Étude mathématique et numérique d'équations hyperboliques non-linéaires : couplage de modèles et chocs non classiques.

Benjamin Boutin

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PIERRE ET MARIE CURIE

École Doctorale de Sciences Mathématiques de Paris Centre (ED386)

Spécialité : Mathématiques Appliquées

présentée par

Benjamin Boutin

pour l'obtention du titre de

Docteur de l'Université Pierre et Marie Curie - Paris 6

Sujet de la thèse :

ÉTUDE MATHÉMATIQUE ET NUMÉRIQUE D'ÉQUATIONS HYPERBOLIQUES NON-LINÉAIRES : COUPLAGE DE MODÈLES & CHOCS NON CLASSIQUES

Organismes d'accueil :

Commissariat à l'Énergie Atomique – Centre de Saclay
DEN/DANS/DM2S/SFME/LETR
Laboratoire Jacques-Louis LIONS – UMR 7598

Soutenue publiquement le 27 novembre 2009
après avis des rapporteurs et devant le jury

François Bouchut	Université Paris Est	*Rapporteur*
Frédéric Coquel	Université Pierre et Marie Curie	*Examinateur*
Jean-François Coulombel	Université Lille 1	*Rapporteur*
Bruno Després	Université Pierre et Marie Curie	*Président*
Francis Filbet	Université Lyon 1	*Examinateur*
Thierry Goudon	INRIA Lille	*Examinateur*
Philippe G. LeFloch	Université Pierre et Marie Curie	*Directeur*
Jacques Segré	CEA de Saclay	*Invité*

Manuscrit conçu et mis en page à l'aide de LaTeX et KOMA-Script,
illustrations réalisées avec PGF, Ti*k*Z et Gnuplot.
Version du 17 novembre 2009.

Remerciements

Je tiens tout d'abord à adresser toute ma reconnaissance à mon directeur de thèse Philippe G. LeFloch. Je le remercie d'avoir accepté de diriger mes travaux et de m'avoir accordé sa confiance au cours de ces trois années.

Je remercie mes deux rapporteurs François Bouchut et Jean-François Coulombel pour l'intérêt qu'ils ont porté à mon manuscrit, ainsi que l'ensemble des membres du jury qui ont accepté d'examiner mon travail et m'ont honoré de leur présence : Bruno Després, Francis Filbet et Thierry Goudon.

J'ai travaillé dans deux laboratoires distincts : au CEA de Saclay SFME/LETR et au Laboratoire Jacques-Louis Lions. Je tiens à en remercier les directions respectives ainsi que l'ensemble des personnels qui contribuent à faire de ces lieux bien plus que des locaux de recherche ! Plus particulièrement, je remercie Daniel Caruge, Eric Royer, Danielle Gallo-Lepage, Jacques Segré, Ketty Drame et Marie-Claude Rolland d'une part, Yvon Maday, Edwige Godleswki, Liliane Ruprecht, Danielle Boulic, Christian David et Khashayar Dadras d'autre part.

Je remercie au passage Annalisa Ambroso de m'avoir accueilli et épaulé durant mes premiers mois de recherche, à commencer par mon stage de M2 au CEA qui aura marqué le point de départ de toute cette aventure.

Ce n'est pas sans émotion que j'adresse également mes remerciements chaleureux à Frédéric Coquel et Frédéric Lagoutière. Leur gentillesse sans égal ainsi que leurs qualités humaines et scientifiques, que chacun sait, m'ont beaucoup apporté !

Je remercie également l'ensemble des membres du groupe de travail « écoulements diphasiques » pour m'avoir offert de participer à leurs réunions hebdomadaires, grands moments de *chocs* et de *détentes*, je cite dans l'ordre alphabétique : Annalisa Ambroso, Christophe Chalons, Frédéric Coquel, Edwige Godlewski, Jean-Marc Hérard, Samuel Kokh, Frédéric Lagoutière, Pierre-Arnaud Raviart, Jacques Segré et Nicolas Seguin.

Je n'oublie pas de remercier également les enseignants avec lesquels j'ai pu échanger durant mon monitorat et plus particulièrement Sidi M. Kaber mon tuteur, mais aussi Muriel, Lahcen et Marie-Aude.

Vient à présent le moment de remercier tous ceux et celles qui ont partagé mon quotidien, pour le pire mais aussi pour le meilleur : au LETR Frédérique, Floraine, Gloria (docteur ès LaTeX), Marie (l'astronaute et le truc bleu), Thierry, Thomas et Yohan ; au LJLL, je remercie

de même tous les thésards ou anciens thésards du 1D01 (et il y en a !) : André, Fiammetta, Florian, Imen, Jian, Kamel, Kirill, Lamjed, Nicolas, Noura, Sever, Sonia, Tobias et Valérie, et parmi les nombreux thésards des autres bureaux, je remercie Alexandra, Alexis, Ange, Évelyne, Hassan, Laurent, Mathieu, Nicolas, Rachida et Thomas. Merci à vous tous et à ceux que j'oublierais.

Ces trois années ne furent heureusement pas seulement mathématiques, aussi je remercie mes amis musiciens et mélomanes pour les concerts partagés devant ou sur la scène : Adrien, Augustin, Aurélie, Cécile, Céline, Deniche, Fabrice, Florent, Jeff, Jérémy, Laguinde, Lo, Tias, Tom, Wédé, et bien sûr, *de par sa chandelle verte*, PU, MU et Bouli à qui nous aurons su sécher *les larmes du ventre tordu*.

Enfin je remercie mes parents, ma sœur, mes grands-parents et toute ma famille pour leur soutien et leurs encouragements malgré l'éloignement.

TABLE DES MATIÈRES

INTRODUCTION GÉNÉRALE

Cette thèse s'intéresse à deux problématiques dans le cadre des systèmes d'équations aux dérivées partielles hyperboliques, toutes deux abordées selon une approche liée à l'étude de l'influence des phénomènes de petites échelles sur les solutions. La première partie concerne le couplage interfacial de deux systèmes hyperboliques de lois de conservation. Dans une seconde partie, nous traitons du calcul numérique de solutions non classiques de lois de conservation.

Dans le contexte des EDP hyperboliques, les solutions développent naturellement des discontinuités, ceci du fait des non-linéarités. L'absence conséquente d'unicité des solutions faibles nécessite l'introduction d'un critère de sélection, le plus souvent construit sur des mécanismes de régularisation qui auront été négligés dans une première approximation. Ces mécanismes peuvent parfois être restitués au travers de propriétés macroscopiques telles que les inégalités d'entropie mais ce n'est pas toujours le cas. Ces préoccupations motivent les études menées dans chacun des deux cadres abordés ici.

Le couplage de deux systèmes d'EDP hyperboliques consiste à résoudre deux lois de conservation séparées par une interface fixe au travers de laquelle une condition de couplage doit être formulée. Elle prend ici la forme d'une condition de transmission ou encore, dans les cas favorables, elle exprime une propriété de continuité de la solution à l'interface. Un comportement résonnant de l'interface peut survenir qui met alors en défaut cette dernière propriété de continuité et peut conduire à des solutions discontinues multiples. Dans ce mémoire, l'approche proposée pour obtenir un critère de sélection dans ce cadre résonnant consiste à plonger le problème sur $\mathbb{R}$ tout entier en conférant à l'interface une structure d'onde stationnaire pour un système hyperbolique augmenté. Cette démarche permet alors d'introduire une régularisation visqueuse du problème couplé et d'en déduire un critère de sélection. En outre, la formulation proposée supporte également un autre mécanisme de régularisation consistant à épaissir les interfaces. Dans ce cas, l'unicité est assurée. Enfin la formulation augmentée autorise aussi la définition d'un couplage multidimensionnel et multicomposante avec éventuel recouvrement des domaines. Un schéma numérique conservatif préservant les équilibres est mis en œuvre pour le calcul des solutions de ce couplage régularisé.

Dans la seconde partie, nous abordons une autre stratégie de sélection des solutions discontinues d'une EDP hyperbolique. Un critère alternatif à la simple viscosité évanescente consiste à tenir compte également des phénomènes d'ordre supérieur. Ainsi, lorsque les différents phénomènes microscopiques qui interviennent sont dans un certain régime, la solution, dite alors *non classique*, diffère de celle obtenue avec le simple critère visqueux. En particulier, elle ne vérifie plus autant d'inégalités d'entropie mais l'influence des phé-

nomènes de petites échelles peut être restituée macroscopiquement par le biais d'autres critères, notamment par ce qu'on appelle la *relation cinétique*. L'étude menée dans cette deuxième partie concerne l'élaboration d'un schéma numérique pour le calcul de ces solutions. Ce schéma s'appuie sur une relation cinétique donnée plutôt que de rechercher une consistance d'ordre élevé avec les phénomènes fins. Il a la particularité d'être à la fois conservatif, « non-oscillant » et de néanmoins très peu diffuser les chocs non classiques.

Première Partie :
Couplage interfacial de deux systèmes hyperboliques

Contexte physique et applicatif. La première Partie de ce manuscrit concerne le couplage interfacial de deux systèmes hyperboliques. Cette thématique de recherche s'inscrit dans les problématiques étudiées au Laboratoire d'Études Thermiques des Réacteurs du CEA-Saclay, plus précisément dans le projet NEPTUNE [78]. Ce projet, auquel participent le CEA, EDF, AREVA-NP et l'IRSN, vise à développer les outils numériques pour la simulation des écoulements fluides dans les réacteurs nucléaires notamment de la thermohydraulique diphasique. L'écoulement du fluide dans le circuit est sujet à différents phénomènes physiques, mettant chacun en jeu des échelles d'espace différentes. Voici des exemples de tels phénomènes qui apparaissent typiquement dans les cœurs de réacteurs nucléaires et les générateurs de vapeur : les interfaces de changement de phase, les interfaces entre milieux poreux et non-poreux et plus généralement les phénomènes physiques variant fortement localement. Les codes de calculs existants ou en développement traitent séparément ces phénomènes et ceci dans des segments séparés du circuit. La démarche numérique adoptée consiste à coupler ces codes en faisant passer l'information jugée utile des uns vers les autres. L'enjeu applicatif consiste à redonner une nouvelle vie aux plateformes de simulation existantes fortement onéreuses lors de leur mise en place et de leur maintenance. Une approche plus mathématique adoptée dans ce travail consiste à étudier en amont le couplage des modèles physiques avant d'envisager la résolution numérique.

L'étude mathématique et numérique du couplage appelle à des questions intéressantes dans le domaine des équations aux dérivées partielles hyperboliques. Elle implique notamment des problèmes liés directement à la perte d'hyperbolicité stricte, aux formulations non conservatives ou encore à la résonance. Dans ce cadre, une collaboration[1] a vu le jour entre le CEA/DM2S et le LJLL. Elle est chargée d'analyser cette problématique de couplage interfacial et de proposer des outils numériques et théoriques adaptés.

Une brève revue du formalisme de couplage

Relation de couplage. Avant de poursuivre, fixons quelques notations. Nous nous intéressons au couplage de deux systèmes de lois de conservation de même taille, portant sur une inconnue $w = w(t,x) \in \mathbb{R}^N$

$$\begin{cases} \partial_t w + \partial_x f_-(w) = 0, & x < 0,\ t > 0, \\ \partial_t w + \partial_x f_+(w) = 0, & x > 0,\ t > 0. \end{cases} \tag{1}$$

À ces lois de conservation vient s'ajouter une condition initiale $w(0,x) = w_0(x)$ à l'instant $t = 0$. Pour espérer avoir un problème bien posé mathématiquement, il est indispensable d'ajouter une information reliant les deux demi-espaces en $x = 0$, c'est cette relation que nous appellerons *relation de couplage*. Une infinité de possibilités s'offre à nous pour le choix d'une telle relation, choix qui peut être guidé par des considérations physiques, notamment

[1] http://www.ann.jussieu.fr/groupes/cea/

une certaine propriété de continuité à l'interface. On peut ainsi considérer la continuité de certaines composantes de la variable, ou d'une transformation non-linéaire de celles-ci, ou encore sélectionner des états stationnaires, c'est-à-dire les solutions qui n'évoluent pas au cours du temps. Nous reviendrons sur ce point ultérieurement.

Couplage par état versus couplage conservatif. Une première approche pour effectuer le couplage consiste à considérer l'interface comme « conservative » – on parle alors de *couplage conservatif* – dans le sens où la solution réalise la continuité du flux à l'interface :

$$f_-(w(t,0^-)) = f_+(w(t,0^+)). \tag{2}$$

En définissant un flux global $f(w,x)$ non homogène en espace

$$f(w,x) = \begin{cases} f_-(w), & x < 0, \\ f_+(w), & x > 0, \end{cases} \tag{3}$$

la formulation globale de (1)-(2) prend alors la forme

$$\partial_t w + \partial_x f(w,x) = 0, \quad x \in \mathbb{R},\ t > 0. \tag{4}$$

De cette manière, l'expression de la conservation de w au voisinage de $x = 0$ restitue la relation (2). On constate que cette approche concerne l'étude des lois de conservation à coefficients discontinus, abordée par exemple dans les travaux de Bachmann [15], Bachmann et Vovelle [16], Bürger et Karlsen [36], Godlewski et Seguin [77], Helluy et Seguin [82], Karlsen *et al.* [92], Seguin et Vovelle [121].

Cette vision des choses ne concerne pas tous les cadres d'applications. Pour certains problèmes physiques, on ne peut exiger la conservation de toutes les inconnues du système. Citons l'exemple des modèles 1D d'écoulement en tuyère à section discontinue. Dans ce cas une perte de charge peut être observée au niveau des discontinuités du milieu, disons en $x = 0$, dont les valeurs empiriques sont rapportées dans des abaques. Un terme source mesure localisé à la discontinuité doit alors être ajouté à l'équation pour en tenir compte :

$$\partial_t w + \partial_x f(w,x) = \mathcal{M}(t)\delta_{x=0}, \tag{5}$$

où $\mathcal{M}(t)$ représente la masse de cette perte de charge.

Un modèle alternatif suggéré par cette observation est celui du *couplage par état*. Dans ce modèle de couplage, on requiert à l'interface la continuité d'un jeu de variables privilégiées, qui peut être l'inconnue w ou une transformation non-linéaire de celle-ci, choisie selon la physique. La relation de couplage attendue à l'interface prend ainsi la forme

$$\theta_-(w(t,0^-)) = \theta_+(w(t,0^+)), \tag{6}$$

où les fonctions $\theta_\pm$ sont deux changements de variables admissibles d'inverses notés dans la suite $\gamma_\pm$ (cf. Figure 1). Cette relation de couplage se réécrit également comme une égalité

des traces pour la variable privilégiée $u = \theta_{\pm}(w)$

$$u(t,0^-) = u(t,0^+). \tag{7}$$

Dans les travaux de Caetano [38], Galié [67], les couplages conservatifs ou par état ont été étudiés notamment dans leur application aux équations d'Euler ou aux modèles fluides diphasiques. D'autres approches y sont proposées permettant par exemple d'envisager le couplage de deux lois de conservation portant sur des variables de dimensions différentes.

Prenons l'exemple du couplage de deux systèmes d'Euler de la dynamique des gaz en coordonnées de Lagrange pour des lois d'état différentes $p = p_{\pm}(\tau,\epsilon)$. Le système prend la forme (1) avec $w = (\tau,v,e)^T$, $f_{\pm}(w) = (-v,p,pv)^T$, où τ est le volume spécifique, v la vitesse et e l'énergie totale. L'interface est systématiquement caractéristique puisque $\lambda = 0$ est toujours valeur propre pour chacun des deux systèmes gauche et droit, mais n'est jamais résonnante, les autres valeurs propres associées à des champs non-linéaires ne pouvant s'annuler. Le problème est ainsi bien posé (hors le cas du vide) : Ambroso *et al.* [10] montrent que le couplage réalisé avec la condition (7) exprimée sur le vecteur $u = (\tau,v,p)^T$ conduit à une unique solution, qui réalise la continuité de la vitesse et de la pression à l'interface

$$v(t,0^-) = v(t,0^+), \quad p(t,0^-) = p(t,0^+). \tag{8}$$

La continuité de τ n'est par contre pas assurée a priori, bien que le flux soit continu $f_-(w(t,0^-)) = f_+(w(t,0^+))$, l'interface étant caractéristique.

Le couplage formulé en coordonnées d'Euler, avec cette fois-ci donc $w = (\rho,\rho v,\rho e)^T$, et $f_{\pm}(w) = (\rho u,\rho v^2+p,(\rho e+p)v)^T$, présente en revanche un caractère résonnant : les valeurs propres associées à des champs vraiment non-linéaires peuvent changer de signe. Les études menées par Chalons *et al.* [49] illustrent la perte d'unicité conséquente (la relation de couplage étant comprise alors au sens affaibli (10) expliqué plus loin). De plus, Ambroso *et al.* [12] observent que le choix des variables $u = (\rho,v,p)^T$, s'il permet de conserver la densité ρ et la quantité de mouvement ρu, n'assure pas la conservation de l'énergie ρe. Mieux que cela, ils montrent que la conservation de l'énergie et la préservation des états de vitesse et de pression constantes sont incompatibles.

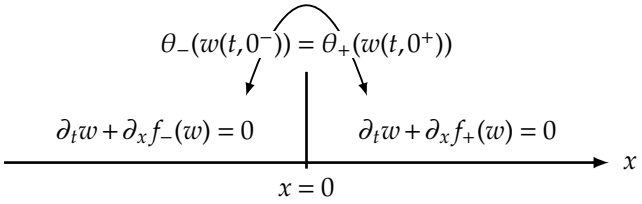


FIGURE *1: Représentation de la problématique de couplage.*

Formulation faible de la relation de couplage. La relation liant les traces de la solution de part et d'autre de l'interface s'interprète en termes de double jeu de conditions aux limites. La trace à gauche $w(t,0^-)$ issue du problème posé sur $\mathbb{R}_-$ sert ainsi de condition de bord pour le problème posé sur $\mathbb{R}_+$ tandis que réciproquement la trace à droite $w(t,0^+)$

issue du problème posé sur $\mathbb{R}_+$ sert de condition de bord pour le problème posé sur $\mathbb{R}_-$. Du fait du caractère hyperbolique des systèmes en considération dans chacun des demi-espaces, les conditions aux limites ne sont pas réalisables *stricto sensu*, et un formalisme affaibli est nécessaire. C'est dans ce cadre affaibli, détaillé ci-après, que Ambroso *et al.* [7; 10; 11], Godlewski *et al.* [73], Godlewski et Raviart [75; 76] ont développé l'étude du couplage.

La formulation affaiblie envisagée dans ces études du couplage provient de l'étude menée par Dubois et LeFloch [65] dans le contexte de l'étude des conditions aux limites pour un problème hyperbolique. La donnée de bord à l'interface, $w(t,x=0)=b(t)$, où $b : t \in \mathbb{R}_+ \mapsto b(t) \in \Omega$, est affaiblie sous la forme d'une condition ensembliste $w(t,0^+) \in \mathcal{O}(b(t))$ où, à chaque instant t fixé, $\mathcal{O}(b(t)) = \left\{\mathcal{W}(0^+,b(t),w), w \in \Omega\right\}$ désigne l'ensemble des traces admissibles en $x=0^+$ des solutions du problème de Riemann avec pour donnée gauche $b(t)$. Ici, $\xi \mapsto \mathcal{W}(\xi,w_L,w_R)$ désigne la solution entropique autosemblable (i.e. ne dépendant que de $\xi = x/t$) du problème hyperbolique en considération et vérifiant la donnée initiale

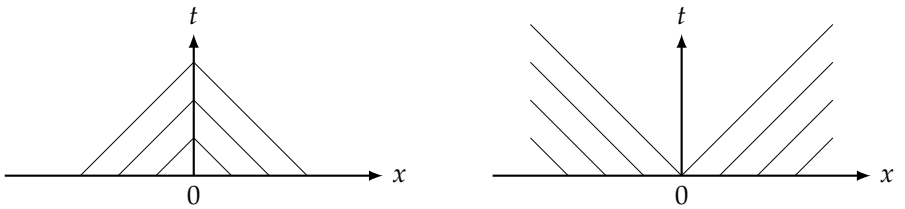
$$w(t=0,x) = \begin{cases} w_L, & x<0, \\ w_R, & x>0. \end{cases} \tag{9}$$

La condition (6) est donc comprise à la manière ensembliste

$$\begin{cases} w(t,0^-) \in \mathcal{O}_-(\theta_-^{-1}(\theta_+(w(t,0^+)))), \\ w(t,0^+) \in \mathcal{O}_+(\theta_+^{-1}(\theta_-(w(t,0^-)))), \end{cases} \tag{10}$$

où $\mathcal{O}_-$ et $\mathcal{O}_+$ sont les ensembles de traces admissibles pour les problèmes de Riemann associés respectivement aux flux f_- et f_+ : $\mathcal{O}_-(b(t)) = \left\{\mathcal{W}_-(0^-,w,b(t)), w \in \Omega\right\}$, $\mathcal{O}_+(b(t)) = \left\{\mathcal{W}_+(0^+,b,w), w \in \Omega\right\}$.

On retiendra que lorsqu'aucune onde n'interagit avec l'interface, la condition affaiblie (10) correspond essentiellement à la condition (6), sinon il y a perte d'unicité. Ce phénomène de *résonance* se produit notamment lorsque les vitesses caractéristiques changent de signe de part et d'autre de l'interface (cf. Figure 2). Ce défaut d'unicité a déjà été observé dans le cas de systèmes hyperboliques sous forme non conservative avec résonance comme étudiés par Goatin et LeFloch [71], Isaacson et Temple [86; 87] ou encore par Temple [129].



(a) $f'_-(w(t,0^-)) > 0$ *et* $f'_+(w(t,0^+)) < 0$ (b) $f'_-(w(t,0^-)) < 0$ *et* $f'_+(w(t,0^+)) > 0$

Figure 2: *Exemples scalaires de situations avec interface résonnante.*

Il faut bien comprendre que la non-unicité observée dans ce cadre est intrinsèque au

problème du couplage et apparaît dans ce formalisme mathématique par la présence de la résonance. Elle est imputable à plusieurs faits.
Tout d'abord l'interface n'est pas vraiment modélisée : qu'elle représente une réalité physique ou pas, la non-unicité révèle la nécessité d'introduire une information supplémentaire. En d'autres termes, la physique « grossière » ne suffit pas à assurer un problème bien posé, et les effets plus « fins » au niveau de l'interface doivent dès lors être pris en compte. Pour le couplage conservatif, différentes approches orientées autour de la formulation d'une condition entropique à l'interface ont été élaborées, citons Audusse et Perthame [13], Bürger *et al.* [37], Diehl [62], Seguin et Vovelle [121] pour l'étude du problème de Cauchy conservatif à flux discontinu.
D'autre part, comme le soulignent Lin et Schecter [113], même en présence d'un critère visqueux, les solutions du problème de Riemann peuvent être multiples (cf. Azevedo et Marchesin [14]). Cette observation se comprend ainsi : ces solutions auto-semblables réalisent l'asymptotique en temps grand des solutions du problème de Cauchy obtenues avec différentes données initiales $w_0(x)$, ces données correspondant toutes à une même donnée de bord du problème de Riemann (w_L, w_R) : $w_L = \lim_{-\infty} w_0(x)$ et $w_R = \lim_{+\infty} w_0(x)$. C'est ce point qui motivera dans la suite l'emploi particulier de l'approximation parabolique à la Dafermos dans ce cadre résonnant.

États stationnaires. On appelle états stationnaires les solutions invariantes au cours du temps. Il apparaît que les constantes sur $\mathbb{R}$ ne sont pas solutions en général pour le problème conservatif (4) dès que $f_- \neq f_+$, puisqu'au contraire les seuls états stationnaires pour cette équation sont ceux satisfaisant à la relation de couplage conservatif (2). De la même manière, les états stationnaires pour le problème avec chargement mesure (5) sont ceux pour lesquels le saut $[f(w(x),x)]_{x=0^-}^{x=0^+}$ égale précisément le chargement $\mathcal{M}$ du terme source. Ainsi, dans le cas d'une solution w de la forme

$$w(t,x) = \begin{cases} w_L, & x < 0, \\ w_R, & x > 0, \end{cases} \tag{11}$$

en choisissant précisément

$$\mathcal{M} = f_+(w_L) - f_-(w_R), \tag{12}$$

les solutions à w constant, $w_L = w_R$, seront états stationnaires du couplage (5). Cette méthode de couplage s'étend naturellement à des chargements mesure dépendant du temps $\mathcal{M}(t)$. Elle a notamment été développée et mise en œuvre numériquement dans les travaux de Galié (cf. Chap.2 et 3 [67]) et Godlewski [72].

Enfin, par un choix adapté de $\mathcal{M}(t)$, on peut dès lors s'assurer que les solutions vérifiant la relation (6), i.e. telles que $u = \theta_\pm(w)$ vérifie (7), seront des états stationnaires du couplage (5). Pour cela, on répercute le changement de variable $\theta_\pm$ dans l'équation qui prend la forme

$$A_0(u,x)\partial_t u + A_1(u,x)\partial_x u = 0, \tag{13}$$

où les matrices A_0 et A_1 sont formellement obtenues à partir des changements de variables

$\gamma_\pm$ et des flux $f_\pm$

$$\begin{aligned} A_0(u,x) &= \nabla\gamma_\pm(u), \quad \pm x > 0, \\ A_1(u,x) &= \nabla f_\pm(\gamma_\pm(u))\nabla\gamma_\pm(u), \quad \pm x > 0. \end{aligned} \tag{14}$$

C'est sur cette base que nous allons mener notre étude.

Contributions

Dans ce travail, le couplage est reformulé comme un système d'EDP augmenté. Une variable supplémentaire v dite « fonction couleur » permet de s'affranchir de la description géométrique de la position de l'interface et de poser le problème sur $\mathbb{R}$ tout entier. Cette variable satisfait à une équation d'évolution triviale et décrira ainsi l'onde stationnaire présente à l'interface. Reformulé dans la variable $u \in \mathbb{R}^N$ qui attend la relation de couplage (7), le système augmenté s'écrit alors

$$\begin{cases} A_0(u,v)\,\partial_t u + A_1(u,v)\,\partial_x u = 0, \\ \partial_t v = 0. \end{cases} \tag{15}$$

En choisissant comme donnée initiale, pour v, une fonction de type Heaviside dont la discontinuité est placée précisément à l'interface, la relation de couplage est alors restituée au travers des invariants de Riemann pour cette onde stationnaire supplémentaire. Pour fixer les idées, cette variable prend par exemple ses valeurs dans $[0,1]$ et chacune des deux valeurs 0 et 1 permet de retrouver respectivement les modèles couplés gauche et droit tandis que les valeurs intermédiaires permettent de passer de manière continue de l'un vers l'autre. On peut alors typiquement penser pour la définition des fonctions $A_0(u,v)$ et $A_1(u,v)$ à des objets de la forme

$$\begin{aligned} A_0(u,v) &= (1-v)\nabla\gamma_-(u) + v\nabla\gamma_+(u), \\ A_1(u,v) &= (1-v)\nabla f_-(\gamma_-(u))\nabla\gamma_-(u) + v\nabla f_+(\gamma_+(u))\nabla\gamma_+(u). \end{aligned} \tag{16}$$

En effet, tant que $A_1(u,v)$ reste inversible, les composantes de u sont les invariants de Riemann pour l'onde de vitesse nulle associée à la discontinuité en v. Leur continuité, attendue lorsque (15) forme un système hyperbolique, restaure ainsi la condition de couplage (7). En revanche, si 0 est valeur propre de $A_1(u,v)$, le système augmenté n'est plus que faiblement hyperbolique, au sens où la base de diagonalisation n'existe plus localement. C'est ainsi que se manifeste la résonance. Cela se produit par exemple dans le cas scalaire lorsque les quantités $f'_-(u)$ et $f'_+(u)$ n'ont pas le même signe.

La formulation (15) présente trois intérêts principaux que nous exploiterons dans la suite.

1. Elle permet de doter l'interface de mécanismes visqueux et de rendre ainsi compte des phénomènes de petites échelles susceptibles de lui conférer une structure et d'obtenir des critères d'admissibilité pour les solutions.

2. Elle suggère un formalisme d'interface épaissie ou régularisée.

3. Elle s'étend aisément au cas multidimensionel et multicomposante.

Nous précisons ci-après les apports sur chacune de ses orientations avant de revenir enfin aux détails techniques.

Structure de l'interface. Dans le contexte du couplage par état, on ne dispose pas de notion d'entropie naturelle à l'interface. Un moyen de récupérer des critères de sélection des solutions consiste donc à retourner au niveau détaillé des mécanismes dissipatifs pour avoir accès à la structure interne de l'interface. Par analogie avec les approches visqueuses menant aux inégalités d'entropie dans un cadre plus habituel, la formulation (15) peut faire l'objet d'une régularisation parabolique. Nous employons la régularisation proposée par Dafermos [57] afin de concentrer l'étude sur les états stationnaires en temps grand dont nous avons déjà souligné l'importance

$$\begin{cases} A_0(u^\epsilon, v^\epsilon)\partial_t u^\epsilon + A_1(u^\epsilon, v^\epsilon)\partial_x u^\epsilon = \epsilon t\, \partial_x\Big(B_0(u^\epsilon, v^\epsilon)\partial_x u^\epsilon\Big), \\ \partial_t v^\epsilon = \epsilon^2 t\, \partial_{xx} v^\epsilon. \end{cases} \tag{17}$$

Cette formulation permet de donner un sens au produit non conservatif (cf. Dal Maso *et al.* [60], LeFloch et Liu [99], LeFloch [100]) et est en mesure de traiter la résonance malgré l'absence de base de diagonalisation pour le système faiblement hyperbolique. Ce type de régularisation, particulièrement adapté à l'analyse des solutions auto-semblables, a fait l'objet de plusieurs études notamment par Lin et Schecter [113], Schecter [118; 119], Schecter et Szmolyan [120], Slemrod [124].

Interface épaisse. Le modèle augmenté dans le régime d'interface infiniment mince ne permet pas d'assurer l'unicité des solutions relativement à la donnée initiale en général. Le système peut admettre plusieurs solutions (auto-semblables) notamment lorsque la résonance se produit. Certaines de ces solutions pourraient être instables sous perturbation de la donnée initiale. En revanche, le modèle EDP régularisé (dans le régime d'interface épaisse) prend la forme d'un système hyperbolique quasi-linéaire avec termes source réguliers, qui a la propriété d'être bien posé. Pour formuler ce modèle, on utilise une fonction couleur v régularisée, modélisant une interface « épaisse », et w solution de

$$\partial_t w + \partial_x f(w, v) = \partial_v f(w, v)\partial_x v. \tag{18}$$

Cette équation n'est rien d'autre qu'une régularisation de la formulation (5) vis à vis de l'inhomogénéité spatiale du flux et du terme source, et on dispose d'un résultat d'unicité pour les solutions entropiques à la Kružkov dans le cas scalaire [94].

Une nouvelle approche numérique est développée pour cette résolution. Elle repose sur une stratégie « équilibre », avec reconstruction de la solution discrète dans chaque maille, afin de rendre compte du terme source. Cette étape de reconstruction rend difficile une éventuelle estimation BV uniforme (elle demeure en fait inconnue) et les résultats de convergence reposent sur les inégalités d'entropies dans le cadre de solutions à valeur mesures de DiPerna [64]. Cette méthode a déjà fait ses preuves dans les travaux antérieurs

de Cockburn *et al.* [50], Coquel et LeFloch [55]. Par ailleurs, la solution numérique obtenue converge bien vers l'unique solution du problème régularisé considéré, et n'est plus sensible au flux numérique employé. Nous voyons dans le Chapitre 1 qu'une approche plus naïve révèle une sensibilité au flux numérique utilisé.

Couplage multidimensionnel et multicomposante. La formulation (15) permet de quitter le cadre académique d'une interface ponctuelle en scalaire et d'envisager de traiter des situations plus complexes. En effet, l'essentiel de la géométrie du problème est restitué par la définition de la variable de couleur v. On peut dès lors se permettre de considérer un couplage entre, disons $L+1$ domaines $(\mathcal{D}_l)_{0\leq l\leq L}$ formant une partition finie de $\mathbb{R}^N$, de frontières suffisamment régulières. Il suffit pour cela de travailler avec une fonction couleur à valeurs vectorielles $v : \mathbb{R}^N \to \mathbb{R}^L$. Un exemple de configuration envisageable est donné sur la Figure 3 avec $N = 2$ et $L = 2$.

Les conditions de couplage sont alors définies en considérant les traces relativement à la normale à l'interface. Nous développons dans ce cadre, les interfaces une fois encore épaissies et régularisées, un schéma équilibre analogue au précédent.

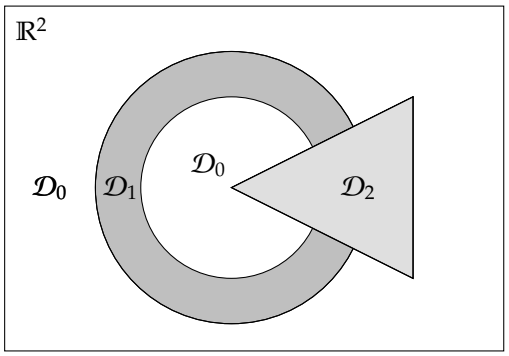


FIGURE 3: *Exemple de configuration de couplage multidimensionnel et multicomposante.*

Résumé détaillé de la première Partie

Résultat d'existence pour le problème de Riemann couplé. Un premier travail porte sur l'analyse du couplage de deux lois de conservation scalaires à une interface spatiale fixe, couplage formalisé par le recollement de deux demi-problèmes de Riemann avec condition de bord. La condition algébrique de saut prescrite (au sens faible) au niveau de l'interface est ainsi reformulée comme une double relation de compatibilité sous la forme ensembliste (10). L'étude exhaustive des solutions du problème de Riemann couplé est alors menée dans le Chapitre 1. Nous obtenons par un procédé constructif le résultat suivant :

Théorème 1.10 p.47 : *Supposons que les fonctions flux f_- et f_+ sont de classe C^1, alors dans le cas scalaire $N = 1$ le problème de Riemann couplé (1)-(10) admet une solution autosemblable, non nécessairement unique.*

Les solutions sont obtenues par une description géométrique des ensembles de traces en considération à partir du graphe des deux fonctions de flux. Cette description s'apparente à

la construction des solutions classiques pour le problème de Riemann, admissibles au sens d'Oleinik, à partir de l'enveloppe convexe ou concave de la fonction flux. Cette construction permettra en outre de révéler des possibles cas de non-unicité de la solution, lorsque la résonance est présente. Cette non-unicité peut être de plusieurs types : existence d'un continuum de solutions toutes continues à l'interface, solutions différentes discontinues à l'interface. Des exemples de tels cas de non-unicité seront construits et mis à l'épreuve d'approches numériques différentes, ce qui, on le verra, révèle une sensibilité de la solution numérique calculée au schéma employé.

Analyse du couplage résonnant par la méthode de Dafermos. Dans le Chapitre 2, nous proposons une seconde reformulation mathématique du problème de couplage. Une variable additionnelle v dite « fonction couleur » est introduite afin de passer d'une formulation géométrique ($x < 0$ et $x > 0$) à une formulation EDP du problème sur le domaine tout entier ($x \in \mathbb{R}$). Le caractère mal posé du problème qui apparaissait précédemment au travers de la résonance est alors perçu dans la possible perte de stricte hyperbolicité du système EDP augmenté. Dans tout ce chapitre, le travail est effectué dans le cadre du couplage de deux systèmes et la variable $u \in \mathbb{R}^N$ représente la variable continue à l'interface. Le problème s'écrit alors sous la forme (15).

Cette formulation étant non conservative et seulement faiblement hyperbolique, l'analyse par approximation visqueuse qui sera employée se révèle contourner ces difficultés. Nous étudions dès lors le phénomène de résonance qui survient à l'interface de couplage en suivant l'analyse proposée par Dafermos [57; 58], qui consiste à observer le comportement des états stationnaires autosemblables en temps grand, après une approximation visqueuse (parabolique) de la solution. Le système approché prend la forme (19) suivante

$$\begin{cases} \Big(-\xi A_0(u^\epsilon, v^\epsilon) + A_1(u^\epsilon, v^\epsilon) \Big) u^\epsilon_\xi = \epsilon \, \Big(B_0(u^\epsilon, v^\epsilon) u^\epsilon_\xi \Big)_\xi \\ -\xi v^\epsilon_\xi = \epsilon^2 v^\epsilon_{\xi\xi} \end{cases} \tag{19}$$

Cette approche est destinée à capturer les états pertinents pour le problème couplé après l'avoir reformulé comme un problème EDP posé sur le domaine tout entier. Il remédie par ailleurs à l'absence d'une base de vecteurs propres pour (19), pour déterminer les solutions lorsque la résonance se produit. À noter que les états stationnaires ne sont pas comme habituellement les simples solutions constantes dans le contexte du couplage et une structure non-triviale à l'interface vient s'ajouter. On se propose de l'étudier également.

Cette analyse permet d'obtenir de nouveau un résultat d'existence pour des données gauche et droite suffisamment proches dans le cas systèmes et sans hypothèse de proximité dans le cas scalaire. Le principe de démonstration s'appuie en premier lieu sur une formule de représentation implicite de la solution u^ϵ comme combinaison d'ondes caractéristiques. Des estimations contrôlant les termes d'interactions entre ces différentes ondes permettent alors d'appliquer des arguments de point fixe. Cette méthode est introduite par LeFloch et Tzavaras [112], Tzavaras [132] et développée par Joseph et LeFloch [88; 89; 90] et nécessite ici de modifier la définition des espaces de points fixe. Dans ces études antérieures les

coefficients d'interaction à estimer étaient de deux types : linéaires et quadratiques ; les estimations étaient alors la conséquence directe de la stricte hyperbolicité supposée du système qui permettait aux ondes caractéristiques d'avoir des supports essentiellement disjoints. Le contexte présent nécessite en revanche de prendre en compte les interactions nouvelles entre ces ondes habituelles du système (correspondant à sa partie strictement hyperbolique) et l'onde artificielle introduite par le couplage qui met justement en défaut la stricte hyperbolicité du système complet et représente le cadre résonnant.

On démontrera finalement que la solution obtenue à la limite $\epsilon \to 0$ est solution entropique, dans un cadre $L^1 \cap BV$, des équations couplées sur chaque demi-espace, l'interface n'étant pas à ce stade explicitement traitée. On a le théorème suivant :

Théorème 2.21 p.99 : *Sous des hypothèses de proximité des données, la suite u^ϵ converge simplement vers $u = \theta_\pm(w)$ à variation bornée, où w autosemblable est solution entropique sur chaque demi-espace de*

$$\begin{cases} \partial_t w + \partial_x f_-(w) = 0, & x < 0,\ t > 0, \\ \partial_t w + \partial_x f_+(w) = 0, & x > 0,\ t > 0. \end{cases}$$

Structure interne des interfaces dans les solutions Riemann-Dafermos. Dans le Chapitre 3, nous nous intéressons à l'étude des propriétés à l'interface de la solution construite au Chapitre 2. À savoir, peut-on obtenir à l'interface une caractérisation des traces de la solution ? À quel point la condition de couplage attendue est-elle réalisée ? Le processus d'approximation visqueuse a-t-il réduit la non-unicité des solutions ? Nous traitons pour le moment le problème lorsque $\theta_\pm = Id$.

Un principe de renormalisation de la solution est mis en œuvre (cf. Figure 4) afin de révéler la structure asymptotique de la solution à l'interface. Dans le cas scalaire nous obtenons une caractérisation « entropique » des solutions, à une couche limite près, dont nous obtenons l'équation à la manière de celle des profils d'ondes progressives. Le résultat principal est le suivant :

Théorème 3.1 p.117 : *Soient f_- et f_+ deux fonctions $C^1(\mathbb{R})$ telles que $f'_\pm$ sont lipschitziennes et soit $u^\epsilon \in L^\infty(\mathbb{R}) \cap BV(\mathbb{R})$ la solution de (19) avec des données de Riemann u_L et u_R et $U^\epsilon(y) = u^\epsilon(\epsilon y)$, alors les faits suivants sont réalisés.*

1. *Il existe $U \in C^2(\mathbb{R})$ limite de la suite $(U^\epsilon)_{\epsilon>0}$ lorsque ϵ tend vers 0. Elle satisfait l'équation de profile visqueux*

$$\left(\frac{1-V}{2} f'_-(U) + \frac{1+V}{2} f'_+(U)\right) U_y = U_{yy}, \tag{20}$$

et admet des limites à l'infini $U_{-\infty}$ et $U_{+\infty}$ qui satisfont à

$$\min(u_L, u_R) \le U_{\pm\infty} \le \max(u_L, u_R). \tag{21}$$

2. *De plus, q_- et q_+ étant des flux d'entropies liés respectivement à f_- et f_+ pour une entropie*

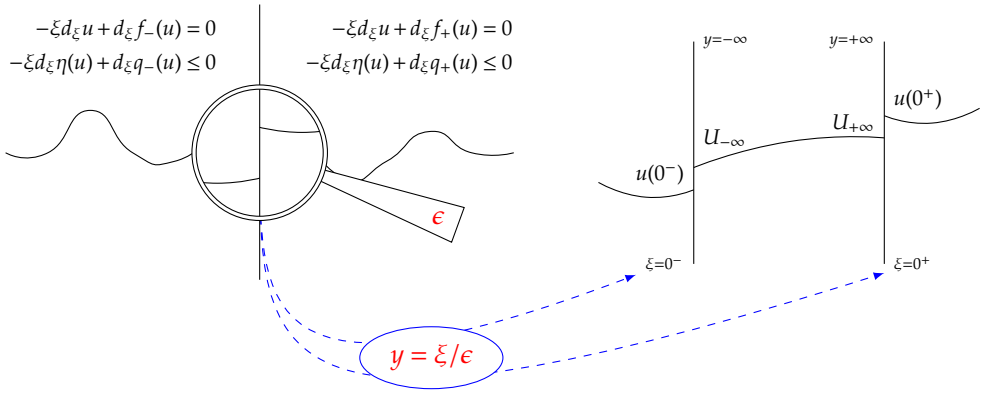


FIGURE 4: *Principe de renormalisation de la solution révélant la structure à l'interface.*

strictement convexe, et $u \in L^\infty(\mathbb{R}) \cap BV(\mathbb{R})$ *étant la limite de* u^ϵ*, on a*

$$\begin{aligned} f_-(u(0^-)) = f_-(U_{-\infty}), \quad q_-(u(0^-)) \ge q_-(U_{-\infty}), \\ f_+(U_{+\infty}) = f_+(u(0^+)), \quad q_+(U_{+\infty}) \ge q_+(u(0^+)). \end{aligned} \tag{22}$$

3. *Si de plus* $f_-'(U_{-\infty}) < 0$ *ou* $f_+'(U_{+\infty}) > 0$*, alors* U *est constante et en particulier* $U_{-\infty} = U_{+\infty}$*.*

Nous obtenons des contraintes supplémentaires qui discriminent certaines solutions dans le plan des données de Riemann (u_ℓ, u_r). La Figure 5 représente la carte des solutions admissibles dans les différents secteurs et leur structure. L'unicité est établie dans de nombreux cas et il y a sinon quatre solutions au plus. Rappelons que nous partions d'une infinité non-dénombrable de solutions ! En particulier, seule la solution constituée de deux ondes de détente séparées par l'état intermédiaire $c/2$ est par exemple admise dans le secteur le plus en haut à gauche. Dans le triangle central où quatre solutions différentes sont possibles, il y a un choc pour le modèle de gauche, un choc pour le modèle de droite, une discontinuité localisée à l'interface et enfin une solution constituée de deux chocs séparés par un état uniquement déterminé.

Modèle de couplage à interface épaissie et schéma équilibre. Le Chapitre 4 présente un schéma équilibre pour traiter le couplage de deux lois de conservation scalaires monodimensionnelles à travers une interface épaissie. Par épaissie, il faut entendre ici un mécanisme de régularisation de la procédure de couplage à une interface mince. Le bénéfice est comme nous le verrons une approximation numérique insensible au raffinement du maillage, dans un contexte naturel de solutions où l'unicité est de mise. En particulier, la condition de couplage sera comprise comme un équilibre à conserver, i.e. la donnée des solutions stationnaires du problème. Le schéma proposé est un schéma de type volumes finis avec une délocalisation de la variable couleur v représentant l'interface. Nous obtenons un résultat de convergence du schéma proposé [34] qui nécessite la mise en œuvre dans un cadre L^∞ (et non BV) de la méthode proposée par Coquel, Diehl, Merkle et Rohde [53]. Cette méthode fait appel à la notion de solutions mesures entropiques (au sens de DiPerna [64]) dont nous démontrons qu'elles vérifient suffisamment d'inégalités d'entropie pour

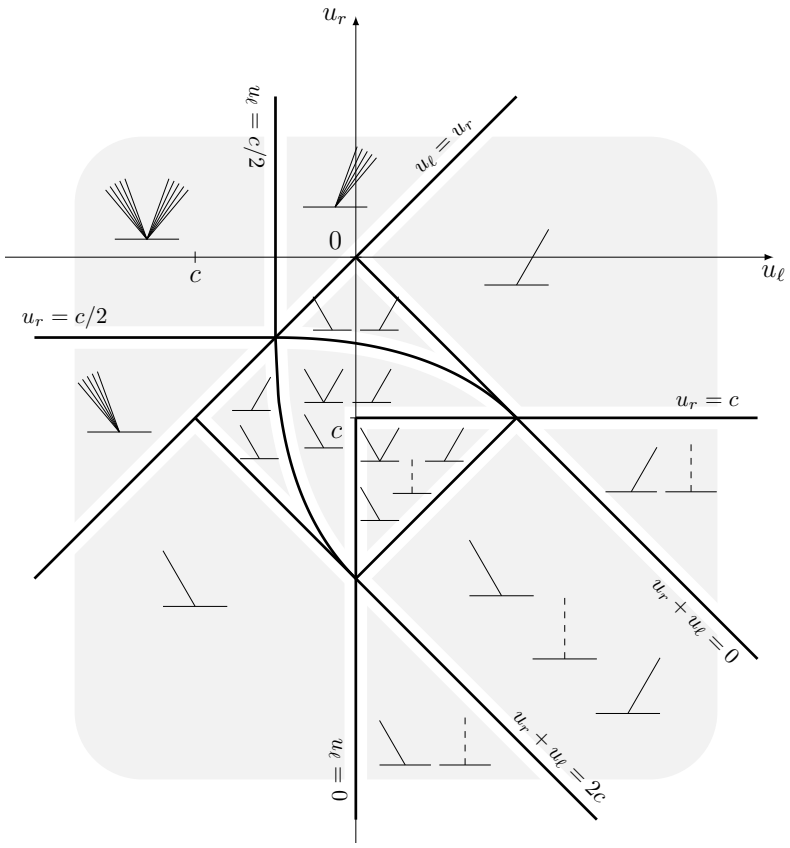


FIGURE 5: *Solutions de Riemann-Dafermos pour le couplage par état. Cas de deux flux quadratiques convexes de la forme* $f_-(u) = u^2/2$, $f_+(u) = (u-c)^2/2$, $c < 0$.

retrouver le résultat dans le cadre initial L^∞. Nous utilisons alors ce schéma pour révéler la sensibilité des solutions à la forme de l'interface dans des cas résonnants (cf Figure 6).

Une classe de schémas équilibres multi-dimensionnels. Le Chapitre 5 étend la construction et les résultats de convergence du chapitre précédent au cas du couplage de plusieurs lois de conservation scalaires multi-dimensionnelles. Les résultats de convergence et la propriété équilibre sont conservées pour une classe de schémas construits à partir d'une combinaison convexe de schémas quasi-1D. Les exemples présenteront des schémas basés sur une combinaison naturelle héritée de considérations géométriques. La Figure 7 présente un test effectué avec les trois domaines représentés sur la Figure 3 ($t = 3.7$ à gauche et $t = 6.0$ à droite).

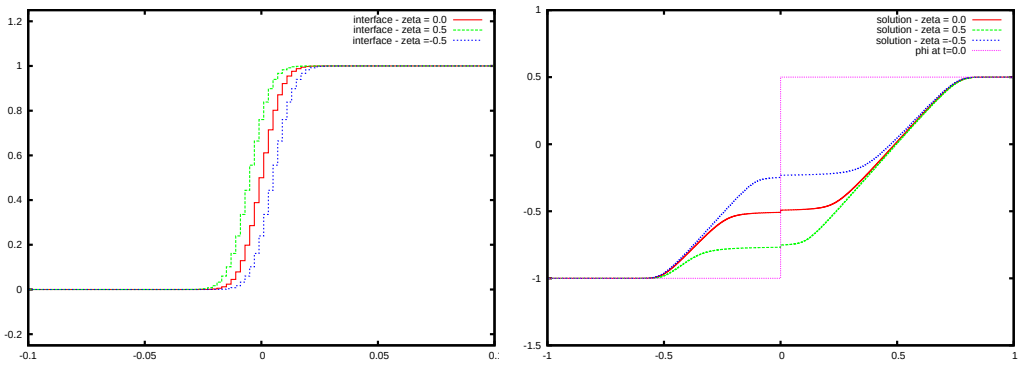


FIGURE 6: *Cas résonnant d'une double détente. Trois interfaces épaisses v différentes (à gauche) et les différentes solutions correspondantes (à droite) pour une même donnée initiale.*

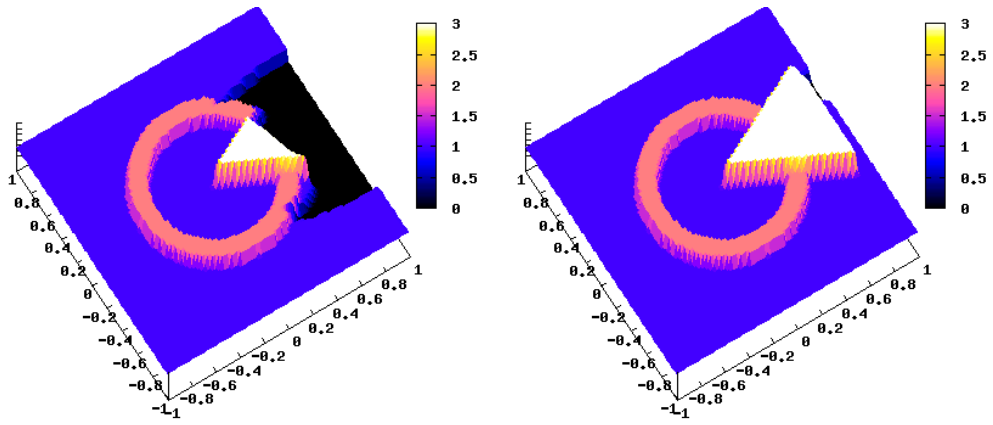


FIGURE 7: *Exemple bi-dimensionnel de couplage pour trois domaines. Onde de choc traversant un motif à deux inhomogénéités (solution à l'instant $t = 3.7$ à gauche et à $t = 6.0$ à droite).*

Seconde Partie :
Solutions non classiques de lois de conservation

Historiquement, les solutions non classiques de lois de conservation sont apparues dans les travaux de physique des matériaux, dans le but de modéliser la dynamique de propagation des ondes de *transition de phase* (cf. Slemrod [122; 123], Truskinovsky [130; 131], LeFloch [101], Abeyaratne et Knowles [1; 2], puis Hayes et LeFloch [79; 80; 81], LeFloch et Thanh [108; 109; 110; 111]). Plus récemment, de telles solutions ont encore été observées dans la modélisation de transport de populations par Benzoni-Gavage et Colombo [27], Colombo et Rosini [52] ou d'écoulements diphasiques en milieu poreux hétérogène par Cancès [39; 40; 41]. Pour une lecture complète sur le sujet, le lecteur pourra se référer à la monographie de LeFloch [102].

