THÈSE DE DOCTORAT DE L'ÉCOLE POLYTECHNIQUE

SPÉCIALITÉ : MATHÉMATIQUES APPLIQUÉES

PRÉSENTÉE PAR

Christophe CHALONS

POUR OBTENIR LE TITRE DE
DOCTEUR DE L'ÉCOLE POLYTECHNIQUE

Sujet de la thèse

BILANS D'ENTROPIE DISCRETS DANS L'APPROXIMATION NUMÉRIQUE DES CHOCS NON CLASSIQUES. APPLICATION AUX ÉQUATIONS DE NAVIER-STOKES MULTI-PRESSION 2D ET À QUELQUES SYSTÈMES VISCO-CAPILLAIRES

soutenue le 25 novembre 2002

devant le jury composé de :

M. Rémi ABGRALL, Rapporteur,
M. Grégoire ALLAIRE, Président du jury,
M. François BOUCHUT, Rapporteur,
M. Frédéric COQUEL,
M. Philippe LEFLOCH, Directeur de thèse,
M. Benoît PERTHAME

à Ludivine,

REMERCIEMENTS

L'aboutissement d'une recherche doctorale se fait rarement sans l'aide et le soutien de plusieurs personnes. Ce fut pour moi un plaisir et une source indéniable d'enrichissement de ne pas faire exception à cette règle. Je tiens à remercier les personnes qui ont été impliquées de près ou de loin dans cette formidable aventure.

Mes premiers remerciements sont adressés à la DGA. Elle a entièrement financé cette thèse.

P. Morice, J.P. Veuillot d'une part, et V. Giovangigli d'autre part m'ont acceuilli dans leur équipe, au DSNA de l'ONERA et au CMAP de l'Ecole Polytechnique. Les conditions de travail y ont été idéales, je les en remercie vivement.

J'ai bénéficié au cours de ma thèse d'un encadrement exceptionnel. P. LeFloch, F. Coquel et C. Marmignon m'ont guidé et soutenu sans cesse pendant ces trois années. Ils ont toujours répondu présents dans les moments difficiles et m'ont fait partager leur passion pour les EDP, l'analyse numérique et le calcul scientifique. Je resterai à jamais admiratif devant leur connaissance de ces domaines. Par ailleurs, leurs grandes qualités humaines sont à la base de la très bonne ambiance et du respect qui ont caractérisé notre collaboration pendant ces trois années. Je tiens à préciser qu'ils ont été (et sont toujours) à mes yeux bien plus que mes "chefs", pour reprendre le terme utilisé parmi les thésards. Je leur dois une grande partie de l'aboutissement de ce travail. Un grand merci pour tout.

Une quatrième personne a joué un rôle déterminant durant cette période. Ma femme Ludivine m'a toujours soutenu dans les moments difficiles que comporte une thèse. Sa patience, sa tolérance et ses encouragements m'ont été très précieux. Je lui en suis très reconnaissant.

Je tiens également à remercier vivement messieurs Abgrall et Bouchut d'avoir accepté de rédiger un rapport sur ce travail, messieurs Allaire et Perthame d'avoir accepté de participer à mon jury. C'est pour moi un honneur de les associer à mon travail.

Enfin, je tiens à saluer les doctorants, les post-doctorants et les ingénieurs du DSNA et du CMAP avec qui j'ai partagé (entre autres) de nombreux repas et pauses café. Grâce à eux, l'ambiance de travail à toujours été très agréable.

Table des matières

Première partie

Introduction

Le travail de recherche effectué au cours de cette thèse se décompose en deux parties principales. Il concerne l'analyse et l'approximation numérique du système des équations de Navier-Stokes à plusieurs pressions indépendantes (première partie), ainsi que de quelques systèmes d'équations aux dérivées partielles développant des chocs non classiques ou des transitions de phase (deuxième partie). Les systèmes considérés sont tous motivés par des applications physiques concrètes et les problèmes abordés s'inscrivent dans un véritable challenge numérique, non encore complètement résolu à ce jour, à savoir contrôler la dissipation (d'entropie) numérique associée à une discrétisation donnée. Nous considérons ici des discrétisations associées à des systèmes sous forme non conservative (première partie) et conservative (deuxième partie).

Première partie (chapitres II 1 à II 4 et III 1 à III 4)

Dans la première partie de ce manuscrit, nous nous intéressons à l'approximation numérique des solutions du système des équations de Navier-Stokes à plusieurs pressions indépendantes, encore appelées équations de Navier-Stokes à plusieurs entropies spécifiques. La somme des pressions partielles du système détermine la pression totale du fluide considéré. Des énergies internes et des viscosités sont naturellement associées à chacune des pressions du système. Avec des notations classiques, l'énergie totale du fluide s'écrit alors, en une dimension d'espace, sous la forme suivante :

$$\rho E = \frac{(\rho u)^2}{2\rho} + \sum_{i=1}^{N} \rho\varepsilon_i,$$

N représentant le nombre de pressions considérées.

Comme nous le constaterons dans les chapitres à venir, le système d'équations aux dérivées partielles étudié ne constitue dans les faits ni plus ni moins qu'une extension naturelle des équations de Navier-Stokes usuelles obtenues en considérant une seule loi de pression. Nous montrerons par exemple que sous des hypothèses de modélisation classiques, l'opérateur du premier ordre extrait du système considéré est hyperbolique, les champs caractéristiques associés aux valeurs propres correspondantes étant soit vraiment non linéaires, soit linéairement dégénérés. Ces propriétés sont également partagées par le système des équations d'Euler (opérateur du premier ordre extrait des équations de Navier-Stokes usuelles).

En regard de ces propriétés agréables et classiques, une grande différence oppose le cadre usuel au cadre qui nous préoccupe : les systèmes considérés ici s'écrivent naturellement sous forme non conservative et n'admettent pas, sauf hypothèse de modélisation restrictive, de formulation équivalente sous forme complètement conservative. Autrement dit, tout changement de variables admissible transforme le système initial en un système équivalent (pour les solutions régulières) dans lequel subsistent toujours des produits non conservatifs (PNC) impliquant l'inconnue et ses dérivées premières (PNC du premier ordre) et/ou l'inconnue et ses dérivées secondes (PNC du second ordre). Les systèmes en considération trouvent, en une dimension d'espace, la forme condensée suivante :

$$\partial_t \mathbf{u} + \mathcal{A}(\mathbf{u})\partial_x \mathbf{u} = \partial_x(\mathcal{D}(\mathbf{u})\partial_x \mathbf{u}), \quad \mathbf{u}(x,t) \in I\!R^P, \quad x \in I\!R, \quad t > 0, \tag{1}$$

où $\mathbf{u}$ représente le vecteur des inconnues et $\mathcal{A}$ une matrice non jacobienne de $I\!R^{P\times P}$, c'est-à-dire telle qu'il n'existe pas de fonction flux $\mathbf{f} : I\!R^P \to I\!R^P$ satisfaisant $\mathcal{A} = D\mathbf{f}$. Notons que par souci de simplicité, l'ensemble des produits non conservatifs du système (1) a été concentré sur l'opérateur du premier ordre, le membre de droite étant de ce fait sous forme complètement conservative. Nous adopterons cette écriture dans la présentation générale par souci de simplicité.

Les systèmes considérés dans cette première partie ont un intérêt de toute première importance puisqu'ils interviennent dans de nombreuses applications physiques :
- en turbulence compressible par exemple, les modèles à deux équations de transport du type $k-\varepsilon$, $k-l$, $k-\omega$ (description du champ turbulent à l'aide d'une échelle unique, voir Berthon [10]) ou les modèles à échelles multiples (description de la turbulence à l'aide de plusieurs échelles, voir ci-après) rentrent dans le formalisme proposé. Dans ces modèles, la pression laminaire est en effet distinguée de la (des) pression(s) turbulente(s).
- en physique des plasmas également, les modèles décrivant les écoulements de gaz réactifs faiblement ionisés en déséquilibre thermochimique (voir Coquel et Marmignon [34], [35], [36]) rentrent dans le cadre de travail considéré. Dans ce contexte, la température des électrons est distincte de celle des autres espèces. Nous sommes donc amenés à considérer plusieurs pressions indépendantes, la pression du gaz d'électrons et la pression des espèces dites lourdes.
- en physique des mélanges de fluides complexes, les modèles décrits par Truesdell [137] ainsi que ceux récemment introduits par Lagoutière [79] s'inscrivent naturellement dans le présent contexte.

Le chapitre III 4 propose d'appliquer les considérations théoriques et numériques développées dans la première partie de ce manuscrit aux modèles multi-échelle de turbulence compressible homogène isotrope.

Précisons dès à présent que l'hypothèse de grands nombres de Reynolds sous-jacente à une telle modélisation rend la dynamique du système (1) principalement dictée par son opérateur convectif (opérateur du premier ordre extrait). L'approximation numérique des solutions du système (1) consiste donc, pour un niveau de raffinement réaliste et loin de toute paroi, à approcher numériquement les solutions du système hyperbolique non conservatif extrait, à savoir :

$$\partial_t \mathbf{u} + \mathcal{A}(\mathbf{u})\partial_x \mathbf{u} = 0, \quad \mathbf{u}(x,t) \in I\!R^P, \quad x \in I\!R, \quad t > 0. \tag{2}$$

La problématique des systèmes hyperboliques sous forme non conservative a fait l'objet de nombreuses études, tant du point de vue théorique (définition du produit non conservatif $\mathcal{A}(\mathbf{u})\partial_x \mathbf{u}$) que du point de vue numérique. Nous proposons de rappeler brièvement les principales motivations de ces études, afin d'introduire le plus naturellement possible les travaux effectués dans la première partie de ce manuscrit.

Systèmes hyperboliques sous forme non conservative

• Du point de vue théorique, le système (2) étant hyperbolique, ses solutions développent des discontinuités en temps finis (voir Lax [84]). Il est indispensable d'entendre

le système (2) en un sens affaibli permettant de définir une notion de solution discontinue. La validité d'une condition d'entropie est nécessaire pour assurer l'unicité des solutions ainsi définies.
Lorsque le système (2) est un système de lois de conservation, c'est-à-dire lorsqu'il existe une fonction flux $\mathbf{f} : I\!R^P \to I\!R^P$ telle que $\mathcal{A} = D\mathbf{f}$, la notion de solution faible au sens des distributions est utilisée pour définir les solutions du système (2). Des résultats d'existence et d'unicité ont alors été obtenus pour des données initiales à variation totale bornée et suffisamment proches d'un état constant (voir Liu [100], [103], Glimm [55], Lax [83], [84]). Nous renvoyons à l'ouvrage de LeFloch [88] pour un exposé de ces résultats ainsi que des extensions.
Lorsque la matrice $\mathcal{A}$ n'est pas la jacobienne d'un flux, la théorie des distributions est inopérante puisque le produit non conservatif $\mathcal{A}(\mathbf{u})\partial_x\mathbf{u}$ n'a pas de sens classique au niveau d'une discontinuité de la solution $\mathbf{u}$. Il s'agit de calculer le produit d'une fonction de type Heavyside (fonction créneau) et d'une masse de Dirac (dérivée au sens des distributions d'une fonction créneau) évaluées sur la discontinuité. Les singularités étant concentrées au même point (point de la discontinuité), ce produit n'a pas de définition unique au sens des distributions. Pour surmonter cette difficulté, Dal Maso, LeFloch et Murat [42] ont introduit (pour les fonctions à variation totale bornée) une notion de solution faible pour les systèmes hyperboliques sous forme non conservative. Elle permet de donner une définition au produit non conservatif $\mathcal{A}(\mathbf{u})\partial_x\mathbf{u}$ et généralise la notion de solution faible au sens des distributions pour les systèmes conservatifs. De manière plus précise, la définition de ce produit non conservatif est basée sur la donnée d'une famille de chemins $\phi : [0,1] \times I\!R^P \times I\!R^P \to I\!R^P$ satisfaisant la propriété suivante :

$$\phi(0,\mathbf{u},\mathbf{v}) = \mathbf{u}, \quad \phi(1,\mathbf{u},\mathbf{v}) = \mathbf{v}, \quad \text{pour tout } (\mathbf{u},\mathbf{v}) \in I\!R^P \times I\!R^P.$$

Sous des hypothèses de consistance et de régularité que nous ne détaillerons pas ici (voir [42]), Dal Maso, LeFloch et Murat définissent alors dans l'espace des mesures de Borel le produit non conservatif $\mathcal{A}(\mathbf{u})\partial_x\mathbf{u}$ au point x_0 d'une discontinuité séparant deux états constants $\mathbf{u}_0$ et $\mathbf{u}_1$ en posant :

$$[\mathcal{A}(\mathbf{u})\partial_x\mathbf{u}]_\phi = \int_0^1 \mathcal{A}(\phi(s,\mathbf{u}_0,\mathbf{u}_1))\frac{\partial\phi}{\partial s}(s,\mathbf{u}_0,\mathbf{u}_1)ds\ \delta_{x_0}. \tag{3}$$

On montre sans aucune difficulté que cette définition du produit $\mathcal{A}(\mathbf{u})\partial_x\mathbf{u}$ est, dans le cas d'un produit conservatif (c'est-à-dire lorsque $\mathcal{A} = D\mathbf{f}$ pour une fonction flux $\mathbf{f}$), indépendante de ϕ et consistante avec le cadre usuel des distributions. Par ailleurs et à titre d'illustration, notons que la définition (3) coïncide avec celle introduite par Volpert [145] lorsque ϕ correspond à la famille des lignes droites :

$$\phi(s,\mathbf{u},\mathbf{v}) = \mathbf{u} + s(\mathbf{v}-\mathbf{u}), \quad \forall\ \mathbf{u},\mathbf{v} \in I\!R^P, \quad \forall\ s \in [0,1].$$

De même que dans le cadre conservatif, la définition (3) conduit à des résultats d'existence et d'unicité des solutions faibles du système (2) (voir Dal Maso, LeFloch et Murat [42], Liu [92], LeFloch et Tzavaras [97], LeFloch [91]).

La définition (3) du produit non conservatif $\mathcal{A}(\mathbf{u})\partial_x\mathbf{u}$ dépend fortement de la famille de chemins ϕ considérée et supposée connue. Une question naturelle concerne l'obtention d'une telle famille de chemins. Selon LeFloch [90] et Sainsaulieu [115], la définition de ϕ repose sur l'information, non encore utilisée, contenue dans le système convectif-diffusif (1) dont dérive (2). Ainsi, une solution choc du système (2), caractérisée avec des notations classiques par le triplet $(\sigma, \mathbf{u}_0, \mathbf{u}_1)$, est définie comme la limite lorsque les termes diffusifs sont négligés d'une solution onde progressive du système (1), c'est-à-dire une solution régulière vérifiant :

$$\begin{aligned} &\mathbf{u}(x,t) = \overline{\mathbf{u}}(\xi), \quad \xi = x - \sigma t, \\ &\lim_{\xi \to -\infty} \overline{\mathbf{u}}(\xi) = \mathbf{u}_0, \quad \lim_{\xi \to +\infty} \overline{\mathbf{u}}(\xi) = \mathbf{u}_1. \end{aligned} \tag{4}$$

Dans [90], LeFloch montre comment obtenir une famille de chemins ϕ consistante avec une telle définition.

Nous adopterons cette définition du produit non conservatif $\mathcal{A}(\mathbf{u})\partial_x \mathbf{u}$ dans la première partie de ce manuscrit, et renvoyons par exemple le lecteur à Colombeau et Leroux [31] pour une autre théorie sur les produits non conservatifs.

Pour conclure, il est important de noter que contrairement au cadre conservatif usuel, la définition adoptée ci-dessus du produit non conservatif $\mathcal{A}(\mathbf{u})\partial_x \mathbf{u}$ dépend de la forme du tenseur de dissipation du système (1). Les solutions onde progressive, et plus précisément les états de sortie $\mathbf{u}_1 = \mathbf{u}_1(\sigma, \mathbf{u}_0)$ associés à de telles solutions, dépendent en effet de la forme du tenseur diffusif du système en considération, comme l'illustrent de nombreuses contributions sur ce sujet : Berthon et Coquel [11], Raviart et Sainsaulieu [112], Sainsaulieu [115], Karni [76] .... Cette propriété est une caractéristique des systèmes hyperboliques sous forme non conservative. Nous la retrouverons dans un autre contexte, celui des systèmes conservatifs de type hyperbolique dont les champs ne sont pas vraiment non linéaires, ou de type hyperbolique-elliptique. Ces modèles décrivent par exemple la dynamique des transitions de phase (voir deuxième partie du manuscrit pour plus de précisions). La principale difficulté dans l'approximation numérique des solutions discontinues du système (2) consistera donc à respecter le plus fidèlement possible la forme du tenseur de dissipation intervenant dans (1), afin d'obtenir les solutions onde de choc associées. Dans le cas contraire, on s'expose à la capture de solutions arbitrairement éloignées de la solution attendue, comme l'illustre la figure 1.

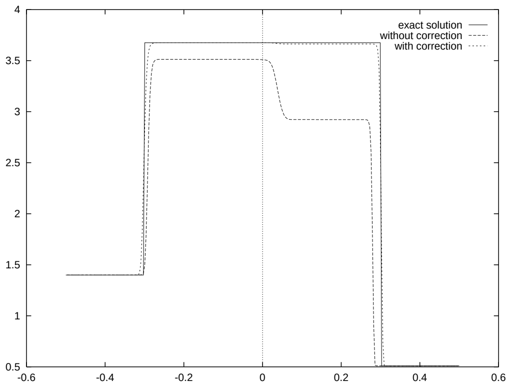

FIG. 1 – *Pression partielle p_1 obtenue avec deux schémas distincts*

Le cas test correspondant à cette simulation est détaillé dans le chapitre II 2.

• Du point de vue numérique, l'approximation des solutions d'un système hyperbolique non conservatif du type (2) reste un problème essentiellement ouvert à ce jour. La principale raison provient de l'absence de résultats généraux de convergence pour les schémas numériques associés. En effet, les méthodes numériques adaptées à de tels systèmes étant sous forme non conservative, le cadre général donné par le théorème de Lax-Wendroff est inopérant. En d'autres termes, nous ne sommes plus en mesure de garantir l'obtention d'une solution discrète en accord avec la solution attendue par raffinement de maillage, même pour une méthode numérique convergente vers une certaine fonction limite. Cette caractéristique des schémas numériques sous forme non conservative s'inscrit dans un cadre plus général que celui qui nous préoccupe ici, puisque même lorsque le système (2) admet une formulation équivalente conservative, de tels schémas s'avèrent être inadaptés. A titre d'exemple, Hou et LeFloch [69] ont montré dans le cas d'une loi de conservation scalaire, que les schémas non conservatifs contenant de la viscosité numérique (Lax-Friedrichs, Godunov, ...) ne convergeait pas, en général, vers la solution attendue du problème considéré. Plus récemment, Berthon [10] a mis en évidence les mêmes observations en approchant par la méthode de Godunov les équations de Navier-Stokes usuelles sous forme non conservative (*i.e.* en privilégiant par exemple la loi d'évolution sur l'entropie spécifique au détriment de la loi de conservation sur l'énergie totale, celle-ci jouant alors le rôle d'entropie de Lax du système). Dans ce cas encore, la solution exacte attendue et la solution numérique ne coïncident pas, et ce en dépit de propriétés d'équivalence au niveau continu des solutions des systèmes considérés. Par ailleurs, l'erreur numérique entre ces deux solutions augmente avec le temps (en un sens à préciser, voir Hou et LeFloch [69], Berthon [10]).

Il ressort de ces études que l'erreur entre solution exacte et solution numérique peut être associée à la dissipation (d'entropie) du schéma utilisé. Dit autrement, cette erreur est en étroite relation avec l'opération de projection inhérente au schéma utilisé (projection L^2 pour la méthode de Godunov par exemple). Ces observations sont corroborées par les travaux de LeFloch et Liu [92] : ils montrent la convergence de la méthode de Glimm [55], **exempte de toute étape de projection**, lorsqu'elle est appliquée à un système sous forme non conservative, celui-ci admettant ou non une formulation équivalente conservative (pourvu que les solutions coïncident au niveau continu).
Dans le cadre conservatif ($\mathcal{A} = D\mathbf{f}$), plusieurs études ont néammoins permis d'obtenir, après correction, des schémas numériques sous forme non conservative qui convergent vers la solution attendue. Hou et Lefloch [69] obtiennent de tels schémas dans le contexte d'une loi de conservation scalaire en corrigeant localement un schéma "vraiment" non conservatif à l'aide d'un schéma conservatif. Le schéma obtenu est encore sous forme non conservative et converge effectivement vers la solution physique par raffinement de maillage. Pour les systèmes, Di Perna [45] a prouvé la convergence d'un schéma hybride Lax-Friedrichs - Glimm non conservatif vers la solution attendue.
Dans le cadre vraiment non conservatif ($\mathcal{A} \neq D\mathbf{f}$), les études menées dans Berthon [10] sont celles qui se rapprochent le plus de la problématique dictée par la présente recherche doctorale. Elles concernent l'approximation des solutions des équations de Navier-Stokes à deux pressions indépendantes, qui est un cas particulier ($N = 2$) du système qui nous préoccupe ici ($N \geq 2$). Berthon propose de corriger les méthodes d'approximation classiques de *"shock-capturing"* (méthode de Godunov, méthode de

Roe, ...) en leur ajoutant une étape de **projection non linéaire** à la fin de chaque itération en temps. Les résultats numériques obtenus par une telle stratégie en attestent le bien-fondé (voir Berthon [10] pour plus de détails).

Les schémas numériques développés dans la première partie de ce manuscrit pour approcher numériquement les solutions des équations de Navier-Stokes multi-pression peuvent être vus comme une extension au cas $N \geq 2$ des études menées dans Berthon [10], en proposant de nouvelles techniques de correction complètement linéaires. La recherche de nouvelles procédures a été motivée par l'extension de la problématique à plusieurs dimensions d'espace. Ces travaux sont à la base d'une technique d'implicitation que nous illustrerons sur des problèmes véritablement bi-dimensionnels.

Soulignons que dans notre contexte, la connaissance du problème de Riemann associé à (2) (voir chapitre II 1) aurait pu nous inciter à approcher numériquement les solutions de ce système par la méthode de Glimm. Cette méthode, exempte de toute étape de projection, suffirait à un tel propos. Nous n'avons volontairement pas adopté une telle stratégie, afin notamment de faciliter l'extension de l'approche à des considérations multi-dimensionnelles.

Travaux effectués

Nous débutons la première partie de ce manuscrit en établissant l'existence et l'unicité d'une solution du problème de Riemann pour l'opérateur du premier ordre (2) associé au système des équations de Navier-Stokes à N pressions indépendantes ($N \geq 2$). Une étude des solutions onde progressive de ce système est nécessaire. Ces solutions caractérisent les chocs admissibles du système hyperbolique sous-jacent et permettent de définir le produit non conservatif $\mathcal{A}(\mathbf{u})\partial_x\mathbf{u}$ selon (3), la famille de chemins $\phi(., \mathbf{u}_0, \mathbf{u}_1)$ coïncidant avec le profil de l'onde progressive joignant les points critiques $\mathbf{u}_0$ et $\mathbf{u}_1$ et se déplaçant à la vitesse σ (voir (4)).

Dans le chapitre II 1, nous montrons l'existence et l'unicité (à une translation près) des solutions onde progressive du système pour des lois de pression de gaz parfaits polytropiques et dans le cas de viscosités constantes. De manière attendue, l'état de sortie $\mathbf{u}_1 = \mathbf{u}_1(\sigma, \mathbf{u}_0)$ dépend de la forme du tenseur de dissipation considéré, par l'intermédiaire du rapport des viscosités $\{\mu_i/\mu_N\}_{i=1,\dots,N-1}$. Nous montrons également que l'état de sortie $\mathbf{u}_1(\sigma, \mathbf{u}_0)$ peut être paramétré par sa densité ρ_1, celle-ci étant solution d'un problème algébrique scalaire. Dès lors, σ et $\mathbf{u}_0$ étant fixés, l'obtention de $\mathbf{u}_1$ se fait par l'implémentation d'un simple algorithme de Newton et ne nécessite plus l'intégration numérique des EDO gouvernant les solutions ondes progressives. Ceci constitue un résultat très agréable dans la pratique, et nouveau dans le contexte des systèmes sous forme non conservative. La propriété de monotonie démontrée du taux de compression ρ_1/ρ_0 par rapport au nombre de Mach relatif permet alors d'établir l'existence et l'unicité (loin du vide) d'une solution du problème de Riemann associé au système hyperbolique non conservatif considéré.

Les chapitres suivants concernent l'approximation numérique des solutions onde progressive.

Le chapitre II 2 propose une extension directe au cas $N \geq 2$ des travaux menés dans Berthon [10] ($N = 2$). Ainsi, après avoir mis en évidence et justifié les erreurs numériques inacceptables générées par des méthodes de type Godunov, nous proposons de corriger ces méthodes classiques à l'aide d'une étape de projection **non linéaire** permettant de forcer au niveau discret la validité des relations de proportionnalité existantes au niveau continu entre les entropies du système. De manière plus précise et avec des notations classiques, ces relations trouvent la forme suivante :

$$\mu_N T_i \{\partial_t \rho S_i + \partial_x \rho S_i u\} = \mu_i T_N \{\partial_t \rho S_N + \partial_x \rho S_N u\}, \quad i = 1, ..., N-1. \tag{5}$$

En particularisant ces relations à des solutions onde progressive, nous obtenons des relations de saut généralisées traduisant de manière exacte la dépendance vis à vis du rapport des viscosités des solutions onde de choc associées. Leur validité discrète est donc cruciale. Dans (5), l'utilisation d'entropies S_i assurant des propriétés de convexité permet d'obtenir des propriétés de stabilité (principes du maximum, positivité des pressions, inégalités d'entropie, ...) à l'issue de l'étape de correction. Celle-ci consiste en la résolution, à chaque pas de temps, d'un problème algébrique **non linéaire** de taille $(N-1)$, ce qui dans la pratique peut s'avérer coûteux.

Les chapitres II 3 et II 4 proposent d'utiliser respectivement des entropies non convexes et les énergies internes partielles pour respecter la notion de proportionnalité entre chacun des bilans d'entropie. Il s'agit alors, avec des notations claires, d'imposer au niveau discret les relations suivantes, équivalentes à (5) pour les solutions régulières :

$$\mu_N \frac{\rho^{\gamma_i - 1}}{\gamma_i - 1} \{\partial_t \rho s_i + \partial_x \rho s_i u\} = \mu_i \frac{\rho^{\gamma_N - 1}}{\gamma_N - 1} \{\partial_t \rho s_N + \partial_x \rho s_N u\}, \quad i = 1, ..., N-1, \tag{6}$$

avec $s_i = p_i / \rho^{\gamma_i}$ pour tout $i = 1, ..., N$, et :

$$\mu_N \{\partial_t \rho \varepsilon_i + \partial_x \rho \varepsilon_i u + p_i \partial_x u\} = \mu_i \{\partial_t \rho \varepsilon_N + \partial_x \rho \varepsilon_N u + p_N \partial_x u\}, \quad i = 1, ..., N-1. \tag{7}$$

Les algorithmes proposés sont basés sur l'utilisation de solveurs de Riemann approchés appropriés et sont justifiés théoriquement par la satisfaction de propriétés de stabilité identiques à celles obtenues dans le chapitre II 2. Le bénéfice est néammoins significatif : l'utilisation de ces nouvelles variables conduit l'étape de correction à être **explicite**. Elle ne nécessite plus la résolution numérique d'un problème algébrique non linéaire. Outre le gain de temps calcul, l'existence de ces algorithmes explicites (astreints à une restriction CFL classique) permet d'envisager une formulation implicite en temps.

Les chapitres suivants font l'objet d'une introduction plus détaillée en début de partie III.

Le chapitre III 1 décrit une méthode de type volumes finis par Relaxation pour résoudre le système des équations d'Euler multi-pression intervenant dans les algorithmes proposés. Cette méthode numérique vérifie les propriétés de stabilité citées précédemment. Par ailleurs, son utilisation s'avère être particulièrement bien adaptée à notre propos, puisqu'elle autorise un découplage naturel des entropies du système relaxé. Cette caractéristique trouvera toute son importance dans les chapitres suivants. Les chapitres III 2 et III 3 proposent une implicitation des schémas numériques de type prédicteur-correcteur proposés. Il s'agit d'impliciter la méthode de Relaxation décrite dans le chapitre précédent d'une part, et les étapes de correction correspondantes d'autre part. Tirant profit du découplage des entropies du système relaxé, nous proposons un algorithme implicite extrêmement peu coûteux. La stabilité a été testée sur des cas tests bi-dimensionnels (écoulement supersonique à Mach 10 autour d'un cylindre par exemple).

Enfin, le chapitre III 4 propose, en utilisant une méthode de pas fractionnaires, d'appliquer ces considérations aux modèles multi-échelle de turbulence compressible homogène isotrope. Les termes de conductivité thermique ainsi que les termes sources sont pris en compte. Une suite naturelle aux présents travaux consisterait à étendre le traitement des contributions thermiques à des problèmes à plusieurs dimensions d'espace.

Deuxième partie (chapitres IV 1 à IV 3)

Dans la deuxième partie de ce manuscrit, nous nous intéressons à l'approximation numérique des solutions discontinues de systèmes non linéaires de lois de conservation. Par opposition aux équations de Navier-Stokes à plusieurs pressions indépendantes étudiées dans la première partie, les systèmes considérés ici s'écriront donc naturellement sous forme conservative.

De manière plus précise, nous considérons la problématique des systèmes hyperboliques de lois de conservation admettant un ou plusieurs champs caractéristiques non vraiment non linéaires (non VNL) et non linéairement dégénérés (non LD), mais aussi celle des systèmes de lois de conservation mixtes de type hyperbolique-elliptique.

L'absence d'une propriété du type "système hyperbolique dont tous les champs sont soit VNL, soit LD" se traduit par l'importance prise par les petites échelles introduites par des opérateurs classiques de régularisation (viscosité, capillarité) dans la dynamique du système, et ce même à un niveau de description purement convectif. Concrètement, les discontinuités de type choc (ou transitions de phase) du système hyperbolique (ou hyperbolique-elliptique) sous-jacent dépendront intrinsèquement de l'opérateur de régularisation choisi. On peut dès lors faire le lien avec la première partie de ce manuscrit consacrée aux équations de Navier-Stokes à plusieurs pressions indépendantes. Ces dernières équations étant sous forme non conservative, les ondes de choc associées sont sensibles à la régularisation considérée (voir ci-dessus). Au niveau numérique, la première et la deuxième partie ont donc en commun une difficulté majeure, à savoir être capable de restituer des discontinuités de type choc (ou transitions de phase) en accord avec la forme du tenseur de dissipation considéré.

Un grand nombre de modèles répondant à des applications physiques concrètes rentrent dans le cadre de travail proposé dans cette deuxième partie :
- en élastodynamique non linéaire par exemple, de nombreux matériaux sont gouvernés par des lois d'état non convexes, de telle sorte que le système hyperbolique de lois de conservation correspondant n'est plus VNL. Dans ce cas, cela signifie que les matériaux peuvent exister sous différentes phases (chacune étant associée à une région de convexité de la loi d'état) correspondant à différents arrangements entre les atomes.
- en dynamique des fluides compressibles également, les changements de phase liquide-vapeur sont modélisés par des lois de pression généralement non convexes, comme par exemple celle introduite par van der Waals. Les modèles correspondants sont soit hyperboliques non VNL (la pression est monotone mais non convexe), soit hyperbolique-elliptique (la pression n'est plus monotone).

Dans toutes ces situations (et dans bien d'autres encore), les termes de diffusion et de dispersion (viscosité et capillarité) jouent un rôle fondamental dans la dynamique des discontinuités. Mentionnons que cette dépendance est connue et a été étudiée à plusieurs reprises : Abeyaratne et Knowles [1], [2], LeFloch [87], Hayes et LeFloch [69], [64], [65], Truskinovsky [138], [139], Schulze et Shearer [117], Slemrod [123], ... De nombreuses autres études ont été menées et à ce titre nous renvoyons par exemple le lecteur à l'ouvrage [88].

Afin de préciser la contribution apportée à ce sujet au cours de la thèse, nous présentons dans un premier temps le cadre général dans lequel se situe notre étude. Nous spécifierons ensuite les principaux travaux effectués.

Problématique générale : rappels

Considérons un système de lois de conservation de la forme

$$\partial_t \mathbf{u} + \partial_x \mathbf{f}(\mathbf{u}) = 0, \quad \mathbf{u}(x,t) \in I\!R^N, \quad x \in I\!R, \quad t > 0, \tag{8}$$

où, de manière classique, $\mathbf{u}$ représente le vecteur des inconnues conservatives et f : $I\!R^N \to I\!R^N$ une application régulière appelée fonction flux. Nous supposerons pour fixer les idées que le système (8) est strictement hyperbolique, c'est-à-dire que la matrice jacobienne $D\mathbf{f}(\mathbf{u})$ est diagonalisable dans $I\!R$ et admet N valeurs propres réelles distinctes $\lambda_1(\mathbf{u}) < ... < \lambda_N(\mathbf{u})$.

Nous rappelons que les solutions du système (8) sont généralement discontinues et que par ailleurs, elles ne sont pas uniquement déterminées par la connaissance de la donnée initiale associée au système (8). Dans ce contexte, il est classique d'ajouter dans le membre de droite de (8) un opérateur de régularisation $\mathcal{R}$, piloté par un petit paramètre ϵ représentant la taille des petites échelles introduites et destiné à tendre vers zéro. Nous obtenons le système suivant :

$$\partial_t \mathbf{u}^\epsilon + \partial_x \mathbf{f}(\mathbf{u}^\epsilon) = \partial_x \mathcal{R}(\epsilon \partial_x \mathbf{u}^\epsilon, \epsilon^2 \partial_{xx} \mathbf{u}^\epsilon, ...), \quad \mathbf{u}(x,t) \in I\!R^N, \quad x \in I\!R, \quad t > 0, \tag{9}$$

permettant de régulariser les solutions du système (8), et dans la limite $\epsilon \to 0$ de sélectionner la solution physique de ce système.

Soit (U, F) un couple d'applications régulières de $I\!R^N \to I\!R$ (appelées entropie et flux d'entropie) satisfaisant la relation de compatibilité $DF = DUD\mathbf{f}$, et tel que l'application U soit strictement convexe. Sous des hypothèses classiques sur $\mathcal{R}$ ($\mathcal{R}$ est dissipative pour U), la solution limite $\mathbf{u} = \lim_{\epsilon \to \mathbf{0}} \mathbf{u}^\epsilon$ obtenue est une solution faible (au sens des distributions) du système (8), compatible avec la validité de l'inégalité d'entropie suivante :

$$\partial_t U(\mathbf{u}) + \partial_x F(\mathbf{u}) \leq 0. \tag{10}$$

• Lorsque tous les champs du système (8) sont VNL (ce qui correspond à une propriété de type convexité sur le flux $\mathbf{f}$) ou LD, les solutions faibles du problème de Riemann (8) pour la donnée initiale

$$\mathbf{u}(x,0) = \begin{cases} \mathbf{u}_L & \text{si} \quad x < 0, \\ \mathbf{u}_R & \text{si} \quad x > 0, \end{cases} \tag{11}$$

avec $\mathbf{u}_L$ et $\mathbf{u}_R \in I\!R^N$, sont sélectionnées de manière **unique** par la validité de l'inégalité d'entropie (10), au moins pour des valeurs de $\mathbf{u}_L$ et $\mathbf{u}_R$ suffisamment proches l'une de l'autre, c'est-à-dire telles que $|\mathbf{u}_R - \mathbf{u}_L| < C$ où C est une constante à préciser (voir LeFloch [88] pour plus de détails). En d'autres termes, l'inégalité d'entropie (10) sélectionne de manière unique la solution physique de (8). Notons que cette solution coïncide avec la solution **classique** proposée par Lax [83], [84], et basée sur le critère de sélection suivant :

un j-choc connectant deux états $\mathbf{u}_0$ et $\mathbf{u}_1$ est admissible si et seulement si

$$\lambda_j(\mathbf{u}_0) \geq \sigma \geq \lambda_j(\mathbf{u}_1), \tag{12}$$

où $\sigma = \sigma(\mathbf{u}_0, \mathbf{u}_1)$ représente la vitesse de propagation du choc. Un choc **classique** au sens de Lax est donc **compressif** en ce sens que les caractéristiques associées aux deux états de la discontinuité rentrent dans le choc (Figure 2).

On montre par ailleurs que la solution **classique** obtenue par le principe de sélection

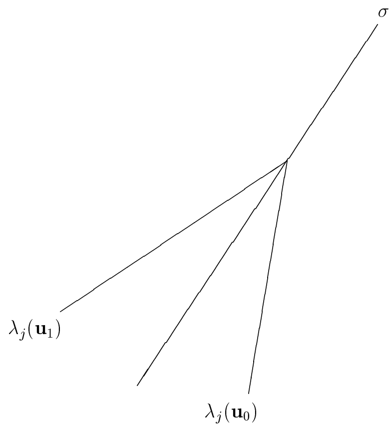

FIG. 2 – *Choc compressif*

(10), ou de manière équivalente par le principe de sélection (12), satisfait également le critère de sélection des solutions faibles introduit par Liu [102] :

un j-choc connectant deux états $\mathbf{u}_0$ et $\mathbf{u}_1$ est admissible si et seulement si

$$\sigma(\mathbf{u}_0, \mathbf{u}) \geq \sigma(\mathbf{u}_0, \mathbf{u}_1), \tag{13}$$

pour tout $\mathbf{u}$ appartenant à la courbe de Hugoniot (issue de $\mathbf{u}_0$) et situé entre $\mathbf{u}_0$ et $\mathbf{u}_1$.

Les critères de sélection des solutions faibles par l'inégalité d'entropie (10), par l'inégalité de Lax (12) ou par les inégalités de Liu (13) sont donc équivalents, lorsque tous les champs caractéristiques du système (8) sont vraiment non linéaires ou linéairement dégénérés. Notons que l'on retrouve ce résultat lorsque (8) n'est constitué que d'une seule équation scalaire ($N = 1$) avec une fonction flux convexe (le champ caractéristique de l'équation (8) est alors VNL). Dans ce cas, une fois encore, l'inégalité d'entropie (10) est suffisante pour sélectionner une unique solution faible du problème de Riemann (8)-(11). Cette solution faible peut être également et de manière équivalente caractérisée par la validité des inégalités de Lax, qui s'écrivent ici, à la traversée d'une discontinuité séparant deux états constants u_0 et u_1 :

$$f'(u_0) \geq \sigma \geq f'(u_1), \tag{14}$$

(avec $\sigma = \sigma(u_0, u_1) = (f(u_1) - f(u_0))/(u_1 - u_0)$ d'après les relations de saut de Rankine-Hugoniot), ou par la validité des inégalités d'Oleinik :

$$\frac{f(u) - f(u_0)}{u - u_0} \geq \frac{f(u_1) - f(u_0)}{u_1 - u_0}, \quad \forall\, u \in [\min(u_0, u_1), \max(u_0, u_1)], \tag{15}$$

dont le critère de Liu (13) constitue une généralisation au cas des systèmes ($N > 1$).

• Lorsque l'un au moins des champs caractéristiques du système (8) n'est ni VNL, ni LD (situation qui nous intéresse ici), l'inégalité d'entropie (10) n'est plus suffisante pour sélectionner les solutions faibles. En conséquence, même pour des états $\mathbf{u}_L$ et $\mathbf{u}_R$ suffisamment proches (au sens $|\mathbf{u}_R - \mathbf{u}_L| < C$), le problème de Riemann (8)-(11) complété par l'inégalité d'entropie (10) n'admet pas une unique solution. Afin de surmonter cette difficulté, Lax [83], [84] propose de sélectionner les solutions du problème de Riemann (8)-(11) par la validité des inégalités (12). On remarque que ce critère de sélection est indépendant de toute constante physique éventuellement présente dans le système physique dont dérive (8). En conséquence, la solution faible sélectionnée ne peut (sauf cas particulier) coïncider avec celle obtenue par une régularisation du type (9). Les paramètres de viscosité et de capillarité étant déterminants dans de nombreuses applications physiques, nous proposons, afin d'en garder le souvenir, de définir les solutions onde de choc du système (8) comme limites d'une solution onde progressive du système régularisé (9), lorsque les petites échelles sont négligées. Une discontinuité séparant deux états constants $\mathbf{u}_0$ et $\mathbf{u}_1$ et se déplaçant à la vitesse σ sera donc admissible si et seulement s'il existe une application régulière

$$\mathbf{u}(x,t) = \overline{\mathbf{u}}(\xi), \quad \xi = x - \sigma t,$$

solution du système (9), et telle que

$$\lim_{\xi \to -\infty} \overline{\mathbf{u}}(\xi) = \mathbf{u}_0, \quad \lim_{\xi \to +\infty} \overline{\mathbf{u}}(\xi) = \mathbf{u}_1.$$

Dans le contexte qui nous intéresse ici, l'existence et l'unicité des solutions onde progressive du système (9) ont été largement étudiées ces dernières années, et ce pour plusieurs opérateurs de régularisation $\mathcal{R}$: nous renvoyons par exemple le lecteur à Bona et Schonbek [16], Jacobs, McKinney et Shearer [71], Hayes et LeFloch [69], Bedjaoui et LeFloch [7], ..., pour une équation scalaire ($N = 1$) et à Schulze et Shearer [117], Slemrod [123], Abeyaratne et Knowles [1], [2], Truskinovsky [138], [139], LeFloch [86], Bedjaoui et LeFloch [8], [9], ..., pour un système de lois de conservation (typiquement $N = 2$).
L'analyse onde progressive montre que les solutions onde de choc ainsi définies sont compatibles avec l'inégalité d'entropie (10) mais ne vérifient pas, en général, les inégalités de Lax (12). Dans ce cas, de tels chocs sont dits **non classiques** et sont par définition "**sous-compressifs**", puisque les caractéristiques associées aux états gauche et droit du choc considéré se déplaçant à la vitesse σ le traversent (Figure 3).
Rappelons que l'approximation numérique des chocs non classiques est difficile puis-

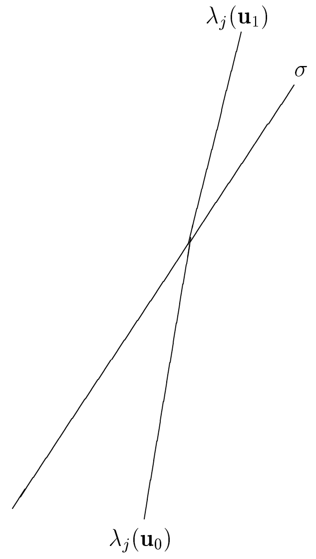

FIG. 3 – *Choc sous-compressif*

qu'ils dépendent des mécanismes de régularisation et que la forme de l'opérateur de régularisation considéré doit être respectée. Soulignons également que l'inégalité d'entropie (10) ne caractérisant pas la solution recherchée, forcer sa validité au niveau discret n'est plus suffisant pour justifier le bien-fondé d'une méthode numérique donnée. Contrairement au cadre usuel de champs VNL (ou LD), le théorème de Lax-Wendroff n'apparaît donc plus totalement adapté à notre propos. Nous verrons aux chapitres IV 1 et IV 2 comment respecter au mieux ces contraintes sur des exemples précis.
Le critère de sélection par analyse onde progressive introduit ci-dessus sélectionne les solutions onde de choc non classiques admissibles, non plus par la validité de l'inégalité d'entropie (10) mais plutôt par la validité d'une égalité d'entropie du type :

$$\partial_t U(\mathbf{u}) + \partial_x F(\mathbf{u}) = \mathcal{M}, \tag{16}$$

où $\mathcal{M}$ désigne une mesure de Borel négative et bornée. Cette mesure est déterminée par l'analyse onde progressive du système (9). Un choc non classique possède sa propre dissipation d'entropie, déterminée par le critère de sélection. Différents opérateurs de régularisation génèrent a priori différentes mesures $\mathcal{M}$.
Le principe de sélection introduit ci-dessus peut aussi être réinterprété dans le formalisme des fonctions cinétiques : pour un état gauche $\mathbf{u}_L$ donné, l'analyse onde progressive définit un état droit $\mathbf{u}_R = \phi(\mathbf{u}_L)$ et une vitesse de propagation σ déterminée par les relations de Rankine-Hugoniot :

$$-\sigma(\phi(\mathbf{u}_L) - \mathbf{u}_L) + (\mathbf{f}(\phi(\mathbf{u}_L)) - \mathbf{f}(\mathbf{u}_L)) = 0,$$

de telle sorte que la discontinuité :

$$\mathbf{u}(x,t) = \begin{cases} \mathbf{u}_L & \text{si} \quad x < \sigma t, \\ \mathbf{u}_R & \text{si} \quad x > \sigma t, \end{cases} \tag{17}$$

soit une solution non classique admissible du système (8). La consistance avec l'inégalité d'entropie (10) au sens des distributions implique

$$D := -\sigma(U(\phi(\mathbf{u}_L)) - U(\mathbf{u}_L)) + (F(\phi(\mathbf{u}_L)) - F(\mathbf{u}_L)) \leq 0.$$

Une fois de plus, différents opérateurs de régularisation génèrent des fonctions cinétiques ϕ possiblement distinctes. Réciproquement, différentes fonctions cinétiques sont associées à différentes régularisations. Ainsi, le concept de fonction cinétique semble bien adapté pour traduire simplement les effets d'un opérateur de régularisation donné, et assurer l'unicité d'une solution faible entropique du système (8). Notons que la connaissance d'une fonction cinétique ϕ détermine la dissipation d'entropie D à travers un choc non classique, σ étant déterminé par les relations de Rankine-Hugoniot habituelles. La mesure de Borel $\mathcal{M}$ intervenant dans (16) est dès lors explicitement connue (comme une fonction de $\mathbf{u}_L$) à travers une telle discontinuité par la donnée de la fonction cinétique ϕ. Nous renvoyons le lecteur à LeFloch [88] pour une revue sur les fonctions cinétiques. D'un point de vue numérique, celles-ci seront considérées dans les chapitres IV 1, IV 2, et IV 3.

Pour conclure ces rappels, nous précisons que les considérations présentées ici dans le contexte des systèmes hyperboliques restent valables pour les systèmes mixtes de type hyperbolique-elliptique, en ce sens que l'inégalité d'entropie (10) ne suffit pas à garantir l'unicité des solutions faibles. De plus, de tels systèmes génèrent (par régularisation visco-capillaire) des discontinuités non classiques sous-compressives correspondant à des transitions de phase subsoniques.

Travaux effectués

La principale motivation des travaux effectués concerne l'approximation numérique des chocs non classiques introduits ci-dessus, ainsi que l'étude de leur comportement (stabilité, dépendance par rapport à l'opérateur de régularisation, comportement asymptotique, interaction, ...).

Dans la littérature, il existe plusieurs approches permettant de traiter numériquement les modèles donnant naissance à des solutions non classiques. Les deux "principales" suivent les caractérisations mises en évidence ci-dessus. La première approche propose de représenter le plus fidèlement possible l'opérateur de régularisation considéré au niveau discret. Les discrétisations obtenues contiennent des termes de diffusion et de dispersion qui permettent d'obtenir numériquement des chocs sous-compressifs (voir Hayes et LeFloch [64], Zhong, LeFloch et Rohde [94], Fan [46], [47], Fan et Slemrod [50], Chalons et LeFloch [24], [25], ...). La deuxième approche propose d'imposer la vitesse de propagation d'une onde non classique au niveau discret par l'intermédiaire d'une fonction cinétique supposée connue. Cette approche est communément répertoriée sous le nom de "sharp interface approach" (interfaces raides) et est associée aux méthodes de type level-set, suivi de front et Glimm (voir Hou, LeFloch et Rosakis [70], LeFloch [86], Abeyaratne et Knowles [3], [2], Truskinovsky [141], [142], Chalons et LeFloch [26], [27], ...).
Les chocs non classiques étant dépendants de l'opérateur de régularisation considéré, les solutions numériques obtenues par la première approche se trouvent être très sensibles à la dissipation numérique, pas toujours maîtrisée, du schéma utilisé. En ce sens, la deuxième approche semble mieux adaptée à notre propos. Remarquons que celle-ci

suppose en contrepartie la connaissance explicite de la fonction cinétique associée à une régularisation donnée, et donc à un problème donné.

Dans ce manuscrit, nous mettons en œuvre les deux techniques pour approcher et étudier le comportement des solutions de systèmes conservatifs ayant un intérêt physique et générant des chocs non classiques sous-compressifs.
Ainsi,

(i) les chapitres IV 1 et IV 2 proposent une approximation numérique du problème (8)-(10), d'une part lorsque l'on fait le choix $N = 1$ avec une fonction flux non convexe (cas académique) et d'autre part lorsque l'on considère les deux lois de conservation modélisant la dynamique des fluides de van der Waals. Les schémas numériques présentés sont basés sur une discrétisation précise des régularisations visco-capillaires considérées, et leur stabilité est assurée par la consistance discrète en un sens à préciser avec l'inégalité d'entropie (10). À l'aide de ces schémas et suivant les travaux de Hayes et LeFloch [64], nous étudions pour les deux modèles considérés le comportement des chocs non classiques et des transitions de phase observées (dépendance par rapport aux coefficients de viscosité et de capillarité, tracé des fonctions cinétiques correspondantes, ...).

(ii) le chapitre IV 3 utilise la méthode de Glimm pour discrétiser le système (8). Le solveur de Riemann considéré pour implémenter cette méthode est associé à une fonction cinétique donnée. Tirant profit des propriétés de stabilité et de précision satisfaites par la méthode de Glimm (convergence, discontinuités raides, ...), nous étudions la stabilité et le comportement asymptotique en temps (pour des fonctions périodiques) des chocs non classiques associés aux fonctions cinétiques utilisées. Pour ces études, nous considérons une loi de conservation scalaire avec une fonction flux non convexe et un système de deux lois de conservation intervenant en élastodynamique non linéaire.

Deuxième partie

Analyse numérique des équations de Navier-Stokes à plusieurs pressions indépendantes

Chapitre 1

The Riemann problem for the multi-pressure Euler system

Ce chapitre fait l'objet d'un article en préparation.

The Riemann problem for the multi-pressure Euler system

Chalons C.[1,2] and Coquel F.[3]

Résumé

We prove existence and uniqueness of Riemann solutions for the Euler equations closed by N independent constitutive pressure laws. This PDE model stands as a natural asymptotic system for the multi-pressure Navier-Stokes equations in the regime of an infinite Reynolds number. Due to the inherent lack of conservation form in the viscous regularization, the limit system exhibits measure valued source terms concentrated on shock discontinuities. These non positive bounded measures, called kinetic relations after [15], are known to provide a suitable tool to encode the small scale sensitiveness in the singular limit. Here considering N independent polytropic pressure laws, we show that these kinetic relations can be derived when solving a simple algebraic problem which governs the endpoints of the underlying viscous shock profiles for any given but prescribed ratios of viscosities in the viscous perturbations. The analysis based on travelling wave solutions will allow us to introduce the asymptotic Euler system in the setting of piecewise Lipschitz-continuous functions so as to study the Riemann problem.

1.1 Introduction

The present work aims at studying a particular class of solutions of the following non linear hyperbolic system with viscous perturbations :

$$\left\{\begin{array}{l} \partial_t \rho_\epsilon + \partial_x(\rho u)_\epsilon = 0, \\ \partial_t(\rho u)_\epsilon + \partial_x(\rho u^2 + \sum_{i=1}^N p_i)_\epsilon = \epsilon \partial_x((\sum_{i=1}^N \mu_i)\partial_x u)_\epsilon, \\ \partial_t(\rho\varepsilon_i)_\epsilon + \partial_x(\rho\varepsilon_i u)_\epsilon + (p_i)_\epsilon \partial_x u_\epsilon = \epsilon(\mu_i)_\epsilon(\partial_x u)^2_\epsilon, \quad i = 1, ..., N, \end{array}\right. \tag{1.1}$$

in the asymptotic regime of a vanishing parameter $\epsilon \to 0^+$, *i.e.* in the regime of an infinite Reynolds number. This PDE system indeed stands as a natural extension of the usual Navier-Stokes equations when considering N independent partial internal energies. The N independent pressure laws p_i are given hereafter polytropic gas definitions.

This work extends a study by Berthon [10] in the case $N = 2$ for finite Reynolds numbers (*i.e.* for fixed $\epsilon > 0$), motivated by some turbulent models from the physics (the so-called $k - \epsilon$ or associated models). The present extension to values $N > 2$ is motivated by more sophisticated turbulent modelling where several turbulent temperatures k_i with $i = 1, ..., N$ are distinguished for the sake of a better accuracy. The PDE system (1.1) can also be understood as a simplified form of complex mixture flow modelling : namely advection equations for species mass fractions are in order. Such

1. O.N.E.R.A., B.P. 72, 29 avenue de la Division Leclerc, 92322 Châtillon Cedex, France.
E-mail: Christophe.Chalons@onera.fr.
2. Centre de Mathématiques Appliquées & Centre National de la Recherche Scientifique, U.M.R. 7641, École Polytechnique, 91128 Palaiseau Cedex, France.
E-mail: chalons@cmapx.polytechnique.fr
3. Laboratoire JLL & Centre National de la Recherche Scientifique, U.M.R. 7598, Boîte courrier 187, 75252 Paris Cedex 05, France.
E-mail: coquel@ann.jussieu.fr

models are discussed for instance by Truesdell [137] or more recently by Lagoutière [79]. All these models are oftenly addressed in practical issues in the regime of large Reynolds numbers (*i.e.* small values of ϵ). Since we consider an infinite domain (*i.e.* without boundary), we are classically led to neglect the small scale effects coming with small values of ϵ in the solutions of (1.1) : that is to say one has to pass in the limit $\epsilon \to 0$. When a single pressure law p_1 is addressed, this procedure formally gives rise to the usual 3×3 Euler equations for governing the conservative unknowns : density ρ, momentum ρu and total energy $\rho E := (\rho u)^2/2\rho + \rho\varepsilon_1$. A full convergence result is still an open question. In the extended framework of N independent pressure laws p_i, the situation turns out to be one step further difficult because of the apparent lack of conservation form. As underlined in Chalons [19] (see also Berthon [10] and below for a brief discussion), the PDE system (1.1) cannot recast in full conservation form without restrictive modelling assumptions. Let us briefly put forward the consequences of such a negative issue.
As already claimed, the first order underlying system obtained when formally setting $\epsilon = 0$ in (1.1) is seen hereafter to be hyperbolic with two extreme genuinely non linear fields. By this property, we thus expect families of smooth solutions $\{\mathbf{u}_\epsilon\}_{\epsilon>0}$ of (1.1) to converge generally speaking to discontinuous functions as ϵ goes to zero. Because of the lack of a conservation form for (1.1), the difficulty is to understand these limit functions as solutions of a first order limit system. Our main objective here is to derive such a limit system when restricting ourselves to piecewise Lipschitz-continuous limit functions. A notion of weak solutions for such functions will naturally arise following the pionnering works by LeFloch [90]. The main result of this paper is the existence and uniqueness of a particular class of such weak solutions of practical importance : namely solutions of the Riemann problem coming with the limit system. Non linear hyperbolic systems in non conservation form have received some important contributions over the past decade. LeFloch, Dal Maso and Murat [42] (see also LeFloch [90]) have in particular proposed a notion of fixed family of paths to define ambiguous products, extending the work by Volpert [145]. Equipped with this definition, they have been able to propose a notion of weak solutions for such systems and have solved the Riemann problem for general systems when considering two sufficiently close initial states. Of course and for two fixed initial states, the solution under consideration heavily depends on the particular family of paths under consideration. LeFloch has proved in [90] that the choice of a particular family of paths can be performed on the ground of an analysis of the travelling wave solutions of a parabolic regularization of the hyperbolic system under consideration. Because of the non conservative nature of the first order system, the derived family of paths generally varies when considering non proportionnal viscous tensors. This sensitiveness has been exemplified for instance by Raviart and Sainsaulieu [112] within the frame of distinct averaged multi-phase flows models and also by Berthon [10] for various closures of turbulent models. Such a sensitiveness has been proved to also occur in completely different settings : mixed hyperbolic-elliptic systems (see Bedjaoui and LeFloch [9], Chalons and LeFloch [25], ...) or even hyperbolic systems in pure conservation form but with a loss of non linearity (see LeFloch [90], [87], Hayes and LeFloch [69], [65] Chalons and LeFloch [24], [26] ...). Here, we adopt this natural strategy to define shock solutions of the limit first order system as limits of travelling wave solutions of (1.1) when $\epsilon \to 0$. Using the Lasalle invariance principle, Berthon and Coquel [14] have proved the existence in

the large of travelling wave solutions for (1.1). Their analysis proves by abstract geometrical considerations the existence of an unique exit state $\mathbf{u}_+$ reached in the future by a travelling wave solution connecting some given state $\mathbf{u}_-$ in the past, a suitable speed of propagation σ being prescribed. The existence of the triple $(\sigma; \mathbf{u}_-, \mathbf{u}_+)$ then naturally serves to define a shock solution (see [90], [112], [10] and below). But due to the non conservative nature of the dynamical system governing the heteroclinic solution under consideration, barely little is known about the exit state $\mathbf{u}_+$ except its existence and the property that its precise form heavily depends on the N ratios of the viscosities $\mu_i / \sum_{j=1}^{N} \mu_j$. Such a dependence simply reflects the sensitiveness we have already reported. In other words, shock solutions are known to exist as regularization depending limit functions, but without precise informations about this dependency. Of course it is out of reach to tackle the study of the Riemann problem without a minimum of informations about shock solutions.

In the present work, we show how to get sufficient informations from the analysis of an auxiliary companion ODE problem. Besides its dimensionless form, the auxiliary ODE system turns out to be linear with constant coefficients: the N viscosity ratios dictate these coefficients, the thermodynamic closures being tacitly fixed. Its solution, therefore depending on these N ratios, will then serve to solve a non linear algebraic scalar problem that characterizes the critical points of the original dynamical system associated with (1.1). From the proposed analysis and roughly speaking, a left state $\mathbf{u}_-$ being prescribed, the state $\mathbf{u}_+$ will come out as a natural function of the speed of propagation σ when just solving a scalar equation for different values of σ (the N viscosity ratios being fixed). On the one hand, the proposed analysis will allow to derive a convenient form of the limit system based on N so-called kinetic functions, after LeFloch [88] (see also [15] for the setting of hyperbolic systems in non conservation form). Roughly speaking, these N kinetic functions stand for the singular limits as ϵ goes to zero of the N entropy dissipation rates entering the N entropy equations associated with (1.1). On the second hand, we will be able to define a notion of shock curves in the physical phase space, which extends naturally the setting of a single pressure law. Rarefaction curves will be characterized without difficulty since their derivation relies on particular smooth solutions for the limit system. We shall be in a position to determine the projection of these two families of curves on the convenient plane with total pressure $(\sum_{i=1}^{N} p_i)$ and velocity (u) as independent coordinates. The strict monotonicity property of these projections and their asymptotic behaviours will allow us to prove existence and uniqueness of a solution to the Riemann problem under the only classical requirement of staying away from vacuum. Of course Riemann solutions will be sensitive with respect to the N viscosity ratios since they are generally made of shock solutions. This small scale sensitiveness will be seen to be encoded in the limit system in the definition of the N kinetic functions.

1.2 Some basic properties of the PDE model

In this section, we precise the required closure assumptions and state useful properties of the solutions of the system (1.1) when choosing $\epsilon = 1$ for simplicity in the notations:

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + \sum_{i=1}^N p_i) = \partial_x ((\sum_{i=1}^N \mu_i) \partial_x u), \\ \partial_t \rho \varepsilon_i + \partial_x \rho \varepsilon_i u + p_i \partial_x u = \mu_i (\partial_x u)^2, \quad i = 1, ..., N. \end{cases} \tag{1.2}$$

The N independent partial pressure laws p_i, $i = 1, ..., N$, entering (1.2) find the following polytropic ideal gas definitions:

$$p_i = (\gamma_i - 1)\rho\varepsilon_i, \quad i = 1, ..., N,$$

where the N adiabatic exponents γ_i are constant real numbers with $\gamma_i > 1$. Then, the viscosity coefficients μ_i are assumed to be non negative real numbers but with the requirement that

$$\mu := \sum_{i=1}^N \mu_i > 0.$$

The unknown $\mathbf{u} = (\rho, \rho u, \{\rho\varepsilon_i\}_{i=1,...,N})$ of (1.2) is associated with the following natural phase space:

$$\Omega := \{\mathbf{u} = (\rho, \rho u, \{\rho\varepsilon_i\}_{i=1,...,N}) \in \mathbb{R}^{N+2},\ \rho > 0,\ \rho u \in \mathbb{R},\ \rho\varepsilon_i > 0,\ i = 1, ..., N\}, \tag{1.3}$$

from which the vacuum is excluded for simplicity. We first state the following easy result which proof is left to the reader (see Berthon [10]):

Proposition 1 *The first order underlying system in (1.2):*

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + \sum_{i=1}^N p_i) = 0, \\ \partial_t \rho \varepsilon_i + \partial_x \rho \varepsilon_i u + p_i \partial_x u = 0, \quad i = 1, ..., N, \end{cases} \tag{1.4}$$

is hyperbolic over Ω and admits the following three distinct eigenvalues:

$$u - c, \quad u, \quad u + c, \quad c^2 = \sum_{i=1}^N \frac{\gamma_i p_i}{\rho},$$

where u has N order of multiplicity. Moreover, the eigenvalues $u - c$ and $u + c$ are associated with genuinely non linear fields, while u is associated with linearly degenerate fields. In addition, the latter admits the velocity u and the total pressure $p := \sum_{i=1}^N p_i$ as Riemann invariants.

By the non linearity of the two extreme fields, solutions of the hyperbolic system (1.4) are known to generally develop discontinuities in finite time. The system (1.4) must be thus understood in a weaker sense, but the non conservative products in (1.4) are (in general) ambiguous in the usual sense of the distributions. Notice that the intermediate fields are naturally associated with discontinuous solutions, namely the so-called contact discontinuities propagating with constant speed u by the linear degeneracy of these fields (see also subsection 1.4.2 for related details). The continuity of the velocity u thus precludes the occurence of ambiguities in all the non conservative products entering (1.4) (but also in (1.2)). The situation is completely different when addressing discontinuities coming with the two extreme fields: namely the so-called shock discontinuities. Indeed and across shocks, the velocity u and the partial pressures p_i will be seen below to be necessarily discontinuous. Therefore, passing into the limit $\epsilon \to 0$ in (1.1) will necessarily result into non trivial Borel measures in both the left and right hand sides of the N last equations. This last observation thus motivates us to study for existence additional conservation laws for smooth solutions of (1.2). The next statement gives the existence of solely one additional non trivial conservation law. Indeed, the N supplementary balance equations exhibited below dot not give rise to conservation laws unless specific assumptions on the adiabatic exponents and/or on the viscosity coefficients are made (see Berthon [10] and Chalons [19] for the details).

Proposition 2 *Smooth solutions of (1.2) obey the following conservation law on the total energy ρE :*

$$\partial_t \rho E + \partial_x(\rho E + p)u = \partial_x(\mu u \partial_x u), \tag{1.5}$$

where we have set :

$$\rho E = \frac{(\rho u)^2}{2\rho} + \rho\varepsilon, \quad \rho\varepsilon = \sum_{i=1}^{N} \rho\varepsilon_i, \tag{1.6}$$

and

$$p = \sum_{i=1}^{N} p_i = \sum_{i=1}^{N} (\gamma_i - 1)\rho\varepsilon_i.$$

Moreover, smooth solutions of (1.2) obey the following entropy balance equations on the specific entropies s_i :

$$\partial_t \rho s_i + \partial_x \rho s_i u = \mu_i \frac{\gamma_i - 1}{\rho^{\gamma_i - 1}} (\partial_x u)^2, \quad i = 1, ..., N, \tag{1.7}$$

where we have set :

$$s_i = \frac{p_i}{\rho^{\gamma_i}}, \quad i = 1, ..., N. \tag{1.8}$$

The $(N+1)$ laws (1.5) and (1.7) are the only additional non trivial equations, up to some classical non linear transforms in the $\{s_i\}_{i=1,...,N}$.

The $(N+2)$ PDE system (1.2) cannot be thus given an admissible change of variables so as to recast in full conservation form. The reported properties of the shock solutions clearly make the equations (1.7), once rescaled in time and space by the parameter ϵ, to give birth to non trivial measures in the limit $\epsilon \to 0$ as soon as $\mu_i > 0$. In other words and whatever are the variables under consideration, passing to the limit in the rescaled system (1.1) will result in the need for dealing with measures concentrated on shock discontinuities. According to the motivations discussed in the introduction, we propose to define the shock solutions when passing into the limit of travelling wave solutions of system (1.1) as ϵ goes to zero. This procedure will allow for defining the measures entering the limit first order system coming from (1.1) as $\epsilon \to 0$. As it is well-known, a travelling wave solution of system (1.2) immediately gives, once rescaled by ϵ, a corresponding travelling wave solution of (1.1) (see hereafter). As an immediate consequence, it thus suffices to study the travelling wave solutions of (1.2).

1.3 Travelling wave solutions

Let us now consider a travelling wave solution of (1.2), that is to say a smooth solution of (1.2), we denote $\mathbf{u}$, under the special form:

$$\left\{ \mathbf{u}(x,t) = \mathbf{u}(\xi), \quad \xi = x - \sigma t, \right. \tag{1.9}$$

where $\sigma \in I\!R$ stands for the speed of propagation. Such a solution has to obey the following usual asymptotic behaviour:

$$\lim_{\xi \to -\infty} \mathbf{u}(\xi) = \mathbf{u}_-, \quad \lim_{\xi \to +\infty} \mathbf{u}(\xi) = \mathbf{u}_+, \tag{1.10}$$

for two constant states $\mathbf{u}_-$ and $\mathbf{u}_+$ in Ω. When considering the system (1.2), the solution $\mathbf{u}(\xi)$ under consideration must satisfy the following non linear $(N+2)$-ODE system:

$$\left\{ \begin{array}{l} -\sigma d_\xi \rho + d_\xi \rho u = 0, \\ -\sigma d_\xi \rho u + d_\xi(\rho u^2 + p) = d_\xi(\mu d_\xi u), \\ -\sigma d_\xi \rho \varepsilon_i + d_\xi \rho \varepsilon_i u + p_i d_\xi u = \mu_i (d_\xi u)^2, \quad i = 1, ..., N, \end{array} \right. \tag{1.11}$$

where d_ξ represents the classical derivative operator. Observe from the first ODE in (1.11) that the relative mass flux $m = \rho(\xi)(u(\xi) - \sigma)$ is an algebraic invariant given by $m = \rho_-(u_- - \sigma)$. In the classical setting of a single pressure law, this algebraic invariant cannot be zero (see below) and is positive (respectively negative) for viscous shock profiles associated with the first field (respectively the third field). Classical considerations within this setting naturally extend to present one to prove that it suffices to restrict ourselves to the study of viscous profiles for shocks coming with the first field. Indeed, viscous shock profiles for the third field will be recovered when exchanging the role of the exit states $\mathbf{u}_-$ and $\mathbf{u}_+$ while reversing the sign of ξ in the forthcoming analysis. On the ground of this observation, let us now introduce a convenient reformulation of the dynamical system (1.11) when considering positive values of m:

Lemma 1 *Let be given a fixed state $\mathbf{u}_-$ in Ω and a speed σ with $m = \rho_-(u_- - \sigma) > 0$. Then, travelling wave solutions of (1.11) equivalently obey the following $(N+1)$-ODE system:*

$$\left\{ \begin{array}{l} d_\xi \tau = \frac{1}{\mu m} \mathcal{F}(\tau, \{s_i\}_{i=1,...,N}), \\ d_\xi s_i = (\gamma_i - 1)\tau^{\gamma_i - 1} \frac{\mu_i}{\mu^2 m} \mathcal{F}^2(\tau, \{s_i\}_{i=1,...,N}), \quad i = 1, ..., N, \end{array} \right. \tag{1.12}$$

where we have set:

$$\mathcal{F}(\tau, \{s_i\}_{i=1,...,N}) = m^2(\tau - \tau_-) + (p(\tau, \{s_i\}_{i=1,...,N}) - p_-). \tag{1.13}$$

Consequently, critical points $\mathbf{u}_+$ of the dynamical system (1.12) are entirely characterized by:

$$\mathcal{F}(\tau_+, \{(s_i)_+\}_{i=1,...,N}) = 0. \tag{1.14}$$

Before addressing the proof of this lemma, notice that (1.14) just reflects the conservation form of the PDE governing the momentum ρu. Since

$$p(\tau, \{s_i\}_{i=1,\dots,N}) = \sum_{i=1}^{N} s_i \tau^{-\gamma_i},$$

with $\gamma_i > 1$, the vector field entering the dynamical system (1.12) is Lipschitz-continuous over the following open subset:

$$\mathcal{D} = \{\mathbf{v} := (\tau, \{s_i\}_{i=1,\dots,N}) \in I\!R^{N+1},\ \tau > 0\}. \tag{1.15}$$

By the well-known continuation theorem, we thus have existence and uniqueness of a trajectory $\mathbf{v}(\xi)$ with $\mathbf{v}(0) = \mathbf{v}_0$ in $\mathcal{D}$ over the maximal interval $[0, \xi^+_{max}(\mathbf{v}_0))$, the associated maximal time ξ^+_{max} being finite if and only if $\tau(\xi^+_{max}) = 0$.

Proof Since by assumption a travelling wave is a smooth solution of (1.2), it must also satisfy the N independent balance equations (1.7) for governing the specific entropies s_i. Consequently, a travelling wave solution equivalently solves:

$$\left\{\begin{array}{l} -\sigma d_\xi \rho + d_\xi \rho u = 0, \\ -\sigma d_\xi \rho u + d_\xi(\rho u^2 + p) = d_\xi(\mu d_\xi u), \\ -\sigma d_\xi \rho s_i + d_\xi \rho s_i u = \mu_i(\gamma_i - 1)\tau^{(\gamma_i - 1)}(d_\xi u)^2, \quad i = 1, ..., N. \end{array}\right.$$

Next, the algebraic invariant $m = \rho(\xi)(u(\xi) - \sigma)$ gives $m d_\xi \tau = d_\xi u$ so that we get after easy calculations:

$$\left\{\begin{array}{l} m^2 d_\xi \tau + d_\xi p = m d_\xi(\mu d_\xi \tau), \\ m d_\xi s_i = m^2 \mu_i(\gamma_i - 1)\tau^{(\gamma_i - 1)}(d_\xi \tau)^2, \quad i = 1, ..., N. \end{array}\right. \tag{1.16}$$

Integrating the first equation of (1.16) from $-\infty$ to ξ, we arrive at:

$$\mu m d_\xi \tau = \mathcal{F}(\tau, \{s_i\}_{i=1,\dots,N}) = m^2(\tau - \tau_-) + (p(\tau, \{s_i\}_{i=1,\dots,N}) - p_-),$$

so that (1.16) easily reexpresses as (1.12). This completes the proof.

The main result of this section is:

Theorem 1 *Let be given a fixed state $\mathbf{u}_-$ in Ω and let us consider a speed σ subject to the following condition:*

$$\frac{u_- - \sigma}{c_-} > 1, \quad c_-^2 = \sum_{i=1}^{N} \gamma_i (p_i \tau)_-. \tag{1.17}$$

Then (1.17) is a necessary and sufficient condition for existence and uniqueness (up to some translation) of a travelling wave solution.

Remark 1 Observe that this statement extends immediately to viscous profiles for shocks coming with the third field but when asking:

$$\frac{u_+ - \sigma}{c_+} < -1, \quad c_+^2 = \sum_{i=1}^{N} \gamma_i (p_i \tau)_+.$$

In other words and for any given prescribed mass flux $m = \rho_-(u_- - \sigma)$ with $m > (\rho c)_-$, the dynamical system (1.12) admits an unique orbit $\mathcal{O}^-$ connecting $\mathbf{u}_-$ in the past and a (unique) state $\mathbf{u}_+$ in the future. The proof of this claim will follow from the developments proposed in the next three subsections. The first one will establish the existence of exactly one orbit, say $\mathcal{O}^-$, that can connect $\mathbf{u}_-$ in the past. This orbit will be seen to be necessarily compressive in its maximal interval of existence: $d_\xi \tau < 0$ for all finite $\xi < \xi^+_{max}$ with $\xi^+_{max} \in \overline{I\!R}$. This strong but expected property will allow us in the second subsection to derive an auxiliary linear ODE problem with constant coefficients. By construction, its solution will describe all the states $(\tau, \{s_i(\tau)\}_{i=1,\dots,N})$ along $\mathcal{O}^-$. Equipped with this parametrization, we shall then solve the algebraic equation $\mathcal{F}(\tau, \{s_i(\tau)\}_{i=1,\dots,N}) = 0$ that governs all the critical points of the dynamical system (1.12). We shall get the existence of an unique relevant $\tau_+ > 0$ so as to define a state $\mathbf{u}_+$ to be connected by $\mathcal{O}^-$ in the future.

1.3.1 Heteroclinic orbits are compressive

Here, the main result is given by the following statement:

Proposition 3 *The state $\mathbf{u}_-$ can be connected in the past by orbits of the dynamical system (1.12) if and only if the Lax condition (1.17) is satisfied. Under this condition, there exists at most one heteroclinic orbit connecting $\mathbf{u}_-$, which in addition is such that the mapping $\xi \to \tau(\xi)$ is necessarily strictly decreasing in the maximal interval of existence.*

In other words, viscous shock profiles are necessarily compressive as soon as they exist. In the usual setting of a single pressure law, the property (1.17) is known to be connected to the familiar Lax shock conditions expressed on the two exit states $\mathbf{u}_-$ and $\mathbf{u}_+$ of a heteroclinic solution. It is worthy to briefly prove for validity this connection in the present extended framework. As seen below, the above statement will follow from the well-known center manifold theorem (see [79] for instance). Indeed, the linearization of the vector field entering the definition of the dynamical system (1.12) under study, when evaluated at the critical point $\mathbf{u}_-$, is easily seen to admit only one non trivial eigenvalue:

$$\lambda(\mathbf{u}_-) = \frac{1}{\mu m}\partial_\tau \mathcal{F}(\mathbf{u}_-) = \frac{(\rho c)^2_-}{\mu m}((\frac{m}{\rho c})^2_- - 1). \tag{1.18}$$

A necessary and sufficient condition for the state $\mathbf{u}_-$ to be connected in the past is then $\lambda(\mathbf{u}_-) > 0$, but this condition is nothing else (1.17). A next and important consequence of the center manifold theorem states that if a critical point $\mathbf{u}_+$ is connected in the future, then necessarily this exit state obeys $\lambda(\mathbf{u}_+) < 0$, that is to say:

$$\frac{u_+ - \sigma}{c_+} < 1, \quad c_+^2 = \sum_{i=1}^N \gamma_i (p_i \tau)_+. \tag{1.19}$$

The two inequalities (1.17) and (1.19) thus express nothing but the familiar Lax shock inequalities on the exit states:

$$u_+ - c_+ < \sigma < u_- - c_-. \tag{1.20}$$

We now prove the above result.

Proof An associated eigenvector with the eigenvalue (1.18) is given by e_τ, the unit vector in the τ direction. By the center manifold theorem, the manifold $\mathcal{W}_{\mathbf{u}}(\mathbf{u}_-)$, made of the totality of the orbits connecting $\mathbf{u}_-$ in the past, is necessarily tangent to e_τ at $\mathbf{u}_-$. As a consequence, there exists exactly two orbits connecting $\mathbf{u}_-$ in the past, say $\mathcal{O}^-$ and $\mathcal{O}^+$, with the following distinctive properties, valid for $|\tau - \tau_-|$ sufficiently small :

$$\tau < \tau_- \text{ for } \tau \in \mathcal{O}^-, \quad \text{while} \quad \tau > \tau_- \text{ for } \tau \in \mathcal{O}^+. \tag{1.21}$$

Let us then prove that the orbit $\mathcal{O}^+$ cannot give rise to a heteroclinic solution. In that aim, it suffices to prove that the following domain is positively invariant for the dynamical system (1.12) :

$$\mathcal{D} = \{\mathbf{u} \in \Omega, \tau > \tau_-, s_i \geq (s_i)_- \text{ for } i \in [1, N], s_{i_0} > (s_{i_0})_- \text{ for some } i_0 \in [1, N]\}, \tag{1.22}$$

while it cannot contain a critical point. The conclusion will directly follow since by definition $\mathcal{O}^+$ enters this region. Indeed since by assumption $\mu := \sum_{i=1}^N \mu_i > 0$, there exists at least one index $i_0 \in [1, N]$ with $\mu_{i_0} > 0$ so that $d_\xi s_{i_0} > 0$, that is to say $s_{i_0}(\xi) > (s_{i_0})_-$.
The required properties are an easy consequence of the next estimate :

$$\mathcal{F}(\tau, \{s_i\}_{i=1,\dots,N}) > 0, \text{ for all } \mathbf{u} \in \mathcal{D}.$$

Let us first observe that the mapping $\mathcal{F}(\tau, \{s_i\}_{i=1,\dots,N})$ strictly increases with each of the specific entropies ($\partial_{s_i}\mathcal{F} = \tau^{-\gamma_i}$), hence any state $\mathbf{u} \in \mathcal{D}$ (*i.e* with $s_i > (s_{i_0})_-$ for some $i_0 \in [1, N]$) obeys :

$$\mathcal{F}(\tau, \{(s_i)_-\}_{i=1,\dots,N}) < \mathcal{F}(\tau, \{s_i\}_{i=1,\dots,N}).$$

Let us then conclude when establishing that

$$0 = \mathcal{F}(\tau_-, \{(s_i)_-\}_{i=1,\dots,N}) < \mathcal{F}(\tau, \{(s_i)_-\}_{i=1,\dots,N}) \text{ for all } \tau > \tau_-. \tag{1.23}$$

This estimate follows when observing that the mapping $\mathcal{F}(\tau, \{s_i\}_{i=1,\dots,N})$ is strictly convex in the τ variable, therefore for all $\tau > \tau_-$:

$$\partial_\tau \mathcal{F}(\tau, \{(s_i)_-\}_{i=1,\dots,N}) > \partial_\tau \mathcal{F}(\tau_-, \{(s_i)_-\}_{i=1,\dots,N}) = (\rho c)_-^2 ((\frac{m}{\rho c})_-^2 - 1).$$

But since the relative mach number must meet the condition (1.17), the mapping $\mathcal{F}(\tau, \{(s_i)_-\}_{i=1,\dots,N})$ strictly increases with τ and (1.23) immediately follows. These considerations therefore preclude $\mathcal{O}^+$ to connect some critical point in the future. The orbit $\mathcal{O}^-$ stays the only possible candidate which we now prove to be necessarily compressive in its maximal interval of existence. Indeed and by definition, $\mathcal{O}^-$ is necessarily made of states with $\tau < \tau_-$ and $\mathcal{F}(\tau, \{(s_i)\}_{i=1,\dots,N}) < 0$ for $|\tau - \tau_-|$ sufficiently small (in view of $\partial_\tau \mathcal{F}(\tau_-, \{(s_i)_-\}_{i=1,\dots,N}) > 0$, see (1.18)). Let us then observe that the mapping $\mathcal{F}$ stays necessarily negative for all values of ξ smaller than the maximal time of existence of the orbit $\mathcal{O}^-$. Indeed, states at which the mapping $\mathcal{F}$ vanishes are by definition critical points of the dynamical system (1.12) and as it is well-known, such states cannot be reached in finite time. This concludes the proof.

In view of Proposition 3, the application $\xi \to \tau(\xi)$ maps the maximal interval of existence $]-\infty, \xi^+_{max}[$ of the orbit connecting $\mathbf{u}_-$ in the past onto $]\tau_{min}, \tau_-[$ for some $\tau_{min} \geq 0$. Its strict monotonicity allows for defining the inverse mapping $\tau \to \xi(\tau)$, from $]\tau_{min}, \tau_-[$ onto $]-\infty, \xi^+_{max}[$ and thus for understanding any given function of ξ as a function of the covolume τ. Specializing this property to each of the specific entropies s_i with $i \in [1, N]$, the N mappings $\tau \in]\tau_{min}, \tau_-[\to s_i(\tau)$ defined from the orbit $\mathcal{O}^-$ of the dynamical system (1.12) are easily seen to obey the following linear ODE system with variable coefficients:

$$d_\tau s_i = (\gamma_i - 1)\tau^{(\gamma_i-1)}\frac{\mu_i}{\mu}\mathcal{F}(\tau, \{s_i\}_{i=1,...,N}), \quad i = 1, ..., N, \tag{1.24}$$

where τ plays the role of a time like variable. Notice that the above system stays meaningfull when evaluated at time $\tau = \tau_-$. Let us thus study for times $\tau \leq \tau_-$ the existence of a solution of this non autonomous system with initial data $s_i(\tau_-) = (s_i)_-$, $i = 1, ..., N$ without reference to the time τ_{min} (*i.e.* without reference to the possibly finite maximal time ξ^+_{max} of existence of $\mathcal{O}^-$). Let us indeed observe that the solution under consideration exists for all positive time τ since the coefficients of (1.24) are then Lipschitz-continuous. Our purpose is to prove for existence a first positive time $\tau_\star < \tau_-$ such that

$$\mathcal{F}(\tau_\star, \{s_i(\tau_\star)\}_{i=1,...,N}) = 0, \tag{1.25}$$

with the property that the solutions $\{s_i(\tau)\}_{i=1,...,N}$ stays larger than $\{s_i(\tau_-)\}_{i=1,...,N}$ for all $\tau \in [\tau_\star, \tau_-]$. Observe that the algebraic equation (1.25) expresses nothing but the conservation of the total impulsion. The existence of such a time $\tau_\star > 0$ will be seen to imply that the orbit $\mathcal{O}^-$ necessarily connects in the future a critical point of the dynamical system (1.12): namely $\mathbf{u}_+ := (\tau_\star, \{s_i(\tau_\star)\}_{i=1,...,N})$.

1.3.2 Reduction to an algebraic problem

We prove below that the solution we seek for, can be actually inferred from the solution of a linear ODE problem with constant coefficients. Equipped with its solution explicitly given by the Duhamel representation formula, we shall then prove that the required time $\tau_\star$ exists and just solves a non linear algebraic equation equivalent to (1.25). More precisely, each of the specific entropies s_i will be rebuilt from the following dimensionless variables:

$$\theta_i = \frac{1}{(\gamma_i - 1)}\frac{(p_i - (p_i)_-)}{m^2\tau_-}, \quad i = 1, ..., N, \tag{1.26}$$

together with:

$$x = -\ln(\frac{\tau}{\tau_-}). \tag{1.27}$$

Observe that $\theta_i = 0$ when $p_i = (p_i)_-$ while $x \geq 0$ when $\tau \leq \tau_-$. x will play the role of a dimensionless time like variable while the dimensionless vector $\mathbf{\Theta} = \{\theta_i\}_{i=1,...,N}$

will denote the solution of the following linear ODE system with constant coefficients derived from the non autonomous system (1.24) at the end of the present section:

$$\begin{cases} d_x\mathbf{\Theta} - \mathcal{M}\mathbf{\Theta} = \alpha\mathbf{a} + (1-\exp(-x))\mathbf{b}, \\ \mathbf{\Theta}(x=0) = 0. \end{cases} \tag{1.28}$$

Here, the constant vectors of $I\!R^N$ entering the right hand side read:

$$\mathbf{a} = {}^T(\{\beta_i/(\gamma_i - 1)\}_{i=1,...,N}), \quad \mathbf{b} = {}^T(\{\mu_i/\mu\}_{i=1,...,N}), \tag{1.29}$$

where the dimensionless coefficients β_i with $i \in \{1, ..., N\}$ stand for the following ratios of sound speeds:

$$\beta_i = \frac{\gamma_i(p_i\tau)_-}{c_-^2}, \quad i = 1, ..., N, \quad \alpha = (\frac{(\rho c)_-}{m})^2. \tag{1.30}$$

α clearly coincides with the squarred inverse of the relative Mach number. Next in (1.28), $\mathcal{M} = (\mathcal{M}_{ij})_{i,j=1,...,N}$ stands for a $N \times N$ matrix with constant real coefficients given by:

$$\mathcal{M}_{ij} = \gamma_i\delta_{ij} - \mu_i(\gamma_j - 1)/\mu, \tag{1.31}$$

where δ denotes the Kronecker symbol. We shall briefly report some interesting properties of this matrix later on, together with some comments about the role of the dimensionless coefficients involved in (1.28). Let us just underline that in our forthcoming developments, the reduced number α will serve as a natural parameter in the solution of (1.28) to represent all the possible speeds of propagation σ that satisfy the Lax inequality (1.17), the left state $\mathbf{u}_-$ being tacitly fixed. As a consequence, α will belong to $[0, 1]$.
Let us now turn deriving the algebraic problem when recalling that the solution of the linear ODE problem (1.28) obeys the Duhamel formula:

$$\mathbf{\Theta}(x) = \alpha\int_0^x \exp[\mathcal{M}(x-y)]dy\ \mathbf{a} + \int_0^x [1-\exp(-y)]\exp[\mathcal{M}(x-y)]dy\ \mathbf{b}. \tag{1.32}$$

On the ground of this solution, we define the mapping $x \to \mathbf{u}(x)$ setting for $x > 0$:

$$\begin{gathered} \tau(x) = \tau_-\exp(-x), \quad u(x) = u_- - m\tau_-(1-\exp(-x)), \\ p_i(x) = (p_i)_- + m^2\tau_-(\gamma_i - 1)\theta_i(x), \quad i = 1, ..., N. \end{gathered} \tag{1.33}$$

Equipped with this mapping (which we do not claim to belong to the phase space Ω for all x), the next issue is concerned with the existence of positive roots x of the following algebraic problem:

$$I(x) = \sum_{i=1}^N p_i(x) + m^2\tau(x) = I_- := \sum_{i=1}^N (p_i)_- + m^2\tau_-. \tag{1.34}$$

The smallest positive root, which we denote $x(\alpha)$, will naturally serve to provide the existence of a critical point $\mathbf{u}_+$ which can be connected by the orbit $\mathcal{O}^-$ in the future, as soon as $\mathbf{u}(x) \in \Omega$ for all $x \in [0, x(\alpha)]$. Indeed, observe that $\tau_\star = \tau(x(\alpha))$ is nothing but the value entering the equation (1.25). This problematic reexpresses equivalently in terms of the dimensionless variables $\{\theta_i\}_{i=1,...,N}$ as follows: prove the existence of a smallest positive root to:

$$\sum_{i=1}^{N}(\gamma_i - 1)\theta_i(x) - (1 - \exp(-x)) = 0. \tag{1.35}$$

Taking advantage of the Duhamel representation formula (1.32), the above algebraic equation just recasts:

$$\begin{aligned}\psi(x,\alpha) := \alpha < \kappa, \int_0^x \exp[\mathcal{M}(x-y)]dy\ \mathbf{a} >\\ + < \kappa, \int_0^x [1 - \exp(-y)]\exp[\mathcal{M}(x-y)]dy\ \mathbf{b} > -(1 - \exp(-x)) = 0,\end{aligned}$$

where we have introduced $\kappa = {}^T(\{\gamma_i - 1\}_{i-1,...,N})$. Let us recall that the parameter α belongs to $[0,1]$. Then, setting

$$\phi(x) =< \kappa, \int_0^x \exp[\mathcal{M}(x-y)]dy\ \mathbf{a} >,$$

we have to prove the existence of positive roots of:

$$\psi(x,\alpha) = \alpha\phi(x) + \psi(x,0) = 0, \tag{1.36}$$

and select the smallest one, say $x(\alpha)$. This will be the matter of the next section.

Lemma 2 *Let be given a fixed state $\mathbf{u}_- \in \Omega$ and the associated reduced numbers $(\beta_i)_{i=1,...,N}$ in $[0,1]^N$ with $\sum_{i=1}^N \beta_i = 1$. Then assume that for α in $[0,1]$, the non linear algebraic problem (1.35) admits at least one positive solution. Denoting $x(\alpha)$ the smallest positive root, assume in addition that:*

$$\theta_i(x) \geq 0 \textit{ for all } x \in [0, x(\alpha)] \textit{ and all } i = 1, ..., N. \tag{1.37}$$

Then the orbit $\mathcal{O}^-$ connecting $\mathbf{u}_-$ in the past, connects a state $\mathbf{u}_+$ in future times given by:

$$\begin{aligned}&\tau_+ = \tau_- \exp(-x(\alpha)), \quad u_+ = u_- - m\tau_-(1 - \exp(-x(\alpha))),\\ &(p_i)_+ = (p_i)_- + m^2\tau_-(\gamma_i - 1)\theta_i(x(\alpha),\alpha), \quad i = 1, ..., N.\end{aligned} \tag{1.38}$$

The proof of this result is postponed at the end of the present section.

Let us observe that the dimensionless system (1.28) and the algebraic equation are kept unchanged when reversing the sign of the relative Mach number m. This invariance property thus ensures that an analysis of the travelling wave solutions associated with the third field will result in exactly the same reduced problem (provided that the role of the exit states $\mathbf{u}_-$ and $\mathbf{u}_+$ are exchanged). In other words, solving in $x(\alpha,\beta)$ the non linear equation (1.36) for all $\alpha \in [0,1]$ and all dimensionless vector β in the following subset of $\mathbb{R}^N$:

$$\mathbf{B} = \{\beta \in \mathbb{R}^N / \ \beta_i \geq 0, \ \ 1 \leq i \leq N, \quad \sum_{i=1}^{N}\beta_i = 1\}, \tag{1.39}$$

will suffice to generate a complete family of travelling waves for either the first or the third non linear field. Let us emphasize that these travelling wave solutions heavily depend on the N viscosity ratios μ_i/μ, since does the matrix $\mathcal{M}$ and the vector $\mathbf{b}$ in (1.28). Let us just mention that whatever are these ratios, the matrix $\mathcal{M}$ is a diagonally dominant M-matrix (*i.e.* with positive diagonal entries and non positive extra diagonal terms).

Let us now prove Lemma 2.

Proof For any given x in $]0, x(\alpha)[$, let us define a state $\mathbf{u}(x)$ in Ω when considering under the assumption (1.37) :

$$\tau(x) = \tau_- \exp(-x), \quad u(x) = u_- - m\tau_-(1 - \exp(-x)),$$

$$p_i(x) = (p_i)_- + m^2\tau_-(\gamma_i - 1)\theta_i(x) \geq (p_i)_-, \quad i = 1, ..., N.$$

Such a family of states can obviously be reparametrized in terms of the covolume and by construction the mapping $\tau \to \mathbf{u}(\tau)$, $\tau \in]\tau_- \exp(-x(\alpha)), \tau_-[$, is nothing but the solution of (1.24) with $s_i(\tau_-) = (s_i)_-$ for all $i = 1, ..., N$. Notice that for all $\tau \in]\tau_- \exp(-x(\alpha)), \tau_-[$, we have $\mathcal{F}(\tau, \{s_i(\tau)\}_{i=1,...,N}) < 0$ while

$$\mathcal{F}(\tau_- \exp(-x(\alpha)), \{s_i(\tau_- \exp(-x(\alpha)))\}_{i=1,...,N}) = 0.$$

Then, for all $\tau \in]\tau_- \exp(-x(\alpha)), \tau_-[$,

$$\frac{d\xi}{d\tau} = \frac{\mu m}{\mathcal{F}(\tau, \{s_i(\tau)\}_{i=1,...,N})} < 0,$$

and the well-defined mapping $\xi \to \tau(\xi)$ maps $]\tau_+, \tau_-[$ on $]-\infty, \xi^+_{max}[$ with $\xi^+_{max} \in \overline{I\!R}$. Then for all $\xi \in]-\infty, \xi^+_{max}[$, we have defined a mapping $\xi \to \mathbf{u}(\xi)$ as the solution of the dynamical system (1.12). But using the fact that $\mathcal{F}(\tau_+, \{s_i(\tau_+)\}_{i=1,...,N}) = 0$, then necessarily $\xi^+_{max} = +\infty$. This concludes the proof.

Let us now conclude this section when deriving the linear ODE system from the non autonomous system (1.24).

Proof By virtue of the second principle of the thermodynamics, we have :

$$\frac{\tau^{(1-\gamma_i)}}{(\gamma_i - 1)} ds_i = d\varepsilon_i + p_i d\tau, \quad i = 1, ..., N,$$

so that (1.24) is easily seen to reexpress equivalently as :

$$d_\tau \varepsilon_i + p_i = \frac{\mu_i}{\mu}\mathcal{F}(\tau, \{s_i\}_{i=1,...,N}), \quad i = 1, ..., N.$$

But using the definition of $\mathcal{F}$ we get :

$$d_\tau \varepsilon_i + p_i = \frac{\mu_i}{\mu}(m^2(\tau - \tau_-) + (p - p_-)), \quad i = 1, ..., N. \tag{1.40}$$

Consequently, (1.40) reads, taking advantage of the identities $\varepsilon_i = p_i\tau/(\gamma_i - 1)$ for all $i = 1, ..., N$:

$$\frac{1}{\gamma_i - 1}\tau d_\tau p_i + \frac{\gamma_i}{\gamma_i - 1}p_i - \frac{\mu_i}{\mu}\sum_{j=1}^{N} p_j = -\frac{\mu_i}{\mu}\sum_{j=1}^{N}(p_j)_- + \frac{\mu_i}{\mu}m^2(\tau - \tau_-). \tag{1.41}$$

These equations then strongly suggest to consider the admissible change of variable $x = -\ln(\frac{\tau}{\tau_-})$ so as to recast (1.41) under the form of a linear system of ODE in the x variable:

$$\frac{1}{\gamma_i - 1}d_x p_i - \frac{\gamma_i}{\gamma_i - 1}p_i + \frac{\mu_i}{\mu}\sum_{j=1}^{N} p_j = \frac{\mu_i}{\mu}\sum_{j=1}^{N}(p_j)_- - \frac{\mu_i}{\mu}m^2\tau_- \exp(-x)(1 - \exp(x)).$$

Next invoking the definition (1.26) of the dimensionless variables $\{\theta_i\}_{i=1,...,N}$, we easily get the equivalent form:

$$\begin{cases} d_x\theta_i - \gamma_i\theta_i + \frac{\mu_i}{\mu}\sum_{j=1}^{N}(\gamma_j - 1)\theta_j = \alpha\frac{\beta_i}{\gamma_i - 1} + \frac{\mu_i}{\mu}(1 - \exp(-x)), \quad i = 1, ..., N, \\ \theta_i(x = 0) = 0, \end{cases} \tag{1.42}$$

which is nothing but the linear ODE problem (1.28).

1.3.3 Solvability of the algebraic problem

In this section we prove that the algebraic problem (1.36) can be solved for all $\alpha \in [0, 1]$. Notice that for all these values of α, $x = 0$ is always solution. But as soon as $\alpha < 1$, Proposition 3 gives the existence of values of the covolume $\tau(\xi) < \tau_-$ for ξ sufficiently negative. Since the orbit $\mathcal{O}^-$ has been proved to be necessarily compressive, this obviously discards $x = 0$ (*i.e.* $\tau(0) = \tau_-$) to be the solution of interest when $\alpha < 1$. By contrast, $x = 0$ is expected to be the admissible solution when considering $\alpha = 1$. Depending on the left state $\mathbf{u}_-$, the algebraic problem (1.36) may admit up to two distinct positive solutions. According to Lemma 2, we are led to select the smallest positive one, say $x(\alpha)$, when $\alpha < 1$ and we prove below that this selection procedure satisfies the positiveness assumption (1.37) stated in Lemma 2. We shall have thus proved that for any given state $\mathbf{u}_-$ and speed σ prescribed according to (1.17), there exists an unique heteroclinic solution, namely $\mathcal{O}^-$, connecting in the future the state $\mathbf{u}_+$ given in (1.38).

Proposition 4 *Let be given $\alpha \in [0, 1[$, then there exists at least one positive root of $\psi(x, \alpha) = 0$. Denoting $x(\alpha)$ the smallest positive root, we have for all $x \in [0, x(\alpha)]$:*

$$\theta_i(x, \alpha) \geq 0, \quad i = 1, ..., N. \tag{1.43}$$

The next statement proves that the above derivation principle allows for defining a continuous function $\alpha \to x(\alpha)$ on the interval $[0, 1]$:

Proposition 5 *The function $\alpha \to x(\alpha)$ satisfies the following transversality considition:*

$$\{\partial_x\psi\}(x(\alpha), \alpha) > 0, \quad \alpha \in [0, 1). \tag{1.44}$$

The mapping $\alpha \to x(\alpha)$ is thus continuous on the interval $[0, 1[$ with moreover

$$\lim_{\alpha \to 1^-} x(\alpha) = 0.$$

Let us first prove Proposition 4.

Proof Let us first fix $\alpha \in [0,1[$ and prove that $\frac{\partial \psi}{\partial x}(0,\alpha) < 0$. Since $\psi(0,\alpha) = 0$, we shall have $\psi(x,\alpha) < 0$ for $x > 0$, x sufficiently small. We shall then establish that $\psi(x,\alpha)$ cannot keep negative values for all $x > 0$. This property will thus prove the existence of a first positive root of $\psi(x,\alpha) = 0$.
The proof heavily relies on the observation that the ODE system (1.42) naturally writes as follows:

$$\frac{d\theta_i}{dx} - \gamma_i\theta_i = \alpha\frac{\beta_i}{\gamma_i - 1} - \frac{\mu_i}{\mu}\psi(x,\alpha), \quad i = 1,...,N. \tag{1.45}$$

Let us thus begin by proving that $\frac{\partial \psi}{\partial x}(0,\alpha) < 0$. A direct calculation gives:

$$\frac{\partial \psi}{\partial x}(0,\alpha) = \sum_{i=1}^{N}(\gamma_i - 1)\frac{\partial \theta_i}{\partial x}(0,\alpha) - 1,$$

so that, in view of (1.45) and since $\psi(0,\alpha) = 0$, $\theta_i(0,\alpha) = 0$ for all $i = 1,...,N$, we have:

$$\frac{\partial \theta_i}{\partial x}(0,\alpha) = \alpha\frac{\beta_i}{\gamma_i - 1},$$

with little abuse in the notations. As a consequence, since $\sum_{i=1}^{N}\beta_i = 1$, we get the required estimate for all $0 \leq \alpha < 1$:

$$\frac{\partial \psi}{\partial x}(0,\alpha) = \alpha(\sum_{i=1}^{N}\beta_i) - 1 = \alpha - 1 < 0.$$

Next, in order to prove that the mapping $x \to \psi(x,\alpha)$ admits at least one positive root for all $0 \leq \alpha < 1$, let us assume that $\psi(x,\alpha) < 0$ for all $x \in I\!R^{+,\star}$ to finally rise a contradiction. Under this assumption, the equivalent form (1.45) of the ODE system immediately implies the following inequality for all $i = 1,...,N$:

$$\frac{d\theta_i}{dx} - \gamma_i\theta_i > \alpha\frac{\beta_i}{\gamma_i - 1},$$

so that we get:

$$\frac{d}{dx}(\theta_i \exp(-\gamma_i x)) > \alpha\frac{\beta_i}{\gamma_i - 1}\exp(-\gamma_i x).$$

Integrating from 0 to x thus yields the positiveness properties:

$$\theta_i(x,\alpha) > \alpha\frac{\beta_i}{\gamma_i(\gamma_i - 1)}(\exp(\gamma_i x) - 1) \geq 0, \tag{1.46}$$

for all $\alpha \in [0,1)$. To rise the contradiction for all $\alpha \in [0,1)$ (including $\alpha = 0$), let us establish the next identity:

$$\sum_{i=1}^{N}\theta_i(x,\alpha) = \exp(x)[1 - \exp(-x)][\alpha\sum_{i=1}^{N}(\frac{\beta_i}{\gamma_i - 1}) + \frac{1}{2}(1 - \exp(-x))]. \tag{1.47}$$

This follows by summing the N equations in (1.45) to obtain the ODE:

$$\frac{d}{dx}(\sum_{i=1}^{N}\theta_i) - (\sum_{i=1}^{N}\theta_i) = \alpha\sum_{i=1}^{N}\frac{\beta_i}{\gamma_i - 1} + (1 - \exp(-x)),$$

which once integrated from 0 to x gives (1.47). Then, the positiveness property (1.46) of each of the θ_i easily implies the following inequality:

$$(\gamma_m - 1)\sum_{i=1}^{N}\theta_i(x,\alpha) \leq \sum_{i=1}^{N}(\gamma_i - 1)\theta_i(x,\alpha), \tag{1.48}$$

where $\gamma_m = \min_{i=1,\dots,N}\gamma_i > 1$. Invoking (1.47), this inequality implies the following lower bound, valid for all $\alpha \in [0,1)$:

$$\begin{aligned}\tfrac{1}{2}[1-\exp(-x)][2\alpha(\gamma_m - 1)\sum_{i=1}^{N}(\tfrac{\beta_i}{\gamma_i - 1})\exp(x) + (\gamma_m - 1)\exp(x) - (\gamma_m + 1)]\\ \leq \psi(x,\alpha).\end{aligned} \tag{1.49}$$

To finally rise the contradiction, it suffices to notice that the left hand side in (1.49) becomes positive for x sufficiently large. As a consequence, there exists necessarily at least one positive root and the function $\alpha \to x(\alpha)$ is well-defined for all $\alpha \in [0,1)$.
To conclude, let us choose $\alpha \in [0,1)$ so that for all x in the non empty interval $(0, x(\alpha))$, we have by construction $\psi(x,\alpha) < 0$. Inequalities (1.46) are still valid and yield the expected estimates (1.43).

We now establish the required continuity property.

Proof Let us first derive the transversality result (1.44). In that aim, let us observe that the function $x \to \psi(x,\alpha)$ is $\mathcal{C}^\infty(\mathbb{R}^+)$ since is $x \to \theta(x,\alpha)$, and let us evaluate the x partial derivative of this mapping on $x(\alpha)$:

$$\frac{\partial\psi}{\partial x}(x(\alpha),\alpha) = \sum_{i=1}^{N}(\gamma_i - 1)\frac{\partial\theta_i}{\partial x}(x(\alpha),\alpha) - \exp(-x(\alpha)).$$

Let us then recall that the ODE system (1.42) reads equivalently:

$$(\frac{d\theta_i}{dx} - \gamma_i\theta_i)(x(\alpha)) = \alpha\frac{\beta_i}{\gamma_i - 1} - \frac{\mu_i}{\mu}\psi(x(\alpha),\alpha),$$

so that, with a little abuse in the notations:

$$\frac{\partial\theta_i}{\partial x}(x(\alpha),\alpha) = \alpha\frac{\beta_i}{\gamma_i - 1} + \gamma_i\theta_i(x(\alpha),\alpha), \quad i = 1,\dots,N.$$

Invoking now the definitions (1.26) and (1.30) for θ_i and β_i, $i = 1,\dots,N$, we have:

$$\begin{aligned}\frac{\partial\psi}{\partial x}(x(\alpha),\alpha) &= \alpha(\sum_{i=1}^{N}\beta_i) + \sum_{i=1}^{N}\gamma_i(\gamma_i - 1)\theta_i(x(\alpha),\alpha) - \exp(-x(\alpha))\\ &= \sum_{i=1}^{N}\frac{\gamma_i p_i(x(\alpha),\alpha)}{m^2\tau_-} - \exp(-x(\alpha)) + (\alpha - \frac{c_-^2}{m^2\tau_-^2}),\end{aligned}$$

where the last term vanishes by the definition of α. Using the identity $\exp(-x(\alpha)) = \tau(\alpha)/\tau_-$, we arrive at the following convenient form:

$$\frac{\partial \psi}{\partial x}(x(\alpha), \alpha) = \exp(-x(\alpha))(\{\frac{\sum_{i=1}^N \gamma_i p_i}{m^2 \tau}\}(x(\alpha), \alpha) - 1).$$

But on $x(\alpha)$, that is to say on the exit state $\mathbf{u}_+(\alpha)$ of the heteroclinic solution connecting $\mathbf{u}_-$ in the past for the relative mach number $\alpha^{-1/2}$, we necessarily have:

$$\{\frac{m^2 \tau}{\sum_{i=1}^N \gamma_i p_i}\}(x(\alpha), \alpha) = \frac{m^2}{(\rho c)_+^2} < 1,$$

as soon as $\alpha < 1$. This estimate is nothing but the Lax inequality already put forward in (1.19). We have therefore proved the required inequality: namely $\frac{\partial \psi}{\partial x}(x(\alpha), \alpha) > 0$ for all $\alpha \in [0, 1)$. Noticing that the mapping $(x, \alpha) \in I\!R^+ \times [0, 1) \to \psi(x, \alpha)$ achieves the same smoothness as the mappings $\theta_i(x, \alpha)$, the function $\alpha \to x(\alpha)$ is necessarily continuous on $[0, 1)$ in view of the transversality result (1.44).
Let us next prove that this function admits a limit as α goes to 1, and that this limit is necessarily null. In that aim, let be given $\alpha < 1$ so that for all $x \in [0, x(\alpha)]$, $\psi(x, \alpha) \leq 0$. For all these values of x, we observe that the estimates (1.46) are valid, so that the next lower bound holds:

$$g(x, \alpha) := \alpha \sum_{i=1}^N \frac{\beta_i}{\gamma_i}(\exp(\gamma_i x) - 1) - (1 - \exp(-x)) \leq \psi(x, \alpha). \tag{1.50}$$

But, direct calculations prove that the function $g(x, \alpha)$ in strictly convex in the x variable ($\alpha > 0$) with $g(0, \alpha) = 0$, while $\partial_x g(0, \alpha) = \alpha - 1 \leq 0$. As a consequence, this function admits exactly two zero located in 0 and $\overline{x}(\alpha)$. Clearly, $\alpha \to \overline{x}(\alpha)$ is continuous on $(0, 1]$ with $\lim_{\alpha \to 1^-} \overline{x}(\alpha) = 0$. Indeed, observe that $\partial_x g(0, 1) = 0$. Equipped with this auxiliary function, we necessarily deduce from the estimate (1.50) and the definition of $x(\alpha)$ (*i.e.* $\psi(x, \alpha) < 0$ for $x \in (0, x(\alpha))$) the following inequality:

$$x(\alpha) \leq \overline{x}(\alpha), \quad \text{for all } \alpha \in (0, 1].$$

This completes the proof.

1.3.4 Additional properties

This section aims at establishing further properties of the function $x(\alpha)$ for $\alpha \in [0, 1]$, which will be of primary importance in the forthcoming analysis of the Riemann problem. All the proofs are postponed at the end of this section. Here, the main result is:

Proposition 6 *The function $\alpha \to x(\alpha)$ is continuously differentiable on $(0, 1)$ and decreases in the large. More precisely, there exists α_c, $0 \leq \alpha_c < 1$, such that $x(\alpha)$ is strictly decreasing on $]\alpha_c, 1[$ while necessarily $x(\alpha) = x(0)$ for all $0 \leq \alpha \leq \alpha_c$.*
In other words, the compression rate τ_-/τ_+ grows monotonically with the relative Mach number. One would have expected a strict monotonicity property but existence and uniqueness of the Riemann solution will follow without the need for proving that

$\alpha_c = 0$. This might not seem too surprising since Smith has proved uniqueness in the setting $N = 1$ provided that the compression rate does not depart too much from monotonicity (see the medium condition [129]). The next result states that the compression rate stays bounded in the regime of an infinite relative Mach number ($\alpha = 0$). More precisely :

Lemma 3 *Let us define* $\gamma_m = \min_{i=1,...,N} \gamma_i$ *and* $\gamma_M = \max_{i=1,...,N} \gamma_i$. *Then, the mapping* $\alpha \to x(\alpha)$ *obeys the following estimates on the maximal compression rate :*

$$\frac{\gamma_m - 1}{\gamma_m + 1} \leq \min_{\alpha \in [0,1]} \exp(-x(\alpha)) = \exp(-x(0)) \leq \frac{\gamma_M - 1}{\gamma_M + 1}. \tag{1.51}$$

Let us notice that the bounds (1.51) are optimal since when considering identical adiabatic exponents, *i.e.* $\gamma = \gamma_i$ for all $i = 1, ..., N$, these inequalities express nothing but the identity :

$$\exp(-x(0)) = \frac{\gamma - 1}{\gamma + 1}, \tag{1.52}$$

which is well-known in the usual setting of a single pressure law ($N = 1$). The proof of Lemma 3 will easily yield under the assumption of N identical adiabatic exponents :

$$\exp(-x(\alpha)) = \frac{\gamma - 1}{\gamma + 1}(1 + 2\alpha \sum_{i=1}^{N} \frac{\beta_i}{\gamma_i - 1}),$$

(see indeed (1.58) below), so that in this simplified setting, $\alpha \to x(\alpha)$ is strictly decreasing. Such a property naturally suggests the last result of this section :

Proposition 7 *Let us assume that* $(\gamma_M - \gamma_m)$ *is sufficiently small. Then, the mapping* $\alpha \to x(\alpha)$ *is strictly decreasing on* $[0, 1]$.

We now prove Proposition 6.

Proof Let us first observe that the mapping $\alpha \to x(\alpha)$ is smooth on $(0, 1)$ as a by product of the implicit functions theorem. Indeed the transversality result (1.44) we have just proved shows that $\frac{\partial \psi}{\partial x}(x, \alpha) > 0$ for $|x - x(\alpha)|$ sufficiently small and $\alpha \in (0, 1)$.

Let us now turn establishing the expected monotonicity property. The proof will rely on the equivalent formulation (1.36) of the algebraic problem to be solved, namely :

$$\psi(x, \alpha) = \alpha\phi(x) + \psi(x, 0) = 0. \tag{1.53}$$

The smoothness of the function $\alpha \to \psi(x(\alpha), \alpha)$ for $\alpha \in (0, 1)$ allows to consider :

$$\begin{aligned} \frac{d}{d\alpha}\psi(x(\alpha), \alpha) &= \phi(x(\alpha)) + \frac{\partial \psi}{\partial x}(x(\alpha), \alpha) \times x'(\alpha) \\ &= -\frac{1}{\alpha}\psi(x(\alpha), 0) + \frac{\partial \psi}{\partial x}(x(\alpha), \alpha) \times x'(\alpha), \end{aligned}$$

since by definition, $x(\alpha)$ solves (1.53). We therefore get for all $\alpha \in (0, 1)$:

$$\frac{\partial \psi}{\partial x}(x(\alpha), \alpha) \times x'(\alpha) = \frac{1}{\alpha}\psi(x(\alpha), 0). \tag{1.54}$$

Let us recall that $\frac{\partial\psi}{\partial x}(x(\alpha),\alpha) > 0$. As a consequence, the expected decreasing property holds for all α such that $x(\alpha) < x(0)$ since we necessarily have $\psi(x(\alpha), 0) < 0$. Indeed, recall that $x(0)$ is the first positive root of $\psi(x, 0) = 0$. Let us now prove that there exists an non empty interval $(\alpha_c, 1)$ where $x(\alpha) < x(0)$. The existence of such a non trivial interval follows from $x(0) > 0 = x(1)$ and the continuity of the function $\alpha \to x(\alpha)$ on $[0, 1]$. Let us then consider the largest open interval $(\alpha_c, 1)$ where $x(\alpha) < x(0)$, *i.e.* with $x(\alpha_c) = x(0)$. We have just proved that the function $\alpha \to x(\alpha)$ strictly decreases on this maximal interval and maps $[\alpha_c, 1]$ onto $[x(1) = 0, x(0)]$.
Let us assume that $\alpha_c > 0$ and prove that necessarily $x(\alpha) = x(0)$ for all $\alpha \in [0, \alpha_c]$. In that aim, observe first that the identity $x(\alpha_c) = x(0)$ implies, since $x(\alpha_c)$ solves the algebraic problem (1.53) :

$$\psi(x(\alpha_c), \alpha_c) = \alpha_c\phi(x(0)) + \psi(x(0), 0) = 0. \tag{1.55}$$

But by construction, $\psi(x(0), 0) = 0$ so that $\alpha_c > 0$ gives in addition that $\phi(x(0)) = 0$. These two identities necessarily make $x(0)$ to be a positive solution of the algebraic problem (1.53) for any given value of $\alpha \in (0, 1)$. We therefore infer that :

$$x(\alpha) \leq x(0), \quad \text{for all } \alpha \in (0, \alpha_c), \tag{1.56}$$

since again by construction $x(\alpha)$ is the first positive solution of (1.53). To conclude, let us assume $x(\alpha) < x(0)$ for some given $\alpha \in (0, \alpha_c)$ to rise a contradiction. To this purpose, recall that the function $\alpha \to x(\alpha)$ have been shown to map $[\alpha_c, 1]$ onto $[x(1) = 0, x(0)]$, so that its continuity proves for existence $\underline{\alpha} \in (\alpha_c, 1)$ such that $x(\underline{\alpha}) = x(\alpha)$. Let us then prove that necessarily $\phi(x(\underline{\alpha})) = 0$. This follows from :

$$\psi(x(\underline{\alpha}), \underline{\alpha}) = \underline{\alpha}\phi(x(\underline{\alpha})) + \psi(x(\underline{\alpha}), 0) = \alpha\phi(x(\underline{\alpha})) + \psi(x(\underline{\alpha}), 0) = \psi(x(\alpha), \alpha) = 0,$$

with $\alpha < \underline{\alpha}$. But $\phi(x(\underline{\alpha})) = 0$ implies in turn that $\psi(x(\underline{\alpha}), 0) = 0$. Therefore $x(\underline{\alpha})$ solves $\phi(x, 0) = 0$ so that $x(\underline{\alpha}) > x(0)$ since $x(0)$ is the first positive root. Hence we would arrive at $x(\alpha) = x(\underline{\alpha}) > x(0)$ and this estimate is in contradiction with (1.56). We have just proved that necessarily $x(\alpha) = x(0)$ for all $\alpha \in [0, \alpha_c]$. This concludes the proof.

Let us then prove Lemma 3.

Proof Let us choose $\alpha \in [0, 1[$. Arguments, identical to those developed in the proof of Proposition 4, ensure for validity the estimates (1.48) et (1.49) for all $x \in [0, x(\alpha)]$. Plugging $x(\alpha)$ in (1.49) yields :

$$\begin{aligned}\tfrac{1}{2}[1 - \exp(-x(\alpha))][2\alpha(\gamma_m - 1)\sum_{i=1}^{N}(\tfrac{\beta_i}{\gamma_i - 1})\exp(x(\alpha)) + (\gamma_m - 1)\exp(x(\alpha)) - (\gamma_m + 1)]\\ \leq \psi(x(\alpha), \alpha) = 0.\end{aligned} \tag{1.57}$$

Symmetrical steps apply to prove the next inequality :

$$\begin{aligned}\psi(x(\alpha), \alpha) = 0 \leq\\ \tfrac{1}{2}[1 - \exp(-x(\alpha))][2\alpha(\gamma_M - 1)\sum_{i=1}^{N}(\tfrac{\beta_i}{\gamma_i - 1})\exp(x(\alpha)) + (\gamma_M - 1)\exp(x(\alpha)) - (\gamma_M + 1),]\end{aligned}$$

where have set $\gamma_M = \max_{i=1,\ldots,N} \gamma_i$. These inequalities easily yield the following rough estimates:

$$\frac{\gamma_m - 1}{\gamma_m + 1}(1 + 2\alpha \sum_{i=1}^{N} \frac{\beta_i}{\gamma_i - 1}) \leq \exp(-x(\alpha)) \leq \frac{\gamma_M - 1}{\gamma_M + 1}(1 + 2\alpha \sum_{i=1}^{N} \frac{\beta_i}{\gamma_i - 1}), \tag{1.58}$$

so that we have:

$$\frac{\gamma_m - 1}{\gamma_m + 1} \leq \min_{\alpha \in [0,1]} \exp(-x(\alpha)) \leq \frac{\gamma_M - 1}{\gamma_M + 1}.$$

This concludes the proof.

To conclude, let us establish Proposition 7.

Proof Let us consider without restriction that $\gamma_m = \gamma_N$. Then, let us observe from the equation (1.47), established in the course of the proof of Proposition 4 that:

$$\theta_N(x, \alpha) = \exp(x)[1 - \exp(-x)][\alpha \sum_{i=1}^{N} (\frac{\beta_i}{\gamma_i - 1}) + \frac{1}{2}(1 - \exp(-x))] - \sum_{i=1}^{N-1} \theta_i(x, \alpha). \tag{1.59}$$

Let us use this identity to reexpress the algebraic problem (1.36) to be solved as follows:

$$\psi(x, \alpha) = \sum_{i=1}^{N-1} (\gamma_i - \gamma_N)\theta_i(x, \alpha) - T_N(x, \alpha) = 0, \tag{1.60}$$

where we have set:

$$T_N(x, \alpha) = \frac{(\gamma_N + 1)}{2}[1 - \exp(-x)][1 - \frac{\gamma_N - 1}{\gamma_N + 1}(1 + 2\alpha \sum_{i=1}^{N} \frac{\beta_i}{\gamma_i - 1}) \exp(x)]. \tag{1.61}$$

To get a convenient expression for the derivative $x'(\alpha)$, let us differentiate with respect to α the formulation (1.60) when evaluated on $x(\alpha)$:

$$\begin{gathered} \partial_x \psi(x(\alpha), \alpha) \times x'(\alpha) = \\ \frac{\partial T_N}{\partial \alpha}(x(\alpha), \alpha) - \sum_{i=1}^{N-1} (\gamma_i - \gamma_N) \frac{\partial \theta_i}{\partial \alpha}(x(\alpha), \alpha). \end{gathered} \tag{1.62}$$

Easy calculations give the following estimate valid for all $x > 0$:

$$\begin{aligned} \frac{\partial T_N}{\partial \alpha}(x, \alpha) &= \frac{1}{2}(\gamma_N + 1)[1 - \exp(-x)][-2\frac{\gamma_N - 1}{\gamma_N + 1} \sum_{i=1}^{N} \frac{\beta_i}{\gamma_i - 1}] \exp(x) \\ &= -(\gamma_N - 1) \exp(x)[1 - \exp(-x)] \sum_{i=1}^{N} (\frac{\beta_i}{\gamma_i - 1}) < 0. \end{aligned} \tag{1.63}$$

The right hand side of (1.62) thus clearly suggests that $x'(\alpha)$ actually stays negative for sufficiently small $|\gamma_i - \gamma_N|$, since we have already proved $\partial_x \psi(x(\alpha), \alpha) > 0$ for all $\alpha \in (0, 1)$. To get this result, let us first recall that for any given set of adiabatic exponents $\{\gamma_i\}_{i=1,\ldots,N}$ we have proved the existence of a value $\alpha_c < 1$ such that $x(\alpha)$ strictly decreases for $\alpha \in (\alpha_c, 1]$. Note that α_c depends continuously on the parameters

γ_i entering the definition of the ODE problem (1.28) since does its solution given by the Duhamel formula and hence the mapping $x(\alpha, \{\gamma_i\}_{i=1,...,N})$. The adiabatic exponent γ_N being fixed, there necessarily exists some $\epsilon > 0$ solely depending on γ_N such that $\alpha_c < 1 - \epsilon$ provided that $|\gamma_i - \gamma_N|$ are small enough. This would otherwise rise a contradiction with the property that $\alpha_c = 0$ when the N adiabatic exponents are identical. Consequently, we just have to prove $x'(\alpha) < 0$ for all $\alpha \leq 1 - \epsilon$. Since $x(\alpha)$ decreases in the large, note that $x(\alpha) > x(1-\epsilon) > 0$ so that $\frac{\partial T_N}{\partial \alpha}(x(\alpha), \alpha) > \epsilon_0 > 0$ for all the α under consideration. Here ϵ_0 solely depends on ϵ.
Next, let us use the Duhamel representation formula (1.32) to obtain:

$$\frac{\partial \mathbf{\Theta}}{\partial \alpha}(x(\alpha), \alpha) = \int_0^{x(\alpha)} \exp[\mathcal{M}(x(\alpha) - y)]dy \ \mathbf{a},$$

where $\mathbf{\Theta} = \{\theta_i\}_{i=1,...,N}$. We thus deduce the following crude estimate, valid for all $\alpha \in [0,1]$:

$$\begin{aligned} \max_{1 \leq i \leq N} |\frac{\partial \theta_i}{\partial \alpha}(x(\alpha), \alpha)| &\leq ||\int_0^{x(\alpha)} \exp[\mathcal{M}(x(\alpha) - y)]dy||_\infty \times ||\mathbf{a}||_\infty \\ &\leq \sup_{x \in [0, x(0)]} ||\int_0^{x} \exp[\mathcal{M}(x - y)]dy||_\infty \times ||\mathbf{a}||_\infty, \end{aligned}$$

with classical notations for the involved vector and matrix norms. Since $x(0)$ is known to be bounded in view of Lemma 3, we thus deduce that $|\sum_{i=1}^{N-1}(\gamma_i - \gamma_N)\frac{\partial \theta_i}{\partial \alpha}(x(\alpha), \alpha)|$ can be made smaller than ϵ_0 when choosing the γ_i sufficiently close to γ_N. This concludes the proof.

1.4 The asymptotic regime and shock solutions

The existence and uniqueness result of Theorem 1 for travelling wave solutions of (1.2) allows us in this section to tackle the asymptotic regime of a vanishing parameter ϵ in (1.1) when defining the notion of shock solutions for the limit system.

1.4.1 Shock solutions

We first deal with shock solutions. Then, rephrazing classical considerations (see Smoller [130] but also LeFloch [90]), starting from a given travelling wave solution $\mathbf{u}(\xi)$ of (1.2) connecting with speed σ a state $\mathbf{u}_-$ to $\mathbf{u}_+$ and a given fixed $\epsilon > 0$, the rescaled function $\mathbf{u}_\epsilon(\xi) = \mathbf{u}(\xi/\epsilon)$ is nothing else a travelling wave solution of (1.1) (*i.e.* with rescaled viscosity coefficents $\epsilon\mu_i$ for $i = 1, ..., N$) connecting the same two states $\mathbf{u}_-$ and $\mathbf{u}_+$ and propagating at the same speed σ. Next and since the proposed rescaling does not affect the (bounded) total variation of $\mathbf{u}$, *i.e.* $TV(\mathbf{u}_\epsilon) = TV(\mathbf{u})$ for all $\epsilon > 0$, the family of travelling wave solutions $\{\mathbf{u}_\epsilon\}_{\epsilon>0}$ is easily seen to strongly converge in L^1_{loc} as ϵ goes to zero to the following step function:

$$\mathbf{u}(x,t) = \begin{cases} \mathbf{u}_- & \text{if} \quad x < \sigma t, \\ \mathbf{u}_+ & \text{if} \quad x > \sigma t. \end{cases} \tag{1.64}$$

Following LeFloch [90] (see also Raviart-Sainsaulieu [112]), these considerations thus motivate a natural notion of shock solutions for the system (1.1) in the limit $\epsilon = 0$, referred hereafter as to the limit system.

