\left(\sigma^{\star,exp}\frac{q^{\star,exp}}{\rho^\star}\right),\end{aligned} \tag{3.36c}$$

$$q^\star - \beta\Delta t\nabla\cdot\sigma^* = q^{\star,exp} - \beta\Delta t\frac{1}{\varepsilon}\nabla p^\star, \tag{3.36d}$$

$$E^\star = E^{\star,exp} - \beta\Delta t\nabla\cdot\left(\frac{\gamma p^\star}{(\gamma-1)\rho^\star}q^\star\right) + \varepsilon\beta\Delta t\nabla\cdot\left(\sigma^\star\frac{q^\star}{\rho^\star}\right) + \frac{\beta\Delta t\lambda}{R}\Delta\left(\frac{p^\star}{\rho^\star}\right), \tag{3.36e}$$

Step 2:

$$\begin{aligned}W^{n+1,exp} = W^n &- \Delta t\left((\beta-1)\nabla\cdot F_e(W^n) + (2-\beta)\left(\nabla\cdot F_e(W^\star)\right)\right)\\ &+ \Delta t\left((1-\beta)\left(\nabla\cdot F_i(W^\star) - G(W^\star)\right)\right).\end{aligned} \tag{3.37a}$$

$$\rho^{n+1} = \rho^{n+1,exp}, \tag{3.37b}$$

$$\begin{aligned}\frac{\varepsilon}{\gamma-1}p^{n+1} - \beta^2\Delta t^2\nabla\cdot\left(\frac{h^{n+1,exp}}{\rho^{n+1}}\nabla p^{n+1}\right) - \frac{\varepsilon\beta\Delta t\lambda}{R}\Delta\left(\frac{p^{n+1}}{\rho^{n+1}}\right) &= \varepsilon E^{n+1,exp}\\ - \varepsilon k^{n+1,exp} - \varepsilon\beta\Delta t\nabla\cdot\left(\frac{h^{n+1,exp}}{\rho^{n+1}}\left(q^{n+1,exp} + \beta\Delta t\nabla\cdot\sigma^{n+1,exp}\right)\right)&\\ + \varepsilon^2\beta\Delta t\nabla\cdot\left(\sigma^{n+1,exp}\frac{q^{n+1,exp}}{\rho^{n+1}}\right),&\end{aligned} \tag{3.37c}$$

$$q^{n+1} - \beta\Delta t\nabla\cdot\sigma^{n+1} = q^{n+1,exp} - \frac{\beta\Delta t}{\varepsilon}\nabla p^{n+1}, \tag{3.37d}$$

$$E^{n+1} = E^{n+1,exp} - \beta\Delta t\nabla\cdot\left(\frac{\gamma p^{n+1}}{(\gamma-1)\rho^{n+1}}q^{n+1}\right) + \frac{\beta\Delta t\lambda}{R}\Delta\left(\frac{p^{n+1}}{\rho^{n+1}}\right). \tag{3.37e}$$

where $k^l = k(W^l)$, $h^l = h(W^l) = \gamma(E^l - k^l)$ and $\sigma^l = \sigma(W^l)$ for $l =$"$\star, exp$","$n+1, exp$".

3.3.2 Order 2 space discretization in one dimension

3.3.2.1 Second order space accuracy for the explicit flux F_e

In order to extend the space accuracy to second order, we classically use the MUSCL technique [54] and so a piecewise linear reconstruction of W_j^n:

$$\widehat{W}_j^n(x) = W_j^n + \alpha_j^n(x - x_j), \tag{3.38}$$

where α_j^n is a limited slope and is computed for each component using a minmod limiter:

$$\alpha_j^n = minmod\left(\frac{W_{j+1}^n - W_j^n}{\Delta x}, \frac{W_j^n - W_{j-1}^n}{\Delta x}\right)$$

where the limiter is defined as

$$minmod(a,b) = \frac{1}{2}(sign(a) + sign(b))\min(|a|,|b|) = \begin{cases} a & if\ |a| < |b|, ab > 0, \\ b & if\ |b| < |a|, ab > 0, \\ 0 & otherwise. \end{cases}$$

It picks out the flattest slope when they have the same sign.
This piecewise linear reconstruction is used for defining the explicit numerical flux at the interfaces. Using the notations introduced for the Order 1 AP scheme

$$(\mathcal{F}_e)_{j+1/2}^n := \frac{F_e(W_{j+1,-}^n) + F_e(W_{j,+}^n)}{2\Delta x} - (\mathcal{D}_e)_{j+1/2}^n (W_{j+1,-}^n - W_{j,+}^n), \tag{3.39}$$

where $(\mathcal{D}_e)_{j+1/2}^n = \frac{1}{2}\max\left(|u_{j,+}^n|, |u_{j+1,-}^n|\right)$ and

$$W_{j,\pm}^n = \widehat{W}_j^n(x_j \pm \frac{\Delta x}{2}) = W_j^n \pm \frac{\Delta x}{2}, \sigma_j^n.$$

3.3.2.2 Second order space accuracy for the remaining terms

In the momentum (3.15e) and in the energy equations (3.15f) the implicit flux F_i is discretized with a centered solver which ensures second order accuracy in space so no reconstruction is needed. Indeed, using a Taylor expansion in terms of $\psi(x_i)$ for $\psi(x_{i+1})$ and $\psi(x_{i-1})$, we have:

$$\psi(x_{i+1}) = \psi(x_i) + \Delta x \partial_x \psi(x_i) + \frac{\Delta x^2}{2}\partial_{xx}^2\psi(x_i) + \frac{\Delta x^3}{6}\partial_{xxx}^3\psi(x_i) + \mathcal{O}(\Delta x^4), \tag{3.40}$$

$$\psi(x_{i-1}) = \psi(x_i) - \Delta x \partial_x \psi(x_i) + \frac{\Delta x^2}{2}\partial_{xx}^2\psi(x_i) - \frac{\Delta x^3}{6}\partial_{xxx}^3\psi(x_i) + \mathcal{O}(\Delta x^4). \tag{3.41}$$

Then, substracting (3.41) to (3.40) and dividing by $2\Delta x$ yields

$$\frac{\partial \psi}{\partial x}(x_j) = \frac{\psi_{i+1} - \psi_{i-1}}{2\Delta x} + \mathcal{O}(\Delta x^2),$$

where $\psi_k = \psi(x_k)$ for $k = i+1, i-1$ and we set $\psi = p$ and $\psi = \frac{\gamma p}{(\gamma-1)\rho}q$.
Moreover, the viscous term $\partial_x\sigma(x_j) = \partial_x(\frac{4\mu}{3}\partial_x u)(x_j) = \frac{4\mu}{3}\partial_{xx}^2 u(x_j)$ is discretized by

$$\frac{4\mu}{3}\frac{u_{i+1} - 2u_i + u_{i-1}}{\Delta x^2}.$$

Adding (3.40) and (3.41), dividing by Δx^2 and rearranging the terms gives:

$$\frac{\partial^2 \psi}{\partial x^2}(x_j) = \frac{\psi_{i+1} - 2\psi_i + \psi_{i-1}}{\Delta x^2} + \mathcal{O}(\Delta x^2),$$

with $\psi_k = \psi(x_k)$ for $k = i+1, i, i-1$. And so, setting $\psi = u$ and $\psi = \left(\frac{p}{\rho}\right)$ shows that the discretizations of $(\partial_x\sigma)_j$ and $\left(\partial_{xx}^2\left(\frac{p}{\rho}\right)\right)_j$ are second order accurate.

Next, we show that the discretization of the following flux operators

$$\left(\partial_x\left(\frac{h}{\rho}\partial_x p\right)\right)_j, \quad \left(\partial_x\left(\sigma\frac{q}{\rho}\right)\right)_j,$$

are also second order accurate. Setting for the first flux term $\phi = \dfrac{h}{\rho}$, $\psi = p$ and for the second one $\phi = u$ and $\psi = \frac{4\mu}{3}u$, there discretization reads

$$\frac{1}{\Delta x}\left(\frac{1}{2}(\phi_{i+1}+\phi_i)\frac{\psi_{i+1}-\psi_i}{\Delta x} - \frac{1}{2}(\phi_i+\phi_{i-1})\frac{\psi_i-\psi_{i-1}}{\Delta x}\right), \tag{3.42}$$

with $\psi_k = \psi(x_k)$ and $\phi_k = \phi(x_k)$ for $k = i+1, i, i-1$. Using the Taylor expansions for ψ_{i+1} (3.40) and ψ_{i-1} (3.41) and in terms of ϕ_i for ϕ_{i+1} and ϕ_{i-1} we have:

$$\begin{aligned}
\frac{\psi_{i+1}-\psi_i}{\Delta x} &= \partial_x\psi_i + \frac{\Delta x}{2}\partial^2_{xx}\psi_i + \frac{\Delta x^2}{6}\partial^3_{xxx}\psi_i + \mathcal{O}(\Delta x^3),\\
\frac{\psi_i-\psi_{i-1}}{\Delta x} &= \partial_x\psi_i - \frac{\Delta x}{2}\partial^2_{xx}\psi_i + \frac{\Delta x^2}{6}\partial^3_{xxx}\psi_i + \mathcal{O}(\Delta x^3),\\
\frac{1}{2}(\phi_{i+1}+\phi_i) &= \phi_i + \frac{\Delta x}{2}\partial_x\phi_i + \frac{\Delta x^2}{4}\partial^2_{xx}\phi_i + \mathcal{O}(\Delta x^3),\\
\frac{1}{2}(\phi_i+\phi_{i-1}) &= \phi_i - \frac{\Delta x}{2}\partial_x\phi_i + \frac{\Delta x^2}{4}\partial^2_{xx}\phi_i + \mathcal{O}(\Delta x^3).
\end{aligned}$$

Then, injecting those expressions into (3.42) gives,

$$\begin{aligned}
D(\partial_x(\phi\partial_x\psi))_j &= \frac{1}{\Delta x}\left(\frac{1}{2}(\phi_{i+1}+\phi_i)\frac{\psi_{i+1}-\psi_i}{\Delta x} - \frac{1}{2}(\phi_i+\phi_{i-1})\frac{\psi_i-\psi_{i-1}}{\Delta x}\right)\\
&= \frac{1}{\Delta x}\left(\Delta x\partial_x\psi_i\partial_x\phi_i + \Delta x\partial^2_{xx}\psi_i\phi_i + \mathcal{O}(\Delta x^3)\right)\\
&= \partial_x(\phi\partial_x\psi)(x_j) + \mathcal{O}(\Delta x^2).
\end{aligned}$$

This, concludes the extension to second order accuracy in space.

3.3.2.3 Order 2 implicit upwinding

For the Order 2 AP scheme, it is sufficient to add numerical diffusion on the implicit flux F_i only at the end of the second step. Adding it in both steps would imply a higher computational cost for similar results. As done for the Order 1 AP scheme, we compute the L^2 stable solution $W_j^{n+1,L2}$ with (3.36)-(3.37), (3.39) and then add numerical dissipation on the conservative variables:

$$\begin{aligned}
W_j^{n+1} = \begin{pmatrix}\rho_j^{n+1,L2}\\ q_j^{n+1,L2}\\ E_j^{n+1,L2}\end{pmatrix} &+ \frac{\beta\Delta t}{\Delta x}\left((\mathcal{D}_i)^n_{j+1/2}(\tilde{W}^{n+1}_{j+1,-} - \tilde{W}^{n+1}_{j,+})\right)\\
&- \frac{\beta\Delta t}{\Delta x}\left((\mathcal{D}_i)^n_{j-1/2}(\tilde{W}^{n+1}_{j,-} - \tilde{W}^{n+1}_{j-1,+})\right).
\end{aligned} \tag{3.43}$$

where $(\mathcal{D}_i)^n_{j+1/2} = 1/2 \max\left(|\lambda_i(W^n_{j+1,-})|, |\lambda_i(W^n_{j,+})|\right)$ and $\tilde{W}^{n+1}_{j,\pm} = W^{n+1}_j \pm \frac{\Delta x}{2}\sigma^n_j$. Let us note that, as mentioned in Section 2.4, IMEX methods of order higher than one for hyperbolic problems cannot be TVD nor L^∞ stable for unconstrained time steps [40, 29]. Thus, the scheme given by (3.36), (3.37) in time and (3.39), (3.43) in space, is still L^2 stable but the oscillations are reduced thanks to the upwinding on the implicit part. We now refer to this scheme as the Order 2 $L^{2,stab}$ AP scheme.

3.3.2.4 Some details for the extension to the two-dimensional case

For the explicit operators the reconstruction procedure is simply carried out dimension by dimension where the linear polynomial writes

$$\widehat{W}^n_{i,j}(x,y) = W^n_{i,j} + \alpha^n_{x_{i,j}}(x - x_{i,j}) + \alpha^n_{y_{i,j}}(y - y_{i,j}), \tag{3.44}$$

and the limited slope $\alpha^n_{x_{i,j}} = minmod(\frac{W_{i+1,j}-W_{i,j}}{\Delta x}, \frac{W_{i,j}-W_{i-1,j}}{\Delta x})$. The slope in y direction is computed in the same manner.
Concerning the implicit viscous stress tensor, we also prove the second order accuracy of the crossed derivatives. Indeed, their discretization reads:

$$\frac{1}{2\Delta y}\left(\frac{\psi_{i+1,j+1} - \psi_{i-1,j+1}}{2\Delta x} + \frac{\psi_{i+1,j-1} - \psi_{i-1,j-1}}{2\Delta x}\right) \tag{3.45}$$

Conducting a Taylor expansion in terms of $\psi(x_i, y_j)$ up to order four on the variables $\psi(x_{i+1}, y_{j+1})$ and $\psi(x_{i-1}, y_{j+1})$ we have:

$$\begin{aligned}
\psi(x_{i+1}, y_{j+1}) &= \psi(x_i, y_j) + \partial_x\psi(x_i, y_j)\Delta x + \partial y\psi(x_i, y_j)\Delta y \\
&+ \frac{1}{2}\partial^2_{xx}\psi(x_i, y_j)\Delta x^2 + \frac{1}{2}\partial^2_{yy}\psi(x_i, y_j)\Delta y^2 + \partial^2_{xy}\psi(x_i, y_j)\Delta x\Delta y \\
&+ \frac{1}{6}\partial^3_{xxx}\psi(x_i, y_j)\Delta x^3 + \frac{1}{6}\partial^3_{yyy}\psi(x_i, y_j)(\Delta y)^3 + \frac{1}{2}\partial^3_{xxy}\psi(x_i, y_j)\Delta x^2\Delta y \\
&+ \frac{1}{2}\partial^3_{yyx}\psi(x_i, y_j)\Delta x\Delta y^2 + \mathcal{O}(\Delta x^4) + \mathcal{O}(\Delta y^4),
\end{aligned} \tag{3.46}$$

$$\begin{aligned}
\psi(x_{i-1}, y_{j+1}) &= \psi(x_i, y_j) + \partial_x\psi(x_i, y_j)(-\Delta x) + \partial y\psi(x_i, y_j)\Delta y \\
&+ \frac{1}{2}\partial^2_{xx}\psi(x_i, y_j)\Delta x^2 + \frac{1}{2}\partial^2_{yy}\psi(x_i, y_j)(\Delta y)^2 + \partial^2_{xy}\psi(x_i, y_j)(-\Delta x)\Delta y \\
&+ \frac{1}{6}\partial^3_{xxx}\psi(x_i, y_j)(-\Delta x)^3 + \frac{1}{6}\partial^3_{yyy}\psi(x_i, y_j)\Delta y^3 + \frac{1}{2}\partial^3_{xxy}\psi(x_i, y_j)\Delta x^2\Delta y \\
&+ \frac{1}{2}\partial^3_{xyy}\psi(x_i, y_j)(-\Delta x)\Delta y^2 + \mathcal{O}(\Delta x^4) + \mathcal{O}(\Delta y^4).
\end{aligned} \tag{3.47}$$

And so:

$$\begin{aligned}
\psi_{i+1,j+1} - \psi_{i-1,j+1} &= 2\partial_x\psi(x_i, y_j)\Delta x - 2\partial^2_{xy}\psi(x_i, y_j)\Delta x\Delta y \\
&+ \frac{1}{3}\partial^3_{xxx}\psi(x_i, y_j)\Delta x^3 + \partial^3_{xyy}\psi(x_i, y_j)\Delta x\Delta y^2 + \mathcal{O}(\Delta x^4) + \mathcal{O}(\Delta y^4),
\end{aligned} \tag{3.48}$$

Similarly,

$$\begin{aligned}
\psi_{i+1,j-1} - \psi_{i-1,j-1} &= 2\partial_x\psi(x_i, y_j)\Delta x - 2\partial^2_{xy}\psi(x_i, y_j)\Delta x\Delta y \\
&+ \frac{1}{3}\partial^3_{xxx}\psi(x_i, y_j)\Delta x^3 + \partial^3_{xyy}\psi(x_i, y_j)\Delta x\Delta y^2 + \mathcal{O}(\Delta x^4) + \mathcal{O}(\Delta y^4).
\end{aligned} \tag{3.49}$$

And thus, adding (3.48) and (3.49) and dividing by $4\Delta y\Delta x$ we get

$$\partial^2_{yx}\psi(x_i,y_j) = \frac{1}{2\Delta y}\left(\frac{\psi_{i+1,j+1}-\psi_{i-1,j+1}}{2\Delta x}+\frac{\psi_{i+1,j-1}-\psi_{i-1,j-1}}{2\Delta x}\right) + \frac{\mathcal{O}(\Delta x^4)+\Delta y^4)}{\Delta y\Delta x}. \quad (3.50)$$

Using the same arguments as for the one-dimensional case, second order accuracy is also guaranteed for the discretization of the remaining terms.

3.4 Two dimensional numerical results for the Euler and Navier-Stokes equations

In this part, we present several numerical test cases which show the good behavior of our AP schemes. Results are shown depending on the numerical test, for the Order 2 L^2 and $L^{2,stab}$ AP schemes where we recall that in the $L^{2,stab}$ AP scheme we applied an implicit upwinding (see Section 3.3.2.3) to reduce the appearing oscillations. We give below their corresponding discretization:

— The **Order 2 L^2 AP scheme** is given by the discretization (3.36), (3.37) in time and (3.39) in space.

— The **Order 2 $L^{2,stab}$ AP scheme** is given by the discretization (3.36), (3.37) in time and (3.39), (3.43) in space.

The numerical test cases chosen are in two dimensions and involve the Full Euler equations or the Navier-Stokes equations in both compressible and incompressible regimes. Here is the list of the 2D numerical tests with their respective description:

1. **A 2D Riemann problem** (Section 3.4.1): A classical benchmark for two dimensional gas dynamics. It assesses the correct performance of the scheme for the full Euler equations in presence of contact discontinuities and shock waves. The simulations are run with the Order 2 L^2 and $L^{2,stab}$ AP schemes.
2. **An Explosion problem** (Section 3.4.2): An axi-symmetric flow is considered, for which a reference solution can be easily computed. The solution exhibits a contact discontinuity, a rarefaction wave, and a shock wave. The simulations are run considering the Euler equations with the Order 2 L^2 and $L^{2,stab}$ AP schemes and considering the Navier-Stokes equations ($\mu \neq 0$ and $\lambda = 0$) with the Order 2 L^2 AP scheme and the same scheme but with an explicit discretization of the viscous terms.
3. **The Gresho vortex vortex problem** (Section 3.4.3): A known stationary solution of both incompressible and compressible Euler equations. We illustrate the asymptotic properties of our Order 2 L^2 AP scheme in the incompressible regime for various small values of the Mach number. We also present the evolution of the kinetic energy for the Order 2 L^2 AP scheme when an unlimited slope is used instead of the minmod limiter.

4. **Smooth Gresho vortex: numerical convergence** (Section 3.4.4): Numerical convergence tests are run for the Euler equations on the Smooth Gresho vortex problem. The correct convergence rates are observed for the Order 2 L^2 and $L^{2,stab}$ AP schemes independently of independently of the Mach number regime. Moreover, as expected the errors do not depend on the Mach number for the Order 2 L^2 scheme whereas this not the case for the Order 2 $L^{2,stab}$ scheme.

5. **First problem of Stokes** (Section 3.4.5): An incompressible viscous fluid (with $\mu \neq 0$ and $\lambda = 0$) is considered for which an analytical solution is known. Simulations are run with the Order 2 L^2 AP scheme for various values of the viscosity coefficient.

6. **Double shear layer: Incompressible solution** (Section 3.4.6): This test case is used to validate the asymptotic consistency of the Order 2 L^2 AP scheme for well-prepared initial data. The vorticty contours plots are shown at different times for the full Euler system and for the Navier-Stokes system ($\mu \neq 0$ and $\lambda = 0$).

7. **Heat conduction** (Section 3.4.7): This test case is a simple test in which convection is created mainly by heat conduction ($\mu \neq 0$ and $\lambda \neq 0$). Results are shown for both Order 2 AP schemes.

8. **Lid-driven cavity flow: steady state incompressible solution** (Section 3.4.8): This test case consists of a square cavity filled with a viscous fluid ($\mu \neq 0$ and $\lambda = 0$) in an incompressible regime. The top moving wall brings the fluid into motion and a steady state is reached. Depending on the Reynolds number (inversely proportional to the viscosity coefficient), we can observe the formation of vortices. Results are given for various values of the Reynolds number with the Order 2 L^2 AP scheme and compare the results given by the Order 2 $L^{2,stab}$ AP scheme for a high Reynolds number.

If not mentioned, the specific gas constant R is set to 1 ($c_v = 2.5$), the adiabatic constant γ to 1.4 and the time step size is given by

$$\Delta t^n = C \frac{1}{\dfrac{\gamma \max_{i,j} |u^n_{i,j}|}{\Delta x} + \dfrac{\gamma \max_{i,j} |v^n_{i,j}|}{\Delta y}}, \tag{3.51}$$

where the constant $C = 0.45$.

3.4.1 2D Riemann problem

We consider a two-dimensional Riemann problem introduced in [72]. We set $\Omega = [-0.5, 0.5]^2$ and transmissive boundary conditions ($\partial W/\partial n = 0$). The initial data consists in four constant states defined in four quadrants. The initial constant states

are the following

$$(\rho, u, v, p)(0, x, y) = \begin{cases} (\ 1, 0.726, 0, 1\ ) & \text{if } x \leq 0\,, y > 0, \\ (\ 0.5313, 0, 0, 0.4\ ) & \text{if } x \geq 0\,, y > 0, \\ (\ 0.8, 0, 0, 1\ ) & \text{if } x \leq 0\,, y \leq 0, \\ (\ 1, 0, 0.726, 1\ ) & \text{if } x \geq 0\,, y \leq 0. \end{cases} \tag{3.52}$$

We set $\varepsilon = 1$, $\mu = \lambda = 0$ and the final time is $t_{end} = 0.25$.
This configuration is referred to as 2DR98 *"configuration F"* in [72] and *"configuration 12"* in [53]. For polytropic gas we can find between 15 and 19 different configurations such that, with constant initial states in each quadrant, a single elementary wave appears at each interface. The chosen configuration is constituted of two shocks moving respectively towards the right and top of the domain and two steady contact discontinuities in the bottom left part of the domain. The physical Mach number ranges between 0 and 1.14 in all the domain and for all times.
In Figure 3.4, we display the density contour plots and isolines at the final time computed with the Order 2 L^2 AP scheme (top) and the Order 2 $L^{2,stab}$ AP scheme (bottom). They are in good agreement with the reference solutions [72] and [53] (see Figure 3.3). With both schemes the contact discontinuities are preserved and do not move with time. Furthermore, the interface computed with the L^2 AP scheme (when no implicit diffusion is added) is much sharper. However, looking at the isolines we see that the L^2 AP presents some spurious oscillations (top right) that do not appear when the stabilization procedure is applied (bottom right).

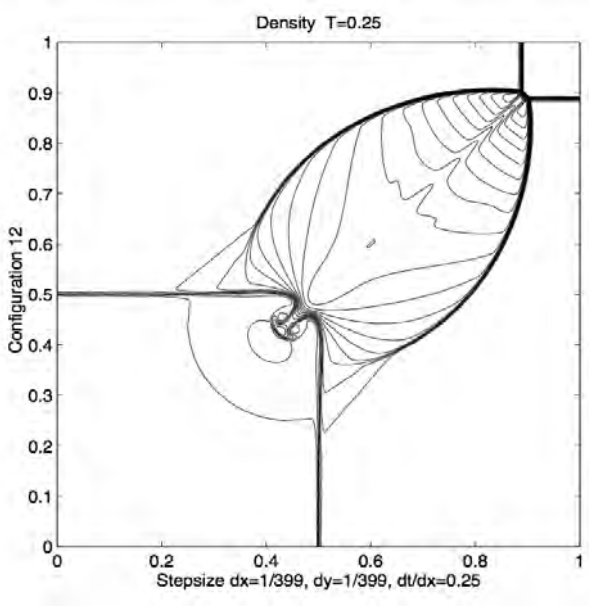


Figure 3.3 – 2D Riemann problem (Section 3.4.1): Density isolines. Reference solution *"configuration 12"* in [53].

In Figure 3.5, we present on the top left, the Mach number distribution calculated

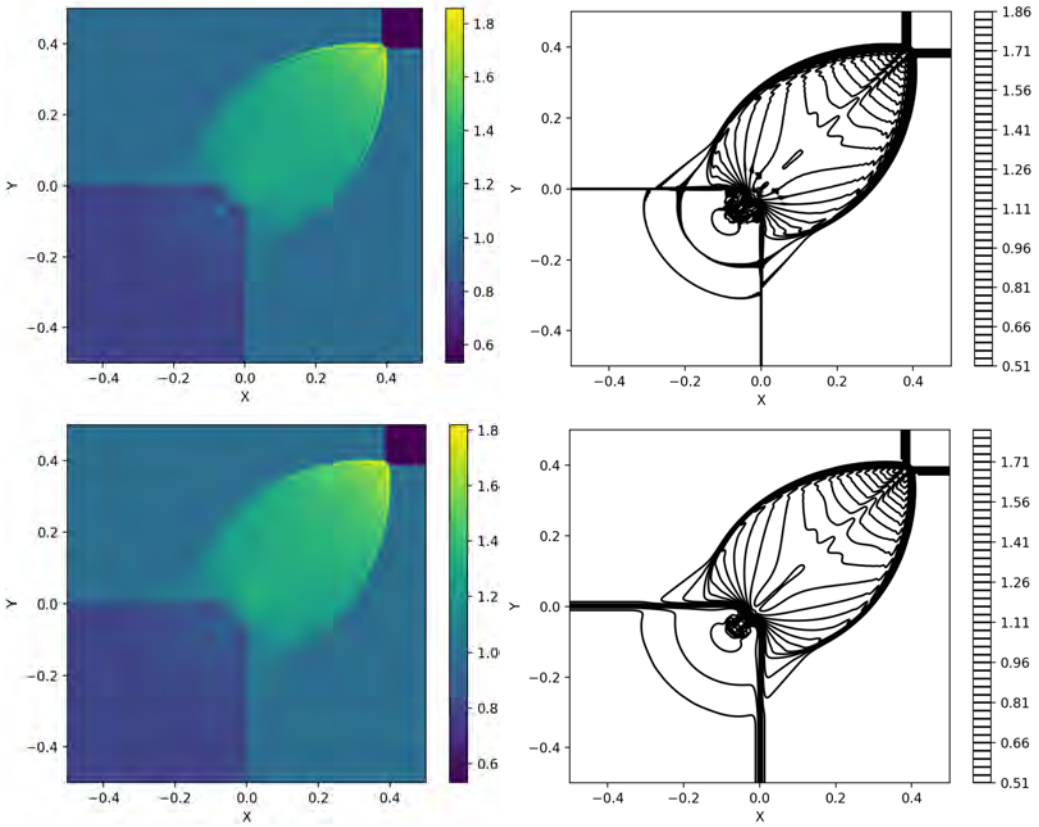


Figure 3.4 – 2D Riemann problem (Section 3.4.1): Density contour plots (left) and density isolines (right). Top panels: Order 2 L^2 AP scheme, bottom panels: Order 2 $L^{2,stab}$ AP scheme.

at each point by

$$M = \sqrt{\varepsilon}\frac{|U|}{c} = \frac{\sqrt{u^2+v^2}}{\sqrt{\frac{\gamma p}{\varepsilon\,\rho}}}. \tag{3.53}$$

We also confirm that the velocities (bottom) have stayed constant along the contact discontinuities.