Les solutions non classiques d'une loi de conservation

$$\partial_t u + \partial_x f(u) = 0, \quad u \equiv u(t,x) \in \mathbb{R},\ x \in \mathbb{R},\ t > 0. \tag{23}$$

sont des solutions qui sortent du cadre habituel compressif, et violent notamment les conditions d'admissibilité au sens de Liu [114]. Les discontinuités dites non classiques ne sont alors pas astreintes à satisfaire aux conditions de compressibilité de Lax et sont seulement *sous-compressives*. Elles ne satisfont *a fortiori* donc plus à toutes les inégalités d'entropie et concernent essentiellement des flux non-convexes (ou des champs caractéristiques non vraiment non-linéaires dans le cas de systèmes). De telles solutions apparaissent par exemple avec la prise en compte de phénomènes régularisants d'ordre supérieur à deux. De manière générale, elles découlent de l'étude des solutions limites d'une régularisation prenant la forme

$$\partial_t u^\epsilon + \partial_x f(u^\epsilon) = \mathcal{R}^\epsilon(u^\epsilon, \partial_x u^\epsilon, \partial_{xx} u^\epsilon, \ldots), \tag{24}$$

dans laquelle l'opérateur $\mathcal{R}^\epsilon$ prétend restituer la physique pertinente « oubliée » dans la limite du premier ordre (23). Dans la série de papiers de Bedjaoui et LeFloch [22; 23; 24; 25], les auteurs explorent par exemple les solutions ondes progressives *limite diffusion-dispersion*, solutions de la régularisation de (23) suivante :

$$\partial_t u + \partial_x f(u) = \beta \partial_x (b(u)\partial_x u) + \gamma \partial_x (c_1(u)\partial_x (c_2(u)\partial_x u)). \tag{25}$$

Ils mettent alors en évidence la pertinence d'une *relation cinétique* caractérisant les solutions discontinues, lorsque les paramètres β et γ tendent simultanément vers zéro. Cette relation coïncide avec la condition d'admissibilité d'Oleinik pour des chocs d'amplitude suffisamment petite et est fonction de la limite du ratio γ/β^2 uniquemment. Elle coïncide également avec les solutions classiques lorsque ce ratio est asymptotiquement infini, autrement dit lorsque la diffusion domine. Plus récemment par exemple, un résultat remarquable est obtenu par les mêmes auteurs [26] dans le cas d'une diffusion plus singulière prenant la forme $\partial_x(|\partial_x|^p \partial_x u)$ où p est un réel positif donné : des chocs non classiques d'amplitude

arbitrairement petite peuvent être obtenus pour les cas les plus singuliers $p > 1/3$, là encore, les relations cinétiques sont calculables explicitement pour certaines valeurs du ratio γ/β^2.

Les solutions non classiques satisfont à une inégalité d'entropie héritée de la forme de l'opérateur de régularisation $\mathcal{R}^\epsilon$,

$$\partial_t U(u) + \partial_x F(u) \leq 0, \tag{26}$$

où (U,F) est un couple d'entropie - flux d'entropie particulier, satisfaisant à la relation de compatibilité $\nabla F = \nabla U \, \nabla f$ avec U strictement convexe. Cette inégalité d'entropie ne suffit pas, dans le cas d'un flux non-convexe, à assurer l'unicité, comme on l'a dit plus haut. Une *relation cinétique* vient compléter l'inégalité d'entropie pour permettre de restituer cette unicité. Elle peut prendre plusieurs formes : on peut quantifier la dissipation d'entropie d'une discontinuité de vitesse donnée (cf. Hayes et LeFloch [79; 80; 81]) ou plus simplement imposer à un choc non-compressif de vérifier une relation algébrique donnée

$$u_+ = \varphi^\flat(u_-), \tag{27}$$

entre les traces à gauche et à droite de la discontinuité. Cette caractérisation est, *a posteriori* et sous des hypothèses naturelles de forme sur la fonction $\varphi^\flat$, suffisamment discriminante pour restituer l'unicité en général. La série de papiers de Amadori *et al.* [3], Baiti *et al.* [17; 18; 19] traite du caractère bien posé des différents solveurs non classiques ainsi envisagés.

D'un point de vue numérique, les solutions non classiques de lois de conservation sont mal calculées par les méthodes habituelles. Un schéma de Godunov employant pour la définition de ses flux le solveur non classique exact ne parvient par exemple qu'à calculer les solutions classiques. Ces méthodes ne perçoivent pas suffisamment les phénomènes de petites échelles qui interviennent dans la régularisation $\mathcal{R}^\epsilon$ et qui pourraient être restitués par exemple à travers la relation cinétique. Plus précisément, c'est la diffusion numérique introduite dans la phase de projection du schéma qui dissimule ces phénomènes de petites échelles. Des approches ont été élaborées pour résoudre le problème, citons les schémas d'ordre élevé (cf. Chalons et LeFloch [46; 47], Hayes et LeFloch [80], LeFloch *et al.* [103], LeFloch et Mohammadian [104], Zhong *et al.* [133]) dans lesquels on recherche une consistance d'ordre supérieur du schéma numérique avec la régularisation (24) via l'étude de son équation équivalente. Une autre classe de solveurs, particulièrement efficace et qui sera employée pour fournir des solutions de références, est fournie par les schémas de type Glimm utilisant un solveur de Riemann non classique exact faisant intervenir directement la relation cinétique sous la forme (27) (cf. Chalons [42; 43; 44], Chalons et LeFloch [48]). Néanmoins, les schémas de type Glimm présentent l'inconvénient de ne pas être conservatifs, tandis que les schémas d'ordre élevé le sont mais génèrent des solutions généralement oscillantes.

Synthèse des résultats obtenus

Dans cette seconde Partie du manuscrit, nous proposons un schéma numérique de type volumes finis adapté au calcul des solutions non classiques définies par une relation cinétique. À la différence des méthodes antérieures, ce schéma est à la fois conservatif et s'accorde parfaitement à la formulation (27) de la relation cinétique sans passer par une montée en ordre. Il évite donc à la fois les défauts des schémas d'ordre élevés et ceux des schémas de type Glimm. Mieux que cela, il combine leurs qualités.

Dans le Chapitre 6, nous nous intéressons au cas d'un flux scalaire concave-convexe (ou convexe-concave) sans changement de monotonie (on a typiquement en tête les flux de la forme $f(u) = u^3 + u$). Cette étude a fait l'objet de la publication [29] avec Chalons, Lagoutière et LeFloch. Le schéma prend la forme d'un schéma de flux conservatif

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x}(f_{j+1/2}^n - f_{j-1/2}^n), \quad j \in \mathbb{Z}, \tag{28}$$

dans lequel les flux $f_{j+1/2}^n$ sont déduits d'une étape « virtuelle » de reconstruction des mailles sujettes à un changement de phase. C'est au cours de cette reconstruction que la relation cinétique prescrite intervient de manière à introduire les discontinuités non classiques « perdues » dans la projection de la solution numérique sur le maillage. Des résultats numériques viendront illustrer la convergence numérique du schéma et la bonne restitution de la fonction cinétique. De plus, à la manière des schémas de Glimm, le schéma calcule les chocs non classiques sans diffusion. La Figure 8 présente la comparaison entre la solution non classique obtenue par LeFloch et Rohde [105] à l'aide d'un schéma d'ordre élevé et la solution calculée par le schéma du Chapitre 6.

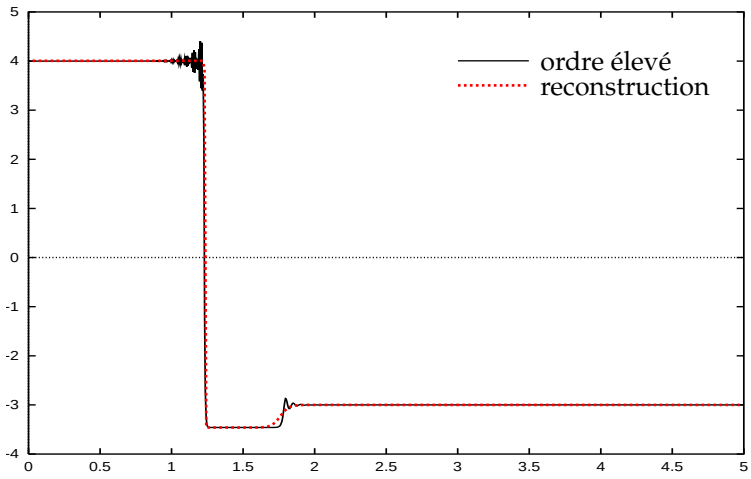


Figure 8: *Comparaison des solutions non classiques obtenues par un schéma d'ordre élevé (noir) et par le schéma de reconstruction (rouge).*

Dans le Chapitre 7, nous adaptons ce schéma au cas d'un flux scalaire concave-convexe

en levant l'hypothèse de monotonie du flux (on permet donc typiquement les flux de la forme $f(u) = u^3 - u$). Ce cas diffère du précédent dans la mesure où les discontinuités non classiques n'ont plus alors une vitesse signée et la définition des flux doit en tenir compte. Néanmoins, le principe de reconstruction précédemment introduit peut, dans ses grandes lignes, être conservé et une adaptation intuitive des flux, adaptée au signe de la vitesse de la discontinuité non classique, permet encore une fois d'obtenir un schéma satisfaisant.

Première partie

Couplage de systèmes de lois de conservation hyperboliques

1 Existence result for the coupling problem of two scalar conservation laws with Riemann initial data

Ce chapitre est consacré à l'étude de deux lois de conservations scalaires couplées à travers une interface fixe située par exemple en $x = 0$. Chacune de ces lois de conservation a sa propre fonction de flux (régulière) et est formulée sur un demi-espace, disons $x < 0$ et $x > 0$. À l'interface $x = 0$, on impose une condition de couplage dont l'objectif est d'assurer dans un sens faible la continuité d'une variable choisie qui peut différer de l'inconnue conservative (ou des flux). Nous obtenons l'existence d'une solution au problème de Riemann couplé en utilisant une approche constructive. Cette dernière permet en particulier d'éclairer des caractéristiques remarquables du problème, comme la non-unicité des solutions continues ou discontinues (à l'interface $x = 0$). Le comportement d'un schéma numérique est également examiné.

Ce travail a été réalisé en collaboration avec Christophe Chalons et Pierre-Arnaud Raviart et a été soumis pour publication [30].

Existence result for the coupling problem of two scalar conservation laws with Riemann initial data

Benjamin Boutin, Christophe Chalons & Pierre-Arnaud Raviart

Abstract

This paper is devoted to the coupling problem of two scalar conservation laws through a fixed interface located for instance at $x = 0$. Each scalar conservation law is associated with its own (smooth) flux function and is posed on a half-space, namely $x < 0$ or $x > 0$. At interface $x = 0$ we impose a coupling condition whose objective is to enforce in a weak sense the continuity of a prescribed variable, which may differ from the conservative unknown (and the flux functions as well). We prove existence of a solution to the coupled Riemann problem using a constructive approach. The latter allows in particular to highlight interesting features like non uniqueness of both continuous and discontinuous (at interface $x = 0$) solutions. The behavior of some numerical scheme is also investigated.

Introduction

The coupling of partial differential equations is of increasing interest in the applied mathematics community, and of course of increasing importance for industrial applications. Such a coupling arises for instance in the simulation of nuclear reactors when different two-phase flow codes are used[1]. In these codes, multiple modelling scales are applied to describe the flow. For instance, different thermal-hydraulic models can be used for each reactor component to take into account its specific behavior, or small scale models can be used, locally, to obtain a better resolution. When these models are put side to side, we face the problem of coupling. In addition to the definition of each model, such a problematic requires to be supplemented with an interfacial model in order to precise the nature of the information that must be exchanged at the coupling interface. This interfacial model may be formulated for instance when imposing the continuity of a given set of variables. It generally strongly affects the whole solution and must therefore be defined in order to achieve a physically coherent description of the whole operating device under consideration.
Let us mention that similar situations appear in the modelling of networks and traffic flows which have received a certain interest in the last few years. We refer for instance the reader to [20, 21, 51], [68, 69, 83], and the references therein.

In this paper, we are interested in the one dimensional coupling problem of two scalar conservation laws through a fixed interface, say $x = 0$, and more precisely in the resolution of the coupled Riemann problem. Each scalar conservation law is associated with its own (smooth) flux function f_α, $\alpha = L, R$ and is posed on a half-space, namely $x < 0$ ($\alpha = L$) or $x > 0$

[1] The authors of the present paper are involved in a joint research program on multiphase flows between CEA (French center fo nuclear research) and University Pierre et Marie Curie-Paris6 (Laboratoire Jacques-Louis Lions) in the frame of the Neptune project [78]. See for instance [5, 6, 8, 10, 11, 49] and the references therein

($\alpha = R$). At the coupling interface, we assume without further details that it is physically relevant to impose the continuity of a given function $v_\alpha = v_\alpha(u)$ of the solution u, meaning that u is expected to satisfy

$$v_\alpha(u(0_-,t)) = v_\alpha(u(0_+,t)),\ t > 0. \tag{1.1}$$

Note that this continuity constraint will be understood in a weak sense (inspired by [65]) and that the v-variable generally depends on α. This approach is fairly general and referred to as the *state coupling method*. See for instance [73, 76], but also [11, 49] and the references therein. It does significantly differ from the *flux coupling method* where the continuity of the flux is imposed at the interface ($v_\alpha(u) = f_\alpha(u)$). See for instance [16, 86, 93] and [85].
This paper gives the first result of global existence of a solution to the coupled Riemann problem in this context of state coupling, using a constructive approach. Except the smoothness hypothesis, no specific assumption is made on the flux functions f_α, $\alpha = L, R$. It is worth noticing right now that the solution of the coupled Riemann problem can be either continuous or discontinuous in the v-variable at the coupling interface. In the first situation, the coupling condition (1.1) is satisfied in the classical sense, while in the other one, it is satisfied in a weak sense only (to be precised hereafter). In addition, the solution to the Riemann problem is shown to be not necessarily unique, since in particular, a 1-parameter family of continuous solutions at the coupling interface may exist for the same Riemann initial data.

The outline of the paper is as follows. In Section 1, we introduce the general framework of the state coupling method. Section 2 is devoted to the main result of this paper, namely the existence of a solution to the coupled Riemann problem. First of all, a geometrical description of the sets of admissible traces at the interface is given. Then, a characterisation is given for the solutions satisfying the coupling condition either in the strong sense (the so-called v-continuous solutions) or in the weak sense (the so-called v-discontinuous solutions). At last, we deduce the existence of at least one self-similar solution to any coupled Riemann problem. Several situations of non-uniqueness are exhibited, and a first case of coupling of scalar conservation laws "with phase change" is treated theoretically. Section 3 is devoted to numerical simulations, using both a relaxation and a Godunov scheme as a building block for the derivation of the numerical strategy.

1.1 The state coupling method

Let $f_\alpha : \mathbb{R} \to \mathbb{R}$, $\alpha = L, R$, be two C^1 functions; given a function $u_0 : \mathbb{R} \to \mathbb{R}$, we want to find a function $u : (x,t) \to u(x,t) \in \mathbb{R}$ solution of

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f_L(u) = 0,\ x < 0,\ t > 0 \tag{1.2}$$

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f_R(u) = 0,\ x > 0,\ t > 0 \tag{1.3}$$

and satisfying the initial condition

$$u(x,0) = u_0(x),\ x \in \mathbb{R} \tag{1.4}$$

together with coupling constraints at $x = 0$ that we now define. Let $\theta_\alpha : \mathbb{R} \to \mathbb{R}$, $\alpha = L, R$, be two *strictly monotone* C^1 functions; we require the function u to satisfy "as far as possible" the continuity constraint

$$\theta_L^{-1}(u(0_-,t)) = \theta_R^{-1}(u(0_+,t)),\ t > 0. \tag{1.5}$$

Setting

$$v(x,t) = \begin{cases} \theta_L^{-1}(u(x,t)),\ x < 0 \\ \\ \theta_R^{-1}(u(x,t)),\ x > 0, \end{cases} \tag{1.6}$$

this constraint must be understood in the weak sense, following Dubois et LeFloch [65], Godlewski et Raviart [76] and Godlewski *et al.* [73]

$$\begin{cases} v(0_-,t) \in \widetilde{O}_L(v(0_+,t)) \\ \\ v(0_+,t) \in \widetilde{O}_R(v(0_-,t)). \end{cases} \tag{1.7}$$

Let us recall the definition of the sets $\widetilde{O}_L(v_d)$ and $\widetilde{O}_R(v_g)$. Denoting by $w_\alpha(\frac{x}{t}; u_g, u_d)$ the self-similar solution of the Riemann problem

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f_\alpha(u) = 0,\ x \in \mathbb{R},\ t > 0$$

$$u(x,0) = \begin{cases} u_g,\ x < 0 \\ u_d,\ x > 0 \end{cases}$$

and setting

$$z_\alpha(\frac{x}{t}; v_g, v_d) = \theta_\alpha^{-1}(w_\alpha(\frac{x}{t}; \theta_\alpha(v_g), \theta_\alpha(v_d)))$$

we have

$$\begin{cases} \widetilde{O}_L(v_d) = \{z_L(0_-; v, v_d);\ v \in \mathbb{R}\} \\ \\ \widetilde{O}_R(v_g) = \left\{z_R(0_+; v_g, v);\ v \in \mathbb{R}\right\}. \end{cases}$$

1.2 Solving the coupled Riemann problem.

We consider the coupled Riemann problem which corresponds to the initial condition

$$u_0(x) = \begin{cases} u_g,\ x < 0 \\ u_d,\ x > 0. \end{cases} \tag{1.8}$$

We set

$$v_g = \theta_L^{-1}(u_g),\ v_d = \theta_R^{-1}(u_d).$$

When the flux functions f_α, $\alpha = L, R$ are *strictly convex*, we are able to exhibit all the solutions of this coupled Riemann problem (1.2)-(1.4),(1.7),(1.8). This is indeed the goal of this section.

1.2.1 Preliminaries.

Let us first recall and derive some preliminary results. Given a function $f \in C^1(\mathbb{R})$, we consider the scalar equation

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0. \tag{1.9}$$

We denote by $w(\frac{x}{t}; u_g, u_d)$ the solution of the Riemann problem for (1.9) corresponding to the initial condition (1.8). In the case of a general flux function f, $w(\frac{x}{t}; u_g, u_d)$ consists of a *composite wave* composed of shock and rarefaction subwaves. It is constructed in the following way.
(i) For $u_d > u_g$, we introduce the lower convex envelope function f_c of f in the interval $[u_g, u_d]$. This interval is divided into rarefaction subintervals where the function f is strictly convex (so that $f_c = f$) separated by shock subintervals where the function f_c is affine (and the graph of f is located above the corresponding chord). Then $w(\frac{x}{t}; u_g, u_d)$ is made of a sequence of rarefaction waves in the rarefaction subintervals and shock waves in the shock subintervals. These waves are bordered on the left by the constant state u_g and on the right by the constant state u_d. These constant states are the only constant states which appear in the solution of the Riemann problem.
(ii) For $u_d < u_g$, we introduce the upper concave envelope function f^c of f in the interval $[u_d, u_g]$. Again this interval is divided into rarefaction subintervals where the function f is strictly concave (so that $f = f^c$) separated by shock intervals where f^c is affine (and the graph of f is located under the corresponding chord. Then, the solution of the Riemann problem then has the same structure as in the case (i).

We shall say that such a composite wave has a nonnegative (resp. nonpositive) speed if all of its subwaves (shocks or rarefactions) have nonnegative (resp. nonpositive) speeds. In order to characterize the composite waves $w(.; u_g, u_d)$ whose speeds are nonnegative or nonpositive, we determine the minimal and maximal speeds σ_{min} and σ_{max} of such a composite wave. Denote by $I(u_g, u_d)$ the closed interval whose end points are u_g and u_d. Then, we can state

Lemma 1.1. *We have*

$$\sigma_{min} = \min_{u \in I(u_g, u_d)} \frac{f(u) - f(u_g)}{u - u_g} \tag{1.10}$$

and

$$\sigma_{max} = \max_{u \in I(u_g, u_d)} \frac{f(u) - f(u_d)}{u - u_d} \tag{1.11}$$

with the following convention

$$\frac{f(u) - f(u_a)}{u - u_a} = f'(u_a) \text{ for } u = u_a,\ a = g, d. \tag{1.12}$$

Proof. Let us check (1.10) for instance. We begin by observing that σ_{min} is the speed of the left boundary of the fan of the composite wave $w(.;u_g,u_d)$. Assume first $u_d > u_g$. If the left subwave of $w(.;u_g,u_d)$ is a shock that connects u_g and a state u_1, its speed is

$$\sigma_{min} = \frac{f(u_1)-f(u_g)}{u_1-u_g}.$$

On the other hand, it is clear geometrically (cf. Fig. 1.1a) that we have for all $u \in (u_g,u_d]$

$$\frac{f(u)-f(u_g)}{u-u_g} > \frac{f(u_1)-f(u_g)}{u_1-u_g}.$$

Hence, using the convention (1.12), we obtain

$$\sigma_{min} = \min_{u\in[u_g,u_d]} \frac{f(u)-f(u_g)}{u-u_g}. \tag{1.13}$$

If this left subwave is a rarefaction, σ_{min} is the speed of the left side of the rarefaction fan which is given by

$$\sigma_{min} = f'(u_g)$$

Again, it is obvious geometrically (cf. Fig. 1.1b) that we have for all $u \in (u_g,u_d]$

$$f'(u_g) < \frac{f(u)-f(u_g)}{u-u_g}.$$

Therefore (1.13) still holds.

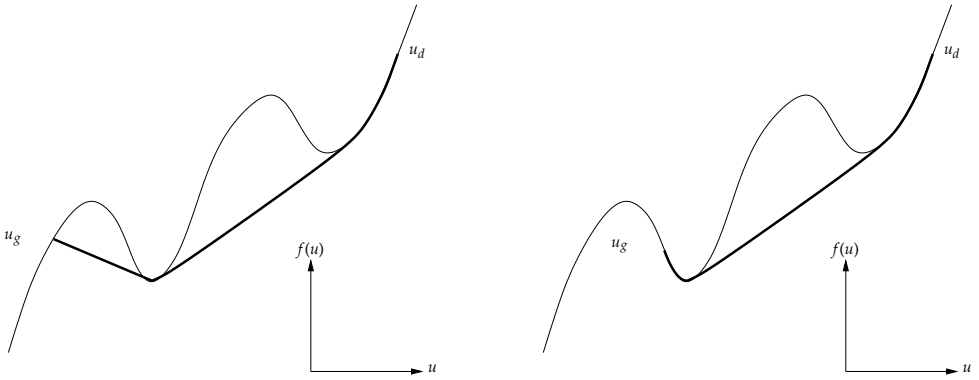


Figure 1.1: $u_g < u_d$ and the left subwave is a shock (a) or a rarefaction (b)

Consider next the case where $u_g > u_d$. Using Fig. 1.2 and a fairly similar analysis, one can check that

$$\sigma_{min} = \min_{u\in[u_d,u_g]} \frac{f(u)-f(u_g)}{u-u_g}.$$

This proves (1.10). The property (1.11) is established exactly in the same way. □

As a consequence of Lemma 1, we obtain that a (composite) wave $w(.;u_g,u_d)$ has a

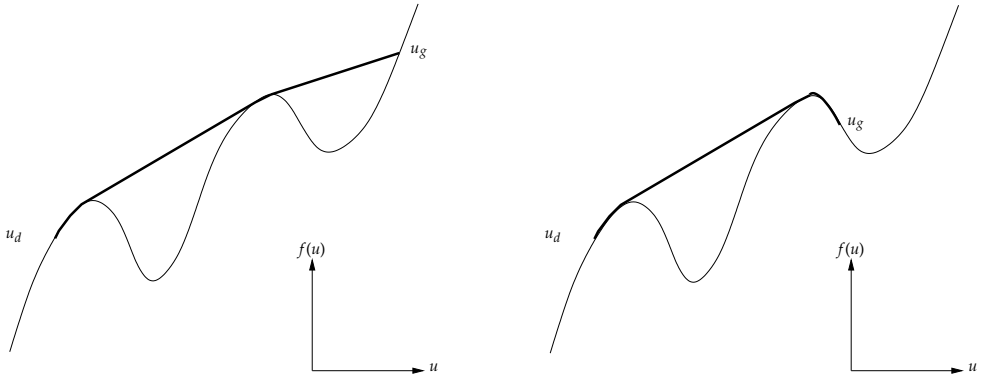


Figure 1.2: $u_g > u_d$ and the left subwave is a shock (a) or a rarefaction (b)

nonnegative speed if and only if

$$\min_{u\in I(u_g,u_d)} \frac{f(u)-f(u_g)}{u-u_g} \geq 0$$

and a nonpositive speed if and only if

$$\max_{u\in I(u_g,u_d)} \frac{f(u)-f(u_d)}{u-u_d} \leq 0.$$

In other words, we get

$$\begin{cases} w(0_-;u_g,u_d)=u_g \Longleftrightarrow \min\limits_{u\in I(u_g,u_d)} \dfrac{f(u)-f(u_g)}{u-u_g} \geq 0, \\ \\ w(0_+;u_g,u_d)=u_d \Longleftrightarrow \max\limits_{u\in I(u_g,u_d)} \dfrac{f(u)-f(u_d)}{u-u_d} \leq 0. \end{cases} \tag{1.14}$$

Now, given a state u_0, we look for the set $\mathcal{E}^+(u_0)$ (resp. $\mathcal{E}^-(u_0)$) of all states $u \neq u_0$ which can be connected to u_0 *on the left* (resp. *on the right*) by a nontrivial (composite) wave $w(\cdot;u,u_0)$ (resp. $w(\cdot;u_0,u)$) whose speed is nonnegative (resp. nonpositive). The above results yield

Lemma 1.2. *We have*

$$\mathcal{E}^+(u_0) = \left\{ u\neq u_0;\ \min_{v\in I(u_0,u)} \frac{f(v)-f(u)}{v-u} \geq 0 \right\} \tag{1.15}$$

and

$$\mathcal{E}^-(u_0) = \left\{ u\neq u_0;\ \max_{v\in I(u_0,u)} \frac{f(v)-f(u)}{v-u} \leq 0 \right\}. \tag{1.16}$$

It remains to give a geometric characterization of the conditions (1.15) and (1.16). This is easily done when the flux function f is either *monotone* or *strictly convex*.

Example 2.1. the case of a monotone flux function. If the function f is strictly increasing

so that

$$\frac{f(v)-f(u)}{v-u} > 0 \text{ for all } v \neq u,$$

we obtain

$$\mathcal{E}^+(u_0) = \mathbb{R} \setminus \{u_0\}, \quad \mathcal{E}^-(u_0) = \emptyset$$

while if the function f is strictly decreasing, we find

$$\mathcal{E}^+(u_0) = \emptyset, \quad \mathcal{E}^-(u_0) = \mathbb{R} \setminus \{u_0\}. \square$$

Example 2.2. the case of a strictly convex flux function. When the function f is strictly convex, we denote by $\bar{u}$ the *sonic state* of f characterized by $f'(\bar{u}) = 0$ with the convention that $\bar{u} = -\infty$ (resp. $\bar{u} = +\infty$) if the function f is strictly increasing (resp. strictly decreasing). With the state u_0, we associate the state $\tilde{u}_0$ defined by

$$\begin{cases} f(\widetilde{u_0}) = f(u_0),\ \tilde{u}_0 \neq u_0 \quad \text{if } \bar{u} \text{ exists,} \\ \\ \tilde{u}_0 = \bar{u} = -\infty \quad \text{if } f \text{ is strictly increasing,} \\ \\ \tilde{u}_0 = \bar{u} = +\infty \quad \text{if } f \text{ is strictly decreasing.} \end{cases} \tag{1.17}$$

Lemma 1.3. *Assume that the function f is strictly convex. Then*

$$\mathcal{E}^+(u_0) = \{u \neq u_0;\ u \geq \max(\bar{u}, \tilde{u}_0)\} \tag{1.18}$$

and

$$\mathcal{E}^-(u_0) = \{u \neq u_0;\ u \leq \min(\bar{u}, \tilde{u}_0)\}. \tag{1.19}$$

Proof. Let us check for instance the property (1.18). Given $u \in \mathbb{R}$, we define the function g by

$$g(v) = \frac{f(v)-f(u)}{v-u}, v \neq u, \quad g(u) = f'(u).$$

Since

$$g'(v) = \frac{f(u)-f(v)+f'(v)(v-u)}{(v-u)^2}$$

and by the strict convexity of f

$$f(u)-f(v)+f'(v)(v-u) > 0, \quad v \neq u,$$

this function g is strictly increasing. Hence we obtain

$$\min_{v \in I(u_0,u)} g(v) = \begin{cases} g(u_0) = \dfrac{f(u_0)-f(u)}{u_0-u} & \text{if } u_0 < u \\ \\ g(u) = f'(u) & \text{if } u_0 > u \end{cases}$$

so that

$$\min_{v\in I(u_0,u)} g(v) \geq 0 \Longleftrightarrow \begin{cases} f(u) \geq f(u_0) & \text{if } u_0 < u \\ \\ f'(u) \geq 0 & \text{if } u_0 > u. \end{cases}$$

Now the condition $f(u) \geq f(u_0)$ for $u_0 < u$ holds trivially if $u_0 \geq \bar{u}$ but means $u \geq \tilde{u}_0$ if $u_0 \leq \bar{u}$. On the other hand, the condition $f'(u) \geq 0$ means $u \geq \bar{u}$. The property (1.18) is then proved. □

In the above examples, $\mathcal{E}^{\pm}(u_0)$ is an interval or the whole real line, the state u_0 being excluded. In the case of a general flux function, $\mathcal{E}^{\pm}(u_0)$ consists of an interval or a union of disjoint intervals (cf. Fig. 1.3).

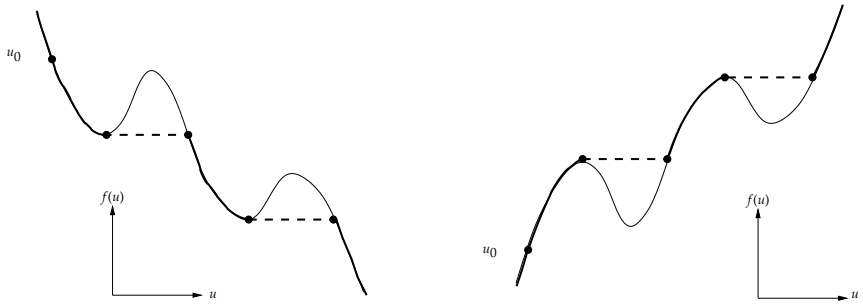


Figure 1.3: In bold-face, the sets $\mathcal{E}^-(u_0)$ (a) and $\mathcal{E}^+(u_0)$ (b), u_0 being excluded in both cases, and the other circle points being included

We next give another useful characterization of the sets $\mathcal{E}^{\pm}(u_0)$

Lemma 1.4. *We have*

$$\mathcal{E}^+(u_0) = \{u = w(0_-; y, u_0),\, y \in \mathbb{R};\, u \neq u_0\} \tag{1.20}$$

and

$$\mathcal{E}^-(u_0) = \{u = w(0_+; u_0, y),\, y \in \mathbb{R};\, u \neq u_0\}. \tag{1.21}$$

Proof. Let $u = w(0_-; y, u_0) \neq u_0$ for some $y \in \mathbb{R}$. Then, u is connected to u_0 by a wave $w(\cdot; u, u_0)$ whose speed is nonnegative, i.e., $u \in \mathcal{E}^+(u_0)$. Conversely, if $u \in \mathcal{E}^+(u_0)$, then $u = w(0_-; u, u_0) \neq u_0$ which proves (1.20). The property (1.21) is established in a similar way. □

In the sequel, we will make use of the following sets:

$$\begin{cases} \mathcal{F}^+(u_0) = \{u = w(0_+; y, u_0),\, y \in \mathbb{R};\, u \neq u_0\} \\ \\ \mathcal{F}^-(u_0) = \{u = w(0_-; u_0, y),\, y \in \mathbb{R};\, u \neq u_0\}. \end{cases} \tag{1.22}$$

that we now characterize.

Lemma 1.5. *We have*

$$\mathcal{F}^+(u_0) = \left\{u \neq u_0;\, \frac{f(v) - f(u)}{v - u} > 0\ \forall v \in I(u, u_0),\, v \neq u\right\}. \tag{1.23}$$

and

$$\mathcal{F}^-(u_0) = \left\{ u \neq u_0;\ \frac{f(v)-f(u)}{v-u} < 0\ \forall v \in I(u,u_0),\ v \neq u \right\}. \tag{1.24}$$

Proof. Let us check (1.23). We first prove

$$\mathcal{F}^+(u_0) = \{u \in \mathcal{E}^+(u_0);\ w(0_-;u,u_0) = w(0_+;u,u_0)\}.$$

Indeed, let $u \in \mathcal{F}^+(u_0)$; clearly u is connected to u_0 on the left by a wave whose speed is nonnegative (cf. Fig. 1.4) so that $u \in \mathcal{E}^+(u_0)$. In addition, we have

$$u = w(0_-;u,u_0) = w(0_+;u,u_0). \tag{1.25}$$

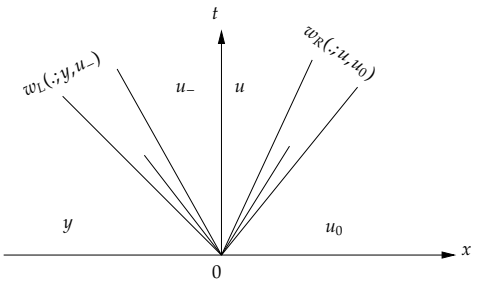


Figure 1.4: $u \in \mathcal{F}^+(u_0)$, i.e. $u = w(0_+;y,u_0)$, $y \in \mathbb{R}$.

Conversely, if $u \in \mathcal{E}^+(u_0)$ satisfies (1.25), u belongs obviously to $\mathcal{F}^+(u_0)$. We next show that $u \in \mathcal{E}^+(u_0)$ satisfies (1.25) if and only if

$$\frac{f(v)-f(u)}{v-u} > 0 \quad \forall v \in I(u,u_0),\ v \neq u.$$

Observe that the equality $w(0_-;u,u_0) = w(0_+;u,u_0)$ holds if and only if the left subwave of $w(\cdot;u,u_0)$ is not a stationary shock. Since $u \in \mathcal{E}^+(u_0)$, we already know from Lemma 2 that (1.15) holds and therefore

$$f'(u) \geq 0, \quad \frac{f(u)-f(u_0)}{u-u_0} \geq 0.$$

Then, if we assume $u > u_0$, it is clear geometrically (cf. Fig. 1.5) that we must have

$$f(v) < f(u) \quad \forall v \in [u_0,u).$$

Indeed, we have a stationary shock if and only if it exists a state $u_1 \in [u_0,u)$ such that $f(u) = f(u_1)$. Hence, there does not exist such a stationary shock if and only if

$$\frac{f(v)-f(u)}{v-u} > 0 \quad \forall v \in [u_0,u).$$

Similarly, for $u < u_0$, a stationary shock does not exist if and only if

$$f(v) < f(u) \quad \forall v \in (u,u_0]$$

or equivalently

$$\frac{f(v)-f(u)}{v-u} > 0 \quad \forall v \in (u, u_0].$$

This proves (1.23). The characterization (1.24) of $\mathcal{F}^-(u_0)$ is obtained analogously. □

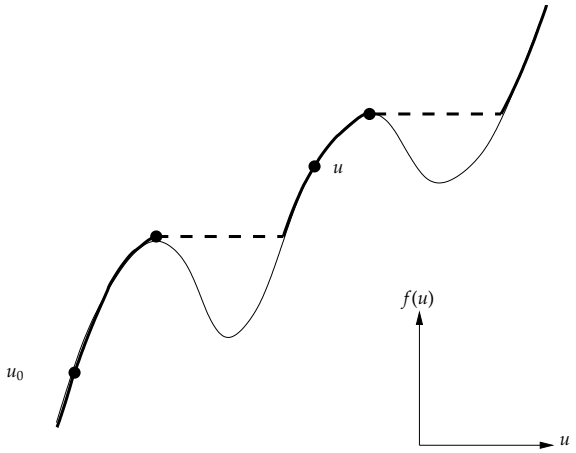


Figure 1.5: In bold-face, the set $\mathcal{F}^+(u_0)$, u_0 being excluded, and the other circle points being included (note the difference with respect to Fig. 1.3b). Here $u \in \mathcal{F}^+(u_0)$.

Example 2.1. (contd.) If the flux function f is strictly increasing, we have

$$\mathcal{F}^+(u_0) = \mathcal{E}^+(u_0) = \mathbb{R} \setminus \{u_0\}, \quad \mathcal{F}^-(u_0) = \mathcal{E}^-(u_0) = \emptyset$$

while for a strictly decreasing function f

$$\mathcal{F}^+(u_0) = \mathcal{E}^+(u_0) = \emptyset, \quad \mathcal{F}^-(u_0) = \mathcal{E}^-(u_0) = \mathbb{R} \setminus \{u_0\}. \, \square$$

Example 2.2. (contd.) Here we can state

Lemma 1.6. *Assume that the function f is strictly convex. Then*

$$\mathcal{F}^+(u_0) = \{u \neq u_0;\ u \geq \max(\bar{u}, \tilde{u}_0),\ u \neq \tilde{u}_0\} \tag{1.26}$$

and

$$\mathcal{F}^-(u_0) = \{u \neq u_0;\ u \leq \min(\bar{u}, \tilde{u}_0),\ u \neq \tilde{u}_0\}. \tag{1.27}$$

Proof. We check for instance the property (1.26). It follows from (1.20) that we have to restrict ourselves to the states $u \geq \max(\bar{u}, \tilde{u}_0)$. Assume first $u \geq \bar{u}$ so that $\bar{u} \geq \tilde{u}_0$. Then, we observe that, for $u \geq \bar{u}$, we have indeed

$$\frac{f(v)-f(u)}{v-u} > 0 \quad \forall v \in I(u_0, u),\ v \neq u$$

and thus $u \in \mathcal{F}^-(u_0)$. Consider next the case $\bar{u} \geq u_0$ for which $\tilde{u}_0 \geq \bar{u}$. For $u \geq \tilde{u}_0$, we obtain

$$\frac{f(v)-f(u)}{v-u} > 0 \quad \forall v \in [u_0, u)$$

if and only if $u > \tilde{u}_0$ (cf. Fig. 1.6) which proves (1.26). □

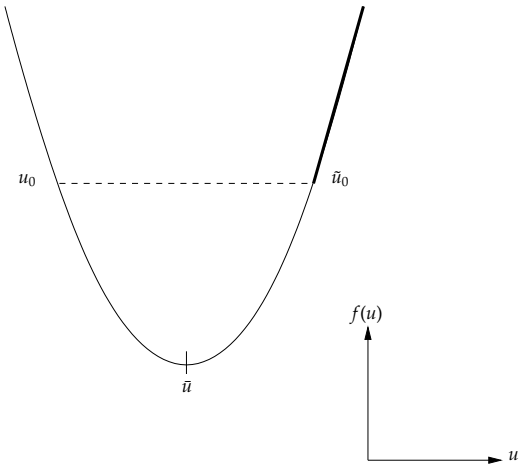


Figure 1.6: In bold-face, the set $\mathcal{F}^+(u_0)$, $\tilde{u}_0$ being excluded.

Now, let $\theta \in C^1(\mathbb{R})$ be a strictly monotone function; only for the sake of convenience, we will assume that θ satisfies $\theta' > 0$ and maps $\mathbb{R}$ onto itself. We set $\tilde{f}(v) = f(\theta(v))$ and we denote by $z(\frac{x}{t}; v_g, v_d)$ the solution of the Riemann problem expressed in the variable $v = \theta^{-1}(u)$, i.e.,

$$z(\frac{x}{t}; v_g, v_d) = \theta^{-1}(w(\frac{x}{t}; \theta(v_g), \theta(v_d)).$$

With a given state v_0 we associate the sets of states

$$\begin{cases} \widetilde{\mathcal{E}}^+(v_0) = \{v = z(0_-; y, v_0),\ y \in \mathbb{R}; v \neq v_0\} \\ \\ \widetilde{\mathcal{F}}^+(v_0) = \{v = z(0_+; y, v_0),\ y \in \mathbb{R}; v \neq v_0\} \end{cases} \tag{1.28}$$

and

$$\begin{cases} \widetilde{\mathcal{E}}^-(v_0) = \{v = z(0_+; v_0, y),\ y \in \mathbb{R}; v \neq v_0\} \\ \\ \widetilde{\mathcal{F}}^-(v_0) = \{v = z(0_-; v_0, y),\ y \in \mathbb{R}; v \neq v_0\}. \end{cases} \tag{1.29}$$

Using Lemmas 2 and 5, we have

$$\begin{cases} \widetilde{\mathcal{E}}^+(v_0) = \left\{v \neq v_0;\ \min_{w \in I(v_0, v)} \dfrac{\tilde{f}(w)-\tilde{f}(v)}{w-v} \geq 0\right\} \\ \\ \widetilde{\mathcal{F}}^+(v_0) = \left\{v \neq v_0;\ \dfrac{\tilde{f}(w)-\tilde{f}(v)}{w-v} > 0 \quad \forall w \in I(v_0, v), w \neq v\right\} \end{cases} \tag{1.30}$$

and

$$\begin{cases} \widetilde{\mathcal{E}}^-(v_0) = \left\{ v \neq v_0;\ \max\limits_{w \in I(v_0,v)} \dfrac{\tilde{f}(w) - \tilde{f}(v)}{w - v} \leq 0 \right\} \\ \\ \widetilde{\mathcal{F}}^-(v_0) = \left\{ v \neq v_0;\ \dfrac{\tilde{f}(w) - \tilde{f}(v)}{w - v} < 0 \quad \forall w \in I(v_0, v),\, w \neq v \right\}. \end{cases} \tag{1.31}$$

Example 2.1. (contd.) If the function f is strictly increasing, we have

$$\widetilde{\mathcal{E}}^+(v_0) = \widetilde{\mathcal{F}}^+(v_0) = \mathbb{R} \setminus \{v_0\}, \quad \widetilde{\mathcal{E}}^-(v_0) = \widetilde{\mathcal{F}}^-(v_0) = \emptyset$$

while if f is strictly decreasing

$$\widetilde{\mathcal{E}}^+(v_0) = \widetilde{\mathcal{F}}^+(v_0) = \emptyset, \quad \widetilde{\mathcal{E}}^-(v_0) = \widetilde{\mathcal{F}}^-(v_0) = \mathbb{R} \setminus \{v_0\}. \ \square$$

Example 2.2. (contd.) When the function f is strictly convex, we denote by $\bar{v} = \theta^{-1}(\bar{u})$ the sonic state of $\tilde{f}$. Given a state v_0, we set $u_0 = \theta(v_0)$ and $\tilde{v}_0 = \theta^{-1}(\tilde{u}_0)$. Then we obtain

$$\widetilde{\mathcal{E}}^+(v_0) = \{v \neq v_0;\, v \geq \max(\bar{v}, \tilde{v}_0)\}, \quad \widetilde{\mathcal{F}}^+(v_0) = \left\{ v \in \widetilde{\mathcal{E}}^+(v_0);\, v \neq \tilde{v}_0 \right\}$$

and

$$\widetilde{\mathcal{E}}^-(v_0) = \{v \neq v_0;\, v \leq \min(\bar{v}, \tilde{v}_0)\}, \quad \widetilde{\mathcal{F}}^-(v_0) = \left\{ v \in \widetilde{\mathcal{E}}^-(v_0);\, v \neq \tilde{v}_0 \right\}. \ \square$$

1.2.2 v-continuous solutions.

Let us now look for all possible *self-similar* solutions $u = u(\frac{x}{t})$ of the coupled Riemann problem. Again, for convenience, we assume that both functions θ_L and θ_R are strictly increasing and map $\mathbb{R}$ onto itself. We begin with those self-similar solutions which are *v-continuous at the interface* (in the strong sense), i.e., which satisfy

$$v(0_-) = v(0_+) = v(0) \tag{1.32}$$

where v is defined from u as in (1.6), or equivalently which satisfy the constraint

$$v(0) \in \left(\left\{ v_g \right\} \cup \widetilde{\mathcal{F}}_L^-(v_g) \right) \cap \left(\{v_d\} \cup \widetilde{\mathcal{F}}_R^+(v_d) \right). \tag{1.33}$$

Hence, besides the trivial solution corresponding to $v(0) = v_g = v_d$ [2], we obtain three types of v-continuous solutions.

(i) The first type of v-continuous solution. If

$$v(0) = v_d \in \widetilde{\mathcal{F}}_L^-(v_g),$$

the solution of the coupled Riemann problem coincides with the solution $z_L(\cdot; v_g, v_d)$ of the L-Riemann problem: it consists of a (composite) L-wave whose speed is nonpositive. Such

[2] which is excluded since, once for all, we have supposed $v_g \neq v_d$.

a solution is characterized by

$$\frac{\tilde{f_L}(v)-\tilde{f_L}(v_d)}{v-v_d}<0 \quad \forall v\in I(v_g,v_d),\ v\neq v_d.$$

(ii) The second type of v-continuous solution. If

$$v(0)=v_g\in\widetilde{\mathcal{F}}_R^{+}(v_d),$$

the solution of the coupled Riemann problem coincides with the solution $z_R(\cdot;v_g,v_d)$ of the R-Riemann problem: it consists of a (composite) R-wave whose speed is nonnegative. Such a solution is characterized by

$$\frac{\tilde{f_R}(v)-\tilde{f_R}(v_g)}{v-v_g}>0 \quad \forall v\in I(v_g,v_d),\ v\neq v_g.$$

(iii) The third type of v-continuous solution. The general case is indeed obtained by choosing

$$v(0)\in\widetilde{\mathcal{F}}_L^{-}(v_g)\cap\widetilde{\mathcal{F}}_R^{+}(v_d)$$

Obviously, this requires the condition

$$\widetilde{\mathcal{F}}_L^{-}(v_g)\cap\widetilde{\mathcal{F}}_R^{+}(v_d)\neq\emptyset.$$

Then a solution of the coupled Riemann problem coincides with $z_L(\cdot;v_g,v(0))$ in the domain $(x<0,t>0)$ and with $z_R(\cdot;v(0),v_d)$ in the domain $(x>,t>0)$. It consists of two (composite) waves: a L-wave whose speed is nonpositive and a R-wave whose speed is nonnegative. Such a solution is characterized by the conditions

$$\begin{cases} \dfrac{\tilde{f_L}(v)-\tilde{f_L}(v(0))}{v-v(0)}<0 \quad \forall v\in I(v_g,v(0)),\ v\neq v(0) \\[2ex] \dfrac{\tilde{f_R}(v)-\tilde{f_R}(v(0))}{v-v(0)}>0 \quad \forall v\in I(v(0),v_d),\ v\neq v(0). \end{cases}$$

We thus find a *one-parameter family* of solutions depending on the parameter $v(0)\in\widetilde{\mathcal{F}}_L^{-}(v_g)\cap\widetilde{\mathcal{F}}_R^{-}(v_d)$.

Let us notice that a solution of type (i) or type (ii) is that of a classical L or R-Riemann problem and may be viewed as a *quasi trivial* solution of this coupling problem.

We now apply these results to the case where both flux functions f_L and f_R are either strictly monotone or strictly convex.

Example 2.3. The case of strictly monotone flux functions.

(a) Suppose first that the functions f_L and f_R are strictly decreasing so that

$$\widetilde{\mathcal{F}}_L^-(v_g) = \mathbb{R} \setminus \left\{v_g\right\}, \quad \widetilde{\mathcal{F}}_R^+(v_d) = \emptyset.$$

Then, clearly the solution of type (i) alone is admissible.
(b) Similarly, if the functions f_L and f_R are strictly increasing so that

$$\widetilde{\mathcal{F}}_L^-(v_g) = \emptyset, \quad \widetilde{\mathcal{F}}_R^+(v_d) = \mathbb{R} \setminus \{v_d\},$$

the solution of type (ii) alone is admissible.
(c) Suppose next that f_L is strictly increasing and f_R is strictly decreasing. We have $\widetilde{\mathcal{F}}_L^-(v_g) = \widetilde{\mathcal{F}}_R^+(v_d) = \emptyset$. Then, none of the existence conditions of a v-continuous solution holds: there does not exist any v-continuous solution of the coupled Riemann problem (except the trivial solution corresponding to $v(0) = v_g = v_d$).
(d) If f_L is strictly decreasing and f_R is strictly increasing , we have

$$\widetilde{\mathcal{F}}_L^-(v_g) = \mathbb{R} \setminus \left\{v_g\right\}, \quad \widetilde{\mathcal{F}}_R^+(v_d) = \mathbb{R} \setminus \{v_d\}.$$

Hence any above condition of existence of a v-continuous solution holds: there exists a one-parameter family of solutions of type (iii) depending on the parameter $v(0) \in \mathbb{R}$. Clearly this family contains the solution of type (i) and that of type (ii). Hence the coupled Riemann problem has an infinite number of v-continuous solutions and it is enough to specify $v(0)$ for determining the unique corresponding solution.

Example 2.4. The case of strictly convex flux functions with sonic states.
Here we assume that f_α, $\alpha = L, R$, is a strictly convex function and possesses a sonic state $\bar{u}_\alpha$. We set: $\bar{v}_\alpha = \theta_\alpha^{-1}(\bar{u}_\alpha)$. With the pair (v_g, v_d), we associate the pair $(\tilde{v}_g, \tilde{v}_d)$ defined by

$$\begin{cases} \tilde{f}_L(\tilde{v}_g) = \tilde{f}_L(v_g),\ \tilde{v}_g \neq v_g \text{ if } v_g \neq \bar{v}_L \\ \\ \tilde{v}_g = \bar{v}_L \text{ if } v_g = \bar{v}_L \end{cases}, \quad \begin{cases} \tilde{f}_R(\tilde{v}_d) = \tilde{f}_R(v_d),\ \tilde{v}_d \neq v_d \text{ if } v_d \neq \bar{v}_R \\ \\ \tilde{v}_d = \bar{v}_R \text{ if } v_d = \bar{v}_R. \end{cases}$$

Using the results of Example 2.2, we thus have

$$\widetilde{\mathcal{F}}_L^-(v_g) = \left\{v \neq v_g;\, v < \min(\bar{v}_L, \tilde{v}_g),\, v \neq \tilde{v}_g\right\}$$

$$\widetilde{\mathcal{F}}_R^+(v_d) = \{v \neq v_d;\, v > \max(\bar{v}_R, \tilde{v}_d),\, v \neq \tilde{v}_d\}.$$

(a) If

$$v_d \in \widetilde{\mathcal{F}}_L^-(v_g) \Leftrightarrow v_d \leq \min(\bar{v}_L, \tilde{v}_g),\, v_d \neq \tilde{v}_g,$$

there exists a solution of type (i) (a L-wave) to the coupled Riemann problem. This is the only solution of this kind.
(b) If

$$v_g \in \widetilde{\mathcal{F}}_R^+(v_g) \Leftrightarrow v_g \geq \max(\bar{v}_R, \tilde{v}_d),\, v_g \neq \tilde{v}_d,$$

there exists a solution of type (ii) (a R-wave) to the coupled Riemann problem. This is the only solution of this type.
(c) When

$$\widetilde{\mathcal{F}}_L^-(v_g) \cap \widetilde{\mathcal{F}}_R^+(v_d) \neq \emptyset \Leftrightarrow \max(\bar{v}_R, \tilde{v}_d) \leq \min(\bar{v}_L, \tilde{v}_g),$$

we can construct a family of solutions of type (iii) (a L-wave followed by a R-wave) depending on the parameter $v(0) \in [\max(\bar{v}_R, \tilde{v}_d), \min(\bar{v}_L, \tilde{v}_g)]$. They are the only solutions of type (iii).

It is worthwile to notice that, given a pair (v_g, v_d), we may have v-continuous solutions of several types. For instance, if

$$v_d \leq \max(\bar{v}_R, \tilde{v}_d) \leq \min(\bar{v}_L, \tilde{v}_g)$$

solutions of types (i) and (iii) are valid.

1.2.3 v-discontinuous solutions.

We next look for the self-similar solutions of the coupled Riemann problem which are *v-discontinuous at the interface* $x = 0$. Setting

$$v_- = v(0_-), \quad v_+ = v(0_+),$$

the coupling constraints (1.7) read here

$$\begin{cases} v_- \in \widetilde{O}_L(v_+) \Leftrightarrow v_- = z_L(0_-; v_-, v_+), \\ \\ v_+ \in \widetilde{O}_R(v_-) \Leftrightarrow v_+ = z_R(0_+; v_-, v_+). \end{cases}$$

Since we assume $v_- \neq v_+$, $z_L(\cdot; v_-, v_+)$ and $z_R(\cdot; v_-, v_+)$ are both non trivial waves. The coupling constraints mean that $z_L(\cdot; v_-, v_+)$ is a wave with a nonnegative speed while $z_R(\cdot; v_-, v_+)$ is a wave with a nonpositive speed.

On the other hand, any solution of the coupled Riemann problem consists necessarily of a L-wave whose speed is nonpositive and a R-wave whose speed is nonnegative. In other words, the wave $z_L(\cdot; v_g, v_-)$ has a nonpositive speed while $z_R(\cdot; v_+, v_d)$ has a non negative speed (see Fig. 1.7).