Definition 1 *The step function (1.64) is said to be a* 1*-shock solution of the limit system if and only if*

$$m = \rho_-(u_- - \sigma) > (\rho c)_-, \tag{1.65}$$

and a 3*-shock solution if and only if*

$$m = \rho_+(u_+ - \sigma) < -(\rho c)_+. \tag{1.66}$$

Let us recall that inequality (1.65) or (1.66) is indeed a necessary (and sufficient) condition for the existence of a travelling wave solution and thus of the limit function under consideration. The following properties for shock solutions are directly inherited from the underlying travelling waves:

Corollary 1 *By construction, a* 1*-shock solution obeys the following Lax shock conditions:*

$$u_+ - c_+ < \sigma < u_- - c_-,$$

and the next N inequalities on the specific entropies:

$$(s_i)_+ \geq (s_i)_- \text{ for } i = 1, ..., N, \tag{1.67}$$

these inequalities being strict for all index i such that $\mu_i > 0$.
Similarly, a 3*-shock solution satisfies:*

$$u_+ + c_+ < \sigma < u_- + c_-,$$

together with for all index $i = 1, ..., N$:

$$(s_i)_+ \leq (s_i)_-, \tag{1.68}$$

with strict inequality as soon as $\mu_i > 0$.
In addition, both types of shock solutions necessarily obey the following Rankine-Hugoniot conditions:

$$-\sigma[\rho] + [\rho u] = 0, \tag{1.69}$$

$$-\sigma[\rho u] + [\rho u^2 + \sum_{i=1}^{N} p_i] = 0, \tag{1.70}$$

together with the additional jump relation:

$$-\sigma[\rho E] + [(\rho E + \sum_{i=1}^{N} p_i)u] = 0, \tag{1.71}$$

which equivalently takes the form of the next Hugoniot-like equation:

$$(\varepsilon_+ - \varepsilon_-) + \frac{1}{2}(p_+ + p_-)(\tau_+ - \tau_-) = 0. \tag{1.72}$$

Proof Smooth solutions of (1.2) are known to obey the additional conservation law (1.5) for governing the total energy, hence also a travelling wave solution. Then integrating the associated ODE form of (1.5) for $\xi \in \mathbb{R}$ just gives (1.71). Next, arguments from the classical 3×3 Euler setting exactly apply to infer from equation (1.71) on the ground of (1.69) and (1.70):

$$m\{(\varepsilon_+ - \varepsilon_-) + \frac{1}{2}(p_+ + p_-)(\tau_+ - \tau_-)\} = 0,$$

when understanding p as a (total) pressure and ε as a (total) internal energy. We just get the Hugoniot equation since m cannot be zero.

The reported properties stand as direct extensions of the inequalities met by the shock solutions of the usual 3×3 Euler setting (*i.e.* with $N = 1$). But by deep contrast with this purely conservative setting, a central discrepancy arises from the fact that shock solutions, in the present extended framework, do by definition heavily depend on the choice of the viscosity coefficients μ_i for $i = 1, ..., N$ who gave birth to the underlying travelling wave solutions. Recall that up to this stage, all these coefficients have been tacitly assumed to be fixed but actually arbitrary non negative values can be prescribed under the requirement of a positive sum. The form of the linear reduced ODE problem (1.28)-(1.29) clearly shows that its solution entirely depends on the ratios μ_i/μ for $i = 1, ..., N$ (the adiabatic exponents γ_i being fixed) and so does necessarily the solution of the algebraic equation (1.36). In other words, for a given state $\mathbf{u}_-$ and a velocity σ prescribed under the Lax condition (1.17), the exit state $\mathbf{u}_+$ entering the definition of the step function (1.64) generally differs when considering distinct viscosity ratios. With this respect, shock solutions for the limit system under consideration are sensitive to the small scale properties of a particular diffusive regularization. Such a sensitiveness is actually classical in the setting of hyperbolic systems in non conservation form (see [112], [10], ...) and was thus expected to take place within the present frame. The travelling wave analysis we have proposed actually allows for precisely encoding this small scale sensitiveness when explicitly defining the exact jump in each of the specific entropies in place of the rather vague inequalities (1.67) and (1.68). Indeed, an arbitrary set of non negative viscosities with positive sum being prescribed, then the uniqueness of the exit state $\mathbf{u}_+$ understood as a function of a given left state $\mathbf{u}_-$ and a relevant speed of propagation σ allows us to give the following definition.

Definition 2 *Let be given fixed viscosities $\mu_i \geq 0$ for $i = 1, ..., N$ with $\sum_{i=1}^N \mu_i > 0$. Then there exists N associated non negative smooth functions κ_i of $\alpha \in [0, 1]$ and $\beta \in \mathcal{B}_N^+$, such that for the 1-shock solution connecting $\mathbf{u}_-$ to $\mathbf{u}_+$ with speed σ:*

$$s_i(\mathbf{u}_+) - s_i(\mathbf{u}_-) = \kappa_i(\alpha(\mathbf{u}_-, \sigma), \beta(\mathbf{u}_-)), \quad i = 1, ..., N, \tag{1.73}$$

while considering the 3-shock solution which connects $\mathbf{u}_-$ to $\mathbf{u}_+$ with speed σ, one has:

$$s_i(\mathbf{u}_-) - s_i(\mathbf{u}_+) = \kappa_i(\alpha(\mathbf{u}_+, \sigma), \beta(\mathbf{u}_+)), \quad i = 1, ..., N. \tag{1.74}$$

The N functions we have introduced will be referred as to kinetic functions after [15] (see also LeFloch [88]). The N relations in (1.73) obviously serve as a definition formula for each of the kinetic functions, the N identities entering (1.74) being then

a direct consequence of the independence property of the reduced ODE system (1.28) with respect to the sign of the relative Mach number (see Section 1.3).
The following statement summarizes the generalized jump conditions satisfied by a shock solution.

Lemma 4 *Let be given a fixed set of viscosities with positive sum and consider the N associated kinetic functions κ_i. Then a shock solution in the sense of Definition 1 solves the following generalized jump conditions :*

$$\begin{aligned}&-\sigma(\rho_+ - \rho_-) + ((\rho u)_+ - (\rho u)_-) = 0,\\&-\sigma((\rho u)_+ - (\rho u)_-) + ((\rho u^2 + p)_+ - (\rho u^2 + p)_-) = 0,\\&-\sigma((\rho E)_+ - (\rho E)_-) + ((\rho E + p)u_+ - (\rho E + p)u_-) = 0,\\&-\sigma((\rho s_i)_+ - (\rho s_i)_-) + ((\rho s_i u)_+ - (\rho s_i u)_-) = |\rho_-(u_- - \sigma)|\kappa_i(\alpha, \beta), \quad i = 1, ..., N-1,\end{aligned} \tag{1.75}$$

together with :

$$-\sigma((\rho s_N)_+ - (\rho s_N)_-) + ((\rho s_N u)_+ - (\rho s_N u)_-) = |\rho_-(u_- - \sigma)|\kappa_N(\alpha, \beta). \tag{1.76}$$

To conclude notice that such kinetic functions are solely defined when $\alpha \in [0, 1]$ but with the property that $\kappa_i(1, \beta) = 0$ for all reduced vector $\beta \in \mathbf{B}$. On the ground of this observation, we choose to extend the domain of definition of these N kinetic functions for all values of $\alpha \in I\!R^+$ when setting by convention for all $\beta \in \mathbf{B}$:

$$\kappa_i(\alpha, \beta) = 0 \text{ for all } \alpha > 1, \quad i = 1, ..., N. \tag{1.77}$$

Such a convention is indeed natural because of the following next result :

Proposition 8 *Let us consider a step function (1.64) with a non zero relative mass flux $m = \rho_-(u_- - \sigma) \in (-(\rho c)_+, (\rho c)_-)$ (i.e. with $\alpha > 1$). Then, this step function cannot solve the generalized jump conditions (1.75)-(1.76) unless $\mathbf{u}_+ = \mathbf{u}_-$.*

In other words, there is no non trivial step functions solutions of (1.75)-(1.76) for a non zero m with $\alpha > 1$. The proof given below will highlight that the validity of (1.77) is indeed necessary to get the conclusion.

Proof By our convention, $\kappa_i(\alpha, \beta) = 0$ for all $i = 1, ..., N$ so that our assumption of a non zero relative mass flux gives necessarily $(s_i)_+ = (s_i)_-$ for all $i = 1, ..., N$. So turning considering the Rankine-Hugoniot condition on the momentum, this one easily recasts as :

$$m^2(\tau_+ - \tau_-) + \sum_{i=1}^{N} s_i(\tau_+^{-\gamma_i} - \tau_-^{-\gamma_i}) = 0.$$

Of course $\tau_+ = \tau_-$ is always solution, but let us assume that this equation can be solved for other values $\tau_+ \neq \tau_-$. Such a non trivial solution must in addition solve the Hugoniot-like equation (1.72) since $m \neq 0$ by assumption (see the proof of Corollary 1). But this Hugoniot-like condition reads :

$$\varepsilon(\tau_+, s_i) - \varepsilon(\tau_-, s_i) + \frac{1}{2}(p(\tau_+, s_i) + p(\tau_-, s_i))(\tau_+ - \tau_-) = 0,$$

where $p(\tau, s_i) = -\partial_\tau \varepsilon(\tau, s_i)$. As a consequence, the above Hugoniot-like condition reads:

$$(\tau_+ - \tau_-)[\frac{1}{2}(p(\tau_+, s_i) + p(\tau_-, s_i)) - 1/\{(\tau_+ - \tau_-)\int_{\tau_-}^{\tau_+} p(v, s_i)dv\}] = 0.$$

But by the strict convexity of the total pressure in the τ variable, the terms inside the brackets is necessarily non zero as soon as $\tau_+ \neq \tau_-$. So this equation cannot be solved for values of τ distinct from τ_-. This completes the proof.

1.4.2 Contact discontinuities

In the classical setting of a single pressure law, shocks are known to coexist with another type of discontinuity: namely the so-called contact discontinuities associated with the linearly degenerate intermediate fields. As expected, such discontinuities arise in the present framework but by contrast to shock solutions, they do not give birth to ambiguities in the limit system. Actually, they are seen below to be already natural weak solutions of parabolic regularization (1.1) of the hyperbolic system under consideration. To that purpose, let us consider the following step function propagating at constant speed u:

$$\mathbf{u}(x,t) = \begin{cases} {}^T(\rho_-, \rho_- u, \{(\rho s_i)_-\}_{i=1,...,N}), & x < u\ t, \\ {}^T(\rho_+, \rho_+ u, \{(\rho s_i)_+\}_{i=1,...,N}), & x > u\ t, \end{cases} \tag{1.78}$$

under the requirement of a constant total pressure:

$$\sum_{i=1}^{N} p_i(\tau_-, (s_i)_-) = \sum_{i=1}^{N} p_i(\tau_+, (s_i)_+). \tag{1.79}$$

The continuity of both the velocity u and the total pressure of course stems from the property of these two quantities to be Riemann invariants for the intermediate fields of the extracted hyperbolic system (1.4) (see Proposition 1 and the associated discussion). Let us then observe that this step function is a weak solution of (1.1) whatever is the rescaling parameter $\epsilon > 0$. Since the velocity u is constant, one simply has to observe that the following identities are valid in the usual sense of the distributions:

$$\begin{cases} \partial_t \rho + u\partial_x \rho = 0, \\ (\partial_t \rho u + u\partial_x \rho u) + \partial_x \sum_{i=1}^N p_i = 0, \\ \partial_t \rho s_i + u\partial_x \rho s_i = 0, \quad i = 1, ..., N. \end{cases} \tag{1.80}$$

Indeed and under the requirement of a constant total pressure, the above PDE are nothing but N transport equations at constant speed u. Since the function (1.78)–(1.79) is solution of (1.1) for all $\epsilon > 0$, it obviously remains a weak solution of the limit system. Such a solution will be called a contact discontinuity. Note that since the relative Mach flux $\rho(u - \sigma)$ coming with an arbitrary function step in the form (1.78) is identically zero, the associated reduced number α achieves its value in $\overline{I\!R}$, namely $\alpha = +\infty$. So that, the convention for extending the domain of definition of the N kinetic functions implies that these all achieve a zero value. Hence, the N last equations in (1.80) can also read under the form of the following trivial identities free from ambiguities:

$$\rho_-(u_- - \sigma)((s_i)_+ - (s_i)_-) = \rho_-(u_- - \sigma)\kappa_i(\infty, \beta), \quad i = 1, ..., N. \tag{1.81}$$

These identities just express that the jump in each of the specific entropies s_i is arbitrary. Notice in addition that contact discontinuities trivially also satisfy in the sense of the distributions the following conservation law for the total energy

$$\partial_t \rho E + \partial_x(\rho E + \sum_{i=1}^{N} p_i)u = 0, \tag{1.82}$$

since the relative mass flux $m = 0$ (see the proof of Corollary 1). Summarizing the above observation, we state:

Lemma 5 *Let be given a fixed set of viscosities with positive sum and consider the N associated kinetic functions κ_i. Then a contact discontinuity solution solves the following jump conditions:*

$$\begin{aligned}
&-\sigma(\rho_+ - \rho_-) + ((\rho u)_+ - (\rho u)_-) = 0,\\
&-\sigma((\rho u)_+ - (\rho u)_-) + ((\rho u^2 + p)_+ - (\rho u^2 + p)_-) = 0,\\
&-\sigma((\rho E)_+ - (\rho E)_-) + ((\rho E + p)u_+ - (\rho E + p)u_-) = 0,\\
&-\sigma((\rho s_i)_+ - (\rho s_i)_-) + ((\rho s_i u)_+ - (\rho s_i u)_-) = |\rho_-(u_- - \sigma)|\kappa_i(\alpha, \beta) = 0, \quad i = 1, ..., N-1,
\end{aligned} \tag{1.83}$$

together with:

$$-\sigma((\rho s_N)_+ - (\rho s_N)_-) + ((\rho s_N u)_+ - (\rho s_N u)_-) = |\rho_-(u_- - \sigma)|\kappa_N(\alpha, \beta) = 0. \tag{1.84}$$

1.4.3 The limit first order system

In this subsection, we focus ourselves on piecewise Lipschitz-continuous limit functions $\mathbf{u}$ obtained as limits of families of solutions $\{\mathbf{u}\}_{\epsilon>0}$ of (1.1) as ϵ goes to zero. On the ground of the N kinetic functions derived in subsection 1.4.1, we introduce with little abuse in the notations N bounded non negative Borel measures $\kappa_i(\mathbf{u})$ defined from any given Lipschitz-continuous function $\mathbf{u}$ as follows. In the zone of smoothness of the function $\mathbf{u}$, these measures naturally identically vanishe while across a discontinuity in $\mathbf{u}$ separating the states $\mathbf{u}_-$ and $\mathbf{u}_+$ and propagating at speed σ:

$$\kappa_i(\mathbf{u}) = \kappa_i(\alpha, \beta)\delta_{x-\sigma t}, \quad i = 1, ..., N, \tag{1.85}$$

where depending on the sign of the relative mass flux $m = \rho_-(u_- - \sigma)$, the dimensionless quantities α and $\beta = \{\beta_i\}_{i=1,\dots,N}$ are built from $\mathbf{u}_-$ $(m \geq 0)$ or $\mathbf{u}_+$ $(m < 0)$. Let us recall that by convention $\kappa_i(\alpha, \beta) = 0$ as soon as $\alpha \in [1, +\infty)$: *i.e.* when $m \in [-(\rho c)_+, (\rho c)_-]$. Consequently and when addressing contact discontinuities, *i.e.* with $m = 0$, the measures under consideration are identically zero. These measures are therefore only non trivial on shock discontinuities, and by construction they stay bounded and positive for all parameter $\alpha \in [0, 1]$. Equipped with these bounded measures, we then propose after Berthon et al. [15]:

Definition 3 *A piecewise Lipschitz-continuous funtion is said to be a solution of (1.1) in the asymptotic regime $\epsilon \to 0$, if and only if it obeys in the sense of the distributions the following first order system with measures source terms :*

$$\begin{cases} \partial_t\rho + \partial_x\rho u = 0, \\ \partial_t\rho u + \partial_x(\rho u^2 + \sum_{i=1}^N p_i) = 0, \\ \partial_t\rho E + \partial_x(\rho E + \sum_{i=1}^N p_i)u = 0 \\ \partial_t\rho s_i + \partial_x\rho s_i u = \kappa_i(\mathbf{u}), \quad i = 1, ..., N-1, \end{cases} \tag{1.86}$$

together with the additional law

$$\partial_t\rho s_N + \partial_x\rho s_N u = \kappa_N(\mathbf{u}). \tag{1.87}$$

The relevance of the proposed definition simply stems from the Lemmae 4 and 5.

1.5 The Riemann problem

This section aims at proving that the Riemann problem for the first order limit system (1.86)-(1.87) admits an unique solution for arbitrary states $\mathbf{u}_L$ and $\mathbf{u}_R$ in Ω, provided that these two states do not yield vacuum. Occurence of vacuum will be given a precise meaning hereafter, which will be seen to naturally extend the no-vacuum condition met in the setting of a single pressure law. To shade some light on the forthcoming developments, let us first comment the eigenstructure of the limit system (1.86). As it is well-known, this eigenstructure has to be considered within the frame of smooth solutions, that is to stay when all the measures identically vanishes. With this respect, the resulting system in conservation form is just equivalent to the first order underlying system (1.4) we have studied in the second section. The system (1.86) then shares exactly the same hyperbolicity properties stated in Proposition 1. Classical considerations from the respective properties of the fields under consideration indicate that the Riemann solution is at most made of four distinct states, namely $\mathbf{u}_L$, $\mathbf{u}_L^\star$, $\mathbf{u}_R^\star$ and $\mathbf{u}_R$, separated by three simple waves. The property of genuine non linearity of the two extreme fields is responsible for the property that the extreme waves are indeed simple, while the intermediate fields associated with the same eigenvalue can only result in a contact discontinuity. This waves pattern thus coincides with the one met in the case of a single pressure law. We shall thus adopt similar steps in the analysis of the Riemann solution when deriving at first a notion of shock curves and of rarefaction curves for the two extreme fields. Then, the property of a constant total pressure $p = \sum_{i=1}^N p_i$ and constant velocity u valid for the intermediate contact discontinuity will naturally suggest us to determine the projection of these two families of curves in the (p, u) plane. The respective monotonicity properties and the asymptotic behaviours of these projections will ensure existence and uniqueness of an intersection point : namely the total pressure and velocity at the expected contact discontinuity.

1.5.1 Shock curves

For any given fixed state $\mathbf{u}_L \in \Omega$, let us first define the 1-shock curve attached to the left state $\mathbf{u}_L$,

$$\mathcal{S}_1(\mathbf{u}_L) = \{\mathbf{u} \in \Omega, \mathbf{u}_L \ \underline{\text{1-shock}}\ \mathbf{u}\},$$

as the subset of Ω made of the totality of states $\mathbf{u}$ which can be connected on the right to $\mathbf{u}_L$ by an 1-shock solution in the sense of Definition 1.
Similarly and for a given state $\mathbf{u}_R \in \Omega$, let us then introduce the 3-shock curve attached to the right state $\mathbf{u}_R$,

$$\mathcal{S}_3(\mathbf{u}_R) = \{\mathbf{u} \in \Omega, \mathbf{u} \ \underline{\text{3-shock}} \ \mathbf{u}_R\},$$

made by definition of the states $\mathbf{u}$ which can be connected on the left to $\mathbf{u}_R$ by an 3-shock solution.
Motivated by the general considerations given at the beginning of this section, we propose to determine the projection of these two subsets on the (p, u) plane. We shall give a complete derivation when restricting ourselves to $\mathcal{S}_1(\mathbf{u}_L)$, the required characterization of $\mathcal{S}_3(\mathbf{u}_R)$ will follow along the same steps.
By the derivation of the family of 1-shock solutions with $\mathbf{u}_L$ as a left state, α defined in (1.30) serves as a natural parameter along $\mathcal{S}_1(\mathbf{u}_L)$:

$$\mathcal{S}_1(\mathbf{u}_L) = \{\mathbf{u}(\alpha) = (\tau(\alpha), u(\alpha), p(\alpha), \{s_i(\alpha)\}_{i=1,\dots,N-1}) \text{ with } \alpha \in [0,1]\},$$

but the following easy statement proves that the total pressure p stands for another admissible parameter, actually more relevant to our purpose :

Lemma 6 *The total pressure is a smooth strictly decreasing function of α in $[0,1]$ which maps $[0,1]$ onto $[p_L, +\infty]$.*
Proof Using the notations of the previous section, the total pressure obeys by construction along $\mathcal{S}_1(\mathbf{u}_L)$:

$$p(\alpha) = p_L + \frac{\rho_L c_L^2}{\alpha}(1 - \frac{\tau(\alpha)}{\tau_L}),$$

where the function $\tau(\alpha) = \tau_L \exp(-x(\alpha))$ has been shown to smoothly increase in the large for α in $[0,1]$ from some τ_{min} in $(0, \tau_L)$ to τ_L while being necessarily strictly increasing for α close to 1 (*i.e.* when $\tau(\alpha)$ is close to τ_L). The required conclusion then easily follows.

Equipped with this result, we then get the required characterization of the projection of $\mathcal{S}_1(\mathbf{u}_L)$ onto the (p, u) plane.

Proposition 9 *Let be given $\mathbf{u}_L$ in Ω. States $\mathbf{u}$ along $\mathcal{S}_1(\mathbf{u}_L)$ are given as a smooth function*

$$\mathbf{u}_1(p) = \{\tau_1(p), u_1(p), p, \{(s_i)_1(p)\}_{i=1,\dots,N-1}\}$$

of p in $[p_L, +\infty)$. Moreover, the velocity u_1 is a smooth strictly decreasing function which maps $[p_L, +\infty)$ onto $(-\infty, u_L]$.

Proof Let us first observe that the function τ is a smooth decreasing (in the large) function of p in $[p_L, +\infty)$ with range $(\tau_{min}, \tau_L]$ for some τ_{min} in $(0, \tau_L)$, while being necessarily strictly decreasing for p close to p_L.
Let us then prove that for all $p \geq p_L$, the velocity reads :

$$u_1(p) = u_L - ((\tau_L - \tau_1(p))(p - p_L))^{1/2}. \tag{1.88}$$

Indeed, the two identities

$$u_1(p) - u_L = m(\tau_1(p) - \tau_L), \quad p - p_L = m^2(\tau_L - \tau_1(p)),$$

readily implies that:

$$(u_1(p) - u_L)^2 = (p - p_L)(\tau_L - \tau_1(p)) \geq 0, \quad p \geq p_L.$$

But by the Lax condition (1.17) for a 1-shock solution, $m > (\rho c)_L$ while $\tau_1(p) \leq \tau_L$ for all $p \geq p_L$, we thus necessarily have $u_1(p) \leq u_L$ and hence the representation formula (1.88). The conclusion easily follows from the properties of the function τ_1 we have just reported.

To conclude this paragraph, let us state concerning the projection of the shock curve $\mathcal{S}_3(\mathbf{u}_R)$ onto the (p, u) plane:

Proposition 10 *For any given fixed $\mathbf{u}_R$ in Ω, states $\mathbf{u}$ along $\mathcal{S}_3(\mathbf{u}_R)$ are given as a smooth function*

$$\mathbf{u}_3(p) = \{\tau_3(p), u_3(p), p, \{(s_i)_3(p)\}_{i=1,\dots,N-1}\}$$

of p in $[p_R, +\infty)$, where the function u_3 strictly increases with p, and with range $[u_R, +\infty)$.

Proof Identical steps are in order except that the Lax condition (1.17) for a 3-shock solution says $m < -(\rho c)_R$ and now imply the following representation formula:

$$u_3(p) = u_R + ((\tau_R - \tau_3(p))(p - p_R))^{1/2}, \quad p \geq p_R.$$

The conclusion then follows.

1.5.2 Rarefaction waves and rarefaction curves

By definition, a rarefaction wave is a smooth solution of the asymptotic system (1.2) which is self-similar in the variable $\xi = x/t$. The smoothness assumption thus makes a rarefaction wave to be a particular solution, say $\mathbf{w}(\xi)$, of the system (1.86) with identically vanishing measures:

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + \sum_{i=1}^N p_i(\rho, s_i)) = 0, \\ \partial_t \rho s_i + \partial_x \rho s_i u = 0, \quad i = 1, ..., N. \end{cases} \tag{1.89}$$

Standard arguments from hyperbolic systems in full conservation form (see for instance [56], [130]) thus immediately apply to give that the self-similar solution under consideration must solve an $(N+2)$ ODE system in the form:

$$d_\xi \mathbf{w}(\xi) = r_i(\mathbf{w}(\xi)), \tag{1.90}$$

for some right eigenvector $r_i(\mathbf{u})$ of (1.89), while satisfying for the same index i along the orbit of (1.90):

$$\lambda_i(\mathbf{w}(\xi)) = \xi. \tag{1.91}$$

Smoothness of the solution of (1.90) is known to imply the non linearity of the field under consideration (since in view of (1.91), we must have $d_\xi \lambda_i(\mathbf{w}(\xi)) = \nabla\lambda_i.r_i(\mathbf{w}(\xi)) = 1$). As a consequence, only the two extreme fields, namely either $i = 1$ or $i = 3$ (actually genuinely non linear) can give rise to rarefaction wave solution.
The following result states that each of the two $(N+2)$ ODE system (1.90) of interest, *i.e.* with $i = 1$ or $i = 3$, admits $(N+1)$ algebraic invariants with linearly independent gradients (*i.e.* a full set of Riemann invariants, see again [56] or [130]).These are straightforward extensions of the two invariants known in the setting of a single pressure law and their existence allows for defining in the present framework, the notion of (two) families of rarefaction curves.

Lemma 7 *For some given state $\mathbf{u}_L \in \Omega$, let us define the* 1*-rarefaction curve attached to the left state $\mathbf{u}_L$ as the following subset of Ω:*

$$\mathcal{R}_1(\mathbf{u}_L) = \{\mathbf{u} \in \Omega, \mathbf{u}_L \ \underline{\textit{1-rarefaction}}\ \mathbf{u}\},$$

namely the set of right states $\mathbf{u}$ which can be connected to $\mathbf{u}_L$ by an admissible 1*-rarefaction wave. Then $\mathbf{u}$ belongs to $\mathcal{R}_1(\mathbf{u}_L)$ if and only if:*

$$u - c(\rho, \{s_i\}_{i=1,...,N}) > u_L - c_L, \quad \textit{as soon as } \mathbf{u} \neq \mathbf{u}_L, \tag{1.92}$$

and

$$s_i = (s_i)_L, \quad i = 1, ..., N, \tag{1.93}$$

$$u + g_L(\rho, \{(s_i)_L\}_{i=1,...,N}) = u_L + g_L(\rho_L, \{(s_i)_L\}_{i=1,...,N}), \tag{1.94}$$

where by definition:

$$g_L(\rho, \{s_i\}_{i=1,...,N}) = \int_{\rho_L}^{\rho} c(v, \{s_i\}_{i=1,...,N}) \frac{dv}{v},$$

with in particular $g_L(\rho_L, \{(s_i)_L\}_{i=1,...,N}) = 0$.
Symmetrically, define the 3*-rarefaction curve attached to a given right state $\mathbf{u}_R \in \Omega$ by:*

$$\mathcal{R}_3(\mathbf{u}_R) = \{\mathbf{u} \in \Omega, \mathbf{u} \ \underline{\textit{3-rarefaction}}\ \mathbf{u}_R\},$$

namely the set of left states $\mathbf{u}$ which can be connected to $\mathbf{u}_R$ by an admissible 3*-rarefaction wave. Then $\mathbf{u} \in \mathcal{R}_3(\mathbf{u}_R)$ if and only if:*

$$u + c(\rho, \{s_i\}_{i=1,...,N}) < u_R + c_R, \quad \textit{as soon as } \mathbf{u} \neq \mathbf{u}_R,$$

and

$$s_i = (s_i)_R, \quad i = 1, ..., N,$$

$$u - g_R(\rho, \{(s_i)_R\}_{i=1,...,N}) = u_R - g_R(\rho_R, \{(s_i)_R\}_{i=1,...,N}),$$

where by definition:

$$g_R(\rho, \{s_i\}_{i=1,...,N}) = \int_{\rho_R}^{\rho} c(v, \{s_i\}_{i=1,...,N}) \frac{dv}{v}.$$

Proof We shall only be concerned here with the characterization of the set of states $\mathcal{R}_1(\mathbf{u}_L)$. Invoking the special form of the smooth solution under consideration, the N advection equations governing each of the specific entropies (away from vacuum) :

$$\partial_t s_i + u\partial_x s_i = 0, \quad i = 1, ..., N,$$

just recast as :

$$(u(\xi) - \xi)d_\xi s_i(\xi) = 0, \quad i = 1, ..., N.$$

But since the identity (1.91) cannot hold for the intermediate (linearly degenerate) field, necessarily $u(\xi) \neq \xi$ and henceforth $d_\xi s_i(\xi) = 0$ for all $i = 1, ..., N$. This is nothing but the N Riemann invariants listed in (1.93).
Consequently, the total pressure p is solely a function of the density ρ for a rarefaction wave solution and the last Riemann invariant (1.94) just follows from the well-known theory of the isentropic Euler equations (see [56]).
To conclude the proof, let us recall that the condition (1.92) is immediately inferred from the genuine non linearity of the field under consideration (*i.e.* $d_\xi \lambda_1(\mathbf{w}(\xi)) = 1$) under the requirement that $\xi(\mathbf{u}) > \xi(\mathbf{u}_L)$: indeed $\mathbf{u}$ on the right is by definition connected to $\mathbf{u}_L$ on the left.

Equipped with this result, we are now in position to determine the projections of these two families of rarefaction curves onto the total pressure-velocity plane. This is precisely the matter of the next statement.

Proposition 11 *For any given state $\mathbf{u}_L \in \Omega$, the total pressure p serves as a parameter along $\mathcal{R}_1(\mathbf{u}_L)$ when defining each state $\mathbf{u} \in \mathcal{R}_1(\mathbf{u}_L)$ from a smooth function $\mathbf{u}_1(p) = (\tau_1(p), u_1(p), p, \{s_i\}_{i=1,...,N-1})$ of p in $(0, p_L]$ where the velocity u_1 is a smooth strictly decreasing function, mapping $(0, p_L]$ onto $[u_L, u_{max}(\mathbf{u}_L))$ for some $u_{max}(\mathbf{u}_L)$ in $(u_L, +\infty)$.*
Similarly and for any given $\mathbf{u}_R \in \Omega$, states $\mathbf{u}$ along $\mathcal{R}_3(\mathbf{u}_R)$ are given as a smooth function $\mathbf{u}_3(p) = (\tau_3(p), u_3(p), p, \{s_i\}_{i=1,...,N-1})$ of p in $(0, p_R]$, u_3 being a smooth strictly increasing function which maps $(0, p_R]$ onto $(u_{min}(\mathbf{u}_R), u_R]$ for some $u_{min}(\mathbf{u}_R)$ in $(-\infty, u_R)$.

Proof The proof we propose only addresses the 1-rarefaction curve $\mathcal{R}_1(\mathbf{u}_L)$. As far as $\mathcal{R}_3(\mathbf{u}_R)$ is concerned, the result will follow from simple adaptations of the arguments we now develop. Since the N specific entropies s_i are Riemann invariants, we first observe that the density ρ serves as a natural parameter along $\mathcal{R}_1(\mathbf{u}_L)$ which domain is actually restricted by the condition (with little abuse in the notations) :

$$\{u - c\}(\rho) \geq \{u - c\}(\rho_L). \tag{1.95}$$

Indeed, observing that

$$\frac{d}{d\rho}\{u - c\}(\rho) = -\frac{1}{2\rho c}\sum_{i=1}^{N} \gamma_i(\gamma_i + 1)(s_i)_L \rho^{\gamma_i - 1} < 0,$$

we infer that ρ must be kept in $(0, \rho_L]$ for (1.95) to hold true. Thus, the Riemann invariant (1.94) defines the velocity u as a smooth strictly decreasing function of the density which mapps $(0, \rho_L]$ onto $[u_L, u_{max})$ where by definition, u_{max} denotes a real number in $[u_L, +\infty[$ given by:

$$u_{max} = u_L + \lim_{\rho \to 0^+} \int_{\rho}^{\rho_L} c(v, \{(s_i)_L\}_{i=1,...,N}) \frac{dv}{v}. \tag{1.96}$$

The boundedness of u_{max} clearly holds by invoking the specific assumptions on the adiabatic exponents $\gamma_i > 1$, $i = 1, ..., N$.
Next and again because the N specific entropies are Riemann invariants, the total pressure

$$p(\rho) = \sum_{i=1}^{N} (s_i)_L \rho^{\gamma_i}$$

is a smoothly strictly increasing function of ρ in $(0, \rho_L]$ and with range $(0, p_L]$ (since $\gamma_i > 1$ for all $i = 1, ..., N$). It thus follows that along $\mathcal{R}_1(\mathbf{u}_L)$, the covolume τ is a smooth decreasing function $\tau_1(p)$ of $p \in (0, p_L]$ while the velocity u is a smooth decreasing function $u_1(p)$ mapping $(0, p_L]$ onto $[u_L, u_{max})$ with u_{max} given by (1.96). This concludes the proof.

1.5.3 Existence and uniqueness

Considering an arbitrary fixed initial data $(\mathbf{u}_L, \mathbf{u}_R)$ throughout this section, we are now in a position to define the following two curves of interest:

$$\mathcal{C}_1(\mathbf{u}_L) = \mathcal{S}_1(\mathbf{u}_L) \cup \mathcal{R}_1(\mathbf{u}_L), \quad \text{and} \quad \mathcal{C}_3(\mathbf{u}_R) = \mathcal{S}_3(\mathbf{u}_R) \cup \mathcal{R}_3(\mathbf{u}_R).$$

By construction,

$$\mathcal{C}_1(\mathbf{u}_L) = \{\mathbf{u} \in \Omega, \mathbf{u}_L \ \underline{\text{1-wave}}\ \mathbf{u}\},$$

is nothing but the set of states $\mathbf{u}$ which can be connected on the right to the left state $\mathbf{u}_L$ by a simple 1-wave (either an admissible shock or an admissible rarefaction). Then,

$$\mathcal{C}_3(\mathbf{u}_R) = \{\mathbf{u} \in \Omega, \mathbf{u}\ \underline{\text{3-wave}}\ \mathbf{u}_R\},$$

consists in states $\mathbf{u}$ which can be connected on the left to the right state $\mathbf{u}_R$ by an admissible simple 3-wave.
The projection of these two curves onto the (p, u) plane then have the following distinctive properties:

Proposition 12 *Along the curve $\mathcal{C}_1(\mathbf{u}_L)$, the velocity u_1 is a smooth strictly decreasing function of the total pressure p mapping $(0, +\infty)$ onto $(-\infty, u_{max}(\mathbf{u}_L))$ for some $u_{max}(\mathbf{u}_L)$ in $(u_L, +\infty)$ while along the curve $\mathcal{C}_3(\mathbf{u}_R)$, the velocity u_3 smoothly strictly increases with p in $(0, +\infty)$ and with range $(u_{min}(\mathbf{u}_R), +\infty)$ for some $u_{min}(\mathbf{u}_R)$ in $(-\infty, u_R)$.*

Proof We just have to check that the corresponding shock curves and rarefaction curves achieve a smooth connection. Focusing ourselves on $\mathcal{C}_1(\mathbf{u}_L)$, let us prove that $\mathcal{S}_1(\mathbf{u}_L)$ and $\mathcal{R}_1(\mathbf{u}_L)$ are actually tangent at $p = p_L$. Indeed, invoking the smooth representation formula (1.94) valid along $\mathcal{R}_1(\mathbf{u}_L)$, we first observe that the velocity u_1 when understood as a (smooth) function of the density admits a left derivative at $\rho = \rho_L$ explicitly given by $-c(\rho_L, \{(s_i)_L\}_{i=1,...,N})/\rho_L$ while the form of the total pressure

$p(\rho) = \sum_{i=1}^{N} \rho^{\gamma_i}(s_i)_L$ along $\mathcal{R}_1(\mathbf{u}_L)$ allows to compute $p'(\rho_L) = c^2(\rho_L, \{(s_i)_L\}_{i=1,\dots,N})$. Consequently, the velocity function $u_1(p)$ admits a left derivative at $p = p_L$ given by $-1/(\rho c)_L$.
Next turning considering the companion representation formula (1.88) along $\mathcal{S}_1(\mathbf{u}_L)$, we have in one hand and for all $p > p_L$:

$$\frac{u_1(p) - u_1(p_L)}{p - p_L} = -(\frac{\tau_1(p_L) - \tau_1(p)}{p - p_L})^{1/2}, \tag{1.97}$$

while on the second hand and again for all $p > p_L$:

$$\frac{\tau_1(p) - \tau_1(p_L)}{p - p_L} = -\frac{1}{m^2(p)} = -\frac{\alpha(p)}{(\rho c)_L^2}. \tag{1.98}$$

Since the covolume is a $\mathcal{C}^1$ function of the total pressure for $p \in [p_L, +\infty)$, (1.97) and (1.98) imply that the velocity function $u_1(p)$ admits a right derivative at $p = p_L$ given by $-1/(\rho c)_L$ (since $\alpha(p_L) = 1$). This is nothing but the required result. Similar arguments apply to prove the smooth connexion of $\mathcal{S}_3(\mathbf{u}_R)$ and $\mathcal{R}_3(\mathbf{u}_R)$.

As a direct consequence of the reported distinct monotonicity properties of the velocity functions $u_1(p)$ and $u_3(p)$ and their asymptotic behaviours, the respective projections of the curves $\mathcal{C}_1(\mathbf{u}_L)$ and $\mathcal{C}_3(\mathbf{u}_R)$ must necessarily intersect at a given pair $(u_\star, p_\star)$ as soon as

$$u_{min}(\mathbf{u}_R) < u_{max}(\mathbf{u}_L).$$

Therefore, the Riemann problem admits at least one solution given with clear notations by :

$$\mathbf{u}_L \ \underline{\text{1-wave}} \ \mathbf{u}_1(p_\star) \ \underline{\text{contact}} \ \mathbf{u}_3(p_\star) \ \underline{\text{3-wave}} \ \mathbf{u}_R.$$

with

$$\mathbf{u}_1(p_\star) = (\tau_1(p_\star), u_\star, p_\star, \{(s_i)_1(p_\star)\}_{i=1,\dots,N}),$$

and

$$\mathbf{u}_3(p_\star) = (\tau_3(p_\star), u_\star, p_\star, \{(s_i)_3(p_\star)\}_{i=1,\dots,N}).$$

But the strict monotonicity properties of the velocity functions under consideration imply that the pair $(u_\star, p_\star)$ is necessarily unique as soon as it exists.
Next turning considering the case $u_{max}(\mathbf{u}_L) \leq u_{min}(\mathbf{u}_R)$, this instance rises the same difficulty as in the setting of the usual Euler equations (*i.e.* with $N = 1$), namely the occurence of vacuum. In such an issue, it is known that the Riemann problem has no solution unless one specifically allows for a "void" state. We shall not address this particular construction and we refer the interested reader to Smith [129] and the references therein.
We have therefore proved :

Theorem 2 *The Riemann problem with initial data $(\mathbf{u}_L, \mathbf{u}_R)$ admits an unique solution away from vacuum, i.e. if*

$$u_{min}(\mathbf{u}_R) < u_{max}(\mathbf{u}_L),$$

where these two particular velocities have been defined in Proposition 11.

Chapitre 2

Numerical approximation of the Navier-Stokes equations with several independent specific entropies

Ce chapitre fait l'objet d'un article soumis pour publication : voir [20]

Numerical approximation of the Navier-Stokes equations with several independent specific entropies

C. Chalons [1, 2] and F. Coquel [3]

Résumé

The present work is devoted to the numerical approximation of the solutions of convective-diffusive systems which can be understood as natural extensions of the classical Navier-Stokes (NS) equations. Solutions of these systems are governed by N ($N \geq 2$) independent pressure laws or equivalently N independent specific entropies. Several models from the physics actually enter the present framework. These extended forms of the usual NS equations (i.e. when $N = 1$) are seen to share many properties but with the very difference that, generally speaking, they cannot be recast in full conservation form. This property is known to make the definition of the endpoints of the Travelling Wave (TW) solutions sensitive with respect to the shape of the diffusive tensor. Motivated by the work of Berthon [10] where $N = 2$, here we exhibit $(N-1)$ Generalized Rankine-Hugoniot (GRH) conditions that reflect this sensitivity. We underline the reason why classical algorithms can only fail in satisfying these GRH relations and we show how to enforce them for validity at the discrete level when extending [10] to $N \geq 2$ arbitrary general pressure laws. Several numerical evidences are proposed.

2.1 The mathematical model and some basic properties

In this work, we are interested in the numerical approximation of the solutions of the following convective-diffusive system (in one space dimension for simplicity) :

$$\left\{ \begin{array}{l} \partial_t \rho + \partial_x \rho u = 0, \qquad x \in I\!R, \quad t > 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + \sum_{i=1}^N p_i) = \partial_x ((\sum_{i=1}^N \mu_i) \partial_x u), \\ \partial_t \rho \epsilon_i + \partial_x (\rho \epsilon_i u) + p_i \partial_x u = \mu_i (\partial_x u)^2, \quad i = 1, ..., N, \end{array} \right. \tag{2.1}$$

where with classical notations, $\rho, \rho u$ and $\{\rho \epsilon_i\}_{i=1,...,N}$ respectively denote the density, the momentum and N independent internal energies. Before addressing the required closure relations, let us stress that (2.1) is nothing but a natural extension of the usual NS equations: namely when considering a single pressure law. Here, N ($N \geq 2$) independent pressure laws $p_i(\rho, \rho\epsilon_i)$ are involved via N independent governing equations for each of the $\rho\epsilon_i$. In (2.1) and for $i = 1, \ldots, N$, μ_i ($\mu_i \geq 0$) denotes a smooth mapping standing for the viscosity law associated with the internal energy $\rho\epsilon_i$. It will be always assumed that $\mu := \sum_{i=1}^N \mu_i > 0$. Depending on the choice of the closure equations for defining each of the μ_i, several models from the physics enter the present framework. Let us quote for instance compressible turbulence models like the $k - \epsilon$ model (see [13]), and more generally multi-scale models (see [19]) for which

1. O.N.E.R.A., B.P. 72, 29 avenue de la Division Leclerc, 92322 Châtillon Cedex, France.
E-mail address: chalons@onera.fr.
2. Centre de Mathématiques Appliquées & Centre National de la Recherche Scientifique, U.M.R. 7641, École Polytechnique, 91128 Palaiseau Cedex, France.
E-mail address: chalons@cmap.polytechnique.fr.
3. Laboratoire JLL, Boîte courrier 187, 75252 Paris Cedex 05, France.
E-mail address: coquel@ann.jussieu.fr.

a laminar pressure and some turbulent pressures are involved, or models coming from the plasma physics (see [34], [41]).
Now considering the required closure equations for defining each of the N pressure laws, we shall assume that according to the second principle of the thermodynamics:

$$-T_i(\tau,\epsilon_i)ds_i = d\epsilon_i + p_i(\tau,\epsilon_i)d\tau, \quad \tau = 1/\rho, \tag{2.2}$$

defines a strictly convex smooth mapping $(\tau,\epsilon_i) \to s_i(\tau,\epsilon_i) > 0$ for any given $i \in \{1,..,N\}$ where the temperature T_i is classically assumed to stay positive. As a well known consequence, the well defined mapping $(\tau,s_i) \to \varepsilon_i(\tau,s_i)$ is then strictly convex, as well as the mapping $(\rho,\rho\varepsilon_i) \to \{\rho s_i\}(\rho,\rho\varepsilon_i) := \rho s_i(1/\rho,\rho\varepsilon_i/\rho)$ and $(\rho,\rho s_i) \to \{\rho\varepsilon_i\}(\rho,\rho s_i) := \rho\varepsilon_i(1/\rho,\rho s_i/\rho)$ with some little abuse in the notations. Each energy law is assumed to obey in addition the well-known Weyl's assumptions (see [57] for the details). We shall use in particular the following asymptotic conditions:

$$\lim_{\rho s_i\to 0^+}\rho\varepsilon_i(\rho,\rho s_i) = +\infty \text{ and } \lim_{\rho s_i\to+\infty}\rho\varepsilon_i(\rho,\rho s_i) = 0 \ (\textit{i.e.} \lim_{\rho\varepsilon_i\to 0^+}\rho s_i(\rho,\rho\varepsilon_i) = +\infty).$$

Notice that all the above assumptions are quite classical within the frame of the usual NS equations (i.e. with $N = 1$). Owing to these assumptions, our first statement highlights the relationships with the usual setting, when defining the following natural phase space for (2.1):

$$\Omega = \{\mathbf{v} := (\rho,\rho u,\{\rho\epsilon_i\}_{1\le i\le N}) \in I\!R^{N+2}/\rho > 0,\ \rho u \in I\!R,\ \rho\epsilon_i > 0,\ 1 \le i \le N\}. \tag{2.3}$$

Lemma 8
The first order extracted system from (2.1) is hyperbolic over Ω, with the following increasingly ordered eigenvalues:

$$\lambda_1(\mathbf{v}) = u - c \le \lambda_2(\mathbf{v}) = \ldots = \lambda_{N+1}(\mathbf{v}) = u, \le \lambda_{N+2}(\mathbf{v}) = u + c,\ c^2(\mathbf{v}) = \sum_{i=1}^{N} c_i^2(\mathbf{v}),$$

where each partial sound speed is such that $c_i^2(\mathbf{v}) := (\partial_\rho p_i)_{s_i} > 0$. The $1-$ and $(N+2)-$ fields are genuinely nonlinear, all the other fields being linearly degenerate.
The proof of this easy result is omitted.
The system (2.1) naturally writes in non conservation form. This rises the question of the existence of an admissible change of variables that recasts (2.1) in full conservation form. With this respect, the next claim gives all the additional non trivial equations for the smooth solutions of (2.1) (see [19] or [10] when $N = 2$):
Theorem 3
Smooth solutions of (2.1) obey the following conservation law:

$$\partial_t\rho E + \partial_x(\rho E + \sum_{i=1}^{N} p_i)u = \partial_x((\sum_{i=1}^{N}\mu_i)u\partial_x u), \tag{2.4}$$

where the total energy reads $\rho E = (\rho u)^2/2\rho + \sum_{i=1}^{N}\rho\epsilon_i$. In addition, such solutions satisfy:

$$\partial_t\rho s_i + \partial_x\rho s_i u = -\frac{\mu_i}{T_i}(\partial_x u)^2, \quad i = 1,...,N. \tag{2.5}$$

Up to classical nonlinear transformations in the $\{s_i\}_{1\le i\le N}$, these are the only additional non trivial equations (without special assumptions on the $\{\mu_i\}_{1\le i\le N}$).

As a consequence, the system (2.1) cannot be recast generally speaking in full conservation form. However, let us rephrase the above result when investigating suitable change of variables for (2.1). Obviously, the three variables ρ, ρu and ρE are natural candidates since they are governed by conservation laws. Next and for completeness, let us consider without restriction the set $\{\rho s_i\}_{1\le i\le N-1}$ to define $\mathbf{u} := (\rho, \rho u, \rho E, \{\rho s_i\}_{1\le i\le N-1})$ as an admissible unknown. Then, smooth solutions of (2.1) equivalently satisfy :

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \qquad x \in I\!R, \quad t > 0, \\ \partial_t \rho u + \partial_x(\rho u^2 + \sum_{i=1}^N p_i) = \partial_x((\sum_{i=1}^N \mu_i)\partial_x u), \\ \partial_t \rho E + \partial_x(\rho E + \sum_{i=1}^N p_i)u = \partial_x((\sum_{i=1}^N \mu_i)u\partial_x u), \\ \partial_t \rho s_i + \partial_x \rho s_i u = -\frac{\mu_i}{T_i}(\partial_x u)^2, \quad i = 1, ..., N-1. \end{cases} \tag{2.6}$$

Let us notice that the first order underlying system in (2.6) is fully conservative. For conciseness in the forthcoming developments, we right away introduce the following condensed notation for (2.6):

$$\partial_t \mathbf{u} + \partial_x \mathbf{f}(\mathbf{u}) = \mathcal{R}(\mathbf{u}, \partial_x(\mathcal{B}(\mathbf{u})\partial_x \mathbf{u})).$$

The definition of the tensor $\mathcal{B}$ follows when writing $(\partial_x u)^2 = \partial_x(u\partial_x u) - u\partial_{xx}u$. Now, let us emphasize that whitout further condition, smooth solutions of (2.6) automatically obey with equality the additional PDE :

$$\partial_t \rho s_N(\mathbf{u}) + \partial_x \rho s_N(\mathbf{u})u = -\frac{\mu_N}{T_N}(\partial_x u)^2. \tag{2.7}$$

Notice that as a consequence of the convexity properties we have put forward at the beginning of this section, the mapping $\mathbf{u} \to \rho s_N(\mathbf{u})$ is strictly convex. Hence, the pair $(\rho s_N, \rho s_N u)$ is nothing but a Lax entropy pair but with a prescribed rate of entropy dissipation. To assess this issue, let us observe after Berthon [10] that despite the identities (2.5) might be understood as N fully independent equations, these do actually evolve proportionally according to :

Corollary 2
Smooth solutions of (2.6) obey :

$$\mu_N T_i\{\partial_t \rho s_i + \partial_x \rho s_i u\} = \mu_i T_N\{\partial_t \rho s_N(\mathbf{u}) + \partial_x \rho s_N(\mathbf{u})u\}, \quad i = 1, ..., N-1. \tag{2.8}$$

(2.8) expresses nothing but a trivial cancellation property in $(\partial_x u)^2$. However, this observation turns out to be crucial for our numerical purposes. Indeed when applied to TW solutions of (2.6), the identities (2.8) can be understood as $(N-1)$ Generalized Rankine-Hugoniot (GRH) conditions (see [19]). To exemplify this claim, let us temporarily adopt the following closure for the viscosity laws : $\mu_i(\mathbf{u}) = \mu_i^0 T_i$ where for each $i = 1, \ldots, N$, μ_i^0 denotes a given positive real number (see also [10] for complementary examples). Let us recall that such a closure can be inferred from the classical transport theory for dilute gases. Then, (2.8) are easily seen to boil down to :

$$\{\partial_t \rho(\frac{\mu_N}{\mu}s_i - \frac{\mu_i}{\mu}s_N(\mathbf{u})) + \partial_x \rho(\frac{\mu_N}{\mu}s_i - \frac{\mu_i}{\mu}s_N(\mathbf{u}))u\} = 0, \quad i = 1, ..., N-1. \tag{2.9}$$

Therefore, (2.9) clearly provide us with $(N-1)$ jump conditions that, with the three classical Rankine-Hugoniot conditions respectively for ρ, ρu and ρE, yield a complete

set of jump relations for defining the endpoints of a TW solution of (2.6). Let us underline that such endpoints clearly depend on the ratios of the viscosities.
More generally, the proportionality in the evolution in time of all the entropy balance equations (as put forward in (2.8)) expresses a well-known property of general convective-diffusive systems in non conservation form. Namely, the endpoints of their TW solutions entirely depend on the shape of the diffusive tensor under consideration (see [42] and [85]). As underlined in the next section, this key property will make the numerical approximation of the TW solutions of (2.8) a non classical issue (see also [87]). In the present work, we show how to enforce numerically for validity the GRH relations (2.8).

2.2 Numerical approximation

In this section, we present a suitable numerical method for approximating the solutions of (2.1). This method can be understood as a prediction-correction strategy. In the prediction step, we just propose to solve the equivalent form (2.6) with appropriate initial data. Such an approach sounds quite natural but however will be seen to fail in preserving the validity of the GRH conditions at the discrete level. Highlighting the roots of such a failure will naturally motivate the correction step which precisely aims at restoring the validity of (2.8) at each time step.
Let be given a constant time step Δt and a constant space step Δx. Introducing $x_{j+1/2} = (j+1/2)\Delta x$ for $j \in Z$ and $t^n = n\Delta t$ for $n \in N$, the approximate solution of (2.6), say $\mathbf{u}_\lambda(x,t)$, with $\mathbf{u}_0$ as initial data is classically sought as a piecewise constant function on each slab $C_j^n = [x_{j-1/2}; x_{j+1/2}) \times [t^n; t^{n+1})$ for $(n,j) \in N \times Z$, and we set

$$\mathbf{u}_\lambda(x,t) = \mathbf{u}_j^n \quad \text{for all} \quad (x,t) \in C_j^n, \quad j \in Z, \quad n \in N.$$

Here λ stands for the ratio $\lambda = \Delta t/\Delta x$, and for completeness we set

$$\mathbf{u}_j^0 = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{u}_0(x)dx.$$

Assuming that the approximate solution $\mathbf{u}_\lambda(x,t^n)$ is known, we propose to advance it to the next time level t^{n+1} according to a classical splitting stategy between the convective and diffusive parts of (2.6). We now describe these two steps into details.

First step ($t^n \to t^{n+1=}$) This step consists in solving the first order underlying system in (2.6), that is to say:

$$\partial_t \mathbf{u} + \partial_x \mathbf{f}(\mathbf{u}) = 0, \tag{2.10}$$

with $\mathbf{u}_\lambda(x,t^n)$ as initial data, and supplemented with the following Lax entropy inequality :

$$\partial_t \rho s_N(\mathbf{u}) + \partial_x \rho s_N(\mathbf{u}) u \leq 0. \tag{2.11}$$

Let us underline that (2.11) is actually needed for uniqueness of the weak solutions of (2.10). Generically and for any given $\mathbf{u}_L$ and $\mathbf{u}_R$ in the phase space, we then denote by $\mathbf{u}(., \mathbf{u}_L, \mathbf{u}_R)$ the self similar solution of the Riemann problem (2.10)-(2.11) with $\mathbf{u}(x,0) = \mathbf{u}_L$ if $x < 0$ and $\mathbf{u}(x,0) = \mathbf{u}_R$ if $x > 0$ as initial data.

To define $\mathbf{u_j^{n+1=}}$, we propose to use the celebrated Godunov method under the usual CFL condition

$$\lambda \max_{\mathbf{u}}(|\lambda_i(\mathbf{u})|, i = 1, ..., N+2) \leq \frac{1}{2}, \tag{2.12}$$

where the maximum is taken over all the $\mathbf{u}$ under consideration. As a consequence, it is well known that $\mathbf{u}_j^{n+1=}$ can be given an updating formula in full conservation form and obeys in addition a discrete cell entropy inequality. More precisely we have:

Theorem 4

Under the CFL condition (2.12), the Godunov method applied to (2.10)-(2.11) with $\mathbf{u}_\lambda(x, t^n)$ *as initial data leads to*

$$\mathbf{u}_j^{n+1=} = \mathbf{u}_j^n - \lambda\Delta\mathbf{f}_{j+1/2}^n, \quad \Delta\mathbf{f}_{j+1/2}^n = \mathbf{f}_{j+1/2}^n - \mathbf{f}_{j-1/2}^n, \quad j \in Z, \tag{2.13}$$

where we have classically set $\mathbf{f}_{j+1/2}^n = \mathbf{f}(\mathbf{u}(0; \mathbf{u}_j^n, \mathbf{u}_{j+1}^n))$ *for all* $j \in Z$.
In addition, the following cell entropy inequality holds true:

$$(\rho s_N)_j^{n+1=} := \rho s_N(\mathbf{u}_j^{n+1=}) \leq (\rho s_N)_j^n - \lambda\Delta\{\rho s_N u\}_{j+1/2}^n, \tag{2.14}$$

with $\{\rho s_N u\}_{j+1/2}^n = \{\rho s_N u\}(\mathbf{u}(0; \mathbf{u}_j^n, \mathbf{u}_{j+1}^n))$, *as well as the maximum principles on the specific entropies*

$$(s_i)_j^{n+1=} \leq \max((s_i)_{j-1}^n, (s_i)_j^n, (s_i)_{j+1}^n), \quad i = 1, ..., N. \tag{2.15}$$

Let us recall that the inequality (2.14) comes from the entropy selection principle (2.11) together with the Jensen inequality associated with the averaging procedure in the Godunov method. As a consequence, such an inequality is strict generally speaking.

To conclude this first step, observe that one could have used a suitable approximate Riemann solver instead of the exact one for solving the Cauchy problem (2.10)-(2.11). Indeed, the forthcoming analysis entirely stays valid provided that natural stability properties (like consistency, average conservation, average entropy dissipation, ...) are assumed to hold true for the approximate Riemann solver under consideration. We refer the reader to [19] for such considerations. Let us now address the second step.

Second step ($t^{n+1=} \to t^{n+1-}$) This step takes into account the diffusive part of (2.6). It amounts to solve

$$\partial_t\mathbf{u} = \mathcal{R}(\mathbf{u}, \partial_x(\mathcal{B}(\mathbf{u})\partial_x\mathbf{u})), \tag{2.16}$$

with $\mathbf{u}_\lambda(x, t^{n+1=})$ as initial data. Notice that the smooth solutions of this system automatically obey the following additional equation:

$$\partial_t\rho s_N(\mathbf{u}) = -\frac{\mu_N}{T_N}(\partial_x u)^2. \tag{2.17}$$

Several numerical techniques are at hand in the litterature to provide us with an updated solution, say $\mathbf{u}_\lambda(x, t^{n+1-})$, which satisfies (2.16) as well as (2.17) in the sense of the finite differences. Here and after [10], we propose the following implicit discretization, to avoid a too restrictive CFL-like condition:

$$\begin{cases} \rho_j^{n+1-} = \rho_j^{n+1=}, \\ (\rho u)_j^{n+1-} = (\rho u)_j^{n+1=} + \Delta t\overline{\partial_x(\mu\partial_x u)}_j^{n+1-}, \\ (\rho E)_j^{n+1-} = (\rho E)_j^{n+1=} + \Delta t\overline{\partial_x(\mu u\partial_x u)}_j^{n+1-}, \\ (\rho s_i)_j^{n+1-} = (\rho s_i)_j^{n+1=} - \Delta t\overline{\frac{\mu_i}{T_i}(\partial_x u)^2}_j^{n+1-}, \quad i = 1, ..., N-1. \end{cases} \tag{2.18}$$

Here the finite differences formulae read:

$$\begin{cases} \overline{\partial_x(\mu\partial_x u)}_j^{n+1-} & = \frac{1}{\Delta x^2}(\mu^n_{j+1/2}(M^n u_{j+1} - M^n u_j) - \mu^n_{j-1/2}(M^n u_j - M^n u_{j-1})), \\ \overline{\partial_x(\mu u\partial_x u)}_j^{n+1-} & = \frac{1}{2\Delta x^2}(\mu^n_{j+1/2}((M^n u_{j+1})^2 - (M^n u_j)^2) - \mu^n_{j-1/2}((M^n u_j)^2 - (M^n u_{j-1})^2)), \\ \overline{\frac{\mu_i}{T_i}(\partial_x u)^2}_j^{n+1-} & = \frac{1}{\overline{T_{ij}}^{n+1-}}((\mu_i)^n_{j+1/2}(M^n u_{j+1} - M^n u_j)^2 + (\mu_i)^n_{j-1/2}(M^n u_j - M^n u_{j-1})^2), \end{cases}$$

where $M^n X = (X^{n+1-} + X^{n+1=})/2$ and for all i from 1 to N:

$$\overline{T_{ij}}^{n+1-} = \begin{cases} -\frac{\varepsilon_i(\tau_j^{n+1-},(s_i)_j^{n+1-}) - \varepsilon_i(\tau_j^{n+1-},(s_i)_j^{n+1=})}{(s_i)_j^{n+1-} - (s_i)_j^{n+1=}} & \text{if} \quad (s_i)_j^{n+1-} \neq (s_i)_j^{n+1=}, \\ T_i(\tau_j^{n+1-}, (s_i)_j^{n+1-}) & \text{otherwise.} \end{cases}$$

Let us emphasize that since the density is kept unchanged, solving (2.18) just amounts to solve a linear problem in $M^n u$, with a positive definite symmetric matrix. The unknowns $(\rho E)_j^{n+1-}$ and $\{(\rho s_i)_j^{n+1-}\}_{i=1,\dots,N-1}$ then follow with explicit formulae. We do not discuss this step into details and refer the reader to [10] for more informations as well as for the proof of the following statement.

Lemma 9

The updating formulae (2.18) lead to

$$(\rho s_N)_j^{n+1-} := \{\rho s_N\}(\mathbf{u}_j^{n+1-}) = (\rho s_N)_j^{n+1=} - \Delta t \overline{\frac{\mu_N}{T_N}(\partial_x u)^2}_j^{n+1-}. \tag{2.19}$$

Moreover, the following maximum principles hold true:

$$(s_i)_j^{n+1-} \le (s_i)_j^{n+1=} \le \max((s_i)^n_{j-1}, (s_i)^n_j, (s_i)^n_{j+1}), \quad i = 1, ..., N. \tag{2.20}$$

Thus, at the end of this second step, (2.16) and (2.17) can be given a consistent definition in the sense of finite differences. To summarize the prediction step (*i.e.* the first step and the second step), we have with clear notations

$$\begin{aligned} \rho_j^{n+1-} &= \rho_j^n - \lambda\Delta\{\rho u\}^n_{j+1/2} \\ (\rho u)_j^{n+1-} &= (\rho u)_j^n - \lambda\Delta\{\rho u^2 + \textstyle\sum_{i=1}^N p_i\}^n_{j+1/2} + \Delta t \overline{\partial_x(\mu\partial_x u)}_j^{n+1-} \\ (\rho E)_j^{n+1-} &= (\rho E)_j^n - \lambda\Delta\{(\rho u^2 + \textstyle\sum_{i=1}^N p_i)u\}^n_{j+1/2} + \Delta t \overline{\partial_x(\mu u\partial_x u)}_j^{n+1-} \\ (\rho s_i)_j^{n+1-} &= (\rho s_i)_j^n - \lambda\Delta\{\rho s_i u\}^n_{j+1/2} - \Delta t \overline{\tfrac{\mu_i}{T_i}(\partial_x u)^2}_j^{n+1-}, \quad i = 1, ..., N-1, \end{aligned}$$

$$(\rho s_N)_j^{n+1-} \le (\rho s_N)_j^n - \lambda\Delta\{\rho s_N u\}^n_{j+1/2} - \Delta t \overline{\frac{\mu_N}{T_N}(\partial_x u)^2}_j^{n+1-}, \tag{2.21}$$

where as a consequence of (2.14), the inequality is strict generally speaking. The classical approach would consist in choosing $\mathbf{u}_j^{n+1} = \mathbf{u}_j^{n+1-}$. But, because of the strict inequality in (2.21) the cancellation property, responsible for (2.8), is lost and the GRH conditions can no longer be satisfied. Obviously, such a failure persists when using a suitable approximate Riemann solver for (2.10)–(2.11). In the next section, numerical evidences will show that the corresponding approximate solutions grossly disagree with the exact ones. This discussion then clearly motivates the need for a correction step to define the updated solution $\mathbf{u}^h(x, t^{n+1})$ so that the required GRH conditions hold. We propose a correction step which extends the procedure introduced in [10] to arbitrary $N \geq 2$ general pressure laws and we prove its relevance.

Correction step ($t^{n+1-} \to t^{n+1}$) We keep unchanged the updated conservative variables, i.e.:

$$\rho_j^{n+1} = \rho_j^{n+1-}, \quad (\rho u)_j^{n+1} = (\rho u)_j^{n+1-}, \quad (\rho E)_j^{n+1} = (\rho E)_j^{n+1-},$$

and we propose to define the N entropies $\{(\rho s_i)_j^{n+1}\}_{i=1,...,N-1}$ as the solutions of the following non linear algebraic problem:

$$\begin{gathered} \overline{\mu_i}\overline{T_N}_j^{n+1-}(\{\rho s_N\}(\mathbf{u}_j^{n+1}) - (\rho s_N)_j^n + \lambda\Delta\{\rho s_N u\}_{j+1/2}^n) \\ = \\ \overline{\mu_N}\overline{T_i}_j^{n+1-}((\rho s_i)_j^{n+1} - (\rho s_i)_j^n + \lambda\Delta\{\rho s_i u\}_{j+1/2}^n), \\ \\ i = 1, ..., N-1. \end{gathered} \tag{2.22}$$

Here $\overline{\mu_i}$ is given a natural consistent definition for all $i = 1, ..., N$. Notice that (2.22) is nothing but a discrete version of the GRH conditions (2.8). The numerical fluxes and the definitions of $\overline{T_i}_j^{n+1-}$ entering (2.22) are directly inherited from the prediction step. The following result assesses the relevance of the correction step.

Theorem 5
The algebraic problem (2.22) admits a single solution $\{(\rho s_i)_j^{n+1}\}_{i=1,...,N-1}$ which obeys in addition the following entropy inequalities

$$\begin{aligned} &(\rho s_i)_j^{n+1} \leq (\rho s_i)_j^n - \lambda\Delta\{\rho s_i u\}_{j+1/2}^n, \quad i = 1, ..., N-1, \\ &(\rho s_N)_j^{n+1} := \{\rho s_N\}(\mathbf{u}_j^{n+1}) \leq (\rho s_N)_j^n - \lambda\Delta\{\rho s_N u\}_{j+1/2}^n, \end{aligned} \tag{2.23}$$

as well as the maximum principles on the specific entropies

$$(s_i)_j^{n+1} \leq \max((s_i)_{j-1}^n, (s_i)_j^n, (s_i)_{j+1}^n), \quad i = 1, ..., N, \quad j \in Z. \tag{2.24}$$

Proof
Let us introduce the following convenient notations

$$(\rho s_N)_j^\star = (\rho s_N)_j^n - \lambda\Delta\{\rho s_N u\}_{j+1/2}^n (\neq \{\rho s_N\}(\mathbf{u}_j^{n+1=})),$$

$$\Gamma_{m,n} = \overline{\mu_m}\overline{T_n}_j^{n+1-} / \overline{\mu_n}\overline{T_m}_j^{n+1-} \geq 0, \quad m = 1, ..., N, \quad , n = 1, ..., N.$$

Let us observe that these quantities are explicitly known at this stage. The first step of the proof consists in recasting (2.22) as follows:

$$\begin{gathered} (\rho s_i)_j^{n+1} = \rho s_i((\rho s_{N-1})_j^{n+1}) = (\rho s_i)_j^{n+1=} + \Gamma_{i,N-1}((\rho s_{N-1})_j^{n+1} - (\rho s_{N-1})_j^{n+1=}), \\ i = 1, ..., N-2, \\ \\ \{\rho s_N\}(\rho_j^{n+1}, (\rho u)_j^{n+1}, (\rho E)_j^{n+1}, \{\rho s_i((\rho s_{N-1})_j^{n+1})\}_{i=1,...,N-2}, (\rho s_{N-1})_j^{n+1}) - (\rho s_N)_j^\star \\ = \\ \Gamma_{N,N-1}((\rho s_{N-1})_j^{n+1} - (\rho s_{N-1})_j^{n+1=}). \end{gathered} \tag{2.25}$$

These identities express the property that the unknowns $(\rho s_i)_j^{n+1}$ for $i = 1, ..., N-2$ can be parametrized by $(\rho s_{N-1})_j^{n+1}$ so that the algebraic problem (2.22) can be solved with $(\rho s_{N-1})_j^{n+1}$ as the only scalar unknown. Thus, existence and uniqueness of the $\{(\rho s_i)_j^{n+1}\}_{i=1,...,N-1}$ as the solutions of (2.25) is equivalent to prove the existence of a non positive real number $\mathcal{M}_j \leq 0$ (namely $\mathcal{M}_j = (\rho s_{N-1})_j^{n+1} - (\rho s_{N-1})_j^{n+1=}$) such that

$$\begin{gathered}(\rho s_i)_j^{n+1} = (\rho s_i)_j^{n+1=} + \Gamma_{i,N-1}\mathcal{M}_j, \quad i = 1, ..., N-1, \\ f(\mathcal{M}_j) = \{\rho s_N\}(\rho_j^{n+1}, (\rho u)_j^{n+1}, (\rho E)_j^{n+1}, \{(\rho s_i)_j^{n+1=} + \Gamma_{i,N-1}\mathcal{M}_j\}_{i=1,...,N-1}) - \\ \Gamma_{N,N-1}\mathcal{M}_j - (\rho s_N)_j^{\star} = 0. \end{gathered} \tag{2.26}$$

Now, since $\partial_{\rho s_i}\{\rho s_N\} = -T_i/T_N < 0$, the mapping $\mathbf{u} \to \rho s_N(\mathbf{u})$ is strictly decreasing with respect to each of the entropies $\{\rho s_i\}_{i=1,...,N-1}$, so that the mapping $\mathcal{M} \to f(\mathcal{M})$ is also strictly decreasing thanks to the positivity of $\Gamma_{m,n}$. Arguing of the asymptotic Weyl's assumptions, let us prove that $\{\rho s_N\}(\mathcal{M}_j) \to +\infty$ (*i.e.* $f(\mathcal{M}_j) \to +\infty$) for a first sufficiently large negative value of $\mathcal{M}_j$, say $\overline{\mathcal{M}_j}$. In that aim, notice that

$$\rho s_i(\mathcal{M}_j) := (\rho s_i)_j^{n+1=} + \Gamma_{i,N-1}\mathcal{M}_j$$

approaches 0^+ for negative enough values of $\mathcal{M}_j$ for each $i \in \{1, ..., N-1\}$. By the asymptotic condition

$$\lim_{\rho s_i \to 0^+} \rho\varepsilon_i(\rho, \rho s_i) = +\infty,$$

there exists necessarily a first value of $\mathcal{M}_j$, namely $\overline{\mathcal{M}_j}$, solution of

$$(\rho E)_j^{n+1} - \frac{((\rho u)_j^{n+1})^2}{2\rho_j^{n+1}} = \sum_{i=1}^{N-1} \rho\varepsilon_i(\rho_j^{n+1}, \rho s_i(\mathcal{M}_j)),$$

that is to say that $\rho\varepsilon_N(\mathcal{M}_j) = 0$ (note that $\rho\varepsilon_N(0) = \rho\varepsilon_N(\rho_j^{n+1=}, (\rho s_N)_j^{n+1=}) > 0$.) But the asymptotic condition

$$\lim_{\rho\varepsilon_N \to 0^+} \rho s_N(\rho, \rho\varepsilon_N) = +\infty$$

just says that

$$\lim_{\mathcal{M}_j \to \overline{\mathcal{M}_j}^+} \rho s_N(\mathcal{M}_j) = +\infty.$$

Note that we have proved in passing that necessarily $\rho s_i(\mathcal{M}_j) > 0$ for all $\overline{\mathcal{M}_j} < \mathcal{M}_j \leq 0$ (since otherwise $f(\mathcal{M}_j) = +\infty$).
Let us next observe the estimate:

$$\begin{aligned} f(0) &= \{\rho s_N\}(\rho_j^{n+1}, (\rho u)_j^{n+1}, (\rho E)_j^{n+1}, \{(\rho s_i)_j^{n+1=}\}_{i=1,...,N-1}) - (\rho s_N)_j^{\star} \\ &= \{\rho s_N\}(\rho_j^{n+1-}, (\rho u)_j^{n+1-}, (\rho E)_j^{n+1-}, \{(\rho s_i)_j^{n+1=}\}_{i=1,...,N-1}) - (\rho s_N)_j^{\star} \leq 0, \end{aligned}$$

as an immediate consequence of (2.21): indeed, notice that thanks to (2.20) and the decreasing property of $\{\rho s_N\}$ with respect to ρs_i, the following inequality holds:

$$\{\rho s_N\}(\rho_j^{n+1-}, (\rho u)_j^{n+1-}, (\rho E)_j^{n+1-}, \{(\rho s_i)_j^{n+1=}\}_{i=1,...,N-1}) \leq \{\rho s_N\}(\mathbf{u}_j^{n+1-}).$$

Since $f(0) \leq 0$ with $\lim_{\mathcal{M}_j \to \overline{\mathcal{M}_j}^+} f(\mathcal{M}_j) = +\infty$, we necessarily get from the strict monotonicity of the continuous function f the existence of an unique $\mathcal{M}_j$, solution of (2.26), with the property that

$$\overline{\mathcal{M}_j} < \mathcal{M}_j \leq 0,$$

so that necessarily :

$$(\rho s_i)_j^{n+1} = (\rho s_i)_j^{n+1=} + \Gamma_{i,N-1}\mathcal{M}_j > 0, \quad i = 1, ..., N-1.$$

Turning now establishing the maximum principles (2.24), the entropy inequalities we have just obtained rewrite $(\rho s_i)_j^{n+1} \leq (\rho s_i)_j^{n+1=}$ that is to say $(s_i)_j^{n+1} \leq (s_i)_j^{n+1=}$ for all $i = 1, ..., N-1$ since $\rho_j^{n+1} = \rho_j^{n+1=}$. Invoking now (2.15), we obtain the required conclusion for $i = 1, ..., N-1$. Concerning now ρs_N, we have just proved $(\rho s_N)_j^{n+1} \leq (\rho s_N)_j^{\star}$ (but $(s_N)_j^{n+1} \geq (s_N)_j^{n+1=}$), which is classically sufficient to obtain the required maximum principle in the Godunov method setting. This concludes the proof.

2.3 Numerical experiments

In this section, we give some numerical evidences to illustrate the relevance of the prediction-correction scheme we have proposed. To that purpose, we deal with N internal energies of polytropic ideal gases (thermally and calorically perfect) associated with N constant adiabatic exponents $\gamma_i > 1$ for all i=1,...,N. Hence, the total energy reads $\rho E = (\rho u)^2/2\rho + \sum_{i=1}^{N} p_i/(\gamma_i - 1)$. For simplicity we consider constant viscosity laws with a Reynolds number equal to 10^5.
We consider as initial data two constant states, called left state and right state in the following, separated by a discontinuity at $x = 0$ and we approximate the solution on a uniform grid with $\Delta x = 1/300$.

Experiment 1 We set $N = 3$, $(\gamma_1, \gamma_2, \gamma_3) = (1.4, 1.6, 1.8)$ and $(\mu_2/\mu_1, \mu_3/\mu_1) = (1., 1.)$, while the left and right states of the initial data read: $(\rho, u, p_1, p_2, p_3)_{left} = (4., 1., 1.2, 1.4, 1.6)$, $(\rho, u, p_1, p_2, p_3)_{right} = (2.5568, -1.4305, 0.5162, 0.5103, 0.4999)$.

Experiment 2 We set $N = 3$, $(\gamma_1, \gamma_2, \gamma_3) = (1.4, 1.4, 1.4)$ and $(\mu_2/\mu_1, \mu_3/\mu_1) = (1., 1.)$ while the left and right states of the initial data read: $(\rho, u, p_1, p_2, p_3)_{left} = (3., 1., 1., 1.2, 1.4)$, $(\rho, u, p_1, p_2, p_3)_{right} = (2.6529, -1.1153, 0.8160, 0.9844, 1.1528)$.

In the Figures 2.1 and 2.2, we compare some of the corresponding pressure profiles for the exact solutions (see [19] for the derivation of exact solutions) together with the numerical solutions generated by the scheme both with and without the correction step. As expected, we observe that the classical approach (without correction) fails in capturing the correct solution while the correction step provides us with a numerical solution in good agreement with the exact one. We refer the reader to [19] for additional numerical experiments.

To conclude this paper, let us emphasize that several distinct methods have been actually developed and analysed (see [19]). Their design principle is always focused at preserving the validity of the GRH conditions we have put forward in (2.8). In

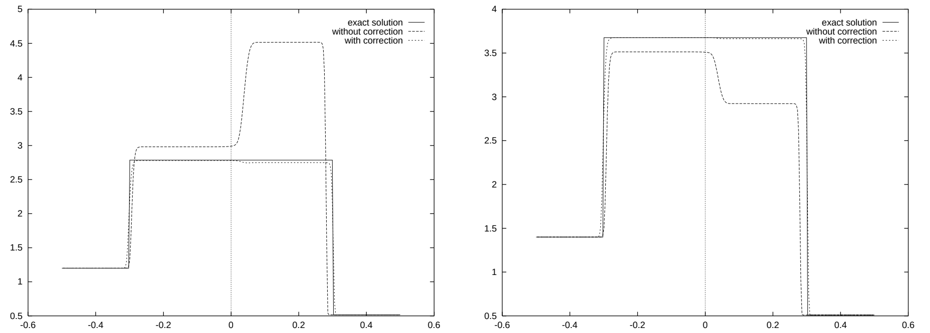

FIG. 2.1 – *Experiment 1 - Pressures 1 and 2*

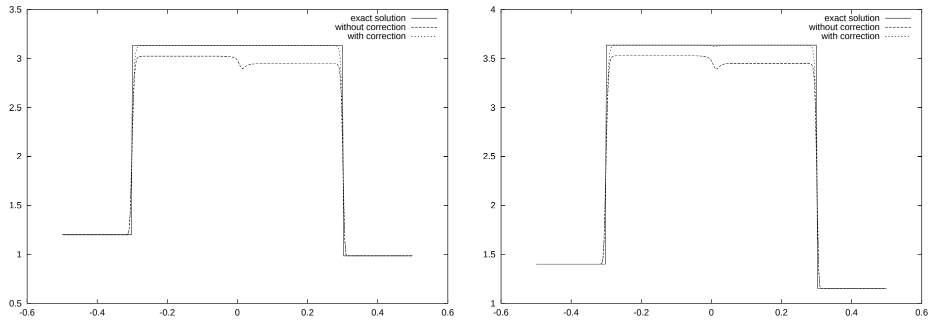

FIG. 2.2 – *Experiment 2 - Pressures 2 and 3*

addition, they all satisfy the same stability estimates but they come with a different cost of numerical evaluation. Here, we have restricted ourselves to the most natural one (along the lines of the first section) but at the expense of the largest numerical cost (we have to solve the nonlinear problem (2.25)). In [19], fully linear correction steps are proposed. In addition, invoking the rotational invariance of the equations under consideration, these methods have been extended to problems with two space variables considering realistic viscosity laws and boundary conditions. We refer the reader to [19].

Chapitre 3

Navier-Stokes equations with several independent pressure laws and explicit predictor-corrector schemes

Ce chapitre fait l'objet d'un article en préparation.

Navier-Stokes equations with several independent pressure laws and explicit predictor-corrector schemes

Chalons C.[1,2] and Coquel F.[3]

Résumé

This work is concerned with the numerical capture of stiff viscous shock solutions of Navier-Stokes equations for complex compressible materials, in the regime of large Reynolds numbers. After [10] and [20], a relevant numerical capture is known to require the satisfaction of an extended set of non classical Rankine-Hugoniot conditions due to the non conservation form of the governing PDE model. Here, we show how to enforce their validity at the discrete level without the need for solving local non linear algebraic problems. Non linearities are bypassed when introducing new averaging techniques which are proved to satisfy all the desirable stability properties when invoking suitable approximate Riemann solutions. A relaxation procedure is proposed to that purpose with the benefit of a fairly simple overall numerical method.

3.1 Introduction

In this paper, we are interested in the numerical approximation of the solutions of the following system in non conservation form :

$$\left\{ \begin{array}{l} \partial_t \rho^\epsilon + \partial_x (\rho u)^\epsilon = 0, \\ \partial_t (\rho u)^\epsilon + \partial_x (\rho u^2 + \sum_{i=1}^N p_i)^\epsilon = \epsilon \partial_x (\sum_{i=1}^N \mu_i \partial_x u^\epsilon), \\ \partial_t p_i^\epsilon + \partial_x (p_i u)^\epsilon + (\gamma_i - 1) p_i^\epsilon \partial_x u^\epsilon = \epsilon \mu_i (\partial_x u^\epsilon)^2, \quad i = 1, ..., N, \end{array} \right. \tag{3.1}$$

for a given fixed but small $\epsilon > 0$ under the assumption of a large Reynolds number. This PDE system readily stands as a natural extension of the usual Navier-Stokes equations when a single pressure is involved in the momentum equation. Here, N independent pressure laws occur and are governed by their own PDE, in complete symmetry with the equation governing a single pressure. Then the sum of all these pressures, $p := \sum_{i=1}^N p_i$, enters the momentum equation. Several PDE models from the physics can be identified under this form. Berthon [10] was the first to recognize that the so-called two transport models (namely the $k - \epsilon$ or related models) from the compressible turbulence setting must be actually understood within the present framework for extended Navier-Stokes equations with $N = 2$. Chalons [19] has then shown that more sophisticated turbulent models (the so-called muti-scale models) also naturally fall within the present frame with $N \geq 2$ arbitrary pressures. Let us mention that other models from the physics also enter the proposed setting : namely the

1. O.N.E.R.A., B.P. 72, 29 avenue de la Division Leclerc, 92322 Châtillon Cedex, France.
E-mail: Christophe.Chalons@onera.fr.

2. Centre de Mathématiques Appliquées & Centre National de la Recherche Scientifique, U.M.R. 7641, École Polytechnique, 91128 Palaiseau Cedex, France.
E-mail: chalons@cmapx.polytechnique.fr

3. Laboratoire JLL & Centre National de la Recherche Scientifique, U.M.R. 7598, Boîte courrier 187, 75252 Paris Cedex 05, France.
E-mail: coquel@ann.jussieu.fr

multi-fluid models introduced by Lagoutière [79] but also other complex gas mixtures analyzed by Truesdell [137].

Here we assume that each of the independent pressures p_i follows a polytropic law with its own adiabatic exponent $\gamma_i > 1,\ i = 1, ..., N$. Next, the viscosity coefficients μ_i entering the system (3.1) denote non negative real numbers but with the requirement of a positive sum, namely $\mu := \sum_{i=1}^{N} \mu_i > 0$.

As a consequence of the thermodynamic closure equations, it can be easily checked (for the details, see [10] with $N = 2$ and [19] with $N > 2$) that the first order underlying system in (3.1) is non linear hyperbolic over the following natural phase space :

$$\Omega = \{\mathbf{V} := (\rho, \rho u, \{p_i\}_{i=1,...,N}) \in I\!R^{N+2} / \rho > 0, \rho u \in I\!R, \{p_i > 0\}_{i=1,...,N}\},$$

from which vacuum is (and will be hereafter) excluded for simplicity in the discussion. Three distinct eigenvalues are in order, namely $u - c < u < u + c$ where the sound speed follows from $c^2 := \sum_{i=1}^{N} \gamma_i p_i / \rho > 0$. The two extreme fields are seen to be genuinely non linear from the assumption $\gamma_i > 1,\ i = 1, ..., N$. By this property, the solutions of the hyperbolic system with viscous perturbations (3.1) are known to generally develop stiff zones of transition often referred in the literature as to viscous shock profiles (see Gilbarg [53] for instance). The numerical approximation of these solutions is of primary importance in the present work. The assumption of a large Reynolds number makes this issue to be particularly challenging since the PDE model (3.1) naturally writes in non conservation form. Actually, Berthon [10] and Chalons [19] have proved that (3.1) cannot recast equivalently for smooth solutions in full conservation form, unless restrictive modelling assumptions are specifically addressed. The following statement proved in [19] gives the existence of solely a non trivial additional conservation law :

Proposition 13 *Smooth solutions of (3.1) satisfy the following conservation law :*

$$\partial_t (\rho E)^\epsilon + \partial_x (\rho H u)^\epsilon = \epsilon \partial_x (\mu u^\epsilon \partial_x u^\epsilon), \tag{3.2}$$

where the total energy ρE and the total enthalpy ρH respectively read :

$$\rho E = \frac{(\rho u)^2}{2\rho} + \sum_{i=1}^{N} \frac{p_i}{\gamma_i - 1}, \quad \rho H = \rho E + p.$$

These solutions also obey the following N entropy balance equations :

$$\partial_t s_i^\epsilon + u^\epsilon \partial_x s_i{}^\epsilon = \epsilon \mu_i \frac{\gamma_i - 1}{\rho^{\epsilon \gamma_i}} (\partial_x u^\epsilon)^2 \geq 0, \quad i = 1, ..., N, \tag{3.3}$$

where the specific entropies are given by $s_i = p_i / \rho^{\gamma_i}$ for $i = 1, ..., N$.

Note from (3.3) that classical non linear transforms in the s_i yield further additional balance equations for governing $\varphi(s_1, ..., s_N)$ where $\varphi : I\!R^N \to I\!R$ denotes any given arbitrary smooth function. But without specific assumptions on the γ_i or the μ_i (see [19] for the details), none of them gives rise to an additional conservation law. The lack of a fully conservative reformulation of (3.1) has been shown in [10] when $N = 2$ and then in [19] for arbitrary $N \geq 2$ to be responsible for the negative issue that classical splitting methods for (3.1) produce approximate solutions which grossly disagree with the expected exact ones. The origin of such a failure can be understood when addressing the equivalent form :

$$\left\{\begin{array}{l} \partial_t\rho^\epsilon + \partial_x(\rho u)^\epsilon = 0, \\ \partial_t(\rho u)^\epsilon + \partial_x(\rho u^2 + p)^\epsilon = \epsilon\partial_x(\mu\partial_x u^\epsilon), \\ \partial_t(\rho E)^\epsilon + \partial_x(\rho H u)^\epsilon = \epsilon\partial_x(\mu u^\epsilon\partial_x u^\epsilon), \\ \partial_t s_i^\epsilon + u^\epsilon\partial_x {s_i}^\epsilon = \epsilon\mu_i\frac{\gamma_i-1}{\rho^{\epsilon\gamma_i}}(\partial_x u^\epsilon)^2, \quad i=1,...,N-1, \end{array}\right. \tag{3.4}$$

so that smooth solutions of (3.4) must obey the following additional balance equation :

$$\partial_t s_N^\epsilon + u^\epsilon\partial_x {s_N}^\epsilon = \epsilon\mu_N\frac{\gamma_N-1}{\rho^{\epsilon\gamma_N}}(\partial_x u^\epsilon)^2. \tag{3.5}$$

Here the specific entropy s_N must be obviously understood as a non linear function of the unknown $\mathbf{u} := (\rho, \rho u, \rho E, \{s_i\}_{i=1,...,N-1})$. After Berthon [10] and Chalons [19], it is essential to notice that the $(N-1)$ entropy production rates in (3.4) must be kept in balance with the last one in (3.5) according to :

$$\mu_N\frac{\rho^{\epsilon\gamma_i}}{\gamma_i-1}\{\partial_t s_i^\epsilon + u^\epsilon\partial_x s_i^\epsilon\} = \mu_i\frac{\rho^{\epsilon\gamma_N}}{\gamma_N-1}\{\partial_t s_N(\mathbf{u}^\epsilon) + u^\epsilon\partial_x s_N(\mathbf{u}^\epsilon)\}, \quad i=1,...,N-1. \tag{3.6}$$

The key issue is to notice that these identities are by contrast to (3.3) no longer ϵ-weighted but are instead driven by the ratios of the viscosities μ_i/μ_N. Notice that again, the specific entropy s_N entering (3.6) has to be understood as a non linear function of $\mathbf{u}$, with $\mathbf{u}$ solution of (3.4) so as to satisfy simultaneously the $(N-1)$ balance relations (3.6). Here stays the main difficulty. With this observation, turning considering the numerical issue, one has to evolve the approximate solution $\mathbf{u}_h$ according to (3.4) so that as a byproduct, s_N evaluated on $\mathbf{u}_h$ updates according to (3.6). For mesh scale of practical importance, viscous shock profiles stay under resolved and this lack of resolution implies in turn that the $(N-1)$ entropy production rates in (3.4) are under estimated. For prescribed updates of $(\rho, \rho u, \rho E)$, such under estimations can be seen in turn to imply that the N^{th} production rate in (3.5) is over estimated (see the proofs in Berthon [10], Chalons [19], Chalons and Coquel [20]). The specific entropies $(s_i)_h$ with $i=1,...,N-1$ and the non linear function $s_N(\mathbf{u}_h)$ cannot thus evolve in the balance put forward in (3.6). A failure in satisfying (3.6) has been proved to result in large errors between approximate and exact solutions while the authors in [10], [19] and [20] have established that solving (3.4) when enforcing for validity the $(N-1)$ relations (3.6) yields approximate solutions in fairly good agreement with exact ones. Roughly speaking, the numerical method proposed in these works can be understood as a predictor-corrector strategy which consists first in approximating the solutions of (3.4) and then, in a second step, in correcting the updated unknown $\mathbf{u}_h$ so as to satisfy the following equivalent form of (3.4) :

$$\left\{\begin{array}{l} \partial_t\rho^\epsilon + \partial_x(\rho u)^\epsilon = 0, \\ \partial_t(\rho u)^\epsilon + \partial_x(\rho u^2 + p)^\epsilon = \epsilon\partial_x(\mu\partial_x u^\epsilon), \\ \partial_t(\rho E)^\epsilon + \partial_x(\rho H u)^\epsilon = \epsilon\partial_x(\mu u^\epsilon\partial_x u^\epsilon), \\ \mu_N\frac{\rho^{\epsilon\gamma_i}}{\gamma_i-1}\{\partial_t s_i^\epsilon + u^\epsilon\partial_x s_i^\epsilon\} = \mu_i\frac{\rho^{\epsilon\gamma_N}}{\gamma_N-1}\{\partial_t s_N(\mathbf{u}^\epsilon) + u^\epsilon\partial_x s_N(\mathbf{u}^\epsilon)\}, \quad i=1,...,N-1. \end{array}\right. \tag{3.7}$$

Let us underline that the above description of the predictor-corrector strategy is rough. Indeed and to prove the relevance of this procedure, namely the positivity of all the independent pressures together with entropy inequalities, it was necessary to deal with non linear transforms in the s_i so as to obey the following convexity properties:

Lemma 10 *Let be given for $i = 1, ..., N$, N smooth strictly decreasing functions f_i : $I\!R^+ \to I\!R^+$ such that:*

$$\frac{f_i''(x)}{f_i'(x)} < -\frac{\alpha - 1}{\alpha} x, \quad \textit{for all } x > 0, \tag{3.8}$$

where we have set $\alpha = \max_{1 \le i \le N} \gamma_i$. Let us then define the N entropies $\rho S_i = \rho f_i(s_i)$. Then the following convexity properties are in order:

(i) the mapping

$$\{\rho E\}(\rho, \rho u, \{\rho S_i\}_{1 \le i \le N}) = \frac{(\rho u)^2}{2\rho} + \sum_{i=1}^{N} \frac{\rho^{\gamma_i}}{\gamma_i - 1} f_i^{-1}(\frac{\rho S_i}{\rho}) \tag{3.9}$$

is strictly convex,

(ii) the mapping in (3.9) is invertible with respect to any given entropy, say S_N, and the mapping

$$\{\rho S_N\}(\rho, \rho u, \rho E, \{\rho S_i\}_{1 \le i \le N-1}) = \rho f_N(\rho E - \frac{(\rho u)^2}{2\rho} - \sum_{i=1}^{N} \frac{\rho^{\gamma_i}}{\gamma_i - 1} f_i^{-1}(\frac{\rho S_i}{\rho})) \tag{3.10}$$

is strictly convex.

The proof of this technical result easily follows from the proof given in [10] with $N = 2$. Observe that choosing a particular family of N functions f_i just amounts to favour a given complete set of equations of state when defining the independent temperatures by:

$$T_i = -\frac{\rho^{\gamma_i - 1}}{\gamma_i - 1} (f_i^{-1})'(\frac{\rho S_i}{\rho}), \quad i = 1, ..., N. \tag{3.11}$$

For a given fixed family, the $(N-1)$ balance equations (3.6) equivalently rewrite

$$\mu_N T_i \{\partial_t (\rho S_i)^\epsilon + \partial_x (\rho S_i u)^\epsilon\} = \mu_i T_N \{\partial_t (\rho S_N)^\epsilon + \partial_x (\rho S_N u)^\epsilon\}, \quad i = 1, ..., N-1,$$

where ρS_N has to be understood as the non linear function given in (3.10). Let us underline that the condition $\alpha = \max_{1 \le i \le N} \gamma_i > 1$ precludes the natural choice $f_i = Id$ for any given $i \in \{1, ..., N\}$. Non linear transforms in the s_i have been thus considered in [10] and [20] for the sake of stability properties but at the expense of the need for solving a non linear algebraic problem at the correction step.

In the present work, we prove that the non linear algebraic problem can be bypassed when choosing $f_i = Id$ for all $i = 1, ..., N$ while still achieving the required stability properties. To circumvent the lack of convexity, we shall introduce suitable averaging procedures which no longer rely on the classical Lebesgue measure involved in the usual Godunov approach. Such averaging procedures will be easily seen to receive fairly simple evaluations when considering classical approximate Riemann solvers for the hyperbolic underlying system in (3.4). We shall in particular derive such a suitable Riemann solver when approximating the solutions of the first order underlying system in (3.4) by those of a relaxation system.

3.2 Classical averagings and stability estimates

The present section aims at briefly describing a variant of the predictor-corrector procedure developed in [20] when revisiting the classical averaging procedures in terms of stability estimates. This variant will turn out to be more convenient for our forthcoming numerical purposes and the discussion we propose below will naturally motivate the need for other suitable averaging techniques. Here, we propose to approximate the smooth solutions of (3.1) when addressing the equivalent formulation :

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + p) = \epsilon \partial_x (\mu \partial_x u), \\ \partial_t s_i + u \partial_x s_i = \epsilon \mu_i \frac{\gamma_i - 1}{\rho^{\gamma_i}} (\partial_x u)^2, \quad i = 1, ..., N, \end{cases} \tag{3.12}$$

that is to say when the total energy is understood as a non linear function of the unknown $\mathbf{u} := (\rho, \rho u, \{s_i\}_{i=1,...,N})$, so that smooth solutions of (3.12) obey the following non trivial additional equation :

$$\partial_t \rho E + \partial_x \rho H u = \epsilon \partial_x (\mu u \partial_x u).$$

The reason for considering the total energy as a non linear function of the unknown will arise naturally in the next section. Note that for simplicity in the notations, the superscript ϵ has been (and will be hereafter) omitted.
Let be given $\mathbf{u}_h(x, t^n)$ some approximate solution of (3.12) at time t^n. This approximate solution will be classically assumed to be piecewise constant with :

$$\mathbf{u}_h(x, t^n) = \mathbf{u}_j^n, \quad \text{for all} \quad x \in (x_{j-1/2}, x_{j+1/2}), \quad j \in Z,$$

where for simplicity $(x_{j+1/2} - x_{j-1/2}) = \Delta x > 0$. In order to advance it to the next time level t^{n+1}, we first propose two steps based on a splitting strategy between the convective and the diffusive part. Despite this procedure sounds quite natural, we aim at proving that a non linear correction step is needed when invoking classical Riemann solvers for solving the convective part. The relevance of the correction step will only follow when focusing on entropy satisfying Riemann solvers.

Step 1 ($t^n \to t^{n+1=}$)
Solving the underlying first order system in (3.12) clearly amounts to consider weak solutions of the following non linear hyperbolic system :

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + p) = 0, \\ \partial_t s_i + u \partial_x s_i = 0, \quad i = 1, ..., N, \end{cases} \tag{3.13}$$

but under this form, the candidate for defining a Lax entropy pair, namely the total energy, has no natural convexity property. This issue seems to rise a first difficulty. Indeed, weak solutions need to be selected according to some entropy criterion. But other suitable criteria, equivalent (in the present setting) to the requirement of an entropy inequality, are actually at hand to uniquely define Riemann solutions of (3.13) (see indeed Lax [84]). The very difficulty actually arises from the fact that we shall need a control over the discrete energy understood as a function of the approximate solution in order to propose a relevant numerical procedure. The required control will be seen to be directly inferred from a discrete energy like inequality. As it is well-known, such a control heavily relies on the convexity property of the total energy with respect to the (discrete) unknown. This naturally motivates us to solve instead of (3.13) the following Cauchy problem with initial data $\overline{\mathbf{w}}_h(x,t^n) := (\rho, \rho u, \{\rho f_i(s_i)\}_{i=1,...,N})_h(x,t^n)$:

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x(\rho u^2 + p) = 0, \\ \partial_t \rho S_i + \partial_x \rho S_i u = 0, \quad i = 1, ..., N, \end{cases} \tag{3.14}$$

where we have considered a fixed family of N non linear functions f_i given under the requirement (3.8). As a consequence, the total energy ρE is convex in the new unknown $\overline{\mathbf{w}}$, and the pair $(\rho E, \rho H u)$ can now serve as a Lax entropy pair for selecting the weak solutions of (3.14) according to :

$$\partial_t \rho E(\overline{\mathbf{w}}) + \partial_x \rho H u(\overline{\mathbf{w}}) \leq 0. \tag{3.15}$$

It is worthy to noticing that entropy weak solutions of (3.14) actually obey in the usual sense of the distributions the following transport equations for each of the specific entropies :

$$\partial_t s_i + u \partial_x s_i = 0, \quad \mathcal{D}'. \tag{3.16}$$

Indeed, it can be easily checked from the Rankine-Hugoniot conditions associated with (3.14) that the solutions u and S_i (so that s_i) cannot be discontinuous simultaneously. So that the non conservative products entering (3.16) are never ambiguous. The transport equations property, valid for all $i = 1, ..., N$, is actually the main reason for rejecting the total energy as a non linear function of the unknown and will be used intensively in the next section to introduce new averaging techniques. In order to solve the Cauchy problem (3.14)-(3.15) for small times, we make use of the celebrated Godunov method so as to consider a sequence of non interacting exact Riemann solutions under the CFL restriction

$$\frac{\Delta t}{\Delta x} \max_{\overline{\mathbf{w}}}(|u - c|(\overline{\mathbf{w}}), |u + c|(\overline{\mathbf{w}})) \leq \frac{1}{2}, \tag{3.17}$$

for all the $\overline{\mathbf{w}}$ under consideration. For the times under consideration, the solution is then obtained as the superposition of these Riemann solutions. With this solution, let us now invoke the classical averagings in order to produce a piecewise constant update :

$$\overline{\mathbf{w}}_j^{n+1=} := \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \overline{\mathbf{w}}_h(x, t^n + \Delta t) dx, \quad j \in Z. \tag{3.18}$$

Note that by the monotonicity assumptions on each of the f_i, the piecewise approximate solution $\mathbf{u}_h(x, t^{n+1=})$ is uniquely recovered from (3.18). But from the non linearity in the f_i, observe that the resulting update of each of the s_i is actually such that :

$$(s_i)_j^{n+1=} := \{s_i\}(\overline{\mathbf{w}}_j^{n+1=}) \neq \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \{s_i\}(\overline{\mathbf{w}}_h(x, t^n + \Delta t)) dx =$$

$$(s_i)_j^n - \frac{\Delta t}{\Delta x}(u^\star)^{j+1/2,-}((s_i)_{j+1}^n - (s_i)_j^n) - \frac{\Delta t}{\Delta x}(u^\star)^{j-1/2,+}((s_i)_j^n - (s_i)_{j-1}^n), \quad j \in Z, \tag{3.19}$$

where $(u^\star)^{j+1/2}$ denotes the speed of the contact discontinuity in the Riemann problem at the interface $x_{j+1/2}$. The identity in the right-hand side of (3.19) easily results from the transport equations (3.16). As a consequence, we infer that in general:

$$(s_i)_j^{n+1=} - (s_i)_j^n + \frac{\Delta t}{\Delta x}(u^\star)_-^{j+1/2}((s_i)_{j+1}^n - (s_i)_j^n) + \frac{\Delta t}{\Delta x}(u^\star)_+^{j-1/2}((s_i)_j^n - (s_i)_{j-1}^n) \neq 0, \tag{3.20}$$

and it is not known if $(s_i)_j^{n+1=}$ solves a discrete version of (3.16). Here the benefit of averaging each of the ρS_i instead of the s_i is that the following in-cell total energy inequality holds under the CFL restriction (3.17):

$$\{\rho E\}(\overline{\mathbf{w}}_j^{n+1=}) \leq (\rho E)_j^n - \frac{\Delta t}{\Delta x}\Delta\{(\rho E + p)u\}_{j+1/2}^n, \quad j \in Z, \tag{3.21}$$

by the well-known Jensen inequality applied to the convex mapping $\{\rho E\}(\overline{\mathbf{w}})$, together with the following discrete conservation laws:

$$(\rho S_i)_j^{n+1=} = (\rho S_i)_j^n - \frac{\Delta t}{\Delta x}\Delta\{\rho S_i u\}_{j+1/2}^n, \quad j \in Z, \tag{3.22}$$

for some quite classical definitions of the numerical flux functions.

Step 2 ($t^{n+1=} \to t^{n+1-}$)
Let us consider $\overline{\mathbf{w}}_h(x, t^{n+1=})$ the solution of (3.14)-(3.15) at time Δt. This one now naturally serves as initial data for solving the viscous perturbations for times $t \in [0, \Delta t]$:

$$\begin{cases} \partial_t \rho = 0, \\ \partial_t \rho u = \epsilon \partial_x(\mu \partial_x u), \\ \partial_t \rho S_i = -\epsilon \frac{\mu_i}{T_i}(\partial_x u)^2, \quad i = 1, ..., N, \end{cases} \tag{3.23}$$

for the temperatures T_i defined in (3.11), that is to say since the momentum equation is decoupled from the entropies:

$$\begin{cases} \partial_t \rho = 0, \\ \partial_t \rho u = \epsilon \partial_x(\mu \partial_x u), \\ \partial_t s_i = \epsilon \mu_i \frac{\gamma_i - 1}{(\rho)^{\gamma_i}}(\partial_x u)^2, \quad i = 1, ..., N. \end{cases} \tag{3.24}$$

The numerical approximation of the solution of (3.23), or equivalently of (3.24), has received a particular attention in [20]. Extending the work by Berthon in [10], these authors have proved that convenient finite differences formulae can be proposed so that the total energy, understood as a function of the updated unknown, evolves in time consistently with:

$$\partial_t \rho E = \epsilon \partial_x(\mu u \partial_x u), \tag{3.25}$$

while the N specific entropies obey by construction a discrete version of:

$$\mu_N \frac{\rho^{\gamma_i}}{\gamma_i - 1}\partial_t s_i = \mu_i \frac{\rho^{\gamma_N}}{\gamma_N - 1}\partial_t s_N, \quad i = 1, ..., N-1. \tag{3.26}$$

More precisely, the solution of (3.24) at time Δt is approximated by:

$$\begin{cases} \rho_j^{n+1-} = \rho_j^{n+1=}, \\ (\rho u)_j^{n+1-} = (\rho u)_j^{n+1=} + \Delta t \overline{\partial_x(\mu \partial_x u)}_j^{n+1-}, \\ (s_i)_j^{n+1-} = (s_i)_j^{n+1=} + \Delta t \mu_i \frac{\gamma_i - 1}{(\rho_j^{n+1-})^{\gamma_i}} \overline{(\partial_x u)^2}_j^{n+1-}, \quad i = 1, ..., N, \end{cases} \tag{3.27}$$

where the precise definition of the discrete operator $\overline{\partial_x(\mu\partial_x u)}_j^{n+1-}$, given in [20] (see also [10]), results from a conservative implicit in time method for solving u. Once this implicit scheme is solved, the discrete operator $\overline{(\partial_x u)^2}_j^{n+1-}$ can receive an explicit definition so that in one hand $\overline{(\partial_x u)^2}_j^{n+1-} \geq 0$, while on the second hand:

$$\{\rho E\}(\rho_j^{n+1-}, (\rho u)_j^{n+1-}, \{(s_i)_j^{n+1-}\}_{i=1,\dots,N}) = (\rho E)_j^{n+1=} + \Delta t \overline{\partial_x(\mu u \partial_x u)}_j^{n+1-}, \quad (3.28)$$

where the resulting discrete quantity $\overline{\partial_x(\mu u \partial_x u)}_j^{n+1-}$ is a consistent and conservative discretization of the expected exact operator (see [20] for the details). Let us emphasize that the N last updating formulae in (3.27) easily imply for all $i = 1, \dots, N-1$:

$$\mu_N \frac{(\rho_j^{n+1-})^{\gamma_i}}{\gamma_i - 1}\{(s_i)_j^{n+1-} - (s_i)_j^{n+1=}\} = \mu_i \frac{(\rho_j^{n+1-})^{\gamma_N}}{\gamma_N - 1}\{(s_N)_j^{n+1-} - (s_N)_j^{n+1=}\}, \quad (3.29)$$

which is nothing but a consistent discrete form of (3.26). To conclude this brief presentation, it is worthy to noticing that on the ground of the discrete operators we have just introduced, we immediately get from the updating formulae (3.27) equivalent discrete formulae for approximating the system (3.23):

$$(\rho S_i)_j^{n+1-} = (\rho S_i)_j^{n+1=} - \Delta t \frac{\mu_i}{\overline{T}_{ij}^{n+1-}} \overline{(\partial_x u)^2}_j^{n+1-}, \quad i = 1, \dots N,$$

where

$$\overline{T}_{ij}^{n+1-} = \frac{(\rho_j^{n+1-})^{\gamma_i - 1}}{\gamma_i - 1} \begin{cases} (f_i^{-1})'(f_i((s_i)_j^{n+1=})) \text{ if } f_i((s_i)_j^{n+1-}) = f_i((s_i)_j^{n+1=}), \\ \frac{f_i^{-1} o f_i((s_i)_j^{n+1-}) - f_i^{-1} o f_i((s_i)_j^{n+1=})}{f_i((s_i)_j^{n+1-}) - f_i((s_i)_j^{n+1=})} \text{ otherwise.} \end{cases}$$

As a consequence of the equivalence, we have:

$$\{\rho E\}(\rho_j^{n+1-}, (\rho u)_j^{n+1-}, \{(\rho S_i)_j^{n+1-}\}_{i=1,\dots,N}) = (\rho E)_j^{n+1=} + \Delta t \overline{\partial_x(\mu u \partial_x u)}_j^{n+1-}, \quad (3.30)$$

while on the same time, for $i = 1, \dots, N-1$:

$$\mu_N \overline{T}_{ij}^{n+1-}\{(\rho S_i)_j^{n+1-} - (\rho S_i)_j^{n+1=}\} = \mu_i \overline{T}_{Nj}^{n+1-}\{(\rho S_N)_j^{n+1-} - (\rho S_N)_j^{n+1=}\}. \quad (3.31)$$

These remarks will be useful below.

Step 3 ($t^{n+1-} \to t^{n+1}$) : Correction procedure
As they stand, the first two steps we have just described provide us with a formally consistent approximation of the solutions of (3.1). But these two steps cannot result in a relevant numerical method since by construction, the L^1 norm of the total energy is (strictly) decreasing in time by virtue of the next inequality:

$$\{\rho E\}(\overline{\mathbf{w}}_j^{n+1-}) \leq (\rho E)_j^n - \frac{\Delta t}{\Delta x}\Delta\{(\rho E + p)u\}_{j+1/2}^n + \Delta t \overline{\partial_x(\mu u \partial_x u)}_j^{n+1-}, \quad (3.32)$$

as a direct consequence of the inequality (3.20) and the equality (3.28). A correction procedure for restoring the conservation of the total energy at the discret level is therefore needed. Despite that inequality (3.32) is responsible for this need, the same inequality (3.32) actually authorizes the correction step to take place, that is to say when preserving the positivity of each of the entropies. Our main motivation here, after [10] and [20], is to propose a correction step so that the final updated unknown satisfies the $(N-1)$ identities (3.31) at the discrete level while preserving the conservation of the total energy. In that aim, we are led to keep unchanged the definitions of the density and the momentum :

$$\rho_j^{n+1} = \rho_j^{n+1-}, \quad (\rho u)_j^{n+1} = (\rho u)_j^{n+1-}, \quad j \in Z,$$

while redefining the N entropies so that in the one hand :

$$(\rho E)_j^{n+1} = (\rho E)_j^n - \frac{\Delta t}{\Delta x}\Delta\{(\rho E + p)u\}_{j+1/2}^n + \Delta t \overline{\partial_x(\mu u \partial_x u)}_j^{n+1-},$$

while on the second hand we arrive at a consistent discretization of :

$$\mu_N \frac{\rho^{\gamma_i}}{\gamma_i - 1}\{\partial_t s_i + u\partial_x s_i\} = \mu_i \frac{\rho^{\gamma_N}}{\gamma_N - 1}\{\partial_t s_N(\mathbf{u}) + u\partial_x s_N(\mathbf{u})\}, \quad i = 1, ..., N-1.$$

In spite of the validity of the $(N-1)$ relations (3.29) :

$$\mu_N \frac{(\rho_j^{n+1-})^{\gamma_i}}{\gamma_i - 1}\{(s_i)_j^{n+1-} - (s_i)_j^{n+1=}\} = \mu_i \frac{(\rho_j^{n+1-})^{\gamma_N}}{\gamma_N - 1}\{(s_N)_j^{n+1-} - (s_N)_j^{n+1=}\},$$

the inequalities (3.20) :

$$(s_i)_j^{n+1=} \neq (s_i)_j^n - \frac{\Delta t}{\Delta x}(u^\star)^{j+1/2,-}((s_i)_{j+1}^n - (s_i)_j^n) - \frac{\Delta t}{\Delta x}(u^\star)^{j-1/2,+}((s_i)_j^n - (s_i)_{j-1}^n),$$

preclude the specific entropies s_i from being suitable candidates. We are thus necessarily left with the non linear form of the entropies S_i which by definition obey similar balance relations :

$$\mu_N T_i\{\partial_t \rho S_i + \partial_x \rho S_i u\} = \mu_i T_N\{\partial_t \rho S_N + \partial_x \rho S_N u\}, \quad i = 1, ..., N-1. \tag{3.33}$$

Note that both set of equalities (3.31) and (3.22) now make possible the derivation of a correction step consistent with (3.33). Indeed, it suffices to redefine $(\rho S_i)_j^{n+1}$ as the solution of the following $(N-1)$ equations :

$$\mu_N \overline{T_i}_j^{n+1-}\{(\rho S_i)_j^{n+1} - (\rho S_i)_j^{n+1=}\} = \mu_i \overline{T_N}_j^{n+1-}\{(\rho S_N)_j^{n+1} - (\rho S_N)_j^{n+1=}\},$$

with $i = 1, ..., N-1$, supplemented with the following non linear equation :

$$\{\rho E\}(\rho_j^{n+1}, (\rho u)_j^{n+1}, \{(\rho S_i)_j^{n+1}\}_{i=1,\dots,N}) = (\rho E)_j^{n+1}. \tag{3.34}$$

Invoking the energy inequality $(\rho E)_j^{n+1-} \leq (\rho E)_j^{n+1}$, a proof similar to the one proposed in [20] gives the existence of a unique solution with the property that

$$0 < (\rho S_i)_j^{n+1} \leq (\rho S_i)_j^{n+1=}, \quad i = 1, ..., N. \tag{3.35}$$

In practice, an iterative procedure is required to solve (3.34). As a consequence of these estimates, each of the partial pressures stays positive as soon as $\rho_j^{n+1} > 0$ while the following entropy inequalities are met (see (3.22)) :

$$(\rho S_i)_j^{n+1} - (\rho S_i)_j^n + \frac{\Delta t}{\Delta x}\Delta\{\rho S_i u\}_{j+1/2}^n \leq 0, \quad i = 1, ..., N.$$

To conclude this section, let us again underline that the crucial stability estimates (3.35) only hold true because of the energy inequality (3.32). The validity of this inequality relies on using classical averagings for non linear versions of the specific entropies s_i. We show in the next section how to suitably average directly the s_i so as to recover the required energy inequality. The benefit will be a fully explicit correction step by opposition to (3.34) while still achieving the same stability properties.

3.3 A new averaging procedure

The predictor-corrector method we now introduce comes along the main lines of the three steps we have described in the previous section. Essentially, the first step devoted to the approximation of the underlying non linear hyperbolic system will receive a suitable modification so as to end up with a fully explicit but still relevant correction step. The second step will remain unchanged.
Here assuming again a piecewise constant approximate solution $\mathbf{u}_h(x,t^n)$ to be given at the date t^n, this one is advanced to the next intermediate time level $t^{n+1=}$ when solving the Cauchy problem (3.14) in the $\overline{\mathbf{w}} := (\rho, \rho u, \{\rho S_i\}_{i=1,\dots,N})$ variable. So that again the total energy plays the role of a Lax entropy for selecting the weak solutions under consideration. The solution $\overline{\mathbf{w}}(x,t)$ of this problem is obtained under the CFL condition (3.17) as the superposition of non interacting Riemann solutions for times t less than Δt. Here instead of projecting onto the constants the ρS_i at time $t^{n+1=} = t^n + \Delta t$, we suggest to consider suitable averagings of $s_i(\overline{\mathbf{w}}(x,t^{n+1=}))$. Note that the classical projections onto the constants:

$$(s_i)_j^{n+1=} = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} s_i(\overline{\mathbf{w}}(x,t^{n+1=}))dx, \quad i = 1,\dots,N,$$

cannot provide us with a relevant definition to make the total energy to decrease simply because of the lack of convexity of the mapping $\mathbf{u} \to \{\rho E\}(\mathbf{u})$. To bypass this lack of convexity, we instead suggest to consider the following averaging procedures.

Definition 4 *Under the CFL condition (3.17), the solution $\overline{\mathbf{w}}$ of the Cauchy problem (3.14) is given at time $t^{n+1=}$ the following averaging formulae over each cell $(x_{j-1/2}, x_{j+1/2})$, $j \in Z$:*

$$\begin{aligned}
\rho_j^{n+1=} &= \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{\rho\}(\overline{\mathbf{w}}(x,t^{n+1=}))dx, \\
(\rho u)_j^{n+1=} &= \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{\rho u\}(\overline{\mathbf{w}}(x,t^{n+1=}))dx,
\end{aligned} \tag{3.36}$$

while the specific entropies are projected onto the constants according to:

$$(s_i)_j^{n+1=} = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{s_i\}(\overline{\mathbf{w}}(x,t^{n+1=})) \times \frac{\rho^{\gamma_i}}{\overline{\rho^{\gamma_i}}_j^{n+1=}}(\overline{\mathbf{w}}(x,t^{n+1=}))dx, \quad j \in Z, \tag{3.37}$$

where we have set:

$$\overline{\rho^{\gamma_i}}_j^{n+1=} = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{\rho^{\gamma_i}\}(\overline{\mathbf{w}}(x,t^{n+1=}))dx.$$

In other words, both the density and the momentum receive classical averaging formulae based on the Lebesgue measure, while by contrast the specific entropies are averaged when invoking probability-like measures: *i.e.* positive Borel measures with unit total mass. Consequently, the positivity of each of the specific entropies is preserved by construction. A well-known consequence of the classical averagings (3.36) is that both density and momentum can be given an equivalent updating formula in full conservation form:

$$\begin{aligned} \rho_j^{n+1=} &= \rho_j^n - \tfrac{\Delta t}{\Delta x}\Delta\{\rho u\}_{j+1/2}^n, \\ (\rho u)_j^{n+1=} &= (\rho u)_j^n - \tfrac{\Delta t}{\Delta x}\Delta\{\rho u^2+p\}_{j+1/2}^n, \end{aligned} \tag{3.38}$$

where classically $\{\rho u\}_{j+1/2}^n = \{\rho u\}(\overline{\mathbf{w}}(0^+,\mathbf{u}_j^n,\mathbf{u}_{j+1}^n))$ with a similar definition for $\{\rho u^2+p\}_{j+1/2}^n$. By contrast, the original projection formulae (3.37) cannot give rise to a discrete conservation form for governing the specific entropies since these variables already obey at the PDE level the following equation:

$$\partial_t s_i + u\partial_x s_i = 0, \quad \mathcal{D}', \quad i=1,...,N. \tag{3.39}$$

We shall prove later on that the averaging formulae (3.37) actually yield a discrete version of (3.39) under the form:

$$(s_i)_j^{n+1=} = (s_i)_j^n - \frac{\Delta t}{\Delta x}v_i^{j-1/2,+}\{(s_i)_j^n-(s_i)_{j-1}^n\} - \frac{\Delta t}{\Delta x}v_i^{j+1/2,-}\{(s_i)_{j+1}^n-(s_i)_j^n\}, \tag{3.40}$$

where the discrete velocity $v_i^{j+1/2*}$ will be given a consistent definition.
The unusual projection formulae (3.37) will be given further insight hereafter but they are primarily dictated by the following result:

Lemma 11 *Let the approximate solution be advanced to the next time level $t^{n+1=}$ according to (3.36)-(3.37). Then, under the CFL condition (3.17), the following in-cell energy inequalities are satisfied for all $j \in Z$:*

$$\{\rho E\}(\mathbf{u}_j^{n+1=}) - (\rho E)_j^n + \lambda\Delta\{\rho\mathcal{H}u\}^+(\mathbf{u}_j^n,\mathbf{u}_{j+1}^n) \le 0, \tag{3.41}$$

where classically:

$$\{\rho\mathcal{H}u\}^+(\mathbf{u}_j^n,\mathbf{u}_{j+1}^n) = \{(\rho E+p)u\}(\overline{\mathbf{w}}(0^+,\mathbf{u}_j^n,\mathbf{u}_{j+1}^n)). \tag{3.42}$$

Proof Classical considerations (see [57]) imply from the Lax entropy inequality (3.15) the following in-cell inequality:

$$\frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}}\{\rho E\}(\overline{\mathbf{w}}(x,t^{n+1=}))dx \le (\rho E)_j^n - \lambda\Delta\{\rho\mathcal{H}u\}^+(\mathbf{u}_j^n,\mathbf{u}_{j+1}^n), \tag{3.43}$$

where the numerical energy flux is given by (3.42). Let us now notice that the definitions of $(s_i)_j^{n+1=}$ and $\overline{\rho^{\gamma_i}}_j^{n+1=}$ for all $i=1,...,N$ lead to

$$\begin{gathered} \int_{x_{j-1/2}}^{x_{j+1/2}}\{\rho E\}(\overline{\mathbf{w}}(x,t^{n+1=}))dx = \\ \int_{x_{j-1/2}}^{x_{j+1/2}}\{\frac{(\rho u)^2}{2\rho}\}(\overline{\mathbf{w}}(x,t^{n+1=}))dx + \sum_{i=1}^N\int_{x_{j-1/2}}^{x_{j+1/2}}\frac{\{p_i\}}{\gamma_i-1}(\overline{\mathbf{w}}(x,t^{n+1=}))dx = \\ \int_{x_{j-1/2}}^{x_{j+1/2}}\{\frac{(\rho u)^2}{2\rho}\}(\overline{\mathbf{w}}(x,t^{n+1=}))dx + \sum_{i=1}^N\frac{(s_i)_j^{n+1=}}{\gamma_i-1}\times\int_{x_{j-1/2}}^{x_{j+1/2}}\{\rho^{\gamma_i}\}(\overline{\mathbf{w}}(x,t^{n+1=}))dx. \end{gathered} \tag{3.44}$$

In the one hand, the convexity of the mapping $(\rho, \rho u) \to (\rho u)^2/2\rho$ ensures from the well-known Jensen's inequality :

$$\frac{((\rho u)_j^{n+1=})^2}{2\rho_j^{n+1=}} \le \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{\frac{(\rho u)^2}{2\rho}\}(\overline{\mathbf{w}}(x, t^{n+1=}))dx. \tag{3.45}$$

In the second hand, the convexity of the mapping $\rho \to \rho^{\gamma_i}$ ($\gamma_i > 1$) gives again because of the Jensen's inequality :

$$(\rho_j^{n+1=})^{\gamma_i} \le \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{\rho^{\gamma_i}\}(\overline{\mathbf{w}}(x, t^{n+1=}))dx, \quad i = 1, ..., N.$$

Since the averaging formulae (3.37) preserve the positivity of the specific entropies, we deduce the required estimate :

$$\begin{aligned} \{\rho E\}(\mathbf{u}_j^{n+1=}) &:= \frac{((\rho u)_j^{n+1=})^2}{2\rho_j^{n+1=}} + \sum_{i=1}^{N} \frac{(\rho_j^{n+1=})^{\gamma_i}}{\gamma_i - 1}(s_i)_j^{n+1=} \\ &\le \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{\rho E\}(\overline{\mathbf{w}}(x, t^{n+1=}))dx. \end{aligned} \tag{3.46}$$

This completes the proof.
The energy inequality (3.41) thus makes relevant the unusual averagings (3.36)-(3.37) to define a discrete method for approximating the weak solutions of the non linear hyperbolic system (3.14) in the first step of our splitting strategy. We are now in a position to address the second step devoted to the viscous perturbations. The second step described in the previous section is kept unchanged. Using the same notations, we thus deduce the following statement :

Corollary 3 *The first two steps in the splitting method yield the following updating formulae :*

$$\left\{ \begin{array}{l} \rho_j^{n+1-} = \rho_j^n - \frac{\Delta t}{\Delta x}\Delta\{\rho u\}_{j+1/2}^n, \\ (\rho u)_j^{n+1-} = (\rho u)_j^n - \frac{\Delta t}{\Delta x}\Delta\{\rho u^2 + p\}_{j+1/2}^n + \Delta t\overline{\partial_x(\mu\partial_x u)}_j^{n+1-}, \\ (s_i)_j^{n+1-} = (s_i)_j^{n+1=} + \Delta t\frac{\gamma_i - 1}{(\rho_j^{n+1-})^{\gamma_i}}\overline{(\partial_x u)^2}_j^{n+1-}, \quad i = 1, ..., N, \end{array} \right. \tag{3.47}$$

where

$$\overline{(\partial_x u)^2}_j^{n+1-} \ge 0, \quad \text{and} \quad \overline{\partial_x(\mu\partial_x u)}_j^{n+1-}, \tag{3.48}$$

are given consistent finite differences formulae in [20]. Under the CFL condition (3.17), the next energy-like inequality holds on each cell $(x_{j-1/2}, x_{j+1/2})$ *:*

$$\{\rho E\}(\mathbf{u}_j^{n+1-}) \le (\rho E)_j^n - \lambda\Delta\{\rho\mathcal{H}u\}^+(\mathbf{u}_j^n, \mathbf{u}_{j+1}^n) + \Delta t\overline{\partial_x(\mu u\partial_x u)}_j^{n+1-}, \tag{3.49}$$

where the last term in (3.49) receives by construction a discrete formula in conservation form.