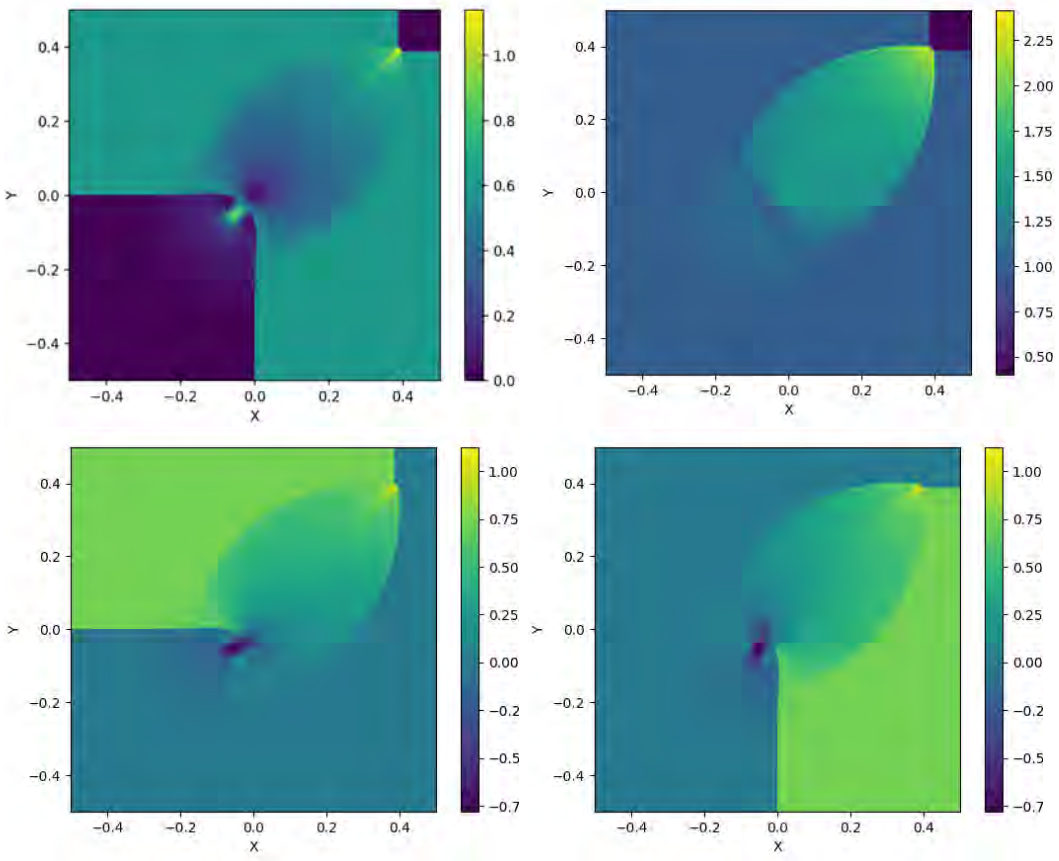


Figure 3.5 – 2D Riemann problem (Section 3.4.1): Physical Mach number (top left), pressure (top right), u velocity (bottom left) and v velocity (bottom right) contour plots at time $t = 0.25$ with the Order 2 L^2 AP scheme for 400×400 points.

3.4.2 Explosion problem

We consider the explosion problem proposed by [78] where we set $\Omega = [-1, 1]^2$ and Dirichlet boundary conditions given by the initial condition (3.54) during all the simulation. Initially the fluid is at rest, with a higher density and pressure at the center of the domain, inside the circle of radius 0.4 and centered in $(0, 0)$. The initial states are given by

$$(\rho, p)(0, x, y) = \begin{cases} (1,\ 1) & \text{if } r < 0.4, \\ (0.125,\ 0.1) & \text{otherwise,} \end{cases} \qquad (u, v)(0, x, y) = (0, 0), \tag{3.54}$$

where $r = \sqrt{x^2 + y^2}$, $\varepsilon = 1$, $\lambda = 0$, $t_{end} = 0.25$ and different values of μ. In the inviscid case, when $\mu = 0$, since the flow is axi-symmetric, the reference solution can be computed solving a one-dimensional inhomogeneous system (see Chapter 17 in [78] for more details). We solve it using an order two explicit scheme with a Rusanov-type solver on a refined grid ($N = 5000$).

In Figure 3.6, we display the results for the Order 2 L^2 and $L^{2,stab}$ AP schemes on a 100×100 grid for the inviscid case. On the top left, we show the physical Mach number distribution for the Order 2 L^2 AP scheme calculated at each point by (3.53). The other three are one-dimensional radial cuts along the x-axis for respectively the density, u velocity and pressure at time $t = 0.25$ for both schemes. We see that both schemes are able to correctly catch the shock front going towards the outside ($x = 0.8$), the contact discontinuity (around $x = 0.6$) and rarefecation wave propagating towards the origin. Around the shock, the L^2 scheme presents some oscillations and they are only slightly reduced by the $L^{2,stab}$ scheme. Since the stabilization proposed is not sufficient for this problem, a solution would be to apply the MOOD procedure proposed in Section 2.4.5.

In Figure 3.7, we validate the correct implicit treatment of the viscous part. We compare the results given by the Order 2 L^2 AP scheme (squares) against the Order 2 L^2 AP scheme but with an explicit discretization of the viscous part (crosses). Results are displayed for $\mu = 10^{-2}$ and 10^{-3}. Looking at the profiles, we observe that in both cases we obtain the same results independently of the discretization. Moreover, Table 3.1 illustrates the advantage on the time step size of treating implicitly the viscous part. For $\mu = 10^{-2}$, the implicit discretization allows a time step that is 5.6 times bigger than with an explicit treatment.

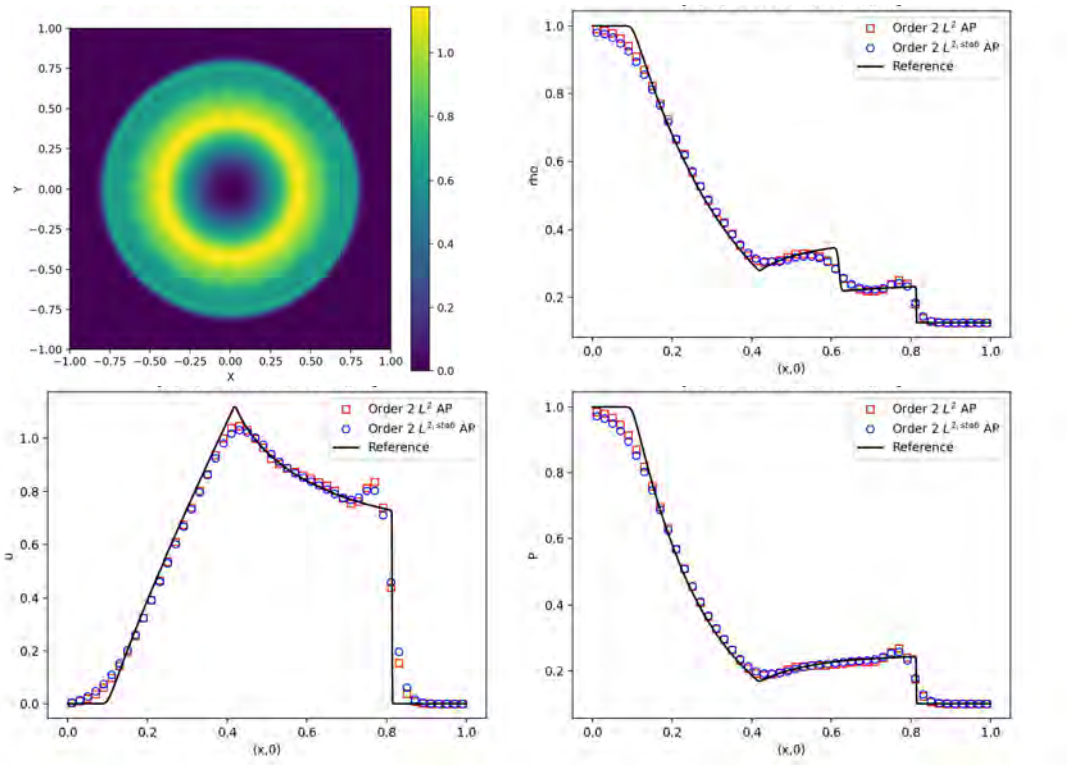


Figure 3.6 – Explosion problem (Section 3.4.2): Comparison of the Order 2 L^2 and $L^{2,stab}$ AP schemes with the reference solution for the Euler equations on a 100×100 grid. Top left: Mach number distribution (with the L^2 scheme). Others: one-dimensional radial cuts along the x-axis for respectively the density (top right), the component u of the velocity (bottom left) and the pressure (bottom right) at time $t = 0.25$.

Scheme	Number of times steps $\mu = 0$	$\mu = 10^{-3}$	$\mu = 10^{-2}$
Order 2 L^2 AP	73	69	58
Order 2 L^2 AP (visc. exp)	73	96	330

Table 3.1 – Explosion problem (Section 3.4.2): Number of time steps for various viscous regimes on a 101×101 grid. Comparison between an implicit and explicit discretization of the viscous flux for the Order 2 L^2 AP scheme.

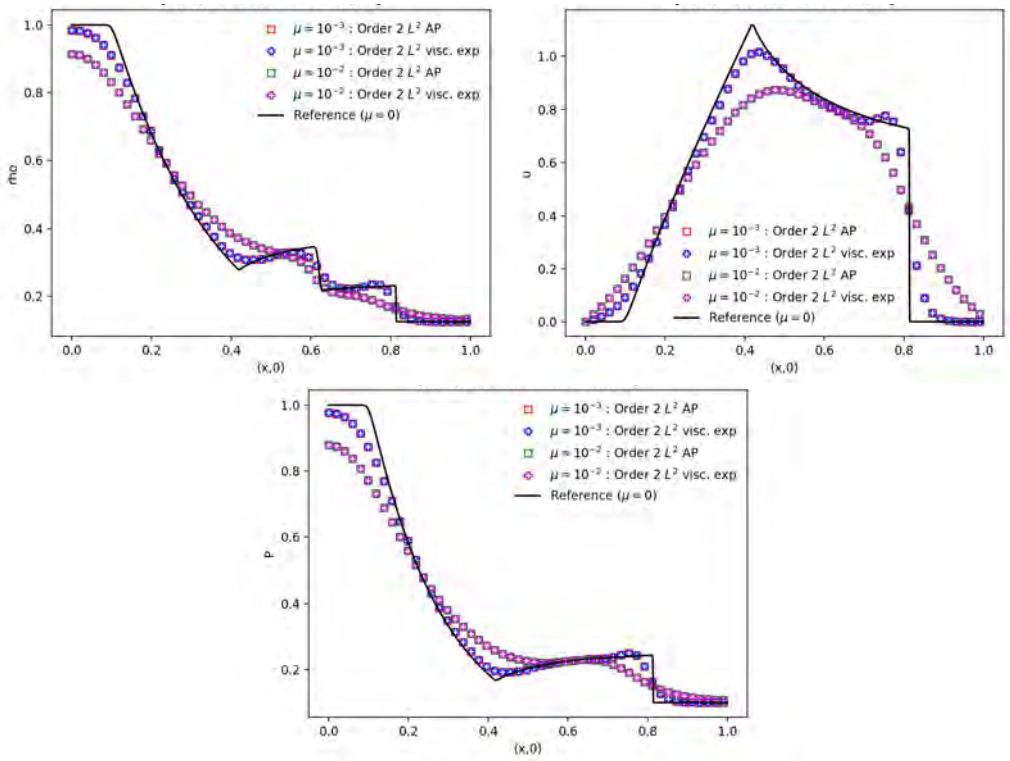


Figure 3.7 – Explosion problem (Section 3.4.2): u velocity (top left) , v velocity (top right) and pressure (bottom) one-dimensional cuts at time $t = 0.25$ and 100×100 points. Comparison between an implicit and explicit discretization of the viscous flux for the Order 2 L^2 AP scheme.

3.4.3 Gresho vortex: AP properties

We solve the Gresho vortex problem that is a known stationary solution of both incompressible and compressible Euler equations. Here we consider the modified setup [62] used to check the ability of the numerical scheme to handle low Mach number flows. The solution, written in polar coordinates, reads

$$\begin{aligned} u_\phi(r) &= \begin{cases} 5r & 0 \le r < 0.2, \\ 2 - 5r & 0.2 \le r < 0.4, \\ 0 & r \ge 0.4, \end{cases} \\ p(r) &= \begin{cases} p_0 + 12.5r^2 & 0 \le r < 0.2, \\ p_0 + 12.5r^2 + 4[1 - 5r - ln(0.2) + ln(r)] & 0.2 \le r < 0.4, \\ p_0 - 2 + 4ln(2) & r \ge 0.4, \end{cases} \\ \rho(r) &= 1, \end{aligned} \tag{3.55}$$

where $u_\phi(r)$ is the angular velocity, $r = \sqrt{(x-0.5)^2 + (y-0.5)^2}$ is the radius on the computational domain $\Omega = [0,1] \times [0,1]$ and

$$p_0 = \frac{\rho}{\gamma M^2}, \tag{3.56}$$

is expressed in terms of the Mach number. The density is constant ($\rho = 1$) and the divergence free velocity field can be obtained from u_ϕ as

$$(u, v) = u_\phi.(-\sin(\phi), \cos(\phi)) \qquad \text{with } \phi = \arctan\left(\frac{y - y_0}{x - x_0}\right).$$

At last, we set $\varepsilon = 1$, $\mu = \lambda = 0$ and periodic boundary conditions.
In Figure 3.8, we show the physical Mach number distribution with the the Order 2 L^2 AP scheme for different values of M after a full turn of the vortex at $t = 0.4\pi$. We see that our scheme is able to preserve the initial distribution (first subfigure) independently of the Mach number regime unlike classical discretizations for which the dissipation is related to M.
To further check the asymptotic accuracy of our scheme we show in Figure 3.9 the ratio between the kinetic energy at each time step $k(t)$ and the initial kinetic energy $k(0)$ for the Mach numbers $M = 10^{-1}, 10^{-2}, 10^{-3}$ and two grid resolutions 40×40 and 80×80. The results are given for the Order 2 L^2 AP scheme with no limiter in space (left) and with the *minmod* limiter (right). We see in the graphs that for a same grid resolution the lines for the different Mach numbers are overlapping which shows that the loss of kinetic energy is independent of the chosen Mach regime. Comparing the two subfigures we see that the loss is mostly due to the limiter in space, 0.9825 (unlimited) against 0.92 (limited) for the coarser grid. The Order 2 unlimited scheme consists in replacing in the linear reconstruction (3.38) the limited slopes $\alpha^n_{x_{i,j}}$ and $\alpha^n_{y_{i,j}}$ by

$$\begin{aligned} \alpha^n_{x_{i,j}} &= \frac{1}{2}\left(\frac{W^n_{i+1,j} - W^n_{i,j}}{\Delta x} + \frac{W^n_{i,j} - W^n_{i-1,j}}{\Delta x}\right), \\ \alpha^n_{y_{i,j}} &= \frac{1}{2}\left(\frac{W^n_{i,j+1} - W^n_{i,j}}{\Delta y} + \frac{W^n_{i,j} - W^n_{i,j-1}}{\Delta y}\right). \end{aligned} \tag{3.57}$$

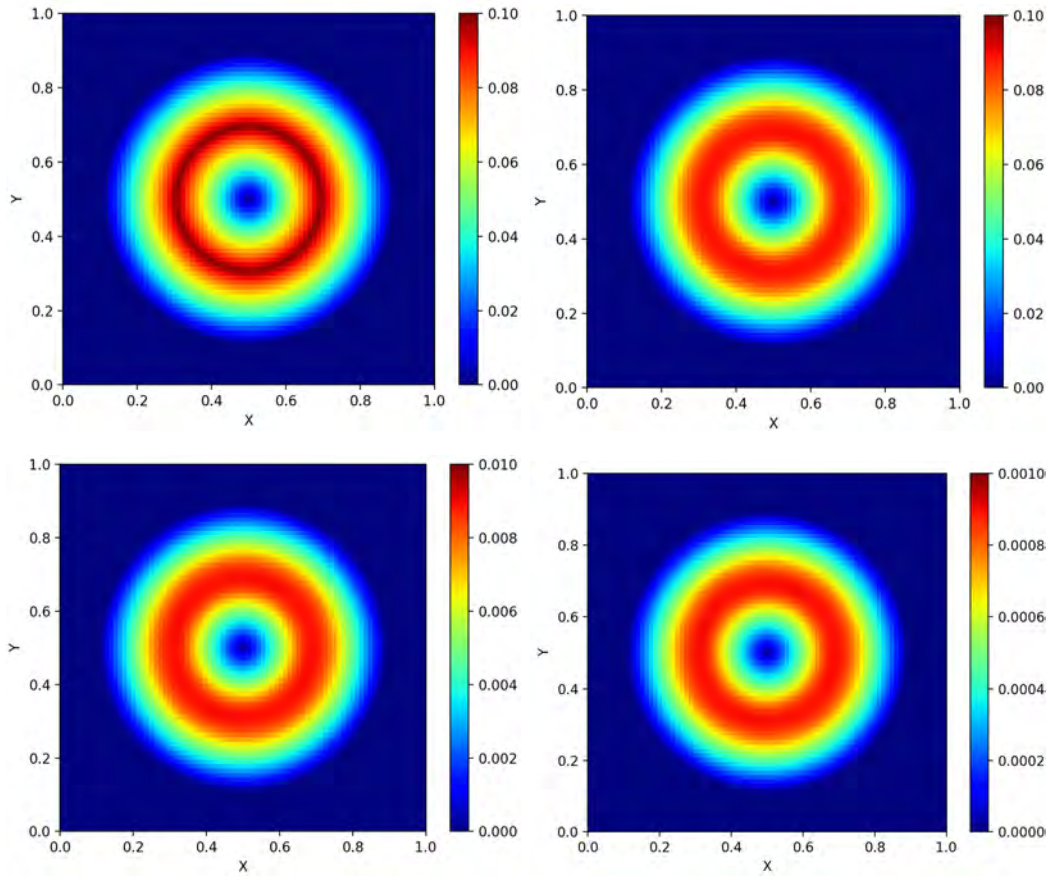


Figure 3.8 – Gresho vortex (Section 3.4.3): Initial Mach number distribution for $M = 10^{-1}$ (top left) and at time $T = 0.4\pi$ with the Order 2 L^2 AP scheme and 80×80 points for $M = 10^{-1}$ (top right), $M = 10^{-2}$ (bottom left) and $M = 10^{-3}$ (bottom right).

In Figure 3.10, we focus on the pressure profile in the x and y direction and compare it against the initial condition. We can see that, even for $M = 10^{-3}$, our scheme is able to capture the pressure perturbations and does not show any oscillation. It is worth mentioning that even on a coarser grid, (Figure 3.11), the initial distribution is maintained (left figure). Let us note that when using the limiter minmod, it is important to correctly initialize the data (here it is correctly initialized choosing an odd number of cells). If not, the calculated slopes may not preserve the circular symmetry of the problem and disrupt the initial solution (middle figure). As illustrated on the right subfigure, this is not the case for an unlimited slope (3.57). In Figure 3.12, the contour plots show that the density (left) remains constant except for some oscillations around $r = 0.4$. The middle and right subfigures are the velocity distribution in x and y direction respectively.

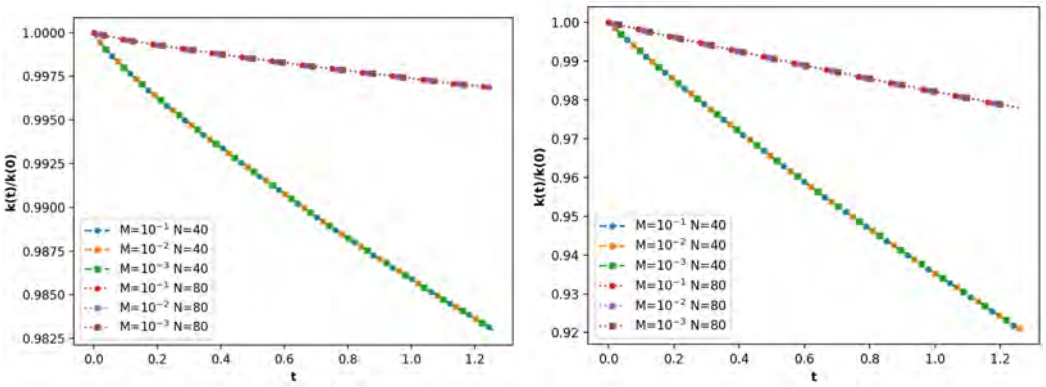
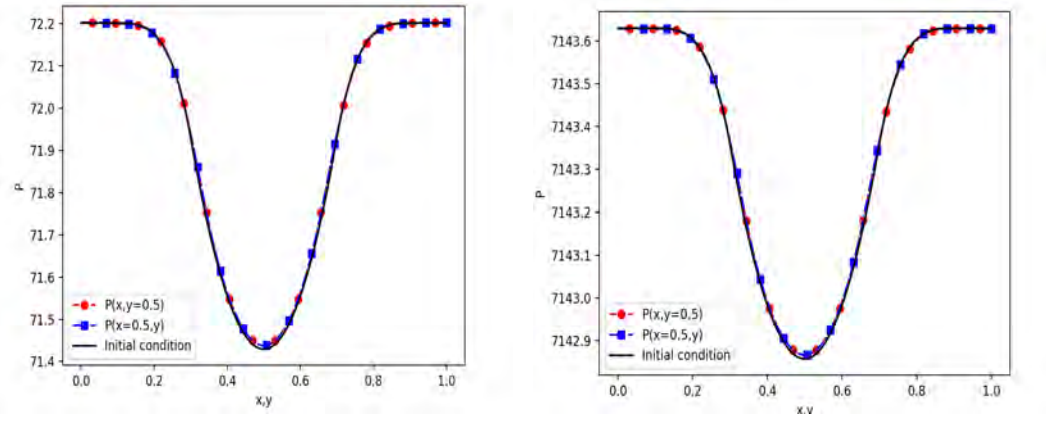


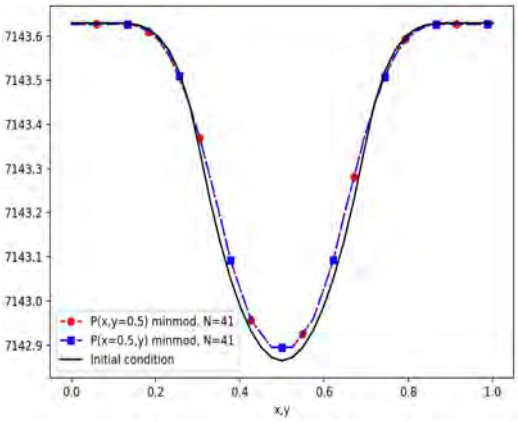
Figure 3.9 – Gresho vortex (Section 3.4.3): Evolution of the kinetic energy with the Order 2 L^2 AP scheme for the Mach numbers $M = 10^{-1}, 10^{-2}, 10^{-3}$ and $N \times N$ points. Left: Order 2 unlimited AP scheme, Right: Order 2 limited AP scheme.



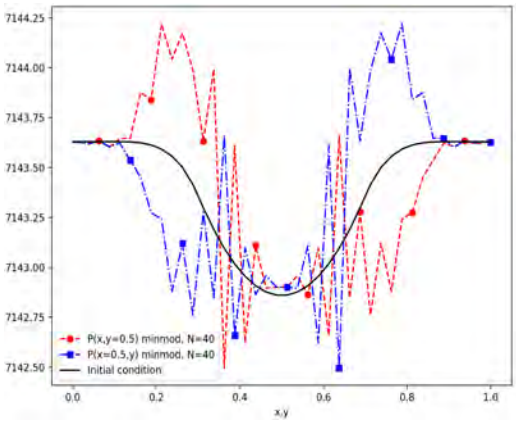
(a) $M = 10^{-1}$ (b) $M = 10^{-2}$

(c) $M = 10^{-3}$

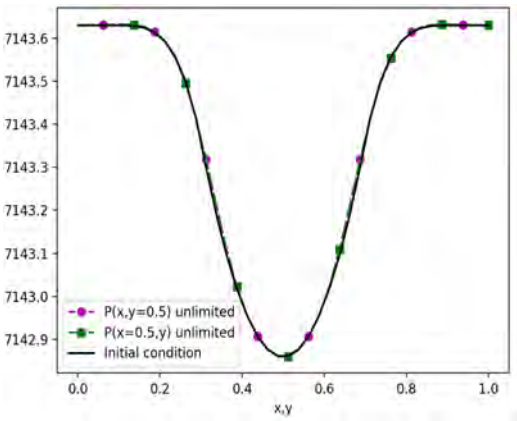
Figure 3.10 – Gresho vortex (Section 3.4.3): Pressure profile in the x and y direction at time $T = 0.4\pi$ against the initial profile for the Mach numbers $M = 10^{-1}, 10^{-2}, 10^{-3}$.



(a) limited, $N = 41$



(b) limited, $N = 40$



(c) unlimited, $N = 40$

Figure 3.11 – Gresho vortex (Section 3.4.3): Pressure profile in the x and y direction at time $T = 0.4\pi$ against the initial profile for the Mach number $M = 10^{-2}$. Left and middle panels: Limited scheme with 41×41 versus 40×40 points. Right panel: Unlimited scheme with 40×40 points.

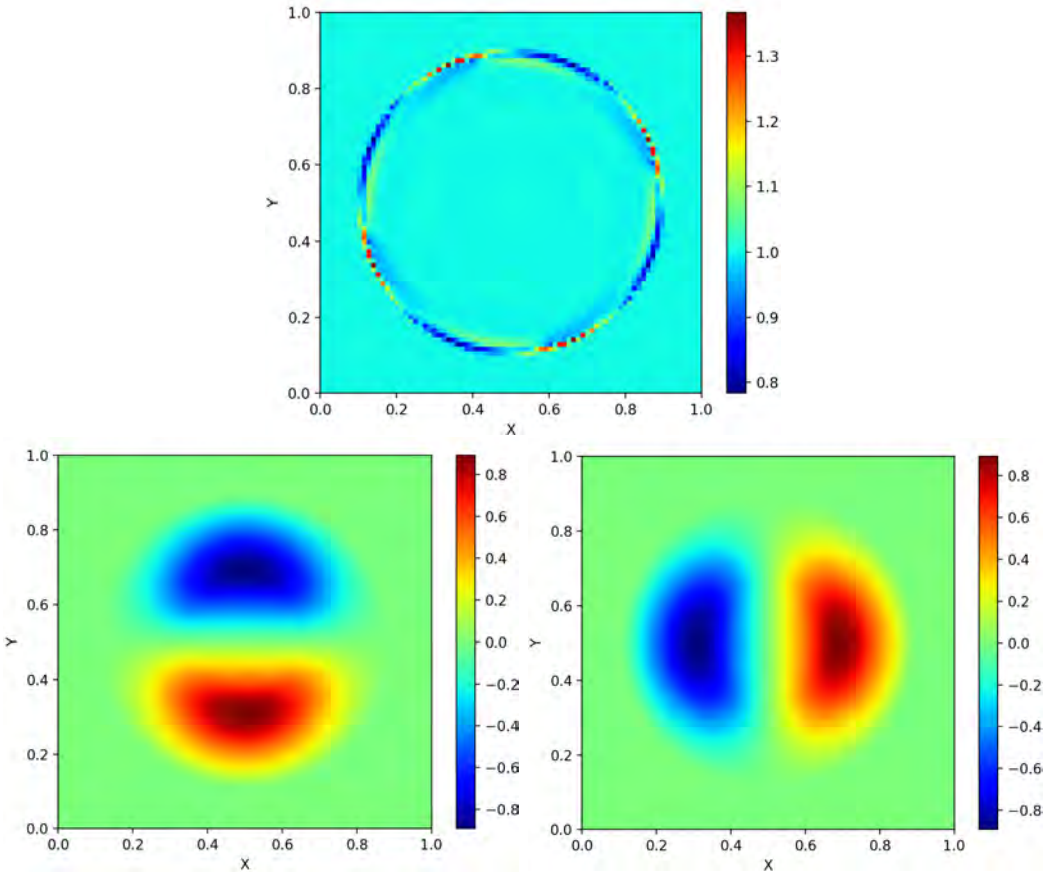


Figure 3.12 – Gresho vortex (Section 3.4.3): Density (top) and velocity (left: u, right: v) contours at time $T = 0.4\pi$ for the Mach number $M = 10^{-3}$ with 80×80 points.