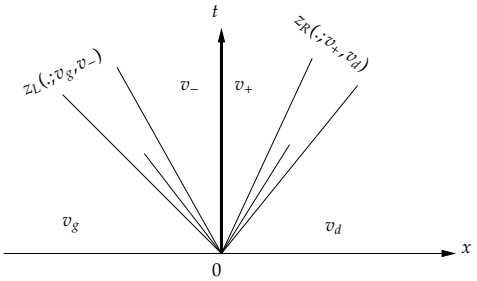


Figure 1.7: A v-discontinuous solution to the coupled Riemann problem

Let us then state

Lemma 1.7. *One of the two following situations holds:*
(i) $v_- = v_g \Rightarrow \tilde{f}_L'(v_g) \geq 0$;
(ii) $v_- \neq v_g \Rightarrow \tilde{f}_L'(v_-) = 0$ *and the right subwave of* $z_L(\cdot; v_g, v_-)$ *is a rarefaction.*

In the case (ii), it is worthwile to notice that v_- is a sonic state of the right rarefaction subwave of $z_L(\cdot; v_g, v_-)$.

Proof. We begin by proving the lemma when $\theta_L = id$, i.e., when $v = u$ is the conservative variable. Since $w_L(\cdot; u_-, u_+)$ has a nonnegative speed, we have

$$\min_{u \in I(u_-, u_+)} \frac{f_L(u) - f_L(u_-)}{u - u_-} \geq 0$$

which implies $f_L'(u_-) \geq 0$. If we assume $u_- = u_g$, we obtain $f_L'(u_g) \geq 0$. Assume next $u_- \neq u_g$. Then $w_L(\cdot; u_g, u_-)$ has a nonpositive speed so that

$$\max_{u \in I(u_g, u_-)} \frac{f_L(u) - f_L(u_-)}{u - u_-} \leq 0$$

which yields $f_L'(u_-) \leq 0$. Hence we find $f_L'(u_-) = 0$. As a consequence, the right subwave of $w_L(\cdot; u_g, u_-)$ is either a rarefaction with u_- as a sonic state or a stationary shock. But, since $w_L(0_-; u_g, u_-) = u_-$, a stationary shock is not allowed. This proves the lemma when $v = u$.

Let us now turn to the general case of a nonconservative variable v. Since $\tilde{f}_L'(v) = f_L'(\theta_L(v))\theta_L'(v)$ and $\theta_L'(v) > 0$, the above properties (i) and (ii) become respectively

$$v_- = v_g \Rightarrow \tilde{f}_L'(v_g) \geq 0$$

$$v_- \neq v_g \Rightarrow \tilde{f}_L'(v_-) = 0$$

and the proof is complete. □

Similarly, one can state

Lemma 1.8. *One of the two following situations holds:*
(i) $v_+ = v_d \Rightarrow \tilde{f}_R'(v_d) \leq 0$;
(ii) $v_+ \neq v_d \Rightarrow \tilde{f}_R'(v_+) = 0$ *and the left subwave of* $z_R(\cdot; v_+, v_d)$ *is a rarefaction.*

As a consequence of Lemmas 7 and 8, we find that the self-similar v-discontinuous solutions of the coupled Riemann problem are necessarily of the four following types.

(i) The first type of v-discontinuous solution (see Fig. 1.8). It consists of a stationary discontinuity with $v_- = v_g$ and $v_+ = v_d$. Such a solution exists if and only if we have

$$\tilde{f}_L'(v_g) \geq 0 \geq \tilde{f}_R'(v_d)$$

together with the coupling conditions which read here

$$\min_{v\in I(v_g,v_d)} \frac{\tilde{f}_L(v)-\tilde{f}_L(v_g)}{v-v_g} \geq 0,$$

$$\max_{v\in I(v_g,v_d)} \frac{\tilde{f}_R(v)-\tilde{f}_R(v_d)}{v-v_d} \leq 0.$$

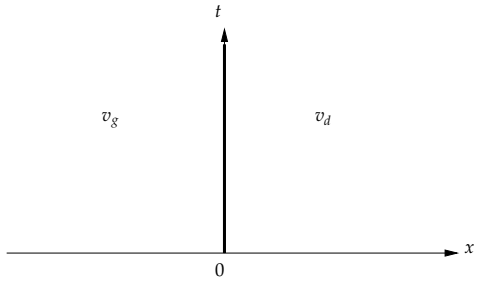


Figure 1.8: The first type of v-discontinuous solution : a stationary discontinuity.

(ii) The second type of v-discontinuous solution (see Fig. 1.9). It consists of a L-wave whose right subwave is a rarefaction with v_- as a sonic state followed by a stationary discontinuity with $v_+ = v_d$. Such a solution exists under the following conditions. On the one hand, we have

$$\tilde{f}_R'(v_d) \leq 0$$

and there exists a sonic state $\bar{v}_L$ of $\tilde{f}_L$, $\bar{v}_L \neq v_g$, such that

$$\max_{v\in I(v_g,\bar{v}_L)} \frac{\tilde{f}_L(v)-\tilde{f}_L(\bar{v}_L)}{v-\bar{v}_L} = \tilde{f}_L'(\bar{v}_L) = 0$$

and $v_- = \bar{v}_L$. On the other hand, we require the associated coupling conditions

$$\min_{v\in I(\bar{v}_L,v_d)} \frac{\tilde{f}_L(v)-\tilde{f}_L(\bar{v}_L)}{v-\bar{v}_L} \geq 0,$$

$$\max_{v\in I(\bar{v}_L,v_d)} \frac{\tilde{f}_R(v)-\tilde{f}_R(v_d)}{v-v_d} \leq 0.$$

(iii) The third type of v-discontinuous solution (see Fig. 1.10). It consists of a stationary discontinuity with $v_- = v_g$ followed by a R-wave whose left subwave is a rarefaction with v_+ as a sonic state. This solution exists under the following conditions. On the one hand, we have

$$\tilde{f}_L'(v_g) \geq 0$$

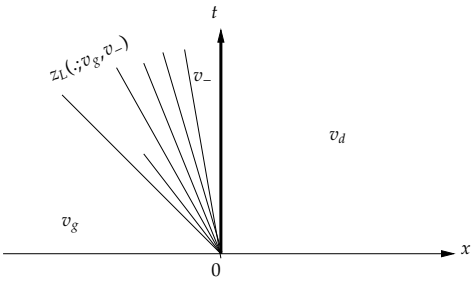


Figure 1.9: The second type of v-discontinuous solution : a L-wave whose right subwave is a rarefaction, followed by a stationary discontinuity.

and there exists a sonic state $\bar{v}_R$ of $\tilde{f}_R$, $\bar{v}_R \neq v_d$ such that

$$\min_{v\in I(\bar{v}_R,v_d)} \frac{\tilde{f}_R(v)-\tilde{f}_R(\bar{v}_R)}{v-\bar{v}_R} = \tilde{f}_R'(\bar{v}_R) = 0$$

and $v_+ = \bar{v}_R$. On the other hand, the associated coupling conditions read

$$\min_{v\in I(v_g,\bar{v}_R)} \frac{\tilde{f}_L(v)-\tilde{f}_L(v_g)}{v-v_g} \geq 0,$$

$$\max_{v\in I(v_g,\bar{v}_R)} \frac{\tilde{f}_R(v)-\tilde{f}_R(\bar{v}_R)}{v-\bar{v}_R} \leq 0.$$

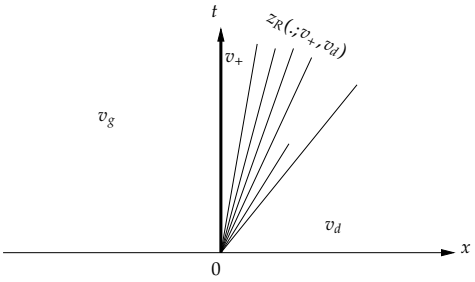


Figure 1.10: The third type of v-discontinuous solution : a stationary discontinuity followed by a R-wave whose left subwave is a rarefaction.

(iv) The fourth type of v-discontinuous solution (see Fig. 1.11). It consists of a L-wave whose right subwave is a rarefaction with v_- as a sonic state followed by a stationary discontinuity and a R-wave whose left subwave is a rarefaction with v_+ as a sonic state. For obtaining such a solution, the following conditions hold: there exist sonic states $\bar{v}_L \neq v_g$ and $\bar{v}_R) \neq v_d$ of $\tilde{f}_L$ and $\tilde{f}_R$ respectively such that

$$\max_{v\in I(v_g,\bar{v}_L)} \frac{\tilde{f}_L(v)-\tilde{f}_L(\bar{v}_L)}{v-\bar{v}_L} = \tilde{f}_L'(\bar{v}_L) = 0,$$

$$\min_{v\in I(\bar v_R, v_d)} \frac{\tilde f_R(v)-\tilde f_R(\bar v_R)}{v-\bar v_R} = \tilde f_R'(\bar v_R) = 0$$

and $v_- = \bar v_L, v_+ = \bar v_R$. In addition, we require the coupling conditions

$$\min_{v\in I(\bar v_L, \bar v_R)} \frac{\tilde f_L(v)-\tilde f_L(\bar v_L)}{v-\bar v_L} \geq 0$$

and

$$\max_{v\in I(\bar v_L, \bar v_R)} \frac{\tilde f_R(v)-\tilde f_R(\bar v_R)}{v-\bar v_R} \leq 0.$$

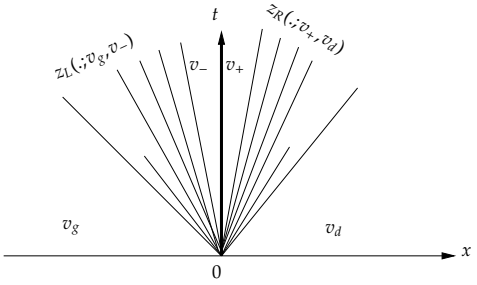


Figure 1.11: The fourth type of v-discontinuous solution : a L-wave whose right subwave is a rarefaction, followed by a stationary discontinuity, itself followed by a R-wave whose left subwave is a rarefaction.

Again we apply the above results to the cases where both flux functions are either strictly monotone or strictly convex.

Example 2.3. The case of strictly monotone flux functions. (contd.)
(a) If the functions f_L and f_R are strictly decreasing, we have $\tilde f_L' \leq 0$ and $\tilde f_R' \leq 0$ and no v-discontinuous solution can exist. This is obvious for solutions of types (i) and (iii). On the other hand, due to the coupling conditions, solutions of types (ii) and (iv) are not admissible. For instance, in the case of a solution of type (ii), the first coupling condition implies the existence of a sonic state $\bar v_L$ such that $\bar v_L = v_d$, i.e., $v_- = v_+$ which is clearly excluded.
(b) If f_L and f_R are strictly increasing, a similar analysis shows again that there cannot exist any v- discontinuous solution.
(c) If f_L is strictly increasing and f_R is strictly decreasing, only the v-discontinuous solution of type (i), i.e., a stationary discontinuity, is admissible. Indeed, such a solution is clearly admissible. On the other hand, a solution of type (ii) cannot exist since the condition

$$\max_{v\in I(v_g, \bar v_L)} \frac{\tilde f_L(v)-\tilde f_L(\bar v_L)}{v-\bar v_L} = \tilde f_L'(\bar v_L) = 0$$

implies $\bar v_L = v_g$ and therefore $v_- = v_g$ so that the L-wave does not exist. Using similar arguments, one can check that the solutions of types (iii) and (iv) are also excluded.
(d) If f_L is strictly decreasing and f_R is strictly increasing, no v-discontinuous solution may exist since the coupling conditions are never satisfied.

To summarize, a v-discontinuous solution exists only when f_L is strictly increasing and f_R is strictly decreasing . This is a stationary discontinuity. □

Example 2.4.(contd.) Consider again the case where both flux functions f_L and f_R are strictly convex and possess sonic states $\bar{u}_L$ and $\bar{u}_R$ respectively. Since, for $\alpha = L, R$, the function θ_α is assumed to satisfy $\theta'_\alpha > 0$, the function $\tilde{f}_\alpha$ has a unique sonic state $\bar{v}_\alpha = \theta_\alpha^{-1}(\bar{u}_\alpha)$ and is strictly decreasing in $(-\infty, \bar{v}_\alpha)$ (resp. strictly increasing in $(\bar{v}_\alpha, +\infty)$). We introduce again the states $\tilde{v}_g$ and $\tilde{v}_d$ defined above. Note that, in this strictly convex case, the sonic state $\bar{v}_L \neq v_g$ of $\tilde{f}_L$ satisfies the condition

$$\max_{v \in I(v_g, \bar{v}_L)} \frac{\tilde{f}_L(v) - \tilde{f}_L(v_g)}{v - v_g} = \tilde{f}'_L(\bar{v}_L) = 0$$

if and only if $v_g < \bar{v}_L$ or equivalently $\tilde{f}'_L(v_g) < 0$. Similarly, the sonic state $\bar{v}_R \neq v_d$ of $\tilde{f}_R$ satisfies the condition

$$\min_{v \in I(\bar{v}_R, v_d)} \frac{\tilde{f}_R(v) - \tilde{f}_R(v_d)}{v - v_d} = \tilde{f}'_R(\bar{v}_R) = 0$$

if and only if $v_d > \bar{v}_R$ or equivalently $\tilde{f}'_R(v_d) > 0$.

On the other hand, for $v_a = v_g \geq \bar{v}_L$ or $v_a = \bar{v}_L$, a coupling condition of the form

$$\min_{I(v_a, v_b)} \frac{\tilde{f}_L(v) - \tilde{f}_L(v_a)}{v - v_a} \geq 0$$

holds if and only if $v_b \geq \tilde{v}_a$ where

$$\tilde{v}_a = \begin{cases} \tilde{v}_g & \text{if } v_a = v_g \\ \bar{v}_L & \text{if } v_a = \bar{v}_L. \end{cases}$$

Similarly, for $v_b = v_d \leq \bar{v}_R$ or $v_b = \bar{v}_R$, a coupling condition of the form

$$\max_{I(v_a, v_b)} \frac{\tilde{f}_R(v) - \tilde{f}_R(v_b)}{v - v_b} \leq 0$$

holds if and only if $v_a \leq \tilde{v}_b$ where

$$\tilde{v}_b = \begin{cases} \tilde{v}_d & \text{if } v_b = v_d \\ \bar{v}_R & \text{if } v_b = \bar{v}_R. \end{cases}$$

Then, it is an easy matter to check that a v-discontinuous solution exists in the following situations.

(a) If

$$\bar{v}_L \leq v_g \leq \tilde{v}_d, \quad \tilde{v}_g \leq v_d \leq \bar{v}_R,$$

we obtain a v-discontinuous solution of type (i), i.e., a stationary discontinuity.

(b) If

$$v_g < \bar{v}_L < v_d \leq \bar{v}_R,$$

we find a v-discontinuous solution of type (ii), i.e., a L-wave followed by a stationary discontinuity.
(c) If

$$\bar{v}_L \leq v_g < \bar{v}_R < v_d,$$

we obtain a v-discontinuous solution of type (iii), i.e., a a stationary discontinuity followed by a R-wave.
(d) If

$$v_g < \bar{v}_L < \bar{v}_R < v_d,$$

we find a v-discontinuous solution of type (iv), i.e., a L-wave followed by a stationary discontinuity and a R-wave.
Note that each case (i)-(iv) is disclosed from the others and each v-discontinuous solution is uniquely defined. □

1.2.4 Solution of the coupled Riemann problem.

We are now able to solve the coupled Riemann problem for all pair (u_g, u_d) or (v_g, v_d). We begin with the cases where the flux functions f_α, $\alpha = L, R$, are either strictly monotone or strictly convex.

Example 2.3.(contd.) We first assume that f_α, $\alpha = L, R$ is a strictly monotone function. Combining the above results, we obtain the following conclusions.
(a) *The functions f_L and f_R are strictly decreasing.* The solution is v-continuous: it is a L-wave.
(b) *The functions f_L and f_R are strictly increasing.* The solution is v-continuous: it is a R-wave.
(c) *The function f_L is strictly increasing and the function f_R strictly decreasing.* The solution is v- discontinuous: it is a stationary discontinuity.
(d) *The function f_L is strictly decreasing and the function f_R is strictly increasing.* The solutions are v- continuous and form a one-parameter family depending on thre parameter $v(0) \in \mathbb{R}$. For $v(0) \neq v_g, v_d$, we obtain a L-wave followed by a R-wave. For $v(0) = v_d$, we obtain a L-wave while, for $v(0) = v_g$, we get a R-wave.

To summarize, in this case, the coupled Riemann problem has always a solution. This solution is unique except in the subcase (d). Note that, as in [76], one could have obtained directly the above results by using a method of characteristics.

Example 2.4.(contd.) Assume now that the flux functions f_L and f_R are strictly convex and possess sonic states $\bar{u}_L$ and $\bar{u}_R$ respectively.Let us check that the coupled Riemann problem has at least one solution. First of all, we already know from the results of section 2.1.2 that a v- continuous solution exists in the following cases.
(a) For $v_d \leq \min(\bar{v}_L, \tilde{v}_g)$, $\quad v_d \neq \tilde{v}_g$, the solution is a L-wave.
(b) For $v_g \geq \max(\bar{v}_R, \tilde{v}_d)$, $\quad v_g \neq \tilde{v}_d$, the solution is a R-wave.

(c) If $\max(\bar{v}_R, \tilde{v}_d) \leq \min(\bar{v}_L, \tilde{v}_g)$, we obtain a family of v-continuous solutions consisting of a L-wave followed by a R-wave and depending on the parameter $v(0) \in [\max(\bar{v}_R, \tilde{v}_d), \min(\bar{v}_L, \tilde{v}_g)]$.

It remains to exhibit a v-discontinuous solution when a v- continuous one does not exist, i.e., when the pair (v_g, v_d) satisfies the conditions

$$\begin{cases} v_d > \min(\bar{v}_L, \tilde{v}_g) \\ v_g < \max(\bar{v}_R, \tilde{v}_d) \\ \max(\bar{v}_R, \tilde{v}_d) > \min(\bar{v}_L, \tilde{v}_g). \end{cases} \tag{1.34}$$

In fact, it is convenient to distinguish the following cases:

$$(v_g \geq \bar{v}_L, v_d \leq \bar{v}_R),\ (v_g \geq \bar{v}_L, v_d > \bar{v}_R),\ (v_g < \bar{v}_L, v_d \leq \bar{v}_R),\ (v_g < \bar{v}_L, v_d > \bar{v}_R).$$

(d) For $(v_g \geq \bar{v}_L, v_d \leq \bar{v}_R)$, the conditions (1.34) become respectively

$$v_d < \tilde{v}_g, \quad v_g < \tilde{v}_d, \quad \tilde{v}_d > \tilde{v}_g.$$

This case is therefore characterized by

$$\bar{v}_L \leq v_g < \tilde{v}_d, \quad \tilde{v}_g < v_d \leq \bar{v}_R.$$

Then, applying the results of section 2.1.3, we obtain that the solution of the coupled Riemann problem is a stationary discontinuity.
(e) For $(v_g \geq \bar{v}_L, v_d > \bar{v}_R)$, the conditions (1.34) read

$$v_d > \tilde{v}_g, \quad v_g < \bar{v}_R, \quad \bar{v}_R > \tilde{v}_g$$

so that this case is characterized by

$$\bar{v}_L \leq v_g < \bar{v}_R < v_d.$$

This implies that the solution is a L-wave followed by a stationary discontinuity.
(f) For $(v_g < \bar{v}_L, v_d \leq \bar{v}_R)$, (1.34) gives

$$v_d > \bar{v}_L, \quad v_g < \tilde{v}_R, \quad \tilde{v}_d > \bar{v}_L.$$

This leads us to the characterization

$$v_g < \bar{v}_L < v_d \leq \bar{v}_R$$

and we obtain a solution consisting of a stationary discontinuity followed by a R-wave.
(g) For $(v_g < \bar{v}_L, v_d > \bar{v}_R)$, the conditions (1.34) become

$$v_d > \bar{v}_L, \quad v_g < \bar{v}_R, \quad \bar{v}_R > \bar{v}_L$$

and therefore

$$v_g < \bar{v}_L < \bar{v}_R < v_d.$$

We find a solution consisting of a L-wave followed by a stationary discontinuity and a R-wave.

Observe that, in each case (d)-(g), the conditions (1.34) are exactly the conditions obtained in the previous section which ensure the existence and uniqueness of a v-discontinuous solution. We thus have proved

Theorem 1.9. *Assume that the functions f_L and f_R are strictly convex and possess sonic states. Then the coupled Riemann problem has at least one solution. The solution is unique except in the case (c) where there exists a one-parameter family of v-continuous solutions.* □

We pass to the general case of arbitrary flux functions. The situation is not as simple as in the above examples due to the possible presence of several sonic states. The purpose of the remaining part of this section is to prove

Theorem 1.10. *Assume that the flux functions f_L and f_R are C^1 functions. Then the coupled Riemann problem has at least one self-similar solution.*

We know already that we can construct a v-continuous solution in the following cases:

$$v_d \in \widetilde{\mathcal{F}}_L^-(v_g), \quad v_g \in \widetilde{\mathcal{F}}_R^+(v_d), \quad \widetilde{\mathcal{F}}_L^-(v_g) \cap \widetilde{\mathcal{F}}_R^+(v_d) \neq \emptyset.$$

It remains to construct at least one v-discontinuous solution of the coupled Riemann problem when

$$v_d \notin \widetilde{\mathcal{F}}_L^-(v_g), \quad v_g \notin \widetilde{\mathcal{F}}_R^+(v_d), \quad \widetilde{\mathcal{F}}_L^-(v_g) \cap \widetilde{\mathcal{F}}_R^+(v_d) = \emptyset. \tag{1.35}$$

We begin with the following remarks. The condition $v_d \notin \widetilde{\mathcal{F}}_L^-(v_g)$ means that $z_L(\cdot; v_g, v_d)$ possesses a nontrivial subwave whose speed is nonnegative. Otherwise, we would get $z_L(\frac{x}{t}; v_g, v_d) = v_d$ for all $x \geq 0$ and therefore $v_d \in \widetilde{\mathcal{F}}_L^-(v_g)$. Hence, we have

$$v_L(0_-) \overset{def}{=} z_L(0_-; v_g, v_d) \neq v_d.$$

Similarly, the condition $v_g \notin \widetilde{\mathcal{F}}_R^+(v_d)$ means that $z_R(\cdot; v_g, v_d)$ possesses a nontrivial subwave whose speed is nonpositive so that

$$v_R(0_+) \overset{def}{=} z_R(0_+; v_g, v_d) \neq v_g.$$

Note that the hypotheses (1.35) imply $v_L(0_-) \neq v_R(0_+)$. Otherwise the function

$$z(\frac{x}{t}; v_g, v_d) = \begin{cases} z_L(\frac{x}{t}; v_g, v_L(0_-)), & \frac{x}{t} < 0 \\ \\ z_R(\frac{x}{t}; v_R(0_+), v_d), & \frac{x}{t} > 0 \end{cases}$$

would be a v-continuous solution of the coupled Riemann problem. On the other hand, we have either $v_L(0_-) = v_g$ (resp. $v_R(0_+) = v_d$) or $v_L(0_-)$ (resp. $v_R(0_+)$) is a sonic state of $\tilde{f}_L$

(resp. $\tilde{f}_R$). Hence, it appears fairly natural to consider the function

$$z(\frac{x}{t};v_g,v_d) = \begin{cases} z_L(\frac{x}{t};v_g,v_d), & x<0 \\ \\ z_R(\frac{x}{t};v_g,v_d), & x>0 \end{cases} \tag{1.36}$$

as a possible solution of the coupled Riemann problem. Indeed, we can state

Lemma 1.11. *Assume the hypotheses (1.35) together with*

$$\begin{cases} v_L(0_-) < v_R(0_+) & if\ v_g < v_d \\ \\ v_L(0_-) > v_R(0_+) & if\ v_g > v_d. \end{cases} \tag{1.37}$$

Then (1.36) is a solution of the coupled Riemann problem.

Proof. We have only to check the coupling conditions which read here

$$\min_{v\in I(v_L(0_-),v_R(0_+))} \frac{\tilde{f}_L(v)-\tilde{f}_L(v_L(0_-))}{v-v_L(0_-)} \geq 0$$

and

$$\max_{v\in I(v_L(0_-),v_R(0_+))} \frac{\tilde{f}_R(v)-\tilde{f}_R(v_R(0_+))}{v-v_R(0_+)} \geq 0.$$

Assume for instance $v_d > v_g$. Since, in that case, $z_L(\cdot;v_g,v_d)$ and $z_R(\cdot;v_g,v_d)$ are monotonically increasing functions, we have by (1.37)

$$v_g \leq v_L(0_-) < v_R(0_+) \leq v_d.$$

Now, we observe that $z_L(\cdot;v_L(0_-),v_d)$ has a nonnegative speed, i.e.,

$$\min_{v\in I(v_L(0_-),v_d)} \frac{\tilde{f}_L(v)-\tilde{f}_L(v_L(0_-))}{v-v_L(0_-)} \geq 0$$

which implies the first coupling condition. On the other hand, $z_R(\cdot;v_g,v_R(0_+))$ has a nonpositive speed, i.e.,

$$\max_{v\in I(v_g,v_R(0_+))} \frac{\tilde{f}_R(v)-\tilde{f}_R(v_R(0_+))}{v-v_R(0_+)} \leq 0$$

which yields the second coupling condition. The case $v_g > v_d$ is analyzed in the same way. □

Note that the proof of the above lemma only uses the first two hypotheses (1.35). Observe that this proof fails if the conditions (1.37) do not hold. It remains to construct a solution of the coupled Riemann problem when either

$$v_g < v_d \quad \text{and} \quad v_L(0_-) > v_R(0_+) \tag{1.38}$$

or

$$v_g > v_d \quad \text{and} \quad v_L(0_-) < v_R(0_+). \tag{1.39}$$

Assume first (1.38). Let us then check that there exists at least one sonic state of $\tilde{f}_L$ in $[v_g, v_R(0_+)]$. It is here convenient to work with the conservative variable u : setting $u_g = \theta_L(v_g)$, $u_R(0_+) = \theta_L(v_R(0_+))$, we introduce the lower convex envelope of f_L in the interval $[u_g, u_R(0_+)]$. This envelope function cannot be strictly decreasing. Otherwise, $w_L(\cdot; u_g, u_R(0_+))$ and therefore $z_L(\cdot; v_g, v_R(0_+))$, would be a wave whose speed is negative. One then could exhibit a v-continuous solution of the coupled Riemann problem, namely

$$z(\frac{x}{t}; v_g, v_d) = \begin{cases} z_L(\frac{x}{t}; v_g, v_R(0_+)), & \frac{x}{t} < 0 \\ \\ z_R(\frac{x}{t}; v_R(0_+), v_d), & \frac{x}{t} > 0. \end{cases}$$

Hence the above envelope function has either a unique minimum which is a sonic state of f_L or an interval of minima which contains such sonic states (at least the end points of this interval). Denote by $\bar{u}_-$ the smallest of all sonic states of both f_L and its lower convex envelope in $[u_g, u_R(0_+)]$. Then, $\bar{v}_- = \theta_L^{-1}(\bar{u}_-)$ is a sonic state of $\tilde{f}_L$ in $[v_g, v_R(0_+)]$. In the same way, there exists at least one sonic state of both f_R and its lower convex envelope in the interval $[u_L(0_-), u_d]$ and we denote by $\bar{u}_+$ the largest of all such sonic states. Then, $\bar{v}_+ = \theta_R^{-1}(\bar{u}_+)$ is a sonic state of $\tilde{f}_R$ in $[v_L(0_-), v_d]$. Now, it appears natural to consider the functions

$$z(\frac{x}{t}; v_g, v_d) = \begin{cases} z_L(\frac{x}{t}; v_g, \bar{v}_-), & \frac{x}{t} < 0 \\ \\ z_R(\frac{x}{t}; v_R(0_+), v_d), & \frac{x}{t} > 0 \end{cases} \tag{1.40}$$

and

$$z(\frac{x}{t}; v_g, v_d) = \begin{cases} z_L(\frac{x}{t}; v_g, v_L(0_-)), & \frac{x}{t} < 0 \\ \\ z_R(\frac{x}{t}; \bar{v}_+, v_d), & \frac{x}{t} > 0 \end{cases} \tag{1.41}$$

as possible candidates to the solution of the coupled Rieman problem. In fact, we can state

Lemma 1.12. *Assume the hypotheses (1.35) and (1.38). Then (1.40) and (1.41) are solutions of the coupled Riemann problem.*

Proof. Let us show that (1.40) is indeed solution. Again, we have to check the associated coupling conditions

$$\bar{v}_- = z_L(0_-; \bar{v}_-, v_R(0_+)) \Leftrightarrow \bar{u}_- = w_L(0_-; \bar{u}_-, u_R(0_+))$$

and

$$v_R(0_+) = z_R(0_+; \bar{v}_-, v_R(0_+)) \Leftrightarrow u_R(0_+) = w_R(0_+; \bar{u}_-, u_R(0_+)).$$

The first coupling condition holds since, by construction, $w_L(\cdot; \bar{u}_-, u_R(0_+))$ is a monotonically increasing function in $[\bar{u}_-, u_R(0_+)]$ and the corresponding wave has a nonnegative speed.

Consider next the second coupling condition. We know that the lower convex envelope of f_R in $[u_g, u_d]$ is a monotonically decreasing function in the interval $[u_g, u_R(0_+)]$ and is strictly convex in an interval $[u_R(0_+), u_R(0_+)+\varepsilon]$, $\varepsilon > 0$ small enough (cf. Fig. 1.12). Then, as $u_g < \bar{u}_- < u_R(0_+)$, it is clear geometrically that the lower convex envelope of f_R in the interval $[\bar{u}_-, u_R(0_+)]$ is a monotonically decreasing function so that $u_R(0_+) = w_R(0_+; \bar{u}_-, u_R(0_+))$ and our assertion is proved.
By using similar arguments, one can prove that (1.41) is also solution. □

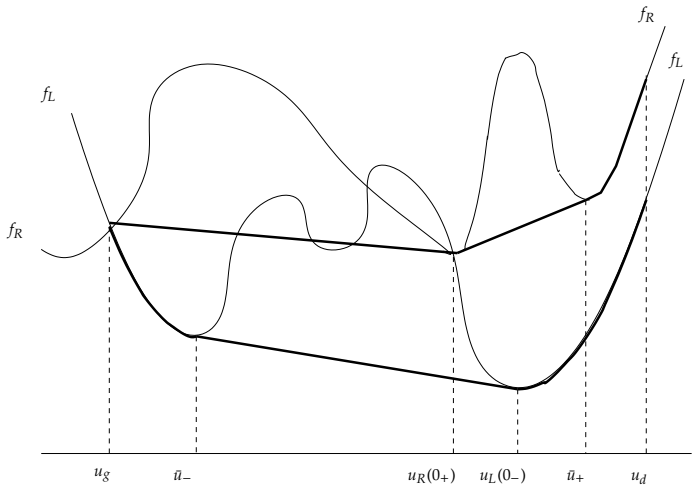


Figure 1.12: A typical example where $v_d \notin \widetilde{\mathcal{F}}_L^-(v_g)$, $\quad v_g \notin \widetilde{\mathcal{F}}_R^+(v_d)$, $\quad \widetilde{\mathcal{F}}_L^-(v_g) \cap \widetilde{\mathcal{F}}_R^+(v_d) = \emptyset$.

Remark. At first glance it would seem natural to consider the function

$$z(\frac{x}{t}; v_g, v_d) = \begin{cases} z_L(\frac{x}{t}; v_g, \bar{v}_-), & \frac{x}{t} < 0 \\ \\ z_R(\frac{x}{t}; \bar{v}_+, v_d), & \frac{x}{t} > 0 \end{cases}$$

as a possible solution of the coupled Riemann problem. However, this is not true since one can easily check that the coupling conditions

$$\bar{v}_- = z_L(0_-; \bar{v}_-, \bar{v}_+), \quad \bar{v}_+ = z_R(0_+; \bar{v}_-, \bar{v}_+)$$

are not satisfied in general (cf. Fig. 1.12). □

We can also state the analogue of Lemma 10 whose proof follows the same lines as above.

Lemma 1.13. *Assume the hypotheses (1.35) and (1.39). Then the coupled Riemann problem has at least two v-discontinuous solutions.*

Theorem 3 is now an obvious consequence of Lemmas 10 and 11.

A natural question now arises: when the coupled Riemann problem posesses several solutions, does there exist any "reasonable" criterion based on entropy or stability arguments for choosing the "right solution"? As a first step in this direction, we conjecture that, if a

v-continuous solution exists, the eventual v-discontinuous solutions should be considered as parasitic ones.

1.2.5 The coupled Riemann problem for two conservation laws "with phase change".

One can extend the above results to the case where the flux functions f_α, $\alpha = L, R$, are only piecewise C^1. For simplicity, we will restrict ourselves in this section to *continuous* functions f_α which satisfy the following properties:
(i) f_α is a C^1 strictly increasing function in the intervals $(-\infty, a_\alpha)$ and $(b_\alpha, +\infty)$, $a_\alpha < b_\alpha$;
(ii) f_α is constant in the interval $[a_\alpha, b_\alpha]$.
One can think of each flux function f_α as modeling a diphasic behavior: the states $u < a_\alpha$ and $u > b_\alpha$ correspond to different phases while the states $u \in [a_\alpha, b_\alpha]$ correspond to a mixture of the two phases.

Again for simplicity, we will restrict ourselves to the u-coupling method. Before constructing the solution of the coupled Riemann problem, let us recall the properties of the solution $w(\cdot; u_g, u_d)$ of the usual Riemann problem associated with such a function $f = f_\alpha$ [3]. By introducing the lower convex envelope (resp. the upper concave envelope) of f between the states u_g and u_d if $u_g < u_d$ (resp. $u_g > u_d$), it is a simple matter to check the following properties of $w(\cdot; u_g, u_d)$: (i) the associated (composite) wave has a nonnegative speed;
(ii) the function $x \to w(\frac{x}{t}; u_g, u_d)$ is continuous at $x = 0$ in the following cases

$$\begin{cases} u_g < a, & u_d \in \mathbb{R} \\ u_g = a, & u_d < a \\ u_g > b, & u_d \in \mathbb{R} \\ u_g = b, & u_d > b; \end{cases} \tag{1.42}$$

(iii) the function $x \to w(\frac{x}{t}; u_g, u_d)$ is discontinuous at $x = 0$ in the following cases

$$\begin{cases} u_g, u_d \in [a, b] \\ u_g \in [a, b), & u_d > b \\ u_g \in (a, b], & u_d < a. \end{cases} \tag{1.43}$$

In the first case, $w(\cdot; u_g, u_d)$ consists of a stationary shock while, in the last two cases, $w(\cdot; u_g, u_d)$ is a composite wave whose left subwave is a stationary shock.

Let us now consider the coupled Riemann problem. Instead of establishing general results for piecewise C^1 flux functions, it is here far simpler to use a direct approach. Since the function f_L is monotonically increasing, the solution of the coupled Riemann problem cannot include a L-wave. Therefore a solution consists of a possible stationary shock wave connecting u_g and u_+ and a R-wave connecting u_+ and u_d.
Assume first $u_g = u_+$ so that the solution is continuous at the interface $x = 0$. Then, using (1.42), we know that this is indeed the case if and only if the pair (u_g, u_d) satisfies one of the

[3] we drop the subscript α for simplicity.

following properties

$$\begin{cases} u_g < a_R, & u_d \in \mathbb{R} \\ u_g = a_R, & u_d < a_R \\ u_g > b_R, & u_d \in \mathbb{R} \\ u_g = b_R, & u_d > b_R. \end{cases} \tag{1.44}$$

Assume next $u_g \neq u_+$. This occurs if and only if, on the one hand, the coupling conditions hold and, on the other hand, $w_R(\cdot; u_+, u_d)$ is either a trivial wave (i.e., $u_d = u_+$) or a wave whose speed is positive. Since the function f_L is monotonically increasing, the first coupling condition

$$\min_{u \in I(u_g, u_+)} \frac{f_L(u) - f_L(u_g)}{u - u_g} \geq 0$$

holds trivially. The second coupling condition

$$\max_{u \in I(u_g, u_+)} \frac{f_R(u) - f_R(u_+)}{u - u_+} \leq 0$$

means that the wave $w_R(\cdot; u_g, u_+)$ has a nonpositive speed. Hence $w_R(\cdot; u_g, u_+)$ is necessarily a stationary shock or equivalently (cf. property (iii) above) we have

$$u_g, u_+ \in [a_R, b_R].$$

If $u_+ = u_d$, we thus have an admissible stationary shock for the solution of the coupled Riemann problem as soon as

$$u_g, u_d \in [a_R, b_R], \quad u_g \neq u_d.$$

Consider next the case $u_+ \neq u_d$. For the speed of the wave $w_R(\cdot; u_+, u_d)$ to be positive, it follows from (1.42) that we must have either

$$u_+ = a_R, \quad u_d < a_R$$

or

$$u_+ = b_R, \quad u_d > b_R.$$

In both cases, one can easily check that the speed of the wave is indeed positive.

As a conclusion, we obtain that the coupled Riemann problem has a unique solution. This solution is u-continuous at the interface $x = 0$ in the cases (1.44) and is u-discontinuous otherwise, i.e., when either

$$u_g, u_d = u_+ \in [a_R, b_R], \quad u_g \neq u_d \tag{1.45}$$

or

$$\begin{cases} u_g \in (a_R, b_R], & u_+ = a_R, \quad u_d < a_R \\ u_g \in [a_R, b_R), & u_+ = b_R, \quad u_d > b_R. \end{cases} \tag{1.46}$$

This result is easily extended to the case of a v-coupling method. It may be viewed as

a generalization of the results of Example 2.3 when both flux functions f_α are strictly increasing.

1.3 Numerical experiments

Our objective in this section is to illustrate numerically the theoretical results we obtained in the previous sections. For that, the following configurations will be considered :

- the case of two strictly monotone flux functions (example 2.3 above),
- the case of two strictly convex flux functions (example 2.4 above),
- a particular configuration where two discontinuous and none continuous (at the coupling interface) solutions are admissible,
- a particular configuration where several discontinuous solutions and continuous solutions exist,
- and the coupling of two conservation laws "with phase change".

The situations leading to several admissible solutions (continuous or discontinuous at interface) are of particular interest since different numerical schemes may capture different solutions. We begin with a brief description of the proposed numerical strategy and then present some numerical results.

1.3.1 Numerical strategy

We consider a finite volume approach. Let Δx and Δt denote the uniform space and time steps and $C_{j+1/2}$ be the cells defined by $C_{j+1/2} = (x_j, x_{j+1})$ with $x_j = j\Delta x$ and whose centers are $x_{j+1/2} = (j+1/2)\Delta x$ for all $j \in \mathbb{Z}$. We set $\lambda = \Delta t/\Delta x$ and $t_n = n\Delta t$ for $n \in \mathbb{N}$. The approximate solution is assumed to be piecewise constant on each cell $C_{j+1/2}$ and at each time t^n and the corresponding value is denoted $u^n_{j+1/2}$.
To begin with, we set as usual

$$u^0_{j+1/2} = \frac{1}{\Delta x}\int_{C_{j+1/2}} u_0(x)dx, \quad j \in \mathbb{Z},$$

where u_0 denotes a given initial condition of the coupling problem.
Then, let G_α, $\alpha = L, R$ be two two-point numerical flux functions that we assume to be consistant with f_α, $\alpha = L, R$. We propose the following update formula for $u^{n+1}_{j+1/2}$:

$$\begin{aligned} u^{n+1}_{j-1/2} &= u^n_{j-1/2} - \lambda(G^n_{L,j} - G^n_{L,j-1}), \qquad j \le 0,\ n \ge 0, \\ u^{n+1}_{j+1/2} &= u^n_{j+1/2} - \lambda(G^n_{R,j+1} - G^n_{R,j}), \qquad j \ge 0,\ n \ge 0, \end{aligned} \tag{1.47}$$

with $G^n_{\alpha,j} = G_\alpha(u^n_{j-1/2}, u^n_{j+1/2})$ for $j \neq 0$. In other words, this consists in a classical finite volume scheme outside of the interface, and only both fluxes $G^n_{L,0}$ and $G^n_{R,0}$ remain to be precised in order to define the numerical coupling procedure at the interface. Following

the previous works Godlewski et Raviart [76], Godlewski *et al.* [73], Ambroso *et al.* [7] (see also Ambroso *et al.* [9; 10; 11]), we set

$$\begin{aligned} G^n_{L,0} &= G_L(u^n_{-1/2}, \theta_L(v^n_{1/2})), \\ G^n_{R,0} &= G_R(\theta_R(v^n_{-1/2}), u^n_{1/2}), \end{aligned} \tag{1.48}$$

where ghost states $v^n_{\pm 1/2}$ are obtained as

$$\begin{aligned} v^n_{-1/2} &= \theta_L^{-1}(u^n_{-1/2}), \\ v^n_{1/2} &= \theta_R^{-1}(u^n_{1/2}). \end{aligned} \tag{1.49}$$

Note from now on that for convenience, we will restrict ourselves to the simple case $\theta_L = \theta_R = id$, so that condition (1.5) reads $u(t,0^-) = u(t,0^+)$ and the ghost states at the interface are simply

$$\begin{aligned} v^n_{-1/2} &= u^n_{-1/2}, \\ v^n_{1/2} &= u^n_{1/2}. \end{aligned} \tag{1.50}$$

At last and as far as the numerical flux functions G_α, $\alpha = L, R$ are concerned, we will consider the celebrated Godunov scheme :

$$G_\alpha(u,v) = \begin{cases} \min_{w\in[u,v]} f_\alpha(w), & u \le v, \\ \max_{w\in[v,u]} f_\alpha(w), & v < u, \end{cases} \tag{1.51}$$

and a relaxation scheme (see for instance Lattanzio et Serre [98]) defined by :

$$G_\alpha(u,v) = \frac{1}{2}\Big(f_\alpha(u) + f_\alpha(v)\Big) + \frac{a(u,v)}{2}(u-v) \quad \text{with} \quad a(u,v) = \max_{[\min(u,v),\max(u,v)]} |f'|. \tag{1.52}$$

1.3.2 Numerical results

Let us now present the numerical tests and results.

Test 1. The case of strictly monotone flux functions.
As Riemann initial data, we take

$$u_0(x) = \begin{cases} u_g & \text{if} \quad x < 0, \\ u_d & \text{if} \quad x > 0, \end{cases}$$

with $u_g = -2$ and $u_d = 2$ and we consider the following cases :

(a) $f_L(u) = -u$ and $f_R(u) = -2u$: the unique solution is continuous at the coupling interface.

(b) $f_L(u) = u$ and $f_R(u) = 2u$: the unique solution is continuous at the coupling interface.

(c) $f_L(u) = u$ and $f_R(u) = -2u$: there is no continuous solution but a unique discontinuous solution.

(d) $f_L(u) = -u$ and $f_R(u) = 2u$: there is no discontinuous solution and a continuum of continuous solutions.

Numerical results, obtained for both relaxation and Godunov approaches, are presented on Fig.1.13. These results are in agreement with the above theoretical results. Note that in the last case both numerical schemes capture the same continuous solution.

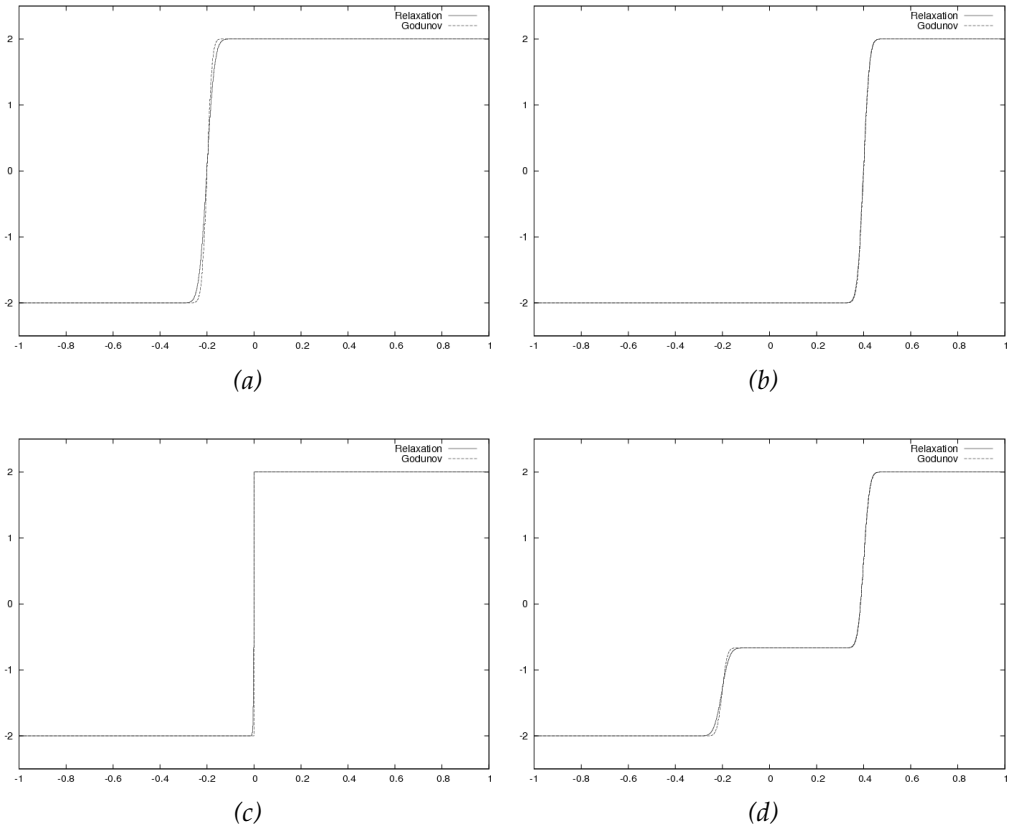


Figure 1.13: Monotone fluxes. 1000 pts. $t = 0.2$. $u_g = -2$, $u_d = 2$.

Test 2. The case of strictly convex flux functions with sonic states.
We consider two different cases, according to the relative position of the sonic points $\bar{v}_L$ and $\bar{v}_R$.

Test 2.1. $f_L(u) = u^2/2$, $f_R(u) = (u-1)^2/2$.
The sonic points are $\bar{v}_L = 0$ and $\bar{v}_R = 1$ and we consider the following four Riemann problems :

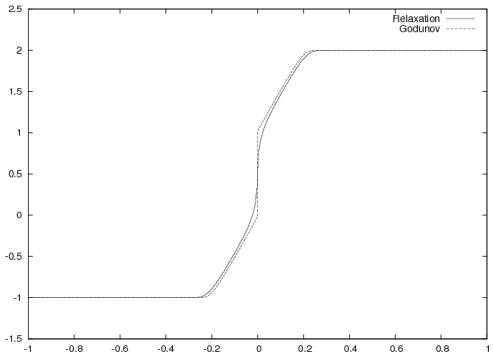
(a) $u_g = -1$ and $u_d = 2$: the unique solution is a discontinuous solution of the fourth type (composite wave consisting of a L-wave, a discontinuity and a R-wave).

(b) $u_g = 2$ and $u_d = -1$: the unique solution is a discontinuous solution of the first type.

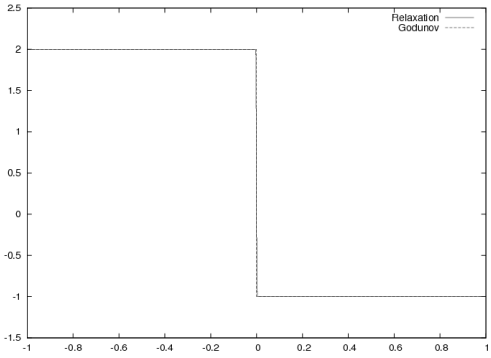
(c) $u_g = 0.5$ and $u_d = -2$: the solutions are a continuous solution of the first type and a discontinuous solution of the first type.

(d) $u_g = -1$ and $u_d = -1.5$: the solutions are a continuous solution of the first type and a discontinuous solution of the second type (that is not a monotonous solution).

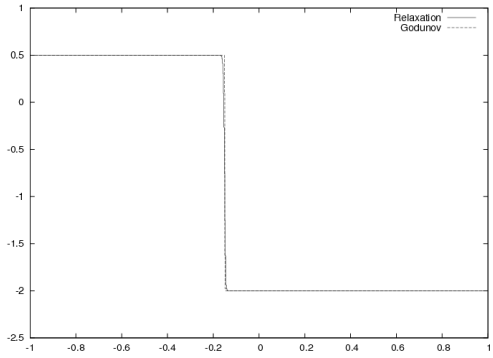
Numerical solutions are presented on Fig.1.14. We observe that both numerical schemes select the continuous solutions when more are presents (cases (c) and (d)).



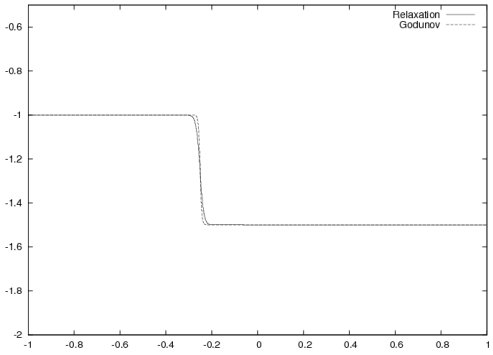
(a) $u_g = -1$, $u_d = 2$



(b) $u_g = 2$, $u_d = -1$



(c) $u_g = 0.5$, $u_d = -2$



(d) $u_g = -1$, $u_d = -1.5$

Figure 1.14: Convex fluxes. 1000 pts. $t = 0.2$

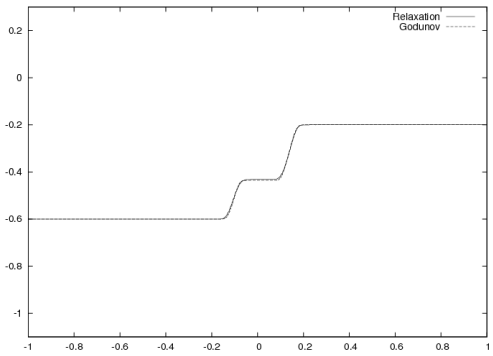
Test 2.2. $f_L(u) = u^2/2$, $f_R(u) = (u+1)^2/2$.
The sonic points are $\bar{v}_L = 0$ and $\bar{v}_R = -1$ and we consider the following four Riemann problems. The numerical solutions are presented on Fig.1.15.

(a) $u_g = -0.6$ and $u_d = -0.2$: the solutions are a continuous solution of the first, second or third type, and a discontinuous solution of the fourth type. Both numerical schemes capture the continuous solution of the third type, here with two rarefaction waves connecting a constant state, that slightly differs for both schemes.

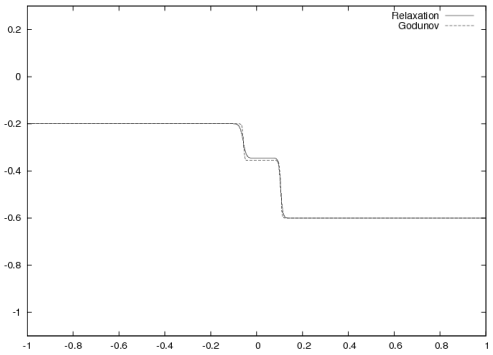
(b) $u_g = -0.2$ and $u_d = -0.6$: the solutions are a continuous solution of the first, second or third type, and a discontinuous solution of the fourth type. Both numerical schemes capture the continuous solution of the third type, here with two shock waves connecting a constant state, that slightly differs for both schemes.

(c) $u_g = 0.5$ and $u_d = 1$: the solutions are a continuous solution of the second type, and a discontinuous solution of the third type. Both numerical schemes capture the unique continuous solution of the second type.

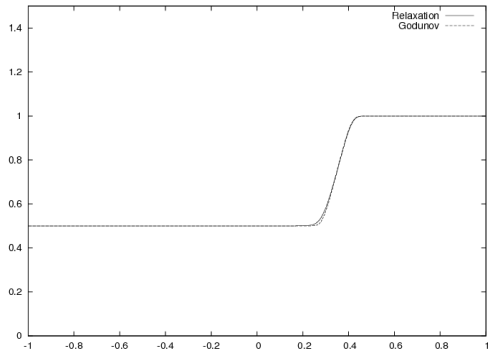
(d) $u_g = 1$ and $u_d = -1.5$: the solutions are a continuous solution of the first, second or third type, and a discontinuous solution of the first type. Both numerical schemes capture the continuous solution of the second type.