Let us indeed recall that the discrete formulae (3.47) are such that by construction, the total energy obeys with equality the following identity in conservation form (see [20] and [10] for the details) :

$$\{\rho E\}(\rho_j^{n+1-}, (\rho u)_j^{n+1-}, \{(s_i)_j^{n+1-}\}_{i=1,...,N}) = (\rho E)_j^{n+1=} + \Delta t\overline{\partial_x(\mu u\partial_x u)}_j^{n+1-}.$$

With this respect, the inequality (3.49) just follows from inequality (3.41).
The proof of Lemma 11 clearly indicates that the energy inequality (3.41) holds strictly generally speaking so that also (3.49). The formulae (3.36)-(3.37)-(3.47) thus make the L^1 norm of the total energy to decrease (strictly) in time. Here again, a correction step to enforce the conservation of the total energy is required. Following the guidelines of the rough description of the method, we are led to keep unchanged the updating formulae for the density and the momentum :

$$\rho_j^{n+1} = \rho_j^{n+1-}, \quad (\rho u)_j^{n+1} = (\rho u)_j^{n+1-}, \quad \text{for all } j \in Z, \tag{3.50}$$

while defining the N specific entropies $(s_i)_j^{n+1}$ as the solutions of the following $(N-1)$ equations :

$$\mu_N \frac{(\rho_j^{n+1})^{\gamma_i}}{\gamma_i - 1}\{(s_i)_j^{n+1} - (s_i)_j^{n+1=}\} = \mu_i \frac{(\rho_j^{n+1})^{\gamma_N}}{\gamma_N - 1}\{(s_N)_j^{n+1} - (s_N)_j^{n+1=}\}, \tag{3.51}$$

supplemented with

$$\sum_{i=1}^{N} \frac{(\rho_j^{n+1})^{\gamma_i}}{\gamma_i - 1}(s_i)_j^{n+1} = (\rho E)_j^{n+1} - \frac{((\rho u)_j^{n+1})^2}{2\rho_j^{n+1}}. \tag{3.52}$$

Here the total energy $(\rho E)_j^{n+1}$ is defined by

$$(\rho E)_j^{n+1} := (\rho E)_j^n - \lambda\Delta\{\rho\mathcal{H}u\}^+(\mathbf{u}_j^n, \mathbf{u}_{j+1}^n) + \Delta t\overline{\partial_x(\mu u\partial_x u)}_j^{n+1-}, \tag{3.53}$$

so as to restore its conservation property at the discrete level. Let us underline that the algebraic problem (3.51)-(3.52) is linear so that the next easy result, which proof is left to the reader, holds :

Proposition 14 *The linear system (3.51) in the unknown $\{(s_i)_j^{n+1}\}_{1\leq i\leq N}$ admits an unique solution explicitly given by :*

$$\begin{gathered}(s_i)_j^{n+1} = (s_i)_j^{n+1=} + \\ \frac{\mu_i(\gamma_i - 1)}{(\rho_j^{n+1})^{\gamma_i}\sum_{l=1}^N \mu_l}((\rho E)_j^{n+1} - \frac{((\rho u)_j^{n+1})^2}{2\rho_j^{n+1}} - \sum_{l=1}^{N}\frac{(\rho_j^{n+1})^{\gamma_l}}{\gamma_l - 1}(s_l)_j^{n+1=}),\end{gathered} \tag{3.54}$$

for all $i = 1, \ldots, N$.
The correction procedure is thus fully explicit by contrast to the technique developed in the previous section. But here again, as expected, the discrete energy inequality (3.49) is essentially responsible for proving the relevance of the present correction step :

Theorem 6 *Under the CFL condition (3.17), the solution (3.54) satisfies the following maximum principles :*

$$(s_i)_j^{n+1} \geq (s_i)_j^{n+1=} \geq \min((s_i)_{j-1}^n, (s_i)_j^n, (s_i)_{j+1}^n), \quad j \in Z, \quad i = 1, ..., N. \tag{3.55}$$

As a consequence, the numerical method (3.36)-(3.37)-(3.47)-(3.50)-(3.53)-(3.54) preserves the positivity of the specific entropies and hence the positivity of the partial pressures :

$$(p_i)_j^{n+1} := (\rho_j^{n+1})^{\gamma_i}(s_i)_j^{n+1} \geq 0, \quad i = 1, ..., N.$$

Proof Note first that

$$\frac{((\rho u)_j^{n+1})^2}{2\rho_j^{n+1}} + \sum_{l=1}^{N} \frac{(\rho_j^{n+1})^{\gamma_l}}{\gamma_l - 1}(s_l)_j^{n+1=} \leq \{\rho E\}(\mathbf{u}_j^{n+1-}),$$

since we have $(s_i)_j^{n+1=} \leq (s_i)_j^{n+1-}$ by the positivity property of $\overline{(\partial_x u)^2}_j^{n+1-}$ reported in the Corollary 3. As a consequence, the energy inequality (3.49) valid under the CFL condition (3.17) just recasts as:

$$\frac{((\rho u)_j^{n+1})^2}{2\rho_j^{n+1}} + \sum_{l=1}^{N} \frac{(\rho_j^{n+1})^{\gamma_l}}{\gamma_l - 1}(s_l)_j^{n+1=} \leq \{\rho E\}(\mathbf{u}_j^{n+1-}) \leq (\rho E)_j^{n+1},$$

by definition of $(\rho E)_j^{n+1}$. Invoking the explicit formula (3.54) for defining each of the specific entropies $(s_i)_j^{n+1}$, the positivity of the viscosity coefficients easily ensures that

$$(s_i)_j^{n+1} \geq (s_i)_j^{n+1=}.$$

The proof will be completed under some suitable CFL condition when proving for validity the convex decomposition (3.40). This will be the matter of the Proposition 16 given hereafter.
The following statement summarizes the main properties of the prediction-correction method we have just proposed.

Corollary 4 *The method (3.36)-(3.37)-(3.47)-(3.50)-(3.53)-(3.54) provides an approximate discrete solution which satisfies the following identities:*

$$\begin{cases} \rho_j^{n+1} = \rho_j^n - \frac{\Delta t}{\Delta x}\Delta\{\rho u\}_{j+1/2}^n, \\ (\rho u)_j^{n+1} = (\rho u)_j^n - \frac{\Delta t}{\Delta x}\Delta\{\rho u^2 + p\}_{j+1/2}^n + \Delta t\overline{\partial_x(\mu\partial_x u)}_j^{n+1-}, \\ (\rho E)_j^{n+1} = (\rho E)_j^n - \frac{\Delta t}{\Delta x}\Delta\{(\rho E + p)u\}_{j+1/2}^n + \Delta t\overline{\partial_x(\mu u\partial_x u)}_j^{n+1-}, \\ \mu_N \dfrac{(\rho_j^{n+1})^{\gamma_i}}{\gamma_i - 1}\{(s_i)_j^{n+1} - (s_i)_j^{n+1=}\} = \mu_i \dfrac{(\rho_j^{n+1})^{\gamma_N}}{\gamma_N - 1}\{(s_N)_j^{n+1} - (s_N)_j^{n+1=}\}, \end{cases} \tag{3.56}$$

where with $i = 1, ..., N$:

$$(s_i)_j^{n+1=} := (s_i)_j^n - \frac{\Delta t}{\Delta x}v_i^{j-1/2,+}\{(s_i)_j^n - (s_i)_{j-1}^n\} - \frac{\Delta t}{\Delta x}v_i^{j+1/2,-}\{(s_i)_{j+1}^n - (s_i)_j^n\}, \tag{3.57}$$

for some consistent averaged values $v_i^{j+1/2}$ of the velocity u defined in (3.67) below. In addition, the following N entropy-like inequalities are valid under the CFL condition (3.17):

$$(s_i)_j^{n+1} - (s_i)_j^n + \frac{\Delta t}{\Delta x}v_i^{j-1/2,+}\{(s_i)_j^n - (s_i)_{j-1}^n\} + \frac{\Delta t}{\Delta x}v_i^{j+1/2,-}\{(s_i)_{j+1}^n - (s_i)_j^n\} \geq 0. \tag{3.58}$$

In other words, the proposed method yields a consistent discretization of the equivalent system (3.7) based on the $(N-1)$ generalized jump conditions (3.6). Let us emphasize that inequalities (3.58) are just a discrete version of the following entropy inequalities:

$$\partial_t s_i + u\partial_x s_i = \mu_i \frac{\gamma_i - 1}{\rho^{\gamma_i}}(\partial_x u)^2 \geq 0, \quad i = 1, ..., N.$$

We mention that discrete version of the inequalities in the ρS_i are not known at the present moment. But let us stress that the satisfaction of the maximum principles (3.55) on each of the specific entropies s_i are known after Tadmor [135] (see also Khobalatte and Perthame [78]) to hold as a consequence of an infinite number of classical entropy inequalities.
Let us now turn defining the discrete velocities entering the identities (3.57). In that aim, it is convenient to adopt an Harten, Lax, Van Leer type of analysis (see [61]) when introducing the following half averagings at a given interface $x_{j-1/2}$ under the CFL (3.17) :

$$(s_i)^{n+1=}_{j+1/2,L} = \frac{2}{\Delta x}\int_{x_j}^{x_{j+1/2}} s_i(\overline{\mathbf{w}}(x,t^{n+1=})) \times \frac{\rho^{\gamma_i}}{\overline{\rho^{\gamma_i}}^{n+1=}_{j+1/2,L}}(\overline{\mathbf{w}}(x,t^{n+1=}))dx \quad j \in Z, \tag{3.59}$$

where we have set :

$$\overline{\rho^{\gamma_i}}^{n+1=}_{j+1/2,L} = \frac{2}{\Delta x}\int_{x_j}^{x_{j+1/2}} \{\rho^{\gamma_i}\}(\overline{\mathbf{w}}(x,t^{n+1=}))dx.$$

Symmetrically we define :

$$(s_i)^{n+1=}_{j+1/2,R} = \frac{2}{\Delta x}\int_{x_{j+1/2}}^{x_{j+1}} s_i(\overline{\mathbf{w}}(x,t^{n+1=})) \times \frac{\rho^{\gamma_i}}{\overline{\rho^{\gamma_i}}^{n+1=}_{j+1/2,R}}(\overline{\mathbf{w}}(x,t^{n+1=}))dx \quad j \in Z, \tag{3.60}$$

where :

$$\overline{\rho^{\gamma_i}}^{n+1=}_{j+1/2,R} = \frac{2}{\Delta x}\int_{x_{j+1/2}}^{x_{j+1}} \{\rho^{\gamma_i}\}(\overline{\mathbf{w}}(x,t^{n+1=}))dx.$$

Notice that the averaging procedures (3.37) for defining each of the $(s_i)^{n+1=}_j$ equivalently rewrite in terms of the above half averages as follows :

$$(s_i)^{n+1=}_j = \frac{\overline{\rho^{\gamma_i}}^{n+1=}_{j-1/2,R} \times (s_i)^{n+1=}_{j-1/2,R} + \overline{\rho^{\gamma_i}}^{n+1=}_{j+1/2,L} \times (s_i)^{n+1=}_{j+1/2,L}}{\overline{\rho^{\gamma_i}}^{n+1=}_{j-1/2,R} + \overline{\rho^{\gamma_i}}^{n+1=}_{j+1/2,L}}, \tag{3.61}$$

since by construction

$$\overline{\rho^{\gamma_i}}^{n+1=}_j = \frac{1}{2}(\overline{\rho^{\gamma_i}}^{n+1=}_{j-1/2,R} + \overline{\rho^{\gamma_i}}^{n+1=}_{j+1/2,L}). \tag{3.62}$$

Equipped with these notations, let us state :
Proposition 15 *Under the CFL condition (3.17), the half averages (3.59) and (3.60) respectively read :*

$$\overline{\rho^{\gamma_i}}^{n+1=}_{j+1/2,L} \times (s_i)^{n+1=}_{j+1/2,L} = \overline{\rho^{\gamma_i}}^{n+1=}_{j+1/2,L} \times (s_i)^n_j - 2\frac{\Delta t}{\Delta x} w_i^{j+1/2,-}((s_i)^n_{j+1} - (s_i)^n_j),$$

$$\overline{\rho^{\gamma_i}}^{n+1=}_{j+1/2,R} \times (s_i)^{n+1=}_{j+1/2,R} = \overline{\rho^{\gamma_i}}^{n+1=}_{j+1/2,R} \times (s_i)^n_{j+1} - 2\frac{\Delta t}{\Delta x} w_i^{j+1/2,+}((s_i)^n_{j+1} - (s_i)^n_j), \tag{3.63}$$

where the discrete quantities $w_i^{j+1/2}$ are given by

$$w_i^{j+1/2,\pm} = w_i^{\pm}(\mathbf{u}^n_j, \mathbf{u}^n_{j+1}) = \int_0^{(u^\star_{j+1/2})^\pm} \rho^{\gamma_i}(\xi)d\xi, \tag{3.64}$$

with

$$(u^\star_{j+1/2})^+ = \max(u^\star_{j+1/2}, 0), \quad (u^\star_{j+1/2})^- = \min(u^\star_{j+1/2}, 0).$$

Here $\xi \to \rho(\xi)$ denotes the density function in the self-similar exact Riemann solution at the interface $x_{j+1/2}$ while $u^\star_{j+1/2}$ is the speed of propagation of the contact discontinuity. By contruction, the mappings $(\mathbf{u}_L, \mathbf{u}_R) \in \Omega^2 \to w_i^\pm(\mathbf{u}_L, \mathbf{u}_R) \in I\!R$ are locally Lipschitz-continuous with the property that

$$w_i^\pm(\mathbf{u}, \mathbf{u}) = \rho^{\gamma_i} u^\pm \quad \textit{for all} \quad \mathbf{u} \in \Omega.$$

In the definition (3.64), it is tacitly assumed that the Riemann solution under consideration is made of three distinct waves so that the velocity $u^\star$ of the intermediate contact wave is clearly well-defined. In order to extend its definition to the case of a Riemann solution made of a single wave (either a 1-wave or a 3-wave with classical definitions), we propose the following natural convention. When $\mathbf{u}_j$ and $\mathbf{u}_{j+1}$ are connected by a 1-wave, we set $u^\star_{j+1/2} = u^n_{j+1}$ while when connected by a 3-wave, we choose $u^\star_{j+1/2} = u^n_j$. This convention ensures the reported Lipschitz-continuity.

Proof Under the restriction CFL (3.17), the solution $\overline{\mathbf{w}}(x,t)$ coincides for all $(x,t) \in (x_j, x_{j+1}) \times (t^n, t^{n+1=})$ with the exact Riemann solution $\overline{\mathbf{w}}((x - x_{j+1/2})/(t - t^n), \mathbf{u}^n_j, \mathbf{u}^n_{j+1})$. Up to some shift in space and time, each of the specific entropies is easily seen to be given as the following step function propagating with speed $u^\star = u^\star_{j+1/2}$, namely the speed of the contact discontinuity :

$$s_i(x,t) = (s_i)^n_j + ((s_i)^n_{j+1} - (s_i)^n_j)\mathbf{H}(x - u^\star t), \tag{3.65}$$

where $\mathbf{H}$ denotes the usual Heavyside function. Let us restrict ourselves to a positive value $u^\star$. The case of a non positive value will follow from straightforward adaptations. Under this assumption, we clearly get when evaluating the formula definition (3.59) :

$$(s_i)^{n+1=}_{j+1/2,L} = (s_i)^n_j$$

while

$$\overline{\rho^{\gamma_i}}^{n+1=}_{j+1/2,R} \times (s_i)^{n+1=}_{j+1/2,R} =$$
$$\frac{2}{\Delta x}\int_0^{\Delta x/2} \rho^{\gamma_i}\Big(\frac{x}{\Delta t}\Big) \times \{(s_i)^n_j + ((s_i)^n_{j+1} - (s_i)^n_j)\mathbf{H}(x - u^\star \Delta t)\}dx =$$
$$\frac{2\Delta t}{\Delta x}\{(s_i)^n_j \int_0^{u^\star} \rho^{\gamma_i}(\xi)d\xi + (s_i)^n_{j+1}\int_{u^\star}^{\frac{\Delta x}{2\Delta t}} \rho^{\gamma_i}(\xi)d\xi\}.$$

We thus easily deduce that

$$\overline{\rho^{\gamma_i}}^{n+1=}_{j+1/2,R} \times (s_i)^{n+1=}_{j+1/2,R} =$$
$$(s_i)^n_{j+1} \times \frac{2\Delta t}{\Delta x}\int_0^{\frac{\Delta x}{2\Delta t}} \rho^{\gamma_i}(\xi)d\xi - \frac{2\Delta t}{\Delta x}\int_0^{u^\star} \rho^{\gamma_i}(\xi)d\xi \times ((s_i)^n_{j+1} - (s_i)^n_j),$$

where

$$\overline{\rho^{\gamma_i}}^{n+1=}_{j+1/2,R} = \frac{2\Delta t}{\Delta x}\int_0^{\frac{\Delta x}{2\Delta t}} \rho^{\gamma_i}(\xi)d\xi.$$

The conclusion easily follows.

We are now in position to establish :

Proposition 16 *Under the CFL condition (3.17), the updating formulae (3.37) read:*

$$(s_i)_j^{n+1=} = (s_i)_j^n - \frac{\Delta t}{\Delta x} v_i^{j-1/2,+}((s_i)_j^n - (s_i)_{j-1}^n) - \frac{\Delta t}{\Delta x} v_i^{j+1/2,-}((s_i)_{j+1}^n - (s_i)_j^n), \quad (3.66)$$

where we have set:

$$v_i^{j+1/2,\pm} = v_i^{\pm}(\mathbf{u}_j^n, \mathbf{u}_{j+1}^n) = \int_0^{(u_{j+1/2}^\star)^\pm} \rho^{\gamma_i}(\xi) d\xi / \overline{\rho^{\gamma_i}}_j^{n+1=}. \quad (3.67)$$

By contruction, the mappings $(\mathbf{u}_L, \mathbf{u}_R) \in \Omega^2 \to v_i^{\pm}(\mathbf{u}_L, \mathbf{u}_R) \in I\!R$ are locally Lipschitz-continuous with the property that

$$v_i^{\pm}(\mathbf{u}, \mathbf{u}) = u^{\pm} \quad \textit{for all} \quad \mathbf{u} \in \Omega.$$

Proof Invoking (3.61) and (3.62), equalities (3.63) readily imply

$$2\overline{\rho^{\gamma_i}}_j^{n+1=} \times (s_i)_j^{n+1=} = \overline{\rho^{\gamma_i}}_{j-1/2,R}^{n+1=} \times (s_i)_j^n - 2\frac{\Delta t}{\Delta x} w_i^{j-1/2,+}((s_i)_j^n - (s_i)_{j-1}^n)$$

$$+\overline{\rho^{\gamma_i}}_{j+1/2,L}^{n+1=} \times (s_i)_j^n - 2\frac{\Delta t}{\Delta x} w_i^{j+1/2,-}((s_i)_{j+1}^n - (s_i)_j^n),$$

that is to say, again thanks to (3.62)

$$(s_i)_j^{n+1=} = (s_i)_j^n - \frac{\Delta t}{\Delta x} \frac{w_i^{j-1/2,+}}{\overline{\rho^{\gamma_i}}_j^{n+1=}}((s_i)_j^n - (s_i)_{j-1}^n) - \frac{\Delta t}{\Delta x} \frac{w_i^{j+1/2,-}}{\overline{\rho^{\gamma_i}}_j^{n+1=}}((s_i)_{j+1}^n - (s_i)_j^n).$$

This concludes the proof.

The purpose of the following sections is to describe a relevant approximate Riemann solver which will yield a fairly simple evaluation of the discrete velocities defined in (3.67). As far as the exact Riemann solver is concerned, we may end up with somehow cumbersome formulae essentially because of the possible rarefaction waves. Our main objective is thus to suitably get rid of such waves when introducing approximate Riemann solutions systematically made of discontinuities separating constant states.

3.4 A Relaxation model

In order to meet the above requirement, we suggest to approximate the solutions of the system (3.14)-(3.15) by those of a suitable extended first order system with singular perturbation. Motivated by the work of Liu [101] and Chen, Levermore and Liu [28], such a system aims at restoring not only the original PDE model but also its entropy inequality in the regime of an infinite relaxation parameter. Here the extended first order system will have the property of being non linear hyperbolic with only linearly degenerate fields. Linear degeneracy is responsible for Riemann solutions uniquely made of discontinuities. Here and by contrast to the approach proposed by Jin and Xin [74], most of the non linearities in the original PDE model are kept for the sake of a better accuracy. Indeed, the relaxation system admits as natural solutions the contact discontinuities of the original PDE. Motivated by the work of Suliciu [131] (see also [32]), we suggest to suitably modify the total pressure law entering the original model since this law turns to concentrate all the genuine non linearities. The proposed approach has deep relationships with relaxation techniques already introduced in the literature, but is actually different in its design principle. Let us mention [32] in the 3×3 Euler setting and also [18] for a general approach devoted to flux vector splitting methods. Here we no longer consider the pressure as an independent variable equipped with its own PDE, but we keep this pressure as a non linear function of the unknown when suitably shadowing the exact role played by the specific volume τ. The strategy we propose actually considerably simplify the analysis of the relaxation model. The associated finite volumes method will actually turn to be algebraically equivalent to the ones described in [32], [18]. In the Lagrangian setting, the corresponding finite volumes method also coincides with the scheme proposed by Desprès [43] via another formalism (see also Gallice [51]). All these contributions can be understood as a fruitful reinterpretation of a scheme proposed by Toro [136] but which can be actually traced back to Viviand and Veuillot [144] in the mid seventies.

3.4.1 The model and its associated relaxation energy

We propose to approximate the entropy weak solutions of (3.14)-(3.15) by those of the following non linear first order system with singular perturbation :

$$\begin{cases} \partial_t \rho_\lambda + \partial_x (\rho u)_\lambda = 0, \\ \partial_t (\rho u)_\lambda + \partial_x (\rho u^2 + \Pi)_\lambda = 0, \\ \partial_t (\rho S_i)_\lambda + \partial_x (\rho S_i u)_\lambda = 0, \quad i = 1, ..., N, \\ \partial_t (\rho \mathcal{T})_\lambda + \partial_x (\rho \mathcal{T} u)_\lambda = \lambda \rho_\lambda (\tau_\lambda - \mathcal{T}_\lambda), \end{cases} \tag{3.68}$$

with the following closure equation :

$$\begin{gathered} \Pi = \Pi(\mathcal{T}, \{S_i\}_{1 \le i \le N}) = p(\mathcal{T}, \{S_i\}_{1 \le i \le N}) + a^2(\mathcal{T} - \tau), \quad \tau = \frac{1}{\rho}, \\ p(\mathcal{T}, \{S_i\}_{1 \le i \le N}) := \sum_{i=1}^{N} p_i(\mathcal{T}, S_i). \end{gathered} \tag{3.69}$$

From this definition, the equilibrium system (3.14) is clearly formally recovered in the regime of an infinite relaxation parameter $\lambda > 0$ since we get $\mathcal{T} = \tau$. In this loose sense, solutions of the relaxation system (3.68) may serve to approximate the solutions of the equilibrium system (3.14). But to prevent the proposed relaxation procedure from instabilities in the asymptotic regime $\lambda = +\infty$, it is well-known after the work by Liu [101] and Chen, Levermore and Liu [28], that the equilibrium system and the relaxation system must meet some compatibility conditions : eigenvalues of both systems must be properly interlaced. These so-called Whitham conditions will be seen hereafter to be satisfied for properly prescribed values of the positive real number a entering the definition of the relaxation pressure law (3.69).
Setting $\mathbf{U} =^T (\rho, \rho u, \{\rho S_i\}_{i=1,...,N}, \rho\mathcal{T})$, the relaxation system (3.68)–(3.69) is associated with the following natural phase space :

$$\Omega_{\mathbf{U}} = \{\mathbf{U} \in \mathbb{R}^{N+3}, \rho > 0, \ \rho S_i > 0, \ i = 1, ..., N, \ \rho\mathcal{T} > 0\}.$$

For simplicity in the notations, (3.68) will be given the next condensed form when $\lambda = 0$:

$$\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = 0, \tag{3.70}$$

where $\mathbf{F}$ finds an immediate definition. The first statement motivates the interest of the proposed relaxation procedure :

Proposition 17 *Let be given $a > 0$. Then the first order system (3.70) is hyperbolic for all $\mathbf{U}$ in $\Omega_{\mathbf{U}}$ with the following distinct eigenvalues :*

$$\lambda_1(\mathbf{U}) = u - a\tau = u - \frac{a}{\rho}, \quad \lambda_2(\mathbf{U}) = u, \quad \lambda_3(\mathbf{U}) = u + a\tau = u + \frac{a}{\rho}.$$

The eigenvalue $\lambda_2(\mathbf{U})$ has $(N+1)$ order of multiplicity while $\lambda_1(\mathbf{U})$ and $\lambda_3(\mathbf{U})$ are simple. All the fields under consideration are linearly degenerate.

As a consequence of the linear degeneracy of all the fields, the Riemann problem associated with (3.70) will be seen below to be trivially solvable. In addition, contact discontinuity solutions of the equilibrium system will be seen to be exactly preserved by the proposed relaxation procedure. These two highly desirable properties have actually dictated the derivation model (3.68)-(3.69).

Proof Easy calculations left to the reader give the required results. Here we simply highlight that the property of linear degeneracy of the field associated with the eigenvalues $\lambda_i(\mathbf{U})$ for $i = 1$ and $i = 3$ can only hold for sound speed $c = c(\rho, \{S_i\}_{i=1,\dots,N})$ in the form $c = a/\rho$ where a is a positive real constant ($a = 0$ precludes the existence of a basis of eigenvectors). In that aim, observe that when considering general closure thermodynamics relations (*i.e.* for general sound speed given by $c^2 = -\partial_\tau p$), an eigenvector associated with the eigenvalue $\lambda_\epsilon(\mathbf{U}) = u + \epsilon c$ (with $\epsilon = -1$ or $\epsilon = 1$) is given by $r_\epsilon(u) = (1, \epsilon c/\rho, \{0\}_{i=1,\dots,N})$. Then, asking $\nabla_{\mathbf{U}}\lambda_\epsilon(\mathbf{U}).r_\epsilon(\mathbf{U}) = \epsilon(\partial_\rho c + c/\rho) = 0$ to be valid for all $\mathbf{U}$ is nothing but the expected property.

The next statement proves that the relaxation system (3.68)–(3.69) enters the general framework proposed by Chen, Levermore and Liu [28] :

Theorem 7
Let us introduce the relaxation energy defined for all $\mathbf{U}$ in $\Omega_{\mathbf{U}}$ by :

$$\{\rho\Sigma\}(\mathbf{U}) = \frac{1}{2}\frac{(\rho u)^2}{\rho} + \rho\phi(\mathcal{T}, \{S_i\}_{1\le i\le N}), \tag{3.71}$$

where

$$\phi(\mathcal{T}, \{S_i\}_{1\le i\le N}) = \sum_{i=1}^{N}\varepsilon_i(\mathcal{T}, S_i) + \frac{1}{2a^2}\{\Pi^2(\mathcal{T}, \{S_i\}_{1\le i\le N}) - (\sum_{i=1}^{N} p_i(\mathcal{T}, S_i))^2\}. \tag{3.72}$$

This relaxation energy is consistent with the equilibrium energy in the sense that :

$$\rho\Sigma(\rho, \rho u, \{\rho S_i\}_{1\le i\le N}, \rho\tau) = \rho E, \quad \{(\rho\Sigma + \Pi)u\}(\rho, \rho u, \{\rho S_i\}_{1\le i\le N}, \rho\tau) = (\rho E + p)u.$$

In addition, smooth solutions of (3.68)–(3.69) satisfy the following energy equation for all $\lambda \ge 0$:

$$\partial_t\{\rho\Sigma\}(\mathbf{U}_\lambda) + \partial_x\{(\rho\Sigma + \Pi)u\}(\mathbf{U}_\lambda) = -\lambda\rho_\lambda(a^2 - \{\rho c\}^2(\mathbf{U}_\lambda))(\tau_\lambda - \mathcal{T}_\lambda)^2, \tag{3.73}$$

which yields an energy inequality under the following Whitham condition :

$$a^2 > \{\rho c\}^2(\mathcal{T}, \{S_i\}_{1\le i\le N}) = -\sum_{i=1}^{N}\{\partial_\tau p_i\}(\mathcal{T}, S_i), \tag{3.74}$$

for all the $\mathcal{T}$ and $\{S_i\}_{1\le i\le N}$ under consideration. Moreover, $(\rho,\ \rho u,\ \{\rho S_i\}_{1\le i\le N})$ being fixed, for all T-compact $\mathcal{K}$ including $\mathcal{T} = \tau$ so as to consider a finite value $a^2(\mathcal{K})$ in the Whitham condition (3.74), then the following minimization principle holds :

$$\rho E = \frac{1}{2}\frac{(\rho u)^2}{\rho} + \sum_{i=1}^{N}\rho\varepsilon_i(\tau, S_i) = \min_{\mathcal{T}\in\mathcal{K}}\{\{\rho\Sigma\}(\mathbf{U})\}. \tag{3.75}$$

In other words and under the Whitham condition (3.74), the relaxation energy spontaneously decreases in L^1 norm in the relaxation procedure. Moreover, its minimum coincides with the equilibrium energy ρE. With this respect, the relaxation energy plays the role of an entropy compatible with the relaxation procedure in the sense of Chen, Levermore and Liu [28]. According to the general framework proposed by these authors, one should have required the relaxation energy to be in addition strictly convex in $\mathbf{U}$. But here, such a convexity property is not useful since we do not need for a Lax entropy selection principle for the weak solutions of the PDE system (3.68)–(3.69). Indeed, all its fields are linearly degenerate and such a strong property is known (see [57], [118]) to imply that the equation (3.73), valid for smooth solutions, is still valid for weak solutions with equality in the sense of the distributions. The validity of (3.73) in $\mathcal{D}'$ together with the reported minimization principle will suffice for our forthcoming purposes.

Lemma 12 *Smooth solutions of (3.68)-(3.69) obey the following pressure-like equation for all $\lambda \ge 0$:*

$$\partial_t\Pi(\mathbf{U}_\lambda) + u_\lambda\partial_x\Pi(\mathbf{U}_\lambda) + a^2\tau_\lambda\partial_x u_\lambda = \lambda(a^2 - \{\rho c\}^2(\mathbf{U}_\lambda)) \times (\tau_\lambda - \mathcal{T}_\lambda), \tag{3.76}$$

where $\{\rho c\}^2(\mathbf{U})$ is defined with little abuse in the notation in (3.74).

Proof The validity of (3.76) stems from the property that smooth solutions of (3.68) satisfy the following transport equations :

$$\begin{cases} \partial_t S_i + u\partial_x S_i = 0, \quad i = 1, ..., N, \\ \partial_t\mathcal{T} + u\partial_x\mathcal{T} = \lambda(\tau - \mathcal{T}), \end{cases} \tag{3.77}$$

where the subscript λ has been (and will be from now on) omitted for simplicity in the notations. We thus easily infer the following identity :

$$\partial_t p(\mathcal{T}, \{S_i\}_{1\le i\le N}) + u\partial_t p(\mathcal{T}, \{S_i\}_{1\le i\le N}) = \lambda(\{\sum_{i=1}^{N}\partial_\tau p_i\}(\mathcal{T}, S_i))(\tau - \mathcal{T}). \tag{3.78}$$

Next and by the density equation, the specific volume τ is known to be solution of :

$$\partial_t\tau + u\partial_x\tau - \tau\partial_x u = 0,$$

so that we deduce :

$$\partial_t a^2(\mathcal{T} - \tau) + u\partial_x a^2(\mathcal{T} - \tau) + a^2\tau\partial_x u = \lambda a^2(\tau - \mathcal{T}). \tag{3.79}$$

The required pressure like equation (3.76) just follows from the definition (3.69) of $\Pi(\mathcal{T}, \{S_i\}_{1\leq i\leq N})$ when summing (3.78) and (3.79).

With this result, let us prove Theorem 7.

Proof Classical considerations imply from the density and momentum equations that smooth solutions obey the following equation for governing the kinetic energy :

$$\partial_t \rho \frac{u^2}{2} + \partial_x \rho \frac{u^3}{2} + u\partial_x \Pi = 0. \tag{3.80}$$

In order to derive an equation in conservation form, let us notice that the additional pressure-like equation (3.76) yields after easy manipulations :

$$\partial_t \rho \frac{\Pi^2}{2a^2} + \partial_x \rho u \frac{\Pi^2}{2a^2} + \Pi \partial_x u = \lambda \rho \Pi (1 - \frac{\{\rho c\}^2}{a^2})(\tau - \mathcal{T}), \tag{3.81}$$

so as to arrive at :

$$\partial_t \rho \{\frac{u^2}{2} + \frac{\Pi^2}{2a^2}\} + \partial_x (\rho u \{\frac{u^2}{2} + \frac{\Pi^2}{2a^2}\} + \Pi u) = \lambda \rho \Pi (1 - \frac{\{\rho c\}^2}{a^2})(\tau - \mathcal{T}). \tag{3.82}$$

The energy involved in (3.82) is obviously not consistent with the equilibrium total energy so that (3.82) must receive extra contributions as follows. Invoking again the transport equations (3.77), any given smooth nonlinear combination of $\mathcal{T}$ and the $\{S_i\}_{i=1,\ldots,N}$, say

$$\theta(\mathcal{T}, \{S_i\}_{i=1,\ldots,N}) = \sum_{i=1}^{N} \varepsilon_i(\mathcal{T}, S_i) - \frac{(\sum_{i=1}^{N} p_i(\mathcal{T}, S_i))^2}{2a^2},$$

is easily seen to satisfy :

$$\partial_t \rho \theta(\{S_i\}_{i=1,\ldots,N}, I) + \partial_x \rho \theta(\{S_i\}_{i=1,\ldots,N}, I) u = -\lambda \rho \sum_{i=1}^{N} p_i(\mathcal{T}, S_i)(1 - \frac{\{\rho c\}^2}{a^2})(\tau - \mathcal{T}). \tag{3.83}$$

Therefore, summing (3.82) and (3.83) gives in view of the definition of the relaxation energy :

$$\partial_t \rho \Sigma + \partial_x (\rho \Sigma u + \Pi u) = \lambda \rho (1 - \frac{\{\rho c\}^2}{a^2}) \times (\Pi - \sum_{i=1}^{N} p_i(\mathcal{T}, S_i))(\tau - \mathcal{T}),$$

which is nothing but the required inequality (3.73) when invoking the definition (3.69) of the relaxation pressure law.

Let us now turn establishing the minimization principle (3.75). Let $(\rho, \rho u, \{\rho S_i\}_{1\leq i\leq N})$ being fixed and let us consider a fixed compact $\mathcal{K} = [\mathcal{T}_{min}, \mathcal{T}_{max}]$ with the property that $\mathcal{T} = \tau \in \mathcal{K}$. Then, easy calculations yield for all $\mathcal{T} \in \mathcal{K}$:

$$\{\partial_{\mathcal{T}} \rho \Sigma\}(\mathbf{U}) = \rho \partial_{\mathcal{T}} \phi(\mathcal{T}, \{S_i\}_{1\leq i\leq N}) = (a^2(\mathcal{K}) - \{\rho c\}^2(\mathcal{T}, \{S_i\}_{1\leq i\leq N})) \times (\mathcal{T} - \tau), \tag{3.84}$$

where $a^2(\mathcal{K})$ denotes some finite values satisfying the Whitham condition on $\mathcal{K}$. As a consequence, this first order derivative exactly vanishes once on $\mathcal{K}$: namely for $\mathcal{T} = \tau$. But the second $\mathcal{T}$-derivative evaluated at this solution gives :

$$\{\partial^2_{\mathcal{T}\mathcal{T}}\rho\Sigma\}(\rho, \rho u, \{\rho S_i\}_{1\le i\le N}, \rho\tau) = (a^2(\mathcal{K}) - \{\rho c\}^2(\tau, \{S_i\}_{1\le i\le N})) > 0, \qquad (3.85)$$

again because of the Whitham condition. This suffices to prove that $\mathcal{T} = \tau$ is a local minimum on $\mathcal{K}$. Since the first derivative vanishes only once and this with a positive second derivative, the endpoints $\mathcal{T}_{min}$ and $\mathcal{T}_{max}$ of $\mathcal{K}$ cannot yield lower values of $\rho\Sigma$. Therefore the local minimum $\mathcal{T} = \tau$ is necessarily a global one. This completes the proof since by construction one has :

$$\{\rho\Sigma\}(\rho, \rho u, \{\rho S_i\}_{1\le i\le N}, \rho\tau) = \rho E.$$

Equipped with these stability results valid under the Whitham condition (3.74), we now turn describing a relevant finite volumes method for approximating the entropy weak solutions of the non linear hyperbolic system (3.14)-(3.15).

3.4.2 The numerical method

Assuming a piecewise constant approximate solution $\mathbf{u}_h$ to be given at time t^n, this one is evolved to the next time level $t^n + \Delta t$ when approximating the solution of the Cauchy problem (3.14)-(3.15) with initial data $\mathbf{u}_h(x, t^n)$ by the solution of the relaxation system (3.68) when prescribing its initial data $\mathbf{U}_h(x, t^n)$ at equilibrium from $\mathbf{u}_h(x, t^n)$ as follows :

$$\mathbf{U}_h(x, t^n) =^T (\rho_h(x, t^n), (\rho u)_h(x, t^n), \{(\rho S_i)_h(x, t^n)\}_{i=1,\ldots N}, (\rho\mathcal{T})_h(x, t^n)), \qquad (3.86)$$

where by definition :

$$(\rho\mathcal{T})_h(x, t^n) = 1, \quad \text{for all} \quad x \in I\!R.$$

The precise choice of the constant a entering the definition of the relaxation pressure Π in (3.69) will be described later on. Its definition will be of course made to meet the Whitham condition (3.74) but some useful refinements will be performed under the convenient CFL restriction :

$$\frac{\Delta t}{\Delta x} \max_{\mathbf{U}}(|u - a\tau|(\mathbf{U}), |u + a\tau|(\mathbf{U})) \le \frac{1}{2}, \qquad (3.87)$$

for all the $\mathbf{U}$ under consideration. Following Jin and Xin [74], we adopt the now classical two steps procedure when first solving the frozen relaxation problem : *i.e.* when choosing $\lambda = 0$, and when then projecting back its frozen solution onto the equilibrium manifold. Strictly speaking, this procedure consists in solving the following Cauchy problem for small times t (*i.e.* under the CFL condition (3.87)) :

$$\begin{cases} \partial_t\rho + \partial_x(\rho u) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u^2 + \Pi) = 0, \\ \partial_t(\rho S_i) + \partial_x(\rho S_i u) = 0, \quad i = 1, ..., N, \\ \partial_t(\rho\mathcal{T}) + \partial_x(\rho\mathcal{T} u) = 0, \end{cases} \qquad (3.88)$$

with initial data $\mathbf{U}_h(x, t^n)$. Note again that no entropy selection principle is needed to select the weak solution $\mathbf{U}_h(x,t)$ under consideration since all the fields of (3.88) are linearly degenerate. Since $\lambda = 0$, this solution generally no longer belongs to the equilibrium manifold, and this motivates the following pointwise in x projection step at time $t^n + \Delta t$:

$$\mathcal{T}_h(x, t^n + \Delta t) = \tau_h(x, t^n + \Delta t), \quad \text{for all} \quad x \in I\!R. \tag{3.89}$$

Such an identity is formally equivalently obtained when solving the following ODE system with singular perturbation in the regime of an infinite relaxation parameter λ:

$$\begin{cases} \partial_t \rho_\lambda = 0, \\ \partial_t (\rho u)_\lambda = 0, \\ \partial_t (\rho S_i)_\lambda = 0, \quad i = 1, ..., N, \\ \partial_t (\rho \mathcal{T})_\lambda = \lambda \rho_\lambda (\tau_\lambda - \mathcal{T}_\lambda). \end{cases} \tag{3.90}$$

This two steps procedure (actually a convenient variant of it) will be seen to yield a entropy consistent finite volumes method for approximating the weak solutions of (3.14)-(3.15) under a relevant choice of the parameter a.

3.4.3 Classical averagings in the relaxation approach

Let us now propose a precise description of the variant we suggest, motivated by an adaptative in space choice of a. A localized definition of the parameter a requires a brief discussion to prove that a relevant finite volumes method can be already derived in the setting of the classical averagings: note that both the density and the momentum have received in the previous section this type of averages. We shall conclude this section when exhibiting the exact Riemann solution for the Relaxation system. Its form will be used in the next section. To motivate an adaptive choice of a, let us observe that the relaxation energy inequality (3.73) with $\lambda > 0$ strongly suggests to choose the smallest value of a that satisfies the Whitham condition (3.74), in order to lower the numerical rate of dissipation measured here in terms of the discrete energy rate of dissipation. Here and after [39], we show that it is actually possible to optimize locally in space the choice of a under the CFL restriction (3.87). To that purpose, it is convenient to adopt a Harten, Lax and Van Leer type of analysis. Here, we suggest to define $\mathbf{U}_h(x,t)$ with $t \in (t^n, t^n + \Delta t)$ locally in space when setting:

$$\mathbf{U}_h(x,t) = \mathbf{U}_{a_{j+1/2}}(\frac{x - x_{j+1/2}}{t - t^n}; \mathbf{U}_j^n, \mathbf{U}_{j+1}^n), \quad (x,t) \in (x_j, x_{j+1}) \times (t^n, t^n + \Delta t), \tag{3.91}$$

where the self-similar function $\mathbf{U}_{a_{j+1/2}}$ denotes the exact Riemann solution of (3.88) at interface $x_{j+1/2}$. Here the relaxation pressure law is locally defined by

$$\Pi = \Pi(\mathcal{T}, \{S_i\}_{1 \le i \le N}) = p(\mathcal{T}, \{S_i\}_{1 \le i \le N}) + a_{j+1/2}^2 (\mathcal{T} - \tau),$$

for some relevant definition of $a_{j+1/2}$ given later on, but solely depending on the two states $\mathbf{U}_j^n$ and $\mathbf{U}_{j+1}^n$. This localized definition is obviously performed for the sake of minimizing the numerical dissipation. But consequently, the function $\mathbf{U}_h(x,t)$ we have defined along these lines is no longer the exact solution of the Cauchy problem since now a discontinuity takes place in the definition of the pressure law at each x_j.

Then the exact solution cannot be the superposition of the Riemann solutions $\mathbf{U}_{a_{j+1/2}}$. Nevertheless, the property that $\mathbf{U}_h(x,t^n)$ is at equilibrium (*i.e.* with $\mathcal{T}_j^n = \tau_j^n$ for all j) will allow us to check that, under the CFL condition (3.87), the proposed function $\mathbf{U}_h(x,t)$ still serves to define a conservative finite volumes method for approximating the entropy weak solutions of the equilibrium system (3.14)-(3.15). To prove this claim, let us consider the following half averagings of the function $\mathbf{U}_h(x,t)$:

$$\begin{aligned} \mathbf{U}^{n+1-}_{j+1/2,L} &= \tfrac{2}{\Delta x}\int_{x_j}^{x_{j+1/2}} \mathbf{U}_h(x,t^n+\Delta t)dx, \\ \mathbf{U}^{n+1-}_{j+1/2,R} &= \tfrac{2}{\Delta x}\int_{x_{j+1/2}}^{x_{j+1}} \mathbf{U}_h(x,t^n+\Delta t)dx. \end{aligned} \tag{3.92}$$

We then have the following result:

Lemma 13 *Under the CFL condition (3.87), the averagings (3.92) meet the following equivalent form*

$$\begin{aligned} \mathbf{U}^{n+1-}_{j+1/2,L} &= \mathbf{U}_j^n - 2\tfrac{\Delta t}{\Delta x}(\mathbf{F}_{a_{j+1/2}}(\mathbf{U}_{a_{j+1/2}}(0^+;\mathbf{U}_j^n,\mathbf{U}_{j+1}^n)) - \mathbf{F}_{a_{j+1/2}}(\mathbf{U}_j^n)), \\ \mathbf{U}^{n+1-}_{j+1/2,R} &= \mathbf{U}_{j+1}^n - 2\tfrac{\Delta t}{\Delta x}(\mathbf{F}_{a_{j+1/2}}(\mathbf{U}_{j+1}^n) - \mathbf{F}_{a_{j+1/2}}(\mathbf{U}_{a_{j+1/2}}(0^+;\mathbf{U}_j^n,\mathbf{U}_{j+1}^n))), \end{aligned} \tag{3.93}$$

where we have used the condensed notation $\mathbf{F}$ *introduced in (3.70). In addition, we have:*

$$\mathbf{F}_{a_{j+1/2}}(\mathbf{U}_j^n) = ((\rho u)_j^n, (\rho u^2+\Pi)_j^n, \{(\rho S_i u)_j^n\}_{i=1,\dots,N}, (\rho\mathcal{T}u)_j^n), \tag{3.94}$$

with

$$\Pi(\mathbf{U}_j^n) = \sum_{i=1}^N (p_i)_j^n, \quad \mathcal{T}(\mathbf{U}_j^n) = \tau_j^n.$$

The same property holds for $\mathbf{F}_{a_{j+1/2}}(\mathbf{U}_{j+1}^n)$.
In other words, each of the fluxes $\mathbf{F}_{a_{j+1/2}}(\mathbf{U}_j^n)$ stays free from a particular choice of $a_{j+1/2}$.

Proof The restriction of the function $\mathbf{U}_h(x,t)$ to $(x_j,x_{j+1})\times(t^n,t^n+\Delta t)$ coincides with the exact Riemann solution of the Cauchy problem (3.88) at interface $x_{j+1/2}$. Thus by construction we have:

$$\partial_t \mathbf{U}_{a_{j+1/2}} + \partial_x \mathbf{F}_{a_{j+1/2}}(\mathbf{U}_{a_{j+1/2}}) = 0, \quad (x,t)\in(x_j,x_{j+1})\times(t^n,t^n+\Delta t).$$

Classical considerations (see for instance [57]) then yield, with little abuse in the notation:

$$\frac{2}{\Delta x}\int_{x_j}^{x_{j+1/2}} \mathbf{U}_{a_{j+1/2}}(x,t^n+\Delta t)dx - \mathbf{U}_j^n + \frac{2}{\Delta x}(\Delta t\,\mathbf{F}_{a_{j+1/2}}(\mathbf{U}_{a_{j+1/2}}(0^+;\mathbf{U}_j^n,\mathbf{U}_{j+1}^n)) -$$

$$\int_{t^n}^{t^n+\Delta t} \mathbf{F}_{a_{j+1/2}}(\mathbf{U}_{a_{j+1/2}}(\frac{x_j - x_{j+1/2}}{t-t^n};\mathbf{U}_j^n,\mathbf{U}_{j+1}^n))dt) = 0.$$

But in view of the CFL restriction (3.87) one clearly has:

$$\mathbf{U}_{a_{j+1/2}}(\frac{x_j - x_{j+1/2}}{t-t^n};\mathbf{U}_j^n,\mathbf{U}_{j+1}^n) = \mathbf{U}_j^n,$$

for all $t \in (t^n, t^n + \Delta t)$. The required first identity in (3.93) follows and the second one is derived along the same lines.
To conclude, it suffices to observe that by construction the states $\mathbf{U}_j^n$ are at equilibrium and hence the identity $\mathbf{F}_{a_{j+1/2}}(\mathbf{U}_j^n) = \mathbf{F}(\mathbf{U}_j^n)$ whatever $a_{j+1/2}$ is.
Equipped with this result and following [61], we are led to propose the following updating:

$$\mathbf{U}_j^{n+1-} = \frac{1}{2}(\mathbf{U}_{j+1/2,L}^{n+1-} + \mathbf{U}_{j-1/2,R}^{n+1-}), \tag{3.95}$$

so as to get the updating of the equilibrium approximate solution:

$$\mathbf{u}_j^{n+1} = \begin{cases} \rho_j^{n+1} = \rho_j^{n+1-}, \\ (\rho u)_j^{n+1} = (\rho u)_j^{n+1-}, \\ (\rho S_i)_j^{n+1} = (\rho S_i)_j^{n+1-}, \quad i = 1, ..., N. \end{cases} \tag{3.96}$$

In agreement with the rough description of the algorithm, the method is advanced to the next time level when defining

$$\mathbf{U}_j^{n+1} = (\mathbf{u}_j^{n+1}, 1), \quad j \in Z.$$

Corollary 5 *Under the CFL restriction (3.87), the equilibrium approximate solution updates according to the following consistent finite volumes method:*

$$\mathbf{u}_j^{n+1} = \mathbf{u}_j^n - \frac{\Delta t}{\Delta x}\Delta \mathbf{f}_{j+1/2}^n, \tag{3.97}$$

where the numerical flux $\mathbf{f}_{j+1/2}^n$ is given from the first $(N+2)$ components of the next Lipschitz-continuous and consistent numerical flux function:

$$(\mathbf{f}_{j+1/2}^n, .) = (\mathbf{F}_{a_{j+1/2}}(\mathbf{U}_{a_{j+1/2}}(0^+; \mathbf{U}_j^n, \mathbf{U}_{j+1}^n)). \tag{3.98}$$

The relevance of the proposed finite volumes method will be assessed in the next paragraph when properly prescribing the value of the parameter a at each interface $x_{j+1/2}$ so as to meet the Whitham condition (3.74). Let us conclude the present one when deriving a fairly simple form of expression for the numerical flux function (3.98). The simplicity in the evaluation of this flux comes from the property that the Riemann problem for (3.88) is trivially solved according to:
Lemma 14 *Let be given $\mathbf{U}_L$ and $\mathbf{U}_R$ two states in $\Omega_{\mathbf{U}}$ at equilibrium (i.e $\mathcal{T}_L = \tau_L$ and $\mathcal{T}_R = \tau_R$). Let us then define $a := a(\mathbf{U}_L, \mathbf{U}_R) > 0$, the parameter entering the relaxation pressure law (3.69), large enough such that the following natural ordering holds:*

$$\lambda_1(\mathbf{U}_L) = u_L - a\tau_L < u^\star < \lambda_3(\mathbf{U}_R) = u_R + a\tau_R, \tag{3.99}$$

where $u^\star$ stands for some intermediate velocity given by:

$$u^\star = \frac{1}{2}(u_L + u_R) - \frac{1}{2a}(p_R - p_L).$$

Then, the self-similar solution $\mathbf{U}_a(x/t; \mathbf{U}_L, \mathbf{U}_R)$ of the Cauchy problem (3.88) for the initial data

$$\mathbf{U}_0(x) = \begin{cases} \mathbf{U}_L & \text{if} \quad x < 0, \\ \mathbf{U}_R & \text{if} \quad x > 0, \end{cases} \tag{3.100}$$

is systematically made of at most four constant states separated by three discontinuities as follows :

$$\mathbf{U}(x/t;\mathbf{U}_L,\mathbf{U}_R)=\begin{cases}\mathbf{U}_L & \text{if } \frac{x}{t}<\lambda_1(\mathbf{U}_L),\\ \mathbf{U}_1 & \text{if } \lambda_1(\mathbf{U}_L)<\frac{x}{t}<\lambda_2(\mathbf{U}_1),\\ \mathbf{U}_2 & \text{if } \lambda_2(\mathbf{U}_2)<\frac{x}{t}<\lambda_3(\mathbf{U}_R),\\ \mathbf{U}_R & \text{if } \lambda_3(\mathbf{U}_R)<\frac{x}{t},\end{cases}$$

with

$$\lambda_2(\mathbf{U}_1)=\lambda_2(\mathbf{U}_2)=u^\star.$$

The intermediate states in this Riemann solution are defined from :

$$\begin{array}{l} u_1=u_2=u^\star,\\ (S_i)_1=(S_i)_L, \quad 1\le i\le N,\\ (S_i)_2=(S_i)_R, \quad 1\le i\le N,\\ \mathcal{T}_1=\tau_L,\\ \mathcal{T}_2=\tau_R,\end{array}$$

together with

$$\rho_1=a/(u^\star-\lambda_1(\mathbf{U}_L)), \quad \rho_2=a/(\lambda_3(\mathbf{U}_R)-u^\star),$$

so that under the ordering condition (3.99), both ρ_1 and ρ_2 are positive. As a consequence of these formulae, the intermediate states $\mathbf{U}_1$ and $\mathbf{U}_2$ are in $\Omega_{\mathbf{U}}$, and are associated with the following total pressure :

$$\Pi(\mathbf{U}_1)=p_L+a^2(\tau_L-\tau_1)=p_R+a^2(\tau_R-\tau_2)=\Pi(\mathbf{U}_2).$$

Proof Since all the fields are linearly degenerate, the Riemann solution we seek for is uniquely made of discontinuities, the i^{th} one propagating at the characteristic speed λ_i of the field under consideration. These discontinuities separate at most four constant states $\mathbf{U}_L$, $\mathbf{U}_1$, $\mathbf{U}_2$ and $\mathbf{U}_R$, two neighbouring states being solutions of the Rankine-Hugoniot conditions associated with the system (3.88). Once the parameter a obeys the natural ordering condition (3.99) in the three wave velocities, these jump conditions are easily seen to yield the conclusion.

3.4.4 The total energy stability estimate

Let us now address the relevant definition of the parameter $a_{j+1/2}$ entering the numerical flux function (3.98) at each interface $x_{j+1/2}$. To this purpose, let us define from the Riemann solution $\mathbf{U}_{a_{j+1/2}}(.;\mathbf{U}_j^n,\mathbf{U}_{j+1}^n)$ the following self-similar function :

$$\mathbf{u}_{a_{j+1/2}}(x,t):={}^T(\rho,\rho u,\{\rho S_i\}_{i=1,\dots,N})_{a_{j+1/2}}\Big(\frac{x-x_{j+1/2}}{t-t^n}\Big),$$
$$(x,t)\in(x_j,x_{j+1})\times(t^n,t^n+\Delta t).$$

Observe that the following half averagings :

$$\begin{aligned}\mathbf{u}^{n+1}_{j+1/2,L} &= \tfrac{2}{\Delta x}\int_{x_j}^{x_{j+1/2}}\mathbf{u}_{a_{j+1/2}}(x,t^n+\Delta t)dx,\\ \mathbf{u}^{n+1}_{j+1/2,R} &= \tfrac{2}{\Delta x}\int_{x_{j+1/2}}^{x_{j+1}}\mathbf{u}_{a_{j+1/2}}(x,t^n+\Delta t)dx,\end{aligned} \tag{3.101}$$

just restore the proposed update for the equilibrium solution $\mathbf{u}_j^{n+1}$ from the next identity:

$$\mathbf{u}_j^{n+1} = \frac{1}{2}(\mathbf{u}_{j+1/2,L}^{n+1} + \mathbf{u}_{j-1/2,R}^{n+1}). \tag{3.102}$$

Equipped with this suitable function, we now state the main result of this section:

Theorem 8 *Let us assume that $a_{j+1/2} := a_{j+1/2}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n)$ is chosen large enough so as to satisfy the ordering condition (3.99) together with the following Whitham conditions:*

$$\begin{aligned}
&\sum_{i=1}^{N} \frac{\partial p_i}{\partial \tau}(\tau_j^n, (S_i)_j^n) + a_{j+1/2}^2 > 0, \quad \sum_{i=1}^{N} \frac{\partial p_i}{\partial \tau}((\tau_1)_{j+1/2}^n, (S_i)_j^n) + a_{j+1/2}^2 > 0,\\
&\sum_{i=1}^{N} \frac{\partial p_i}{\partial \tau}(\tau_{j+1}^n, (S_i)_{j+1}^n) + a_{j+1/2}^2 > 0, \quad \sum_{i=1}^{N} \frac{\partial p_i}{\partial \tau}((\tau_2)_{j+1/2}^n, (S_i)_{j+1}^n) + a_{j+1/2}^2 > 0,
\end{aligned} \tag{3.103}$$

where $(\tau_1)_{j+1/2}^n$ and $(\tau_2)_{j+1/2}^n$ denote the specific volumes in the intermediate states of the Riemann solution $\mathbf{U}_{a_{j+1/2}}(., \mathbf{U}_j^n, \mathbf{U}_{j+1}^n)$. Then, under the CFL restriction (3.87), the following energy-like inequalities hold:

$$\begin{aligned}
&\frac{2}{\Delta x}\int_{x_j}^{x_{j+1/2}} \{\rho E\}(\mathbf{u}_{a_{j+1/2}}(x, t^n + \Delta t))dx \leq (\rho E)_j^n - 2\frac{\Delta t}{\Delta x}(\{(\rho\Sigma + \Pi)u\}_{j+1/2}^n - (\rho H u)_j^n),\\
&\frac{2}{\Delta x}\int_{x_{j+1/2}}^{x_{j+1}} \{\rho E\}(\mathbf{u}_{a_{j+1/2}}(x, t^n + \Delta t))dx \leq (\rho E)_{j+1}^n - 2\frac{\Delta t}{\Delta x}((\rho H u)_{j+1}^n - \{(\rho\Sigma + \Pi)u\}_{j+1/2}^n),
\end{aligned} \tag{3.104}$$

where by definition,

$$\{(\rho\Sigma + \Pi)u\}_{j+1/2}^n = \{(\rho\Sigma + \Pi)u\}(\mathbf{U}_{a_{j+1/2}}(0^+; \mathbf{U}_j^n, \mathbf{U}_{j+1}^n)).$$

Observe that for sufficiently large values of a, $(\tau_1)_{j+1/2}^n$ (respectively $(\tau_2)_{j+1/2}^n$) becomes arbitrarily close to τ_j^n (respectively τ_{j+1}^n). As an immediate consequence, it is easy to check that if a is chosen large enough, the Whitham conditions (3.99) (as well as the ordering property (3.103)) necessarily hold. We refer the reader to Bouchut [18] for simplified Whitham conditions which are more convenient to deal with from a numerical point of view.

Proof We prove below the first inequality in (3.104). The companion one will follow from simple adaptations. Let us first recall that the Riemann solution $\mathbf{U}_{a_{j+1/2}}(.; \mathbf{U}_j^n, \mathbf{U}_{j+1}^n)$ is obtained when choosing $\lambda = 0$ in the relaxation procedure. So that it obeys in the sense of the distributions the following total energy additional conservation law in view of Theorem 7:

$$\partial_t\{\rho\Sigma\}(\mathbf{U}_{a_{j+1/2}}) + \partial_x\{(\rho\Sigma + \Pi)u\}(\mathbf{U}_{a_{j+1/2}}) = 0.$$

Let us indeed recall that the property of linear degeneracy of all the fields implies the validity of the above conservation law not only for smooth but also weak solutions of

(3.88). Now classical arguments give:

$$\frac{2}{\Delta x}\int_{x_j}^{x_{j+1/2}}\{\rho\Sigma\}(\mathbf{U}_{a_{j+1/2}}(x,t^n+\Delta t))dx =$$

$$(\rho E)_j^n - \frac{2}{\Delta x}(\Delta t\{(\rho\Sigma+\Pi)u\}_{j+1/2}^n - \int_{t^n}^{t^n+\Delta t}\{(\rho\Sigma+\Pi)u\}(\mathbf{U}_{a_{j+1/2}}(x_j,t))dt). \tag{3.105}$$

But exactly the same arguments as those developed in the proof of Lemma 13 apply to prove that under the CFL restriction (3.87):

$$\int_{t^n}^{t^n+\Delta t}\{(\rho\Sigma+\Pi)u\}(\mathbf{U}_{a_{j+1/2}}(x_j,t))dt = \Delta t((\rho\Sigma+\Pi)u)_j^n = \Delta t(\rho H u)_j^n,$$

since and again $\mathbf{U}_j^n$ is by construction at equilibrium.
Let us next prove that the following inequality for all $x \in (x_j, x_{j+1})$:

$$\rho E(\mathbf{u}_{a_{j+1/2}}(x,t^n+\Delta t)) \leq \rho\Sigma(\mathbf{U}_{a_{j+1/2}}(x,t^n+\Delta t)). \tag{3.106}$$

This inequality will then easily imply the required entropy-like inequality from (3.105). To prove (3.106), we first observe that for all $x \in (x_j, x_{j+1/2}+\lambda_1(\mathbf{U}_j^n)\Delta t)$, one trivially has:

$$\{\rho E\}(\mathbf{u}_{a_{j+1/2}}(x,t^n+\Delta t)) = \{\rho\Sigma\}(\mathbf{U}_{a_{j+1/2}}(x,t^n+\Delta t)) = (\rho E)_j^n.$$

Let us now prove that for all x in $\mathcal{I}_1 = (x_{j+1/2}+\lambda_1(\mathbf{U}_j^n)\Delta t, x_{j+1/2}+u^\star\Delta t)$, namely when $\mathbf{U}_{a_{j+1/2}}(x,t^n+\Delta t) = (\mathbf{U}_1)_{j+1/2}^n$, that

$$\rho E(\mathbf{u}_{a_{j+1/2}}(x,t^n+\Delta t)) \leq \rho\Sigma(\mathbf{U}_{a_{j+1/2}}(x,t^n+\Delta t)). \tag{3.107}$$

To that purpose, let us define the following compact $\mathcal{K} = [\mathcal{T}_{min}, \mathcal{T}_{max}]$ where $\mathcal{T}_{min} = \min(\tau_j^n, (\tau_1)_{j+1/2}^n)$ and $\mathcal{T}_{max} = \max(\tau_j^n, (\tau_1)_{j+1/2}^n)$, so that both $\mathcal{T}(x,t^n+\Delta t) = \tau_j^n$ and $\tau(x,t^n+\Delta t) = (\tau_1)_{j+1/2}^n$ in the Riemann solution belongs to this compact for all $x \in \mathcal{I}_1$. But, the first two discrete Whitham conditions just express that $a_{j+1/2}^2$ satisfies the requirement for applying the minimization principle stated in Theorem 7 over the compact $\mathcal{K}$ for any given fixed $x \in \mathcal{I}_1$. This indeed follows from the convexity of each of the partial pressure laws in τ, which ensures the monotonicity of $\tau \to \partial_\tau \sum_{i=1}^N p_i$. The validity of (3.107) thus follows. Exactly in the same way, one can prove that:

$$\rho E(\mathbf{u}_{a_{j+1/2}}(x,t^n+\Delta t)) \leq \rho\Sigma(\mathbf{U}_{a_{j+1/2}}(x,t^n+\Delta t)),$$

for all $x \in (x_{j+1/2}+u^\star\Delta t, x_{j+1/2}+\lambda_3(\mathbf{U}_{j+1}^n)\Delta t)$ while:

$$\{\rho E\}(\mathbf{u}_{a_{j+1/2}}(x,t^n+\Delta t)) = \{\rho\Sigma\}(\mathbf{U}_{a_{j+1/2}}(x,t^n+\Delta t)) = (\rho E)_{j+1}^n,$$

when $x \in (x_{j+1/2}+\lambda_3(\mathbf{U}_{j+1}^n)\Delta t, x_{j+1})$. This completes the proof.

3.4.5 The new averaging procedure in the relaxation approach

This section aims at briefly exemplifying the averaging techniques introduced in Definition 4 within the frame of the present relaxation procedure. Let us from now on notice that the Riemann solutions associated with the relaxation system share with the exact equilibrium Riemann solutions the property that the specific entropies are just the superposition of step functions, each propagating with the speed of a contact discontinuity. Consequently, the representation formula (3.65) applies and exactly the same definitions (3.67) of the discrete velocities $v_i^{j+1/2,\pm}$ follow to get the next updating formulae :

$$(s_i)_j^{n+1=} = (s_i)_j^n - \frac{\Delta t}{\Delta x} v_i^{j-1/2,+}((s_i)_j^n - (s_i)_{j-1}^n) - \frac{\Delta t}{\Delta x} v_i^{j+1/2,-}((s_i)_{j+1}^n - (s_i)_j^n). \quad (3.108)$$

Here the very benefit comes from the very simple structure of each of the Riemann solutions which makes fairly straightforward the evaluation of

$$v_i^{j+1/2,\pm} = v_i^{\pm}(\mathbf{u}_j^n, \mathbf{u}_{j+1}^n) = \int_0^{(u^\star_{j+1/2})^{\pm}} \rho^{\gamma_i}(\xi)d\xi / \overline{\rho^{\gamma_i}}_j^{n+1=}. \quad (3.109)$$

The relevance of the Relaxation approach with respect to this new averaging technique is assessed by this last statement :

Proposition 18 *Under the CFL restriction (3.87), the following energy-like inequalities hold :*

$$\{\rho E\}(\mathbf{u}_j^{n+1=}) - (\rho E)_j^n + \lambda\Delta\{\rho\mathcal{H}u\}_{j+1/2}^n \leq 0, \quad j \in Z, \quad (3.110)$$

where with the notations of the previous section, the numerical energy flux reads :

$$\{\rho\mathcal{H}u\}_{j+1/2}^n = \{(\rho\Sigma + \Pi)u\}(\mathbf{U}_{a_{j+1/2}}(0^+; \mathbf{U}_j^n, \mathbf{U}_{j+1}^n)).$$

Proof As an immediate consequence of the inequalities (3.104), we get under the CFL restriction (3.87) :

$$\frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{\rho E\}(\overline{\mathbf{w}}(x, t^n + \Delta t))dx \leq (\rho E)_j^n - \frac{\Delta t}{\Delta x}\Delta\{\rho\mathcal{H}u\}_{j+1/2}^n,$$

where we have set :

$$\overline{\mathbf{w}}(x, t^n + \Delta t) = \begin{cases} \mathbf{u}_{a_{j-1/2}}(x, t^n + \Delta t), & x \in (x_{j-1/2}, x_j), \\ \mathbf{u}_{a_{j+1/2}}(x, t^n + \Delta t), & x \in (x_j, x_{j+1/2}). \end{cases}$$

Recalling the identities (3.101)-(3.102), we have thus arrived at exactly the same form of inequality (3.43) which is the starting point of the proof of the Lemma 11. The conclusion follows when applying exactly the same steps.

Observe that after the correction step coming with the present relaxation approach, the specific entropies receive exactly the same formula as the one proposed in (3.54). The positiveness of these corrected specific entropies is here again ensured by the energy inequalities (3.110). To conclude, let us emphasize that Corollary 4 applies to the present method exactly in the same lines.

3.5 Numerical experiments

This section aims at illustrating that the new averaging procedures we have described succeed in capturing numerical solutions of (3.1) in fairly good agreement with exact ones. We consider constant viscosity laws $\{\mu_i\}_{i=1,\dots,N}$ with a Reynolds number equal to 10^5 which is typical in aerodynamic. The numerical experiments we propose consist in two computations performed when considering Riemann initial data:

$$\mathbf{V}(x,0) = \begin{cases} \mathbf{V}_L & \text{if} \quad x < 0, \\ \mathbf{V}_R & \text{if} \quad x > 0, \end{cases} \tag{3.111}$$

for some given $\mathbf{V}_L$ and $\mathbf{V}_R$ in the phase space Ω, and when dealing with an uniform mesh made of 300 points. We then compare the corresponding numerical solutions together with the exact ones. To assess the importance of the correction step, we also display the approximate solutions obtained by a classical splitting approach (without correction procedure).

Experiment 1

We set $N = 3$ and choose

γ_1	γ_2	γ_3	μ_2/μ_1	μ_3/μ_1
1.4	1.6	1.4	1.	1.

while the left and right states entering (3.111) write as follows:

Experiment 1

$\mathbf{V}_L$	density	velocity	pressure 1	pressure 2	pressure 3
	3.	2.	1.5	1.	1.

$\mathbf{V}_R$	density	velocity	pressure 1	pressure 2	pressure 3
	2.6848	-2.1586	1.1930	0.7086	0.7649

Let us observe that despite the exact solution exhibits large pressure jumps, the numerical solution produced by our scheme fully agrees with the exact one, contrary to the approximate solution given by the classical splitting approach.

Experiment 2

We now deal with ratios of the viscosity laws $\{\mu_i/\mu_1\}_{i=2,3}$ large with respect to 1. More precisely, we consider:

γ_1	γ_2	γ_3	μ_2/μ_1	μ_3/μ_1
1.4	1.6	1.6	100.	100.

The left and right states entering (3.111) read as follows:

Experiment 2

$\mathbf{V}_L$

density	velocity	pressure 1	pressure 2	pressure 3
1.	1.	1.	1.	0.6

$\mathbf{V}_R$

density	velocity	pressure 1	pressure 2	pressure 3
1.4291	-0.5477	0.2277	0.4355	0.0929

Once more, let us notice the excellent agreement between the solutions given by our procedure and the exact ones.

Experiment 1

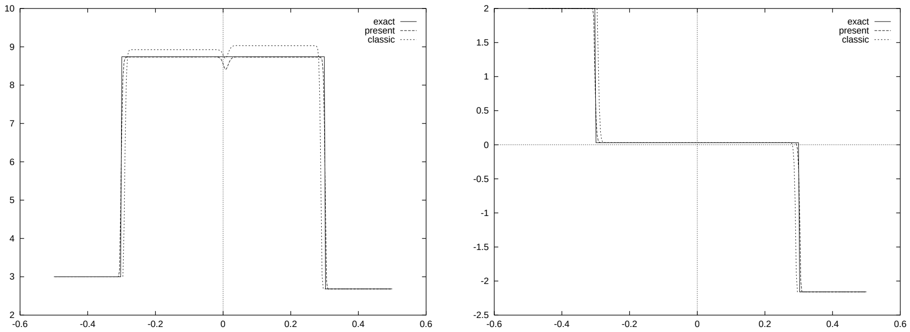

FIG. 3.1 – *Density, Velocity*

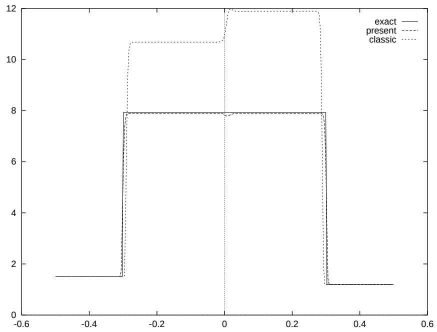

FIG. 3.2 – *Pressure 1*

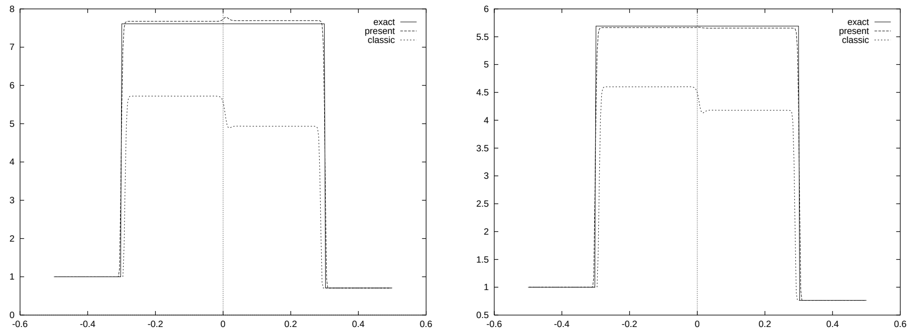

FIG. 3.3 – *Pressure 2, Pressure 3*

Experiment 2

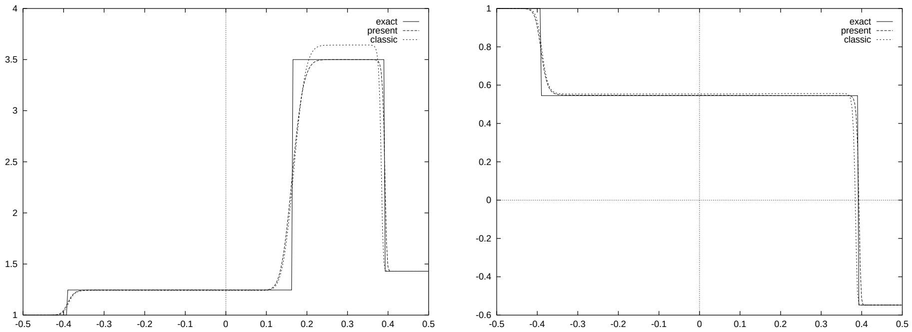

FIG. 3.4 – *Density, Velocity*

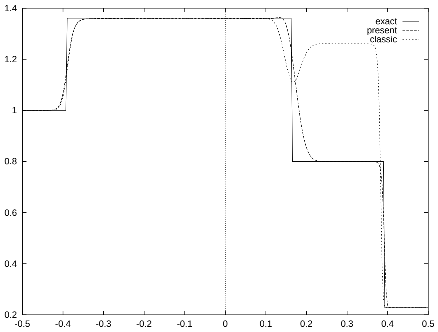

FIG. 3.5 – *Pressure 1*

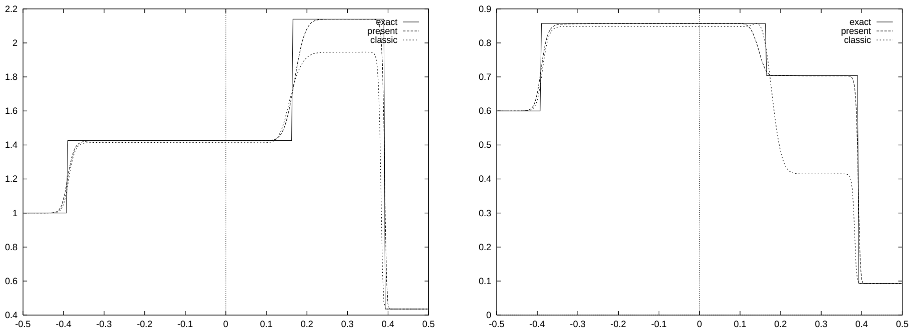

FIG. 3.6 – *Pressure 2, Pressure 3*

Chapitre 4

Euler equations with several independent pressure laws and entropy satisfying explicit projection schemes

Ce chapitre fait l'objet d'un article en préparation.

Euler equations with several independent pressure laws and entropy satisfying explicit projection schemes

Chalons C.[1,2] and Coquel F.[3]

Résumé

This work aims at numerically approximating the entropy weak solutions of Euler-like systems asymptotically recovered from the multi-pressure Navier-Stokes equations in the regime of an infinite Reynolds number. The non conservation form of the limit PDE model makes the shock solutions to be sensitive with respect to the underlying small scales. Here we propose to encode these small scale effects via a set of generalized jump conditions expressed in terms of the independent internal energies. The interest in considering internal energies stems from the presence of solely first order non conservative products by contrast to other variables. These non conservative products are defined in the now classical sense proposed by Dal Maso, LeFloch and Murat [42]. We show how to enforce the generalized jump conditions at the discrete level with a fairly simple numerical procedure. This method is proved to satisfy a full set of stability estimates and to produce approximate solutions in good agreement with exact Riemann solutions.

4.1 Introduction

This work is devoted to the numerical approximation of the solutions of the Navier-Stokes equations with several independent pressure laws, in the regime of an infinite Reynolds number. Here, $N \geq 2$ pressure laws enter the PDE model via N independent governing equations. The sum of the N partial pressures determines the total pressure of the fluid. The systems under consideration are nothing but a natural extension of the classical Navier-Stokes equations (*i.e.* with $N = 1$) and several models from the physics (plasma physics, turbulence) actually enter the present framework (see for instance [10], [19]). In many situations, these models come with a very large Reynolds number. Therefore and away from some boundaries, it sounds natural to neglect the small scales due to the viscous perturbations so as to consider some asymptotic Euler system in the limit of an infinite Reynolds number. However, in spite of the fact that these models exhibit close relashionships with the usual setting, a very difference arises when addressing the case $N \geq 2$. Namely the systems under consideration are then generally in non conservation form in the sense that non conservative products will always persist through any given admissible change of variables. Such non conservative products will involve either the unknown with its first order derivative or the unknown

1. O.N.E.R.A., B.P. 72, 29 avenue de la Division Leclerc, 92322 Châtillon Cedex, France.
E-mail: Christophe.Chalons@onera.fr.
2. Centre de Mathématiques Appliquées & Centre National de la Recherche Scientifique, U.M.R. 7641, École Polytechnique, 91128 Palaiseau Cedex, France.
E-mail: chalons@cmapx.polytechnique.fr
3. Laboratoire JLL & Centre National de la Recherche Scientifique, U.M.R. 7598, Boîte courrier 187, 75252 Paris Cedex 05, France.
E-mail: coquel@ann.jussieu.fr

with its second order derivative, but generally speaking both types of products do coexist. The systems under study take the following condensed form :

$$\partial_t \mathbf{V}^\epsilon + \mathbf{A}(\mathbf{V}^\epsilon)\partial_x \mathbf{V}^\epsilon = \epsilon \mathcal{R}(\mathbf{V}^\epsilon, \partial_x(\mathcal{B}(\mathbf{V}^\epsilon)\partial_x \mathbf{V}^\epsilon)), \tag{4.1}$$

where the rescaling parameter $\epsilon > 0$ stands as the inverse of the Reynolds number. The lack of full conservation form forbids us from using the classical distributions theory for defining the limit solutions as $\epsilon \to 0$. Several theories have been developed to handle the singular limit and we quote for instance the works by LeFloch, Dal Maso and Murat [42], LeFloch [90], Colombeau and Leroux [31], and more recently by Berthon et al. [15] (see also [21]). For our forthcoming numerical purposes, the problematic discussed hereafter turns out to fall in a very natural way within the mathematical framework proposed by LeFloch [90]. Indeed, this author suggests to define shock solutions for the limiting system as the limit of travelling wave solutions of (4.1) as the rescaling parameter ϵ goes to zero (see also Raviart-Sainsaulieu [112]). Such a natural approach allows the definition of the singular non conservative products in the limit $\epsilon \to 0$ in terms of a fixed family of paths which coincide with the graph of the travelling wave solutions. We shall adopt this approach in the present work. But let us emphasize that all the works we have reported highlight a deep discrepancy with respect to the purely conservative setting with genuine non linearity, namely the definition of shock solutions is sensitive with respect to the small scales present in the viscous perturbations. Sensitivity occurs in the sense that non proportional diffusive tensors give rise to distinct shock solutions. Put in other words, the lack of an admissible change of variables which recast the system in full conservation form makes the definition of the endpoints of the travelling wave solutions sensitive with respect to the shape of the diffusive tensor. This sensitiveness in the definition of discontinuous solutions is also known in other settings. Let us quote hyperbolic systems in conservation form but with non linear fields that are not genuinely non linear, or mixed hyperbolic-elliptic systems (again in conservation form) describing phase transformations (see [88], [65], [25], ...).

Due to the sensitiveness we have just put forward, the numerical approximation of shock solutions associated with (4.1) in the asymptotic regime $\epsilon \to 0$ is a challenging issue. The very difficulty stays in the fact that the precise shape of the diffusive tensor must be carefully accounted for. In terms of numerical methods, this means that the numerical viscosity and the exact one must fit in order to capture numerical solutions in agreement with the expected ones. Recent works have been devoted to this matter. In [10], Berthon considers the numerical approximation of the solutions of the Navier-Stokes equations (*i.e.* with fixed $\epsilon > 0$) when addressing two pressure laws ($N = 2$). He proves that classical approaches like splitting strategies between the convective and the diffusive parts for approximating the solutions of the system under study actually fail. They produce approximate solutions which exhibit large errors with the exact ones, and the main reason comes from a too large numerical dissipation in the procedure with respect to the expected one. In a loose sense, the precise shape of the diffusive tensor in (4.1) is not exactly taken into account. To overcome this difficulty, Berthon introduces a correction step, referred as a "non linear projection" step, to enforce for validity at the discrete level some generalized jump conditions. These conditions express, in terms of entropies, the dependence of the shock solutions with respect to the exact diffusive tensor when encoding all the corresponding small scales effects. In practice, this correction step consists in solving a non linear and scalar algebraic problem on each computational cell. The resulting numerical scheme (prediction step + correction step), named "non linear projection scheme", obeys all the required positivity preserving and stability properties (entropy inequalities, maximum principles, ...), and we refer the reader to [10] for some numerical experiments illustrating the correct design of this strategy. In [20], Chalons and Coquel propose an extension of this scheme when dealing with $N > 2$ independent pressure laws. In this context, $(N-1)$ distinct generalized jump conditions that reflect the sensitiveness of the solutions to the diffusive tensor (still in terms of entropies) are enforced for validity at the discrete level by the mean of a non linear correction step. Once more, the procedure succeeds in capturing numerical solutions in agreement with exact ones while satisfying the same stability properties as in the setting $N = 2$. But again at the expense of the need for solving a succession of local non linear algebraic problems. To overcome this drawback, we propose in this work to enforce for validity the generalized jump conditions when expressed in terms of the internal energies instead of the entropies. The benefit of dealing with internal energies is that the correction step becomes fully explicit due to the linear dependence of the total energy with respect to the partial internal energies. As a consequence, we no longer have to solve local non linear algebraic problems. In addition, this new approach shares with the one proposed in [20] (and also in [10]) the same positivity preserving and stability properties.

The present paper is organized as follows. The first section describes the PDE systems under study with a special emphasis put on the consequences of the lack of an equivalent conservation form. In particular, we highlight that shock solutions associated with these systems (in the asymptotic regime $\epsilon \to 0$) actually depend on the precise shape of the corresponding diffusive tensors, via the ratios of the viscosities $\mu_i/(\sum_{j=1}^{N} \mu_j)$ with $i = 1, ..., N$. To this purpose, we put forward the so-called generalized jump conditions and we introduce an equivalent formulation (for smooth solutions) of (4.1) with second order operator in full conservation form. The interest in considering this equivalent form stems from the fact that it entirely encodes the small scales effects via the ratios of the viscosity laws. The second section is devoted to the study of the asymptotic regime $\epsilon \to 0$, when exhibiting the Euler like system which governs the limiting solutions of (4.1). With this aim, we deal with the two equivalent formulations of the previous section and for defining the non conservative products entering the underlying first order systems, we rely on the theory of paths introduced by Dal Maso, LeFloch and Murat [42] (see also LeFloch [90]). Section 4.4 presents a suitable numerical method for approximating the weak solutions of the limiting systems on the basis of the Godunov method. Its design principle is focused on rightly dealing with the exact diffusive tensor, while ensuring stability properties at the expense of an interesting numerical cost. Section 4.5 generalizes this scheme in the Approximate Riemann Solvers framework proposed by Harten, Lax and Van Leer [61]. To conclude, section 4.6 gives some numerical evidences of the correct design of the scheme.

4.2 The mathematical setting

In this paper, we are interested in the numerical approximation of the solutions of the following system in non conservation form :

$$\left\{\begin{array}{l}\partial_t \rho^\epsilon + \partial_x \rho u^\epsilon = 0, \\ \partial_t \rho u^\epsilon + \partial_x (\rho u^{\epsilon 2} + \sum_{i=1}^N p_i^\epsilon) = \epsilon \partial_x (\sum_{i=1}^N \mu_i \partial_x u^\epsilon), \\ \partial_t \rho \varepsilon_i^\epsilon + \partial_x \rho \varepsilon_i u^\epsilon + p_i^\epsilon \partial_x u^\epsilon = \epsilon \mu_i (\partial_x u^\epsilon)^2, \quad i = 1, ..., N,\end{array}\right. \tag{4.2}$$

in the regime of a vanishing rescaling parameter $\epsilon > 0$. The PDE system readily stands as a natural extension of the usual Navier-Stokes equations when a single pressure is involved in the momentum equation. Here, N independent pressure laws occur and are governed via N internal energies $\rho\varepsilon_i$. Several models from the physics actually enter the present framework since their associated PDE system take the form (4.2) but for a prescribed number N of pressure laws. We refer the reader to [19] for a detailed presentation. The required closure equations are as follows. The internal energies are assumed to obey the second principle of the thermodynamics, *i.e.* $\rho\varepsilon_i$ with $i \in \{1, ..., N\}$ is associated with an entropy ρs_i solution of

$$- T_i ds_i = d\varepsilon_i + p_i d\tau, \quad \tau = 1/\rho, \tag{4.3}$$

such that the mapping $(\tau, s_i) \to \varepsilon_i(\tau, s_i)$ is strictly convex. Then we have from (4.3),

$$p_i(\tau, s_i) = -\partial_\tau \varepsilon_i(\tau, s_i), \quad T_i(\tau, s_i) = -\partial_{s_i} \varepsilon_i(\tau, s_i),$$

where the temperature T_i is classically assumed to stay positive. As a well known consequence, the well defined mapping $(\tau, \varepsilon_i) \to s_i(\tau, \varepsilon_i)$ is strictly convex and so is also, with some little abuse in the notation, the mapping $(\rho, \rho\varepsilon_i) \to \{\rho s_i\}(\rho, \rho\varepsilon_i) := \rho s_i(\frac{1}{\rho}, \frac{\rho\varepsilon_i}{\rho})$. Each pressure law p_i will be assumed in addition to obey the general Weyl's conditions for real gases (see [57] for the details). For simplicity, the viscosity coefficients μ_i will be assumed to be non negative real constants such that $\sum_{i=1}^N \mu_i > 0$. In what follows, we shall tacitly assumed that $\mu_N > 0$ and this is actually the case up to some relabelling in the N entropies.
To shorten the notations, the system (4.2) is given the following condensed form :

$$\partial_t \mathbf{V}^\epsilon + \mathbf{A}(\mathbf{V}^\epsilon) \partial_x \mathbf{V}^\epsilon = \epsilon \mathcal{R}(\mathbf{V}^\epsilon, \partial_x(\mathcal{B}(\mathbf{V}^\epsilon) \partial_x \mathbf{V}^\epsilon)), \tag{4.4}$$

where the unknown $\mathbf{V}^\epsilon$ is associated with the following natural phase space :

$$\Omega = \{\mathbf{V} = (\rho, \rho u, \{\rho\varepsilon_i\}_{i=1,\dots,N}) \in I\!R^{N+2} / \rho > 0, \rho u \in I\!R, \{\rho\varepsilon_i > 0\}_{i=1,\dots,N}\}.$$

The definition of the tensor $\mathcal{B}$ in (4.4) follows when writing

$$\mu_i (\partial_x u)^2 = \mu_i (\partial_x (u \partial_x u) - u \partial_{xx} u). \tag{4.5}$$

Let us underline that the non conservative products entering the definition of (4.4) involve not only the unknown $\mathbf{V}$ with its first derivative but also the unknown with its second derivative as expressed by the identity (4.5). In other words, both the first order operator and second order perturbation naturally occur in non conservation form. Let us first state some elementary properties of the non linear first order system. By virtue of the convexity of the $\varepsilon_i(\tau, s_i)$ mappings with $i \in \{1, ..., N\}$, we have the following statement which easy proof is left to the reader :

Lemma 15 *The first order underlying system in (4.2) is hyperbolic over the phase space Ω and admits the following three distinct eigenvalues :*

$$u - c, \quad u, \quad u + c \quad \text{with} \quad c^2 = \sum_{i=1}^{N} c_i^2, \quad c_i^2 = \partial^2_{\tau\tau}\varepsilon_i(\tau, s_i) > 0, \tag{4.6}$$

where u has N order of multiplicity.
Moreover, under the Weyl's assumptions, the two extreme fields are genuinely non linear while the intermediate fields associated with the eigenvalue u are linearly degenerate. In addition, the latter admits the velocity u and the total pressure $\sum_{i=1}^{N} p_i$ as two independent Riemann invariants.

The system under consideration exhibits close relationships with the usual setting of the Navier-Stokes equations with a single pressure law, but naturally writes in non conservation form. This rises the question of the existence of additional conservation laws for smooth solutions of (4.2) and this is precisely the matter of the next statement :

Proposition 19 *Smooth solutions of (4.2) satisfy the following conservation law :*

$$\partial_t \rho E^\epsilon + \partial_x \rho H u^\epsilon = \epsilon \partial_x ((\sum_{i=1}^{N} \mu_i) u^\epsilon \partial_x u^\epsilon), \tag{4.7}$$

where the total energy ρE and the total enthalpy ρH respectively read :

$$\rho E = \frac{(\rho u)^2}{2\rho} + \sum_{i=1}^{N} \rho\varepsilon_i, \quad \rho H = \rho E + \sum_{i=1}^{N} p_i.$$

These solutions also obey the following N entropy balance equations :

$$\partial_t \rho s_i^\epsilon + \partial_x \rho s_i u^\epsilon = -\frac{\mu_i}{T_i^\epsilon} (\partial_x u^\epsilon)^2, \quad i = 1, ..., N. \tag{4.8}$$

Note from (4.8) that classical non linear tranformations in the s_i yield further additional balance equations for governing $\varphi(s_1, ..., s_N)$ where $\varphi : \mathbb{R}^N \to \mathbb{R}$ denotes any given arbitrary smooth function. Nevertheless and without specific assumptions on the thermodynamic closure equations, none of these additional equations boils down to a non trivial additional conservation law. With this respect, the system (4.2) is genuinely in non conservation form. In order to put forward a deep consequence of the lack of conservation form, it is useful to illustrate the nature of the non trivial conservation laws coming with restrictive modelling assumptions. In that aim, we briefly discuss in the next subsection a particular setting for which the system (4.2) recasts in full conservation form but with the striking property that the limit system obtained in the asymptotic regime $\epsilon \to 0$ does encode an explicit souvenir of the vanishing viscous perturbation. Namely, the available additional conservation laws explicitly involve the ratios of the viscosity coefficients. Such a souvenir exactly expresses the property that the N independent entropies cannot achieve independent jumps for shock solutions but on the contrary jumps that obey proportionality relations, the latter being precisely dictated by the ratios of the viscosities. In the general setting for (4.2), such proportionality conditions will obviously persist but in a weaker form and will be referred as to generalized jump conditions. We propose to point out this deep property on a simple example for its primary importance in the numerical approximation of the solutions of (4.2).

4.2.1 A simplified setting

Let us temporarily choose $N = 2$ when considering a polytropic closure relation for defining each of the two independent pressure laws. Let us assume in addition the same constant adiabatic exponent, say $\gamma > 1$, for the two pressures so that:

$$p_i = (\gamma - 1)\rho\varepsilon_i, \quad i = 1, 2. \tag{4.9}$$

Let us then complete the above equations of state when considering the following specific entropies

$$s_i = p_i/\rho^\gamma, \quad i = 1, 2. \tag{4.10}$$

The interest of such a simplified setting finds its root in the following statement:

Proposition 20 *Assuming the modelling assumptions (4.9)–(4.10), the smooth solutions of (4.2) obey equivalently the following system in full conservation form:*

$$\left\{\begin{array}{l} \partial_t \rho^\epsilon + \partial_x \rho u^\epsilon = 0, \\ \partial_t \rho u^\epsilon + \partial_x (\rho u^{\epsilon 2} + \sum_{i=1}^2 p_i^\epsilon) = \epsilon \partial_x (\sum_{i=1}^2 \mu_i \partial_x u^\epsilon), \\ \partial_t \rho E^\epsilon + \partial_x (\rho E^\epsilon + \sum_{i=1}^2 p_i^\epsilon) u^\epsilon = \epsilon \partial_x (\sum_{i=1}^2 \mu_i u^\epsilon \partial_x u^\epsilon), \\ \partial_t \rho (s_1 - \frac{\mu_1}{\mu_2} s_2)^\epsilon + \partial_x \rho (s_1 - \frac{\mu_1}{\mu_2} s_2) u^\epsilon = 0. \end{array}\right. \tag{4.11}$$

Proof
This result is a direct consequence of proposition 19 when noticing that (4.8) gives in the present simplified setting:

$$\frac{1}{\mu_i}\Big\{\partial_t \rho s_i^\epsilon + \partial_x \rho s_i u^\epsilon\Big\} = \epsilon \frac{\gamma - 1}{\rho^{\epsilon \gamma - 1}} (\partial_x u^\epsilon)^2, \quad i = 1, 2. \tag{4.12}$$

Then substrating these two equations readily yields the last non trivial conservation law in (4.11). To conclude it suffices to notice that the mapping $\mathbf{V} \in \Omega \to^T (\rho, \rho u, \rho E, \rho(s_1 - \frac{\mu_2}{\mu_1} s_2))$ yields an admissible change of variables.
As an immediate (but formal) consequence of proposition 20, we next state:

Corollary 6 *Assuming that smooth solutions $\{\mathbf{V}^\epsilon\}_{\epsilon > 0}$ of (4.11) converge strongly (in relevant topologies) as ϵ goes to 0 to a limit function $\mathbf{V}$, then $\mathbf{V}$ is a solution in sense of the distributions of the following non linear hyperbolic system:*

$$\left\{\begin{array}{l} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + \sum_{i=1}^2 p_i) = 0, \\ \partial_t \rho E + \partial_x (\rho E + \sum_{i=1}^2 p_i) u = 0, \\ \partial_t \rho (s_1 - \frac{\mu_1}{\mu_2} s_2) + \partial_x \rho (s_1 - \frac{\mu_1}{\mu_2} s_2) u = 0, \end{array}\right. \tag{4.13}$$

supplemented with the Lax entropy inequality :

$$\partial_t \rho h(s_i) + \partial_x \rho h(s_i) u \leq 0, \quad h(s_i) = -\ln(s_i), \quad i = 1, 2. \tag{4.14}$$

In the zone of smoothness of the limit solution, $\mathbf{V}$ *satisfies in addition to the conservation laws for* $\rho, \rho u$ *and* ρE *:*

$$\partial_t \rho s_i + \partial_x \rho s_i u = 0, \quad i = 1, 2, \tag{4.15}$$

while and by contrast across shocks (i.e. discontinuities associated with the two extreme fields, see lemma 15), the limit solution must obey with classical notations the following Rankine-Hugoniot conditions :

$$\left\{ \begin{array}{l} m = \rho^-(u^- - \sigma) = \rho^+(u^+ - \sigma), \\ m(u^+ - u^-) + (\sum_{i=1}^2 p_i^+ - \sum_{i=1}^2 p_i^-) = 0, \\ m(E^+ - E^-) + ((\sum_{i=1}^2 p_i u)^+ - (\sum_{i=1}^2 p_i u)^-) = 0, \\ (s_1^+ - s_1^-) = \frac{\mu_1}{\mu_2}(s_2^+ - s_2^-), \end{array} \right. \tag{4.16}$$

supplemented with the entropy condition :

$$m(h(s_i^+) - h(s_i^-)) < 0, \quad i = 1, 2. \tag{4.17}$$

The smooth mapping h entering the inequalities (4.14) and (4.17) is considered for the sake of convexity of the mapping $(\rho, \rho\varepsilon_i) \to \{\rho h(s_i)\}(\rho, \rho\varepsilon_i)$.

Proof

Assuming a zone of smoothness for $\mathbf{V}$, classical manipulations on the momentum and total energy equations yield (see [57] in the setting of the usual Euler equations) :

$$\partial_t \rho(\varepsilon_1 + \varepsilon_2) + \partial_x \rho(\varepsilon_1 + \varepsilon_2) u + \sum_{i=1}^{2} p_i \partial_x u = 0,$$

that is to say, in view of (4.3) :

$$\partial_t \rho(s_1 + s_2) + \partial_x \rho(s_1 + s_2) u = 0,$$

and henceforth the expected identities (4.15) when invoking the last conservation law in (4.13). Turning now considering shock solutions, we just underline that the Rankine-Hugoniot conditions coming with (4.13) clearly give :

$$-\sigma[\rho(s_1 - \frac{\mu_1}{\mu_2} s_2)] + [\rho(s_1 - \frac{\mu_1}{\mu_2} s_2) u] = 0,$$

which is easily seen to boil down to the expected one in (4.16) by virtue of a non zero relative mass flux m for the discontinuities under consideration (again see [57] for the details).

Let us briefly rephraze the above results. First the conservation of the entropies (4.15) in the zone of smoothness of the limit solution is a well-known and expected property. The breakdown of conservation (4.17) for shock solutions is also well-known but the striking novelty here stays in the property that the jumps in the two specific entropies must be kept in balance according to the ratio of the two viscosities. This unusual Rankine-Hugoniot condition obviously implies that the shock solutions propagating with speed σ and separating two constant states $\mathbf{V}^-$ and $\mathbf{V}^+$ have the property that $\mathbf{V}^+$ does not only depend on σ and $\mathbf{V}^-$ but also on the ratio of the viscosities μ_1/μ_2 : *i.e* $\mathbf{V}^+ := \{\mathbf{V}^+\}(\sigma, \mathbf{V}^-, \mu_1/\mu_2)$. In other words, modifying the ratio μ_1/μ_2 does modify $\mathbf{V}^+$, σ and $\mathbf{V}^-$ being fixed. Such a dependency is actually natural and illustrates a general property met by shock solutions for hyperbolic systems in non conservation form (see [85], [90], [112] and below).

4.2.2 The general setting

Let us turn back to the general setting for which (4.2) is genuinely in non conservation form. Actually and as shown below, this system can be given two convenient equivalent forms for smooth solutions with the following distinctive properties. A first formulation which allows to recast the first order operator in conservation form but not its second order counterpart. Then a second formulation where by contrast the second order operator achieves a conservation form to the detriment of the first order one. Both formulations are deeply motivated by the need to track closely the ratios of the viscosity coefficients in the governing equations. More precisely, both formulations intend to encode the property that the rates of dissipation governing each of the entropies ρs_i, $i \in \{1, ..., N\}$, in their evolution in time, are actually kept in balance according to such ratios.
The first formulation readily follows from proposition 19 when considering the admissible change of variables given by :

$$\mathbf{U} := \mathbf{U}(\mathbf{V}) = (\rho, \rho u, \{\rho s_i\}_{i=1,\dots,N}), \tag{4.18}$$

so that (4.2) is given the following equivalent form :

$$\partial_t \mathbf{U}^\epsilon + \partial_x \mathbf{F}(\mathbf{U}^\epsilon) = \epsilon \mathcal{R}(\mathbf{U}^\epsilon, \partial_x(\mathcal{C}(\mathbf{U}^\epsilon)\partial_x \mathbf{U}^\epsilon)). \tag{4.19}$$

Here and with clear notations, the flux function $\mathbf{F} : \Omega \to I\!R^{N+2}$ reads $\mathbf{F}(\mathbf{U}) = (\rho u, \rho u^2 + \sum_{i=1}^N p_i, \{\rho s_i u\}_{i=1,\dots,N})$ while the diffusive tensor $\mathcal{C}(\mathbf{U}^\epsilon)$ finds a clear definition (see indeed (4.5)). Let us then underline that the N entropy balance equations (4.8) entering the definition of (4.19) do not evolve independently but actually proportionally according to :

$$\tfrac{\mu_N}{T_N^\epsilon}\{\partial_t \rho s_i^\epsilon + \partial_x \rho s_i u^\epsilon\} = \tfrac{\mu_i}{T_i^\epsilon}\{\partial_t \rho s_N^\epsilon + \partial_x \rho s_N u^\epsilon\} \quad 1 \le i \le N-1. \tag{4.20}$$

These identities simply reflect a (trivial) cancellation property in $(\partial_x u^\epsilon)^2$ we have already used within the simplified setting. Of course, (4.20) is nothing but a natural extension of the non trivial conservation law built on the ratio of the viscosities in the system (4.11). Next, let us observe that the smooth solutions of (4.2) obviously satisfy the additional conservation law (4.7) for governing the total energy ρE :

$$\partial_t\{\rho E\}(\mathbf{U}^\epsilon) + \partial_x\{\rho H u\}(\mathbf{U}^\epsilon) = \epsilon \partial_x(\sum_{i=1}^N \mu_i u^\epsilon \partial_x u^\epsilon). \tag{4.21}$$

The equalities (4.20) and (4.21) stay at the basis of the work [20] (see also [19]) where a natural generalization of the system (4.11) to the general setting is proposed. Such an extension has been proved to be successful in the numerical approximation of the solutions of (4.19) in the limit $\epsilon \to 0$ which we again emphasize to be heavily driven by the property that entropy rates of dissipation must be kept in balance according to (4.20). At the cornerstone of this success is a non linear correction procedure for (4.19) which aims at enforcing for validity at the discrete level the $(N-1)$ identities in (4.20) while ensuring the conservation of the total energy in (4.21). Roughly speaking, the non linearities in the procedure just reflect the non linearities in the total energy ρE when expressed in terms of the entropies $\{\rho s_i\}_{1\leq i\leq N}$ (see [20], [19] for the details). Observing by contrast that the total energy ρE does linearly depend on the internal energies $\{\rho\varepsilon_i\}_{1\leq i\leq N}$ strongly promotes to introduce an alternative formulation to (4.19), now built on the internal energies but which mimics the key equalities (4.20). The second principle of the thermodynamics (4.3) obviously allows for this, since it equivalently recasts (4.20) as:

$$\mu_N\{\partial_t\rho\varepsilon_i^\epsilon + \partial_x\rho\varepsilon_i u^\epsilon + p_i^\epsilon\partial_x u^\epsilon\} = \mu_i\{\partial_t\rho\varepsilon_N^\epsilon + \partial_x\rho\varepsilon_N u^\epsilon + p_N^\epsilon\partial_x u^\epsilon\}, \qquad 1\leq i\leq N-1. \tag{4.22}$$

Therefore and since the viscosities are constant real numbers, we get:

$$\partial_t\rho(\varepsilon_i - \tfrac{\mu_i}{\mu_N}\varepsilon_N)^\epsilon + \partial_x\rho(\varepsilon_i - \tfrac{\mu_i}{\mu_N}\varepsilon_N)u^\epsilon + (p_i - \tfrac{\mu_i}{\mu_N}p_N)^\epsilon\partial_x u^\epsilon = 0, \qquad 1\leq i\leq N-1. \tag{4.23}$$

These identities then suggest us to introduce:

Lemma 16 *The following mapping:*

$$\mathbf{W} := \mathbf{W}(\mathbf{V}) = (\rho, \rho u, \rho E, \{\rho w_i\}_{i=1,\dots,N-1}), \qquad \rho w_i = \rho(\varepsilon_i - \tfrac{\mu_i}{\mu_N}\varepsilon_N), \quad 1\leq i\leq N-1, \tag{4.24}$$

yields an admissible change of variables which equivalently recasts the system (4.2) under the following form:

$$\partial_t\mathbf{W}^\epsilon + \mathbf{B}(\mathbf{W}^\epsilon)\partial_x\mathbf{W}^\epsilon = \partial_x(\mathcal{D}(\mathbf{W}^\epsilon)\partial_x\mathbf{W}^\epsilon), \tag{4.25}$$

where $\mathbf{B}$ *and* $\mathcal{D}$ *find clear definitions from (4.23) and (4.7).*

This result simply follows when noticing that:

$$\{\rho\varepsilon_i\}(\mathbf{W}) = \rho w_i + \frac{\mu_i}{\sum_{j=1}^N \mu_j}\Big(\rho E - \frac{(\rho u)^2}{2\rho} - \sum_{j=1}^{N-1}\rho w_j\Big), \quad 1\leq i\leq N-1,$$

while of course

$$\{\rho\varepsilon_N\}(\mathbf{W}) = \rho E - \frac{(\rho u)^2}{2\rho} - \sum_{j=1}^{N-1}\rho\varepsilon_j(\mathbf{W}).$$

Let us again emphasize that the interest in considering the equivalent form (4.25) over the original one (4.2) stems from the fact that the balance to be kept in the entropy rates of dissipation is built in but when expressed in terms of the internal energies: namely (4.22). With this respect, the form (4.25) will be shown to be particularly suited for our numerical purposes. Since we are interested in the asymptotic regime of a vanishing parameter ϵ, we propose to make meaningfull in the next section the identities (4.22) and thus the change of variables (4.24) in the limit $\epsilon \to 0$.

4.3 The asymptotic regime

This section aims at exhibiting the asymptotic first order system for governing limit solutions of (4.2) in the regime $\epsilon \to 0$. This system will arise as a natural extension of the limit model (4.13) obtained under simplifying assumptions. We also extend here the equivalence lemma 16 between the two formulations (4.2) and (4.25) to a class of discontinuous functions which naturally arise in the limit $\epsilon \to 0$. Here, we specifically focus ourselves on piecewise Lipschitz continuous functions.
Let us first motivate the forthcoming developments when formally addressing the asymptotic regime $\epsilon \to 0$ in the non linear system (4.2). As it is well known, solutions of the underlying hyperbolic system may develop discontinuities in finite time. As a consequence, families of smooth solutions $\{\mathbf{V}^\epsilon\}_{\epsilon>0}$ fail generally speaking to converge to smooth functions and this rises the question of the singular limit of $\mathbf{A}(\mathbf{V}^\epsilon)\partial_x\mathbf{V}^\epsilon$. Assuming a piecewise Lipschitz continuous limit function, such a product coincides with a bounded Borel measure. Similarly and obviously, the ϵ-weighted non conservative product $\epsilon(\partial_x u^\epsilon)^2$ does not vanish generally with ϵ but rather converges to a non negative bounded measure, say $\mathcal{N}$. We are thus led to consider the following system for governing the limit solution $\mathbf{V}$:

$$\begin{cases} \partial_t\rho + \partial_x\rho u = 0, \\ \partial_t\rho u + \partial_x(\rho u^2 + \sum_{i=1}^N p_i) = 0, \\ \partial_t\rho\varepsilon_i + \partial_x\rho\varepsilon_i u + [p_i\partial_x u]_\phi = \mu_i\mathcal{N}, \quad i = 1, ..., N. \end{cases} \tag{4.26}$$

In order to define both the limits $[p_i\partial_x u]_\phi$ and $\mathcal{N}$, we adopt the framework developed by LeFloch [90], Dal Maso-LeFloch-Murat [42]. Following these authors, the non conservative products $[p_i\partial_x u]_\phi$ for $i \in \{1, ..., N\}$ coincide with the classical products $p_i\partial_x u$ in the zones of smoothness of $\mathbf{V}$. Then across a discontinuity propagating with speed σ and separating two states $\mathbf{V}^-$ and $\mathbf{V}^+$, these products are defined as the following Borel measures:

$$[p_i\partial_x u]_\phi = \int_0^1 p_i(\phi(s; \mathbf{V}^-, \mathbf{V}^+))\partial_s u(\phi(s; \mathbf{V}^-, \mathbf{V}^+))ds\, \delta_{x-\sigma t}, \quad 1 \le i \le N, \tag{4.27}$$

where ϕ denotes a fixed family of paths to be prescribed. Such a family is subject to some consistency and smoothness conditions detailed in [42]. Equipped with the fixed family of paths entering (4.27), we next suggest to define $\mathcal{N}$ as a bounded non negative Borel measure concentrated on the discontinuities of $\mathbf{V}$ and given by:

$$\mathcal{N} = \int_0^1 |\partial_s u(\phi(s; \mathbf{V}^-, \mathbf{V}^+))|^2 ds\, \delta_{x-\sigma t}. \tag{4.28}$$

With clear notations, we shall use in the sequel the following condensed form:

$$\partial_t\mathbf{V} + [\mathbf{A}(\mathbf{V})\partial_x\mathbf{V}]_\phi = \mathbf{N}(\mathbf{V}), \tag{4.29}$$

with

$$\mathbf{N}(\mathbf{V}) = (\{N_i(\mathbf{V})\}_{i=1,...,N}), \quad N_i(\mathbf{V}) = \mu_i \mathcal{N}(\mathbf{V}), \quad i = 1, ..., N. \tag{4.30}$$

Next turning considering the formulation (4.25), this one extends to the present framework of piecewise Lipschitz continuous solutions when considering for a fixed family of paths, say ψ, the following formulation :

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + \sum_{i=1}^N p_i) = 0, \\ \partial_t \rho E + \partial_x (\rho E + \sum_{i=1}^N p_i) u = 0, \\ \partial_t \rho w_i + \partial_x \rho w_i u + [(p_i - \frac{\mu_i}{\mu_N} p_N) \partial_x u]_\psi = 0, \quad i = 1, ..., N-1, \end{cases} \tag{4.31}$$

or again the following condensed form :

$$\partial_t \mathbf{W} + [\mathbf{B}(\mathbf{W}) \partial_x \mathbf{W}]_\psi = 0, \tag{4.32}$$

supplemented by the Lax entropy inequalities :

$$\partial_t \rho s_i + \partial_x \rho s_i u \leq 0, \quad i = 1, ..., N. \tag{4.33}$$

Our purpose is to prove that both extended forms (4.29) and (4.32) still continue to be equivalent for the discontinuous solutions under consideration provided that both family of paths ϕ and ψ are characterized by the study of their Travelling Waves (TW) solutions. Indeed and after LeFloch [90] (see also Raviart-Sainsaulieu [112] and below) the required family of paths can be defined so as to coincide with the graph of the TW (up to some parametrization).
To assess the relevance of solely invoking TW solutions, it is worth noticing that the velocity u can achieve non trivial jumps only across the discontinuities associated with the first and last fields (see indeed lemma 15). In other words, only these two fields which give rise to shock solutions (versus contact solutions for the intermediate ones) may induce non trivial measures in (4.27) and (4.28). As a consequence, these shock solutions can be defined as the limit as ϵ goes to zero of associated TW solutions of (4.2) (see [85], [112]). The benefit of making use of TW solutions directly comes from the equivalence lemma 16 valid for smooth solutions when recalling (see below) that TW solutions are nothing but smooth solutions of (4.2). In other words and thanks to lemma 16, TW solutions of (4.2) and (4.25) are just identical for smooth solutions. We thus immediately infer :

Lemma 17 *Assuming that the two families of paths ϕ and ψ, respectively in (4.29) and (4.32), are defined according to the graph of TW solutions of (4.2) (or equivalently (4.25)), then both formulations admit the same shock solutions and are therefore equivalent.*

From now on, the limit systems (4.29) and (4.32) will be always dealt with when closed according to the above family of paths.
For the sake of completeness, we briefly recall here the definition of shock solutions for (4.29) (and hence (4.32)) from TW solutions. A Travelling Wave is a smooth solution of (4.2) of the special form $\mathbf{V}_\epsilon(x,t) = \mathcal{V}_\epsilon(x - \sigma t) = \mathcal{V}_\epsilon(\xi)$ which satisfies $\lim_{\xi \to \pm\infty} \mathcal{V}_\epsilon(\xi) =$

$\mathbf{V}^{\pm}$. Next, introducing the function $\mathcal{V} : I\!R \to I\!R$ defined by $\mathcal{V}(\frac{\xi}{\epsilon}) = \mathcal{V}_\epsilon(\xi)$, then $\mathcal{V}$ is solution of the following ODE problem, free from the rescaling parameter ϵ:

$$\begin{cases} -\sigma d_\xi \rho + d_\xi \rho u = 0, \\ -\sigma d_\xi \rho u + d_\xi (\rho u^2 + \sum_{i=1}^N p_i) = d_\xi (\sum_{i=1}^N \mu_i d_\xi u), \\ -\sigma d_\xi \rho \varepsilon_i + d_\xi \rho \varepsilon_i u + p_i d_\xi u = \mu_i (d_\xi u)^2,\ i = 1, ..., N, \end{cases} \tag{4.34}$$

with $\lim_{\xi \to \pm\infty} \mathcal{V}(\xi) = \mathbf{V}^{\pm}$. Existence and uniqueness (up to some translation) of solution to the above problem have been proved in [19] under the classical condition that the relative Mach number is larger than 1. Due to the non conservation form met by these equations, let us emphasize that the critical point $\mathbf{V}^+$ not only depends on σ and $\mathbf{V}^-$ but also on the ratios of the viscosities $\{\mu_i / \sum_{i=1}^N \mu_i\}_{i=1,...,N}$.
Equipped with these smooth solutions, we then observe that for all $\epsilon > 0$:

$$||d_\xi \mathcal{V}_\epsilon||_{L^1(I\!R)} = ||d_\xi \mathcal{V}||_{L^1(I\!R)} < \infty,$$

so that the family $(\mathcal{V}_\epsilon)_{\epsilon > 0}$ converges in the L^1_{loc} topology to the following discontinuous function:

$$\mathbf{V}(x,t) = \begin{cases} \mathbf{V}^- & \text{if} \quad x < \sigma t \\ \mathbf{V}^+ & \text{if} \quad x > \sigma t, \end{cases} \tag{4.35}$$

which by definition is said to be a shock solution of (4.29), after LeFloch [90] (see also [112]). We are now in position to exhibit the jump conditions satisfied by the shock solution (4.35) under consideration. In that aim, let us deduce from (4.34) after integration along the profile:

$$\begin{array}{l} m := \rho^-(u^- - \sigma) = \rho^+(u^+ - \sigma), \\ m(u^+ - u^-) + (\sum_{i=1}^N p_i^+ - \sum_{i=1}^N p_i^-) = 0, \\ m(\varepsilon_i^+ - \varepsilon_i^-) + \int_{I\!R} p_i d_\xi u d\xi = \mu_i \int_{I\!R} (d_\xi u)^2 d\xi. \end{array} \tag{4.36}$$

Next and to recover the required relation (4.27)-(4.28), it suffices to define the family of paths ϕ when setting

$$\phi(s; \mathbf{V}^-, \mathbf{V}^+) = \mathcal{V}(\Phi(s)), \tag{4.37}$$

where Φ is an increasing smooth bijection from $]0,1[$ onto $I\!R$. The same steps exactly apply for defining the shock solutions of (4.32) and again since TW solutions do coincide, both family of shock solutions are identical. Besides this equivalence result, let us conclude this section with the following easy result which proof is left to the reader:

Corollary 7 *Let us consider a weak solution* $\mathbf{V}$ *of (4.32) (i.e. of (4.29)). In the zone of smoothness of* $\mathbf{V}$*, one has:*

$$\partial_t \rho \varepsilon_i + \partial_x \rho \varepsilon_i u + p_i \partial_x u = 0, \quad i = 1, ..., N, \tag{4.38}$$

i.e.

$$\partial_t \rho s_i + \partial_x \rho s_i u = 0, \quad i = 1, ..., N, \tag{4.39}$$

while shock solutions obey:

$$m(\varepsilon_i^+ - \varepsilon_i^-) + \int_0^1 p_i(\phi(s; \mathbf{V}^-, \mathbf{V}^+)) \partial_s u(\phi(s; \mathbf{V}^-, \mathbf{V}^+)) ds =$$

$$\frac{\mu_i}{\mu_N} \Big\{ m(\varepsilon_N^+ - \varepsilon_N^-) + \int_0^1 p_N(\phi(s; \mathbf{V}^-, \mathbf{V}^+)) \partial_s u(\phi(s; \mathbf{V}^-, \mathbf{V}^+)) ds \Big\}, \tag{4.40}$$

$$i = 1, ..., N-1,$$

with

$$m(s_i^+ - s_i^-) < 0, \quad i = 1, ..., N. \tag{4.41}$$

The identities (4.38) and (4.40) are clearly nothing but the extensions of (4.15) and the corresponding one in (4.16) respectively, when expressed in the general setting in terms of internal energies instead of entropies. The Rankine-Hugoniot like conditions (4.40) will be referred in the sequel as to generalized jump conditions.

4.4 Exact Riemann solver and the generalized jump conditions

We now address the numerical approximation of the weak solutions of (4.32)–(4.33). The key issue turns out to enforce at the discrete level each of the N entropy rates of dissipation to be kept in balance according to the ratios of the viscosities while ensuring the discrete conservation of the density, the momentum and the total energy. The system (4.32) has been precisely introduced to fulfill this requirement. We shall prove below, on the basis of the Godunov method, that the equivalence with the original system (4.29) is lost at the discrete level. Approximating piecewise Lipschitz continuous solutions via (4.29) actually results in large errors between discrete and exact solutions. The roots of such a failure will be clearly identified with the numerical dissipation in the method which produces entropy dissipations at the wrong rates. This observation will naturally lead us to enforce the required balance in the entropies dissipation rates at the discrete level. Doing so, we shall end up with a finite volume approximation of the equivalent system (4.32), which is fully explicit, satisfies several stability properties under a classical CFL condition and produces approximate solutions in close agreement with the exact ones.
The derivation of the method and its properties is performed here on the basis of the Godunov method, and it will find a fairly natural extension in the next section devoted to Approximate Riemann Solvers. Let us first briefly report here some useful properties of the exact Riemann solution for (4.29) that are in order to prove the forthcoming statements.
Existence (away from vacuum) of an unique solution for the Riemann problem for (4.29) has been proved in [15] (see also [19]). As strongly suggested by lemma 15, the Riemann solution is made of at most three simple waves (under the Weyl's assumptions) separating four constant states: namely $\mathbf{V}_L$, $\mathbf{V}_L^\star$, $\mathbf{V}_R^\star$ and $\mathbf{V}_R$ where with classical notations $\mathbf{V}_L$ and $\mathbf{V}_R$ denote the two constant states in the Riemann initial data. The Riemann solution, we denote $\mathbf{w}(x/t, \mathbf{V}_L, \mathbf{V}_R)$ in the sequel, is shown below to obey maximum principles on each of the specific entropies. This result extends a property proved by Tadmor in the classical Euler setting (*i.e.* $N = 1$) and will play an important role hereafter.

Lemma 18 *The self similar solution* $\mathbf{w}(x/t, \mathbf{V}_L, \mathbf{V}_R)$ *of the Cauchy problem (4.29) with initial data* $\mathbf{V}_0(x) = \mathbf{V}_L$ *if* $x < 0$ *and* $\mathbf{V}_R$ *otherwise is such that:*

$$s_i(., \mathbf{V}_L, \mathbf{V}_R) \leq \max(s_{iL}, s_{iR}), \quad i = 1, ..., N.$$

Proof
Using standard arguments (see [57]) it can be easily proved that each of the specific entropies s_i are Riemann invariants for the first and third genuine non linear fields and jump freely across the contact discontinuity. Consequently, s_i can achieve distinct values from s_{iL} and s_{iR} only across shock solutions associated with these two fields and therefore at most two distinct values: namely $s_{iL}^\star$ and $s_{iR}^\star$. But across 1-shock solutions, the relative mass flux m in (4.41) is positive (see [57] in the classical Euler setting) so that $s_{iL}^\star < s_{iL}$, while across 3-shock solutions, m is negative so that $s_{iR}^\star < s_{iR}$. This concludes the proof.

Let us now address the numerical approximation of the weak solutions of (4.32)–(4.33). For simplicity in the notations, we restrict ourselves to uniform cartesian discretization of $I\!R_x \times I\!R_t^+$ defined by a constant time step Δt and a constant space step Δx. Introducing $x_{j+1/2} = (j+1/2)\Delta x$ for $j \in Z$ and $t^n = n\Delta t$ for $n \in N$, the cartesian grids under consideration then read $\cup_{j\in Z, n\in N}\mathcal{C}_j^n$ where $\mathcal{C}_j^n := [x_{j-1/2}, x_{j+1/2}) \times [t^n, t^{n+1})$. The approximate solution of the Cauchy problem (4.29) with $\mathbf{V}_0$ as initial data, we denote $\mathbf{V}_\lambda(x,t)$ in the sequel with $\lambda = \Delta t/\Delta x$, is classically sought as a piecewise constant function at each time level t^n,

$$\mathbf{V}_\lambda(x,t^n) := \mathbf{V}_j^n, \quad \text{for all } x \in [x_{j-1/2}, x_{j+1/2}), \quad n > 0, \quad j \in Z, \tag{4.42}$$

with when $n = 0$:

$$\mathbf{V}_j^0 = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{V}_0(x)dx, \quad \text{for all } j \in Z. \tag{4.43}$$

Assuming that the approximate solution $\mathbf{V}_\lambda(x,t^n)$ is known at a given time $t^n \geq 0$, this one is then defined for $t \in [t^n, t^{n+1})$ as the solution of the Cauchy problem (4.29) with $\mathbf{V}_\lambda(x,t^n)$ as initial data. Provided that Δt is chosen small enough, *i.e.* under the CFL restriction:

$$\lambda \max_{\mathbf{V}} \rho(\mathbf{A}(\mathbf{V})) \leq \frac{1}{2}, \tag{4.44}$$

where the maximum is taken over all the $\mathbf{V}$ under consideration, $\mathbf{V}_\lambda(x,t)$ with $t \in [t^n, t^{n+1})$ is nothing but the juxtaposition of a sequence of non interacting adjacent Riemann solutions $\mathbf{w}((x - x_{j+1/2})/(t-t^n); \mathbf{V}_j^n, \mathbf{V}_{j+1}^n)$, centered at each cell interface $x_{j+1/2}$. Let us then consider the L^2-projection of this solution at time t^{n+1-} onto piecewise constant functions:

$$\mathbf{V}_j^{n+1-} = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{V}_\lambda(x, t^{n+1-})dx, \quad j \in Z. \tag{4.45}$$

Notice that the averages (4.45) can be given the convenient equivalent form:

$$\begin{cases} \rho_j^{n+1-} = \rho_j^n - \lambda\Delta\{\rho u\}_{j+1/2}^n, \\ (\rho u)_j^{n+1-} = (\rho u)_j^n - \lambda\Delta\{\rho u^2 + \sum_{i=N}^N p_i\}_{j+1/2}^n, \end{cases} \tag{4.46}$$

while the internal energies $\rho\varepsilon_i$ for $i \in \{1, \ldots, N\}$ update according to:

$$(\rho\varepsilon_i)_j^{n+1-} = (\rho\varepsilon_i)_j^n - \lambda\Delta\{\rho\varepsilon_j u\}_{j+1/2}^n + \lambda < (N_i - [p_i\partial_x u]_\phi)(\mathbf{V}_\lambda(x,t)), \mathcal{C}_j^n > . \tag{4.47}$$

All the numerical fluxes entering the formulae (4.46)–(4.47), say generically $X^n_{j+1/2}$ at the cell interface $x_{j+1/2}$, classically find the definition $X^n_{j+1/2} = \{X\}(\mathbf{w}(0^+; \mathbf{V}^n_j, \mathbf{V}^n_{j+1}))$ under the CFL restriction (4.44) while with clear notations, one has:

$$< (N_i - [p_i\partial_x u]_\phi)(\mathbf{V}_\lambda(x,t)), \mathcal{C}^n_j >=$$
$$-\int\int_{\mathcal{C}^n_j} \{p_i\partial_x u\}(\mathbf{V}_\lambda(x,t))dxdt + \sum_{shocks\, \in\, \mathcal{C}^n_j} \rho^-(u^- - \sigma)(\varepsilon_i^+ - \varepsilon_i^-).$$

Here the first contribution comes from the fact that the Borel measure $\mathcal{N}_i$ in (4.30) identically vanishes everywhere except at shock locations while the second contribution clearly results from the jump conditions (4.36) across propagating shocks in $\mathcal{C}^n_j$. According to the usual Godunov's procedure, one would update the approximate solution at time t^{n+1} when defining $\mathbf{V}_\lambda(x, t^{n+1}) = \mathbf{V}^{n+1-}_j$ for all $x \in [x_{j-1/2}, x_{j+1/2})$ and $j \in Z$. However, such an updating formula would prevent the L^1 norm of the total energy to be preserved with time because of the next statement:

Proposition 21 *Under the CFL condition (4.44), the following inequality holds for all $j \in Z$:*

$$\{\rho E\}(\mathbf{V}^{n+1-}_j) \leq \{\rho E\}(\mathbf{V}^n_j) - \lambda\Delta\{\rho H u\}(\mathbf{w}(0^+; \mathbf{V}^n_j, \mathbf{V}^n_{j+1})), \tag{4.48}$$

such an inequality being strict generally speaking.