3.4.4 Smooth Gresho vortex: numerical convergence

In order to validate the second order accuracy of our scheme we propose to consider a smooth version of the Gresho vortex introduced by [77]. A smoothed velocity profile is proposed with an angular velocity that this time is twice continuously differentiable. Then, the pressure profile is calculated with this new velocity in order to get a stationary vortex. The solution, written in polar coordinates, reads

$$\begin{aligned} u_\phi(r) &= \begin{cases} 75r^2 - 250r^3 & 0 \le r < 0.2, \\ -4 + 60r - 225r^2 + 250r^3 & 0.2 \le r < 0.4, \\ 0 & r \ge 0.4, \end{cases} \\ p(r) &= \begin{cases} p_0 + 1406.5r^4 - 7500r^5 + (10416 + \frac{2}{3})r^6 & 0 \le r < 0.2, \\ p_0 + p_2(r) & 0.2 \le r < 0.4, \\ p_0 + p_2(r) & r \ge 0.4, \end{cases} \\ \rho(r) &= 1, \end{aligned} \tag{3.58}$$

where $u_\phi(r)$ is the angular velocity, $r = \sqrt{(x-0.5)^2 + (y-0.5)^2}$ is the radius on the computational domain $\Omega = [0,1] \times [0,1]$,

$$p_0 = \frac{\rho}{\gamma M^2},$$

and

$$\begin{aligned} p_2(r) = 65.8843399322788 - 480r + 2700r^2 - (9666 + \frac{2}{3})r^3 + 20156.25r^4 \\ - 22500r^5 + (10416 + \frac{2}{3})r^6 + 16ln(r). \end{aligned} \tag{3.59}$$

We set $\varepsilon = 1$, $\mu = \lambda = 0$ and periodic boundary conditions everywhere.
To assess the numerical order of accuracy, we compute the L^1 errors for several uniform meshes. The error $L1_w$ for a variable w is computed as the ratio between the error of the scheme at the final time T and the exact solution (here the initial configuration):

$$L1_w = \frac{|w(x,y,T) - w(x,y,0)|_{L1}}{|w(x,y,0)|_{L1}} = \frac{\sum_{i,j} |w_{i,j}(T) - w_{i,j}(0)|}{\sum_{i,j} |w_{i,j}(0)|}. \tag{3.60}$$

We present in Table 3.2 for each conservative variable w and various Mach regimes, the errors $L1_w$ and convergence rates for the Order 2 L^2 AP scheme. The solution is smooth so, for the computation there is no limiter in space, the unlimited slope is given by (3.57).
The errors for the density and momentum are as expected independent of the Mach number regime and the desired convergence rates are reached with a small drop on the density. The errors for the energy are very small and they reach for the lowest Mach number regimes the tolerance of the linear solver used for the pressure equation.

M	N	$L1_\rho$	EOC_ρ	$L1_{\rho u}$	$EOC_{\rho u}$	$L1_{\rho v}$	$EOC_{\rho v}$	$L1_E$	EOC_E
10^{-1}	20	5.592e-02	-	1.448e-01	-	1.448e-01	-	2.163e-04	-
	40	1.637e-02	1.77	1.637e-02	2.26	3.029e-02	2.26	5.620e-05	1.95
	80	4.072e-03	2.00	4.739e-03	2.68	4.739e-03	2.68	1.932e-05	1.54
	160	1.028e-03	1.99	8.882e-04	2.42	8.882e-04	2.42	7.671e-06	1.33
10^{-2}	20	5.586e-02	-	1.451e-01	-	1.451e-01	-	8.704e-06	-
	40	1.636e-02	1.77	2.986e-02	2.28	2.986e-02	2.28	1.106e-06	2.98
	80	4.074e-03	2.01	4.831e-03	2.63	4.831e-03	2.63	1.279e-07	3.11
	160	1.026e-03	1.99	8.759e-04	2.46	8.759e-04	2.46	2.145e-08	2.58
10^{-3}	20	5.588e-02	-	1.451e-01	-	1.451e-01	-	9.579e-06	-
	40	1.637e-02	1.77	2.988e-02	2.28	2.988e-02	2.28	3.986e-07	4.59
	80	4.075e-03	2.01	4.823e-03	2.63	4.823e-03	2.63	8.008e-08	2.32
	160	1.026e-03	1.99	8.759e-04	2.46	8.759e-04	2.46	1.118e-08	2.84

Table 3.2 – Smooth Gresho vortex (Section 3.4.4): Convergence table at $T = 0.4\pi$ and $N \times N$ points for the Order 2 L^2 AP scheme. Errors given for the Mach numbers $M = 10^{-1}, 10^{-2}$ and 10^{-3}.

Let us now compare the L^1 error plots of the Order 2 L^2 AP scheme and the Order 2 $L^{2,stab}$ AP scheme. The second one being the scheme where we added an upwinding on the implicit part related to the Mach number (3.43) to reduce the oscillations.
In Figures 3.13-3.15, we show the results for the density and momentums in the x and y direction. In each case, the Order 2 scheme with no stabilization (left), reaches the expected converge rates and the errors do not depend on the Mach number regime, the error lines are overlapping. When we add the stabilization (right), we observe that the expected convergence rates are finally reached for finer meshes independently of the asymptotic regime. Moreover, the accuracy of the scheme depends on the Mach number which is consistent since the upwinding is higher as the Mach number decreases. In particular, from Figure 3.14 we have that for 320×320 points, taking $M_1 = 10^{-1}$ and $M_3 = 10^{-3}$ the ratio $L1_{\rho u}(M_1)/L1_{\rho u}(M_3) = \mathcal{O}(M_3/M_1) = \mathcal{O}(10^{-2})$.
Concerning the error plots for the energy and pressure in Figure 3.16, the error becomes very small and reaches the tolerance of the solver used for the pressure equation. For these variables, the errors for the Order 2 L^2 AP scheme also depend on the Mach number regimes.

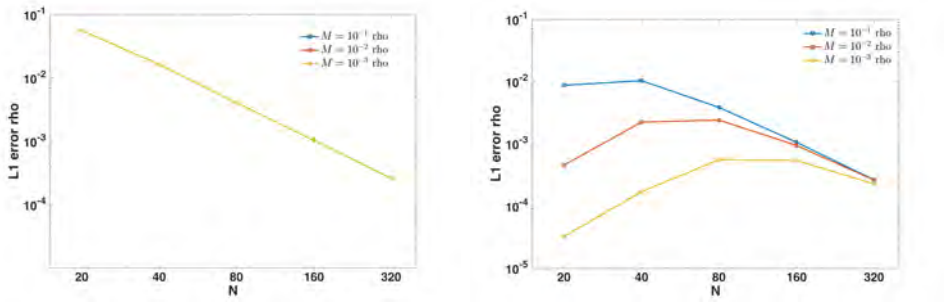


Figure 3.13 – Smooth Gresho vortex (Section 3.4.4): L^1 errors for ρ. Order 2 L^2 AP scheme (left) versus Order 2 $L^{2,stab}$ AP scheme (right).

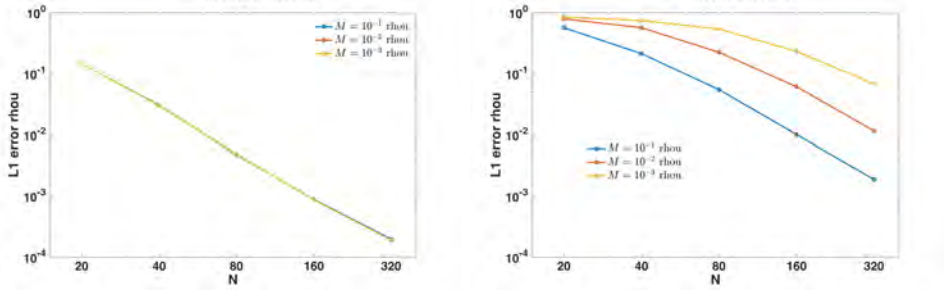


Figure 3.14 – Smooth Gresho vortex (Section 3.4.4): L^1 errors for ρu. Order 2 L^2 AP scheme (left) and Order 2 $L^{2,stab}$ AP scheme (right).

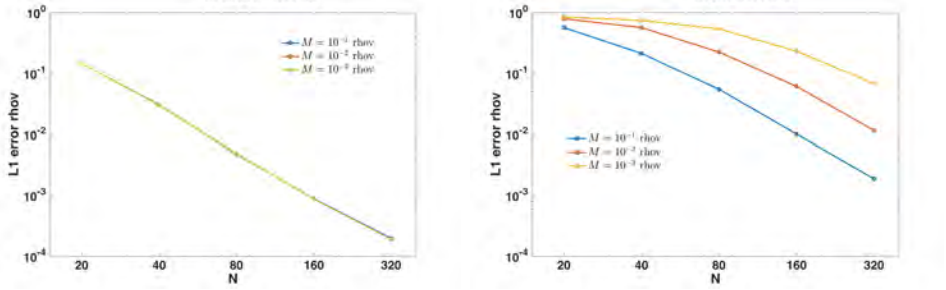


Figure 3.15 – Smooth Gresho vortex (Section 3.4.4): L^1 errors for ρv. Order 2 L^2 AP scheme (left) and Order 2 $L^{2,stab}$ AP scheme (right).

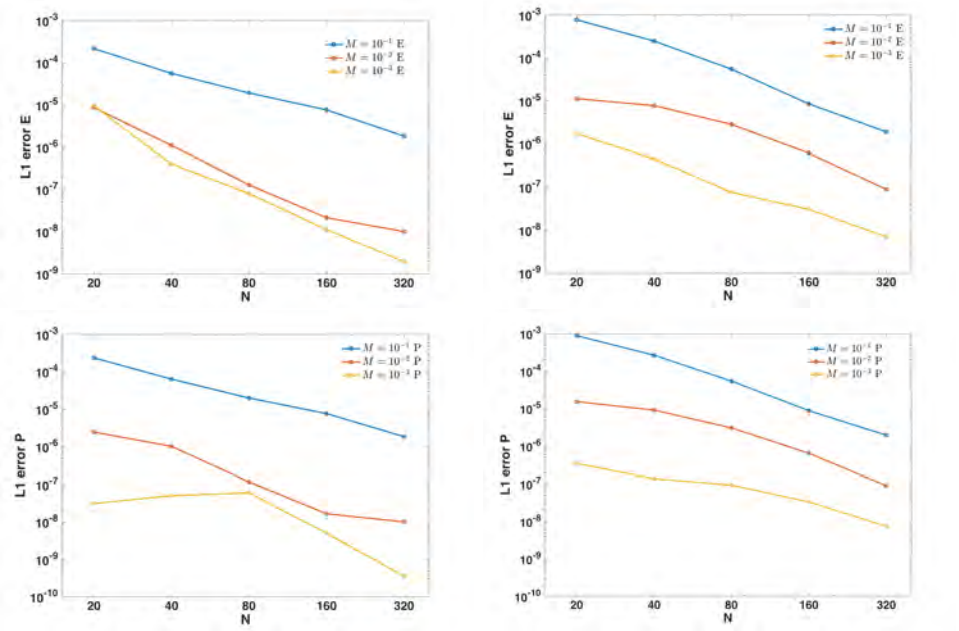


Figure 3.16 – Smooth Gresho vortex (Section 3.4.4): L^1 errors for E and p. Order 2 L^2 AP scheme (left) and Order 2 $L^{2,stab}$ AP scheme (right).

3.4.5 First problem of Stokes

We simulate a test problem for which an analytical solution is known. In Stokes's first problem [74] we consider an incompressible viscous fluid in a semi-infinite plate. The fluid is at rest and then suddenly a constant velocity is set at $y = 0$ (see definition sketch in Figure 3.17). The fluid is brought into motion by the action of the viscous stress at the bottom. Here we set $\Omega = [0, 1] \times [0, 2]$ and the initial data are given by:

$$\rho(0, x, y) = 1, \quad u(0, x, y) = \begin{cases} U & \text{if } y = 0, \\ 0 & \text{otherwise,} \end{cases} \quad v(0, x, y) = 0, \quad p(0, x, y) = 1. \tag{3.61}$$

with $U = 1$, $\varepsilon = 10^{-6}$ and $\lambda = 0$. For a semi-infinite plate, the exact solution of the problem is

$$u(x, y, t) = U \left(1 - erf \left(\frac{y}{2\sqrt{\gamma \mu}} \right) \right), \tag{3.62}$$

with constant density, pressure and velocity v.

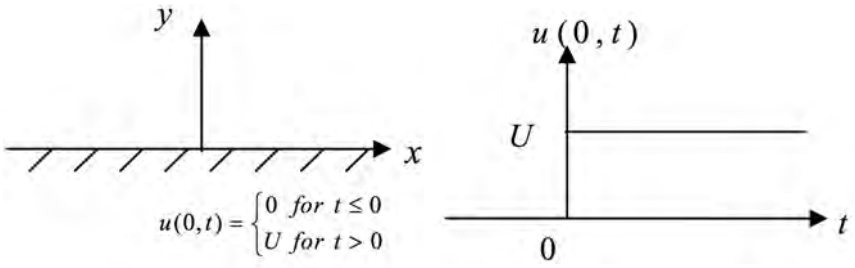


Figure 3.17 – Stokes' first problem (Section 3.4.5): Definition sketch of the u velocity at the wall [22].

For the simulations we consider periodic boundary conditions on the x direction and Dirichlet boundary conditions on the y direction. In particular we set the exact solution given by (3.62) for the velocity u and the initial condition for the other variables.
We run the simulations with the Order 2 L^2 AP scheme until the final time $T = 30$ with 5×100 points and for the viscosity coefficients $\mu = 10^{-1}$, $\mu = 10^{-2}$ and $\mu = 10^{-3}$. In Figure 3.18, we show for each coefficient the u velocity contour plot at the final time (left) and compare our solution against the exact solution at different time levels (right). We compare it plotting the u velocity versus the wall distance. As expected, the disturbance caused by the impulsive motion of the boundary diffuses into the fluid as time progresses and faster when μ is bigger.
In Figure 3.19, we present the density, the component of the velocity v and the pressure profiles for $\mu = 10^{-2}$ at the final time $T = 30$ and we compare it with the incompressible solution. We see that our scheme is able to preserve the constant states of the limit model for each variable up to a satisfying error precision.

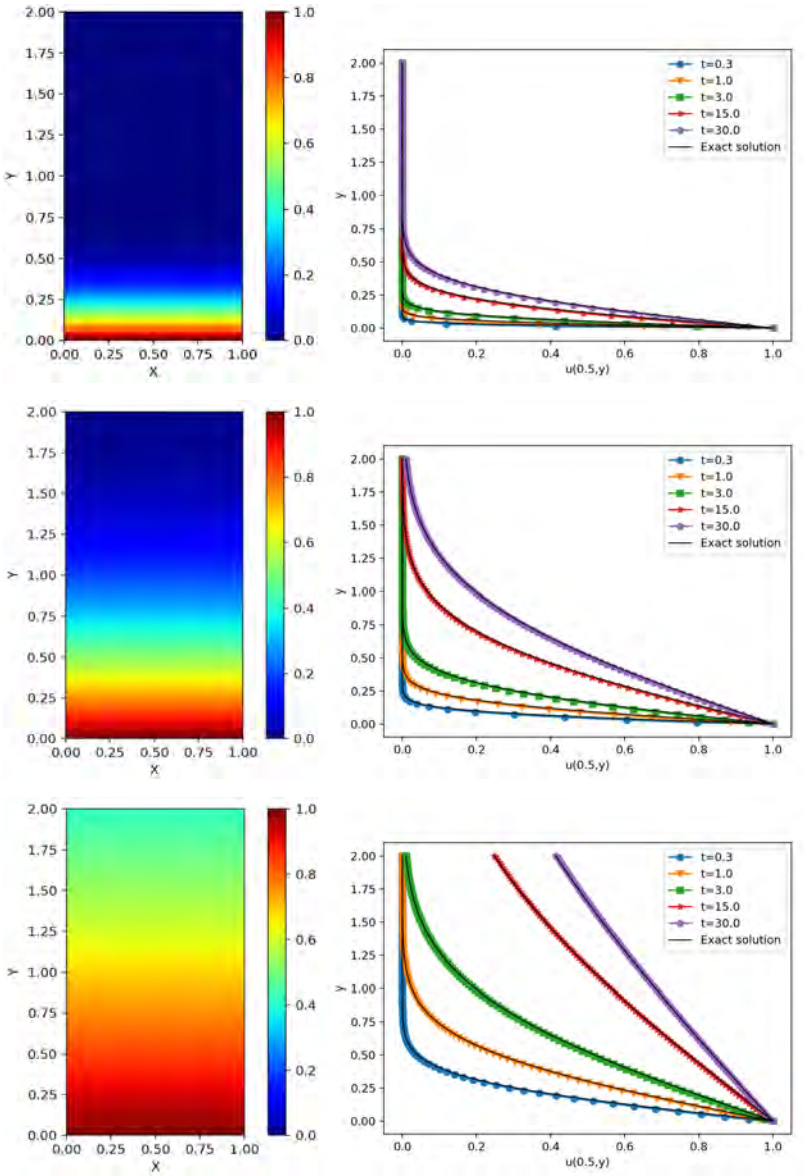


Figure 3.18 – Stokes' first problem (Section 3.4.5): Results for the viscosity coefficients $\mu = 10^{-3}$ (top), $\mu = 10^{-2}$ (middle) and $\mu = 10^{-1}$ (bottom) with 5×100 points. Left: u-distribution contour plot at $T = 30$. Right: Comparison of the u velocity versus the wall distance against the exact solution at times $t = 0.3, 1.0, 15$ and 30.

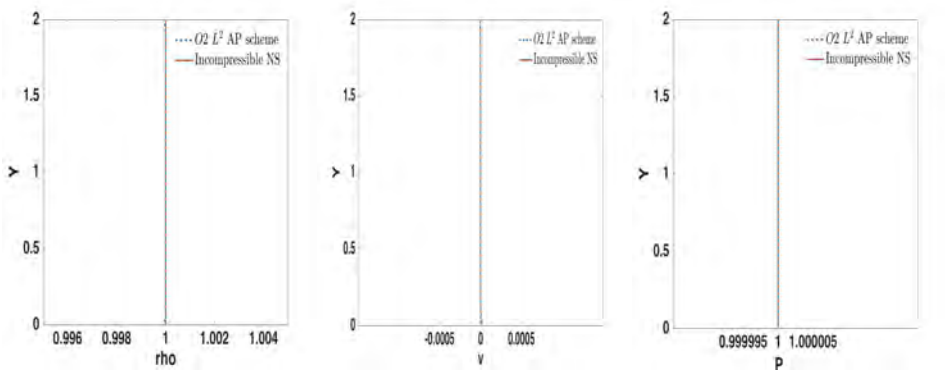


Figure 3.19 – Stokes' first problem (Section 3.4.5): Density (left) , v velocity (middle) and pressure (right) profiles for $\mu = 10^{-2}$ at time $t = 30$ with 5×100 points. The three variables stay constant across time.

3.4.6 Double shear layer: Incompressible solution

We consider a test case studied in [7] which consists of a double shear layer in a periodic domain. It is used to validate the asymptotic consistency of our scheme since for small values of ε we can compare our results with a reference solution computed solving the incompressible Navier-Stokes equations. We set $\Omega = [0,1]^2$ and periodic boundary conditions everywhere. The initial data are well prepared to the incompressible regime (divergence free velocity field and constant pressure) and are given by:

$$\begin{aligned} &\rho(0,x,y) = 1, \qquad u(0,x,y) = \begin{cases} \tanh(30(y-0.25)) & \text{if } y < 0.5, \\ \tanh(30(0.75-y)) & \text{otherwise,} \end{cases} \\ &v(0,x,y) = 0.05\sin(2\pi x), \qquad p(0,x,y) = 1, \end{aligned} \tag{3.63}$$

where $\lambda = 0$. The shear layer is initially perturbed by a vertical velocity of small amplitude. Then, each of the layers will evolve into large vortices and will be thinned between those rolls. One relevant quantity is the vorticity w

$$w = \partial_x v - \partial_y u,$$

which we compute using a second order difference approximation.

$$w_{i,j} = \frac{v_{i+1,j} - v_{i-1,j}}{2\Delta x} - \frac{u_{i,j+1} - u_{i,j-1}}{2\Delta y}. \tag{3.64}$$

As a reference solution for the incompressible Euler equations we can take the results given by [7] for 128×128 and 256×256 grid points in Figures 1 and 2 respectively at times $t = 0.4, 0.8, 1.2$ and 1.8 (see Figure 3.20 for a snapshot of Figure 2 from [7]).
In Figure 3.21, we show the contour vorticity plots for the full Euler equations, i.e., $\mu = 0$, at time $t = 1.2$ for decreasing values of ε. We observe that for large values of ε the scheme does not capture the incompressible solution while for $\varepsilon = 10^{-3}$ and 10^{-6} the results are in very good agreement. Therefore, for the following simulations we set $\varepsilon = 10^{-3}$.
In Figure 3.22, we present the vorticity evolution for the full Euler equations on two grid resolutions. We see that the main structure of the incompressible solution is captured. On the coarser grid (top), we observe on the thinned layer between the rolls the smearing of the solution while by using a finer grid (bottom) the smearing is significantly minimized.
In Figure 3.23, we present on the finer grid the contour plots of the physical variables at time $t = 1.2$ with no viscosity. On the top left corner we see that the initial constant distribution for the density is maintained except for some local oscillations. For the pressure (top right corner), the fluctuations on the domain are of order of the square Mach number, the Mach number varying in the range 0 to 0.0288. At the bottom, we find the velocity components u (left) and v (right);
We consider now the Navier-Stokes equations where we set the viscosity coefficient $\mu = 2.10^{-4}$. In Figure 3.24, the fundamental structure of the vorticity field is relatively similar to the Euler case (see Figure 3.22) with a slightly more diffusive structure and no smearing even on the coarser grid.

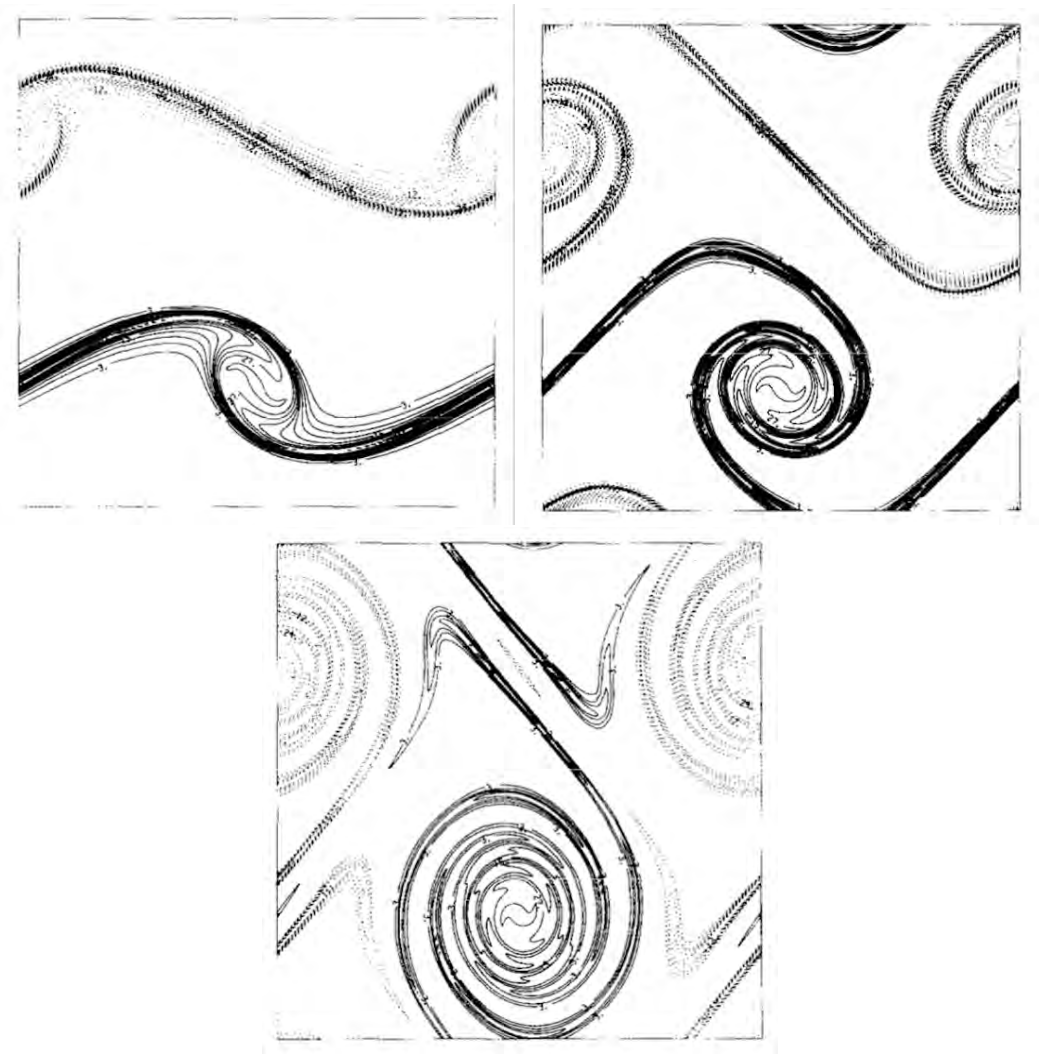


Figure 3.20 – Double shear layer (Section 3.4.6): Reference solution [7]: Vorticity contours for the full Euler equations at times $t = 0.8$ (top left), $t = 1.2$ (top right) and $t = 1.8$ (bottom) on a 256×256 grid.

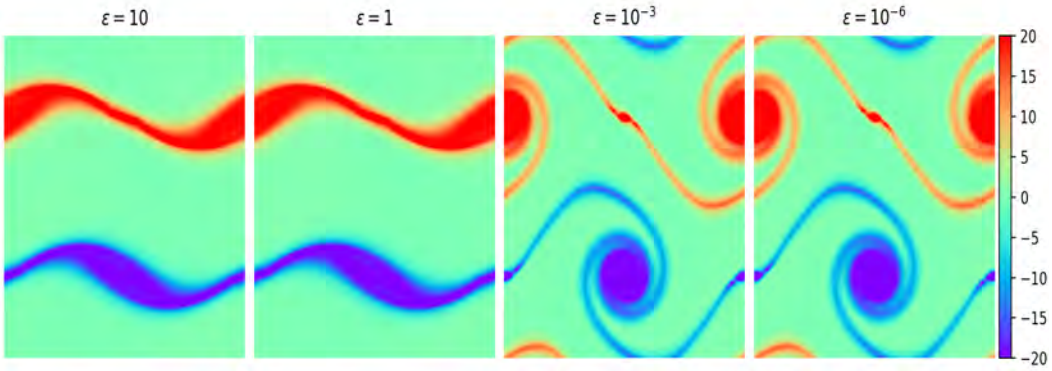


Figure 3.21 – Double shear layer (Section 3.4.6): Vorticity contours for the full Euler equations, i.e., $\mu = 0$, with the Order 2 L^2 AP scheme on a 128×128 grid at time $t = 1.2$ for decreasing values of ε.