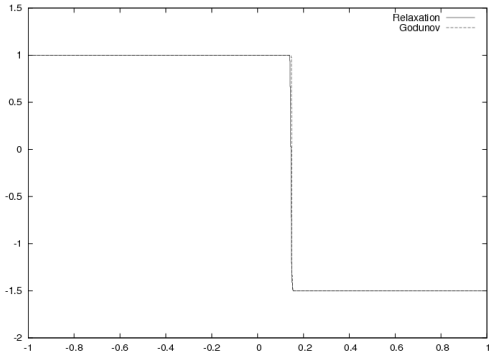
(a) $u_g = -0.6$, $u_d = -0.2$



(b) $u_g = -0.2$, $u_d = -0.6$



(c) $u_g = 0.5$, $u_d = 1$



(d) $u_g = 1$, $u_d = -1.5$

Figure 1.15: Convex fluxes. 1000 pts. $t = 0.2$

Test 3. A particular configuration where only two discontinuous solutions are admissible.
The flux functions f_L and f_R are defined from the derivatives f_L' and f_R' given by

$$f_L'(u) = (u+1)(u+\frac{1}{10})(u-1),$$

$$f_R'(u) = -(u+\frac{1}{2})(u-\frac{2}{5})(u-\frac{3}{2}).$$

The sonic points are thus $u_L^- = -1$, $u_L^0 = -1/10$ and $u_L^+ = 1$ for the left flux f_L and $u_R^- = -1/2$, $u_R^0 = 2/5$ and $u_R^+ = 3/2$ for the right flux f_R. The Riemann initial data is such that $u_g = -1.25$ and $u_d = 1.75$. This configuration is such that $u_g \notin \widetilde{\mathcal{F}}_R^+(u_d)$, $u_d \notin \widetilde{\mathcal{F}}_L^-(u_g)$. Moreover the set $\widetilde{\mathcal{F}}_L^-(u_g) \cap \widetilde{\mathcal{F}}_R^+(u_d)$ is empty and therefore there is no continuous solution. Fig. 1.16 represents both fluxes and in bold-face the sets $\widetilde{\mathcal{F}}_R^+(u_d)$ and $\widetilde{\mathcal{F}}_L^-(u_g)$.

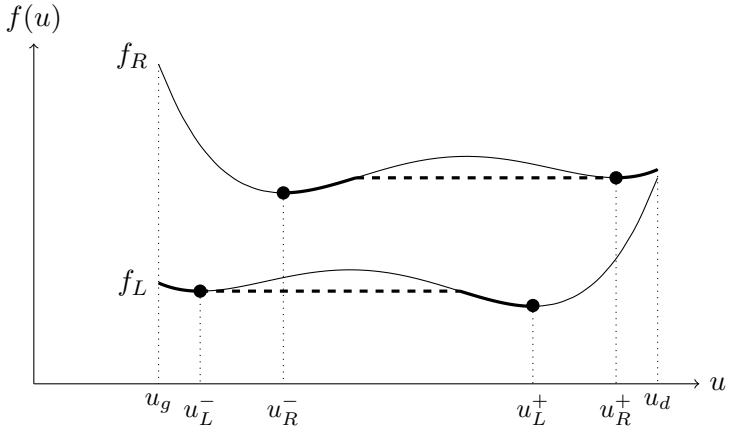


Figure 1.16: A case where $u_g \notin \widetilde{\mathcal{F}}_R^+(u_d)$, $u_d \notin \widetilde{\mathcal{F}}_L^-(u_g)$, and $\widetilde{\mathcal{F}}_R^+(u_d) \cap \widetilde{\mathcal{F}}_L^-(u_g) = \emptyset$.

The proposed coupled Riemann problem admits only discontinuous solutions, each one of the third type. More precisely, they are

(a) a L-wave (whose speed is nonpositive) connecting u_g to u_L^+, followed by a R-wave (whose speed is nonnegative) connecting u_R^+ to u_d,

(b) a L-wave (whose speed is nonpositive) connecting u_g to u_L^-, followed by a R-wave (whose speed is nonnegative) connecting u_R^- to u_d.

Numerical solutions are presented on Fig. 1.17. We observe that the Godunov scheme captures the solution (b), while the relaxation scheme captures the solution (a).

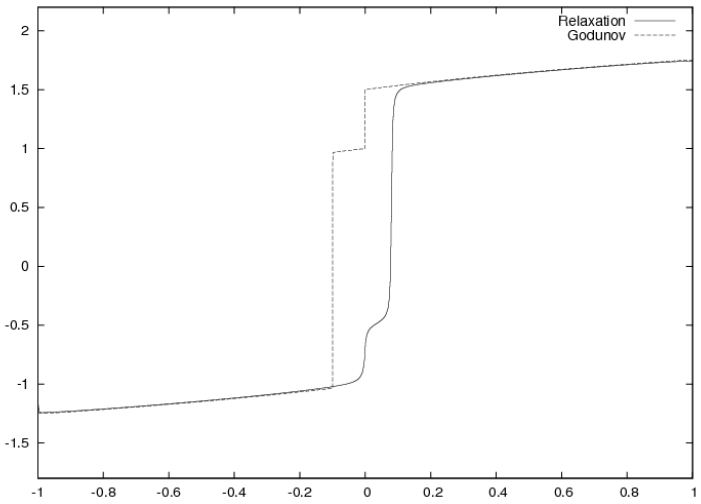


Figure 1.17: Multiple disc. solutions. 10000 pts. $u_L = -1.25$, $u_R = 1.75$, $t = 1.5$

Test 4. A particular configuration where several discontinuous solutions are admissible, and continuous solutions also exist.
Here again, both fluxes are obtained from the derivatives given by

$$f_L'(u) = (u+1)(u+\frac{1}{2})(u-1),$$

$$f_R'(u) = -(u-\frac{5}{4})(u-\frac{3}{4})(u+\frac{3}{4}).$$

We take $u_g = -1.5$ and $u_d = 1.75$ so that $u_g \notin \widetilde{\mathcal{F}}_R^+(u_d)$, $u_d \notin \widetilde{\mathcal{F}}_L^-(u_g)$, but the set $\widetilde{\mathcal{F}}_L^-(u_g) \cap \widetilde{\mathcal{F}}_R^+(u_d)$ is not empty. Fig. 1.18 represents both fluxes and in bold-face the sets $\widetilde{\mathcal{F}}_R^+(u_d)$ and $\widetilde{\mathcal{F}}_L^-(u_g)$. We can see that both numerical schemes capture the same continuous solution.

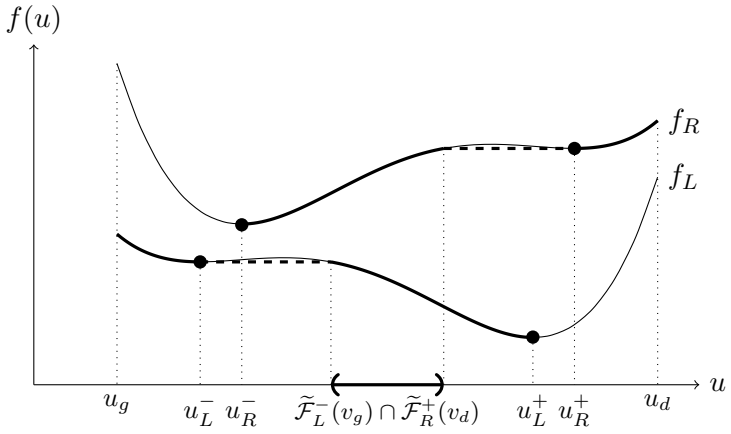


Figure 1.18: A case where $u_g \notin \widetilde{\mathcal{F}}_R^+(u_d)$, $u_d \notin \widetilde{\mathcal{F}}_L^-(u_g)$, and $\widetilde{\mathcal{F}}_R^+(u_d) \cap \widetilde{\mathcal{F}}_L^-(u_g) \neq \emptyset$.

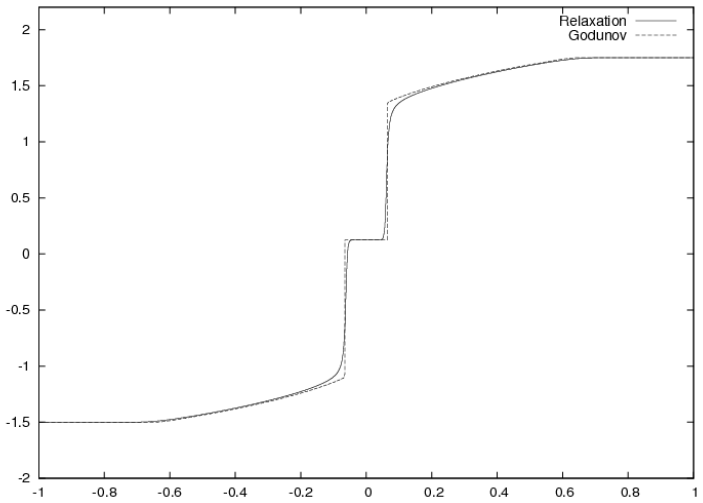


Figure 1.19: Multiple disc. solutions. 10000 pts. $t = 0.5$. $u_L = -1.5$, $u_R = 1.75$.

Test 5. The coupling of two conservation laws "with phase change"
The rest of this section is devoted to the coupling of two scalar conservation laws "with phase change". We consider the following fluxes

$$f_L(u) = \begin{cases} -\dfrac{u^2-1}{2}, & u < -1, \\ 0, & -1 \le u \le 1, \\ \dfrac{u^2-1}{2}, & u > 1, \end{cases} \qquad f_R(u) = f_L(u-1/2) - 1,$$

that are represented on Fig. 1.20.

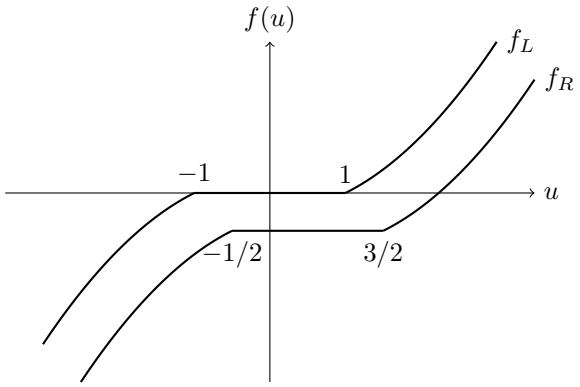


Figure 1.20: Scalar flux for the coupling "with phase change".

The numerical solutions are shown on Fig.1.21 for both relaxation and Godunov schemes and these match with the expected theoretical results. They correspond to the solution for the right-problem, that consists here in a shock transition wave connecting the left-state of the Riemann problem to the point of C^1-discontinuity ($u = 3/2$), a constant part and finally a rarefaction wave to the right-state.

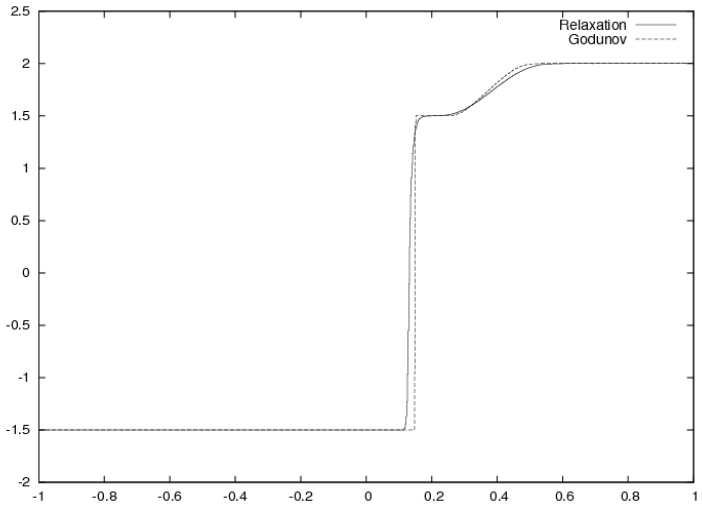


Figure 1.21: Phase change. 1000 pts. $u_L = -1.5$, $u_R = 2.0$, $t = 0.3$.

2 Self-similar viscous approximations for thin interfaces

Nous proposons une approche pour le couplage d'équations hyperboliques non-linéaires, basée sur une formulation par un système EDP augmenté. Un effet de résonance est naturellement présent dans la formulation augmentée proposée, due à la possible perte d'hyperbolicité globale. De façon importante, la non-unicité des solutions faibles est observée pour le problème avec une donnée initiale. Nous explorons des approximations visqueuses auto-semblables pour le problème de Riemann et établissons un théorème d'existence sous des hypothèses structurelles assez générales. La clef de voûte de ce chapitre réside dans l'estimation des interactions non-linéaires pour les solutions du système hyperbolique augmenté, et notamment des interactions concernant les ondes résonnantes.

Le travail présenté dans ce chapitre a été réalisé en collaboration avec Frédéric Coquel et Philippe G. LeFloch et fait l'objet de la publication en préparation [32]. La Section 2.5 est une annexe absente de l'article.

Coupling nonlinear hyperbolic equations (I). Self-similar viscous approximations for thin interfaces

Abstract

We propose a framework based on a formulation of the coupling of nonlinear hyperbolic equations via an augmented system of PDE's. A resonance effect naturally takes place in the proposed augmented formulation due to the possible failure of global hyperbolicity. Importantly, non-uniqueness of weak solutions is observed for the initial value problem. We investigate self-similar vanishing viscosity approximations for the Riemann problem and we establish an existence theorem under fairly general structural assumptions. The key contribution of this paper consists of the nonlinear interaction estimates for the solutions of the augmented hyperbolic system, including wave interactions involving resonant waves.

2.1 Introduction

2.1.1 Self-similar regularizations of resonant systems

We are interested in the following class of hyperbolic systems of $N+1$ partial differential equations

$$\begin{aligned} A_0(u,v)\,\partial_t u + A_1(u,v)\,\partial_x u &= 0,\\ \partial_t v &= 0, \end{aligned} \tag{2.1}$$

where $u = u(t,x) \in \mathcal{U} \subset \mathbb{R}^N$, $v = v(t,x) \in \mathbb{R}$ with $x \in \mathbb{R}$ and $t \geq 0$ are the main unknowns. Here, the smooth mappings $A_0, A_1 : \mathcal{U} \times \mathbb{R} \to \mathbb{R}^{N\times N}$ are matrix-valued and A_0 is invertible while and A_1 is invertible except at some point $(u^\star, v^\star) \in \mathcal{U} \times \mathbb{R}$, that is

$$\begin{aligned} &\det A_0(u,v) \neq 0,\\ &\det A_1(u,v) \neq 0, \quad (u,v) \neq (u^\star, v^\star),\\ &\det A_1(u^\star, v^\star) = 0. \end{aligned}$$

Note that the first equation is formally equivalent to the nonconservative system with variable coefficients

$$\partial_t u + A_0^{-1}(u,v)A_1(u,v)\partial_x u = 0.$$

Moreover, we assume $A_0^{-1}(u,v)A_1(u,v)$ to be $\mathbb{R}-$diagonalisable with all eigenvalues $\lambda_i(u,v)$ differents (hyperbolicity assumption). However, the noninvertibiliy of $A_1(u^\star, v^\star)$ implies there exists a (unique) index $m \in \{1,\dots,N\}$ such that

$$\lambda_m(u^\star, v^\star) = 0. \tag{2.2}$$

As a consequence, the system (2.1) is "weakly hyperbolic", only. Our objective, precisely, is to study this system in this resonant context.

Dafermos [57–59] advocated the use of self-similar regularizations to capture the whole structure of wave fans within the solution of a Riemann problem. It consists in searching solutions asymptotic in large time, that would be obtained as self-similar ones i.e. solutions that depends only on $\xi = x/t$. In the variable (x,t), it can be done be introducing in (2.1) small viscosity terms depending on a multiplicative coefficient t and of $B_0(u,v)$ a viscosity matrix.

$$\begin{array}{c} A_0(u^\epsilon,v^\epsilon)\partial_t u^\epsilon + A_1(u^\epsilon,v^\epsilon)\partial_x u^\epsilon = \epsilon\ t\partial_x\Big(B_0(u^\epsilon,v^\epsilon)\partial_x u^\epsilon\Big) \\ \partial_t v^\epsilon = \epsilon^p t\,\partial_{xx} v^\epsilon \end{array} \qquad x\in\mathbb{R}, t>0.$$

The factor ϵ represents a small term e^{-T} for a large characteritic time T. In the variable ξ, equations satisfied by a self-similar viscous solutions $(u^\epsilon(\xi), v^\epsilon(\xi))$ write then

$$\begin{array}{c} \Big(-\xi A_0(u^\epsilon,v^\epsilon) + A_1(u^\epsilon,v^\epsilon)\Big)u^\epsilon_\xi = \epsilon\ \Big(B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi\Big)_\xi \\ -\xi v^\epsilon_\xi = \epsilon^p v^\epsilon_{\xi\xi} \end{array} \qquad \xi\in\mathbb{R}, \tag{2.3}$$

while the coupling formula (2.10) writes

$$u^\epsilon(0^-) = u^\epsilon(0^+).$$

An outline of this paper is as follows. Section 2.2 presents the case of a scalar unknown u. We first get the existence of a solution to the viscous self-similar Riemann problem (2.3). Then for vanishing viscosity (ϵ tends to 0), we obtain a self-similar solution to (2.1) that would be moreover an entropy solution in each half-space. This global result of existence does not need any smallness assumption on the data or the both coupled models.
Sections 2.3 and 2.4 treat the more general case of N-dimensional systems ($u\in\mathbb{R}^N$). Under smallness assumption on the Riemann data and proximity of both coupled models, we obtain the existence of the solution. Section 2.3 presents the main estimates that justify the construction made in Section 2.4.

2.1.2 Motivations

This study is in line with the coupling problematic where one interests in two strictly hyperbolic half-problems respectively written on $\mathbb{R}^+_t\times\mathbb{R}^-_x$ and $\mathbb{R}^+_t\times\mathbb{R}^+_x$

$$\begin{array}{ll} \partial_t w + \partial_x f_-(w) = 0, & x<0, t>0, \\ \partial_t w + \partial_x f_+(w) = 0, & x>0, t>0. \end{array} \tag{2.4}$$

Besides a supplemented initial data, some coupling condition must be prescribed at the *fixed interface* $\{x=0\}$ for all positive time t. Thus the problem takes naturally the form of boundary problem for each half-problem, and the question is to formulate these boundary conditions in order to make the global problem well-posed. Following Godlewski and

Raviart [73, 76] one wishes that

$$\theta_-(w(0^-,t)) = \theta_+(w(0^+,t)), \quad t>0, \tag{2.5}$$

where θ_-, θ_+ are two invertible functions of $\mathbb{R}^N$, with respective inverse functions named $\gamma_- = \theta_-^{-1}$ and $\gamma_+ = \theta_+^{-1}$. These functions are introduced in order to treat different couplings. For example, with $\theta_- = \theta_+ = \mathrm{Id}$ one recover the continuity of w at the interface, while with $\theta_\pm = f_\pm$ the flux should be continuous at the interface so that it becomes conservative.

In this coupling condition (2.5), equality is expected in a strong sense, as long as the interface is not characteristic for the left half-problem neither for the right one. Otherwise, following the theory of boundary problems for hyperbolic systems, developped by Dubois and LeFloch [65] and studied for the coupling problematic by Godlewski and Raviart [73, 76] and Ambroso et al. [7], this condition can only be achieved in a weak sense and writes then

$$\begin{aligned} w(0^+,t) &\in \mathcal{O}_+\Big(\theta_+ \circ \theta_-^{-1}(w(0^-,t))\Big), \\ w(0^-,t) &\in \mathcal{O}_-\Big(\theta_- \circ \theta_+^{-1}(w(0^+,t))\Big). \end{aligned} \tag{2.6}$$

where $\mathcal{O}_+(b)$ (resp. $\mathcal{O}_-(b)$) is the set of admissible trace at $\xi = 0^+$ (resp. $\xi = 0^-$) for the self-similar Riemann problem for flux f_+ (resp. f_-) with prescribed trace b at 0^- (resp. 0^+)

$$\begin{aligned} \mathcal{O}_+(b) &= \Big\{\mathcal{W}_+(0^+,b,u), u\in\Omega\Big\}, \\ \mathcal{O}_-(b) &= \Big\{\mathcal{W}_-(0^-,u,b), u\in\Omega\Big\}, \end{aligned}$$

$\mathcal{W}_+(\cdot,b,u)$ being the self-similar solution of the following Cauchy problem

$$\partial_t w + \partial_x f_+(w) = 0, \quad x\in\mathbb{R}, t>0, \qquad w(x,0) = \begin{cases} b, & x<0, \\ u, & x>0, \end{cases}$$

and similarly $\mathcal{W}_-(\cdot,u,b)$ being the self-similar solution of the following Cauchy problem

$$\partial_t w + \partial_x f_-(w) = 0, \quad x\in\mathbb{R}, t>0, \qquad w(x,0) = \begin{cases} u, & x<0, \\ b, & x>0. \end{cases}$$

While the interface is not characteristic, the conditions (2.6) give (2.5) back, otherwise (2.5) is not achieved in a strong sense anymore. The question is thus to understand what occurs, particularly in term of existence and of unicity of solutions satisying (2.6). Under this form, the question of the nonlinear coupling of two boundary problems seems to be complex. We propose to extend the problem by suppressing the interface and defining a new problem on the whole space $\mathbb{R}_t^+ \times \mathbb{R}_x$. The first step is in the change of variable

$$u_- = \theta_-(w), \quad u_+ = \theta_+(w), \tag{2.7}$$

that permits to rewrite both half-problems as

$$\partial_t \gamma_\pm(u_\pm) + \partial_x f_\pm(\gamma_\pm(u)) = 0, \quad \pm x > 0, \quad t > 0, \tag{2.8}$$

with coupling conditions

$$\begin{aligned} u(0^+,t) &\in O_+(u(0^-,t)), \\ u(0^-,t) &\in O_-(u(0^+,t)), \end{aligned} \tag{2.9}$$

where $O_+(b)$ (and similarly $O_-(b)$) is the following set of admissible trace at $\xi = 0^+$

$$O_+(b) = \Big\{ W_+(0^+, b, \mathrm{u}), \mathrm{u} \in \mathcal{U} \Big\},$$

and $W_+(\cdot, b, u)$ is the self-similar solution of the following Cauchy problem

$$\partial_t \gamma_+(u) + \partial_x f_+(\gamma_+(u)) = 0, \quad x \in \mathbb{R}, t > 0, \qquad u(x,0) = \begin{cases} b, & x < 0, \\ \mathrm{u}, & x > 0. \end{cases}$$

Outside the resonance phenomenon, this reformulation gives simple continuity of u at interface

$$u(0^-,t) = u(0^+,t). \tag{2.10}$$

We propose then to study the solutions of (2.8)-(2.9) through the solutions of the following Cauchy problem

$$\begin{aligned} A_0(u,v)\partial_t u + A_1(u,v)\partial_x u &= 0, \\ \partial_t v &= 0, \end{aligned} \quad x \in \mathbb{R}, t > 0. \tag{2.11}$$

where the "color function" v takes his value in $[-1,1]$, the value $v = -1$ corresponding to the left half-problem and the value $v = 1$ to the right half-problem, in the sense that one requires following consistance properties on A_0 and A_1 together with the coupling problem (2.4)-(2.5)

$$\begin{aligned} A_0(u,\pm 1) &= \nabla \gamma_\pm(u), \\ A_1(u,\pm 1) &= \nabla f_\pm\Big(\gamma_\pm(u)\Big) \nabla \gamma_\pm(u). \end{aligned} \tag{2.12}$$

The definition of $A_0(u,v)$ (respectively $A_1(u,v)$) for $v \in]-1,1[$ should connect by a smooth way $A_0(u,-1)$ to $A_0(u,1)$ (resp. $A_1(u,-1)$ to $A_1(u,1)$). Moreover A_0 is invertible and $A_0^{-1}A_1$ is everywhere $\mathbb{R}$-diagonalisable what property extends naturaly strict hyperbolicity of original hyperbolic half-problems. The system (2.11) is then completed with initial data

$$\begin{aligned} u_0(x,0) = u_0(x) &= \theta_\pm(w_0(x)), \quad \pm x > 0 \\ v_0(x,0) = v_0(x) &= \pm 1, \quad \pm x > 0 \end{aligned} \tag{2.13}$$

Formally, the interface of coupling is said to be characteristic at a point $(u^\star, v^\star) \in \mathcal{U} \times [-1,1]$ as soon as the matrix $A_1(u^\star, v^\star)$ admits the eigenvalue zero. Indeed, as long as $A_1(u,v)$ is an invertible matrix, the u vector gives the N Riemann invariants corresponding to the eigenvalue 0 of the $(N+1)\times(N+1)$ system (2.11). For the intial data v_0, across the discontinuity of $v(x,t)$ localised at $\{x = 0, t > 0\}$, the solution (u,v) satisfies constancy of

the Riemann invariants of the linearly degenerate field associate with eigenvalue 0, what means simply (2.10) is satisfied. Nevertheless, as $A_1(u,v)$ is not locally invertible, (2.10) is not satisfied anymore a priori.

2.2 Existence theory for scalar conservation laws

2.2.1 Riemann problem with diffusion

The problem under consideration here is nonconservative in nature, although it reduces to a conservative system when the component v (see below) equals ± 1. As explained in Section 2.1, we seek a self-similar function u which is the limit of solutions u^ϵ to the problem

$$\begin{aligned} \Big(-\xi A_0(u^\epsilon,v^\epsilon)+A_1(u^\epsilon,v^\epsilon)\Big)u^\epsilon_\xi &= \epsilon\left(B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi\right)_\xi,\\ -\xi v^\epsilon_\xi &= \epsilon^p\, v^\epsilon_{\xi\xi}, \end{aligned} \tag{2.14}$$

supplemented with Riemann initial data

$$\begin{aligned} &u^\epsilon(-\infty)=u_L, \quad u^\epsilon(+\infty)=u_R,\\ &v^\epsilon(-\infty)=-1, \quad v^\epsilon(+\infty)=1. \end{aligned} \tag{2.15}$$

In (2.14), the maps A_0 and A_1 are smooth real functions, that satisfy the following consistency condition with the underlying coupling problem, in sense that there exists some constants c_1,c_2,c_3, such that

$$\begin{aligned} &A_0(u,-1)=\gamma_-'(u), && A_0(u,1)=\gamma_+'(u),\\ &A_1(u,-1)=(f_-\circ\gamma_-)'(u), && A_1(u,1)=(f_+\circ\gamma_+)'(u),\\ &0<c_1\le A_0(u,v), && 0<c_2\le B_0(u,v)\le c_3. \end{aligned} \tag{2.16}$$

Moreover, using the notation $\mathcal{U}=[\min(u_L,u_R),\max(u_L,u_R)]$, we introduce the Lipschitz constants ω_0, ω_1 of A_0 and A_1, respectively, i.e.

$$\left|A_j(\tilde u,\tilde v)-A_j(u,v)\right|\le\omega_j\left(|\tilde u-u|+|\tilde v-v|\right)$$

for all $(\tilde u,\tilde v),(u,v)\in\mathcal{U}\times[-1,1]$ and $j=0,1$.

Observe that the first equation in (2.14) can be equivalently rewritten as

$$\Big(-\xi+\lambda(u^\epsilon,v^\epsilon)\Big)G(u^\epsilon,v^\epsilon)\,B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi=\epsilon\,\left(B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi\right)_\xi, \tag{2.17}$$

where we set

$$\lambda(u,v):=\frac{A_1(u,v)}{A_0(u,v)},\quad G(u,v):=\frac{A_0(u,v)}{B_0(u,v)}.$$

Furthermore, our assumptions (2.16) imply that we can find $M,\Lambda>0$ such that

$$|\lambda(u,v)|\le\Lambda<M,\qquad (u,v)\in\mathcal{U}\times[-1,1].$$

This uniform bound is connected with the property of finite speed of propagation for the limiting equation. We will handle the problem under consideration on the bounded interval $[-M,M]$ with the boundary conditions in (2.15) imposed at the end points $\pm M$ and M, and next let M tend to infinity.

The solution will be sought in the function space

$$\mathcal{H} = \mathcal{H}([-M,M]) := \Big\{u \in C([-M,M]); u(\xi) \in \mathcal{U}\Big\}.$$

Proposition 2.1 (Existence for Riemann problem with diffusion)**.** *For each $\epsilon > 0$ the problem (2.14) admits smooth solutions $u^\epsilon \in \mathcal{H}([-M,M])$ and $v^\epsilon \in C(\mathbb{R})$ given by the (implicit) representation formulas*

$$\begin{aligned} v^\epsilon(\xi) &= -1 + 2\frac{\displaystyle\int_{-\infty}^{\xi} e^{-\zeta^2/2\epsilon^p}\,d\zeta}{\displaystyle\int_{-\infty}^{+\infty} e^{-\zeta^2/2\epsilon^p}\,d\zeta}, \\ u^\epsilon(\xi) &= u_L + (u_R - u_L)\frac{\displaystyle\int_{-M}^{\xi} e^{-h^\epsilon(u^\epsilon;\zeta)/\epsilon}\, B_0^{-1}(u^\epsilon,v^\epsilon)\,d\zeta}{\displaystyle\int_{-M}^{M} e^{-h^\epsilon(u^\epsilon;\zeta)/\epsilon}\, B_0^{-1}(u^\epsilon,v^\epsilon)\,d\zeta}, \end{aligned} \tag{2.18}$$

with

$$h^\epsilon(u^\epsilon;\xi) := \int_\alpha^\xi \Big(\omega - \lambda\,(u^\epsilon,v^\epsilon)\Big)\, G(u^\epsilon,v^\epsilon)\,d\omega.$$

Moreover, u^ϵ and v^ϵ are monotone bounded continuous functions with uniformly bounded total variation

$$TV(u^\epsilon) \le |u_R - u_L|, \qquad TV(v^\epsilon) \le 2.$$

Proof. Solving the second equation in (2.14) is immediate. On the other hand, we can rewrite the problem (2.14) as

$$\begin{aligned} B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi &= \varphi, \\ \Big(-\xi + \lambda(u^\epsilon,v^\epsilon)\Big)\, G(u^\epsilon,v^\epsilon)\,\varphi &= \epsilon\varphi_\xi. \end{aligned}$$

Given $\tilde{u} \in \mathcal{H}$ we consider the solution $u^\epsilon(\tilde{u};\xi)$ of the "linearized" problem

$$\begin{aligned} B_0(\tilde{u},v^\epsilon)u^\epsilon_\xi &= \varphi, \\ \Big(-\xi + \lambda(\tilde{u},v^\epsilon)\Big)\, G(\tilde{u},v^\epsilon)\,\varphi &= \epsilon\varphi_\xi, \end{aligned}$$

together with the boundary conditions

$$u^\epsilon(-M) = u_L, \qquad u^\epsilon(M) = u_R.$$

The solution is explicitely given by

$$\begin{aligned} u^\epsilon(\tilde{u};\xi) &= u_L + (u_R - u_L)\frac{\displaystyle\int_{-M}^{\xi} \varphi(\tilde{u};\zeta)B_0^{-1}(\tilde{u},v^\epsilon)\,d\zeta}{\displaystyle\int_{-M}^{M} \varphi(\tilde{u};\zeta)B_0^{-1}(\tilde{u},v^\epsilon)\,d\zeta},\\ \varphi(\tilde{u};\xi) &= \exp\Big(-h^\epsilon(\tilde{u};\xi)/\epsilon\Big),\\ h^\epsilon(\tilde{u};\xi) &= \int_\alpha^\xi \Big(\omega - \lambda(\tilde{u},v^\epsilon)\Big)G(\tilde{u},v^\epsilon)\,d\omega, \end{aligned} \tag{2.19}$$

in which $\alpha \in [-M,M]$ is arbitrary. The above formulas determine a map T^ϵ that takes $\tilde{u} \in \mathcal{H}$ to the function $u^\epsilon(\tilde{u};\cdot) \in \mathcal{H}$. We need to find a fixed-point of T^ϵ.

The uniform bounds on $\lambda(\tilde{u},v)$ and $G(\tilde{u},v)$ (for all $\tilde{u} \in \mathcal{H}$ and $v \in [-1,1]$) allow us to choose $\alpha_\epsilon \in [-M,M]$ so that

$$h^\epsilon(\tilde{u};\xi) \geq 0, \quad \xi \in [-M,M]; \qquad h^\epsilon(\tilde{u};\alpha_\epsilon) = 0.$$

Consequently, for all $\xi \in [-M,M]$ and for some constant c_4 we have

$$0 \leq h^\epsilon(\tilde{u};\xi) \leq c_4,$$

so that

$$c_3^{-1}\exp(-c_4/\epsilon) \leq \varphi(\tilde{u},\xi)B_0^{-1}(\tilde{u},v^\epsilon) \leq c_2^{-1}.$$

We also obtain the uniform bound

$$\begin{aligned} \left|\frac{d}{d\xi}u_\epsilon(\tilde{u};\xi)\right| &\leq |u_R - u_L|\frac{\varphi(\tilde{u};\xi)B_0^{-1}(\tilde{u},v^\epsilon)}{\displaystyle\int_{-M}^{M}\varphi(\tilde{u};\zeta)B_0^{-1}(\tilde{u},v^\epsilon)\,d\zeta}\\ &\leq |u_R - u_L|\frac{c_2^{-1}}{2Mc_3^{-1}\exp(-c_4/\epsilon)} \leq |u_R - u_L|\frac{c_3\,\exp(c_4/\epsilon)}{2Mc_2}. \end{aligned}$$

The latter bound above being independent of $\tilde{u}$, we deduce that the family T_ϵ is equicontinuous and its image is relatively compact in $\mathcal{H}$. Since this image is a convex closed subset of the Banach space $C([-M,M])$, Schauder fixed-point theorem applies and ensures that T^ϵ admits a fixed point. Hence, there exists $u^\epsilon \in \mathcal{H}$ such that $T^\epsilon(u^\epsilon) = u^\epsilon$, and the representation formula (2.18) holds. The uniform total variation bounds follow directly from (2.18). □

2.2.2 Passage to the limit

We now establish two lemmas.

Lemma 2.2 (Existence of a pointwise limit)**.** *After extracting a subseqeuence if necessary, the sequence u^ϵ converges to a limiting function $u \in BV([-M,M])$ such that*

$$u^\epsilon(\xi) \to u(\xi), \qquad \xi \in [-M,M],$$

which satisfies, in the sense of distributions,

$$\begin{aligned} -\xi\,\frac{d}{d\xi}\gamma_-(u)+\frac{d}{d\xi}f_-(\gamma_-(u))=0, \qquad \xi<0,\\ -\xi\,\frac{d}{d\xi}\gamma_+(u)+\frac{d}{d\xi}f_+(\gamma_+(u))=0, \qquad \xi>0. \end{aligned} \tag{2.20}$$

Lemma 2.3 (Entropy inequalities)**.** *The limit u given by the above lemma also satisfies, in the sense of distributions,*

$$\begin{aligned} -\xi\,\frac{d}{d\xi}\eta(\gamma_-(u))+\frac{d}{d\xi}q_-(\gamma_-(u))\le 0, \qquad \xi<0,\\ -\xi\,\frac{d}{d\xi}\eta(\gamma_+(u))+\frac{d}{d\xi}q_+(\gamma_+(u))\le 0, \qquad \xi>0, \end{aligned} \tag{2.21}$$

for all convex entropy functions η and associated entropy flux $q'_\pm=\eta' f'_\pm$.

Proof of Lemma 2.2. In view of Lemma 2.1, Helly's compactness theorem applies and ensures the existence of a pointwise limit $(u,v)\in BV_{\rm loc}$ as ϵ tends to 0. Let $\phi\in C_0^\infty((\theta,M))$ be compactly supported test function, where $\theta>0$ is given. In the integral form, (2.14) reads

$$-\int_0^M \xi\,A_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi\,\phi+\int_0^M A_1(u^\epsilon,v^\epsilon)u^\epsilon_\xi\,\phi=\epsilon\int_0^M\Big(B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi\Big)_\xi\phi,$$

that is

$$-\int_0^M \xi\,A_0(u^\epsilon,1)u^\epsilon_\xi\,\phi+\int_0^M A_1(u^\epsilon,1)u^\epsilon_\xi\,\phi+\Omega=\epsilon\int_0^M\Big(B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi\Big)_\xi\phi,$$

where

$$\Omega:=\int_0^M\xi\Big(A_0(u^\epsilon,1)-A_0(u^\epsilon,v^\epsilon)\Big)u^\epsilon_\xi\,\phi-\int_0^M\Big(A_1(u^\epsilon,1)-A_1(u^\epsilon,v^\epsilon)\Big)u^\epsilon_\xi\,\phi.$$

Using (2.16) we can write

$$-\int_0^M\xi\,\frac{d}{d\xi}\gamma_+(u^\epsilon)\,\phi+\int_0^M\frac{d}{d\xi}f_+(\gamma_+(u^\epsilon))\,\phi+\Omega=\epsilon\int_0^M\Big(B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi\Big)_\xi\phi,$$

The quantity Ω vanishes with ϵ, since

$$\begin{aligned} |\Omega|&\le M\int_\theta^M\omega_0|1-v^\epsilon|\,|u^\epsilon_\xi|\,|\phi|\;+\int_\theta^M\omega_1|1-v^\epsilon|\,|u^\epsilon_\xi|\,|\phi|\\ &\le(M\,\omega_0+\omega_1)\,|1-v^\epsilon(\theta)|\,\|\phi\|_\infty TV(u^\epsilon), \end{aligned}$$

where the total variation term $TV(u^\epsilon)$ remains bounded and $|1-v^\epsilon(\theta)|$ tends to 0. On the other hand, we have

$$\left|\epsilon\int_0^M\Big(B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi\Big)_\xi\phi\right|=\left|\epsilon\int_0^M\Big(B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi\Big)\,\phi_\xi\right|\le\epsilon\,\|\phi_\xi\|_\infty\,c_3TV(u^\epsilon)$$

which also converges to zero. Letting ϵ tend to 0 we conclude that

$$\int_0^M \left(-\xi \frac{d}{d\xi}\gamma_+(u) + \frac{d}{d\xi} f_+(\gamma_+(u))\right)\phi = 0.$$

The same arguments apply on the interval $[-M,0]$, by using test-functions supported in the interval $(-M,\theta)$, with $\theta < 0$. Hence, we obtain

$$\int_{-M}^0 \left(-\xi \frac{d}{d\xi}\gamma_-(u) + \frac{d}{d\xi} f_-(\gamma_-(u))\right)\phi = 0,$$

and the desired identity (2.20) follows. □

Proof of Lemma 2.3. Let $\phi \in C_0^\infty((\theta,M))$ be a non-negative test-function, with $\theta > 0$. Multiplying (2.14) by $\eta'(C(u,1))\phi$, we get

$$\begin{aligned}
&-\int_0^M \xi A_0(u^\epsilon,v^\epsilon)u_\xi^\epsilon \eta'(C(u^\epsilon,1))\,\phi + \int_0^M A_1(u^\epsilon,v^\epsilon)u_\xi^\epsilon \eta'(C(u^\epsilon,1))\,\phi \\
&= \epsilon \int_0^M \left(B_0(u^\epsilon,v^\epsilon)u_\xi^\epsilon\right)_\xi \eta'(C(u^\epsilon,1))\,\phi
\end{aligned}$$

Observing that

$$\begin{aligned}
&\left|\int_0^M \xi\,(A_0(u^\epsilon,v^\epsilon) - A_0(u^\epsilon,1))\,u_\xi^\epsilon \eta'(C(u^\epsilon,1))\,\phi\right| \\
&\le M\,\omega_0|1-v^\epsilon(\theta)|\;TV(u^\epsilon)\;\|\eta'(C(\cdot,1))\|_\infty\|\phi\|_\infty,
\end{aligned}$$

and similarly for the coefficient A_1 we see that the left-hand side in the former identity is equivalent (modulo terms that tend to zero with ϵ) to

$$\begin{aligned}
&-\int_0^M \xi\partial_u(C(u^\epsilon,1))\frac{d}{d\xi}u^\epsilon \eta'(C(u^\epsilon,1))\,\phi + \int_0^M A_1(u^\epsilon,1)\frac{d}{d\xi}u^\epsilon \eta'(C(u^\epsilon,1))\,\phi \\
&= -\int_0^L \xi \frac{d}{d\xi}\eta(\gamma_+(u^\epsilon))\,\phi + \int_0^M \frac{d}{d\xi} q_+(\gamma_+(u^\epsilon))\,\phi.
\end{aligned}$$

The right-hand side can be rewritten in the form

$$\begin{aligned}
&\epsilon \int_0^M \left(B_0(u^\epsilon,v^\epsilon)u_\xi^\epsilon\right)_\xi \eta'(C(u^\epsilon,1))\,\phi \\
&= -\epsilon \int_0^M B_0(u^\epsilon,v^\epsilon)\left(u_\xi^\epsilon\right)^2 A_0(u^\epsilon,1)\eta''(C(u^\epsilon,1))\,\phi \\
&\quad -\epsilon \int_0^M B_0(u^\epsilon,v^\epsilon)u_\xi^\epsilon \eta'(C(u^\epsilon,1))\,\phi_\xi,
\end{aligned}$$

in which the first term is non-positive and the second one tends to 0. Thus, letting $\epsilon \to 0$ we obtain

$$-\int_0^M \xi \frac{d}{d\xi}\eta(\gamma_+(u))\,\phi + \int_0^M \frac{d}{d\xi} q_+(\gamma_+(u))\,\phi \le 0.$$

The arguments in the half-space $\xi < 0$ are similar. □

2.2.3 Riemann problem for the hyperbolic system

In view of the boundary condition (2.15) we require that

$$u(\xi) = \begin{cases} u_L, & \xi \le -M, \\ u_R, & \xi \ge M. \end{cases} \tag{2.22}$$

The conclusions in Lemmas 2.2 and 2.3 actually hold in the intervals $(-\infty, -M)$ and on $(M, +\infty)$. In addition, an interface condition for the solution u at the end points $\xi = -M$ and $\xi = M$ is now provided, and is necessary to ensure that both (2.20) and (2.21) extend in the sense of distributions on $(0, +\infty)$ and on $(-\infty, 0)$ by the continuation from (2.22).

Lemma 2.4. *The solutions u^ϵ converge uniformly towards u_R (repectively u_L) on the interval $(\Lambda, M]$ (resp. $(-M, -\Lambda[$).*

Proof. Let $\xi \in \left(\frac{\Lambda+M}{2}, M\right)$ be given. According to (2.18) we have

$$\begin{aligned} |u^\epsilon(\xi) - u_R| &= |u_L - u_R| \frac{\displaystyle\int_\xi^M e^{-h^\epsilon(u^\epsilon)/\epsilon} B_0^{-1}(u^\epsilon, v^\epsilon)}{\displaystyle\int_{-M}^M e^{-h^\epsilon(u^\epsilon)/\epsilon} B_0^{-1}(u^\epsilon, v^\epsilon)} \\ &\le |u_L - u_R| \frac{\displaystyle\int_{\frac{\Lambda+M}{2}}^M e^{-h^\epsilon(u^\epsilon)/\epsilon} B_0^{-1}(u^\epsilon, v^\epsilon)}{\displaystyle\int_{-M}^M e^{-h^\epsilon(u^\epsilon)/\epsilon} B_0^{-1}(u^\epsilon, v^\epsilon)}. \end{aligned} \tag{2.23}$$

We use here the constant $\alpha \in [-\Lambda, \Lambda]$ as the lower integration bound for both h^ϵ and

$$h := \xi \mapsto \int_\alpha^\xi \Big(\omega - \lambda(u(\omega), v(\omega))\Big) G(u(\omega), v(\omega))\, d\omega.$$

Hence, $h(\xi) \ge 0$. Moreover, $h^\epsilon(u^\epsilon, \cdot)$ converges uniformly to h, with

$$\begin{aligned} &\left|h^\epsilon(u^\epsilon, \xi) - h(\xi)\right| \\ &= \left| \int_\alpha^\xi \Big(\lambda(u^\epsilon(\omega), v^\epsilon(\omega)) - \lambda(u(\omega), v(\omega))\Big) G(u^\epsilon(\omega), v^\epsilon(\omega))\, d\omega \right| \\ &\le \left(\frac{1}{c}\omega_1 + \|A_1\|_\infty \frac{\omega_0}{c^2}\right) \|G\|_\infty \int_{-M}^M \Big(|u^\epsilon(\omega) - u(\omega)| + |v^\epsilon(\omega) - v(\omega)|\Big) d\omega \\ &\le \left(\frac{1}{c}\omega_1 + \|A_1\|_\infty \frac{\omega_0}{c^2}\right) \|G\|_\infty \Big(\|u^\epsilon - u\|_1 + \|v^\epsilon - v\|_1\Big) \to 0 \end{aligned}$$

The uniform convergence of $h^\epsilon(u^\epsilon, \cdot)$ towards a positive continuous function h such that $h(\alpha) = 0$ insures that there exists ϵ_0 sufficiently small, together with $A > B > 0$, $\eta > 0$ such that for all $\epsilon < \epsilon_0$

$$\begin{aligned} h^\epsilon(u^\epsilon, \xi) &\ge A, \qquad \frac{\Lambda + M}{2} \le \xi \le M, \\ h^\epsilon(u^\epsilon, \xi) &\le B, \qquad \xi \in [-M, M],\ |\xi - \alpha| < \eta. \end{aligned}$$

Thus, (2.23) implies

$$|u^\epsilon(\xi) - u_R| \leq |u_R - u_L| \frac{\frac{M-\Lambda}{2} e^{-A/\epsilon} c_2^{-1}}{\eta e^{-B/\epsilon} c_3^{-1}} = |u_R - u_L| \frac{M-\Lambda}{2\eta} \frac{c_3}{c_2} e^{-(A-B)/\epsilon},$$

and u^ϵ converges uniformly towards u_R on the open interval $(\frac{\Lambda+M}{2}, M)$. The same argument leads to the uniform convergence of u^ϵ towards u_L on $(-M, \frac{\Lambda-M}{2})$. □

We summarize our results in Lemmas 2.2, 2.3 and 2.4, as follows.

Theorem 2.5 (The Riemann problem or the coupling of two scalar equations)**.** *Up to extracting a subsequence, the solutions u^ϵ converge pointwise to a function $u \in BV(\mathbb{R})$,*

$$u^\epsilon(\xi) \to u(\xi), \qquad \xi \in \mathbb{R},$$

which satisfies the conservation laws and entropy inequalities

$$\begin{aligned} -\xi \frac{d}{d\xi} \gamma_-(u) + \frac{d}{d\xi} f_-(\gamma_-(u)) &= 0, \qquad \xi < 0, \\ -\xi \frac{d}{d\xi} \gamma_+(u) + \frac{d}{d\xi} f_+(\gamma_+(u)) &= 0, \qquad \xi > 0, \end{aligned}$$

$$\begin{aligned} -\xi \frac{d}{d\xi} \eta(\gamma_-(u)) + \frac{d}{d\xi} q_-(\gamma_-(u)) &\leq 0, \qquad \xi < 0, \\ -\xi \frac{d}{d\xi} \eta(\gamma_+(u)) + \frac{d}{d\xi} q_+(\gamma_+(u)) &\leq 0, \qquad \xi > 0, \end{aligned}$$

for all convex entropy pairs, together with the boundary conditions

$$u(-\infty) = u_L, \qquad u(+\infty) = u_R.$$

Equivalently, in terms of the function w in (2.7), we have established

$$\begin{aligned} -\xi \frac{d}{d\xi} w + \frac{d}{d\xi} f_-(w) &= 0, \qquad -\xi \frac{d}{d\xi} \eta(w) + \frac{d}{d\xi} q_-(w) \leq 0, \qquad \xi < 0, \\ -\xi \frac{d}{d\xi} w + \frac{d}{d\xi} f_+(w) &= 0, \qquad -\xi \frac{d}{d\xi} \eta(w) + \frac{d}{d\xi} q_+(w) \leq 0, \qquad \xi > 0, \end{aligned}$$

with

$$w(-\infty) = w_L, \qquad w(+\infty) = w_R.$$

Later in Section 5, we will also investigate the behavior of the solution at the interface $\xi = 0$.

2.3 Existence theory for systems

2.3.1 Terminology and notation

This section follows Tzavaras [132] and Joseph and LeFloch [88], and adapts their technique to cover the class of nonconservative and resonant systems under consideration:

$$\begin{aligned} -\xi A_0(u^\epsilon, v^\epsilon) u^\epsilon_\xi + A_1(u^\epsilon, v^\epsilon) u^\epsilon_\xi &= \epsilon \Big(B_0(u^\epsilon, v^\epsilon) u^\epsilon_\xi \Big)_\xi, \\ -\xi v^\epsilon_\xi &= \epsilon^p v^\epsilon_{\xi\xi}. \end{aligned} \tag{2.24}$$

We refer to the problem (2.24) as the *diffusive Riemann problem*. We will establish that, provided $u_L, u_R \in \mathcal{B}(\delta_1)$ with δ_1 sufficiently smaller than δ_0 and under some structural hypothesis on A_0 and A_1, this system admits a smooth, self-similar solution $u_\epsilon = u_\epsilon(x/t) \in \mathcal{B}(\delta_0)$, which has uniformly bounded total variation

$$TV(u_\epsilon) + TV(v_\epsilon) \leq C, \quad C > 0. \tag{2.25}$$

Solutions to (2.24) will be given under the form of an asymptotic expansion in terms of the "wave strengths" of the associated Riemann problem (2.1). The uniform estimate (2.25) will then be the key to the convergence analysis ($\epsilon \to 0$) and will allow us to establish the existence of the Riemann solution and to derive his properties.

This general strategy was first proposed for $B_0 = I$ and $A_0 = I$ and a general existence result was proven by Tzavaras [132] (for the conservative systems) and by LeFloch and Tzavaras [112] (for the nonconservative systems (2.1)). It was then extended by Joseph and LeFloch in a series of papers [88–91]. For many other results on self-similar limits (viscosity-capillarity terms, large data results, mixed-type model, etc) see Slemrod [66], Slemrod and Tzavaras [125], Fan and Slemrod [66], and LeFloch and Rohde [106].