The proof of this result is postponed at the end of the present section. Rephrazing the inequality (4.48), the formulae (4.46)–(4.47) make the L^1-norm of the total energy to strictly decrease with time and as a consequence, cannot provide us with a relevant numerical method for approximating the discontinuous solutions under consideration. Numerical evidences of such a failure, given in section 4.6, clearly indicate that the approximate solutions derived from (4.46)–(4.47) grossly disagree with the exact ones. One would be therefore tempted to enforce for validity the conservation of the discrete total energy when defining for all j in Z:

$$\begin{aligned}(\rho E)^{n+1-}_j &:= \tfrac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{\rho E\}(\mathbf{V}_\lambda(x, t^{n+1-}))dx \\ &= (\rho E)^n_j - \lambda\Delta\{\rho H u\}(\mathbf{w}(0^+; \mathbf{V}^n_j, \mathbf{V}^n_{j+1})),\end{aligned} \tag{4.49}$$

while keeping unchanged the definition of ρ^{n+1-}_j, $(\rho u)^{n+1-}_j$ and up to some renumbering, the first $(N-1)$ internal energies $(\rho\varepsilon_i)^{n+1-}_j$. Here, the last internal energy would thus be redefined according to:

$$\{\rho\varepsilon_N\}(\overline{\mathbf{V}}^{n+1-}_j) := (\rho E)^{n+1-}_j - \frac{((\rho u)^{n+1-}_j)^2}{2\rho^{n+1-}_j} - \sum_{i=1}^{N-1} (\rho\varepsilon_i)^{n+1-}_j, \tag{4.50}$$

where we have set:

$$\overline{\mathbf{V}}^{n+1-}_j =^T (\rho^{n+1-}_j,\ (\rho u)^{n+1-}_j,\ (\rho E)^{n+1-}_j,\ \{(\rho\varepsilon_i)^{n+1-}_j\}_{1\leq i\leq N-1}). \tag{4.51}$$

Doing so, one would then end up with a failure of (4.22) and then of the generalized jump conditions (4.40) at the discrete level as expressed in the next statement (which proof is also postponed at the end of this section) :

Proposition 22 *Under the CFL condition (4.44), the following inequalities hold true for all $i = 1, ..., N$ and all $j \in Z$:*

$$\mu_N\{(\rho\varepsilon_i)_j^{n+1-} - (\rho\varepsilon_i)_j^n + \lambda\Delta\{\rho\varepsilon_i u\}_{j+1/2}^n + \lambda < [p_i\partial_x u]_\phi(\mathbf{V}_\lambda(x,t)), \mathcal{C}_j^n > \}$$
$$\leq$$
$$\mu_i\{\{\rho\varepsilon_N\}(\overline{\mathbf{V}}_j^{n+1-}) - (\rho\varepsilon_N)_j^n + \lambda\Delta\{\rho\epsilon_N u\}_{j+1/2}^n + \lambda < [p_N\partial_x u]_\phi(\mathbf{V}_\lambda(x,t)), \mathcal{C}_j^n > \}, \tag{4.52}$$

these inequalities being strict generally speaking.

In order to restore the conservation of the total energy at each time level while preserving the validity of the generalized jump conditions, we are therefore led to keep unchanged the updated values of both the density and momentum :

$$\rho_j^{n+1} = \rho_j^{n+1-}, \quad (\rho u)_j^{n+1} = (\rho u)_j^{n+1-}, \quad \text{for all } j \in Z, \tag{4.53}$$

while defining the N internal energies $(\rho\varepsilon_i)_j^{n+1}$ as the solutions of the following $(N-1)$ relations :

$$\mu_N\{(\rho\varepsilon_i)_j^{n+1} - (\rho\varepsilon_i)_j^n + \lambda\Delta\{\rho\epsilon_j u\}_{j+1/2}^n + \lambda < [p_i\partial_x u]_\phi(\mathbf{V}_\lambda(x,t)), \mathcal{C}_j^n > \}$$
$$=$$
$$\mu_i\{(\rho\varepsilon_N)_j^{n+1} - (\rho\varepsilon_N)_j^n + \lambda\Delta\{\rho\epsilon_N u\}_{j+1/2}^n + \lambda < [p_N\partial_x u]_\phi(\mathbf{V}_\lambda(x,t)), \mathcal{C}_j^n > \},$$

supplemented with :

$$\{\rho E\}_j^{n+1} := \{\rho E\}_j^n - \lambda\Delta\{\rho H u\}_{j+1/2}^n = \frac{((\rho u)_j^{n+1})^2}{2\rho_j^{n+1}} + \sum_{i=1}^N (\rho\varepsilon_i)_j^{n+1}. \tag{4.54}$$

Let us underline that the above $(N-1)$ identities are nothing but the discrete version of (4.22). Let us further notice that (4.30) readily implies for each $j \in Z$:

$$\mu_N < N_i(\mathbf{V}_\lambda(x,t)), \mathcal{C}_j^n >= \mu_i < N_N(\mathbf{V}_\lambda(x,t)), \mathcal{C}_j^n >, \tag{4.55}$$

so that the $\{(\rho\varepsilon_i)_j^{n+1}\}_{1\leq i\leq N}$ equivalently solve :

$$\begin{cases} \mu_N\{(\rho\varepsilon_i)_j^{n+1} - (\rho\varepsilon_i)_j^{n+1-}\} = \mu_i\{(\rho\varepsilon_N)_j^{n+1} - (\rho\varepsilon_N)_j^{n+1-}\}, \ 1 \leq i \leq N-1, \\ \displaystyle\sum_{i=1}^N (\rho\varepsilon_i)_j^{n+1} = (\rho E)_j^{n+1} - \frac{((\rho u)_j^{n+1})^2}{2\rho_j^{n+1}}. \end{cases} \tag{4.56}$$

As an immediate consequence, we have :

Proposition 23 *The linear system (4.56) in the unknown $\{(\rho\varepsilon_i)_j^{n+1}\}_{1\leq i\leq N}$ admits an unique solution explicitly given by :*

$$(\rho\varepsilon_i)_j^{n+1} = (\rho\varepsilon_i)_j^{n+1-} + \frac{\mu_i}{\sum_{l=1}^N \mu_l}((\rho E)_j^{n+1} - \frac{((\rho u)_j^{n+1})^2}{2\rho_j^{n+1}} - \sum_{l=1}^N (\rho\varepsilon_l)_j^{n+1-}), \tag{4.57}$$

for all $i = 1, \ldots, N$.

The last result of this section proves the correct design of the scheme we propose, when stating its stability properties.

Theorem 9 *Under the CFL condition (4.44), the numerical method (4.45)-(4.53)-(4.57) preserves the positivity of the internal energies, that is to say* $(\rho\varepsilon_i)_j^n \geq 0$ *for* $i = 1, \dots, N$ *and all* $j \in Z$*. In addition, it obeys for* $i = 1, ..., N$ *the following entropy inequalities :*

$$(\rho s_i)_j^{n+1} := \{\rho s_i\}(\rho_j^{n+1}, (\rho\varepsilon_i)_j^{n+1}) \leq (\rho s_i)_j^n - \lambda\Delta\{\rho s_i u\}(\mathbf{w}(0^+; \mathbf{V}_j^n, \mathbf{V}_{j+1}^n)), \quad (4.58)$$

and achieves the next maximum principles on the specific entropies :

$$(s_i)_j^{n+1} \leq \max((s_i)_{j-1}^n, (s_i)_j^n, (s_i)_{j+1}^n).$$

We conclude this section with the proofs of the results stated in this section. Let us first establish the proposition 21.

Proof of proposition 21

Arguing about the convexity of the mapping $(\rho, \rho u) \mapsto \frac{(\rho u)^2}{2\rho}$ and using the definition of the averages (4.45) at time t^{n+1-}, the Jensen's inequality immediately implies

$$\begin{gathered}\{\rho E\}(\mathbf{V}_j^{n+1-}) := \frac{((\rho u)_j^{n+1-})^2}{2\rho_j^{n+1-}} + \sum_{i=1}^N (\rho\varepsilon_i)_j^{n+1-} \leq \\ \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{\frac{(\rho u)^2}{2\rho}\}(\mathbf{V}_\lambda(x, t^{n+1-}))dx + \sum_{i=1}^N (\rho\varepsilon_i)_j^{n+1-} = \\ \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{\rho E\}(\mathbf{V}_\lambda(x, t^{n+1-}))dx.\end{gathered} \quad (4.59)$$

Next, weak solutions of the system (4.29) are known to obey the following conservation law for the total energy :

$$\partial_t \rho E + \partial_x(\rho E + \sum_{i=1}^N p_i)u = 0, \quad \text{in } \mathcal{D}'. \quad (4.60)$$

Classical arguments then apply to get the desired identity :

$$\frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{\rho E\}(\mathbf{V}_\lambda(x, t^{n+1-}))dx = (\rho E)_j^n - \lambda\Delta\{\rho H u\}_{j+1/2}^n. \quad (4.61)$$

Let us now turn proving the next proposition.

Proof of proposition 22

First and by construction, the definition of the first $(N-1)$ internal energies have been kept unchanged, *i.e.* :

$$(\rho\varepsilon_i)_j^{n+1-} = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} (\rho\varepsilon_i)(\mathbf{V}_\lambda(x, t^{n+1-}))dx, \quad 1 \leq i \leq N-1, \quad (4.62)$$

so that (4.47) apply, that is to say for $i \in \{1, \ldots, N-1\}$:

$$(\rho\varepsilon_i)_j^{n+1-} - (\rho\varepsilon_i)_j^n + \lambda\Delta\{\rho\varepsilon_i u\}_{j+1/2}^n + \lambda < [p_i \partial_x u]_\phi(\mathbf{V}_\lambda(x,t)), \mathcal{C}_j^n >= \\ \lambda < N_i(\mathbf{V}_\lambda(x,t)), \mathcal{C}_j^n > . \tag{4.63}$$

Let us now check that the last internal energy obeys the following inequality:

$$\{\rho\varepsilon_N\}(\overline{\mathbf{V}}_j^{n+1-}) \geq \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} (\rho\varepsilon_N)(\mathbf{V}_\lambda(x, t^{n+1-}))dx. \tag{4.64}$$

To that purpose, it suffices to notice that the total energy inequality (4.48) established in proposition 21 just recasts:

$$\tfrac{((\rho u)_j^{n+1-})^2}{2\rho_j^{n+1-}} + \sum_{i=1}^{N-1} (\rho\varepsilon_i)_j^{n+1-} + \tfrac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{\rho\varepsilon_N\}(\mathbf{V}_\lambda(x, t^{n+1-}))dx \leq (\rho E)_j^{n+1-}, \tag{4.65}$$

that is from the definition (4.50):

$$\{\rho\varepsilon_N\}(\overline{\mathbf{V}}_j^{n+1-}) := (\rho E)_j^{n+1-} - \tfrac{((\rho u)_j^{n+1-})^2}{2\rho_j^{n+1-}} - \sum_{i=1}^{N-1} (\rho\varepsilon_i)_j^{n+1-} \geq \\ \tfrac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} (\rho\varepsilon_N)(\mathbf{V}_\lambda(x, t^{n+1-}))dx.$$

This gives the required inequality (4.64) which exactly rewrites, when considering the formula (4.47) for $i = N$:

$$\{\rho\varepsilon_N\}(\overline{\mathbf{V}}_j^{n+1-}) - (\rho\varepsilon_N)_j^n + \lambda\Delta\{\rho\varepsilon_N u\}_{j+1/2}^n + \lambda < [p_N \partial_x u]_\phi(\mathbf{V}_\lambda(x,t)), \mathcal{C}_j^n > \\ \geq \lambda < N_N(\mathbf{V}_\lambda(x,t)), \mathcal{C}_j^n > . \tag{4.66}$$

Let us notice that generally speaking, such an inequality is strict. The positivity of the viscosity coefficients μ_i then easily yields the conclusion from (4.63) and (4.66), in view of the identities (4.55).

We next establish the theorem 9.

Proof of theorem 9
Since by construction, the total energy $(\rho E)_j^{n+1}$ finds the conservative discrete formulation (4.54), we immediately infer from the energy inequality proved in proposition 21:

$$\{\rho E\}(\mathbf{V}_j^{n+1-}) = \frac{((\rho u)_j^{n+1-})^2}{2\rho_j^{n+1-}} + \sum_{i=1}^{N} (\rho\varepsilon_i)_j^{n+1-} \leq (\rho E)_j^{n+1}. \tag{4.67}$$

As a consequence, invoking the explicit formulae (4.57) for defining each of the internal energies $(\rho\varepsilon_i)_j^{n+1}$, the positivity of the viscosity coefficients easily ensures that:

$$(\rho\varepsilon_i)_j^{n+1} \geq (\rho\varepsilon_i)_j^{n+1-} \quad \text{for all } i \in \{1, \ldots, N\}. \tag{4.68}$$

But, $(\rho\varepsilon_i)_j^{n+1-}$ stays by construction positive. Indeed, this quantity is nothing but the average (see indeed (4.45)) of the exact internal energy $(\rho\varepsilon_i)$ evaluated on Riemann solutions under the CFL restriction (4.44). Such solutions stay within the admissible phase space and this gives the required conclusion.
Turning now establishing the entropy inequalities (4.58), the weak solutions of (4.29) obey in view of the previous section :

$$\partial_t \rho s_i + \partial_x \rho s_i u \leq 0, \quad \mathcal{D}', \tag{4.69}$$

so that using classical arguments, we arrive at :

$$\frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{\rho s_i\}(\mathbf{V}_\lambda(x,t^{n+1-}))dx \leq (\rho s_i)_j^n - \lambda\Delta\{\rho s_i u\}(\mathbf{w}(0^+;\mathbf{V}_j^n,\mathbf{V}_{j+1}^n)).$$

But, by invoking the convexity of the mapping $(\rho,\rho\varepsilon_i)\mapsto\{\rho s_i\}(\rho,\rho\varepsilon_i)$, we get from the Jensen's inequality :

$$\{\rho s_i\}(\rho_j^{n+1-},(\rho\varepsilon_i)_j^{n+1-}) \leq \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{\rho s_i\}(\mathbf{V}_\lambda(x,t^{n+1-}))dx.$$

Now it is sufficient to notice that :

$$\frac{\partial \rho s_i}{\partial \rho\varepsilon_i} = -\frac{1}{T_i} < 0,$$

to get the required conclusion from the inequality (4.68).
To prove the validity of the maximum principles on the specific entropies, let us first recall from the beginning of this section, that the Riemann solution at each cell interface $x_{j+1/2}$ satisfies :

$$s_i(\frac{x-x_{j+1/2}}{t-t^n},\mathbf{V}_j^n,\mathbf{V}_{j+1}^n) \leq \max((s_i)_j^n,\ (s_i)_{j+1}^n). \tag{4.70}$$

Equipped with these inequalities valid for all $j\in Z$ and all $i=1,...,N$, we deduce from the latter considerations that under the CFL condition (4.44) :

$$\begin{aligned} s_i(\rho_j^{n+1},(\rho\varepsilon_i)_j^{n+1}) := \tfrac{(\rho s_i)_j^{n+1}}{\rho_j^{n+1}} \leq \tfrac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \tfrac{\{\rho s_i\}(\mathbf{V}_\lambda(x,t^{n+1-}))}{\rho_j^{n+1}}dx \leq \\ \max((s_i)_j^{n+1},\ (s_i)_j^n,\ (s_i)_{j+1}^n)\tfrac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \tfrac{\{\rho\}(\mathbf{V}_\lambda(x,t^{n+1-}))}{\rho_j^{n+1}}dx, \end{aligned} \tag{4.71}$$

arguing about the positivity of the density in each of the exact Riemann solutions. This is nothing but the required conclusion by definition of ρ_j^{n+1}.

4.5 Approximate Riemann Solvers and the generalized jump conditions

This section aims at generalizing the method described in the previous section when considering Approximate Riemann Solvers. The basic motivation here stems from the lack of a precise knowledge of the required kinetic relations for general pressure and viscosity laws. The extension we propose primarily intends to take advantage

of the celebrated framework proposed by Harten-Lax-Van Leer [61]. This framework have indeed received a considerable attention over the past two decades and several valuable methods, ranging from kinetic solvers [110], [57] to Approximate Riemann Solvers with a large variety of intermediate states [113], [61], [56] enter in (see also [19] and the references therein). Our purpose is to directly inherit from all these suitable methods. We show here that all of them may naturally apply to approximate the solutions of (4.2), despites that they were primarily developed in the setting of systems of (pure) conservation laws. Indeed, these methods are seen in this section to naturally and easily extend to the present setting when enforcing for validity at the discrete level the equalities (4.22) and then the generalized jump conditions (4.40) while preserving each of their respective stability and precision properties.
We first give a rough description of the extension and then define suitable Approximate Riemann Solvers. With the notations introduced in the first section, let us consider for some given rescaling parameter $\epsilon > 0$ the following viscosity form of the system under study :

$$\partial_t \mathbf{U}^\epsilon + \partial_x \mathbf{F}(\mathbf{U}^\epsilon) = \epsilon \mathcal{R}(\mathbf{U}^\epsilon, \partial_x(\mathcal{C}(\mathbf{U}^\epsilon)\partial_x \mathbf{U}^\epsilon)), \tag{4.72}$$

so that the smooth solutions of (4.72) obey the following additional conservation law :

$$\partial_t\{\rho E\}(\mathbf{U}^\epsilon) + \partial_x(\rho E^\epsilon + \sum_{i=1}^N p_i^\epsilon)u^\epsilon = \epsilon\partial_x(\sum_{i=1}^N \mu_i u^\epsilon \partial_x u^\epsilon). \tag{4.73}$$

Let be given $\mathbf{U}_h(x,t^n)$ some approximate solution of (4.72) at time t^n. We propose to evolve it to the next time level t^{n+1} into two steps that are classically based on a splitting of operators.

Step 1 The first order (convective) operator, extracted from (4.72), is solved for small times $t \in [0, \Delta t]$ with $\mathbf{U}_h(x,t^n)$ as initial data, that is :

$$\left\{ \begin{array}{l} \partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = 0, \\ \mathbf{U}(x,0) = \mathbf{U}_h(x,t^n). \end{array} \right. \tag{4.74}$$

Note that for uniqueness of weak solutions, the above Cauchy problem must be supplemented with an entropy selection principle we express here thanks to the total energy (indeed strictly convex in the variable $\mathbf{U}$) :

$$\partial_t\{\rho E\}(\mathbf{U}) + \partial_x(\rho E + \sum_{i=1}^N p_i)u \leq 0. \tag{4.75}$$

Obviously, solving the Cauchy problem (4.74)-(4.75) will be precisely the matter of classical Approximate Riemann Solvers.

Step 2 Let $\mathbf{U}_h(x,t^{n+1-})$ denote the solution of (4.74)-(4.75) at time Δt. This one now naturally serves as initial data to solve the second order (viscosity) operator in (4.72) for times $t \in [0, \Delta t]$:

$$\left\{ \begin{array}{l} \partial_t \mathbf{U}^\epsilon = \epsilon \mathcal{R}(\mathbf{U}^\epsilon, \partial_x(\mathcal{C}(\mathbf{U}^\epsilon)\partial_x \mathbf{U}^\epsilon)), \\ \mathbf{U}^\epsilon(x,0) = \mathbf{U}_h(x,t^{n+1-}). \end{array} \right. \tag{4.76}$$

Note that the total energy evolves according to

$$\partial_t \{\rho E\}(\mathbf{U}^\epsilon) = \epsilon \partial_x (\sum_{i=1}^{N} \mu_i u^\epsilon \partial_x u^\epsilon).$$

Let us now recast these $(N+2)$ equations of (4.76) in the $\mathbf{V}$ variable:

$$\begin{cases} \partial_t \rho^\epsilon = 0, \\ \partial_t \rho u^\epsilon = \epsilon \partial_x (\sum_i \mu_i \partial_x u^\epsilon), \\ \partial_t \rho \varepsilon_i^\epsilon = \epsilon \mu_i (\partial_x u^\epsilon)^2, \quad \text{for } i = 1, \dots, N, \end{cases} \tag{4.77}$$

so that for any given $\epsilon > 0$, we infer the following $(N+1)$ equations:

$$\begin{cases} \partial_t \rho^\epsilon = 0, \\ \partial_t \rho u^\epsilon = \epsilon \partial_x (\sum_i \mu_i \partial_x u^\epsilon), \\ \mu_N \partial_t \rho \varepsilon_i^\epsilon = \mu_i \partial_t \rho \varepsilon_N^\epsilon, \quad \text{for } i = 1, \dots, N-1, \end{cases} \tag{4.78}$$

we must supplement with

$$\partial_t \{\rho E\}(\rho^\epsilon, \rho u^\epsilon, \{\rho \varepsilon_i^\epsilon\}_{1 \le i \le N}) = \epsilon \partial_x (\sum_i \mu_i u^\epsilon \partial_x u^\epsilon). \tag{4.79}$$

Since we do not want to resolve the small scales with ϵ, we consider (4.78)-(4.79) in the limit $\epsilon \to 0$ to get the final form of the problem to be solved in this second step:

$$\begin{cases} \partial_t \rho = 0, \\ \partial_t \rho u = 0, \\ \mu_N \partial_t \rho \varepsilon_i = \mu_i \partial_t \rho \varepsilon_N, \quad \text{for } i = 1, \dots, N-1, \\ \partial_t \{\rho E\}(\rho, \rho u, \{\rho \varepsilon_i\}_{1 \le i \le N}) = 0. \end{cases} \tag{4.80}$$

With the benefit of the vanishing parameter $\epsilon > 0$, note that the corresponding solution is explicitly given by:

$$\rho(x,t) = \rho_h(x, t^{n+1-}), \quad \rho u(x,t) = (\rho u)_h(x, t^{n+1-}),$$

$$\rho E(x,t) = (\rho E)_h(x, t^{n+1-}),$$

$$\rho \epsilon_i(x,t) = (\rho \epsilon_i)_h(x, t^{n+1-}) + \tfrac{\mu_i}{\mu_N}(\rho \epsilon_N(x,t) - (\rho \epsilon_N)_h(x, t^{n+1-})).$$

Let us underline the very similarity of these identities with the discrete versions derived within the frame of the exact Riemann solver (see (4.56) for instance). This observation stays at the very basis of the extension we now describe in its details. Let us first introduce some notations while briefly stating some classical assumptions (see [61], [56]) about approximate Riemann solutions for (4.74)-(4.75), which will be of constant use in this section.

Approximate Riemann Solvers for (4.74)-(4.75)

Let be given a family $\mathbf{w}^a : I\!R \times \Omega^2 \to \Omega$ of Lipschitz continuous functions in the self-similar variable $\xi = x/t$. This family, intending to serve as approximate Riemann solutions for (4.74)-(4.75), is subject to the following four consistency properties:

(H_1) $\mathbf{w}^a(.,\mathbf{U},\mathbf{U})=\mathbf{U}$, for all $\mathbf{U}\in\Omega$.

(H_2) For all bounded subset $\mathcal{B}\in\Omega$, there exists a constant C such that for all pairs $(\mathbf{U}_L^1,\mathbf{U}_R^1)$ and $(\mathbf{U}_L^2,\mathbf{U}_R^2)$ in $\mathcal{B}$,

$$||\mathbf{w}^a(.,\mathbf{U}_L^1,\mathbf{U}_R^1)-\mathbf{w}^a(.,\mathbf{U}_L^2,\mathbf{U}_R^2)||\leq C(||\mathbf{U}_L^2-\mathbf{U}_L^1||+||\mathbf{U}_R^2-\mathbf{U}_R^1||).$$

(H_3) There exists an unique smaller and non negative real $\sigma_m^a(\mathbf{U}_L,\mathbf{U}_R)$ (the maximum speed at which propagate approximate waves) such that

$$\begin{cases} \mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R)=\mathbf{U}_L & \text{if} \quad \xi<-\sigma_m^a(\mathbf{U}_L,\mathbf{U}_R), \\ \mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R)=\mathbf{U}_R & \text{if} \quad \xi>\sigma_m^a(\mathbf{U}_L,\mathbf{U}_R). \end{cases}$$

(H_4) For all λ satisfying the CFL condition

$$\lambda\sigma_m^a(\mathbf{U}_L,\mathbf{U}_R)\leq\frac{1}{2}, \tag{4.81}$$

the following average conservation property on $\mathbf{w}^a$ holds true:

$$\int_{-1/2\lambda}^{1/2\lambda}\mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R)d\xi=\frac{1}{2\lambda}(\mathbf{U}_L+\mathbf{U}_R)-(\mathbf{F}(\mathbf{U}_R)-\mathbf{F}(\mathbf{U}_L)). \tag{4.82}$$

Equipped with this family of approximate Riemann solutions, let us then consider after Harten-Lax-Van Leer [61] the following two averages :

$$\begin{aligned} \overline{\mathbf{U}}_L^a &:= \overline{\mathbf{U}}_L^a(\lambda;\mathbf{U}_L,\mathbf{U}_R)=2\lambda\textstyle\int_{-1/2\lambda}^{0}\mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R)d\xi, \\ \overline{\mathbf{U}}_R^a &:= \overline{\mathbf{U}}_R^a(\lambda;\mathbf{U}_L,\mathbf{U}_R)=2\lambda\textstyle\int_{0}^{1/2\lambda}\mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R)d\xi, \end{aligned} \tag{4.83}$$

for any given pair $(\mathbf{U}_L,\mathbf{U}_R)$ in Ω^2.

Step 1 in details

Let now be given a piecewise constant approximate solution $\mathbf{U}_\lambda(x,t^n)$ for (4.74). The classical approach for updating this solution at the next time level $t^n+\Delta t$ would result, in full agreement with [61], when considering:

$$\mathbf{U}_\lambda(x,t^n+\Delta t)=\frac{1}{2}(\overline{\mathbf{U}}_L^a(\lambda;\mathbf{U_j^n},\mathbf{U_{j+1}^n})+\overline{\mathbf{U}}_R^a(\lambda;\mathbf{U_{j-1}^n},\mathbf{U_j^n})),\ x\in(x_{j-1/2},x_{j+1/2}), \tag{4.84}$$

under the CFL like condition

$$\lambda\sigma_m^a(\mathbf{U}_j^n,\mathbf{U}_{j+1}^n)\leq\frac{1}{2},\quad \lambda=\frac{\Delta t}{\Delta x}. \tag{4.85}$$

Here we propose to keep the definition of $\rho_\lambda(x,t^n+\Delta t)$ and $(\rho u)_\lambda(x,t^n+\Delta t)$ according to (4.84), that is

$$\rho_j^{n+1-}=\rho_\lambda(x,t^n+\Delta t),\quad (\rho u)_j^{n+1-}=(\rho u)_\lambda(x,t^n+\Delta t), \tag{4.86}$$

but to average each of the internal energies $\rho\varepsilon_i$ with $i \in \{1, ..., N\}$ instead of the corresponding entropy ρs_i. Namely, we introduce in full symmetry with (4.83)-(4.84) and for all $i = 1, ..., N$,

$$(\rho\varepsilon_i)_j^{n+1-} = \frac{1}{2}(\overline{\rho\varepsilon_i}_L^a(\lambda; \mathbf{U_j^n}, \mathbf{U_{j+1}^n}) + \overline{\rho\varepsilon_i}_R^a(\lambda; \mathbf{U_{j-1}^n}, \mathbf{U_j^n})), \quad \text{for all } j \in Z, \tag{4.87}$$

where we have set:

$$\begin{aligned} \overline{\rho\varepsilon_i}_L^a &= \overline{\rho\varepsilon_i}_L^a(\lambda; \mathbf{U}_L, \mathbf{U}_R) = 2\lambda \textstyle\int_{-1/2\lambda}^{0} \{\rho\varepsilon_i\}(\mathbf{w}^a(\xi, \mathbf{U}_L, \mathbf{U}_R))d\xi, \\ \overline{\rho\varepsilon_i}_R^a &= \overline{\rho\varepsilon_i}_R^a(\lambda; \mathbf{U}_L, \mathbf{U}_R) = 2\lambda \textstyle\int_{0}^{1/2\lambda} \{\rho\varepsilon_i\}(\mathbf{w}^a(\xi, \mathbf{U}_L, \mathbf{U}_R))d\xi. \end{aligned} \tag{4.88}$$

Notice that as a consequence of the proposed averages, the entropy ρs_i for $i \in \{1, ..., N\}$ actually advances in time according to

$$\{\rho s_i\}_j^{n+1-} := \{\rho s_i\}(\rho_j^{n+1-}, \{\rho\varepsilon_i\}_j^{n+1-}) \leq \frac{1}{2}(\overline{\rho s_i}_L^a + \overline{\rho s_i}_R^a), \tag{4.89}$$

by Jensen inequality. So, (4.86) and (4.89) yield the final formula for defining $\mathbf{U}_j^{n+1-}$ at the end of the first step. The updating procedure we propose is thus distinct from the usual one and is motivated at generalizing the approach developed in the previous section (see indeed formula (4.45) and remark 2 below). But let us stress at this stage that our approach preserves discrete energy inequalities (see the selection principle (4.75)) as soon as there are satisfied using the standard averages (4.84). Before addressing this issue, let us recall that the weak consistency property (H_4) is well known to give birth to some mapping $\mathcal{F}^a : \Omega^2 \to I\!R^{N+2}$ so that (4.83) by construction recasts in

$$\begin{aligned} \overline{\mathbf{U}}_R^a(\lambda; \mathbf{U}_L, \mathbf{U}_R) &= \mathbf{U}_R - 2\lambda(\mathbf{F}(\mathbf{U}_R) - \mathcal{F}^a(\mathbf{U}_L, \mathbf{U}_R)), \\ \overline{\mathbf{U}}_L^a(\lambda; \mathbf{U}_L, \mathbf{U}_R) &= \mathbf{U}_L - 2\lambda(\mathcal{F}^a(\mathbf{U}_L, \mathbf{U}_R) - \mathbf{F}(\mathbf{U}_L)). \end{aligned} \tag{4.90}$$

Each of these two identities equivalently defines the so-called numerical flux function $\mathcal{F}^a$ which is easily seen from (H_1)-(H_2) to be consistent in the usual sense of the Lax-Wendroff theorem (see [56]) and with the notations of the last section. Thus and as a consequence of (4.86)-(4.90), both the density and the momentum find an expected conservative updating formula

$$\rho_j^{n+1-} = \rho_j^n - \lambda(\mathcal{F}_\rho^a(\mathbf{U_j^n}, \mathbf{U_{j+1}^n}) - \mathcal{F}_\rho^a(\mathbf{U_{j-1}^n}, \mathbf{U_j^n})),$$

$$(\rho u)_j^{n+1-} = (\rho u)_j^n - \lambda(\mathcal{F}_{\rho u}^a(\mathbf{U_j^n}, \mathbf{U_{j+1}^n}) - \mathcal{F}_{\rho u}^a(\mathbf{U_{j-1}^n}, \mathbf{U_j^n})),$$

where clear notations have been used. Such a conclusion obviously fails for the internal energies since their governing equations are not in conservation form. It is however expected that the averages (4.87) actually yield some relevant discrete forms of the associated partial differential equations. In the case of the exact Riemann solver but for the pure conservative system (4.74)-(4.75), we indeed get with clear notations:

$$(\rho\varepsilon_i)_j^{n+1-} = (\rho\varepsilon_i)_j^n - \lambda\Delta(\rho\varepsilon_i u)_{j+1/2}^n - \lambda < p_i \partial_x u, \mathcal{C}_j^n > . \tag{4.91}$$

Remark 2
By contrast with (4.91), adopting the usual procedure for advancing in time the discrete solution, *i.e.* averaging each of the entropies ρs_i for $i \in \{1, ..., N\}$, would have yield, within the setting of the exact Riemann solver,

$$\{\rho\varepsilon_i\}(\rho_j^{n+1-}, (\rho s_i)_j^{n+1-}) \neq (\rho\varepsilon_i)_j^n - \lambda\Delta(\rho\varepsilon_i u)_{j+1/2}^n - \lambda < p_i\partial_x u, \mathcal{C}_j^n >,$$

because of the Jensen inequality. This apparent lack of consistency with the underlying partial differential equations therefore strongly promotes the updating strategy we have put forward in this section.
Without further details about the family of approximate Riemann solutions, it seems rather untractable to give an equivalent expression form for the averages (4.87). Let us however underline that such averages can be actually easily evaluated when considering the hierarchy of Approximate Riemann Solvers with k intermediate states $(0 < k < N+1)$ separated by discontinuities. We refer the reader to [19] for the details.
Let us now considering the existence of in-cell energy inequalities for our procedure. After Harten-Lax-Van Leer, we classically now ask the family of approximate Riemann solutions to obey the following weak consistency property with the Lax inequality (4.75):

(H_5) For all λ satisfying the CFL condition (4.81), the following average dissipation property on $\{\rho E\}(\mathbf{w}^a)$ holds true:

$$\begin{aligned}&\textstyle\int_{-1/2\lambda}^{1/2\lambda}\{\rho E\}(\mathbf{w}^a(\xi, \mathbf{U}_L, \mathbf{U}_R))d\xi \leq \\ &\tfrac{1}{2\lambda}(\{\rho E\}(\mathbf{U}_L) + \{\rho E\}(\mathbf{U}_R)) - (\{\rho Hu\}(\mathbf{U}_R) - \{\rho Hu\}(\mathbf{U}_L)).\end{aligned} \tag{4.92}$$

Then, defining

$$\begin{aligned}\overline{\{\rho E\}}_L^a &= \overline{\{\rho E\}}_L^a(\lambda; \mathbf{U}_L, \mathbf{U}_R) = 2\lambda \textstyle\int_{-1/2\lambda}^{0}\{\rho E\}(\mathbf{w}^a(\xi, \mathbf{U}_L, \mathbf{U}_R))d\xi, \\ \overline{\{\rho E\}}_R^a &= \overline{\{\rho E\}}_R^a(\lambda; \mathbf{U}_L, \mathbf{U}_R) = 2\lambda \textstyle\int_{0}^{1/2\lambda}\{\rho E\}(\mathbf{w}^a(\xi, \mathbf{U}_L, \mathbf{U}_R))d\xi,\end{aligned} \tag{4.93}$$

a direct consequence of the above consistency property is the expected easy following result (see [56]).

Lemma 19 *Let $\{\rho\mathcal{H}u\}^-$ and $\{\rho\mathcal{H}u\}^+$ be two $\Omega^2 \to I\!R$ mappings defined by:*

$$\begin{aligned}\{\rho\mathcal{H}u\}^-(\mathbf{U}_L, \mathbf{U}_R) &= \{\rho Hu\}(\mathbf{U}_R) + \tfrac{1}{2\lambda}(\overline{\{\rho E\}}_R^a(\lambda; \mathbf{U}_L, \mathbf{U}_R) - \{\rho E\}(\mathbf{U}_R)), \\ \{\rho\mathcal{H}u\}^+(\mathbf{U}_L, \mathbf{U}_R) &= \{\rho Hu\}(\mathbf{U}_L) - \tfrac{1}{2\lambda}(\overline{\{\rho E\}}_L^a(\lambda_L; \mathbf{U}_L, \mathbf{U}_R) - \{\rho E\}(\mathbf{U}_L)),\end{aligned} \tag{4.94}$$

for all $(\mathbf{U}_L, \mathbf{U}_R) \in \Omega^2$.
Assume that the CFL condition (4.81) and hypothesis (H) are valid.
Then, we have for all $(\mathbf{U}, \mathbf{U}_L, \mathbf{U}_R) \in \Omega^3$:

$$\{\rho\mathcal{H}u\}^-(\mathbf{U}_L, \mathbf{U}_R) \leq \{\rho\mathcal{H}u\}^+(\mathbf{U}_L, \mathbf{U}_R),$$

and

$$\{\rho\mathcal{H}u\}^-(\mathbf{U}, \mathbf{U}) = \{\rho\mathcal{H}u\}^+(\mathbf{U}, \mathbf{U}) = \{\rho Hu\}(\mathbf{U}).$$

Let us then state the main result of the step 1.

Proposition 24 *Let the approximate solution be advanced according to (4.86)-(4.89) under the CFL restriction (4.85). Then provided that the entropy consistency condition* (H_5) *is met, the following in-cell energy inequalities are satisfied for all* $j \in Z$ *:*

$$\{\rho E\}(\mathbf{U}_j^{n+1-}) - (\rho E)_j^n + \lambda\Delta(\{\rho\mathcal{H}u\}^+(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n) \leq 0. \tag{4.95}$$

In addition, we have :

$$(\rho s_i)_j^{n+1-} - (\rho s_i)_j^n + \lambda\Delta\{\rho s_i u\}_{j+1/2}^n \leq 0, \tag{4.96}$$

for all $i \in \{1, ..., N\}$.

The proof is postponed at the end of the present section. It will clearly indicate that energy inequalities identical to (4.95) but with the energy numerical flux $\{\rho\mathcal{H}u\}^-$ also hold. For definiteness in the forthcoming developments, we choose without restriction the flux $\{\rho\mathcal{H}u\}^+$. Observe in addition that these energy inequalities are in complete analogy with the ones obtained in the setting of the exact Riemann solver (see proposition 21).

Restoring the conservation of the total energy while enforcing for validity the generalized jump conditions is now the aim of the second step in the method.

Step 2 in details
Following the guidelines of the rough description of the method, we are led to set :

$$\rho_j^{n+1} = \rho_j^{n+1-}, \quad (\rho u)_j^{n+1} = (\rho u)_j^{n+1-}, \quad \text{for all } j \in Z, \tag{4.97}$$

while defining the N internal energies $(\rho\varepsilon_i)_j^{n+1}$ as the solutions of the following linear system :

$$\left\{\begin{array}{l} \mu_N\{(\rho\varepsilon_i)_j^{n+1} - (\rho\varepsilon_i)_j^{n+1-}\} = \mu_i\{(\rho\varepsilon_N)_j^{n+1} - (\rho\varepsilon_N)_j^{n+1-}\}, \ 1 \leq i \leq N-1, \\ \displaystyle\sum_{i=1}^{N}(\rho\varepsilon_i)_j^{n+1} = (\rho E)_j^{n+1} - \frac{((\rho u)_j^{n+1})^2}{2\rho_j^{n+1}}. \end{array}\right. \tag{4.98}$$

In order to restore the conservation of the total energy, we propose to advance the total energy $(\rho E)_j^{n+1}$ in time according to :

$$(\rho E)_j^{n+1} := (\rho E)_j^n - \lambda\Delta\{\rho\mathcal{H}u\}^+(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n). \tag{4.99}$$

As a consequence of these definitions we have :

Proposition 25 *The linear system (4.98) in the unknown* $\{(\rho\varepsilon_i)_j^{n+1}\}_{1\leq i\leq N}$ *admits an unique solution explicitly given by :*

$$(\rho\varepsilon_i)_j^{n+1} = (\rho\varepsilon_i)_j^{n+1-} + \frac{\mu_i}{\sum_{l=1}^N \mu_l}((\rho E)_j^{n+1} - \frac{((\rho u)_j^{n+1})^2}{2\rho_j^{n+1}} - \sum_{l=1}^{N}(\rho\varepsilon_l)_j^{n+1-}), \tag{4.100}$$

for all $i = 1, \ldots, N$.

This statement highlights the complete symmetry in the updating formulae coming either with the exact Riemann solver considered in the previous section or with the extensions we propose (see proposition 23).

Remark 3
As soon as the Approximate Riemann Solver under consideration allows for a discrete form of the partial differential equations governing $\{\rho\varepsilon_i\}_{i=1,\dots,N}$ (see remark 2), then (4.98) reads

$$\begin{gathered}\mu_N\{(\rho\varepsilon_i)_j^{n+1} - (\rho\varepsilon_i)_j^n + \lambda\Delta\{\rho\epsilon_j u\}_{j+1/2}^n + \lambda < p_i\partial_x u, \mathcal{C}_j^n > \}\\ = \\ \mu_i\{(\rho\varepsilon_N)_j^{n+1} - (\rho\varepsilon_N)_j^n + \lambda\Delta\{\rho\epsilon_N u\}_{j+1/2}^n + \lambda < p_N\partial_x u, \mathcal{C}_j^n > \},\end{gathered}$$

which is nothing but a discrete version of (4.22).

In addition, these extensions share with the exact Riemann solver the following desirable stability properties:

Theorem 10 *Under the CFL condition (4.85), the numerical method (4.86)-(4.87)-(4.97)-(4.99)-(4.100) preserves the positivity of the internal energies, that is to say* $(\rho\varepsilon_i)_j^{n+1} \geq 0$ *for* $i = 1,\dots,N$ *and all* $j \in Z$. *Moreover, it obeys for* $i = 1,...,N$ *and all* $j \in Z$ *the following entropy inequalities:*

$$(\rho s_i)_j^{n+1} := \{\rho s_i\}(\rho_j^{n+1}, (\rho\varepsilon_i)_j^{n+1}) \leq (\rho s_i)_j^n - \lambda\Delta\{\rho s_i u\}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n). \tag{4.101}$$

Assuming in addition that the family of approximate Riemann solutions under consideration obeys the following weak form of the maximum principles on the specific entropies $\{s_i\}_{i=1,\dots,N}$:

$$\begin{aligned}\int_{-1/2\lambda}^0\{\rho s_i\}(\mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R))d\xi &\leq \max((s_i)_L,(s_i)_R)\int_{-1/2\lambda}^0\{\rho\}(\mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R))d\xi,\\ \int_0^{1/2\lambda}\{\rho s_i\}(\mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R))d\xi &\leq \max((s_i)_L,(s_i)_R)\int_0^{1/2\lambda}\{\rho\}(\mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R))d\xi,\end{aligned} \tag{4.102}$$

then we have for all $i = 1,...,N$:

$$(s_i)_j^{n+1} \leq \max((s_i)_{j-1}^n, (s_i)_j^n, (s_i)_{j+1}^n), \quad j \in Z.$$

We conclude this section when giving the proofs of the results we have claimed true. To that purpose, we previously state and prove the following lemma.

Lemma 20 *Assume that the Approximate Riemann Solver* $\mathbf{w}^a$ *obeys* (4.102). *Then, with clear notations, the entropy inequalities:*

$$\begin{aligned}\{\rho s_i\}(\overline{\rho}_L^a, \overline{\rho\varepsilon_i}_L^a) &\leq \{\rho s_i\}(\rho_L, (\rho\varepsilon_i)_L) - 2\lambda(\mathcal{F}_{\rho s_i}^a(\mathbf{U}_L,\mathbf{U}_R) - (\rho s_i u)_L),\\ \{\rho s_i\}(\overline{\rho}_R^a, \overline{\rho\varepsilon_i}_R^a) &\leq \{\rho s_i\}(\rho_R, (\rho\varepsilon_i)_R) - 2\lambda((\rho s_i u)_R - \mathcal{F}_{\rho s_i}^a(\mathbf{U}_L,\mathbf{U}_R)),\end{aligned} \tag{4.103}$$

hold true for all $i=1,...,N$, as well as the maximum principles:

$$\begin{aligned}\{s_i\}(\overline{\rho}^a_L,\overline{\rho\varepsilon_i}^a_L) &:= \tfrac{\{\rho s_i\}(\overline{\rho}^a_L,\overline{\rho\varepsilon_i}^a_L)}{\overline{\rho}^a_L} \leq \max((s_i)_L,(s_i)_R),\\ \{s_i\}(\overline{\rho}^a_R,\overline{\rho\varepsilon_i}^a_R) &:= \tfrac{\{\rho s_i\}(\overline{\rho}^a_R,\overline{\rho\varepsilon_i}^a_R)}{\overline{\rho}^a_R} \leq \max((s_i)_L,(s_i)_R).\end{aligned} \tag{4.104}$$

Proof of lemma 20
First, arguing about the convexity of the mapping $(\rho,\rho\varepsilon_i)\mapsto\{\rho s_i\}(\rho,\rho\varepsilon_i)$ for all $i=1,...,N$, Jensen's inequality leads to

$$\begin{aligned}\{\rho s_i\}(\overline{\rho}^a_L,\overline{\rho\varepsilon_i}^a_L) &\leq 2\lambda\textstyle\int_{-1/2\lambda}^{0}\{\rho s_i\}(\mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R))d\xi,\\ \{\rho s_i\}(\overline{\rho}^a_R,\overline{\rho\varepsilon_i}^a_R) &\leq 2\lambda\textstyle\int_{0}^{1/2\lambda}\{\rho s_i\}(\mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R))d\xi.\end{aligned} \tag{4.105}$$

Thus, entropy inequalities (4.103) are immediately obtained provided we mention (4.90). In addition, maximum principles (4.104) are also an immediate consequence of (4.105) thanks to (4.102).

We now follow by establishing proposition 24, related to the step 1.

Proof of proposition 24
First, invoking the convexity of the mapping $(\rho,\rho u)\mapsto\frac{(\rho u)^2}{2\rho}$ and the definition (4.84)-(4.86)-(4.87), we obtain:

$$\begin{gathered}\{\rho E\}(\mathbf{U}_j^{n+1-}) := \tfrac{((\rho u)_j^{n+1-})^2}{2\rho_j^{n+1-}} + \sum_{i=1}^N(\rho\varepsilon_i)_j^{n+1-} \leq\\ \tfrac{1}{2}\big(\tfrac{(\overline{\rho u}^a_L)^2}{2\overline{\rho}^a_L}+\tfrac{(\overline{\rho u}^a_R)^2}{2\overline{\rho}^a_R}\big)+\sum_{i=1}^N(\rho\varepsilon_i)_j^{n+1-},\end{gathered}$$

so that Jensen's inequality and definitions (4.83) lead to

$$\begin{gathered}\{\rho E\}(\mathbf{U}_j^{n+1-}) \leq \sum_{i=1}^N(\rho\varepsilon_i)_j^{n+1-}+\\ \tfrac{1}{2}\Big(2\lambda\int_{-1/2\lambda}^{0}\{\tfrac{(\rho u)^2}{2\rho}\}(\mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R))d\xi+2\lambda\int_0^{1/2\lambda}\{\tfrac{(\rho u)^2}{2\rho}\}(\mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R))d\xi\Big).\end{gathered}$$

Arguing now about the averages (4.87)-(4.88), we easily obtain

$$\begin{gathered}\{\rho E\}(\mathbf{U}_j^{n+1-}) \leq\\ \tfrac{1}{2}\Big(2\lambda\int_{-1/2\lambda}^{0}\{\rho E\}(\mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R))dx+2\lambda\int_0^{1/2\lambda}\{\rho E\}(\mathbf{w}^a(\xi,\mathbf{U}_L,\mathbf{U}_R))dx\Big),\end{gathered}$$

that is to say

$$\{\rho E\}(\mathbf{U_j^{n+1,-}}) \leq \frac{1}{2}\big(\overline{\{\rho E\}}^a_L(\lambda;\mathbf{U^n_j},\mathbf{U^n_{j+1}})+\overline{\{\rho E\}}^a_R(\lambda;\mathbf{U^n_{j-1}},\mathbf{U^n_j})\big),$$

or more precisely

$$\{\rho E\}(\mathbf{U}_j^{n+1-}) \leq (\rho E)_j^n - \lambda(\{\rho\mathcal{H}u\}^+(\mathbf{U}_j^n,\mathbf{U}_{j+1}^n)-\{\rho\mathcal{H}u\}^-(\mathbf{U}_{j-1}^n,\mathbf{U}_j^n)),$$

using the definitions (4.94) of $\{\rho\mathcal{H}u\}^{\pm}$. To conclude with the desired energy inequality (4.95), we juste have to recall the validity of $\{\rho\mathcal{H}u\}^{-} \leq \{\rho\mathcal{H}u\}^{+}$ stated in the lemma 19.

Concerning now the entropy inequalities (4.96), the convexity of the mapping $\{\rho s_i\}$ for all $i = 1, ..., N$ and definitions (4.84)-(4.86)-(4.87) of ρ_j^{n+1-} and $\{(\rho\varepsilon_i)_j^{n+1,-}\}_{i=1,...,N}$, ensure that

$$\{\rho s_i\}(\rho_j^{n+1-}, (\rho\varepsilon_i)_j^{n+1,-}) \leq \tfrac{1}{2}(\{\rho s_i\}(\overline{\rho}_L^a, \overline{\rho\varepsilon_i}_L^a) + \{\rho s_i\}(\overline{\rho}_R^a, \overline{\rho\varepsilon_i}_R^a)). \tag{4.106}$$

The required conclusion is then easily obtained by applying inequalities (4.103) of lemma 20.

Now, it remains to give the proof of theorem 10.

Proof of theorem 10

Since by construction, the total energy $(\rho E)_j^{n+1}$ finds the conservative discrete formulation (4.99), we immediately infer from (4.95) :

$$\{\rho E\}(\mathbf{U}_j^{n+1-}) = \frac{((\rho u)_j^{n+1-})^2}{2\rho_j^{n+1-}} + \sum_{i=1}^{N}(\rho\varepsilon_i)_j^{n+1-} \leq (\rho E)_j^{n+1}. \tag{4.107}$$

As a consequence, invoking the explicit formulae (4.100) for defining each of the internal energies $(\rho\varepsilon_i)_j^{n+1}$, the positivity of the viscosity coefficients easily ensures that :

$$(\rho\varepsilon_i)_j^{n+1} \geq (\rho\varepsilon_i)_j^{n+1-} \quad \text{for all } i \in \{1, \ldots, N\}. \tag{4.108}$$

But, $(\rho\varepsilon_i)_j^{n+1-}$ stays by construction positive. Indeed, this quantity is nothing but the average (see indeed (4.87)-(4.88)) of the exact internal energy $(\rho\varepsilon_i)$ evaluated on approximate Riemann solutions under the CFL restriction (4.85). Such solutions stay within the admissible phase space and this gives the required conclusion.

Turning now establishing the entropy inequalities (4.101), we first recall that

$$\frac{\partial\rho s_i}{\partial\rho\varepsilon_i} = -\frac{1}{T_i} < 0,$$

to get for all $i = 1, ..., N$ and thanks to (4.108) :

$$(\rho s_i)_j^{n+1} = \{\rho s_i\}(\rho_j^{n+1}, (\rho\varepsilon_i)_j^{n+1}) \leq \{\rho s_i\}(\rho_j^{n+1}, (\rho\varepsilon_i)_j^{n+1,-}).$$

As an immediate consequence of (4.106) and for concluding, we now just quote lemma 20, namely (4.103) for the expected entropy inequalities and (4.104) for the maximum principles.

4.6 Numerical experiments

In this section, we perform numerical experiments for illustrating the correct design of the scheme we have described in the previous section. To that purpose, we deal with polytropic ideal gases associated with N constant adiabatic exponents $\gamma_i > 1$ for all $i = 1, ..., N$. As a consequence, observe in particular that the total energy now reads:

$$\rho E = \frac{(\rho u)^2}{2\rho} + \sum_{i=1}^{N} \frac{p_i}{\gamma_i - 1}.$$

In addition and for simplicity, we use constant viscosity laws $\{\mu_i\}_{i=1,\dots,N}$ with a Reynolds number equal to 10^5 which is typical in aerodynamic.
The system (4.72) is solved with the following Riemann initial data:

$$\mathbf{U}(x,0) = \begin{cases} \mathbf{U}_L & \text{if} \quad x < 0, \\ \mathbf{U}_R & \text{if} \quad x > 0, \end{cases} \tag{4.109}$$

$\mathbf{U}_L$ and $\mathbf{U}_R$ to be prescribed in the phase space. We consider a 300-point mesh and plot the corresponding numerical solutions (see figures below) together with the exact ones and the ones given by a classical approach. Here "classical" means that we do not enforce for validity (4.22) (and then the generalized jump conditions (4.40)) from a numerical point of view. This is the very difference with the one proposed in the previous section. To conclude, let us precise that we have used the Relaxation method described in [19] as a suitable Approximate Riemann Solver (see previous section). We also refer the reader to [19] for additional numerical experiments.

Experiment 1

We set $N = 3$ and choose

γ_1	γ_2	γ_3	μ_2/μ_1	μ_3/μ_1
1.4	1.4	1.4	1.	1.

while the left and right states entering (4.109) write as follows:

Experiment 1

	density	velocity	pressure 1	pressure 2	pressure 3
$\mathbf{U}_L$	2.	1.	1.	1.2	1.4

	density	velocity	pressure 1	pressure 2	pressure 3
$\mathbf{U}_R$	1.2838	-1.3990	0.1850	0.1305	0.2550

Since the adiabatic exponents are equal, it turns out that the system under study admits an equivalent full conservation form. More precisely, this system is nothing but the usual 3×3 Euler equations (with the pressure $p := \sum_{i=1}^{N} p_i$) supplemented with $(N-1)$ additional independent transport equations (we refer the reader to [10] for such considerations). As a consequence, the classical approach should provide numerical solutions of ρ, ρu and ρE in full agreement with the exact ones. The difference may be located on the partial pressure laws only. Concerning our new approach, it turns out that each component agrees with the exact solution (see below).

Experiment 2

The adiabatic exponents are now different and the ratios of the viscosity laws no longer necessarily equal to 1:

γ_1	γ_2	γ_3	μ_2/μ_1	μ_3/μ_1
1.2	1.4	1.6	1.	100.

The left and right states entering (4.109) read as follows:

Experiment 2

	density	velocity	pressure 1	pressure 2	pressure 3
$\mathbf{U}_L$	2.	2.	1.5	1.	1.

	density	velocity	pressure 1	pressure 2	pressure 3
$\mathbf{U}_R$	3.3244	-1.2999	1.4254	0.8667	0.0537

Once more, we observe that our new method provides us with good numerical results compared to the classical approach, in spite of the fact that this test case is difficult (the pressure jumps are very large).

Experiment 1

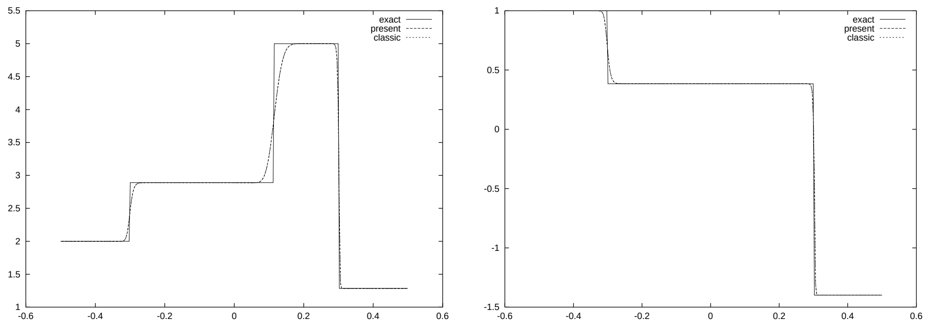

FIG. 4.1 – *Density, Velocity*

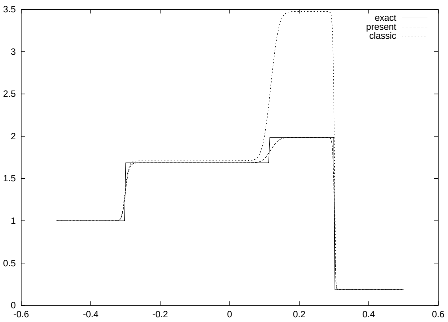

FIG. 4.2 – *Pressure 1*

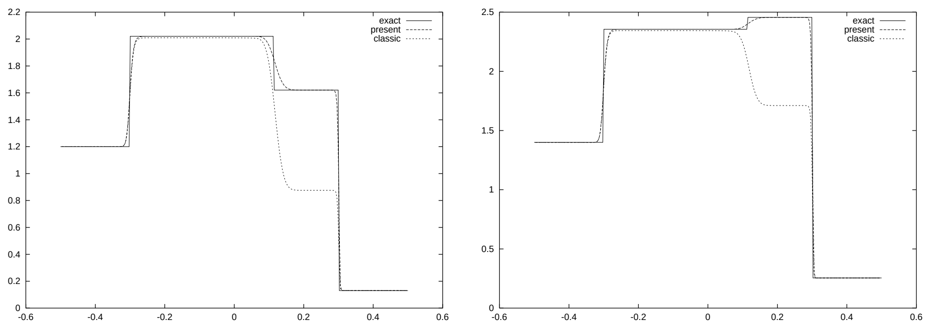

FIG. 4.3 – *Pressure 2, Pressure 3*

Experiment 2

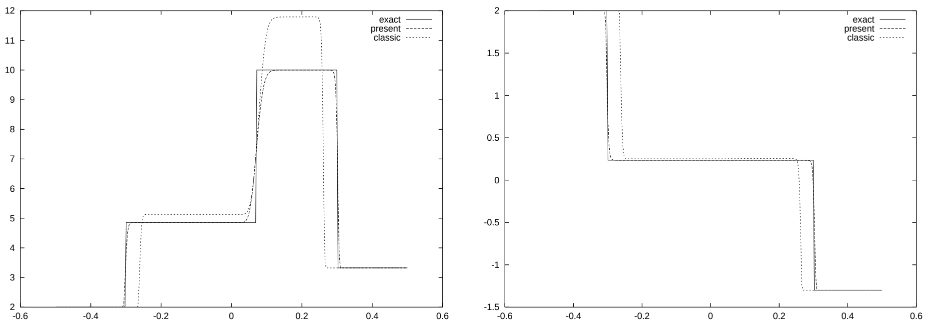

FIG. 4.4 – *Density, Velocity*

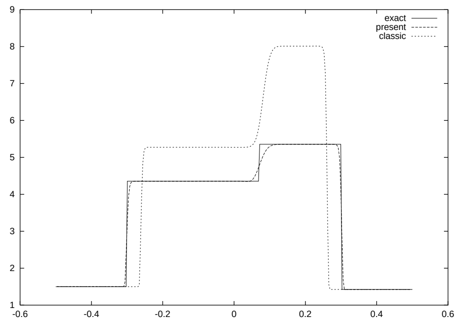

FIG. 4.5 – *Pressure 1*

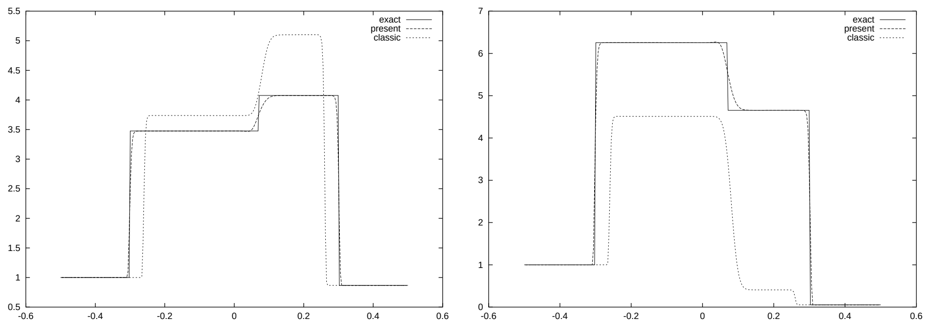

FIG. 4.6 – *Pressure 2, Pressure 3*

Troisième partie

Calcul scientifique

Cette partie du mémoire est consacrée au Calcul Scientifique. Elle est motivée par l'approximation numérique des solutions stationnaires à deux variables d'espace des équations de Navier-Stokes multi-pression.
La nature bi-dimensionnelle du problème considéré ne posera pas de difficultés de principe car nous nous bornerons à exploiter la propriété d'invariance par rotation des équations. Il est bien connu qu'une telle propriété permet de n'avoir à considérer qu'une formulation quasi-1D des équations, pour déduire à des fins d'approximation numérique une méthode bi-dimensionnelle consistante. Les arguments sous-tendant une telle démarche sont excessivement classiques et nous avons délibérément choisi de ne pas les rapporter dans ce mémoire. Nous rappellerons en temps utile quelques références qui nous ont aidés lors de la conception du code de calcul 2D. Soulignons simplement que d'autres approches, réellement bi-dimensionnelles, existent : citons en particulier les méthodes de "fluctuation-splitting" (voir la contribution récente [5] ainsi que les références citées).
La principale difficulté à laquelle nous aurons à répondre est le caractère stationnaire des solutions bi-dimensionnelles envisagées. Conformément à une approche universellement adoptée, nous proposons d'approcher les solutions stationnaires comme limites en temps infini de solutions instationnaires du système d'EDP considéré. Cette stratégie, dite de marche en temps, évite la perte de la propriété d'hyperbolicité dans la formulation stationnaire de l'opérateur du premier ordre sous-jacent. Puisque seules les solutions stationnaires nous intéressent, il convient de s'affranchir de toute notion de restriction CFL lors de la marche en temps pour parvenir le plus rapidement possible vers une solution indépendante du temps. Le "plus rapidement possible" doit être compris ici en terme de temps CPU. Il s'agit dès lors de concilier coût d'évaluation et absence de restriction CFL. Notre principale préoccupation sera de répondre à cette exigence. Une immense littérature existe dans le contexte des équations de Navier-Stokes usuelles. Nous avons choisi d'étendre à notre cadre une technique très éprouvée dans la pratique consistant à linéariser une formulation totalement implicite en temps d'une variante des méthodes de volumes finis discutées dans les précédents chapitres. Une telle approche relève d'un certain nombre d'heuristiques (dont la justification rigoureuse semble être hors de portée) mais s'avère performante dans les applications. Dans les faits, la stratégie développée ci-après reposera sur l'inversion, à chaque itération, de systèmes linéaires de petite taille (à matrice tridiagonale par blocs 3×3 et à matrices tridiagonales scalaires). Nous montrerons qu'il sera possible de procéder à l'itération suivante dès que la densité et l'énergie interne totale sont positives à l'itération courante (chacune des entropies restant alors nécessairement positive).

Plan de la seconde partie
L'objectif de ce paragraphe est de procéder à quelques observations générales permettant de motiver les principaux choix algorithmiques présidant la définition de la méthode d'intégration en temps discutée dans les chapitres suivants. L'énoncé de ces choix permettra d'introduire le plan général de cette seconde partie. Puisque nous exploiterons l'invariance par rotation des équations, nous mènerons toute la discussion

dans le contexte d'une seule variable d'espace. En reprenant les notations du chapitre 2 (partie II), il s'agit d'approcher les solutions du système suivant :

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + \sum_{i=1}^N p_i) = \epsilon \partial_x ((\sum_{i=1}^N \mu_i) \partial_x u), \\ \partial_t \rho E + \partial_x (\rho E + \sum_{i=1}^N p_i) u = \epsilon \partial_x ((\sum_{i=1}^N \mu_i) u \partial_x u), \\ \partial_t \rho S_i + \partial_x \rho S_i u = -\epsilon \frac{\mu_i}{T_i} (\partial_x u)^2, \quad i = 1, ..., N-1. \end{cases} \tag{110}$$

Nous préciserons dès le prochain chapitre les relations de fermeture considérées, mais rappelons que les solutions régulières de (110) vérifient sans condition supplémentaire la loi d'évolution suivante :

$$\partial_t \rho S_N + \partial_x \rho S_N u = -\epsilon \frac{\mu_N}{T_N} (\partial_x u)^2. \tag{111}$$

Proposons une interprétation heuristique des développements menés lors des précédents chapitres. Les solutions considérées présentent des zones de transition très rapide (profils de choc visqueux dans le régime d'un grand nombre de Reynolds, *i.e.* $\epsilon << 1$) dont les états de sortie s'avèrent difficiles à capturer numériquement. Cette difficulté est essentiellement liée au fait que la zone de transition n'étant pas résolue, chacun des bilans d'entropie $\epsilon \frac{\mu_i}{T_i} (\partial_x u)^2$ est largement sous-évalué au niveau discret lorsque le pas de discrétisation en espace est réaliste (le nombre de Reynolds ramené à ce pas est trop grand). Ces observations ont alors motivé l'introduction de la formulation équivalente suivante :

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + \sum_{i=1}^N p_i) = \epsilon \partial_x ((\sum_{i=1}^N \mu_i) \partial_x u), \\ \partial_t \rho E + \partial_x (\rho E + \sum_{i=1}^N p_i) u = \epsilon \partial_x ((\sum_{i=1}^N \mu_i) u \partial_x u), \\ \partial_t \rho S_i + \partial_x \rho S_i u = \frac{\mu_i}{\mu_N} \frac{T_N}{T_i} \{ \partial_t \rho S_N + \partial_x \rho S_N u \}, \quad i = 1, ..., N-1, \end{cases} \tag{112}$$

puisque le rapport des viscosités impliqué dans chacune des $(N-1)$ dernières relations de proportionnalité est indépendant du nombre de Reynolds. Cette remarque permet d'appréhender le succès de la formulation équivalente (112) dans l'approximation numérique des solutions considérées. Répondant à notre motivation générale, il nous faut proposer une implicitation linéarisée en temps du système équivalent (112). Celui-ci peut être compris comme la perturbation visqueuse d'un système hyperbolique dont la forme se déduit en annulant formellement ϵ dans (112) (et non (110), voir les chapitres consacrés au régime asymptotique $\epsilon \to 0$). Comme il est connu (voir chapitre 2), une réponse adéquate repose sur la connaissance d'un solveur de Riemann approché, de type Roe par exemple, pour le système hyperbolique considéré. Or, nous ne connaissons pas un tel solveur même si la solution exacte du problème de Riemann a été mise en évidence au chapitre II 1. Nous devons donc recourir à une stratégie de prédiction-correction conforme à celles développées dans les précédents chapitres.
L'étape de prédiction repose sur la résolution du problème de Riemann pour le système de lois de conservation :

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + \sum_{i=1}^N p_i) = 0, \\ \partial_t \rho E + \partial_x (\rho E + \sum_{i=1}^N p_i) u = 0, \\ \partial_t \rho S_i + \partial_x \rho S_i u = 0, \quad i = 1, ..., N-1, \end{cases} \tag{113}$$

l'étape de correction visant ensuite à forcer la validité des $(N-1)$ relations de proportionnalité :

$$\partial_t \rho S_i + \partial_x \rho S_i u = \frac{\mu_i}{\mu_N}\frac{T_N}{T_i}\{\partial_t \rho S_N + \partial_x \rho S_N u\}, \quad i = 1, ..., N-1. \tag{114}$$

Notre principal objectif est donc d'impliciter en temps de manière découplée chacune de ces étapes. Les chapitres 1 et 2 sont consacrés à l'exposé d'une technique d'intégration implicite en temps des équations (113) et la méthode associée aux relations (114) est décrite au chapitre 3.

De manière à affiner nos objectifs, rappelons qu'il est possible d'associer une linéarisation de type Roe au système hyperbolique (113) (voir Berthon [10] pour une construction dans le cas $N=2$ s´étendant aisément au cas N quelconque). Toutefois, les techniques d'implicitation usuelles de la méthode de Roe (*cf.* chapitre 2 pour les détails) nécessitent la résolution d'un système linéaire dont la matrice est tridiagonale par blocs $(N+2)\times(N+2)$. Il semble inutilement coûteux d'avoir à procéder à une telle inversion puisque la mise à jour résultante de chacune des N entropies sera détruite à l'occasion de l'étape de correction. L'idée naturelle est donc de chercher à découpler les $(N-1)$ équations gouvernant les entropies des trois premières équations, appelées dans la suite bloc hydrodynamique.
Le principal obstacle est la dépendance explicite de la pression totale relativement à toutes les entropies. De sorte à contourner cette dépendance, nous développons au chapitre 1 une stratégie de Relaxation selon laquelle la pression totale est relaxée au profit d'une nouvelle variable d'évolution. Dans une formulation totalement explicite en temps, le schéma résultant sera identique à la méthode de volumes finis déduite d'une méthode de Relaxation voisine, introduite au chapitre II 3. La méthode de volumes finis associée capture exactement par construction les solutions discontinuité de contact stationnaire. Dans le contexte des équations de Navier-Stokes, l'importance de cette propriété de précision est montrée dans [143]. Rappelons simplement qu'au chapitre II 3, la pression totale relaxée dépendait de toutes les entropies et que pour cette raison, elle ne permet pas de proposer une implicitation répondant à notre exigence de découplage. Disposant de notre nouvelle approche, nous proposerons au chapitre 2 une stratégie d'implicitation en temps découplée entre le bloc hydrodynamique et le bloc des équations de transport des entropies. Le découplage opéré résultera en l'inversion d'un problème linéaire de matrice tridiagonale par blocs 3×3, et en l'inversion de $(N-1)$ systèmes linéaires associés à la même matrice tridiagonale scalaire. Par construction, la technique de mise à jour des entropies respectera leur positivité et ce sans restriction CFL. Ce chapitre conclura la description de l'implicitation en temps de l'étape de prédiction.
Les champs de densité, d'impulsion et d'énergie totale étant connus à l'issue de cette étape, nous proposerons au chapitre 3 de corriger chacune des N entropies selon une intégration implicite en temps, découplée, des relations de proportionnalité (114). Il s'agira de résoudre $(N-1)$ systèmes linéaires associés à une même matrice tridiagonale scalaire. La technique proposée préservera la positivité des entropies dès que la densité et l'énergie interne totale restent positives à l'issue de l'étape de prédiction, et ce pour un très faible coût d'évaluation.
Pour conclure cette partie, le chapitre 4 applique les développements menés ci-dessus aux modèles multi-échelle de turbulence. Nous adoptons une stratégie de splitting

d'opérateurs. La prise en compte des lois de Fourier se fait par un algorithme implicite linéarisé en temps assurant la positivité des grandeurs concernées. Concernant la prise en compte des termes sources, nous montrons que l'implicitation linéarisée proposée s'intègre explicitement en temps. Les formules de mise à jour obtenues garantissent une fois de plus la positivité des grandeurs impliquées. Précisons que par manque de temps, aucune simulation numérique n'a pu être effectuée sur ce modèle.

Nous avons implémenté l'algorithme proposé dans le code CELHYO (voir par exemple [33], [35], [37], [38]) développé à l'ONERA. Le volume de programmation nécessaire a été important car les algorithmes développés dans les sections suivantes sont dédiés à une problématique nouvelle et non initialement prévue dans le code. Nous avons toutefois bénéficié, outre de la structure bi-dimensionnelle du code, d'un certain nombre de sous programmes tels que l'approximation (explicite et implicite) des contributions visqueuses sur l'impulsion et l'énergie totale, ou encore les algorithmes de résolution de problèmes linéaires (GMRES en l'occurence). Nous les avons utilisés sans modification.

Chapitre 1

Méthode de Relaxation pour les équations d'Euler multi-pression

1.1 Introduction

Ce chapitre est consacré à la description d'une méthode de type volumes finis, explicite en temps, pour l'approximation numérique des solutions faibles du système suivant :

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + \sum_{i=1}^N p_i) = 0, \\ \partial_t \rho E + \partial_x (\rho E + \sum_{i=1}^N p_i) u = 0, \\ \partial_t \rho S_i + \partial_x \rho S_i u = 0, \quad i = 1, ..., N-1. \end{cases} \tag{1.1}$$

Nous considérons ici N lois de pressions générales répondant aux lois d'état complètes suivantes. Les pressions partielles $\{p_i\}_{i=1,...,N}$ dérivent d'énergies internes $(\tau, S_i) \to \varepsilon_i(\tau, S_i) > 0$ supposées régulières et strictement convexes pour tout $i = 1, ..., N$, et telles que

$$-T_i(\tau, S_i) = \frac{\partial \varepsilon_i}{\partial S_i}(\tau, S_i), \quad -p_i(\tau, S_i) = \frac{\partial \varepsilon_i}{\partial \tau}(\tau, S_i), \quad i = 1, ..., N. \tag{1.2}$$

De manière classique, les températures T_i et les pressions partielles p_i sont supposées positives. Ainsi, les applications $(\tau, \varepsilon_i) \to S_i(\tau, \varepsilon_i)$ sont strictement convexes pour tout $i = 1, ..., N$, tout comme, avec un léger abus de notations, les applications $(\rho, \rho\varepsilon_i) \to \rho S_i(\rho, \rho\varepsilon_i) = \rho S_i(1/\rho, \rho\varepsilon_i/\rho)$ et $(\rho, \rho S_i) \to \rho\varepsilon_i(\rho, \rho S_i) = \rho\varepsilon_i(1/\rho, \rho S_i/\rho)$ (voir [57] pour les détails). Remarquons également que par convexité des énergies internes ε_i pour tout $i = 1, ..., N$, nous avons $\partial p_i / \partial \tau < 0$.
La loi de fermeture sur l'énergie totale est donnée par :

$$\rho E = \frac{(\rho u)^2}{2\rho} + \sum_{i=1}^N \rho\varepsilon_i(1/\rho, \rho S_i/\rho). \tag{1.3}$$

Pour simplifier les démonstrations qui vont suivre, nous adoptons les hypothèses techniques suivantes : les pressions partielles $\{p_i\}_{i=1,...,N}$ sont telles que pour tout $i = 1, ..., N$:

$$\begin{cases} \frac{\partial^2 p_i}{\partial \tau^2}(\tau, S_i) > 0, \\ \\ \lim_{\tau\to 0} p_i(\tau, S_i) = +\infty, \quad \lim_{\tau\to +\infty} p_i(\tau, S_i) = 0. \end{cases}$$

Elles satisfont donc aux hypothèses de Weyl, hypothèses classiques dans le cadre usuel $N = 1$.
Le système hyperbolique (1.1) est naturellement complété par l'inégalité d'entropie

$$\partial_t \rho S_N + \partial_x \rho S_N u \leq 0, \tag{1.4}$$

où l'application ρS_N des inconnues est strictement convexe. Avec des notations claires, nous introduisons la forme condensée suivante pour (1.1)-(1.4) :

$$\begin{aligned} &\partial_t \mathbf{u} + \partial_x \mathbf{f}(\mathbf{u}) = 0, \\ &\partial_t \rho S_N(\mathbf{u}) + \partial_x \rho S_N(\mathbf{u}) u \leq 0. \end{aligned} \tag{1.5}$$

L'espace des états associé à (1.1) est donné par :

$$\Theta = \{\mathbf{u} \in I\!R^{N+2}, \rho > 0, \{\rho S_i > 0\}_{i=1,...,N-1}, \rho E - \frac{(\rho u)^2}{2\rho} - \sum_{i=1}^{N-1} \rho \varepsilon_i(1/\rho, S_i) > 0\}. \tag{1.6}$$

Ainsi que précisé dans l'introduction générale de la seconde partie du mémoire, notre objectif est d'approcher les solutions faibles du système (1.5) par celles d'un système de Relaxation autorisant un découplage naturel entre le bloc hydrodynamique gouvernant $\rho, \rho u, \rho E$ et les $(N-1)$ équations de transport relatives aux entropies. Dans les faits, il nous est apparu plus aisé d'introduire un tel système de Relaxation pour un système compagnon à (1.5), obtenu en échangeant le rôle de l'entropie de Lax ρS_N et l'énergie totale (voir sections 1.2 et 1.3). Nous serons alors en mesure de proposer une méthode de volumes finis pour (1.5), où le découplage escompté subsiste une fois rétablis les rôles respectifs de l'énergie totale et de l'entropie ρS_N. Nous conclurons l'exposé de la méthode d'approximation en déduisant de l'ensemble de la construction proposée l'existence (à l'aide du théorème des fonctions implicites) d'un modèle de Relaxation directement consistant avec le système (1.5), et conduisant à un schéma strictement identique.

1.2 Un système compagnon

Nous proposons d'approcher les solutions faibles du système (1.1) par l'intermédiaire de l'approximation des solutions faibles du système suivant :

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + \sum_{i=1}^{N} p_i) = 0, \\ \partial_t \rho S_i + \partial_x \rho S_i u = 0, \quad i = 1, ..., N, \end{cases} \tag{1.7}$$

complété par l'inégalité d'entropie

$$\partial_t \rho E + \partial_x (\rho E + \sum_{i=1}^{N} p_i) u \leq 0. \tag{1.8}$$

Remarquons en effet que l'application $\mathbf{U} \to \rho E(\mathbf{U})$ est strictement convexe en conséquence des hypothèses thermodynamiques précédentes.
Contrairement à (1.5), ce système privilégie la conservation de l'entropie ρS_N au détriment de l'énergie totale ρE. Ainsi, ρE sera considérée comme une fonction des inconnues du système et le couple $(\rho E, (\rho E + \sum_{i=1}^N p_i)u)$ jouera le rôle d'un couple d'entropie de Lax du système. Notons que si le système (1.1) et le système compagnon (1.7) possèdent les mêmes solutions régulières, leurs solutions choc respectives sont évidemment différentes. Une technique de correction sera alors nécessaire : proposée dans [32], elle rétablira sans calcul supplémentaire l'énergie totale et l'entropie ρS_N dans leur rôle respectif.
Nous aurons besoin dans la suite des notations supplémentaires suivantes. En posant $\mathbf{U} :=^T (\rho, \rho u, \{\rho S_i\}_{i=1,\dots,N})$, l'espace des états admissibles du système (1.7) est

$$\Omega = \{\mathbf{U} \in I\!R^{N+2}, \rho > 0, \{\rho S_i > 0\}_{i=1,\dots,N}\}. \tag{1.9}$$

Pour simplifier les notations, nous introduisons la forme condensée suivante pour le système (1.7) :

$$\partial_t \mathbf{U} + \partial_x \mathbf{F}(\mathbf{U}) = 0, \tag{1.10}$$

où $\mathbf{F}$ trouve une définition naturelle.

1.3 Un modèle de Relaxation pour le système compagnon

Notre objectif est de proposer une méthode d'approximation des solutions faibles du système (1.10) permettant pour un faible coût d'évaluation de respecter les contraintes mises en avant au début de ce chapitre. Il est connu que les non linéarités du système considéré rendent difficilement conciliables l'ensemble de ces exigences dans le contexte de méthodes classiques (voir [57] pour une revue). Nous nous proposons ici d'étendre à notre problématique une stratégie récente reposant sur l'approximation des solutions concernées par celles d'un autre système, plus grand, mais par construction aux non linéarités moindres. Ce type de système, dit système de Relaxation, est censé restaurer dans la limite d'un paramètre de Relaxation infini, non seulement les équations initiales, mais également la condition d'entropie associée. Cette technique a été motivée par les travaux pionniers de Liu [101], puis ceux de Chen, Levermore et Liu [28] et a été appliquée dans un premier temps par Jin et Xin [74] pour des raisons d'approximation numérique. Dans cette approche, toutes les non linéarités du système initial sont relaxées et une telle démarche conduit à définir une méthode de type Lax-Friedrichs connue pour être trop dissipative dans la pratique. Des travaux plus récents, dévolus aux équations d'Euler (voir par exemple Coquel et al. [32]) ont montré qu'une Relaxation sélective des non linéarités conduit à une méthode de faible coût d'approximation tout en conservant de bonnes propriétés de résolution (capture exacte des contacts stationnaires). Motivée par les travaux de Suliciu [131], cette approche suggère de relaxer uniquement la définition de la loi de pression (voir également Coquel et Perthame [39] pour une Relaxation de l'énergie). Elle donne une interprétation complémentaire des schémas de Desprès [43] dans le cadre Lagrangien (voir également Gallice [51]) et a été depuis également analysée par Bouchut [18] dans un cadre plus général. Elle peut être comprise comme une justification d'une méthode due à Toro [136] indépendamment proposée par Cambier, Veuillot et Viviand [144] dès 1978.

Le principe est ici de proposer un système de Relaxation dont tous les champs sont linéairement dégénérés (résolution immédiate du problème de Riemann) et possédant une entropie de Lax coïncidant avec l'énergie totale dans la limite d'un paramètre de Relaxation infini. On montre que cette entropie vérifie un équivalent du théorème H de la théorie cinétique des gaz. Étendant de manière naturelle l'approche proposée dans [32] au présent contexte des équations d'Euler multi-pression, nous suggérons de relaxer la somme des pressions partielles $p = \sum_{i=1}^{N} p_i$, au profit d'une nouvelle variable Π, assujettie à sa propre loi d'évolution. Le système relaxé obtenu trouve la forme suivante :

$$\left\{\begin{array}{l} \partial_t\rho + \partial_x\rho u = 0, \\ \partial_t\rho u + \partial_x(\rho u^2 + \Pi) = 0, \\ \partial_t\rho S_i + \partial_x\rho S_i u = 0, \quad i = 1, ..., N, \\ \partial_t\Pi + u\partial_x\Pi + a^2\tau\partial_x u = \lambda(\sum_{i=1}^{N} p_i(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i) - \Pi), \end{array}\right. \tag{1.11}$$

où l'on a posé

$$I = \Pi + a^2\tau, \quad \tau = \frac{1}{\rho}.$$

La constante réelle a représente un paramètre du modèle de Relaxation devant satisfaire des conditions de stabilité explicitées ultérieurement (condition de Whitham). L'application

$$\begin{array}{rcl} (I\!R_+^\star)^{N+1} & \rightarrow & I\!R_+^\star \\ (\{S_i\}_{i=1,...,N}, I) & \mapsto & \mathcal{T}(\{S_i\}_{i=1,...,N}, I) \end{array}$$

sera également précisée par la suite, mais soulignons dès à présent qu'elle vérifiera par construction la propriété d'équivalence

$$\Pi = \sum_{i=1}^{N} p_i(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i) \Leftrightarrow \mathcal{T}(\{S_i\}_{i=1,...,N}, I) = \tau. \tag{1.12}$$

Nous remarquons formellement que dans la limite du paramètre de Relaxation λ ($\lambda > 0$) infini, limite que nous appelons comme il est d'usage "équilibre", nous retrouvons dans (1.11) le système initial (1.7) puisqu'alors $\Pi = \sum_{i=1}^{N} p_i(\tau, S_i)$. En ce sens, les solutions du système (1.11) réalisent une approximation des solutions du système d'équilibre (1.7).
Le système relaxé (1.11) peut être réécrit sous forme complètement conservative :

$$\left\{\begin{array}{l} \partial_t\rho + \partial_x\rho u = 0, \\ \partial_t\rho u + \partial_x(\rho u^2 + \Pi) = 0, \\ \partial_t\rho S_i + \partial_x\rho S_i u = 0, \quad i = 1, ..., N, \\ \partial_t\rho I + \partial_x\rho I u = \lambda\rho(\sum_{i=1}^{N} p_i(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i) - \Pi). \end{array}\right. \tag{1.13}$$

Cette forme, valable pour les solutions régulières, sera privilégiée dans la suite. En posant $\overline{\mathbf{U}} =^T (\rho, \rho u, \{\rho S_i\}_{i=1,...,N}, \rho I)$, l'espace des états admissibles de ce système est

$$\overline{\Omega} = \{\overline{\mathbf{U}} \in I\!R^{N+3}, \rho > 0, \{\rho S_i > 0\}_{i=1,...,N}, \rho I > 0\}.$$

Nous introduisons la forme condensée suivante pour l'opérateur du premier ordre extrait de (1.13) :

$$\partial_t\overline{\mathbf{U}} + \partial_x\overline{\mathbf{F}}(\overline{\mathbf{U}}) = 0, \tag{1.14}$$

où $\overline{\mathbf{F}}$ trouve une définition immédiate.

Proposition 1

Supposons $a > 0$. Le système du premier ordre (1.14) est hyperbolique sur $\overline{\Omega}$. Il admet les valeurs propres

$$\lambda_1 = u - a\tau = u - \frac{a}{\rho}, \quad \lambda_2 = u, \quad \lambda_3 = u + a\tau = u + \frac{a}{\rho}.$$

La valeur propre λ_2 est de multiplicité $N+1$ et les valeurs propres λ_1 et λ_3 sont d'ordre 1. Ces valeurs propres sont toutes associées à des champs linéairement dégénérés.

Démonstration

La preuve de cet énoncé résulte de calculs simples laissés au lecteur. Notre objectif est ici de vérifier que la propriété de dégénérescence linéaire des champs associés aux valeurs propres $\lambda_i(\overline{\mathbf{U}})$ pour $i = 1$ et $i = 3$ ne vaut que pour des vitesses du son $c = c(\rho, \{S_i\}_{i=1,\ldots,N})$ de la forme $c = a/\rho$ où a est une constante réelle positive (a ne peut pas être nul sous peine de perdre l'existence d'une base de diagonalisation). En privilégiant le système de variables admissibles $(\rho, u, \{S_i\}_{i=1,\ldots,N})$, il suffit de remarquer que dans le cas d'une thermodynamique générale, un vecteur propre associé à la valeur propre $\lambda_\epsilon(\overline{\mathbf{U}}) = u + \epsilon c$ (avec $\epsilon = -1$ ou $\epsilon = 1$) est donné par $r_\epsilon(u) = (1, \epsilon c/\rho, \{0\}_{i=1,\ldots,N})$. Dès lors, exiger $\nabla_{\overline{\mathbf{U}}}\lambda_\epsilon(\overline{\mathbf{U}}).r_\epsilon(\overline{\mathbf{U}}) = \epsilon(\partial_\rho c + c/\rho) = 0$ pour tout état $\overline{\mathbf{U}}$ impose la propriété recherchée. Ceci justifie la technique de Relaxation introduite.

En conséquence de la propriété de dégénérescence linéaire de tous les champs, la solution du problème de Riemann du système (1.14) est explicitement connue, comme l'illustre le théorème suivant dont le résultat étend de manière naturelle celui obtenu au titre des équations d'Euler usuelles (voir Coquel et al. [32]).

Théorème 1

Soient $\overline{\mathbf{U}}_L$ et $\overline{\mathbf{U}}_R$ deux états constants de $\overline{\Omega}$.
Supposons que la constante a vérifie la condition

$$\begin{gathered}\lambda_1(\overline{\mathbf{U}}_L) = u_L - a\tau_L < u^\star < \lambda_3(\overline{\mathbf{U}}_R) = u_R + a\tau_R, \\ u^\star = \tfrac{1}{2}(u_L + u_R) - \tfrac{1}{2a}(\Pi_R - \Pi_L), \\ \Pi_L = I_L - a^2\tau_L, \quad \Pi_R = I_R - a^2\tau_R.\end{gathered} \tag{1.15}$$

Alors, la solution faible auto-semblable $\overline{\mathbf{U}}(x,t;\overline{\mathbf{U}}_L,\overline{\mathbf{U}}_R) := \overline{\mathbf{U}}(x/t;\overline{\mathbf{U}}_L,\overline{\mathbf{U}}_R)$ du système (1.14) pour la donnée initiale

$$\overline{\mathbf{U}}_0(x) = \begin{cases} \overline{\mathbf{U}}_L & si \quad x < 0, \\ \overline{\mathbf{U}}_R & si \quad x > 0, \end{cases} \tag{1.16}$$

est constituée de quatre états constants :

$$\overline{\mathbf{U}}(x/t;\overline{\mathbf{U}}_L,\overline{\mathbf{U}}_R) = \begin{cases} \overline{\mathbf{U}}_L & si \quad \frac{x}{t} < \lambda_1(\overline{\mathbf{U}}_L), \\ \overline{\mathbf{U}}_1 & si \quad \lambda_1(\overline{\mathbf{U}}_L) < \frac{x}{t} < \lambda_2(\overline{\mathbf{U}}_1), \\ \overline{\mathbf{U}}_2 & si \quad \lambda_2(\overline{\mathbf{U}}_2) < \frac{x}{t} < \lambda_3(\overline{\mathbf{U}}_R), \\ \overline{\mathbf{U}}_R & si \quad \lambda_3(\overline{\mathbf{U}}_R) < \frac{x}{t}, \end{cases}$$

avec

$$\lambda_2(\overline{\mathbf{U}}_1) = \lambda_2(\overline{\mathbf{U}}_2) = u^\star.$$

En posant

$$\begin{aligned}
&u_1 = u_2 = u^\star,\\
&\rho_1 = a/(u_2 - \lambda_1(\overline{\mathbf{U}}_L)),\\
&\rho_2 = a/(\lambda_3(\overline{\mathbf{U}}_R) - u_2),\\
&\Pi^\star = \Pi_1 = \Pi_2 = \tfrac{1}{2}(\Pi_L + \Pi_R) - \tfrac{a}{2}(u_R - u_L),\\
&(S_i)_1 = (S_i)_L, \quad 1 \leq i \leq N,\\
&(S_i)_2 = (S_i)_R, \quad 1 \leq i \leq N,\\
&I_1 = \Pi_1 + \tfrac{a^2}{\rho_1} = I_L,\\
&I_2 = \Pi_2 + \tfrac{a^2}{\rho_2} = I_R,
\end{aligned}$$

les états constants sont caractérisés de la façon suivante :

$$\overline{\mathbf{U}}_1 = \begin{pmatrix} \rho_1 \\ (\rho u)_1 \\ (\rho S_1)_1 \\ \vdots \\ (\rho S_N)_1 \\ (\rho I)_1 \end{pmatrix}, \quad \overline{\mathbf{U}}_2 = \begin{pmatrix} \rho_2 \\ (\rho u)_2 \\ (\rho S_1)_2 \\ \vdots \\ (\rho S_N)_2 \\ (\rho I)_2 \end{pmatrix}.$$

De plus, $\overline{\mathbf{U}}_1$ *et* $\overline{\mathbf{U}}_2$ *sont dans* $\overline{\Omega}$.

Démonstration

Les champs étant tous linéairement dégénérés, la solution du problème de Riemann considéré est systématiquement composée de discontinuités se déplaçant à la vitesse caractéristique du champ concerné et séparant au plus quatre états constants : $\overline{\mathbf{U}}_L$, $\overline{\mathbf{U}}_1$, $\overline{\mathbf{U}}_2$, $\overline{\mathbf{U}}_R$. L'examen des relations de saut suffit à conclure en remarquant que la condition (1.15) portant sur la constante a assure que les trois ondes composant la solution du problème de Riemann sont toujours ordonnées selon le sens naturel :

$$\sigma_1 = \lambda_1(\overline{\mathbf{U}}_L) < \sigma_2 = u^\star < \sigma_3 = \lambda_3(\overline{\mathbf{U}}_R).$$

En conséquence des formules définissant les densités intermédiaires ρ_1 et ρ_2, ces dernières sont par construction positives dès que a vérifie (1.15).

Nous verrons plus tard que par construction le modèle de Relaxation préserve les discontinuités de contact stationnaires du modèle d'équilibre et ce quelque soit le choix de a. Toutefois, il est connu d'après les travaux de Liu [101] et de Chen, Levermore et Liu [28] (voir également Bouchut [18]), que a, paramètre du modèle de Relaxation, doit être astreint à vérifier certaines conditions indispensables à la stabilité de la procédure de Relaxation considérée dans le régime asymptotique $\lambda = +\infty$. Ces conditions, dites conditions sous-caractéristiques ou d'entrelacements, expriment dans notre contexte que la vitesse du son a/ρ du système relaxé doit être nécessairement plus grande que la vitesse du son c associée au modèle d'équilibre. Nous retrouverons cette condition dans le paragraphe suivant dévolu à l'analyse de la consistance de la condition d'entropie avec le procédé de Relaxation proposé.
Pour conclure cette section, nous proposons de décrire brièvement la procédure numérique que nous développerons par la suite. Cette procédure est classique dans le

contexte des méthodes de Relaxation (voir par exemple Jin et Xin [74]). La solution du problème de Riemann de l'opérateur du premier ordre extrait du système relaxé étant explicitement connue, le schéma numérique est basé sur la méthode de Godunov. Supposons donc connue à l'instant t^n une solution discrète

$$\mathbf{U}_{\Delta x}(x,t^n) = {}^T(\rho_{\Delta x}(x,t^n), (\rho u)_{\Delta x}(x,t^n), \{(\rho S_i)_{\Delta x}(x,t^n)\}_{i=1,...N}), \tag{1.17}$$

de (1.7). De manière classique, nous supposons cette solution constante sur chaque intervalle $[x_{j-1/2}, x_{j+1/2}[$ avec $x_{j+1/2} = (j+1/2)\Delta x$. Δx représente le pas de discrétisation en espace, et j parcourt l'ensemble des entiers relatifs.
L'évolution en temps de cette approximation jusqu'à la date t^{n+1} se fait alors en deux étapes :

Premier pas : Evolution en temps $(t^n \to t^{n+1-})$
Lors de ce premier pas, nous résolvons pour des temps suffisamment petits le système (1.14) avec la donnée initiale à l'équilibre suivante, déduite de $\mathbf{U}_{\Delta x}(x,t^n)$:

$$\overline{\mathbf{U}}_{\Delta x}(x,t^n) = {}^T(\rho_{\Delta x}(x,t^n), (\rho u)_{\Delta x}(x,t^n), \{(\rho S_i)_{\Delta x}(x,t^n)\}_{i=1,...N}, (\rho I)_{\Delta x}(x,t^n)), \tag{1.18}$$

avec

$$(\rho I)_{\Delta x}(x,t^n) = \rho_{\Delta x}(x,t^n)\sum_{i=1}^{N}(p_i)_{\Delta x}(x,t^n) + a^2.$$

Le choix de a sera mené à chaque pas de temps, conformément à une stratégie décrite ultérieurement. Sous la condition CFL classique

$$\frac{\Delta t}{\Delta x}\max_{i=1,3}|\lambda_i(\overline{\mathbf{U}})| \leq \frac{1}{2}, \tag{1.19}$$

et lorsque la valeur de a est choisie constante en espace, la solution exacte du problème de Cauchy considéré est constituée de la juxtaposition des solutions des problèmes de Riemann posés à chaque interface $x_{j+1/2}$. Ces problèmes de Riemann sont explicitement résolus d'après le précédent théorème. Nous utilisons donc dans la pratique une formulation à trois points de type Godunov pour obtenir par projection L^2 une solution notée $\overline{\mathbf{U}}_{\Delta x}(x,t^{n+1-})$, constante par morceaux sur chaque intervalle $[x_{j-1/2}, x_{j+1/2}[$, $j \in Z$. Cette solution réactualisée n'a aucune raison d'appartenir à la variété d'équilibre, ce qui justifie l'introduction du second pas.

Deuxième pas : Relaxation $(t^{n+1-} \to t^{n+1})$
Ce deuxième pas consiste en une projection de la solution $\overline{\mathbf{U}}_{\Delta x}(x,t^{n+1-})$, obtenue à l'issue du pas précédent, sur la variété d'équilibre $\lambda = +\infty$. Plus précisément, nous définissons

$$\overline{\mathbf{U}}_{\Delta x}(x,t^{n+1}) =$$
$${}^T(\rho_{\Delta x}(x,t^{n+1}), (\rho u)_{\Delta x}(x,t^{n+1}), \{(\rho S_i)_{\Delta x}(x,t^{n+1})\}_{i=1,...N}, (\rho I)_{\Delta x}(x,t^{n+1})), \tag{1.20}$$

avec

$$\begin{aligned}
&\rho_{\Delta x}(x,t^{n+1}) = \rho_{\Delta x}(x,t^{n+1-}),\\
&(\rho u)_{\Delta x}(x,t^{n+1}) = (\rho u)_{\Delta x}(x,t^{n+1-}),\\
&(\rho S_i)_{\Delta x}(x,t^{n+1}) = (\rho S_i)_{\Delta x}(x,t^{n+1-}), \quad i = 1,...,N,
\end{aligned}$$

et

$$(\rho I)_{\Delta x}(x,t^{n+1}) = \rho_{\Delta x}(x,t^{n+1-}) \sum_{i=1}^N (p_i(\tau,S_i))_{\Delta x}(x,t^{n+1-}) + a^2.$$

Ces formules sont équivalentes à résoudre le système d'équations différentielles ordinaires suivant pour $\lambda \to +\infty$,

$$\begin{cases} \partial_t \rho = 0, \\ \partial_t \rho u = 0, \\ \partial_t \rho S_i = 0, \quad i = 1,...,N, \\ \partial_t \rho I = \lambda \rho (\sum_{i=1}^N p_i(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i) - \Pi), \end{cases} \tag{1.21}$$

avec comme donnée initiale la solution $\overline{\mathbf{U}}_{\Delta x}(x,t^{n+1-})$.
Pour achever la description de l'algorithme, nous posons :

$$\mathbf{U}_{\Delta x}(x,t^{n+1}) =^T (\rho_{\Delta x}(x,t^{n+1}), (\rho u)_{\Delta x}(x,t^{n+1}), \{(\rho S_i)_{\Delta x}(x,t^{n+1})\}_{i=1,...N}). \tag{1.22}$$

Cet algorithme peut donc être compris comme une stratégie de splitting d'opérateurs appliquée au système (1.13), en considérant dans un premier pas les phénomènes convectifs puis dans un second pas le terme source de Relaxation dans la limite $\lambda = +\infty$.

1.4 Entropie de Lax et premières propriétés

En vue de justifier rigoureusement le bien-fondé de l'algorithme décrit ci-dessus, notre objectif est maintenant de construire un couple d'entropie de Lax $(\rho\Sigma, \rho\Sigma u + \Pi u)$ pour le système relaxé (1.14), coïncidant avec le couple d'entropie de Lax $(\rho E, \rho E u + \sum_{i=1}^N p_i u)$ du système d'équilibre (1.7) dans le régime d'un paramètre de Relaxation infini. De manière à proposer un relèvement du couple d'entropie de Lax à l'équilibre, il convient dans un premier temps de définir le relèvement $\mathcal{T}(\{S_i\}_{i=1,...,N}, I)$ du covolume τ intervenant dans la condition d'équilibre (1.12). Son obtention repose sur le résultat suivant.

Lemme 1
Pour tous $\{S_i > 0\}_{i=1,...,N}$ et $I > 0$ fixés, l'application f définie par

$$\begin{array}{rcl} I\!R_+^\star & \to & I\!R \\ \tau & \mapsto & f(\tau) = \sum_{i=1}^N p_i(\tau, S_i) + a^2\tau - I, \end{array} \tag{1.23}$$

est convexe en τ. De plus,

$$\lim_{\tau\to+\infty} f(\tau) = +\infty, \quad \lim_{\tau\to 0} f(\tau) = +\infty.$$

Démonstration
Chacune des pressions p_i étant par hypothèse convexe par rapport à sa première variable, la convexité de l'application f est immédiate. Par ailleurs, la positivité des pressions implique l'inégalité $f(\tau) \geq a^2\tau - I$. Ainsi, $\lim_{\tau\to+\infty} f(\tau) = +\infty$. La limite en 0 est quant à elle déduite de l'hypothèse asymptotique $\lim_{\tau\to 0} p_i(\tau, S_i) = +\infty$ pour tout $i = 1,...,N$.

En conséquence de ce résultat, pour tous $\{S_i > 0\}_{i=1,\ldots,N}$ et $I > 0$ fixés, l'application f définie par (1.23) s'annule au plus deux fois sur $\mathbb{R}_+^\star$. Par construction, nous serons dans la suite systématiquement en présence de deux racines distinctes. Par définition, nous distinguerons la plus grande de ces racines en la notant $\mathcal{T}(\{S_i\}_{i=1,\ldots,N}, I)$. Elle vérifie :

$$\begin{cases} f(\mathcal{T}(\{S_i\}_{i=1,\ldots,N}, I)) = 0, \quad \mathcal{T}(\{S_i\}_{i=1,\ldots,N}, I) > 0, \\ \sum_{i=1}^N \frac{\partial p_i}{\partial \tau}(\mathcal{T}(\{S_i\}_{i=1,\ldots,N}, I), S_i) + a^2 > 0. \end{cases} \tag{1.24}$$

Il importe de remarquer que lorsque la variable I est prise à l'équilibre, c'est-à-dire lorsque $I = \sum_{i=1}^N p_i(\tau, S_i) + a^2\tau$, et que la constante a est choisie suffisamment grande de sorte à satisfaire l'inégalité suivante (dite condition de Whitham) :

$$\sum_{i=1}^N \frac{\partial p_i}{\partial \tau}(\tau, S_i) + a^2 > 0,$$

nous avons par définition $\mathcal{T}(\{S_i\}_{i=1,\ldots,N}, I) = \tau$. La condition de Whitham exprime la condition d'entrelacement sur les valeurs propres rapportée dans la section précédente. Munis de cette nouvelle définition, nous considérons l'application

$$(\tau, u, \{S_i\}_{i=1,\ldots,N}, I) \mapsto \Sigma(\tau, u, \{S_i\}_{i=1,\ldots,N}, I),$$

en posant

$$\Sigma(\tau, u, \{S_i\}_{i=1,\ldots,N}, I) = \frac{1}{2}u^2 + \theta(\tau, \{S_i\}_{i=1,\ldots,N}, I), \tag{1.25}$$

$$\theta(\tau, \{S_i\}_{i=1,\ldots,N}, I) = \frac{\Pi^2}{2a^2} + \phi(\{S_i\}_{i=1,\ldots,N}, I), \quad \Pi = I - a^2\tau, \tag{1.26}$$

et

$$\phi(\{S_i\}_{i=1,\ldots,N}, I) = \sum_{i=1}^N \varepsilon_i(\mathcal{T}(\{S_i\}_{i=1,\ldots,N}, I), S_i) - \frac{(\sum_{i=1}^N p_i(\mathcal{T}(\{S_i\}_{i=1,\ldots,N}, I), S_i))^2}{2a^2}. \tag{1.27}$$

Avec un léger abus de notation, nous définissons également l'application

$$(\rho, \rho u, \{\rho S_i\}_{i=1,\ldots,N}, \rho I) \mapsto \rho\Sigma(\rho, \rho u, \{\rho S_i\}_{i=1,\ldots,N}, \rho I),$$

en posant

$$\rho\Sigma(\rho, \rho u, \{\rho S_i\}_{i=1,\ldots,N}, \rho I) = \rho\Sigma(\tau, u, \{S_i\}_{i=1,\ldots,N}, I), \quad \tau = \frac{1}{\rho}. \tag{1.28}$$

Remarquons que dès que $\mathcal{T}(\{S_i\}_{i=1,\ldots,N}, I) = \tau$ (*i.e.* lorsque I est à l'équilibre) et lorsque a est choisi suffisamment grand (voir ci-dessus), il est clair que

$$\rho\Sigma(\rho, \rho u, \{\rho S_i\}_{i=1,\ldots,N}, \rho I) = \frac{(\rho u)^2}{2\rho} + \sum_{i=1}^N \rho\varepsilon_i(\tau, S_i) = \rho E.$$

A l'équilibre, l'énergie relaxée $\rho\Sigma$ coïncide donc avec l'énergie totale ρE.
Le résultat suivant exprime le fait que cette énergie relaxée est bien compatible au

sens de la théorie proposée par Chen, Levermore et Liu [28] avec la procédure de Relaxation proposée.

Théorème 2

(i) *Les solutions régulières du système (1.14) (i.e. lorsque $\lambda = 0$) vérifient sans condition supplémentaire la loi de conservation suivante sur $\rho\Sigma$:*

$$\partial_t \rho\Sigma + \partial_x(\rho\Sigma + \Pi)u = 0. \tag{1.29}$$

(ii) *Les solutions régulières du système (1.13) (i.e. avec $\lambda > 0$) vérifient sans condition supplémentaire la loi d'évolution suivante sur $\rho\Sigma$:*

$$\partial_t \rho\Sigma + \partial_x(\rho\Sigma + \Pi)u = -\frac{\lambda}{a^2}\rho(\sum_{i=1}^{N} p_i(\mathcal{T}(\{S_i\}_{i=1,\dots,N}, I), S_i) - \Pi)^2 \leq 0. \tag{1.30}$$

Démonstration

La démonstration de ce théorème s'obtient en manipulant les équations des différents systèmes considérés.

Dans un premier temps, les lois de conservation de la masse et de la quantité de mouvement conduisent à

$$\partial_t u + u\partial_x u + \tau\partial_x \Pi = 0, \quad \tau = \frac{1}{\rho}.$$

Une multiplication par u implique

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

équivalent en vertu de la conservation de la masse à :

$$\partial_t \rho \frac{u^2}{2} + \partial_x \rho \frac{u^3}{2} + u\partial_x \Pi = 0. \tag{1.31}$$

D'autre part, les lois de conservation de la masse et d'évolution de la variable I du système (1.13) conduisent successivement à

$$\partial_t \Pi + u\partial_x \Pi + a^2\tau\partial_x u = \lambda(\sum_{i=1}^{N} p_i(\mathcal{T}(\{S_i\}_{i=1,\dots,N}, I), S_i) - \Pi),$$

et

$$\partial_t \rho \frac{\Pi^2}{2a^2} + \partial_x \rho u \frac{\Pi^2}{2a^2} + \Pi\partial_x u = \lambda\rho\frac{\Pi}{a^2}(\sum_{i=1}^{N} p_i(\mathcal{T}(\{S_i\}_{i=1,\dots,N}, I), S_i) - \Pi). \tag{1.32}$$

Enfin, il est clair que pour toute fonction $\phi(\{S_i\}_{i=1,\dots,N}, I)$ régulière, nous avons

$$\partial_t \rho\phi(\{S_i\}_{i=1,\dots,N}, I) + \partial_x \rho\phi(\{S_i\}_{i=1,\dots,N}, I)u =$$

$$\lambda\rho\frac{\partial \phi}{\partial I}(\{S_i\}_{i=1,\dots,N}, I)(\sum_{i=1}^{N} p_i(\mathcal{T}(\{S_i\}_{i=1,\dots,N}, I), S_i) - \Pi), \tag{1.33}$$

d'après la validité des équations

$$\begin{cases} \partial_t\rho + \partial_x\rho u = 0, \\ \partial_t\rho S_i + \partial_x\rho S_i u = 0, \quad i = 1, ..., N, \\ \partial_t\rho I + \partial_x\rho I u = \lambda\rho(\sum_{i=1}^N p_i(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i) - \Pi). \end{cases}$$

L'addition des équations (1.31)-(1.32)-(1.33) permet d'écrire

$$\partial_t\rho\Sigma + \partial_x(\rho\Sigma + \Pi)u =$$

$$\lambda\rho(\sum_{i=1}^N p_i(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i) - \Pi) \times (\frac{\partial\phi}{\partial I}(\{S_i\}_{i=1,...,N}, I) + \frac{\Pi}{a^2}).$$

Afin de préciser le second membre de cette équation, remarquons que par définition de $\mathcal{T}(\{S_i\}_{i=1,...,N}, I)$, nous avons

$$\frac{\partial\mathcal{T}}{\partial I} = \frac{1}{\sum_{i=1}^N \frac{\partial p_i}{\partial\tau}(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i) + a^2}. \tag{1.34}$$

Par dérivation de (1.27), nous obtenons

$$\frac{\partial\phi}{\partial I} = \frac{\partial\mathcal{T}}{\partial I} \times (\sum_{i=1}^N \frac{\partial\varepsilon_i}{\partial\tau}(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i) -$$

$$\frac{\sum_{i=1}^N p_i(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i) \times \sum_{i=1}^N \frac{\partial p_i}{\partial\tau}(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i)}{a^2}),$$

c'est-à-dire

$$\frac{\partial\phi}{\partial I} = -\frac{\sum_{i=1}^N p_i(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i)}{a^2} \times$$

$$\frac{\partial\mathcal{T}}{\partial I} \times (\sum_{i=1}^N \frac{\partial p_i}{\partial\tau}(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i) + a^2),$$

puisque par définition,

$$p_i(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i) = -\frac{\partial\varepsilon_i}{\partial\tau}(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i).$$

D'après (1.34), nous obtenons :

$$\frac{\partial\phi}{\partial I} = -\frac{\sum_{i=1}^N p_i(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i)}{a^2}, \tag{1.35}$$

de telle sorte que :

$$\partial_t\rho\Sigma + \partial_x(\rho\Sigma + \Pi)u = -\frac{\lambda}{a^2}\rho(\sum_{i=1}^N p_i(\mathcal{T}(\{S_i\}_{i=1,...,N}, I), S_i) - \Pi)^2.$$

Un raisonnement identique en posant $\lambda = 0$ permet d'obtenir pour le système (1.14)

$$\partial_t\rho\Sigma + \partial_x(\rho\Sigma + \Pi)u = 0.$$

Lemme 2
L'application Σ définie par (1.25) est strictement convexe en I.

Démonstration
Par définition, nous avons

$$\frac{\partial \Sigma}{\partial I} = \frac{\partial \theta}{\partial I} = \frac{\partial}{\partial I}(\frac{\Pi^2}{2a^2} + \phi(\{S_i\}_{i=1,\dots,N}, I)),$$

avec rappelons le $\Pi = I - a^2\tau$. Ainsi, (1.35) entraîne

$$\frac{\partial \Sigma}{\partial I} = \frac{I - a^2\tau}{a^2} - \frac{\sum_{i=1}^N p_i(\mathcal{T}(\{S_i\}_{i=1,\dots,N}, I), S_i)}{a^2},$$

et, par définitions de Π et $\mathcal{T}(\{S_i\}_{i=1,\dots,N}, I)$,

$$\frac{\partial \theta}{\partial I} = \frac{\Pi - \sum_{i=1}^N p_i(\mathcal{T}(\{S_i\}_{i=1,\dots,N}, I), S_i)}{a^2} = \mathcal{T}(\{S_i\}_{i=1,\dots,N}, I) - \tau.$$

Nous obtenons

$$\frac{\partial^2 \Sigma}{\partial I^2} = \frac{\partial \mathcal{T}}{\partial I}.$$

D'après (1.34), l'application Σ est donc strictement convexe en I en vertu de (1.24).

Corollaire 1
L'application $\rho\Sigma$ définie par (1.28) est strictement convexe en ρI.

Démonstration
D'après (1.28), il est clair que

$$\frac{\partial \rho\Sigma}{\partial \rho I} = \frac{\partial \Sigma}{\partial I}, \quad \frac{\partial^2 \rho\Sigma}{\partial (\rho I)^2} = \frac{1}{\rho}\frac{\partial^2 \Sigma}{\partial I^2}.$$

L'application $\rho\Sigma$ est de ce fait strictement convexe en ρI en vertu du lemme précédent et par positivité de ρ.

De manière plus précise, nous énonçons les deux résultats suivants.
Théorème 3
L'application Σ définie par (1.25)-(1.26)-(1.27) est strictement convexe.

Corollaire 2
L'application $\rho\Sigma$ définie par (1.28) est strictement convexe.

Ces résultats n'étant pas indispensables pour la suite de nos propos, nous omettons volontairement leurs démonstrations. Nous sommes néammoins en mesure d'affirmer que par définition, le couple $(\rho\Sigma, \rho\Sigma u + \Pi u)$ est un couple d'entropie de Lax pour le système relaxé (1.14) puisque la loi de conservation (1.29) est satisfaite. De ce fait, il est important de souligner que les champs du système (1.14) étant linéairement dégénérés, les solutions faibles de ce système vérifient sans hypothèse supplémentaire la loi de conservation (1.29) sur $\rho\Sigma$ au sens des distributions.

Pour conclure cette section, nous établissons le résultat suivant comparant l'énergie $\rho\Sigma$ du système relaxé (1.14) et l'énergie ρE du système à l'équilibre (1.10).

Théorème 4
Soit $(\rho, \rho u, \{\rho S_i\}_{i=1,\ldots,N}, \rho I)$ un $(N+3)$-uplet de $\overline{\Omega}$. Supposons que la constante a vérifie la condition de Whitham

$$\sum_{i=1}^{N} \frac{\partial p_i}{\partial \tau}(\tau, S_i) + a^2 > 0, \quad \tau = 1/\rho. \tag{1.36}$$

Posons

$$I^{eq} := I^{eq}(\tau, \{S_i\}_{i=1,\ldots,N}) = \sum_{i=1}^{N} p_i(\tau, S_i) + a^2\tau. \tag{1.37}$$

Alors sur son domaine de définition l'énergie relaxée $\rho\Sigma$ est convexe en ρI et atteint son minimum sur la position d'équilibre ρI^{eq} :

$$\rho\Sigma(\rho, \rho u, \{\rho S_i\}_{i=1,\ldots,N}, \rho I^{eq}) \leq \rho\Sigma(\rho, \rho u, \{\rho S_i\}_{i=1,\ldots,N}, \rho I), \tag{1.38}$$

avec de plus

$$\rho\Sigma(\rho, \rho u, \{\rho S_i\}_{i=1,\ldots,N}, \rho I^{eq}) = \rho E(\rho, \rho u, \{\rho S_i\}_{i=1,\ldots,N}).$$

Remarquons que les identités de ce théorème sont équivalentes à

$$\Sigma(\tau, u, \{S_i\}_{i=1,\ldots,N}, I^{eq}) \leq \Sigma(\tau, u, \{S_i\}_{i=1,\ldots,N}, I), \tag{1.39}$$

avec de plus

$$\Sigma(\tau, u, \{S_i\}_{i=1,\ldots,N}, I^{eq}) = \rho E(\rho, \rho u, \{\rho S_i\}_{i=1,\ldots,N})/\rho.$$

Démonstration
En vertu de la remarque précédente, il suffit d'établir la validité de (1.39).
Rappelons que par construction, dès que $I = I^{eq}$, nous avons $\mathcal{T}(\{S_i\}_{i=1,\ldots,N}, I^{eq}) = \tau$ puisque la condition de Whitham (1.36) est satisfaite, et

$$\Sigma(\tau, u, \{S_i\}_{i=1,\ldots,N}, I^{eq}) = \frac{1}{2}u^2 + \sum_{i=1}^{N} \varepsilon_i(\tau, S_i).$$

Par ailleurs, on rappelle (voir démonstration du lemme 2) que

$$\frac{\partial \Sigma}{\partial I} = \frac{\partial \theta}{\partial I} = \mathcal{T}(\{S_i\}_{i=1,\ldots,N}, I) - \tau.$$

L'application Σ étant strictement convexe par rapport à la variable I d'après le lemme 2, son minimum en I est réalisé lorsque $\mathcal{T}(\{S_i\}_{i=1,\ldots,N}, I) = \tau$. L'application $I \to \mathcal{T}(\{S_i\}_{i=1,\ldots,N}, I)$ étant strictement croissante (voir (1.34) et la définition de $\mathcal{T}(\{S_i\}_{i=1,\ldots,N}, I)$), ce minimum ne peut être atteint que pour $I = I^{eq}$. L'inégalité (1.39) est démontrée.

1.5 Schéma par Relaxation

Nous donnons maintenant une description précise de la méthode numérique proposée en Section 1.3 pour approcher numériquement les solutions du problème de Cauchy (1.7)-(1.8), pour la donnée initiale

$$\mathbf{U}(x,0) = \mathbf{U}_0(x). \tag{1.40}$$

Soient Δt et Δx les pas de temps et d'espace choisis constants par souci de simplicité. Les solutions approchées $\mathbf{U}_{\Delta x}(x,t) : I\!R \times I\!R^+_\star \to \Omega$ du système (1.7)-(1.8)-(1.40) seront cherchées sous la forme de fonctions constantes par morceaux, telles que

$$\mathbf{U}_{\Delta x}(x,t) = \mathbf{U}_j^n = ^T(\rho_j^n, (\rho u)_j^n, \{(\rho S_i)_j^n\}_{i=1,\dots,N}) \tag{1.41}$$

si $x \in [x_{j-1/2}, x_{j+1/2}[$, $j \in Z$ et $t \in [t^n, t^{n+1}[$, avec $t^n = n\Delta t$, $n \in N$.
A l'instant $t = 0$, nous introduisons

$$\mathbf{U}_j^0 = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} {}^T(\rho_0(x), (\rho u)_0(x), \{(\rho S_i)_0(x)\}_{i=1,\dots,N})dx, \quad j \in Z. \tag{1.42}$$

Supposons connue la solution approchée $\mathbf{U}_{\Delta x}(x,.)$ à une date $t^n \geq 0$. Nous suggérons d'avancer cette solution à l'instant suivant t^{n+1} en adoptant la stratégie générale décrite précédemment mais en faisant choix d'une valeur de a définie relativement à chaque interface $x_{j+1/2}$. Notre objectif est en effet de minimiser la dissipation numérique de la méthode en choisissant la plus petite valeur de a vérifiant des conditions de Whitham locales explicitées ci-après. Cette volonté repose sur l'heuristique suivante. Un développement de Chapman-Enskog mené autour de la variété d'équilibre du présent modèle de Relaxation (voir Bouchut [18], et aussi [28], [39] pour les détails dans le cas 3×3) conduit à devoir considérer une perturbation du système d'équilibre qui s'avère être visqueuse dès que la condition de Whitham est satisfaite. La viscosité apparente croît avec a, ce qui suggère fortement de minimiser cette valeur localement. Fixons une arête $x_{j+1/2}$ et associons lui une valeur $a_{j+1/2} = a(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n)$. Relativement à cette valeur, nous définissons le système de Relaxation suivant, où le paramètre de Relaxation λ est choisi nul conformément à la description du premier pas de la méthode :

$$\partial_t \overline{\mathbf{U}} + \partial_x \overline{\mathbf{F}}_{a_{j+1/2}}(\overline{\mathbf{U}}) = 0. \tag{1.43}$$

La dépendance de la fonction flux en $a_{j+1/2}$ intervient bien évidemment dans la définition de la pression relaxée :

$$\Pi = I - a_{j+1/2}^2 \tau, \quad \tau = 1/\rho. \tag{1.44}$$

L'observation centrale aux développements qui vont suivre est la suivante. A un état donné $\mathbf{U}$ dans Ω, associons l'état relaxé $\overline{\mathbf{U}}$ dans $\overline{\Omega}$ défini par :

$$\overline{\mathbf{U}} = {}^T(\mathbf{U}, (\rho I) = \rho \sum_{i=1}^N p_i(\mathbf{U}) + a_{j+1/2}^2), \tag{1.45}$$

et évaluons la fonction flux $\overline{\mathbf{F}}_{a_{j+1/2}}$ sur cet état à l'équilibre (*i.e.* avec $\Pi(\overline{\mathbf{U}}) = \sum_{i=1}^N p_i(\mathbf{U})$) pour obtenir l'identité suivante :

$$\overline{\mathbf{F}}_{a_{j+1/2}}(\overline{\mathbf{U}}) = {}^T(\mathbf{F}(\mathbf{U}), (\rho I u) = (\rho \sum_{i=1}^N p_i(\mathbf{U}) + a_{j+1/2}^2)u). \tag{1.46}$$

En d'autres termes, les $(N+2)$ premières composantes de cette fonction flux sont indépendantes du choix du coefficient a défini relativement à l'arête $x_{j+1/2}$. Cette remarque évidente associée au choix de la restriction CFL (1.19) ($\leq 1/2$) permet, ainsi que nous le verrons, de localiser le choix de a à chaque interface tout en définissant une méthode de volumes finis conservative et consistante pour la réactualisation de $\mathbf{U}_j^{n+1}$.
Munis du système de Relaxation (1.43), formons un problème de Riemann à l'interface $x_{j+1/2}$ en considérant les données gauche et droite suivantes :

$$\overline{\mathbf{U}}_{j,+}^n = {}^T(\mathbf{U}_j^n, (\rho I)_{j,+}^n = \rho_j^n \sum_{i=1}^N p_i(\mathbf{U}_j^n) + a_{j+1/2}^2), \tag{1.47}$$

$$\overline{\mathbf{U}}_{j+1,-}^n = {}^T(\mathbf{U}_{j+1}^n, (\rho I)_{j+1,-}^n = \rho_{j+1}^n \sum_{i=1}^N p_i(\mathbf{U}_{j+1}^n) + a_{j+1/2}^2). \tag{1.48}$$

Désignons par $\overline{\mathbf{U}}_{a_{j+1/2}}((x-x_{j+1/2})/(t-t^n); \overline{\mathbf{U}}_{j,+}^n, \overline{\mathbf{U}}_{j+1,-}^n)$ la solution auto-semblable du problème de Riemann considéré, donnée conformément au théorème 1. Suivant le formalisme proposé par Harten, Lax et Van Leer [61] pour les méthodes de volumes finis, nous introduisons les projections L^2 à l'instant t^{n+1-} de cette solution, respectivement sur les demi-cellules $(x_j, x_{j+1/2})$ et $(x_{j+1/2}, x_{j+1})$:

$$\begin{aligned}
\overline{\mathbf{U}}_{j+1/2,L}^{n+1-} &= \tfrac{2}{\Delta x}\int_{x_j}^{x_{j+1/2}} \overline{\mathbf{U}}_{a_{j+1/2}}((x-x_{j+1/2})/(t^{n+1-}-t^n); \overline{\mathbf{U}}_{j,+}^n, \overline{\mathbf{U}}_{j+1,-}^n)dx,\\
\overline{\mathbf{U}}_{j+1/2,R}^{n+1-} &= \tfrac{2}{\Delta x}\int_{x_{j+1/2}}^{x_{j+1}} \overline{\mathbf{U}}_{a_{j+1/2}}((x-x_{j+1/2})/(t^{n+1-}-t^n); \overline{\mathbf{U}}_{j,+}^n, \overline{\mathbf{U}}_{j+1,-}^n)dx.
\end{aligned} \tag{1.49}$$

Des considérations très classiques permettent de réexprimer ces moyennes sous la forme suivante :

$$\begin{aligned}
\overline{\mathbf{U}}_{j+1/2,L}^{n+1-} &= \overline{\mathbf{U}}_{j,+}^n - 2\tfrac{\Delta t}{\Delta x}(\overline{\mathbf{F}}_{a_{j+1/2}}^g(\overline{\mathbf{U}}_{j,+}^n, \overline{\mathbf{U}}_{j+1,-}^n) - \overline{\mathbf{F}}_{a_{j+1/2}}(\overline{\mathbf{U}}_{j,+}^n)),\\
\overline{\mathbf{U}}_{j+1/2,R}^{n+1-} &= \overline{\mathbf{U}}_{j+1,-}^n - 2\tfrac{\Delta t}{\Delta x}(\overline{\mathbf{F}}_{a_{j+1/2}}(\overline{\mathbf{U}}_{j+1,-}^n) - \overline{\mathbf{F}}_{a_{j+1/2}}^g(\overline{\mathbf{U}}_{j,+}^n, \overline{\mathbf{U}}_{j+1,-}^n)),
\end{aligned} \tag{1.50}$$

où la fonction flux numérique impliquée trouve la définition standard :

$$\overline{\mathbf{F}}_{a_{j+1/2}}^g(\overline{\mathbf{U}}_{j,+}^n, \overline{\mathbf{U}}_{j+1,-}^n) = \overline{\mathbf{F}}_{a_{j+1/2}}(\overline{\mathbf{U}}_{a_{j+1/2}}(0^+; \overline{\mathbf{U}}_{j,+}^n, \overline{\mathbf{U}}_{j+1,-}^n)). \tag{1.51}$$

Avec ces définitions, introduisons alors la moyenne suivante définie relativement à chaque cellule $(x_{j-1/2}, x_{j+1/2})$ à l'instant t^{n+1-} :

$$\overline{\mathbf{U}}_j^{n+1-} = \frac{1}{2}(\overline{\mathbf{U}}_{j+1/2,L}^{n+1-} + \overline{\mathbf{U}}_{j-1/2,R}^{n+1-}). \tag{1.52}$$

Notons que cette définition ne peut pas être comprise comme résultant de la moyenne d'une juxtaposition de solutions de problèmes de Riemann sans interaction, même sous la condition CFL (1.19). En effet, la définition du système de Relaxation ainsi que celle des données initiales associées changent en $x = x_j$, puisque a est défini relativement à chaque interface $x_{j+1/2}$ (c'est-à-dire localement sur chaque intervalle (x_j, x_{j+1})).