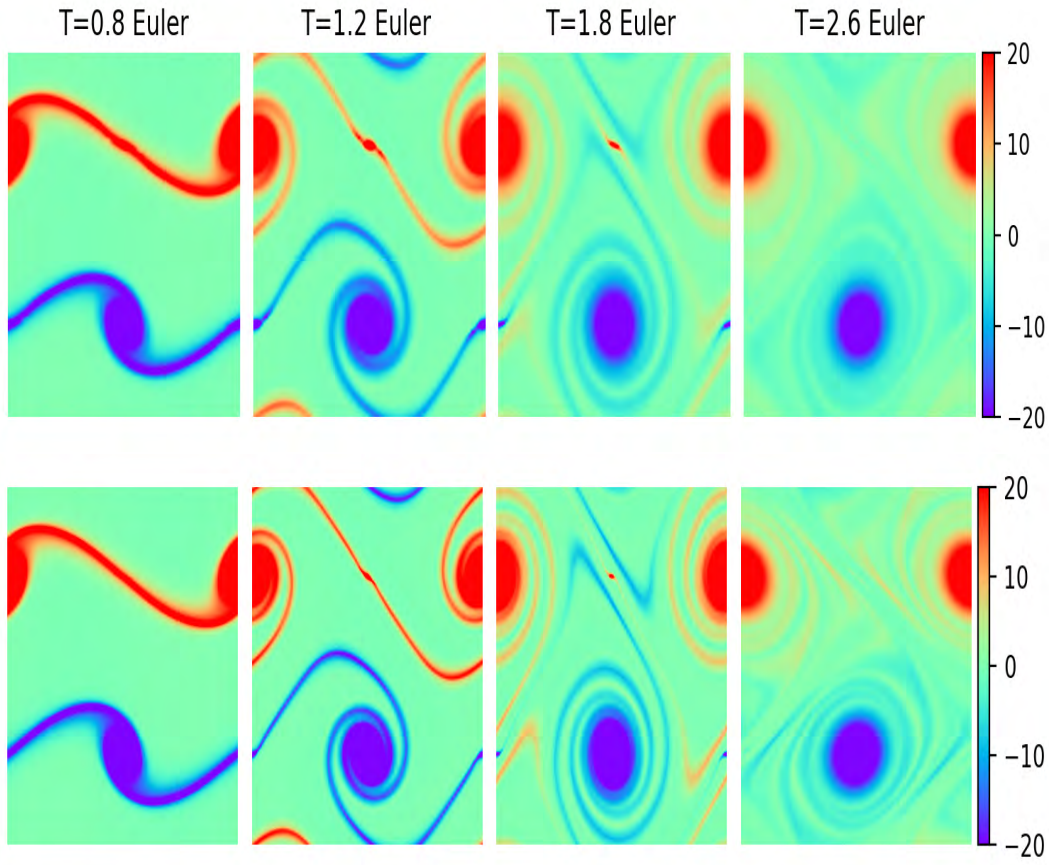


Figure 3.22 – Double shear layer (Section 3.4.6): Vorticity contours for the full Euler equations with the Order 2 L^2 AP scheme setting $\varepsilon = 10^{-3}$ at times $t = 0.8$, 1.2, 1.8 and 2.6 on a 128×128 (top) and a 256×256 grid.

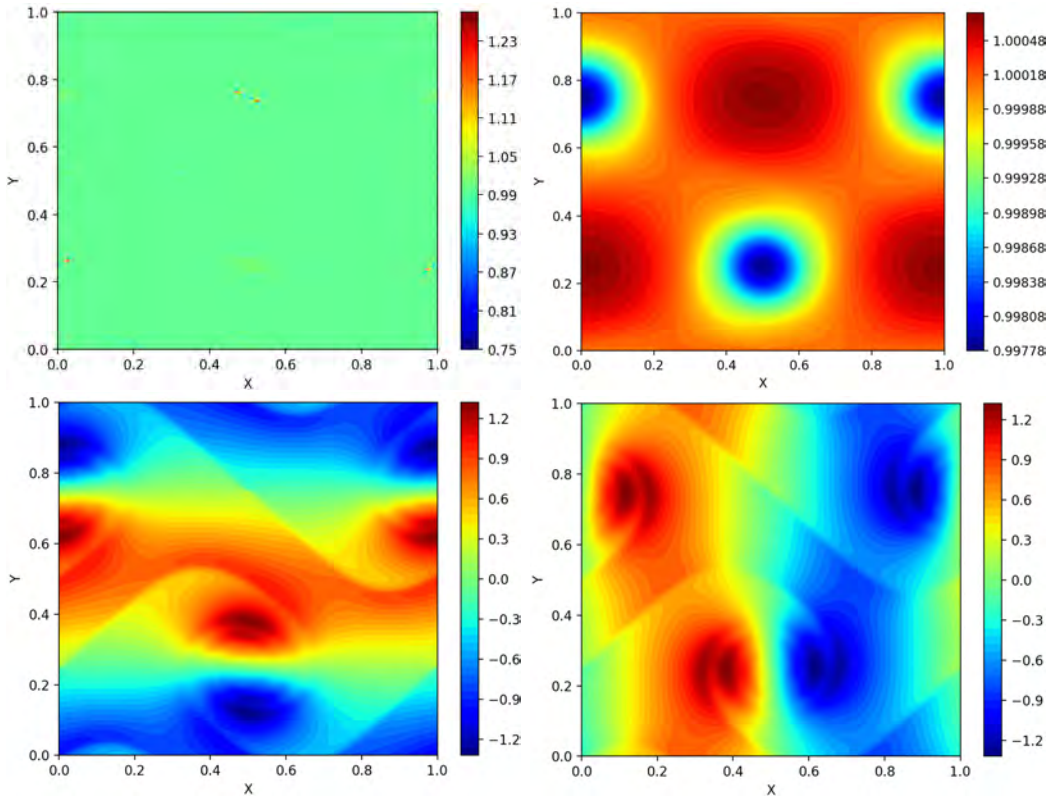


Figure 3.23 – Double shear layer (Section 3.4.6): Contour plot of the physical variables ρ (top left), p top right), u (bottom left) and v bottom right) for the full Euler equations with the Order 2 L^2 AP scheme setting $\varepsilon = 10^{-3}$ at time $t = 1.2$ on a 256×256 grid.

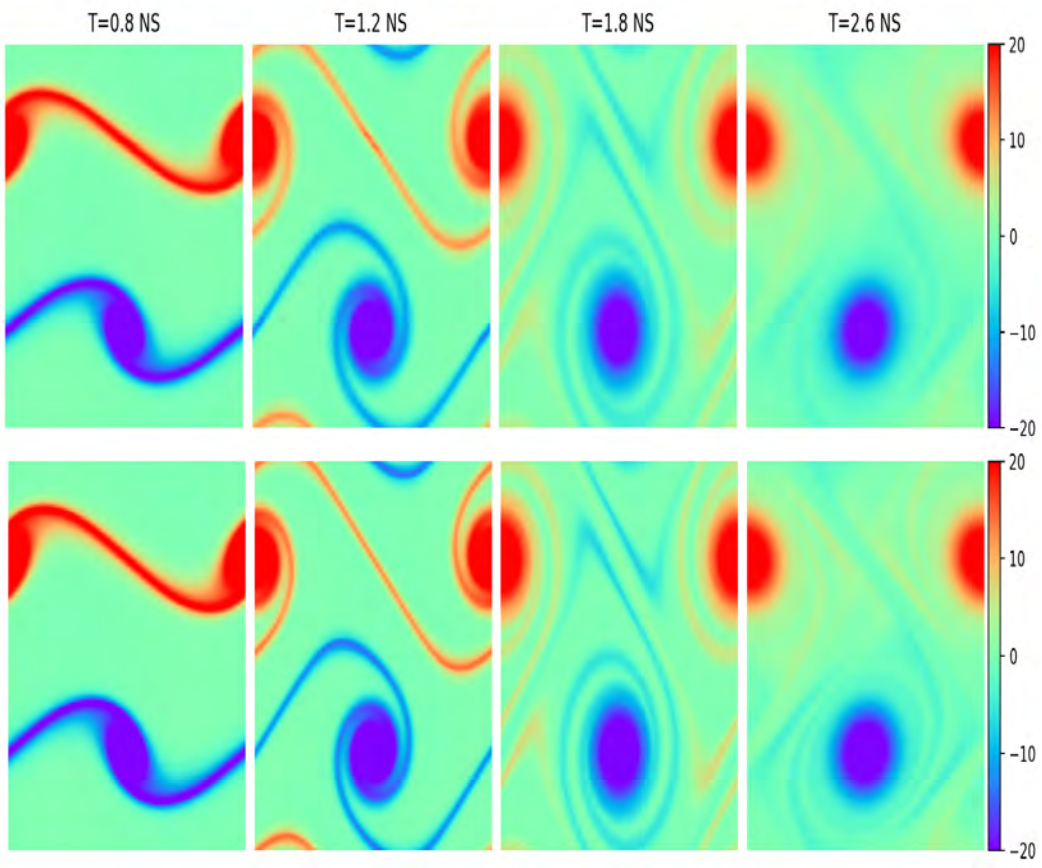


Figure 3.24 – Double shear layer (Section 3.4.6): Vorticity contours with the Order 2 L^2 AP scheme for the full Navier-Stokes equations setting $\varepsilon = 10^{-3}$ and $\mu = 2.10^{-4}$ at times $t = 0.8$, 1.2, 1.8 and 2.6 on a 128×128 (top) and a 256×256 grid.

3.4.7 Heat conduction

In order to validate our scheme in the presence of heat conduction, we consider the following simple problem initially proposed by [36] in one dimension. Following the same setting as is [13], we consider a fluid initially at rest, with a higher density at the center of the domain, inside the circle of radius 0.2 and centered in $(0, 0)$. The initial states are given by

$$\rho(0, x, y) = \begin{cases} 2 & \text{if } r < 0.2, \\ 0.5 & \text{otherwise,} \end{cases} \qquad (u, v)(0, x, y) = (0, 0), \quad p(0, x, y) = 1, \qquad (3.65)$$

where $r = \sqrt{x^2 + y^2}$, $\varepsilon = 1$, $\lambda = 10^{-2}$, $\mu = 10^{-2}$ and $t_{end} = 1$. We set transmissive boundary conditions ($\partial W/\partial n = 0$) everywhere.
Figure 3.25, displays one-dimensional cuts of the temperature and heat flux along the x axis for $y = 0$. The simulations are run with the Order 2 L^2 and $L^{2,stab}$ AP schemes until $t = 1.0$. Since we start with a velocity equal to $(0, 0)$, we have chosen a smaller time step until $t = 0.01$ (related to the eigenvalues of the jacobian matrix associated to the inviscid flux i.e $|U| \pm c$) and then the usual restriction related only to the fluid velocity U.

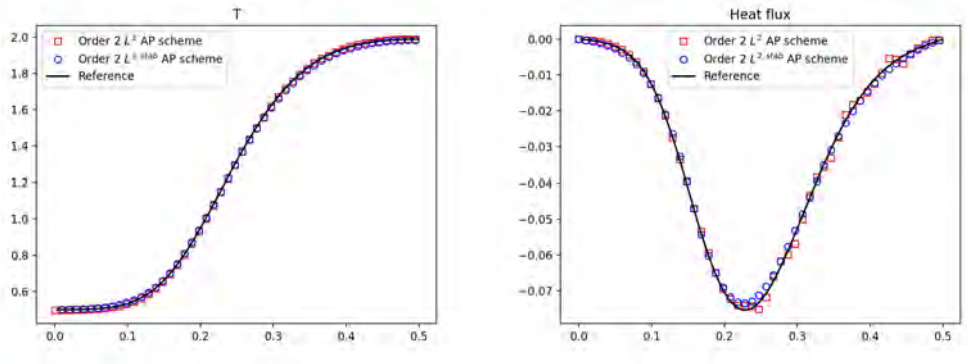


Figure 3.25 – Heat conduction (Section 3.4.7): Temperature and heat flux at time $t = 1.0$ with the Order 2 L^2 and $L^{2,stab}$ AP schemes for 101×101 points.

3.4.8 Lid-driven cavity flow: steady state incompressible solution

The following test case is well-known for the Navier-Stokes equations in the low Mach number regime. We consider a fluid in $\Omega = [0,1]^2$ with constant density and pressure and where the velocity is set to zero in all the domain except on the upper boundary where $(u,v) = (1,0)$. The other three walls are stationary and we impose a no slip boundary condition $(u,v) = (0,0)$. Then we expect the creation of a primary vortex at the center of the cavity and secondary vortices on the bottom corners as the Reynolds number increases (see sketch in Figure 3.26). The initial data are the following:

$$\rho(0,x,y) = 1, \qquad u(0,x,y) = \begin{cases} 0 & \text{if } y < 1, \\ 1 & \text{if } y = 1, \end{cases} \qquad v(0,x,y) = 0, \qquad p(0,x,y) = 1, \tag{3.66}$$

with $\varepsilon = 10^{-5}$, no heat conduction $\lambda = 0$ and μ given by the Reynolds number R_e. The Reynolds number is defined by $R_e = \rho_0 U_0 L/\mu$, where $\rho_0 = 1$, $U_0 = 1$ and $L = 1$ are respectively the characteristic density, velocity and lenght. Thus, we set $\mu = 1/R_e$. Moreover, in order to run the simulations we also impose a condition on the pressure: $\partial p/\partial n = 0$ on all the walls.

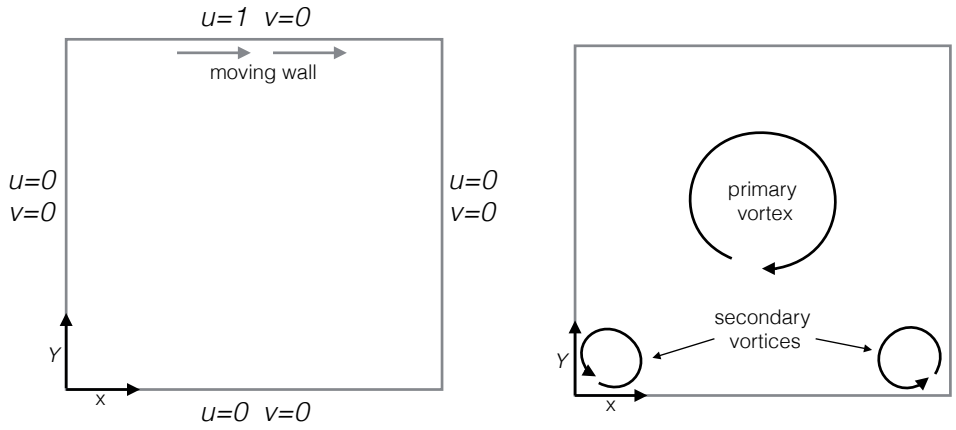


Figure 3.26 – Lid driven cavity flow (Section 3.4.8): Geometry for the problem and expected vortex formation.

The simulation is run for different Reynolds numbers until a steady state is reached. For each simulation we show the final Mach number $M = \sqrt{\varepsilon}|U|/c$ distribution, the u velocity contours along with the streamlines. We also compare the u and v velocity profiles along the lines $y = 0$ and $x = 0$ respectively with the reference solution given in [81] for the incompressible Navier-Stokes equations.
In Figure 3.27, we show the results from top to bottom for $R_e = 100$, $R_e = 400$ and $R_e = 1000$ at times $t = 20.0$, $t = 30.0$ and $t = 30.0$ respectively. On the left row, we observe the Mach number varying in the range $[0, 0.0027]$. On the middle row, we see as expected a primary vortex moving towards the cavity center and the formation of secondary vortices on the bottom corners as the Reynold number

increases. On the right row, the velocity profiles are in good agreement with the reference solution [81].

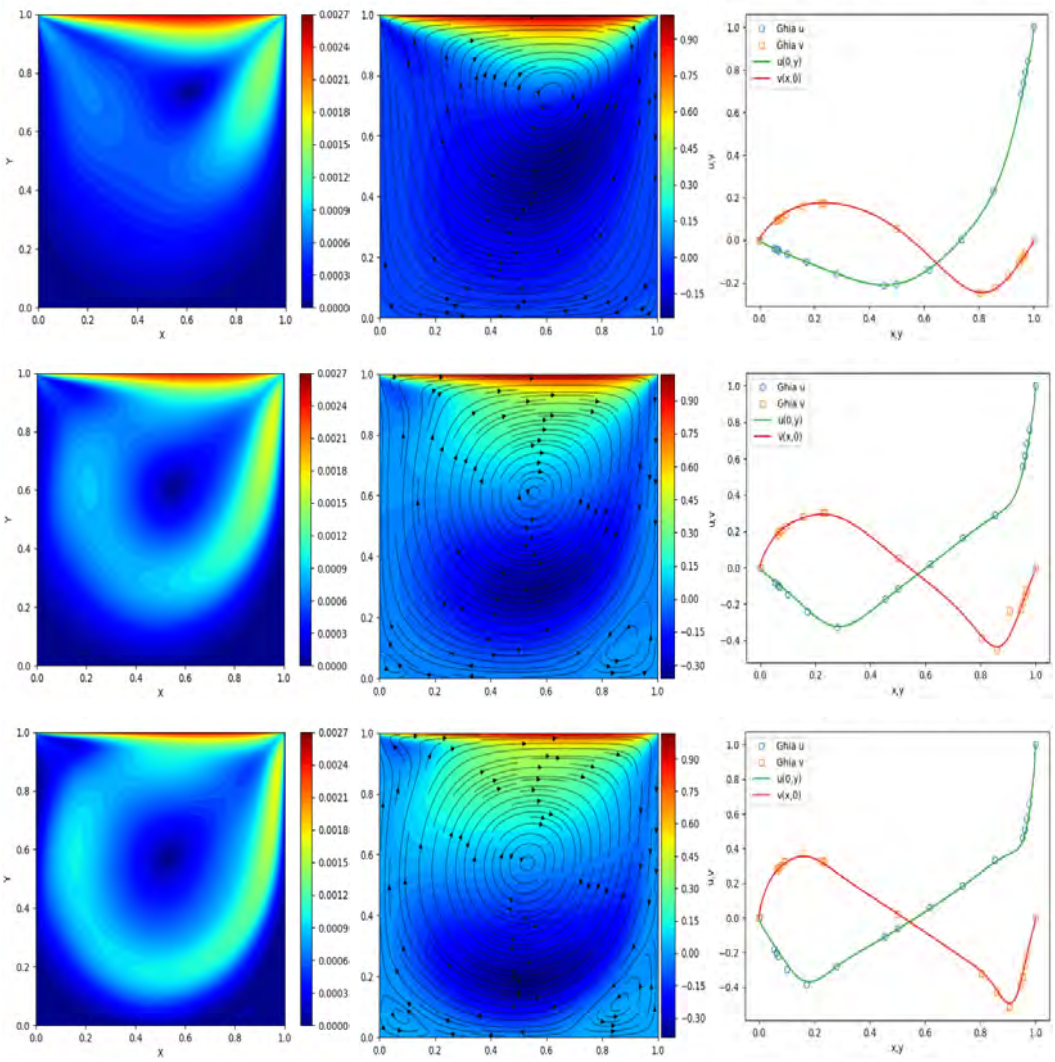


Figure 3.27 – Lid driven cavity flow (Section 3.4.8): Results with the Order 2 L^2 AP scheme for various Reynolds numbers $R_e = 100$ (top), $Re = 400$ (middle) and $Re = 1000$ (bottom) on a 100×100 grid. Mach number contours (left), u velocity contours with velocity streamlines (middle) and velocity profiles compared with the reference solution [81] (right).

In Figure 3.28, we show the contour plots of the density, v velocity and pressure setting the Reynolds number to $R_e = 1000$. We are again in agreement with the expected results. For the density we observe again the structure of the primary vortex and for the pressure some oscillations appear on the top corners. It is interesting to note that when adding the upwinding on the implicit part, the oscillations vanish (see Figure 3.29, bottom plot). This is at the cost of having a more diffusive scheme. We illustrate it showing the velocity streamlines for two grid resolutions. Indeed,

looking at the two top figures in Figure 3.29, we observe that we do not recover the structure of the secondary vortices on the coarser grid.

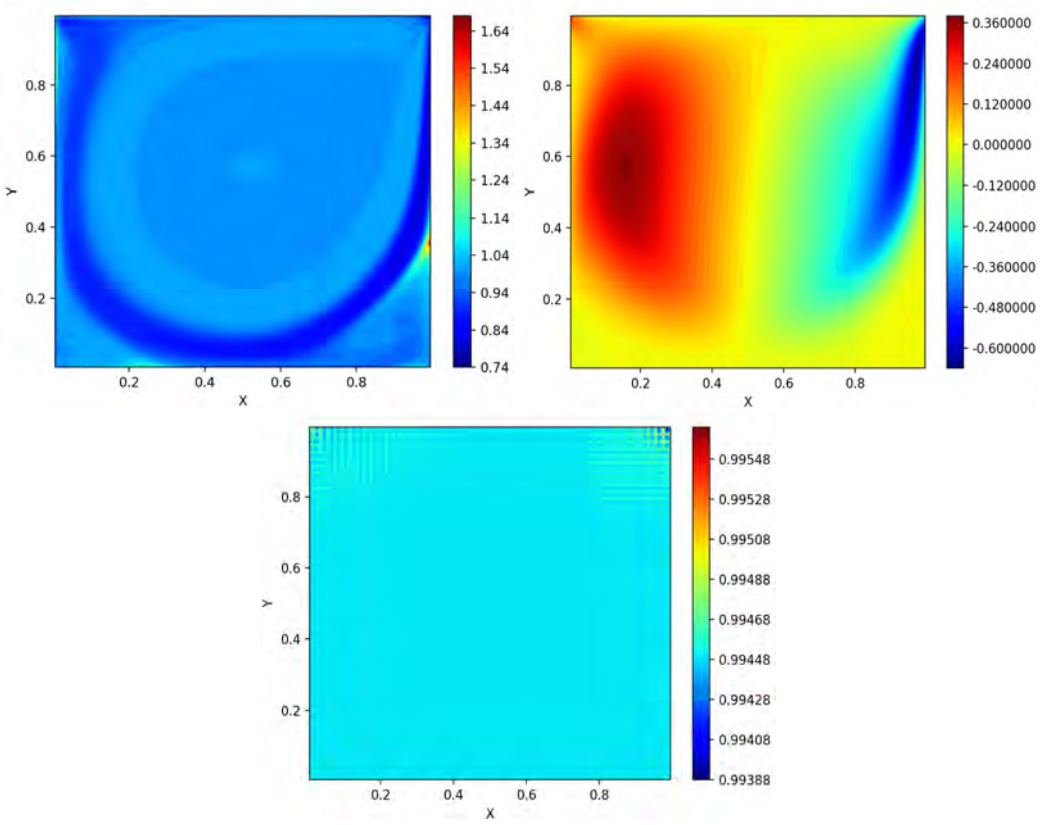


Figure 3.28 – Lid driven cavity flow (Section 3.4.8): Results for $Re = 1000$ at $t = 30.0$ with the Order 2 L^2 AP scheme on a 100×100 mesh. Contours for the density (top left), v velocity (top right) and pressure (bottom).

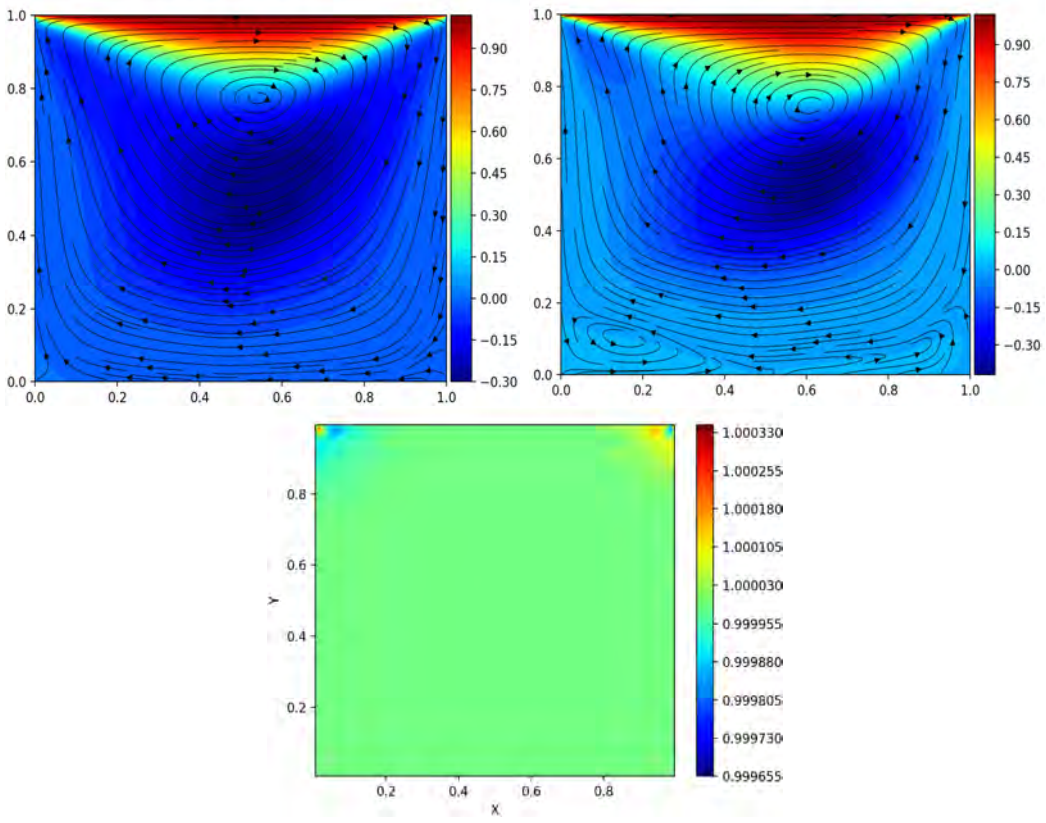


Figure 3.29 – Lid driven cavity flow (Section 3.4.8): Results for $Re = 1000$ at $t = 30.0$ with the Order 2 $L^{2,stab}$ AP scheme. u velocity contours for 100×100 points (top left) versus 200×200 points (top right) and pressure contour for 100×100 points (bottom).

Chapter 4

Conclusion and perspectives

4.1 Conclusion

In this thesis, we have developed and analyzed asymptotic preserving finite volume schemes for the Euler and Navier-Stokes equations in the low Mach number limit.

We presented in Chapter 1, the difficulties related to the simulation of low Mach number flows. When the acoustic waves are very fast compared to the fluid velocity, classical explicit schemes used for compressible fluids fail. They are not able to capture the correct asymptotic behavior of the flow and their stability is ensured under a very restrictive condition of type C.F.L. related to the Mach number. We then presented the interest and principles of asymptotic preserving (AP) schemes for multiscale models. In our context, they allow us to be consistent with the limit model when the Mach number tends to zero and stable independently of the Mach number. The scaling of the Euler and Navier-Stokes equations was also discussed.

Then, we studied and developed in Chapter 2, AP schemes for the full Euler model. We first studied the continuous system and detailed the formal low Mach number limit. In the limit, the incompressible Euler model is obtained and its reformulation yields an explicit equation for the pressure. We focused on the construction of a Finite Volume AP scheme based on IMEX (implicit-explicit) methods where the flux is split into two parts: one treated explicitly and one treated implicitly. The goal of this work was first to analyze the asymptotic properties of existing flux splittings in order to correctly choose the one for developing such schemes. We studied for some flux splittings, the hyperbolicity of the implicit and explicit flux operators, a necessary condition for the asymptotic stability, the asymptotic consistency formally passing to the limit in the semi-discretizations and the preservation of contact discontinuities. For the chosen flux splitting, we proposed a new linear AP scheme based on the non linear semi-discretization proposed in [11]. In particular, we proposed a linearization of the pressure equation obtained reformulating the semi-discretization. We also proved the asymptotic consistency and the preservation of contact discontinuities of the linear scheme. For the discretization in space, we used a Rusanov-type solver for the explicity part and a centered solver for the implicit part leading to an L^2 stable scheme. We also proposed adding an upwinding in the implicit part as well for obtaining a more stable scheme. Moreover, conducting a Fourier linear stability analysis we proved the L^2 stability of both proposed schemes under a C.F.L. condition related to the fluid velocity and independent of the Mach number. Additionally, it emphasized that the upwinding on the implicit

part improved the stability. We compared on a set of numerical tests our order 1 schemes against the nonlinear scheme proposed in [11] and thus showed the good behavior and robustness of our schemes. In the next section, we proposed an order 2 extension based on a Runge-Kutta IMEX scheme in time and MUSCL techniques in space. Based on the work [29] for the isentropic Euler equations, we were able to reduce the spurious oscillations classically observed for high order IMEX schemes. For that, we relied on a MOOD type procedure based on physical detection criteria and the construction of a TVD AP scheme more precise than the order 1. We concluded this chapter with numerical tests assessing the increase in accuracy and low oscillating properties of the proposed optimized scheme and its performance on a large range of Mach number regimes.