For each $u \in \mathcal{U}$ and $v \in [-1,1]$, we denote by

$$\lambda_1(u,v) < \ldots < \lambda_N(u,v)$$

the real and distinct eigenvalues of the $N \times N$ matrix $A(u,v) = A_1(u,v)A_0^{-1}(u,v)$ and by $l_1(u,v), \ldots, l_N(u,v)$ and $r_1(u,v), \ldots, r_N(u,v)$ corresponding basis of left- and right-eigenvectors respectively. These vectors are normalized so that $l_i(u,v) \cdot r_j(u,v) = 0$ if $i \neq j$ and $l_i(u,v) \cdot r_i(u,v) = 1$. Remember we are interested in solutions u taking values in a small neighborhood of a given state in $u^\star \in \mathbb{R}^N$ and, without loss of generality, we can choose $u^\star = 0$ so that this set is the ball $\mathcal{U} := \mathcal{B}(\delta_0)$ centered at the origin with (small) radius δ_0.
Thus we may assume (by reducing δ_0 if necessary) that the wave speeds $\lambda_i(u,v)$ are sufficiently close to the constants $\lambda_i(0,0)$ and, in particular, are uniformly separated for all $u \in \mathcal{B}(\delta_0)$ in the sense that, for some constants

$$-M < \underline{\Lambda}_1 < \overline{\Lambda}_1 < \underline{\Lambda}_2 < \ldots < \underline{\Lambda}_N < \overline{\Lambda}_N < M,$$

we have

$$\underline{\Lambda}_i := \lambda_i(0,0) - O(\delta_0), \quad \overline{\Lambda}_i := \lambda_i(0,0) + O(\delta_0) \tag{2.26}$$

and

$$\underline{\Lambda}_i \le \lambda_i(u,v) \le \overline{\Lambda}_i, \quad u \in \mathcal{B}(\delta_0), v \in [-1,1]. \tag{2.27}$$

Let m be the index that corresponds to the resonant wave, i.e. such that there exists some $v^\star \in [-1,1]$ such that $\lambda_m(u^\star, v^\star) = 0$, also with $\underline{\Lambda}_m < 0 < \overline{\Lambda}_m$. In addition, for δ_0 sufficiently small the vectors $r_i(u,v)$ are sufficiently close to $r_i(0,0)$ thus we can assume that

$$\begin{aligned} &l_i(u_1,v)\cdot r_i(u_2,v) \ge 1-\delta_0, \quad u_1,u_2 \in \mathcal{B}(\delta_0), v \in [-1,1], \\ &|l_i(u_1,v)\cdot r_j(u_2,v)| \le \delta_0, \quad u_1,u_2 \in \mathcal{B}(\delta_0), i \ne j, v \in [-1,1]. \end{aligned} \tag{2.28}$$

In (2.24), the matrix $B_0 = B_0(u,v)$ is assumed to be non-degenerate and to depend smoothly upon u and v. We will treat the case that the diffusion matrix $B(u,v) = B_0(u,v)A_0^{-1}(u,v)$ is sufficient close to the identity matrix, that is, for some given matrix norm and $\eta > 0$ sufficiently small

$$\sup_{\substack{u\in\mathcal{B}(\delta_0)\\ v\in[-1,1]}} |B(u,v) - I| \le \eta.$$

To handle this arbitrary diffusion matrix $B(u,v)$, the following *generalized eigenvalue problem* is relevant

$$\begin{aligned} &\big(-\xi\,\mathrm{Id} + A(u,v)\big)\widehat{r_i}(u,v,\xi) = \mu_i(u,v,\xi)\,B(u,v)\widehat{r_i}(u,v,\xi), \\ &\widehat{l_i}(u,v,\xi)\cdot\big(-\xi\,\mathrm{Id} + A(u,v)\big) = \mu_i(u,v,\xi)\widehat{l_i}(u,v,\xi)\cdot B(u,v). \end{aligned} \tag{2.29}$$

We impose the following normalization to generalized left- and right-eigenvectors:

$$\begin{aligned} &\widehat{r_i}(u,v,\xi)\cdot\widehat{r_i}(u,v,\xi) = 1, \\ &\widehat{l_i}(u,v,\xi)\cdot B(u,v)\widehat{r_j}(u,v,\xi) = 0 \quad \text{if } i \ne j, \\ &\widehat{l_i}(u,v,\xi)\cdot B(u,v)\widehat{r_i}(u,v,\xi) = 1. \end{aligned}$$

Multiplying the first equation in (2.29) on the left by $\widehat{r_i}(u,v,\xi)$ and rearranging terms, we get

$$\mu_i(u,v,\xi) = \big(-\xi + \widehat{\lambda}_i(u,v,\xi)\big)d_i(u,v,\xi), \tag{2.30}$$

where

$$\begin{aligned} &\widehat{\lambda}_i(u,v,\xi) := \widehat{r_i}(u,v,\xi)\cdot A(u,v)\widehat{r_i}(u,v,\xi), \\ &1/d_i(u,v,\xi) := \widehat{r_i}(u,v,\xi)\cdot B(u,v)\widehat{r_i}(u,v,\xi). \end{aligned} \tag{2.31}$$

Clearly, in the special case where $B(u,v) = I$, we find

$$\mu_i(u,v,\xi) = -\xi + \lambda_i(u,v), \quad \widehat{r_i}(u,v,\xi) = r_i(u,v), \quad \widehat{l_i}(u,v,\xi) = l_i(u,v).$$

So, by continuity, when B gets closer to the identity matrix, the coefficients $d_i(u,v,\xi)$ and $\widehat{\lambda}_i(u,v,\xi)$ get closer to 1 and $\lambda_i(u,v)$ respectively. In consequence, under the assumption

$|B(u,v)-I|<\eta$ with η sufficiently small and by increasing the gaps $\overline{\Lambda}_i-\underline{\Lambda}_i$ if necessary, we can always assume that

$$\begin{aligned}\underline{\Lambda}_i - O(\eta) \le \widehat{\lambda}_i(u,v,\xi) \le \overline{\Lambda}_i + O(\eta),\\ 1 - O(\eta) \le d_i(u,v,\xi) \le 1 + O(\eta),\end{aligned}$$

for $u\in\mathcal{B}(\delta_0)$, $v\in[-1,1]$, $\xi\in[-M,M]$. Furthermore, we find following estimates:

Lemma 2.6. *The ξ-derivatives of the generalized eigenvectors and eigenvalues satisfy*

$$\begin{aligned}|\partial_\xi \widehat{r}_i(u,v,\xi)| &= O(\eta),\\ \partial_\xi \mu_i(u,v,\xi) &= -1 + O(\eta).\end{aligned} \tag{2.32}$$

Proof. Omitting for convenience the explicit dependence in variables (u,v,ξ), and differentiating the first equation of (2.29) with respect to the variable ξ, it appears that

$$(-\xi I + A)\partial_\xi \widehat{r}_i - \mu_i B \partial_\xi \widehat{r}_i = \widehat{r}_i + \partial_\xi \mu_i B \widehat{r}_i.$$

Multiplying previous equation on the left by $\widehat{l}_i$, we get

$$\begin{cases}(\mu_i-\mu_j)\widehat{l}_i\cdot B\partial_\xi\widehat{r}_j = \widehat{l}_i\cdot\widehat{r}_j = O(\eta), & j\neq i,\\ \partial_\xi\mu_i = -\widehat{l}_i\cdot\widehat{r}_i = -1+O(\eta), & j=i.\end{cases}$$

By decomposing the vector $\partial_\xi\widehat{r}_i=\sum_j\alpha_{ij}\widehat{r}_j$, we find

$$\alpha_{ij} = \widehat{l}_j\cdot B\partial_\xi\widehat{r}_i = O(\eta), \quad i\neq j,$$

thus

$$\partial_\xi\widehat{r}_i = O(\eta) + \alpha_{ii}\widehat{r}_i,$$

then using the ξ-derivative of the normalization's equality of $\widehat{r}_i$, we finally get

$$0 = \widehat{r}_i\cdot\partial_\xi\widehat{r}_i = O(\eta)+\alpha_{ii},$$

and (2.32) follows. □

In the sequel, another quantity will be usefull to describe the smallness assumptions on the modeling data:

Remark 2.1. The quantity

$$\nu := \left\|\widehat{l}_i\cdot\partial_v(B\widehat{r}_j)\right\|_{L^\infty} \tag{2.33}$$

is representative of the closeness of both left- and right-models in consideration and will assumed to be sufficiently small in the sequel.

The following example, borrowed from gaz dynamics, illustrates how the closeness of models can be explicitely related to this quantity ν.

Example 2.1 (p-system). Let consider the coupling of two 2×2 systems with two different

pressure laws $\tau \mapsto p_\pm(\tau)$

$$\partial_t \tau - \partial_x u = 0,$$
$$\partial_t u + \partial_x p_\pm(\tau) = 0.$$

The matrix of the corresponding fluxes have jacobian

$$\nabla F_\pm = \begin{pmatrix} 0 & -1 \\ p'_\pm(\tau) & 0 \end{pmatrix}.$$

We consider now the case that the coupling is based on the choice $B = \mathrm{Id}$ and

$$A(\tau, v) = \frac{1+v}{2}\nabla F_+ + \frac{1-v}{2}\nabla F_- = \begin{pmatrix} 0 & -1 \\ \frac{1+v}{2}p'_+(\tau) + \frac{1-v}{2}p'_-(\tau) & 0 \end{pmatrix}.$$

Hence, the eigenvalues, left- and right-eigenvectors of the system equal respectively

$$\lambda_\pm(\tau, v) = \pm\sqrt{-\frac{1+v}{2}p'_+(\tau) - \frac{1-v}{2}p'_-(\tau)},$$
$$\widehat{r}_\pm(\tau, v) = \begin{pmatrix} 1 & \mp\sqrt{-\frac{1+v}{2}p'_+(\tau) - \frac{1-v}{2}p'_-(\tau)} \end{pmatrix}^T,$$
$$\widehat{l}_\pm(\tau, v) = \begin{pmatrix} \pm\sqrt{-\frac{1+v}{2}p'_+(\tau) - \frac{1-v}{2}p'_-(\tau)} & 1 \end{pmatrix}.$$

A short calculation gives

$$\left|\widehat{l_i} \cdot \partial_v(B\widehat{r_j})\right| \leq \left|\frac{p'_+(\tau) - p'_-(\tau)}{2\sqrt{-\frac{1+v}{2}p'_+(\tau) - \frac{1-v}{2}p'_-(\tau)}}\right| \leq \left|\frac{p'_+(\tau) - p'_-(\tau)}{2\min(\sqrt{-p'_+(\tau)}, \sqrt{-p'_-(\tau)})}\right|.$$

Finally we obtain

$$\nu \leq C\|p'_+ - p'_-\|_\infty.$$

2.3.2 Equations for the characteristic coefficients

We complete (2.24) with following boundary conditions, inherited from Riemann initial data

$$\begin{aligned} u^\epsilon(-M) &= u_L, \quad u^\epsilon(M) = u_R, \\ v^\epsilon(-M) &= -1, \quad v^\epsilon(M) = 1. \end{aligned} \tag{2.34}$$

We describe here the asymptotic expansion for the solution of the diffusive Riemann problem (2.24)-(2.34). To handle arbitrary diffusion matrix, the decomposition must be modified and, instead, be based on the eigenvectors defined earlier in (2.29). We first solve explicitly the equation concerning v^ϵ:

$$\psi^\epsilon(\xi) := v^\epsilon_\xi(\xi) = 2\frac{e^{-\frac{\xi^2}{2\epsilon^p}}}{\int_{-M}^{M} e^{-\frac{x^2}{2\epsilon^p}}\,dx}. \tag{2.35}$$

Note that in the case $p = +\infty$, we formally get the Dirac solution $\psi = v_\xi = 2\delta_{\xi=0}$. In other words the transition from $v_L = -1$ to $v_R = 1$ at the interface is then discontinous.

Equation (2.24) rewrites

$$\Big(-\xi I + A(u^\epsilon, v^\epsilon)\Big)A_0(u^\epsilon, v^\epsilon)u^\epsilon_\xi = \epsilon\Big(B(u^\epsilon, v^\epsilon)A_0(u^\epsilon, v^\epsilon)u^\epsilon_\xi\Big)_\xi. \tag{2.36}$$

The resolution is based on the decomposition of the gradient $A_0(u^\epsilon, v^\epsilon)u^\epsilon_\xi(\xi)$ on the basis of eigenvectors $\widehat{r}_j(u^\epsilon(\xi), v^\epsilon(\xi), \xi)$, that is, we set

$$\begin{aligned} a^\epsilon_j(\xi) &:= \widehat{l}_j(u^\epsilon, v^\epsilon, \xi)B(u^\epsilon, v^\epsilon)A_0(u^\epsilon, v^\epsilon)u^\epsilon_\xi(\xi), \\ A_0(u^\epsilon, v^\epsilon)\,u^\epsilon_\xi(\xi) &= \sum_{j=1}^{N} a^\epsilon_j(\xi)\widehat{r}_j(u^\epsilon, v^\epsilon, \xi), \end{aligned} \tag{2.37}$$

where the a^ϵ_j's are referred to as the *characteristic coefficients* of $A_0(u^\epsilon, v^\epsilon)u^\epsilon_\xi$. Dropping (until further notice) the explicit dependence in ϵ, the right-hand side of (2.36) equals

$$\begin{aligned} \Big(B(u,v)\,A_0(u,v)u_\xi\Big)_\xi = &\sum_j a'_k B(u,v)\widehat{r}_j(u,v,\cdot) \\ &+ \sum_{j,k} a_j a_k D_u\Big(B\widehat{r}_j\Big)(u,v,\cdot)\,A_0^{-1}(u,v)\widehat{r}_k(u,v,\cdot) \\ &+ \sum_j a_j \partial_v\Big(B\widehat{r}_j\Big)(u,v,\cdot)\,v_\xi + \sum_j a_j B(u,v)\,\partial_\xi\widehat{r}_j(u,v,\cdot). \end{aligned}$$

For any solution of (2.24), we have therefore

$$\begin{aligned} &\sum_j \Big(\epsilon a'_j B(u,v)\widehat{r}_j(u,v,\cdot) - a_j\Big(-\xi + A(u,v)\Big)\widehat{r}_j(u,v,\cdot)\Big) \\ &= -\epsilon\sum_{j,k} a_j a_k D_u\Big(B\widehat{r}_j\Big)(u,v,\cdot)\,A_0^{-1}(u,v)\widehat{r}_k(u,v,\cdot) - \epsilon\sum_j a_j\partial_v\Big(B\widehat{r}_j\Big)(u,v,\cdot)\,\psi \\ &\quad - \epsilon\sum_j a_j B(u,v)\,\partial_\xi\widehat{r}_j(u,v,\cdot). \end{aligned}$$

Multiplying the above by each $\widehat{l}_i(u,v,\cdot)$ for $i = 1,\dots,N$ and relying on equation (2.29) we find a *coupled system of N differential equations* that the characteristic coefficients a_i should satisfy:

$$a'_i - \frac{\mu_i(u,v,\cdot)}{\epsilon}a_i = \eta L_i(u,v,\cdot) + Q_i(u,v,\cdot) + S_i(u,v,\cdot), \tag{2.38a}$$

where linear, quadratic and source terms are respectively

$$\begin{aligned} L_i(u,v,\cdot) &:= \sum_j \pi_{ij}(u,v,\cdot)a_j, \\ Q_i(u,v,\cdot) &:= \sum_{j,k} \kappa_{ijk}(u,v,\cdot)a_j a_k, \\ S_i(u,v,\cdot) &:= \sum_j \sigma_{ij}(u,v,\cdot)a_j\psi, \end{aligned} \tag{2.38b}$$

and their coefficients are precisely defined by

$$\begin{aligned}
\pi_{ij}(u,v,\cdot) &:= -\eta^{-1}\left(\widehat{l_i}\cdot B\,\partial_\xi \widehat{r_j}\right)(u,v,\cdot),\\
\kappa_{ijk}(u,v,\cdot) &:= -\left(\widehat{l_i}\cdot D_u(B\widehat{r_j})A_0^{-1}\widehat{r_k}\right)(u,v,\cdot),\\
\sigma_{ij}(u,v,\cdot) &:= \left(\widehat{l_i}\cdot \partial_v(B\widehat{r_j})\right)(u,v,\cdot).
\end{aligned} \tag{2.38c}$$

Equation (2.24) is then equivalent to (2.37)-(2.38). Observe that by (2.32) and (2.33), (2.38) has the form

$$a_i' - \frac{1}{\epsilon}\mu_i(u,v,\cdot)a_i = O(\eta)\sum_j |a_j| + O(1)\sum_{j,k}|a_j||a_k| + O(\nu)\sum_j |a_j||\psi|. \tag{2.39}$$

Let first observe the principal part of the equation (2.38), that is, given some function $u = u(y)$ (which at this stage need not to be a solution to (2.38)) we consider the following *decoupled system of linear equations* for $i = 1,\dots,N$:

$$\varphi_i^{\star\prime} - \frac{\mu_i(u,v,\cdot)}{\epsilon}\varphi_i^\star = 0. \tag{2.40}$$

Their solutions are multiples of *fundamental normalized solutions* associated with the function u

$$\varphi_i^\star := \frac{e^{-g_i/\epsilon}}{\displaystyle\int_{-M}^{M} e^{-g_i/\epsilon}\,dy} \qquad \text{with } g_i(y) := -\int_{\rho_i}^{y}\mu_i(u,v,\cdot)(x)\,dx. \tag{2.41}$$

Here, the ρ_i's are constants chosen so that the functions g_i are non-negative (what will be explained in Section 2.3.3). Clearly, the functions $\varphi_i^\star$ are strictly positive and chosen so that their total integral equals 1. We will search the complete solutions a_i of (2.38a) under the form

$$a_i = \tau_i\varphi_i^\star + \theta_i \tag{2.42}$$

where τ_i refers to a wave strengh and with θ_i small w.r.t τ_i. When solving the equation (2.38a) for a given right-hand side and considering again first-order terms, we are naturally led to consider the following *linear wave interaction coefficients* $J_{j\to i}$, the *quadratic wave interaction coefficients* $F_{jk\to i}$ and the *resonant quadratic interaction coefficients* $J^\psi_{j\to i}$ respectively defined by

$$J_{j\to i}(y) := \varphi_i^\star(y)\int_{c_i}^{y}\frac{\varphi_j^\star(x)}{\varphi_i^\star(x)}\,dx, \tag{2.43}$$

$$F_{j,k\to i}(y) := \varphi_i^\star(y)\int_{c_i}^{y}\frac{\varphi_j^\star\,\varphi_k^\star}{\varphi_i^\star}\,dx, \tag{2.44}$$

$$J^\psi_{j\to i}(y) := \varphi_i^\star(\xi)\int_{c_i}^{\xi}\psi(x)\frac{\varphi_j^\star(x)}{\varphi_i^\star(x)}\,dx \tag{2.45}$$

for some constants $c_i \in [\underline{\Lambda}_i, \overline{\Lambda}_i]$ independent of ϵ. By studying these coefficients, we will gain useful information on the possible growth of the total variation of solutions to (2.38):

roughly speaking, $J_{j\to i}$ bounds the influence of the j-th family on the i-th family, $F_{j,k\to i}$ bounds the contribution on the i-th family due to interactions between waves of the j-th and k-th characteristic families and $J^{\psi}_{j\to i}$ bounds the influence of the j-th family on the i-th family "through" the coupling wave ψ. In order to estimate these all coefficients, we first study useful properties of linearized waves measures.

2.3.3 Linearized wave measures

The results of this section are already given by Joseph and LeFloch [88] and presented here for convenience in the lecture.

We give us now some speed range $[\lambda_{\min}, \lambda_{max}]$ and we present here canonical properties of objects constructed through formula (2.41) for this speed range. We introduce first a space of "almost linear at infinity" functions. Observe the μ_i will belong to such spaces for their respective speed ranges $[\underline{\Lambda}_i, \overline{\Lambda}_i]$.

Definition 2.1 ($\mathcal{L}$ spaces)**.** We call a function $h : [-M,M] \mapsto \mathbb{R}$ to be an element of $\mathcal{L}$ if there exists two functions $d, \lambda \in L^1([-M,M],\mathbb{R})$, and two positive reals $d_{\min}, d_{\max}$ such as

$$h(x) = d(x)\Big(\lambda(x) - x\Big), \quad x \in [-M,M]$$

with

$$0 < d_{\min} \le d(x) \le d_{\max},$$
$$-M < \lambda_{\min} \le \lambda(x) \le \lambda_{\max} < M.$$

Lemma 2.7. *Let $h : [-M,M] \mapsto \mathbb{R}$ an element of $\mathcal{L}$. Being given $y \in [-M,M]$, we set*

$$g(x) = -\int_y^x h(x')\,dx', \qquad x \in [-M,M],$$

then g is Lipschitz continuous and achieves its global minimum at some (non-unique) point $\rho \in [\lambda_{\min}, \lambda_{\max}]$ (obviously independent of y) that realizes $\lambda(\rho) = \rho$.

Proof. Using definition 2.1 of $\mathcal{L}$ space we get

$$h(x) > 0, \qquad x < \lambda_{\min},$$
$$h(x) < 0, \qquad x > \lambda_{\max},$$

g being continuous, decreasing on $[-M, \lambda_{\min}]$ and increasing on $[\lambda_{\max}, M]$ (because $g' = -h$), we deduce g achieves its global minimum at some (non-unique) point $\rho \in [\lambda_{\min}, \lambda_{\max}]$. Moreover $h(\rho) = -g' = 0$, that means $\lambda(\rho) = \rho$. □

We now introduce new notations suggested by formula (2.41) to get useful algebraic prop-

erties. For $x, y \in [-M, M]$ and $h \in L^1([-M, M], \mathbb{R})$:

$$\phi(y, x; h) := \exp\left(\frac{1}{\epsilon}\int_y^x h(x')\,dx'\right),$$
$$I(y; h) := \int_{-M}^{M} \phi(y, x'; h)\,dx',$$
$$\varphi(x; h) := \frac{\phi(y, x; h)}{I(y; h)}.$$

That is, $x \mapsto \varphi(x; h)$ is the wave measure, of unit total mass, solution of the differential equation

$$\varphi' - \frac{h}{\epsilon}\varphi = 0.$$

Lemma 2.8 (Algebraic properties of the mapping ϕ)**.** *For general functions $h, \tilde{h}$ in $L^1([-M, M], \mathbb{R})$ and $x, y, z \in [-M, M]$ we have the following three algebraic properties:*

$$\frac{\phi(x, y; h)}{\phi(x, y; \tilde{h})} = \phi(x, y; h - \tilde{h}),$$
$$\frac{\phi(x, y; h)}{\phi(z, y; h)} = \frac{\phi(y, z; h)}{\phi(y, x; h)} = \phi(x, z; h),$$
$$\phi(x, y; h) = \phi(y, x; -h) = \frac{1}{\phi(y, x; h)}.$$

Note that, thanks to these properties, the previous definition of $\varphi(x; h)$ is really independent of the variable y what justifies the notation. Moreover a judicious choice of y could be useful in the following, typically, when the function h belong to $\mathcal{L}$, by taking $y = \rho$ given by Lemma 2.7, we obtain in the exponential defining ϕ a negative argument that vanishes only at points where g is minimised.

Now, fix $\mu_1, \mu_2 \in \mathcal{L}$ and $\varphi_1 := \varphi(\cdot; \mu_1)$ and $\varphi_2 := \varphi(\cdot; \mu_2)$ the solutions of differential equation with respectively $h = \mu_1$ and $h = \mu_2$, we note ρ_1 and ρ_2 the minimisation points of respective associated functions "g" defined in Lemma 2.7. Moreover we give us some unspecified point $c \in [\lambda_{\min}, \lambda_{\max}]$ independent of ϵ. We then want to control the linear interaction coefficient

$$J_{\varphi_2 \to \varphi_1}(y) = \varphi_1(y) \int_c^y \frac{\varphi_2(x)}{\varphi_1(x)}\,dx \tag{2.46a}$$

characterising the first order (linear) influence of φ_2 on φ_1. Using Lemma 2.8 we find those two following equivalent expressions, both useful in the sequel according to the sign of $\mu_2 - \mu_1$,

$$J_{\varphi_2 \to \varphi_1}(y) = \frac{I(\rho_1; \mu_1)}{I(\rho_2; \mu_2)} \phi(\rho_2, \rho_1; \mu_2)\,\varphi_1(y) \int_c^y \phi(\rho_1, x; \mu_2 - \mu_1)\,dx,$$
$$J_{\varphi_2 \to \varphi_1}(y) = \varphi_2(y) \int_c^y \phi(x, y; \mu_1 - \mu_2)\,dx.$$

In order to estimate those coefficients, we also need for more information on the asymptotical behaviour of appearing quantities as ϵ tends to 0. The two following lemmas will give it.

Lemma 2.9 (Asymptotic behavior of ϕ)**.** *Let $[x,y]$ be an interval of $[-M,M]$, (with $x<y$), and let h be a continuous function on $[x,y]$. If h is strictly positive bounded, say we have $h(x')\geq h_{\min}>0$ on $[x,y]$, then the following integral is at most linear in ϵ*

$$\int_x^y \phi(x',y,-h)\,dx' \leq \frac{\epsilon}{h_{\min}}, \tag{2.47}$$

Lemma 2.10 (Asymptotic behavior of $I(\rho;h)$)**.** *For a function $h\in\mathcal{L}$, and ρ defined through Lemma 2.7, the integrals $I(\rho;h)$ checks*

$$c\epsilon \leq I(\rho;h) \leq 2M. \tag{2.48}$$

Proof of Lemma 2.9. Let suppose $h(x')\geq h_{\min}>0$ then, for all $x<x'<y$, we have

$$-\int_{x'}^y h(t)\,dt \leq -h_{\min}(y-x'),$$

thus

$$\int_x^y \phi(x',y;h)\,dx' \leq \int_x^y e^{-h_{\min}(y-x')/\epsilon}\,dx' \leq \frac{\epsilon}{h_{\min}}\int_0^\infty e^{-x'}\,dx' \leq \frac{\epsilon}{h_{\min}},$$

and the result follows. □

Proof of Lemma 2.10. Because of the definition of ρ, the argument of the exponential defining ϕ is everywhere nonpositive so $\phi\leq 1$ and $I(\rho;h)\leq 2M$. Moreover at the point ρ, the primitive of h is locally Lipschitz continuous, so there exists a sufficiently small η and a constant $c>0$ such that

$$0\leq -\int_\rho^x h(x')\,dx' \leq \frac{1}{c}|x-\rho|, \qquad |x-\rho|<\eta.$$

Then

$$I(\rho;h)\geq \int_{-\eta}^{\eta} e^{-\frac{1}{c\epsilon}|x|}\,dx = 2c\epsilon\int_0^{\frac{\eta}{c\epsilon}} e^{-x'}\,dx' \geq c\epsilon.$$

□

Note that in Lemma 2.9, if $h_{\min}=0$, then the considered integral remains bounded as ϵ vanishes.

The following result Lemma 2.11 shows that, on any compact subset of the complement set $[\lambda_{\min},\lambda_{\max}]^c$, the mass of the linearized wave measures tends to zero. In the limit, all the mass of the wave measure $\varphi(\cdot;h)$ is concentrated on the interval $[\lambda_{\min},\lambda_{\max}]$.

Lemma 2.11 (Behavior of linearized wave measures)**.** *For all $h\in\mathcal{L}$ the function $\varphi(\cdot;h)$ satisfy the estimates*

$$0\leq\varphi(x;h)\leq O(1/\epsilon)\begin{cases} e^{-(x-\lambda_{\min})^2 d_{\min}/2\epsilon}, & -M<x<\lambda_{\min},\\ 1, & x\in[\lambda_{\min},\lambda_{\max}],\\ e^{-(x-\lambda_{\max})^2 d_{\min}/2\epsilon}, & \lambda_{\max}<x<M. \end{cases} \tag{2.49}$$

Proof. For $x \geq \lambda_{\max}$ we have, with ρ defined through Lemma 2.7, and d and λ given through Definition 2.1 of $\mathcal{L}$ space

$$\begin{aligned} -\int_{\rho}^{x} h(y)\,dy &= \int_{\lambda_{\max}}^{x} d(y)\Big(y-\lambda(y)\Big)dy + \int_{\rho_i}^{\lambda_{\max}} d(y)\Big(y-\lambda(y)\Big)dy \\ &\geq \int_{\lambda_{\max}}^{x} d(y)\Big(y-\lambda_{\max}\Big)dy \geq \frac{(x-\lambda_{\max})^2 d_{\min}}{2}, \end{aligned}$$

while a similar argument for $x < \lambda_{\min}$ gives

$$-\int_{\rho}^{x} h(y)\,dy \geq \int_{\lambda_{\min}}^{x} d(y)\Big(y-\lambda_{\min}\Big)dy \geq \frac{(x-\lambda_{\min})^2 d_{\min}}{2}.$$

Finally, the desired conclusion follows from definition of $\varphi(\cdot;h)$ and from Lemma 2.10. □

2.3.4 Interaction coefficients

We present now some revisited results about linear and quadratic interaction coefficients, already treated by Joseph and LeFloch [88], in Lemma 2.12. A new result that concerns resonant wave interaction coefficients will complete this analysis in Lemma 2.14.

Lemma 2.12 (Interaction coefficients estimate)**.** *The linear interaction coefficient $J_{j\to i}$ defined in (2.43) satisfy the following estimate for all $i, j = 1,\dots,N$ and $y \in [-M,M]$:*

$$\left|J_{j\to i}(y)\right| \leq \begin{cases} O(\epsilon)\,(\varphi_i^\star(y)+\varphi_j^\star(y)), & i \neq j, \\ 2M\varphi_i^\star(y), & i = j. \end{cases} \tag{2.51}$$

Moreover, the quadratic wave interaction coefficients defined in (2.44) satisfy

$$|F_{j,k\to i}(y)| \leq C\Big(\varphi_i^\star(y)+\varphi_j^\star(y)+\varphi_k^\star(y)\Big). \tag{2.52}$$

Lemma 2.13 (Localization of interaction)**.** *By choosing δ small enough, for all $i \neq j$ there exists positive constants C and D independent of ϵ such that*

$$\left\|\frac{\varphi_i^\star}{\varphi_j^\star}\right\|_{L^\infty([\underline{\Lambda}_j,\overline{\Lambda}_j])} \leq Ce^{-D/\epsilon}. \tag{2.53}$$

Lemma 2.14 (Resonant quadratic interaction coefficient)**.** *Given $\psi \in L^1$ and $i \neq j$, then one has*

$$\left|J^{\psi}_{j\to i}(y)\right| \leq O(1)\,\|\psi\|_1(\varphi_j^\star(y)+\varphi_i^\star(y)), \qquad \epsilon > 0.$$

Proof of Lemma 2.12. The case $i = j$ is obvious, so we only need to consider the case $i \neq j$. For definiteness we suppose that $j > i$, the proof for $j < i$ being similar, except the permutation of cases.

First, using Lemmas 2.9 and 2.10, we get in the region $y > c_i$:

$$\left|\varphi_i^\star(y)\int_{c_i}^{y}\frac{\varphi_j^\star(x)}{\varphi_i^\star(x)}\,dx\right| = \varphi_j^\star(y)\int_{c_i}^{y}\phi(x,y;\mu_i-\mu_j)\,dx \le \frac{\epsilon}{\underline{\Lambda}_j-\overline{\Lambda}_i}\varphi_j^\star(y).$$

On the other hand, in the region $y < c_i$ we have

$$\left|\varphi_i^\star(y)\int_{c_i}^{y}\frac{\varphi_j^\star(x)}{\varphi_i^\star(x)}\,dx\right| = \varphi_i^\star(y)\frac{I_i}{I_j}\phi(\rho_j,\rho_i;\mu_j)\int_{y}^{c_i}\phi(\rho_i,x,\mu_j-\mu_i)dx$$
$$= \varphi_i^\star(y)\frac{I_i}{I_j}\phi(\rho_j,\rho_i;\mu_j)\phi(\rho_i,c_i;\mu_j-\mu_i)\int_{y}^{c_i}\phi(c_i,x,\mu_j-\mu_i)dx.$$

But, by Lemma 2.9

$$\int_{y}^{c_i}\phi(c_i,x,\mu_j-\mu_i)dx \le \frac{\epsilon}{\underline{\Lambda}_j-\overline{\Lambda}_i}$$

and, by an easy computation,

$$\phi(\rho_j,\rho_i;\mu_j) \le e^{-(\underline{\Lambda}_j-\rho_i)^2/(2\epsilon(1+\eta))},$$
$$\phi(\rho_i,c_i;\mu_j-\mu_i) \le e^{(\overline{\Lambda}_j-\underline{\Lambda}_i)|c_i-\rho_i|/\epsilon},$$
$$I_i/I_j = O(1/\epsilon).$$

Using these observations we also get for all $y < c_i$

$$\left|\varphi_i^\star(y)\int_{c_i}^{y}\frac{\varphi_j^\star(x)}{\varphi_i^\star(x)}\,dx\right| \le Ce^{-\beta_{ij}/\epsilon}\varphi_i^\star(y),$$

with

$$\beta_{ij} = -(\overline{\Lambda}_j-\underline{\Lambda}_i)|c_i-\rho_i| + \frac{(\underline{\Lambda}_j-\rho_i)^2}{2(1+\eta)}$$
$$\ge -(\overline{\Lambda}_j-\underline{\Lambda}_i)(\overline{\Lambda}_i-\underline{\Lambda}_i) + \frac{1}{2(1+\eta)}(\underline{\Lambda}_j-\overline{\Lambda}_i)^2$$

When δ_0 tends to 0, $(\overline{\Lambda}_j-\underline{\Lambda}_i)$ remains bounded, $(\overline{\Lambda}_i-\underline{\Lambda}_i)$ vanishes, while $(\underline{\Lambda}_j-\overline{\Lambda}_i)$ tends to $\lambda_j(0,0)-\lambda_i(0,0) \neq 0$, so assuming δ_0 small enough, we can suppose each quantity β_{ij} is positive. The first desired result (2.51) therefore follows.
Using the inequality $|a\,b| \le (a^2+b^2)/2$ we have

$$|F_{j,k\to i}| \le \frac{1}{2}(|F_{j,j\to i}| + |F_{k,k\to i}|).$$

So, we only need to consider the coefficients of the form $F_{jj\to i}$, that is,

$$G_{i,j}(y) := \frac{\phi(\rho_i,y;\mu_i)}{I_j^2}\left|\int_{c_i}^{y}\frac{\phi(\rho_j,x;\mu_j)^2}{\phi(\rho_i,x;\mu_i)}\,dx\right|.$$

To avoid any distinction between the cases $y > c_i$ and $y < c_i$, the integral would be noted $\int_{[c_i,y]}$ (this is allowed by the positivity of integrand). Clearly, when $j = i$ we have

$$G_{i,i} \leq \varphi_i,$$

because

$$\int_{[c_i,y]} \phi(\rho_i, x; \mu_i)\, dx \leq I_i.$$

So, we now suppose $i \neq j$, then

$$\begin{aligned} G_{i,j}(y) &= \frac{\phi(\rho_i, y; \mu_i)}{I_j^2} \int_{[c_i,y]} \phi(\rho_i, x; -\mu_i)\, \phi(\rho_j, x; \mu_j)^2\, dx \\ &= \frac{\phi(\rho_j, y; \mu_j)}{I_j^2} \int_{[c_i,y]} \phi(x, y; \mu_i - \mu_j)\, \phi(\rho_j, x; \mu_j)\, dx \\ &\leq \frac{\phi(\rho_j, y; \mu_j)}{I_j^2} \int_{[c_i,y]} \phi(x, y; \mu_i - \mu_j)\, dx, \end{aligned}$$

thanks to the judicious choice of ρ_j that gives $\phi(\rho_j, x; \mu_j) \leq 1$. Note the important use of this L^∞ estimate that will default in the future. Finally we have

$$G_{i,j}(y) \leq \frac{1}{I_j} J_{j\to i}(y)$$

which, together with (2.51) and Lemma 2.10 completes this proof and estimates on linear and quadratic wave interaction coefficients.

□

Proof of Lemma 2.13. Fix $y \in [\underline{\Lambda}_j, \overline{\Lambda}_j]$ and recall that

$$\varphi_i^\star(y) \leq \frac{C}{\epsilon} \exp\left(\frac{1}{\epsilon} \int_{\rho_i}^{y} (\lambda_i(t) - t)\, dt\right).$$

If $y \geq \rho_i$ then

$$\begin{aligned} \int_{\rho_i}^{y} (\lambda_i(t) - t)\, dt \leq \int_{\rho_i}^{y} (\overline{\Lambda}_i - t)\, dt &\leq \frac{1}{2}\left((\overline{\Lambda}_i - \rho_i)^2 - (\overline{\Lambda}_i - y)^2\right) \\ &\leq \frac{1}{2}\left((\overline{\Lambda}_i - \underline{\Lambda}_i)^2 - (\overline{\Lambda}_i - y)^2\right), \end{aligned}$$

while, if $y \leq \rho_i$ then

$$\begin{aligned} \int_{\rho_i}^{y} (\lambda_i(t) - t)\, dt = \int_{y}^{\rho_i} (t - \lambda_i(t))\, dt &\leq \int_{y}^{\rho_i} (t - \underline{\Lambda}_i)\, dt \\ &\leq \frac{1}{2}\left((\underline{\Lambda}_i - \rho_i)^2 - (\underline{\Lambda}_i - y)^2\right) \leq \frac{1}{2}\left((\underline{\Lambda}_i - \overline{\Lambda}_i)^2 - (\underline{\Lambda}_i - y)^2\right). \end{aligned}$$

If $i < j$ then for all $\epsilon > 0$, $y \le \rho_i$ and it follows

$$\int_{\rho_i}^{y} (\lambda_i(t) - t)\, dt \le \frac{1}{2}\Big((\overline{\Lambda}_i - \underline{\Lambda}_i)^2 - (\overline{\Lambda}_i - \underline{\Lambda}_j)^2\Big),$$

while if $j < i$ then for all $\epsilon > 0$, $y \ge \rho_i$ and we obtain

$$\int_{\rho_i}^{y} (\lambda_i(t) - t)\, dt \le \frac{1}{2}\Big((\overline{\Lambda}_i - \underline{\Lambda}_i)^2 - (\underline{\Lambda}_i - \overline{\Lambda}_j)^2\Big),$$

In all cases, we thus can write

$$\varphi_i^\star(y) \le \frac{C}{\epsilon} \exp\Big(\frac{1}{2\epsilon}\Big(\ell_i^2 - \Delta_{ij}^2\Big)\Big),$$

where $\ell_i = \overline{\Lambda}_i - \underline{\Lambda}_i$ represents the width of the i-th wave and Δ_{ij} the gap between two different waves $\Delta_{ij} = \min(|\overline{\Lambda}_i - \underline{\Lambda}_j|, |\overline{\Lambda}_j - \underline{\Lambda}_i|)$.
Moreover, we have

$$\frac{1}{\varphi_j^\star(y)} \le 2M \exp\left(-\frac{1}{\epsilon}\int_{\rho_j}^{y} (\lambda_j(t) - t)\, dt\right) \le 2M e^{(\overline{\Lambda}_j - \underline{\Lambda}_j)^2/\epsilon} \le 2M e^{\ell_j^2/\epsilon}.$$

So finally,

$$\frac{\varphi_i^\star(y)}{\varphi_j^\star(y)} \le 2MC\epsilon^{-1} e^{\frac{\ell_j^2 + \ell_i^2 - \Delta_{ij}^2}{\epsilon}},$$

Choosing δ_0 small enough, we can assure positivity of the all quantities $\Delta_{ij}^2 - \ell_j^2 + \ell_i^2$, and the result follows, the multiplicative coefficient ϵ^{-1} can be forgotten simply by reducing the exponential factor D. □

Remark 2.2. For the case of quadratic interaction coefficients that involves the coupling waves ψ, the previous method does not work any more because it would require an L^∞ estimate on ψ that obviously fails. However, we present in Lemma 2.14 an important result to control the interaction coefficient $J_{j\to i}^{\psi}$. Observe this quantity is considered as a linear interaction coefficient of the i-waves on the j-wave through the coupling wave ψ rather than as a quadratic interaction coefficient of $\varphi_i^\star$ and ψ onto $\varphi_j^\star$. This wave do not play the same part than other waves of the hyperbolic system in u, because it transcribes only the effects of the coupling.

Proof of Lemma 2.14. By integration by part, denoting $\Psi \in L^\infty$ a primitive of ψ, (with so $\|\Psi\|_\infty \le \|\psi\|_1 < \infty$), we obtain

$$\begin{aligned} J_{j\to i}^{\psi}(y) &= \varphi_j^\star(y)\left[\Psi(x)\frac{\varphi_j^\star(x)}{\varphi_i^\star(x)}\right]_{c_i}^{y} - \varphi_i^\star(y)\int_{c_j}^{y} \Psi(x)\frac{d}{dx}\left(\frac{\varphi_j^\star(x)}{\varphi_i^\star(x)}\right)dx \\ &= \varphi_j^\star(y)\Psi(y) - \varphi_i^\star(y)\Psi(c_i)\frac{\varphi_j^\star(c_i)}{\varphi_i^\star(c_i)} - \varphi_i^\star(y)\int_{c_i}^{y} \Psi(x)\frac{d}{dx}\left(\frac{\varphi_j^\star(x)}{\varphi_i^\star(x)}\right)dx. \end{aligned}$$

The explicit formula for the $\varphi_k^\star$ gives

$$\frac{d}{dx}\left(\frac{\varphi_j^\star(x)}{\varphi_i^\star(x)}\right) = \frac{d}{dx}\left(\frac{I_i}{I_j}\exp\Big(\frac{1}{\epsilon}\int_{\rho_j}^x(\lambda_j(t)-t)\,dt - \frac{1}{\epsilon}\int_{\rho_i}^x(\lambda_j(t)-t)\,dt\Big)\right)$$
$$= \frac{1}{\epsilon}\big(\lambda_j(x)-\lambda_i(x)\big)\frac{\varphi_j^\star(x)}{\varphi_i^\star(x)}.$$

Consequently,

$$\left|J_{j\to i}^{\psi}(y)\right| \le \|\Psi\|_\infty\left(\varphi_j^\star(y)+\varphi_i^\star(y)\left\|\frac{\varphi_j^\star}{\varphi_i^\star}\right\|_{L^\infty([\underline{\Lambda}_i,\overline{\Lambda}_i])} + \frac{\|\lambda_i-\lambda_j\|_\infty}{\epsilon}\left|\varphi_i^\star(y)\int_{c_i}^y\frac{\varphi_j^\star(x)}{\varphi_i^\star(x)}\,dx\right|\right).$$

Thus, by Lemma 2.13, and by the knowledge (Lemma 2.12) about binary interaction terms for two different waves, to know

$$\left|J_{j\to i}(y)\right| \le O(\epsilon)\Big(\varphi_j^\star(y)+\varphi_i^\star(y)\Big),$$

we get

$$\left|J_{j\to i}^{\psi}(y)\right| \le \|\psi\|_1\Big(\varphi_j^\star(y)+\varphi_i^\star(y)Ce^{-D/\epsilon}+O(1)(\varphi_j^\star(y)+\varphi_i^\star(y))\Big).$$

The result follows. □

Remark 2.3. The method just employed could be used to review Lemma 2.12 directly, noticing $\|\varphi_k^\star\|_1 = 1$ and $J_{j\to i}^{\varphi_k^\star} = F_{j,k\to i}$, and on the other side, $\|\varphi_j^\star\|_1 = 1$ and $J_{k\to i}^{\varphi_j^\star} = F_{j,k\to i}$, we deduce simultaneously

$$\left|F_{j,k\to i}(y)\right| \le O(1)\Big(\varphi_j^\star(y)+\varphi_i^\star(y)\Big),$$
$$\left|F_{j,k\to i}(y)\right| \le O(1)\Big(\varphi_k^\star(y)+\varphi_i^\star(y)\Big).$$

And by this way (2.52) follows.

2.4 Construction of the entropy solution

2.4.1 Correction vector for a given strength

Let $C_0(\mathbb{R})$ stands for continuous functions that decay to zero as $|\xi| \to +\infty$ and define the following weighted sup-norm for $\theta \in [C_0(\mathbb{R})]^N$

$$\|\theta\| = \sum_{k=1}^N \sup_{\xi\in\mathbb{R}}\frac{|\theta_k(\xi)|}{\sum_{h=1}^N\varphi_h^\star(\xi)}. \tag{2.54}$$

Roughly speaking, the correction θ we introduce in the development (2.42), recall

$$a_i = \tau_i\varphi_i^\star + \theta_i,$$

to solve the equation is allowed to present shocks (i.e. singularities as ϵ goes to 0) at most at points where the $\varphi_i^\star$ present themselves shocks, but their strength must stay at most of same order as strength of each $\varphi_i^\star$. Thus we search θ in the Banach space

$$E = \Big\{\theta = (\theta_1, \dots, \theta_N) \in [C_0(\mathbb{R})]^N \ : \ \|\theta\| < \infty\Big\}, \tag{2.55}$$

with the previously introduced weighted sup-norm $\|\cdot\|$. For $\delta > 0$, let B_δ be the 0-centered ball with radius δ in $\mathbb{R}^N$

$$B_\delta = \Big\{\tau \in \mathbb{R}^N \ : \ |\tau| \le \delta\Big\}, \tag{2.56}$$

and for $\tau \in B_\delta$

$$\mathcal{F} = \Big\{\theta \in E \ : \ |\theta_k(\xi)| \le A(\eta|\tau| + |\tau|^2 + \nu|\tau|) \sum_h \varphi_h^\star(\xi),\ k = 1, \dots, N\Big\}, \tag{2.57}$$

where A is a positive constants to be adjusted later and $\nu = \|\sigma\|_\infty$ (remember Remark 2.1, this parameter ν is though to be characteristic of the closeness of both coupled models).

The set $\mathcal{F}$ is a closed bounded subset of E in the weighted norm $\|\cdot\|$. The quadratic quantity $|\tau|^2$ is already present in Tzavaras [132] and comes from quadratic interactions between the $\tau_i\varphi_i^\star$ waves. Note however that the presence of the coupling wave ψ (of unit total mass), with its strength ν, enforces the subset $\mathcal{F}$ to contain correction waves that comes from interactions between ψ and the $\tau_i\varphi_i^\star$, and of strength at most $\nu|\tau|$ relative to the $\varphi_h^\star$.

Now, e will define, for a given strength τ the correction $\theta(\tau;\cdot)$. Let define the map T that takes $u \in \bar\Omega$ where

$$\bar\Omega = \Big\{u \in C^0([-M, M]), \|u(\cdot) - u_L\|_\infty \le \varsigma\Big\}, \tau \in B_\delta$$

and $\theta \in \mathcal{F}$ to the vector-valued function $T(u, \tau, \theta)$ whose components are given for $k = 1, \dots, N$ by:

$$\begin{aligned} T_k(u,\tau,\theta)(\xi) &= \eta\, \varphi_k^\star(\xi) \int_{c_k}^{\xi} \frac{1}{\varphi_k^\star(x)} \sum_i \pi_{ik}(x)\Big(\tau_i\varphi_i^\star(x) + \theta_i(x)\Big)dx \\ &+ \varphi_k^\star(\xi) \int_{c_k}^{\xi} \frac{1}{\varphi_k^\star(x)} \sum_{i,j} \kappa_{ijk}(x)\Big(\tau_i\varphi_i^\star(x) + \theta_i(x)\Big)\Big(\tau_j\varphi_j^\star(x) + \theta_j(x)\Big)dx \\ &+ \varphi_k^\star(\xi) \int_{c_k}^{\xi} \frac{1}{\varphi_k^\star(x)} \sum_i \sigma_{ik}(x)\Big(\tau_i\varphi_i^\star(x) + \theta_i(x)\Big)\psi(x)\,dx. \end{aligned} \tag{2.58}$$

This map T has the following contraction properties:

Lemma 2.15. *There exists positive constants A, η, δ_0 and ν such that for $\delta < \delta_0$:*

1. *$T : \bar\Omega \times B_\delta \times \mathcal{F} \to \mathcal{F}$ is well defined.*
2. *There exists $0 < \alpha < 1$, such that*

$$\Big\|T(u,\tau,\theta) - T(u,\tau,\hat\theta)\Big\| \le \alpha\Big\|\theta - \hat\theta\Big\|, \qquad \theta, \hat\theta \in \mathcal{F},$$

and for any $u \in \bar{\Omega}$, $\tau \in B_\delta$. Therefore $T(u,\tau,\cdot) : \mathcal{F} \to \mathcal{F}$ is a uniform contraction.

3. *There exists a positive constant C, depending on μ but independent of δ, such that*

$$\left\| T(u,\tau,\theta) - T(u,\hat{\tau},\theta) \right\| \leq C(\eta+\nu+\delta)|\tau-\hat{\tau}|, \qquad \tau,\hat{\tau} \in B_\delta$$

and for any $u \in \bar{\Omega}$, $\theta \in \mathcal{F}$.

We deduce from this lemma the following existence result of a correction $\theta(\tau;\cdot)$

Proposition 2.16. *Given $u \in \bar{\Omega}, \tau \in B_\delta$, there exists a unique $\theta(\tau;\cdot) \in \mathcal{F}$, i.e. in the class of functions satifsying*

$$|\theta_k(\tau;\cdot)| \leq A(\eta|\tau| + |\tau|^2 + \nu|\tau|) \sum_h \varphi_h^\star, \qquad |\tau| \leq \delta,\ k = 1,\dots,N, \tag{2.59}$$

solution of the fixed point equation $T(u,\tau,\theta) = \theta$. Moreover, there exists a constant C independent of δ such that $\theta(\cdot;\tau)$ satisfies

$$|\theta_k(\tau;\cdot) - \theta_k(\hat{\tau};\cdot)| \leq C(\eta+\nu+\delta)|\tau-\hat{\tau}| \sum_h \varphi_h^\star, \qquad \tau,\hat{\tau} \in B_\delta. \tag{2.60}$$

Proof of Lemma 2.15. We follow [132], but need to introduce the additional coupling wave $\nu\psi$.

- First of all, we show T keeps stable the subset $\mathcal{F}$, using the definitions (2.57), (2.58), and the different definitions of interaction coefficients (2.43), (2.44) and (2.45), we get

$$\begin{aligned}
&|T_k(u,\tau,\theta)(\xi)| \\
&\leq \eta\|\pi\|_\infty \sum_i J_{i\to k}(\xi)\Big(|\tau| + A\left(\eta|\tau| + |\tau|^2 + \nu|\tau|\right)\sum_j 1\Big) \\
&\quad + \|\kappa\|_\infty \sum_{ij} F_{ij\to k}(\xi)\Big(|\tau|^2 + 2|\tau|A\left(\eta|\tau| + |\tau|^2 + \nu|\tau|\right)\sum_l 1 + A^2\left(\eta|\tau| + |\tau|^2 + \nu|\tau|\right)^2 \sum_{lm} 1\Big) \\
&\quad + \|\sigma\|_\infty \sum_i J^\psi_{i\to k}(\xi)\Big(|\tau| + A\left(\eta|\tau| + |\tau|^2 + \nu|\tau|\right)\sum_j 1\Big).
\end{aligned}$$

  Thus, by using Lemmas 2.12 and 2.14 we get

$$\begin{aligned}
&|T_k(u,\tau,\theta)(\xi)| \\
&\leq \eta\|\pi\|_\infty\, C_1 N^2\Big(|\tau| + A\left(\eta|\tau| + |\tau|^2 + \nu|\tau|\right)\Big)\sum_h \varphi_h^\star(\xi) \\
&\quad + \|\kappa\|_\infty\, C_2 N^4\Big(|\tau|^2 + 2A|\tau|\left(\eta|\tau| + |\tau|^2 + \nu|\tau|\right) + A^2\left(\eta|\tau| + |\tau|^2 + \nu|\tau|\right)^2\Big)\sum_h \varphi_h^\star(\xi) \\
&\quad + \|\sigma\|_\infty\, C_3 N^2\Big(|\tau| + A\left(\eta|\tau| + |\tau|^2 + \nu|\tau|\right)\Big)\sum_h \varphi_h^\star(\xi) \\
&\leq C\left(1 + A(\eta+\nu+\delta)\right)^2(\eta|\tau| + |\tau|^2 + \nu|\tau|)\sum_h \varphi_h^\star(\xi),
\end{aligned}$$

where C is a constant depending only on N the dimension of the space, on $\|\pi\|_\infty$, $\|\kappa\|_\infty$ and of the constants C_1, C_2, C_3. A necessary condition to get the stability of the subset $\mathcal{F}$ by $T(u,\tau,\cdot)$ is also

$$C(1+A(\eta+\nu+\delta))^2 \leq A$$

A way to get this inequality is for example, fixing $A = 4C$, to choose η, ν and δ together such that

$$\eta+\nu+\delta \leq 1/4C.$$

- T is an uniform contraction relative to the variable $\theta \in \mathcal{F}$, because with the same tools we have

$$\begin{aligned}
&|T_k(u,\tau,\theta)(\xi) - T_k(u,\tau,\hat{\theta})(\xi)| \\
&\leq \eta\|\pi\|_\infty \sum_j \|\theta-\hat{\theta}\| \sum_i J_{i\to k}(\xi) \\
&\quad + \|\kappa\|_\infty \sum_{ij} \Big(2|\tau|\,\|\theta-\hat{\theta}\| \sum_l F^k_{il}(\xi) + 2A\left(\eta|\tau| + |\tau|^2 + \nu|\tau|\right)\|\theta-\hat{\theta}\| \sum_{lm} F_{lm\to k}(\xi)\Big) \\
&\quad + \|\sigma\|_\infty \sum_j \|\theta-\hat{\theta}\| \sum_i J^\psi_{i\to k}(\xi).
\end{aligned}$$

Thus, by using Lemmas 2.12 and 2.14 we get

$$\begin{aligned}
|T_k(u,\tau,\theta)(\xi) - T_k(u,\tau,\hat{\theta})(\xi)| &\leq \eta\|\pi\|_\infty\, C_1N^2\|\theta-\hat{\theta}\| \sum_h \varphi^\star_h(\xi) \\
&\quad + \|\kappa\|_\infty\, C_2N^4\left(|\tau| + A\left(\eta|\tau| + |\tau|^2 + \nu|\tau|\right)\right)\|\theta-\hat{\theta}\| \sum_h \varphi^\star_h(\xi) \\
&\quad + \|\sigma\|_\infty\, C_3N^2\|\theta-\hat{\theta}\| \sum_h \varphi^\star_h(\xi) \\
&\leq C(\eta+\nu+\delta)(1+A\delta)\|\theta-\hat{\theta}\| \sum_h \varphi^\star_h(\xi).
\end{aligned}$$

Finally, we obtain the point (2) of Proposition 2.15 with $\alpha = C(\eta+\nu+\delta)(1+A\delta)$ by choosing $\eta+\nu+\delta$ sufficiently small to assure that $\alpha < 1$.

- The last point concerns a partial lipschitz property of T relative to the variable τ.

$$\begin{aligned}
&|T_k(u,\tau,\theta)(\xi) - T_k(u,\hat{\tau},\theta)(\xi)| \\
&\leq \eta\|\pi\|_\infty|\tau-\hat{\tau}| \sum_i J_{i\to k}(\xi) \\
&\quad + \|\kappa\|_\infty \sum_{ij} \Big(|\tau_i\hat{\tau}_j - \tau_j\hat{\tau}_i| F_{ij\to k}(\xi) + 2|\tau-\hat{\tau}| A\delta(\eta+\delta+\nu) \sum_l F_{il\to k}(\xi)\Big) \\
&\quad + \|\sigma\|_\infty|\tau-\hat{\tau}| \sum_i J^\psi_{i\to k}(\xi).
\end{aligned}$$

Thus, by using Lemmas 2.12 and 2.14 we get

$$\begin{aligned}|T_k(u,\tau,\theta)(\xi)-T_k(u,\hat{\tau},\theta)(\xi)| &\le \eta\|\pi\|_\infty C_1 N|\tau-\hat{\tau}|\sum_h \varphi_h^\star(\xi)\\ &\quad+\|\kappa\|_\infty C_2 N^2\Big(2\delta|\tau-\hat{\tau}|+2AN\delta|\tau-\hat{\tau}|\Big)\sum_h \varphi_h^\star(\xi)\\ &\quad+\|\sigma\|_\infty C_3 N|\tau-\hat{\tau}|\sum_h \varphi_h^\star(\xi)\\ &\le C(\eta+\nu+\delta)|\tau-\hat{\tau}|\sum_h \varphi_h^\star(\xi),\end{aligned}$$

and the point (3) of Proposition 2.15 follows.