Toutefois, la définition de la solution discrète à l'équilibre $\mathbf{U}_j^{n+1}$, donnée conformément au second pas de la stratégie générale à partir de la connaissance de (1.52) trouve une formulation conservative consistante au sens usuel. La valeur réactualisée à l'instant t^{n+1} de la solution discrète à l'équilibre s'écrit sur chaque cellule $(x_{j-1/2}, x_{j+1/2})$:

$$\mathbf{U}_{\Delta x}(x, t^{n+1}) = \mathbf{U}_j^{n+1} = \begin{cases} \rho_j^{n+1} = \rho_j^{n+1-}, \\ (\rho u)_j^{n+1} = (\rho u)_j^{n+1-}, \\ (\rho S_i)_j^{n+1} = (\rho S_i)_j^{n+1-}, \quad i = 1, ..., N. \end{cases} \tag{1.53}$$

Remarquons que d'après (1.52) cette mise à jour se réécrit trivialement sous la forme :

$$\mathbf{U}_j^{n+1} = \frac{1}{2}(\mathbf{U}_{j+1/2,L}^{n+1} + \mathbf{U}_{j-1/2,R}^{n+1}), \tag{1.54}$$

où par définition nous avons posé pour tout $j \in Z$:

$$\begin{aligned} \mathbf{U}_{j+1/2,L}^{n+1} &= ((\rho)_{j+1/2,L}^{n+1-}, (\rho u)_{j+1/2,L}^{n+1-}, \{(\rho S_i)_{j+1/2,L}^{n+1-}\}_{i=1,...,N}), \\ \mathbf{U}_{j+1/2,R}^{n+1} &= ((\rho)_{j+1/2,R}^{n+1-}, (\rho u)_{j+1/2,R}^{n+1-}, \{(\rho S_i)_{j+1/2,R}^{n+1-}\}_{i=1,...,N}). \end{aligned} \tag{1.55}$$

Or, l'examen des formules de demi-moyennes (1.55) rentrant dans la définition (1.54) conduit à considérer

$$\overline{\mathbf{F}}_{a_{j-1/2}}(\overline{\mathbf{U}}_{j,-}^n) = {}^T(\mathbf{F}(\mathbf{U}_j^n), (\rho I u)_{j,-}^n), \tag{1.56}$$

et

$$\overline{\mathbf{F}}_{a_{j+1/2}}(\overline{\mathbf{U}}_{j,+}^n) = {}^T(\mathbf{F}(\mathbf{U}_j^n), (\rho I u)_{j,+}^n), \tag{1.57}$$

en vertu de l'identité (1.46). Ces considérations conduisent immédiatement à réexprimer la valeur réactualisée $\mathbf{U}_j^{n+1}$ donnée en (1.54) selon la formulation conservative suivante :

$$\mathbf{U}_j^{n+1} = \mathbf{U}_j^n - \frac{\Delta t}{\Delta x}(\mathbf{F}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n) - \mathbf{F}(\mathbf{U}_{j-1}^n, \mathbf{U}_j^n)), \tag{1.58}$$

où le flux numérique $\mathbf{F}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n) \in I\!R^{N+2}$ est défini pour tout $j \in Z$ par (1.51) :

$$\mathbf{F}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n) =^T (\overline{F}_\rho^g, \overline{F}_{\rho u}^g, \{\overline{F}_{\rho S_i}^g\}_{i=1,...,N})_{a_{j+1/2}}(\overline{\mathbf{U}}_{j,+}^n, \overline{\mathbf{U}}_{j+1,-}^n). \tag{1.59}$$

L'algorithme permettant de définir la suite $(\mathbf{U}_j^{n+1})_{j\in Z}$ rentre donc dans le cadre des schémas explicites à trois points sous forme conservative.
Nous montrons maintenant que le présent algorithme proposé pour approcher numériquement les solutions de (1.7)-(1.8)-(1.40) partage avec la méthode de Godunov exacte (*i.e.* appliquée directement sur le système (1.7) et non sur le système relaxé (1.14)) bon nombre des propriétés de stabilité et de précision attendues. En effet, les solutions des problèmes de Riemann $\overline{\mathbf{U}}_{a_{j+1/2}}((x - x_{j+1/2})/(t - t^n); \overline{\mathbf{U}}_{j,+}^n, \overline{\mathbf{U}}_{j+1,-}^n)$ vérifient aisément pour tout $j \in Z$ les principes du maximum suivants (voir théorème 1) :

$$\min((S_i)_j^n, (S_i)_{j+1}^n) \leq S_i((x - x_{j+1/2})/(t - t^n); \overline{\mathbf{U}}_{j,+}^n, \overline{\mathbf{U}}_{j+1,-}^n) \leq \max((S_i)_j^n, (S_i)_{j+1}^n),$$

$$\rho((x - x_{j+1/2})/(t - t^n); \overline{\mathbf{U}}_{j,+}^n, \overline{\mathbf{U}}_{j+1,-}^n) > 0,$$

de telle sorte que la mise à jour (1.53) entraîne les principes du maximum

$$\min((S_i)^n_{j-1},(S_i)^n_j,(S_i)^n_{j+1}) \leq (S_i)^{n+1}_j \leq \max((S_i)^n_{j-1},(S_i)^n_j,(S_i)^n_{j+1}), \tag{1.60}$$

pour tout $j \in Z$ et tout $i = 1, ..., N$. Nous montrons ci-dessous que l'algorithme définit un schéma entropique au sens des volumes finis (voir l'inégalité d'entropie (1.62) sur ρE). La seule propriété de précision perdue concerne la capture exacte des solutions ondes de choc stationnaires (la condition de Whitham s'y oppose) alors que les contacts stationnaires sont exactement résolus. Dans l'énoncé qui suit ainsi que dans les développements ultérieurs, nous omettrons la dépendance du coefficient a relativement à l'indice d'arête, lorsque cela ne prête pas à confusion.

Théorème 5
Supposons que la constante $a := a_{j+1/2}(\mathbf{U}^n_j, \mathbf{U}^n_{j+1})$ choisie localement à chaque interface $x_{j+1/2}$ et le couple $(\mathbf{U}_L, \mathbf{U}_R) := (\mathbf{U}^n_j, \mathbf{U}^n_{j+1})$ satisfont pour tout $j \in Z$ la condition (1.15) et les conditions de Whitham suivantes :

$$\begin{aligned} &\sum_{i=1}^N \frac{\partial p_i}{\partial \tau}(\tau_L,(S_i)_L) + a^2 > 0, \quad \sum_{i=1}^N \frac{\partial p_i}{\partial \tau}(\tau_1,(S_i)_L) + a^2 > 0,\\ &\sum_{i=1}^N \frac{\partial p_i}{\partial \tau}(\tau_R,(S_i)_R) + a^2 > 0, \quad \sum_{i=1}^N \frac{\partial p_i}{\partial \tau}(\tau_2,(S_i)_R) + a^2 > 0, \end{aligned} \tag{1.61}$$

où nous avons utilisé les définitions du théorème 1 pour $\tau_1 = 1/\rho_1$ et $\tau_2 = 1/\rho_2$. Alors, sous la condition CFL (1.19), le schéma numérique (1.58) satisfait les propriétés suivantes :

(i) Propriétés de stabilité :
- *Stabilité L^1 : $\mathbf{U}^{n+1}_j \in \Omega$ pour tout $j \in Z$.*
- *Inégalité d'entropie :*

$$\begin{aligned} (\rho E)^{n+1}_j := \rho E(\mathbf{U}^{n+1}_j) \leq \\ (\rho E)^n_j - \frac{\Delta t}{\Delta x}(F_{\rho\Sigma}(\mathbf{U}^n_j,\mathbf{U}^n_{j+1}) - F_{\rho\Sigma}(\mathbf{U}^n_{j-1},\mathbf{U}^n_j)), \quad j \in Z, \end{aligned} \tag{1.62}$$

où le flux numérique d'énergie est donné par la fonction flux consistante :

$$F_{\rho\Sigma}(\mathbf{U}^n_j,\mathbf{U}^n_{j+1}) = \{(\rho\Sigma + \Pi)u\}_{a_{j+1/2}}(\overline{\mathbf{U}}_{a_{j+1/2}}(0^+;\overline{\mathbf{U}}^n_{j,+},\overline{\mathbf{U}}^n_{j+1,-})). \tag{1.63}$$

- *Principes du maximum :*

$$\begin{aligned} \min((S_i)^n_{j-1},(S_i)^n_j,(S_i)^n_{j+1}) \leq (S_i)^{n+1}_j \leq \max((S_i)^n_{j-1},(S_i)^n_j,(S_i)^n_{j+1}),\\ i = 1, ..., N, \quad j \in Z. \end{aligned} \tag{1.64}$$

(ii) Propriété de précision :
- *Les discontinuités de contact stationnaires du système (1.7) sont exactement résolues.*

L'inégalité (1.62) exprime la consistance du schéma numérique proposé avec l'inégalité d'entropie (1.8).
Remarquons que les covolumes intermédiaires intervenant dans (1.61) :

$$\tau_1 = \tau_1(a) = \tau_L + \frac{1}{a^2}(\sum_{i=1}^{N}(p_i)_L - \Pi^\star), \quad \tau_2 = \tau_2(a) = \tau_R + \frac{1}{a^2}(\sum_{i=1}^{N}(p_i)_R - \Pi^\star),$$

avec

$$\Pi^\star = \Pi^\star(a) = \frac{1}{2}\sum_{i=1}^{N}((p_i)_L + (p_i)_R) - \frac{a}{2}(u_R - u_L),$$

peuvent être rendus arbitrairement proches respectivement de τ_L et τ_R, quitte à prendre a suffisamment grand. En conséquence, les conditions de Whitham (1.61) portant sur les états intermédiaires τ_1 et τ_2 peuvent toujours être satisfaites. Il en est de même pour l'hypothèse (1.15) portant sur l'ordre naturel des ondes. Nous ne détaillons pas ici d'algorithme permettant d'optimiser le calcul de a selon (1.15)-(1.61). Soulignons cependant qu'il est possible d'obtenir une formulation affaiblie de ces conditions de stabilité, plus simple à manipuler et à imposer numériquement. Nous renvoyons le lecteur à Bouchut [18] pour une description dans le cas d'une loi de gaz parfait polytropique.

La démonstration du théorème (5) utilise la validité du résultat suivant :
Lemme 3
Soient $\overline{\mathbf{U}}_L$ et $\overline{\mathbf{U}}_R$ deux états de $\overline{\Omega}$ à l'équilibre, c'est-à-dire tels que

$$\begin{cases} (\rho I)_L = \rho_L \sum_{i=1}^{N} p_i(1/\rho_L, (\rho S_i)_L/\rho_L) + a^2, \\ (\rho I)_R = \rho_R \sum_{i=1}^{N} p_i(1/\rho_R, (\rho S_i)_R/\rho_R) + a^2. \end{cases}$$

En reprenant les notations du théorème 1, supposons que la constante a satisfait la condition (1.15) et les conditions de Whitham (1.61).
Soit $\overline{\mathbf{U}}(x/t; \overline{\mathbf{U}}_L, \overline{\mathbf{U}}_R) = (\rho, \rho u, \{\rho S_i\}_{i=1,\dots,N}, \rho I)(x/t; \overline{\mathbf{U}}_L, \overline{\mathbf{U}}_R)$ la solution auto-semblable du problème de Riemann (1.14) pour la donnée initiale (1.16).
Alors, pour tout $(x,t) \in I\!R \times I\!R^+$

$$\rho E((\rho, \rho u, \{\rho S_i\}_{i=1,\dots,N})(x/t; \overline{\mathbf{U}}_L, \overline{\mathbf{U}}_R)) \leq \rho\Sigma(\overline{\mathbf{U}}(x/t; \overline{\mathbf{U}}_L, \overline{\mathbf{U}}_R)). \tag{1.65}$$

Démonstration du Lemme 3
Pour tout $(x,t) \in I\!R \times I\!R^+$, la solution $\overline{\mathbf{U}}(x/t; \overline{\mathbf{U}}_L, \overline{\mathbf{U}}_R)$ du problème de Riemann (1.14) pour la donnée initiale (1.16) est, d'après le théorème 1, telle que

$$(\{S_i\}_{i=1,\dots,N}, I)(x/t; \overline{\mathbf{U}}_L, \overline{\mathbf{U}}_R)$$

vaut soit $(\{(S_i)_L\}_{i=1,\dots,N}, I_L)$, soit $(\{(S_i)_R\}_{i=1,\dots,N}, I_R)$. Ainsi, la condition de Whitham (1.61) étant satisfaite, $\mathcal{T}(\{S_i\}_{i=1,\dots,N}, I)$ est bien défini pour tout (x,t) et vaut respectivement τ_L ou τ_R, puisque I_L et I_R sont par hypothèse à l'équilibre. De manière plus précise et en reprenant les notations du théorème 1, $\overline{\mathbf{U}}(x/t; \overline{\mathbf{U}}_L, \overline{\mathbf{U}}_R)$ est telle que :
(i) pour tout couple (x,t) tel que $x/t \in]-\infty, \lambda_1(\overline{\mathbf{U}}_L)[$,

$$\overline{\mathbf{U}}(x/t; \overline{\mathbf{U}}_L, \overline{\mathbf{U}}_R) = (\rho_L, (\rho u)_L, \{(\rho S_i)_L\}_{i=1,\dots,N}, (\rho I)_L).$$

L'état $\overline{\mathbf{U}}_L$ étant à l'équilibre, nous avons

$$\rho E(\rho_L, (\rho u)_L, \{(\rho S_i)_L\}_{i=1,\dots,N}) = \rho\Sigma(\rho_L, (\rho u)_L, \{(\rho S_i)_L\}_{i=1,\dots,N}, (\rho I)_L).$$

Sur cet intervalle, (1.65) est donc satisfaite avec égalité.
(ii) pour tout couple (x,t) tel que $x/t \in \,]\lambda_1(\overline{\mathbf{U}}_L), \lambda_2(\overline{\mathbf{U}}_1)[$,

$$\overline{\mathbf{U}}(x/t; \overline{\mathbf{U}}_L, \overline{\mathbf{U}}_R) = (\rho_1, \rho_1 u^\star, \{\rho_1 (S_i)_L\}_{i=1,\dots,N}, \rho_1 I_L).$$

Remarquons tout d'abord que $\rho\Sigma(\rho_1, \rho_1 u^\star, \{\rho_1 (S_i)_L\}_{i=1,\dots,N}, \rho_1 I_L)$ existe et est bien défini puisque $\rho_1 > 0$ et que par construction $\mathcal{T}(I_L, \{(S_i)_L\}_{i=1,\dots,N}) = \tau_L$. Posons

$$I_{1L}^{eq} = I_{1L}^{eq}(\tau_1, \{(S_i)_L\}_{i=1,\dots,N}) = \sum_{i=1}^{N} p_i(\tau_1, (S_i)_L) + a^2 \tau_1, \quad \tau_1 = \frac{1}{\rho_1}. \tag{1.66}$$

La condition de Whitham (1.61) étant satisfaite, $\rho\Sigma(\rho_1, \rho_1 u^\star, \{\rho_1 (S_i)_L\}_{i=1,\dots,N}, \rho_1 I_{1L}^{eq})$ existe et est bien défini. En effet, $\mathcal{T}(\{(S_i)_L\}_{i=1,\dots,N}, I_{1L}^{eq})$ est bien défini et vaut τ_1 d'après (1.66). Il s'en suit par définition de $\rho\Sigma$ que

$$\rho\Sigma(\rho_1, \rho_1 u^\star, \{\rho_1 (S_i)_L\}_{i=1,\dots,N}, \rho_1 I_{1L}^{eq}) = \rho E(\rho_1, \rho_1 u^\star, \{\rho_1 (S_i)_L\}_{i=1,\dots,N}). \tag{1.67}$$

Nous avons montré que l'application $I \to \mathcal{T}(\{(S_i)_L\}_{i=1,\dots,N}, I)$ est définie en I_L et en I_{1L}^{eq}. La convexité de l'application f définie au lemme 1 montre qu'elle admet nécessairement deux racines distinctes pour tout

$$I \in \; [\min(I_L, I_{1L}^{eq}), \max(I_L, I_{1L}^{eq})],$$

puisque tel est le cas aux bornes de cet intervalle. En conséquence, $\mathcal{T}(\{(S_i)_L\}_{i=1,\dots,N}, I)$ (et donc $\Sigma(\rho_1, \rho_1 u^\star, \{\rho_1 (S_i)_L\}_{i=1,\dots,N}, \rho_1 I)$) est bien défini sur l'intervalle précédent et vérifie

$$\mathcal{T}(\{(S_i)_L\}_{i=1,\dots,N}, I) \in \; [\min(\tau_L, \tau_1), \max(\tau_L, \tau_1)].$$

Par ailleurs, l'application

$$I \to \rho_1 \Sigma(\rho_1, \rho_1 u^\star, \{\rho_1 (S_i)_L\}_{i=1,\dots,N}, \rho_1 I)$$

étant définie sur l'intervalle $[\min(I_L, I_{1L}^{eq}), \max(I_L, I_{1L}^{eq})]$, elle y est convexe d'après le lemme 2. Le théorème 4 précise que le minimum est atteint sur la valeur de I réalisant l'équilibre, à savoir $I = I_{1L}^{eq}$. Nous obtenons :

$$\rho\Sigma(\rho_1, \rho_1 u^\star, \{\rho_1 (S_i)_L\}_{i=1,\dots,N}, \rho_1 I_{1L}^{eq}) \leq \rho\Sigma(\rho_1, \rho_1 u^\star, \{\rho_1 (S_i)_L\}_{i=1,\dots,N}, \rho_1 I_L).$$

L'inégalité (1.65) est donc satisfaite d'après (1.67) pour tout couple (x,t) tel que $x/t \in \,]\lambda_1(\overline{\mathbf{U}}_L), \lambda_2(\overline{\mathbf{U}}_1)[$.
De la même manière qu'en **(ii)** (respectivement en **(i)**) on prouve que (1.65) est satisfaite pour tout couple (x,t) tel que $x/t \in]\lambda_2(\overline{\mathbf{U}}_1), \lambda_3(\overline{\mathbf{U}}_R)[$ (respectivement pour tout couple (x,t) tel que $x/t \in]\lambda_3(\overline{\mathbf{U}}_R), +\infty[$). Ceci conclut la démonstration du lemme.

Démonstration du théorème 5
La propriété de stabilité L^1 est vraie par construction : rappelons que les densités intermédiaires ρ_1 et ρ_2 sont positives dès que a vérifie (1.15) et que chaque entropie

spécifique reste positive en vertu des principes du maximum (1.60) mis en évidence précédemment.
Etablissons maintenant l'inégalité d'entropie (1.62). De sorte à alléger le système de notations introduit lors de l'exposé détaillé de la méthode, écrivons pour tout (x,t) appartenant à $(x_j, x_{j+1}) \times (t^n, t^{n+1-})$:

$$(\mathbf{U}, \rho I)(x,t) = \overline{\mathbf{U}}_{a_{j+1/2}}(x,t) := \overline{\mathbf{U}}_{a_{j+1/2}}((x - x_{j+1/2})/(t-t^n); \overline{\mathbf{U}}^n_{j,+}, \overline{\mathbf{U}}^n_{j+1,-}),$$

avec par construction (voir (1.55) et (1.49)) :

$$\mathbf{U}^{n+1}_{j+1/2,L} = \frac{2}{\Delta x}\int_{x_j}^{x_{j+1/2}} \mathbf{U}(x, t^{n+1-})dx, \tag{1.68}$$

$$\mathbf{U}^{n+1}_{j+1/2,R} = \frac{2}{\Delta x}\int_{x_{j+1/2}}^{x_{j+1}} \mathbf{U}(x, t^{n+1-})dx. \tag{1.69}$$

Le lemme précédent assure que pour tout $x \in (x_j, x_{j+1})$ l'inégalité suivante a lieu :

$$\rho E(\mathbf{U}(x, t^{n+1-})) \leq \rho\Sigma(\overline{\mathbf{U}}_{a_{j+1/2}}(x, t^{n+1-})).$$

Moyennons cette inégalité sur chaque demi-cellule $(x_j, x_{j+1/2})$ et $(x_{j+1/2}, x_{j+1})$ pour obtenir respectivement :

$$\tfrac{2}{\Delta x}\int_{x_j}^{x_{j+1/2}} \rho E(\mathbf{U}(x, t^{n+1-}))dx \leq \tfrac{2}{\Delta x}\int_{x_j}^{x_{j+1/2}} \rho\Sigma(\overline{\mathbf{U}}_{a_{j+1/2}}(x, t^{n+1-}))dx, \tag{1.70}$$

et

$$\tfrac{2}{\Delta x}\int_{x_{j+1/2}}^{x_{j+1}} \rho E(\mathbf{U}(x, t^{n+1-}))dx \leq \tfrac{2}{\Delta x}\int_{x_{j+1/2}}^{x_{j+1}} \rho\Sigma(\overline{\mathbf{U}}_{a_{j+1/2}}(x, t^{n+1-}))dx. \tag{1.71}$$

En invoquant les relations (1.68) et (1.69), la stricte convexité de l'application $\mathbf{U} \to \rho E(\mathbf{U})$ permet de déduire par l'inégalité de Jensen :

$$\rho E(\mathbf{U}^{n+1}_{j+1/2,L}) \leq \tfrac{2}{\Delta x}\int_{x_j}^{x_{j+1/2}} \rho\Sigma(\overline{\mathbf{U}}_{a_{j+1/2}}(x, t^{n+1-}))dx, \tag{1.72}$$

et

$$\rho E(\mathbf{U}^{n+1}_{j+1/2,R}) \leq \tfrac{2}{\Delta x}\int_{x_{j+1/2}}^{x_{j+1}} \rho\Sigma(\overline{\mathbf{U}}_{a_{j+1/2}}(x, t^{n+1-}))dx. \tag{1.73}$$

Donnons une formulation équivalente à chacune des moyennes intervenant dans le membre de droite de ces inégalités. A cette fin, rappelons que la propriété du système de Relaxation (1.14) de ne posséder que des champs linéairement dégénérés implique que toute loi de conservation supplémentaire vérifiée par les solutions régulières de ce système est encore vérifiée (*i.e.* avec égalité à zéro) au sens des distributions pour les solutions faibles (voir Godlewski et Raviart [57] ou Serre [118] pour ce résultat). En conséquence, la solution du problème de Riemann considéré vérifie au sens faible :

$$\partial_t \rho\Sigma + \partial_x(\rho\Sigma + \Pi)u = 0, \tag{1.74}$$

de sorte qu'en projection sur chaque demi-cellule $(x_j, x_{j+1/2})$ et $(x_{j+1/2}, x_{j+1})$, la propriété d'équilibre de la donnée initiale conduit à :

$$\begin{aligned}(\rho\Sigma)^{n+1-}_{j+1/2,L} &:= \tfrac{2}{\Delta x}\int_{x_j}^{x_{j+1/2}} \rho\Sigma(\overline{\mathbf{U}}_{a_{j+1/2}}(x,t^{n+1-}))dx = \\ (\rho E)^n_j &- 2\tfrac{\Delta t}{\Delta x}(\overline{F}^g_{\rho\Sigma,a_{j+1/2}}(\overline{\mathbf{U}}^n_{j,+},\overline{\mathbf{U}}^n_{j+1,-}) - \overline{F}_{\rho\Sigma,a_{j+1/2}}(\overline{\mathbf{U}}^n_{j,+})),\end{aligned} \tag{1.75}$$

$$\begin{aligned}(\rho\Sigma)^{n+1-}_{j+1/2,R} &:= \tfrac{2}{\Delta x}\int_{x_{j+1/2}}^{x_{j+1}} \rho\Sigma(\overline{\mathbf{U}}_{a_{j+1/2}}(x,t^{n+1-}))dx = \\ (\rho E)^n_{j+1} &- 2\tfrac{\Delta t}{\Delta x}(\overline{F}_{\rho\Sigma,a_{j+1/2}}(\overline{\mathbf{U}}^n_{j+1,-}) - \overline{F}^g_{\rho\Sigma,a_{j+1/2}}(\overline{\mathbf{U}}^n_{j,+},\overline{\mathbf{U}}^n_{j+1,-})).\end{aligned} \tag{1.76}$$

La fonction flux numérique intervenant dans ces deux identités vérifie par construction la définition (1.63) donnée dans l'énoncé. Avec des notations claires, la fonction flux exact $\overline{F}_{\rho\Sigma,a_{j+1/2}}$ trouve quant à elle une définition immédiate à partir de celle du flux $\overline{\mathbf{F}}_{a_{j+1/2}}$. En observant à l'instar des relations (1.56) et (1.57) que la propriété d'équilibre vérifiée par $\overline{\mathbf{U}}^n_{j,-}$ et $\overline{\mathbf{U}}^n_{j,+}$ pour tout $j \in Z$ se traduit par l'égalité suivante :

$$\overline{F}_{\rho\Sigma,a_{j-1/2}}(\overline{\mathbf{U}}^n_{j,+}) = \overline{F}_{\rho\Sigma,a_{j+1/2}}(\overline{\mathbf{U}}^n_{j,-}) = \{(\rho E + \sum_{i=1}^N p_i)u\}(\mathbf{U}^n_j),$$

nous déduisons de (1.72)–(1.73) et (1.75)–(1.76) l'inégalité suivante vérifiée sur la cellule $(x_{j-1/2}, x_{j+1/2})$:

$$\begin{gathered}\tfrac{1}{2}(\rho E(\mathbf{U}^{n+1}_{j+1/2,L}) + \rho E(\mathbf{U}^{n+1}_{j-1/2,R})) \leq \\ (\rho E)^n_j - \tfrac{\Delta t}{\Delta x}(\overline{F}^g_{\rho\Sigma,a_{j+1/2}}(\overline{\mathbf{U}}^n_{j,+},\overline{\mathbf{U}}^n_{j+1,-}) - \overline{F}^g_{\rho\Sigma,a_{j-1/2}}(\overline{\mathbf{U}}^n_{j-1,+},\overline{\mathbf{U}}^n_{j,-})).\end{gathered}$$

Pour conclure à la validité de l'inégalité d'énergie, il suffit d'invoquer la relation (1.54) :

$$\mathbf{U}^{n+1}_j = \frac{1}{2}(\mathbf{U}^{n+1}_{j+1/2,L} + \mathbf{U}^{n+1}_{j-1/2,R}),$$

ainsi que l'inégalité de Jensen.

Concernant enfin la propriété de précision (ii), rappelons que l'algorithme proposé consiste en une méthode de Godunov appliquée au système relaxé (1.14), la donnée initiale à l'équilibre considérée étant déduite de la donnée initiale du système (1.7) (voir (1.45) par exemple). La méthode de Godunov capturant de manière exacte les contacts stationnaires, il suffit de prouver que les contacts stationnaires du système à l'équilibre (1.7) sont aussi des contacts stationnaires pour la formulation relaxée (1.14). Considérons $\mathbf{U}_L$ et $\mathbf{U}_R$ deux états constants (appartenant à Ω) d'une discontinuité de contact stationnaire du système (1.7) associée à la valeur propre u. Les relations de Rankine-Hugoniot impliquent $\mathbf{F}(\mathbf{U}_L) = \mathbf{F}(\mathbf{U}_R)$ de telle sorte que $u_L = u_R = 0$, et $\sum_{i=1}^N (p_i)_L = \sum_{i=1}^N (p_i)_R$. Notons $\overline{\mathbf{U}}_L$ et $\overline{\mathbf{U}}_R$ les deux états constants à l'équilibre de $\overline{\Omega}$ associés à $\mathbf{U}_L$ et $\mathbf{U}_R$, c'est-à-dire tels que

$$(\rho I)_L = \rho_L \sum_{i=1}^N (p_i)_L + a^2, \quad (\rho I)_R = \rho_R \sum_{i=1}^N (p_i)_R + a^2,$$

ou encore

$$\Pi_L = \sum_{i=1}^{N} (p_i)_L, \quad \Pi_R = \sum_{i=1}^{N} (p_i)_R.$$

Les formules donnant la solution du problème de Riemann (voir théorème 1) montrent que $u^\star = u_L = u_R = 0$, $\Pi^\star = \Pi_L = \Pi_R$. $\overline{\mathbf{U}}_L$ et $\overline{\mathbf{U}}_R$ sont donc deux états constants appartenant à $\overline{\Omega}$ d'une discontinuité de contact stationnaire du système (1.14), associée à la valeur propre u. Ceci conclut la démonstration.

1.6 Application aux équations d'Euler multi-pression

1.6.1 Définition du schéma

Dans cette section, nous montrons comment déduire de la méthode de Relaxation précédemment décrite, une méthode de volumes finis consistante pour l'approximation des solutions faibles des équations d'Euler multi-pression (1.5), *i.e.* :

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + \sum_{i=1}^{N} p_i) = 0, \\ \partial_t \rho E + \partial_x (\rho E + \sum_{i=1}^{N} p_i) u = 0, \\ \partial_t \rho S_i + \partial_x \rho S_i u = 0, \quad i = 1, ..., N-1, \end{cases}$$

complétées par l'inégalité d'entropie

$$\partial_t \rho S_N(\mathbf{u}) + \partial_x \rho S_N(\mathbf{u}) u \leq 0.$$

La technique utilisée s'appuie sur une idée proposée par Coquel et Perthame [40] (voir également Coquel et al. [32]).
La solution approchée de (1.5), notée $\mathbf{u}_{\Delta x}(x,t) : I\!R \times I\!R_\star^+ \to \Theta$, avec $\mathbf{u_0}$ comme donnée initiale, est une fois de plus considérée constante par morceaux sur chaque pavé $[x_{j-1/2}, x_{j+1/2}[\times[t^n, t^{n+1}[$ pour $(n,j) \in N \times Z$, et nous posons

$$\mathbf{u}_{\Delta x}(x,t) = \mathbf{u}_j^n, \quad \text{pour tout } (x,t) \in [x_{j-1/2}, x_{j+1/2}[\times[t^n, t^{n+1}[. \tag{1.77}$$

Nous posons par ailleurs $\mathbf{u}_j^0 = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{u}_0(x) dx$ pour tout $j \in Z$.
Supposons connue la suite $(\mathbf{u}_j^n = (\rho_j^n, (\rho u)_j^n, (\rho E)_j^n, \{(\rho S_i)_j^n\}_{i=1,...,N-1}))_{j\in Z}$, et définissons la suite correspondante $(\mathbf{U}_j^n = (\rho_j^n, (\rho u)_j^n, \{(\rho S_i)_j^n\}_{i=1,...,N}))_{j\in Z}$ de Ω à l'aide de la loi de fermeture (1.3). Nous proposons alors d'avancer la solution $\mathbf{u}_{\Delta x}(x,t)$ jusqu'à l'instant t^{n+1} en deux étapes :

Premier pas : Evolution en temps $(t^n \to t^{n+1-})$
Ce premier pas consiste à faire évoluer en temps la suite $(\mathbf{U}_j^n)_{j\in Z}$ en utilisant la méthode de Relaxation décrite dans la section précédente. En reprenant les notations correspondantes, nous obtenons une mise à jour $(\mathbf{U}_j^{n+1-})_{j\in Z}$ définie par

$$\mathbf{U}_j^{n+1-} = \mathbf{U}_j^n - \frac{\Delta t}{\Delta x}(\mathbf{F}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n) - \mathbf{F}(\mathbf{U}_{j-1}^n, \mathbf{U}_j^n)), \quad j \in Z, \tag{1.78}$$

où la fonction flux numérique est donnée par (1.59). D'après le théorème 5, cette mise à jour conservative est consistante avec le système (1.7) et possède toutes les propriétés de stabilité requises. Elle est en particulier, d'après (1.62), consistante avec l'inégalité d'énergie (1.8) :

$$\rho E(\mathbf{U}_j^{n+1-}) \le (\rho E)_j^n - \tfrac{\Delta t}{\Delta x}(F_{\rho\Sigma}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n) - F_{\rho\Sigma}(\mathbf{U}_{j-1}^n, \mathbf{U}_j^n)), \tag{1.79}$$

où la fonction flux d'énergie numérique est donnée par (1.63).

Second pas : Conservation de l'énergie totale $(t^{n+1-} \to t^{n+1})$
L'objectif de ce deuxième pas est de définir la suite $(\mathbf{u}_j^{n+1})_{j\in Z}$ en forçant la conservation de l'énergie totale ρE par opposition à l'inégalité (1.79). Pour cela, nous proposons de garder inchangées les variables suivantes :

$$\rho_j^{n+1} = \rho_j^{n+1-}, \quad (\rho u)_j^{n+1} = (\rho u)_j^{n+1-}, \quad (\rho S_i)_j^{n+1} = (\rho S_i)_j^{n+1-}, \quad i = 1, ..., N-1, \tag{1.80}$$

et de redéfinir l'énergie totale $(\rho E)_j^{n+1}$ en posant

$$(\rho E)_j^{n+1} = (\rho E)_j^n - \tfrac{\Delta t}{\Delta x}(F_{\rho\Sigma}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n) - F_{\rho\Sigma}(\mathbf{U}_{j-1}^n, \mathbf{U}_j^n)). \tag{1.81}$$

Nous avons ainsi défini une évolution conservative et consistante à trois points de la suite $(\mathbf{u}_j^n)_{j\in Z}$. Nous établissons le résultat suivant mettant en évidence les propriétés de cette mise à jour explicite.

Théorème 6
Sous les hypothèses du théorème 5, le schéma numérique (1.78)-(1.80)-(1.81) satisfait les propriétés suivantes :

(i) Propriétés de stabilité :
- *Stabilité* L^1 : $\mathbf{u}_j^{n+1} \in \Theta$ *pour tout* $j \in Z$.
- *Inégalité d'entropie :*

$$\begin{aligned}(\rho S_N)_j^{n+1} &:= \rho S_N(\mathbf{u}_j^{n+1}) \le \\ &(\rho S_N)_j^n - \tfrac{\Delta t}{\Delta x}(F_{\rho S_N}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n) - F_{\rho S_N}(\mathbf{U}_{j-1}^n, \mathbf{U}_j^n)), \quad j \in Z.\end{aligned} \tag{1.82}$$

- *Principes du maximum :*

$$\min((S_i)_{j-1}^n, (S_i)_j^n, (S_i)_{j+1}^n) \le (S_i)_j^{n+1} \le \max((S_i)_{j-1}^n, (S_i)_j^n, (S_i)_{j+1}^n), \quad i = 1, ..., N-1, \tag{1.83}$$

$$(S_N)_j^{n+1} \le \max((S_N)_{j-1}^n, (S_N)_j^n, (S_N)_{j+1}^n), \quad j \in Z. \tag{1.84}$$

(ii) Propriété de précision :
- *Les discontinuités de contact stationnaires du système (1.5) sont exactement résolues.*

Démonstration du théorème 6
Tout d'abord, les principes du maximum (1.83) sont immédiats compte-tenu de (1.80) et du théorème 5.
Concernant la propriété de précision, toute discontinuité de contact stationnaire pour

le système (1.5) est une discontinuité de contact stationnaire pour le système relaxé (1.14). En effet, dans le cas d'une discontinuité de contact, (1.4) est vraie avec une égalité à zéro (la dissipation d'entropie correspondante est nulle). Or les mises à jour (1.80) et (1.81) correspondent à une projection L^2 de type Godunov des variables conservées $(\rho, \rho u, \rho\Sigma, (\rho S_i)_{i=1,...,N})$ du système (1.14). La méthode de Godunov préservant les discontinuités de contact stationnaires, la propriété de précision attendue est vérifiée.
Afin d'établir l'inégalité d'entropie (1.82), remarquons que la loi d'énergie

$$\rho E = \frac{(\rho u)^2}{2\rho} + \sum_{i=1}^{N} \rho\varepsilon_i(1/\rho, \rho S_i/\rho)$$

définit implicitement l'application $\mathbf{u} \to \rho S_N(\mathbf{u})$ et implique en la dérivant partiellement par rapport à ρE :

$$1 = \frac{\partial \varepsilon_N}{\partial S_N}(\frac{1}{\rho}, \frac{\rho S_N}{\rho}) \times \frac{\partial \rho S_N}{\partial \rho E},$$

c'est-à-dire $\frac{\partial \rho S_N}{\partial \rho E} = -\frac{1}{T_N} < 0$ d'après (1.2). L'application $\mathbf{u} \to \rho S_N(\mathbf{u})$ est donc une application décroissante de ρE. Ainsi, la validité de (1.79), qui se réexprime d'après la définition (1.81) sous la forme :

$$\rho E(\mathbf{U}_j^{n+1-}) \leq (\rho E)_j^{n+1},$$

entraîne

$$\rho S_N(\rho_j^{n+1}, (\rho u)_j^{n+1}, (\rho E)_j^{n+1}, \{(\rho S_i)_j^{n+1}\}_{i=1,...,N-1}) \leq$$

$$\rho S_N(\rho_j^{n+1}, (\rho u)_j^{n+1}, \rho E(\mathbf{U}_j^{n+1-}), \{(\rho S_i)_j^{n+1}\}_{i=1,...,N-1}) = (\rho S_N)_j^{n+1-}, \quad j \in Z. \tag{1.85}$$

Le membre de droite de cette inégalité coïncide en effet par construction avec la quantité $(\rho S_N)_j^{n+1-}$ d'après (1.80). Nous obtenons l'inégalité attendue en invoquant (1.78) :

$$\rho S_N(\mathbf{u}_j^{n+1}) \leq (\rho S)_j^n - \frac{\Delta t}{\Delta x}(F_{\rho S_N}(\mathbf{U}_j^n, \mathbf{U}_{j+1}^n) - F_{\rho S_N}(\mathbf{U}_{j-1}^n, \mathbf{U}_j^n)).$$

Par ailleurs, puisque $\rho_j^{n+1} = \rho_j^{n+1-}$, (1.85) entraîne $(S_N)_j^{n+1} \leq (S_N)_j^{n+1-}$. La validité du principe du maximum (1.84) est assurée puisque d'après le théorème 5, $(S_N)_j^{n+1-} \leq \max((S_N)_{j-1}^n, (S_N)_j^n, (S_N)_{j+1}^n)$. De plus, la décroissance de l'application ε_N par rapport à la variable S_N ($\partial\varepsilon_N/\partial S_N = -T_N < 0$) entraîne

$$0 \leq \varepsilon_N(\rho_j^{n+1-}, (S_N)_j^{n+1-}) = \varepsilon_N(\rho_j^{n+1}, (S_N)_j^{n+1-}) \leq \varepsilon_N(\rho_j^{n+1}, (S_N)_j^{n+1}).$$

On en déduit :

$$\rho_j^{n+1}\varepsilon_N(\rho_j^{n+1}, (S_N)_j^{n+1}) = (\rho E)_j^{n+1} - \frac{((\rho u)_j^{n+1})^2}{2\rho_j^{n+1}} - \sum_{i=1}^{N-1} \rho_j^{n+1}\varepsilon_i(1/\rho_j^{n+1}, (S_i)_j^{n+1}) \geq 0.$$

Le vecteur $\mathbf{u}_j^{n+1}$ est donc dans Θ puisque la positivité des grandeurs ρ_j^{n+1} et $(\rho S_i)_j^{n+1}$ pour tout $i = 1, ..., N-1$ est assurée par construction. Ceci conclut la démonstration du théorème.

1.6.2 Un modèle de Relaxation équivalent

L'objectif de ce paragraphe est de montrer que l'algorithme de la méthode de volumes finis précédente s'appuyant sur une méthode de Relaxation pour le système compagnon (1.14), peut être réinterprétée directement comme une méthode de Relaxation pour le système Euler multi-pression, compatible au sens de Chen, Levermore et Liu [28] avec le couple d'entropie de Lax $(\rho S_N, \rho S_N u)$. Dans ce nouveau modèle de Relaxation, l'énergie $\rho\Sigma$ devient une inconnue conservative, et un relèvement $\rho\mathcal{S}_\mathcal{N}$ de l'entropie ρS_N est défini de manière implicite à partir de la construction précédente. Cette réinterprétation sera utile aux développements suivants consacrés à l'implicitation en temps de la présente technique d'approximation numérique.

Le premier résultat suivant montre qu'il est possible d'inverser le relèvement de l'énergie $\rho\Sigma$ relativement à ρS_N de sorte à définir ultérieurement le relèvement de $\rho\mathcal{S}_\mathcal{N}$.

Lemme 4
L'application $\overline{\mathbf{U}} = {}^T(\rho, \rho u, \{\rho S_i\}_{i=1,\dots,N}, \rho I) \to \rho\Sigma(\overline{\mathbf{U}})$ *est inversible par rapport à la variable* ρS_N.

Démonstration
Il suffit de montrer que l'application

$$^T(\tau, u, \{S_i\}_{i=1,\dots,N}, I) \to \Sigma(\tau, u, \{S_i\}_{i=1,\dots,N}, I)$$

est inversible en S_N. Rappelons que cette application est définie par :

$$\Sigma(\tau, u, \{S_i\}_{i=1,\dots,N}, I) = \frac{u^2}{2} + \frac{(I - a^2\tau)^2}{2a^2} + \phi(\{S_i\}_{i=1,\dots,N-1}, S_N, I),$$

où

$$\phi(\{S_i\}_{i=1,\dots,N-1}, S_N, I) = \varepsilon(\mathcal{T}, \{S_i\}_{i=1,\dots,N}) - \frac{p^2(\mathcal{T}, \{S_i\}_{i=1,\dots,N})}{2a^2},$$

et avec

$$\mathcal{T} = \mathcal{T}(\{S_i\}_{i=1,\dots,N-1}, S_N, I), \quad \varepsilon = \sum_{i=1}^{N} \varepsilon_i(\mathcal{T}, S_i), \quad p = \sum_{i=1}^{N} p_i(\mathcal{T}, S_i).$$

Un calcul direct donne :

$$\frac{\partial\Sigma}{\partial S_N} = (\frac{\partial\varepsilon}{\partial\mathcal{T}} - \frac{p}{a^2}\frac{\partial p}{\partial\mathcal{T}})\frac{\partial\mathcal{T}}{\partial S_N} + (\frac{\partial\varepsilon}{\partial S_N}_{|\mathcal{T}} - \frac{p}{a^2}\frac{\partial p}{\partial S_N}_{|\mathcal{T}})$$

c'est-à-dire

$$\frac{\partial\Sigma}{\partial S_N} = -\frac{p}{a^2}(a^2 + \frac{\partial p}{\partial\mathcal{T}})\frac{\partial\mathcal{T}}{\partial S_N} + (\frac{\partial\varepsilon}{\partial S_N}_{|\mathcal{T}} - \frac{p}{a^2}\frac{\partial p}{\partial S_N}_{|\mathcal{T}}).$$

Or par construction, $\mathcal{T}$ est solution de

$$I = p(\mathcal{T}, \{S_i\}_{i=1,\dots,N}) + a^2\mathcal{T},$$

de sorte que

$$(a^2 + \frac{\partial p}{\partial\mathcal{T}})\frac{\partial\mathcal{T}}{\partial S_N} = -\frac{\partial p}{\partial S_N}_{|\mathcal{T}}.$$

Nous obtenons donc

$$\frac{\partial \Sigma}{\partial S_N} = \frac{\partial \varepsilon}{\partial S_N}_{|\mathcal{T}} = -T_N(\mathcal{T}, S_N) < 0, \tag{1.86}$$

ce qui conclut la preuve.

Munis de ce résultat, nous définissons le changement de variable admissible $\overline{\mathbf{u}}(\overline{\mathbf{U}}) = (\rho, \rho u, \rho\Sigma, \{\rho S_i\}_{i=1,...,N-1}, \rho I)$ et considérons l'application $\rho\mathcal{S}_{\mathcal{N}}(\overline{\mathbf{u}})$ définie implicitement à partir de la définition de $\rho\Sigma(\overline{\mathbf{U}})$. Avec un léger abus de notations, le relèvement $\mathcal{T}$ du covolume τ sera à l'avenir compris comme une fonction de $\overline{\mathbf{u}}$. Nous avons alors le résultat suivant :

Proposition 2
Considérons le modèle de Relaxation suivant pour le système Euler multi-pression :

$$\begin{cases} \partial_t\rho + \partial_x\rho u = 0, \\ \partial_t\rho u + \partial_x(\rho u^2 + \Pi) = 0, \\ \partial_t\rho\Sigma + \partial_x(\rho\Sigma + \Pi)u = 0, \\ \partial_t\rho S_i + \partial_x\rho S_i u = 0, \quad i = 1, ..., N-1, \\ \partial_t\rho I + \partial_x\rho I u = \lambda\rho(p(\overline{\mathbf{u}}) - \Pi). \end{cases} \tag{1.87}$$

Les solutions régulières de ce modèle vérifient l'inégalité d'entropie suivante :

$$\partial_t\rho\mathcal{S}_{\mathcal{N}} + \partial_x\rho\mathcal{S}_{\mathcal{N}}u = -\lambda\frac{\rho}{a^2 T_N(\mathcal{T}, \mathcal{S}_{\mathcal{N}})}(p(\overline{\mathbf{u}}) - \Pi)^2 < 0. \tag{1.88}$$

En d'autres termes, le système (1.87) constitue bien un modèle de Relaxation du système (1.5) au sens de Chen, Levermore et Liu [28].

Démonstration
De manière à obtenir le résultat attendu, comprenons la loi de conservation sur $\rho\Sigma$ lorsque cette énergie est fonction de la variable $\overline{\mathbf{U}}$. Des calculs identiques à ceux proposés lors de la démonstration du théorème 2 conduisent aux relations (1.31) et (1.32) à partir du nouveau modèle de Relaxation, soit donc en les superposant :

$$\partial_t(\frac{\rho u^2}{2} + \frac{\rho\Pi^2}{2a^2}) + \partial_x(\frac{\rho u^2}{2} + \frac{\rho\Pi^2}{2a^2} + \Pi)u = \lambda\frac{\rho\Pi}{a^2}(p - \Pi).$$

En soustrayant cette identité à la loi de conservation gouvernant $\rho\Sigma(\overline{\mathbf{U}})$, nous obtenons avec les notations de la preuve précédente :

$$\rho(\partial_t\phi + u\partial_x\phi) = -\lambda\frac{\rho\Pi}{a^2}(p - \Pi).$$

En tenant compte des équations de transport satisfaites par les $(N-1)$ premières entropies spécifiques, il vient :

$$\rho\frac{\partial\phi}{\partial I}(\partial_t I + u\partial_x I) + \rho\frac{\partial\phi}{\partial S_N}(\partial_t\mathcal{S}_{\mathcal{N}} + u\partial_x\mathcal{S}_{\mathcal{N}}) = -\lambda\frac{\rho\Pi}{a^2}(p - \Pi).$$

Des arguments semblables à ceux développés lors de la preuve du théorème 2 montrent que $\frac{\partial \phi}{\partial I}$ est toujours donné par la formule (1.35), alors que $\frac{\partial \phi}{\partial S_N} = \frac{\partial \Sigma}{\partial S_N} = -T_N$ en vertu de la preuve précédente. Nous venons donc d'établir :

$$-\rho T_N(\partial_t \mathcal{S}_\mathcal{N} + u\partial_x \mathcal{S}_\mathcal{N}) = \lambda \frac{\rho}{a^2}(p - \Pi)^2,$$

ce qui n'est autre que le résultat recherché.

L'équivalence de la méthode de volumes finis associée à ce nouveau modèle de Relaxation, obtenue conformément à la stratégie générale à deux pas ($\lambda = 0$ puis $\lambda = +\infty$), avec la méthode construite dans les précédents paragraphes résulte de la propriété de dégénérescence linéaire de tous les champs. En effet, rappelons-le, une loi de conservation satisfaite par les solutions régulières d'un système présentant cette propriété, reste vraie au sens des distributions. La solution du problème de Riemann pour le nouveau modèle est en particulier la même que celle du modèle (1.14) et en conséquence la méthode numérique explicite est strictement identique. Elle possède dès lors les mêmes propriétés de stabilité. Il nous reste juste à préciser les valeurs de Σ à la traversée des ondes constituant cette solution. Ceci fait l'objet du résultat suivant concluant la présente section.

Théorème 7
Soit $\overline{\Theta}$ l'espace des états admissibles du système (1.87) avec $\lambda = 0$ défini par

$$\overline{\Theta} = \{\overline{\mathbf{u}} \in I\!R^{N+3}, \rho > 0, \{\rho S_i > 0\}_{i=1,\dots,N-1}, \rho E - \frac{(\rho u)^2}{2\rho} - \sum_{i=1}^{N-1} \rho\varepsilon_i(1/\rho, \rho S_i/\rho) > 0, \rho I > 0\}, \tag{1.89}$$

et $(\overline{\mathbf{u}}_L, \overline{\mathbf{u}}_R)$ deux états de $\overline{\Theta}$.
Supposons que la constante a vérifie la condition

$$\lambda_1(\overline{\mathbf{u}}_L) = u_L - a\tau_L < u^\star < \lambda_3(\overline{\mathbf{u}}_R) = u_R + a\tau_R,$$

$$u^\star = \tfrac{1}{2}(u_L + u_R) - \tfrac{1}{2a}(\Pi_R - \Pi_L), \tag{1.90}$$

$$\Pi_L = I_L - a^2\tau_L, \quad \Pi_R = I_R - a^2\tau_R.$$

Alors, la solution faible auto-semblable $\overline{\mathbf{u}}(x,t;\overline{\mathbf{u}}_L,\overline{\mathbf{u}}_R) := \overline{\mathbf{u}}(x/t;\overline{\mathbf{u}}_L,\overline{\mathbf{u}}_R)$ du système (1.87) avec $\lambda = 0$ pour la donnée initiale

$$\overline{\mathbf{u}}_0(x) = \begin{cases} \overline{\mathbf{u}}_L & si \quad x < 0, \\ \overline{\mathbf{u}}_R & si \quad x > 0, \end{cases} \tag{1.91}$$

est constituée de quatre états constants :

$$\overline{\mathbf{u}}(x/t;\overline{\mathbf{u}}_L,\overline{\mathbf{u}}_R) = \begin{cases} \overline{\mathbf{u}}_L & si \quad \frac{x}{t} < \lambda_1(\overline{\mathbf{u}}_L), \\ \overline{\mathbf{u}}_1 & si \quad \lambda_1(\overline{\mathbf{u}}_L) < \frac{x}{t} < \lambda_2(\overline{\mathbf{u}}_1), \\ \overline{\mathbf{u}}_2 & si \quad \lambda_2(\overline{\mathbf{u}}_2) < \frac{x}{t} < \lambda_3(\overline{\mathbf{u}}_R), \\ \overline{\mathbf{u}}_R & si \quad \lambda_3(\overline{\mathbf{u}}_R) < \frac{x}{t}, \end{cases}$$

avec

$$\lambda_2(\overline{\mathbf{u}}_1) = \lambda_2(\overline{\mathbf{u}}_2) = u^\star.$$

En posant

$$\begin{aligned}
&u_1 = u_2 = u^\star,\\
&\rho_1 = a/(u_2 - \lambda_1(\overline{\mathbf{u}}_L)),\\
&\rho_2 = a/(\lambda_3(\overline{\mathbf{u}}_R) - u_2),\\
&\Pi^\star = \Pi_1 = \Pi_2 = \tfrac{1}{2}(\Pi_L + \Pi_R) - \tfrac{a}{2}(u_R - u_L),\\
&(S_i)_1 = (S_i)_L, \quad 1 \le i \le N-1,\\
&(S_i)_2 = (S_i)_R, \quad 1 \le i \le N-1,\\
&I_1 = \Pi_1 + \tfrac{a^2}{\rho_1} = I_L,\\
&I_2 = \Pi_2 + \tfrac{a^2}{\rho_2} = I_R,\\
&\Sigma_1 = \Sigma_L + (\Pi_L u_L - \Pi^\star u^\star)/a,\\
&\Sigma_2 = \Sigma_R - (\Pi_R u_R - \Pi^\star u^\star)/a,
\end{aligned}$$

les états constants sont caractérisés de la façon suivante :

$$\overline{\mathbf{u}}_1 = \begin{pmatrix} \rho_1 \\ (\rho u)_1 \\ (\rho\Sigma)_1 \\ (\rho S_1)_1 \\ \vdots \\ (\rho S_{N-1})_1 \\ (\rho I)_1 \end{pmatrix}, \quad \overline{\mathbf{u}}_2 = \begin{pmatrix} \rho_2 \\ (\rho u)_2 \\ (\rho\Sigma)_2 \\ (\rho S_1)_2 \\ \vdots \\ (\rho S_{N-1})_2 \\ (\rho I)_2 \end{pmatrix}.$$

De plus, les états $\overline{\mathbf{u}}_1$ *et* $\overline{\mathbf{u}}_2$ *sont dans* $\overline{\Theta}$.

1.7 Validation numérique

Nous proposons de tester la méthode de Relaxation présentée pour approcher les solutions du système (1.5). Pour cela, nous utilisons des lois de pression de gaz parfaits polytropiques, les coefficients adiabatiques associés étant notés $\{\gamma_i > 1\}_{i=1,\dots,N}$. La loi de fermeture sur l'énergie totale (1.3) se réécrit

$$\rho E = \frac{(\rho u)^2}{2\rho} + \sum_{i=1}^{N} \frac{p_i}{\gamma_i - 1}.$$

Nous effectuons dans cette section des expériences numériques en une dimension d'espace pour plusieurs cas tests de tube à choc. La discontinuité séparant les deux états constants (états gauche et droit) de la donnée initiale est située à l'abscisse $x = 0$.

Après avoir présenté ces cas tests, nous donnons les résultats numériques obtenus pour plusieurs maillages uniformes en espace, et à l'ordre 1 d'approximation. Nous utilisons pour ces calculs les modèles à 1, 2 ou 3 lois de pression. Nous comparons systématiquement la solution donnée par la méthode de Relaxation proposée avec celle obtenue par une méthode de Roe développée par Berthon [10]. Les solutions discrètes obtenues par ces deux méthodes sont identiques pour les cas tests considérés. Notons toutefois que la méthode de Relaxation proposée est entropique, contrairement à la méthode de Roe (sauf à faire appel à des techniques de correction proposées par exemple par Harten).

Expérience 1 : tube à choc de Sod

Pour ce premier cas test, nous faisons le choix d'une seule loi de pression polytropique associée à un coefficient adiabatique fixé à 1.4. Le système étudié n'est autre que le système des équations d'Euler habituelles. Les conditions initiales sont celles du tube à choc de Sod. Les valeurs initiales, exprimées en variables primitives, sont rappelées dans le tableau suivant.

Expérience 1

	densité	vitesse	pression
État gauche	1.	0.	1.
État droit	0.125	0.	0.1

L'examen des résultats ne laisse apparaître aucune différence notable entre les méthodes de Roe et de Relaxation.
A titre d'illustration, nous effectuons deux expériences supplémentaires en prenant successivement $N = 2$ et $N = 3$. Nous comparons les solutions numériques avec les solutions exactes pour obtenir une nouvelle fois un bon accord. Nous précisons les coefficients adiabatiques ainsi que les états gauche et droit de la donnée initiale, et validons numériquement la méthode de Relaxation.

Expérience 2

N	γ_1	γ_2
2	1.6	1.4

Expérience 2

	densité	vitesse	pression 1	pression 2
État gauche	1.	1.	0.6	1.
État droit	1.0553	-0.8863	0.3277	0.1633

Expérience 3

N	γ_1	γ_2	γ_3
3	1.6	1.4	1.4

Expérience 3

État gauche

densité	vitesse	pression 1	pression 2	pression 3
1.	1.	0.6	1.	1.

État droit

densité	vitesse	pression 1	pression 2	pression 3
1.1583	-1.0020	0.2450	0.2890	0.4023

Présentation des résultats

Expérience 1 : tube à choc de Sod

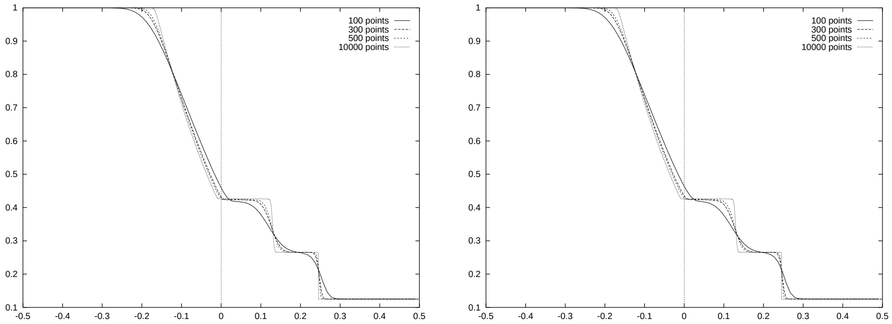

FIG. 1.1 – *Densité : Méthode de Roe, Méthode de Relaxation*

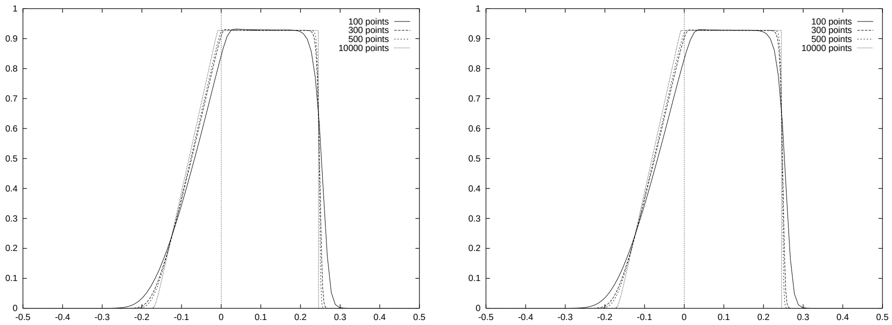

FIG. 1.2 – *Vitesse : Méthode de Roe, Méthode de Relaxation*

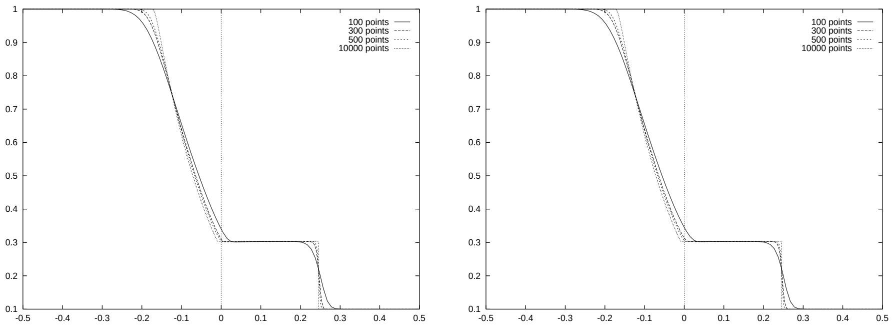

FIG. 1.3 – *Pression : Méthode de Roe, Méthode de Relaxation*

Expérience 2 : Cas 2 pressions

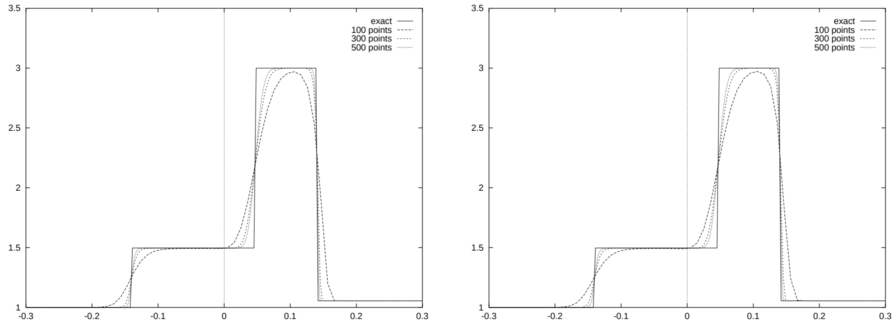

FIG. 1.4 – *Densité : Méthode de Roe, Méthode de Relaxation*

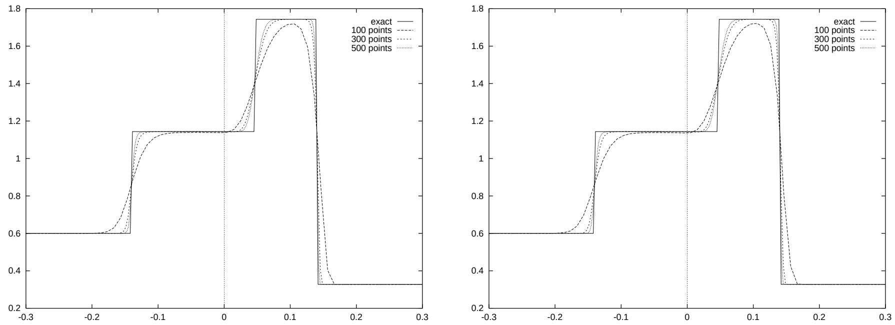

FIG. 1.5 – *Pression 1 : Méthode de Roe, Méthode de Relaxation*

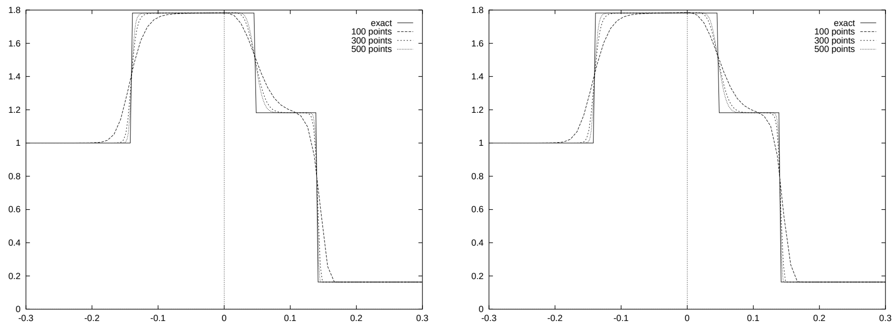

FIG. 1.6 – *Pression 2 : Méthode de Roe, Méthode de Relaxation*

Expérience 3 : Cas 3 pressions

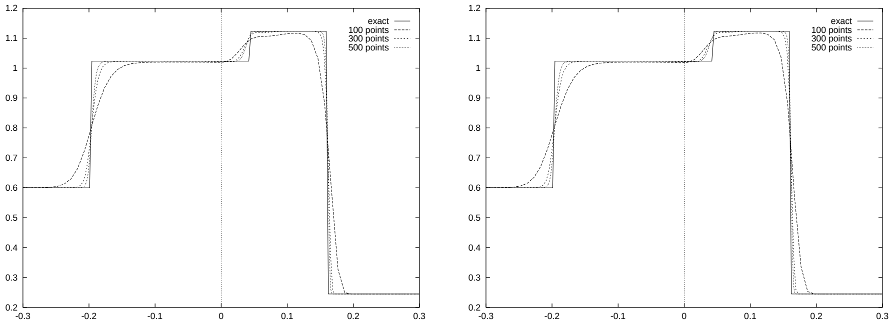

FIG. 1.7 – *Pression 1 : Méthode de Roe, Méthode de Relaxation*

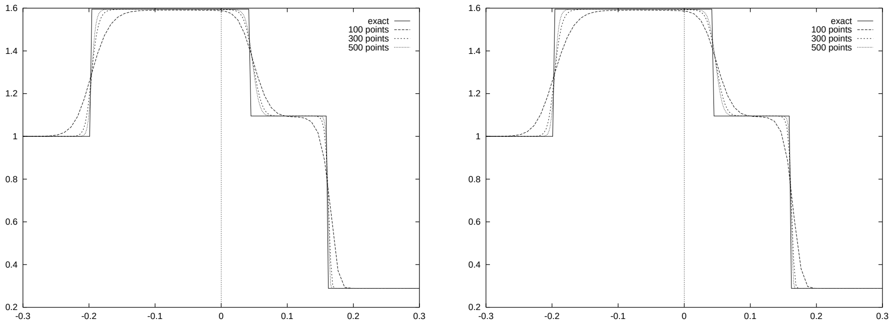

FIG. 1.8 – *Pression 2 : Méthode de Roe, Méthode de Relaxation*

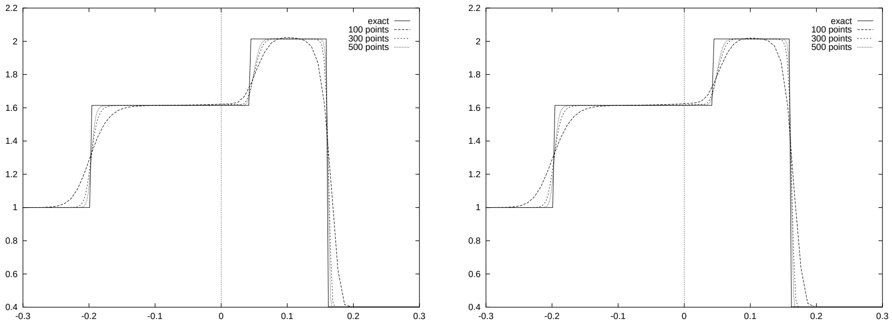

FIG. 1.9 – *Pression 3 : Méthode de Roe, Méthode de Relaxation*

Chapitre 2

Implicitation de la méthode de Relaxation

L'objectif de ce chapitre est de proposer une implicitation en temps de la méthode de Relaxation décrite dans le chapitre précédent pour approcher les solutions stationnaires du système des équations d'Euler à plusieurs pressions indépendantes (1.5).

Motivés par la mise en place d'une technique d'implicitation robuste et peu coûteuse, nous proposons de tirer profit du découplage naturel, obtenu par construction, des entropies du système (1.87) lorsque $\lambda = 0$. Par découplage, nous entendons qu'une fois le choix de la vitesse a effectué conformément aux conditions de Whitham à chaque pas de temps, l'évolution des variables ρ, ρu, $\rho\Sigma$ et ρI est totalement indépendante de l'évolution de chacune des entropies spécifiques. Ce découplage résulte de la procédure de Relaxation sur la pression totale. Il ne vaut évidemment pas pour le système à l'équilibre (1.5) puisque les pressions partielles p_i dépendent intrinsèquement des entropies S_i. Il apparait donc naturel de vouloir impliciter en temps de manière découplée le bloc transport des entropies :

$$\{\partial_t \rho S_i + \partial_x \rho S_i u = 0, \quad i = 1, ..., N-1, \tag{2.1}$$

et le bloc hydrodynamique :

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + \Pi) = 0, \\ \partial_t \rho\Sigma + \partial_x (\rho\Sigma + \Pi) u = 0, \\ \partial_t \rho I + \partial_x \rho I u = 0, \quad I = \Pi + a^2/\rho. \end{cases} \tag{2.2}$$

Nous renvoyons le lecteur à la fin de la section 2.1.2 pour une discussion précise sur la détermination de la valeur du paramètre a.

La technique que nous développons ci-dessous nous permettra d'assurer la validité d'un principe du maximum pour chacune des entropies spécifiques S_i, et ce sans restriction CFL. Elle reposera sur la résolution de $(N-1)$ systèmes linéaires à structure tridiagonale scalaire en une dimension d'espace (pentadiagonale en deux dimensions d'espace). La matrice à inverser sera toujours la même. La principale difficulté concernera l'implicitation en temps du bloc hydrodynamique. Une implicitation "naturelle" de ces équations, basée sur la réinterprétation du flux numérique de la méthode de

Godunov dans le formalisme des flux numériques associés à une méthode de Roe, sera dans un premier temps proposée. La robustesse de l'algorithme sera illustrée numériquement. Néammoins, une analyse plus fine établira que cette approche "naturelle" n'autorise pas l'obtention de solutions rigoureusement convergées en temps. La mise en évidence des raisons de ce défaut de convergence permettra alors de déduire une technique de correction simple de cette implicitation "naturelle". L'expérience numérique menée pour des problèmes en deux dimensions d'espace montrera que l'algorithme obtenu est stable, et qu'il permet le calcul de solutions parfaitement stationnaires du système Euler multi-pression. Soulignons que l'algorithme proposé présente un coût d'évaluation très avantageux par rapport à une approche classique reposant sur une implicitation du système Euler multi-pression complet.

De manière à débuter la description de la procédure, nous supposons connue à l'instant t^n une solution discrète à l'équilibre

$$\mathbf{u}_j^n = (\rho_j^n, (\rho u)_j^n, (\rho E)_j^n, \{(\rho S_i)_j^n\}_{i=1,\dots,N-1}), \quad j \in Z,$$

de vecteurs de Θ et nous définissons la solution relaxée associée :

$$\overline{\mathbf{u}}_j^n = (\rho_j^n, (\rho u)_j^n, (\rho E)_j^n, \{(\rho S_i)_j^n\}_{i=1,\dots,N-1}, (\rho I)_j^n), \quad j \in Z,$$

où

$$(\rho I)_j^n = \rho_j^n \sum_{i=1}^{N} p_i(1/\rho_j^n, (\rho S_i)_j^n/\rho_j^n) + a^2, \quad j \in Z.$$

Il s'agit de proposer une mise à jour implicite en temps de cette solution.

2.1 Implicitation "naturelle" de la méthode de Relaxation

Rappelons brièvement la stratégie de remise à jour de la solution discrète dans le cadre d'une formulation explicite en temps. Celle-ci est composée de deux pas, respectivement un pas d'évolution en temps où le paramètre de Relaxation λ est choisi nul (*i.e.* en l'absence du terme de Relaxation) suivi immédiatement d'une projection sur la variété d'équilibre (*i.e.* avec formellement $\lambda = +\infty$). Cette stratégie donne entière satisfaction pour une formulation explicite en temps. Il semble dès lors naturel d'adopter une démarche identique dans l'établissement d'une formulation implicite en temps : une implicitation en temps de l'étape d'évolution avec $\lambda = 0$ conduisant à des valeurs réactualisées de ρ, ρu, ρE, $\{\rho S_i\}_{i=1,\dots,N}$, ρI à une date t donnée, puis de manière à redémarrer l'algorithme, un écrasement à cette même date de la valeur de ρI en posant :

$$\rho I = \rho \sum_{i=1}^{N} p_i(1/\rho, \rho S_i/\rho) + a^2.$$

Ici, il faut évidemment définir à cette date t une nouvelle valeur de a suffisamment grande conformément aux conditions de Whitham, fonctions de la solution à l'équilibre réactualisée à cet instant. Nous renvoyons le lecteur à la fin de la sous-section 2.1.2 pour une brève discussion sur le choix de a. L'objectif de cette section est de proposer une telle implicitation avec le souci de découpler les blocs hydrodynamique et transport des entropies. L'expérience numérique montrera que l'algorithme implicite du bloc hydrodynamique partage des propriétés de robustesse comparable à celle

d'une méthode de Van Leer dans la résolution des équations d'Euler 3×3. Toutefois, nous montrerons que l'implicitation proposée ne peut pas conduire à une solution parfaitement convergée en temps. L'analyse de ce défaut permettra alors de corriger simplement cette première technique d'implicitation pour obtenir un algorithme partageant la même stabilité avec le bénéfice de converger vers une solution rigoureusement stationnaire.
Nous commençons par détailler l'implicitation de l'étape d'évolution en temps puis concluons avec la mise en évidence du défaut de convergence.

2.1.1 Implicitation du bloc transport des entropies

Motivations et choix algorithmiques

Notre objectif est de proposer un algorithme inconditionnellement stable, *i.e.* sans restriction CFL, assurant un principe du maximum pour chacune des entropies spécifiques. En conséquence de cela, nous serons assurés de la positivité de chacune de ces entropies, ainsi que de la préservation des constantes. Rappelons qu'au niveau continu, la validité d'un principe du maximum repose sur l'utilisation conjointe de la densité ρ et de l'entropie ρS_i (ici i est fixé) :

$$\partial_t \rho + \partial_x \rho u = 0, \tag{2.3}$$

$$\partial_t \rho S_i + \partial_x \rho u S_i = 0. \tag{2.4}$$

pour obtenir classiquement l'équation de transport suivante (et donc le principe du maximum recherché) :

$$\partial_t S_i + u \partial_x S_i = 0.$$

Remarquons que cette dernière identité vaut également au sens des distributions pour les solutions Lipschitz-continues par morceaux du système considéré, puisque la vitesse u et l'entropie spécifique S_i ne sont jamais discontinues simultanément. Il n'y a donc pas d'ambiguïté dans le produit non conservatif $u\partial_x S_i$.
Envisageons maintenant la validité des manipulations précédentes dans un contexte discret en temps et en espace. A cette fin, adoptons formellement dans une première étape une écriture conservative totalement implicite en temps pour considérer les analogues discrets de (2.3) et (2.4) suivants :

$$\rho_j^{n+1} = \rho_j^n - \frac{\Delta t}{\Delta x}((\rho u)_{j+1/2}^{n+1} - (\rho u)_{j-1/2}^{n+1}), \tag{2.5}$$

$$(\rho S_i)_j^{n+1} = (\rho S_i)_j^n - \frac{\Delta t}{\Delta x}((\rho u S_i)_{j+1/2}^{n+1} - (\rho u S_i)_{j-1/2}^{n+1}), \tag{2.6}$$

où conformément à la méthode de Relaxation proposée, la formule de Larrouturou [81] a lieu :

$$(\rho u S_i)_{j+1/2}^{n+1} = (\rho u)_{j+1/2,+}^{n+1} (S_i)_j^{n+1} + (\rho u)_{j+1/2,-}^{n+1} (S_i)_{j+1}^{n+1}, \tag{2.7}$$

avec la notation générique :

$$X_- = \min(X, 0), \quad X_+ = \max(X, 0), \quad \forall\, X \in I\!R. \tag{2.8}$$

Des manipulations algébriques simples conduisent à déduire de (2.5) et (2.6), avec des notations claires :

$$\begin{aligned} \rho_j^n (S_i)_j^{n+1} &= \rho_j^n (S_i)_j^n - \tfrac{\Delta t}{\Delta x}(\rho u)_{j+1/2}^{n+1}((S_i)_{j+1/2}^{n+1} - (S_i)_j^{n+1}) \\ &- \tfrac{\Delta t}{\Delta x}(\rho u)_{j-1/2}^{n+1}((S_i)_j^{n+1} - (S_i)_{j-1/2}^{n+1}), \end{aligned} \tag{2.9}$$

soit en vertu de (2.7), la formule suivante :

$$\begin{aligned} (S_i)_j^{n+1} &= (S_i)_j^n - \tfrac{\Delta t}{\Delta x}\tfrac{(\rho u)_{j+1/2,-}^{n+1}}{\rho_j^n}((S_i)_{j+1}^{n+1} - (S_i)_j^{n+1}) \\ &- \tfrac{\Delta t}{\Delta x}\tfrac{(\rho u)_{j-1/2,+}^{n+1}}{\rho_j^n}((S_i)_j^{n+1} - (S_i)_{j-1}^{n+1}). \end{aligned} \tag{2.10}$$

Nous montrerons plus loin (voir théorème 8) que sous une forme algébriquement identique, un principe du maximum a nécessairement lieu sans restriction CFL. La validité de (2.10) repose clairement sur l'identité $(\rho S_i)_j^{n+1} = \rho_j^{n+1} \times (S_i)_j^{n+1}$ où les formules de mises à jour de ρ_j^{n+1} et $(\rho S_i)_j^{n+1}$ sont liées par la formule de Larrouturou.
Dans la pratique, il est d'une part illusoire pour des raisons de coût calcul d'envisager une mise à jour de la densité ρ_j^{n+1} répondant à la formule complètement implicite (2.5). Une technique de linéarisation en temps classique lui sera préférée. D'autre part, nos efforts portent sur un découplage du bloc hydrodynamique et du bloc transport des entropies. Il semble difficile dans ces conditions d'opérer une linéarisation en temps des formules (2.5) et (2.6) permettant de garantir une formule algébriquement équivalente à la relation (2.10). Dans les faits, il s'avère difficile de proposer un algorithme assurant la constance éventuelle de l'entropie S_i, ce qui traduit une violation du principe du maximum.
Dans ces conditions, nous proposons d'opérer un découplage complet entre les blocs considérés en résolvant les équations de transport (*i.e.* en l'absence de ρ) :

$$\{\partial_t S_i + u\partial_x S_i = 0, \quad i = 1, ..., N-1. \tag{2.11}$$

La nature non conservative (et non ambiguë) de ces équations n'est pas rédhibitoire pour les applications que nous envisageons ici. Nous renvoyons en particulier le lecteur aux travaux de De Vuyst [44] qui illustrent l'intérêt de cette formulation dans d'autres applications. Nous proposons ci-après un algorithme conduisant naturellement à la satisfaction d'un principe du maximum sans restriction CFL, et possédant la propriété de n'avoir à inverser qu'un seul système linéaire tridiagonal scalaire pour toutes les entropies S_i, $i = 1, ..., N-1$.

Définition de l'algorithme

De manière à motiver l'introduction de l'algorithme implicite en temps, rappelons tout d'abord la discrétisation explicite des équations de transport (2.11) résultant d'une approche décentrée. Un champ de vitesse noté $\{u^*\}_{j+1/2}^n$ étant supposé donné à chaque interface $x_{j+1/2}$, la méthode de Godunov appliquée au système (2.11) s'écrit ici :

$$(S_i)_j^{n+1} = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \{S_i\}(\overline{\mathbf{u}}_{\Delta x}(x, t^{n+1-}))dx, \quad i = 1, ..., N-1, \quad j \in Z.$$

En intégrant (2.11) sous une condition CFL usuelle (≤ 1) et sur chaque pavé

$$(x_{j-1/2}, x_{j+1/2}) \times (t^n, t^{n+1}),$$

nous obtenons de manière classique la mise à jour décentrée et explicite suivante :

$$\begin{gathered}(S_i)_j^{n+1} = (S_i)_j^n - \tfrac{\Delta t}{\Delta x}\times \\ (\{u^*\}_{j-1/2,+}^n\{(S_i)_j^n - (S_i)_{j-1}^n\} + \{u^*\}_{j+1/2,-}^n\{(S_i)_{j+1}^n - (S_i)_j^n\}).\end{gathered} \tag{2.12}$$

Nous proposons une implicitation linéarisée de (2.12) en figeant à l'instant t^n les vitesses moyennes $\{u^*\}_{j+1/2}^n$. Nous obtenons pour tout $i = 1, ..., N-1$ et pour tout $j \in Z$ la mise à jour implicite suivante :

$$\begin{gathered}(S_i)_j^{n+1} = (S_i)_j^n - \tfrac{\Delta t}{\Delta x}\times \\ (\{u^*\}_{j-1/2,+}^n\{(S_i)_j^{n+1} - (S_i)_{j-1}^{n+1}\} + \{u^*\}_{j+1/2,-}^n\{(S_i)_{j+1}^{n+1} - (S_i)_j^{n+1}\}),\end{gathered} \tag{2.13}$$

permettant d'évaluer les quantités $\{(S_i)_j^{n+1}\}_{j\in Z}$. La mise en œuvre numérique correspondante consiste en la résolution de $(N-1)$ systèmes linéaires (un par entropie) dont la matrice globale notée A possède une structure tridiagonale à coefficients scalaires. De manière plus précise et d'après (2.13), une ligne courante j de la matrice A est constituée des éléments a_j, b_j et c_j vérifiant

$$\begin{aligned} a_j &= -\tfrac{\Delta t}{\Delta x}\{u^*\}_{j-1/2,+}^n, \\ b_j &= 1 + \tfrac{\Delta t}{\Delta x}(\{u^*\}_{j-1/2,+}^n - \{u^*\}_{j+1/2,-}^n), \\ c_j &= \tfrac{\Delta t}{\Delta x}\{u^*\}_{j+1/2,-}^n. \end{aligned} \tag{2.14}$$

La matrice A, identique pour chaque entropie S_i, est clairement une M-matrice à diagonale strictement dominante (par M-matrice nous entendons classiquement une matrice à diagonale positive et à termes extra-diagonaux négatifs ou nuls). Sous des conditions aux limites *adhoc* (non précisées ici) le système linéaire de taille finie résultant est donc inversible, et ce indépendamment du choix du pas de temps. Ceci assure l'existence et l'unicité des $(S_i)_j^{n+1}$. Une analyse plus fine des propriétés de cette matrice permet d'établir la stabilité de la méthode en exhibant les principes du maximum recherchés :

Théorème 8
La mise à jour implicite (2.13) des inconnues $\{(S_i)_j^{n+1}\}_{j\in Z}$ *pour tout* $i = 1, ..., N-1$ *admet une unique solution vérifiant les principes du maximum suivants :*

$$\begin{gathered}\min_{k\in Z}(S_i)_k^n \leq (S_i)_j^{n+1} \leq \max_{k\in Z}(S_i)_k^n, \\ i = 1, ..., N-1, \quad j \in Z.\end{gathered} \tag{2.15}$$

Démonstration
Définissons pour tout $n \in N$ et pour tout $i = 1, ..., N-1$ le vecteur $S_i^n = ((S_i)_j^n)_{j\in Z}$.

La mise à jour implicite (2.13) se réexprime sous la forme $AS_i^{n+1} = S_i^n$. Les coefficients de la matrice tridiagonale A étant donnés par (2.14), il est clair que le vecteur $v = (v_j)_{j\in Z}$ défini par $v_j = 1$ pour tout $j \in Z$ est un vecteur propre de la matrice A associé à la valeur propre 1.
Considérons maintenant les vecteurs $m_i = (m_{i,j})_{j\in Z}$ définis pour tout $i = 1, ..., N-1$ par $m_{i,j} = \min_{k\in Z}(S_i)_k^n$ pour tout $j \in Z$. Nous avons en particulier $m_i = \min_{k\in Z}(S_i)_k^n \times v$, de telle sorte que les vecteurs m_i sont également des vecteurs propres de la matrice A associés à la valeur propre 1.
Ainsi, il est clair que :

$$A(S_i^{n+1} - m_i) = AS_i^{n+1} - Am_i = AS_i^{n+1} - m_i = S_i^n - m_i, \quad i = 1, ..., N-1.$$

Par ailleurs, la matrice A étant une M-matrice (à diagonale dominante) son inverse est une matrice positive : tous ses coefficients sont positifs ou nuls (voir [95] pour plus de détails). Les vecteurs $S_i^n - m_i$ étant par construction trivialement positifs, il en est donc de même pour les vecteurs $S_i^{n+1} - m_i$, ce qui signifie que

$$\min_{k\in Z}(S_i)_k^n \leq (S_i)_j^{n+1}, \quad i = 1, ..., N-1, \quad j \in Z.$$

Un raisonnement similaire mené avec le vecteur $M_i = (M_{i,j})_{j\in Z}$ défini pour tout $i = 1, ..., N-1$ par $M_{i,j} = \max_{k\in Z}(S_i)_k^n$, $j \in Z$ permet d'obtenir la majoration attendue, à savoir

$$(S_i)_j^{n+1} \leq \max_{k\in Z}(S_i)_k^n, \quad i = 1, ..., N-1, \quad j \in Z.$$

Ceci conclut la démonstration.

2.1.2 Implicitation du bloc hydrodynamique

Motivations et choix algorithmiques

Par construction, la formulation explicite en temps de la méthode de volumes finis associée au bloc hydrodynamique repose sur la fonction flux numérique de Godunov. Or, il est connu que cette fonction flux n'est que Lipschitz-continue et non différentiable. Rappelons en effet que Larrouturou a établi qu'une méthode de volumes finis capturant exactement des discontinuités stationnaires (ici en l'occurence les discontinuités de contact) ne peut pas être différentiable. Nous avons volontairement privilégié cette propriété de précision (pour des raisons déjà évoquées) au détriment de la propriété de différentiabilité pourtant fort agréable dans le contexte d'une implicitation linéarisée en temps. Puisque pour des raisons de coût, il est indispensable de procéder à une telle linéarisation, il importe d'isoler les défauts de différentiabilité dans la fonction flux considérée, ce qui semble difficile en l'état. De manière à contourner cette difficulté, nous montrons dans cette section que la méthode de Godunov ici développée peut être réinterprétée de manière équivalente comme une méthode de Roe (pour le système de Relaxation (1.87) avec $\lambda = 0$ et non pas bien sûr pour le système d'équilibre (1.1)). L'existence d'une réécriture du flux basée sur une linéarisation de Roe permet alors classiquement d'isoler les défauts de régularité en nous ramenant à des algorithmes implicites linéarisés largement éprouvés dans d'autres applications (voir par exemple les références classiques [108], [147], ...). La réinterprétation en termes de méthode de Roe est suggérée par une observation due à Mehlman [104], consistant à reconnaître que lorsque localement la solution d'un problème de Riemann pour le système des équations d'Euler n'est composée que de discontinuités (choc-contact-choc), alors la méthode de Godunov associée est localement équivalente à une méthode de Roe. Dans notre contexte, la validité de cette réinterprétation est uniforme puisque par la propriété de dégénérescence linéaire partagée par tous les champs du système (1.87) avec $\lambda = 0$, toutes les ondes composant la solution du problème de Riemann sont systématiquement des discontinuités. Nous établissons ce résultat de manière rigoureuse ci-dessous.

Définition de l'algorithme

Nous définissons au préalable quelques notations nécessaires pour la suite.
Soient $\overline{\mathbf{u}}_L$ et $\overline{\mathbf{u}}_R$ deux états de $\overline{\Theta}$. Nous introduisons les vecteurs $\overline{\mathbf{w}}_L = \overline{\mathbf{w}}(\overline{\mathbf{u}}_L)$ et $\overline{\mathbf{w}}_R = \overline{\mathbf{w}}(\overline{\mathbf{u}}_R)$ de $I\!R^4$ définis par

$$\begin{aligned} \overline{\mathbf{w}}_L = \overline{\mathbf{w}}(\overline{\mathbf{u}}_L) = (\rho_L, (\rho u)_L, (\rho E)_L, (\rho I)_L), \\ \overline{\mathbf{w}}_R = \overline{\mathbf{w}}(\overline{\mathbf{u}}_R) = (\rho_R, (\rho u)_R, (\rho E)_R, (\rho I)_R), \end{aligned} \tag{2.16}$$

et définissons la solution auto-semblable $\overline{\mathbf{w}}(x/t; \overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)$ du système (2.2) pour la donnée initiale

$$\overline{\mathbf{w}}_0(x) = \begin{cases} \overline{\mathbf{w}}_L & \text{si} \quad x < 0, \\ \overline{\mathbf{w}}_R & \text{si} \quad x > 0. \end{cases} \tag{2.17}$$

En reprenant les notations du théorème 7, nous avons donc

$$\overline{\mathbf{w}}(x/t; \overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R) = \begin{cases} \overline{\mathbf{w}}_L & \text{si} \quad \frac{x}{t} < \lambda_1(\overline{\mathbf{w}}_L), \\ \overline{\mathbf{w}}_1 & \text{si} \quad \lambda_1(\overline{\mathbf{w}}_L) < \frac{x}{t} < \lambda_2(\overline{\mathbf{w}}_1), \\ \overline{\mathbf{w}}_2 & \text{si} \quad \lambda_2(\overline{\mathbf{w}}_2) < \frac{x}{t} < \lambda_3(\overline{\mathbf{w}}_R), \\ \overline{\mathbf{w}}_R & \text{si} \quad \lambda_3(\overline{\mathbf{w}}_R) < \frac{x}{t}, \end{cases} \tag{2.18}$$

avec

$$\overline{\mathbf{w}}_1 = (\rho_1, (\rho u)_1, (\rho\Sigma)_1, (\rho I)_1), \quad \overline{\mathbf{w}}_2 = (\rho_2, (\rho u)_2, (\rho\Sigma)_2, (\rho I)_2).$$

Soit $(\overline{\mathbf{w}}_j^n)_{j\in Z}$ la suite des vecteurs de $I\!R^4$ à l'équilibre déduits de $(\overline{\mathbf{u}}_j^n)_{j\in Z}$ et définis par

$$\overline{\mathbf{w}}_j^n = \overline{\mathbf{w}}(\overline{\mathbf{u}}_j^n) = (\rho_j^n, (\rho u)_j^n, (\rho E)_j^n, (\rho I)_j^n), \quad j \in Z,$$

$$(\rho I)_j^n = \rho_j^n \textstyle\sum_{i=1}^N p_i(1/\rho_j^n, (\rho S_i)_j^n/\rho_j^n) + a^2.$$

La méthode de Godunov appliquée au système (2.2) se réécrit alors sous la condition CFL ≤ 1 sous la forme conservative suivante :

$$\overline{\mathbf{w}}_j^{n+1} = \overline{\mathbf{w}}_j^n - \frac{\Delta t}{\Delta x}(\overline{\mathbf{f}}^g(\overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n) - \overline{\mathbf{f}}^g(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n)), \quad j \in Z, \tag{2.19}$$

où le flux numérique $\overline{\mathbf{f}}^g(\overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n)$ est défini pour tout $j \in Z$ par

$$\overline{\mathbf{f}}^g(\overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n) = \overline{\mathbf{f}}^e(\overline{\mathbf{w}}(0^+; \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n)). \tag{2.20}$$

Le flux exact $\overline{\mathbf{f}}^e$ du système (2.2) trouve quant à lui la définition

$$\overline{\mathbf{f}}^e(\rho, \rho u, \rho E, \rho I) = (\rho u, \rho u^2 + \Pi, (\rho E + \Pi)u, \rho I u), \quad I = \Pi + a^2/\rho.$$

De manière à réinterpréter le flux numérique de la méthode de Godunov (2.20) dans le formalisme des flux numériques associés à une méthode de Roe, la première étape de notre raisonnement consiste à réexprimer de manière équivalente ce flux numérique selon le résultat suivant :

Proposition 3
Soient $\overline{\mathbf{u}}_L$ et $\overline{\mathbf{u}}_R$ deux états de $\overline{\Theta}$ et $\overline{\mathbf{w}}_L$ et $\overline{\mathbf{w}}_R$ les vecteurs correspondants de $I\!R^4$ définis par (2.16). Soit $\overline{\mathbf{w}}(x/t; \overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)$ la solution auto-semblable définie par (2.18) du problème de Riemann (2.2) pour la donnée initiale (2.17).
Alors, le flux numérique $\overline{\mathbf{f}}^g(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)$ associé à la méthode de Godunov et défini par $\overline{\mathbf{f}}^g(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R) = \overline{\mathbf{f}}^e(\overline{\mathbf{w}}(0^+; \overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R))$ trouve la formulation équivalente suivante :

$$\begin{gathered}\overline{\mathbf{f}}^g(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R) = \tfrac{1}{2}(\overline{\mathbf{f}}^e(\overline{\mathbf{w}}_L) + \overline{\mathbf{f}}^e(\overline{\mathbf{w}}_R)) - \\ \tfrac{1}{2}(|\lambda_1(\overline{\mathbf{w}}_L)|(\overline{\mathbf{w}}_1 - \overline{\mathbf{w}}_L) - |\lambda_2(\overline{\mathbf{w}}_1)|(\overline{\mathbf{w}}_2 - \overline{\mathbf{w}}_1) - |\lambda_3(\overline{\mathbf{w}}_R)|(\overline{\mathbf{w}}_R - \overline{\mathbf{w}}_2)).\end{gathered} \tag{2.21}$$

Démonstration
Soit $\overline{\mathbf{w}}(x/t; \overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)$ la solution auto-semblable définie par (2.18) du problème de Riemann (2.2) pour la donnée initiale (2.17). Nous définissons les quantités

$$\overline{\mathbf{w}}_l = \frac{2}{\Delta x} \int_{-\frac{\Delta x}{2}}^{0} \overline{\mathbf{w}}(x/\Delta t; \overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R) dx,$$

$$\overline{\mathbf{w}}_r = \frac{2}{\Delta x} \int_{0}^{\frac{\Delta x}{2}} \overline{\mathbf{w}}(x/\Delta t; \overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R) dx.$$

La solution du problème de Riemann n'étant composée que d'états constants (voir (2.18)), nous obtenons facilement

$$\begin{gathered}\overline{\mathbf{w}}_l = \overline{\mathbf{w}}_L - \tfrac{2\Delta t}{\Delta x} \times \\ (\lambda_1^-(\overline{\mathbf{w}}_L)(\overline{\mathbf{w}}_1 - \overline{\mathbf{w}}_L) + \lambda_2^-(\overline{\mathbf{w}}_1)(\overline{\mathbf{w}}_2 - \overline{\mathbf{w}}_1) + \lambda_3^-(\overline{\mathbf{w}}_R)(\overline{\mathbf{w}}_R - \overline{\mathbf{w}}_2)),\end{gathered}$$

et

$$\begin{gathered}\overline{\mathbf{w}}_r = \overline{\mathbf{w}}_R - \tfrac{2\Delta t}{\Delta x} \times \\ (\lambda_1^+(\overline{\mathbf{w}}_L)(\overline{\mathbf{w}}_1 - \overline{\mathbf{w}}_L) + \lambda_2^+(\overline{\mathbf{w}}_1)(\overline{\mathbf{w}}_2 - \overline{\mathbf{w}}_1) + \lambda_3^+(\overline{\mathbf{w}}_R)(\overline{\mathbf{w}}_R - \overline{\mathbf{w}}_2)),\end{gathered}$$

où nous avons de nouveau utilisé les notations génériques (2.8).
En intégrant le système (2.2) sur les rectangles $(-\Delta x/2, 0) \times (0, \Delta t)$ (respectivement $(0, \Delta x/2) \times (0, \Delta t)$), des arguments classiques conduisent par ailleurs à :

$$\overline{\mathbf{w}}_l = \overline{\mathbf{w}}_L - \frac{2\Delta t}{\Delta x}(\overline{\mathbf{f}}^g(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R) - \overline{\mathbf{f}}^e(\overline{\mathbf{w}}_L)),$$

et

$$\overline{\mathbf{w}}_r = \overline{\mathbf{w}}_R - \frac{2\Delta t}{\Delta x}(\overline{\mathbf{f}}^e(\overline{\mathbf{w}}_R) - \overline{\mathbf{f}}^g(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)).$$

Il est ainsi possible d'obtenir deux nouvelles expressions du flux numérique de la méthode de Godunov, à savoir

$$\overline{\mathbf{f}}^g(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R) = \overline{\mathbf{f}}^e(\overline{\mathbf{w}}_L) +$$

$$\lambda_1^-(\overline{\mathbf{w}}_L)(\overline{\mathbf{w}}_1 - \overline{\mathbf{w}}_L) - \lambda_2^-(\overline{\mathbf{w}}_1)(\overline{\mathbf{w}}_2 - \overline{\mathbf{w}}_1) - \lambda_3^-(\overline{\mathbf{w}}_R)(\overline{\mathbf{w}}_R - \overline{\mathbf{w}}_2),$$

et

$$\overline{\mathbf{f}}^g(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R) = \overline{\mathbf{f}}^e(\overline{\mathbf{w}}_R) -$$

$$\lambda_1^+(\overline{\mathbf{w}}_L)(\overline{\mathbf{w}}_1 - \overline{\mathbf{w}}_L) - \lambda_2^+(\overline{\mathbf{w}}_1)(\overline{\mathbf{w}}_2 - \overline{\mathbf{w}}_1) - \lambda_3^+(\overline{\mathbf{w}}_R)(\overline{\mathbf{w}}_R - \overline{\mathbf{w}}_2).$$

Enfin, puisque

$$|\lambda_i| = \lambda_i^+ - \lambda_i^-,$$

nous réexprimons ce flux en prenant la demi-somme des deux expressions précédentes pour obtenir :

$$\overline{\mathbf{f}}^g(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R) = \tfrac{1}{2}(\overline{\mathbf{f}}^e(\overline{\mathbf{w}}_L) + \overline{\mathbf{f}}^e(\overline{\mathbf{w}}_R)) -$$

$$\tfrac{1}{2}(|\lambda_1(\overline{\mathbf{w}}_L)|(\overline{\mathbf{w}}_1 - \overline{\mathbf{w}}_L) - |\lambda_2(\overline{\mathbf{w}}_1)|(\overline{\mathbf{w}}_2 - \overline{\mathbf{w}}_1) - |\lambda_3(\overline{\mathbf{w}}_R)|(\overline{\mathbf{w}}_R - \overline{\mathbf{w}}_2)).$$

Ceci conclut la démonstration.

Pour comprendre ce flux de Godunov dans le formalisme des méthodes de Roe, il reste donc à montrer l'existence d'une matrice de Roe notée $\overline{\mathbf{A}}(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)$ satisfaisant les trois propriétés de consistance suivantes (voir [113] pour plus de détails) :

$$\begin{array}{l} 1.\ \overline{\mathbf{A}}(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R) \quad \text{est diagonalisable à valeurs propres réelles,} \\ 2.\ \overline{\mathbf{A}}(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)(\overline{\mathbf{w}}_R - \overline{\mathbf{w}}_L) = \overline{\mathbf{f}}^e(\overline{\mathbf{w}}_R) - \overline{\mathbf{f}}^e(\overline{\mathbf{w}}_L), \\ 3.\ \overline{\mathbf{A}}(\overline{\mathbf{w}}, \overline{\mathbf{w}}) = \nabla \overline{\mathbf{f}}^e(\overline{\mathbf{w}}), \end{array} \tag{2.22}$$

ainsi que

$$\begin{array}{l} 4.\ |\overline{\mathbf{A}}|(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)(\overline{\mathbf{w}}_R - \overline{\mathbf{w}}_L) = \\ (|\lambda_1(\overline{\mathbf{w}}_L)|(\overline{\mathbf{w}}_1 - \overline{\mathbf{w}}_L) + |\lambda_2(\overline{\mathbf{w}}_1)|(\overline{\mathbf{w}}_2 - \overline{\mathbf{w}}_1) + |\lambda_3(\overline{\mathbf{w}}_R)|(\overline{\mathbf{w}}_R - \overline{\mathbf{w}}_2)). \end{array} \tag{2.23}$$

En effet, dès que cette dernière relation est vérifiée, la proposition précédente montre que la fonction flux de Godunov s'écrit comme une fonction flux de Roe :

$$\overline{\mathbf{f}}^g(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R) = \tfrac{1}{2}(\overline{\mathbf{f}}^e(\overline{\mathbf{w}}_L) + \overline{\mathbf{f}}^e(\overline{\mathbf{w}}_R)) - \tfrac{1}{2}|\overline{\mathbf{A}}|(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)(\overline{\mathbf{w}}_R - \overline{\mathbf{w}}_L). \tag{2.24}$$

L'existence d'une telle linéarisée de Roe résulte du résultat suivant.

Théorème 9

Soient $\overline{\mathbf{u}}_L$ et $\overline{\mathbf{u}}_R$ deux états de $\overline{\Theta}$ et $\overline{\mathbf{w}}_L$ et $\overline{\mathbf{w}}_R$ les vecteurs correspondants de $\mathbb{R}^4$ définis par (2.16). En reprenant les notations de la proposition précédente et du théorème 7, nous définissons les vecteurs colonnes R_1, R_2^1, R_2^2, et R_3 en posant

$$R_1 = \begin{pmatrix} 1 \\ u_L - a/\rho_L \\ I_L \\ E_L + \Pi_L/\rho_L - au^\star/\rho_L \end{pmatrix}, \quad R_2^1 = \begin{pmatrix} 1 \\ u^\star \\ \Pi^\star \\ 0 \end{pmatrix},$$

$$R_2^2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad R_3 = \begin{pmatrix} 1 \\ u_R + a/\rho_R \\ I_R \\ E_R + \Pi_R/\rho_R + au^\star/\rho_R \end{pmatrix},$$

de telle sorte que

$$\begin{aligned} &\overline{\mathbf{w}}_1 - \overline{\mathbf{w}}_L = (\rho_1 - \rho_L)R_1, \\ &\overline{\mathbf{w}}_2 - \overline{\mathbf{w}}_1 = (\rho_2 - \rho_1)R_2^1 + (\rho_2\Sigma_2 - \rho_1\Sigma_1)R_2^2, \\ &\overline{\mathbf{w}}_R - \overline{\mathbf{w}}_2 = (\rho_R - \rho_2)R_3. \end{aligned} \tag{2.25}$$

Alors, ces vecteurs forment une base de $I\!R^4$ dès que a vérifie la condition de séparation des valeurs propres (1.90).
De plus, la matrice $\overline{\mathbf{A}}(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)$ définie à partir de sa forme diagonalisée par

$$\begin{gathered} \overline{\mathbf{A}}(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R) = B(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R) \begin{pmatrix} \lambda_1(\overline{\mathbf{w}}_L) & 0 & 0 & 0 \\ 0 & \lambda_2(\overline{\mathbf{w}}_1) & 0 & 0 \\ 0 & 0 & \lambda_2(\overline{\mathbf{w}}_2) & 0 \\ 0 & 0 & 0 & \lambda_3(\overline{\mathbf{w}}_R) \end{pmatrix} B^{-1}(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R), \\ B(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R) = [R_1, R_2^1, R_2^2, R_3](\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R), \end{gathered} \tag{2.26}$$

vérifie les quatre propriétés de consistance (2.22)-(2.23).

Démonstration

En utilisant les valeurs explicites (voir théorème 7) des vecteurs $\overline{\mathbf{w}}_1$ et $\overline{\mathbf{w}}_2$, il est facile de vérifier que les trois relations définitions (2.25) conduisent aux vecteurs colonnes proposés. Rappelons que la condition (1.90) sur a impose aux densités intermédiaires d'être strictement positives. Des calculs simples exploitant l'ensemble des invariants vérifiés par la solution du problème de Riemann (voir théorème 7) montrent que ces quatre vecteurs sont nécessairement linéairement indépendants. En conséquence, la matrice B^{-1} existe et la linéarisée proposée est bien définie.
Ensuite, la troisième propriété de consistance est satisfaite, puisqu'un calcul direct montre que les vecteurs propres de la jacobienne du flux exact $\nabla\overline{\mathbf{f}}^e(\overline{\mathbf{w}})$ évalués en $\overline{\mathbf{w}} = (\rho, \rho u, \rho\Sigma, \rho I)$ sont précisément $(1, u - a/\rho, I, \Sigma + \Pi/\rho - au/\rho), (1, u, \Pi, 0), (0, 0, 0, 1)$ et $(1, u + a/\rho, I, \Sigma + \Pi/\rho + au/\rho)$ respectivement associés aux valeurs propres $\lambda_1(\overline{\mathbf{w}})$, $\lambda_2(\overline{\mathbf{w}})$ (valeur propre d'ordre 2), $\lambda_3(\overline{\mathbf{w}})$. Nous rappelons en effet que $\Pi = I - a^2/\rho$.
La première propriété étant vraie par construction, il reste à démontrer la validité des propriétés 2 et 4.
Par construction, les vecteurs R_1, R_2^1, R_2^2, et R_3 sont des vecteurs propres de la matrice $\overline{\mathbf{A}}(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)$.
Ainsi, d'après (2.25) :

$$\begin{aligned} \overline{\mathbf{A}}(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)(\overline{\mathbf{w}}_R - \overline{\mathbf{w}}_L) &= \overline{\mathbf{A}}(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)(\overline{\mathbf{w}}_1 - \overline{\mathbf{w}}_L) + \overline{\mathbf{A}}(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)(\overline{\mathbf{w}}_2 - \overline{\mathbf{w}}_1) \\ &+ \overline{\mathbf{A}}(\overline{\mathbf{w}}_L, \overline{\mathbf{w}}_R)(\overline{\mathbf{w}}_R - \overline{\mathbf{w}}_2) \\ &= \lambda_1(\overline{\mathbf{w}}_L)(\overline{\mathbf{w}}_1 - \overline{\mathbf{w}}_L) + \lambda_2(\overline{\mathbf{w}}_1)(\overline{\mathbf{w}}_2 - \overline{\mathbf{w}}_1) \\ &+ \lambda_3(\overline{\mathbf{w}}_R)(\overline{\mathbf{w}}_R - \overline{\mathbf{w}}_2). \end{aligned}$$

Or, les valeurs propres $\lambda_1(\overline{\mathbf{w}}_L)$, $\lambda_2(\overline{\mathbf{w}}_1)$, $\lambda_3(\overline{\mathbf{w}}_R)$ ne sont autres que les vitesses de propagation des discontinuités composant la solution du problème de Riemann. En invoquant les conditions de Rankine-Hugoniot associées, nous déduisons de l'identité précédente :

$$\begin{aligned}\overline{\mathbf{A}}(\overline{\mathbf{w}}_L,\overline{\mathbf{w}}_R)(\overline{\mathbf{w}}_R-\overline{\mathbf{w}}_L) &= \overline{\mathbf{f}}^e(\overline{\mathbf{w}}_1)-\overline{\mathbf{f}}^e(\overline{\mathbf{w}}_L)+\overline{\mathbf{f}}^e(\overline{\mathbf{w}}_2)-\overline{\mathbf{f}}^e(\overline{\mathbf{w}}_1)\\ &+ \overline{\mathbf{f}}^e(\overline{\mathbf{w}}_R)-\overline{\mathbf{f}}^e(\overline{\mathbf{w}}_2)\\ &= \overline{\mathbf{f}}^e(\overline{\mathbf{w}}_R)-\overline{\mathbf{f}}^e(\overline{\mathbf{w}}_L).\end{aligned}$$

La deuxième propriété de consistance est donc démontrée.
De la même manière et puisque $|\overline{\mathbf{A}}|(\overline{\mathbf{w}}_L,\overline{\mathbf{w}}_R)$ vaut par définition :

$$|\overline{\mathbf{A}}|(\overline{\mathbf{w}}_L,\overline{\mathbf{w}}_R)=B\begin{pmatrix}|\lambda_1(\overline{\mathbf{w}}_L)| & 0 & 0 & 0\\ 0 & |\lambda_2(\overline{\mathbf{w}}_1)| & 0 & 0\\ 0 & 0 & |\lambda_2(\overline{\mathbf{w}}_2)| & 0\\ 0 & 0 & 0 & |\lambda_3(\overline{\mathbf{w}}_R)|\end{pmatrix}B^{-1}, \tag{2.27}$$

les vecteurs R_1, R_2^1, R_2^2, et R_3 sont des vecteurs propres de la matrice $|\overline{\mathbf{A}}|(\overline{\mathbf{w}}_L,\overline{\mathbf{w}}_R)$. Nous obtenons donc la quatrième propriété d'après les relations définitions (2.25), ce qui conclut la démonstration.

Quelques rappels sur l'implicitation linéarisée d'une méthode de Roe et choix de a

Les brefs rappels donnés ci-dessous sont utiles pour préciser le choix du paramètre a entrant dans la définition du système de Relaxation. Ils ne sont en aucun cas une présentation exhaustive des techniques d'implicitation linéarisée associées à la méthode de Roe. Il existe en effet dans la littérature au moins deux implicitations linéarisées de la méthode de Roe (voir par exemple [147]), l'une conservative, l'autre non, sauf à convergence de l'algorithme en temps vers une solution stationnaire. Puisque notre objectif ne concerne que les solutions stationnaires, nous décrivons brièvement l'une d'entre elles, à savoir celle non conservative (dans la phase instationnaire) car elle permet de mieux appréhender les raisons heuristiques du succès dans les applications pratiques de la linéarisation proposée. Nous suivons ici une présentation très classiquement adoptée dans l'immense littérature consacrée au sujet.
Introduisons la notation $\overline{\mathbf{A}}^n_{j+1/2}=\overline{\mathbf{A}}(\overline{\mathbf{w}}^n_j,\overline{\mathbf{w}}^n_{j+1})$. La deuxième propriété de (2.22) implique

$$\overline{\mathbf{f}}^e(\overline{\mathbf{w}}^n_{j+1})-\overline{\mathbf{f}}^e(\overline{\mathbf{w}}^n_j)=\overline{\mathbf{A}}^n_{j+1/2}(\overline{\mathbf{w}}^n_{j+1}-\overline{\mathbf{w}}^n_j),$$

de telle sorte que (2.19) peut être réécrit, grâce à (2.24), sous la forme :

$$\begin{gathered}\overline{\mathbf{w}}^{n+1}_j=\overline{\mathbf{w}}^n_j-\tfrac{\Delta t}{2\Delta x}\times\\ ((\overline{\mathbf{A}}^n_{j-1/2}+|\overline{\mathbf{A}}|^n_{j-1/2})(\overline{\mathbf{w}}^n_j-\overline{\mathbf{w}}^n_{j-1})+(\overline{\mathbf{A}}^n_{j+1/2}-|\overline{\mathbf{A}}|^n_{j+1/2})(\overline{\mathbf{w}}^n_{j+1}-\overline{\mathbf{w}}^n_j)).\end{gathered} \tag{2.28}$$

En définissant les matrices $\overline{\mathbf{A}}^+(\overline{\mathbf{w}}_L,\overline{\mathbf{w}}_R)$ et $\overline{\mathbf{A}}^-(\overline{\mathbf{w}}_L,\overline{\mathbf{w}}_R)$ par

$$\overline{\mathbf{A}}^\pm(\overline{\mathbf{w}}_L,\overline{\mathbf{w}}_R)=B\begin{pmatrix}\lambda_1^\pm(\overline{\mathbf{w}}_L) & 0 & 0 & 0\\ 0 & \lambda_2^\pm(\overline{\mathbf{w}}_1) & 0 & 0\\ 0 & 0 & \lambda_2^\pm(\overline{\mathbf{w}}_2) & 0\\ 0 & 0 & 0 & \lambda_3^\pm(\overline{\mathbf{w}}_R)\end{pmatrix}B^{-1}, \tag{2.29}$$

nous obtenons facilement

$$|\overline{\mathbf{A}}| = \overline{\mathbf{A}}^{+} - \overline{\mathbf{A}}^{-}, \quad \overline{\mathbf{A}} = \overline{\mathbf{A}}^{+} + \overline{\mathbf{A}}^{-}.$$

Ainsi, (2.28) devient

$$\begin{aligned} \overline{\mathbf{w}}_j^{n+1} &= \overline{\mathbf{w}}_j^n - \tfrac{\Delta t}{\Delta x}\times \\ (\overline{\mathbf{A}}^{+n}_{j-1/2}(\overline{\mathbf{w}}_j^n - \overline{\mathbf{w}}_{j-1}^n) &+ \overline{\mathbf{A}}^{-n}_{j+1/2}(\overline{\mathbf{w}}_{j+1}^n - \overline{\mathbf{w}}_j^n)), \end{aligned} \tag{2.30}$$

où l'on a posé

$$\overline{\mathbf{A}}^{+n}_{j+1/2} = \overline{\mathbf{A}}^{+}(\overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n), \quad \overline{\mathbf{A}}^{-n}_{j+1/2} = \overline{\mathbf{A}}^{-}(\overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n), \quad j \in Z.$$

L'implicitation linéarisée de (2.30) est classiquement obtenue en figeant à l'instant t^n les matrices $\overline{\mathbf{A}}^{+n}_{j-1/2}$ et $\overline{\mathbf{A}}^{-n}_{j+1/2}$. Ainsi, la suite $(\overline{\mathbf{w}}_j^{n+1})_{j\in Z}$ est solution de :

$$\begin{aligned} \overline{\mathbf{w}}_j^{n+1} &= \overline{\mathbf{w}}_j^n - \tfrac{\Delta t}{\Delta x}\times \\ (\overline{\mathbf{A}}^{+n}_{j-1/2}(\overline{\mathbf{w}}_j^{n+1} - \overline{\mathbf{w}}_{j-1}^{n+1}) &+ \overline{\mathbf{A}}^{-n}_{j+1/2}(\overline{\mathbf{w}}_{j+1}^{n+1} - \overline{\mathbf{w}}_j^{n+1})), \end{aligned} \tag{2.31}$$

où rappelons-le, $\overline{\mathbf{w}}_j^n = (\rho_j^n, (\rho u)_j^n, (\rho E)_j^n, (\rho I)_j^n)$ pour tout $j \in Z$ et tout $n \in N$. La mise en œuvre numérique consiste en la résolution d'un système linéaire

$$\mathcal{A}\overline{\mathbf{w}}_j^{n+1} = \overline{\mathbf{w}}_j^n \tag{2.32}$$

où la matrice $\mathcal{A}$ possède une structure tridiagonale par blocs, chaque bloc représentant une matrice de taille 4×4 dans notre cas.
Plus précisément, une ligne courante j de la matrice est constituée des blocs élémentaires $\mathbf{a}_j, \mathbf{b}_j$ et $\mathbf{c}_j$ donnés par :

$$\begin{aligned} \mathbf{a}_j &= -\tfrac{\Delta t}{\Delta x}\overline{\mathbf{A}}^{+n}_{j-1/2}, \\ \mathbf{b}_j &= Id + \tfrac{\Delta t}{\Delta x}(\overline{\mathbf{A}}^{+n}_{j-1/2} - \overline{\mathbf{A}}^{-n}_{j+1/2}), \\ \mathbf{c}_j &= \tfrac{\Delta t}{\Delta x}\overline{\mathbf{A}}^{-n}_{j+1/2}. \end{aligned} \tag{2.33}$$

Pour achever la descritpion de l'implicitation du bloc hydrodynamique, il convient de préciser à chaque pas de temps le choix du paramètre a entrant dans la définition de la méthode de Relaxation. A cette fin, reformulons l'algorithme d'implicitation proposé en adoptant une écriture incrémentale en temps classique (voir [147]). Il est facile de constater que la résolution du problème linéaire (2.32) est équivalente à celle du problème incrémental suivant :

$$\begin{aligned} \mathcal{A}\Delta\overline{\mathbf{w}}_j^{n+1/2} &= \overline{\mathbf{r}}(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n), \quad \Delta\overline{\mathbf{w}}_j^{n+1/2} = \frac{\overline{\mathbf{w}}_j^{n+1} - \overline{\mathbf{w}}_j^n}{\Delta t}, \\ \overline{\mathbf{r}}(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n) &:= -\tfrac{1}{\Delta x}\{\overline{\mathbf{f}}^g(\overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n) - \overline{\mathbf{f}}^g(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n)\}, \end{aligned} \tag{2.34}$$

puisque

$$\overline{\mathbf{f}}^g(\overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n) - \overline{\mathbf{f}}^g(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n) = (\overline{\mathbf{A}}^{+n}_{j-1/2}(\overline{\mathbf{w}}_j^n - \overline{\mathbf{w}}_{j-1}^n) + \overline{\mathbf{A}}^{-n}_{j+1/2}(\overline{\mathbf{w}}_{j+1}^n - \overline{\mathbf{w}}_j^n)). \tag{2.35}$$

Pour obtenir cette identité, nous avons évidemment supposé que la valeur de a entrant dans la définition de la matrice $\mathcal{A}$ et du membre de droite (appelé communément résidu explicite) est identique. Toutefois, l'expérience de la mécanique des fluides numérique indique qu'il est souvent préférable d'utiliser une matrice $\mathcal{A}$ provenant de la linéarisation d'une fonction flux numérique dissipative dans un souci de robustesse, tout en conservant une définition du résidu explicite associée à une fonction flux choisie pour ses qualités de résolution. De sorte qu'à convergence en temps, la solution stationnaire est décrite par cette dernière fonction flux. Conformément à cette heuristique, nous évaluons la partie résidu explicite à chaque pas de temps en faisant choix d'une valeur de a locale satisfaisant au plus près les conditions de Whitham exhibées au théorème 5. Concernant la définition de la matrice $\mathcal{A}$, nous avons adopté un choix de a global résultant du maximum des valeurs précédentes.

Ceci achève la description de l'implicitation du bloc hydrodynamique (2.2). Pour conclure cette section, rapportons brièvement les considérations classiques visant à justifier formellement la linéarisation proposée. Les identités (2.33) suggèrent que la matrice tridiagonale par blocs ainsi obtenue pourrait être associée, via un changement de base (conduisant aux variables caractéristiques), à une M-matrice : matrice à diagonale positive (dominante) et à termes extra-diagonaux négatifs. Ceci ne vaut bien entendu que si les parties positive $\overline{\mathbf{A}}^+{}_{j+1/2}$ et négative $\overline{\mathbf{A}}^-{}_{j+1/2}$ se diagonalisent pour tout j dans la même base. Ceci est faux en toute rigueur mais l'algorithme fonctionne remarquablement bien dans la pratique.

2.1.3 Projection sur la variété d'équilibre et défaut de convergence

Les algorithmes d'implicitation des blocs hydrodynamique et transport des entropies permettent de remettre à jour la solution discrète relaxée $(\rho, \rho u, \rho E, \{\rho S_i\}_{i=1,\ldots,N-1}, \rho I)$ et donc de redéfinir à l'issue de cette réactualisation une nouvelle valeur de ρI à l'équilibre à l'instant considéré :

$$\rho I = \rho \sum_{i=1}^{N} p_i(1/\rho, \rho S_i/\rho) + a^2, \tag{2.36}$$

pour un nouveau choix de a tel qu'il vient d'être discuté. Les expériences de validation du bloc hydrodynamique menées dans un même code de calculs ont montré qu'il était possible d'adopter une stratégie de montée en nombre CFL comparable à celle d'une méthode de Van Leer linéairement implicitée (dans le cadre des équations d'Euler usuelles, voir la présentation des résultats numériques). Comme il est communément admis, cette observation indique une bonne stabilité de l'algorithme d'implicitation proposé.

Toutefois, l'historique de décroissance de la norme L^2 des résidus explicites fait apparaître après quelques décades de décroissance une valeur limite, synonyme de non convergence vers une solution parfaitement stationnaire en temps ($||\overline{\mathbf{r}}(\overline{\mathbf{w}}^n_{j-1}, \overline{\mathbf{w}}^n_j, \overline{\mathbf{w}}^n_{j+1})|| = 0$, condition de convergence en temps que nous préciserons ci-dessous). L'objectif de cette section est d'expliquer pourquoi l'implicitation en temps du bloc hydrodynamique ne peut pas en règle générale conduire à des solutions rigoureusement stationnaires. La mise en évidence de l'origine de cette impossibilité permettra d'introduire une technique de correction simple et stable (comme le montrent les expériences numériques). Il suffit, ainsi que nous le verrons, de restreindre la discussion au cadre des équations d'Euler 3×3 puisque seule est incriminée ici l'implicitation du bloc hydrodynamique.

Précisons dès maintenant notre critère de convergence en temps vers une solution stationnaire. Dans le cadre classique, une solution est dite stationnaire si la norme L^2 de tous les résidus explicites est nulle. L'inversibilité de la matrice $\mathcal{A}$ (sous des conditions aux limites *adhoc*) implique en conséquence que tous les incréments en temps sont nuls. Toutefois dans notre contexte, seule la convergence en temps vers un stationnaire des variables densité, quantité de mouvement et énergie totale nous intéresse dans la pratique. Selon notre stratégie d'implicitation linéarisée, l'évolution en temps de ces variables est cependant couplée à celle de la variable de Relaxation ρI, variable sur laquelle il sera difficile, sinon impossible, d'exprimer une notion de convergence en temps (voir ci-dessous). Dès lors et de manière à définir un critère de convergence sur ρ, ρu et ρE, nous exigerons non seulement des résidus associés de converger vers zéro, mais également des incréments en temps de ces mêmes variables de converger vers zéro. Ce critère peut paraître exigeant mais il semble malaisé d'en proposer un plus faible lorsque l'une des variables d'évolution, ici ρI, linéairement couplée aux autres via le problème linéaire (2.34) sera vue ne pas pouvoir converger en temps. Avec des notations claires, nous exigeons donc :

$$\overline{\mathbf{r}}_\rho(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n) = \overline{\mathbf{r}}_{\rho u}(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n) = \overline{\mathbf{r}}_{\rho\Sigma}(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n) = 0,$$

et

$$\Delta\rho_j^{n+1/2} = \Delta\rho u_j^{n+1/2} = \Delta\rho\Sigma_j^{n+1/2} = 0.$$

Par opposition aux variables ρ, ρu et ρE, la variable de Relaxation ρI connaît deux mises à jour distinctes au cours de chaque itération : une prédiction via le système de Relaxation puis une correction par projection sur la variété d'équilibre. Supposons dans une première étape que les inconnues ρ, ρu, ρE deviennent solutions stationnaires du système de Relaxation. Montrons alors formellement que la variable ρI ne peut pas être solution stationnaire de ce système. Nous en déduirons par contradiction dans une seconde étape que les inconnues ρ, ρu, ρE ne peuvent pas non plus être en général solutions stationnaires du système les gouvernant.

Première étape Il suffit de supposer que les quatre variables composant le système de Relaxation convergent vers une solution stationnaire, et d'aboutir à une contradiction. La discussion proposée est formelle puisqu'elle repose sur une analyse de consistance au sens des volumes finis. Si la convergence en temps avait lieu, la solution discrète relaxée représenterait une approximation constante par morceaux d'une solution du système

$$\begin{cases} \partial_x \rho u = 0, \\ \partial_x(\rho u^2 + \Pi) = 0, \\ \partial_x(\rho\Sigma u + \Pi u) = 0, \\ \partial_x \rho I u = 0, \quad I = \Pi + a^2/\rho, \end{cases} \tag{2.37}$$

puisque l'hypothèse d'annulation de tous les résidus explicites revient à exiger l'annulation de tous les incréments en temps. D'après la formule (2.35), ceux-ci prennent chacun la forme d'une différence de flux consistante au sens des volumes finis avec l'opérateur concerné dans (2.37). Notons ensuite que la variable de Relaxation ρI ne peut être qu'à l'équilibre en conséquence de la projection sur la variété d'équilibre à chaque itération et de l'hypothèse d'indépendance relativement (aux itérations) au temps. Ceci implique l'identité :

$$I(x) = p(x) + a^2\tau(x). \tag{2.38}$$

De ce fait, la solution stationnaire considérée est nécessairement une approximation d'une solution de :

$$\begin{cases} \partial_x \rho u = 0, \\ \partial_x(\rho u^2 + p) = 0, \\ \partial_x(\rho E u + p u) = 0, \\ \partial_x \rho p u + a^2 \partial_x u = 0. \end{cases} \tag{2.39}$$

Or, la validité des trois premières équations conduit sous l'hypothèse d'une solution régulière, la pression p à satisfaire l'identité suivante :

$$u\partial_x p - \tau\frac{\partial p}{\partial \tau}\partial_x u = 0, \tag{2.40}$$

version stationnaire de la loi d'évolution bien connue de la pression :

$$\partial_t p + u\partial_x p - \tau\frac{\partial p}{\partial \tau}\partial_x u = 0.$$

Rapprochons l'identité (2.40) de la quatrième équation du système (2.39), clairement équivalente à :

$$u\partial_x p + a^2\tau\partial_x u = 0.$$

La soustraction de ces deux EDO pour la pression p conduit à :

$$(a^2 + \frac{\partial p}{\partial \tau})\tau\partial_x u = 0.$$

La constance en x de $(\rho u)(x)$ fait qu'en règle générale $\tau\partial_x u$ est non nul, d'où nécessairement une contradiction avec la condition de Whitham qui exige $a^2 + \partial_\tau p > 0$.