In Chapter 3, we proposed an extension of the constructed AP schemes for the Navier-Stokes equations allowing us to simulate viscous flows as well. First, we detailed the formal low Mach number limit in the continuous model. At the limit, due to the temperature variations we do not recover the incompressible Navier-Stokes equations. The incompressibility constraint is retrieved only neglecting the effects of heat conduction. We also mentioned the stability constraints related to the presence of diffusive terms. In the next section, we presented the semi-discretiztion proposed that is based on the same flux splitting used in Chapter 2 for the Euler part ensuring the asymptotic stability and on an implicit treatment of the diffusive terms allowing us to be efficient also in highly viscous regimes. To obtain a simple linear scheme, we proposed following the same strategy as for the Euler equations, a linearization of the the viscous terms in the pressure equation. For the new proposed scheme the asymptotic consistency was also studied on the semi-discretization. For the space discretization, a Rusanov-type solver was used for the explicit part and a centered solver was proposed for the implicit part. We provided details on the full discretization in two dimensions and in particular on the discretization of the viscous stress tensor. Afterwards, the scheme was extended to second order accuracy and a stabilization procedure based on a upwinding of the implicit Euler flux was proposed. In the last section, we presented a number of two dimensional numerical tests for a wide range of Mach numbers assessing the good behavior and asymptotic properties of our order 2 schemes both for the full Euler and the Navier-Stokes equations.

To conclude, in this thesis we proposed a new simple linear asymptotic preserving method based on Finite Volume IMEX methods for the Euler and Navier-Stokes equations. Conducting an analyses of existing works, we were able to choose the better suited flux splitting for our problem and prove the asymptotic consistency in the low Mach number limit of the proposed semi-discretization. Moreover, throughout the choice made for the space discretization based on previous theoretical results and the L^2 linear stability analysis conducted for the full Euler case, we proved the asymptotic stability and robustness of the constructed order 1 schemes. Additionally, our method was extended up to second order accuracy and for the Euler case

we were able to ensure low oscillatory properties using a correction procedure.

4.2 Perspectives

In future works, we aim to extend and improve the MOOD procedure proposed for the Euler equations to the Navier-Stokes system. We are also interested in developing high-order accurate asymptotic methods for the Euler or Navier-Stokes equations, with a focus on efficient implementation in 3D using parallelization techniques. This will enable us to simulate complex physical phenomena.
Furthermore, we have noticed that in some cases, the upwinding used in the implicit part of the Euler flux to improve the stability can lead to excessive diffusion. This issue requires further investigation in order to appropriately choose the diffusion coefficient depending on the situation.
Finally, we believe that the performance of asymptotic preserving schemes for multiscale models can be enhanced by employing domain decomposition techniques. By dividing the computational domain into different regions, each with its own better suited numerical approach, we can achieve better performance. Indeed, as we have mentioned in this thesis, the parameter ε, when it plays the role of the physical Mach number, can vary in the space-time domain. In regions where ε is of order 1, conventional explicit schemes are less constrained by the stability condition on the time step, they are more cost-effective and less diffusive compared to asymptotic preserving schemes. In regions where ε is close to 0, discretizing the asymptotic model seems to be the most appropriate solution. However, in intermediate regions where a classical discretization becomes expensive and the limit model is not yet valid, a discretization of the multiscale model with asymptotic preserving schemes is relevant (see Figure 4.1). The combination of these different approaches using domain decomposition techniques should result in highly efficient schemes.

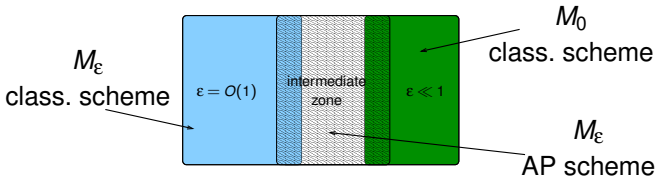


Figure 4.1 – Multiscale models and domain decomposition in combination with AP schemes.

Appendix A

Maple: Our linear L^2 and L^∞ schemes

```
> restart : with(LinearAlgebra) : with(linalg) : with(plots) :
> #Flux and Matrix of the Euler system
> F := Vector(3) : F[1] := rho·u : F[2] := rho·u·u + c^2·rho/(ga·epsilon) : F[3] := epsilon·rho·u^3/2
      + c^2·rho·u/(ga − 1) : F;
```

$$\begin{bmatrix} \rho\, u \\ \rho\, u^2 + \frac{c^2\,\rho}{ga\,\epsilon} \\ \frac{\epsilon\,\rho\,u^3}{2} + \frac{c^2\,\rho\,u}{ga-1} \end{bmatrix} \tag{1}$$

```
> Id := Matrix(3, shape = identity) :
> Df := Matrix(3) :
> Df[1,2] := 1 : Df[2,1] := (ga − 3)·u^2/2 : Df[2,2] := (3 − ga)·u : Df[2,3] :=
      (ga − 1)/epsilon :
> Df[3,1] := −c^2·u/(ga − 1) + (ga − 2)/2 ·epsilon·u^3 : Df[3,2] := c^2/(ga − 1) + (3 − 2·ga)/2
      ·epsilon·u^2 : Df[3,3] := ga·u :
> Df;
```

$$\begin{bmatrix} 0 & 1 & 0 \\ \frac{(ga-3)\,u^2}{2} & (3-ga)\,u & \frac{ga-1}{\epsilon} \\ -\frac{c^2\,u}{ga-1} + \frac{(ga-2)\,\epsilon\,u^3}{2} & \frac{c^2}{ga-1} + \frac{(3-2\,ga)\,\epsilon\,u^2}{2} & ga\,u \end{bmatrix} \tag{2}$$

```
> # Eigenvalues and eigenvectors
> L := simplify(Eigenvalues(Df)); Lord := Vector(3) : Lord[1] := L[3] : Lord[2] := L[1] :
      Lord[3] := L[2] : Lord;
```

$$L := \begin{bmatrix} u \\ \frac{c\sqrt{\epsilon} + \epsilon\,u}{\epsilon} \\ \frac{-c\sqrt{\epsilon} + \epsilon\,u}{\epsilon} \end{bmatrix} \tag{3}$$

$$\begin{bmatrix} \frac{-c\sqrt{\epsilon} + \epsilon\,u}{\epsilon} \\ u \\ \frac{c\sqrt{\epsilon} + \epsilon\,u}{\epsilon} \end{bmatrix} \tag{3}$$

```
> U := Vector(3);
```

$$U := \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{4}$$

```
> U[1] := rho : U[2] := rho·u : U[3] := epsilon·rho·u^2/2 + rho·c^2/(ga·(ga − 1)) : U;
```

$$\begin{bmatrix} \rho \\ \rho\, u \\ \frac{\epsilon\,\rho\,u^2}{2} + \frac{\rho\,c^2}{ga\,(ga-1)} \end{bmatrix} \tag{5}$$

```
> # Matrix of the eigenvectors
> P := Matrix(3) : P[1,2] := 1 : P[2,2] := u : P[3,2] := epsilon/2 ·u^2 : P[1,1] := 1 : P[1,
      3] := 1 :
> P[2,1] := u − c/sqrt(epsilon) : P[2,3] := u + c/sqrt(epsilon) :
> P[3,1] := epsilon·u^2/2 − sqrt(epsilon)·u·c + c^2/(ga − 1) :
> P[3,3] := epsilon·u^2/2 + sqrt(epsilon)·u·c + c^2/(ga − 1) :
> P;
```

$$\begin{bmatrix} 1 & 1 & 1 \\ u - \frac{c}{\sqrt{\epsilon}} & u & u + \frac{c}{\sqrt{\epsilon}} \\ \frac{\epsilon\,u^2}{2} - \sqrt{\epsilon}\,u\,c + \frac{c^2}{ga-1} & \frac{\epsilon\,u^2}{2} & \frac{\epsilon\,u^2}{2} + \sqrt{\epsilon}\,u\,c + \frac{c^2}{ga-1} \end{bmatrix} \tag{6}$$

```
> #1/M := c sur·(u sqrt(epsilon))
> Ps := Matrix(3) : Ps[1] := P[1] : Ps[2,1] := u·(1 − 1/M) : Ps[2,2] := u : Ps[2,3] := u
      ·(1 + 1/M) : Ps[3,1] := epsilon·u^2·(1/2 − 1/M + 1/((ga − 1)·M^2)) : Ps[3,2] :=
```

$$\frac{\text{epsilon}\cdot u^2}{2} : Ps[3,3] := \text{epsilon}\cdot u^2\cdot\left(\frac{1}{2}+\frac{1}{M}+\frac{1}{(ga-1)\cdot M^2}\right) : Ps;$$

$$\begin{bmatrix} 1 & 1 & 1 \\ u\left(1-\frac{1}{M}\right) & u & u\left(1+\frac{1}{M}\right) \\ \epsilon u^2\left(\frac{1}{2}-\frac{1}{M}+\frac{1}{(ga-1)M^2}\right) & \frac{\epsilon u^2}{2} & \epsilon u^2\left(\frac{1}{2}+\frac{1}{M}+\frac{1}{(ga-1)M^2}\right) \end{bmatrix} \quad (7)$$

> $Pinv := MatrixInverse(P);$

$$Pinv := \begin{bmatrix} \frac{u\left(\epsilon\, ga\, u-\epsilon u+2c\sqrt{\epsilon}\right)}{4c^2} & -\frac{\sqrt{\epsilon}\left(u\sqrt{\epsilon}\, ga-u\sqrt{\epsilon}+c\right)}{2c^2} & \frac{ga-1}{2c^2} \\ \frac{-\epsilon u^2 ga+\epsilon u^2+2c^2}{2c^2} & \frac{\epsilon u(ga-1)}{c^2} & -\frac{ga-1}{c^2} \\ -\frac{u\left(-\epsilon\, ga\, u+2c\sqrt{\epsilon}+\epsilon u\right)}{4c^2} & -\frac{\sqrt{\epsilon}\left(u\sqrt{\epsilon}\, ga-u\sqrt{\epsilon}-c\right)}{2c^2} & \frac{ga-1}{2c^2} \end{bmatrix} \quad (8)$$

> $Psinv := Matrix(3) : Psinv[1,1] := \frac{1}{4}\cdot\left((ga-1)\cdot M^2+2\cdot M\right);$

$$Psinv_{1,1} := \frac{(ga-1)M^2}{4}+\frac{M}{2} \quad (9)$$

> $Psinv := MatrixInverse(Ps);$

$$Psinv := \begin{bmatrix} \frac{M(M\, ga-M+2)}{4} & -\frac{M(M\, ga-M+1)}{2u} & \frac{(ga-1)M^2}{2u^2\epsilon} \\ -\frac{1}{2}M^2 ga+\frac{1}{2}M^2+1 & \frac{M^2(ga-1)}{u} & -\frac{(ga-1)M^2}{u^2\epsilon} \\ \frac{M(M\, ga-M-2)}{4} & -\frac{M(M\, ga-M-1)}{2u} & \frac{(ga-1)M^2}{2u^2\epsilon} \end{bmatrix} \quad (10)$$

> $test := Id - convert(multiply(P, Pinv), Matrix) :$

> $[simplify(test[3,1]), simplify(test[3,2]), simplify(test[3,3])];$

$$[0,0,0] \quad (11)$$

> $DD := convert(multiply(Pinv, convert(multiply(Df, P), Matrix)), Matrix) :$

> $[simplify(DD[3,3]-Lord[3]), simplify(DD[1,1]-Lord[1]), simplify(DD[2,2]-Lord[2])];$

$$[0,0,0] \quad (12)$$

> $[simplify(DD[1,2]), simplify(DD[1,3]), simplify(DD[2,1]), simplify(DD[2,3]), simplify(DD[3,1]), simplify(DD[3,2])];$

$$[0,0,0,0,0,0] \quad (13)$$

> #stabilité L2 du schema de Ruzanov ou LF explicite pour le systeme linearisé dans le cas u>0,

alpha=dt sur dx

> $A := simplify\left(Id - I\cdot\text{alpha}\cdot\sin(\text{phi})\cdot DD + \left(u+\frac{c}{\text{sqrt(epsilon)}}\right)\cdot\text{alpha}\cdot(\cos(\text{phi})-1)\cdot Id\right);$

$$A := \left[\left[\frac{(-I\alpha\sin(\phi)\, u+\cos(\phi)\,\alpha\, u-\alpha u+1)\sqrt{\epsilon}+c\,(I\sin(\phi)+\cos(\phi)-1)\,\alpha}{\sqrt{\epsilon}}, 0, 0\right], \left[0, -\frac{(I\sin(\phi)\,\alpha\, u-\cos(\phi)\,\alpha\, u+\alpha u-1)\sqrt{\epsilon}-\alpha c(\cos(\phi)-1)}{\sqrt{\epsilon}}, 0\right], \left[0,0,-\frac{(I\sin(\phi)\,\alpha\, u-\cos(\phi)\,\alpha\, u+\alpha u-1)\sqrt{\epsilon}+c\,(I\sin(\phi)-\cos(\phi)+1)\,\alpha}{\sqrt{\epsilon}}\right]\right] \quad (14)$$

> $simplify\left(A[1,1]-1+I\cdot\text{alpha}\cdot\sin(\text{phi})\cdot\left(u-\frac{c}{\text{sqrt(epsilon)}}\right)-\left(u+\frac{c}{\text{sqrt(epsilon)}}\right)\cdot\text{alpha}\cdot(\cos(\text{phi})-1)\right);$

$$0 \quad (15)$$

> $simplify(Eigenvalues(A)) :$

#il ne trouve pas les valeurs propres alors qu'elles sont sur la diagonales....

> $solve(CharacteristicPolynomial(A, \text{lambda})=0, \text{lambda}) :$#ne marche pas non plus

>

> # Explicit flux and matrix

> $Fe := Vector(3) : Fe[1] := \text{rho}\cdot u : Fe[2] := \text{rho}\cdot u^2 : Fe[3] := \frac{\text{epsilon}\cdot\text{rho}\cdot u^3}{2} : Fe;$

$$\begin{bmatrix} \rho\, u \\ \rho\, u^2 \\ \frac{\epsilon\,\rho\, u^3}{2} \end{bmatrix} \quad (16)$$

> $Dfe := Matrix(3) : Dfe[1,2] := 1 : Dfe[2,1] := -u^2 : Dfe[2,2] := 2\cdot u : Dfe[3,1] := -\text{epsilon}\cdot u^3 : Dfe[3,2] := \frac{3\cdot\text{epsilon}}{2}\cdot u^2 : Dfe;$

$$\begin{bmatrix} 0 & 1 & 0 \\ -u^2 & 2u & 0 \\ -\epsilon u^3 & \frac{3\epsilon u^2}{2} & 0 \end{bmatrix} \quad (17)$$

> $Eigenvalues(Dfe);$

(18)

$$\begin{bmatrix} 0 \\ u \\ u \end{bmatrix} \quad (18)$$

> $Ve := convert(multiply(Dfe, U), Vector) : [simplify(Ve[1] - Fe[1]), simplify(Ve[2] - Fe[2]), simplify(Ve[3] - Fe[3])];$

$$[0, 0, 0] \quad (19)$$

> $Ae := simplify(convert(multiply(Pinv, convert(multiply(Dfe, P), Matrix)), Matrix));$

$$Ae := \begin{bmatrix} \frac{u}{2} & 0 & -\frac{u}{2} \\ -\frac{-u\sqrt{\epsilon}+c}{\sqrt{\epsilon}} & u & \frac{u\sqrt{\epsilon}+c}{\sqrt{\epsilon}} \\ -\frac{u}{2} & 0 & \frac{u}{2} \end{bmatrix} \quad (20)$$

>

>

> # a partir de là alpha=deltat sur deltax

> $Bew := (1 - alpha \cdot u \cdot (1 - cos(phi))) \cdot Id - I \cdot alpha \cdot sin(phi) \cdot Dfe;$ #matrice fourier explicite

$$Bew := \left[\left[1 - \alpha\, u\,(1-\cos(\phi)),\ -I\,\alpha\sin(\phi),\ 0\right],\ \left[I\,\alpha\sin(\phi)\,u^2,\ 1-\alpha\,u\,(1-\cos(\phi)) - 2\,I\,\alpha\sin(\phi)\,u,\ 0\right],\ \left[I\,\alpha\sin(\phi)\,\epsilon\,u^3,\ -\frac{3\,I}{2}\,\alpha\sin(\phi)\,\epsilon\,u^2,\ 1-\alpha\,u\,(1-\cos(\phi))\right]\right] \quad (21)$$

>

>

> $Be := simplify(convert(Multiply(Pinv, convert(Multiply(Bew, P), Matrix)), Matrix));$

$$Be := \left[\left[-\frac{I\,\alpha\sin(\phi)\,u}{2} + \cos(\phi)\,\alpha\,u - \alpha\,u + 1,\ 0,\ \frac{I}{2}\,\alpha\sin(\phi)\,u\right],\ \left[\frac{I\sin(\phi)\,\alpha\,(-u\sqrt{\epsilon}+c)}{\sqrt{\epsilon}},\ -I\,\alpha\sin(\phi)\,u + \cos(\phi)\,\alpha\,u - \alpha\,u + 1,\ \frac{-I\sin(\phi)\,\alpha\,(u\sqrt{\epsilon}+c)}{\sqrt{\epsilon}}\right],\ \left[\frac{I}{2}\,\alpha\sin(\phi)\,u,\ 0,\ -\frac{I\,\alpha\sin(\phi)\,u}{2} + \cos(\phi)\,\alpha\,u - \alpha\,u + 1\right]\right] \quad (22)$$

>

> $Bes := Matrix(3) : Bes[1,1] := -\frac{I \cdot sin(phi) \cdot cfl}{2} - cfl \cdot (1 - cos(\phi)) + 1 : Bes[1,3] := \frac{I}{2}\sin(\phi) \cdot cfl :$

> $Bes[2,1] := -I \cdot sin(phi) \cdot cfl \cdot \left(1 - \frac{1}{M}\right) : Bes[2,2] := -I \cdot sin(\phi) \cdot cfl + 1 - cfl \cdot (1 - cos(\phi)) : Bes[2,3] := -I \cdot sin(phi) \cdot cfl \cdot \left(1 + \frac{1}{M}\right) :$

> $Bes[3,1] := \frac{I}{2} \cdot \sin(\phi) \cdot cfl : Bes[3,3] := -\frac{I \cdot sin(phi) \cdot cfl}{2} - cfl \cdot (1 - cos(\phi)) + 1 :$

> $Bes;$

$$\left[\left[-\frac{I\sin(\phi)\,cfl}{2} - cfl\,(1-\cos(\phi)) + 1,\ 0,\ \frac{I}{2}\sin(\phi)\,cfl\right],\ \left[-I\sin(\phi)\,cfl\left(1-\frac{1}{M}\right),\ -I\sin(\phi)\,cfl + 1 - cfl\,(1-\cos(\phi)),\ -I\sin(\phi)\,cfl\left(1+\frac{1}{M}\right)\right],\ \left[\frac{I}{2}\sin(\phi)\,cfl,\ 0,\ -\frac{I\sin(\phi)\,cfl}{2} - cfl\,(1-\cos(\phi)) + 1\right]\right] \quad (23)$$

>

> $Eigenvalues(Bes); Eigenvalues(Be);$

$$\begin{bmatrix} cfl\cos(\phi) - cfl + 1 \\ -I\sin(\phi)\,cfl + cfl\cos(\phi) - cfl + 1 \\ -I\sin(\phi)\,cfl + cfl\cos(\phi) - cfl + 1 \end{bmatrix}$$

$$\begin{bmatrix} \cos(\phi)\,\alpha\,u - \alpha\,u + 1 \\ -I\,\alpha\sin(\phi)\,u + \cos(\phi)\,\alpha\,u - \alpha\,u + 1 \\ -I\,\alpha\sin(\phi)\,u + \cos(\phi)\,\alpha\,u - \alpha\,u + 1 \end{bmatrix} \quad (24)$$

> $Becfl1 := Matrix(3) : Becfl1[1,1] := -\frac{I \cdot sin(phi)}{2} + \cos(\phi) : Becfl1[1,3] := \frac{I}{2}\sin(\phi) :$

> $Becfl1[2,1] := -I \cdot sin(phi) \cdot \left(1 - \frac{1}{M}\right) : Becfl1[2,2] := -I \cdot \sin(\phi) + \cos(\phi) : Becfl1[2,3] := -I \cdot sin(phi) \cdot \left(1 + \frac{1}{M}\right) :$

> $Becfl1[3,1] := \frac{I}{2}\sin(\phi) : Becfl1[3,3] := -\frac{I \cdot sin(phi)}{2} + \cos(\phi) :$

> $Eigenvalues(Becfl1);$

$$\begin{bmatrix} \cos(\phi) \\ -I\sin(\phi) + \cos(\phi) \\ -I\sin(\phi) + \cos(\phi) \end{bmatrix} \quad (25)$$

```
> Biw := Matrix(3): Biws := Matrix(3): Biw[1,1] := 1: Biws := 1:
> Biw[2,1] := -alpha/epsilon * I*sin(phi)/(1 + alpha^2*c^2/epsilon * 2*(1-cos(phi))) * ((ga-1)*epsilon*u^2/2 - I*alpha
      *sin(phi)*u*(-c^2 + ga*(ga-1)*epsilon*u^2/2)):
> Biw[2,2] := 1 + I*alpha*sin(phi)/(epsilon*(1 + alpha^2*c^2/epsilon * 2*(1-cos(phi)))) * ((ga-1)*epsilon*u + I*alpha
      *sin(phi)*(c^2 - ga*(ga-1)*epsilon*u^2)):
> Biw[2,3] := -I*alpha*sin(phi)/(epsilon*(1 + alpha^2*c^2/epsilon * 2*(1-cos(phi)))) * (ga - 1 - alpha*I*sin(phi)*ga
      *(ga-1)*u):
> Biw[3,1] := I*alpha*sin(phi)/(ga-1) * (c^2*u - (-c^2*alpha*I*sin(phi)/epsilon + ga*u)/(1 + alpha^2*c^2/epsilon * 2*(1-cos(phi)))
      *((ga-1)*epsilon*u^2/2 - I*alpha*sin(phi)*u*(-c^2 + ga*(ga-1)*epsilon*u^2/2))):
> Biw[3,2] := I*alpha*sin(phi)/(ga-1) * (-c^2 + (-c^2*alpha*I*sin(phi)/epsilon + ga*u)/(1 + alpha^2*c^2/epsilon * 2*(1-cos(phi)))
      *((ga-1)*epsilon*u + I*alpha*sin(phi)*(c^2 - ga*(ga-1)*epsilon*u^2))):
> Biw[3,3] := 1 - I*alpha*sin(phi)*(-c^2*alpha*I*sin(phi)/epsilon + ga*u)/(1 + alpha^2*c^2/epsilon * 2*(1-cos(phi)))
      *(1 - alpha*I*sin(phi)*ga*u):
> Bi := simplify(convert(Multiply(Pinv, convert(Multiply(Biw, P), Matrix)), Matrix)):
> Eigenvalues(Bi):
> Bis := Matrix(3):
```

```
> temp := (cfl, M) -> -1/(2*(-1 - 2*cfl^2/M^2*(1-cos(phi)))) * (ga*cfl^2*(cos(phi)^2 - 1) + 2 - I
      *cfl*sin(phi) + cfl/M * (-cfl/M*(cos(phi) - 1)*((I*cfl*sin(phi)*ga - 2)*cos(phi) + I*cfl
      *sin(phi)*ga + 2) + cfl*cos(phi)^2*(I*sin(phi)*cfl/M^2 - (ga + 1)) - 2*I*cfl^2/M^2*cos(phi)
      *sin(phi) + (cfl^2/M^2 + 2)*I*sin(phi) + cfl*(ga + 1))):
> simplify(temp(alpha*u, sqrt(epsilon)*u/c) - Bi[1,1]);
                                  0                                  (26)
> Bis[1,1] := -1/(2*(-1 - 2*cfl^2/M^2*(1-cos(phi)))) * (ga*cfl^2*(cos(phi)^2 - 1) + 2 - I*cfl
      *sin(phi) + cfl/M * (-cfl/M*(cos(phi) - 1)*((I*cfl*sin(phi)*ga - 2)*cos(phi) + I*cfl
      *sin(phi)*ga + 2) + cfl*cos(phi)^2*(I*sin(phi)*cfl/M^2 - (ga + 1)) - 2*I*cfl^2/M^2*cos(phi)
      *sin(phi) + (cfl^2/M^2 + 2)*I*sin(phi) + cfl*(ga + 1))):
> simplify(Bis[1,1] - temp(cfl, M));
                                  0                                  (27)
> Bi[1,2]; Bi[1,3];
                                  0

   1/(sqrt(epsilon)*(4*alpha^2*c^2*cos(phi) - 4*alpha^2*c^2 - 2*epsilon)) * (alpha*sin(phi)*((alpha*sin(phi)*ga*u + 1)*u*epsilon^(3/2)
      + c*alpha*(I*c*alpha*(cos(phi) - 1)*(cos(phi) + 1)*u*ga*sqrt(epsilon) + I*c^2*alpha*cos(phi)^2 - 2*I*c^2*alpha*cos(phi)
      - (ga - 1)*epsilon*u*sin(phi) + I*c^2*alpha)))                  (28)
> temp := (cfl, M) -> 1/(2*(-1 - 2*cfl^2/M^2*(1-cos(phi)))) * cfl*sin(phi)*(cfl*sin(phi)*ga + I
      + cfl/M*(-I*cfl*sin(phi)^2*ga/M + I*cfl*(cos(phi) - 1)^2/M^2 - (ga - 1)*sin(phi))):
> simplify(Bi[1,3] - temp(alpha*u, sqrt(epsilon)*u/c));
                                  0                                  (29)
```

```
> Bis[1,3] := 1/(2·(-1 - 2·cfl^2/M^2·(1-cos(phi)))) · cfl·sin(phi)·(cfl·sin(phi)·ga + I + cfl/M·(
    - I·cfl·sin(phi)^2·ga/M + I·cfl·(cos(phi)-1)^2/M^2 - (ga-1)·sin(phi))):
> simplify(Bis[1,3] - temp(cfl, M));
                                  0                                  (30)
> Bi[2,1];
1/sqrt(ε)·(2α^2 c^2 cos(φ) - 2α^2 c^2 - ε)·(α sin(φ)·(-(α sin(φ) ga u + 1) u ε^(3/2) + c·(   (31)
  -c α·(I u ga α cos(φ)^2 - I u ga α + sin(φ))·sqrt(ε) + I c^2 α^2 cos(φ)^2 - 2 I c^2 α^2 cos(φ)
  + I c^2 α^2 + sin(φ) α ε u + I ε))))
> temp := (cfl, M) -> 1/(-1 - 2·cfl^2/M^2·(1-cos(phi)))·cfl·sin(phi)·(-(cfl·sin(phi)·ga + I)
    + 1/M·(cfl·sin(phi)/M·(I·cfl·ga·sin(phi) - 1) + I·cfl^2/M^2·(1-cos(phi))^2 + sin(phi)·cfl
    + I)):
> simplify(Bi[2,1] - temp(alpha·u, sqrt(epsilon)·u/c));
                                  0                                  (32)
> Bis[2,1] := 1/(-1 - 2·cfl^2/M^2·(1-cos(phi)))·cfl·sin(phi)·(-(cfl·sin(phi)·ga + I) + 1/M
    ·(cfl·sin(phi)/M·(I·cfl·ga·sin(phi) - 1) + I·cfl^2/M^2·(1-cos(phi))^2 + sin(phi)·cfl + I)):
> simplify(Bis[2,1] - temp(cfl, M));
                                  0                                  (33)
> Bi[2,2]; Bis[2,2] := 1:
                                  1                                  (34)
> Bi[2,3];
-1/sqrt(ε)·(2α^2 c^2 cos(φ) - 2α^2 c^2 - ε)·(α sin(φ)·((α sin(φ) ga u + 1) u ε^(3/2)          (35)
  + c·(c α·(I u ga α cos(φ)^2 - I u ga α + sin(φ))·sqrt(ε) + I c^2 α^2 cos(φ)^2
  - 2 I c^2 α^2 cos(φ) + I c^2 α^2 + sin(φ) α ε u + I ε)))
```

```
> temp := (cfl, M) -> -cfl·sin(phi)/(-1 - 2·cfl^2/M^2·(1-cos(phi)))·(cfl·sin(phi)·ga + I + 1/M·(cfl/M
    ·sin(phi)·(1 - I·cfl·ga·sin(phi)) + I·cfl^2/M^2·(1-cos(phi))^2 + sin(phi)·cfl + I)):
> simplify(Bi[2,3] - temp(alpha·u, sqrt(epsilon)·u/c));
                                  0                                  (36)
> Bis[2,3] := -cfl·sin(phi)/(-1 - 2·cfl^2/M^2·(1-cos(phi)))·(cfl·sin(phi)·ga + I + 1/M·(cfl/M·sin(phi)
    ·(1 - I·cfl·ga·sin(phi)) + I·cfl^2/M^2·(1-cos(phi))^2 + sin(phi)·cfl + I)):
> simplify(Bis[2,3] - temp(cfl, M));
                                  0                                  (37)
> Bi[3,1];
-1/sqrt(ε)·(4α^2 c^2 cos(φ) - 4α^2 c^2 - 2ε)·(α sin(φ)·(-(α sin(φ) ga u + 1) u ε^(3/2) + c α·(   (38)
  -I c α·(cos(φ) - 1)·(cos(φ) + 1) u ga sqrt(ε) + I c^2 α cos(φ)^2 - 2 I c^2 α cos(φ) - (ga
  - 1) ε u sin(φ) + I c^2 α)))
> temp := (cfl, M) -> -cfl·sin(phi)/(2·(-1 - 2·cfl^2/M^2·(1-cos(phi))))·(-(cfl·sin(phi)·ga + I) + cfl/M
    ·(I·sin(phi)^2·ga·cfl/M + I·cfl/M^2·(1-cos(phi))^2 - (ga-1)·sin(phi))):
> simplify(Bi[3,1] - temp(alpha·u, sqrt(epsilon)·u/c));
                                  0                                  (39)
> Bis[3,1] := -cfl·sin(phi)/(2·(-1 - 2·cfl^2/M^2·(1-cos(phi))))·(-(cfl·sin(phi)·ga + I) + cfl/M
    ·(I·sin(phi)^2·ga·cfl/M + I·cfl/M^2·(1-cos(phi))^2 - (ga-1)·sin(phi))):
> simplify(Bis[3,1] - temp(cfl, M));
                                  0                                  (40)
> Bi[3,2]; Bi[3,3];
                                  0
```