□

Proof of Proposition 2.16. The inequality (2.59) is a direct consequence of the contraction mapping theorem, that previous proposition ensures to apply ($\alpha<1$). Let $u\in\bar{\Omega}$ be given, and $\tau,\hat{\tau}\in B_\delta$, then

$$\begin{aligned}\theta(\tau)-\theta(\hat{\tau}) &= T(u,\tau,\theta(\tau))-T(u,\hat{\tau},\theta(\hat{\tau}))\\ &= \Big(T(u,\tau,\theta(\tau))-T(u,\tau,\theta(\hat{\tau}))\Big)+\Big(T(u,\tau,\theta(\hat{\tau}))-T(u,\hat{\tau},\theta(\hat{\tau}))\Big).\\ \|\theta(\tau)-\theta(\hat{\tau})\| &\le \Big\|T(u,\tau,\theta(\tau))-T(u,\tau,\theta(\hat{\tau}))\Big\|+\Big\|T(u,\tau,\theta(\hat{\tau}))-T(u,\hat{\tau},\theta(\hat{\tau}))\Big\|\\ &\le \alpha\Big\|\theta(\tau)-\theta(\hat{\tau})\Big\|+C(\eta+\nu+\delta)|\tau-\hat{\tau}|.\end{aligned}$$

Hence,

$$\|\theta(\tau)-\theta(\hat{\tau})\|\le\frac{C}{1-\alpha}(\eta+\nu+\delta)|\tau-\hat{\tau}|,$$

and (2.60) ensues. □

2.4.2 Strength vector for given Riemann data

Fix a left-state vector $u_L\in\mathbb{R}^N$ and $u\in\bar{\Omega}=\Big\{u\in C^0([-M,M]),\|u(\cdot)-u_L\|_\infty\le\varsigma\Big\}$. Being given $\tau\in B_\delta$, we previously construct a unique $\theta(\tau,\cdot)\in\mathcal{F}$ such that $T(u,\tau,\theta)=\theta$. The question is now to link the vector $\tau\in B_\delta$ together with the boundary data u_L,u_R. Consider the following operator:

$$S(\tau)=u_L+A_0(u,v)^{-1}\sum_k\int_{-M}^{M}\Big[\tau_k\varphi_k^\star(\xi)+\theta_k(\tau,\xi)\Big]\widehat{r_k}(u(\xi),v(\xi),\xi)d\xi.$$

Lemma 2.17. *There exists positive constants δ and r such that $P:B_r(u_L)\times\bar{\Omega}\times B_\delta\to B_\delta$ defined by the vector*

$$P(u_R,u,\tau)=A_0(u,v)(u_R-u_L)-\sum_k\int_{-M}^{M}\theta_k(\tau,\xi)\widehat{r_k}(u(\xi),v(\xi),\xi)\,d\xi.$$

has the property that there exists a constant $0 < \alpha < 1$ *such that*

$$|P(u_R, u, \tau) - P(u_R, u, \hat{\tau})| \leq \alpha |\tau - \hat{\tau}|, \qquad \tau, \hat{\tau} \in B_\delta$$

for any $u_R \in B_r(u_L), u \in \bar{\Omega}$*; that is* $P(u_R, u, \cdot)$ *is a uniform contraction on* B_δ*.*

Proposition 2.18. *Assume* $u \in \bar{\Omega}$*. There exists positive constants* r *and* δ *such that:*

1. *Given* $u_R \in B_r(u_L)$ *there exists a unique solution of the equation* $S(\tau) = u_R$ *with* $\tau \in B_\delta$*.*
2. *For each* $u \in \bar{\Omega}$ *and* $\epsilon > 0$ *the inverse map* $S^{-1} : B_r(u_L) \to B_\delta$ *is well defined and satisfies*

$$|S^{-1}(u_R)| \leq \gamma |u_R - u_L|, \tag{2.61}$$

where γ *is a constant which depends on* ς*, but is independent of* $u \in \bar{\Omega}$ *and of* ϵ*.*

Proof of Lemma 2.17. Let $u_R \in B_r(u_L)$, $u \in \bar{\Omega}$ and $\tau \in B_\delta$,

$$\begin{aligned} |P(u_R, u, \cdot)| &\leq \|A_0\| |u_R - u_L| + N^2 R A(\eta|\tau| + |\tau|^2 + \nu|\tau|) \\ &\leq \Big(\|A_0\| r + RN^2 A(\eta\delta + \delta^2 + \nu\delta) \Big) \end{aligned}$$

$P(u_R, u, B_\delta) \subset B_\delta$ is realised if

$$\Big(\|A_0\| r + RN^2 A(\eta\delta + \delta^2 + \nu\delta) \Big) \leq \delta$$

that is by choosing r, η, δ and ν such that

$$RN^2 A(\eta + \delta + \nu)\delta \leq \delta/2,$$

$$\|A_0\| \beta r \leq \delta/2.$$

that is to say

$$\eta + \delta + \nu \leq 1/2RN^2 A,$$

$$r \leq \delta/2\|A_0\|.$$

Given τ and $\hat{\tau}$ in B_δ

$$P(u_R, u, \tau) - P(u_R, u, \hat{\tau}) = \sum_k \int_{-M}^{M} [\theta_k(\xi, \tau) - \theta_k(\xi, \hat{\tau})] \widehat{r_k}(u(\xi), v(\xi), \xi) d\xi,$$

$$\begin{aligned} |P(u_R, u, \tau) - P(u_R, u, \hat{\tau})| &\leq \sum_k \int_{-M}^{M} |\theta_k(\xi, \tau) - \theta_k(\xi, \hat{\tau})| \, |\widehat{r_k}(u(\xi), v(\xi), \xi)| d\xi \\ &\leq RNC(\eta + \delta + \nu)|\tau - \hat{\tau}| \sum_k \int_{-M}^{M} \varphi_k^\star(\xi) d\xi \\ &\leq N^2 C(\eta + \delta + \nu)|\tau - \hat{\tau}|. \end{aligned}$$

Provided that

$$\alpha := N^2 C(\eta + \delta + \nu) < 1, \tag{2.62}$$

then $P(u_R,u,\cdot)$ is a uniform contraction on B_δ. □

Proof of Proposition 2.18. Let u_L be fixed. The equation $S(\tau)=u_R$ takes the form

$$\begin{aligned}&A_0(u,v)(u_R-u_L)\\&=\sum_k \tau_k\int_{-M}^{M}\varphi_k^\star(\xi)\widehat{r_k}(u(\xi),v(\xi),\xi)d\xi+\sum_k\int_{-M}^{M}\theta_k(\tau,\xi)\widehat{r_k}(u(\xi),v(\xi),\xi)d\xi,\end{aligned}$$

in other words τ solves the equation

$$A_0(u,v)(u_R-u_L)=C(u,v)\,\tau+\sum_k\int_{-M}^{M}\theta_k(\tau,\xi)\widehat{r_k}(u(\xi),v(\xi),\xi)d\xi, \tag{2.63}$$

where $C(u,v)$ is the matrix whose k-th columns is given by

$$\int_{-M}^{M}\varphi_k^\star(\xi)\widehat{r_k}(u(\xi),v(\xi),\xi)d\xi,\qquad k=1,\dots,N.$$

This matrix has the important property it is invertible for any $u\in\bar\Omega$ and the inverse matrix $C(u,v)^{-1}$ is uniformly bounded (cf [132])

$$|C(u,v)^{-1}|\le\beta,\qquad u\in\bar\Omega. \tag{2.64}$$

In order to solve the equation $S(\tau)=u_R$, observe that solutions of (2.63) are also fixed points of the map $\tau\mapsto C(u,v)^{-1}P(u_R,u,\tau)$, whose existence are ensured by Lemma 2.17. As a consequence, given $u_R\in B_r(u_L)$, there exists a unique fixed point τ of $P(u_R,u,\cdot)$ in the ball B_δ. Moreover it also satisfies

$$\begin{aligned}|\tau|&\le\left|A_0(u,v)C(u,v)^{-1}\right||u_R-u_L|+\left|C(u,v)^{-1}\right|\sum_k\int_{-M}^{M}|\theta_k(\xi,\tau)|\widehat{r_k}(u(\xi),v(\xi),\xi)|d\xi\\&\le\|A_0\|\beta|u_R-u_L|+\beta RAN^2(\eta|\tau|+|\tau|^2+\nu|\tau|)\\&\le\|A_0\|\beta|u_R-u_L|+1/2|\tau|.\end{aligned}$$

Thus, $|\tau|\le 2\|A_0\|\beta|u_R-u_L|$, which finally implies (2.61). □

2.4.3 Riemann problem

We search for a solution of (2.36) under the form (2.37) satisfying the boundary conditions (2.34) and where v is known by (2.35).

Theorem 2.19. *Under the smallness assumptions, there exists $u^\epsilon\in\bar\Omega$ solution of (2.34)-(2.36). Moreover there exists a constant K independent of ϵ such that following inequalities hold for all ϵ,*

$$\begin{aligned}&TV(u^\epsilon)\le K|u_R-u_L|,\\&\epsilon|u^\epsilon_\xi|\le K.\end{aligned} \tag{2.65}$$

Proof. $\varsigma > 0$ is choosen so that conditions of eigenvalue separation are fulfilled on $\bar{\Omega}$. Fix u_L and $u \in \bar{\Omega}$. For ϵ fixed, we construct z as

$$z(\xi) = u_L + A_0^{-1}(u,v) \int_{-M}^{\xi} \sum_{j=1}^{n} \Big(\tau_j \varphi_j^\star(\zeta) + \theta_j(\tau;\zeta)\Big) \widehat{r_j}(u(\zeta), v(\zeta), \zeta)\, d\zeta$$

by following steps:

1. Each $\varphi_j^\star$ is constructed as the fundamental wave measure from (2.40), remember

$$\varphi_j^{\star\prime} - \frac{\mu_j(u,v,\cdot)}{\epsilon} \varphi_j^\star = 0.$$

2. For each τ small enough we can get, through Proposition 2.16, a correction $\theta(\tau,\cdot)$ so that $a_j = \tau_j \varphi_j^\star + \theta_j$ is solution of (2.38a)

$$a_j' - \frac{\mu_j(u,v,\cdot)}{\epsilon} a_j = \eta L_j(u,v,\cdot) + Q_j(u,v,\cdot) + S_j(u,v,\cdot).$$

3. The vector of strength τ is then chosen, through Proposition 2.18, as a solution of $S(\tau) = u_R \in B_r(u_L)$. This way the solution $\tilde{u}$ of

$$A_0(u,v)\, \tilde{u}_\xi(\xi) = \sum_j a_j(\xi) \widehat{r_j}(u,v,\xi)$$

satisfying $\tilde{u}(-M) = u_L$, satisfies moreover $\tilde{u}(M) = u_R$.

These steps allow us to construct an operator $\mathcal{T} : \bar{\Omega} \to E, u \mapsto z$, and $\mathcal{T}(u) = z \in \bar{\Omega}$. We only need to get a fixed point result on $\mathcal{T}$ to get the solution u of the whole problem, and then sufficiently strong estimates to ensures existence of the limit as ϵ tends to 0. □

We then have:

Lemma 2.20. *The function v^ϵ converges towards the sign function denoted "*sgn*", and more precisely one has for $c > 0$*

$$\left\| v^\epsilon - \mathrm{sgn} \right\|_{L^\infty(\mathbb{R}\setminus[-c,c])} = o(\epsilon). \tag{2.66}$$

Proof. Indeed the formula (2.35) implies v^ϵ takes the form

$$v^\epsilon(\xi) = -1 + 2 \frac{\int_{-M}^{\xi} e^{-\frac{x^2}{2\epsilon^p}} dx}{\int_{-M}^{M} e^{-\frac{x^2}{2\epsilon^p}} dx}.$$

Fix $\xi > c > 0$, so that

$$\left| v^\epsilon(x) - 1 \right| \le 2 \frac{\int_{c}^{M} e^{-\frac{x^2}{2\epsilon^p}} dx}{\int_{-M}^{M} e^{-\frac{x^2}{2\epsilon^p}} dx} \le 2 \frac{M e^{-\frac{c^2}{2\epsilon^p}}}{\epsilon^{p/2} \int_{-M/\epsilon^{p/2}}^{M/\epsilon^{p/2}} e^{-\frac{y^2}{2}} dy} \le \frac{C}{\epsilon^{p/2}} e^{-\frac{c^2}{2\epsilon^p}}.$$

For $\xi < -c < 0$, by the same procedure, we get

$$\left|v^\epsilon(x)+1\right| \le \frac{C}{\epsilon^{p/2}} e^{-\frac{c^2}{2\epsilon^p}},$$

the lemma is therefore proved. □

Theorem 2.21 (Convergence to an entropy solution out of the interface)**.** *The sequence u^ϵ converges pointwise towards $u \in BV$, satisfying*

$$\begin{aligned} -\xi \frac{d}{d\xi}\gamma_-(u) + \frac{d}{d\xi} f_-(\gamma_-(u)) = 0, \qquad \mathcal{D}'(\mathbb{R}_-), \\ -\xi \frac{d}{d\xi}\gamma_+(u) + \frac{d}{d\xi} f_+(\gamma_+(u)) = 0, \qquad \mathcal{D}'(\mathbb{R}_+). \end{aligned} \tag{2.67}$$

Let $\eta_\pm = \eta_\pm(u) \in \mathbb{R}^N$ be two entropy functions compatible with the viscosity matrix in the sense that

$$\nabla^2 \eta_\pm(u) B_{0\pm}(u) \ge 0, \qquad u \in \mathcal{U}.$$

Then following entropy inequalities are satisfied,

$$\begin{aligned} -\xi \frac{d}{d\xi}\eta_-(\gamma_-(u)) + \frac{d}{d\xi} q_-(\gamma_-(u)) \le 0, \qquad \mathcal{D}'(\mathbb{R}_-), \\ -\xi \frac{d}{d\xi}\eta_+(\gamma_+(u)) + \frac{d}{d\xi} q_+(\gamma_+(u)) \le 0, \qquad \mathcal{D}'(\mathbb{R}_+). \end{aligned} \tag{2.68}$$

Proof. Let O be a neighborhood of 0, the solution u^ϵ of (2.24) is obtained as a smooth function so that, under the consistancy hypothesis (2.12), we have outside O

$$\begin{aligned} A_0(u^\epsilon, v^0)\frac{du^\epsilon}{d\xi} &= \frac{d}{d\xi}\gamma_\pm(u^\epsilon), \\ A_1(u^\epsilon, v^0)\frac{du^\epsilon}{d\xi} &= \frac{d}{d\xi} f_\pm(\gamma_\pm(u^\epsilon)). \end{aligned}$$

Let $\phi \in C_c^\infty(\mathbb{R}_- \setminus O)$ be a test-function with a compact support included in $\mathbb{R}_- \setminus O$, then (2.24) implies

$$\begin{aligned} &\int_{\mathbb{R}} \Big(-\xi\, d_\xi \gamma_-(u^\epsilon) + \frac{d}{d\xi} f_-(\gamma_-(u^\epsilon))\Big)\phi \, d\xi \\ &= \int_{\mathbb{R}} -\xi \, \Big(A_0(u^\epsilon, v^0) - A_0(u^\epsilon, v^\epsilon)\Big) u^\epsilon_\xi \phi \, d\xi \\ &+ \int_{\mathbb{R}} \Big(A_1(u^\epsilon, v^0) - A_1(u^\epsilon, v^\epsilon)\Big) u^\epsilon_\xi \phi \, d\xi \\ &+ \int_{\mathbb{R}} \epsilon \Big(B_0(u^\epsilon, v^\epsilon) u^\epsilon_\xi\Big)_\xi \phi \, d\xi. \end{aligned}$$

Moreover, thanks to Lemma 2.20 and using Lipschitz inequalities, we have

$$\left|\int_{\mathbb{R}} -\xi\Big(A_0(u^\epsilon,v^0)-A_0(u^\epsilon,v^\epsilon)\Big)u^\epsilon_\xi\phi\,d\xi\right| \le o(\epsilon)\mathrm{Lip}(A_0)\|\xi\phi\|_\infty TV(u^\epsilon),$$
$$\left|\int_{\mathbb{R}} \Big(A_1(u^\epsilon,v^0)-A_1(u^\epsilon,v^\epsilon)\Big)u^\epsilon_\xi\phi\,d\xi\right| \le o(\epsilon)\mathrm{Lip}(A_1)\|\phi\|_\infty TV(u^\epsilon),$$
$$\epsilon\left|\int_{\mathbb{R}} \Big(B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi\Big)_\xi\phi\,d\xi\right| = \epsilon\left|\int_{\mathbb{R}} B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi\phi_\xi\,d\xi\right| \le \epsilon\|B_0\|_\infty\|\phi_\xi\|_\infty TV(u^\epsilon).$$

Thus, as ϵ tends to 0, we get the weak formulation for the limit u

$$\int_{\mathbb{R}}\Big(-\xi\,\frac{d}{d\xi}\gamma_-(u)+\frac{d}{d\xi}f_-(\gamma_-(u))\Big)\varphi\,d\xi=0.$$

By a similar method we get for $\varphi\in C_c^\infty(\mathbb{R}_+\setminus O)$

$$\int_{\mathbb{R}}\Big(-\xi\,\frac{d}{d\xi}\gamma_+(u)+\frac{d}{d\xi}f_+(\gamma_+(u))\Big)\varphi\,d\xi=0.$$

Entropy inequalities are obtained using first the consistancy hypothesis (2.12), that give outside O

$$\nabla\eta_\pm(u^\epsilon)\cdot A_0(u^\epsilon,v^0)u^\epsilon_\xi=\frac{d}{d\xi}\eta_\pm(\gamma_\pm(u^\epsilon)),$$
$$\nabla\eta_\pm(u^\epsilon)\cdot A_1(u^\epsilon,v^0)u^\epsilon_\xi=\frac{d}{d\xi}q_\pm(\gamma_\pm(u^\epsilon)).$$

Let $\phi\in C_c^\infty(\mathbb{R}_-\setminus O)$ be a non-negative test function with a compact support included in $\mathbb{R}_-\setminus O$, then (2.24) implies

$$\begin{aligned}
&\int_{\mathbb{R}}\Big(-\xi\,\frac{d}{d\xi}\eta_-(\gamma_-(u^\epsilon))+\frac{d}{d\xi}q_-(\gamma_-(u^\epsilon))\Big)\phi\,d\xi\\
&=\int_{\mathbb{R}}-\xi\,\nabla\eta_-(u^\epsilon)\cdot\Big(A_0(u^\epsilon,v^0)-A_0(u^\epsilon,v^\epsilon)\Big)u^\epsilon_\xi\phi\,d\xi\\
&+\int_{\mathbb{R}}\nabla\eta_-(u^\epsilon)\cdot\Big(A_1(u^\epsilon,v^0)-A_1(u^\epsilon,v^\epsilon)\Big)u^\epsilon_\xi\phi\,d\xi\\
&+\int_{\mathbb{R}}\epsilon\nabla\eta_-(u^\epsilon)\cdot\Big(B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi\Big)_\xi\phi\,d\xi.
\end{aligned}$$

With similar arguments as previously, the first and the second terms of right hand side tends to 0 as ϵ tends to 0. Moreover, after reporting the ξ-derivative on $\phi\nabla\eta_-(u^\epsilon)$, the last term equals

$$\begin{aligned}
&-\int_{\mathbb{R}}\epsilon\phi_\xi\nabla\eta_-(u^\epsilon)\cdot B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi\,d\xi\\
&-\int_{\mathbb{R}}\epsilon\phi\nabla^2\eta_-(u^\epsilon)\cdot B_0(u^\epsilon,v^0)|u^\epsilon_\xi|^2\,d\xi\\
&+\int_{\mathbb{R}}\epsilon\phi\nabla^2\eta_-(u^\epsilon)\cdot(B_0(u^\epsilon,v^0)-B_0(u^\epsilon,v^\epsilon))|u^\epsilon_\xi|^2\,d\xi,
\end{aligned}$$

that satisfy following estimates

$$\left|\int_{\mathbb{R}} \epsilon\phi_\xi \nabla\eta_-(u^\epsilon)\cdot B_0(u^\epsilon,v^\epsilon)u^\epsilon_\xi\, d\xi\right| \le K\epsilon TV(u^\epsilon),$$
$$\left|\int_{\mathbb{R}} \epsilon\phi \nabla^2\eta_-(u^\epsilon)\cdot (B_0(u^\epsilon,v^0)-B_0(u^\epsilon,v^\epsilon))|u^\epsilon_\xi|^2\, d\xi\right|$$
$$\le K\mathrm{Lip}(B_0)\|\nabla^2\eta_-\|_\infty\|\phi\|_\infty TV(u^\epsilon)\|v^\epsilon - v^0\|_{L^\infty(\mathbb{R}\setminus O)}.$$

However the quantity $\int_{\mathbb{R}} \epsilon\phi\nabla^2\eta_-(u^\epsilon)\cdot B_0(u^\epsilon,v^0)|u^\epsilon_\xi|^2\, d\xi$ is not ensure to vanish as ϵ tends to 0, but it converges towards a positive value under the hypothesis $\nabla^2\eta_\pm B_{0\pm}(u)\ge 0$. It follows the following weak formulation of the entropy inequality on $\mathbb{R}_-$

$$\int_{\mathbb{R}}\Big(-\xi\,\frac{d}{d\xi}\eta_-(\gamma_-(u))+\frac{d}{d\xi}q_-(\gamma_-(u))\Big)\phi\, d\xi \le 0.$$

The same method permits to get the entropy inequality on $\mathbb{R}_+$. □

2.5 Appendix : Precised estimates on resonant interaction coefficients

We now obtain local estimates on resonant interaction coefficients $J^{\psi}_{j\to i}$ previously studied (see Lemma 2.14), in the manner of local estimates on linear interaction coefficients $J_{j\to i}(y)$ (see Lemma 2.23). We recall that (the absolute value of) those coefficients takes the form

$$J^{\psi}_{j\to i}(y) := \left|\varphi_i^\star(y)\int_{c_i}^{y}\psi(x)\frac{\varphi_j^\star(x)}{\varphi_i^\star(x)}\,dx\right|,$$

where the $\varphi_k^\star$'s $\in L^1$ are the fundamental measure waves and $\psi \in L^1$ is the coupling wave. Those waves are given through formulae (2.35) and (2.41) and we consider in the sequel a more general context for the coupling modeling where the thickness of the interface is given by a supplemented parameter $\kappa > 0$ satisfying a certain scaling with ϵ. Ommiting for convenience the truncation of the domain to the bounded interval $(-M,M)$, we thus consider as coupling wave the function $\psi = \psi^\kappa$ given by

$$\psi^\kappa(y) = \frac{e^{-y^2/2\kappa}}{\sqrt{2\pi\kappa}}.$$

Remember that thanks to Lemma 2.11, the k-th wave essentially "live" inside the interval $[\underline{\Lambda}_k, \overline{\Lambda}_k]$ in sense that it decreases exponentially out of this interval as ϵ goes to 0, while it may tend to infinity at some isolated points in $[\underline{\Lambda}_k, \overline{\Lambda}_k]$, keeping unit total mass. The coupling wave ψ besides has a similar behaviour and converges to the Dirac function as κ goes to 0.

The following estimates should make clear the result obtained earlier (Lemma 2.14).

Proposition 2.22. *Suppose $\kappa = o(\epsilon)$ and $\kappa = O(\epsilon^2)$, then resonant interaction coefficients are bounded, uniformly in ϵ and κ, as follows*

$$J^{\psi}_{j\to i}(y) \le O(1)\sum_h \varphi_h^\star(y), \quad i,j = 1,\dots,N. \tag{2.69}$$

Remark : For convenience, in different following proofs and statements, we will not take rigorous care about varied constants, and they may implicitely be modified: the multiplicative ones, named C, the exponential ones, named D. Note moreover we choose here as integration constant for the resonant wave $c_m = 0$ without loss of generality.

We first derive local estimates for linear interaction (already treated in Lemma 2.14), and next we treat local resonant interaction coefficients estimates, for the different possible configurations, according to the nature of involved waves.

Lemma 2.23 (Precised estimates on linear interaction coefficients)**.** *Suppose $i < j$, then*

$$J_{j\to i}(y) \le \begin{cases} Ce^{-D/\epsilon}\varphi_i^\star(y), & y \le c_i, \\ \dfrac{\epsilon}{\underline{\Lambda}_j - \overline{\Lambda}_i}\varphi_j^\star(y), & y \ge c_i. \end{cases}$$

Suppose now $j < i$, then

$$J_{j\to i}(y) \le \begin{cases} \dfrac{\epsilon}{\underline{\Lambda}_i - \overline{\Lambda}_j}\varphi_j^\star(y), & y \le c_i, \\ Ce^{-D/\epsilon}\varphi_i^\star(y), & y \ge c_i. \end{cases}$$

Proof. The proof relies on Lemmas 2.9 and 2.10 and follows the proof of Lemma 2.14. □

2.5.1 Self-influence of a wave

If $i = j$, then the estimate is immediat,

$$J^\psi_{i\to i}(y) = \varphi_i^\star(y)\int_{c_i}^{y}\psi(x)dx \le \varphi_i^\star(y)\,\|\psi\|_1.$$

2.5.2 Influence of a non-resonant wave on another non-resonant wave

Consider the most general case where $i, j \neq m$ and $i \neq j$, then both waves $\varphi_i^\star$ and $\varphi_j^\star$ essentially live out of the interface: $0 \notin [\underline{\Lambda}_i, \overline{\Lambda}_i] \cup [\underline{\Lambda}_j, \overline{\Lambda}_j]$. The resonant interaction coefficient satisfies estimates of Lemma 2.25, proved using following tool lemma

Lemma 2.24. *Let $\alpha < \beta$ be two reals, and define*

$$P_\pm(\alpha,\beta) = (2\pi\kappa)^{-1/2}\int_\alpha^\beta \exp\left(-\frac{1}{2\kappa}\left(x \pm \frac{\kappa}{\epsilon}\Delta_{ij}\right)^2\right)dx.$$

Then $P_\pm$ satisfies following estimates, as ϵ and κ tend to 0, under the assumption $\kappa = o(\epsilon)$.

- *If $\alpha \le 0 \le \beta$, then*
$$P_\pm(\alpha,\beta) \le 1.$$
- *If $0 < \alpha < \beta$, then*
$$P_\pm(\alpha,\beta) \le \frac{1}{2}\exp\left(-\frac{1}{2\kappa}\left(\alpha - \frac{\kappa}{\epsilon}\Delta_{ij}\right)^2\right).$$
- *If $\alpha < \beta < 0$, then*
$$P_\pm(\alpha,\beta) \le \frac{1}{2}\exp\left(-\frac{1}{2\kappa}\left(\beta + \frac{\kappa}{\epsilon}\Delta_{ij}\right)^2\right).$$

Lemma 2.25. *Suppose $\kappa = o(\epsilon)$ and $\kappa = O(\epsilon^2)$, and fix $0 < \eta < \min(|\underline{\Lambda}_i|, |\overline{\Lambda}_i|)$. Then there exist $C, D > 0$ independent of ϵ and of κ such that*

- *if $j > i > m$ ($c_i > 0$), then*

$$J^\psi_{j\to i}(y) \le \begin{cases} \varphi_i^\star(y)Ce^{-D/\epsilon}, & y \le \eta, \\ \varphi_i^\star(y)Ce^{-\eta^2/2\kappa}, & \eta \le y \le c_i, \\ \varphi_j^\star(y)Ce^{-\eta^2/2\kappa}, & \eta \le c_i \le y. \end{cases} \tag{2.70}$$

- *if $j < i < m$ ($c_i < 0$), then*

$$J^{\psi}_{j\to i}(y) \le \begin{cases} \varphi^{\star}_j(y) C e^{-\eta^2/2\kappa}, & y \le c_i, \\ \varphi^{\star}_i(y) C e^{-\eta^2/2\kappa}, & c_i \le y \le -\eta, \\ \varphi^{\star}_i(y) C e^{-D/\epsilon}, & -\eta \le y. \end{cases} \tag{2.71}$$

- *if $j > i$ and $i < m$ ($c_i < 0$), then*

$$J^{\psi}_{j\to i}(y) \le \begin{cases} \varphi^{\star}_i(y) C e^{-\eta^2/2\kappa}, & y \le c_i, \\ \varphi^{\star}_j(y) C e^{-\eta^2/2\kappa}, & c_i \le y \le -\eta, \\ \varphi^{\star}_j(y) C, & -\eta \le y \le \eta, \\ \varphi^{\star}_j(y) C e^{-D/\epsilon}, & \eta \le y. \end{cases} \tag{2.72}$$

- *if $j < i$ and $i > m$ ($c_i > 0$), then*

$$J^{\psi}_{j\to i}(y) \le \begin{cases} \varphi^{\star}_i(y) C e^{-D/\epsilon}, & y \le -\eta, \\ \varphi^{\star}_i(y) C, & -\eta \le y \le \eta, \\ \varphi^{\star}_i(y) C e^{-\eta^2/2\kappa}, & \eta \le y \le c_i, \\ \varphi^{\star}_j(y) C e^{-\eta^2/2\kappa}, & c_i \le y. \end{cases} \tag{2.73}$$

Proof of Lemma 2.24. The proof relies on the estimate on the rest integral

$$\int_x^{\infty} e^{-t^2}\, dt \le \frac{\sqrt{\pi}}{2} e^{-x^2}, \quad x > 0,$$

and on the total mass evaluation

$$\int_{-\infty}^{\infty} e^{-t^2}\, dt = \sqrt{\pi}.$$

Using the change of variable $t = (2\kappa)^{-1/2}\left(x \pm \frac{\kappa}{\epsilon}\Delta_{ij}\right)$, where, under the assumption $\kappa = o(\epsilon)$, the term κ/ϵ goes to 0, the considered expression rewrites also

$$P_{\pm}(\alpha,\beta) = \pi^{-1/2} \int_{(2\kappa)^{-1/2}\left(\alpha \pm \frac{\kappa}{\epsilon}\Delta_{ij}\right)}^{(2\kappa)^{-1/2}\left(\beta \pm \frac{\kappa}{\epsilon}\Delta_{ij}\right)} e^{-t^2}\, dt.$$

As ϵ and κ tend to 0, both bounds of integration tend to $\pm\infty$ according to the sign of α and β. □

We now prove Lemma 2.25, using the result of Lemma 2.24 with α and β equal to c_i and to y, or inversely.

Proof of Lemma 2.25. According to the relative sign of $\mu_i - \mu_j$ and $y - c_i$ (under the same principle as in Lemma 2.23), we treat the expression of $J^{\psi}_{j\to i}(y)$:

1. For $j > i$ and $c_i \le y$, one writes

$$J^{\psi}_{j\to i}(y) = \varphi^{\star}_j(y)\left|\int_{c_i}^{y} \frac{\varphi^{\star}_i(y)\varphi^{\star}_j(x)}{\varphi^{\star}_i(x)\varphi^{\star}_j(y)}\psi(x)\,dx\right| = \varphi^{\star}_j(y)\int_{c_i}^{y}\phi(x,y,\mu_i-\mu_j)\psi(x)\,dx,$$

where the integral factor is less than

$$(2\pi\kappa)^{-1/2}\int_{c_i}^{y}\exp\left(-\frac{1}{\epsilon}(\underline{\Lambda}_j-\overline{\Lambda}_i)(y-x)-\frac{1}{2\kappa}x^2\right)dx \le \exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(y-\frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right)P_-(c_i,y).$$

By Lemma 2.24, we therefore obtain the following cases:

a) if $0 < c_i \le y$, then the integral factor is less than

$$\frac{1}{2}\exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(y-\frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right)\exp\left(-\frac{1}{2\kappa}\left(c_i-\frac{\kappa}{\epsilon}\Delta_{ij}\right)^2\right),$$

therefore $J^{\psi}_{j\to i}(y)$ is less than $\varphi^{\star}_j(y)Ce^{-\eta^2/2\kappa}$.

b) if $c_i < 0 < \eta \le y$, then the integral factor is less than

$$\exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(y-\frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right),$$

therefore $J^{\psi}_{j\to i}(y)$ is less than $\varphi^{\star}_j(y)Ce^{-D/\epsilon}$.

c) if $c_i \le y < -\eta < 0$, then the integral factor is less than

$$\frac{1}{2}\exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(y-\frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right)\exp\left(-\frac{1}{2\kappa}\left(y-\frac{\kappa}{\epsilon}\Delta_{ij}\right)^2\right),$$

therefore $J^{\psi}_{j\to i}(y)$ is less than $\varphi^{\star}_j(y)Ce^{-\eta^2/2\kappa}$.

d) finally if $c_i < -\eta < y < \eta$, then both previous upper bounds for the integral factor do not go to 0, and therefore $J^{\psi}_{j\to i}(y)$ is only less than $\varphi^{\star}_j(y)C$.

2. For $i > j$ and $y \le c_i$, one uses the same expression as in previous case

$$J^{\psi}_{j\to i}(y) = \varphi^{\star}_j(y)\int_{y}^{c_i}\phi(x,y,\mu_i-\mu_j)\psi(x)\,dx,$$

where the integral factor is less than

$$(2\pi\kappa)^{-1/2}\int_{y}^{c_i}\exp\left(-\frac{1}{\epsilon}(\underline{\Lambda}_i-\overline{\Lambda}_j)(x-y)-\frac{1}{2\kappa}x^2\right)dx \le \exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(-y-\frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right)P_+(y,c_i).$$

Thus by Lemma 2.24, we obtain the following cases:

a) If $y \le c_i < 0$, then the integral factor is less than

$$\frac{1}{2}\exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(-y-\frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right)\exp\left(-\frac{1}{2\kappa}\left(c_i+\frac{\kappa}{\epsilon}\Delta_{ij}\right)^2\right),$$

therefore $J^{\psi}_{j\to i}(y)$ is less than $\varphi_j^\star(y)Ce^{-\eta^2/2\kappa}$.

b) If $y \le \eta < 0 < c_i$, then the integral factor is less than

$$\exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(-y-\frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right),$$

therefore $J^{\psi}_{j\to i}(y)$ is less than $\varphi_j^\star(y)Ce^{-D/\epsilon}$.

c) If $0 < \eta \le y \le c_i$, then the integral factor is less than

$$\frac{1}{2}\exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(-y-\frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right)\exp\left(-\frac{1}{2\kappa}\left(y+\frac{\kappa}{\epsilon}\Delta_{ij}\right)^2\right),$$

therefore $J^{\psi}_{j\to i}(y)$ is less than $\varphi_j^\star(y)Ce^{-\eta^2/2\kappa}$.

d) finally if $-\eta < y < \eta < c_i$, then both previous upper bounds for the integral factor do not go to 0, and therefore $J^{\psi}_{j\to i}(y)$ is only less than $\varphi_j^\star(y)C$.

3. For $j > i$ and $y \le c_i$, one writes

$$\begin{aligned} J^{\psi}_{j\to i}(y) &= \varphi_i^\star(y)\frac{\varphi_j^\star(c_i)}{\varphi_i^\star(c_i)}\left|\int_y^{c_i}\frac{\varphi_i^\star(c_i)\varphi_j^\star(x)}{\varphi_i^\star(x)\varphi_j^\star(c_i)}\psi(x)\,dx\right| \\ &= \varphi_i^\star(y)\frac{\varphi_j^\star(c_i)}{\varphi_i^\star(c_i)}\int_y^{c_i}\phi(c_i,x,\mu_j-\mu_i)\psi(x)\,dx. \end{aligned}$$

On the one hand, the coefficient $\varphi_j^\star(c_i)/\varphi_i^\star(c_i)$ has already been treated during the proof of Lemma 2.14 and satisfies a uniform estimate of the form

$$\frac{\varphi_j^\star(c_i)}{\varphi_i^\star(c_i)} \le Ce^{-D/\epsilon}.$$

On the other hand, the integral part $\int_y^{c_i}\phi(c_i,x,\mu_j-\mu_i)\psi(x)\,dx$ is less than

$$\begin{aligned} &(2\pi\kappa)^{-1/2}\int_y^{c_i}\exp\left(-\frac{1}{\epsilon}(\underline{\Lambda}_j-\overline{\Lambda}_i)(c_i-x)-\frac{1}{2\kappa}x^2\right)dx \\ &\le \exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(c_i-\frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right)P_-(y,c_i) \end{aligned}$$

Thus by Lemma 2.24, we obtain the following cases:

a) If $y \le c_i < 0$, we get for the integral part the upper bound

$$\frac{1}{2}\exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(c_i - \frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right)\exp\left(-\frac{1}{2\kappa}\left(c_i - \frac{\kappa}{\epsilon}\Delta_{ij}\right)^2\right),$$

therefore $J^{\psi}_{j\to i}(y)$ is less than $\varphi_i^\star(y)Ce^{-\eta^2/2\kappa}$.

b) If $y \le -\eta < 0 < c_i$, we get for the integral part the upper bound

$$\exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(c_i - \frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right),$$

therefore $J^{\psi}_{j\to i}(y)$ is less than $\varphi_i^\star(y)Ce^{-D/\epsilon}$.

c) If $0 < \eta \le y \le c_i$, we get for the integral part the upper bound

$$\frac{1}{2}\exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(c_i - \frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right)\exp\left(-\frac{1}{2\kappa}\left(y - \frac{\kappa}{\epsilon}\Delta_{ij}\right)^2\right),$$

therefore $J^{\psi}_{j\to i}(y)$ is less than $\varphi_i^\star(y)Ce^{-\eta^2/2\kappa}$.

d) Finally, if $-\eta < y < \eta < c_i$, then both previous upper bounds for the integral factor do not go to 0, and therefore $J^{\psi}_{j\to i}(y)$ is only less than $\varphi_i^\star(y)Ce^{-D/\epsilon}$.

4. For $i > j$ and $c_i \le y$, one uses the same expression as in previous case

$$J^{\psi}_{j\to i}(y) = \varphi_i^\star(y)\frac{\varphi_j^\star(c_i)}{\varphi_i^\star(c_i)}\int_{c_i}^{y}\phi(c_i, x, \mu_j - \mu_i)\psi(x)\,dx.$$

By the same kind of computation as in previous cases, the integral part $\int_{c_i}^{y}\phi(c_i, x, \mu_j - \mu_i)\psi(x)\,dx$ is less than

$$\begin{aligned}&(2\pi\kappa)^{-1/2}\int_{c_i}^{y}\exp\left(-\frac{1}{\epsilon}(\underline{\Lambda}_i - \overline{\Lambda}_j)(x - c_i) - \frac{1}{2\kappa}x^2\right)dx\\ &\le \exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(-c_i - \frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right)P_+(c_i, y)\end{aligned}$$

Thus by Lemma 2.24, we obtain the following cases:

a) If $0 < c_i \le y$, we get for the integral part the upper bound

$$\frac{1}{2}\exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(-c_i - \frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right)\exp\left(-\frac{1}{2\kappa}\left(c_i + \frac{\kappa}{\epsilon}\Delta_{ij}\right)^2\right),$$

therefore $J^{\psi}_{j\to i}(y)$ is less than $\varphi_i^\star(y)Ce^{-\eta^2/2\kappa}$.

b) If $c_i < 0 < \eta \le y$, we get for the integral part the upper bound

$$\varphi_i^\star(y)\exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(-c_i - \frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right),$$

therefore $J^{\psi}_{j\to i}(y)$ is less than $\varphi^{\star}_i(y)Ce^{-D/\epsilon}$.

c) If $c_i \le y \le -\eta < 0$, we get for the integral part the upper bound

$$\frac{1}{2}\exp\left(-\frac{\Delta_{ij}}{\epsilon}\left(-c_i-\frac{\kappa}{2\epsilon}\Delta_{ij}\right)\right)\exp\left(-\frac{1}{2\kappa}\left(y+\frac{\kappa}{\epsilon}\Delta_{ij}\right)^2\right),$$

therefore $J^{\psi}_{j\to i}(y)$ is less than $\varphi^{\star}_i(y)Ce^{-\eta^2/2\kappa}$.

d) Finally, if $-\eta < y < \eta < c_i$, then both previous upper bounds for the integral factor do not go to 0, and therefore $J^{\psi}_{j\to i}(y)$ is only less than $\varphi^{\star}_i(y)Ce^{-D/\epsilon}$.

□

2.5.3 Influence of the resonant wave on non-resonant waves

If $j = m$ and $i \neq j$, we first observe that the behaviour of $J^{\psi}_{m\to i}(y)$, has three "reasons" to be large (through the quantity $\varphi^{\star}_m\psi/\varphi^{\star}_i$). This explains the particular weakness of the following estimate, especially for $0 \in [y, c_i)$, however sufficient to get the general expected result.

Lemma 2.26. *Given $0 < \eta \le \min(-\underline{\Lambda}_m, \overline{\Lambda}_m)$, there exist $C, D > 0$ independent of ϵ and of κ such that*

- *if $i > m$, then*

$$J^{\psi}_{m\to i}(y) \le \begin{cases} \varphi^{\star}_i(y)\,Ce^{-D/\epsilon}e^{-\eta^2/2\kappa}, & y \ge c_i, \\ \varphi^{\star}_m(y)\,C\epsilon e^{-\eta^2/2\kappa}, & \eta \le y \le c_i, \\ \varphi^{\star}_m(y), & y \le \eta. \end{cases} \tag{2.74}$$

- *if $i < m$, then*

$$J^{\psi}_{m\to i}(y) \le \begin{cases} \varphi^{\star}_i(y)\,Ce^{-D/\epsilon}e^{-\eta^2/2\kappa}, & y \le c_i, \\ \varphi^{\star}_m(y)\,C\epsilon e^{-\eta^2/2\kappa}, & c_i \le y \le -\eta, \\ \varphi^{\star}_m(y), & -\eta \le y. \end{cases} \tag{2.75}$$

Proof. We assume $i > m$, the case $i < m$ using the same line of argument. Using L^{∞} estimate for ψ and our knowledge on linear interaction coefficients between $\varphi^{\star}_m$ and $\varphi^{\star}_i$ (see Lemma 2.23 just above), we get for $y \ge c_i$,

$$\begin{aligned} J^{\psi}_{m\to i}(y) &\le \psi(\eta)\left|\varphi^{\star}_i(y)\int_{c_i}^{y}\frac{\varphi^{\star}_m(x)}{\varphi^{\star}_i(x)}\,dx\right| \\ &\le \frac{e^{-\eta^2/2\kappa}}{\sqrt{2\pi\kappa}}\varphi^{\star}_i(y)Ce^{-D/\epsilon}, \end{aligned}$$

and for $\eta \le y \le c_i$

$$\begin{aligned} J^{\psi}_{m\to i}(y) &\le \psi(\eta)\left|\varphi^{\star}_i(y)\int_{c_i}^{y}\frac{\varphi^{\star}_m(x)}{\varphi^{\star}_i(x)}\,dx\right| \\ &\le \frac{e^{-\eta^2/2\kappa}}{\sqrt{2\pi\kappa}}\varphi^{\star}_m(y)\frac{\epsilon}{\overline{\Lambda}_i-\underline{\Lambda}_m}. \end{aligned}$$

On the other hand, for $y \le \eta$ we use the L^1-character of ψ and the negativity of $(\lambda_m - \lambda_i)$:

$$\begin{aligned} J^{\psi}_{m\to i}(y) &\le \varphi^{\star}_m(y) \int_y^{c_i} \frac{\varphi^{\star}_m(x)\varphi^{\star}_i(y)}{\varphi^{\star}_i(x)\varphi^{\star}_m(y)}\psi(x)\,dx \\ &\le \varphi^{\star}_m(y) \int_y^{c_i} \exp\left(\int_y^x (\lambda_m - \lambda_i)/\epsilon\right)\psi(x)\,dx \le \varphi^{\star}_m(y)\int_y^{c_i}\psi(x)\,dx \le \varphi^{\star}_m(y). \end{aligned}$$

□

Remark 2.4. This result is though to be optimal in the case where $\varphi^{\star}_m(y)$ is an upper bound of $J^{\psi}_{m\to i}(y)$, i.e. when the integration interval (c_i, y) contains or approaches the interface 0.

2.5.4 Influence of a non-resonant wave on the resonant wave

If $i = m$ and $j \neq i$, then the coefficient $J^{\psi}_{j\to m}(y)$ satisfies following estimates

Lemma 2.27. *Given $0 < \eta \le \min(-\underline{\Lambda}_m, \overline{\Lambda}_m)$, there exist $C, D > 0$ independent of ϵ and of κ such that*

- *if $j > m$, then*

$$J^{\psi}_{j\to m}(y) \le \begin{cases} \varphi^{\star}_m(y)\,Ce^{-D/\epsilon}e^{-\eta^2/2\kappa} + \varphi^{\star}_m(y)\,Ce^{-D/\epsilon} & y \le -\eta, \\ \varphi^{\star}_m(y)\,Ce^{-D/\epsilon} & -\eta \le y \le \eta, \\ \varphi^{\star}_j(y)\,C\epsilon\, e^{-\eta^2/2\kappa} + \varphi^{\star}_m(y)\,Ce^{-D/\epsilon} & y \ge \eta. \end{cases} \tag{2.76}$$

- *if $j < m$, then*

$$J^{\psi}_{j\to m}(y) \le \begin{cases} \varphi^{\star}_j(y)\,C\epsilon\, e^{-\eta^2/2\kappa} + \varphi^{\star}_m(y)\,Ce^{-D/\epsilon} & y \le -\eta, \\ \varphi^{\star}_m(y)\,Ce^{-D/\epsilon} & -\eta \le y \le \eta, \\ \varphi^{\star}_m(y)\,Ce^{-D/\epsilon}e^{-\eta^2/2\kappa} + \varphi^{\star}_m(y)\,Ce^{-D/\epsilon} & y \ge \eta. \end{cases} \tag{2.77}$$

Proof. We assume $j > m$, the case $j < m$ using the same line of argument. Suppose first $-\eta \le y \le \eta$, using Lemma 2.13 we obtain

$$J^{\psi}_{j\to m}(y) \le \varphi^{\star}_m(y)\left\|\frac{\varphi^{\star}_j}{\varphi^{\star}_m}\right\|_{L^\infty([\underline{\Lambda}_m, \overline{\Lambda}_m])} \|\psi\|_{L^1} \le \varphi^{\star}_m(y) Ce^{-D/\epsilon}.$$

Now, for both cases $y \ge \eta$ and $y \le -\eta$, we cut the integral in two parts, one that contains the point 0 and uses the property $\psi \in L^1$ and the other one that uses the exponential decreasing

of ψ out of zero. Then for $y \geq \eta > 0$

$$\begin{aligned} J^{\psi}_{j\to m}(y) &\leq \varphi^{\star}_m(y) \int_0^{\eta} \frac{\varphi^{\star}_j(x)}{\varphi^{\star}_m(x)} \psi(x)\, dx + \varphi^{\star}_m(y) \int_{\eta}^{y} \frac{\varphi^{\star}_j(x)}{\varphi^{\star}_m(x)} \psi(x)\, dx \\ &\leq \varphi^{\star}_m(y) \left\| \frac{\varphi^{\star}_j}{\varphi^{\star}_m} \right\|_{L^{\infty}([\underline{\Lambda}_m, \overline{\Lambda}_m])} \|\psi\|_{L^1} + \psi(\eta) \varphi^{\star}_m(y) \int_0^{y} \frac{\varphi^{\star}_j(x)}{\varphi^{\star}_m(x)}\, dx \\ &\leq \varphi^{\star}_m(y) C e^{-D/\epsilon} + \frac{e^{-\eta^2/2\kappa}}{\sqrt{2\pi\kappa}} \varphi^{\star}_i(y) \frac{\epsilon}{\underline{\Lambda}_j - \overline{\Lambda}_m}. \end{aligned}$$

By the same way, for $y \leq -\eta < 0$

$$\begin{aligned} J^{\psi}_{j\to m}(y) &\leq \varphi^{\star}_m(y) \int_{-\eta}^{0} \frac{\varphi^{\star}_j(x)}{\varphi^{\star}_m(x)} \psi(x)\, dx + \varphi^{\star}_m(y) \int_{y}^{-\eta} \frac{\varphi^{\star}_j(x)}{\varphi^{\star}_m(x)} \psi(x)\, dx \\ &\leq \varphi^{\star}_m(y) \left\| \frac{\varphi^{\star}_j}{\varphi^{\star}_m} \right\|_{L^{\infty}([\underline{\Lambda}_m, \overline{\Lambda}_m])} \|\psi\|_{L^1} + \psi(-\eta) \varphi^{\star}_m(y) \int_{y}^{0} \frac{\varphi^{\star}_j(x)}{\varphi^{\star}_m(x)}\, dx \\ &\leq \varphi^{\star}_m(y) C e^{-D/\epsilon} + \frac{e^{-\eta^2/2\kappa}}{\sqrt{2\pi\kappa}} \varphi^{\star}_m(y) C e^{-D/\epsilon}. \end{aligned}$$

□

3 Resonant interfaces with internal structure

Dans le chapitre précédent, un système EDP augmenté a été introduit afin de coupler ensemble deux équations hyperboliques non-linéaires. Cette formulation nous a permis, au travers d'une apprximation visqueuse auto-semblable à la Dafermos, d'obtenir l'existence de solutions auto-semblables pour le problème de Riemann couplé. Dans le chapitre présent, nous poursuivons cette analyse et explorons la structure interne de l'interface et récupérons un critère de sélection associé au mécanisme de régularisation. Nous concluons ensuite par des exemples illustrant la possible non-unicité des solutions, malgré leur apparente simplicité.

Le travail présenté dans ce chapitre a été réalisé en collaboration avec Frédéric Coquel et Philippe G. LeFloch et fait l'objet de la publication en préparation [33].

Coupling nonlinear hyperbolic equations (II). Resonant interfaces with internal structure

Abstract

In the first paper of this series, an augmented PDE system was introduced in order to couple together two nonlinear hyperbolic equations. This formulation allowed us, via Dafermos self-similar viscosity method, to obtain the existence of self-similar solutions to the coupled Riemann problem. In the present paper, we continue this analysis and investigate the internal structure of the interface and recover selection criteria associated with the regularization mechanism. We then conclude by given evidence that solutions can be non-unique, even in some apparently simple examples.

3.1 Introduction

This paper deals with the analysis of the asymptotical behaviour of the Riemann-Dafermos approximations of the solutions for the coupling problem of two different conservation laws, each one being posed on a fixed half-space, with respective fluxes f_- for $x<0$ and f_+ on $x>0$. The equations in consideration are the following

$$\begin{aligned} &\partial_t u + \partial_x f_-(u) = 0, \quad x<0,\ t>0,\\ &\partial_t u + \partial_x f_+(u) = 0, \quad x>0,\ t>0, \end{aligned} \tag{3.1}$$

together with a coupling condition that links the traces of u on both sides of the interface $x=0$. Here we consider as coupling condition the continuity of u

$$u(t,0^-) = u(t,0^+). \tag{3.2}$$

This problem exhibits difficulties such as resonance at the interface with the non-uniqueness of the solutions as a consequence. Therefore, as studied in a previous paper by Boutin *et al.* [32], we propose to investigate the solution of this problem through its reformulation under a nonconservative extended PDE system, that is moreover nonstrictly hyperbolic, say

$$\begin{aligned} \partial_t u + \left(\frac{1-v}{2} f_-'(u) + \frac{1+v}{2} f_+'(u)\right)\partial_x u = 0,\\ \partial_t v = 0. \end{aligned} \tag{3.3}$$

We pay attention to the problems with Riemann initial conditions

$$u(x,0) = \begin{cases} u_L, & x<0,\\ u_R, & x>0, \end{cases} \qquad v(x,0) = \begin{cases} -1, & x<0,\\ 1, & x>0. \end{cases} \tag{3.4}$$

The Dafermos approximation of the latter consists in investigating the selfsimilar solutions $(u^\epsilon, v^\epsilon) = (u^\epsilon(\xi), v^\epsilon(\xi))$ (where $\xi = x/t$) of the vanishing viscous formulation of (3.3), namely

$$\begin{aligned} -\xi u^\epsilon_\xi + \Big(\frac{1-v^\epsilon}{2} f'_-(u^\epsilon) + \frac{1+v^\epsilon}{2} f'_+(u^\epsilon)\Big) u^\epsilon_\xi &= \epsilon u^\epsilon_{\xi\xi}, \\ -\xi v^\epsilon_\xi &= \epsilon^2 v^\epsilon_{\xi\xi}. \end{aligned} \tag{3.5}$$

This system of ODEs is supplemented with boundary conditions written at $\xi = \pm L$ where L denotes a finite upper bound for the quantity $|f'_-| + |f'_+|$ over the set $[u_L, u_R]$ (i.e. represents the larger speed in the hyperbolic problem in consideration)

$$\begin{aligned} u^\epsilon(-L) &= u_L, \qquad v^\epsilon(-L) = -1, \\ u^\epsilon(L) &= u_R, \qquad v^\epsilon(L) = 1. \end{aligned} \tag{3.6}$$

In our last paper, we obtained the existence of at least one solution u^ϵ for this problem, that converges, as ϵ goes to 0, to $u \in L^\infty \cap BV$ entropy solution for each half-problem (3.1) (in the sense that it satisfies on each open half-space natural entropy inequalities for the corresponding equation). In fact, this result has been successfully achieved for a large class of systems of conservation laws, but we restrict here the study to scalar conservation laws and to the particular form of (3.3).