Deuxième étape Supposons la convergence en temps vers un état stationnaire des variables ρ, ρu et ρE de sorte que ρI ne peut pas devenir indépendant du temps au sens de la discussion précédente. Sur la base de notre critère de convergence, nous obtenons par hypothèse :

$$\overline{\mathbf{r}}_\rho(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n) = \overline{\mathbf{r}}_{\rho u}(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n) = \overline{\mathbf{r}}_{\rho\Sigma}(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n) = 0,$$

et

$$\Delta\rho_j^{n+1/2} = \Delta\rho u_j^{n+1/2} = \Delta\rho\Sigma_j^{n+1/2} = 0,$$

tandis que

$$\overline{\mathbf{r}}_{\rho I}(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n) \neq 0 \quad \text{et} \quad \Delta\rho I_j^{n+1/2} \neq 0.$$

En reprenant les notations (2.33), la structure tridiagonale par blocs 4×4 de la matrice $\mathcal{A}$ implique alors localement en espace (*i.e.* sur une maille courante j) l'identité vectorielle suivante :

$$\mathbf{a}_j \Delta\overline{\mathbf{w}}_{j-1}^{n+1/2} + \mathbf{b}_j \Delta\overline{\mathbf{w}}_j^{n+1/2} + \mathbf{c}_j \Delta\overline{\mathbf{w}}_{j+1}^{n+1/2} = \overline{\mathbf{r}}(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n), \tag{2.41}$$

avec pour $l = j-1, j, j+1$:

$$\Delta\overline{\mathbf{w}}_l^{n+1/2} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \Delta\rho I_l^{n+1/2} \end{pmatrix}, \quad \text{et} \quad \overline{\mathbf{r}}(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \overline{\mathbf{r}}_{\rho I}(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n) \end{pmatrix}. \tag{2.42}$$

Les trois incréments $\{\Delta\rho I_l^{n+1/2}\}_{l=j-1,j,j+1}$ sont donc astreints à être localement solutions d'un système linéaire de quatre équations *a priori* indépendantes, *i.e.* avec des notations claires :

$$\begin{cases} a_j^{k,4}\Delta\rho I_{j-1}^{n+1/2} + b_j^{k,4}\Delta\rho I_j^{n+1/2} + c_j^{k,4}\Delta\rho I_{j+1}^{n+1/2} = 0, \quad k = 1, ..., 3, \\ a_j^{4,4}\Delta\rho I_{j-1}^{n+1/2} + b_j^{4,4}\Delta\rho I_j^{n+1/2} + c_j^{4,4}\Delta\rho I_{j+1}^{n+1/2} = \overline{\mathbf{r}}_{\rho I}(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n), \end{cases} \tag{2.43}$$

de sorte que l'incrément vectoriel $\Delta\rho I^{n+1/2}$ doit simultanément être dans le noyau de trois opérateurs linéaires *a priori* distincts tout en étant solution non triviale (*i.e.* $\Delta\rho I^{n+1/2} \neq 0$) d'un système linéaire avec second membre non nul puisque le résidu $\overline{\mathbf{r}}_{\rho I}$ ne peut être nul. Des tests numériques (non rapportés ici) montrent que les quatre équations (2.43) sont généralement incompatibles de sorte que les inconnues ρ, ρu et $\rho\Sigma$ ne peuvent être solutions stationnaires du système relaxé gouvernant leurs évolutions.

2.2 Correction de l'implicitation "naturelle"

L'objectif de cette section est de remédier au défaut de convergence vers un stationnaire de l'algorithme implicite associé au bloc hydrodynamique, décrit dans la section précédente. Puisque notre intérêt porte sur le système Euler multi-pression, nous exposerons le principe de la stratégie de correction directement dans ce cas. Sa restriction au cas des équations Euler 3×3 s'en deduira immédiatement (en posant $N = 1$).

L'impossibilité de converger vers des valeurs stationnaires en temps de ρ, ρu et ρE résulte du fait que la variable de Relaxation ρI devrait satisfaire à convergence, un ensemble sur-contraint de conditions. Il nous faut donc éliminer ces contraintes généralement incompatibles. Nous montrons formellement ci-dessous qu'un recouplage adéquat de l'étape de projection sur la variété d'équilibre avec la résolution en temps du bloc hydrodynamique permet d'éliminer ces contraintes. Un tel recouplage conduira à exprimer localement sur chaque cellule de calcul l'incrément ρI comme une fonction linéaire des incréments en densité, quantité de mouvement, énergie totale et entropies spécifiques. Les incréments en temps sur les entropies spécifiques ayant été évalués au préalable, la structure tridiagonale par blocs 4×4 associée au bloc hydrodynamique deviendra une structure toujours tridiagonale mais cette fois composée de blocs 3×3, le membre de droite constitué des résidus explicites recevant une modification précisée ci-dessous. Le bénéfice est donc double puisque la convergence vers une solution rigoureusement stationnaire sera assurée (ainsi que le montre l'expérience), et ce au prix de l'inversion d'un problème matriciel de taille moindre.

Nous donnons ici la conséquence formelle de la technique de recouplage envisagée puisque celle-ci doit être menée de manière linéarisée en temps pour des raisons de coût d'évaluation, et ce dans le régime asymptotique d'un paramètre de Relaxation $\lambda \to +\infty$. Une justification de cette procédure réclame des calculs fastidieux ainsi que des hypothèses difficilement vérifiables dans la pratique. Nous ne les rapporterons pas ici dans le détail faute de temps. La remarque concluant cette section donne les grandes lignes directrices de la démarche sous-tendant ces calculs.

Au terme de ces calculs, il s'avère que tout se passe comme s'il fallait prendre en compte l'étape de remise à l'équilibre lors de la résolution du système (2.34). Il s'agit dans les faits d'abandonner le choix $\lambda = 0$ lors de l'évolution en temps de la variable de Relaxation ρI pour lui substituer la résolution implicite linéarisée en temps de l'équation :

$$\partial_t \rho I + \partial_x \rho I u = \lambda \rho (p - \Pi),$$

dans le régime asymptotique $\lambda = +\infty$. Après division par λ et passage à la limite formel sur ce paramètre, une telle démarche conduit à devoir imposer à ρI de vérifier une version linéarisée en temps de la condition d'équilibre :

$$(\rho I)_j^{n+1} = \rho_j^{n+1} \sum_{i=1}^{N} p_i(1/\rho_j^{n+1}, (S_i)_j^{n+1}) + a^2, \quad j \in Z. \tag{2.44}$$

Ici, l'entropie $(S_N)_j^{n+1} := S_N(\rho_j^{n+1}, (\rho u)_j^{n+1}, (\rho\Sigma)_j^{n+1}, \{(S_i)_j^{n+1}\}_{i=1,\dots,N-1})$ est déterminée par la loi de fermeture sur l'énergie totale (1.3). Notons qu'en définissant la fonction $\mathbf{u} \to f(\mathbf{u})$ par

$$f(\mathbf{u}) = \rho \sum_{i=1}^{N} p_i(1/\rho, S_i) + a^2, \quad \mathbf{u} = (\rho, \rho u, \rho\Sigma, \{S_i\}_{i=1,\dots,N-1}),$$

avec $S_N = S_N(\rho, \rho u, \rho\Sigma, \{S_i\}_{i=1,\dots,N-1})$, (2.44) se réécrit $(\rho I)_j^{n+1} = f(\mathbf{u}_j^{n+1})$. Conformément à l'approche adoptée pour l'implicitation du bloc hydrodynamique, la linéarisation envisagée repose sur un développement de Taylor à l'ordre 1 de la fonction f autour de l'instant t^n :

$$\begin{aligned}(\rho I)_j^{n+1} &= (\rho I)_j^n + (\partial_\rho f)_j^n \Delta\rho_j^{n+1/2} + (\partial_{\rho u} f)_j^n \Delta\rho u_j^{n+1/2} \\ &+ (\partial_{\rho\Sigma} f)_j^n \Delta\rho\Sigma_j^{n+1/2} + \textstyle\sum_{i=1}^{N-1} (\partial_{S_i} f)_j^n \Delta(S_i)_j^{n+1/2}.\end{aligned} \tag{2.45}$$

Dès lors, les incréments $\Delta\rho I_j^{n+1/2}$ ne sont plus des inconnues indépendantes mais s'expriment comme des fonctions affines des incréments inconnus $\Delta\rho_j^{n+1/2}$, $\Delta\rho u_j^{n+1/2}$ et $\Delta\rho\Sigma_j^{n+1/2}$. Précisons en effet que les incréments $\Delta(S_i)_j^{n+1/2}$, $i = 1, ..., (N-1)$ intervenant dans (2.45) sont déterminés par l'implicitation découplée du bloc transport des entropies et donc connus à ce niveau de description de l'algorithme. De même, la fonction f et l'ensemble de ses dérivées partielles sont explicitement connues puisqu'évaluées à l'instant t^n.
Sur une maille courante j et de manière à rendre compte de la relation algébrique (2.45) liant les incréments, le système linéaire à résoudre (2.34), déduit de (2.41)-(2.45), s'écrit maintenant sous la forme :

$$\tilde{\mathbf{a}}_j \Delta\mathbf{w}_{j-1}^{n+1/2} + \tilde{\mathbf{b}}_j \Delta\mathbf{w}_j^{n+1/2} + \tilde{\mathbf{c}}_j \Delta\mathbf{w}_{j+1}^{n+1/2} = \tilde{\mathbf{r}}(\overline{\mathbf{w}}_{j-1}^n, \overline{\mathbf{w}}_j^n, \overline{\mathbf{w}}_{j+1}^n),$$

où nous avons posé $\mathbf{w} = (\rho, \rho u, \rho\Sigma)$. Les matrices impliquées $\tilde{\mathbf{a}}_j$, $\tilde{\mathbf{b}}_j$ et $\tilde{\mathbf{c}}_j$ sont déduites de (2.41)-(2.45) et sont définies avec des notations claires, colonnes par colonnes, par

$$\tilde{a}_j^{k,1} = a_j^{k,1} + a_j^{k,4}(\partial_\rho f)_j^n, \quad \tilde{a}_j^{k,2} = a_j^{k,2} + a_j^{k,4}(\partial_{\rho u} f)_j^n, \quad \tilde{a}_j^{k,3} = a_j^{k,3} + a_j^{k,4}(\partial_{\rho\Sigma} f)_j^n, \tag{2.46}$$

pour tout $k = 1, ..., 3$, tandis que les résidus explicites s'écrivent :

$$\tilde{\mathbf{r}}_\rho = \overline{\mathbf{r}}_\rho - a_j^{1,4} \sum_{i=1}^{N-1} (\partial_{S_i} f)_j^n \Delta(S_i)_j^{n+1/2}, \quad \tilde{\mathbf{r}}_{\rho u} = \overline{\mathbf{r}}_{\rho u} - a_j^{2,4} \sum_{i=1}^{N-1} (\partial_{S_i} f)_j^n \Delta(S_i)_j^{n+1/2}, \tag{2.47}$$

et

$$\tilde{\mathbf{r}}_{\rho\Sigma} = \overline{\mathbf{r}}_{\rho\Sigma} - a_j^{3,4} \sum_{i=1}^{N-1} (\partial_{S_i} f)_j^n \Delta(S_i)_j^{n+1/2}. \tag{2.48}$$

Ainsi, l'implicitation du bloc hydrodynamique permettant d'évaluer la suite

$$(\rho_j^{n+1}, (\rho u)_j^{n+1}, (\rho E)_j^{n+1})_{j \in Z},$$

consiste en la résolution d'un système linéaire dont la matrice possède une structure tridiagonale par blocs, chaque bloc étant de taille 3×3. Les contraintes incompatibles qui portaient sur l'incrément $\Delta\rho I_j^{n+1/2}$ ont donc été éliminées au prix d'une modification des matrices élémentaires gouvernant les incréments en ρ, ρu, ρE. En ce sens, aucun argument ne s'oppose plus à la convergence de l'algorithme proposé vers un stationnaire en ces variables. Notons en effet que lorsque les implicitations des blocs transport des entropies et hydrodynamique convergent simultanément vers une solution stationnaire (cette situation est bien évidemment celle qui nous intéresse), l'annulation des incréments en temps $\Delta(S_i)_j^{n+1/2} = 0$ conduit à l'annulation des résidus explicites $\overline{\mathbf{r}}_\rho$, $\overline{\mathbf{r}}_{\rho u}$ et $\overline{\mathbf{r}}_{\rho\Sigma}$.

La remarque concluant cette section a pour but d'indiquer que la procédure proposée n'est en fait qu'une prise en compte de la projection sur la variété d'équilibre (*i.e.* avec $\lambda = +\infty$) mais linéarisée en temps. Il est dès lors indispensable de corriger cette linéarisation de manière à redémarrer l'algorithme à chaque itération sur une solution rigoureusement à l'équilibre, en réactualisant la valeur de ρI localement sur chaque cellule selon (2.44).

De manière à résumer la stratégie d'implicitation en temps des équations d'Euler multi-pression, il s'agit dans un premier temps de résoudre $(N-1)$ systèmes linéaires à structure tridiagonale scalaire pour l'étape de convection des entropies $(S_i)_j^{n+1}$ dont la connaissance est indispensable à la définition des résidus (2.47)-(2.48). Dans un second temps, il s'agit de résoudre un système linéaire dont la matrice possède une structure tridiagonale par blocs de taille 3×3 (pour le calcul des quantités $\rho_j^{n+1}, (\rho u)_j^{n+1}, (\rho E)_j^{n+1}$) et enfin de redéfinir ρI à l'équilibre. Ainsi et *stricto sensu*, l'évolution du bloc hydrodynamique se retrouve être faiblement couplée avec celle du bloc transport des entropies. Ce couplage faible n'intervenant que dans la définition des résidus "explicites" est bien évidemment moins coûteux qu'une approche implicite linéarisée classique consistant à résoudre un système linéaire à structure tridiagonale par blocs de taille $(N+2) \times (N+2)$. Notons enfin que si l'on adoptait une condition

CFL ≤ 1 construite sur la vitesse de convection des entropies (moins restrictive que la condition CFL construite sur les ondes acoustiques), l'évaluation des résidus (2.47)-(2.48) serait complètement explicite puisque les incréments $\Delta(S_i)^{n+1/2}$ connaîtraient une mise à jour explicite stable. Nous avons privilégié ici une convergence vers l'état stationnaire la plus rapide possible, et donc exempte de toute restriction de principe sur le nombre CFL. L'expérience numérique (voir ci-dessous) montre que la correction proposée conduit à une méthode implicite linéarisée robuste permettant d'obtenir une annulation des résidus explicites satisfaite au zéro machine.

Remarque 1 Ce qui suit n'est qu'une justification formelle de la procédure de correction que nous venons d'introduire. Une telle justification requiert des hypothèses d'existence de solutions de problèmes algébriques non linéaires puis de problèmes linéaires avec perturbation singulière dont la vérification *a priori* semble être hors de portée. Par ailleurs et même au prix de ces hypothèses, les calculs permettant d'aboutir à la technique de correction proposée sont fastidieux et le temps nous a manqué pour en faire le détail ici. Nous nous contenterons de présenter les grandes lignes conduisant à ces calculs. En considérant un paramètre de Relaxation $\lambda > 0$ fixé, recoupler l'étape de projection sur la variété d'équilibre avec l'évolution en temps du bloc hydrodynamique revient formellement à devoir résoudre un système algébrique non linéaire localement défini sur chaque cellule de calcul j par les identités vectorielles suivantes :

$$\mathbf{a}_j \Delta \overline{\mathbf{w}}_{j-1}^{n+1/2} + \mathbf{b}_j \Delta \overline{\mathbf{w}}_{j}^{n+1/2} + \mathbf{c}_j \Delta \overline{\mathbf{w}}_{j+1}^{n+1/2} = \overline{\mathbf{r}}(\overline{\mathbf{w}}_{j-1}^{n}, \overline{\mathbf{w}}_{j}^{n}, \overline{\mathbf{w}}_{j+1}^{n}) + \lambda \mathcal{R}_j^{n+1},$$

où le terme de Relaxation s'écrit

$$\mathcal{R}_j^{n+1} = {}^T(0, 0, 0, \rho_j^{n+1} p(\mathcal{T}_j^{n+1}, \{(S_i)_j^{n+1}\}_{i=1,\ldots,N-1}, (\rho \mathcal{S}_{\mathcal{N}})_j^{n+1}) - (\rho\Pi)_j^{n+1}).$$

Ici et conformément à l'analyse du système de Relaxation considéré, nous avons

$$(\rho\Pi)_j^{n+1} = (\rho I)_j^{n+1} - a^2,$$

et il faut comprendre le relèvement $\mathcal{T}$ du covolume τ et le relèvement $\mathcal{S}_{\mathcal{N}}$ de l'entropie S_N comme des fonctions non linéaires de ρ_j^{n+1}, $(\rho u)_j^{n+1}$, $(\rho I)_j^{n+1}$, $(\rho\Sigma)_j^{n+1}$ et $\{(\rho S_i)_j^{n+1}\}_{i=1,\ldots,N-1}$ définies dans la sous-section 1.6.2. Il aurait fallu paramétrer l'éventuelle solution de ce problème non linéaire relativement à λ, ce que nous n'avons pas fait pour simplifier les notations. De manière à contourner les non linéarités précédentes, nous proposons d'opérer une linéarisation en temps du terme de Relaxation $\mathcal{R}$ à λ fixé, et ce à l'instant t^n. Nous sommes dès lors conduit à devoir considérer un problème linéaire en l'incrément $\Delta \overline{\mathbf{w}}_{\lambda}^{n+1/2}$. La matrice associée à ce problème linéaire est tridiagonale par blocs 4×4, la dernière ligne de chaque bloc composant la diagonale dépendant explicitement du paramètre λ. Après division par λ, la dernière ligne de chacun des blocs constituant le système linéaire dépend donc de $1/\lambda$. Suivant la procédure de remise à l'équilibre, nous faisons tendre λ vers l'infini dans ce système linéaire avec perturbation singulière. Des calculs montrent que localement sur chaque cellule de calcul, la ligne où figure le résidu explicite en ρI (lui même divisé par λ) dégénère pour conduire à la relation algébrique (2.45) fixant l'incrément $\Delta \rho I_j^{n+1/2}$ en fonction des incréments des autres variables sur la même cellule. Dès lors, la dernière colonne de chaque bloc liant les incréments en $j-1$, j et $j+1$ conduit à devoir modifier le système limite selon (2.46)-(2.47)-(2.48).

2.3 Résultats numériques

Nous considérons dans cette section l'approximation numérique des solutions stationnaires du système des équations d'Euler multi-pression en deux dimensions d'espace. L'objectif consiste à tester la stabilité de l'implicitation proposée dans les sections précédentes. Le système (1.1) s'écrit sous la forme :

$$\begin{cases} \partial_t \rho + \nabla \cdot \rho \omega = 0, \\ \partial_t \rho \omega + \nabla \cdot \rho \omega \otimes \omega + \nabla \sum_{i=1}^{N} p_i = 0, \\ \partial_t \rho E + \nabla \cdot (\rho E + \sum_{i=1}^{N} p_i) \omega = 0, \\ \partial_t \rho S_i + \nabla \cdot \rho S_i \omega = 0, \quad i = 1, ..., N-1, \end{cases} \tag{2.49}$$

et l'inégalité d'entropie (1.4) sous la forme :

$$\partial_t \rho S_N + \nabla \cdot \rho S_N \omega \leq 0, \tag{2.50}$$

où $\omega = (u, v) \in I\!R^2$ représente le champ de vitesse dans le repère usuel (O, Ox, Oy). Nous considérons une fois de plus des lois de pression de gaz parfaits polytropiques (les coefficients adiabatiques associés sont notés $\gamma_i > 1$, $i = 1, ..., N$) de telle sorte que la loi de fermeture sur l'énergie totale ρE s'écrit :

$$\rho E = \frac{1}{2} \rho |\omega|^2 + \sum_{i=1}^{N} \frac{p_i}{\gamma_i - 1}, \quad |\omega|^2 = u^2 + v^2. \tag{2.51}$$

L'approche que nous adoptons ici exploite très classiquement l'invariance par rotation du système (2.49). Nous nous contenterons juste de rappeler que dans ce contexte, nous avons uniquement besoin de définir une fonction flux numérique "quasi-1D" consistante avec la projection des équations du système dans une direction fixée. La fonction flux numérique dans l'autre direction se déduit simplement de la précédente sous l'action d'une rotation sur le champ de vitesse, suivie de la rotation inverse sur les composantes en impulsion de la fonction flux numérique. Il s'agit là de considérations très classiques qui valent pour la formulation explicite en temps, mais également implicite linéarisée en temps et nous ne rentrons pas plus dans les détails. Nous renvoyons le lecteur à [57] pour une présentation générale dans le contexte de méthodes explicites en temps, et par exemple à l'ouvrage [68] pour une discussion dans le cadre de méthodes implicites linéarisées en temps. Pour fixer les idées, nous donnons la projection dans la direction Ox (sans aucune restriction) du système (2.49)-(2.50) :

$$\begin{cases} \partial_t \rho + \partial_x \rho u = 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + \sum_{i=1}^{N} p_i) = 0, \\ \partial_t \rho v + \partial_x \rho v u = 0, \\ \partial_t \rho E + \partial_x (\rho E + \sum_{i=1}^{N} p_i) u = 0, \\ \partial_t \rho S_i + \partial_x \rho S_i u = 0, \quad i = 1, ..., N-1, \end{cases} \tag{2.52}$$

$$\partial_t \rho S_N + \partial_x \rho S_N u \leq 0. \tag{2.53}$$

Une équation de transport supplémentaire sur la vitesse tangentielle v vient donc s'ajouter aux équations de transport sur les entropies S_i du système (1.1). L'approximation numérique de cette équation est rigoureusement identique à celle des entropies S_i.
L'ordre 2 d'approximation en espace est obtenu par une approche MUSCL (voir [57]) menée ici sur les variables primitives, à savoir la densité, le vecteur vitesse et les pressions partielles $\{p_i\}_{i=1,\ldots,N}$. Le limiteur de pente utilisé est celui proposé par Van Albada et al. ($\phi(r) = (r^2 + r)/(1 + r^2)$ avec des notations classiques).
Afin de tester la stabilité de l'implicitation proposée, nous approchons numériquement plusieurs écoulements stationnaires, supersoniques à l'infini amont, et arrivant sur un cylindre. L'interaction d'un tel écoulement avec le cylindre donne naissance à un fort choc détaché devant celui-ci. Pour cela, nous utilisons un maillage curviligne structuré avec ni points de discrétisation sur la paroi du cylindre et nj points de discrétisation sur l'axe Ox (direction de la ligne d'arrêt). Pour chaque cas test, après avoir précisé les données correspondantes, nous présentons les iso-valeurs des champs de pression et de densité. Les distributions des pressions partielles sur la ligne de point d'arrêt seront également présentées.
Précisons enfin que nous avons utilisé les conditions aux limites classiques suivantes pour effectuer l'ensemble des simulations :

- condition de glissement sur la paroi du cylindre,
- condition de sortie supersonique en aval,
- condition de symétrie sur la ligne d'arrêt.

La prise en compte de ces conditions aux limites au niveau discret a été menée conformément au code de calculs CELHYO développé à l'ONERA. Il s'agit d'une stratégie classique reposant sur l'utilisation d'états miroirs (paroi et condition de symétrie) que nous ne rappellerons pas ici.

Expérience 1

Pour ce premier cas test, nous faisons le choix $N = 1$. Le système (2.49) correspond aux équations d'Euler habituelles. Le maillage utilisé (et représenté sur la figure 2.1) est tel que $ni = 32$ et $nj = 41$. Le coefficient adiabatique γ est quant à lui fixé à 1.2, tandis que le champ amont arrivant sur le cylindre est tel que :

Expérience 1

	ρ	u	v	p
Champ amont	7.032e-4.	4318.	0.	109.

Le nombre de Mach associé à ce calcul est environ égal à 10. Le cas test est donc sévère, comme l'illustre la figure 2.4 mettant en évidence un rapport de pression supérieur à 100 à la traversée du choc détaché composant la solution stationnaire. Dans ce contexte, l'implicitation proposée de la méthode de Relaxation s'est avérée être d'une robustesse supérieure à la méthode de Roe classique, mais inférieure à une méthode de Van Leer linéairement implicitée. La robustesse est ici évaluée relativement à une stratégie de montée en nombre CFL au cours des itérations. Pour une stratégie donnée, nous avons cependant observé pour ce cas test une convergence plus rapide vers la solution stationnaire pour la méthode de Relaxation que pour la méthode de Van Leer. Plus que la robustesse, c'est cette dernière propriété qui nous est apparue importante. Cette propriété est illustrée dans le tableau ci-dessous qui compare à pas de temps donné le nombre d'itérations et le temps CPU nécessaires pour obtenir des résidus explicites L^2 de l'ordre de 10^{-8}. Nous donnons à titre indicatif les performances de la méthode de Relaxation complètement explicite en temps (certes sous une condition CFL incomparable mais en l'absence de toute procédure d'inversion de système linéaire). Nous nous sommes restreints ici à un critère de convergence de 10^{-8} pour rendre ce calcul explicite d'un coût non rédhibitoire.

	Itérations	Temps CPU	Temps CPU / itération
Relaxation explicite	11900	19642.176	1.65
Relaxation implicite	655	3791.925	5.79
Van Leer implicite	725	4950.689	6.828

Nous tenons à préciser que nous ne sommes pas en mesure d'affirmer laquelle des méthodes de Relaxation et de Van Leer est la plus performante en temps calculs. Le nombre de simulations à mener pour de telles conclusions est trop important. Notre propos est simplement d'énoncer les tendances observées pour ce cas test difficile, à savoir : une utilisation de pas de temps plus élevés pour la méthode de Van Leer et une convergence vers l'état stationnaire plus rapide pour la méthode de Relaxation.

A titre d'illustration, nous testons maintenant la méthode en considérant plusieurs pressions indépendantes (*i.e.* $N > 1$). Le code CELHYO de l'ONERA dans lequel la méthode par Relaxation a été implémentée ne traitant pas à l'origine plusieurs lois de pression indépendantes, aucune comparaison de type stabilité ou vitesse de convergence avec une autre méthode n'a pu être effectuée sur ces cas tests.
Nous n'avons pas rencontré de difficulté particulière dans la réalisation des simulations numériques proposées. Soulignons que celles-ci sont relativement sévères puisqu'elles s'accompagnent de taux de compression élevés. Le critère de convergence vers la solution stationnaire a été fixé à 10^{-12} dans tous les cas.

Expérience 2

Pour ce problème, nous considérons $N = 2$ pressions partielles p_1 et p_2 associées respectivement aux coefficients adiabatiques $\gamma_1 = 1.3$ et $\gamma_2 = 1.2$. Le maillage utilisé est le même que pour la simulation précédente, et le champ amont considéré est le suivant :

Expérience 2

	ρ	u	v	p_1	p_2
Champ amont	7.032e-4.	4318.	0.	36.	70.

Le nombre de Mach est toujours environ égal à 10.

Expérience 3

Pour ce cas test, nous considérons trois pressions partielles. Les coefficients adiabatiques vérifient $\gamma_1 = 1.4$, $\gamma_2 = 1.3$ et $\gamma_3 = 1.2$. Le maillage utilisé est tel que $ni = 32$ et $nj = 88$. Le champ arrivant sur le cylindre est le suivant :

Expérience 3

	ρ	u	v	p_1	p_2	p_3
Champ amont	4.018e-4.	964.	0.	10.	26.	40.

Pour ce calcul, le nombre de Mach est approximativement 2.

Présentation des résultats

L'examen des résultats exhibés ci-après indique que les solutions discrètes proposées ne présentent pas de pathologie particulière. La structure de l'écoulement dans le cas de plusieurs pressions indépendantes est largement conforme à l'intuition.

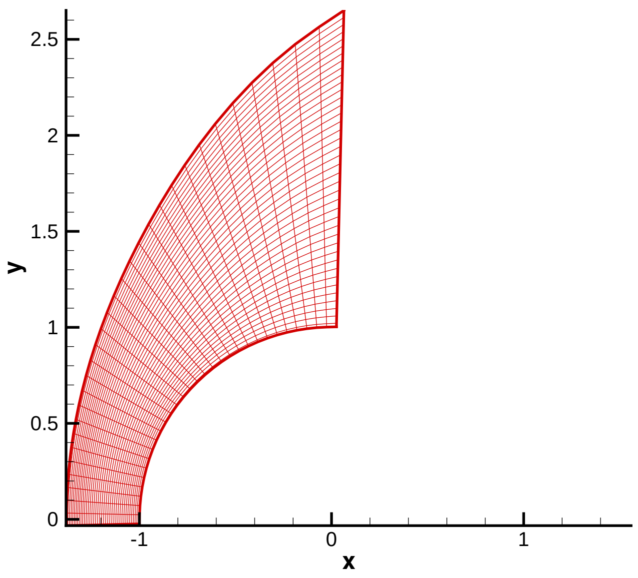

FIG. 2.1 – *Expérience 1. Maillage 32×41*

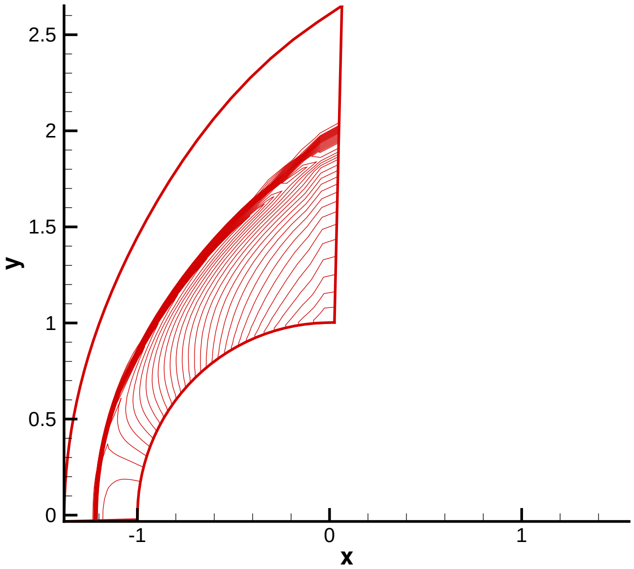

Fig. 2.2 – *Expérience 1. Densité : iso-valeurs*

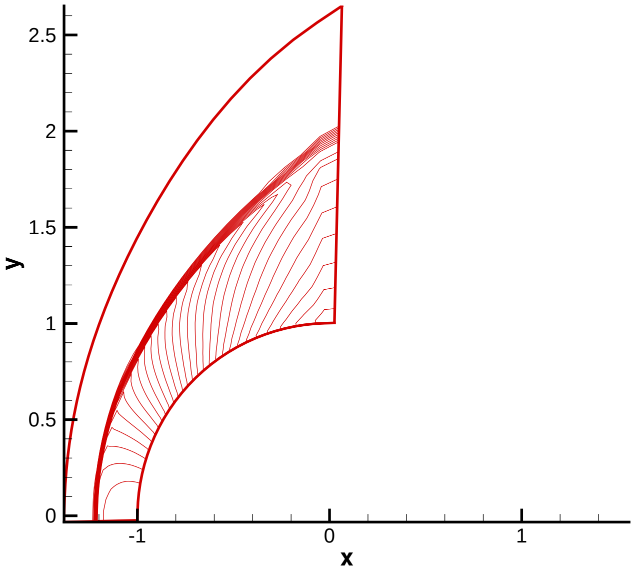

Fig. 2.3 – *Expérience 1. Pression : iso-valeurs*

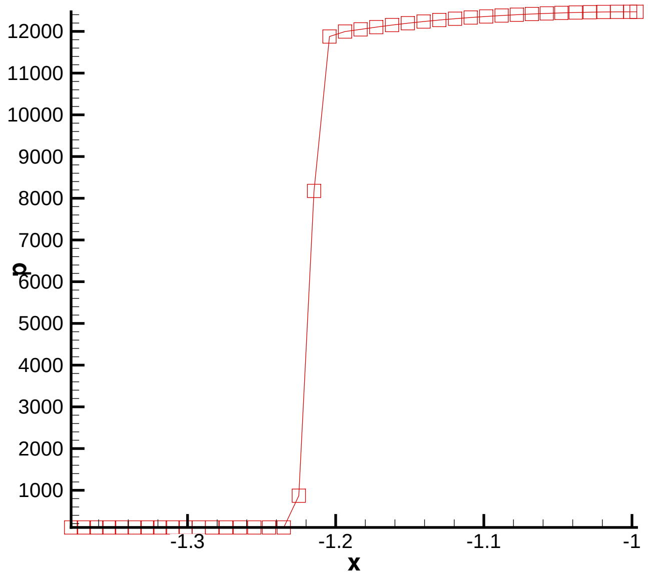

Fig. 2.4 – *Expérience 1. Pression sur la ligne d'arrêt*

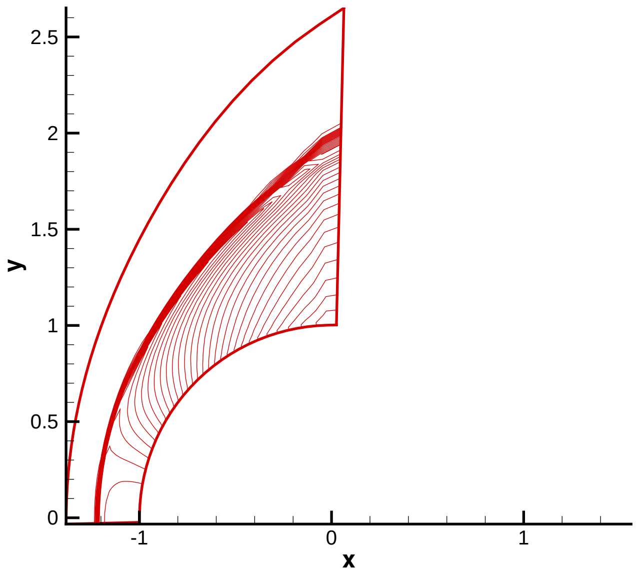

Fig. 2.5 – *Expérience 2. Densité : iso-valeurs*

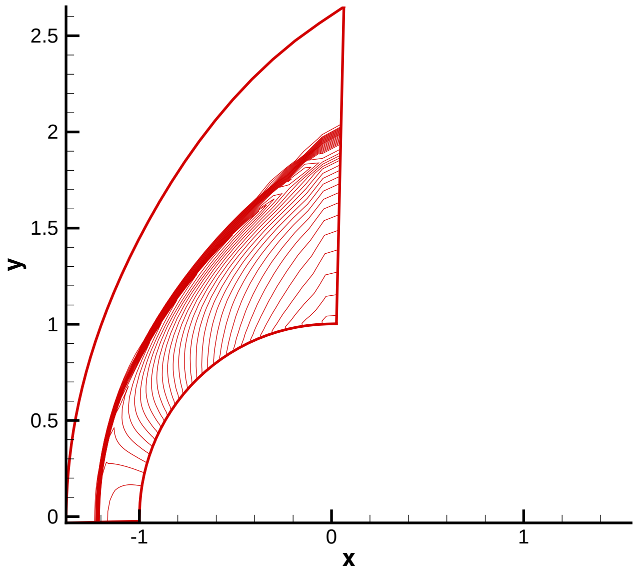

Fig. 2.6 – *Expérience 2. Pression* p_1 *: iso-valeurs*

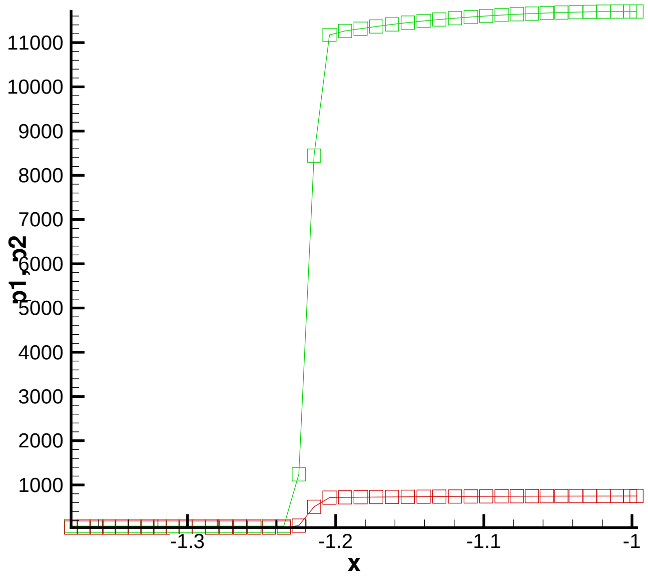

Fig. 2.7 – *Expérience 2. Pressions sur la ligne d'arrêt*

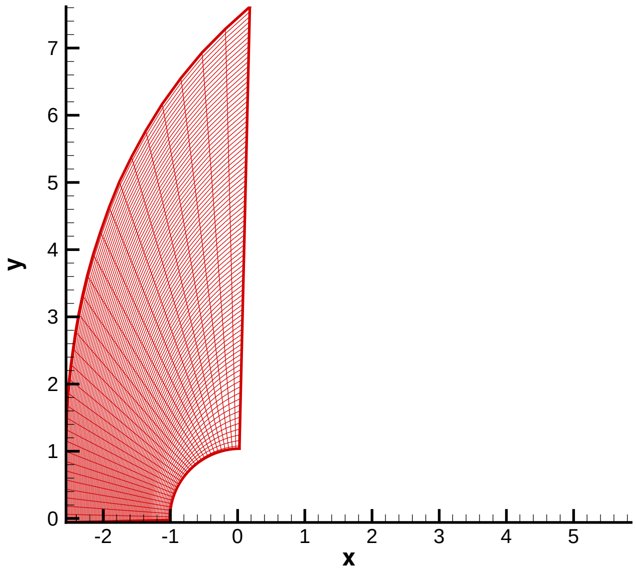

Fig. 2.8 – *Expérience 3. Maillage 32×88*

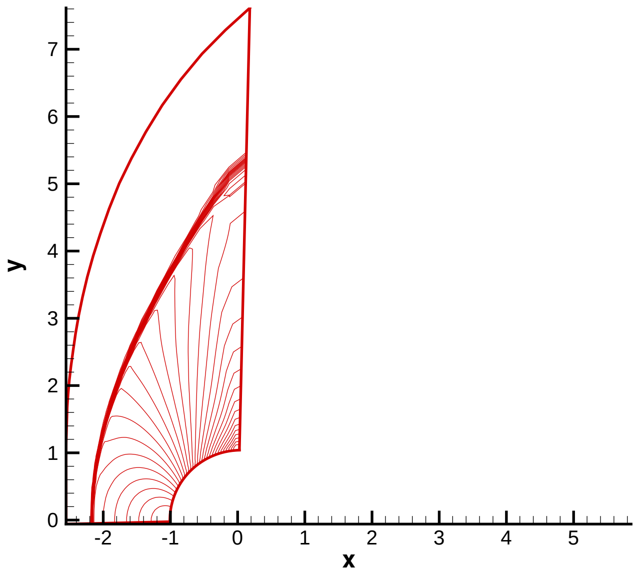

Fig. 2.9 – *Expérience 3. Densité : iso-valeurs*

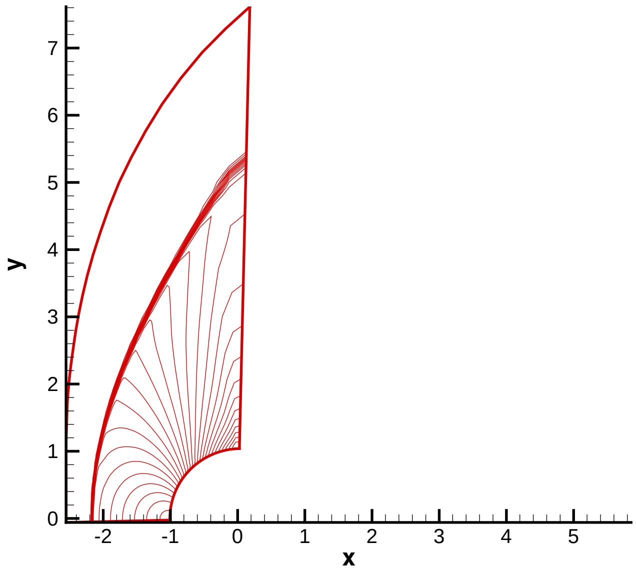

FIG. 2.10 – *Expérience 3. Pression p_2: iso-valeurs*

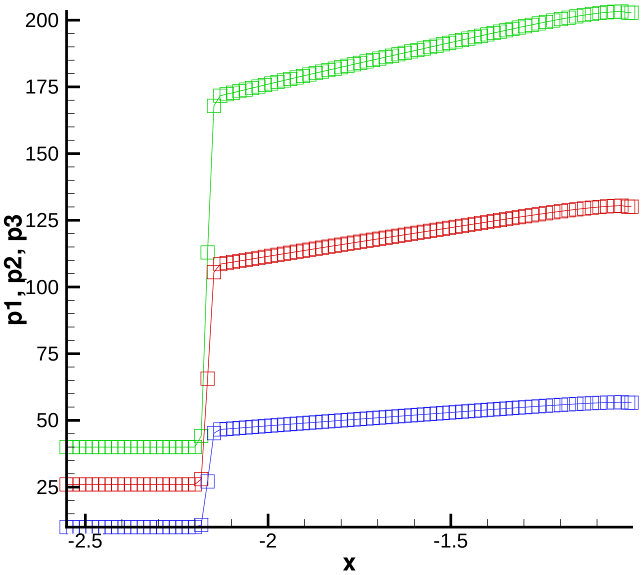

FIG. 2.11 – *Expérience 3. Pressions sur la ligne d'arrêt*

Chapitre 3

Algorithme implicite pour les équations de Navier-Stokes à plusieurs pressions indépendantes

Conformément aux développements menés dans les premiers chapitres, la discrétisation des équations de Navier-Stokes multi-pression requiert deux pas distincts : une étape de prédiction, suivie d'une étape de correction visant à forcer la consistance des relations de saut généralisées au niveau discret. Sur la base des développements proposés au précédent chapitre, nous disposerons aisément d'une implicitation linéarisée de l'étape de prédiction. Il s'agit ici de proposer une implicitation linéarisée de l'étape de correction. Nous motiverons l'introduction d'une nouvelle formulation des relations de proportionnalité entre les bilans d'entropie. Elle sera comprise comme une hybridation de deux formulations utilisées dans un contexte purement explicite (voir chapitres II 2 et II 4). Par construction, cette nouvelle formulation autorisera une mise à jour implicite possédant un coût d'évaluation "quasi-optimal", et préservant la positivité des entropies spécifiques dès que la densité et l'énergie interne totale déduites de l'étape de prédiction sont positives. Par "optimal", nous entendons qu'il s'agira d'impliciter uniquement $(N-1)$ équations de transport en présence d'un terme source connu lors de la correction. Il faudra ensuite résoudre, localement sur chaque cellule, un problème algébrique dont la solution est explicitement donnée dans le cas de N lois de pression polytropiques.

3.1 Considérations algorithmiques

L'objectif de ce paragraphe est de mettre en avant un certain nombre de difficultés dans l'établissement d'une formulation implicite linéarisée en temps des relations de saut généralisées telles que formulées dans les chapitres précédents. L'analyse de ces difficultés, essentiellement liées au souci de conjuguer efficacité et stabilité (positivité des entropies corrigées), conduira à introduire une nouvelle formulation de ces relations de compatibilité, permettant de concilier ces deux contraintes.

3.1.1 A propos des relations de saut généralisées

Les chapitres 2, 4 et 3 (partie II) consacrés au développement de méthodes explicites en temps ont montré toute l'importance du respect au niveau discret des relations de saut généralisées. Différents algorithmes de type prédicteur-correcteur ont été proposés à cet effet, et reposent sur différentes formes (équivalentes au niveau continu) de ces relations de compatibilité. La plus naturelle repose sur une formulation en entropies :

$$\mu_N T_i\{\partial_t \rho S_i + \partial_x \rho S_i u\} = \mu_i T_N\{\partial_t \rho S_N(\mathbf{U}) + \partial_x \rho S_N(\mathbf{U}) u\}, \quad i = 1, ..., N-1. \quad (3.1)$$

Il s'agit dans ce cas de corriger les N entropies conformément aux $(N-1)$ relations (3.1), une fois connues la densité, l'impulsion et l'énergie totale (voir chapitre II 2). Sous cette forme, l'algorithme requiert l'inversion d'une équation non linéaire, sur chaque cellule de calcul et à chaque pas de temps. Dans le chapitre II 4, nous avons montré que la formulation équivalente en énergies internes :

$$\mu_N\{\partial_t \rho\varepsilon_i + \partial_x \rho\varepsilon_i u + p_i \partial_x u\} = \mu_i\{\partial_t \rho\varepsilon_N + \partial_x \rho\varepsilon_N u + p_N \partial_x u\}, \quad i = 1, ..., N-1, \quad (3.2)$$

déduite du second principe de la thermodynamique, conduisait à des solutions discrètes virtuellement identiques aux précédentes mais avec l'avantage d'un algorithme complètement explicite. Dans les deux cas, un argument clef permettant d'assurer la positivité des entropies et des énergies ainsi corrigées repose sur la satisfaction d'une inégalité d'entropie (ou d'énergie totale) à l'issue de l'étape de prédiction. L'examen de la formule de remise à jour des énergies internes (4.57), établie au chapitre II 4, suffit à attester qu'une violation de l'inégalité d'énergie discrète ne peut résulter qu'en une violation de la positivité des énergies internes (dès que l'une d'entre elles est suffisamment petite par exemple).
Toutes ces constructions exploitent de manière centrale la nature explicite en temps de l'avancement de la solution, tant dans l'étape de prédiction (existence d'une inégalité d'entropie) que dans la phase de correction (découplage en espace des valeurs à réactualiser). Soulignons tout d'abord qu'il est illusoire d'espérer satisfaire une inégalité d'entropie discrète à l'issue de l'implicitation linéarisée en temps de l'étape de prédiction. Cette constatation suffit déjà à remettre en cause l'intérêt des deux formulations considérées, puisqu'encore une fois, le respect de la positivité dépend directement de la validité d'une telle inégalité. Même en ignorant cette difficulté, l'implicitation des relations de proportionnalité (3.1) ou (3.2) conduit par définition à devoir exprimer implicitement les dérivées partielles en espace, par des formules aux différences portant sur l'inconnue évaluée à l'instant t^{n+1}. Ces formules doivent dès lors être linéarisées. La nature non conservative de la formulation en énergies complique singulièrement cette tâche, d'autant plus qu'elle couple de manière non triviale les trois vitesses caractéristiques de propagation. Rappelons très sommairement que le succès d'une implicitation linéarisée repose sur une bonne implicitation de chacune des informations caractéristiques. Nous avons donc choisi de ne pas aborder son traitement. Les relations (3.1) présentent l'avantage de n'impliquer qu'une seule vitesse caractéristique en conduisant à des équations de transport, équations que nous savons impliciter (la densité est connue). Toutefois, les non linéarités de l'application $\mathbf{U} \to \rho S_N(\mathbf{U})$ nécessiteraient, pour des raisons de coût, l'utilisation d'une linéarisation complexe qui au mieux conduirait à devoir considérer un système linéaire tridiagonal

par blocs $(N-1) \times (N-1)$. Même en faisant abstraction du coût calcul, il semble difficile de procéder à une linéarisation garantissant la positivité des entropies spécifiques. En ce sens, aucune des deux formulations ne donne satisfaction.

3.1.2 Vers une autre formulation des relations de saut généralisées

L'idée directrice est ici de chercher à hybrider les deux formulations précédentes (en entropies et en énergies), tout en s'efforçant de garantir sous des hypothèses les plus faibles possibles la positivité des entropies spécifiques.

La première étape est de constater que si chacune des énergies internes est inconnue, l'énergie interne totale est en revanche connue à l'issue de l'étape de prédiction, puisque par définition, la densité, l'impulsion et l'énergie totale restent inchangées au cours de l'étape de correction. Nous montrerons ci-dessous que cette propriété permettra de proposer une approximation discrète (en temps-espace) et consistante de l'expression

$$\partial_t \rho\varepsilon + \partial_x \rho\varepsilon u + p\partial_x u, \tag{3.3}$$

entièrement connue avant de procéder à l'étape de correction. Ici ε et p désignent respectivement l'énergie interne totale et la pression totale. La connaissance de cette expression est à la base des développements menés dans le chapitre II 4, et ayant abouti à l'exposé d'une méthode de correction explicite en énergies internes. Comme indiqué au paragraphe précédent, nous devons toutefois préférer à des fins d'implicitation une formulation en équations de transport sur les entropies n'impliquant qu'une seule vitesse caractéristique. Il nous faut donc rapprocher ces équations de l'expression précédente.

L'analyse des équations du système d'EDP considéré montre que ses solutions régulières vérifient

$$\partial_t \rho\varepsilon + \partial_x \rho\varepsilon u + p\partial_x u = \mu(\partial_x u)^2, \tag{3.4}$$

alors que chacune des entropies spécifiques est solution de :

$$\partial_t \rho S_i + \partial_x \rho S_i u = -\frac{\mu_i}{T_i}(\partial_x u)^2, \quad i = 1, ..., N. \tag{3.5}$$

On en déduit les N relations de proportionnalité suivantes :

$$\partial_t \rho S_i + \partial_x \rho S_i u = -\frac{\mu_i}{\mu T_i}(\partial_t \rho\varepsilon + \partial_x \rho\varepsilon u + p\partial_x u), \quad i = 1, ..., N. \tag{3.6}$$

Puisque le second membre est explicitement connu, il pourrait être tentant d'imposer les relations de saut généralisées sous la forme de ces N relations. Notons en effet qu'elles impliquent la validité des relations de saut généralisées (3.1) et donc (3.2). De plus, un tel algorithme ne nécessiterait que l'inversion de N systèmes linéaires résultant d'une implicitation linéarisée des équations de transport entrant dans le membre de gauche. Il y a toutefois deux obstacles. Rappelons tout d'abord que si la matrice impliquée possède une structure de M-matrice (voir chapitre précédent) et qu'à ce titre son inverse est une matrice positive, la positivité des S_i résulterait de l'hypothèse (naturelle d'après (3.4)) d'un membre de droite positif. Mais cette propriété (liée à la satisfaction d'une inégalité d'énergie) n'est assurée que dans le contexte d'une formulation complètement explicite (voir chapitre II 4). D'autre part, une formulation

discrète de ces N relations conduit à devoir considérer $(N+3)$ mises à jour, pour $(N+2)$ inconnues. Rappelons en effet que ρ, ρu et ρE sont déja connues à l'issue de l'étape de prédiction. Il ne peut en toute généralité en résulter qu'un conflit avec la définition de l'énergie totale, et donc avec la conservation de cette dernière quantité (voir ci-après pour les détails).
De manière à contourner ces deux obstacles, nous proposons ci-dessous d'exprimer des extensions directes des relations (3.6), construites sur $(N-1)$ fonctions des entropies spécifiques, à savoir le logarithme des rapports des $(N-1)$ premières sur la dernière. La connaissance des N entropies sera alors déduite de la connaissance de ces $(N-1)$ rapports en invoquant la définition de l'énergie totale (assurant ainsi sa conservation). Prendre le logarithme des rapports résulte évidemment du souci d'assurer la positivité de chacune des entropies spécifiques S_i. Nous montrerons dans les faits qu'à la condition que l'énergie interne totale et la densité sont toutes deux positives à l'issue de l'étape de prédiction, alors chacune des entropies spécifiques restera positive à l'issue de l'étape de correction. Soulignons pour conclure que les conditions de positivité sur la densité et l'énergie interne totale sont des conditions faibles indispensables à la poursuite des calculs (et ce déjà dans le cas d'une seule pression).

3.2 Formulation proposée

Conformément à la discussion précédente, nous proposons ci-dessous une nouvelle formulation des relations de saut généralisées (3.1) et (3.2), qui interviendra de manière centrale dans l'algorithme implicite présenté dans la section suivante.

Proposition 4
Soient $\{X_i\}_{i=1,...,N}$ les quantités définies par

$$X_i = \ln \frac{S_i}{S_N}, \quad i = 1, ..., N-1, \tag{3.7}$$

avec par convention $X_N = 0$. Alors les solutions régulières du système (110) vérifient les $(N-1)$ relations de proportionnalité suivantes :

$$\partial_t \rho X_i + \partial_x \rho X_i u = \tfrac{C_i}{\mu}(\partial_t \rho\varepsilon + \partial_x \rho\varepsilon u + p\partial_x u), \quad i = 1, ..., N-1, \tag{3.8}$$

où l'on a posé :

$$\rho\varepsilon = \sum_{j=1}^{N} \rho\varepsilon_j, \quad p = \sum_{j=1}^{N} p_j, \quad \mu = \sum_{j=1}^{N} \mu_j, \tag{3.9}$$

et

$$C_i = \frac{\mu_N}{T_N S_N} - \frac{\mu_i}{T_i S_i}, \quad i = 1, ..., N-1.$$

Notons que les rapports C_i/μ dépendent par construction des rapports des viscosités et qu'en ce sens, ils restent indépendants de tout paramètre ϵ de changement d'échelle (correspondant typiquement à l'inverse du nombre de Reynolds). Muni de la preuve de ce résultat, il sera facile d'établir en conséquence :
Proposition 5
Aves les notations de l'énoncé précédent, les solutions régulières du système (110)

vérifient de manière équivalente le système d'EDP suivant :

$$\begin{cases} \partial_t\rho + \partial_x\rho u = 0, \\ \partial_t\rho u + \partial_x(\rho u^2 + p) = \partial_x(\mu\partial_x u), \\ \partial_t\rho E + \partial_x(\rho E + p)u = \partial_x(\mu u\partial_x u), \\ \partial_t\rho X_i + \partial_x\rho X_i u = \frac{C_i}{\mu}(\partial_t\rho\varepsilon + \partial_x\rho\varepsilon u + p\partial_x u), \quad i = 1, ..., N-1, \end{cases} \tag{3.10}$$

avec la relation de fermeture :

$$\rho\varepsilon = \rho E - \frac{(\rho u)^2}{2\rho} = \sum_{i=1}^{N} \rho\varepsilon_i(1/\rho, S_i).$$

Démonstration
De manière à être complet, nous détaillons ci-dessous la preuve d'un résultat déjà esquissé.
Dans un premier temps, en sommant sur i les lois d'évolution sur les énergies internes $\{\varepsilon_i\}_{i=1,...,N}$ du système (110) (voir chapitre II 4 par exemple), nous obtenons directement l'égalité

$$\partial_t\rho\varepsilon + \partial_x\rho\varepsilon u + p\partial_x u = \mu(\partial_x u)^2, \tag{3.11}$$

en arguant des définitions (3.9).
Ensuite, les lois d'évolution (3.5) et la conservation de la masse présente dans le système considéré entraînent, pour les solutions régulières, la validité des lois d'évolution suivantes sur les entropies spécifiques S_i, $i = 1, ..., N$:

$$\rho(\partial_t S_i + u\partial_x S_i) = -\frac{\mu_i}{T_i}(\partial_x u)^2, \quad i = 1, ..., N,$$

de telle sorte que les propriétés élémentaires de la fonction logarithme entraînent successivement :

$$\rho(\partial_t \ln S_i + u\partial_x \ln S_i) = -\frac{\mu_i}{T_i S_i}(\partial_x u)^2, \quad i = 1, ..., N-1, \tag{3.12}$$

et

$$\rho(\partial_t \ln S_N + u\partial_x \ln S_N) = -\frac{\mu_N}{T_N S_N}(\partial_x u)^2. \tag{3.13}$$

La soustraction de (3.12) et (3.13) entraîne, d'après (3.7),

$$\rho(\partial_t X_i + u\partial_x X_i) = C_i(\partial_x u)^2, \quad i = 1, ..., N-1, \tag{3.14}$$

c'est-à-dire

$$\partial_t\rho X_i + \partial_x\rho X_i u = C_i(\partial_x u)^2, \quad i = 1, ..., N-1, \tag{3.15}$$

en invoquant une fois de plus la conservation de la masse. Nous obtenons alors les égalités (3.8) attendues en invoquant (3.11). Ceci conclut la démonstration.

3.3 Application à une formulation explicite en temps

L'objectif de ce paragraphe est d'illustrer brièvement la discrétisation des précédentes relations de proportionnalité dans le contexte d'une formulation explicite en temps. L'examen de ces formules permettra une extension naturelle au cadre d'une formulation implicite linéarisée en temps.

Supposons connue à l'instant t^n une approximation des inconnues du système considéré. Celle-ci est avancée à la date suivante t^{n+1} en trois pas comparables à ceux détaillés dans le chapitre II 2. Les deux premiers pas consistent respectivement en une résolution de l'opérateur hyperbolique sous-jacent suivie de sa perturbation dissipative. L'objet du troisième pas est de corriger les prédictions obtenues des entropies partielles conformément aux relations de compatibilité exhibées au paragraphe précédent.

De manière à motiver l'introduction de la discrétisation proposée, il est utile de rappeler brièvement l'algorithme de mise à jour des $(N+2)$ inconnues de notre problème, résultant d'une stratégie de prédiction-correction construite sur les énergies internes partielles telle qu'analysée au chapitre II 4. Répondant aux motivations décrites lors de l'introduction, notre objectif est ici de montrer comment définir naturellement une approximation consistante de l'expression :

$$\{\partial_t \rho\varepsilon + \partial_x \rho\varepsilon u + p\partial_x u\},$$

entièrement connue avant l'étape de correction. A cette fin, rappelons que l'étape de prédiction consiste à projeter sur les constantes la densité, l'impulsion ainsi que les N énergies internes de manière à obtenir les lois d'évolution discrètes suivantes :

$$(\rho\varepsilon_i)_j^{n+1-} - (\rho\varepsilon_i)_j^n + \frac{\Delta t}{\Delta x}\{\Delta\{\rho\varepsilon_i u\}_{j+1/2} + (p_i\partial_x u)_j^n\} = \Delta t \,"\mu_i(\partial_x u)^2",$$

où les produits non conservatifs impliqués trouvent une définition détaillée au chapitre II 4. Nous avons établi qu'en conséquence de ces projections L^2, l'énergie totale n'est pas conservée puisqu'elle décroît en norme L^1 en vertu de l'inégalité suivante :

$$\{\rho E\}_j^{n+1-} := \frac{(\rho u)_j^{n+1-\,2}}{2\rho_j^{n+1-}} + \sum_{i=1}^{N}(\rho\varepsilon_i)_j^{n+1-} \leq (\rho E)_j^n - \frac{\Delta t}{\Delta x}\Delta\{(\rho E + p)u\}_{j+1/2}^n.$$

Ce défaut de conservation a motivé l'introduction d'une étape de correction forçant la conservation de la densité, de l'impulsion et de l'énergie totale, pour redéfinir les N énergies internes comme solutions des $(N-1)$ relations suivantes :

$$\mu_N\{(\rho\varepsilon_i)_j^{n+1} - (\rho\varepsilon_i)_j^{n+1=}\} = \mu_i\{(\rho\varepsilon_N)_j^{n+1} - (\rho\varepsilon_N)_j^{n+1=}\}, \quad i = 1, ..., N-1,$$

complétées de :

$$\frac{(\rho u)_j^{n+1\,2}}{2\rho_j^{n+1}} + \sum_{i=1}^{N}(\rho\varepsilon_i)_j^{n+1} = (\rho E)_j^{n+1} := (\rho E)_j^n - \frac{\Delta t}{\Delta x}\Delta\{(\rho E + p)u\}_{j+1/2}^n.$$

Ici et conformément au chapitre II 4, nous avons posé :

$$(\rho\varepsilon_i)_j^{n+1=} = (\rho\varepsilon_i)_j^n - \frac{\Delta t}{\Delta x}\{\Delta\{\rho\varepsilon_i u\}_{j+1/2}^n + (p_i\partial_x u)_j^n\}, \quad i = 1, ..., N.$$

Des manipulations algébriques simples ont alors conduit aux N réactualisations suivantes :

$$(\rho\varepsilon_i)_j^{n+1} = (\rho\varepsilon_i)_j^{n+1=} + \frac{\mu_i}{\mu}\{(\rho\varepsilon)_j^{n+1} - (\rho\varepsilon)_j^{n+1=}\}, \quad i = 1, ..., N,$$

où nous avons posé successivement et par définition

$$(\rho\varepsilon)_j^{n+1} = (\rho E)_j^{n+1} - \frac{(\rho u)_j^{n+1^2}}{2\rho_j^{n+1}}, \tag{3.16}$$

puis :

$$(\rho\varepsilon)_j^{n+1=} = \sum_{l=1}^{N}(\rho\varepsilon_l)_j^{n+1=} = (\rho\varepsilon)_j^n - \frac{\Delta t}{\Delta x}\{\Delta\{\rho\varepsilon u\}_{j+1/2}^n + \sum_{i=1}^{N}(p_i\partial_x u)_j^n\}.$$

Ces N formules de remise à jour constituent par construction une discrétisation consistante des relations de proportionnalité suivantes :

$$\mu\{\partial_t\rho\varepsilon_i + \partial_x\rho\varepsilon_i u + p_i\partial_x u\} = \mu_i\{\partial_t\rho\varepsilon + \partial_x\rho\varepsilon u + p\partial_x u\}, \quad i = 1, ..., N, \tag{3.17}$$

en vertu de la définition des $(\rho\varepsilon_i)_j^{n+1=}$. Ainsi tout se passe comme si nous disposions de $(N+3)$ formules de mise à jour, respectivement pour $\rho, \rho u, \rho E$ et les N énergies internes, pour seulement $(N+2)$ inconnues. La linéarité de l'énergie totale (comprise comme fonction des inconnues discrètes) relativement à chacune des énergies internes est ici responsable de l'absence de conflit entre les $(N+3)$ égalités précédentes, au sens où l'égalité suivante a bien lieu :

$$(\rho E)_j^{n+1} = \{\rho E\}(\rho_j^{n+1}, (\rho u)_j^{n+1}, \{(\rho\varepsilon_i)_j^{n+1}\}_{1\leq i\leq N}).$$

Munis de ces remarques simples, tentons de répondre aux motivations développées lors de l'introduction de ce chapitre en substituant aux équations gouvernant les énergies internes leurs homologues exprimés en entropies. Arguant du second principe de la thermodynamique, les N relations de proportionnalité (3.17) s'écrivent de manière équivalente sur le plan continu :

$$\partial_t\rho S_i + \partial_x\rho S_i u = \frac{\mu_i}{\mu T_i}\{\partial_t\rho\varepsilon + \partial_x\rho\varepsilon u + p\partial_x u\}, \quad i = 1, ..., N. \tag{3.18}$$

Une discrétisation brutale de ces identités consisterait à envisager la procédure de correction suivante :

$$(\rho S_i)_j^{n+1} = (\rho S_i)_j^{n+1=} + \frac{\mu_i}{\mu(T_i)_j^n}\{(\rho\varepsilon)_j^{n+1} - (\rho\varepsilon)_j^{n+1=}\}, \quad i = 1, ..., N, \tag{3.19}$$

avec par construction :

$$(\rho S_i)_j^{n+1=} = (\rho S_i)_j^n - \frac{\Delta t}{\Delta x}\Delta\{\rho S_i u\}_{j+1/2}^n, \quad i = 1, ..., N.$$

Dans (3.19), $(\rho\varepsilon)_j^{n+1=}$ et $(\rho\varepsilon)_j^{n+1}$ conservent les définitions précédentes.
Toutefois, considérer simultanément les N formules de remise à jour (3.19) nécessiterait de comprendre désormais chaque énergie interne selon :

$$(\rho\varepsilon_i)_j^{n+1} = \{\rho\varepsilon_i\}(\frac{1}{\rho_j^{n+1}}, (S_i)_j^{n+1}), \quad i = 1, ..., N.$$

Les non linéarités impliquées dans ces N formules rentrent en règle générale en conflit avec la définition de l'énergie totale, imposant la satisfaction de :

$$(\rho\varepsilon)_j^{n+1} = (\rho E)_j^{n+1} - \frac{{(\rho u)_j^{n+1}}^2}{2\rho_j^{n+1}} = \sum_{i=1}^{N} (\rho\varepsilon_i)_j^{n+1}.$$

En d'autres termes, privilégier les N formules de mise à jour (3.19) se traduit par une violation de la conservation de l'énergie totale :

$$\{\rho E\}(\rho_j^{n+1}, (\rho u)_j^{n+1}, \{(\rho S_i)_j^{n+1}\}_{1\leq i\leq N}) \neq (\rho E)_j^n - \frac{\Delta t}{\Delta x}\Delta\{(\rho E + p)u\}_{j+1/2}^n.$$

Il serait certes possible de ne considérer que les $(N-1)$ premières mises à jour dans (3.19) et de réactualiser la dernière énergie interne $(\rho\varepsilon_N)_j^{n+1}$ (et donc $(S_N)_j^{n+1}$) en exigeant la conservation de l'énergie totale :

$$(\rho\varepsilon_N)_j^{n+1} = (\rho E)_j^{n+1} - \frac{{(\rho u)_j^{n+1}}^2}{2\rho_j^{n+1}} - \sum_{i=1}^{N-1} \{\rho\varepsilon_i\}(1/\rho_j^{n+1}, (S_i)_j^{n+1}).$$

Toutefois, il n'y a aucune raison pour qu'une telle remise à jour définisse une énergie interne positive dans le contexte des schémas implicites linéarisés en temps. Nous excluons donc cette stratégie.
Ces remarques justifient la formulation des relations de compatibilité construites sur les $(N-1)$ entropies X_i. En effet, une extension directe de la précédente discussion à cette nouvelle formulation conduit à définir les $(N-1)$ formules de mise à jour :

$$(\rho X_i)_j^{n+1} = (\rho X_i)_j^{n+1=} + \frac{(C_i)_j^n}{\mu}\{(\rho\varepsilon)_j^{n+1} - (\rho\varepsilon)_j^{n+1=}\}, \quad i = 1, ..., N-1, \tag{3.20}$$

avec

$$(\rho X_i)_j^{n+1=} = (\rho X_i)_j^n - \frac{\Delta t}{\Delta x}\Delta\{\rho X_i u\}_{j+1/2}^n, \quad i = 1, ..., N-1.$$

De sorte à définir chacune des entropies $(S_i)_j^{n+1}$, il s'agit ensuite de résoudre en l'inconnue $(S_N)_j^{n+1}$ l'équation suivante valable localement sur chaque cellule :

$$\sum_{i=1}^{N} \rho_j^{n+1}\varepsilon_i(1/\rho_j^{n+1}, (S_N)_j^{n+1}\exp((\rho X_i)_j^{n+1}/\rho_j^{n+1})) = (\rho E)_j^{n+1} - \frac{{(\rho u)_j^{n+1}}^2}{2\rho_j^{n+1}}, \tag{3.21}$$

pour obtenir, d'après (3.7) :

$$(\rho S_i)_j^{n+1} = \rho_j^{n+1}(S_N)_j^{n+1}\exp((\rho X_i)_j^{n+1}/\rho_j^{n+1}), \quad i = 1, ..., N-1. \tag{3.22}$$

Nous montrerons plus tard que la solution du problème algébrique considéré existe et est bien positive dès que la densité et l'énergie interne totale le sont. Cette hypothèse est manifestement plus faible que celle exhibée en faisant choix des entropies S_i comme ci-dessus. Qui plus est, nous verrons que sous l'hypothèse de N lois de pression polytropiques, la solution du problème algébrique est explicitement connue.

Afin de convaincre le lecteur du bien-fondé de l'approche explicite que nous venons

de décrire, nous proposons de reprendre la première simulation effectuée dans le chapitre II 2. Nous comparons la solution numérique générée par la présente approche construite sur la nouvelle formulation des relations de compatibilité, à celle obtenue par la technique de correction déduite d' une formulation directe en ρS_i (voir (3.1) et chapitre II 2). Pour le cas test considéré, les énergies internes $\{\varepsilon_i\}_{i=1,\dots,N}$ sont associées à des gaz parfaits polytropiques (les coefficients adiabatiques sont notés $\{\gamma_i\}_{i=1,\dots,N}$), de telle sorte que la nouvelle procédure de correction est explicite (ainsi qu'annoncé). Dans les figures présentées ci-dessous, nous appelons donc correction non linéaire l'approche du chapitre II 2 (puisqu'elle nécessite la résolution d'un problème algébrique non linéaire) et par opposition, correction linéaire la mise à jour explicite (3.20)-(3.21)-(3.22).
Nous faisons par ailleurs le choix $N = 3$ et renvoyons le lecteur au chapitre II 2 pour la définition exacte de la donnée initiale du problème de Riemann considéré.
Comme l'illustre la figure 3.1, les deux approches donnent des résultats tout à fait analogues. Ils sont en bon accord avec la solution exacte attendue et ce contrairement à la solution générée par une technique classique (sans étape de correction). Ceci suffit à illustrer la robustesse de la nouvelle stratégie de correction proposée.

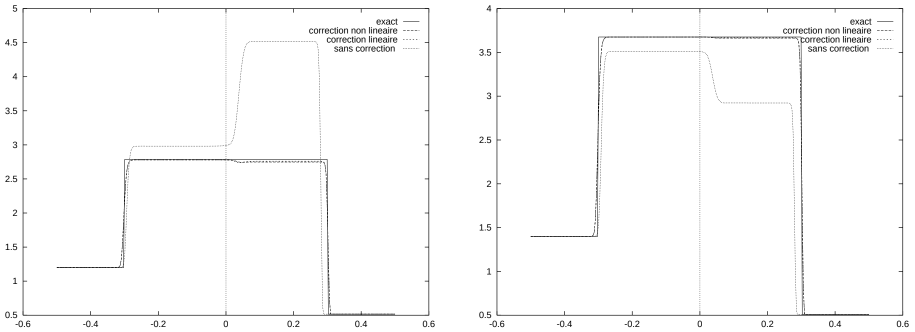

Fig. 3.1 – *Pressions 1 et 2 pour plusieurs corrections*

3.4 Application à une formulation implicite en temps

3.4.1 Algorithme

Nous décrivons ci-dessous une version implicite linéarisée en temps de la précédente étape de correction. Les deux premiers pas, dévolus à l'opérateur du premier ordre puis à l'opérateur du second ordre, reçoivent une formulation implicite linéarisée en temps déduite des précédents chapitres et brièvement résumée ci-après pour conclure. A l'issue du premier pas, nous pouvons évaluer la réactualisation partielle de l'énergie interne totale selon :

$$(\rho\varepsilon)_j^{n+1=} = \sum_{i=1}^{N} \rho\varepsilon_i(\rho_j^{n+1=}, (S_i)_j^{n+1=}), \tag{3.23}$$

où les valeurs prédites $(S_i)_j^{n+1=}$ ont été obtenues conformément à l'implicitation découplée du bloc transport des entropies (voir chapitre précédent pour plus de détails).

Rappelons que l'algorithme proposé préserve la positivité de chacune des entropies de sorte que la formule précédente a bien un sens. Au terme du second pas, nous disposons des valeurs définitivement réactualisées ρ_j^{n+1}, $(\rho u)_j^{n+1}$, $(\rho E)_j^{n+1}$, et sommes donc en mesure de définir à cette même date l'énergie interne totale :

$$(\rho\varepsilon)_j^{n+1} = (\rho E)_j^{n+1} - \frac{((\rho u)_j^{n+1})^2}{2\rho_j^{n+1}}.$$

Ces deux quantités étant connues avant correction des entropies spécifiques S_i, nous suggérons d'impliciter en temps l'algorithme proposé en considérant :

$$\begin{aligned}(\rho X_i)_j^{n+1} - (\rho X_i)_j^n + \lambda\Delta(\{\rho u\}_{j+1/2}^{n,+}(X_i)_j^{n+1} + \{\rho u\}_{j+1/2}^{n,-}(X_i)_{j+1}^{n+1}) = \\ \tfrac{(C_i)_j^n}{\mu}((\rho\varepsilon)_j^{n+1} - (\rho\varepsilon)_j^{n+1=}), \quad i = 1, ..., N-1.\end{aligned} \tag{3.24}$$

Nous avons donc choisi ici une implicitation linéarisée en temps immédiate des formules des flux de Larrouturou :

$$\{\rho X_i u\}_{j+1/2} = \{\rho u\}_{j+1/2}^{n,+} \times (X_i)_j^{n+1} + \{\rho u\}_{j+1/2}^{n,-} \times (X_i)_{j+1}^{n+1}, \tag{3.25}$$

où classiquement :

$$X^+ = \max(X, 0), \quad X^- = \min(X, 0), \quad \forall\ X \in I\!R.$$

Les flux numériques $\{\rho u\}_{j+1/2}^n$ sont associés à la méthode de Relaxation décrite dans le chapitre 1 et sont connus à l'issue du premier pas (puisque rentrant dans la définition du résidu explicite sur la variable ρ).

Notons qu'une analyse simple de consistance au sens des différences finies permet de se convaincre que $\{(\rho\varepsilon)_j^{n+1} - (\rho\varepsilon)_j^{n+1=}\}/\Delta t$ réalise bien une approximation en temps-espace de l'expression $\partial_t\rho\varepsilon + \partial_x\rho\varepsilon u + p\partial_x u$.

La mise en œuvre correspondante pour le calcul des inconnues $\{(X_i)_j^{n+1}\}_{j\in Z}$ consiste, pour chaque $i = 1, ..., N-1$, en la résolution d'un système linéaire dont la matrice possède une structure tridiagonale à coefficients scalaires. De manière plus précise et d'après (3.24), une ligne courante j de cette matrice est constituée des éléments a_j, b_j et c_j vérifiant :

$$\begin{aligned} a_j &= -\tfrac{\Delta t}{\Delta x}\{\rho u\}_{j-1/2}^{n,+}, \\ b_j &= \rho_j^{n+1} + \tfrac{\Delta t}{\Delta x}(\{\rho u\}_{j-1/2}^{n,+} - \{\rho u\}_{j+1/2}^{n,-}), \\ c_j &= \tfrac{\Delta t}{\Delta x}\{\rho u\}_{j+1/2}^{n,-}. \end{aligned}$$

Cette matrice est une M-matrice à diagonale strictement dominante (indépendamment de la condition CFL). Elle est donc inversible sous des conditions aux limites *adhoc*, ce qui garantit l'existence et l'unicité des quantités $(X_i)_j^{n+1}$ ainsi définies. Disposant localement des $(N-1)$ quantités $(X_i)_j^{n+1}$, nous sommes alors en mesure de résoudre localement sur chaque cellule le problème algébrique (3.21) permettant de définir les valeurs réactualisées des entropies selon (3.22) (voir la justification ci-dessous).

Ceci achève la description de l'implicitation de l'étape de correction.

3.4.2 Positivité des entropies

L'énoncé suivant assure l'existence et l'unicité d'une solution S_N strictement positive du problème algébrique (3.21), dès que la densité et l'énergie interne totale sont positives à l'issue du premier pas. En conséquence de cela, chacune des entropies spécifiques reste positive sous cette hypothèse faible à l'issue de l'étape de correction implicite (voir en effet (3.22)). Ceci justifie la stratégie adoptée.

Proposition 6
Soit $(\rho, \rho u, \rho E, \{\rho X_i\}_{i=1,...,N})$ un vecteur donné de $I\!R^{N+3}$ tel que

$$\rho > 0, \quad (\rho E) - \frac{(\rho u)^2}{2\rho} > 0, \tag{3.26}$$

et par convention $\rho X_N = 0$. Soit $S_N \to f(S_N)$ l'application de $I\!R^{+\star} \to I\!R$ définie par

$$f(S_N) = (\rho E) - \frac{(\rho u)^2}{2\rho} - \sum_{i=1}^{N} \rho\varepsilon_i(1/\rho, S_N \exp((\rho X_i)/\rho)). \tag{3.27}$$

Alors, il existe un unique $S_N^\star \in I\!R^{+\star}$ tel que $f(S_N^\star) = 0$.

Nous verrons dans la démonstration ci-dessous que l'application f ainsi définie est strictement concave et strictement croissante sur $I\!R^{+,\star}$. En conséquence, un algorithme de Newton permet de déterminer rapidement et précisément la solution $S_N^\star \in I\!R^{+\star}$ recherchée.

Démonstration
Par dérivation de l'application f définie par (3.27), nous obtenons l'inégalité :

$$f'(S_N) = -\sum_{i=1}^{N} \rho \exp((\rho X_i)/\rho) \frac{\partial \varepsilon_i}{\partial S_i}(1/\rho, S_N \exp((\rho X_i)/\rho)) > 0,$$

puisque par hypothèse,

$$\rho > 0 \quad \text{et} \quad \frac{\partial \varepsilon_i}{\partial S_i}(1/\rho, S_N \exp((\rho X_i)/\rho)) = -T_i(1/\rho, S_N \exp((\rho X_i)/\rho)) < 0.$$

Nous obtenons facilement en dérivant une seconde fois la fonction f l'égalité

$$f''(S_N) = -\sum_{i=1}^{N} \rho \exp(2(\rho X_i)/\rho) \frac{\partial^2 \varepsilon_i}{\partial^2 S_i}(1/\rho, S_N \exp((\rho X_i)/\rho)),$$

c'est-à-dire $f''(S_N) < 0$ par convexité des énergies internes $(\tau, S_i) \to \varepsilon(\tau, S_i)$ pour tout $i = 1, ..., N$. L'application f est donc strictement croissante et strictement concave. Les hypothèses asymptotiques de Weyl vérifiées par les énergies internes $\{\varepsilon_i\}_{i=1,...,N}$ assurent que pour tout $i = 1, ..., N$:

$$\lim_{S_i \to 0} \varepsilon_i(\tau, S_i) = +\infty, \quad \lim_{S_i \to +\infty} \varepsilon_i(\tau, S_i) = 0,$$

de telle sorte que, d'après (3.26) :

$$\lim_{S_N \to 0} f(S_N) = -\infty, \quad \lim_{S_N \to +\infty} f(S_N) = \rho E - \frac{(\rho u)^2}{2\rho} > 0.$$

L'existence et l'unicité de la solution de l'équation $f(S_N) = 0$ sont assurées. Ceci conclut la démonstration.

3.4.3 Le cas polytropique

Nous vérifions ici que sous l'hypothèse de N lois de pression polytropiques, la résolution du problème algébrique local (3.21) fixant la dernière entropie admet une solution explicite. A cette fin, il convient de faire un choix particulier de lois d'état consistant à considérer les entropies spécifiques suivantes :

$$s_i = \frac{p_i}{\rho^{\gamma_i}}, \quad \gamma_i > 1, \quad i = 1, ..., N.$$

Un tel choix nous fait certes perdre la propriété de stricte convexité des énergies internes associées (la preuve du précédent théorème tombe donc) mais autorise une résolution explicite du problème considéré garantissant toujours la positivité des entropies recherchées et sous les mêmes hypothèses. En effet, il suffit de constater que l'énergie interne totale s'écrit :

$$\rho\varepsilon = \sum_{i=1}^{N} \frac{p_i}{\gamma_i - 1} = s_N \times \{\sum_{i=1}^{N} \frac{\rho^{\gamma_i}}{\gamma_i - 1} \exp(X_i)\},$$

de sorte que, lorsque la densité et l'énergie interne totale sont positives, nous avons $s_N > 0$. Ici nous avons évidemment défini $X_i = \ln(s_i/s_N)$.

3.5 Description de l'algorithme implicite complet

Nous résumons brièvement dans cette section la formulation implicite linéarisée en temps de chacun des trois pas constituant une itération (voir chapitre II 2). Notre objectif est de remarquer qu'une fois la première itération accomplie, il n'est plus utile dans la pratique de procéder à la résolution du bloc transport des entropies dans la phase de prédiction (voir chapitre 2 pour plus de détails). En anticipant la discussion, nous pouvons en effet tirer profit du découplage entre l'étape de prédiction et l'étape de correction. Il est ainsi possible d'exploiter la connaissance des entropies corrigées au terme de l'itération précédente de manière à évaluer, lors du premier pas de l'itération courante, la correction des résidus explicites (voir formules (2.47)-(2.48)) ainsi que l'énergie interne totale prédite (voir formule (3.23)). La pratique a montré que cette simplification conduit à un algorithme stable présentant le bénéfice de n'avoir à résoudre à chaque itération que N systèmes linéaires tridiagonaux scalaires (contre $2N - 1$ sinon) et d'un système linéaire tridiagonal par blocs de taille 3×3. Le coût de l'algorithme définitif est en ce sens optimal.

Premier pas : opérateur convectif ($t^n \to t^{n+1=}$)
Nous suggérons d'utiliser lors de ce premier pas l'implicitation linéarisée en temps,

décrite au chapitre 2, du système Euler multi-pression. Conformément à ce chapitre, l'implicitation linéarisée en temps du bloc hydrodynamique requiert une réactualisation des entropies spécifiques de manière à rendre compte de l'implicitation linéarisée de la projection sur la variété d'équilibre (voir la définition des résidus explicites corrigés (2.47)-(2.48)). Après la première (quelques) itération(s), nous suggérons de tirer profit du découplage entre ce premier pas prédictif et le pas correctif en utilisant la connaissance des entropies corrigées à l'issue de l'itération précédente pour évaluer les résidus explicites corrigés (2.47)-(2.48). La pratique justifie un tel mode opératoire.

Deuxième pas : opérateur diffusif ($t^{n+1=} \to t^{n+1-}$)
La procédure décrite dans le chapitre II 2 pour résoudre l'opérateur du second ordre étant déjà implicite, nous pouvons la conserver sans aucune modification. Le coût calcul correspondant est particulièrement intéressant puisqu'il consiste en la résolution d'un unique système linéaire inversible permettant d'évaluer les vitesses discrètes $(u_j^{n+1-})_{j\in Z}$ et dont la matrice en j possède une structure tridiagonale scalaire. Les évolutions implicites ρ_j^{n+1-}, $(\rho u)_j^{n+1-}$, $(\rho E)_j^{n+1-}$ et $(\rho S_i)_j^{n+1-}$, $i = 1, ..., N-1$, sont ensuite explicitement données (voir chapitre II 2 pour les détails). Notons simplement que dans les simulations numériques $2D$ menées dans le contexte du code CELHYO, les deux premières étapes n'en constituent qu'une seule, au prix d'un algorithme d'implicitation de l'opérateur diffusif légèrement modifié (et non discuté ici car classique).

Troisième pas : étape de correction ($t^{n+1-} \to t^{n+1}$)
Elle a fait l'objet de la section précédente.

3.6 Expériences numériques

Le principal objectif de cette section consiste à tester la stabilité de l'implicitation proposée de l'étape de correction, qui est la seule réelle nouveauté de la section précédente. Celle de la méthode de Relaxation intervenant au titre de première étape a déjà été testée (avec succès) dans le chapitre 1.

Nous proposons d'effectuer des simulations numériques bi-dimensionnelles. La description en une dimension d'espace de la procédure numérique s'étend naturellement au cas multi-dimensionnel, en invoquant l'invariance par rotation des équations. Pour plus de détails sur ce point, nous renvoyons le lecteur au chapitre 1 dans lequel une telle stratégie a déjà été adoptée, où plus généralement à [57].
Nous approchons plusieurs écoulements stationnaires autour d'un cylindre, supersoniques à l'infini amont. Pour cela, nous utilisons un maillage structuré avec ni points de discrétisation sur la paroi du cylindre et nj points de discrétisation sur l'axe Ox (direction de la ligne d'arrêt). L'ordre 2 d'approximation en espace est obtenu par une approche MUSCL sur les variables primitives, à savoir la densité, la quantité de mouvement et les pressions partielles $\{p_i\}_{i=1,...,N}$. Le limiteur de pente utilisé est celui proposé par Van Albada et al. ($\phi(r) = (r^2+r)/(1+r^2)$ avec des notations classiques). Pour chaque cas test, après avoir précisé les données initiales correspondantes, nous présentons les iso-valeurs du champ de densité. Les distributions des pressions sur la ligne d'arrêt seront également présentées.

Nous proposons ci-dessous trois simulations numériques : les deux premières sont réalisées dans le régime asymptotique d'un nombre de Reynolds infini, la dernière est

menée pour un nombre de Reynolds (construit sur le rayon du cylindre) de 10^5. Nous nous sommes placés dans le contexte le plus défavorable de rapports de viscosités égaux à l'unité (voir les chapitres précédents et également Berthon [10]).
L'intérêt des deux premières simulations est de tester la robustesse de l'implicitation linéarisée de l'étape de correction indépendamment de l'implicitation linéarisée des contributions diffusives portant sur l'impulsion et l'énergie totale. Notons que si le nombre de Reynolds est infini, la procédure de correction doit être considérée puisque les $(N-1)$ rapports C_i/μ sont indépendants de ce nombre.
Concernant les conditions aux limites relatives aux deux premiers cas tests, nous utilisons les conditions aux limites de type Euler suivantes :

- condition de glissement pour le vecteur vitesse sur la paroi du cylindre,
- condition supersonique pour la sortie,
- condition de symétrie sur la ligne d'arrêt.

La troisième simulation consiste quant à elle en un véritable calcul Navier-Stokes (hormis les lois de Fourier ici négligées), avec les conditions aux limites suivantes :

- condition d'adhérence pour le vecteur vitesse sur la paroi du cylindre,
- condition supersonique pour la sortie,
- condition de symétrie sur la ligne d'arrêt.

Nous utilisons pour ce calcul une loi de viscosité de Sutherland et fixons toutes les viscosités égales :

$$\mu_i = (1.458.10^{-6} T_N^{3/2})/(110.4 + T_N), \quad i = 1, ..., N,$$

de telle sorte qu'une fois de plus, les rapports des viscosités en considération sont tous égaux à 1.

Nous décrivons maintenant les données initiales correspondantes à chacune de ces trois simulations, en notant $\omega = (u, v) \in I\!R^2$ le champ de vitesse dans le repère usuel (O, Ox, Oy).

Expérience 1
Pour ce premier cas test, nous considérons $N = 2$ pressions partielles p_1 et p_2 associées respectivement aux coefficients adiabatiques $\gamma_1 = 1.3$ et $\gamma_2 = 1.2$. Le maillage utilisé (et représenté sur la figure 3.2) est tel que $ni = 32$ et $nj = 41$. Le champ arrivant sur le cylindre est caractérisé par les valeurs suivantes données en unités internationales :

Expérience 1

	ρ	u	v	p_1	p_2
Champ amont	7.032e-4.	4318.	0.	36.	70.

Le nombre de Mach correspondant est environ égal à 10.

Expérience 2
Nous posons $N = 3$, $\gamma_1 = 1.4$, $\gamma_2 = 1.3$ et $\gamma_3 = 1.2$. Le maillage correspondant (voir

figure 3.5) est tel que $(ni, nj) = (32, 88)$. Le champ amont vérifie :

Expérience 2

	ρ	u	v	p_1	p_2	p_3
Champ amont	4.018e-4.	964.	0.	10.	26.	40.

Le nombre de Mach associé est proche de 2.

Expérience 3

Pour ce cas test, nous considérons trois pressions partielles associées aux coefficients adiabatiques $\gamma_1 = 1.4$, $\gamma_2 = 1.3$ et $\gamma_3 = 1.2$. Le maillage est représenté sur la figure 3.8. Il vérifie $ni = 60$ et $nj = 48$. Le champ amont considéré est le suivant :

Expérience 3

	ρ	u	v	p_1	p_2	p_3	T_3
Champ amont	4.018e-4.	1150.	0.	10.	20.	40.	335.23

Le nombre de Mach est environ égal à 2.35 tandis que le nombre de Reynolds pour ce calcul est de l'ordre de 10^5.

Présentation des résultats

Avant de présenter les solutions stationnaires obtenues pour ces différents problèmes, nous tenons à rapporter une difficulté opératoire. A cette fin, mentionnons que notre initialisation de la solution discrète consiste à choisir un champ uniformément constant, prescrit par les conditions d'écoulement à l'infini amont. En conséquence, le choc discret attendu prend naissance à la paroi, puis s'en détache pour venir se stationnariser à sa position finale à l'amont du cylindre. Nous avons constaté expérimentalement (au moins pour l'expérience 3) que tant que le choc discret n'avait pas quitté la paroi, la procédure d'implicitation de l'étape de correction s'est révélée être instable sauf à recourir à des petits nombres CFL (typiquement de l'ordre de 1). Dans la pratique, nous avons choisi de faire converger rapidement l'algorithme vers une solution quasi-stationnaire en désactivant la procédure de correction. Une fois cette première solution obtenue, nous avons brutalement rétabli cette procédure de correction pour de nouveau converger en temps. Nous n'avons alors pas rencontré de difficultés particulières : le nombre CFL retrouve rapidement la valeur qu'il possédait à convergence de la première solution. Soulignons que cette première solution est fort éloignée de la solution finale attendue (comme l'attestent les résultats 1D avec et sans correction). En ce sens, l'implicitation de l'étape de correction semble être robuste. Nous privilégions ici une erreur de codage informatique que nous ne sommes pas parvenus à localiser, même si nous suspectons un interfaçage défectueux de nos routines avec le code CELHYO dans le traitement des conditions aux limites paroi.

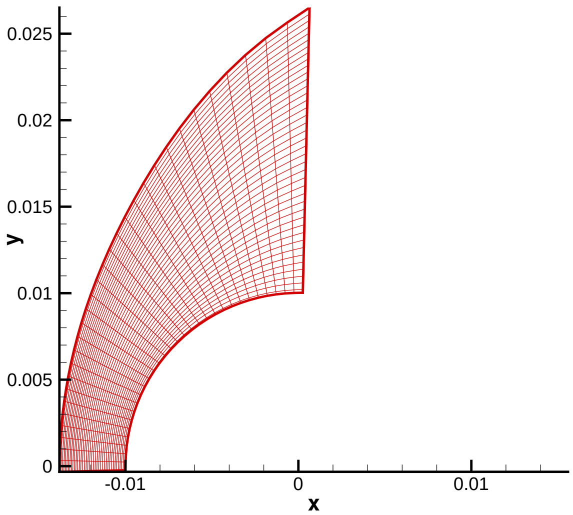

FIG. 3.2 – *Expérience 1. Maillage 32×41*

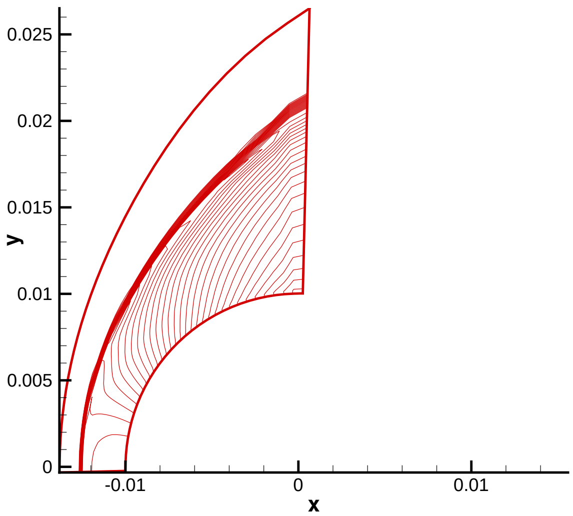

FIG. 3.3 – *Expérience 1. Densité : iso-valeurs*

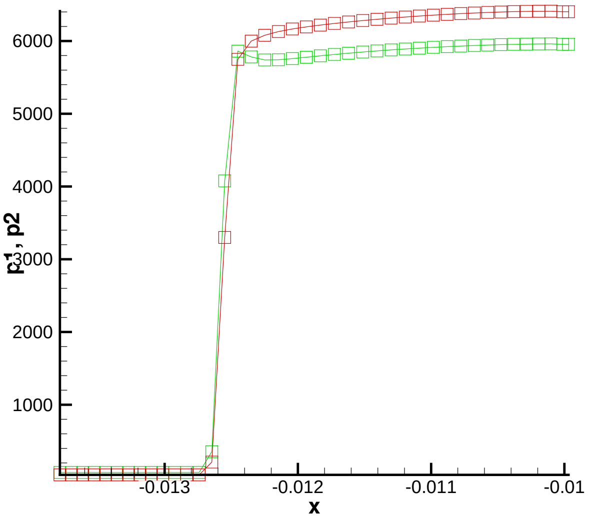

Fig. 3.4 – *Expérience 1. Pressions sur la ligne d'arrêt*

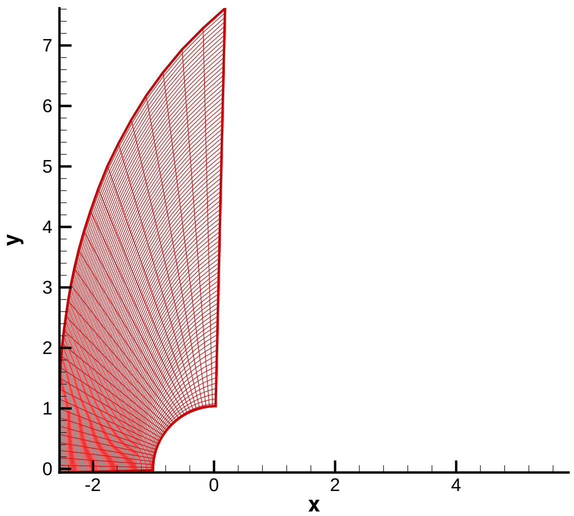

Fig. 3.5 – *Expérience 2. Maillage 32×88*

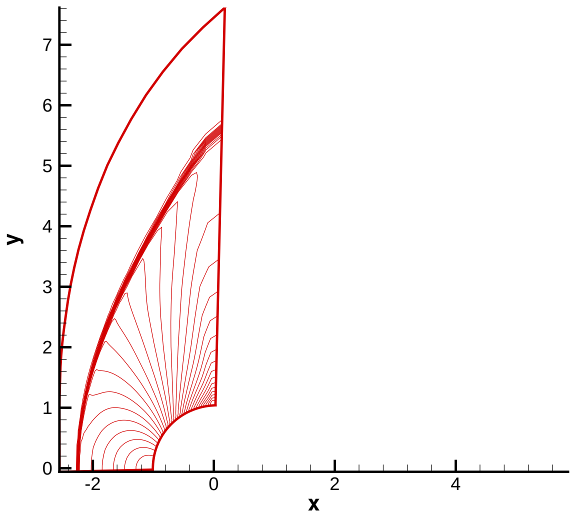

FIG. 3.6 – *Expérience 2. Densité: iso-valeurs*

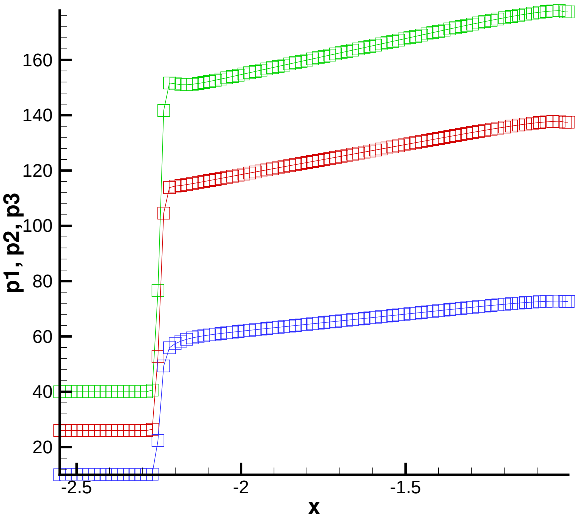

FIG. 3.7 – *Expérience 2. Pressions sur la ligne d'arrêt*

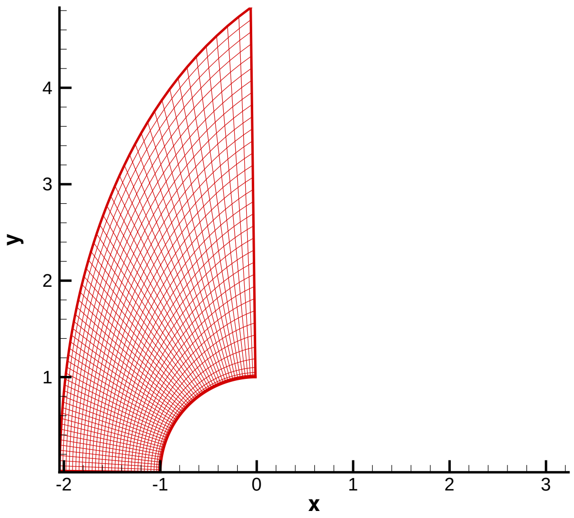

Fig. 3.8 – *Expérience 3. Maillage 60×48*

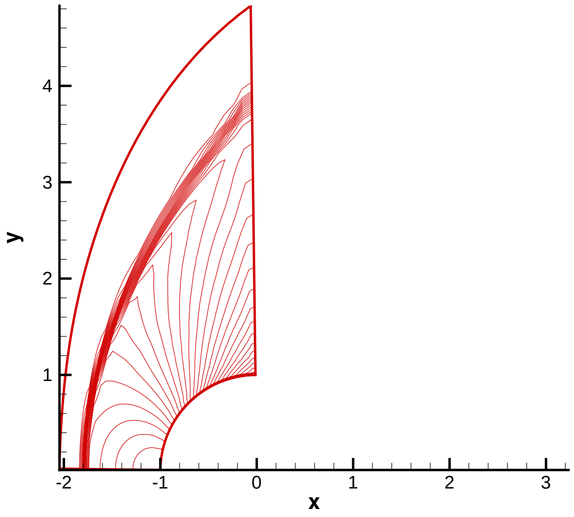

Fig. 3.9 – *Expérience 3. Densité : iso-valeurs*

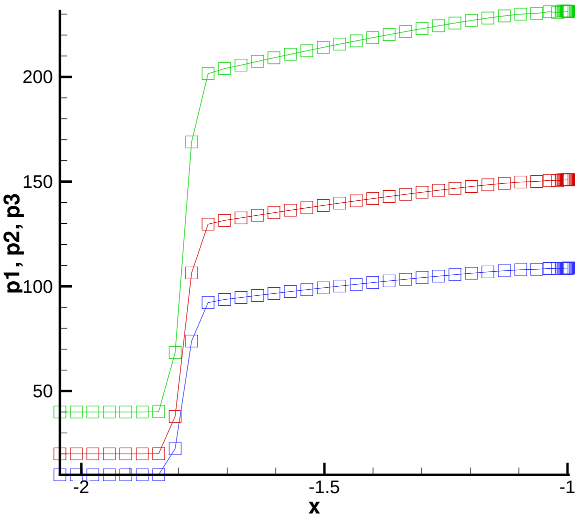

FIG. 3.10 – *Expérience 3. Pressions sur la ligne d'arrêt*

Chapitre 4

Application aux modèles multi-échelle de turbulence compressible homogène isotrope

L'objectif de ce chapitre est d'appliquer les considérations théoriques et numériques développées dans les précédents chapitres à une classe de modèles de turbulence compressible homogène isotrope. Nous considérons les modèles dits multi-échelle. Ces modèles s'inscrivent dans le formalisme des équations de Navier-Stokes à plusieurs pressions indépendantes proposé et étudié jusqu'à présent.

Nous débutons ce chapitre par des rappels sur les écoulements compressibles turbulents, l'objectif étant de situer les modèles multi-échelle dans un cadre général de représentation et de simulation des phénomènes de turbulence compressible (section 4.1). Nous présentons et analysons ensuite les équations décrivant ces modèles, pour s'apercevoir qu'ils constituent une extension naturelle des équations de Navier-Stokes à plusieurs pressions indépendantes. Celles-ci sont complétées par des équations d'évolution sur des grandeurs turbulentes moyennes à préciser (section 4.2). Nous en proposons alors une approximation numérique consistante, basée sur les procédures développées dans les chapitres précédents pour approcher les solutions des équations de Navier-Stokes à plusieurs pressions (section 4.3).

4.1 Quelques brèves notions de turbulence

Dans cette section, nous rappelons quelques notions de turbulence afin de situer dans un contexte général les modèles de turbulence compressible multi-échelle étudiés dans la section suivante. Nous insistons sur le fait que cette section ne constitue pas une présentation, ni même une introduction à la turbulence. Pour cela, nous renvoyons le lecteur aux ouvrages de référence Schiestel [116], Lesieur [98], Piquet [111], Sagault [114], ainsi qu'à la thèse de Gleize [54], ... Cette section est largement inspirée de ces travaux. Nous renvoyons également à l'ouvrage de Mohammadi-Pironneau [106] et à celui de Hirsch [68] pour l'exposé de méthodes d'approximation numérique.

4.1.1 Les écoulements turbulents

La notion de turbulence caractérise certains écoulements de fluide. S'il n'existe pas de définition précise de la turbulence, on peut cependant la caractériser par plusieurs critères. Les plus souvent retenus sont :
. le caractère aléatoire des fluctuations spatiales et temporelles du champ des vitesses,
. la grande capacité de mélange,
. le caractère chaotique de la solution, qui exhibe une très forte dépendance par rapport à la condition initiale et aux conditions aux limites.
Par ailleurs, la turbulence se traduit par un mouvement tourbillonnaire présentant un large spectre de dimensions de tourbillons et un spectre correspondant en fréquence. Ce mouvement peut être conçu comme un enchevêtrement de structures tourbillonnantes dont les vecteurs rotationnels sont orientés dans toutes les directions et sont fortement instationnaires. Les plus gros tourbillons (associés aux basses fréquences) sont en général déterminés par les conditions aux limites de l'écoulement et leurs dimensions sont alors de l'ordre de grandeur du domaine occupé par l'écoulement. Ils extraient de l'énergie cinétique du mouvement moyen pour la fournir aux agitations à grandes échelles (cf figure 4.1), et sont animés des vitesses les plus faibles. L'interaction de ces tourbillons entre eux provoque leurs déformations et étirements. Cet allongement des filets tourbillonnaires conduit à un transfert progressif de l'énergie cinétique à des échelles de plus en plus petites jusqu'à ce que les forces visqueuses deviennent actives et la dissipent. C'est la cascade d'énergie. L'évolution des petits tourbillons (associés aux hautes fréquences) est déterminée par les forces visqueuses. Ils sont le résultat de la dégradation des plus gros tourbillons par l'étirement. Ils sont dissipés par les forces visqueuses et leurs énergies cinétiques sont transformées en énergie interne thermodynamique.

4.1.2 Les principales méthodes de simulation de la turbulence

Les différentes méthodes de simulation de la turbulence peuvent être regroupées en trois grandes catégories : la simulation directe (DNS ou *Direct Numerical Simulation*), la simulation des grandes échelles (LES ou *Large Eddy Simulation*) et les méthodes statistiques (RANS ou *Reynolds Averaged Navier-Stokes*).

• La DNS

La DNS consiste à résoudre directement les équations de Navier-Stokes. Pour un écoulement turbulent, une telle résolution nécessite la prise en compte de la dynamique de toutes les échelles spatio-temporelles de la solution. Cette précision est indispensable pour observer la turbulence de façon détaillée, et obtenir des prédictions intéressantes pour les applications. Cependant, les écoulements turbulents contenant des échelles de tailles très différentes, ces calculs s'avèrent la plupart du temps impossibles à réaliser dans la pratique. Ainsi si l'on désigne par L et l les tailles caractéristiques respectivement la plus grande (largeur d'une plaque plane, d'un canal) et la plus petite (taille des régions tourbillonnaires) du problème, le nombre de points du maillage correspondant est de l'ordre de $(L/l)^d$, d représentant la dimension de l'espace. Les théories de

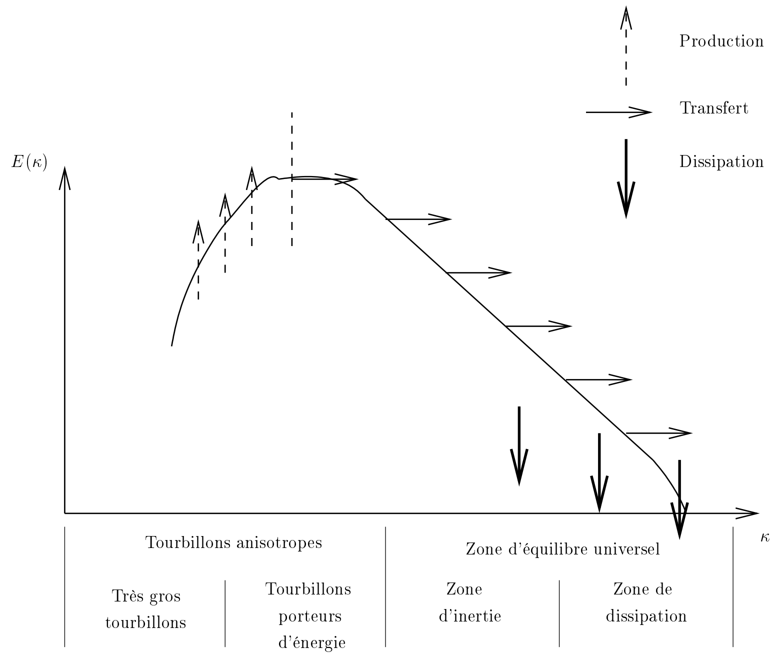

FIG. 4.1 – *Spectre d'énergie turbulente*

cascade turbulente (Kolmogorov par exemple) permettent d'estimer l :

$$l \simeq Re^{-n}, \quad \frac{3}{4} \leq n \leq 1.$$

Les nombres de Reynolds Re typiques en aérodynamique classique étant de l'ordre de 10^6, on vérifie aisément que le nombre de points de discrétisation nécessaires pour représenter les plus petites échelles de la turbulence à grand nombre de Reynolds atteint des valeurs irréalistes en regard des capacités des ordinateurs actuels. Cette approche ne peut donc être utilisée que pour simuler des écoulements simples à faible nombre de Reynolds.

Pour rendre des calculs plus généraux envisageables, il est nécessaire de simplifier l'approche précédente en ne resolvant plus directement la dynamique de toutes les échelles de la solution exacte. Pour cela, il faut introduire un nouveau niveau de description, moins fin, du sytème fluide. Ceci revient à opérer une sélection parmi les échelles. Certaines d'entre elles (les échelles les plus grandes) seront simulées directement lors du calcul alors que d'autres (typiquement les plus petites) devront être modélisées. Les méthodes LES et RANS suivent ce principe.

• **La LES**

L'approche LES est un intermédiaire entre le calcul direct (DNS) et l'approche statistique (RANS). Comme énoncé précédemment, la base de cette méthode est une sélection des échelles. On définit une longueur caractéristique dite *longueur de coupure* : sont appelées grandes échelles ou échelles résolues celles qui sont d'une taille caractéristique plus grande que la longueur de coupure et petites échelles ou échelles sous-mailles les autres. La *longueur de coupure* est donc ici typiquement de l'ordre de la taille de la maille. Ainsi, toutes les échelles du mouvement turbulent sont résolues jusqu'à une taille de l'ordre de la maille, et la dissipation par les tourbillons d'échelles inférieures est modélisée par un modèle de sous-mailles. Cette technique permet de rendre réaliste l'approximation d'écoulements à nombre de Reynolds élevé. Pour plus de détails, nous renvoyons le lecteur à Sagault [114].

• **Les méthodes RANS**

L'approche RANS est la plus utilisée pour les calculs d'ingénierie. Le principe consiste à ne calculer qu'une moyenne statistique de la solution exacte $\mathbf{u}$. Celle-ci est décomposée comme la somme de sa moyenne statistique $<\mathbf{u}>$ et d'une fluctuation $\mathbf{u}'$. Cette dernière n'est pas représentée directement par la réalisation numérique et n'est prise en compte que par le biais d'un modèle de turbulence. Il est ainsi possible de réduire considérablement le nombre de points de discrétisation. En revanche, le caractère statistique de la solution ne permet pas une description fine des mécanismes physiques, et ne donne accès qu'aux propriétés moyennes de l'écoulement.
Parmi les méthodes RANS, il existe plusieurs principes : le principe d'échelle unique et le principe d'échelles multiples.

Le principe d'échelle unique
Dans ce principe, on ne s'intéresse qu'aux propriétés les plus grossières de l'écoulement en évacuant toute considération relative aux interactions entre les différentes échelles. L'avantage est de ne travailler que sur des quantités dynamiques très simples (l'énergie cinétique de la turbulence et le taux de dissipation par exemple). L'inconvénient est de perdre un nombre important d'informations dynamiques du fait de la négligence des interactions entre échelles. On peut citer comme exemple le modèle $k-\epsilon$.

Le principe d'échelles multiples
L'hypothèse précédente consistant à décrire tout le champ turbulent à l'aide d'une échelle unique est dans les faits très discutable (équilibre spectral, ...). Ainsi, l'idée de décrire la turbulence à l'aide de plusieurs échelles, chacune caractéristique de processus physiques bien distincts, apparaît naturellement (cf figure 4.2).
Les modèles utilisant ce principe sont appelés **modèles multi-échelle** (ou modèles à échelles multiples), et constituent la principale motivation des travaux effectués dans la section suivante. Pour plus de détails sur les fondements de ces modèles, nous renvoyons le lecteur à Schiestel [116], Gleize [54], ...

4.2 Les modèles multi-échelle

L'objectif de cette section est de présenter les modèles multi-échelle de turbulence compressible homogène isotrope, et de les réinterpréter comme des équations de Navier-Stokes à plusieurs pressions indépendantes : la pression laminaire et des pressions turbulentes. Ces équations seront complétées par des lois d'évolution sur des grandeurs turbulentes que nous préciserons par la suite.

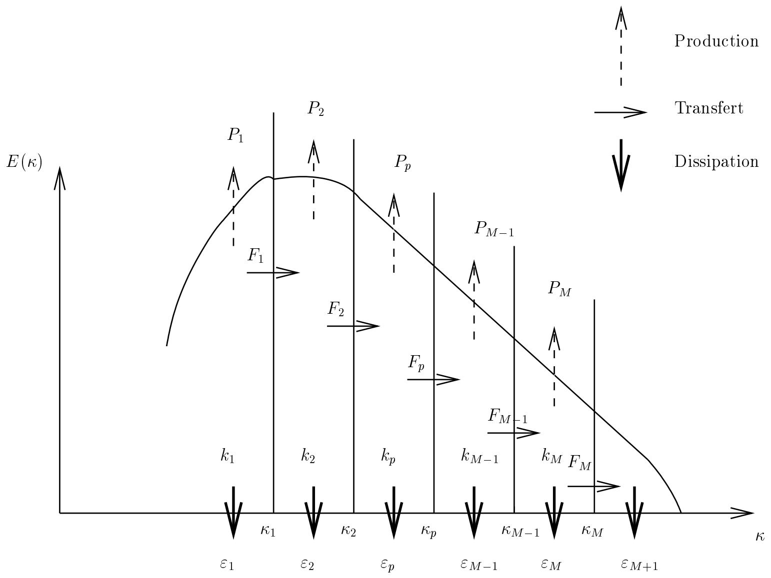

FIG. 4.2 – *Découpage du spectre d'énergie en M zones*

Sous l'hypothèse de grands nombres de Reynolds, les modèles multi-échelle considérés sont décrits par le système d'équations aux dérivées partielles non linéaires suivant :

$$\left\{\begin{array}{l}
\partial_t\rho + \partial_x\rho u = 0, \\[2ex]
\partial_t\rho u + \partial_x(\rho u^2 + p + (\gamma_\tau - 1)\sum_{m=1}^{M}\rho k^{(m)}) = \partial_x((\mu + \sum_{m=1}^{M}\mu_\tau^{(m)})\partial_x u), \\[2ex]
\partial_t\rho E + \partial_x(\rho E + p + (\gamma_\tau - 1)\sum_{m=1}^{M}\rho k^{(m)})u = \\
\partial_x((\mu + \sum_{m=1}^{M}\mu_\tau^{(m)})u\partial_x u) + \partial_x(\kappa\partial_x T) + \sum_{m=1}^{M}\partial_x(\kappa_m\partial_x k^{(m)}), \\[2ex]
\partial_t\rho k^{(m)} + \partial_x\rho k^{(m)}u = \\
P_k^{(m)} + \rho(F^{(m-1)} - F^{(m)}) + \partial_x(\kappa_m\partial_x k^{(m)}), \\[2ex]
\partial_t\rho F^{(m)} + \partial_x\rho F^{(m)}u = \\
C_0^{(m)}\frac{F^{(m)}}{k^{(m)}}P_k^{(m)} + \rho\frac{F^{(m)}}{k^{(m)}}(C_1^{(m)}F^{(m-1)} - C_2^{(m)}F^{(m)}) + \partial_x(\tilde{\kappa}_m\partial_x F^{(m)}), \\
\qquad\qquad\qquad m = 1, ..., M,
\end{array}\right. \tag{4.1}$$

où, de manière classique, ρ, ρu et ρE représentent respectivement la densité, la quantité de mouvement et l'énergie totale du fluide considéré. Notons que par souci de simplicité dans les notations, le système d'équations est présenté ici en une seule dimension d'espace.

Dans ce modèle, M, entier naturel avec $M \geq 1$, représente le nombre d'échelles utilisées pour décrire la turbulence et décomposer le spectre d'énergie correspondant. Chacune de ces échelles étant caractéristique de processus physiques bien distincts, une énergie cinétique turbulente notée $\rho k^{(m)}$, une viscosité turbulente notée $\mu_\tau^{(m)}$, et un

taux de dissipation turbulente noté $\rho F^{(m)}$ leur sont associés pour tout $m = 1, ..., M$. Par convention, nous posons $F^{(0)} = 0$ dans (4.1). Soulignons que chacune des trois quantités $\rho k^{(m)}$, $\mu_\tau^{(m)}$ et $\rho F^{(m)}$ doit être positive ou nulle.
Notons que du point de vue de la modélisation, les énergies cinétiques turbulentes $\rho k^{(m)}$ et les viscosités turbulentes $\mu_\tau^{(m)}$ s'ajoutent, dans les lois de conservation de la quantité de mouvement et de l'énergie totale, respectivement à la pression laminaire p et à la viscosité laminaire μ du fluide.
Les équations considérées dans ce chapitre seront toujours supposées être sous forme adimensionnée. De ce fait, les notations

$$\mu, \{\mu_\tau^{(m)}\}_{m=1,...,M}, \kappa, \{\kappa_m\}_{m=1,...,M}, \{\tilde{\kappa}_m\}_{m=1,...,M},$$

représentent respectivement des nombres de Reynolds et des nombres de Prandtl, que nous continuerons cependant par abus de langage d'appeler viscosités et conductivités thermiques. De même, nous appellerons communément les inconnues ρ, u, p, ... par leur dénomination dimensionnée, à savoir densité, vitesse, pression, ...
Les viscosités turbulentes $\{\mu_\tau^{(m)}\}_{m=1,...,M}$ sont habituellement définies par

$$\mu_\tau^{(m)} = C_\mu^{(m)} \rho \frac{k^{(m)^2}}{F^{(m)}},$$

où les $\{C_\mu^{(m)}\}_{m=1,...,M}$ sont des constantes positives du modèle. Notons que ces viscosités turbulentes partielles sont positives sous la condition $\rho F^{(m)} \geq 0$. Les conductivités thermiques sont généralement modélisées par

$$\kappa = \frac{\mu}{Pr} + \sum_{m=1}^{M} \frac{\mu_\tau^{(m)}}{Pr_\tau^{(m)}}, \quad m = 1, ..., M, \tag{4.2}$$

$$\kappa_m = \mu + \frac{\mu_\tau^{(m)}}{\alpha_m}, \quad m = 1, ..., M, \tag{4.3}$$

$$\tilde{\kappa}_m = C_0^{(m)} (\gamma_\tau - 1)(\mu + \frac{\mu_\tau^{(m)}}{\alpha_\varepsilon^{(m)}}), \quad m = 1, ..., M, \tag{4.4}$$

où Pr et $Pr_\tau^{(m)}, \alpha_m, \alpha_\varepsilon^{(m)}$ pour tout $m = 1, ..., M$, sont également des constantes positives du modèle à préciser (voir Schiestel [116], Gleize [54] pour leurs choix). La viscosité laminaire μ est choisie constante.
Pour conclure la description du système, il reste à définir le terme de production $P_k^{(m)}$ intervenant dans (4.1) qui trouve classiquement la forme :

$$P_k^{(m)} = -(\gamma_\tau - 1)\rho k^{(m)} \partial_x u + \mu_\tau^{(m)} (\partial_x u)^2, \tag{4.5}$$

et à relier la pression laminaire p aux inconnues du système par la loi de fermeture sur l'énergie totale ρE :

$$\rho E = \frac{(\rho u)^2}{2\rho} + \frac{p}{\gamma - 1} + \sum_{m=1}^{M} \rho k^{(m)}. \tag{4.6}$$

γ désigne une constante strictement supérieure à 1 et représente le coefficient adiabatique associé à p. De même, le coefficient adiabatique γ_τ associé aux énergies cinétiques turbulentes $\rho k^{(m)}$ est strictement supérieur à 1. Il trouve la définition suivante :

$$\gamma_\tau = 1 + \frac{2}{d}, \quad d \geq 1,$$

où d représente la dimension d'espace du système initial dont dérive la formulation "quasi-1D" (4.1).
Remarquons enfin que lorsque l'on fait le choix $M = 1$ (voir Berthon [10]), le modèle (4.1) dégénère trivialement sur le modèle $k - \varepsilon$ classique, F jouant le rôle de ε.

Notre propos consiste maintenant à réinterpréter les modèles multi-échelle considérés dans le formalisme des équations de Navier-Stokes à plusieurs pressions indépendantes, complétées par des lois d'évolution à préciser. Cette analogie nous permettra de mettre à profit les études menées dans la première partie de ce manuscrit, et d'en proposer une approximation numérique consistante (voir section 4.3).

Les équations de conservation de la quantité de mouvement et de l'énergie totale présentes dans (4.1) suggèrent l'introduction des notations

$$p_\tau^{(m)} = (\gamma_\tau - 1)\rho k^{(m)}, \quad m = 1, ..., M, \tag{4.7}$$

permettant de définir la quantité $p + (\gamma_\tau - 1)\sum_{m=1}^{M} \rho k^{(m)}$ comme une somme de pressions partielles : une pression laminaire p et des pressions turbulentes $p_\tau^{(m)}$. Une analyse dimensionnelle nous indique que les quantités $\{(\gamma_\tau - 1)\rho k^{(m)}\}_{m=1,...,M}$ possèdent bien la dimension d'une pression. Précisons que les quantités $\sqrt{k^{(m)}}, k^{(m)}/F^{(m)}, {k^{(m)}}^{3/2}/F^{(m)}$ sont, pour tout $m = 1, ..., M$, respectivement homogènes à des vitesses, des temps et des longueurs caractéristiques de la turbulence. Ainsi, le système (4.1) se réécrit de manière équivalente sous la forme :

$$\left\{\begin{array}{l}
\partial_t\rho + \partial_x\rho u = 0,\\[2ex]
\partial_t\rho u + \partial_x(\rho u^2 + p + \sum_{m=1}^{M} p_\tau^{(m)}) = \partial_x((\mu + \sum_{m=1}^{M} \mu_\tau^{(m)})\partial_x u),\\[2ex]
\partial_t\rho E + \partial_x(\rho E + p + \sum_{m=1}^{M} p_\tau^{(m)})u = \\
\partial_x((\mu + \sum_{m=1}^{M} \mu_\tau^{(m)})u\partial_x u) + \partial_x(\kappa\partial_x T) + \sum_{m=1}^{M} \partial_x(\kappa_m\partial_x k^{(m)}),\\[2ex]
\partial_t p_\tau^{(m)} + \partial_x p_\tau^{(m)} u + (\gamma_\tau - 1)p_\tau^{(m)}\partial_x u = \\
(\gamma_\tau - 1)(\mu_\tau^{(m)}(\partial_x u)^2 + \rho(F^{(m-1)} - F^{(m)}) + \partial_x(\kappa_m\partial_x k^{(m)}),\\[2ex]
\partial_t\rho F^{(m)} + \partial_x\rho F^{(m)} u = \\
C_0^{(m)}\frac{F^{(m)}}{k^{(m)}}P_k^{(m)} + \rho\frac{F^{(m)}}{k^{(m)}}(C_1^{(m)}F^{(m-1)} - C_2^{(m)}F^{(m)}) + \partial_x(\tilde{\kappa}_m\partial_x F^{(m)}),\\
\qquad\qquad\qquad m = 1, ..., M.
\end{array}\right. \tag{4.8}$$

Cette formulation assimile les lois d'évolution des quantités $\{\rho k^{(m)}\}_{m=1,...,M}$ à de véritables lois de pression. Pour aller plus loin dans cette direction, nous indiquons maintenant sans démonstration (voir Berthon [10] dans le cas $M = 1$) que la définition (4.5) et la loi de fermeture sur l'énergie totale (4.6) définissant la pression laminaire p, entraînent la validité, pour les solutions régulières du système (4.8), de la loi d'évolution suivante sur p :

$$\partial_t p + \partial_x pu + (\gamma - 1)p\partial_x u = (\gamma - 1)(\mu(\partial_x u)^2 + \rho F^{(M)} + \partial_x(\kappa\partial_x T)).$$

En remplaçant dans (4.8) la loi de conservation sur l'énergie totale ρE par la loi de pression sur p, nous obtenons la formulation équivalente suivante du modèle de turbulence compressible multi-échelle (4.1) :

$$\begin{cases} \partial_t\rho + \partial_x\rho u = 0, \\ \partial_t\rho u + \partial_x(\rho u^2 + p + \sum_{m=1}^M p_\tau^{(m)}) = \partial_x((\mu + \sum_{m=1}^M \mu_\tau^{(m)})\partial_x u), \\ \partial_t p + \partial_x pu + (\gamma - 1)p\partial_x u = (\gamma - 1)(\mu(\partial_x u)^2 + \rho F^{(M)} + \partial_x(\kappa\partial_x T)), \\ \partial_t p_\tau^{(m)} + \partial_x p_\tau^{(m)} u + (\gamma_\tau - 1)p_\tau^{(m)}\partial_x u = \\ (\gamma_\tau - 1)(\mu_\tau^{(m)}(\partial_x u)^2 + \rho(F^{(m-1)} - F^{(m)}) + \partial_x(\kappa_m\partial_x k^{(m)}), \\ \partial_t\rho F^{(m)} + \partial_x\rho F^{(m)} u = \\ C_0^{(m)}\frac{F^{(m)}}{k^{(m)}}P_k^{(m)} + \rho\frac{F^{(m)}}{k^{(m)}}(C_1^{(m)}F^{(m-1)} - C_2^{(m)}F^{(m)}) + \partial_x(\tilde{\kappa}_m\partial_x F^{(m)}), \\ \qquad m = 1, ..., M. \end{cases} \tag{4.9}$$

Elle implique les deux lois de conservation usuelles sur la masse et la quantité de mouvement ainsi que $(M+1)$ lois de pression indépendantes. Ainsi, nous réinterprétons le modèle (4.1) comme le système des équations de Navier-Stokes à plusieurs pressions indépendantes (aux termes sources près), complété par des lois d'évolution sur les grandeurs turbulentes $\rho F^{(m)}$. Dans le contexte de l'approximation numérique des solutions du système (4.1), il apparaît dès lors qu'une stratégie de splitting d'opérateurs entre la partie convective-diffusive et les autres opérateurs du système (4.9) nous permettra de tirer profit des études réalisées dans les chapitres précédents. Nous détaillerons ces différents points dans la section suivante. Nous renvoyons à [10] pour l'exposé d'une stratégie voisine mais dédiée au cas $M = 1$.

Nous terminons cette section en donnant deux nouvelles formulations équivalentes du système (4.9) pour les solutions régulières.