$$\frac{1}{\sqrt{\epsilon}\left(4\alpha^2 c^2\cos(\phi)-4\alpha^2 c^2-2\epsilon\right)}\left(\left(-\cos(\phi)^2\alpha^2 ga\,u^2+\alpha^2 ga\,u^2+\mathrm{I}\sin(\phi)\alpha u-2\right)\epsilon^{3/2}+c\alpha\left(c\alpha\left(\cos(\phi)-1\right)\left(\left(\mathrm{I}\alpha\sin(\phi)ga\,u-2\right)\cos(\phi)+\mathrm{I}\alpha\sin(\phi)ga\,u+2\right)\sqrt{\epsilon}+\left(\mathrm{I}\sin(\phi)\alpha c^2-u\epsilon(ga+1)\right)\alpha\cos(\phi)^2-2\,\mathrm{I}\,c^2\alpha^2\cos(\phi)\sin(\phi)+\left(\mathrm{I}c^2\alpha^2+2\,\mathrm{I}\epsilon\right)\sin(\phi)+u\alpha\epsilon(ga+1)\right)\right) \tag{41}$$

```
> temp := (cfl, M)->1/(2*(-1-2*cfl^2/M^2*(1-cos(phi)))*(cfl^2*ga*(1-cos(phi)^2)+I
  *sin(phi)*cfl-2
  +1/M*((1/M*(((I*cfl*sin(phi)*ga-2)*cos(phi)+I*cfl*ga*sin(phi)+2)*(cos(phi)
  -1)*cfl)+(I*sin(phi)*cfl/M^2-(ga+1))*cfl*cos(phi)^2-2*I*cfl^2/M^2*cos(phi)*sin(phi)
  +(I*cfl^2/M^2+2*I)*sin(phi)+cfl*(ga+1)))*cfl)):
> simplify(Bi[3,3]-temp(alpha*u,sqrt(epsilon)*u/c));
                                  0                                  (42)
> Bis[3,3] := 1/(2*(-1-2*cfl^2/M^2*(1-cos(phi)))*(cfl^2*ga*(1-cos(phi)^2)+I*sin(phi)
  *cfl-2
  +1/M*((1/M*(((I*cfl*sin(phi)*ga-2)*cos(phi)+I*cfl*ga*sin(phi)+2)*(cos(phi)
  -1)*cfl)+(I*sin(phi)*cfl/M^2-(ga+1))*cfl*cos(phi)^2-2*I*cfl^2/M^2*cos(phi)*sin(phi)
  +(I*cfl^2/M^2+2*I)*sin(phi)+cfl*(ga+1)))*cfl);
```

$$Bis_{3,3} := \frac{1}{-2-\frac{4\,cfl^2\left(1-\cos(\phi)\right)}{M^2}}\left(cfl^2\,ga\left(1-\cos(\phi)^2\right)+\mathrm{I}\sin(\phi)\,cfl-2\right) \tag{43}$$

$$+\frac{1}{M}\left(\left(\frac{cfl\left(\cos(\phi)-1\right)\left(\left(\mathrm{I}\,cfl\sin(\phi)\,ga-2\right)\cos(\phi)+\mathrm{I}\,cfl\sin(\phi)\,ga+2\right)}{M}+cfl\cos(\phi)^2\left(\frac{\mathrm{I}\sin(\phi)\,cfl}{M^2}-ga-1\right)-\frac{2\,\mathrm{I}\,cfl^2\cos(\phi)\sin(\phi)}{M^2}+\left(\frac{\mathrm{I}\,cfl^2}{M^2}+2\,\mathrm{I}\right)\sin(\phi)+cfl\left(ga+1\right)\right)cfl\right)\Bigg)$$

```
> simplify(Bis[3,3]-temp(cfl,M));
```

$$0 \tag{44}$$

```
> Matamps := convert(Multiply(Bis,Bes),Matrix):
> Matamps[1,2]; Matamps[2,2]; Matamps[3,2];
```

$$0$$

$$-\mathrm{I}\sin(\phi)\,cfl+1-cfl\left(1-\cos(\phi)\right)$$

$$0 \tag{45}$$

```
> simplify(Matamps[1,1]+Matamps[3,3]);
```

$$\frac{1}{-2\,cfl^2\cos(\phi)+M^2+2\,cfl^2}\left(-\left(\mathrm{I}\,cfl\sin(\phi)\,ga-M^2\,ga-2\right)cfl^3\cos(\phi)^3+\left(\mathrm{I}\left(cfl\,ga-ga-1\right)cfl\sin(\phi)-cfl\,M^2\,ga+M^2\,ga-6\,cfl+2\right)cfl^2\cos(\phi)^2-\left(\mathrm{I}\,cfl\left(-cfl^2\,ga+M^2-2\,cfl\right)\sin(\phi)+\left(M^2\,ga-6\right)cfl^2+4\,cfl-2\,M^2\right)cfl\cos(\phi)+\mathrm{I}\left(-cfl^3\,ga+\left(ga-1\right)cfl^2+cfl\,M^2-2\,M^2\right)cfl\sin(\phi)+\left(cfl-1\right)\left(\left(M^2\,ga-2\right)cfl^2-2\,M^2\right)\right) \tag{46}$$

```
> simplify(Matamps[1,1]*Matamps[3,3]-Matamps[1,3]*Matamps[3,1]);
```

$$-\frac{1}{-2\,cfl^2\cos(\phi)+M^2+2\,cfl^2}\left(\left(-cfl^3\left(M^2\,ga+1\right)\cos(\phi)^3+\left(\mathrm{I}\left(M^2\,ga+1\right)cfl\sin(\phi)+\left(M^2\,ga+3\right)cfl-1+\left(-ga-1\right)M^2\right)cfl^2\cos(\phi)^2+cfl\left(\mathrm{I}\left(M^2-2\,cfl\right)cfl\sin(\phi)+\left(M^2\,ga-3\right)cfl^2+2\,cfl-M^2\right)\cos(\phi)-\mathrm{I}\left(\left(M^2\,ga-1\right)cfl^2+cfl\,M^2-2\,M^2\right)cfl\sin(\phi)+\left(-M^2\,ga+1\right)cfl^3+\left(-1+\left(ga+1\right)M^2\right)cfl^2+cfl\,M^2-M^2\right)\left(cfl\cos(\phi)-cfl+1\right)\right) \tag{47}$$

```
> lambda := 1-cfl*(1-cos(phi))-I*sin(phi)*cfl;
```

$$\lambda := -\mathrm{I}\sin(\phi)\,cfl+1-cfl\left(1-\cos(\phi)\right) \tag{48}$$

```
> detBs := x->simplify(CharacteristicPolynomial(Matamps,x)); Polyamps := x
  ->Determinant(Matamps-x*Id); Polyamps(1-cfl*(1-cos(phi))-I*cfl*sin(phi));
  simplify(detBs(x)+Polyamps(x));
```

$$detBs := x \mapsto simplify(CharacteristicPolynomial(Matamps,x))$$

$$Polyamps := x \mapsto Determinant(Matamps-x\cdot Id)$$

$$0$$

$$0 \tag{49}$$

```
> Qamps := x->simplify(quo(-Polyamps(x),x-(1-cfl*(1-cos(phi))-I*cfl*sin(phi)),'r
```

```
'));collect(Qamps(x),x);
```

$$Qamps := x \mapsto simplify(quo(-Polyamps(x), x + \mathrm{I}\cdot\sin(\phi)\cdot cfl - 1 + cfl\cdot(1-\cos(\phi)), 'r'))$$

$$x^2 + \frac{1}{-2\,cfl^2\cos(\phi) + M^2 + 2\,cfl^2}\left(\left(-cfl^3\left(-\mathrm{I}\,cfl\sin(\phi)\,ga + M^2\,ga + 2\right)\cos(\phi)\right)^3\right. \tag{50}$$
$$+ cfl^2\left(\mathrm{I}\left(-cfl\,ga + ga + 1\right)cfl\sin(\phi) + \left(M^2\,ga + 6\right)cfl - M^2\,ga - 2\right)\cos(\phi)^2$$
$$+ \left(\mathrm{I}\,cfl\left(-ga\,cfl^2 + M^2 - 2\,cfl\right)\sin(\phi) + \left(M^2\,ga - 6\right)cfl^2 + 4\,cfl - 2\,M^2\right)cfl\cos(\phi)$$
$$- \mathrm{I}\left(-cfl^3\,ga + (ga-1)\,cfl^2 + cfl\,M^2 - 2\,M^2\right)cfl\sin(\phi) + \left(-M^2\,ga + 2\right)cfl^3$$
$$\left. + \left(M^2\,ga - 2\right)cfl^2 + 2\,cfl\,M^2 - 2\,M^2\right)x\Big)$$
$$+ \frac{1}{-2\,cfl^2\cos(\phi) + M^2 + 2\,cfl^2}\left(cfl^4\left(M^2\,ga + 1\right)\cos(\phi)^4 - cfl^3\left(\mathrm{I}\left(M^2\,ga\right.\right.\right.$$
$$+ 1\Big)cfl\sin(\phi) + \left(2\,M^2\,ga + 4\right)cfl + \left(-2\,ga - 1\right)M^2 - 2\Big)\cos(\phi)^3 + cfl^2\left(\mathrm{I}\left(\left(M^2\,ga\right.\right.\right.$$
$$+ 3\Big)cfl + \left(-ga - 1\right)M^2 - 1\Big)cfl\sin(\phi) + 6\,cfl^2 + \left(\left(-2\,ga - 1\right)M^2 - 6\right)cfl + (ga$$
$$+ 2)\,M^2 + 1\Big)\cos(\phi)^2 + cfl\left(\mathrm{I}\left(\left(M^2\,ga - 3\right)cfl^2 + \left(2\,M^2 + 2\right)cfl - 3\,M^2\right)cfl\sin(\phi)\right.$$
$$+ \left(2\,M^2\,ga - 4\right)cfl^3 + \left(\left(-2\,ga - 1\right)M^2 + 6\right)cfl^2 + \left(-2\,M^2 - 2\right)cfl + 2\,M^2\Big)$$
$$\cos(\phi) - \mathrm{I}\left(\left(M^2\,ga - 1\right)cfl^3 + \left(\left(-ga + 1\right)M^2 + 1\right)cfl^2 - 3\,cfl\,M^2 + 2\,M^2\right)cfl\sin(\phi)$$
$$+ \left(-M^2\,ga + 1\right)cfl^4 + \left((2\,ga + 1)\,M^2 - 2\right)cfl^3 + \left(-M^2\,ga + 1\right)cfl^2 - 2\,cfl\,M^2$$
$$\left. + M^2\right)$$

```
> Polys := (x, cfl, M, ga, phi)->x^2 + 1/(M^2+2*cfl^2*(1-cos(phi)))*((-(-I*cfl*sin(phi)*ga
  + M^2*ga + 2)*cfl^3*cos(phi)^3 + ((-cfl*ga + ga + 1)*cfl*sin(phi)*I + (M^2*ga + 6)*cfl
  - M^2*ga - 2)*cfl^2*cos(phi)^2 + cfl*(cfl*(-cfl^2*ga + M^2 - 2*cfl)*sin(phi)*I + (M^2*ga
  - 6)*cfl^2 + 4*cfl - 2*M^2)*cos(phi) - I*(-cfl^3*ga + (ga - 1)*cfl^2 + cfl*M^2 - 2*M^2)
  *cfl*sin(phi) + (-M^2*ga + 2)*cfl^3 + (M^2*ga - 2)*cfl^2 + 2*cfl*M^2 - 2*M^2)*x)
  + 1/(M^2+2*cfl^2*(1-cos(phi)))*(cfl^4*(M^2*ga + 1)*cos(phi)^4 - (cfl*(M^2*ga + 1)
  *sin(phi)*I + (2*M^2*ga + 4)*cfl + (-2*ga - 1)*M^2 - 2)*cfl^3*cos(phi)^3 + (((M^2*ga
  + 3)*cfl + (-ga - 1)*M^2 - 1)*cfl*sin(phi)*I + 6*cfl^2 + ((-2*ga - 1)*M^2 - 6)*cfl
  + (ga + 2)*M^2 + 1)*cfl^2*cos(phi)^2 + cfl*(cfl*((M^2*ga - 3)*cfl^2 + (2*M^2 + 2)*cfl
  - 3*M^2)*sin(phi)*I + (2*M^2*ga - 4)*cfl^3 + ((-2*ga - 1)*M^2 + 6)*cfl^2 + (-2*M^2
  - 2)*cfl + 2*M^2)*cos(phi) - I*((M^2*ga - 1)*cfl^3 + ((-ga + 1)*M^2 + 1)*cfl^2 - 3
  *cfl*M^2 + 2*M^2)*cfl*sin(phi) + (-M^2*ga + 1)*cfl^4 + ((2*ga + 1)*M^2 - 2)*cfl^3 + (
  -M^2*ga + 1)*cfl^2 - 2*cfl*M^2 + M^2);
```

$$Polys := (x, cfl, M, ga, \phi) \mapsto x^2 + \frac{1}{M^2 + 2\cdot cfl^2\cdot(1-\cos(\phi))}\left(\left(-\left(-\mathrm{I}\cdot cfl\cdot\sin(\phi)\cdot ga + M^2\right.\right.\right. \tag{51}$$
$$\left.\left.\left. \cdot ga + 2\right)\cdot cfl^3\cdot\cos(\phi)\right)^3 + \left(\mathrm{I}\cdot\left(-cfl\cdot ga + ga + 1\right)\cdot cfl\cdot\sin(\phi) + \left(M^2\cdot ga + 6\right)\cdot cfl - M^2\cdot ga\right.$$
$$\left. - 2\right)\cdot cfl^2\cdot\cos(\phi)^2 + cfl\cdot\left(\mathrm{I}\cdot cfl\cdot\left(-cfl^2\cdot ga + M^2 - 2\cdot cfl\right)\cdot\sin(\phi) + \left(M^2\cdot ga - 6\right)\cdot cfl^2 + 4\right.$$
$$\cdot cfl - 2\cdot M^2\Big)\cdot\cos(\phi) - \mathrm{I}\cdot\left(-cfl^3\cdot ga + (ga-1)\cdot cfl^2 + cfl\cdot M^2 - 2\cdot M^2\right)\cdot cfl\cdot\sin(\phi) + \Big($$
$$-M^2\cdot ga + 2\Big)\cdot cfl^3 + \left(M^2\cdot ga - 2\right)\cdot cfl^2 + 2\cdot cfl\cdot M^2 - 2\cdot M^2\Big)\cdot x\Big)$$
$$+ \frac{1}{M^2 + 2\cdot cfl^2\cdot(1-\cos(\phi))}\left(cfl^4\cdot\left(M^2\cdot ga + 1\right)\cdot\cos(\phi)^4 - \left(\mathrm{I}\cdot cfl\cdot\left(M^2\cdot ga + 1\right)\right.\right.$$
$$\cdot\sin(\phi) + \left(2\cdot M^2\cdot ga + 4\right)\cdot cfl + \left(-2\cdot ga - 1\right)\cdot M^2 - 2\Big)\cdot cfl^3\cdot\cos(\phi)^3 + \Big(\mathrm{I}\cdot\Big(\Big(M^2\cdot ga$$
$$+ 3\Big)\cdot cfl + \left(-ga - 1\right)\cdot M^2 - 1\Big)\cdot cfl\cdot\sin(\phi) + 6\cdot cfl^2 + \left(\left(-2\cdot ga - 1\right)\cdot M^2 - 6\right)\cdot cfl$$
$$+ (ga + 2)\cdot M^2 + 1\Big)\cdot cfl^2\cdot\cos(\phi)^2 + cfl\cdot\Big(\mathrm{I}\cdot cfl\cdot\Big(\left(M^2\cdot ga - 3\right)\cdot cfl^2 + \left(2\cdot M^2 + 2\right)\cdot cfl - 3$$
$$\cdot M^2\Big)\cdot\sin(\phi) + \left(2\cdot M^2\cdot ga - 4\right)\cdot cfl^3 + \left(\left(-2\cdot ga - 1\right)\cdot M^2 + 6\right)\cdot cfl^2 + \left(-2\cdot M^2 - 2\right)\cdot cfl$$
$$+ 2\cdot M^2\Big)\cdot\cos(\phi) - \mathrm{I}\cdot\Big(\left(M^2\cdot ga - 1\right)\cdot cfl^3 + \left(\left(-ga + 1\right)\cdot M^2 + 1\right)\cdot cfl^2 - 3\cdot cfl\cdot M^2 + 2$$
$$\cdot M^2\Big)\cdot cfl\cdot\sin(\phi) + \left(-M^2\cdot ga + 1\right)\cdot cfl^4 + \left((2\cdot ga + 1)\cdot M^2 - 2\right)\cdot cfl^3 + \left(-M^2\cdot ga + 1\right)$$
$$\cdot cfl^2 - 2\cdot cfl\cdot M^2 + M^2\Big)$$

```
> test := collect(simplify(Polys(x, cfl, M, ga, phi) - Qamps(x)), x);
```

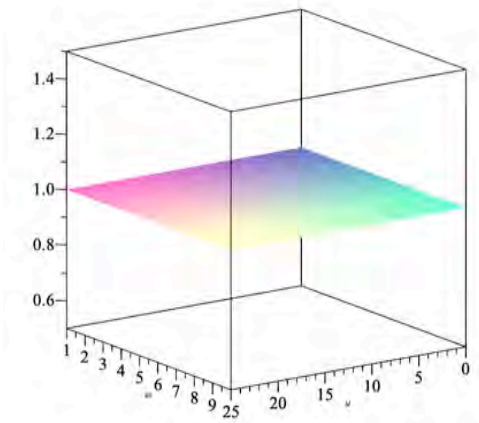
$$test := 0 \tag{52}$$

```
> nphi := 21 : hphi := 2*Pi/nphi : dessin_polys := (M, cfl, ga)
    ->max(seq(seq(evalf(abs(solve(Polys(x, cfl, M, ga, (j-1)*hphi) = 0, x)[i])), i = 1..2), j
  = 1..nphi+1));
```

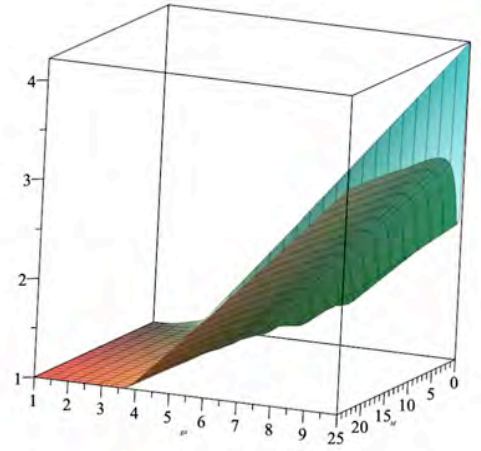
```
Warning, (in dessin_polys) `j` is implicitly declared local
Warning, (in dessin_polys) `i` is implicitly declared local
```

$$dessin\_polys := (M, cfl, ga) \mapsto \max(seq(seq(evalf(|solve(Polys(x, cfl, M, ga, (j-1)\cdot hphi) \tag{53}$$
$$= 0, x)_i|), i = 1..2), j = 1..nphi + 1))$$

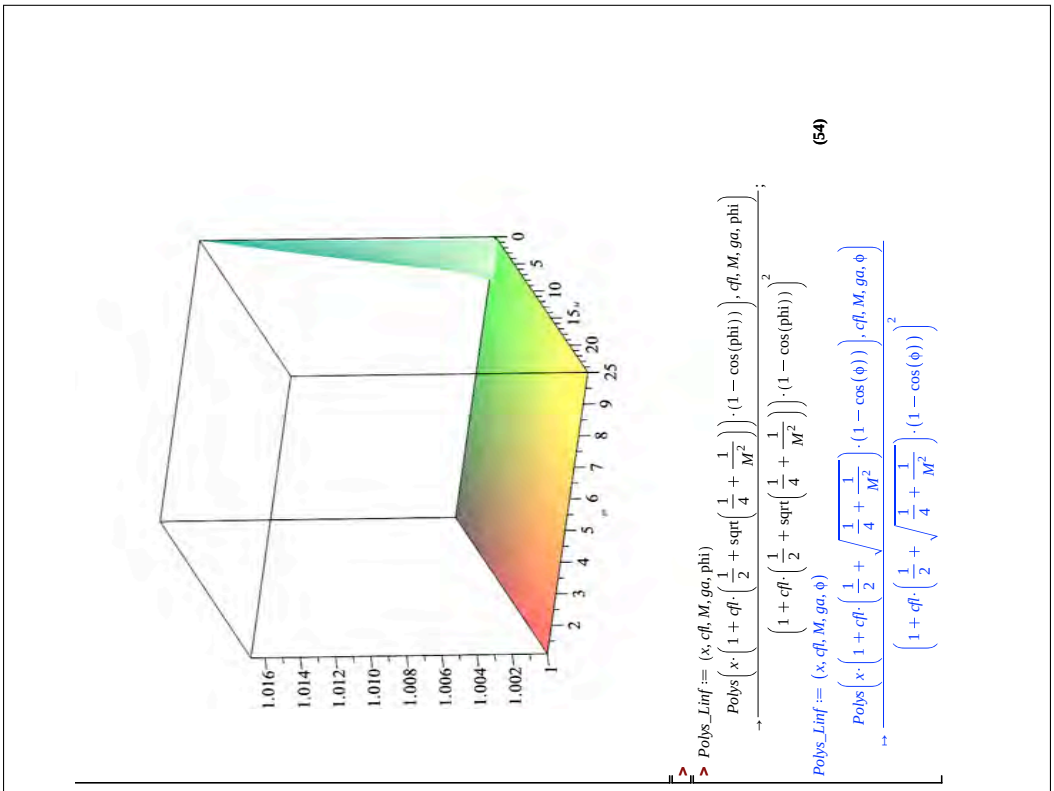
```
> plot3d(dessin_polys(M, 1/ga, ga), M = 0..25, ga = 1..10);
```

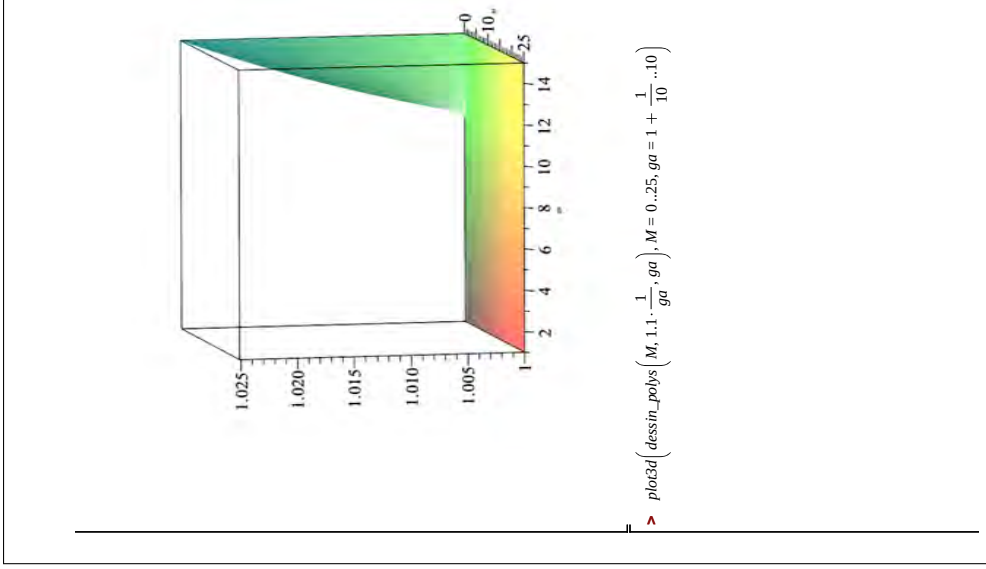



```
> plot3d(dessin_polys(M, 1, ga), M = 0..25, ga = 1..10);
```



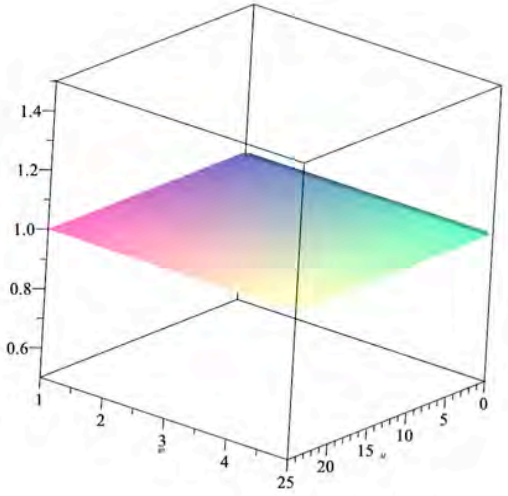

```
> plot3d(dessin_polys(M, 1/ga, ga), M = 0..25, ga = 1..15);
```


Polys_Linf := (x, cfl, M, ga, phi)
Polys(x · (1 + cfl · (1/2 + sqrt(1/4 + 1/M^2))) · (1 − cos(phi))), cfl, M, ga, phi)
→ ----------
(1 + cfl · (1/2 + sqrt(1/4 + 1/M^2))) · (1 − cos(phi)))^2
(54)
Polys_Linf := (x, cfl, M, ga, ϕ)
Polys(x · (1 + cfl · (1/2 + sqrt(1/4 + 1/M^2)) · (1 − cos(ϕ))), cfl, M, ga, ϕ)
↦ ----------
(1 + cfl · (1/2 + sqrt(1/4 + 1/M^2)) · (1 − cos(ϕ)))^2