Once we get this existence result and this minimal consistency with both conservation laws, it remains to characterize the behaviour of the limiting solution at the interface $\xi = 0$. We expect first a consistency with the coupling condition (3.2), in a sense to be precised latter, and then a selection of solutions due to the viscous limiting process. To this end, we proceed to a *blow-up* of the neighborhood of the origin using the change of variable $\xi = \epsilon y$ and setting as new unknowns

$$U^\epsilon(y) = u^\epsilon(\epsilon y), \qquad V^\epsilon(y) = v^\epsilon(\epsilon y). \tag{3.7}$$

This method is thought to reveal the behaviour of u^ϵ at the neighborhood of the interface, that is to say the possible boundary layer that disappear as ϵ tends to 0, but that contains useful information characterizing the link between traces of u at $\xi = 0^-$ and at $\xi = 0^+$.

Remark 3.1. Actually, the choice of a vanishing coefficient ϵ^2 in the viscosity for v^ϵ gives the following limiting profile $V(y)$ for $V^\epsilon(y)$ as ϵ goes to 0

$$V(y) := -1 + 2 \int_{-\infty}^{y} e^{-s^2/2}\, ds \Big/ \int_{-\infty}^{+\infty} e^{-s^2/2}\, ds\,. \tag{3.8}$$

This function $V \in C^\infty(\mathbb{R})$ maps monotonically $\mathbb{R}$ onto $(-1,1)$. This important fact will provide smoothness on U^ϵ and on its limit U. This second order viscosity coefficient is to relate with the linear degenerated character of the field associated to the v component in the hyperbolic system (3.3).

The main result of this paper is the

Theorem 3.1. *Let f_- and f_+ be two functions in $C^1(\mathbb{R})$ such that $f'_\pm$ are Lipschitz continuous, let $u^\epsilon \in L^\infty(\mathbb{R}) \cap BV(\mathbb{R})$ be solution of (3.5)-(3.6) and U^ϵ given by (3.7), then the following facts are satified.*

1. *There exists $U \in C^2(\mathbb{R})$ limit of the sequence $(U^\epsilon)_{\epsilon>0}$ as $\epsilon > 0$. It satisfies, in the sense of distributions, the viscous profile equation*

$$\left(\frac{1-V}{2} f'_-(U) + \frac{1+V}{2} f'_+(U)\right) U_y = U_{yy}, \tag{3.9}$$

and admits limits at infinities $U_{-\infty}$ and $U_{+\infty}$, that satifies

$$\min(u_L, u_R) \le U_{\pm\infty} \le \max(u_L, u_R). \tag{3.10}$$

2. *Moreover, q_- and q_+ being entropy fluxes related respectively to f_- and f_+ for a strictly convex entropy, and $u \in L^\infty(\mathbb{R}) \cap BV(\mathbb{R})$ being the limiting solution of (3.5)-(3.6), we have*

$$\begin{aligned} &f_-(u(0^-)) = f_-(U_{-\infty}), \quad q_-(u(0^-)) \ge q_-(U_{-\infty}), \\ &f_+(U_{+\infty}) = f_+(u(0^+)), \quad q_+(U_{+\infty}) \ge q_+(u(0^+)). \end{aligned} \tag{3.11}$$

3. *If $f'_-(U_{-\infty}) < 0$ or if $f'_+(U_{+\infty}) > 0$, then the internal boundary layer U is constant and thus $U_{-\infty} = U_{+\infty}$.*

Observe importantly we may not get $u(0^-) = u(0^+)$ (or $f_-(u(0^-)) = f_+(u(0^+))$), if U is not a trivial connection from $U_{-\infty}$ to $U_{+\infty}$. Furthermore, the equations (3.11) give the structure of the potential boundary layer on both side of the interface : a zero-speed entropy shock either for the left-, or for the right-, or for both problems at the interface.

The paper organizes as follows.
In a first section, we will obtain Lemma 3.2 and Lemma 3.3 useful to get then the Theorem 3.1 whose proof will be detailled in §3.2.4. In a second section, we will apply these results to the coupling of two Burgers fluxes with different sonic points. We do not generally obtain uniqueness of the solution for a given Riemann problem, but the lack of uniqueness is strongly reduced by this Dafermos analysis that provide therefore a new selection criterion. In order to finally describe the diagrams of solutions in the plane of Riemann data (u_L, u_R), we take advantage of a Laplace analysis, that will filter out most of solutions.

3.2 Construction of the viscous interface profile

3.2.1 Viscous profile equation

Proof of point 1 of Theorem 3.1. As $(u^\epsilon)_{\epsilon>0}$, the sequence $(U^\epsilon)_{\epsilon>0}$ belongs uniformly to $L^\infty(\mathbb{R}) \cap BV(\mathbb{R})$. Each U^ϵ takes value in the (non-oriented) interval $[u_L, u_R]$, is monotone over $\mathbb{R}$ and, due to (3.6), satisfies the boundary conditions $U^\epsilon(-L/\epsilon) = u_L$, $U^\epsilon(L/\epsilon) = u_R$. Let U be the L^1_{loc} limit of (a subsequence) of $(U^\epsilon)_{\epsilon>0}$. The function U is bounded and monotone and

therefore admits limits as ξ tends to $\pm\infty$ denoted in the sequel respectively $U_{\pm\infty}$. As a direct consequence, these limits should satisfy the inequalities (3.10) too. But we point out that they *a priori* differ from u_L and u_R (both limiting processes $\epsilon \to 0$ and $\xi \to \pm\infty$ do not commute).

Moreover, using some structural properties obtained is the previous paper [32], we get $\|u^\epsilon_\xi\|_\infty \leq C/\epsilon$ where C is independant of ϵ. It appears therefore that the derivative U^ϵ_y belongs to L^∞ uniformly in ϵ.

We now write (3.5) at the point $\xi = \epsilon y$, we get

$$-\epsilon y U^\epsilon_y(y) + \left(\frac{1-V^\epsilon(y)}{2} f'_-(U^\epsilon(y)) + \frac{1+V^\epsilon(y)}{2} f'_+(U^\epsilon(y))\right) U^\epsilon_y(y) = U^\epsilon_{yy}(y).$$

Let $\phi \in C^\infty_0(\mathbb{R})$ be a test-function with compact support in $\mathbb{R}$ and consider the variationnal formulation of this ordinary differential equation

$$-\epsilon \int_{\mathbb{R}} y U^\epsilon_y \phi \, dy + \int_{\mathbb{R}} \left(\frac{1-V^\epsilon(y)}{2} f'_-(U^\epsilon(y)) + \frac{1+V^\epsilon(y)}{2} f'_+(U^\epsilon(y))\right) U^\epsilon_y \phi \, dy = -\int_{\mathbb{R}} U^\epsilon_y \phi_y \, dy.$$

The three following contributions are all bounded

$$\left|\int_{\mathbb{R}} y U^\epsilon_y \phi \, dy\right| \leq \|y\phi\|_\infty TV(U^\epsilon),$$
$$\left|\int_{\mathbb{R}} \left(\frac{1-V^\epsilon(y)}{2} f'_-(U^\epsilon(y)) + \frac{1+V^\epsilon(y)}{2} f'_+(U^\epsilon(y))\right) U^\epsilon_y \phi \, dy\right| \leq (\|f'_-\|_\infty + \|f'_+\|_\infty)\|\phi\|_\infty TV(U^\epsilon),$$
$$\left|\int_{\mathbb{R}} U^\epsilon_y \phi_y \, dy\right| \leq \|\phi_y\|_\infty TV(U^\epsilon).$$

Thus, passing to the the limit, U satisfies the weak formulation of (3.9)

$$\int_{\mathbb{R}} \left(\frac{1-V}{2} f'_-(U) + \frac{1+V}{2} f'_+(U)\right) U_y \phi \, dy = -\int_{\mathbb{R}} U_y \phi_y \, dy.$$

This second order equation rewrites also under the following first order form

$$\begin{aligned} U_y &= W, \\ W_y &= \left(\frac{1-V}{2} f'_-(U) + \frac{1+V}{2} f'_+(U)\right) W, \end{aligned} \tag{3.12}$$

in other words, satisfies a system of ordinary differential equations with smooth coefficients (remember $V \in C^\infty(\mathbb{R})$), we obtain *a posteriori* the required smoothness : $U \in C^2(\mathbb{R})$. □

3.2.2 Steady boundary layers

The equation (3.9) is not sufficient to get information at infinities $y \pm \infty$ during the limiting process $\epsilon \to 0$. To this end we will need the following

Lemma 3.2. *Under the assumptions of Theorem 3.1, U satifies the equation*

$$\begin{aligned} U_y &= \Big(\frac{1-V}{2}f_-(U)+\frac{1+V}{2}f_+(U)\Big)-f_+(u(0^+))-\frac{1}{2}\int_\infty^y \Big(f_+(U(s))-f_-(U(s))\Big)V_y(s)\,ds,\\ U_y &= \Big(\frac{1-V}{2}f_-(U)+\frac{1+V}{2}f_+(U)\Big)-f_-(u(0^-))-\frac{1}{2}\int_{-\infty}^y \Big(f_+(U(s))-f_-(U(s))\Big)V_y(s)\,ds. \end{aligned} \tag{3.13}$$

Proof. We get more information on the solution by integrating (3.5) for ξ between ϵy and $a>0$,

$$\epsilon\int_a^{\epsilon y} u^\epsilon_{\xi\xi} = -\int_a^{\epsilon y}\xi u^\epsilon_\xi + \int_a^{\epsilon y}\Big(\frac{1-v^\epsilon}{2}f'_-(u^\epsilon)+\frac{1+v^\epsilon}{2}f'_+(u^\epsilon)\Big)u^\epsilon_\xi.$$

We thus get

$$\begin{aligned} \epsilon u^\epsilon_\xi(\epsilon y)-\epsilon u^\epsilon_\xi(a) = -\epsilon y u^\epsilon(\epsilon y)+\Big(\frac{1-v^\epsilon}{2}f_-(u^\epsilon)+\frac{1+v^\epsilon}{2}f_+(u^\epsilon)\Big)(\epsilon y)-\frac{1}{2}\int_a^{\epsilon y}\Big(f_+(u^\epsilon)-f_-(u^\epsilon)\Big)v^\epsilon_\xi\\ + a\,u^\epsilon(a)+\int_a^{\epsilon y}u^\epsilon-\Big(\frac{1-v^\epsilon}{2}f_-(u^\epsilon)+\frac{1+v^\epsilon}{2}f_+(u^\epsilon)\Big)(a). \end{aligned}$$

Moving a asymptoticaly on 0, as ϵ tends to 0, the quantity $u^\epsilon_\xi(a)$ could diverge, so fix $\delta>0$ and integrate previous equation relative to a on $[0,\delta]$ to get a mean expression. It becomes

$$\begin{aligned} \epsilon u^\epsilon_\xi(\epsilon y) &= \frac{\epsilon}{\delta}\int_0^\delta u^\epsilon_\xi(a)da-\epsilon y u^\epsilon(\epsilon y)+\Big(\frac{1-v^\epsilon}{2}f_-(u^\epsilon)+\frac{1+v^\epsilon}{2}f_+(u^\epsilon)\Big)(\epsilon y)\\ &-\frac{1}{2\delta}\int_0^\delta\int_a^{\epsilon y}\Big(f_+(u^\epsilon)-f_-(u^\epsilon)\Big)v^\epsilon_\xi\,ds\,da+\frac{1}{\delta}\int_0^\delta\Big(a\,u^\epsilon(a)+\int_a^{\epsilon y}u^\epsilon\Big)da\\ &-\frac{1}{\delta}\int_0^\delta\Big(\frac{1-v^\epsilon}{2}f_-(u^\epsilon)+\frac{1+v^\epsilon}{2}f_+(u^\epsilon)\Big)(a)\,da, \end{aligned}$$

that is, rewritten in the profile variables U^ϵ, V^ϵ,

$$\begin{aligned} U^\epsilon_y(y) &= \frac{\epsilon}{\delta}\int_0^\delta u^\epsilon_\xi(a)\,da-\epsilon y\,U^\epsilon(y)+\Big(\frac{1-V^\epsilon(y)}{2}f_-(U^\epsilon(y))+\frac{1+V^\epsilon(y)}{2}f_+(U^\epsilon(y))\Big)\\ &-\frac{1}{2\delta}\int_0^\delta\int_{a/\epsilon}^y\Big(f_+(U^\epsilon)-f_-(U^\epsilon)\Big)V^\epsilon_y\,ds\,da+\frac{1}{\delta}\int_0^\delta\Big(a\,u^\epsilon(a)+\int_a^{\epsilon y}u^\epsilon\,ds\Big)da\\ &-\frac{1}{\delta}\int_0^\delta\Big(\frac{1-v^\epsilon}{2}f_-(u^\epsilon)+\frac{1+v^\epsilon}{2}f_+(u^\epsilon)\Big)(a)\,da \end{aligned}$$

Thus we finally get as ϵ goes to 0

$$\begin{aligned} U_y(y) = \Big(\frac{1-V(y)}{2}f_-(U(y))+\frac{1+V(y)}{2}f_+(U(y))\Big)-\frac{1}{2}\int_\infty^y\Big(f_+(U)-f_-(U)\Big)V_y\,ds\\ -\frac{1}{\delta}\int_0^\delta f_+(u(a))\,da+\frac{1}{\delta}\int_0^\delta au(a)\,da+\frac{1}{\delta}\int_0^\delta\int_a^0 u(s)\,dsda. \end{aligned}$$

Now, let δ tends to 0, the first equation of (3.13) follows. The same work for $\delta<0$ give similarly the second equation of (3.13). □

3.2.3 Entropy boundary layer

Lemma 3.3. *Under the assumptions of Theorem 3.1, $\eta \in C^1(\mathbb{R})$ being a stricly convex entropy function associated to the entropy fluxes q_- and q_+ (respectively for the fluxes f_- and f_+), then U satifies the following entropy inequalities on each half-space*

$$\begin{aligned}
&\eta'(U)U_y - \Big(\frac{1-V}{2}q_-(U) + \frac{1+V}{2}q_+(U)\Big) + q_+(u(0^+)) + \frac{1}{2}\int_{+\infty}^{y}\Big(q_+(U(s)) - q_-(U(s))\Big)V_y(s)\,ds \le 0,\\
&\eta'(U)U_y - \Big(\frac{1-V}{2}q_-(U) + \frac{1+V}{2}q_+(U)\Big) + q_-(u(0^-)) + \frac{1}{2}\int_{-\infty}^{y}\Big(q_+(U(s)) - q_-(U(s))\Big)V_y(s)\,ds \ge 0.
\end{aligned} \tag{3.14}$$

Proof. We first consider the following expression

$$\begin{aligned}
\epsilon(\eta(u^\epsilon))_{\xi\xi} &= \epsilon u^\epsilon_{\xi\xi}\eta'(u^\epsilon) + \epsilon(u^\epsilon_\xi)^2\eta''(u^\epsilon)\\
&= \Big(-\xi + \frac{1-v^\epsilon}{2}f'_-(u^\epsilon) + \frac{1+v^\epsilon}{2}f'_+(u^\epsilon)\Big)\eta'(u^\epsilon)u^\epsilon_\xi + \epsilon\eta''(u^\epsilon)(u^\epsilon_\xi)^2.
\end{aligned}$$

Let be $a > 0$ and $y \in \mathbb{R}$, integrating the latter between a and ϵy, we get

$$\begin{aligned}
(\epsilon\eta(u^\epsilon))_\xi(\epsilon y) - \epsilon(\eta(u^\epsilon))_\xi(a) &= \Big(\frac{1-v^\epsilon}{2}q_-(u^\epsilon) + \frac{1+v^\epsilon}{2}q_+(u^\epsilon)\Big)(\epsilon y)\\
&\quad - \Big(\frac{1-v^\epsilon}{2}q_-(u^\epsilon) + \frac{1+v^\epsilon}{2}q_+(u^\epsilon)\Big)(a) - \frac{1}{2}\int_a^{\epsilon y}\Big(q_+(u^\epsilon) - q_-(u^\epsilon)\Big)v^\epsilon_\xi\,ds\\
&\quad - \int_a^{\epsilon y}\xi(\eta(u^\epsilon))_\xi\,d\xi + \int_a^{\epsilon y}\epsilon\eta''(u^\epsilon)(u^\epsilon_\xi)^2\,d\xi.
\end{aligned}$$

It reads also, using the blow-up functions U^ϵ and V^ϵ

$$\begin{aligned}
(U^\epsilon_y\eta'(U^\epsilon))(y) - \epsilon(\eta(u^\epsilon))_\xi(a) &= \Big(\frac{1-V^\epsilon}{2}q_-(U^\epsilon) + \frac{1+V^\epsilon}{2}q_+(U^\epsilon)\Big)(y)\\
&\quad - \Big(\frac{1-v^\epsilon}{2}q_-(u^\epsilon) + \frac{1+v^\epsilon}{2}q_+(u^\epsilon)\Big)(a) - \frac{1}{2}\int_{a/\epsilon}^{y}\Big(q_+(U^\epsilon) - q_-(U^\epsilon)\Big)V^\epsilon_y\,ds\\
&\quad - \int_a^{\epsilon y}\xi(\eta(u^\epsilon))_\xi\,d\xi + \int_{a/\epsilon}^{y}(U^\epsilon_y)^2\eta''(U^\epsilon)\,d\xi
\end{aligned}$$

Under the assumption $y < a/\epsilon$, we thus obtain

$$\begin{aligned}
&U^\epsilon_y(y)\eta'(U^\epsilon(y)) - \epsilon(\eta(u^\epsilon))_\xi(a) - \Big(\frac{1-V^\epsilon}{2}q_-(U^\epsilon) + \frac{1+V^\epsilon}{2}q_+(U^\epsilon)\Big)(y)\\
&+ \Big(\frac{1-v^\epsilon}{2}q_-(u^\epsilon) + \frac{1+v^\epsilon}{2}q_+(u^\epsilon)\Big)(a) + \frac{1}{2}\int_{a/\epsilon}^{y}\Big(q_+(U^\epsilon) - q_-(U^\epsilon)\Big)V^\epsilon_y\,ds + \int_a^{\epsilon y}\xi(\eta(u^\epsilon))_\xi\,d\xi \quad \le 0.
\end{aligned}$$

Let $\delta > 0$ be fixed, and let choose ϵ small enough to ensure $\delta > \epsilon y$. Then, averaging the

previous inequality for a over $[\epsilon y, \delta]$, it comes

$$U_y^\epsilon(y)\eta'(U^\epsilon(y)) - \frac{1}{\delta-\epsilon y}\int_{\epsilon y}^{\delta} \epsilon u_\xi^\epsilon(a)\eta'(u^\epsilon(a))\,da - \Big(\frac{1-V^\epsilon}{2}q_-(U^\epsilon) + \frac{1+V^\epsilon}{2}q_+(U^\epsilon)\Big)(y) \\ + \frac{1}{\delta-\epsilon y}\int_{\epsilon y}^{\delta}\Big(\frac{1-v^\epsilon}{2}q_-(u^\epsilon) + \frac{1-v^\epsilon}{2}q_+(u^\epsilon)\Big)(a)\,da \\ + \frac{1}{2(\delta-\epsilon y)}\int_{\epsilon y}^{\delta}\int_{a/\epsilon}^{y}\Big(q_+(U^\epsilon) - q_-(U^\epsilon)\Big)V_y^\epsilon\,ds\,da + \frac{1}{\delta-\epsilon y}\int_{\epsilon y}^{\delta}\int_a^{\epsilon y}\xi(\eta(u^\epsilon))_\xi\,d\xi\,da \le 0$$

Now, we let ϵ tends to 0, and thus obtain in the sense of distribution

$$U_y(y)\eta'(U(y)) - \Big(\frac{1-V}{2}q_-(U) + \frac{1+V}{2}q_+(U)\Big)(y) + \frac{1}{\delta}\int_0^{\delta} q_+(u)(a)\,da \\ + \frac{1}{2\delta}\int_0^{\delta}\int_{\infty}^{y}\Big(q_+(U) - q_-(U)\Big)V_y\,ds\,da + \frac{1}{\delta}\int_0^{\delta}\int_a^0 \xi(\eta(u))_\xi\,d\xi\,da \le 0$$

Letting δ tends to 0, we get the entropy inequality of (3.14) that concerns the left-space. By a similar way, using $\delta < 0$, we get the entropy inequality that concerns the right-space. □

3.2.4 Sticking of the traces of the solution to the boundary layer

Leaning on Lemmas 3.2 and 3.3, we are now in position to get the

Proof of point 2 of Theorem 3.1. Let be n an integer and let integrate the first equation of (3.13) on a unit interval in the neighborhood of $+\infty$, say $[n, n+1]$

$$U(n+1) - U(n) = \int_n^{n+1}\left(\frac{1-V(y)}{2}f_-(U(y)) + \frac{1+V(y)}{2}f_+(U(y))\right)dy - f_+(u(0^+)) \\ - \frac{1}{2}\int_n^{n+1}\int_{\infty}^{y}(f_+(U(x)) - f_-(U(x)))V_y(x)\,dx\,dy.$$

Rearranging terms, we thus obtain

$$\int_n^{n+1} f_+(U(y))\,dy \;-\; f_+(u(0^+)) = U(n+1) - U(n) - \int_n^{n+1}\frac{1-V(y)}{2}(f_-(U(y)) - f_+(U(y)))dy \\ + \frac{1}{2}\int_n^{n+1}\int_{\infty}^{y}(f_+(U(x)) - f_-(U(x)))(x)V_y(x)\,dx\,dy.$$

Each term of the right hand side tends to 0 as n tends to infinity, while the left hand side converges towards $f_+(U_{+\infty}) - f_+(u(0^+))$. Thus, the equality $f_+(U_{+\infty}) = f_+(u(0^+))$ follows. With similar arguments, we get from the first inequation of (3.14) the inequality $q_+(U_{+\infty}) \ge q_+(u(0^+))$.

For the same reason, using the second equation of (3.13) and the second inequation of (3.14) integrated on a unit interval in a neighborhood of $-\infty$, we obtain the equality $f_-(U_{-\infty}) = f_-(u(0^-))$ and $q_-(u(0^-)) \ge q_-(U_{-\infty})$. □

3.2.5 Stability properties

Proof of point 3 of Theorem 3.1. From the equation (3.9), we will deduce the global form of the boundary layer U. The equations (3.13) ensure U_y tends to 0 as y goes to $\pm\infty$. Using (3.12), one can solve implicitely U_y as

$$\begin{aligned} U_y(y) &= W(0)\exp\left[\int_0^y \left(\frac{1-V}{2}f'_-(U)+\frac{1+V}{2}f'_+(U)\right)ds\right] \\ &= W(0)\exp\left[\int_0^y f'_+(U)ds\right]\exp\left[\int_0^y \frac{1-V}{2}(f'_-(U)-f'_+(U))ds\right]. \end{aligned}$$

The argument of the second exponential term admits a finite limit as y goes to ∞, because $f'_-(U)-f'_+(U)\in L^\infty(\mathbb{R})$ while $\frac{1-V}{2}$ tends to 0 exponentially. The first exponential factor therefore leads the global quantity to 0. More precisely its behaviour depends closely of the sign of $f'_+(U_{+\infty})$. And we have the following cases.

- If $f'_+(U_{+\infty})<0$, then $\int_0^y f'_+(U)dx \sim y f'_+(U_{+\infty})$ tends to $-\infty$ and $\exp\left[\int_0^y f'_+(U)dx\right]\to 0$.
- If $f'_+(U_{+\infty})>0$, then $\int_0^y f'_+(U)dx \sim y f'_+(U_{+\infty})$ tends to $+\infty$ and $\exp\left[\int_0^y f'_+(U)dx\right]\to +\infty$, but $W=U_y$ should tend to 0: it is possible only if $W(0)$ and thus W equals identically 0, so that finally $U_{-\infty}=U_{+\infty}$.
- For the limiting case $f'_+(U_{+\infty})=0$ the conclusion is trickier without solving more precisely the viscous profile equation (3.9), and we do not deal with it in this paper.

We get similarly the expected result for $y=-\infty$, according to the sign of $f'_-(U_{-\infty})$. □

3.3 Example of the coupling of two Burgers equations

Let study solutions for the coupling of two Burgers' equations with quadratic fluxes with different sonic points 0 and c respectively, say $f_-(u)=u^2/2$ and $f_+(u)=(u-c)^2/2$.

3.3.1 Structure of Riemann-Dafermos solutions

The application of Theorem 3.1, in this case where the fluxes $f_\pm$ are both strictly convex functions, gives here the following informations

- On the one hand at $\xi=0^+$,
  - either $U_{+\infty}=u(0^+)$,
    - if $f'_+(u(0^+))>0$ then $U_{+\infty}=U_{-\infty}$,
    - if $f'_+(u(0^+))<0$ then $U_{+\infty}$ may differ from $U_{-\infty}$,
  - or $U_{+\infty}>u(0^+)$. Then the discontinuity $(U_{+\infty},u(0^+))$ is an entropy shock wave with zero-speed for the right-problem, and thus $U_{+\infty}=2c-u(0^+)$, with $u(0^+)<c$. Moreover, because $f'_+(U_{+\infty})>0$, we have $U_{-\infty}=U_{+\infty}=2c-u(0^+)$.

- On the other hand at $\xi = 0^-$,
  - either $U_{-\infty} = u(0^-)$,
    - if $f'_-(u(0^-)) < 0$ then $U_{+\infty} = U_{-\infty}$,
    - if $f'_-(u(0^-)) > 0$ then $U_{+\infty}$ may differ from $U_{-\infty}$,
  - or $U_{-\infty} < u(0^-)$. Then, the discontinuity $(u(0^-), U_{-\infty})$ is an entropy shock wave with zero-speed for the left-problem, and thus $U_{-\infty} = -u(0^-)$, with $u(0^-) > 0$. Moreover, because $f'_-(U_{-\infty}) < 0$, we have $U_{+\infty} = U_{-\infty} = -u(0^-)$.

We intend now to decline all possible Riemann-Dafermos solutions, in the plane of Riemann data (u_L, u_R).For $c > 0$, the solutions are represented on Figure. **??**, we obtain unicity of the solution in the whole plane, except in two domains where a shock wave for either the left- or the right-flux and a standing wave may appear. For $c < 0$, on Figure. **??**, the unicity is less common. However the worst domain (the central triangle) presents a set of at most four different solutions. Let now give some elements for the comprehension of two examples in this last case $c < 0$.

- In the shock-sector of the upper-left corner (with for example $(u_L, u_R) = (1, 0)$), the single possibility is a shock wave for the right-flux.
- In the sector just below (with for example $(u_L, u_R) = (-3c, 2c)$), this shock wave for the right-flux is not the unique possible solution. The standing wave at the interface is a solution too.

3.3.2 Laplace method

We recall the solution u^ϵ of (3.5)-(3.6) solves implicitely under the form

$$u^\epsilon(\xi) = u_L + (u_R - u_L)\frac{\int_{-L}^{\xi} e^{-\frac{h^\epsilon(s)}{\epsilon}} ds}{\int_{-L}^{L} e^{-\frac{h^\epsilon(s)}{\epsilon}} ds}, \tag{3.15}$$

where we set

$$h^\epsilon(\xi) = \int_0^\xi \Big(s - f'_-(u^\epsilon(s))\frac{1 - v^\epsilon(s)}{2} - f'_+(u^\epsilon(s))\frac{1 - v^\epsilon(s)}{2}\Big) ds. \tag{3.16}$$

Supplemented to the results of Theorem 3.1, this implicit representation formula permits to rule out some solutions, because of necessary monotonicity properties for example, but also using a more sophisticated but classical analysis: the Laplace method. We recall first the basic tool in

Proposition 3.4 (Laplace method). *Let $I = [a, b]$ be a closed interval of $\mathbb{R}$, and let $h \in C^2(I)$ be a real-valued function that admits a unique global minimum over I at $x = a$ and such that $h''(a) > 0$, then*

$$\int_a^b e^{-h(x)/\epsilon} dx \underset{\epsilon \to 0}{\sim} \sqrt{\frac{2\pi\epsilon}{h''(a)}} e^{-h(a)/\epsilon}.$$

In order to apply this last tool, it will be necessary to remove the dependence of h^ϵ in ϵ. Fortunately, we get the next lemma.

Lemma 3.5. *The function h^ϵ converges uniformly to h as ϵ goes to 0, over $(-L,L)$, defined by*

$$h(\xi) = \int_0^\xi \Big(s - f'_-(u(s))\chi_{\mathbb{R}_-}(s) - f'_+(u(s))\chi_{\mathbb{R}_+}(s)\Big)ds. \tag{3.17}$$

Proof. Let be $\xi \in [-L,L]$ and observe

$$\begin{aligned}
|h^\epsilon(\xi) - h(\xi)| &\leq \left|\int_0^\xi f'_-(u^\epsilon)\frac{1-v^\epsilon}{2} - f'_-(u)\chi_{\mathbb{R}_-}ds\right| + \left|\int_0^\xi f'_+(u^\epsilon)\frac{1+v^\epsilon}{2} - f'_+(u)\chi_{\mathbb{R}_+}ds\right| \\
&\leq \int_{-L}^L \Big|f'_-(u^\epsilon(s)) - f'_-(u(s))\Big|\chi_{\mathbb{R}_-}(s)ds + \int_{-L}^L \Big|f'_-(u^\epsilon(s))\Big|\left|\chi_{\mathbb{R}_-}(s) - \frac{1-v^\epsilon}{2}\right|ds \\
&+ \int_{-L}^L \Big|f'_+(u^\epsilon(s)) - f'_+(u(s))\Big|\chi_{\mathbb{R}_+}(s)ds + \int_{-L}^L \Big|f'_+(u^\epsilon(s))\Big|\left|\chi_{\mathbb{R}_+}(s) - \frac{1+v^\epsilon}{2}\right|ds \\
&\leq (\mathrm{Lip}(f'_-) + \mathrm{Lip}(f'_+))\int_{-L}^L |u^\epsilon - u|ds + (\|f'_-\|_\infty + \|f'_+\|_\infty)\int_{-L}^L \left|\chi_{\mathbb{R}_-}(s) - \frac{1-v^\epsilon}{2}\right|ds.
\end{aligned}$$

Finally, u^ϵ and v^ϵ converges in L^1_{loc} respectively to u and to the sign function, therefore this last quantity tends to 0. □

Double rarefaction solutions Let consider as possible limiting Riemann-Dafermos solution a double rarefaction, as represented on Figure 3.1a. It is composed of a rarefaction wave for the left flux f_- connecting the states u_L to an intermediate state denoted $u^\star$, followed by a constant plateau at $u = u^\star$ for $\xi \in [f'_-(u^\star), f'_+(u^\star)]$ and a rarefaction wave for the right flux f_+ connecting $u^\star$ to u_R. The function h is as represented on the Figure 3.1b and admits two minima plateau. A necessary condition during the limiting process in the Laplace analysis to get such a solution u, is that these two plateau should have the same height. In other terms the condition writes

$$h(f'_-(u^\star)) = h(f'_+(u^\star)), \text{ i.e. } u^\star = c/2. \tag{3.18}$$

Proof of (3.18). Such a double rarefaction solution is possible only if $u_R > u^\star > u_L$. We suppose by contradiction, that $\delta = h(f'_-(u^\star)) - h(f'_+(u^\star)) \neq 0$, and without loss of generality that $\delta > 0$. Because of the uniform convergence of h^ϵ towards h, for $\epsilon > 0$ small enough, we have $|h^\epsilon(s) - h(s)| \leq \delta/3$ for any $s \in [-L,L]$. Therefore, we get the inequality

$$e^{-\frac{2\delta}{3\epsilon}}\frac{\int_{-L}^0 e^{-\frac{h(s)}{\epsilon}}ds}{\int_{-L}^L e^{-\frac{h(s)}{\epsilon}}ds} \leq \frac{\int_{-L}^0 e^{-\frac{h^\epsilon(s)}{\epsilon}}ds}{\int_{-L}^L e^{-\frac{h^\epsilon(s)}{\epsilon}}ds} \leq e^{\frac{2\delta}{3\epsilon}}\frac{\int_{-L}^0 e^{-\frac{h(s)}{\epsilon}}ds}{\int_{-L}^L e^{-\frac{h(s)}{\epsilon}}ds}.$$

Using the Proposition 3.4 over $I = [-L, f'_-(u_L)]$, $[f'_-(u^\star), 0]$, $[0, f'_+(u^\star)]$ and at last $[f'_+(u_R), L]$, the Laplace analysis gives then

$$\begin{aligned}
\int_{-L}^0 e^{-\frac{h(s)}{\epsilon}}ds &\sim (f'_-(u^\star) - f'_-(u_L))e^{-\frac{h(f'_-(u^\star))}{\epsilon}}(1 + O(\sqrt{\epsilon})), \\
\int_{-L}^L e^{-\frac{h(s)}{\epsilon}}ds &\sim (f'_+(u_R) - f'_+(u^\star))e^{-\frac{h(f'_+(u^\star))}{\epsilon}}(1 + O(\sqrt{\epsilon})),
\end{aligned}$$

so that finally

$$e^{-\frac{5\delta}{3\epsilon}}\frac{f'_-(u^\star)-f'_-(u_L)}{f'_+(u_R)-f'_+(u^\star)} \lesssim \frac{\int_{-L}^{0} e^{-\frac{h^\epsilon(s)}{\epsilon}}ds}{\int_{-L}^{L} e^{-\frac{h^\epsilon(s)}{\epsilon}}ds} \lesssim e^{-\frac{\delta}{3\epsilon}}\frac{f'_-(u^\star)-f'_-(u_L)}{f'_+(u_R)-f'_+(u^\star)}.$$

Both squeezing quantities tends to 0 as ϵ tends to 0, therefore by (3.15), we have

$$\lim_{\epsilon\to 0} u^\epsilon(0) = u_L \neq u^\star.$$

□

Double shock solutions Let now consider as possible limiting Riemann-Dafermos solution a double shock, as represented on Figure 3.2a. It is composed of a shock wave for the left flux f_- connecting the states u_L to an intermediate state denoted $u^\star$, of speed λ_- given by the Rankine-Hugoniot relation, followed by a another shock wave for the right flux connecting $u^\star$ to u_R, of speed λ_+. The function h is as represented on the Figure 3.2b. It admits two local minima respectively at $\xi = \lambda_-$ and $\xi = \lambda_+$. The following analysis shows that such a solution u is possible only if both minima have the same height, in other words only if

$$h(\lambda_-) = h(\lambda_+) \text{ i.e. } u^\star = \frac{u_R^2 - u_L^2 - 4c^2}{2(u_R - u_L - 4c)}. \tag{3.19}$$

Proof of (3.19). Such a double shock solution is possible only if $u_L > u^\star > u_R$. We suppose by contradiction, that $\delta = h(\lambda_-) - h(\lambda_+) \neq 0$, and without loss of generality that $\delta > 0$. Because of the uniform convergence of h^ϵ towards h, for $\epsilon > 0$ small enough, we have $|h^\epsilon(s) - h(s)| \leq \delta/4$, for any $s \in [-L, L]$. Therefore, we get the inequality

$$0 \leq \frac{\int_{-L}^{0} e^{-\frac{h^\epsilon(s)}{\epsilon}}ds}{\int_{-L}^{L} e^{-\frac{h^\epsilon(s)}{\epsilon}}ds} \leq e^{\frac{\delta}{2\epsilon}}\frac{\int_{-L}^{0} e^{-\frac{h(s)}{\epsilon}}ds}{\int_{-L}^{L} e^{-\frac{h(s)}{\epsilon}}ds}.$$

Because h reaches its global minimum over $[-L, 0]$ at λ_- and its global minimum over $[-L, L]$ at λ_+, there exists $\eta > 0$ such that

$$\begin{aligned} &h(s) \geq h(\lambda_-) = h(\lambda_+) + \delta, \quad s \in [-L, 0], \\ &h(s) \leq h(\lambda_+) + \delta/2, \quad |s - \lambda_+| \leq \eta. \end{aligned}$$

Thus, we obtain

$$\int_{-L}^{0} e^{-\frac{h(s)}{\epsilon}}ds \leq L e^{-\frac{h(\lambda_+)}{\epsilon} - \frac{\delta}{\epsilon}}, \text{ and } \int_{-L}^{L} e^{-\frac{h(s)}{\epsilon}}ds \geq 2\eta e^{-\frac{h(\lambda_+)}{\epsilon} - \frac{\delta}{2\epsilon}},$$

and the first inequality becomes

$$0 \leq \frac{\int_{-L}^{0} e^{-\frac{h^\epsilon(s)}{\epsilon}}ds}{\int_{-L}^{L} e^{-\frac{h^\epsilon(s)}{\epsilon}}ds} \leq e^{-\frac{\delta}{2\epsilon} - \frac{\delta}{\epsilon} + \frac{\delta}{2\epsilon}}\frac{L}{2\eta} = e^{-\frac{\delta}{\epsilon}}\frac{L}{2\eta}.$$

Consequently, by (3.15), we get

$$\lim_{\epsilon \to 0} u^{\epsilon}(0) = u_L \neq u^{\star}.$$

□

Due to the monotonicity of u, this intermediate state $u^{\star}$ given by (3.19) needs to belong to the interval $[u_R, u_L]$. It restricts the domain of existence of a double-shock solution to the states (u_L, u_R) that satisfy

$$u_R - 4c - 2\sqrt{c(5c - 2u_R)} \leq u_L \leq u_R + 2\sqrt{c(2u_R - c)}).$$

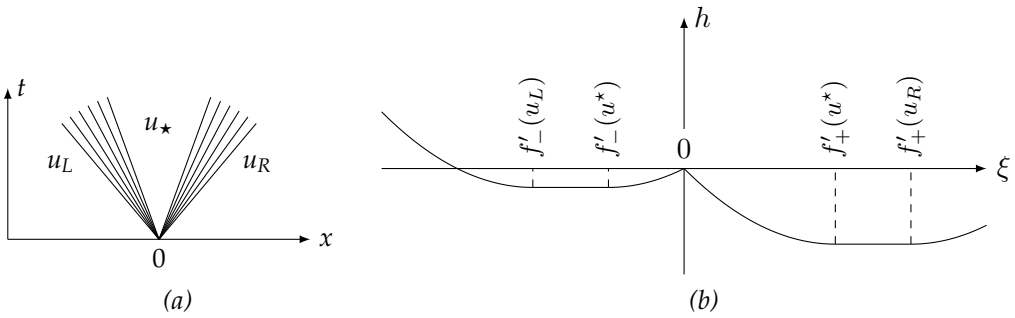


Figure 3.1: Double rarefaction solutions structure

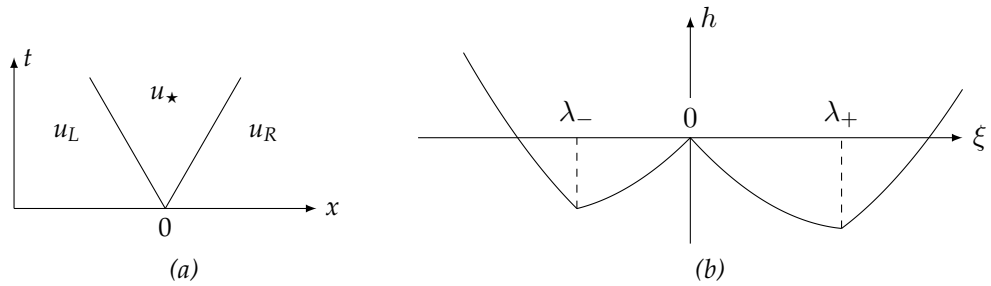


Figure 3.2: Double shock solutions structure

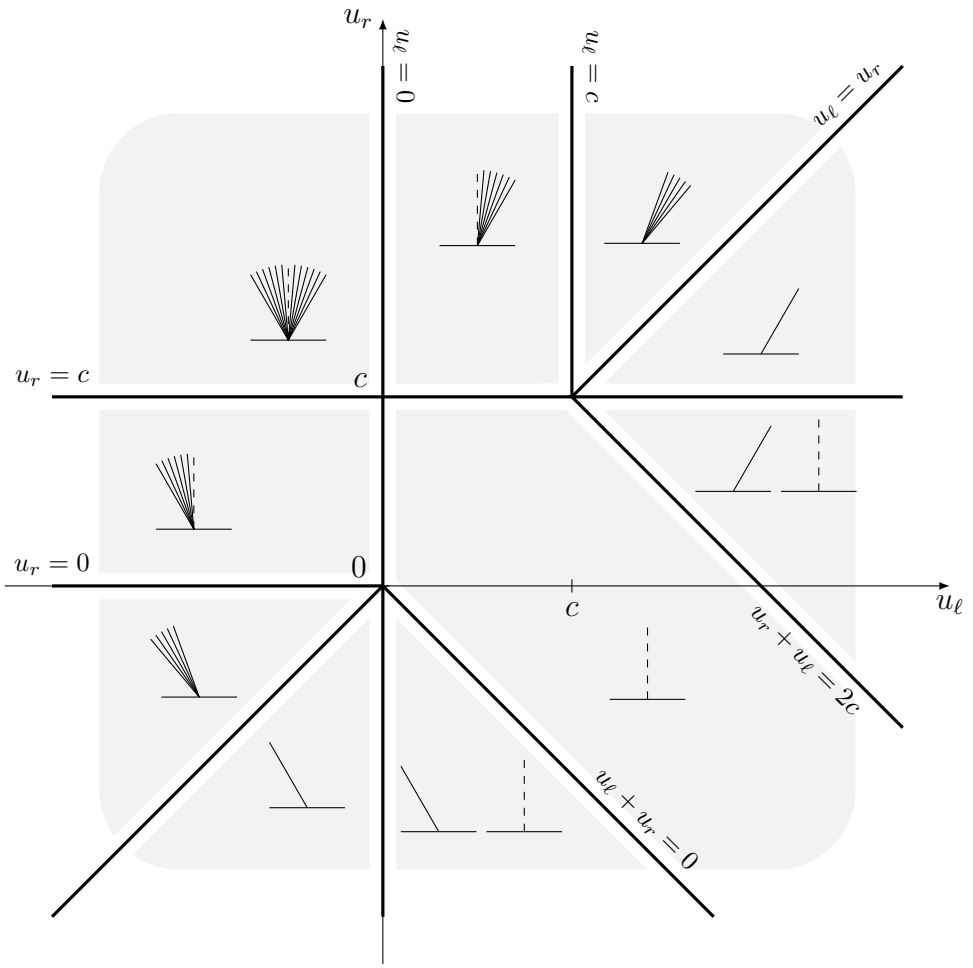


Figure 3.3: Riemann-Dafermos solutions for the state coupling of two convex quadratic fluxes for $c > 0$.

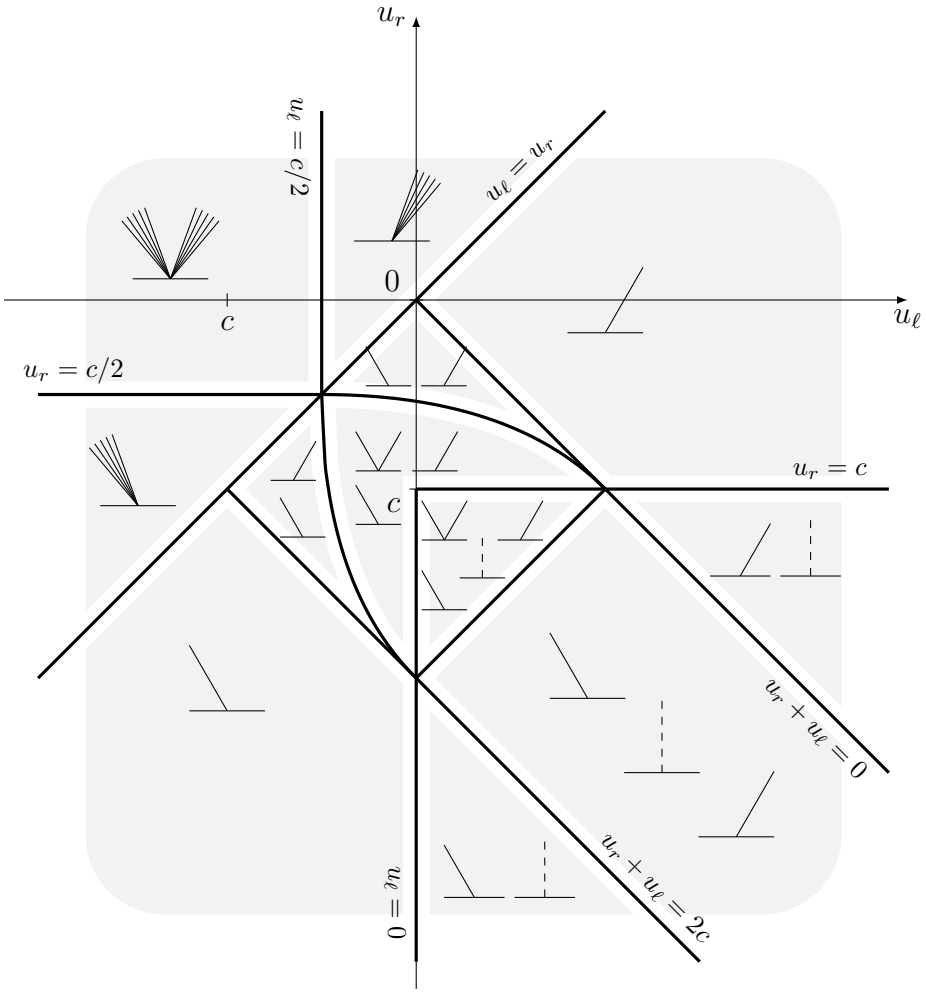


Figure 3.4: Riemann-Dafermos solutions for the state coupling of two convex quadratic fluxes for $c < 0$.

4 A REGULARIZATION METHOD BASED ON THICK INTERFACES

Dans les deux chapitres précédents, nous avons introduit une nouvelle approche pour le couplage d'équations hyperboliques non-linéaires par une formulation EDP augmentée qui présente l'avantage d'éviter une description explicite des interfaces. Nous avons établi sous des hypothèses assez générales l'existence de solutions auto-semblables dans ce régime d'interface mince. D'autre part, on a observé que plusieurs solutions distinctes peuvent apparaître même lorsque l'interface est dotée de mécanismes de régularisation visqueuse. Dans le chapitre présent, nous montrons que cette approche par systeme d'EDP augmenté permet une autre stratégie de régularisation naturelle basée sur des interfaces épaissies. Nous montrons que ce nouveau cadre guarantit l'unicité de la solution du problème de Cauchy, au moins pour des lois de conservation scalaires avec donnée initiale dans L^∞. Ce résultat est obtenu par une approche constructive reposant sur un schéma de volumes finis équilibre.

Le travail présenté dans ce chapitre a été réalisé en collaboration avec Frédéric Coquel et Philippe G. LeFloch et fait l'objet de la publication en préparation [34].

Coupling nonlinear hyperbolic equations (III). A regularization method based on thick interfaces

Abstract

In the first two papers of this series, we introduced a new framework for the coupling of nonlinear hyperbolic equations via an augmented PDE's formulation, which has the advantage to avoid any explicit description of the interfaces. We established a fairly general existence result for self-similar solutions with thin interfaces. In this regime, several distinct solutions may arise even though the interface is endowed with viscous regularizing mechanisms. In the present paper, we show that the proposed augmented PDE's approach naturally allows another regularization strategy based on thick interfaces. We show that this new setting guaranties uniqueness for the solution of the Cauchy problem, at least for scalar conservation laws with initial data in L^∞. This is established by a constructive approach based on a well-balanced finite volume scheme.

4.1 Introduction

4.1.1 Main problematic and notations

In this paper, we are interested in the following system of PDE on the unknown $u := u(t,x) \in \mathbb{R}$ and $v := v(t,x) \in [0,1]$

$$\begin{aligned} \partial_t C_0(u,v) + \partial_x C_1(u,v) - \partial_v C_1(u,v)\partial_x v = 0, \\ \partial_t v = 0, \end{aligned} \tag{4.1}$$

where C_0 and C_1 take the form

$$\begin{aligned} C_0(u,v) &= (1-v)\gamma_-(u) + v\gamma_+(u), \\ C_1(u,v) &= (1-v)f_-(\gamma_-(u)) + vf_+(\gamma_+(u)), \end{aligned}$$

and γ_- and γ_+ are given strictly increasing functions in $C^1(\mathbb{R})$ and f_- and f_+ belong to $C^1(\mathbb{R})$. Together with the problem (4.1), we consider the following initial data

$$\begin{aligned} u(0,x) &= u^0(x), \\ v(0,x) &= v^0(x). \end{aligned} \tag{4.2}$$

A supplemented entropy criterion complete the formulation of this hyperbolic problem. It consists into inequalities of the form

$$\partial_t U(C_0(u,v)) + \partial_x q(u,v) - \partial_v q(u,v)\partial_x v \le 0, \tag{4.3}$$

where U is any convex function and q is the associated flux defined through

$$\partial_u q(u,v) = U'(C_0(u,v))\partial_u C_1(u,v),$$

for example with

$$q(u,v) = \int^u U'(C_0(\theta,v))\partial_\theta C_1(\theta,v)\,d\theta.$$

The motivation for this problem is to propose a thick approach to study the coupling problem of two different scalar conservation laws, say

$$\begin{aligned} &\partial_t w + \partial_x f_-(w) = 0, \quad x < 0, \quad t > 0, \\ &\partial_t w + \partial_x f_+(w) = 0, \quad x > 0, \quad t > 0, \end{aligned} \tag{4.4}$$

together with a coupling condition at the interface $x = 0$

$$\theta_-(w(t,0^-)) = \theta_+(w(t,0^+)), \quad t > 0. \tag{4.5}$$

The functions θ_- and θ_+ correspond then precisely to the inverse functions of γ_- and γ_+ respectively.

The source term in the RHS of (4.1) writes also

$$\partial_v C_1(u,v)\partial_x v = \Big(f_+(\gamma_+(u)) - f_-(\gamma_-(u))\Big)\partial_x v, \tag{4.6}$$

and thus the Jacobian matrix of the system (4.1) reads

$$A(u,v) = \begin{pmatrix} (\partial_u C_0(u,v))^{-1}\partial_u C_1(u,v) & 0 \\ 0 & 0 \end{pmatrix}.$$

It admits two eigenvalues: 0 for the characteristic field corresponding to the second equation, and $\lambda(u,v) := (\partial_u C_0(u,v))^{-1}\partial_u C_1(u,v)$. Observe this second eigenvalue may be zero, both eigenvalues then coincide. It occurs precisely when $\partial_u(f_-(\gamma_-(u)))$ and $\partial_u(f_+(\gamma_+(u)))$ have opposite signs, for some appropriate value of $v \in [0,1]$. The global system is hence only *non-strictly* hyperbolic.

Suppose v is constant, say

$$v(x) = v_\star \in [0,1], \quad x \in \mathbb{R}.$$

The function $C_0(\cdot, v_\star)$ is then an admissible change of variable. In the sequel, we denote $u(\cdot, v_\star)$ his inverse function. It means

$$w = C_0(u, v_\star) \tag{4.7}$$

will be equivalent, with little abuse in the notations, to

$$u = u(w, v_\star).$$

In this variable, the equations (4.1)-(4.3) become then a classical conservation law with its entropy inequalities

$$\begin{aligned} \partial_t w + \partial_x h(w, v_\star) &= 0, \\ \partial_t U(w) + \partial_x F(w, v_\star) &\leq 0, \end{aligned} \tag{4.8}$$

where we setted the flux and the entropy flux

$$\begin{aligned} h(w,v) &= C_1(u(w,v),v), \\ F(w,v) &= q(u(w,v),v), \end{aligned}$$

More generally, if v^0 is smooth, source terms are introduced to take into account the space variation of v in the conservation law and in the entropy inequalities

$$\begin{aligned} s(w,v) &= \partial_v C_1(u(w,v),v)\partial_x v, \\ S(w,v) &= \partial_v q(u(w,v),v)\partial_x v. \end{aligned}$$

The system (4.1)-(4.3) rewrites then under the form

$$\begin{aligned} \partial_t w + \partial_x h(w,v) &= s(w,v), \\ \partial_t U(w) + \partial_x F(w,v) &\leq S(w,v). \end{aligned} \tag{4.9}$$

with initial data

$$w(0,x) = w^0(x) := C_0(u^0, v^0).$$

Because v satisfy a trivial PDE and solves into $v(t,x) = v^0(x)$, we assimilate with a slight abuse in the notation $h(w, v^0(x))$ to $h(w,x)$ and similarly $s(w, v^0(x))$ to $s(w,x)$, etc. Under this point of view, the problem (4.9) belongs to the general framework of scalar hyperbolic equations with nonhomogeneous flux and source term, and the next Kružkov theorem applies.