La première réécriture est basée sur une transformation des lois gouvernant l'évolution des quantités $\{\rho F^{(m)}\}_{m=1,...,M}$ en de véritables lois de transport (en l'absence de conductivité thermique) sur des grandeurs caractéristiques de la turbulence. Plus précisément, en introduisant les quantités $\{S_m^X\}_{m=1,...,M}$ définies par

$$S_m^X = \ln(\frac{F^{(m)}}{k^{(m)C_0^{(m)}}}), \quad m = 1, ..., M, \tag{4.10}$$

nous établissons facilement l'équivalence, pour les solutions régulières, du système

(4.9) avec le système suivant :

$$\left\{\begin{array}{l}
\partial_t\rho + \partial_x\rho u = 0, \\
\\
\partial_t\rho u + \partial_x(\rho u^2 + p + \sum_{m=1}^M p_\tau^{(m)}) = \partial_x((\mu + \sum_{m=1}^M \mu_\tau^{(m)})\partial_x u), \\
\\
\partial_t p + \partial_x p u + (\gamma - 1)p\partial_x u = (\gamma - 1)(\mu(\partial_x u)^2 + \rho F^{(M)} + \partial_x(\kappa\partial_x T)), \\
\\
\partial_t p_\tau^{(m)} + \partial_x p_\tau^{(m)} u + (\gamma_\tau - 1)p_\tau^{(m)}\partial_x u = \\
(\gamma_\tau - 1)(\mu_\tau^{(m)}(\partial_x u)^2 + \rho(F^{(m-1)} - F^{(m)}) + \partial_x(\kappa_m\partial_x k^{(m)})), \\
\\
\partial_t\rho S_m^X + \partial_x\rho S_m^X u = \\
\rho\frac{F^{(m-1)}}{k^{(m)}}(C_1^{(m)} - C_0^{(m)}) - \rho\frac{F^{(m)}}{k^{(m)}}(C_2^{(m)} - C_0^{(m)}) + \\
\frac{1}{F^{(m)}}\partial_x(\tilde{\kappa}_m\partial_x F^{(m)}) - \frac{C_0^{(m)}}{k^{(m)}}\partial_x(\kappa_m\partial_x k^{(m)}), \\
\qquad\qquad m = 1, ..., M.
\end{array}\right. \tag{4.11}$$

Remarquons que lorsque $C_0^{(m)} = \frac{3}{2}$, la quantité S_m^X représente une longueur caractéristique de la turbulence.

La deuxième formulation consiste à introduire les entropies spécifiques s et $\{s^{(m)}\}_{m=1,...,M}$ définies par :

$$s = \frac{p}{\rho^\gamma}, \quad s^{(m)} = \frac{p_\tau^{(m)}}{\rho^{\gamma_\tau}}, \quad m = 1, ..., M, \tag{4.12}$$

de telle sorte que les lois de pression du système (4.11) se réécrivent pour les solutions régulières, et en vertu de la conservation de la masse, sous la forme :

$$\partial_t s + u\partial_x s = \frac{\gamma - 1}{\rho^\gamma}(\mu(\partial_x u)^2 + \rho F^{(M)} + \partial_x(\kappa\partial_x T)),$$

$$\partial_t s^{(m)} + u\partial_x s^{(m)} = \frac{\gamma_\tau - 1}{\rho^{\gamma_\tau}}(\mu_\tau^{(m)}(\partial_x u)^2 + \rho(F^{(m-1)} - F^{(m)}) + \partial_x(\kappa_m\partial_x k^{(m)})),$$
$$m = 1, ..., M,$$

ou encore,

$$\partial_t\rho s + \partial_x\rho s u = \frac{\gamma - 1}{\rho^{\gamma - 1}}(\mu(\partial_x u)^2 + \rho F^{(M)} + \partial_x(\kappa\partial_x T)), \tag{4.13}$$

$$\partial_t\rho s^{(m)} + \partial_x\rho s^{(m)} u = \frac{\gamma_\tau - 1}{\rho^{\gamma_\tau - 1}}(\mu_\tau^{(m)}(\partial_x u)^2 + \rho(F^{(m-1)} - F^{(m)}) + \partial_x(\kappa_m\partial_x k^{(m)})), \tag{4.14}$$
$$m = 1, ..., M.$$

De manière plus générale, nous établissons le lemme suivant :

Lemme 5
Soient $h : I\!R \to I\!R$ et $\{h_m : I\!R \to I\!R\}_{m=1,...,M}$, $(M + 1)$ applications supposées régulières.
Alors, les solutions régulières du système (4.11) vérifient les lois d'évolution suivantes :

$$\partial_t\rho h(s) + \partial_x\rho h(s) u = \frac{\gamma - 1}{\rho^{\gamma - 1}}h'(s)(\mu(\partial_x u)^2 + \rho F^{(M)} + \partial_x(\kappa\partial_x T)), \tag{4.15}$$

et pour tout $m = 1, ..., M$ *:*

$$\partial_t \rho h_m(s^{(m)}) + \partial_x \rho h_m(s^{(m)}) u =$$

$$\frac{\gamma_\tau - 1}{\rho^{\gamma_\tau - 1}} h'_m(s^{(m)})(\mu_\tau^{(m)}(\partial_x u)^2 + \rho(F^{(m-1)} - F^{(m)}) + \partial_x(\kappa_m \partial_x k^{(m)})). \quad (4.16)$$

Nous obtenons alors sans aucune difficulté la formulation équivalente suivante du système (4.11) :

$$\left\{\begin{array}{l} \partial_t \rho + \partial_x \rho u = 0, \\ \\ \partial_t \rho u + \partial_x(\rho u^2 + p + \sum_{m=1}^M p_\tau^{(m)}) = \partial_x((\mu + \sum_{m=1}^M \mu_\tau^{(m)})\partial_x u), \\ \\ \partial_t \rho E + \partial_x(\rho E + p + \sum_{m=1}^M p_\tau^{(m)}) u = \\ \partial_x((\mu + \sum_{m=1}^M \mu_\tau^{(m)}) u \partial_x u) + \partial_x(\kappa \partial_x T) + \sum_{m=1}^M \partial_x(\kappa_m \partial_x k^{(m)}), \\ \\ \partial_t \rho S^{(m)} + \partial_x \rho S^{(m)} u = \\ \frac{\gamma_\tau - 1}{\rho^{\gamma_\tau - 1}} h'_m(s^{(m)})(\mu_\tau^{(m)}(\partial_x u)^2 + \rho(F^{(m-1)} - F^{(m)}) + \partial_x(\kappa_m \partial_x k^{(m)})), \\ \\ \partial_t \rho S_m^X + \partial_x \rho S_m^X u = \\ \rho \frac{F^{(m-1)}}{k^{(m)}}(C_1^{(m)} - C_0^{(m)}) - \rho \frac{F^{(m)}}{k^{(m)}}(C_2^{(m)} - C_0^{(m)}) + \\ \frac{1}{F^{(m)}} \partial_x(\tilde{\kappa}_m \partial_x F^{(m)}) - \frac{C_0^{(m)}}{k^{(m)}} \partial_x(\kappa_m \partial_x k^{(m)}), \\ \qquad\qquad\qquad m = 1, ..., M, \end{array}\right. \quad (4.17)$$

avec par définition :

$$S^{(m)} = h_m(s^{(m)}), \quad m = 1, ..., M.$$

Bien entendu, les solutions régulières de ce système vérifient sans condition supplémentaire la loi d'évolution suivante sur l'entropie spécifique $S = h(s)$, qui doit être comprise comme une fonction du vecteur $\mathbf{v^S}$:

$$\mathbf{v^S} = {}^T(\rho, \rho u, \rho E, \{\rho S^{(m)}\}_{m=1,...,M}, \{\rho S_m^X\}_{m=1,...,M})$$

des inconnues du système (4.17) :

$$\partial_t \rho S(\mathbf{v^S}) + \partial_x \rho S(\mathbf{v^S}) u = \frac{\gamma - 1}{\rho^{\gamma - 1}} h'(h^{-1}(S(\mathbf{v^S})))(\mu(\partial_x u)^2 + \rho F^{(M)} + \partial_x(\kappa \partial_x T)). \quad (4.18)$$

Sous des hypothèses *adhoc* de type convexité sur les applications $\{h_m\}_{m=1,...,M}$ et h (voir [10] pour plus détails), on montre que l'application $\mathbf{v^S} \to \rho S(\mathbf{v^S})$ est strictement convexe, de sorte que le couple $(\rho S(\mathbf{v^S}), \rho S(\mathbf{v^S}) u)$ joue le rôle d'un couple d'entropie de Lax.

4.3 Approximation numérique

L'objectif de cette section est de proposer une méthode numérique consistante pour approcher numériquement les solutions du modèle de turbulence compressible multi-échelle (4.17), pour une donnée initiale $\mathbf{v_0^S}$. Afin d'éviter des restrictions CFL contraignantes, nous proposons un traitement numérique implicite de ce système. Nous aurons pour souci de préserver au niveau discret la positivité éventuelle des différentes quantités intervenant dans le modèle.

Supposons donnés un pas de temps constant Δt et un pas d'espace constant Δx, et définissons la quantité $\lambda = \Delta t/\Delta x$. Nous considérons les solutions approchées $\mathbf{v}_\lambda^{\mathbf{S}}(x,t)$ du système (4.17) constantes par morceaux sur chaque pavé $\mathcal{C}_j^n := [x_{j-1/2}, x_{j+1/2}[\times[t^n, t^{n+1}[$, avec pour tout $j \in Z$, $x_{j+1/2} = (j+1/2)\Delta x$ et pour tout $n \in N$, $t^n = n\Delta t$. Nous posons alors :

$$\mathbf{v}_\lambda^{\mathbf{S}}(x,t^n) = (\mathbf{v}^{\mathbf{S}})_j^n = {}^T(\rho, \rho u, \rho E, \{\rho S^{(m)}\}_{m=1,\dots,M}, \{\rho S_m^X\}_{m=1,\dots,M})_j^n,$$

$$\text{si } (x,t) \in \mathcal{C}_j^n, \quad j \in Z, \quad n \in N,$$

et par définition :

$$(\mathbf{v}^{\mathbf{S}})_j^0 = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{v}_\mathbf{0}^{\mathbf{S}}(x)dx, \quad j \in Z.$$

Supposons connue la solution approchée $\mathbf{v}_\lambda^{\mathbf{S}}(x,t^n)$, $n \geq 0$. Pour avancer en temps cette solution jusqu'à l'instant t^{n+1}, nous adoptons une stratégie de splitting entre les différents opérateurs du système (4.17). La procédure correspondante est composée de trois sous-pas.

Premier pas : opérateur convectif-diffusif ($t^n \to t^{n+1=}$)
Ce premier pas consiste en la résolution de l'opérateur convectif-diffusif extrait du système (4.17) :

$$\begin{cases} \partial_t\rho + \partial_x\rho u = 0, \\ \partial_t\rho u + \partial_x(\rho u^2 + p + \sum_{m=1}^M p_\tau^{(m)}) = \partial_x((\mu + \sum_{m=1}^M \mu_\tau^{(m)})\partial_x u), \\ \partial_t\rho E + \partial_x(\rho E + p + \sum_{m=1}^M p_\tau^{(m)})u = \partial_x((\mu + \sum_{m=1}^M \mu_\tau^{(m)})u\partial_x u), \\ \partial_t\rho S^{(m)} + \partial_x\rho S^{(m)}u = \frac{\gamma_\tau - 1}{\rho^{\gamma_\tau - 1}} h_m'(s^{(m)})\mu_\tau^{(m)}(\partial_x u)^2, \quad m = 1,\dots,M, \\ \partial_t\rho S_m^X + \partial_x\rho S_m^X u = 0, \quad m = 1,\dots,M. \end{cases} \tag{4.19}$$

En invoquant l'équivalence pour les solutions régulières des systèmes (4.17) et (4.11), nous transformons de manière équivalente les équations de (4.19) en considérant le système :

$$\begin{cases} \partial_t\rho + \partial_x\rho u = 0, \\ \partial_t\rho u + \partial_x(\rho u^2 + p + \sum_{m=1}^M p_\tau^{(m)}) = \partial_x((\mu + \sum_{m=1}^M \mu_\tau^{(m)})\partial_x u), \\ \partial_t p + \partial_x p u + (\gamma - 1)p\partial_x u = (\gamma - 1)\mu(\partial_x u)^2, \\ \partial_t p_\tau^{(m)} + \partial_x p_\tau^{(m)} u + (\gamma_\tau - 1)p_\tau^{(m)}\partial_x u = (\gamma_\tau - 1)\mu_\tau^{(m)}(\partial_x u)^2, \quad m = 1,\dots,M, \end{cases} \tag{4.20}$$

complété par les M lois de transport suivantes sur les grandeurs turbulentes S_m^X :

$$\left\{\ \partial_t \rho S_m^X + \partial_x \rho S_m^X u = 0, \quad m = 1, ..., M. \right. \tag{4.21}$$

Les équations (4.21) sont naturellement découplées de celles du système (4.20). Soulignons également que le système (4.20) représente le système des équations de Navier-Stokes à $(M+1)$ pressions indépendantes. Ce système ayant été très largement considéré dans les chapitres précédents, nous tirons profit de ces études en proposant pour ce premier pas une mise à jour implicite des inconnues du système (4.20) coïncidant avec l'approximation, présentée dans le chapitre 3, du système des équations de Navier-Stokes multi-pression. Nous rappelons que cette mise à jour assure la positivité des pressions $p_j^{n+1=}$ et $\{(p_\tau^{(m)})_j^{n+1=}\}_{m=1,...,M}$ (à la condition que la densité et l'énergie interne totale restent positives, condition indispensable à la poursuite des calculs) à l'issue de ce premier pas, et donc des énergies cinétiques turbulentes $\{(\rho k^{(m)})_j^{n+1=}\}_{m=1,...,M}$.
Concernant la mise à jour des grandeurs turbulentes $\{(S_m^X)_j^{n+1=}\}_{m=1,...,M}$, nous proposons la discrétisation classique et implicite linéarisée suivante des équations (4.21) :

$$(\rho S_m^X)_j^{n+1=} - (\rho S_m^X)_j^n + \frac{\Delta t}{\Delta x} \Delta \{\rho S_m^X u\}_{j+1/2}^{n+1=} = 0, \quad m = 1, ..., M, \tag{4.22}$$

où les flux numériques $\{\rho S_m^X u\}_{j+1/2}^{n+1=}$ trouvent la définition suivante :

$$\{\rho S_m^X u\}_{j+1/2}^{n+1=} = \{\rho u\}_{j+1/2}^n \times \left\{ \begin{array}{ll} (S_m^X)_j^{n+1=} & \text{si} \quad \{\rho u\}_{j+1/2}^n > 0, \\ (S_m^X)_{j+1}^{n+1=} & \text{sinon.} \end{array} \right. \tag{4.23}$$

Les flux numériques $\{\rho u\}_{j+1/2}^n$ sont associés à la méthode conservative utilisée pour discrétiser l'équation de conservation de la masse dans (4.20).
D'après (4.22)-(4.23), la mise à jour implicite des grandeurs $\{(S_m^X)_j^{n+1=}\}_{m=1,...,M}$ au cours de ce premier pas consiste en la résolution de M systèmes linéaires dont la matrice globale en j (la même pour tous) possède une structure tridiagonale scalaire. Cette matrice est en fait une M-matrice à diagonale dominante et à coefficients extra-diagonaux négatifs. Elle est donc inversible sous des conditions aux limites *adhoc*, ce qui garantit l'existence et l'unicité des mises à jour implicites ainsi définies.
La positivité des taux de dissipation $\{(F^{(m)})_j^{n+1=}\}_{m=1,...,M}$ à l'issue de ce premier pas est garantie d'après (4.10) puisque les énergies cinétiques turbulentes obtenues sont positives.

Deuxième pas : prise en compte des termes sources $(t^{n+1=} \to t^{n+1-})$
De manière à poursuivre l'évolution en temps de la solution discrète obtenue à l'issue du pas précédent, nous résolvons le problème d'équations différentielles ordinaires suivant, prenant en compte les termes sources du système (4.17) :

$$\left\{ \begin{array}{l} \partial_t \rho = 0, \\ \partial_t \rho u = 0, \\ \partial_t \rho E = 0, \\ \partial_t \rho S^{(m)} = \frac{\gamma_\tau - 1}{\rho^{\gamma_\tau - 1}} h'_m(s^{(m)}) \rho (F^{(m-1)} - F^{(m)}), \quad m = 1, ..., M, \\ \partial_t \rho S_m^X = \rho \frac{F^{(m-1)}}{k^{(m)}} (C_1^{(m)} - C_0^{(m)}) - \rho \frac{F^{(m)}}{k^{(m)}} (C_2^{(m)} - C_0^{(m)}), \quad m = 1, ..., M. \end{array} \right. \tag{4.24}$$

La donnée initiale considérée coïncide avec la solution discrète $\mathbf{v}_\lambda^{\mathbf{S}}(x, t^{n+1=})$ obtenue à l'issue du premier pas.
En invoquant l'équivalence des systèmes (4.17) et (4.1) pour les solutions régulières, la prise en compte des termes sources du modèle multi-échelle revient à résoudre le système d'équations différentielles ordinaires suivant :

$$\begin{cases} \partial_t \rho = 0, \\ \partial_t \rho u = 0, \\ \partial_t \rho E = 0, \\ \partial_t k^{(m)} = F^{(m-1)} - F^{(m)}, \quad m = 1, ..., M, \\ \partial_t F^{(m)} = C_1^{(m)} F^{(m-1)} \frac{F^{(m)}}{k^{(m)}} - C_2^{(m)} \frac{(F^{(m)})^2}{k^{(m)}}, \quad m = 1, ..., M. \end{cases} \tag{4.25}$$

Nous réécrivons de manière équivalente les $2M$ équations non triviales de (4.25) sous la forme :

$$\partial_t k^{(m)} = \lambda_{m-1} k^{(m-1)} - \lambda_m k^{(m)}, \quad m = 1, ..., M, \tag{4.26}$$
$$\partial_t F^{(m)} = \lambda_m^{(1)} F^{(m-1)} - \lambda_m^{(2)} F^{(m)}, \quad m = 1, ..., M, \tag{4.27}$$

avec

$$\lambda_m = \frac{F^{(m)}}{k^{(m)}}, \quad \lambda_m^{(1)} = C_1^{(m)} \frac{F^{(m)}}{k^{(m)}}, \quad \lambda_m^{(2)} = C_2^{(m)} \frac{F^{(m)}}{k^{(m)}}, \quad m = 1, ..., M.$$

Soulignons que l'évolution en temps des énergies cinétiques turbulentes $k^{(m)}$ impose au travers de la conservation de l'énergie totale une évolution en temps de l'énergie interne thermodynamique $p/(\gamma - 1)$. Le système d'EDO (4.25) montre que cette évolution est donnée par

$$\partial_t p = (\gamma - 1) \rho F^{(M)} \geq 0.$$

De manière à préserver la positivité de p, nous proposons un traitement implicite linéarisé des équations considérées après avoir posé, de manière consistante avec (4.25) :

$$\rho_j^{n+1-} = \rho_j^{n+1=}, \quad (\rho u)_j^{n+1-} = (\rho E)_j^{n+1=}, \quad (\rho E)_j^{n+1-} = (\rho E)_j^{n+1=}, \quad j \in Z.$$

L'implicitation proposée fige les échelles de temps $\lambda_m, \lambda_m^{(1)}$, et $\lambda_m^{(2)}$ à la date $t^{n+1=}$. Ainsi, le système d'équations différentielles ordinaires à résoudre peut être génériquement réexprimé par :

$$\partial_t X_m = a_{m-1} X_{m-1} - a_m X_m, \quad m = 1, ..., M, \tag{4.28}$$
$$\partial_t Y_m = a_m^{(1)} Y_{m-1} - a_m^{(2)} Y_m, \quad m = 1, ..., M, \tag{4.29}$$

où les constantes $a_{m-1}, a_m^{(1)}$, et $a_m^{(2)}$ sont supposées connues pour tout $m = 1, ..., M$. Notons que ces constantes sont positives dès que les données initiales $k^{(m)}$ et $F^{(m)}$ le sont. X_m et Y_m représentent respectivement les énergies cinétiques turbulentes $k^{(m)}$ et les taux de dissipation turbulente $F^{(m)}$. Nous obtenons le résultat suivant :

Lemme 6 *Supposons $X_M(t) \geq 0$ pour tous les temps t considérés. Alors, la pression $p(t)$ reste positive pour ces mêmes temps.*

Démonstration
Tenant compte de la constance en temps de ρ et de ρu, la conservation de l'énergie totale

$$\rho E = \frac{(\rho u)^2}{2\rho} + \sum_{m=1}^{M} \rho X_m + \frac{p}{\gamma - 1},$$

conduit à l'EDO suivante :

$$\partial_t p = -(\gamma - 1)\rho \partial_t (\sum_{m=1}^{M} X_m).$$

En vertu de l'implicitation proposée, les X_m sont solutions exactes de (4.28) de sorte que :

$$\partial_t (\sum_{m=1}^{M} X_m) = a_0 X_0 - a_M X_M,$$

où par construction $a_0 = 0$ puisque $\lambda_0 = 0$ (rappelons que par convention $F^{(0)} = 0$). Nous en déduisons donc que $p(t)$ est solution exacte de l'EDO suivante :

$$\partial_t p = (\gamma - 1)\rho a_M X_M.$$

La conclusion suit du fait que l'échelle de temps $a_M = \lambda_M$ a été figée à l'instant $t^{n+1=}$ sur une valeur positive, la densité l'étant par construction car constante et X_M l'étant par hypothèse.

Nous avons ensuite :

Théorème 10
Soient $a_m, 1 \leq m \leq M$, M constantes positives supposées connues.
Le système d'EDO

$$\begin{array}{rcl} \partial_t X_m &=& a_{m-1} X_{m-1} - a_m X_m, \quad 1 \leq m \leq M, \quad a_0 = 0, \\ X_m(t=0) &=& X_m(0) \geq 0, \end{array} \tag{4.30}$$

s'intègre explicitement. Les solutions $X_m(t)$ sont positives pour tout temps t et vérifient la relation de récurrence

$$X_m(t) = \sum_{l=1}^{m} \alpha_l^m \exp(-a_l t), \quad 1 \leq m \leq M,$$

avec, pour $2 \leq m \leq M$,

$$\alpha_l^m = \begin{cases} \frac{a_{m-1}}{a_m - a_l} \alpha_l^{m-1} & \textit{pour} \quad 1 \leq l \leq m-1, \\ X_m(0) - \sum_{l=1}^{m-1} \alpha_l^m & \textit{pour} \quad l = m, \end{cases}$$

et pour $m = 1$,

$$\alpha_1^1 = X_1(0).$$

En particulier, leur évaluation ne nécessite, à chaque instant, que le calcul de M exponentielles.

Démonstration
L'équation du système (4.30) pour $m = 1$ s'écrit

$$\begin{array}{rcl} \partial_t X_1 & = & -a_1 X_1, \\ X_1(t=0) & = & X_1(0) \geq 0. \end{array}$$

Par intégration directe, nous obtenons

$$X_1(t) = \alpha_1^1 \exp(-a_1 t), \quad \alpha_1^1 = X_1(0) \geq 0.$$

En particulier X_1 est à valeurs positives pour tout temps $t \geq 0$.
Supposons que

$$X_{m-1}(t) \geq 0, \quad \forall \quad t \geq 0. \tag{4.31}$$

La multiplication de l'équation générale du système (4.30) par $\exp(a_m t)$ implique

$$\partial_t\{X_m \exp(a_m t)\} = a_{m-1} X_{m-1} \exp(a_m t).$$

Après intégration, il vient

$$X_m(t) = X_m(0)\exp(-a_m t) + a_{m-1}\exp(-a_m t)\int_0^t X_{m-1}(s)\exp(a_m s)ds.$$

Ainsi, puisque toutes les données initiales sont supposées positives, l'hypothèse de récurrence (4.31) implique $X_m(t) \geq 0, \quad \forall \quad t \geq 0$.
Ceci conclut la démonstration du théorème.

De la même manière, nous obtenons le résultat suivant.
Théorème 11
Soient $a_m^{(1)}, a_m^{(2)}$, $1 \leq m \leq M$, $2M$ constantes positives (avec $a_1^{(1)} = 0$) supposées connues.
Le système d'EDO

$$\begin{array}{rcl} \partial_t Y_m & = & a_m^{(1)} Y_{m-1} - a_m^{(2)} Y_m, \quad 1 \leq m \leq M, \\ Y_m(t=0) & = & Y_m(0) \geq 0, \end{array} \tag{4.32}$$

s'intègre explicitement. Les solutions $Y_m(t)$ sont positives pour tout temps t et vérifient la relation de récurrence

$$Y_m(t) = \sum_{l=1}^{m} \beta_l^m \exp(-a_l^{(2)} t), \quad 1 \leq m \leq M,$$

avec, pour $2 \leq m \leq M$,

$$\beta_l^m = \begin{cases} \dfrac{a_m^{(1)}}{a_m^{(2)} - a_l^{(2)}} \beta_l^{m-1} & \textit{pour} \quad 1 \leq l \leq m-1, \\ \\ Y_m(0) - \displaystyle\sum_{l=1}^{m-1} \beta_l^m & \textit{pour} \quad l = m, \end{cases}$$

et pour $m = 1$,

$$\alpha_1^1 = Y_1(0).$$

En particulier, leur évaluation ne nécessite, à chaque pas de temps, que le calcul de M exponentielles.

En reprenant les notations de ces deux théorèmes, nous posons tout naturellement et pour tout $j \in Z$:

$$(k^{(m)})_j^{n+1-} = \sum_{l=1}^{m} (\alpha_l^m)_j \exp(-(a_l)_j \Delta t), \quad m = 1, ..., M,$$

$$(F^{(m)})_j^{n+1-} = \sum_{l=1}^{m} (\beta_l^m)_j \exp(-(a_l^{(2)})_j \Delta t), \quad m = 1, ..., M,$$

avec pour tout $j \in Z$:

$$(a_m)_j = \frac{(F^{(m)})_j^{n+1=}}{(k^{(m)})_j^{n+1=}}, \quad m = 1, ..., M, \quad (a_0)_j = 0,$$

$$(a_m^{(1)})_j = C_1^{(m)} \frac{(F^{(m)})_j^{n+1=}}{(k^{(m)})_j^{n+1=}}, \quad (a_m^{(2)})_j = C_2^{(m)} \frac{(F^{(m)})_j^{n+1=}}{(k^{(m)})_j^{n+1=}}, \quad m = 1, ..., M.$$

L'implicitation proposée assure donc la positivité des énergies cinétiques turbulentes et des taux de dissipation turbulente, en vertu des deux précédents théorèmes. L'intégration explicite résultante favorise son implémentation tout en réduisant son coût calcul.

Troisième pas : prise en compte des termes de conductivité thermique $(t^{n+1-} \to t^{n+1})$

Nous achevons la stratégie de splitting d'opérateurs proposée par la prise en compte des termes de conductivité thermique. Il s'agit de résoudre le problème de Cauchy attendu suivant :

$$\left\{ \begin{array}{l} \partial_t \rho = 0, \\ \partial_t \rho u = 0, \\ \partial_t \rho E = \partial_x(\kappa \partial_x T) + \displaystyle\sum_{m=1}^{M} \partial_x(\kappa_m \partial_x k^{(m)}), \\ \partial_t \rho S^{(m)} = \frac{\gamma_\tau - 1}{\rho^{\gamma_\tau - 1}} h'_m(s^{(m)}) \partial_x(\kappa_m \partial_x k^{(m)}), \quad m = 1, ..., M, \\ \partial_t \rho S_m^X = \frac{1}{F^{(m)}} \partial_x(\tilde{\kappa}_m \partial_x F^{(m)}) - \frac{C_0^{(m)}}{k^{(m)}} \partial_x(\kappa_m \partial_x k^{(m)}), \quad m = 1, ..., M, \end{array} \right. \tag{4.33}$$

pour la donnée initiale obtenue en considérant la solution discrète telle que définie à l'issue du second pas. Nous prenons à nouveau avantage de l'invariance en temps de la densité pour considérer le système équivalent suivant :

$$\left\{ \begin{array}{l} \partial_t \rho = 0, \\ \partial_t \rho u = 0, \\ \frac{\rho}{\gamma - 1} \partial_t T = \partial_x(\kappa \partial_x T), \\ \rho \partial_t k^{(m)} = \partial_x(\kappa_m \partial_x k^{(m)}), \quad m = 1, ..., M, \\ \rho \partial_t F^{(m)} = \partial_x(\tilde{\kappa}_m \partial_x F^{(m)}), \quad m = 1, ..., M. \end{array} \right. \tag{4.34}$$

Nous proposons une intégration en temps implicite de ce système, en adoptant la discrétisation suivante :

$$\begin{cases} \rho_j^{n+1} = \rho_j^{n+1-}, \\ (\rho u)_j^{n+1} = (\rho u)_j^{n+1-}, \\ (\frac{\rho}{\gamma-1})_j^{n+1} T_j^{n+1} = (\frac{\rho}{\gamma-1})_j^{n+1-} T_j^{n+1-} + \Delta t(\overline{\partial_x(\kappa\partial_x T)})_j^{n+1}, \\ \rho_j^{n+1}(k^{(m)})_j^{n+1} = \rho_j^{n+1-}(k^{(m)})_j^{n+1-} + \Delta t(\overline{\partial_x(\kappa_m\partial_x k^{(m)})})_j^{n+1}, \\ \rho_j^{n+1}(F^{(m)})_j^{n+1} = \rho_j^{n+1-}(F^{(m)})_j^{n+1-} + \Delta t(\overline{\partial_x(\tilde\kappa_m\partial_x F^{(m)})})_j^{n+1}, \end{cases} \tag{4.35}$$

où les opérateurs aux différences finies trouvent les expressions consistantes suivantes (voir [10] dans le cas $M = 1$):

$$\begin{cases} (\overline{\partial_x(\kappa\partial_x T)})_j^{n+1} = \frac{1}{\Delta x^2}\times \\ (\kappa_{j+1/2}^{n+1-}(A^nT_{j+1} - A^nT_j) - \kappa_{j-1/2}^{n+1-}(A^nT_j - A^nT_{j-1})), \\ (\overline{\partial_x(\kappa_m\partial_x k^{(m)})})_j^{n+1} = \frac{1}{\Delta x^2}\times \\ ((\kappa_m)_{j+1/2}^{n+1-}(A^n(k^{(m)})_{j+1} - A^n(k^{(m)})_j) - (\kappa_m)_{j-1/2}^{n+1-}(A^n(k^{(m)})_j - A^n(k^{(m)})_{j-1})), \\ (\overline{\partial_x(\tilde\kappa_m\partial_x F^{(m)})})_j^{n+1} = \frac{1}{\Delta x^2}\times \\ ((\tilde\kappa_m)_{j+1/2}^{n+1-}(A^n(F^{(m)})_{j+1} - A^n(F^{(m)})_j) - (\tilde\kappa_m)_{j-1/2}^{n+1-}(A^n(F^{(m)})_j - A^n(F^{(m)})_{j-1})), \end{cases} \tag{4.36}$$

avec l'opérateur A^n défini par :

$$A^nX = \frac{X^{n+1-} + X^{n+1}}{2}.$$

Dans la pratique, il s'agit de résoudre $(2M+1)$ problèmes symétriques définis positifs et découplés portant respectivement sur les variables T, $k^{(m)}$ et $F^{(m)}$. Les matrices correspondantes sont tridiagonales et de type M-matrice à diagonale dominante (sous des conditions aux limites *adhoc*). Les inverses de ces matrices existent donc, ce qui garantit l'existence et l'unicité des mises à jour implicites. Ces inverses possèdent aussi la propriété de ne posséder que des coefficients tous positifs ou nuls. Nous renvoyons le lecteur à [95] pour plus de détails sur ces notions. On montre comme dans [10] que les mises à jour proposées ci-dessus garantissent la positivité des grandeurs réactualisées T, $k^{(m)}$ et $F^{(m)}$, ce qui constitue une propriété de stabilité appréciable. Nous renvoyons le lecteur à [10] pour plus de détails sur le bien fondé (théorique et numérique) d'une telle implicitation dans le cas $M = 1$ (correspondant au modèle $k - \varepsilon$).

Quatrième partie

Chocs non classiques et transitions de phase

La dernière partie de ce manuscrit est consacrée à l'approximation des solutions de systèmes de lois de conservation de type hyperbolique (mais dont les champs possèdent un défaut de vraiment non linéarité) ou hyperbolique-elliptique. Les principales motivations sont données dans l'introduction générale du mémoire. Les travaux sont regroupés en trois chapitres, chacun contenant un résumé, une introduction et une conclusion permettant de mettre en avant les principaux résultats obtenus.

Chapitre 1

A fully discrete scheme for diffusive-dispersive conservation laws

Ce chapitre a fait l'objet d'une publication dans le journal *Numerische Mathematik*: voir [24]

A fully discrete scheme for diffusive-dispersive conservation laws

C. Chalons [1] and P.G. LeFloch [2]

Résumé

We introduce a fully discrete (in both space and time) scheme for the numerical approximation of diffusive-dispersive hyperbolic conservation laws in one-space dimension. This scheme extends an approach by LeFloch and Rohde [94]: it satisfies a cell entropy inequality and, as a consequence, the space integral of the entropy is a decreasing function of time. This is an important stability property, shared by the continuous model as well. Following Hayes and LeFloch [64], we show that the limiting solutions generated by the scheme need not coincide with the classical Oleinik-Kruzkov entropy solutions, but contain nonclassical undercompressive shock waves. Investigating the properties of the scheme, we stress various similarities and differences between the continuous model and the discrete scheme (dynamics of nonclassical shocks, nucleation, etc).

1.1 Introduction

In this paper, we are interested in the numerical approximation of the *limiting* solutions generated by the diffusive-dispersive conservation law [87]:

$$\partial_t u + \partial_x f(u) = \epsilon\,\beta\, U'(u)_{xx} + \epsilon^2\,\gamma\, U'(u)_{xxx}, \qquad u = u_\epsilon^{\beta,\gamma}(x,t), \quad x \in I\!R,\, t > 0, \tag{1.1}$$

when $\epsilon > 0$ tends to zero. Here the flux $f : I\!R \to I\!R$ is a smooth given function and $\beta, \gamma > 0$ are fixed parameters. In the right-hand side of (1.1), $U : I\!R \to I\!R$ is a given, strictly convex function. The following is known concerning the limiting solutions $\lim_{\epsilon\to 0} u_\epsilon^{\beta,\gamma}$. (We refer to the review [87] and the references cited therein.) First of all, they generally depend on the parameters β and γ. Based on simple scaling arguments, one can see that they only depend on the ratio $\delta = \beta/\gamma$. So we define

$$u^\delta := \lim_{\epsilon\to 0} u_\epsilon^{\beta,\gamma}, \tag{1.2}$$

provided the limit exists in some strong topology. Based on (1.1) and

$$\begin{aligned}\partial_t U(u) + \partial_x F(u) \;&=\; \tfrac{\epsilon\beta}{2}\,(U'(u)^2)_x \\ &-\;\; \epsilon\,\beta\, U'(u)_x^2 + \epsilon^2\gamma\, \left(U'(u)\, U'(u)_{xx} - U'(u)_x^2/2\right)_x,\end{aligned}$$

it is easy to check that the function u^δ satisfies the hyperbolic conservation law

$$\partial_t u + \partial_x f(u) = 0, \tag{1.3}$$

1. O.N.E.R.A., B.P. 72, 29 avenue de la Division Leclerc, 92322 Châtillon Cedex, France. E-mail address: chalons@onera.fr.

2. Centre de Mathématiques Appliquées & Centre National de la Recherche Scientifique, U.M.R. 7641, Ecole Polytechnique, 91128 Palaiseau Cedex, France. E-mail address: lefloch@cmap.polytechnique.fr.

and the entropy inequality

$$\partial_t U(u) + \partial_x F(u) \le 0, \tag{1.4}$$

where U is regarded as an "entropy" of (1.3), and $F : \mathbb{R} \to \mathbb{R}$ is the corresponding entropy flux defined by $F'(u) = U'(u)\, f'(u)$. Note also that from (1.4) it follows that

$$\int_{\mathbb{R}} U(u(x,t))\, dx \le \int_{\mathbb{R}} U(u(x,s))\, dx, \qquad t \ge s. \tag{1.5}$$

In this paper, we propose a fully discrete (in space and time) finite difference scheme for the numerical approximation of the solutions u^δ in (1.2). We rely on the approach developed recently by LeFloch and Rohde [94] and based on Tadmor's notion of entropy conservative flux (see [134]) for the hyperbolic part of (1.1). We will introduce here a fully discrete version of the semi-discrete scheme derived in [94]. The high-order accuracy in time is provided by a standard Runge-Kutta technique. In Section 1.2, we can prove that our fully discrete scheme satisfies a cell entropy inequality, which implies that the entropy is a decreasing function of time. See Theorem 13 and Corollary 8.

In Section 1.3, we investigate the properties of our scheme, especially in terms of stability and nucleation. The stability condition derived in Section 1.2 is numerically investigated. We also demonstrate that the scheme admits a nucleation threshold. Above the threshold value, nonclassical solutions violating the standard entropy criterion are observed. The dependence of the threshold in δ is investigated. We recall that these undercompressive, nonclassical solutions play an important role in many models of continuum mechanics when diffusive and dispersive effects are in balance [87].

1.2 A fully discrete scheme

We start from the semi-discrete method derived by LeFloch and Rohde [94]. Consider the following scheme in conservative form (j describing the integers)

$$\frac{d}{dt} u_j(t) = -\frac{1}{h}\,(g_{j+1/2}(t) - g_{j-1/2}(t)), \qquad t \ge 0, \tag{1.6}$$

with $g_{j+1/2}(t) := g(u_{j-1}(t), u_j(t), u_{j+1}(t), u_{j+2}(t))$, the parameter $h > 0$ being the mesh size. The numerical flux has the form $g := g^1 + g^2 + g^3$, where g^1 is consistent with the hyperbolic flux $f(u)$, g^2 is an approximation of $\beta\, h\, U'(u)_x$, and g^2 of $\gamma\, h^2\, U'(u)_{xx}$. More precisely, using the entropy variable

$$v := U'(u), \quad g(v) := f(u), \quad G(v) := F(u),$$

and therefore $v_0 := U'(u_0)$, etc, we define

$$g^1(u_{-1}, u_0, u_1, , u_2) := \int_0^1 g(v_0 + s(v_1 - v_0))\, ds - \frac{1}{12}(v_2 - v_1 - v_0 + v_{-1}) g'(v_0), \tag{1.7}$$

$$g^2(v_0, v_1) := -\frac{\beta}{2}\,(v_1 - v_0), \tag{1.8}$$

and

$$g^3(v_{-1}, v_0, v_1, v_2) := -\frac{\gamma}{6}\,(v_2 - v_1 - v_0 + v_{-1}). \tag{1.9}$$

Note that we are using the *same* notation for the exact flux $g = g(v)$ expressed in the entropy variable and for the numerical flux.

Theorem 11
The scheme (1.6)-(1.7)-(1.8)-(1.9) is conservatif and consistent with the hyperbolic conservation law (1.3). It also satisfies the cell entropy inequality ($t \geq 0$ and j describing the integers)

$$\frac{d}{dt}U(u_j(t)) + \frac{1}{h}\left(G_{j+1/2}(t) - G_{j-1/2}(t)\right) \leq 0, \tag{1.10}$$

where the numerical entropy flux has the form $G := G^1 + G^2 + G^3$, where

$$\begin{aligned} G^1(v_{-1}, v_0, v_1, v_2) \quad &:= \quad \tfrac{(v_0+v_1)}{2} g(v_{-1}, v_0, v_1, v_2) \\ &- \quad \tfrac{1}{2}\Big(\psi(v_0, v_1, v_2) + \psi(v_{-1}, v_0, v_1)\Big) \end{aligned} \tag{1.11}$$

with

$$\psi(v_0, v_1, v_2) := v_1 g(v_1) - G(v_1) + \frac{1}{12}(v_1 - v_0) g'(v_0)\,(v_1 - v_2),$$

$$G^2(v_0, v_1) := -\frac{\beta}{4}\left(v_1^2 - v_0^2\right), \tag{1.12}$$

and

$$G^3(v_{-1}, v_0, v_1, v_2) = -\frac{\gamma}{6}\Big(v_{-1}v_1 + v_0 v_2 - 2\, v_0 v_1\Big). \tag{1.13}$$

Moreover, the equivalent equation of the scheme (1.6)-(1.7)-(1.8)-(1.9), up to the (third-order) terms in h^2, coincide with the continuous model (1.1) provided ϵ is replaced with h.

Proof
It is straightforward to calculate that

$$\frac{d}{dt}U(u_j(t)) + \frac{1}{h}\Big(G_{j+1/2}(t) - G_{j-1/2}(t)\Big) = -D_j(t), \tag{1.14}$$

where

$$D_j(t) = \frac{\beta}{4}\left(|v_j(t) - v_{j-1}(t)|^2 + |v_{j+1}(t) - v_j(t)|^2\right) \geq 0.$$

The second claim in the theorem follows by performing a Taylor expansion in (1.6) and by using the definitions (1.7)-(1.8)-(1.9). □

We now turn to defining our fully discrete scheme, having in mind to generalize Theorem 11, under some CFL stability restriction on the time discretization. To begin with, we analyze a first order time-discretization. Denote by $k > 0$ the size of the regularly spaced time-mesh. Consider the following scheme ($n \geq 0$ and j describing the integers)

$$u_j^{n+1} = u_j^n - \lambda\,(g_{j+1/2}^n - g_{j-1/2}^n), \tag{1.15}$$

where the ratio $\lambda = k/h$ is kept fixed. In (1.15), we use the notation $g_{j+1/2}^n := g(u_{j-1}^n, u_j^n, u_{j+1}^n, u_{j+2}^n)$, and we define $G_{j+1/2}^n$ similarly, etc.

Theorem 12
The scheme (1.15) satisfies the cell entropy inequality

$$\begin{aligned} & U(u_j^{n+1}) - U(u_j^n) + \lambda\,(G_{j+1/2}^n - G_{j-1/2}^n) + \tfrac{\beta\,\lambda}{4}\,(|v_j^n - v_{j-1}^n|^2 + |v_{j+1}^n - v_j^n|^2) \\ & \leq 3\lambda^2\,\|U''\|_\infty\,(\|g'\|_\infty^2\,D_j^{1,n} + \beta^2\,D_j^{2,n} + \gamma^2\,D_j^{3,n}), \end{aligned} \tag{1.16}$$

where

$$D_j^{1,n} := \frac{1}{36}\,(|v_{j-1}^n - v_{j-2}^n|^2 + 73\,|v_j^n - v_{j-1}^n|^2 + 73\,|v_{j+1}^n - v_j^n|^2 + |v_{j+2}^n - v_{j+1}^n|^2),$$

$$D_j^{2,n} := \frac{1}{4}\,(|v_j^n - v_{j-1}^n|^2 + |v_{j+1}^n - v_j^n|^2),$$

and

$$D_j^{3,n} := \frac{1}{18}\,(|v_{j-1}^n - v_{j-2}^n|^2 + |v_j^n - v_{j-1}^n|^2 + |v_{j+1}^n - v_j^n|^2 + |v_{j+2}^n - v_{j+1}^n|^2).$$

Hence, for λ sufficiently small, the entropy is a strictly decreasing function of time, in the sense

$$\sum_{j=-\infty}^{\infty} U(u_j^{n+1}) + K\,\lambda \sum_{j=-\infty}^{\infty} |v_j^n - v_{j-1}^n|^2 \leq \sum_{j=-\infty}^{\infty} U(u_j^n), \tag{1.17}$$

with

$$K := \frac{\beta}{2} - 3\lambda\,\|U''\|_\infty\,(\frac{37}{9}\,\|g'\|_\infty^2 + \frac{\beta^2}{2} + \frac{2}{9}\,\gamma^2).$$

Furthermore, for smooth solutions, the equivalent equation of the scheme (1.15) is

$$\partial_t u + \partial_x f(u) = h\,\beta\,U'(u)_{xx} + h^2\,\gamma\,U'(u)_{xxx} + O(h^3 + k). \tag{1.18}$$

Observe that the equation (1.18) no longer correspond to the continous model (1.1). Namely the time-discretization generates a term of order $O(k)$ which can not be absorbed in the cubic error $O(h^3)$ (except of course if we were to impose the very drastic restriction $k = O(h^3)$.) Therefore a more accurate time-discretization will be necessary.

Proof
By a Taylor expansion in λ we can deduce from (1.15) that

$$U(u_j^{n+1}) - U(u_j^n) + \lambda\,U'(u_j^n)\,(g_{j+1/2}^n - g_{j-1/2}^n) \leq \frac{\lambda^2}{2}\,\|U''\|_\infty\,|g_{j+1/2}^n - g_{j-1/2}^n|^2. \tag{1.19}$$

We determine the right hand side of (1.19) by treating successively the terms in

$$\begin{aligned} \tfrac{1}{2}\,|g_{j+1/2}^n - g_{j-1/2}^n|^2 \quad & \leq \quad 3(|g_{j+1/2}^n - g(v_j^n)|^2 + |g(v_j^n) - g_{j-1/2}^n|^2 \\ & + \quad |g_{j+1/2}^{2,n}|^2 + |g_{j-1/2}^{2,n}|^2 + |g_{j+1/2}^{3,n}|^2 + |g_{j-1/2}^{3,n}|^2). \end{aligned}$$

In view of (1.7)-(1.8)-(1.9) we find

$$\frac{1}{2}\,|g_{j+1/2}^n - g(v_j^n)|^2 \leq \|Dg\|_\infty^2\,(\frac{1}{72}\,|v_j - v_{j-1}|^2 + |v_{j+1} - v_j|^2 + \frac{1}{72}\,|v_{j+2} - v_{j+1}|^2),$$

and the same inequality remains true by replacing $g(v_j^n)$ with $g(v_{j+1}^n)$. On the other hand,

$$|g_{j+1/2}^{2,n}|^2 + |g_{j-1/2}^{2,n}|^2 \le \frac{\beta^2}{4}\,(|v_j - v_{j-1}|^2 + |v_{j+1} - v_j|^2),$$

$$|g_{j+1/2}^{3,n}|^2 + |g_{j-1/2}^{3,n}|^2 \le$$

$$\tfrac{\gamma^2}{18}\,(|v_{j-1} - v_{j-2}|^2 + |v_j - v_{j-1}|^2 + |v_{j+1} - v_j|^2 + |v_{j+2} - v_{j+1}|^2).$$

The desired inequality (1.16) then follows. Finally, summing over all value j and shifting the index $j \mapsto j+1$ or $j \mapsto j-1$ whenever necessary, one deduce (1.17) from (1.16).

The statement about the order of accuracy of the scheme is an immediate consequence of the corresponding properties known for the semi-discrete scheme; see Theorem 11. □

In view of (1.17), the entropy is a decreasing function of time provided K is positive. This leads us to the following CFL-like restriction on λ to ensure the stability of the scheme:

$$6\,\lambda\,\frac{\|U''\|_\infty}{\beta}\,(\frac{37}{9}\,\|g'\|_\infty^2 + \frac{\beta^2}{2} + \frac{2}{9}\,\gamma^2) \le 1. \tag{1.20}$$

Of course, (1.20) also depends on β, γ and the maximum speed $\|g'\|_\infty$, as well as the entropy U.

It is interesting to observe that, in the limit $\lambda \to 0$, that is $k \to 0$ but h kept fixed, the entropy inequality (1.16) implies the semi-discrete one, (1.10), while (1.17) tends to (namely, $K \to \beta/2$)

$$\frac{d}{dt}\sum_{j=-\infty}^{\infty} U(u_j) + \frac{\beta}{2}\sum_{j=-\infty}^{\infty} |v_j(t) - v_{j-1}(t)|^2 \le 0. \tag{1.21}$$

We stress that the stability condition (1.20) is trivially satisfied in the limit $\lambda \to 0$! Indeed, the semi-discrete scheme (1.6)-(1.7)-(1.8)-(1.9) is unconditionnally stable (in the sense that (1.10)-(1.11)-(1.12)-(1.13) holds).

To obtain, as in LeFloch and Rohde [94], a scheme whose equivalent equation coincides with (1.1), we need to increase the order of discretization in time, to the third order, at least. Following Shu [121], we rely on a Runge-Kutta technique. Observe that the (first order in time) scheme can be rewritten in the form

$$\begin{aligned} &(i) \qquad u_j^{n+1} = u_j^n + L(u^n)_j, \\ &(ii) \qquad L(u^n)_j = -\lambda\,(g_{j+1/2}^n - g_{j-1/2}^n). \end{aligned} \tag{1.22}$$

A high-order accurate (in time) scheme is obtained by the Runge-Kutta technique, as follows:

$$\begin{aligned} &(i) \qquad u_j^{n+1} = \textstyle\sum_{k=0}^m (\alpha_k\, u_j^{n-k} + \beta_k\, L(u^{n-k})_j), \\ &(ii) \qquad u_j^{-k} = u_j^0, \quad k = 0, 1, ..., m. \end{aligned} \tag{1.23}$$

In the above, the coefficients α_k and β_k can be determined to have third-order accuracy in time. In our numerical experiments below, we will use

$$\begin{aligned} & m = 3, \\ & \alpha_0 = \tfrac{16}{27}\,,\ \alpha_1 = \alpha_2 = 0\,,\ \alpha_3 = \tfrac{11}{27}, \\ & \beta_0 = \tfrac{16}{9}\,,\ \beta_1 = \beta_2 = 0\,,\ \beta_3 = \tfrac{44}{27}, \end{aligned} \tag{1.24}$$

Note that the coefficients are positive and of unit sum, i.e.,

$$\begin{aligned} & \alpha_k \geq 0, \qquad k = 0, 1, ..., m, \\ & \textstyle\sum_{k=0}^m \alpha_k = 1. \end{aligned} \tag{1.25}$$

Theorem 13
Assume that the coefficients α_k, $k = 0, \dots, m$, satisfy the condition (1.25). Then the fully discrete scheme (1.23) satisfies the following local entropy inequality

$$\begin{aligned} & U(u_j^{n+1}) - \textstyle\sum_{k=0}^m \alpha_k U(u_j^{n-k}) + \lambda \sum_{k=0}^m \beta_k (G_{j+1/2}^{n-k} - G_{j-1/2}^{n-k}) \\ & + \tfrac{\lambda\beta}{4} \textstyle\sum_{k=0}^m \beta_k \left(|v_{j+1}^{n-k} - v_j^{n-k}|^2 + |v_j^{n-k} - v_{j-1}^{n-k}|^2\right) \\ & \leq \lambda^2 \textstyle\sum_{k=0}^m \frac{3\,\beta_k^2}{\alpha_k} \|U''\|_\infty \left(\|g'\|_\infty^2 D_j^{1,n-k} + \beta^2 D_j^{2,n-k} + \gamma^2 D_j^{3,n-k}\right). \end{aligned} \tag{1.26}$$

Hence, the entropy satisfies

$$\sum_{j=-\infty}^{+\infty} U(u_j^{n+1}) + \lambda \sum_{j=-\infty}^{+\infty} \sum_{k=0}^{m} d_k\, |v_{j+1}^{n-k} - v_j^{n-k}|^2 \leq \sum_{j=-\infty}^{+\infty} \sum_{k=0}^{m} \alpha_k U(u_j^{n-k}), \tag{1.27}$$

where

$$d_k := \frac{\beta\,\beta_k}{2} - \frac{3\,\lambda\,\beta_k^2}{\alpha_k}\, \|U''\|_\infty \Big(\|g'\|_\infty^2\, \frac{37}{9} + \frac{\beta^2}{2} + \frac{2\gamma^2}{9}\Big), \tag{1.28}$$

with the convention that $\beta_k/\alpha_k = 0$ when $\alpha_k = \beta_k = 0$. If the coefficients are chosen as in (1.24), then the equivalent equation of the scheme (1.23) is

$$\partial_t u + \partial_x f(u) = h\,\beta\, U'(u)_{xx} + h^2\, \gamma\, U'(u)_{xxx} + O(h^3 + k^3), \tag{1.29}$$

which, up to quadratic terms in h, coincides with the continuous model (1.1) provided h is replaced with ϵ.

Proof
In view of (1.25), the scheme (1.23i) admits the following convex decomposition:

$$u_j^{n+1} = \sum_{k=0}^{m} \alpha_k\, w_j^{n-k}, \tag{1.30}$$

with

$$w_j^{n-k} = u_j^{n-k} + \frac{\lambda\,\beta_k}{\alpha_k}\, L(u^{n-k})_j. \tag{1.31}$$

Since the entropy U is a convex function, we deduce from (1.30) that

$$U(u_j^{n+1}) \le \sum_{k=0}^{m} \alpha_k \, U(w_j^{n-k}).$$

Therefore, using the entropy inequality (1.16) with λ replaced with $\lambda \, \beta_k/\alpha_k$, we obtain immediately the inequality (1.26).

The global entropy inequality (1.27) is a consequence of (1.26), obtained by summing over all j and then shifting the index whenever needed. □

In view of (1.28), it is natural to impose the following CFL-like *stability restriction* on the ratio $\lambda = k/h$:

$$\begin{aligned} &(i) \qquad 6\,\lambda \, \sup_{0 \le k \le m} \tfrac{\beta_k}{\alpha_k} \, \tfrac{\|U''\|_\infty}{\beta} \left(\|g'\|_\infty^2 \, \tfrac{37}{9} + \tfrac{\beta^2}{2} + \tfrac{2\,\gamma^2}{9} \right) \le 1, \\ &(ii) \qquad \beta_k \ge 0. \end{aligned} \tag{1.32}$$

At this juncture, recall that the limiting solution $\lim_{\epsilon \to 0} u_\epsilon^{\beta,\gamma}$ only depends on the ratio $\delta = \beta/\gamma$. So, for a given value of δ, it is natural to choose β so that to minimize the constrain in (1.32), i.e. to minimize

$$\frac{\beta_k}{\alpha_k} \, \frac{\|U''\|_\infty}{\beta} \left(\|g'\|_\infty^2 \, \frac{37}{9} + \frac{\beta^2}{2} + \frac{2\,\beta^2}{9\,\delta^2} \right),$$

with respect to all $\beta > 0$. It is easy to check that we can then replace (1.32) with

$$\begin{aligned} &(i) \qquad 12\,\lambda \, \sup_{0 \le k \le m} \tfrac{\beta_k}{\alpha_k} \, \|U''\|_\infty \left(\|g'\|_\infty^2 \, \tfrac{37}{9} \left(\tfrac{1}{2} + \tfrac{2}{9\,\delta^2} \right) \right)^{1/2} \le 1, \\ &(ii) \qquad \beta_k \ge 0. \end{aligned} \tag{1.33}$$

From Theorem 13 we deduce that the entropy can never exceed the entropy at time $t = 0$.

Corollary 8
Under the restriction (1.32), we have

$$\sum_{j=-\infty}^{\infty} U(u_j^n) \le \sum_{j=-\infty}^{\infty} U(u_j^0). \tag{1.34}$$

Proof
In view of (1.32), the coefficients in (1.28) are positive, so that (1.27) gives us

$$\sum_{j=-\infty}^{+\infty} U(u_j^{n+1}) \le \sum_{j=-\infty}^{+\infty} \sum_{k=0}^{m} \alpha_k \, U(u_j^{n-k}). \tag{1.35}$$

We then proceed by induction on n. When $n = 0$, the inequality (1.35) together with (1.23ii) yields exactly

$$\sum_{j=-\infty}^{\infty} U(u_j^1) \le \sum_{j=-\infty}^{\infty} U(u_j^0),$$

which is (1.34).

If now, for some N,

$$\sum_{j=-\infty}^{\infty} U(u_j^n) \leq \sum_{j=-\infty}^{\infty} U(u_j^0) \qquad \text{for all } n \leq N,$$

then it follows again from (1.35) and (1.23ii) that

$$\sum_{j=-\infty}^{\infty} U(u_j^{N+1}) \leq \sum_{j=-\infty}^{+\infty} \sum_{k=0}^{m} \alpha_k U(u_j^0) = \sum_{j=-\infty}^{\infty} U(u_j^0).$$

This completes the proof of (1.34). □

It is clear that, if the solution is periodic in space, then the inequality (1.27) still holds when we sum up over one period. Thus (1.34) also holds true on one period. This observation will be used in next section, where we will numerically test the validity of this inequality.

1.3 Numerical experiments

In the present section, following Hayes and LeFloch [64], we experiment with the scheme proposed in Section 1.2. We demonstrate now that, in the limit $h \to 0$ (with $\lambda = k/h$ kept fixed), our fully discrete scheme generates nonclassical shock waves. We also check the monotonicity of the entropy as a function of time, in agreement with Corollary 8. On the other hand, we investigate the nucleation property of the scheme by studying the Riemann problem. In particular, we observe that, when the Riemann data remain below some threshold value, the corresponding solutions are always classical. All these properties were also met with the continuous model (1.1); see [87] and the references therein.

Let us recall that finite difference schemes for scalar conservation laws with nonconvex flux may contain some oscillations and, as a consequence, converge to (nonmonotone) nonclassical solutions. See LeFloch [87] for a review. Such a behavior can be avoided by restricting attention to TVD (total variation diminishing) schemes (after Harten). Indeed, in a TVD scheme, the number of local maxima or minima is a decreasing function of time, and as was observed in [69], such a scheme can not generate nonclassical shocks. The TVD approach is not applicable to handle limiting solutions generated by the diffusive-dispersive model (1.1), since the (nonclassical) solutions of interest do not possess monotonicity properties in general.

All the numerical experiments in this section will be done with the model (1.1) and the following choice of flux and entropy:

$$f(u) = u^3 - u, \quad U(u) = u^2/2, \qquad u \in I\!R. \tag{1.36}$$

The coefficients α_k and β_k given by (1.24) will be used. The time step is always taken to be the largest allowed by our CFL stability condition (1.32), excepted when it is mentionned otherwise. The space variable is taken to be $x \in [-1/2, 1/2]$.

Experiments 1 : We start by displaying some typical solutions generated by our scheme. The physical and numerical coefficients are chosen to be $\beta = 5.0$ and $\gamma = 37.5$

. We used 800 mesh points and t is equal to 0.003 and multiplied the condition (1.32) by a factor 20.

In Figure 1.1, we present, after about 8000 iterations of time, a two-wave solution, made of a nonclassical shock preceeded by a classical shock. Observe that the solution is non-monotone, altough it has been generated from a monotone Riemann data:

$$u(x,0) = \begin{cases} u_l & \text{for } x < 0, \\ u_r & \text{for } x > 0. \end{cases} \tag{1.37}$$

Here we have used $u_l = 4$ and $u_r = -3$.

In Figure 1.2, we display a two-wave solution containing a nonclassical shock and a rarefaction wave. Observe that the two waves are not attached, as it would be in a solution containing classical shocks only. Here we have $u_l = 4$ and $u_r = -3$ and about 15000 iterations of time have been used.

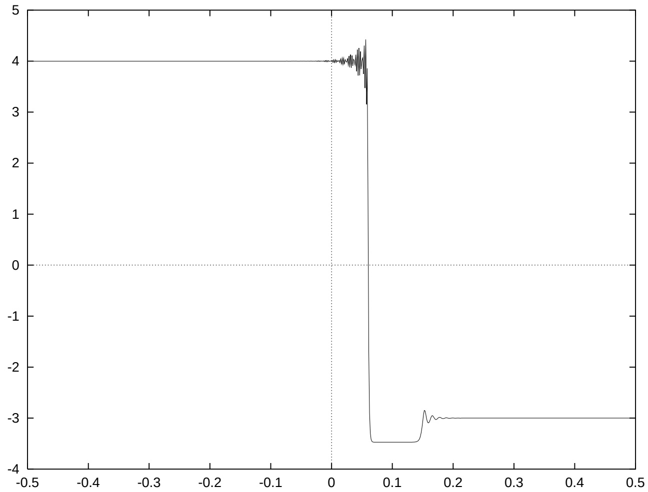

FIG. 1.1 – *A nonclassical shock + a classical shock*

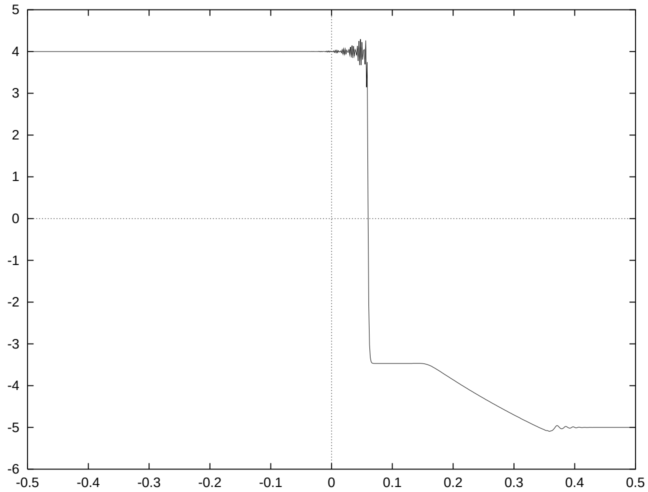

FIG. 1.2 – *A nonclassical shock + a rarefaction*

Experiments 2 : We want now to test the stability condition (1.32) on λ, found analytically in Section 1.2. First of all we demonstrate that the entropy is decreasing in time. Consider the periodic initial data $u(x,0) = -\sin(2\,\pi x))$ for $x \in [-1/2, 1/2]$ and let us solve (1.1) with periodic boundary conditions. The following parameters will be used

$$h = 0.00125, \qquad \beta = 5.0, \qquad \gamma = 18.75.$$

We want to test the validity of (1.34) on one period. The large-time behavior of the solution is displayed in Figure 1.3. We used here about 48000 iterations (with (1.33) as CFL-like condition) in time to reach the time $t = 0.24$. The solution contains a nonclassical shock and a rarefaction wave.

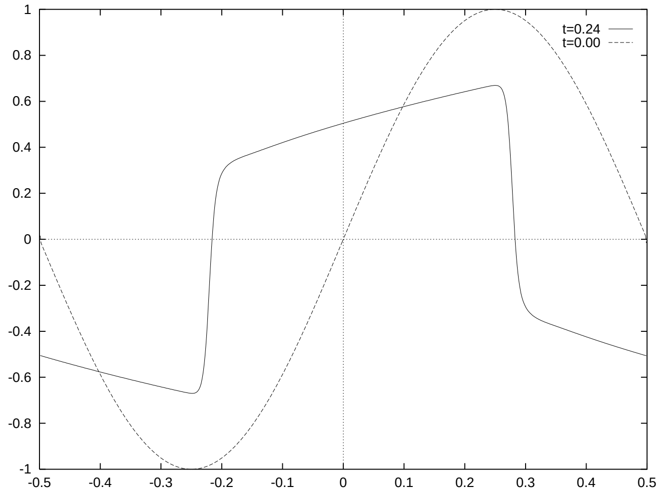

FIG. 1.3 – *A nonclassical shock + a rarefaction*

For λ sufficiently small, we observe that the decrease predicted by Corollary 8 holds. See Figure 1.4.

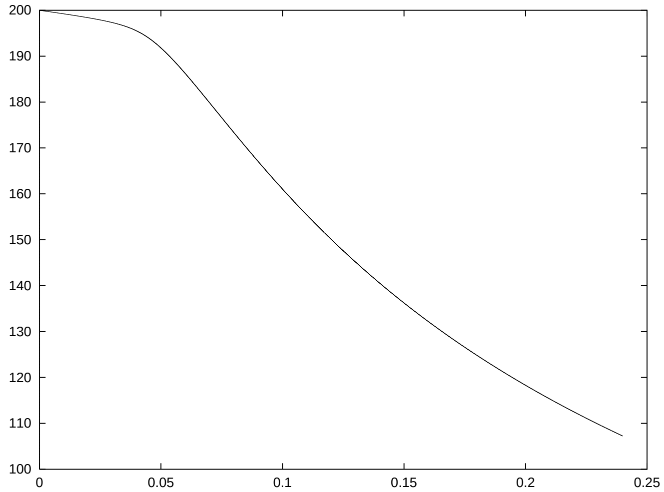

FIG. 1.4 – *Entropy versus time*

However, increasing λ, we note that, while the scheme still converges to nonclassical solutions (having oscillations of larger amplitude, however), it may well increase the entropy. For example, multiplying the time-step (1.32) by a factor 12, we see on Figures 1.5 and 1.6 that, even for small times, the scheme is less stable and the solutions much oscillatory, and (1.34) is violated. In fact, even the local in time inequality (1.27) is violated. However, the entropy remains decreasing in average.

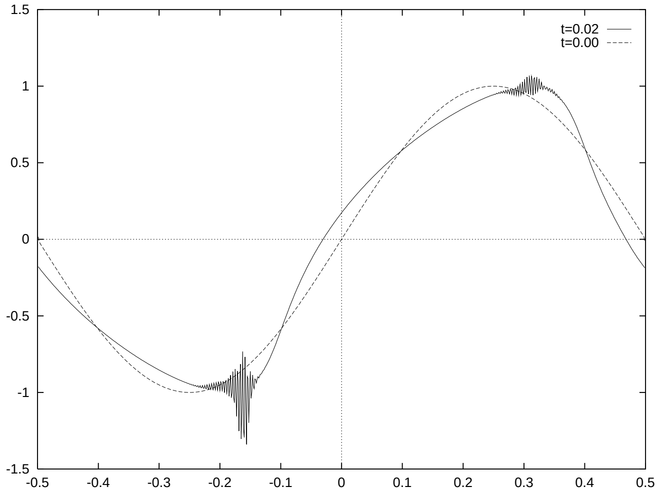

FIG. 1.5 – *Numerical solution with a large CFL*

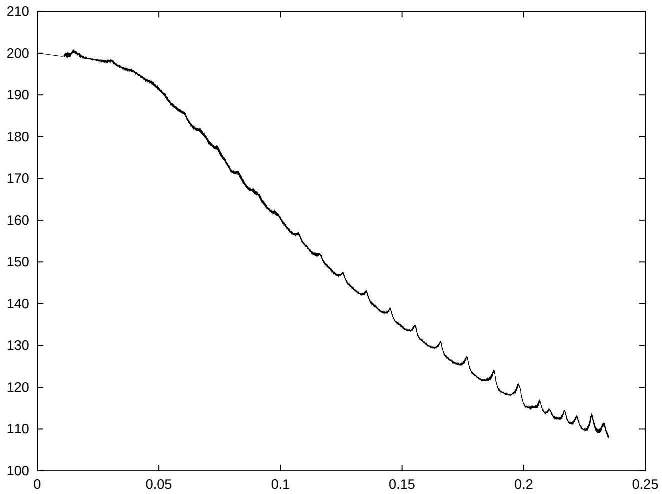

FIG. 1.6 – *Entropy with a large CFL*

Finally, our numerical tests demonstrate that the entropy may well increase when the conditions of the theorems in Section 1.2 are not fullfiled. Thereore the condition (1.32) on λ plays the role of a CFL condition, in a situation where one can not rely on the standard CFL condition (restricting λ in terms of the largest speed $|f'(u)|$ in the problem). Precisely, Figures 1.7 and 1.8 display some numerical solution obtained with our scheme but using now the standard CFL condition. Numerical oscillations are now much more important, and again (1.34) is violated. The time oscillations in Figure 1.8 have significantly larger amplitude than the ones in Figure 1.6. However, it would be interesting to derive further conditions on λ ensuring the stability of the scheme for arbitrary values of the parameters β and γ, precisely like the standard

CFL.

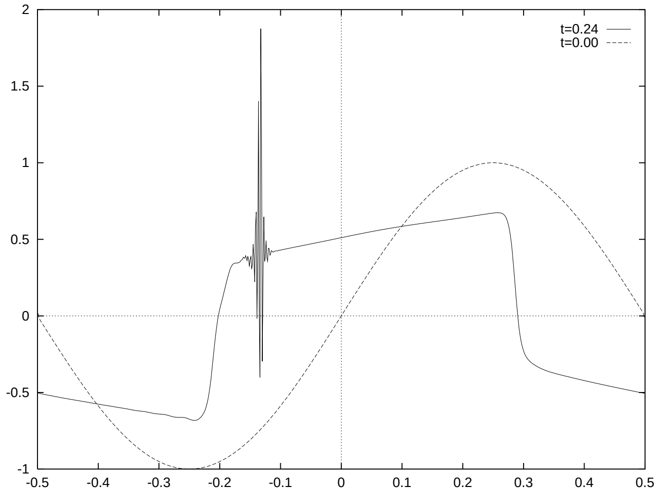

FIG. 1.7 – *Numerical solution with the standard CFL*

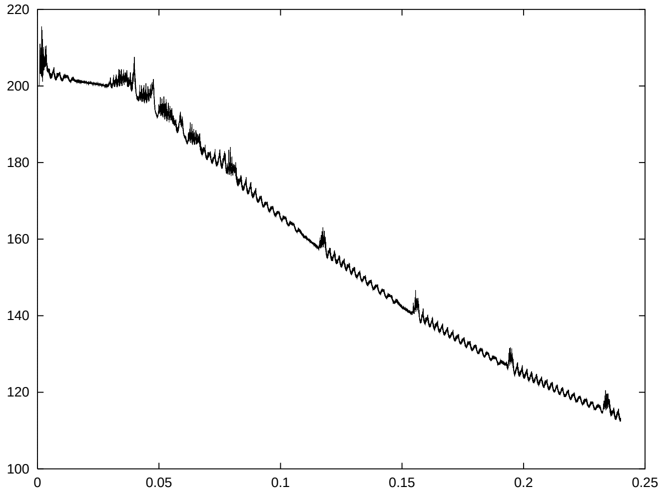

FIG. 1.8 – *Entropy with the standard CFL*

Experiments 3 : We now study how the scheme happens to nucleate nonclassical shocks. The discussion depends on the values on the Riemann data and the ratio of the diffusion to the dispersion. Consider the Riemann data (1.37) with $u_l > 0$ and $u_r < 0$. Motivated by a property of the continuous model (1.1), we claim that there exists a threshold $\psi(\delta)$ such that, for all $u_l < \psi(\delta)$, the Riemnann solution remains completely classical for arbitrary u_r, while for $u_l > \psi(\delta)$ the solution is nonclassical. This is indeed illustated by Figure 1.9, where

$$u_r = -5, \quad \beta = 5 \quad \text{and} \quad \gamma = 37.5.$$

The solution is classical when $u_l = 1$, but nonclassical when $u_l = 3$ or $u_l = 2$. The mesh contains 800 points for the cases $u_l = 3$ and $u_l = 4$, and 1200 points for the case $u_l = 1$. The numerical solutions are displayed at the time $t = 0.003$.

Similarly, there exists a threshold $\phi(u_l)$ such that, for $\delta > \psi(u_l)$, the Riemnann solution remains completely classical for arbitrary u_r, while for $\delta < \psi(u_l)$ the solution is always nonclassical. This is illustated in Figure 1.10, where

$$u_l = 4, \quad u_r = -5, \quad t = 0.003.$$

The solution is classical when $\beta = 30$, but nonclassical when $\beta = 5$ or $\beta = 15$. We used a grid made of 400 points and reduced δ by keeping γ fixed but by decreasing β.

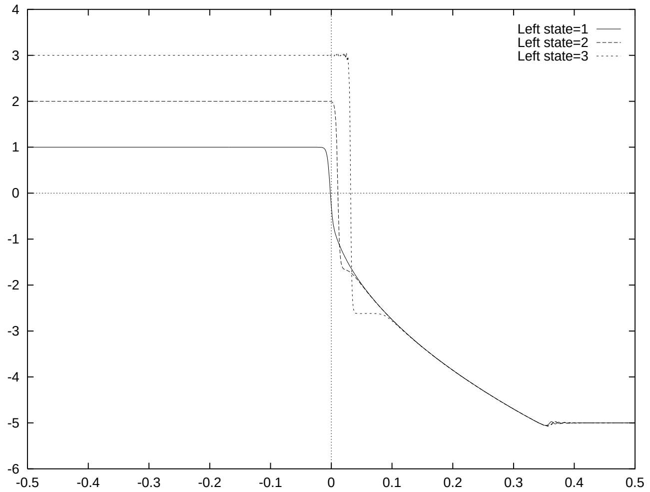

FIG. 1.9 – *Several values of u_l.*

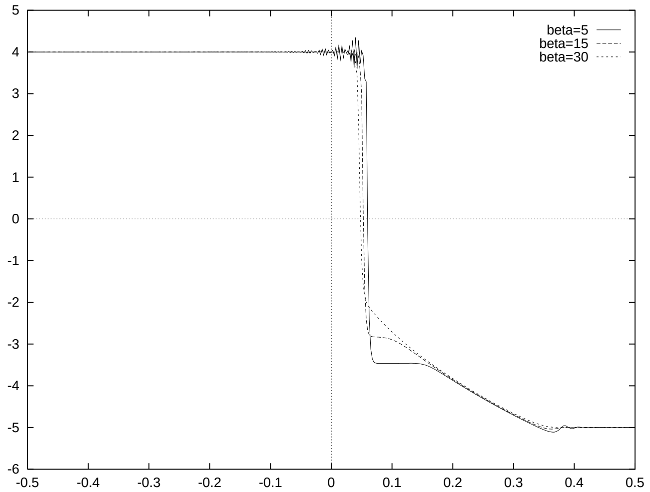

FIG. 1.10 – *Several values of β*

Finally, we claim that the function $\psi(\delta)$ increases when δ increases. To illustrate this point numerically, we consider the following parameters:

$$u_l = 4, \quad u_r = -5, \quad t = 0.003, \quad \beta = 5.$$

We used a mesh made of 800 points, and $\gamma = 1.75$, so that the ratio δ is larger, compared with the test in Figure 1.2. On Figure 1.11, the solution is classical.

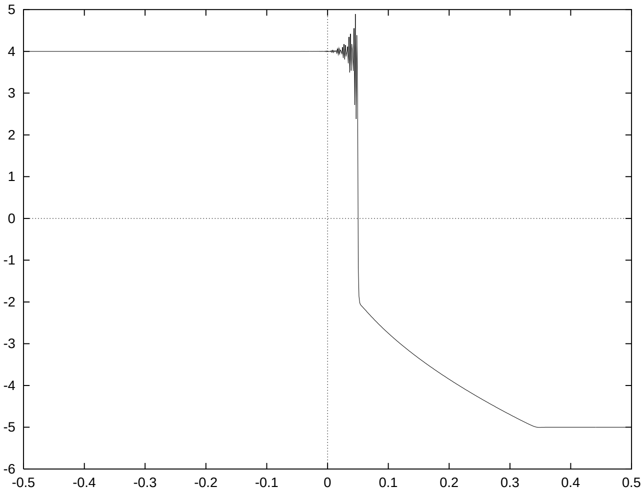

FIG. 1.11 – *Numerical solution for* $\gamma = 1.75$

Compare it with Figure 1.2 which has a nonclassical solution. This illustrates that $\psi(\delta)$ as δ increases.

1.4 Conclusion

To conclude, we emphasize that all the fundamental properties of the continuous model (1.1) are also shared by our numerical scheme. We demonstrated that the limiting solutions may contain undercompressive, nonclassical shocks that depend on the ratio δ. This happens only when the ratio of the diffusion to the dispersion is sufficiently small, or when the amplitude of the solution is sufficiently large. The class of schemes presented in this paper is efficient to compute nonclassical shock waves. For sufficiently small CFL numbers, the schemes satisfy exactly a discrete version of the entropy decay property (1.5); see Figure 1.4. Under a standard CFL condition, mild oscillations are observed in Figure 1.6, while the convergence toward the nonclassical solutions is still ensured.

Finally, we point out that the scheme proposed in Section 1.2 also applies to systems of conservation laws, and that for systems the entropy inequalities in Theorem 13 still hold.

Chapitre 2

High-order entropy conservative schemes and kinetic relations for van der Waals fluids

Ce chapitre a fait l'objet d'une publication dans le journal *Journal of Computational Physics* : voir [25]

High-order entropy conservative schemes and kinetic relations for van der Waals fluids

C. Chalons[1] and P.G. LeFloch[2]

Résumé

We consider the mixed (hyperbolic-elliptic) system of two conservation laws modeling the dynamics of van der Waals fluids. Viscosity and capillarity effects are taken into account. We introduce a new class of semi-discrete *high-order schemes* which are *entropy conservative* (in the sense of Tadmor) when the viscosity is neglected and, otherwise, dissipate the associated mathematical entropy. Our numerical schemes generalize the works by Tadmor [134] and LeFloch and Rohde [94] who proposed second-order and third-order, entropy conservative schemes, respectively.

Following Hayes and LeFloch [64], we demonstrate numerically that balanced viscosity and capillarity terms in van der Waals fluids may generate *nonclassical shock waves* or *subsonic propagating phase transitions.* Such waves are *undercompressive* and do not satisfy standard entropy criteria. They must be characterized by a *kinetic function,* which we determine numerically in this paper from vanishing viscosity and capillarity. The kinetic relation is the right tool to discuss the interplay between the viscosity, capillarity, and discretization parameters in van der Waals fluids.

2.1 Introduction

The dynamics of compressible fluids undergoing liquid-solid or vapor-liquid phase transformations can be modeled by the standard balance laws (mass, momentum) supplemented with a nonconvex equation of state, such as the one introduced by van der Waals. Restricting attention to a model of two conservation laws (the temperature being, formally, kept constant), one knows that, above some (critical) temperature this model is hyperbolic but not globally genuinely nonlinear (the pressure is a decreasing but not a globally convex function of the specific volume); on the other hand, below the critical temperature, the model is a mixed (hyperbolic-elliptic) system of conservation laws (the pressure is decreasing except on some bounded interval). Solutions of such systems of nonlinear P.D.E.'s are generally discontinuous and exhibit several distinct types of propagating waves:

(1) compressive shock waves satisfying the standard Lax or Liu entropy criteria;

(2) rarefaction waves, which are smooth and self-similar solutions;

(3) supersonic phase boundaries, which propagate faster than the characteristic speed;

(4) and (in the mixed type case) stationary phase boundaries.

1. O.N.E.R.A., B.P. 72, 29 avenue de la Division Leclerc, 92322 Châtillon Cedex, France. E-mail: chalons@onera.fr.

2. Centre de Mathématiques Appliquées & Centre National de la Recherche Scientifique, U.M.R. 7641, Ecole Polytechnique, 91128 Palaiseau Cedex, France. E-mail: lefloch@cmap.polytechnique.fr.

Until recently the Riemann problem –for which the initial datum is a single step-function– was solved allowing only stationnary and supersonic phase boundaries plus standard classical waves [119, 62]. Recently, after the works by James [71], Truskinovsky [138, 139], Slemrod [123], Abeyaratne and Knowles [1, 2], LeFloch [86, 87, 88], Shearer et al. [71, 120], and Hayes and LeFloch [69, 64, 65], it became clear that non-stationary, subsonic phase interfaces (in the hyperbolic-elliptic regime) and non-classical shock waves (in the hyperbolic but not genuinely nonlinear regime) should be included when solving the Riemann problem. Indeed, such waves are admissible in the sense that they do arise in viscosity-capillarity limits of the system.

Subsonic phase boundaries and nonclassical shocks have a special flavor: they are not uniquely characterized by the standard Rankine-Hugoniot relations and their unique selection requires as *additional jump relation* called a *kinetic relation.* Recall that there is indeed no universal selection criterion for propagating phase boundaries. The basic reason is that such waves are *undercompressive,* in the sense that –comparing with compressive shocks– fewer characteristics are impinging on the discontinuity.

The numerical approximation of the model under consideration was initiated by Slemrod and followers [128, 6, 121, 73, 29]. Computing kinetic relations to characterize undercompressive waves such as nonclassical shocks and subsonic phase boundaries was first tackled by Hayes and LeFloch [65], who identified the basic issues arising numerically. The present paper is a natural extension of [65].

In Section 2.2 below, we discuss the mathematical properties of the system modeling the dynamics of fluids with viscosity and capillarity effects included. A special emphasis is put on the mathematical entropy inequality, here associated with the total energy.

In Section 2.3, we introduce a new class of semi-discrete, *high-order schemes* which are *entropy conservative* (in the sense of Tadmor) if the viscosity term is neglected and, otherwise, satisfy a *discrete entropy inequality.* Our construction is inspired by Tadmor [133, 134] and LeFloch and Rohde [94] who derived earlier second-order and third-order, entropy conservative schemes, respectively. The main new idea is to treat as an independent variable the derivative of one of the conservative variable, here v_y, where v is the specific volume. An evolution equation for v_y is formulated and discretized in the scheme. An important observation is that an entropy inequality can indeed be derived.

In Sections 2.4 and 2.5, we show that our numerical scheme allows us to compute subsonic phase boundaries and nonclassical shock waves. We consider here the Riemann problem investigated theoretically by Slemrod [125] and Fan [46, 47, 48, 49]. Following Hayes and LeFloch [64], who studied nonconvex scalar conservation laws and phase transition models from nonlinear elasticity theory, we determine numerically the kinetic function associated with van der Waals fluids and we discuss the interplay between the viscosity, capillarity, discretization parameters. Concluding remarks are provided in Section 2.6.

2.2 Viscosity-capillarity model of compressible fluids

2.2.1 The mathematical formulation

The derivation of the equations is based on a variational formulation. See, for instance, the presentation given in Gavrilyuk and Gouin [52]. Let $(y,t) \mapsto \chi(y,t)$ be the Lagrangian description of the fluid motion. That is, by definition, the particle which was initially at the position y is located at the point $\chi(y,t)$ at the time t. The particle velocity u and the specific volume v are defined from χ by

$$u = \chi_t, \qquad v = \chi_y. \tag{2.1}$$

Prescribing an internal energy function of the form

$$e = e(v, v_y),$$

we postulate that the action

$$J(\chi) = \int_0^T \int_\Omega \Big(e(v, v_y) - \frac{u^2}{2}\Big)\, dydt = \int_0^T \int_\Omega \Big(e(\chi_y, \chi_{yy}) - \frac{\chi_t^2}{2}\Big)\, dydt \tag{2.2}$$

is minimal among all "admissible" χ. Here $\Omega \subset I\!R$ is the (bounded) interval initially occupied by the fluid and $[0,T]$ is a given time interval.

Let $g : \Omega \times [0,T] \to I\!R$ be a smooth function with compact support. Replacing in (2.2) χ with $\chi + g$ and keeping the first-order terms in g only, we obtain

$$\begin{aligned} J(\chi+g) = &\ \textstyle\int_0^T \int_\Omega \Big(e(\chi_y + g_y, \chi_{yy} + g_{yy}) - \frac{1}{2}\,(\chi_t + g_t)^2\Big)\, dydt \\ = &\ J(\chi) + \textstyle\int_0^T \int_\Omega \Big(\frac{\partial e}{\partial \chi_y}(\chi_y, \chi_{yy})\, g_y + \frac{\partial e}{\partial \chi_{yy}}(\chi_y, \chi_{yy})\, g_{yy} - \chi_t\, g_t\Big)\, dydt \\ &+ O(|g|^2), \end{aligned}$$

and, after integration par parts,

$$J(\chi+g) = J(\chi) +$$

$$\textstyle\int_0^T \int_\Omega \Big(\Big(-\frac{\partial e}{\partial \chi_y}(\chi_y, \chi_{yy})\Big)_y + \Big(\frac{\partial e}{\partial \chi_{yy}}(\chi_y, \chi_{yy})\Big)_{yy} + \chi_{tt}\Big)\, g\, dydt + O(|g|^2).$$

Since the solution χ should minimize the action J and that g is arbitrary, this formally yields

$$\chi_{tt} + \Big(-\frac{\partial e}{\partial \chi_y}(\chi_y, \chi_{yy}) + \Big(\frac{\partial e}{\partial \chi_{yy}}(\chi_y, \chi_{yy})\Big)_y\Big)_y = 0. \tag{2.3}$$

Returning to the functions u and v, in view of (2.1) we also have $v_t = \chi_{yt} = u_y$. Introducing the pressure term

$$P(v, v_y, v_{yy}) = -\frac{\partial e}{\partial v}(v, v_y) + \Big(\frac{\partial e}{\partial v_y}(v, v_y)\Big)_y, \tag{2.4}$$

we deduce from (2.1) and (2.3) that

$$\begin{aligned} & v_t - u_y = 0, \\ & u_t + P(v, v_y, v_{yy})_y = 0. \end{aligned} \tag{2.5}$$

This completes the derivation of the fluid model for the unknowns v and u.

Next, we take into account the viscosity effects. Denote by $\mu(v)$ the (volume dependent) viscosity coefficient of the fluid, the system (2.5) is replaced with

$$\begin{aligned} & v_t - u_y = 0, \\ & u_t + P(v, v_y, v_{yy})_y = (\mu(v)\, u_y)_y. \end{aligned} \tag{2.6}$$

Finally, defining the total energy by

$$E(v, u, v_y) = e(v, v_y) + \frac{u^2}{2},$$

we find the additional conservation law

$$E(v, u, v_y)_t + (P(v, v_y, v_{yy})\, u)_y = \Big(u_y \frac{\partial e}{\partial v_y}(v, v_y)\Big)_y + (\mu(v)\, u\, u_y)_y - \mu(v)\, u_y^2. \tag{2.7}$$

The energy plays here the role of a mathematical entropy.

It remains to discuss the properties of the internal energy function e. A standard choice in the literature on phase transition dynamics in fluids is to take e to be a quadratic function in v_y. (Linear term should not appear because of the invariance of the energy via the transformation $y \mapsto -y$.) We may assume that, for some positive function $\lambda(v)$, called the *capillarity coefficient,*

$$e(v, v_y) = \epsilon(v) + \lambda(v)\, \frac{v_y^2}{2}. \tag{2.8}$$

Under this condition, the total pressure P can be decomposed into a pressure-term depending only on v and a capillarity term, as follows:

$$P(v, v_y, v_{yy}) = p(v) - \lambda'(v)\, \frac{v_y^2}{2} + (\lambda(v)\, v_y)_y, \qquad p(v) = -\epsilon'(v). \tag{2.9}$$

Based on (2.9), the equations (2.6) take the more familiar form:

$$\begin{aligned} & v_t - u_y = 0, \\ & u_t + p(v)_y = g(\lambda'(v)\, \tfrac{v_y^2}{2} - (\lambda(v)\, v_y)_y g)_y + (\mu(v)\, u_y)_y. \end{aligned} \tag{2.10}$$

On the other hand, using (2.8)-(2.9), the equation (2.7) becomes

$$\begin{gathered} \Big(\epsilon(v) + \tfrac{u^2}{2} + \lambda(v)\, \tfrac{v_y^2}{2}\Big)_t + (p(v)\, u)_y = \\ \Big(u\, (\tfrac{\lambda'(v)}{2}\, v_y^2 - (\lambda(v)\, v_y)_y)\Big)_y + (\mu(v)\, u\, u_y)_y + (u_y\, \lambda(v)\, v_y)_y - \mu(v)\, u_y^2. \end{gathered} \tag{2.11}$$

We point out that this model has been often considered in the mathematical literature on phase transition dynamics, under the simplifying assumption that the capillarity

$\lambda(v) = \lambda_0$ be a positive constant.

2.2.2 Hyperbolicity and decrease of the mathematical entropy

It is immediate to check the following property. Consider the following linearization of the model (2.4)-(2.5) near some constant value v_0:

$$\begin{aligned} &v_t - u_y = 0, \\ &u_t - \tfrac{\partial^2 e}{\partial^2 v}(v_0, 0)\, v_y = 0, \end{aligned} \tag{2.12}$$

obtained by keeping first-order differential terms only. Then (2.12) is a strictly hyperbolic system of P.D.E.'s if and only if the internal energy function satisfies

$$\frac{\partial^2 e}{\partial^2 v}(v_0, 0) > 0. \tag{2.13}$$

Moreover, if (2.13) holds for all v_0, then the total energy $E(v, u, 0) = e(v, 0) + u^2/2$ is a strictly convex function of the conservative variables (v, u). For van der Waals fluids, the hyperbolicity condition holds only in some regions. See Section 2.4 below.

Given some internal energy function $e = e(v, v_y)$ and some nonlinear viscosity $\mu = \mu(v) > 0$, let us consider the corresponding viscosity-capillarity model (2.6), where the pressure $P(v, v_y)$ is defined by (2.4). Consider any solution $(y, t) \mapsto (v, u)(y, t)$ decaying to some constant solution (v^*, u^*) as $|y| \to \infty$ (its first-order space derivative also vanishing at infinity) Then we have

$$\frac{d}{dt}\mathcal{E}(t) \le 0, \tag{2.14}$$

where

$$\mathcal{E}(t) := \int_{I\!R} \Big(e(v(y,t), v_y(y,t)) - e(v^*, 0) - \frac{\partial e}{\partial v}(v^*, 0)\,(v(y,t) - v^*) + \frac{u(y,t)^2}{2} \Big)\, dy.$$

We conclude with some remarks in the hyperbolic regime.

For the example (2.8), the condition (2.13) reads $\epsilon''(v) > 0$. Let us assume the uniform bound

$$0 < \epsilon_0 \le \epsilon''(v) \le \epsilon_1 \qquad \text{for all } v > 0.$$

We have here

$$e(v, v_y) - e(v^*, 0) - \frac{\partial e}{\partial v}(v^*, 0)\,(v - v^*) = \epsilon(v) - \epsilon(v^*) - \epsilon'(v^*)\,(v - v^*) + \lambda(v)\,\frac{v_y^2}{2}.$$

Therefore, assuming also some bounds

$$\lambda_0 \le \lambda(v) \le \lambda_1 \qquad \text{for all } v > 0,$$

(2.14) yields the following a priori bound (for all $t \ge 0$)

$$\begin{aligned} &\textstyle\int_{I\!R} (\epsilon_0\, |v(y,t) - v^*|^2 + |u(y,t) - u^*|^2 + \lambda_0\, v_y^2)\, dy \\ &\le \textstyle\int_{I\!R} (\epsilon_1\, |v(y,0) - v^*|^2 + |u(y,0) - u^*|^2 + \lambda_1\, v_y^2)\, dy. \end{aligned}$$

More generaly, if the energy function satisfies inequalities of the form

$$\begin{aligned}\tfrac{\epsilon_0}{2}\,|v-v^*|^2+\tfrac{\beta_0}{2}\,|v_y|^2 \;&\le\; e(v,v_y)-e(v^*,0)-\tfrac{\partial e}{\partial v}(v^*,0)\,(v-v^*)\\ &\le\; \tfrac{\epsilon_1}{2}\,|v-v^*|^2+\tfrac{\beta_1}{2}\,|v_y|^2,\end{aligned}$$

for all v and v_y under consideration, then the following a priori estimate holds for all times t

$$\begin{aligned}&\int_{I\!R}\Big(\epsilon_0\,|v(y,t)-v^*|^2+|u(y,t)-u^*|^2+\beta_0\,|v_y(y,t)|^2\Big)\,dy\\ &\le \int_{I\!R}\Big(\epsilon_1\,|v(y,0)-v^*|^2+|u(y,0)-u^*|^2+\beta_1\,|v_y(y,0)|^2\Big)\,dy.\end{aligned}$$

2.3 A class of entropy consistent schemes

In this section, we focus on the numerical discretization of the model described in Section 2.2. Our objective is to derive a class of high-order, conservative, finite difference schemes that, additionally, are conservative for the associated mathematical entropy E when the viscosity effects are neglected and satisfy the discrete entropy inequality (2.14).

Our main concern is to design a scheme able to capture the zero viscosity-capillarity limits. Therefore, following Hayes and LeFloch [69], we are going to scale out all of the equation using the discretization parameter denoted by h; see the equations (2.15)-(2.16) and (2.18) below. Then the basic requirement is that the equivalent equation associated with the scheme should coincide up to some high-order term $O(h^3)$, at least, with these reference (continuous) equations.

Comparing rigorously the limiting solutions generated by the schemes and the limiting solutions generated by the zero viscosity-capillarity limits is the main contribution in [69] and this is the strategy pursued in the present work as well. We refer to the discussion in the following sections.

2.3.1 Discretization of the flux and capillarity terms

Consider first the model (2.4)-(2.5), neglecting therefore the viscosity. Introducing the new variables

$$w = h\,v_y, \quad z = h\,u_y,$$

we can consider w as an independent variable. We summarize the set of equations under consideration as follows:

$$\begin{aligned}v_t-u_y&=0,\\ u_t+P_y&=0,\\ w_t-z_y&=0.\end{aligned}\tag{2.15}$$

Recall that P is given by

$$P=-\frac{\partial e}{\partial v}(v,w)+h\,\Big(\frac{\partial e}{\partial v_y}(v,w)\Big)_y.\tag{2.16}$$

Denote by $h>0$ the space step of the discretization and, for all integer j, define the mesh point by $x_j=j\,h$. We denote by $v_j=v_j(t)$, $u_j=u_j(t)$, and $w_j=w_j(t)$

the discrete approximations at the points x_j (with prescribed initial conditions). Calling $u_{j+1/2}$, $P_{j+1/2}$, and $z_{j+1/2}$ some numerical flux terms still to be defined, we are interested in difference schemes in the following conservative form

$$
\begin{aligned}
&h \tfrac{d}{dt} v_j - (u_{j+1/2} - u_{j-1/2}) = 0,\\
&h \tfrac{d}{dt} u_j + P_{j+1/2} - P_{j-1/2} = 0,\\
&h \tfrac{d}{dt} w_j - (z_{j+1/2} - z_{j-1/2}) = 0.
\end{aligned}
\tag{2.17}
$$

We aim at defining the discrete fluxes $u_{j+1/2}$, $P_{j+1/2}$, and $z_{j+1/2}$ in such a way that the scheme (2.17) is also entropy conservative in the sense of [134]. That is, we seek for a discrete version of the energy equation

$$
\begin{aligned}
&\partial_t E + \partial_y F = 0,\\
&E = e(v,w) + \tfrac{u^2}{2}, \qquad F = P(v,w,h\,w_y)\,u - h\,u_y\,\tfrac{\partial e}{\partial v_y}(v,w).
\end{aligned}
\tag{2.18}
$$

With this in mind, we choose

$$
\begin{aligned}
u_{j+1/2} &= \sum_{k=-K}^{K+1} \alpha_k\, u_{j+k},\\
z_{j+1/2} &= \sum_{l=-L}^{L} \beta_l\, (u_{j+1+l} - u_{j+l}),
\end{aligned}
\tag{2.19}
$$

in which the terms u_{j+k}, for instance, are directly given by the second equation in (2.17). Some basic consistency argument forces us to impose

$$
\sum_{k=-K}^{K+1} \alpha_k = \sum_{l=-L}^{L} \beta_l = 1. \tag{2.20}
$$

It remains to specify the expression of the pressure $P_{j+1/2}$. This is done by requiring the condition

$$
\frac{d}{dt} \sum_{j=-\infty}^{+\infty} E_j = 0, \quad E_j = e(v_j, w_j) + \frac{u_j^2}{2}. \tag{2.21}
$$

By differentiation of the above expression for the energy E_j, we easily obtain

$$
h\,\frac{d}{dt} E_j = h\,\frac{\partial e}{\partial v}(v_j, w_j)\,\frac{d}{dt} v_j + h\,\frac{\partial e}{\partial v_y}(v_j, w_j)\,\frac{d}{dt} w_j + h\,u_j\,\frac{d}{dt} u_j.
$$

In view of (2.17) we have

$$
\begin{aligned}
h\,\tfrac{d}{dt} E_j = \;& \tfrac{\partial e}{\partial v}(v_j, w_j)\,(u_{j+1/2} - u_{j-1/2}) - u_j\,(P_{j+1/2} - P_{j-1/2})\\
&+ \tfrac{\partial e}{\partial v_y}(v_j, w_j)\,(z_{j+1/2} - z_{j-1/2}).
\end{aligned}
\tag{2.22}
$$

Using now (2.19) and after integration by parts, for instance writing

$$-\sum_{j=-\infty}^{+\infty} u_j(P_{j+1/2} - P_{j-1/2}) = \sum_{j=-\infty}^{+\infty} P_{j+1/2}(u_{j+1} - u_j),$$

we arrive at

$$h\frac{d}{dt}\sum_{j=-\infty}^{+\infty} E_j = \sum_{j=-\infty}^{+\infty} (P_{j+1/2} + Q_{j+1/2} + R_{j+1/2})\,(u_{j+1} - u_j),$$

with

$$Q_{j+1/2} = \sum_{k=-K}^{K+1} \alpha_k \tfrac{\partial e}{\partial v}(v_{j+1-k}, w_{j+1-k}),$$

$$R_{j+1/2} = \sum_{l=-L}^{L+1} \beta_l \left(-\tfrac{\partial e}{\partial v_y}(v_{j+1-l}, w_{j+1-l}) + \tfrac{\partial e}{\partial v_y}(v_{j-l}, w_{j-l})\right).$$

One sufficient condition for (2.21) to hold is to choose $P_{j+1/2}$ according to

$$P_{j+1/2} = -Q_{j+1/2} - R_{j+1/2}.$$

Using the change variable $k' = 1-k$ and $l' = -l$, one checks easily that the above definition is equivalent to setting

$$\begin{aligned} P_{j+1/2} = & -\sum_{k=-K}^{K+1} \alpha_k \tfrac{\partial e}{\partial v}(v_{j+k}, w_{j+k}) \\ & + \sum_{l=-L}^{L} \beta_l \left(\tfrac{\partial e}{\partial v_y}(v_{j+l+1}, w_{j+l+1}) - \tfrac{\partial e}{\partial v_y}(v_{j+l}, w_{j+l})\right). \end{aligned} \tag{2.23}$$

This completes the derivation of the scheme for (2.4)-(2.5).

2.3.2 Local entropy inequalities

Next we determine some mathematical entropy fluxes for each cell. Replace $u_{j+1/2}$, $P_{j+1/2}$, and $z_{j+1/2}$ in (2.22) by their definition. We obtain

$$h\frac{d}{dt}E_j = \sum_{k=-K}^{K+1} \alpha_k (A^{(1)}_{j,k} - A^{(2)}_{j,k}) + \sum_{l=-L}^{L} \beta_l (B^{(1)}_{j,l} - 2B^{(2)}_{j,l} + B^{(3)}_{j,l}), \tag{2.24}$$

with

$$A^{(1)}_{j,k} = u_{j+k}\tfrac{\partial e}{\partial v}(v_j, w_j) - u_j\tfrac{\partial e}{\partial v}(v_{j-k}, w_{j-k}),$$

$$A^{(2)}_{j,k} = u_{j+k-1}\tfrac{\partial e}{\partial v}(v_j, w_j) - u_j\tfrac{\partial e}{\partial v}(v_{j+1-k}, w_{j+1-k}),$$

$$B^{(1)}_{j,l} = u_{j+l-1}\tfrac{\partial e}{\partial v_y}(v_j, w_j) - u_j\tfrac{\partial e}{\partial v_y}(v_{j+1-l}, w_{j+1-l}),$$

$$B^{(2)}_{j,l} = u_{j+l}\tfrac{\partial e}{\partial v_y}(v_j, w_j) - u_j\tfrac{\partial e}{\partial v_y}(v_{j-l}, w_{j-l}),$$

$$B^{(3)}_{j,l} = u_{j+l+1}\tfrac{\partial e}{\partial v_y}(v_j, w_j) - u_j\tfrac{\partial e}{\partial v_y}(v_{j-1-l}, w_{j-1-l}).$$

Finally, using decompositions of the type

$$a_{j+k} - a_j = (a_{j+k} + a_{j+k-1} + \ldots + a_{j+1}) - (a_{j+k-1} + a_{j+k-2} + \ldots + a_j),$$

we see that each of the terms above admit a conservative form. This allows us to determine easily a discrete entropy flux $F_{j+1/2}$. We conclude that there exist numerical fluxes $F_{j+1/2}$ (formally consistent with the continuous flux F) so that the following discrete conservation law holds:

$$h\frac{d}{dt}\left(e(v_j, w_j) + \frac{u_j^2}{2}\right) + F_{j+1/2} - F_{j-1/2} = 0. \tag{2.25}$$

In conclusion, the formulas (2.17), (2.19), and (2.23) define a scheme for the unknowns v_j and u_j, which is conservative for all of the equations in (2.17), including the discrete energy

$$E_j = e(v_j, w_j) + u_j^2/2. \tag{2.26}$$

In particular we have the important stability property:

$$\frac{d}{dt}\sum_{j=-\infty}^{\infty}\left(e(v_j, w_j) + \frac{u_j^2}{2}\right) = 0. \tag{2.27}$$

2.3.3 Discretization of the viscosity terms

To complete the description of the numerical scheme, we now take into account the viscosity terms, covering the general case of a nonlinear viscosity coefficient $\mu(v)$. It is sufficient to present the construction in the case of the system where the flux and capillarity terms have been formally neglected, that is:

$$\begin{aligned} &v_t = 0,\\ &u_t = h\,(\mu(v)\,u_y)_y,\\ &w_t = 0. \end{aligned} \tag{2.28}$$

Consider the numerical discretization scheme:

$$\begin{aligned} &h\,\tfrac{d}{dt}v_j = 0,\\ &h\,\tfrac{d}{dt}u_j = \mu_{j+1/2}\,q_{j+1/2} - \mu_{j-1/2}\,q_{j-1/2},\\ &h\,\tfrac{d}{dt}w_j = 0, \end{aligned} \tag{2.29}$$

with

$$q_{j+1/2} = \sum_{l=-L}^{L}\beta_l\,(u_{j+l+1} - u_{j+l}). \tag{2.30}$$

Here the numerical values $\mu_{j+1/2}$ are approximations of sufficiently high-order of the nonlinear viscosity $\mu(v)$.

Differentiating the energy E_j defined in (2.26) with respect to t and using (2.29), we obtain

$$h\,\frac{d}{dt}E_j = u_j\,(\mu_{j+1/2}\,q_{j+1/2} - \mu_{j-1/2}\,q_{j-1/2}).$$

Summing over j and integrating by parts, we obtain

$$\begin{aligned}\sum_{j=-\infty}^{+\infty} \tfrac{d}{dt}E_j &= - \sum_{j=-\infty}^{+\infty} \mu_{j+1/2}\, q_{j+1/2}\, (u_{j+1} - u_j) \\ &= - \sum_{j=-\infty}^{+\infty} \sum_{l=-L}^{L} \mu_{j+1/2}\beta_l\, (u_{j+l+1} - u_{j+l})\, (u_{j+1} - u_j).\end{aligned}$$

Setting

$$m_{j+l} = u_{j+l+1} - u_{j+l},$$

it follows

$$\begin{aligned}\sum_{j=-\infty}^{+\infty} \tfrac{d}{dt}E_j &= - \sum_{j=-\infty}^{+\infty} \sum_{l=-L}^{L} \mu_{j+1/2}\beta_l\, m_{j+l}\, m_j \\ &= - \sum_{j=-\infty}^{+\infty} \sum_{k=j-L}^{j+L} \mu_{j+1/2}\beta_{k-j}\, m_k\, m_j.\end{aligned} \tag{2.31}$$

Assume from now on that the coefficients β_l (first introduced in (2.19), then used again in (2.30)) are chosen such that the following (infinite) quadratic form is non-negative:

$$\sum_{j=-\infty}^{+\infty} \sum_{k=j-L}^{j+L} \mu_{j+1/2}\beta_{k-j}\, m_k\, m_j \geq 0 \tag{2.32}$$

for all w_k under consideration. Then from (2.31) we deduce that

$$\sum_{j=-\infty}^{+\infty} \frac{d}{dt}E_j \leq 0, \qquad E_j = e(v_j, w_j) + \frac{u_j^2}{2}. \tag{2.33}$$

This property is to be compared with the equality found in (2.27) for the flux and capillarity terms.

2.3.4 Order of accuracy

The parameters α_k and β_l in (2.19) and (2.30) must be chosen so that (2.20) and (2.33) hold true. We also require that the order of accuracy be sufficiently high so that the equivalent equation associated with the scheme coincides with the original system except for terms of order $O(h^3)$.

For fixed values K and L in (2.19) and (2.30), it is always possible to find some coefficients α_k, $k = -K, \ldots, K+1$, and β_l, $l = -L, \ldots, L$, to guarantee:

$$u_{j+1/2} - u_{j-1/2} = \sum_{k=-K}^{K+1} \alpha_k (u_{j+k} - u_{j+k-1}) = h u_y + 0(h^{2K+3}) \tag{2.34}$$

and

$$
\begin{aligned}
z_{j+1/2} - z_{j-1/2} &= \sum_{l=-L}^{L+1} \beta_l (u_{j+l+1} - 2u_{j+l} + u_{j+l-1}) \\
&= h^2 (u_j)_{yy} + 0(h^{2L+4}) \\
q_{j+1/2} - q_{j-1/2} &= \sum_{l=-L}^{L+1} \beta_l (u_{j+l+1} - 2u_{j+l} + u_{j+l-1}) \\
&= h^2 (u_j)_{yy} + 0(h^{2L+4})
\end{aligned}
\tag{2.35}
$$

Hence, from the definition (2.23) of $P_{j+1/2}$, it is clear that

$$P_{j+1/2} - P_{j-1/2} = -h(\frac{\partial e}{\partial v}(v_j, w_j))_y + 0(h^{2K+3}) + h^2(\frac{\partial e}{\partial v_y}(v_j, w_j))_{yy} + 0(h^{2L+4}),$$

that is

$$P_{j+1/2} - P_{j-1/2} = h\, P(v_j, w_j) + 0(h^{2K+3}) + 0(h^{2L+4}) \tag{2.36}$$

holds.

Based on (2.34)–(2.36) we then find the equivalent equation of the scheme:

$$
\begin{aligned}
& v_t - u_y = 0(h^{2K+2}), \\
& u_t + P_y = h(\mu(v) u_y)_y + 0(h^{2K+2}) + 0(h^{2L+3}), \\
& w_t - z_y = 0(h^{2L+3}).
\end{aligned}
$$

It follows that the optimal choice is obtained for $K = L$, so that we find

$$
\begin{aligned}
& v_t - u_y = 0(h^{2K+2}), \\
& u_t + P_y = h(\mu(v) u_y)_y + 0(h^{2K+2}), \\
& w_t - z_y = 0(h^{2K+3}).
\end{aligned}
\tag{2.37}
$$

In Section 2.5 below we will investigate several choices of parameters $K = L = 1, 2, 3$.

For $K = L = 1$, we find

$$
\begin{aligned}
& (\alpha_{-1}, \alpha_0, \alpha_1, \alpha_2) = \left(\tfrac{-1}{12}, \tfrac{7}{12}, \tfrac{7}{12}, \tfrac{-1}{12}\right), \\
& (\beta_{-1}, \beta_0, \beta_1) = \left(\tfrac{-1}{12}, \tfrac{7}{6}, \tfrac{-1}{12}\right).
\end{aligned}
$$

For $K = L = 2$, we find

$$
\begin{aligned}
& (\alpha_{-2}, \alpha_{-1}, \alpha_0, \alpha_1, \alpha_2, \alpha_3) = \left(\tfrac{1}{60}, \tfrac{-2}{15}, \tfrac{37}{60}, \tfrac{37}{60}, \tfrac{-2}{15}, \tfrac{1}{60}\right), \\
& (\beta_{-2}, \beta_{-1}, \beta_0, \beta_1, \beta_2) = \left(\tfrac{1}{90}, \tfrac{-23}{180}, \tfrac{37}{30}, \tfrac{-23}{180}, \tfrac{1}{90}\right).
\end{aligned}
$$

For $K = L = 3$, we find

$$(\alpha_{-3}, \alpha_{-2}, \alpha_{-1}, \alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4) = \left(\tfrac{-1}{280}, \tfrac{29}{840}, \tfrac{-139}{840}, \tfrac{533}{840}, \tfrac{533}{840}, \tfrac{-139}{840}, \tfrac{29}{840}, \tfrac{-1}{280}\right),$$

$$(\beta_{-3}, \beta_{-2}, \beta_{-1}, \beta_0, \beta_1, \beta_2, \beta_3) = \left(\tfrac{-1}{560}, \tfrac{11}{504}, \tfrac{-779}{5040}, \tfrac{533}{420}, \tfrac{-779}{5040}, \tfrac{11}{504}, \tfrac{-1}{560}\right).$$

Let us finally check the sign property (2.32) when the viscosity μ is constant. Consider the case $K = L = 1$. We have:

$$\begin{aligned} \sum_{j=-\infty}^{+\infty} \sum_{k=-1}^{1} \mu\, \beta_k\, m_{j+k}\, m_j &= -\sum_{j=-\infty}^{+\infty} \tfrac{\mu}{12}(-m_j\, m_{j-1} + 14\, m_j^2 - m_j\, m_{j+1}) \\ &= \tfrac{\mu}{6} \sum_{j=-\infty}^{+\infty} (7\, m_j^2 - m_j\, m_{j+1}). \end{aligned}$$

Since

$$(m_j - m_{j+1})^2 = m_j^2 + m_{j+1}^2 - 2m_j\, m_{j+1} \geq 0,$$

we have

$$\sum_{j=-\infty}^{+\infty} (m_j^2 - m_j\, m_{j+1}) \geq 0,$$

and

$$\frac{\mu}{6} \sum_{j=-\infty}^{+\infty} (7\, m_j^2 - m_j\, m_{j+1}) \geq 0.$$

This indeed implies the desired decreasing property (2.33). By a tedious but rather straighforward calculation, it can be checked that the same result holds in the other two cases $K = L = 2$ and $K = L = 3$.

2.4 Subsonic phase boundaries and nonclassical shocks

In this section we show that the schemes proposed in Section 2.3 allow us to compute subsonic (and supersonic) phase boundaries and nonclassical undercompressive shock waves. We are primary interested in the van der Waals pressure-law. However, it is convenient also to compare it with a cubic pressure-law described below. Throughout, the time-discretization is based on a standard Runge-Kutta approach of sufficiently high-order of accuracy. Except when specified otherwise, all the tests are done with the scheme in Section 2.2 corresponding to $K = L = 2$, so that the scheme is 6th order in space. All of the numerical solutions will be generated from an initial datum of the form

$$(v(x,0), u(x,0)) = \begin{cases} (v_l, u_l) & \text{for } x < 0, \\ (v_r, u_r) & \text{for } x > 0, \end{cases}$$

for some constant Riemann data to be specified. The visocisty and capillarity coefficients μ and λ will be taken to be constant.

2.4.1 Cubic pressure law

The van der Waals can be well approximated by the cubic equation

$$p(v) = -(v-a)^3 + v + b, \qquad v > 0, \tag{2.38}$$

where $a > 0$ and $b > 0$ are constants. In our experiments, for simplicity in the calculations, we take $a = 4$ and $b = 6$. See Figure 2.1a for a representation of the graph of p. Setting

$$v_- = 4 - \frac{\sqrt{3}}{3}, \quad v_+ = 4 + \frac{\sqrt{3}}{3}, \tag{2.39}$$

three distinct regions can be distinguished:

(1) The interval $v \in [0, v_-]$ corresponds to the liquid phase: the system (2.19) is strictly hyperbolic and genuinely nonlinear.

(2) In the interval $v \in [v_-, v_+]$, the system is of elliptic type.

(3) The interval $v \in [v_+, \infty)$ corresponds to the vapor phase: the system (2.19) is strictly hyperbolic and genuinely nonlinear.

The so-called Maxwell stationary phase boundary, by definition, corresponds to a zero entropy dissipation. Here it connects $v = 3$ to $v = 5$ or vice-versa.

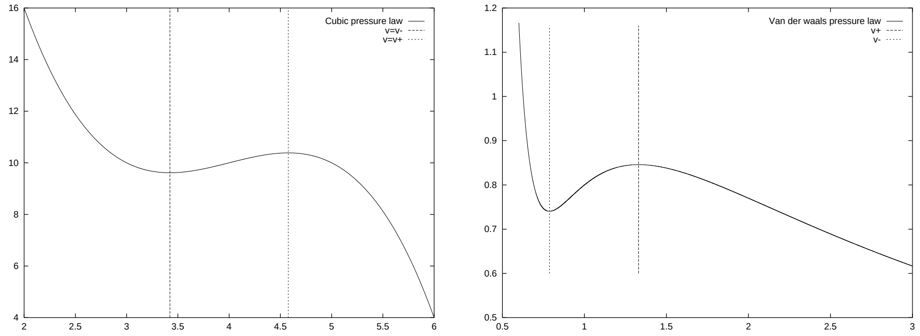

FIG. 2.1 – *(a) Cubic pressure law (b) van der Waals pressure-law*

Test 1 : Propagating phase boundary

Figure 2.2 (v and u components) displays a typical subsonic phase boundaries, here propagating to the left, and preceeded with a rarefaction wave. The dot lines based on the critival values v_- and v_+ limit the hyperbolic and elliptic regions. The data for this test are

$$(\mu, \lambda) = (1, 0.2), \quad (v_l, u_l) = (3, 0), \quad (v_r, u_r) = (5, 2).$$

The mesh contains 600 points and the solution is represented at the time $t = 0.25$.

Test 2 : Stationary phase boundary

Next in Figure 2.3 we used

$$(\mu, \lambda) = (1, 1.5), \quad (v_l, u_l) = (2.8, 0), \quad (v_r, u_r) = (5.2, 0).$$

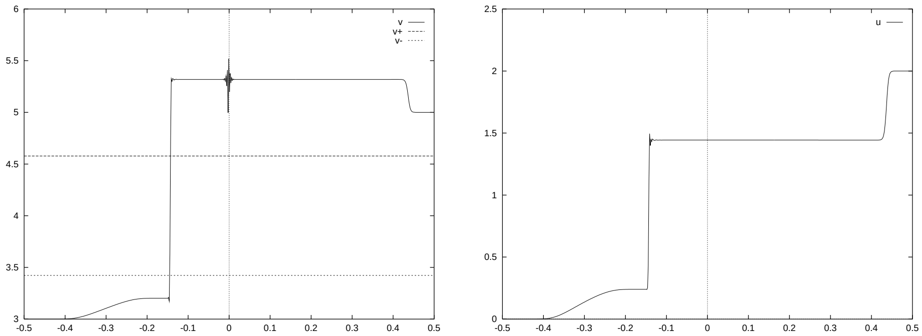

FIG. 2.2 – *Propagating phase boundary (a) volume component (b) velocity component*

The mesh contains 500 points and the solution is displayed at the time $t = 0.15$. We start here with a continuous velocity, which induces simply a stationary phase transition plus two rarefaction waves in each of the characteristic families. Not surprisingly the stationary phase boundary satisfies the Maxwell condition, indeed it connects the values $v = 3$ to $v = 5$.

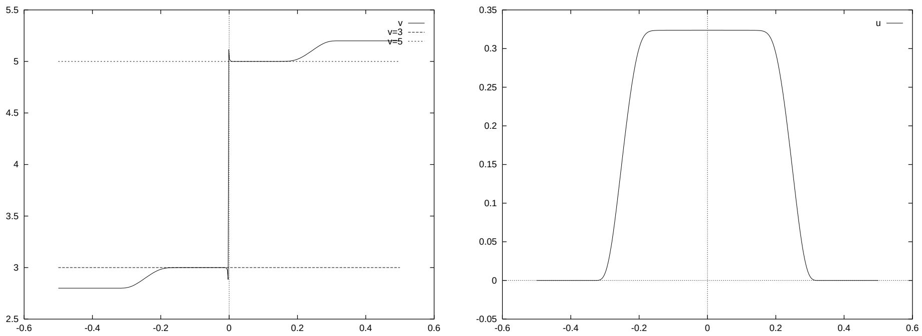

FIG. 2.3 – *Propagating phase boundary (a) volume component (b) velocity component*

Test 3 : Effect of the viscosity and capillarity coefficients

The solution depends on the relative values of the viscosity and capillarity coefficients. Indeed, in Figure 2.4, using

$$\mu = 1, \quad (v_l, u_l) = (2.8, 0), \quad (v_r, u_r) = (5, 2),$$

and successively $\lambda = 0.1, 0.18, 0.25$, we obtain propagating phase boundaries propagating with various subsonic speeds. Here we used 800 mesh points and displayed the solution at the time $t = 0.22$.

Figure 2.5 illustrates that the phase boundaries are truly subsonic: the straight line connecting the two states cuts the graph of the pressure.

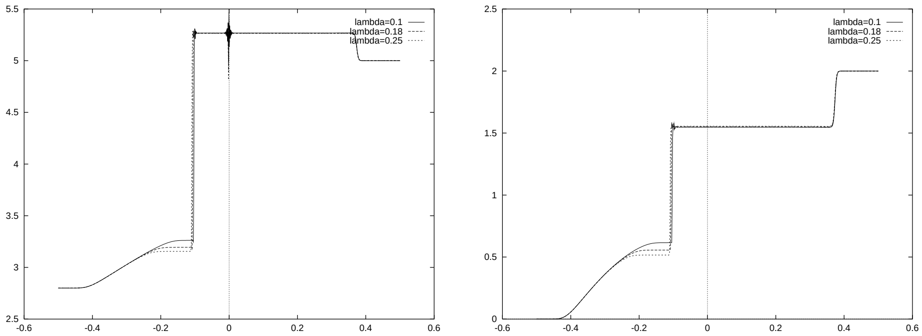

FIG. 2.4 – *Several values of* λ *(a) volume component (b) velocity component*

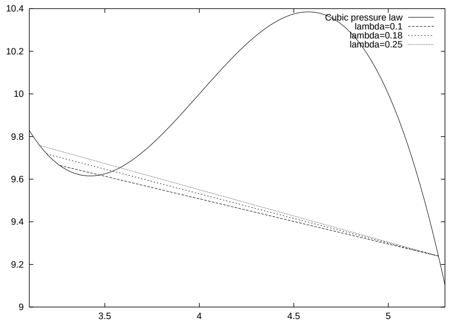

FIG. 2.5 – *Subsonic phase boundaries*

2.4.2 van der Waals pressure-law

In the rest of this section we deal with the well-known van der Waals equation of state, given by

$$p(v,T) = \frac{RT}{v-b} - \frac{a}{v^2}, \tag{2.40}$$

where a, b, R are numerical constants and where the temperature $T > 0$ is fixed. We use here the following standard constants

$$a = 3, \quad b = \frac{1}{3}, \quad R = \frac{8}{3}, \tag{2.41}$$

and, for definiteness, the temperature is taken to be $T = 0.95$, just below the critical temperature $T = 1$. Similarly as in Section 2.4.1 on the cubic model, the system under consideration is hyperbolic and genuinely nonlinear in each on the regions $v < v_-$ and $v > v_+$, where, approximatively,

$$v_- = 0.787, \quad v_+ = 1.330.$$

The Maxwell line corresponds to the phase boundary connecting $v = 0.684$ (liquid) and $v = 1.727$ (vapor). Observe that the pressure is a convex function of v sufficiently small but a concave function for v sufficiently large (See Figure 2.1b).

Above the critical temperature, the model under consideration is always hyperbolic but is not always genuinely nonlinear; see Section 2.4.3.

Test 4 : Propagating phase boundary

Figure 2.6 displays a propagating phase transition obtained from the following data:

$$(\mu, \lambda) = (0.1, 1e-5), \quad (v_l, u_l) = (0.6, -2), \quad (v_r, u_r) = (1.5, 0).$$

We used a mesh with 1000 points and we represent the solution at the time $t = 0.15$.

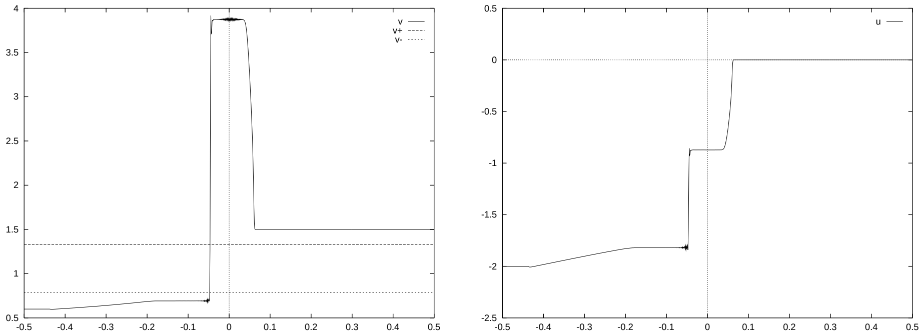

FIG. 2.6 – *Propagating phase boundary (a) volume component (b) velocity component*

Test 5 : Stationary phase boundary

In Figure 2.7, the test was performed with

$$(\mu, \lambda) = (0.1, 1e-3), \quad (v_l, u_l) = (0.684117091, 0), \quad (v_r, u_r) = (1.72700257, 0).$$

The initial jump is very close to the Maxwell stationary phase boundary and, as expected, the scheme keep the Maxwell discontinuity stationary. The mesh contains here 600 points and the solution is displayed at the time $t = 0.20$.

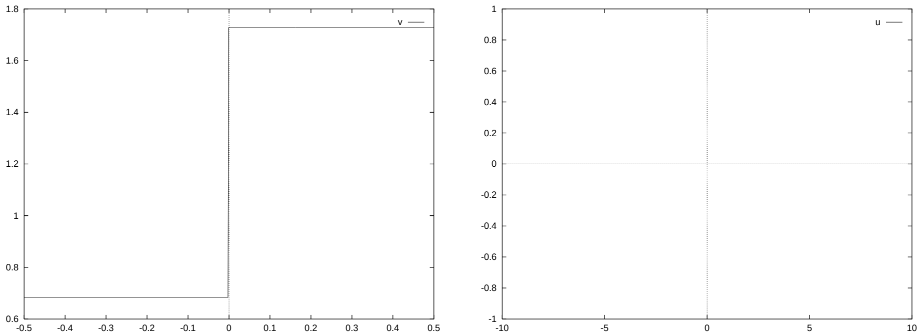

FIG. 2.7 – *Stationary phase boundary (a) volume component (b) velocity component*

Test 6 : Effect of the viscosity and capillarity coefficients

Finally, as in Test 3, we demonstrate that the solutions depend on the relative sizes of the viscosity and capillarity parameters. Let us use

$$\mu = 0.1, \quad (v_l, u_l) = (0.675, -1.45), \quad (v_r, u_r) = (0.75, 0),$$

together with several values of capillarity $\lambda = 10^{-5}, 0.1, 0.75$. See Figure 2.8. The solution contains two (symmetric) propagating phase boundaries with opposite speeds. The mesh contains 800 points and the solution is represented at the time $t = 0.22$.

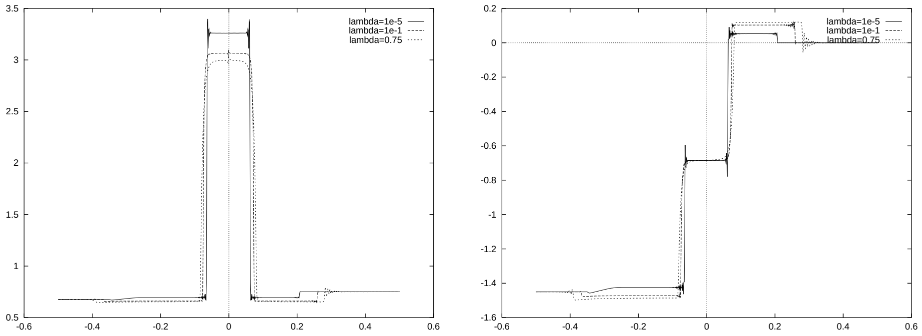

FIG. 2.8 – *Several values of* λ *(a) volume component (b) velocity component*

Figure 2.9 illustrates that the phase boundaries are truly subsonic: the straight line connecting the two states cuts the graph of the pressure.