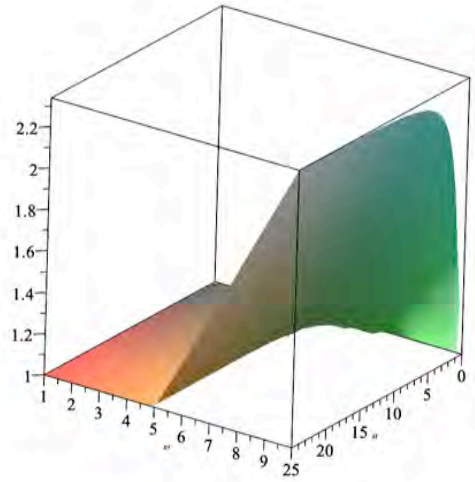

plot3d(dessin_polys(M, 1.1 · 1/ga, ga), M = 0..25, ga = 1 + 1/10..10)

```
> nphi := 21 : hphi := 2·Pi/nphi : dessin_PolysLinf := (M, cfl, ga)
      →max(seq(seq(evalf(abs(solve(Polys_Linf(x, cfl, M, ga, (j − 1)·hphi) = 0, x)[i])), i = 1
      ..2), j = 1..nphi + 1))
Warning, (in dessin_PolysLinf) `j` is implicitly declared local
Warning, (in dessin_PolysLinf) `i` is implicitly declared local
dessin_PolysLinf := (M, cfl, ga) ↦ max(seq(seq(evalf(|solve(Polys_Linf(x, cfl, M, ga, (j − 1)   (55)
    ·hphi) = 0, x)_i|), i = 1..2), j = 1..nphi + 1))
```

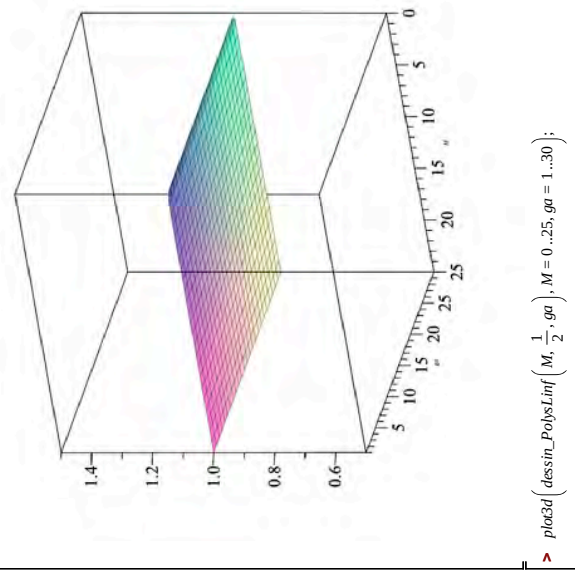
```
> plot3d(dessin_PolysLinf(M, 1, ga), M = 0..25, ga = 1..5);
```



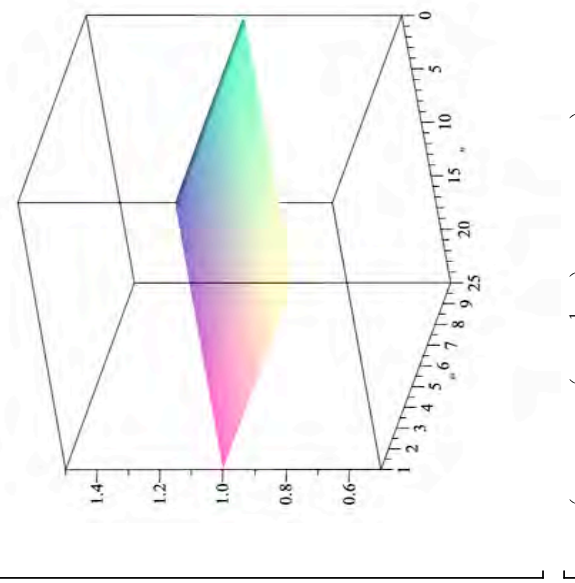

```
> plot3d(dessin_PolysLinf(M, 1, ga), M = 0..25, ga = 1..10);
```



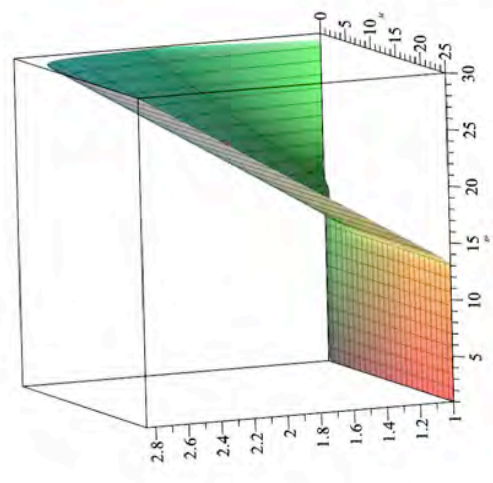

```
> plot3d(dessin_PolysLinf(M, 1/ga, ga), M = 0..25, ga = 1..10);
```




```
plot3d( dessin_PolysLinf( M, 1/2, ga ), M = 0..25, ga = 1..30 );
```




```
plot3d( dessin_PolysLinf( M, 1.7-1/ga, ga ), M = 0..25, ga = 1..30 );
```


0
5
10
15
20
25
30
1
5
10
15
20
25
30
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8

Appendix B

Maple: Modified non linear L^2 scheme proposed in [11]

```
> restart : with(LinearAlgebra) with(linalg) : with(plots) :
> F := Vector(3) : F[1] := rho·u : F[2] := rho·u·u + c^2·rho/(ga·epsilon) : F[3] := epsilon·rho·u^3/2
    + c^2·rho·u/(ga − 1) : F;
```

$$\begin{bmatrix} \rho\, u \\ \rho\, u^2 + \frac{c^2\,\rho}{ga\,\epsilon} \\ \frac{\epsilon\,\rho\,u^3}{2} + \frac{c^2\,\rho\,u}{ga-1} \end{bmatrix} \tag{1}$$

```
> Id := Matrix(3, shape = identity) :
> Df := Matrix(3) :
> Df[1,2] := 1 : Df[2,1] := (ga − 3)·u^2/2 : Df[2,2] := (3 − ga)·u : Df[2,3] := (ga − 1)/epsilon :
> Df[3,1] := −c^2·u/(ga − 1) + (ga − 2)/2 ·epsilon·u^3 : Df[3,2] := c^2/(ga − 1) + (3 − 2·ga)/2
    ·epsilon·u^2 : Df[3,3] := ga·u :
> Df;
```

$$\begin{bmatrix} 0 & 1 & 0 \\ \frac{(ga-3)\,u^2}{2} & (3-ga)\,u & \frac{ga-1}{\epsilon} \\ -\frac{c^2\,u}{ga-1} + \frac{(ga-2)\,\epsilon\,u^3}{2} & \frac{c^2}{ga-1} + \frac{(3-2\,ga)\,\epsilon\,u^2}{2} & ga\,u \end{bmatrix} \tag{2}$$

```
> L := simplify(Eigenvalues(Df));
```

$$L := \begin{bmatrix} u \\ \frac{c\sqrt{\epsilon} + \epsilon\,u}{\epsilon} \\ \frac{-c\sqrt{\epsilon} + \epsilon\,u}{\epsilon} \end{bmatrix} \tag{3}$$

```
> U := Vector(3) : U[1] := rho : U[2] := rho·u : U[3] := epsilon·rho·u^2/2 + rho·c^2/(ga·(ga − 1)) :
  U;
```

$$\begin{bmatrix} \rho \\ \rho\, u \\ \frac{\epsilon\,\rho\,u^2}{2} + \frac{\rho\,c^2}{ga\,(ga-1)} \end{bmatrix} \tag{4}$$

```
> P := Matrix(3) : P[1,2] := 1 : P[2,2] := u : P[3,2] := epsilon/2 ·u^2 : P[1,1] := 1 : P[1,
    3] := 1 :
> P[2,1] := u − c/sqrt(epsilon) : P[2,3] := u + c/sqrt(epsilon) :
> P[3,1] := epsilon·u^2/2 − sqrt(epsilon)·u·c + c^2/(ga − 1) :
> P[3,3] := epsilon·u^2/2 + sqrt(epsilon)·u·c + c^2/(ga − 1) :
> P;
```

$$\begin{bmatrix} 1 & 1 & 1 \\ u - \frac{c}{\sqrt{\epsilon}} & u & u + \frac{c}{\sqrt{\epsilon}} \\ \frac{\epsilon\,u^2}{2} - \sqrt{\epsilon}\,u\,c + \frac{c^2}{ga-1} & \frac{\epsilon\,u^2}{2} & \frac{\epsilon\,u^2}{2} + \sqrt{\epsilon}\,u\,c + \frac{c^2}{ga-1} \end{bmatrix} \tag{5}$$

```
> Pinv := MatrixInverse(P);
```

$$Pinv := \begin{bmatrix} \frac{u\,(\epsilon\,ga\,u - \epsilon\,u + 2\,c\sqrt{\epsilon})}{4\,c^2} & -\frac{\sqrt{\epsilon}\,(u\sqrt{\epsilon}\,ga - u\sqrt{\epsilon} + c)}{2\,c^2} & \frac{ga-1}{2\,c^2} \\ \frac{-\epsilon\,u^2\,ga + \epsilon\,u^2 + 2\,c^2}{2\,c^2} & \frac{\epsilon\,u\,(ga-1)}{c^2} & -\frac{ga-1}{c^2} \\ -\frac{u\,(-\epsilon\,ga\,u + 2\,c\sqrt{\epsilon} + \epsilon\,u)}{4\,c^2} & -\frac{\sqrt{\epsilon}\,(u\sqrt{\epsilon}\,ga - u\sqrt{\epsilon} - c)}{2\,c^2} & \frac{ga-1}{2\,c^2} \end{bmatrix} \tag{6}$$

```
> simplify(convert(multiply(P, Pinv), Matrix));
```

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{7}$$

```
> DD := simplify(convert(multiply(Pinv, convert(multiply(Df, P), Matrix)), Matrix));
```

(8)

$$DD := \begin{bmatrix} -\frac{-u\sqrt{\epsilon}+c}{\sqrt{\epsilon}} & 0 & 0 \\ 0 & u & 0 \\ 0 & 0 & \frac{u\sqrt{\epsilon}+c}{\sqrt{\epsilon}} \end{bmatrix} \quad (8)$$

> $Fe := Vector(3) : Fe[1] := \mathrm{rho} \cdot u : Fe[2] := \mathrm{rho} \cdot u^2 : Fe[3] := \frac{\mathrm{epsilon} \cdot \mathrm{rho} \cdot u^3}{2} : Fe;$

$$\begin{bmatrix} \rho\, u \\ \rho\, u^2 \\ \frac{\epsilon\, \rho\, u^3}{2} \end{bmatrix} \quad (9)$$

> $Dfe := Matrix(3) : Dfe[1,2] := 1 : Dfe[2,1] := -u^2 : Dfe[2,2] := 2 \cdot u : Dfe[3,1] := -\mathrm{epsilon} \cdot u^3 : Dfe[3,2] := \frac{3 \cdot \mathrm{epsilon}}{2} \cdot u^2 : Dfe;$

$$\begin{bmatrix} 0 & 1 & 0 \\ -u^2 & 2\,u & 0 \\ -\epsilon\, u^3 & \frac{3\,\epsilon\, u^2}{2} & 0 \end{bmatrix} \quad (10)$$

> $Eigenvalues(Dfe);$

$$\begin{bmatrix} 0 \\ u \\ u \end{bmatrix} \quad (11)$$

> $Fi := Vector(3) : Fi[2] := \frac{c^2 \cdot \mathrm{rho}}{ga \cdot \mathrm{epsilon}} : Fi[3] := \frac{c^2 \cdot \mathrm{rho} \cdot u}{ga - 1} : Fi;$

$$\begin{bmatrix} 0 \\ \frac{c^2\, \rho}{ga\, \epsilon} \\ \frac{c^2\, \rho\, u}{ga - 1} \end{bmatrix} \quad (12)$$

> $Dfi := Matrix(3) : Dfi[2,1] := \frac{(ga-1) \cdot u^2}{2} : Dfi[3,1] := -\frac{c^2 \cdot u}{ga-1} + \frac{ga \cdot \mathrm{epsilon} \cdot u^3}{2} : Dfi[2,2] := -(ga-1) \cdot u : Dfi[3,2] := \frac{c^2}{ga-1} - \mathrm{epsilon} \cdot ga \cdot u^2 : Dfi[2,3] := \frac{(ga-1)}{\mathrm{epsilon}} : Dfi[3,3] := ga \cdot u : Dfi;$

$$\begin{bmatrix} 0 & 0 & 0 \\ \frac{(ga-1)\, u^2}{2} & -(ga-1)\, u & \frac{ga-1}{\epsilon} \\ -\frac{c^2\, u}{ga-1} + \frac{ga\, u^3\, \epsilon}{2} & \frac{c^2}{ga-1} - \epsilon\, u^2\, ga & ga\, u \end{bmatrix} \quad (13)$$

> $Lai := Eigenvalues(Dfi);$

$$Lai := \begin{bmatrix} 0 \\ \frac{\epsilon\, u + \sqrt{\epsilon^2\, u^2 + 4\, c^2\, \epsilon}}{2\, \epsilon} \\ -\frac{-\epsilon\, u + \sqrt{\epsilon^2\, u^2 + 4\, c^2\, \epsilon}}{2\, \epsilon} \end{bmatrix} \quad (14)$$

> $DDi := simplify(convert(multiply(Pinv, convert(multiply(Dfi, P), Matrix)), Matrix));$

$$DDi := \begin{bmatrix} \frac{u\sqrt{\epsilon}-2\,c}{2\sqrt{\epsilon}} & 0 & \frac{u}{2} \\ \frac{-u\sqrt{\epsilon}+c}{\sqrt{\epsilon}} & 0 & -\frac{u\sqrt{\epsilon}+c}{\sqrt{\epsilon}} \\ \frac{u}{2} & 0 & \frac{u\sqrt{\epsilon}+2\,c}{2\sqrt{\epsilon}} \end{bmatrix} \quad (15)$$

> $DDe := simplify(convert(multiply(Pinv, convert(multiply(Dfe, P), Matrix)), Matrix));$

$$DDe := \begin{bmatrix} \frac{u}{2} & 0 & -\frac{u}{2} \\ -\frac{-u\sqrt{\epsilon}+c}{\sqrt{\epsilon}} & u & \frac{u\sqrt{\epsilon}+c}{\sqrt{\epsilon}} \\ -\frac{u}{2} & 0 & \frac{u}{2} \end{bmatrix} \quad (16)$$

> $DDes := Matrix(3) : DDes[1,1] := \frac{u}{2} : DDes[1,3] := -\frac{u}{2} : DDes[2,1] := u - \frac{c}{\mathrm{sqrt(epsilon)}} : DDes[2,2] := u : DDes[2,3] := u + \frac{c}{\mathrm{sqrt(epsilon)}} : DDes[3,1] := -\frac{u}{2} : DDes[3,3] := \frac{u}{2} : DDes;$

(17)

$$\begin{bmatrix} \frac{u}{2} & 0 & -\frac{u}{2} \\ u-\frac{c}{\sqrt{\epsilon}} & u & u+\frac{c}{\sqrt{\epsilon}} \\ -\frac{u}{2} & 0 & \frac{u}{2} \end{bmatrix} \quad (17)$$

> $DDis := Matrix(3) : DDis[1,1] := \frac{u}{2} - \frac{c}{\mathrm{sqrt}(\mathrm{epsilon})} : DDis[1,3] := \frac{u}{2} : DDis[2,1] := -u + \frac{c}{\mathrm{sqrt}(\mathrm{epsilon})} : DDis[2,3] := -u - \frac{c}{\mathrm{sqrt}(\mathrm{epsilon})} : DDis[3,1] := \frac{u}{2} : DDis[3,3] := \frac{u}{2} + \frac{c}{\mathrm{sqrt}(\mathrm{epsilon})} : DDis;$

$$\begin{bmatrix} \frac{u}{2}-\frac{c}{\sqrt{\epsilon}} & 0 & \frac{u}{2} \\ -u+\frac{c}{\sqrt{\epsilon}} & 0 & -u-\frac{c}{\sqrt{\epsilon}} \\ \frac{u}{2} & 0 & \frac{u}{2}+\frac{c}{\sqrt{\epsilon}} \end{bmatrix} \quad (18)$$

> *# a partir de là alpha=deltat sur deltax*

> $Be := Id - I\cdot\mathrm{alpha}\cdot\sin(\mathrm{phi})\cdot DDes + u\cdot\mathrm{alpha}\cdot(\cos(\mathrm{phi})-1)\cdot Id$;*#matrice fourier explicite*

$$Be := \left[\left[1-\frac{I\,\alpha\,\sin(\phi)\,u}{2}+u\,\alpha\,(\cos(\phi)-1),\ 0,\ \frac{I}{2}\,\alpha\,\sin(\phi)\,u\right],\right.$$
$$\left[-I\,\alpha\,\sin(\phi)\left(u-\frac{c}{\sqrt{\epsilon}}\right),\ 1-I\,\alpha\,\sin(\phi)\,u+u\,\alpha\,(\cos(\phi)-1),\ -I\,\alpha\,\sin(\phi)\left(u+\frac{c}{\sqrt{\epsilon}}\right)\right],$$
$$\left.\left[\frac{I}{2}\,\alpha\,\sin(\phi)\,u,\ 0,\ 1-\frac{I\,\alpha\,\sin(\phi)\,u}{2}+u\,\alpha\,(\cos(\phi)-1)\right]\right] \quad (19)$$

> $Bbarrei := Id + I\cdot\mathrm{alpha}\cdot\sin(\mathrm{phi})\cdot DDis$;*#matrice fourier implicite*

$$Bbarrei := \begin{bmatrix} 1+I\,\alpha\,\sin(\phi)\left(\frac{u}{2}-\frac{c}{\sqrt{\epsilon}}\right) & 0 & \frac{I}{2}\,\alpha\,\sin(\phi)\,u \\ I\,\alpha\,\sin(\phi)\left(-u+\frac{c}{\sqrt{\epsilon}}\right) & 1 & I\,\alpha\,\sin(\phi)\left(-u-\frac{c}{\sqrt{\epsilon}}\right) \\ \frac{I}{2}\,\alpha\,\sin(\phi)\,u & 0 & 1+I\,\alpha\,\sin(\phi)\left(\frac{u}{2}+\frac{c}{\sqrt{\epsilon}}\right) \end{bmatrix} \quad (20)$$

> $Eigenvalues(Be)$;
#sous cfl alpha·u <= 1 les valeurs propres sont toutes plus petites que 1 mais condition ni nécessaire ni suffisante

$$\begin{bmatrix} u\,\alpha\cos(\phi)-u\,\alpha+1 \\ -I\,\alpha\,\sin(\phi)\,u+u\,\alpha\cos(\phi)-u\,\alpha+1 \\ -I\,\alpha\,\sin(\phi)\,u+u\,\alpha\cos(\phi)-u\,\alpha+1 \end{bmatrix} \quad (21)$$

> $Eigenvalues(Bbarrei)$;
 # valeurs propres toutes plus grandes que 1 mais condition ni névessaire ni suffisante

$$\begin{bmatrix} \frac{I\sin(\phi)\,\epsilon\,\alpha\,u+2\,\epsilon+\sqrt{-\sin(\phi)^2\,\alpha^2\,\epsilon^2\,u^2-4\sin(\phi)^2\,\alpha^2\,c^2\,\epsilon}}{2\,\epsilon} \\ \frac{I\sin(\phi)\,\epsilon\,\alpha\,u+2\,\epsilon-\sqrt{-\sin(\phi)^2\,\alpha^2\,\epsilon^2\,u^2-4\sin(\phi)^2\,\alpha^2\,c^2\,\epsilon}}{2\,\epsilon} \\ 1 \end{bmatrix} \quad (22)$$

> *# on pose* $\frac{1}{\mathrm{beta}} = M = \frac{\mathrm{sqrt}(\mathrm{epsilon})\cdot u}{c}$ *et* $\mathrm{alpha}\cdot u = 1,\ u > 0$

> $Bee := Matrix(3,3) : Bee[1,1] := \cos(\mathrm{phi}) - I\cdot\frac{\sin(\mathrm{phi})}{2} : Bee[1,3] := \frac{I\cdot\sin(\mathrm{phi})}{2} : Bee[2,1] := -I\cdot\sin(\mathrm{phi})\cdot(1-\mathrm{beta}) : Bee[2,2] := \cos(\mathrm{phi}) - I\cdot\sin(\mathrm{phi}) : Bee[2,3] := -I\cdot\sin(\mathrm{phi})\cdot(1+\mathrm{beta}) : Bee[3,1] := \frac{I}{2}\cdot\sin(\mathrm{phi}) : Bee[3,3] := \cos(\mathrm{phi}) - \frac{I\cdot\sin(\mathrm{phi})}{2} : Bee;$

$$\begin{bmatrix} \cos(\phi)-\frac{I\sin(\phi)}{2} & 0 & \frac{I}{2}\sin(\phi) \\ -I\sin(\phi)\,(1-\beta) & \cos(\phi)-I\sin(\phi) & -I\sin(\phi)\,(1+\beta) \\ \frac{I}{2}\sin(\phi) & 0 & \cos(\phi)-\frac{I\sin(\phi)}{2} \end{bmatrix} \quad (23)$$

> $Bbarrei := Matrix(3,3) : Bbarrei[1,1] := 1 + I\cdot\sin(\mathrm{phi})\cdot\left(\frac{1}{2}-\mathrm{beta}\right) : Bbarrei[1,3] := \frac{I\cdot\sin(\mathrm{phi})}{2} : Bbarrei[2,1] := I\cdot\sin(\mathrm{phi})\cdot(-1+\mathrm{beta}) : Bbarrei[2,2] := 1 : Bbarrei[2,3] := I\cdot\sin(\mathrm{phi})\cdot(-1-\mathrm{beta}) : Bbarrei[3,1] := \frac{I\cdot\sin(\mathrm{phi})}{2} : Bbarrei[3,3] := 1 + I\cdot\sin(\mathrm{phi})\cdot\left(\frac{1}{2}+\mathrm{beta}\right) : Bbarrei;\ det(Bbarrei);$

$$\begin{bmatrix} 1+I\sin(\phi)\left(\frac{1}{2}-\beta\right) & 0 & \frac{I}{2}\sin(\phi) \\ I\sin(\phi)\,(-1+\beta) & 1 & I\sin(\phi)\,(-1-\beta) \\ \frac{I}{2}\sin(\phi) & 0 & 1+I\sin(\phi)\left(\frac{1}{2}+\beta\right) \end{bmatrix}$$
$$1 + I\sin(\phi) + \sin(\phi)^2\,\beta^2 \quad (24)$$

```
> #inverse de Bbarrei
> temp := Matrix(3) : temp[1, 1] := 1 + I·sin(phi)·(1/2 + beta) : temp[1, 3] :=
  -I·sin(phi)/2 : temp[2, 1] := sin(phi)^2·β^2 + I·sin(phi)·(1 - beta) : temp[2, 2] := 1 + I
  ·sin(phi) + sin(phi)^2·β^2 : temp[2, 3] := sin(phi)^2·β^2 + I·sin(phi)·(1 + beta) : temp[3,
  1] := -I·sin(phi)/2 : temp[3, 3] := 1 + I·sin(phi)·(1/2 - beta) : Bbarreiinv :=
  simplify(1/(1 + I·sin(phi) + sin(phi)^2·β^2)·temp);
```

$$Bbarreiinv := \begin{bmatrix} \frac{2 + (2\,I\,\beta + I)\sin(\phi)}{2\sin(\phi)^2\beta^2 + 2\,I\sin(\phi) + 2} & 0 & \frac{-I\sin(\phi)}{2\sin(\phi)^2\beta^2 + 2\,I\sin(\phi) + 2} \\ -\frac{\sin(\phi)\,(-\sin(\phi)\,\beta^2 + I\,\beta - I)}{1 + I\sin(\phi) + \sin(\phi)^2\beta^2} & 1 & \frac{\sin(\phi)\,(\sin(\phi)\,\beta^2 + I\,\beta + I)}{1 + I\sin(\phi) + \sin(\phi)^2\beta^2} \\ \frac{-I\sin(\phi)}{2\sin(\phi)^2\beta^2 + 2\,I\sin(\phi) + 2} & 0 & \frac{2 + (-2\,I\,\beta + I)\sin(\phi)}{2\sin(\phi)^2\beta^2 + 2\,I\sin(\phi) + 2} \end{bmatrix} \quad (25)$$

```
> simplify(Multiply(Bbarreiinv, Bbarrei));
```