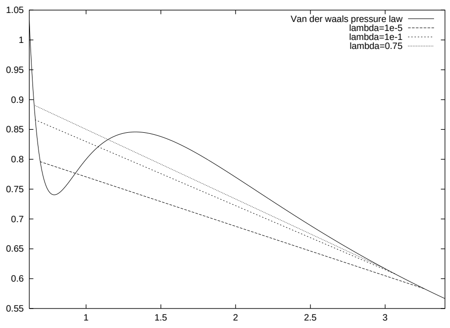

FIG. 2.9 – *Subsonic phase boundaries*

2.4.3 Nonclassical shock waves in the hyperbolic regime

Above the critical temperature (that is $T > 1$ in normalized units), the van der Waals model is strictly hyperbolic, even though not always genuinely nonlinear. In this hyperbolic regime, instead of subsonic phase boundaries, the equations may exhibit nonclassical shock waves. This is illustrated below. Recall that the cubic equation of state in the hyperbolic regime was dealt with by Hayes and LeFloch in [65].

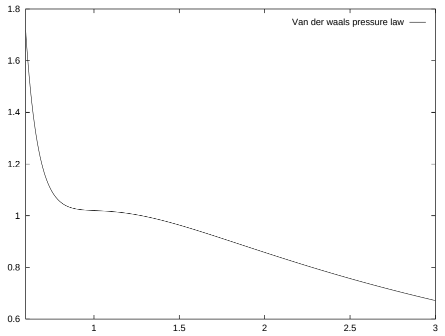

FIG. 2.10 – *van der Waals pressure law – T=1.005*

We choose now $T = 1.005$. Figure 2.10 represents the graph of the corresponding pressure law.

The viscosity coefficient is fixed to be 0.1. The mesh size is 1000 points and $K = L = 2$ as previously. Figure 2.11 shows a nonclassical shock obtained with the following values

$$(v_l, u_l) = (0.8, 0) \quad (v_r, u_r) = (1.5, 1) \quad \lambda = 0.001.$$

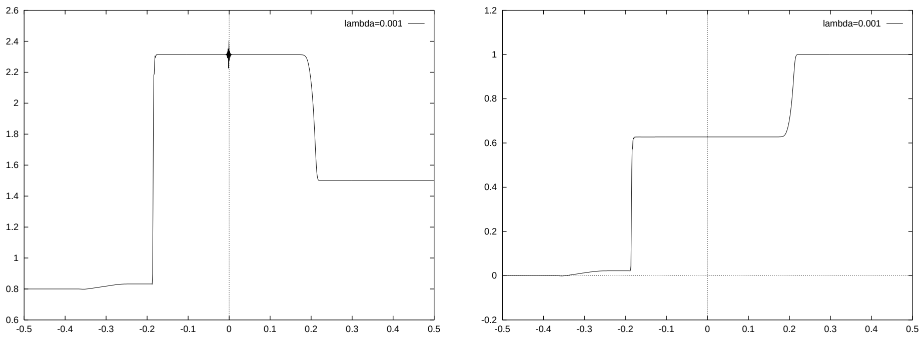

FIG. 2.11 – *Nonclassical shock wave (van der Waals) (a) volume (b) speed*

Figure 2.12 shows the dependence of this solution with respect to the capillarity coefficient.

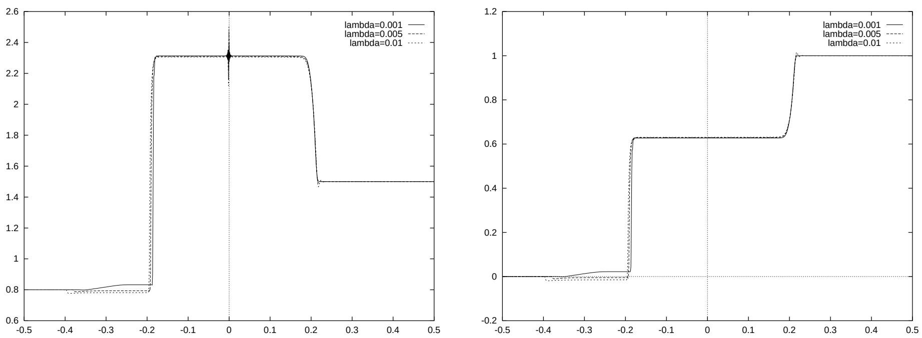

FIG. 2.12 – *Several values of λ (van der Waals law) (a) volume (b) speed*

Figure 2.13 illustrates that shocks are truly non-classical shock: the straight line connecting the two states cuts the graph of the pressure.

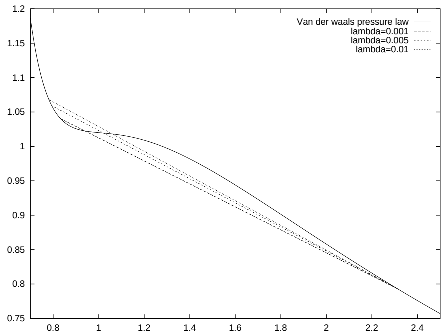

FIG. 2.13 – *Non classical shock on the pressure plot*

2.5 Kinetic functions

To characterize the dynamics of subsonic phase boundaries and nonclassical shock waves, we now determine numerically the kinetic functions associated with the schemes introduced in Section 2.3. Precisely, we compute the right-hand value of the volume as a function of the left-hand state. We also compute the propagation speed of the phase boundaries or the nonclassical shocks. As become clear on the plots below, the kinetic functions depend on the ratio viscosity/capillarity (more precisely on μ^2/λ) and on the specific scheme under consideration as well.

Together with the kinetic functions, we plot also some extremal curves which are known to limit the range of the kinetic functions, especially the curve along which the entropy dissipation vanishes and the (Maxwell) curve along which the shock speed vanishes.

Kinetic functions for several capillarity coefficients

Figure 2.14 concerns the cubic pressure law (2.38). In all of the runs we used the following initial data

$$(v_l, u_l) = (3, 0), \quad v_r = 5.$$

The viscosity is taken to be $\mu = 2$ and the mesh contains 1200 points. We computed three distinct kinetic curves associated with different values of the capillarity coefficient. A point on the curve is associated with a propagating phase boundary found for some given initial velocity u_r. Each curve is obtained by letting u_r describe the interval $[0.6, 5.]$.

Figure 2.15 concerns van der Waals pressure-law (2.40). We choose here the initial data

$$(v_l, u_l) = (0.666, -1.8), \quad v_r = 0.75.$$

The viscosity is $\mu = 0.1$ and the mesh contains 1200 points. Here each curve is obtained for u_r describing the interval $[-1.8, 0.]$.

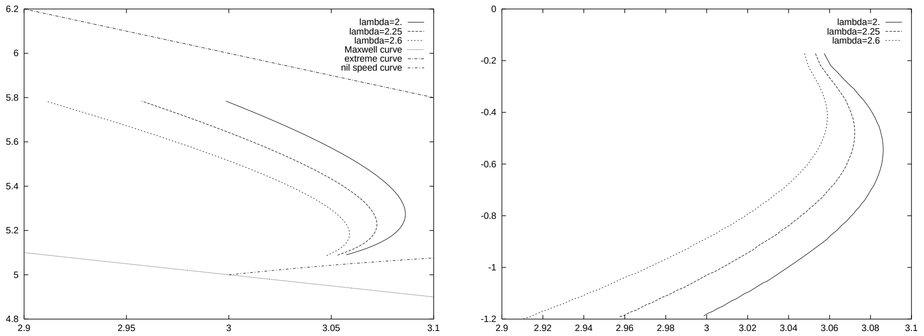

FIG. 2.14 – *Several values of λ (cubic law)* *(a) volume* *(b) speed*

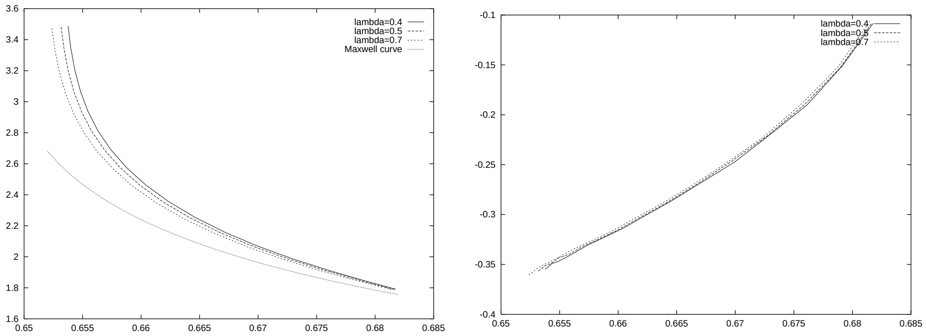

FIG. 2.15 – *Several values of λ (van der Waals)* *(a) volume* *(b) speed*

Kinetic functions for several schemes

Next we study the effect of the order of accuracy of the scheme on the kinetic curves (Cubic law). For definiteness we fix the following constants:

$$(v_l, u_l) = (3, 0), \quad v_r = 5, \quad \lambda = 1.5, \quad \beta = 1.$$

Similarly as above, we compute kinetic curves by letting v_l describe the interval $[0.6, 10]$. Figure 2.16 displays the results for several choices of the parameters K et L.

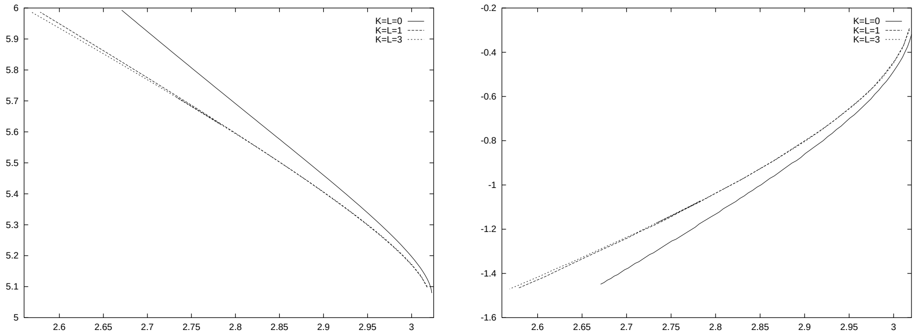

FIG. 2.16 – *Several values of K and L (cubic law)* *(a) volume* *(b) speed*

2.6 Conclusions

In this paper, we dealt with propagating phase boundaries modeled by the isothermal model of compressible fluids governed by van der Waals-type equations of state.

1) Capillarity effects were taken into account by using the derivative of the specific volume, v_y, as an independant variable. We have introduced a new class of entropy conservative numerical schemes in the sense of Tadmor. These schemes are endowed with nonlinear stability properties: the total energy –which plays the role of a mathematical entropy in the sense of Lax– is decreasing in time, the decay being due to the viscosity only. Hence we were able to reproduce at the discrete level an important property satisfied by the physical model. We also established that the proposed schemes may have sufficiently high order of accuracy, so that the corresponding equivalent equation coincide with the continuous model up to the order $O(h^3)$ at least.

2) We demonstrated the existence of propagating subsonic phase boundaries and of nonclassical undercompressive shock waves for the van der Waals model. The proposed schemes successfully computed these undercompressive waves. Non-steady subsonic boundaries (not consistent with the standard Maxwell construction) are not found in thermodynamics textbooks nor in the Riemann solvers derived in [119] and [62]. Subsonic phase boundaries are induced precisely by the capillarity effects, kept in balance with the viscosity effects. The former create oscillations while the latter introduce dissipation in the equations. The small-scale effects are dominant in determining the dynamics of undercompressive waves. The observed oscillations are entirely expected and standard numerical methods such as the TVD schemes are clearly not adapted.

3) Following [65], kinetic relations associated with the schemes were numerically determined. These curves depend on the order of the schemes and also on the relative strength of the viscosity and capillarity coefficients. The kinetic curves allow us to compare the properties of the schemes.

4) The equivalent equation provides only some indication of the behavior of the schemes. For sufficiently small propagation speed a good agreement have been observed. But, for waves with sufficiently large speed or large amplitude, some discrepancy does arise. This is a central difficulty with dissipation-sensitive problem.

5) We have encountered some specific numerical difficulties with the van der Waals pressure law, due to its shape. In normalized units at least and in the range of interest near the inflection point, the subsonic phase boundaries connect states having very different characteristic speeds. Indeed, the sound speed in the liquid tends to infinity as the volume v tends to zero, while for large v in the vapor phase the sound speeds tends to zero. This appears clearly in Figure 2.1b. (This behavior is not found for the cubic law in Figure 2.1a.) This has dramatic consequences from the numerical standpoint. A small error on the liquid state corresponds to a large error on the vapor one. As a consequence, finding numerically the range of left and right-hand states for which subsonic phase boundaries exist has been particularly challenging. Then, nonclassical shock waves and subsonic phase boundaries may also be particularly delicate to observe in practical situations.

6) Finally, we point out that the extension of the present approach for the viscosity-capillarity model of van der Waals fluids based on three conservation laws

(mass, momemtum, energy) should be possible. Many of the properties derived here generalize indeed immediately to this model.

Chapitre 3

Computing undercompressive waves with the random choice scheme. I. Nonclassical shock waves

Ce chapitre fait l'objet d'un article soumis pour publication : voir [26]

Computing undercompressive waves with the random choice scheme. I. Nonclassical shock waves

C. Chalons[1] and P.G. LeFloch[2]

Résumé

For several nonlinear hyperbolic models of interest we investigate the stability and large-time behavior of undercompressive shock waves characterized by a kinetic relation. The latter are considered as interfaces between two materials with distinct constitutive relations. In this first part we study nonclassical entropy solutions to scalar conservation laws with concave-convex flux-function and a non-genuinely nonlinear, strictly hyperbolic model of two conservation laws arising in nonlinear elastodynamics. We use Glimm's random choice scheme but we replace the classical Riemann solver with the *nonclassical* one described recently in [88, 95]. Our numerical experiments demonstrate the great interest of using the random choice scheme for computing nonclassical shock waves which are known to be very sensitive to dissipation and dispersion mechanisms. In this paper, we study carefully various issues related to nonclassical shocks and their stability and instability under perturbations.

3.1 Introduction

We are interested in computing weak solutions of the initial-value problem for one-dimensional, nonlinear systems of conservation laws of the form

$$\partial_t u + \partial_x f(u) = 0, \quad u(x,t) \in I\!R^N,\, x \in I\!R,\, t > 0, \tag{3.1}$$

$$u(x,0) = u_0(x), \qquad x \in I\!R. \tag{3.2}$$

Such systems arise in a broad spectrum of problems in compressible fluid dynamics, nonlinear elastodynamics, etc. Due to the nonlinearity $f : I\!R^N \to I\!R^N$, solutions are generally discontinuous and it is well-known that weak solutions (in the integral sense) are not uniquely determined by their initial data $u_0 : I\!R \to I\!R^N$, and must be constrained by an *entropy inequality* of the form

$$\partial_t U(u) + \partial_x F(u) \leq 0 \tag{3.3}$$

in the weak sense, where (U, F) denotes an entropy-entropy flux pair for the system (3.1), satisfying by definition $DF^T = DU^T Df$. Here, we will focus on solutions containing *undercompressible shock waves:* the number of characteristics (wave modes) impinging on the discontinuity is less than what is usually required for the linearized stability. Undercompressive waves turn out to be not uniquely determined by (3.3). However, under some assumptions to be specified in several examples below, the uniqueness of the *entropy solution* of the problem (3.1)-(3.3) is ensured when a *kinetic*

1. O.N.E.R.A., B.P. 72, 29 avenue de la Division Leclerc, 92322 Châtillon Cedex, France. E-mail: chalons@onera.fr.

2. Centre de Mathématiques Appliquées & Centre National de la Recherche Scientifique, U.M.R. 7641, Ecole Polytechnique, 91128 Palaiseau Cedex, France. E-mail: lefloch@cmap.polytechnique.fr.

relation is added along each undercompressive discontinuity connecting a left-hand state u_- to a right-hand state u_+:

$$u_+ = \varphi^\flat(u_-). \tag{3.4}$$

The *kinetic function* $\varphi^\flat : I\!R^N \to I\!R^N$ is a Lipschitz continuous mapping satisfying the basic conditions

$$-\lambda^\flat(u_-)\,(\varphi^\flat(u_-) - u_-) + f(\varphi^\flat(u_-)) - f(u_-) = 0 \tag{3.5}$$

and

$$-\lambda^\flat(u_-)\,(U(\varphi^\flat(u_-)) - U(u_-)) + F(\varphi^\flat(u_-)) - F(u_-) \leq 0, \tag{3.6}$$

where $\lambda^\flat(u_-)$ denotes the speed of propagation. The nonclassical shocks are also refer as *phase transition* or *interfaces*. For a complete discussion of the notion of kinetic relation we refer to the recent monograph [89] and the references therein.

Our general aim is to investigate numerically the stability and large-time behavior of undercompressive waves using, as a tool, Glimm's random choice scheme [55]. In the present paper, we focus attention on *nonclassical shocks* of strictly hyperbolic systems which fail to be globally genuinely nonlinear in the sense of Lax. That is, the Jacobian matrix $Df(u)$ admits real and distinct eigenvalues $\lambda_j(u)$ and independent eigenvectors $r_j(u)$ $(1 \leq j \leq N)$ and for one wave family at least the product $\nabla\lambda_j(u) \cdot r_j(u)$ does not keep a constant sign. Specifically, we consider the scalar conservation law with cubic flux-function (Section 3.2) and an hyperbolic model arising in nonlinear elastodynamics (Section 3.3). In the second part [27] we will study phase boundaries of an hyperbolic-elliptic model arising in phase dynamics.

Let us recall some basic features of the random choice method. Glimm's scheme [55, 100, 30] is based on an equidistributed sequence $(a_n)_{n=1,2,\cdots}$ of values in the interval $(0,1)$ satisfying, by definition, for each interal $J \subset (0,1)$

$$\frac{1}{m}\,\mathrm{card}\{n\,/1 \leq n \leq m \;\text{ and }\; a_n \in J\} \to \mathrm{meas}(J)$$

as $m \to \infty$. The scheme is based on solving *Riemann problems* corresponding to the piecewise constant initial data:

$$u_0(x) = \begin{cases} u_L, & x < 0, \\ u_R, & x > 0, \end{cases} \tag{3.7}$$

where u_L and u_R are constant states. The Riemann solution has a rather simple form: explicitely, it is made of several shock waves and rarefaction waves separated by constant states. Glimm's scheme proceeds as follows. First, the initial data u_0 in (3.2) is replaced by a piecewise constant approximation $u_0^{\Delta x}$ where $\Delta x > 0$ represents the constant mesh size of a regular mesh $x_k = k\,\Delta x$ for $k = \cdots, -1, 0, 1, \cdots$ At each initial discontinuity a Riemann problem is solved locally, using a classical or a nonclassical Riemann solver. (Such solvers are derived in [88].) At some sufficiently small time Δt satisfying the stability condition

$$\Delta t\,\sup|\lambda_j(u)| < \frac{\Delta x}{2}, \tag{3.8}$$

a value is picked up "randomly" within each local Riemann solutions. This provides us with the new approximation at time Δt. The construction is continued inductively in time until the approximation $u^{\Delta x, \Delta t} = u^{\Delta x, \Delta t}(x, t)$ is determined for all times.

Throughout the present paper, following a proposal by Collela [30] we use van der Corput random sequence defined by

$$a_n = \sum_{k=0}^{m} i_k 2^{-(k+1)},$$

where $n = \sum_{k=0}^{m} i_k 2^k$, $i_k = 0, 1$, denotes the binary expansion of the integers $n = 1, 2, \ldots$. The first few elements of this sequence are

$$\begin{array}{llll} 1 = 1_2, & a_1 = 0.5, & 2 = 10_2, & a_2 = 0.25, \\ 3 = 11_2, & a_3 = 0.75, & 4 = 100_2, & a_4 = 0.125, \\ 5 = 101_2, & a_5 = 0.625, & 6 = 110_2, & a_6 = 0.375, \\ 7 = 111_2, & a_7 = 0.875, & 8 = 1000_2, & a_8 = 0.0625. \end{array}$$

This sequence is well distributed within the interval. For instance one can prove that

$$\begin{array}{ll} a_i < 0.5, & i \text{ even}, \\ a_i > 0.5, & i \text{ odd}. \end{array}$$

Excellent numerical results in the smooth parts of the solution are obtained with this sequence, as we will illustrate below (Test 1).

3.2 Conservation Law with Cubic Flux

To begin with, we consider the nonlinear conservation law

$$\partial_t u + \partial_x u^3 = 0, \quad u(x,t) \in I\!R, \tag{3.9}$$

which is the simplest example of a nonlinear hyperbolic equation which fails to be globally genuinely nonlinear. Following [69, 89] we consider solutions satisfying the conservation law (3.9) in the integral sense, the initial condition

$$u(x, 0) = u_0(x), \qquad x \in I\!R, \tag{3.10}$$

the entropy inequality

$$\partial_t u^2 + \frac{3}{2}\,\partial_x u^4 \leq 0 \tag{3.11}$$

in the integral sense and (for definiteness in our numerical investigations) a kinetic function in the form

$$\varphi^\flat(u) = -\beta\, u, \quad \beta \in [1/2, 1). \tag{3.12}$$

The parameter β is a measure of "how far the solution" is from the classical regime. Precisely, as $\beta \to 1/2$, the Riemann solution converges to the classical entropy solution selected by Oleinik entropy inequalities. The choice $\beta = 1$ is not truly allowed in the general theory of [88], but is a limiting case, associated with the "maximal kinetic relation": The correponding shock waves have zero-entropy dissipation and achieve the largest propagation speed among all admissible nonclassical shocks.

Observe that (3.12) satisfies the standard assumption made on kinetic functions in [89]:

$$\begin{aligned} -u < \varphi^\flat(u) \leq -u/2, \quad u > 0, \\ -u/2 \leq \varphi^\flat(u) < -u, \quad u < 0. \end{aligned}$$

We also set

$$\varphi^\sharp(u) = -u - \varphi^\flat(u).$$

Given constant states u_L and u_R, the *nonclassical Riemann solver* associated with (3.9)-(3.12) and the initial data

$$u_0(x) = \begin{cases} u_L, & x < 0, \\ u_R, & x > 0, \end{cases}$$

is given as follows, assuming for definiteness that $u_L > 0$:

(1) If $u_R \geq u_L$, the solution is a rarefaction wave connecting u_L to u_R.
(2) If $u_R \in [\varphi^\sharp(u_L), u_L)$, the solution is a classical shock wave connecting u_L to u_R.
(3) If $u_R \in (\varphi^\flat(u_L), \varphi^\sharp(u_L))$, the solution contains a nonclassical shock connecting u_L to $\varphi^\flat(u_L)$, followed by a classical shock connecting $\varphi^\flat(u_L)$ to u_R.
(4) If $u_R \leq \varphi^\flat(u_L)$, the solution contains a nonclassical shock connecting u_L to $\varphi^\flat(u_L)$, followed by a rarefaction connecting $\varphi^\flat(u_L)$ to u_R.

Test 1 : Van der Corput sequence

We start by comparing the efficiency of van der Corput sequence versus an "arbitrary" random sequence. Consider the Riemann data

$$u_0(x) = \begin{cases} 2, & x < 0, \\ 4, & x > 0. \end{cases}$$

The solution is made of a single rarefaction wave and the kinetic function is here irrelevant. The mesh contains 300 points and the solution is plotted on Figure 3.1 at the time $t = 0.007$. We observe that van der corput sequence gives much more accurate and regular results than an arbitrary random sequence.

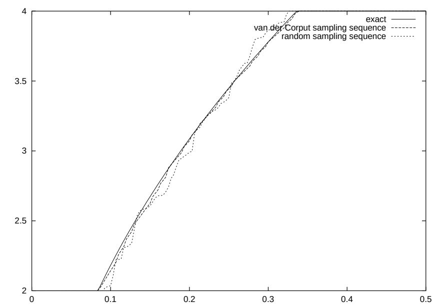

FIG. 3.1 – *Two random sequences*

Test 2 : The Riemann solver

Next, we display the four qualitative behaviors of the Riemann solution for the problem (3.9)-(3.12), with

$$u_0(x) = \begin{cases} u_l, & x < -0.375, \\ u_r, & x > -0.375, \end{cases}$$

and the parameter value $\beta = 2/3$. The mesh contains 300 points and numerical solutions are plotted on Figures 3.2 and 3.3. We observe that the numerical solutions fully agree with the above nonclassical Riemann solver.

• Figure 3.2 (a): *A single classical shock wave.* We plot here the numerical solution at the time $t = 0.05$, corresponding to the data $u_l = 2$ and $u_r = 1$.

• Figure 3.2 (b): *A single rarefaction wave.* We plot the solution obtained with $u_l = 1$ and $u_r = 2$, at the time $t = 0.05$.

• Figure 3.3 (c): *A nonclassical shock wave followed by a classical shock wave.* We plot the numerical solution at the time $t = 0.6$, corresponding to the Riemann data $u_l = 1$ and $u_r = -0.4$.

• Figure 3.3 (d): *A nonclassical shock wave followed by a rarefaction wave.* We plot the numerical solution at the time $t = 0.05$, corresponding to $u_l = 1$ and $u_r = -2$.

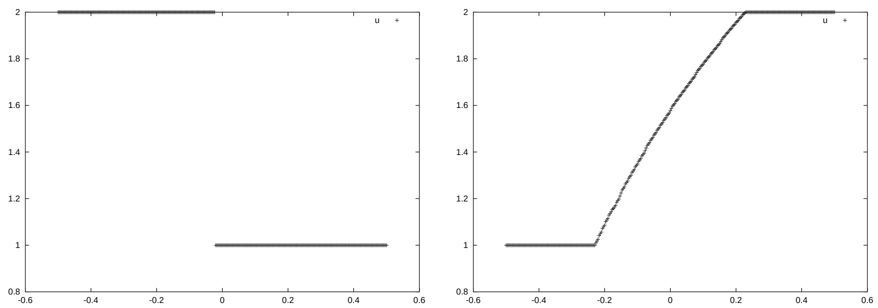

FIG. 3.2 – *(a) and (b)* *Typical solutions : single classical wave*

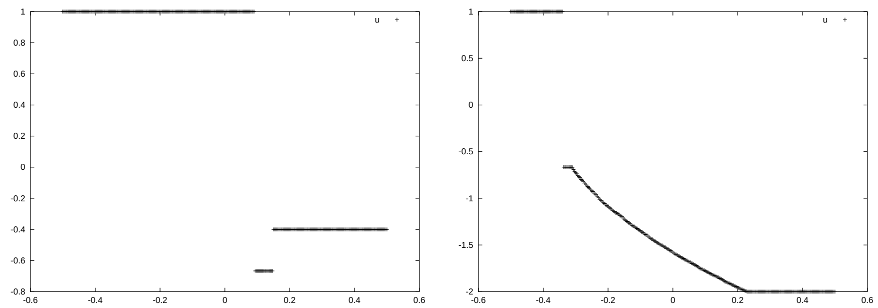

FIG. 3.3 – *(c) and (d)* *Typical solutions : classical wave and nonclassical wave*

Observe that shock waves are represented by sharp discontinuities which is a main feature of the random choice scheme: all values taken by the numerical solutions are actual values taken by the exact solutions. The location of the discontinuities however depend on the random sequence. By constrast, finite differences schemes generate intermediate points within propagating discontinuity and a discrete numerical shock profile.

Test 3 : Splitting of shock waves

We present three related approaches illustrating an important numerical difficulty encountered with nonclassical shock waves: the nonclassical Riemann solver depend continuously upon its end states in the integral norm but *not* in the pointwise sense. Because of this lack of continuity, some sort of "spikes" are sometimes observed in the numerical solution. The Riemann problem associated with

$$u_0(x) = \begin{cases} u_L, & x < -0.375, \\ \varphi^\sharp(u_L), & x > -0.375, \end{cases}$$

can be solved with a single classical shock wave.

• Figure 3.4: *Small perturbations of a classical shock.* If we perturb the above initial data by an arbitrarily small amount $\epsilon > 0$, specifically

$$u_0^\epsilon(x) = \begin{cases} u_L, & x < -0.375, \\ \varphi^\sharp(u_L) - \epsilon, & x > -0.375, \end{cases}$$

the solution now contains a *two-wave, non-monotone pattern,* precisely a nonclassical shock connecting u_L to $\varphi^\flat(u_L)$ followed with a classical shock connecting $\varphi^\flat(u_L)$ to $\varphi^\sharp(u_L) - \epsilon$.

On Figure 3.4, using $\epsilon = 0.01$, $\epsilon = 0.05$, and $\epsilon = 0.1$ successively, we plotted the numerical Riemann solution, for a mesh containing 300 points, at the time $t = 0.18$, and with $\beta = 2/3$. We observe that, as ϵ diminishes, the wave speeds of the two shocks become closer and the solution exhibits a *spike,* while the intermediate value is determined *exactly* by Glimm's scheme. This feature is *not* shared by finite differences scheme and the presence of spikes may be an important source of error in numerical computations performed with finite differences schemes.

• Figure 3.5: *Interaction of a classical shock and a rarefaction.* Second, consider the initial data

$$u_0(x) = \begin{cases} u_L, & x < -0.40, \\ \varphi^\sharp(u_L) + \epsilon, & -0.40 < x < -0.38, \\ \varphi^\sharp(u_L) - \epsilon, & x > -0.38. \end{cases}$$

The solution contains, for small times, a classical shock wave plus a rarefaction wave (with arbitrary small strength) which interact in finite time and generate a nonclassical shock. In Figure 3.5, for $\epsilon = 0.01$, $\epsilon = 0.05$, and $\epsilon = 0.1$ successively, we plot the numerical solution obtained with 300 points, at the time $t = 0.2$ with $\beta = 2/3$. We note that the total variation of the solution *after* the interaction is much *greater* than the total variation of the initial data:

$$TV(u(t)) > TV(u_0).$$

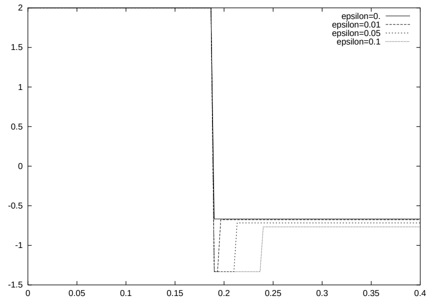

FIG. 3.4 – *Small perturbations of a classical shock*

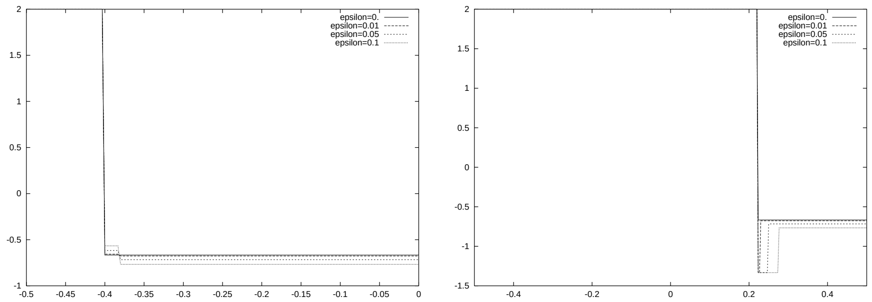

FIG. 3.5 – *Interaction of two classical shocks ($t = 0$ and $t = 0.2$)*

• Figures 3.6-3.7: *Large perturbations.*

This last test is different in nature. We demonstrate here that nonclassical shocks (as well as classical shocks) are stable under large perturbations, in the following sense: when a large perturbation is added at the initial time within a classical shock wave (Figure 3.6) or a nonclassical shock wave (Figure 3.7), the solution *converges asymptotically in time* to the unperturbed initial wave. More precisely, this is true up to a possible *shift* in the location of the discontinuity. In the numerical experiments, the amplitude of the perturbation lies between -0.42 et -0.35. At the time under consideration in Figure 3.6, the unperturbed shock should be located at the point -0.42.

Test 4 : Periodic data

We begin with the initial condition

$$u_0(x) = -0.5 + \sum_{i=0}^{4} \cos(5ix), \quad x \in [-0.5, 0.5]$$

with periodic boundary conditions

$$u(-0.5, t) = u(0.5, t).$$

The averaged of the solution over one period is constant and in our experimentrs equals about 0.61826856. Throughout we use a mesh containing 300 points.

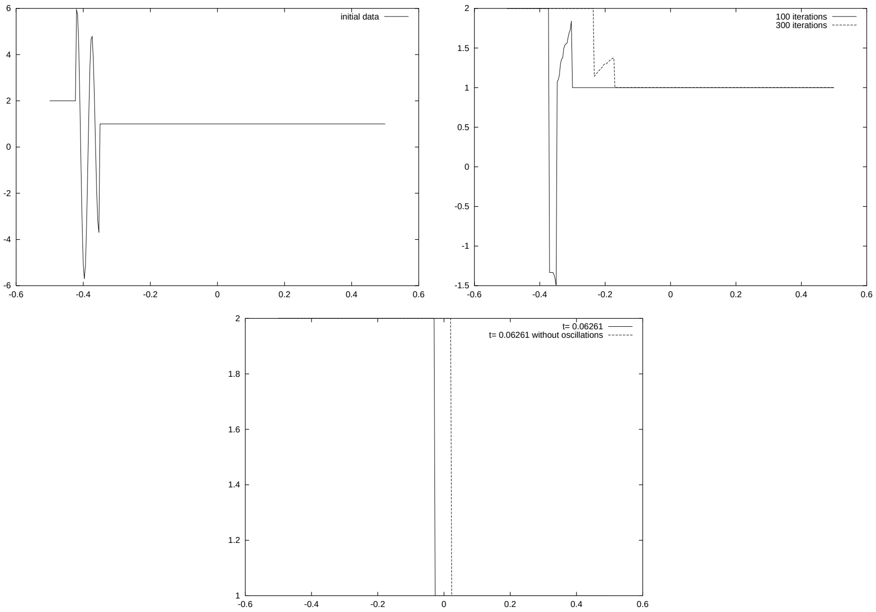

FIG. 3.6 – *Large perturbations (classical shock)*

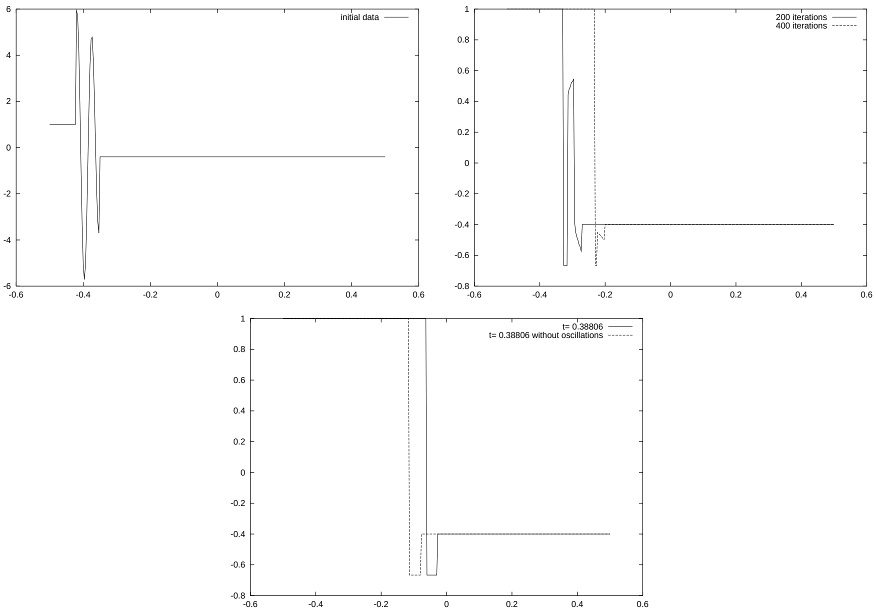

FIG. 3.7 – *Large perturbations (nonclassical shock + classical shock)*

• Figure 3.8: *Classical Riemann solution with positive average.* To begin with, we use the classical kinetic function corresponding to $\beta = 1/2$. On Figure 3.8, we see that the number of phase transitions is *decreasing* in time and that the numerical solution converges to a constant value. In the last figure we recognize the so-called *N-wave* which is well-known for scalar conservation laws with convex flux. Since the average value is positive, for sufficiently large times the solution takes positive values only, and the convex part of the flux only is relevant.

• Figure 3.9: *Nonclassical Riemann solution with positive average.* We use now $\beta = 2/3$. We observe that the number of phase transitions is still *decreasing* and that the solution converges to a constant state. Again, we recognize an *N-wave.* Since the average value is positive, for sufficiently large times the solution takes positive values only, the convex part of the flux only is relevant and the nonclassical solver coincides with the classical solver in this region, and the solution is eventually entirely classical. Observe also that Figures 3.8 and 3.9 are qualitatively similar, the maximal value in the asymptotic nonclassical solution being larger then the one in the classical solution.

• Figure 3.10: *Maximal nonclassical Riemann solution with positive average.* Finally, we use $\beta = 1$ and observe that the numerical solution no longer converges to a constant state. Instead, the number of phase transitions remains *constant* and the solution converges to a piecewise constant function made of *several phase transitions* with coinciding propagation speeds.

• Figure 3.11-3.12: *The case $\beta = 1 - \epsilon$ with ϵ small.* Taking $\epsilon = 0.01$, we observe that the qualitative behavior of the solution is quite analogous to Figure 3.9. However, the convergence is *very slow,* and there is a transient regime during which the solution may appear to *behave like* the one in Figure 3.10. The number of phase transitions is *very slowly decreasing* but the solution does eventually converge to the average value of the initial data.

• Figure 3.13: *Classical Riemann solution with zero average.* In the remaining tests in this series, we consider the periodic initial condition with period 1/2:

$$u_0(x) = \cos(4\pi x),$$

which has *zero average.* Throughout, the mesh contains 300 points and to begin with we take $\beta = 1/2$. We observe here that the number of phase transitions is *decreasing* and that the solution *converges to zero* in a pointwise sense. More precisely, a new feature must be pointed out: there is a transient regime during which the solution exhibits what we will call a *"double N-wave pattern",* which is somewhat made of two standard well-known N-wave profiles with opposite monotonicity. Notice that the period 1/2 is eventually lost as the numerical solution converges eventually to a standard N-wave profile. See also Test 5 below for further mesh refinements.

• Figure 3.14: *Nonclassical Riemann solution with zero average.* Next, we use $\beta = 2/3$. The number of phase transitions is *decreasing* and the solution converges

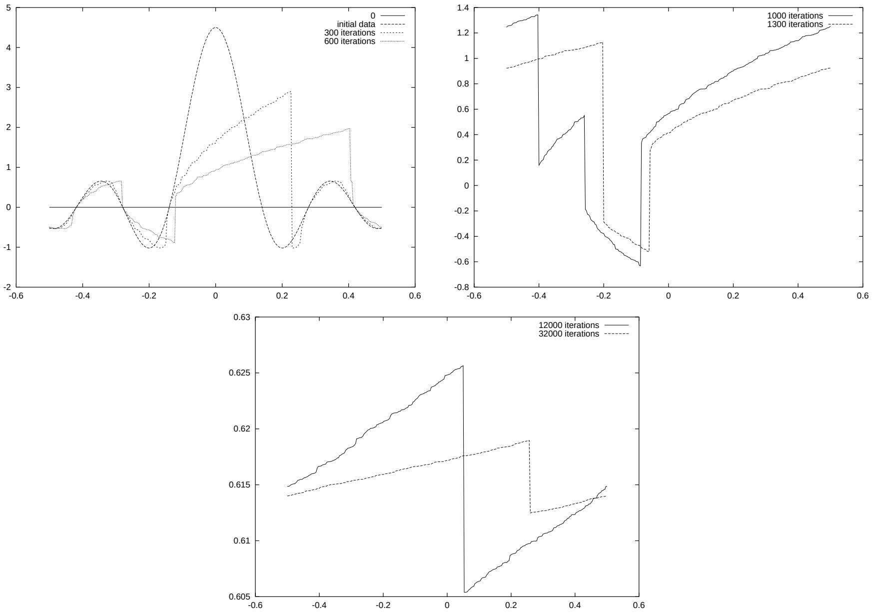

FIG. 3.8 – *Classical Riemann solution with positive average*

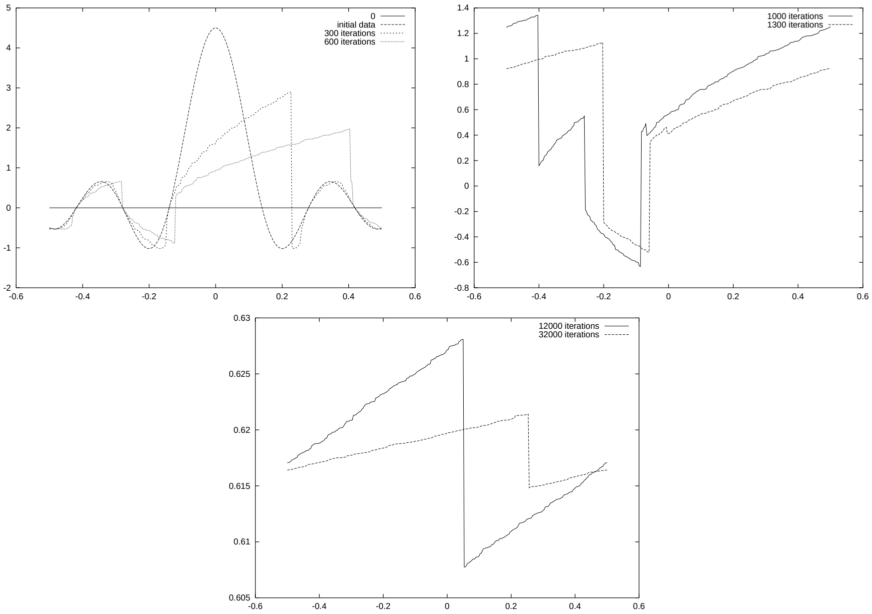

FIG. 3.9 – *Nonclassical Riemann solution with positive average*

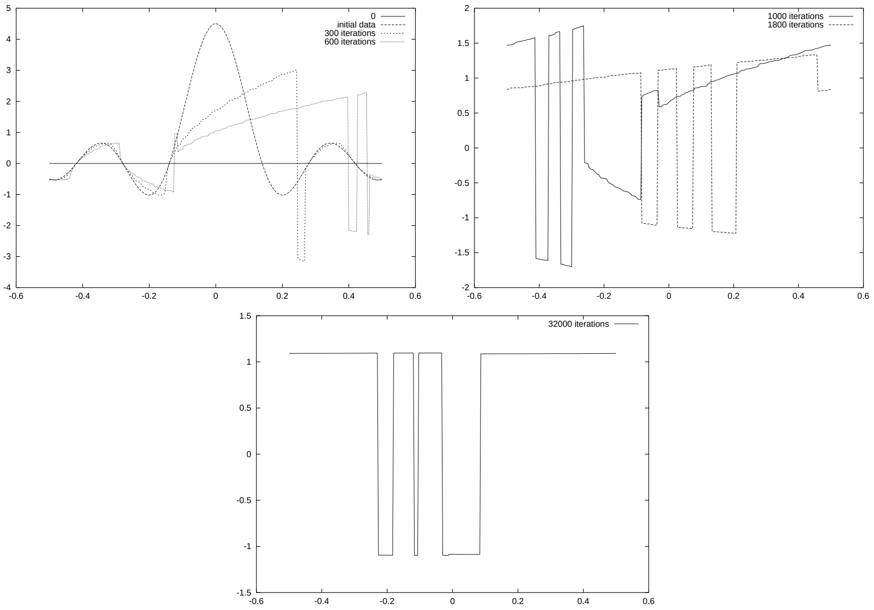

FIG. 3.10 – *Maximal nonclassical Riemann solution with positive average*

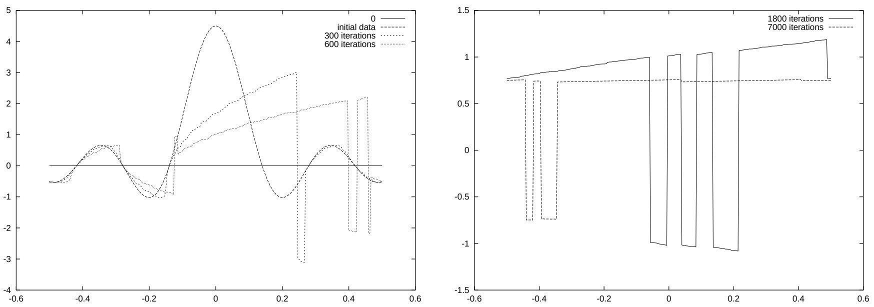

FIG. 3.11 – *The case $\beta = 1 - \epsilon$*

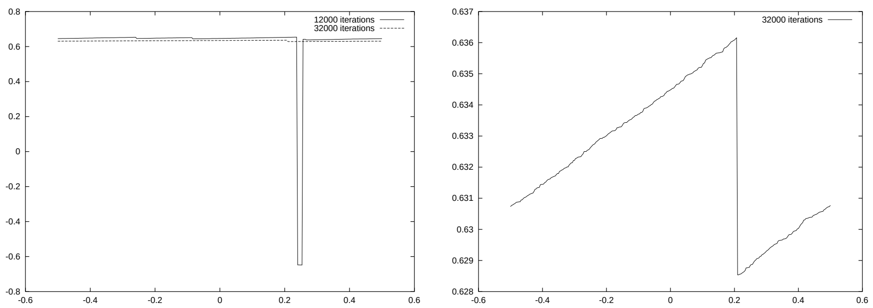

FIG. 3.12 – *The case $\beta = 1 - \epsilon$*

to zero. Again, we observe a double N-wave which eventually disappears and get transformed into a more familiar single N-wave.

• Figure 3.15: *Maximal nonclassical Riemann solution with zero average.* Finally, we use $\beta = 1$ and notice that the number of phase transitions is now *constant* in time. The numerical solution converges to a function made of several phase transitions with coinciding propagation speeds.

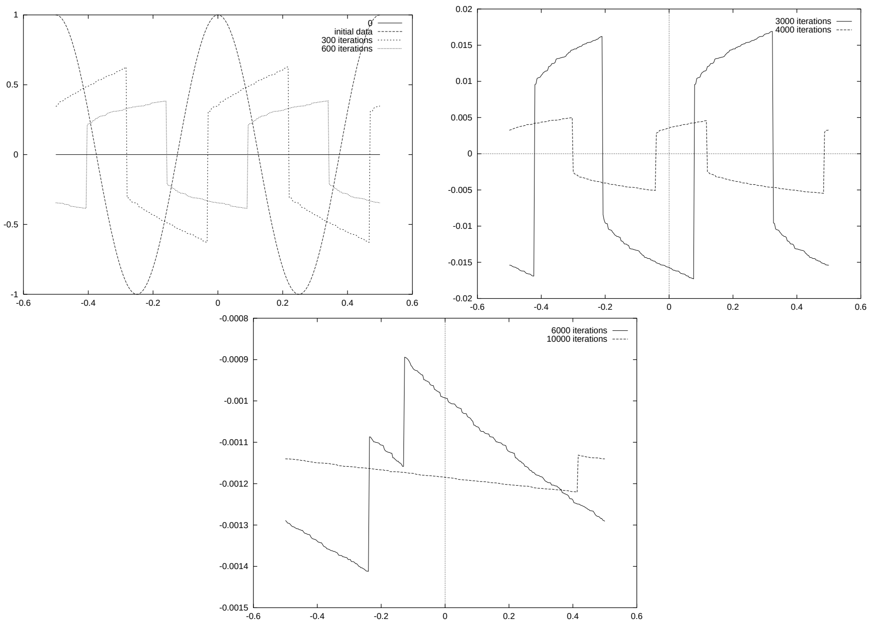

FIG. 3.13 – *Classical Riemann solution with zero average*

Test 5 : Mesh refinements

Two mesh refinement experiments are described here. First, using $\beta = 2/3$ and

$$u_0(x) = \begin{cases} 3, & x < -0.375, \\ -3, & x > -0.375, \end{cases}$$

on Figure 3.16 we plotted the numerical solution corresponding to finer and finer mesh lengths, at the time $t = 0.023$. As expected, the numerical solution approaches the exact solution determined by our nonclassical Riemann solver. This illustrates the convergence property of the random choice scheme to nonclassical entropy solutions.

Second, we plotted on Figure 3.17 the solution with the initial periodic data

$$u_0(x) = \cos(4\pi\, x),$$

for several mesh lengths. As expected, we observe that the numerical solution gets closer and closer to the exact solution (that is 0) pointwisely as the mesh is refined. We do observe a convergence to the double N-wave profile discovered in Test 4.

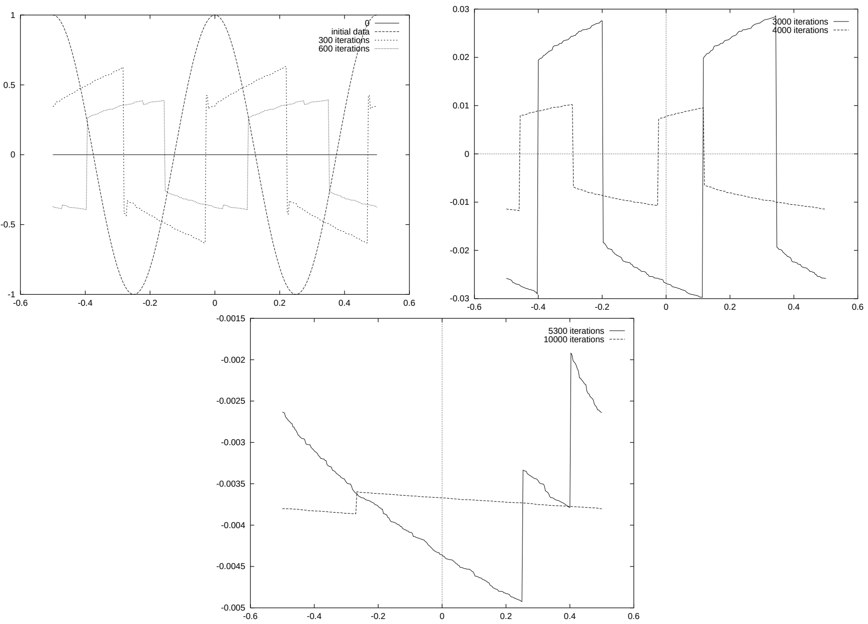

FIG. 3.14 – *Nonclassical Riemann solution with zero average*

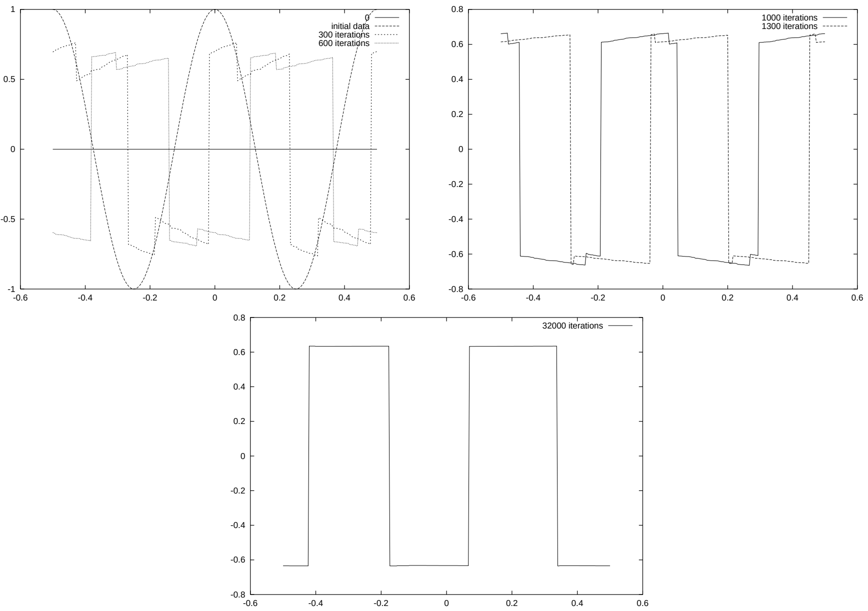

FIG. 3.15 – *Maximal nonclassical Riemann solution with zero average*

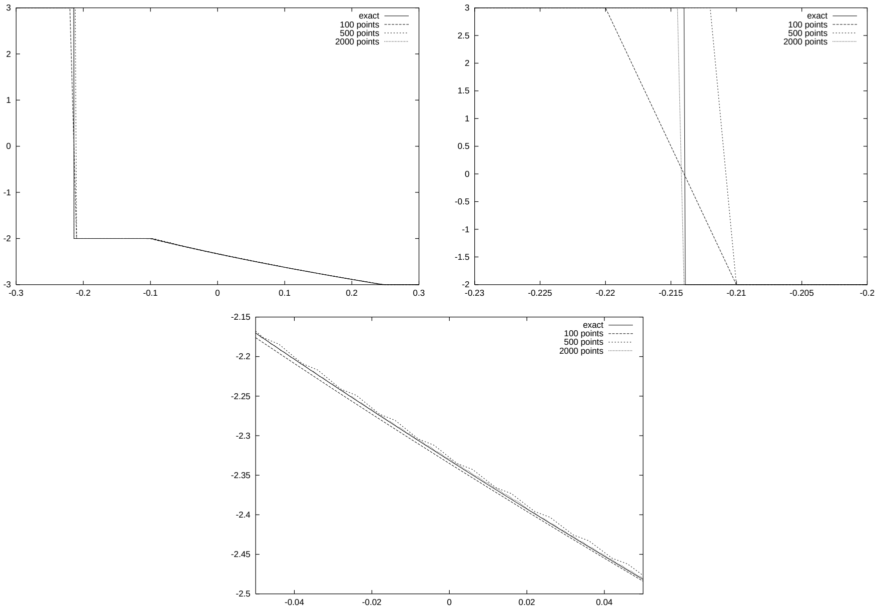

FIG. 3.16 – *Mesh refinement (global solution and zooms)*

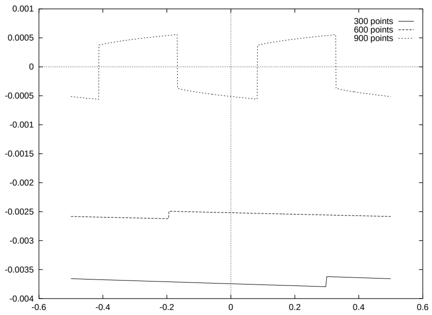

FIG. 3.17 – *Mesh refinement (periodic data)*

3.3 Hyperbolic Model of Nonlinear Elastodynamics

We turn to the following model of two conservation laws for the mass and momentum of some nonlinear material, with prescribed initial data (v_0, w_0),

$$\begin{aligned} &\partial_t v - \partial_x \sigma(w) = 0, \qquad x \in I\!R,\ t > 0, \\ &\partial_t w - \partial_x v = 0, \\ &(v(x,0), w(x,0)) = (v_0(x), w_0(x)), \qquad x \in I\!R, \end{aligned} \tag{3.13}$$

where v and w represent the velocity and the deformation gradient of the material, respectively. The strain-stress function σ is taken to be

$$\sigma(w) = w^3 + m\,w, \tag{3.14}$$

with $m = 2$ for the numerical experiments below (excepted when mentioned otherwise). The Cauchy problem (3.13) is supplemented with the following *kinetic function* for each undercompressive nonclassical shock

$$\varphi^\beta(w) = -\beta\,w, \tag{3.15}$$

where $\beta \in [1/2, 1)$ is a parameter. We refer to [95] for a detailled discussion of the nonclassical Riemann solver corresponding to this problem. We simply recall here that when $\beta = 1/2$ the nonclassical Riemann solver actually coincides with the classical solver characterized, for instance, by Liu entropy criterion. When $\beta > 1/2$, the Riemann solution may contain shock waves which violate Liu criterion, are undercompressive, and satisfy the *kinetic relation*

$$\begin{aligned} w_+ &= \varphi^\beta(w_-) \quad \text{for shock with positive speed,} \\ w_- &= \varphi^\beta(w_+) \quad \text{for shock with negative speed.} \end{aligned} \tag{3.16}$$

As was noticed with scalar conservation laws, the "maximal kinetic relation" corresponding to $\beta = 1$ is not truly allowed in the general theory of [88] but can be regarded as an interesting limiting case. The corresponding shock waves have zero-entropy dissipation and achieve the largest propagation speed (in modulus) among all admissible nonclassical shocks.

Test 6 : Splitting of shock waves

• Figure 3.18: *Small perturbations of a classical shock.* We consider the Riemann initial data

$$(v_0(x), w_0(x)) = \begin{cases} (1, 1+\epsilon), & x < 0.5, \\ (-11, -3), & x > 0.5, \end{cases}$$

and we plot the corresponding numerical solution for several values of ϵ. When $\epsilon = 0$, the exact solution of the Riemann problem is a single classical shock. We see that the structure for arbitrary small values ϵ is entirely different from the one with $\epsilon = 0$. In particular, the L^∞-norm of the solution is *not continuous* as $\epsilon \to 0$. We use a mesh with 600 points by unit interval, $t = 0.4$, and $\beta = 2/3$.

• Figures 3.19-3.20: *Interaction of two classical waves.* As in the scalar case, two incoming waves (of the same family) may interact and can generate a nonclassical

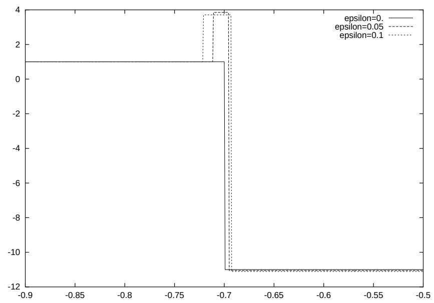

FIG. 3.18 – *Small perturbations of a classical shock (v and w components)*

shock wave. The total variation of the solution increases at the interaction and the w-component becomes non-monotone after the interaction. Of course, it is well-known that, for hyperbolic systems of conservation laws, the total variation may increase at interactions; however, the increase here is not of the order of the product of the strengths of the incoming waves. Specifically, we consider the initial data

$$(v_0(x), w_0(x)) = \begin{cases} (1, 1.2), & x < -0.1, \\ (-1.72814176, -0.337385342), & -0.1 < x < 0.1, \\ (-11, -3), & x > 0.1, \end{cases}$$

and we plot the solutions after various iteration numbers. We use a mesh with 300 points per unit interval.

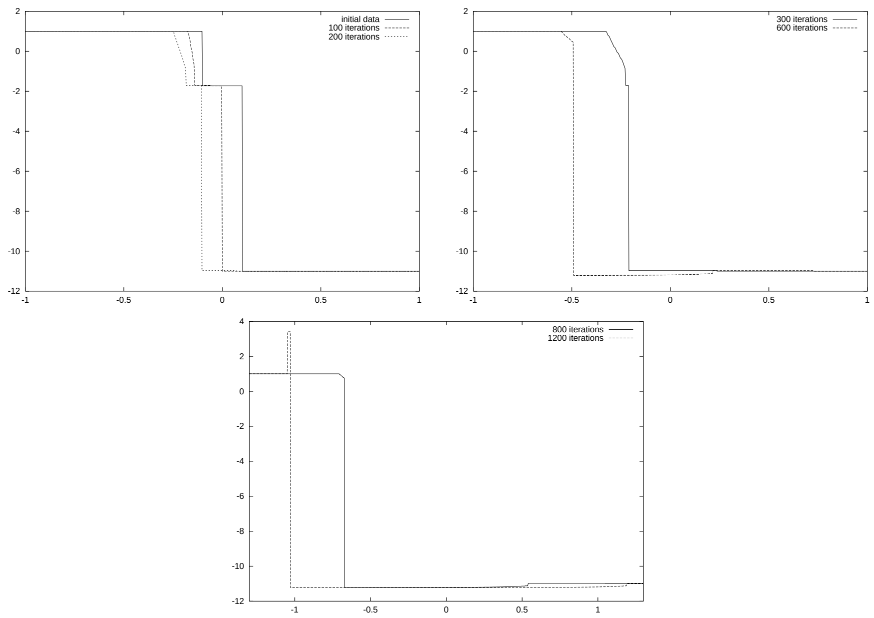

FIG. 3.19 – *Interaction of two classical shocks (v component)*

• Figures 3.21-3.22: *Large perturbations.* Finally, we observe that the nonclassical Riemann solver (as well as the classical one) is stable under possibly large pertur-

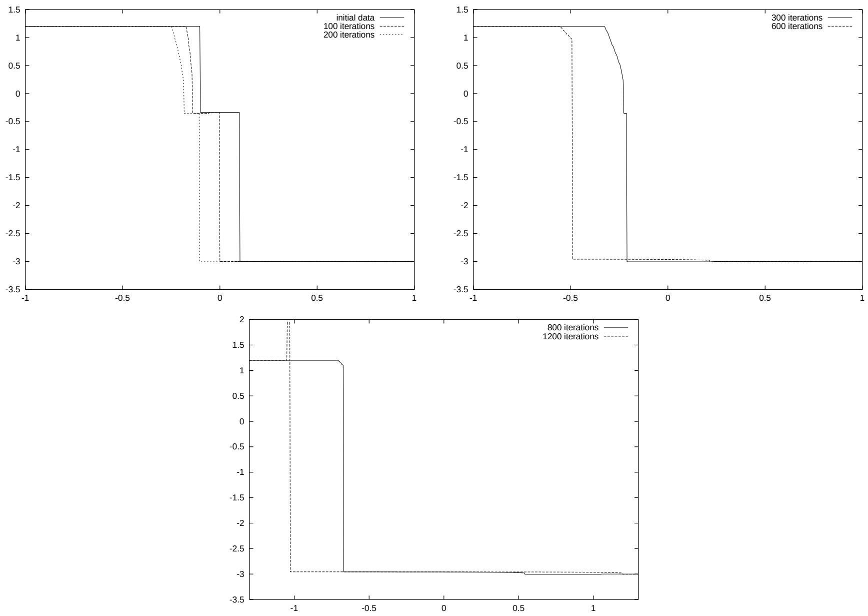

FIG. 3.20 – *Interaction of two classical shocks (w component)*

bations. Oscillations ranging between -0.03 and 0 are introduced in a given initial data. The discontinuity is located at 0 when no oscillations are present. The left-hand state is taken to be $(v_L, w_L) = (0, 1)$ and the right-hand state $(v_R, w_R) = (-6.71190015066, -2)$. We use a mesh containing 200 points per unit interval. We observe that the unperturbed profile is recovered asymptotically in time, but the position of the discontinuity is shifted.

Test 7 : Periodic data

As in the scalar case, we now study the asymptotic behavior of the solutions for several kinetic functions. We use the stress-function given by (3.14), but with now $m = 0.05$, and a mesh containing 1000 points.

• Figures 3.23-3.28: *The classical case.* First, we consider the case $\beta = 1/2$ and the following periodic initial data

$$(v_0(x), w_0(x)) = \begin{cases} (0, 0.4), & x < x_0, \\ (0, -0.2), & x_0 < x < x_1, \\ (0, 0.4), & x > x_1, \end{cases} \tag{3.17}$$

where x_0 and x_1 will be specified shortly. Each of the following computations is performed over the unit interval $[-0.5; 0.5]$ with periodic boundary conditions.

First, we choose $x_0 = 0$ and $x_1 = 0.3$ so that the corresponding average $\overline{w}_0$ of w_0 is positive, precisely $\overline{w}_0 = 0.22$. The components v and w are plotted on Figures 3.23 and 3.24, while Figure 3.25 shows the two Riemann invariants

$$v - \int^w \sqrt{\sigma'(x)}\, dx \quad \text{and} \quad v + \int^w \sqrt{\sigma'(x)}\, dx$$

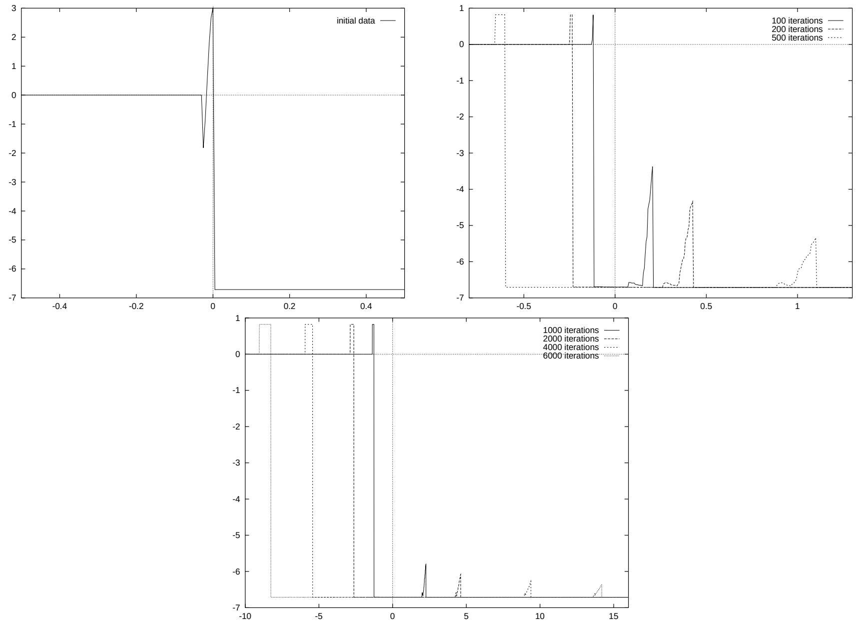

FIG. 3.21 – *Large perturbations (v component)*

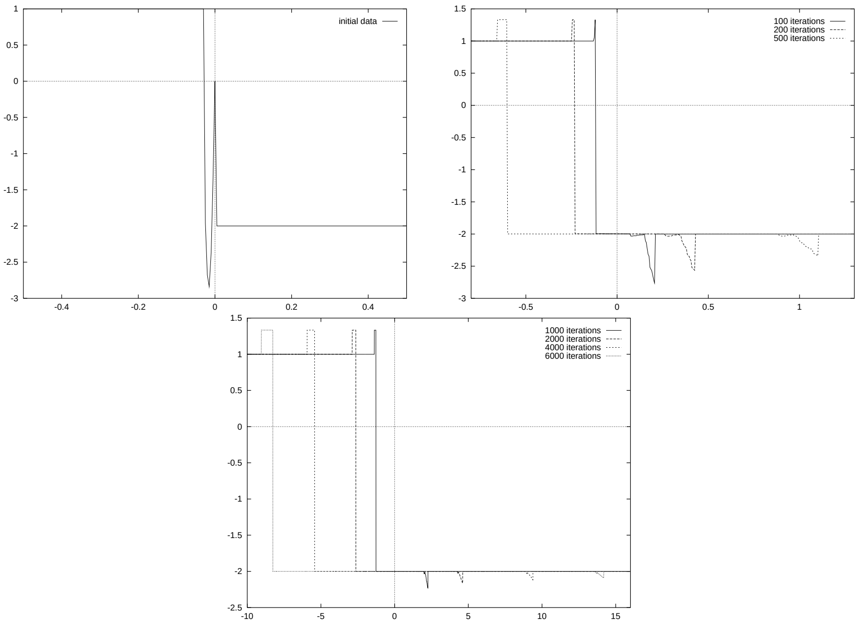

FIG. 3.22 – *Large perturbations (w component)*

after linearization near the average value $\overline{w}_0$, that is,

$$v - \sqrt{\sigma'(\overline{w}_0)}\, w \quad \text{and} \quad v + \sqrt{\sigma'(\overline{w}_0)}\, w.$$

We observe that, as the time evolves, these variables exhibit *two standard N-waves,* propagating at speed with opposite signs.

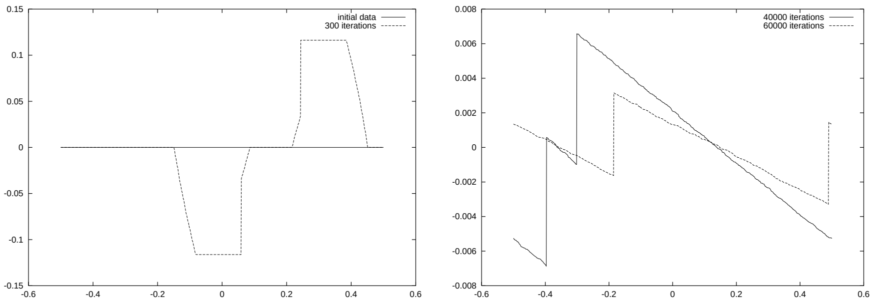

FIG. 3.23 – *Classical Riemann solution (v component)*

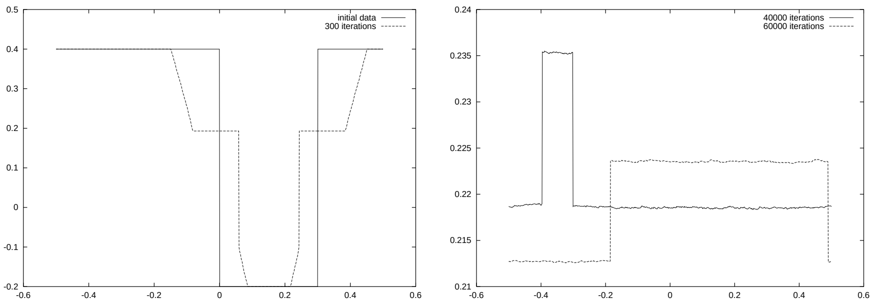

FIG. 3.24 – *Classical Riemann solution (w component)*

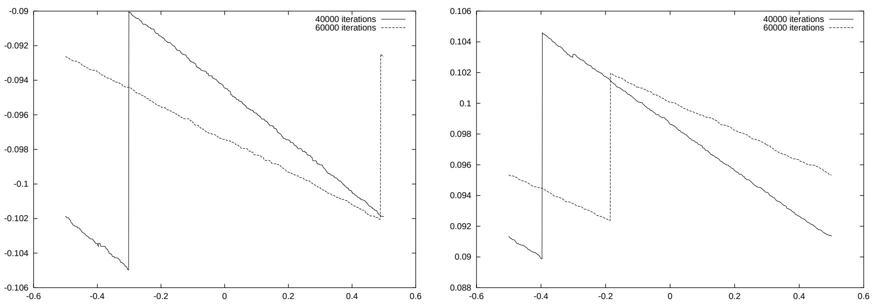

FIG. 3.25 – *Classical Riemann solution (Linearized Riemann invariants)*

Second, we choose $x_0 = -0.3.$ and $x_1 = -0.3 + 2/3$, so that the averages of v_0 and w_0 both equal 0. Figures 3.26 and 3.27 show the components v and w, respectively.

The *Riemann invariants* linearized near 0, that is,

$$v - \sqrt{\sigma'(0)}\, w \quad \text{and} \quad v + \sqrt{\sigma'(0)}\, w,$$

exhibit a *double N-wave pattern;* see Figure 3.28. We recall that the same behavior on the conservative variable was observed in the scalar case. (Compare with Figures 3.13, 3.14 and 3.17.)

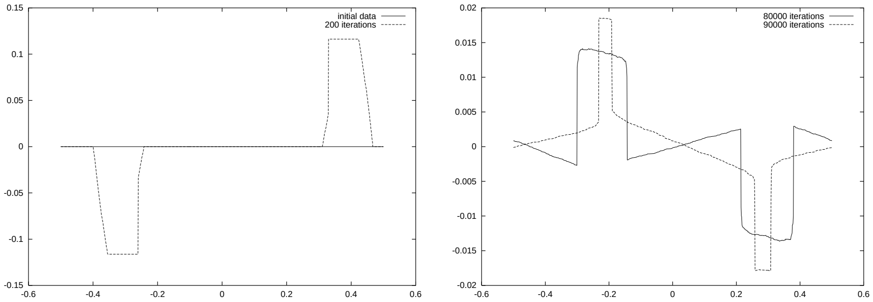

Fig. 3.26 – *Classical Riemann solution (v component)*

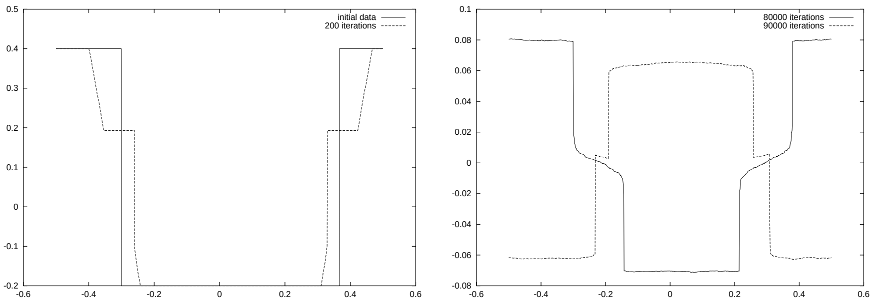

Fig. 3.27 – *Classical Riemann solution (w component)*

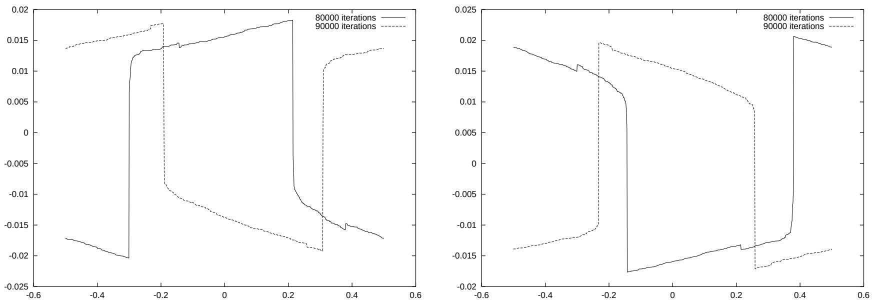

Fig. 3.28 – *Classical Riemann solution (Linearized Riemann invariants)*

• Figures 3.29-3.34: *The nonclassical case.* We now study the nonclassical value $\beta = 3/4$ together with the following periodic initial data

$$(v_0(x), w_0(x)) = \begin{cases} (0.1, 0.4), & x < x_0, \\ (0, -0.2), & x_0 < x < x_1, \\ (0.1, 0.4), & x > x_1. \end{cases}$$

Figures 3.29, 3.30, and 3.31 correspond to the choice $x_0 = 0$ and $x_1 = 3$ (so that the average of w_0 is 0.22), while Figures 3.32, 3.33, and 3.34 show the case $x_0 = -0.3$ and $x_1 = -0.3 - 2/3$ for which the average of w_0 equals zero. The same asymptotic behaviors as in the classical case are observed for the linearized Riemann invariants, that is, *standard N-waves* (Figure 3.31) and *double N-waves* (Figure 3.34), depending on the average of w_0.

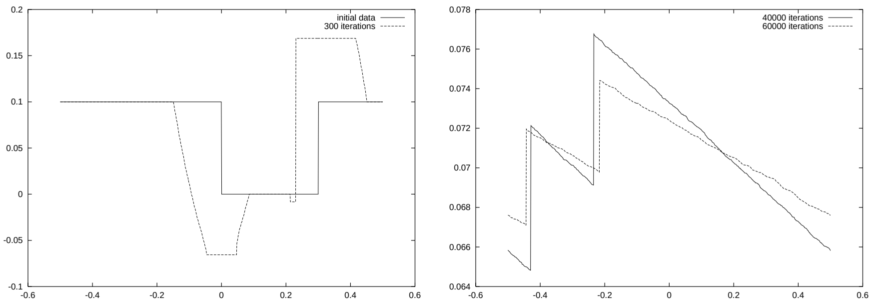

FIG. 3.29 – *Nonclassical Riemann solution (v component)*

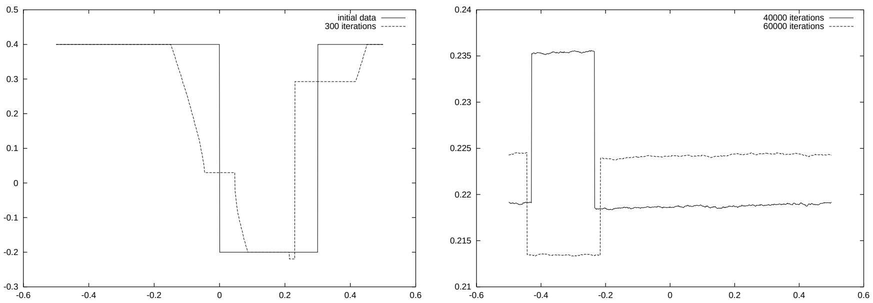

FIG. 3.30 – *Nonclassical Riemann solution (w component)*

• Figures 3.35-3.40: *The maximal nonclassical case.* Here we investigate the maximal choice $\beta = 1$, with the following periodic initial data:

$$(v_0(x), w_0(x)) = \begin{cases} (-0.3, 0.4), & x < -0.3, \\ (0, -0.2), & -0.3 < x < -0.3 + 2/3, \\ (-0.3, 0.4), & x > -0.3 + 2/3. \end{cases}$$

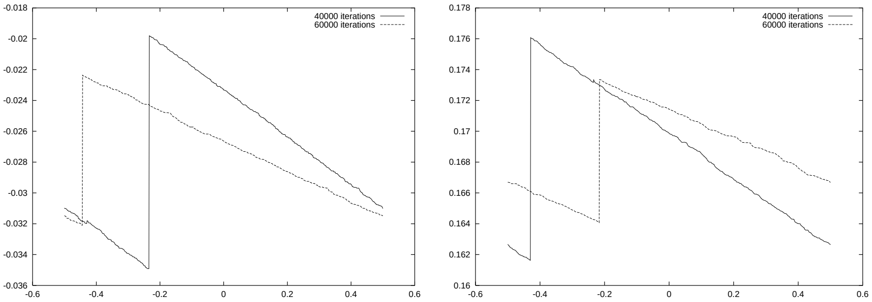

FIG. 3.31 – *Nonclassical Riemann solution (Linearized Riemann invariants)*

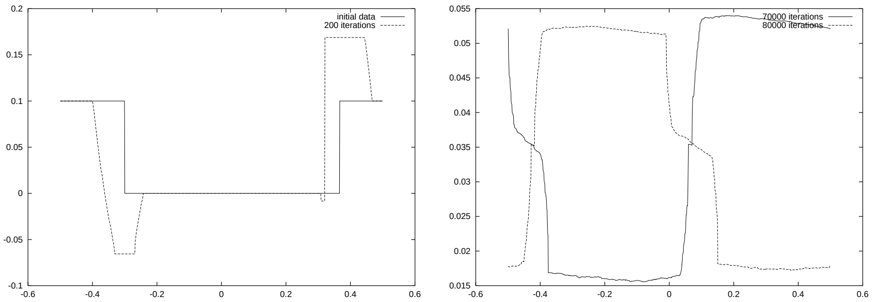

FIG. 3.32 – *Nonclassical Riemann solution (v component)*

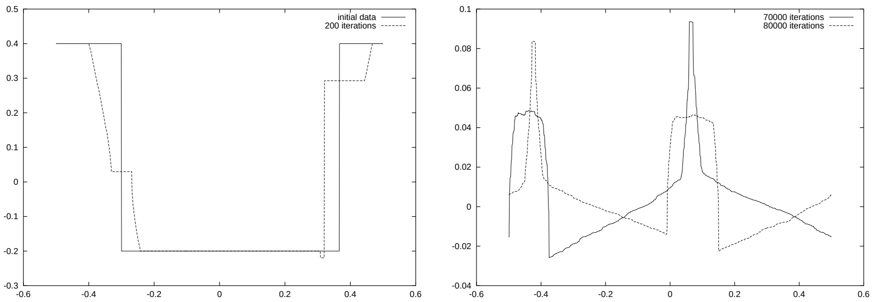

FIG. 3.33 – *Nonclassical Riemann solution (w component)*

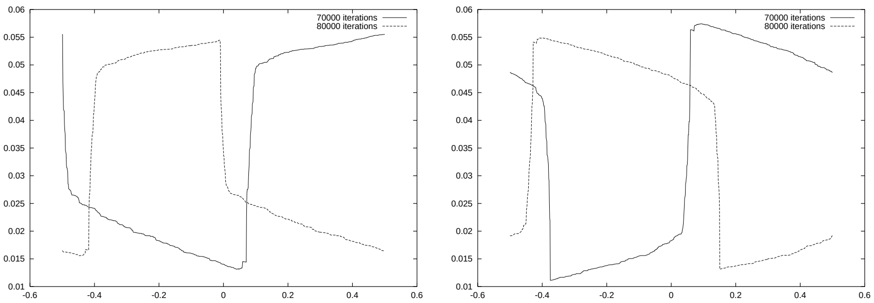

FIG. 3.34 – *Nonclassical Riemann solution (Linearized Riemann invariants)*

Here, the average of the component w equals zero. Figure 3.35 shows the component w after various numbers of iteration. For large times, we observe (as in the scalar case) that the solution tends to a piecewise constant function made of two phase transitions satisfying the prescribed kinetic relation. Since $\beta = 1$, they have the same propagation speed. Observe also that the number of phase transitions is *equal* to the one of initial data.

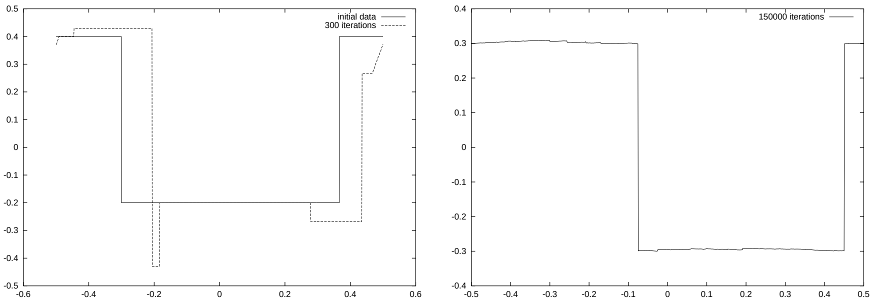

FIG. 3.35 – *Maximal nonclassical Riemann solution (w component)*

It is worth pointing out that another qualitative behavior can be observed with the maximal nonclassical Riemann solver. Indeed, if we consider the periodic initial data

$$(v_0(x), w_0(x)) = \begin{cases} (0, 0.4), & x < -0.3, \\ (0, -0.2), & -0.3 < x < -0.3 + 2/3, \\ (0, 0.4), & x > -0.3 + 2/3, \end{cases}$$

(the average of w_0 is 0 as well as the average of v_0), we observe now a *double N-wave pattern* on the linearized Riemann invariants, and for large times, the solution does not seem to converge to a solution made of phase transitions with equal propagation speeds. Asymptotically, the solution does contain two phase transitions, that is, exactly the *same* number as there was in the initial data.

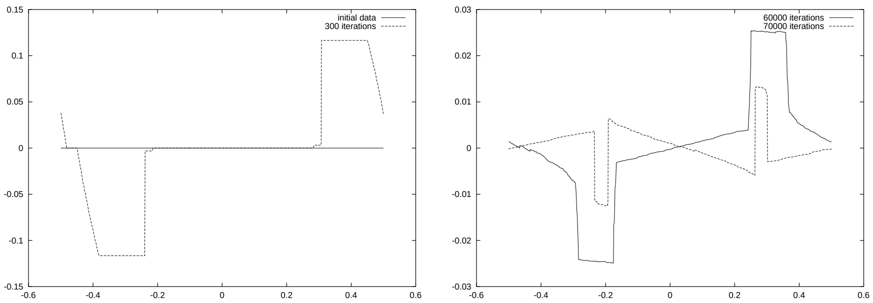

FIG. 3.36 – *Maximal nonclassical Riemann solution (v component)*

We give another test of interest which demonstrates that the number of phase transitions may be asymptotically be *larger* than the one in the initial data. Consider

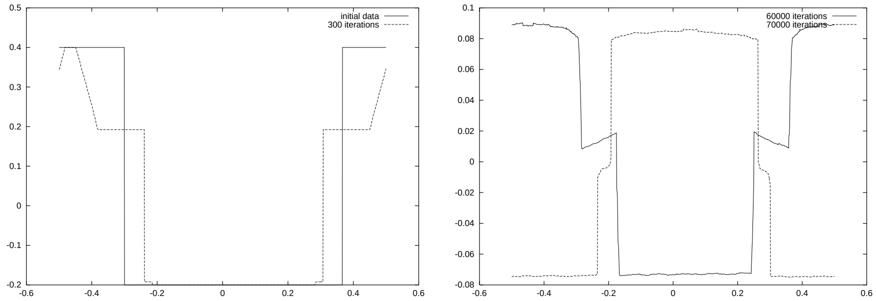

FIG. 3.37 – *Maximal nonclassical Riemann solution (w component)*

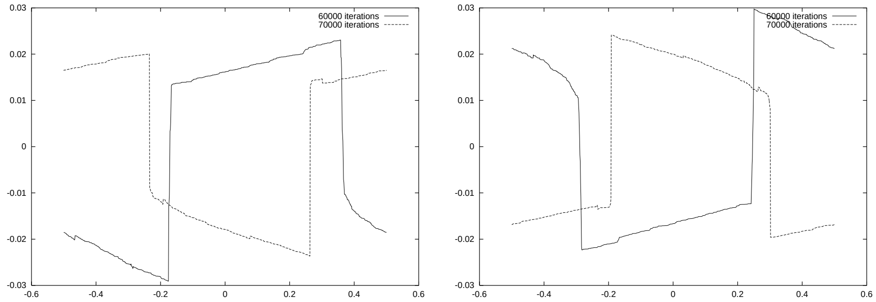

FIG. 3.38 – *Maximal nonclassical Riemann solution (Linearized Riemann invariants)*

the periodic initial data

$$(v_0(x), w_0(x)) = \begin{cases} (-0.3, 0.4), & x < -0.3, \\ (0.15, -0.2), & -0.3 < x < -0.3 + 2/3, \\ (-0.3, 0.4), & x > -0.3 + 2/3. \end{cases}$$

We plotted the component w of the solution on Figure 3.39. For large times, we observe a solution made of *four* phase transitions with coinciding propagation speeds.

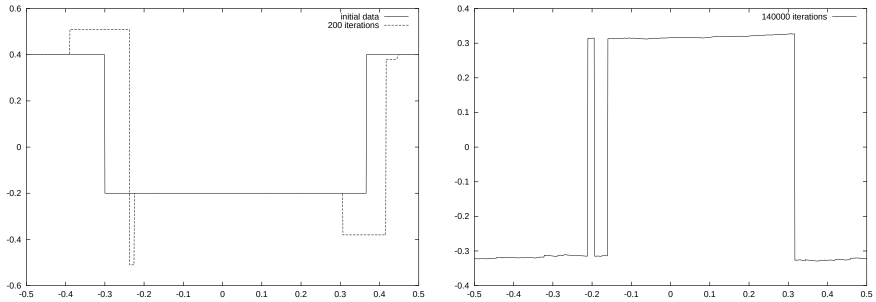

FIG. 3.39 – *Increasing the number of phase transitions (w component)*

To conclude this section, consider a periodic initial data made of four phase transitions, precisely

$$(v_0(x), w_0(x)) = \begin{cases} (0.1, 0.4), & x < -0.4, \\ (0.2, -0.3), & -0.4 < x < -0.1, \\ (0., 0.4), & -0.1 < x < 0.1, \\ (0, -0.2), & 0.1 < x < 0.3, \\ (0.1, 0.4), & x > 0.3. \end{cases}$$

We plotted the component w of the numerical solution on Figure 3.40. Observe that, here, none of the averages of v_0 or w_0 is zero. For sufficiently large times, we obtain a solution made of two phase transitions (propagating with the same speed), so that the number of transitions has decreased with respect to the one in the initial data.

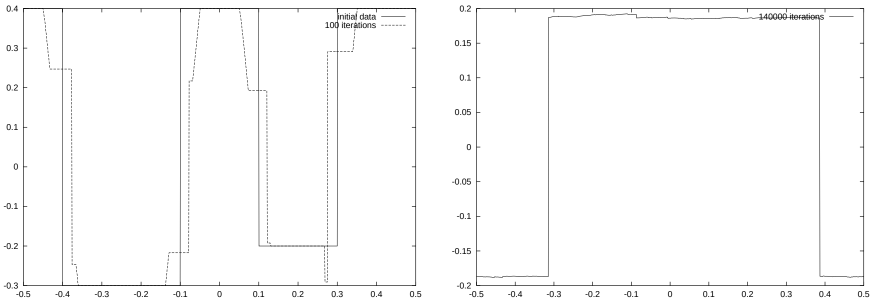

FIG. 3.40 – *Initial data made of four phase transitions (w component)*

Test 8 : Mesh refinements

Consider the initial data

$$(v_0(x), w_0(x)) = \begin{cases} (1, 3), & x < 0, \\ (0.2, -4), & x > 0. \end{cases}$$

We plot the numerical solution for several mesh lengths, at the time $t = 0.06$ and for $\beta = 2/3$. This test illustrates that the random choice method converges very sharply to the exact nonclassical solution of the problem under consideration.

3.4 Conclusions

In this paper, we have studied two hyperbolic models giving rise to nonclassical undercompressive shock waves, namely the scalar conservation laws and the model of elastodynamics when the flux is a cubic function. We have investigated the stability and time-asymptotic properties of classical and nonclassical entropy solutions.

(1) We have demonstrated that Glimm's scheme converges to exact solutions, even when the classical Riemann solver is replaced with a nonclassical Riemann solver.

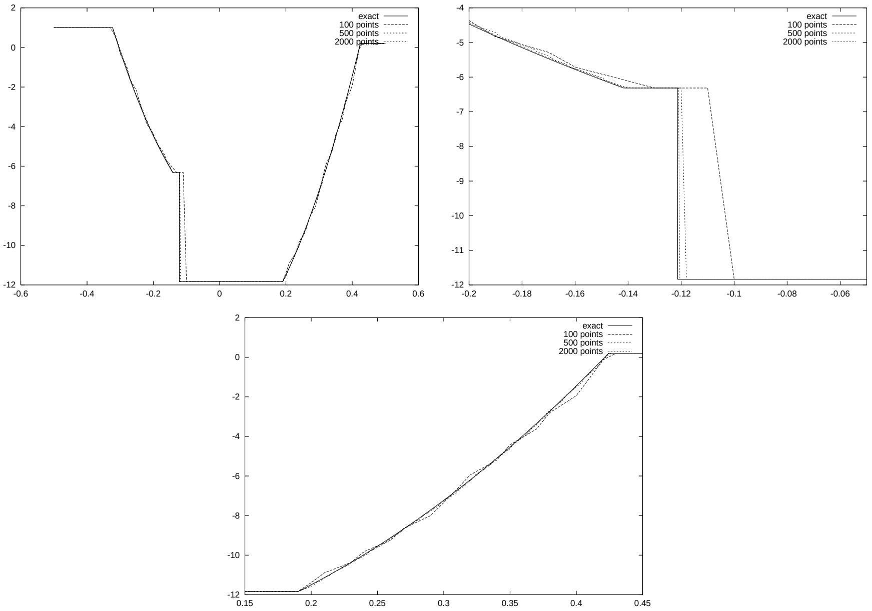

FIG. 3.41 – *System case: convergence (global v-component and zooms)*

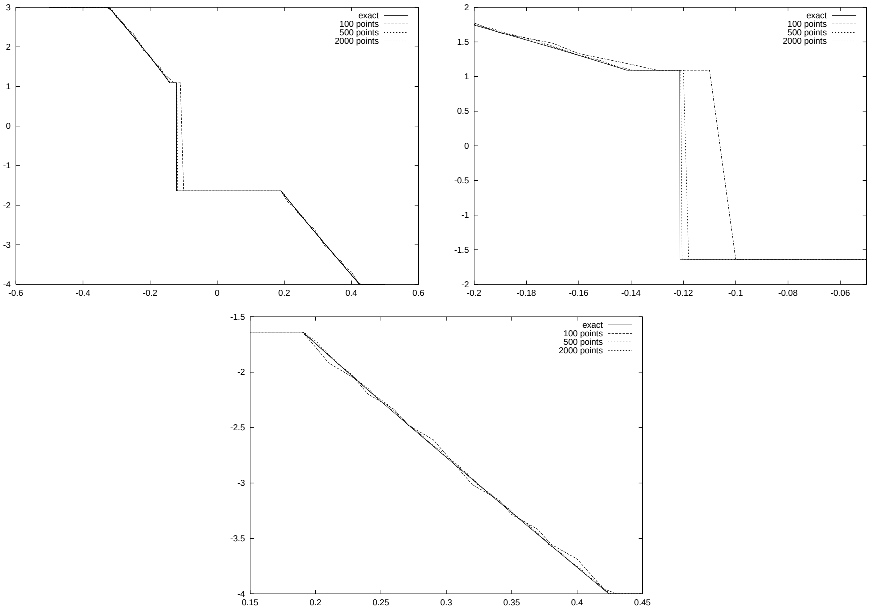

FIG. 3.42 – *System case: convergence (global w-component and zooms)*

(2) It is well-known that for classical entropy solutions the total variation is a non-increasing functional for scalar conservation laws, and, for systems, may increase only by an amount of the order of the incoming wave strength. This is no longer true for nonclassical solutions. We have observed that, under small perturbations, classical shock waves may be *split into a nonclassical shock and a classical one,* which is the reason why the *total variation of nonclassical solutions may increase drastically* as the solution evolves in time.

(3) We have studied the time-asymptotic behavior of periodic solutions. We found that any periodic solution asymptotically converges to a constant state and, more precisely, approaches a well-known *N-wave pattern* when the average of the initial data is non-zero but approaches a *"double N-wave pattern"* if this average equals zero. In the latter, nonclassical double N-waves contain two interfaces propagating with a speed fullfiling the Rankine-Hugoiniot relation and the kinetic relation. These conclusions hold for the conservative variable in the scalar case (Section 3.2) and for the linearized Riemann invariants for the system (Section 3.3). It would be interesting to determine this pattern analytically, but this seems to be very chalenging. It is conceivable that the existence of "double N-wave pattern is typical of classical and nonclassical solutions of general hyperbolic systems which fail to have globally genuinely nonlinear characteristic fields.

(4) Special attention was devoted to the maximal kinetic function, which allows us to exhibit still another interestig behavior of nonclassical entropy solutions. Generally speaking, the numerical solution converges for large times to an oscillating pattern involving only two constant values and a finite number of transitions propagating with the same speed. In addition, this number may decrease as well as increase in time.

A follow-up paper [27] will investigate to which extend our conclusions remain valid for the hyperbolic-elliptic version of (3.13). Additional difficulties arise in this context, for instance the lack of continuous dependence and uniqueness of the nonclassical solver (see [96]), which we will have to address from the numerical standpoint.

To close this paper, we emphasize that Glimm's compare favorably with numerical methods for computing nonclassical shock waves. In particular, by taking into account diffusive and dispersive terms directly within a finite difference scheme, it is possible to compute nonclassical shock waves with arbitary accuracy and to approach a prescribe kinetic relation generated by a zero diffusion-dispersion limit. See [64], [94], [25], [24], [93]. However the accuracy deteriorates with shocks with large strength and with large-time computations.

In contrast, Glimm's scheme

• converges to the correct solution satisfying the prescribed kinetic relation for all nonclassical shock waves, even with arbitrary strength,

• is very flexible and encompass arbitary kinetic functions, which need not be generated by a specific diffusive-dispersive mechanism, but may be determined by experiments (for instance)

• and allows one to determine large-time asymptotics of nonclassical solutions.

Bibliographie

[1] Abeyaratne R. et Knowles J.K., *Kinetic relations and the propagation of phase boundaries in solids*, Arch. Rational Mech. Anal., vol 114, pp 119-154 (1991).

[2] Abeyaratne R. et Knowles J.K., *Implications of viscosity and strain gradient effects for the kinetics of propagating phase boundaries*, SIAM J. Appl. Math., vol 51, pp 1205-1221 (1991).

[3] Abeyaratne R. et Knowles J.K., *On the driving traction acting on a surface of strain discontinuity in a continuum*, J. Mech. Phys. Solids, vol 38, pp 345-360 (1990).

[4] Abgrall R., *Extension de la méthode de Roe aux équations d'Euler multi-espèces*, La recherche aérospatiale, vol 6, pp 31-43 (1988).

[5] Abgrall R., *Toward the ultimate conservative scheme : Following the Quest*, J. Comput. Phys., vol 167 (2), pp 277-315 (2001).

[6] Affouf M. et Caflisch R., *A numerical study of Riemann problem solutions and stability for a system of viscous conservation laws of mixed type*, SIAM J. Appl. Math., vol 51, pp 605-634 (1991).

[7] Bedjaoui N. et LeFloch P.G., *Diffusive-dispersive traveling waves and kinetic relations. I. Nonconvex hyperbolic conservation laws*, J. Differential Equations, vol 178, pp 574-607 (2002).

[8] Bedjaoui N. et LeFloch P.G., *Diffusive-dispersive traveling waves and kinetic relations. III. An hyperbolic model from nonlinear elastodynamics*, Ann. Univ. Ferrara Sc. Mat., vol 47, pp 117-144 (2001).

[9] Bedjaoui N. et LeFloch P.G., *Diffusive-dispersive traveling waves and kinetic relations. II. An hyperbolic-elliptic model of phase transitions*, Proc. Royal Soc. Edinburgh, (2002).

[10] Berthon C., *Contribution à l'analyse numérique des équations de Navier-Stokes compressibles à deux entropies spécifiques. Application à la turbulence compressible*, Thèse de l'Université Paris VI, (1999).

[11] Berthon C. et Coquel F., *A convection-diffusion system with first and second order in a nonconservation form. Part 1 : Existence of traveling wave solutions*, preprint.

[12] Berthon C. et Coquel F., *A convection-diffusion system with first and second order in a nonconservation form. Part 2 : Approximate jump relations*, preprint.

[13] Berthon C. et Coquel F., *Nonlinear projection scheme to solve systems in non conservation form*, in Innovative methods in Computational Fluid Dynamics.

[14] Berthon C. et Coquel F., *Travelling wave solutions for the Navier-Stokes equations with several specific entropies*, work in preparation.

[15] Berthon C., Coquel F. et LeFloch P.G., *Entropy dissipation measure and kinetic relation associated with nonconservative hyperbolic systems*, work in preparation.

[16] Bona J. et Schonbek M.E., *Traveling wave solutions to the Korteweg-de Vries Burgers equation*, Proc. Royal Soc. Edinburgh, vol 101A, pp 207-226 (1985).

[17] Bouchut F., *A reduced stability condition for nonlinear relaxation to conservation laws*, Prépublication 02-13 du DMA (2002).

[18] Bouchut F., *Entropy satisfying flux vector splittings and kinetic BGK models*, To appear in Num. Math. (2002).

[19] Chalons C., Thèse de l'École Polytechnique, (2002).

[20] Chalons C. et Coquel F., *Numerical approximation of the Navier-Stokes equations with several independent specific entropies*, submitted.

[21] Chalons C. et Coquel F., *The Riemann problem for the multi-pressure Euler system*, preprint.

[22] Chalons C. et Coquel F., *Euler equations with several independent pressure laws and entropy satisfying explicit projection schemes*, preprint.

[23] Chalons C. et Coquel F., *Navier-Stokes equations with several independent pressure laws and explicit predictor-corrector schemes*, preprint.

[24] Chalons C. et LeFloch P.G., *A fully discrete scheme for diffusive-dispersive conservation laws*, Numerisch Math., vol 89, pp 493-509 (2001).

[25] Chalons C. et LeFloch P.G., *High-order entropy conservative schemes and kinetic relations for van der Waals fluids*, J. Comput. Phys., vol 167, pp 1-23 (2001).

[26] Chalons C. et LeFloch P.G., *Computing undercompressive waves with the random choice method. I. Nonclassical shock waves*, submitted.

[27] Chalons C. et LeFloch P.G., *Computing undercompressive waves with the random choice method. II. Propagating phase boundaries*, work in preparation.

[28] Chen G.Q., Levermore C.D. et Liu T.P., *Hyperbolic conservation laws with stiff relaxation terms and entropy*, Comm. Pure Appl. Math, vol 48(7), pp 787-830 (1995).

[29] Cockburn B. et Gau H., *A model numerical scheme for the propagation of phase transitions in solids*, SIAM J. Sci. Comput., vol 17, pp 1092-1121 (1996).

[30] Collela P., *Glimm's method for gas dynamics*, SIAM J. Sci. Stat. Comput., vol 3, pp 76-110 (1982).

[31] Colombeau J.F. et Leroux A.Y., *Multiplications of distributions in elasticity and hydrodynamics*, J. Math. Phys., vol 29, pp 315-319 (1988).

[32] Coquel F., Godlewski E., In A., Perthame B. et Rascle P., *Some new Godunov and relaxation methods for two phase flows*, Proceedings of an International conference on Godunov methods : Theory and Applications, Kluwer Academic/Plenum Publishers (2001).

[33] Coquel F. et Joly V., *Développement d'un code de calculs d'écoulements hypersoniques en déséquilibre chimique et vibrationnel*, Rapport Interne 2/2536AY, ONERA (1992).

[34] Coquel F. et Marmignon C., *A Roe-type linearization for the Euler equations for weakly ionized multi-component and multi-temperature gas*, AIAA 12^{th} CFD Conference, San Diego (USA) (1995).

[35] Coquel F. et Marmignon C., *Simulation numérique d'écoulements faiblement ionisés en déséquilibre thermochimique*, Rapport Interne 6/1929AY, ONERA (1993).

[36] Coquel F. et Marmignon C., *A consistent conservative formulation for the Euler system for weakly ionized mixtures of gaz*, Finite volumes for complex applications II, Editors Vilsmeier R., Benkhaldoun F., Henel D., Hermes (1999).

[37] Coquel F. et Marmignon C., *Simulation numérique d'écoulements faiblement ionisés en déséquilibre thermochimique*, Rapport Interne 12/1929AY 024A, ONERA (1995).

[38] Coquel F. et Marmignon C., *Simulation numérique d'écoulements en déséquilibre thermochimique et faiblement ionisés*, Rapport Interne 14/1929AY, ONERA (1997).

[39] Coquel F. et Perthame B., *Relaxation of energy and approximate Riemann solvers for general pressure laws in fluid dynamics*, SIAM J. Math. Anal., vol 35(6), pp 2223-2249 (1998).

[40] Coquel F. et Perthame B., Note non publiée.

[41] Cordier S., Degond P., Markowich P. et Schmeiser C., *Travelling waves analysis and jump relations for Euler-Poisson model in quasi-neutral limit*, Asymptotic Analysis, vol 10, (1995).

[42] Dal Maso G., LeFloch P.G. et Murat F., *Definition and weak stability of a non conservative product*, J. Math. Pures Appli., vol 74, pp 483-548 (1995).

[43] Desprès B., *Lagrangian systems of conservation laws. Invariance properties of Lagrangian systems of conservation laws, approximate Riemann solvers and the entropy condition*, Numer. Math., vol 89, pp 99-134 (2001).

[44] De Vuyst F., *Schémas non-conservatifs et schémas cinétiques pour la simulation numérique d'écoulements hypersoniques non visqueux en déséquilibre thermochimique*, Thèse de l'Université Paris VI, (1994).

[45] DiPerna R.J., *Finite difference scheme for conservation laws*, Comm. Pure and Appl. Math., vol 25, pp 379-450 (1982).

[46] Fan H.T., *A limiting "viscosity" approach to the Riemann problem for materials exhibiting a change of phase II*, Arch. Rational Mech. Anal., vol 116, pp 317-337 (1992).

[47] Fan H.T., *The uniqueness and stability of solution of the Riemann problem for a system of conservation laws of mixed type*, Trans. Amer. Math. Soc., vol 333, pp 913-938 (1992).

[48] Fan H.T., *A vanishing viscosity approach on the dynamics of phase transitions in van der Waals fluids*, J. Differential Equations, vol 103, pp 179-204 (1993).

[49] Fan H.T., *One-phase Riemann problem and wave interactions in systems of conservation laws of mixed type*, SIAM J. Math. Anal., vol 24, pp 840-865 (1993).

[50] Fan H.T. et Slemrod M., *The Riemann problem for systems of conservation laws of mixed type, in "Shock induces transitions and phase structures in general media"*, R. Fosdick, E. Dunn, and H. Slemrod ed., IMA Vol. Math. Appl., Springer-Verlag, vol 52, pp 61-91 (1993).

[51] Gallice G., *Schémas de type Godunov entropiques et positifs pour la dynamique des gaz et la magnétohydrodynamique lagrangiennes*, C. R. Acad. Sci. Paris Sér. I Math., vol 332, pp 1037-1040 (2001).

[52] Gavrilyuk S. et Gouin H., *A new form of governing equations of fluids arising from Hamilton's principle*, Int. J. Eng. Scie., vol 12, pp 1485-1520 (1999).

[53] Gilbarg G., *The existence and limit behavior of the one-dimensional shock layer*, Amer. J. Math., vol 73, pp 256-274 (1951).

[54] Gleize V., *Simulation numérique d'écoulements turbulents compressibles hors équilibre à l'aide de schémas multi-échelle*, Thèse de l'Université d'Aix-Marseille II (1994).

[55] Glimm J., *Solutions in the large time for nonlinear hyperbolic systems of equations*, Comm. Pure Appl. Math., vol 18, pp 697-715 (1965).

[56] Godlewsky E. et Raviart P.A., **Hyperbolic systems of conservation laws**, Ellipse (1991).

[57] Godlewsky E. et Raviart P.A., **Numerical approximation of hyperbolic systems of conservation laws**, Springer (1995).

[58] Godunov S.K., *A difference scheme for numerical computation of discontinuous solutions of equations of fluids dynamics*, Math. Sbornik, vol 47, pp 271-306 (1959).

[59] Harten A., *On the symmetric form of systems of conservation laws with entropy*, J. Comp. Phys., vol 49, pp 151-164 (1983).

[60] Harten A. et Lax P.D., *A random choice finite difference scheme for hyperbolic conservation laws*, SIAM J. Num. Anal., vol 18, pp 289-315 (1981).

[61] Harten A., Lax P.D. et Van Leer B., *On upstream differencing and Godunov-type schemes for hyperbolic conservation laws*, SIAM Rev., vol 25, pp 35-61 (1983).

[62] Hattori H., *The Riemann problem for a van der Waals fluid with the entropy rate admissibility criterion*, Arch. Rational Mech. Anal., vol 92, pp 247-263 (1986).

[63] Hayes B.T. et LeFloch P.G., *Nonclassical shocks and kinetic relations : Scalar conservation laws*, Arch. Rational Mech. Anal., vol 139, pp 1-56 (1997).

[64] Hayes B.T. et LeFloch P.G., *Nonclassical shocks and kinetic relations : Finite difference schemes*, SIAM J. Numer. Anal., vol 35, pp 2169-2194 (1998).

[65] Hayes B.T. et LeFloch P.G., *Nonclassical shocks and kinetic relations : Strictly hyperbolic systems*, SIAM J. Math. Anal., vol 31, pp 941-991 (2000).

[66] Hayes B.T. et LeFloch P.G., *Nonclassical shocks waves and kinetic relations (II)*, Préprint CMAP École Polytechnique (France), vol 357, (1996).

[67] Herard J.M., *Solveur de Riemann approché pour un système hyperbolique non conservatif issu de la turbulence compressible*, Internal EDF report, HE-41/95/009/A (1995).

[68] Hirsch C., **Numerical computation of internal and external flows**, John Wiley and Sons (1990).

[69] Hou T.Y. et LeFloch P.G., *Why nonconservative schemes converge to wrong solutions : error analysis*, Math. of Comp., vol 62(206), pp 497-530 (1994).

[70] Hou T.Y., Rosakis P. et LeFloch P.G., *A level set approach to the computation of twinning and phase transition dynamics*, J. Comput. Phys., vol 150, pp 302-331 (1999).

[71] Jacobs D., McKinney W.R., et Shearer M., *Traveling wave solutions of the modified Korteweg-de Vries Burgers equation*, J. Differential Equations, vol 116, pp 448-467 (1995).

[72] James R.D., *The propagation of phase boundaries in elastic bars*, Arch. Rational Mech. Anal., vol 73, pp 125-158 (1980).

[73] Jin S., *Numerical integrations of systems of conservation laws of mixed type*, SIAM J. Appl. Math., vol 55, pp 1536-1551 (1995).

[74] Jin S. et Xin Z., *The Relaxation Schemes for Systems of Conservation Laws in Arbitrary Space Dimension*, Comm. Pure. Appl. Math., vol 48, pp 235-276 (1995).

[75] Jones W.P. et Launder B.E., *The prediction of laminarization with a two-equation model of turbulence*, Journal Heat Mass Trans., vol 15, pp 301-314 (1972).

[76] Karni S., *Viscous shock profiles and primitive formulations*, SIAM, J. Num. Anal., vol 29 (6), pp 1592-1609 (1992).

[77] Katsoulakis M. et Tzavaras A.E., *Contractive relaxation systems and the scalar multidimensional conservation law*, Comm. in P.D.E., vol 22, pp 195-233 (1997).

[78] Khobalatte B. et Perthame B., *Maximum principle on the entropy and second order kinetic schemes*, Math. of Comp., vol 62, 205, pp 119-131 (1994).

[79] Lagoutière F., *Modélisation mathématique et résolution numérique de problèmes de fluides compressibles à plusieurs constituants*, Thèse de l'Université Paris VI, (2000).

[80] Larrouturou B., *On upwind approximations of multi-dimensional multi-species flows*, Computational methods in applied sciences, Ch. Hirsch (Eds), Elsevier (1992).

[81] Larrouturou B., *How to preserve the mass fractions positivity when computing compressible multi-component flows*, J. Comput. Phys., vol 95, pp 59-84 (1991).

[82] Lascaux P. et Theodor R., **Analyse numérique matricielle appliquée a l'art de l'ingénieur**, Masson (1989).

[83] Lax P.D., *Hyperbolic systems of conservation laws, II*, Comm. Pure Appl. Math., vol 10, pp 537-566 (1957).

[84] Lax P.D., *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*, CBMS Monograph, Society for Industrial and Applied Mathematics, Philadelphia (1973).

[85] LeFloch P.G., *Sur quelques problèmes hyperboliques nonlinéaires*, Thèse d'habilitation de l'Université Paris VI, (1990).

[86] LeFloch P.G., *Propagating phase boundaries: formulation of the problem and existence via the Glimm scheme*, Arch. Rational Mech. Anal., vol 123, pp 153-197 (1993).

[87] LeFloch P.G., *An introduction to nonclassical shocks of systems of conservation laws*, Proceedings of the "International School on Theory and Numerics for Conservation Laws", Freiburg/Littenweiler (Germany), ed. D. Kröner, M. Ohlberger and C. Rohde (20-24 Octobre 1997), Lecture Notes in Computational Science and Engineering, pp 28-72 (1998).

[88] LeFloch P.G., **Hyperbolic Systems of Conservation Laws: The theory of classical and nonclassical shock waves**, E.T.H. Lecture Notes Series, Birkhäuser (2002).

[89] LeFloch P.G., *Computational Methods for propagating phase boundaries*, In "Intergranular and Interphase Boundaries in Materials: iib95", Lisbon, June 1995. Eds. A.C. Ferro, J.P. Conde and M.A. Fortes. Materials Science Forum Vols, vol 207-209, pp 509-515 (1996).

[90] LeFloch P.G., *Shock waves for nonlinear hyperbolic systems in nonconservative form*, Institute for Math. and its Appl., Minneapolis, Preprint # 593, (1989).

[91] LeFloch P.G., *Entropy weak solutions to nonlinear hyperbolic systems in nonconservative form*, Comm. Partial Differential Equations, vol 13, pp 669-727 (1988).

[92] LeFloch P.G. et Liu T.P., *Existence theory for nonlinear hyperbolic systems in nonconservative form*, Forum Math., vol 5, pp 261-280 (1993).

[93] LeFloch P.G., Mercier J.M. et Rohde C., *Fully discrete entropy conservative schemes of arbitrary order*, submitted.

[94] LeFloch P.G. et Rohde C., *High-order schemes, entropy inequalities, and nonclassical shocks*, SIAM J. Numer. Anal., vol 37, pp 2023-2060 (2000).

[95] LeFloch P.G. et Thanh M.D., *Nonclassical Riemann solvers and kinetic relation I. An hyperbolic model of elastodynamics*, Z. Angew. Math. Phys., vol 52, pp 597-619 (2001).

[96] LeFloch P.G. et Thanh M.D., *Nonclassical Riemann solvers and kinetic relation II. An hyperbolic-elliptic model of phase transitions*, Proc. Royal Soc. Edinburgh, vol 131A, pp 1-39 (2001).

[97] LeFloch P.G. et Tzavaras A.E., *Existence theory for the Riemann problem for nonconservative hyperbolic systems*, C. R. Acad. Sci., Paris Sér. I Math., vol 323, pp 347-352 (1996).

[98] Lesieur M., **Turbulence in fluids**, Kluwer Academic Publishers, Dordrecht (1990).

[99] LeVeque R.J., **Numerical Methods for Conservation Laws**, Lectures in Mathematics, ETH Zurich, Birkhäuser (1992).

[100] Liu T.P., *The deterministic version of the Glimm scheme*, Comm. Math. Phys., vol 57, pp 135-148 (1977).

[101] Liu T.P., *Hyperbolic conservation laws with relaxation*, Comm. Math. Phys., vol 108, pp 153-175 (1987).

[102] Liu T.P., *Admissible solutions of hyperbolic conservation laws*, Amer. Math. Soc., vol 30, (1981).

[103] Liu T.P., *The Riemann problem for general systems of conservation laws*, J. Differential Equations, vol 18, pp 218-234 (1975).

[104] Mehlman G., *Étude de quelques problèmes liés aux écoulements en déséquilibre chimique et thermique*, Thèse de l'École Polytechnique, (1991).

[105] Mehlman G., *An Approximate Riemann Solver for Fluid Systems Based on a Shock Curve Decomposition*, Proceedings of the 3$^{\text{rd}}$ International conference on Hyperbolic Problems, Engquist Gustafsson eds., Chartwell Bratt, pp 727-741 (1991).

[106] Mohammadi B. et Pironneau O., **Analysis of the k-epsilon turbulence model**, Research in Applied Mathematics, Masson (1994).

[107] Montigny-Rannou F., *Analyse numérique appliquée à la mécanique des fluides*, Formation professionnelle continue, ONERA (1993 et 1994).

[108] Mulder W.A. et Van Leer B., *Experiments with implicit upwind methods for the Euler equations*, J. Comput. Phys., vol 59, pp 232-246 (1985).

[109] Natalini R., *Convergence to equilibrium for the relaxation approximations of conservation laws*, Comm. Pure Appl. Math., vol 49, pp 1-30 (1996).

[110] Perthame B., *Boltzmann type schemes for gaz dynamics and the entropy property*, SIAM J. Numer. Anal., vol 27, pp 1405-1421 (1990).

[111] Piquet J., *La turbulence et sa modélisation*, Cours de l'E.N.S.T.A., (1983).

[112] Raviart P.A. et Sainsaulieu L., *A non conservative hyperbolic system modelling spray dynamics. Part 2. Existence of travelling waves solutions*, Mathematical Methods and Models in Applied Sciences.

[113] Roe P.L., *Approximate Riemann solvers, parameter vectors and difference schemes*, J. Comp. Phys., vol 43, pp 357-372 (1981).

[114] Sagault P., **Introduction à la simulation des grandes échelles pour les écoulements de fluide compressible**, Springer (1998).

[115] Sainsaulieu L., *Travelling waves solutions of convection-diffusion systems whose convection terms are weakly nonconservative*, SIAM Journal of Applied Math., vol 55, pp 1552-1576 (1995).

[116] Schiestel R., **Modélisation et simulation des écoulements turbulents**, Traités des Nouvelles Technologies, Série mécanique, Editions Hermes (1993).

[117] Schulze S. et Shearer M., *Undercompressive shocks for a system of hyperbolic conservation laws with cubic nonlinearity*, J. Math. Anal. Appl., vol 229, pp 344-362 (1999).

[118] Serre D., **Systèmes de lois de conservation**, Diderot éditeur, Arts et Sciences (1996).

[119] Shearer M., *The Riemann problem for a class of conservation laws of mixed type*, J. Differential Equations, vol 46, pp 426-443 (1982).

[120] Shearer M. et Yang Y., *The Riemann problem for the p-system of conservation laws of mixed type with a cubic nonlinearity*, Proc. A Royal Soc. Edinburgh. 125A, pp 675-699 (1995).

[121] Shu C.W., *Total-variation diminishing time discretization*, SIAM J. Sci. Stat. Comput., vol 9, pp 1073-1084 (1988).

[122] Shu C.W., *A numerical method for systems of conservation laws of mixed type admitting hyperbolic flux splitting*, J. Comput. Phys., vol 100, pp 424-429 (1992).

[123] Slemrod M., *Admissibility criteria for propagating phase boundaries in a van der Waals fluid*, Arch. Rational Mech. Anal., vol 81, pp 301-315 (1983).

[124] Slemrod M., *Lax-Friedrichs and the viscosity-capillarity criterion*, in Physical Mathematics and Nonlinear Partial Differential Equations, S.M. Rankin and J.H. Lightbourne eds., Marcel Dekker, New York, pp 75-84 (1985).

[125] Slemrod M., *A limiting "viscosity" approach to the Riemann problem for materials exhibiting a change of phase I*, Arch. Rational Mech. Anal., pp 327-365 (1989).

[126] Slemrod M., *The viscosity-capillarity criterion for shocks and phase transitions*, Arch. Rational Mech. Anal., vol 83, pp 333-361 (1983).

[127] Slemrod M., *Dynamic phase transitions in a van der Waals fluid*, J. Differential Equations, vol 52, pp 1-23 (1984).

[128] Slemrod M. et Flaherty J.E., *Numerical integration of a Riemann problem for a van der Waals fluid*, in Phase Transformations, C.A. Elias and G. John eds., Elsevier Amsterdam, (1986).

[129] Smith R.G., *The Riemann problem in gas dynamics*, Trans. Amer. Math. Soc., vol 249, Number 1, pp 1-50 (1979).

[130] Smoller J., **Shock waves and reaction-diffusion equation**, Springer-Verlag, New-York (1983).

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

[132] Suliciu I., *On the thermodynamics of fluids with relaxation and phase transitions. II-Phase transitions*, Int. J. Engag. Sci., vol 36, pp 949-971 (1998).

[133] Tadmor E., *Entropy conservative finite element schemes*, in "Numerical methods for Compressible Flows Finite Difference Element and Volume Techniques", Proc. Winter Annual Meeting, Amer. Soc. Mech. Engin, T.E. Tezduyar and T.J.R. Hughes, ed., AMD-Vol. 78, pp 149-158 (1986).

[134] Tadmor E., *The numerical viscosity of entropy stable schemes for systems of conservation laws*, Math. of Comput., vol 49, pp 91-103 (1987).

[135] Tadmor E., *A minimum entropy principle in the gas dynamics equations*, ICASE Report 86-33 (1986).

[136] Toro E.F., **Riemann Solvers and Numerical Methods for Fluid Dynamics**, Springer-Verlag (1999).

[137] Truesdell C., **Rational Thermodynamics**, Springer-Verlag, New-York (1984).

[138] Truskinovsky L., *Dynamics of non-equilibrium phase boundaries in a heat conducting nonlinear elastic medium*, J. Appl. Math. and Mech., (PMM) vol 51, pp 777-784 (1987).

[139] Truskinovsky L., *Kinks versus shocks, in "Shock induced transitions and phase structures in general media"*, R. Fosdick, E. Dunn, and H. Slemrod ed., IMA Vol. Math. Appl. 52, Springer-Verlag (1993).

[140] Truskinovsky L., *About the "normal growth" approximation in the dynamical theory of phase transitions*, Contin. Mech. Thermodyn., vol 6, pp 185-208 (1994).

[141] Truskinovsky L., *Structure of an isothermal phase discontinuity*, Soviet Physics Doklady, vol 30, pp 945-948 (1985).

[142] Truskinovsky L., *Transition to detonation in dynamic phase changes*, Arch. Rat. Mech. Anal., vol 123, pp 375-397 (1994).

[143] Van Leer B., Thomas J.L., Roe P.L. et Newsome R.W., *A comparison of numerical flux formulas for the Euler and Navier-Stokes equations*, AIAA paper, 87-1104-CP (1987).

[144] Viviand H. et Veuillot J.P., *Méthodes pseudo-instationnaires pour le calcul d'écoulements transsoniques*, ONERA Publication 1978-4 (1978).

[145] Volpert A.I., *The space BV and quasilinear equations*, Math. Sbornik, vol 73(115) no 2, pp 225-267 (1967).

[146] Whitham J., **Linear and Nonlinear waves**, Wiley, New-York (1974).

[147] Yee H.C. *Construction of Explicit and Implicit Symmetric TVD Schemes and Their Applications*, J. Comp. Phys., vol 68, pp 151-179 (1987).

[148] Zel'dolvich Y. et Raizer Y.P., **Physics of shock waves and high-temperature hydrodynamic phenomena**, vol 1 et 2, Academic Press (1966).

[149] Zhong X.G., Hou T.Y. et LeFloch P.G., *Computational methods for propagating phase boundaries*, J. Comput. Phys., vol 124, pp 192-216 (1996).

Résumé

La présente recherche doctorale en analyse numérique (et calcul scientifique) aborde le problème du contrôle de la dissipation d'entropie numérique associée à une discrétisation donnée. Cette problématique constitue un véritable challenge numérique non encore complètement résolu à ce jour.

Le travail se décompose en deux parties principales dont les caractéristiques sont réellement différentes.

La première partie concerne l'approximation numérique des solutions (instationnaires en 1D et stationnaires en 2D) du système des équations de Navier-Stokes à plusieurs pressions indépendantes. Ce système est *hyperbolique* et possède des champs *vraiment non linéaires* sous des hypothèses classiques, mais s'écrit naturellement sous forme *non conservative.*

La deuxième partie est dédiée à l'approximation numérique des solutions instationnaires en 1D de quelques systèmes de *lois de conservation* de type soit *hyperbolique* mais dont les champs possèdent un *défaut de vraiment non linéarité*, ou soit mixte *hyperbolique-elliptique.*

Dans toutes ces situations motivées par des applications physiques concrètes, le contrôle de la dissipation d'entropie joue un rôle déterminant dans la caractérisation des solutions recherchées. Les schémas numériques proposés dans ce manuscrit sont obtenus par une analyse fine des bilans d'entropie associés.

Abstract

In this thesis (numerical analysis and scientific computation areas), we are interested in the knowledge of the numerical entropy dissipation associated with a given numerical scheme. It turns out that generally speaking, this problem is still open while being of crucial interest in many applications.

Our study is composed of two distinct parts.

The first one is concerned with the numerical approximation of the (unsteady in 1D and steady in 2D) solutions of the Navier-Stokes equations involving several independent pressure laws. As in the usual setting of a single pressure law, this system is *hyperbolic* with *genuinely non linear* associated fields under classical assumptions. However, it naturally writes in *non conservation* form.

The second one deals with the numerical approximation of the unsteady solutions (in 1D) of several systems of *conservation laws* which are either *hyperbolic* but with *non genuinely non linear* (and *non linearly degenerate*) fields, or mixed *hyperbolic-elliptic.*

Several models from the physics enter the present framework for which the entropy dissipation plays an important role in the selection of the physical solution. Here we propose several relevant numerical schemes designed on the basis of a precise study of the associated entropy dissipation.