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (26)$$

```
> Mat := simplify(convert(Multiply(Bbarreiinv, Bee), Matrix)); Mat[2, 2];
```

$$Mat := \left[\left[\frac{(-1-\beta)\cos(\phi)^2 + (2 + (2\,I\,\beta + I)\sin(\phi))\cos(\phi) - I\sin(\phi) + \beta + 1}{2\,I\sin(\phi) - 2\,\beta^2\cos(\phi)^2 + 2\,\beta^2 + 2},\ 0,\right.\right.$$
$$\left.-\frac{\sin(\phi)\,(I\cos(\phi) + \sin(\phi)\,\beta - I + \sin(\phi))}{2\,I\sin(\phi) - 2\,\beta^2\cos(\phi)^2 + 2\,\beta^2 + 2}\right],$$
$$\left[-\frac{1}{-\beta^2\cos(\phi)^2 + I\sin(\phi) + \beta^2 + 1}\left(\left(I\,\beta^2(-1+\beta)\cos(\phi)^2 + \left(-\sin(\phi)\,\beta^2\right.\right.\right.\right.$$
$$\left.\left.\left.\left.+ I\,\beta - I\right)\cos(\phi) + (2\,\beta - 1)\sin(\phi) - I\left(\beta^2 + 1\right)(-1+\beta)\right)\sin(\phi)\right), \cos(\phi)$$
$$\left.\left.- I\sin(\phi),\ \frac{1}{-\beta^2\cos(\phi)^2 + I\sin(\phi) + \beta^2 + 1}\left(\left(I\,\beta^2(1+\beta)\cos(\phi)^2\right.\right.\right.\right.$$
$$\left.\left.\left.\left.+ \left(\sin(\phi)\,\beta^2 + I\,\beta + I\right)\cos(\phi) + (2\,\beta + 1)\sin(\phi) - I(1+\beta)\left(\beta^2 + 1\right)\right)\sin(\phi)\right)\right],$$
$$\left[-\frac{\sin(\phi)\,(I\cos(\phi) - \sin(\phi)\,\beta - I + \sin(\phi))}{2\,I\sin(\phi) - 2\,\beta^2\cos(\phi)^2 + 2\,\beta^2 + 2},\ 0,\right.$$
$$\left.\left.\frac{(-1+\beta)\cos(\phi)^2 + (2 + (-2\,I\,\beta + I)\sin(\phi))\cos(\phi) - I\sin(\phi) - \beta + 1}{2\,I\sin(\phi) - 2\,\beta^2\cos(\phi)^2 + 2\,\beta^2 + 2}\right]\right]$$
$$\cos(\phi) - I\sin(\phi) \quad (27)$$

```
> a31 := ((-sin(phi)^2/2·(1 - beta) + I·sin(phi)/2·(1 - cos(phi)))/(1 + I·sin(phi) + sin(phi)^2·β^2)) : simplify(a31 - Mat[3,
  1]);
```

$$0 \quad (28)$$

```
> a13 := ((-sin(phi)^2/2·(1 + beta) + I·sin(phi)/2·(1 - cos(phi)))/(1 + I·sin(phi) + sin(phi)^2·β^2)) : simplify(a13 - Mat[1,
  3]);
```

$$0 \quad (29)$$

```
> a33 := -1/(1 + I·sin(phi) + sin(phi)^2·β^2)·(I·sin(phi)·cos(phi)·beta + I·sin(phi)/2·(1
  - cos(phi)) - sin(phi)^2/2·(1 - beta) - cos(phi)) : simplify(a33 - Mat[3, 3]);
```

$$0 \quad (30)$$

```
> a11 :=
  (I·sin(phi)·cos(phi)·(1/2 + beta) + sin(phi)^2/2·(1 + beta) - I·sin(phi)/2 + cos(phi))/(1 + I·sin(phi) + sin(phi)^2·β^2) :
  simplify(a11 - Mat[1, 1]);
```

$$0 \quad (31)$$

```
>
> Q := x↦(a11 - x)·(a33 - x) - a13·a31; collect(expand(Q(x)), x);
```

$$Q := x \mapsto (a11 - x)\cdot(a33 - x) - a13\cdot a31$$
$$x^2 + \left(-\frac{\sin(\phi)^2}{1 + I\sin(\phi) + \sin(\phi)^2\beta^2} + \frac{I\sin(\phi)}{1 + I\sin(\phi) + \sin(\phi)^2\beta^2}\right.$$
$$\left. - \frac{2\cos(\phi)}{1 + I\sin(\phi) + \sin(\phi)^2\beta^2} - \frac{I\sin(\phi)\cos(\phi)}{1 + I\sin(\phi) + \sin(\phi)^2\beta^2}\right)x$$
$$+ \frac{I\sin(\phi)\cos(\phi)^2}{\left(1 + I\sin(\phi) + \sin(\phi)^2\beta^2\right)^2} - \frac{I\sin(\phi)\cos(\phi)}{\left(1 + I\sin(\phi) + \sin(\phi)^2\beta^2\right)^2} \quad (32)$$

```
> # on dessine le max des modules des valeurs propres
> Pdeux := (x, phi, M)->x^2
    + (-M^2*sin(phi)^2 - 2*cos(phi)*M^2 + I*M^2*(sin(phi) - cos(phi)*sin(phi)))/(M^2 + I*sin(phi)*M^2 + sin(phi)^2)*x
    + 1/(M^2 + I*sin(phi)*M^2 + sin(phi)^2)^2*(cos(phi)^2*sin(phi)^2*M^2 + cos(phi)^2*M^4
    + cos(phi)*sin(phi)^2*M^4 + I*(cos(phi)^2*sin(phi)*M^4 - sin(phi)*cos(phi)*M^4
    - sin(phi)^3*cos(phi)*M^2));
```

$$Pdeux := (x, \phi, M) \mapsto x^2 + \frac{\left(-M^2\sin(\phi)^2 - 2\cos(\phi)M^2 + \mathrm{I}M^2(\sin(\phi) - \cos(\phi)\sin(\phi))\right)x}{M^2 + \mathrm{I}\sin(\phi)M^2 + \sin(\phi)^2} + \frac{1}{\left(M^2 + \mathrm{I}\sin(\phi)M^2 + \sin(\phi)^2\right)^2}\left(\cos(\phi)^2\sin(\phi)^2M^2 + \cos(\phi)^2M^4 + \cos(\phi)\sin(\phi)^2M^4 + \mathrm{I}\left(\cos(\phi)^2\sin(\phi)M^4 - \cos(\phi)\sin(\phi)M^4 - \sin(\phi)^3\cos(\phi)M^2\right)\right) \quad (33)$$

```
> simplify(Pdeux(x, phi, 1/beta) - Q(x));
```

$$0 \quad (34)$$

```
> Nphi := 50; hphi := 2*Pi/Nphi;
```

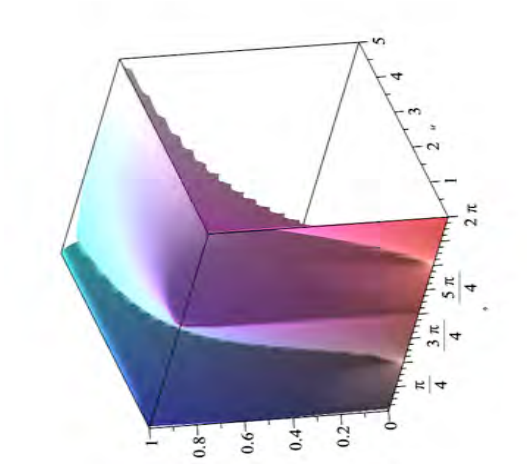
$$Nphi := 50$$

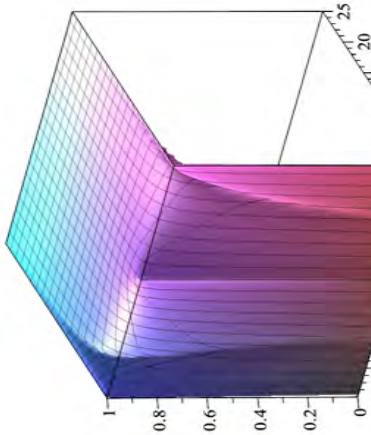
$$hphi := \frac{\pi}{25} \quad (35)$$

```
> val := (phi, M)->[solve(Pdeux(x, phi, M), x)]:
> valmax := (phi, M)->evalf(max(abs(val(phi, M)[1]), abs(val(phi, M)[2]))):
> valt := M->max(seq(abs(solve(Pdeux(x, i*2*Pi/Nphi, M) = 0, x)[j]), j = 1..2), i = 0..Nphi)):
Warning, (in valt) `i` is implicitly declared local
Warning, (in valt) `j` is implicitly declared local
> plot(valt(M), M = 0..25);
```



```
> plot3d(valmax(phi, M), phi = 0..2*Pi, M = 0..25, grid = [100, 100]);
```






```
plot3d(valmax(phi, M), phi = 0..2·Pi, M = 0..5, grid = [100, 100]);
```

Bibliography

[1] Thomas Alazard. "Incompressible limit of the nonisentropic Euler equations with the solid wall boundary conditions". In: *Advances in Differential Equations* 10.1 (2005), pp. 19–44. DOI: 10.57262/ade/1355867894.

[2] Thomas Alazard. "Low Mach Number Limit of the Full Navier-Stokes Equations". In: *Archive for Rational Mechanics and Analysis* 180 (2006), pp. 1–73. DOI: 10.1007/s00205-005-0393-2.

[3] Paola Allegrini and Marie-Hélène Vignal. "Study of a new low oscillatory second order all Mach IMEX finite volume scheme for the full Euler equations". In: *submitted* (2023).

[4] Anthony A. Amsden and Francis H. Harlow. "A simplified MAC technique for incompressible fluid flow calculations". In: *Journal of Computational Physics* 6 (1970), pp. 322–325.

[5] Kiyoshi Asano. "On the incompressible limit of the compressible Euler equation". In: *Japan Journal of Applied Mathematics* 4 (1987), pp. 455–488.

[6] Stavros Avgerinos, Florian Bernard, Angelo Iollo, and Giovanni Russo. "Linearly implicit all Mach number shock capturing schemes for the Euler equations". In: *Journal of Computational Physics* 393 (2019), pp. 278–312. DOI: 10.1016/j.jcp.2019.04.020.

[7] John B. Bell, Phillip Colella, and Harland M. Glaz. "A second-order projection method for the incompressible navier-stokes equations". In: *Journal of Computational Physics* 85 (1989), pp. 257–283.

[8] Hester Bijl and Pieter Wesseling. "A Unified Method for Computing Incompressible and Compressible Flows in Boundary-Fitted Coordinates". In: *Journal of Computational Physics* 141.2 (1998), pp. 153–173. ISSN: 0021-9991. DOI: https://doi.org/10.1006/jcph.1998.5914.

[9] Georgij Bispen, Maria Lukacova-Medvid'ova, and Leonid Yelash. "Asymptotic preserving IMEX finite volume schemes for low Mach number Euler equations with gravitation". In: *Journal of Computational Physics* 335 (2017), pp. 222–248.

[10] Sebastiano Boscarino, Giovanni Russo, and Leonardo Scandurra. "All Mach Number Second Order Semi-implicit Scheme for the Euler Equations of Gas Dynamics". In: *Journal of Scientific Computing* 77 (2017), pp. 850–884.

[11] Walter Boscheri, Giacomo Dimarco, Raphaël Loubère, Maurizio Tavelli, and Marie-Hélène Vignal. "A second order all Mach number IMEX finite volume solver for the three dimensional Euler equations". In: *Journal of Computational Physics* 415 (2020), p. 109486.

[12] Walter Boscheri, Giacomo Dimarco, and Maurizio Tavelli. "An efficient second order all Mach finite volume solver for the compressible Navier-Stokes equations". In: *Computer Methods in Applied Mechanics and Engineering* 374 (2021), p. 113602. ISSN: 0045-7825. DOI: https://doi.org/10.1016/j.cma.2020.113602.

[13] Walter Boscheri, Giacomo Dimarco, and Maurizio Tavelli. "An efficient second order all Mach finite volume solver for the compressible Navier-Stokes equations". In: *Computer Methods in Applied Mechanics and Engineering* 374 (2021), p. 113602. ISSN: 0045-7825. DOI: https://doi.org/10.1016/j.cma.2020.113602.

[14] Walter Boscheri and Lorenzo Pareschi. "High order pressure-based semi-implicit IMEX schemes for the 3D Navier-Stokes equations at all Mach numbers". In: *Journal of Computational Physics* 434 (2021), p. 110206. ISSN: 0021-9991. DOI: https://doi.org/10.1016/j.jcp.2021.110206.

[15] Wen-Tzong Lee Bram van Leer and Philip L. Roe. "Characteristic time-stepping or local preconditioning of the Euler equations". In: 1991. DOI: 10.2514/6.1991-1552.

[16] Christophe Chalons, Mathieu Girardin, and Samuel Kokh. "An all-regime Lagrange-Projection like scheme for the gas dynamics equations on unstructured meshes". In: *Communications in Computational Physics* 20 (2016), pp. 188–233.

[17] Christophe Chalons, Mathieu Girardin, and Samuel Kokh. "Large Time Step and Asymptotic Preserving Numerical Schemes for the Gas Dynamics Equations with Source Terms". In: *SIAM Journal on Scientific Computing* 35.6 (2013), A2874–A2902. DOI: 10.1137/130908671.

[18] Do Hyung Choi and Charles L. Merkle. "Application of time-iterative schemes to incompressible flow". In: *AIAA Journal* 23 (1984), pp. 1518–1524.

[19] Alexandre Joel Chorin. "A numerical method for solving incompressible viscous flow problems". In: *Journal of Computational Physics* 2.1 (1967), pp. 12–26. ISSN: 0021-9991. DOI: https://doi.org/10.1016/0021-9991(67)90037-X.

[20] Stéphane Clain, Steven Diot, and Raphaël Loubère. "A high-order finite volume method for systems of conservation laws—Multi-dimensional Optimal Order Detection (MOOD)". In: *Journal of computational Physics* 230.10 (2011), pp. 4028–4050.

[21] Floraine Cordier, Pierre Degond, and Anela Kumbaro. "An Asymptotic-Preserving all-speed scheme for the Euler and Navier-Stokes equations". In: *Journal of Computational Physics* 231 (2011), pp. 5685–5704.

[22] Iain G. Currie. "Fundamental Mechanics of Fluids, 2nd Ed." In: vol. 9. McGraw Hill, 1993, p. 225.

[23] Raphaël Danchin. "Zero Mach number limit for compressible flows with periodic boundary conditions". In: *American Journal of Mathematics* 124 (2002), pp. 1153–1219.

[24] Pierre Degond, Fabrice Deluzet, Afeintou Sangam, and Marie-Hélène Vignal. "An Asymptotic Preserving scheme for the Euler equations in a strong magnetic field". In: *Journal of Computational Physics* 228.10 (2009), pp. 3540–3558. ISSN: 0021-9991. DOI: https://doi.org/10.1016/j.jcp.2008.12.040.

[25] Pierre Degond, Shi Jin, and Jian-guo Liu. "Mach-number uniform asymptotic-preserving gauge schemes for compressible flows". In: *Bull. Inst. Math. Acad. Sin. (N.S.)* 2 (2007).

[26] Pierre Degond and Min Tang. "All Speed Scheme for the Low Mach Number Limit of the Isentropic Euler Equations". In: *Communications in Computational Physics* 10 (2009). DOI: 10.4208/cicp.210709.210610a.

[27] Stéphane Dellacherie. "Analysis of Godunov type schemes applied to the compressible Euler system at low Mach number". In: *J. Comput. Physics* 229 (2010), pp. 978–1016. DOI: 10.1016/j.jcp.2009.09.044.

[28] Benoît Desjardins, Emmanuel Grenier, Pierre-Louis Lions, and Nader Masmoudi. "Incompressible limit for solutions of the isentropic Navier-Stokes equations with Dirichlet boundary conditions". In: *Journal de Mathématiques Pures et Appliquées* 78.5 (1999), pp. 461–471. ISSN: 0021-7824. DOI: https://doi.org/10.1016/S0021-7824(99)00032-X.

[29] Giacomo Dimarco, Raphaël Loubère, Victor Michel-Dansac, and Marie-Hélène Vignal. "Second-order implicit-explicit total variation diminishing schemes for the Euler system in the low Mach regime". In: *Journal of Computational Physics* 372 (2017), pp. 178–201.

[30] Giacomo Dimarco, Raphaël Loubère, and Marie-Hélène Vignal. "Study of a New Asymptotic Preserving Scheme for the Euler System in the Low Mach Number Limit". In: *SIAM Journal on Scientific Computing* 39.5 (2017), A2099–A2128. DOI: 10.1137/16M1069274.

[31] Giacomo Dimarco and Lorenzo Pareschi. "Asymptotic Preserving Implicit-Explicit Runge–Kutta Methods for Nonlinear Kinetic Equations". In: *SIAM Journal on Numerical Analysis* 51.2 (2013), pp. 1064–1087. DOI: 10.1137/12087606X.

[32] Walid Kheriji Dionysios Grapsas Raphaèle Herbin and Jean-Claude Latché. "An unconditionally stable staggered pressure correction scheme for the compressible Navier-Stokes equations". In: 2016.

[33] Steven Diot, Stéphane Louis Clain, and Raphaël Loubère. "Improved detection criteria for the multi-dimensional optimal order detection (MOOD) on unstructured meshes with very high-order polynomials". In: *Computers & Fluids* 64 (2012), pp. 43–63.

[34] Steven Diot, Raphaël Loubère, and Stephane Clain. "The Multidimensional Optimal Order Detection method in the three-dimensional case: very high-order finite volume method for hyperbolic systems". In: *International Journal for Numerical Methods in Fluids* 73.4 (2013), pp. 362–392.

[35] Changsheng Dou, Song Jiang, and Yaobin Ou. "Low Mach number limit of full Navier-Stokes equations in a 3D bounded domain". In: *Journal of Differential Equations* 258.2 (2015), pp. 379–398. ISSN: 0022-0396. DOI: https://doi.org/10.1016/j.jde.2014.09.017.

[36] Michael Dumbser, Ilya Peshkov, Evgeniy Romenski, and Olindo Zanotti. "High order ADER schemes for a unified first order hyperbolic formulation of Newtonian continuum mechanics coupled with electro-dynamics". In: *Journal of Computational Physics* 348 (2016). DOI: 10.1016/j.jcp.2017.07.020.

[37] Michael Dumbser and Casulli Vincenzo. "A conservative, weakly nonlinear semi-implicit finite volume scheme for the compressible Navier-Stokes equations with general equation of state". In: *Applied Mathematics and Computation* 272 (2016), pp. 479–497. DOI: 10.1016/j.amc.2015.08.042.

[38] M. Elena Vázquez-Cendón Eleuterio F. Toro. "Flux splitting schemes for the Euler equations". In: *Computers & Fluids* 70 (2012). DOI: 10.1016/j.compfluid.2012.08.023.

[39] Robert Eymard, Thierry Gallouët, and Raphaèle Herbin. "Finite volume methods". In: *Handbook of Numerical Analysis* 7 (2000), pp. 713–1018.

[40] Sigal Gottlieb, Chi-Wang Shu, and Eitan Tadmor. "Strong Stability-Preserving High-Order Time Discretization Methods". In: *SIAM Review* 43.1 (2001), pp. 89–112. DOI: 10.1137/S003614450036757X.

[41] Frieder Lörcher Gregor J. Gassner and Claus-Dieter Munz. "A Discontinuous Galerkin Scheme based on a Space-Time Expansion II. Viscous Flow Equations in Multi Dimensions". In: *Journal of Scientific Computing* 34 (2008), pp. 260–286.

[42] Hervé Guillard and Angelo Murrone. "On the behavior of upwind schemes in the low Mach number limit: II. Godunov type schemes". In: *Computers & Fluids* 33 (2004), pp. 655–675.

[43] Hervé Guillard and Cécile Viozat. "On the behaviour of upwind schemes in the low Mach number limit". In: *Computers & Fluids* 28 (1999), pp. 63–86.

[44] Jeffrey Haack, Shi Jin, and Jian-Guo Liu. "An All-Speed Asymptotic-Preserving Method for the Isentropic Euler and Navier-Stokes Equations". In: *Communications in Computational Physics* 12.4 (2012), pp. 955–980. DOI: 10.4208/cicp.250910.131011a.

[45] Francis H. Harlow and Anthony A. Amsden. "Numerical calculation of almost incompressible flow". In: *Journal of Computational Physics* 3 (1968), pp. 80–93.

[46] Changqing Hu and Chi-Wang Shu. "Weighted Essentially Non-oscillatory Schemes on Triangular Meshes". In: *Journal of Computational Physics* 150.1 (1999), pp. 97–127. ISSN: 0021-9991. DOI: https://doi.org/10.1006/jcph.1998.6165.

[47] Shi Jin. "Efficient Asymptotic-Preserving (AP) Schemes For Some Multiscale Kinetic Equations". In: *SIAM Journal on Scientific Computing* 21 (2000). DOI: 10.1137/S1064827598334599.

[48] Kailash C. Karki and Suhas V. Patankar. "Pressure based calculation procedure for viscous flows at all speeds in arbitrary configurations". In: *AIAA Journal* 27 (1988), pp. 1167–1174.

[49] Sergiu Klainerman and Andrew Majda. "Compressible and incompressible fluids". English (US). In: *Communications on Pure and Applied Mathematics* 35.5 (1982), pp. 629–651. ISSN: 0010-3640. DOI: 10.1002/cpa.3160350503.

[50] Sergiu Klainerman and Andrew J. Majda. "Singular limits of quasilinear hyperbolic systems with large parameters and the incompressible limit". In: (1981).

[51] Rupert Klein. "Semi-implicit extension of a Godunov-type scheme based on low Mach number asymptotics". In: *Journal of Computational Physics* 121.2 (1995), pp. 213–237. ISSN: 0021-9991. DOI: https://doi.org/10.1016/S0021-9991(95)90034-9.

[52] Rupert Klein, Nicola Botta, Thomas Schneider, Claus-Dieter Munz, Sabine Roller, Andreas Meister, L Hoffmann, and Thomas Sonar. "Asymptotic adaptive methods for multi-scale problems in fluid mechanics". In: *Journal of Engineering Mathematics* 39 (2001), pp. 261–343.

[53] Peter D. Lax and Xu-Dong Liu. "Solution of Two-Dimensional Riemann Problems of Gas Dynamics by Positive Schemes". In: *SIAM Journal on Scientific Computing* 19.2 (1998), pp. 319–340. DOI: 10.1137/S1064827595291819.

[54] Bram van Leer. "Towards the Ultimate Conservative Difference Scheme V. A Second-order Sequel to Godunov's Method". In: *Journal of Computational Physics* 135 (1997), pp. 229–248. DOI: 10.1006/jcph.1997.5704.

[55] Xue-song Li and Chun-wei Gu. "An All-Speed Roe-type scheme and its asymptotic analysis of low Mach number behaviour". In: *Journal of Computational Physics* 227 (2008), pp. 5144–5159.

[56] Xue-song Li and Chun-wei Gu. "Mechanism of Roe-type schemes for all-speed flows and its application". In: *Computers & Fluids* 86 (2013), pp. 56–70.

[57] Pierre-Louis Lions and Nader Masmoudi. "Incompressible limit for a viscous compressible fluid". In: *Journal de Mathématiques Pures et Appliquées* 77.6 (1998), pp. 585–627. ISSN: 0021-7824. DOI: https://doi.org/10.1016/S0021-7824(98)80139-6.

[58] Pierre-Louis Lions and Nader Masmoudi. "Incompressible limit for a viscous compressible fluid". In: *Journal de mathématiques pures et appliquées* 77.6 (1998), pp. 585–627.

[59] Guy Metivier and Steve Schochet. "Averaging theorems for conservative systems and the weakly compressible Euler equations". In: *Journal of Differential Equations* 187 (2003), pp. 106–183. DOI: 10.1016/S0022-0396(02)00037-2.

[60] Guy Metivier and Steve Schochet. "The Incompressible Limit of the Non-Isentropic Euler Equations". In: *Archive for Rational Mechanics and Analysis* 158 (2001), pp. 61–90. DOI: 10.1007/PL00004241.

[61] Victor Michel-Dansac and Andrea Thomann. "On High-Precision L^∞-stable IMEX Schemes for Scalar Hyperbolic Multi-scale Equations". In: 2021, pp. 79–94. ISBN: 978-3-030-72849-6. DOI: 10.1007/978-3-030-72850-2_4.

[62] Fabian Miczek, Friedrich K. Roepke, and Philipp V. F. Edelmann. "New numerical solver for flows at various Mach numbers". In: *Astronomy and Astrophysics* 576 (2014), pp. 1–16.

[63] Claus-Dieter Munz, Michael Dumbser, and Sabine Roller. "Linearized Acoustic Perturbation Equations for Low Mach Number Flow with Variable Density and Temperature". In: *J. Comput. Phys.* 224.1 (2007), 352?364. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2007.02.022.

[64] Grenier Nicolas, Jean Paul Vila, and Philippe Villedieu. "An accurate low-Mach scheme for a compressible two-fluid model applied to free-surface flows". In: *Journal of Computational Physics* 252 (2013), pp. 1–19. DOI: 10.1016/j.jcp.2013.06.008.

[65] Sebastian Noelle, Georgij Bispen, Koottungal Revi Arun, Maria Lukacova-Medvid'ova, and Claus-Dieter Munz. "A Weakly Asymptotic Preserving Low Mach Number Scheme for the Euler Equations of Gas Dynamics". In: *SIAM J. Sci. Comput.* 36 (2014).

[66] Lorenzo Pareschi and Giovanni Russo. "Implicit-Explicit Runge-Kutta schemes for stiff systems of differential equations". In: vol. 3. 2001, pp. 269–288.

[67] Walid Kheriji Raphaèle Herbin and Jean-Claude Latché. "On some implicit and semi-implicit staggered schemes for the shallow water and Euler equations". In: *ESAIM: M2AN* 48.6 (2014), pp. 1807–1857. DOI: 10.1051/m2an/2014021.

[68] Walid Kheriji Raphaèle Herbin and Jean-Claude Latché. "Pressure correction staggered schemes for barotropic one-phase and two-phase flows". In: *Computers & Fluids* 88 (2013), pp. 524–542. DOI: 10.1016/j.compfluid.2013.09.022.

[69] Ventzislav Rusanov. "The calculation of the interaction of non-stationary shock waves and obstacles". In: *Ussr Computational Mathematics and Mathematical Physics* 1 (1962), pp. 304–320.

[70] Steve Schochet. "The compressible Euler equations in a bounded domain: Existence of solutions and the incompressible limit". In: *Communications in Mathematical Physics* 104.1 (1986), pp. 49–75. DOI: 10.1007/BF01210792.

[71] Steven Schochet. "Fast singular limits of hyperbolic PDEs". In: *Journal of differential equations* 114.2 (1994), pp. 476–512.

[72] Carsten W. Schulz-Rinne, James P. Collins, and Harland M. Glaz. "Numerical solution of the Riemann problem for two-dimensional gas dynamics". In: *SIAM Journal on Scientific and Statistical Computing (Society for Industrial and Applied Mathematics); (United States)* (1993). ISSN: 0196-5204. DOI: 10.1137/0914082.

[73] Evan Andrew Sewall and Danesh K. Tafti. "A time-accurate variable property algorithm for calculating flows with large temperature variations". In: *Computers & Fluids* 37 (2008), pp. 51–63.

[74] George Gabriel Stokes et al. "On the Effect of the Internal Friction of Fluids on the Motion of Pendulums". In: *Transactions of the Cambridge Philosophical Society* 9 (1851), p. 8.

[75] Min Tang. "Second order all speed method for the isentropic Euler equations". In: *Kinetic and Related Models* 5 (2012), pp. 155–184.

[76] Maurizio Tavelli and Michael Dumbser. "A pressure-based semi-implicit space-time discontinuous Galerkin method on staggered unstructured meshes for the solution of the compressible Navier-Stokes equations at all Mach numbers". In: *Journal of Computational Physics* 341 (2016), pp. 341–376.

[77] Andrea Thomann, Markus Zenk, Gabriella Puppo, and Christian Klingenberg. "An All Speed Second Order IMEX Relaxation Scheme for the Euler Equations". In: *Communications in Computational Physics* 28.2 (2020), pp. 591–620.

[78] Eleuterio Toro. "Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction". In: 2009, third edition. DOI: 10.1007/b79761.

[79] Eli Turkel. "Preconditioned methods for solving the incompressible and low speed compressible equations". In: *Journal of Computational Physics* 72.2 (1987), pp. 277–298. ISSN: 0021-9991. DOI: https://doi.org/10.1016/0021-9991(87)90084-2.

[80] Steven J. Ruuth Uri M. Ascher and Raymond J. Spiteri. "Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations". In: *Applied Numerical Mathematics* 25 (1997), pp. 151–167.

[81] Kirti N. Ghia Urmila Ghia and C. T. Shin. "High-Re solutions for incompressible flow using the Navier-Stokes equations and a multigrid method". In: *Journal of Computational Physics* 48 (1982), pp. 387–411.

Analyse et développement de schémas Volumes Finis asymptotiquement préservants dans la limite bas-Mach pour les équations d'Euler et de Navier-Stokes

Résumé

Dans cette thèse, nous développons et étudions des schémas asymptotiquement préservants (AP) pour les équations d'Euler complet et de Navier-Stokes dans le régime des faibles nombres de Mach. Pour les écoulements subsoniques, les ondes acoustiques sont très rapides par rapport à la vitesse du fluide. D'un point de vue numérique, lorsque le nombre de Mach tend vers zéro, les schémas classiques explicites présentent deux inconvénients majeurs : ils ne sont pas consistants à la limite et imposent une contrainte très restrictive sur le pas de temps pour garantir la stabilité du schéma, car ils doivent suivre les ondes acoustiques rapides. Nous proposons un nouveau schéma linéaire asymptotiquement préservant, avec une condition de type C.F.L. indépendante du nombre de Mach, et asymptotiquement consistent, c'est-à-dire qu'il donne une discrétisation du modèle limite lorsque le nombre de Mach est suffisamment petit. De plus, pour les équations de Navier-Stokes, nous choisissons une discrétisation implicite des termes de diffusion, ce qui nous permet d'utiliser de plus grands pas de temps dans les régimes fortement visqueux aussi. Ce type de schéma a été largement étudié dans la littérature, en particulier pour le cas isentropique, mais aussi pour le système d'Euler complet ou de Navier-Stokes avec différentes méthodes. Dans ce travail, nous proposons d'abord un schéma AP d'ordre 1 basé sur une discrétisation IMEX (Implicite-Explicite) en temps et volumes finis colocalisés en l'espace. Une extension d'ordre 2 est également proposée avec l'utilisation d'une procédure MOOD pour détecter et réduire les oscillations apparaissant classiquement avec les schémas d'ordre élevé.

Analysis and development of Finite Volume schemes asymptotically preserving in the low Mach number limit for the Euler and Navier-Stokes equations

Abstract

In this thesis, we develop and study asymptotic preserving (AP) schemes for the compressible Euler and Navier-Stokes equations in the low Mach number regime. For subsonic flows, the acoustic waves are very fast compared to the velocity of the fluid. From a numerical point of view, when the Mach number tends to zero, classical explicit schemes present two major drawbacks: they loose consistency in the limit and impose a very restrictive constraint on the time step to guaranty the stability of the scheme since they have to follow the fast acoustic waves. We propose a new linear asymptotic preserving scheme, with a C.F.L. condition independent of the Mach number, and asymptotically consistent, that is it degenerates into a consistent discretization of the limit model when the Mach number is sufficiently small. Moreover, for the Navier-Stokes equations we choose an implicit discretization of the diffusion terms such that we can use larger time steps in strongly viscous regimes as well. This type of schemes has been widely studied in the literature, in particular for the isentropic case but also for the full Euler or Navier-Stokes system with various methods. In this work, first we propose an AP scheme based on an IMEX (Implicit-Explicit) discretization in time and cell-centered finite volume in space. A low oscillating second order extension is also proposed with the use of a MOOD procedure to detect and reduce the oscillation classically appearing for high-order schemes.