Theorem 4.1 (Kružkov theorem [94]). *Let $h \in C^1(\mathbb{R}\times\mathbb{R})$ be a nonhomogeneous flux, $s \in C^1(\mathbb{R}\times\mathbb{R})$, and $w_0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$, then there exists a unique entropy solution $w \in L^\infty(\mathbb{R}^+, L^1(\mathbb{R}) \cap L^\infty(\mathbb{R}))$ to the Cauchy problem*

$$\begin{aligned} \partial_t w + \partial_x h(w,x) &= s(w,x), \\ \partial_t U(w) + \partial_x F(w,x) &\leq S(w,x), \end{aligned} \tag{4.10}$$

with initial data $w(0,\cdot) = w^0$.

Observe that every steady solution for (4.9) should satisfy

$$\partial_x h(w,v) = s(w,v),$$

or formally in the u variable

$$((1-v)f_-'(\gamma_-(u))\gamma_-'(u) + v f_+'(\gamma_+(u))\gamma_+'(u))\partial_x u = 0.$$

At the numerical level, the difficulty consists in capturing such steady solutions, especially when the coefficient $((1-v)f'_-(\gamma_-)\gamma'_-(u)+vf'_+(\gamma_+)\gamma'_+(u))$ vanishes, that means, when a strict hyperbolicity loss occurs for (4.1). To avoid this problem, and in accordance with the coupling condition (4.5), we propose to enforce the solution with constant u to be stable solutions, even if resonance occurs. The numerical scheme will be designed according to this choice. Let thus define the next notion of equilibrium.

Definition 4.1 (Equilibrium)**.** A discontinuity between two states (w_-,v_-) and (w_+,v_+) of $\mathbb{R}\times[0,1]$, is said to be an equilibrium if and only if it satisfy the relation

$$u(w_-,v_-)=u(w_+,v_+). \tag{4.11}$$

In the context of the coupling problematic, such an equilibrium can be considered as a standing wave for the extended problem in (w,v), with Riemann invariant u.

4.1.2 Outline

In the next section, we introduce an equilibrium scheme for solving (4.1)-(4.2)-(4.3) that would preserve moreover solutions satisfying the equilibrium property $u=cst$. A reconstruction step will be used in order to preserve these equilibrium. We obtain then a L^∞ control on the numerical solution u_h in Theorem 4.3, and for the trivial case $C_0(u,v)=u$ a BV control ensuring the convergence of the scheme to the unique solution of Theorem 4.1. However, in the general case, the sequence $(u_h)_{h>0}$ does not satisfy the usually required BV control, due to the reconstruction step. Following Coquel and LeFloch [55], we thus take advantage of the theory of measure-valued solutions (see Tartar [128] and DiPerna [63, 64]) to establish the convergence (for a subsequence at least) to an entropy measure-valued solution. We recall hereafter some elements to enter this theory and the reader will refer to references above and to works by Szepessy [126], [127] for more precisions on the subject.

4.2 Presentation of the well-balanced scheme

In this section, we design a numerical scheme, that is consistent with the weak formulation of (4.9) in order to benefits the uniqueness of solution obtained through Theorem 4.1, and that preserves moreover steady states such that $u(w,v)$ is a constant.

Let $(x_{j+1/2})_{j\in\mathbb{Z}}$ be a discretization grid for the real line, and $(x_j)_{j\in\mathbb{Z}}$ the interpoints grid, such that $x_{j+1/2}+x_{j-1/2}=2x_j$. Three main variables will be used in the following : w_j^n and u_j^n on the first grid $(x_{j-1/2},x_{j+1/2})$, and $v_{j+1/2}$ representing the inhomogeneity of the model v on the second grid (x_j,x_{j+1}) (see Figure 4.1). The discrete initial data corresponds for example to

$$\begin{aligned} u_j^0&=\frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}}u^0(x)\,dx,\\ v_{j+1/2}^0&=\frac{1}{\Delta x}\int_{x_j}^{x_{j+1}}v^0(x)\,dx, \end{aligned} \tag{4.12}$$

where because the solution v is independent of the time, we simply denote in the sequel $v_{j+1/2} := v^0_{j+1/2}$. We thus define a discrete initial data in the variable w for each half-cell

$$\begin{aligned} w^0_{j+1/2-} &:= C_0(u^0_j, v_{j+1/2}), \\ w^0_{j-1/2+} &:= C_0(u^0_j, v_{j-1/2}). \end{aligned} \tag{4.13}$$

Let define the following piecewise constant functions $v_h = v_h(x)$, $u_h = u_h(t,x)$ such that

$$\begin{aligned} u_h(t,x) &= \sum_{n\geq 0}\sum_{j\in\mathbb{Z}} u^n_j \chi_{[t^n,t^{n+1}[\times[x_{j-1/2},x_{j+1/2}[}(t,x), \\ v_h(x) &= \sum_{j\in\mathbb{Z}} \frac{v_{j-1/2}+v_{j+1/2}}{2}\chi_{[x_{j-1/2},x_{j+1/2}[}(x). \end{aligned} \tag{4.14}$$

where the u^n_j will be an approximation of the mean value of the exact solution over the cell $[x_{j-1/2},x_{j+1/2}[$ at time t^n, and $v_{j+1/2} = v^0_{j+1/2}$. We present now the three main steps in which the scheme consists: reconstruction, advection and projection steps (see Figure 4.1).

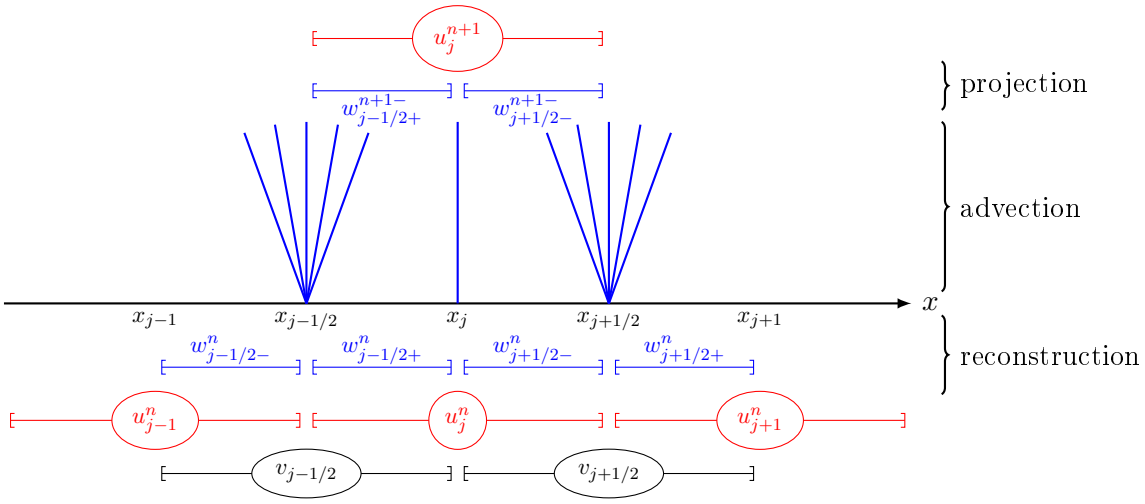


Figure 4.1: Numerical strategy

Reconstruction step This step consists in producing a discontinuity, located at x_j, between the states $(w^n_{j-1/2+}, v_{j-1/2})$ and $(w^n_{j+1/2-}, v_{j+1/2})$, and that will be an equilibrium with the given invariant u^n_j. We define to this end the next reconstructed states $w^n_{j-1/2+}$ and $w^n_{j+1/2-}$ by

$$\begin{aligned} w^n_{j+1/2-} &:= C_0(u^n_j, v_{j+1/2}), \\ w^n_{j-1/2+} &:= C_0(u^n_j, v_{j-1/2}). \end{aligned} \tag{4.15}$$

and a global state

$$w^n_j = C_0\left(u^n_j, \frac{v_{j-1/2}+v_{j+1/2}}{2}\right). \tag{4.16}$$

such that $w^n_j = \frac{1}{2}(w^n_{j+1/2-} + w^n_{j-1/2+})$.

Advection step The advection is realized by solving Riemann problem on each half-cell. Notice that the shift between the respective grids for u_j^n and for $v_{j+1/2}$ is a crucial point because v is continuous at the interface $x_{j+1/2}$, thus the flux $h(w,x)$ is homogeneous and the source term is zero along the cell (x_j, x_{j+1}). Hence, the equation (4.9) reads as a standard conservation law (homogeneous and without source term) (4.8) with $v_\star = v_{j+1/2}$. Moreover the discontinuity at x_j, considered as a standing wave, gives rise to the following definition for numerical fluxes at each intercell

$$\begin{aligned} &h(w^n_{j+1/2,-}, v_{j+1/2}), \text{ at } x_j^+, \\ &H^n_{j+1/2} = H(w^n_{j+1/2,-}, w^n_{j+1/2,+}, v_{j+1/2}), \text{ at } x_{j+1/2}, \\ &h(w^n_{j+1/2,+}, v_{j+1/2}), \text{ at } x_{j+1}^-. \end{aligned} \tag{4.17}$$

In the previous definition, $H(\cdot,\cdot,v)$ is here a two point numerical flux (typically an E-flux or a monotone flux) consistent, for each $v \in [0,1]$, with the flux $h(\cdot,v)$: $H(w,w,v) = h(w,v)$.

To summarize this advection step, under a standard Courant-Friedrichs-Lewy condition

$$\frac{\Delta t}{\Delta x}\left(\max_{j\in\mathbb{Z},w} h'(w, v_{j+1/2})\right) \le 1/2, \tag{4.18}$$

the advection step writes, with $\lambda = \Delta t/\Delta x$,

$$\begin{aligned} w^{n+1-}_{j+1/2,-} &= w^n_{j+1/2,-} - 2\lambda\left(H^n_{j+1/2} - h(w^n_{j+1/2,-}, v_{j+1/2})\right), \\ w^{n+1-}_{j+1/2,+} &= w^n_{j+1/2,+} - 2\lambda\left(h(w^n_{j+1/2,+}, v_{j+1/2}) - H^n_{j+1/2}\right). \end{aligned} \tag{4.19}$$

Moreover, due to the entropy character of the flux, this advection step satisfies the following entropy inequalities

$$\begin{aligned} U(w^{n+1-}_{j+1/2-}) - U(w^n_{j+1/2-}) + 2\lambda\left(F^n_{j+1/2} - F(w^n_{j+1/2-}, v_{j+1/2})\right) \le 0, \\ U(w^{n+1-}_{j+1/2+}) - U(w^n_{j+1/2+}) + 2\lambda\left(F(w^n_{j+1/2+}, v_{j+1/2}) - F^n_{j+1/2}\right) \le 0. \end{aligned} \tag{4.20}$$

Thus, on an entire cell centered on x_j, the total entropy inequality writes

$$\begin{aligned} &\frac{1}{2}\left(U(w^{n+1-}_{j+1/2-}) + U(w^{n+1-}_{j-1/2+})\right) - \frac{1}{2}\left(U(w^n_{j+1/2-}) + U(w^n_{j-1/2+})\right) \\ &+ \lambda\left(F^n_{j+1/2} - F^n_{j-1/2}\right) + \lambda\left(F(w^n_{j-1/2+}, v_{j-1/2}) - F(w^n_{j+1/2-}, v_{j+1/2})\right) \le 0. \end{aligned} \tag{4.21}$$

Projection step We finally recover a constant state u_j^{n+1} over the cell $(x_{j-1/2}, x_{j+1/2})$ by the conservative average in the advected variable

$$w_j^{n+1} = \frac{1}{2}\left(w^{n+1-}_{j-1/2+} + w^{n+1-}_{j+1/2-}\right), \tag{4.22}$$

and then by considering

$$\begin{aligned} u_j^{n+1} &= u\left(w_j^{n+1}, \frac{v_{j-1/2}+v_{j+1/2}}{2}\right) \\ &= u\left(\frac{w_{j-1/2+}^{n+1-}+w_{j+1/2-}^{n+1-}}{2}, \frac{v_{j-1/2}+v_{j+1/2}}{2}\right). \end{aligned} \tag{4.23}$$

In other words, the iteration process from w_j^n to w_j^{n+1} could be summarized into the definition of reconstructed states (4.15) and the update formula

$$\begin{aligned} w_j^{n+1} = u_j^n &- \lambda\left(H_{j+1/2}^n - H_{j-1/2}^n\right) \\ &+ \lambda\left(h(w_{j-1/2+}^n, v_{j-1/2}) - h(w_{j+1/2-}^n, v_{j+1/2})\right). \end{aligned} \tag{4.24}$$

We recognize in this last formulation the classical incoming flux $H_{j-1/2}^n$ and outgoing flux $H_{j+1/2}^n$ and a supplemented source term at x_j consistent with $s(w,v)$. More precisely, one gets in this source term the discrete form of (4.6)

$$\begin{aligned} &h(w_{j-1/2+}^n, v_{j-1/2}) - h(w_{j+1/2-}^n, v_{j+1/2}) = \\ &\qquad \left(f_+(\gamma_+(u_j^n)) - f_-(\gamma_-(u_j^n))\right)(v_{j-1/2} - v_{j+1/2}). \end{aligned} \tag{4.25}$$

Property 4.2 (Equilibrium)**.** *Let be given an initial data (u^0, v^0) for (4.1) such that u is constant, say*

$$u^0(x) = u_\star, \quad x \in \mathbb{R}, \tag{4.26}$$

then for all $t > 0$

$$u_h(t,x) = u_\star, \quad x \in \mathbb{R}. \tag{4.27}$$

Proof. By definition we get first $u_j^0 = u_\star$ for all $j \in \mathbb{Z}$, and thus in both half cells corresponding to the same state $v_{j+1/2}$, it comes $w_{j+1/2,-}^0 = w_{j+1/2,+}^0$. Consequently, the flux writes

$$H_{j+1/2}^0 = H(w_{j+1/2,-}^0, w_{j+1/2,+}^0, v_{j+1/2}) = h(w_{j+1/2,-}^0, v_{j+1/2})$$

by the consistency property. Equations (4.19) thus read simply

$$w_{j+1/2,-}^{1-} = w_{j+1/2,-}^0, \qquad w_{j+1/2,+}^{1-} = w_{j+1/2,+}^0,$$

and then $w_j^1 = u_j^0$. But observe u_j^1 is obtained as the unique solution of

$$\left(1 - \frac{v_{j-1/2}+v_{j+1/2}}{2}\right)\gamma_-(u_j^1) + \frac{v_{j-1/2}+v_{j+1/2}}{2}\gamma_+(u_j^1) = w_j^1, \tag{4.28}$$

precisely the same equation as u_j^0 already satisfies, therefore by unicity $u_j^1 = u_j^0 = u_\star$ and the conclusion follows. □

This equilibrium property ensures in particular that any initial profile $(w_j^0)_{j\in\mathbb{Z}}$ consisting in a Riemann data satisfying the coupling condition (4.5) is a stationary solution.

4.3 Inf-sup estimate

For convenience we first introduce the following notation : $[a_0, a_1, \dots, a_p]$ refers to the convex hull of the at most countable real set $\{a_0, a_1, \dots, a_p\}$, (with $p \in \mathbb{N} \cup \{+\infty\}$). By this way, any inequality of type

$$\min(a_0, a_1, \dots, a_p) \leq w \leq \max(a_0, a_1, \dots, a_p),$$

will equivalently write

$$w \in [a_0, a_1, \dots, a_p].$$

Proposition 4.3 (L^∞ stability)**.** *Under CFL restriction (4.18), for any monotone flux or E-flux H, the numerical solution satisfies the following L^∞ estimate*

$$u_j^n \in \mathcal{U} := [(u_j^0)_{j\in\mathbb{Z}}]. \tag{4.29}$$

Remark 4.1. Observe the Proposition 4.3 furnishes a alternative proof to recover the equilibrium property 4.2.

It follows directly from the Proposition 4.3 the following corollary for w_j^n.

Corollary 4.4. *Under the assumptions of Proposition 4.3, the numerical solution satisfies the following L^∞ estimate*

$$w_j^n \in \mathcal{W} := [(\gamma_-(u_j^0))_{j\in\mathbb{Z}}, (\gamma_+(u_j^0))_{j\in\mathbb{Z}}]. \tag{4.30}$$

Proof of Proposition 4.3. The proof relies directly on the following local L^∞ estimate

$$u_j^{n+1} \in [u_{j-1}^n, u_j^n, u_{j+1}^n]. \tag{4.31}$$

The CFL condition implies during the advection step and thus through Riemann problem solving a first classical L^∞ control

$$\begin{aligned} w_{j-1/2,+}^{n+1-} &\in [w_{j-1/2,-}^n, w_{j-1/2,+}^n], \\ w_{j+1/2,-}^{n+1-} &\in [w_{j+1/2,-}^n, w_{j+1/2,+}^n]. \end{aligned}$$

Using the nondecreasing character of $u(\cdot, v)$ for each $v \in [0,1]$, one gets then

$$\begin{aligned} u(w_{j-1/2,+}^{n+1-}, v_{j-1/2}) &\in [u(w_{j-1/2,-}^n, v_{j-1/2}), u(w_{j-1/2,+}^n, v_{j-1/2})] = [u_{j-1}^n, u_j^n], \\ u(w_{j+1/2,-}^{n+1-}, v_{j+1/2}) &\in [u(w_{j+1/2,-}^n, v_{j+1/2}), u(w_{j+1/2,+}^n, v_{j+1/2})] = [u_{j+1}^n, u_j^n], \end{aligned}$$

Let compare the quantities $u_j^{n+1} = u(w_{j-1/2,+}^{n+1}, v_{j-1/2}) = u(w_{j+1/2,-}^{n+1}, v_{j+1/2})$ to their equivalent before the projection step $\tilde{u}_{j-1/2,+}^n = u(w_{j-1/2,+}^{n+1-}, v_{j-1/2})$ and $\tilde{u}_{j+1/2,-}^n = u(w_{j+1/2,-}^{n+1-}, v_{j+1/2})$. The conservation property (4.22) reads also

$$\begin{aligned} w_j^{n+1} = &\frac{1}{2}\Big((1 - v_{j-1/2})\gamma_-(\tilde{u}_{j-1/2,+}^n) + v_{j-1/2}\gamma_+(\tilde{u}_{j-1/2,+}^n)\Big) \\ &+ \frac{1}{2}\Big((1 - v_{j+1/2})\gamma_-(\tilde{u}_{j+1/2,-}^n) + v_{j+1/2}\gamma_+(\tilde{u}_{j+1/2,-}^n)\Big). \end{aligned}$$

And by definition, one has simultaneously

$$w_j^{n+1} = \frac{1}{2}\Big((1-v_{j-1/2})\gamma_-(u_j^{n+1}) + v_{j-1/2}\gamma_+(u_j^{n+1})\Big) + \frac{1}{2}\Big((1-v_{j+1/2})\gamma_-(u_j^{n+1}) + v_{j+1/2}\gamma_+(u_j^{n+1})\Big).$$

Thus finally,

$$(1-v_{j-1/2})(\gamma_-(\tilde{u}^n_{j-1/2,+}) - \gamma_-(u_j^{n+1})) + v_{j-1/2}(\gamma_+(\tilde{u}^n_{j-1/2,+}) - \gamma_+(u_j^{n+1})) =$$
$$-(1-v_{j+1/2})(\gamma_-(\tilde{u}^n_{j+1/2,-}) - \gamma_-(u_j^{n+1})) - v_{j+1/2}(\gamma_+(\tilde{u}^n_{j+1/2,-}) - \gamma_+(u_j^{n+1})).$$

Because of the nondecreasing character of γ_- and of γ_+, the LHS has the sign of $\tilde{u}^n_{j-1/2,+} - u_j^{n+1}$, while the RHS has the sign of $-\tilde{u}^n_{j+1/2,-} + u_j^{n+1}$, thus we immediately get $u_j^{n+1} \in [\tilde{u}^n_{j-1/2,+}, \tilde{u}^n_{j+1/2,-}]$ and the result follows. □

4.4 Entropy control. Convergence of the scheme

In this section, we prove the convergence of the numerical scheme in the general case of strictly increasing functions γ_- and γ_+. Because of the reconstruction step, the scheme does not satisfy any BV stability in the general case, we thus work in the sense of Young measures and recover uniqueness and strong convergence for the limiting solution through entropy inequalities (Lemma 4.7). However we need first a control of discrete oscillations to obtain these discrete entropy inequalities (Lemma 4.6). This method to obtain the convergence of a numerical scheme in the L^∞ context has been initiated by Coquel and LeFloch [55].

Let $\mu_{t,x}$ be the Young measure associated to u_h as $h \to 0$, i.e. such that for all $a \in C(\mathbb{R})$ the $L^\infty(\mathbb{R}^+ \times \mathbb{R})$ weak star limit

$$a(u_h(\cdot)) \underset{h\to 0}{\rightharpoonup} \bar{a}(\cdot) \tag{4.32}$$

exists, where

$$\bar{a}(t,x) = \int_{\mathbb{R}} a(\cdot) d\mu_{t,x} \equiv \langle \mu_{t,x}, a(\cdot)\rangle, \quad \text{for a.e. } (t,x) \in \mathbb{R}^+ \times \mathbb{R}. \tag{4.33}$$

The main theorem of this paper is the following

Theorem 4.5 (Entropy measure-valued solution and convergence of the scheme)**.** *The Young measure $\mu = \mu_{t,x}$ associated to the uniform L^∞ sequence u_h is an entropy measure-valued solution of (4.1)-(4.2): it satisfies in the weak sense the equation*

$$\partial_t\langle \mu, C_0(\cdot,v)\rangle + \partial_x\langle \mu, C_1(\cdot,v)\rangle - \langle \mu, \partial_v C_1(\cdot,v)\rangle \partial_x v = 0, \tag{4.34}$$

and the weak formulation of entropy inequalities (4.3)

$$\partial_t\langle \mu, U(C_0(\cdot,v))\rangle + \partial_x\langle \mu, q(\cdot,v)\rangle - \langle \mu, \partial_v q(\cdot,v)\rangle \partial_x v \le 0. \tag{4.35}$$

Moreover, the sequence u_h converges strongly in $L^1_{loc}(\mathbb{R}^+ \times \mathbb{R})$ to the unique entropy solution u of

(4.1)-(4.2).

As mentioned above, the proof of this result will be based on discrete entropy inequalities obtained through Lemma 4.7. To get them, a crucial tool will be the

Lemma 4.6 (Weak control on discrete oscillations). *Let $\varphi \in C_0^\infty(\mathbb{R})$ be a positive test function with compact support and let define*

$$\varphi_j = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \varphi(x)\,dx,$$

then, for any fixed time $T = N\Delta t$ there exists a constant $C(T,\varphi)$ independent of Δx such that the numerical solution satisfies the control

$$\Delta x \sum_{n=0}^{N}\sum_{j\in\mathbb{Z}} \left| w_{j+1/2-}^{n+1-} - w_{j-1/2+}^{n+1-} \right|^2 \varphi_j \le C(T,\varphi). \tag{4.36}$$

Lemma 4.7 (Discrete entropy inequality). *Let $\varphi \in C_0^\infty(\mathbb{R}^+\times\mathbb{R})$ be a positive test function with compact support and let define*

$$\varphi_{j+1/2}^n = \frac{1}{\Delta x}\int_{x_j}^{x_{j+1}} \varphi(t^n,x)\,dx,$$

and the piecewise constant function $\varphi_h = \varphi_h(t,x)$ such that

$$\varphi_h(t,x) = \sum_{n\ge 0}\sum_{j\in\mathbb{Z}} \frac{\varphi_{j-1/2}^n + \varphi_{j+1/2}^n}{2} \chi_{[t^n,t^{n+1}[\times[x_{j-1/2},x_{j+1/2}[}(t,x)$$

and the discrete derivatives

$$\begin{aligned}
\Delta_x v_h(x) &= \sum_{n\ge 0}\sum_{j\in\mathbb{Z}} \frac{v_{j+1/2}-v_{j-1/2}}{\Delta x} \chi_{[t^n,t^{n+1}[\times[x_{j-1/2},x_{j+1/2}[}(t,x),\\
\Delta_x \varphi_h(t,x) &= \sum_{n\ge 0}\sum_{j\in\mathbb{Z}} \frac{\varphi_{j+1/2}^n-\varphi_{j-1/2}^n}{\Delta x} \chi_{[t^n,t^{n+1}[\times[x_{j-1/2},x_{j+1/2}[}(t,x),\\
\Delta_t \varphi_h(t,x) &= \sum_{n\ge 1}\sum_{j\in\mathbb{Z}} \frac{\varphi_j^n-\varphi_j^{n-1}}{\Delta t} \chi_{[t^n,t^{n+1}[\times[x_{j-1/2},x_{j+1/2}[}(t,x).
\end{aligned}$$

then

$$\begin{aligned}
&\int_{\Delta t}^{+\infty}\int_{\mathbb{R}} U(C_0(u_h,v_h))\Delta_t\varphi_h\,dx\,dt + \int_{\mathbb{R}} U(C_0(u_h^0,v_h))\varphi_h^0\,dx\\
&+\int_0^{+\infty}\int_{\mathbb{R}} q(u_h,v_h)\Delta_x\varphi_h\,dx\,dt - \int_0^{+\infty}\int_{\mathbb{R}} \partial_v q(u_h,v_h)\varphi_h\Delta_x v_h\,dx\,dt\\
&\ge O(\sqrt{\Delta x}).
\end{aligned} \tag{4.37}$$

An easy tool useful in the proofs of Lemma 4.6 and Lemma 4.7 is the

Lemma 4.8 (Precised convexity). *Let w_-, w_+ be two reals and a function $U \in C^2([w_-,w_+])$ and define both mean values $\tilde{w} = \frac{1}{2}(w_- + w_+)$, and $\tilde{W} = \frac{1}{2}(U(w_-)+U(w_+))$, then the following inequality*

vault

$$\frac{1}{8}\min_{[w_-,w_+]} U''\,|w_+-w_-|^2 \le \tilde{W}-U(\tilde{w}) \le \frac{1}{8}\max_{[w_-,w_+]} U''\,|w_+-w_-|^2.$$

Finally, the convergence of the numerical solution will be obtained through the

Proposition 4.9 (Uniqueness of the mv entropy solution)**.** *Suppose that the Young measure μ associated to the sequence u_h is an entropy mv-solution to (4.1) and let u denote the unique entropy solution of (4.1). Then*

$$\mu_{t,x} = \delta_{u(t,x)}, a.e,$$

i.e. $\mu_{t,x}$ reduces to a Dirac measure concentrated at $u(t,x)$.

Proof of Lemma 4.8. Observe $\tilde{w} = \frac{1}{2}(w_- + w_+)$ is the mean value $\tilde{w} = \int_0^1 w(x)\,dx$ of the piecewise constant function $w : [0,1] \mapsto \mathbb{R}$ defined by

$$w(x) = \begin{cases} w_-, & x \le 1/2, \\ w_+, & x > 1/2. \end{cases}$$

Then using Taylor development of U at point $\tilde{w}$ with an order 2 rest term, we obtain

$$\begin{aligned} U(w(x)) - U(\tilde{w}) - U'(\tilde{w})(w(x)-\tilde{w}) &= \int_0^1 U''(\tilde{w}+s(w(x)-\tilde{w}))(1-s)\,ds(w(x)-\tilde{w})^2 \\ &\ge \frac{1}{2}\min_{[w_-,w_+]} U''\,(w(x)-\tilde{w})^2. \end{aligned}$$

Thus, by integrating over $x \in [0,1]$, we get the inequality

$$\tilde{W}-U(\tilde{w}) \ge \frac{1}{2}\min_{[w_-,w_+]} U'' \int_0^1 (w(x)-\tilde{w})^2\,dx \ge \frac{1}{8}\min_{[w_-,w_+]} U''\,|w_+-w_-|^2.$$

Similarly, for the lower bound, we have

$$U(w(x)) - U(\tilde{w}) - U'(\tilde{w})(w(x)-\tilde{w}) \le \frac{1}{2}\max_{[w_-,w_+]} U''\,(w(x)-\tilde{w})^2,$$

and then

$$\tilde{W}-U(\tilde{w}) \le \frac{1}{8}\max_{[w_-,w_+]} U''\,|w_+-w_-|^2.$$

□

Proof of Lemma 4.6. Observe first that we have

$$\begin{aligned} &F(w^n_{j-1/2+}, v_{j-1/2}) - F(w^n_{j+1/2-}, v_{j+1/2}) = q(u^n_j, v_{j-1/2}) - q(u^n_j, v_{j+1/2}), \\ &= \int_0^1 \partial_v q(u^n_j, v_{j-1/2} + s(v_{j+1/2}-v_{j-1/2}))\,ds\,(v_{j+1/2}-v_{j-1/2}) \\ &= \mathcal{T}(v_{j-1/2}, v_{j+1/2}, u^n_j)(v_{j+1/2}-v_{j-1/2}) \end{aligned} \tag{4.38}$$

where the function $\mathcal{T}$ defined by

$$\mathcal{T}(v_{j-1/2}, v_{j+1/2}, u_j^n) = \int_0^1 \partial_v q(u_j^n, v_{j-1/2} + s(v_{j+1/2} - v_{j-1/2}))\, ds$$

belongs to $L^\infty([0,1]^2 \times \mathcal{U})$. Now introducing the quantities $U(w_j^n)$, the inequation (4.21) rewrites

$$\begin{aligned}
&U(w_j^{n+1}) - U(w_j^n) + \lambda\left(F_{j+1/2}^n - F_{j-1/2}^n\right) \\
&\le U(w_j^{n+1}) - \frac{1}{2}\left(U(w_{j+1/2-}^{n+1-}) + U(w_{j-1/2+}^{n+1-})\right) \\
&\quad + \frac{1}{2}\left(U(w_{j+1/2-}^{n}) + U(w_{j-1/2+}^{n})\right) - U(w_j^n) \\
&\quad - \lambda \mathcal{T}(v_{j-1/2}, v_{j+1/2}, u_j^n)(v_{j+1/2} - v_{j-1/2}).
\end{aligned}$$

The RHS is controlled essentially through Lemma 4.8 because w_j^{n+1} is precisely the mean value of $w_{j+1/2-}^{n+1-}$ and of $w_{j-1/2+}^{n+1-}$, and w_j^n the mean value of $w_{j+1/2-}^{n}$ and of $w_{j-1/2+}^{n}$, the previous inequality becomes

$$\begin{aligned}
&U(w_j^{n+1}) - U(w_j^n) + \lambda\left(F_{j+1/2}^n - F_{j-1/2}^n\right) \\
&\le -\frac{1}{8}\min_{w\in\mathcal{W}} U'' \left|w_{j+1/2-}^{n+1-} - w_{j-1/2+}^{n+1-}\right|^2 \\
&\quad + \frac{1}{8}\max_{w\in\mathcal{W}} U'' \left|w_{j+1/2-}^{n} - w_{j-1/2+}^{n}\right|^2 \\
&\quad + \lambda \|\mathcal{T}\|_{L^\infty([0,1]^2\times\mathcal{U})} \operatorname{Lip}(v)\Delta x
\end{aligned}$$

where $|w_{j+1/2-}^n - w_{j-1/2+}^n|$ reads also

$$\begin{aligned}
|w_{j+1/2-}^n - w_{j-1/2+}^n| &= |\gamma_+(u_j^n) - \gamma_-(u_j^n)|\left|v_{j+1/2} - v_{j-1/2}\right| \\
&\le \|\gamma_+ - \gamma_-\|_{L^\infty(\mathcal{U})} \operatorname{Lip}(v)\Delta x.
\end{aligned} \tag{4.39}$$

Finally there exists a positive constant C such that

$$\begin{aligned}
&U(w_j^{n+1}) - U(w_j^n) + \lambda\left(F_{j+1/2}^n - F_{j-1/2}^n\right) \\
&\le -C\left|w_{j+1/2-}^{n+1-} - w_{j-1/2+}^{n+1-}\right|^2 + C\Delta x^2 + C\Delta x.
\end{aligned}$$

Let sum over all $j \in \mathbb{Z}$ this inequality multiplied by $\varphi_j \Delta x$, we thus get

$$\begin{aligned}
&\int_{\mathbb{R}} U(w_h(x,t^{n+1}))\varphi(x)\,dx - \int_{\mathbb{R}} U(w_h(x,t^n))\varphi(x)\,dx - \Delta t \Delta x \sum_{j\in\mathbb{Z}} F_{j+1/2}^n \frac{\varphi_{j+1} - \varphi_j}{\Delta x} \\
&\le -C\Delta x \sum_{j\in\mathbb{Z}} \left|w_{j+1/2-}^{n+1-} - w_{j-1/2+}^{n+1-}\right|^2 \varphi_j + C(\Delta x^2 + \Delta x)\Delta x \sum_{j\in\mathbb{Z}} \varphi_j.
\end{aligned}$$

Let now sum over all $n = 0$ to $N = \frac{T}{\Delta t}$, it comes

$$\begin{aligned}
&\int_{\mathbb{R}} U(w_h(x,T))\varphi(x)\,dx + C\Delta x \sum_{n=0}^{N}\sum_{j\in\mathbb{Z}} \left|w^{n+1-}_{j+1/2-} - w^{n+1-}_{j-1/2+}\right|^2 \varphi_j \\
&\leq \frac{CT}{\lambda}(\Delta x+1)\|\varphi\|_{L^1(\mathbb{R})} + T\|F\|_\infty \|\varphi\|_{TV(\mathbb{R})} + \int_{\mathbb{R}} U(w^0(x))\varphi(x)\,dx \\
&\leq CT\|\varphi\|_{W^{1,1}} + \int_{\mathbb{R}} U(w^0(x))\varphi(x)\,dx.
\end{aligned}$$

The conclusion follows. □

Proof of Lemma 4.7. We take advantage of the entropy inequalities (4.20) on the cells centered at $x_{j+1/2}$, it writes thus

$$\begin{aligned}
&\frac{1}{2}\left(U(w^{n+1-}_{j+1/2-}) + U(w^{n+1-}_{j+1/2+})\right) - \frac{1}{2}\left(U(w^{n}_{j+1/2-}) + U(w^{n}_{j+1/2+})\right) \\
&+ \lambda\left(F(w^n_{j+1/2+}, v_{j+1/2}) - F(w^n_{j+1/2-}, v_{j+1/2})\right) \leq 0.
\end{aligned}$$

We sum this inequality times $\varphi^n_{j+1/2}\Delta x$ over all $j \in \mathbb{Z}$, and over all n from 0 to $N = T/\Delta t$ to get

$$\begin{aligned}
&\Delta x \sum_{j,n} \frac{1}{2}\left(U(w^{n+1-}_{j+1/2-})\varphi^n_{j+1/2} + U(w^{n+1-}_{j-1/2+})\varphi^n_{j-1/2}\right) \\
&- \Delta x \sum_{j,n} \frac{1}{2}\left(U(w^{n}_{j+1/2-})\varphi^n_{j+1/2} + U(w^{n}_{j-1/2+})\varphi^n_{j-1/2}\right) \\
&+ \lambda\Delta x \sum_{j,n}\left(F(w^n_{j+1/2+}, v_{j+1/2}) - F(w^n_{j+1/2-}, v_{j+1/2})\right)\varphi^n_{j+1/2} \\
&\leq 0.
\end{aligned}$$

And the inequality follows

$$\begin{aligned}
&\Delta x \sum_{j,n} U(w^{n+1}_j)\varphi^n_j - \Delta x \sum_{j,n} U(w^n_j)\varphi^n_j \\
&+ \lambda\Delta x \sum_{j,n}\left(F(w^n_{j+1/2+}, v_{j+1/2}) - F(w^n_{j+1/2-}, v_{j+1/2})\right)\varphi^n_{j+1/2} \\
\leq &\Delta x \sum_{j,n}\left(U(w^{n+1}_j)\frac{\varphi^n_{j+1/2}+\varphi^n_{j-1/2}}{2} - U(w^{n+1-}_{j+1/2-})\frac{\varphi^n_{j+1/2}}{2} - U(w^{n+1-}_{j-1/2+})\frac{\varphi^n_{j-1/2}}{2}\right) \\
&- \Delta x \sum_{j,n}\left(U(w^{n}_j)\frac{\varphi^n_{j+1/2}+\varphi^n_{j-1/2}}{2} - U(w^{n}_{j+1/2-})\frac{\varphi^n_{j+1/2}}{2} - U(w^{n}_{j-1/2+})\frac{\varphi^n_{j-1/2}}{2}\right).
\end{aligned} \tag{4.40}$$

The first term A of the RHS, corresponding to the entropy evolution during the projection

step, rewrites

$$A = \Delta x \sum_{j,n} \left(U(w_j^{n+1}) \frac{\varphi_{j+1/2}^n + \varphi_{j-1/2}^n}{2} - U(w_{j+1/2-}^{n+1-}) \frac{\varphi_{j+1/2}^n}{2} - U(w_{j-1/2+}^{n+1-}) \frac{\varphi_{j-1/2}^n}{2} \right)$$
$$= \Delta x \sum_{j,n} \left(U(w_{j-1/2+}^{n+1-}) - U(w_{j+1/2-}^{n+1-}) \right) \frac{\varphi_{j+1/2}^n - \varphi_{j-1/2}^n}{4}$$
$$+ \Delta x \sum_{j,n} \left(U(w_j^{n+1}) - \frac{U(w_{j-1/2+}^{n+1-}) + U(w_{j+1/2-}^{n+1-})}{2} \right) \frac{\varphi_{j+1/2}^n + \varphi_{j-1/2}^n}{2},$$

notice however we have

$$w_j^{n+1} = \frac{w_{j-1/2+}^{n+1-} + w_{j+1/2-}^{n+1-}}{2},$$

and thus by convexity of U

$$\Delta x \sum_{j,n} \left(U(w_j^{n+1}) - \frac{U(w_{j+1/2-}^{n+1-}) + U(w_{j-1/2+}^{n+1-})}{2} \right) \varphi_j^n \leq 0,$$

so that, using a Cauchy-Schwarz inequality, and setting a smooth function ψ with compact support in $\mathbb{R}$, such that $\varphi(x,t) \neq 0 \Rightarrow \psi(x) = 1$ for all $(t,x) \in \mathbb{R}^+ \times \mathbb{R}$, and its discrete representation over the grid

$$\psi_j = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \psi(x)\, dx,$$

we obtain

$$A \leq \frac{\Delta x \Delta t}{\lambda} \sum_{j,n} \frac{\max|U'|}{4} \left| w_{j-1/2+}^{n+1-} - w_{j+1/2-}^{n+1-} \right| \left| \frac{\varphi_{j+1/2}^n - \varphi_{j-1/2}^n}{\Delta x} \right|$$
$$\leq \frac{\max|U'|}{4\sqrt{\lambda}} \left(\Delta x \sum_{j,n} \left(w_{j-1/2+}^{n+1-} - w_{j+1/2-}^{n+1-} \right)^2 \psi_j \right)^{1/2} \left(\Delta t \Delta x \sum_{j,n} \left(\frac{\varphi_{j+1/2}^n - \varphi_{j-1/2}^n}{\Delta x} \right)^2 \right)^{1/2} \sqrt{\Delta x}.$$

In this last expression, the first sum is controlled using Lemma 4.6 with the test function ψ, and the second one bounded uniformly as Δx goes to 0. Finally the first term of the RHS satisfies

$$A \leq O(\sqrt{\Delta x}).$$

By a similar computation, the second term B of the RHS, corresponding to the entropy

evolution during the reconstruction step, equals

$$B = \Delta x \sum_{j,n} \left(-U(w_j^n) \frac{\varphi_{j+1/2}^n + \varphi_{j-1/2}^n}{2} + U(w_{j+1/2-}^n) \frac{\varphi_{j+1/2}^n}{2} + U(w_{j-1/2+}^n) \frac{\varphi_{j-1/2}^n}{2} \right)$$
$$= \Delta x \sum_{j,n} \left(U(w_{j+1/2-}^n) - U(w_{j-1/2+}^n) \right) \frac{\varphi_{j+1/2}^n - \varphi_{j-1/2}^n}{4}$$
$$+ \Delta x \sum_{j,n} \left(\frac{U(w_{j-1/2+}^n) + U(w_{j+1/2-}^n)}{2} - U(w_j^n) \right) \frac{\varphi_{j+1/2}^n + \varphi_{j-1/2}^n}{2}.$$

Using a Lipschitz control for the first contribution and Lemma 4.8 to treat the second term, we obtain

$$B \leq \frac{\Delta x \Delta t}{\lambda} \sum_{j,n} \frac{\max |U'|}{4} \left| w_{j+1/2-}^n - w_{j-1/2+}^n \right| \left| \frac{\varphi_{j+1/2}^n - \varphi_{j-1/2}^n}{\Delta x} \right|$$
$$+ \Delta x \sum_{j,n} \frac{C}{8} \left| w_{j+1/2-}^n - w_{j-1/2+}^n \right|^2 \varphi_j^n.$$

Hence, by (4.39)

$$B \leq \left(\frac{\max |U'|}{4\lambda} \|\partial_x \varphi\|_{L^1(\mathbb{R}\times[0,T])} + \frac{C}{8\lambda} \|\varphi\|_{L^1(\mathbb{R}\times[0,T])} \right) O(\Delta x) = O(\Delta x).$$

We consider finally the flux contribution in the LHS, say

$$C = \lambda \Delta x \sum_{j,n} \left(F(w_{j+1/2+}^n, v_{j+1/2}) - F(w_{j+1/2-}^n, v_{j+1/2}) \right) \varphi_{j+1/2}^n$$
$$= -\Delta x \Delta t \sum_{j,n} \frac{1}{2} \left(F(w_{j-1/2+}^n, v_{j-1/2}) + F(w_{j+1/2+}^n, v_{j+1/2}) \right) \frac{\varphi_{j+1/2}^n - \varphi_{j-1/2}^n}{\Delta x}$$
$$- \Delta t \sum_{j,n} \left(F(w_{j+1/2-}^n, v_{j+1/2}) - F(w_{j-1/2+}^n, v_{j-1/2}) \right) \frac{\varphi_{j+1/2}^n + \varphi_{j-1/2}^n}{2}.$$

where

$$\frac{1}{2} \left(F(w_{j-1/2+}^n, v_{j-1/2}) + F(w_{j+1/2+}^n, v_{j+1/2}) \right) = \frac{1}{2} \left(q(u_j^n, v_{j-1/2}) + q(u_j^n, v_{j+1/2}) \right)$$
$$= q\left(u_j^n, \frac{v_{j-1/2} + v_{j+1/2}}{2} \right) + O(\Delta x),$$

thanks to Lipschitz character of $q(u_j^n, \cdot)$ and of v. Moreover, using (4.38) and the Lipschitz character of $\partial_v q(u_j^n, \cdot)$ and of v, the last quantity rewrites

$$F(w_{j+1/2-}^n, v_{j+1/2}) - F(w_{j-1/2+}^n, v_{j-1/2}) = \mathcal{T}(v_{j-1/2}, v_{j+1/2}, u_j^n)(v_{j+1/2} - v_{j-1/2})$$
$$= \partial_v q\left(u_j^n, \frac{v_{j-1/2} + v_{j+1/2}}{2} \right) (v_{j+1/2} - v_{j-1/2}) + O(\Delta x).$$

And the global inequality (4.40) becomes

$$\begin{aligned}
&\Delta t\Delta x \sum_{n\geq 1}\sum_{j\in\mathbb{Z}} U(w_j^n)\frac{\varphi_j^n-\varphi_j^{n-1}}{\Delta t} + \Delta x\sum_{j\in\mathbb{Z}} U(w_j^0)\varphi_j^0 \\
&+\Delta t\Delta x \sum_{n\geq 0}\sum_{j\in\mathbb{Z}} q\left(u_j^n, \frac{v_{j-1/2}+v_{j+1/2}}{2}\right)\frac{\varphi_{j+1/2}^n-\varphi_{j-1/2}^n}{\Delta x} \\
&-\Delta t\Delta x \sum_{n\geq 0}\sum_{j\in\mathbb{Z}} \partial_v q\left(u_j^n, \frac{v_{j-1/2}+v_{j+1/2}}{2}\right)\frac{v_{j+1/2}-v_{j-1/2}}{\Delta x}\frac{\varphi_{j+1/2}^n+\varphi_{j-1/2}^n}{2} \\
&\geq O(\sqrt{\Delta x}).
\end{aligned}$$

Rewritten in terms of functions w_h, u_h, v_h and their discrete derivatives, we get the expected inequality (4.37). □

Proof of Proposition 4.9. We will use the following Kružkov's entropies, entropy fluxes, and entropy source terms (where we drop for convenience the dependance in x in the notation)

$$\begin{aligned}
\tilde{U}(u_1,u_2) &= |C_0(u_1,v)-C_0(u_2,v)|,\\
\tilde{q}(u_1,u_2) &= \operatorname{sgn}(u_1-u_2)(C_1(u_1,v)-C_1(u_2,v)),\\
\tilde{s}^a(u_1,u_2) &= \operatorname{sgn}(u_1-u_2)\partial_v C_1(u_1,v)\partial_x v,\\
\tilde{s}^b(u_1,u_2) &= -\operatorname{sgn}(u_1-u_2)\partial_v C_1(u_2,v)\partial_x v.
\end{aligned}$$

Let be μ and ν two entropy measure-valued solutions to (4.1), we thus get the two following statements with previous Kružkov's entropies

$$\begin{aligned}
\partial_t\langle\mu,\tilde{U}(\cdot,\bar{u}_2)\rangle + \partial_x\langle\mu,\tilde{q}(\cdot,\bar{u}_2)\rangle - \langle\mu,\tilde{s}^a(\cdot,\bar{u}_2)\rangle \leq 0, \quad \bar{u}_2\in\mathbb{R},\\
\partial_t\langle\nu,\tilde{U}(\bar{u}_1,\cdot)\rangle + \partial_x\langle\nu,\tilde{q}(\bar{u}_1,\cdot)\rangle - \langle\nu,\tilde{s}^b(\bar{u}_1,\cdot)\rangle \leq 0, \quad \bar{u}_1\in\mathbb{R}.
\end{aligned} \tag{4.41}$$

Let now introduce the tensor product $\mu\otimes\nu = \mu_{t,x}\otimes\nu_{t,x}$, with

$$\langle\mu_{t,x}\otimes\nu_{t,x},\tilde{U}\rangle := \iint_{\mathbb{R}^2}\tilde{U}(\bar{u}_1,\bar{u}_2)d\mu_{t,x}(\bar{u}_1)d\nu_{t,x}(\bar{u}_2).$$

We deduce from (4.41), that it satisfies the inequality

$$\partial_t\langle\mu\otimes\nu,\tilde{U}\rangle + \partial_x\langle\mu\otimes\nu,\tilde{q}\rangle - \langle\mu\otimes\nu,\tilde{s}\rangle \leq 0,$$

where

$$\begin{aligned}
\langle\mu\otimes\nu,\tilde{s}\rangle &= \iint_{\mathbb{R}^2}(\tilde{s}^a(\bar{u}_1,\bar{u}_2)+\tilde{s}^b(\bar{u}_1,\bar{u}_2))d\mu_{t,x}(\bar{u}_1)d\nu_{t,x}(\bar{u}_2)\\
&= \iint_{\mathbb{R}^2}\operatorname{sgn}(\bar{u}_1-\bar{u}_2)\left(\partial_v C_1(\bar{u}_1,v(x))-\partial_v C_1(\bar{u}_2,v(x))\right)\partial_x v(x)d\mu_{t,x}(\bar{u}_1)d\nu_{t,x}(\bar{u}_2).
\end{aligned}$$

Integrating this inequality over $x \in \mathbb{R}$ and $t \in [0,T]$, we thus obtain

$$\int_{\mathbb{R}} \langle \mu \otimes \nu, \tilde{U} \rangle(T,x)\,dx \leq \int_{\mathbb{R}} \langle \mu \otimes \nu, \tilde{U} \rangle(0,x)\,dx + \int_0^T \int_{\mathbb{R}} \langle \mu \otimes \nu, \tilde{s} \rangle(t,x)\,dx\,dt$$

Because μ and ν satisfy the same initial condition, $d\mu_{0,x} = d\nu_{0,x} = \delta_{u^0(x)}$ and

$$\begin{aligned}\langle \mu \otimes \nu, \tilde{U} \rangle(0,x) &= \iint_{\mathbb{R}^2} |C_0(\bar{u}_1, v(x)) - C_0(\bar{u}_2, v(x))| d\mu_{0,x}(\bar{u}_1) d\nu_{0,x}(\bar{u}_2) \\ &= |C_0(u^0(x), v(x)) - C_0(u^0(x), v(x))| = 0,\end{aligned}$$

moreover

$$\begin{aligned}|\langle \mu \otimes \nu, \tilde{s} \rangle(t,x)| &\leq \|\partial_x v\|_\infty \mathrm{Lip}(\partial_v C_1) \iint_{\mathbb{R}^2} |\bar{u}_1 - \bar{u}_2| d\mu_{t,x}(\bar{u}_1) d\nu_{t,x}(\bar{u}_2) \\ &\leq \|\partial_x v\|_\infty \mathrm{Lip}(\partial_v C_1) \mathrm{Lip}(C_0^{-1}) \iint_{\mathbb{R}^2} |C_0(\bar{u}_1, v(x)) - C_0(\bar{u}_2, v(x))| d\mu_{t,x}(\bar{u}_1) d\nu_{t,x}(\bar{u}_2) \\ &\leq \|\partial_x v\|_\infty \mathrm{Lip}(\partial_v C_1) \mathrm{Lip}(C_0^{-1}) \langle \mu \otimes \nu, \tilde{U} \rangle(t,x)\end{aligned}$$

Finally, there exists a positive constant C such that

$$\int_{\mathbb{R}} \langle \mu \otimes \nu, \tilde{U} \rangle(T,x)\,dx \leq C \int_0^T \int_{\mathbb{R}} \langle \mu \otimes \nu, \tilde{U} \rangle(t,x)\,dx\,dt,$$

where a Gronwall lemma gives thus for any $t \in [0,T]$

$$\int_{\mathbb{R}} \langle \mu \otimes \nu, \tilde{U} \rangle(t,x)\,dx \leq \int_{\mathbb{R}} \langle \mu \otimes \nu, \tilde{U} \rangle(0,x)\,dx\, e^{Ct}.$$

and as previously, this quantity is 0 because of the initial data. Naturally the first term is positive because $\tilde{U}$ is positive and thus

$$\langle \mu \otimes \nu, \tilde{U} \rangle(t,x) = 0, \quad \text{a.e.} (t,x) \in [0,T] \times \mathbb{R}.$$

In other words, $\mu_{t,x}$ and $\nu_{t,x}$ coincide with a same Dirac measure with support say $\tilde{u}(t,x)$ that is then a classical solution of (4.1) satisfying entropy inequalities (4.3) and the initial condition (4.2), thus $\tilde{u} = u$ a.e.

□

Proof of Theorem 4.5. The expected inequality is understood in the sense of distributions, i.e. for all φ test function with compact support

$$\begin{aligned}&\int_0^{+\infty} \int_{\mathbb{R}} \langle \mu_{t,x}, U(C_0(\cdot, v)) \rangle \partial_t \varphi\,dx\,dt + \int_{\mathbb{R}} \langle \mu_{0,x}, U(C_0(\cdot, v)) \rangle \varphi^0\,dx \\ &+ \int_0^{+\infty} \int_{\mathbb{R}} \langle \mu_{t,x}, q(\cdot, v) \rangle \partial_x \varphi\,dx\,dt - \int_0^{+\infty} \int_{\mathbb{R}} \langle \mu_{t,x}, \partial_v q(\cdot, v) \rangle \varphi \partial_x v\,dx\,dt \geq 0\end{aligned}$$

This inequality for the Young measure $\mu_{t,x}$ will be obtained as the direct consequence of

Lemma 4.7. It remains to obtain the following entropy inequality:

$$\liminf_{h\to 0}\Big(\int_0^{+\infty}\int_{\mathbb{R}} U(C_0(u_h,v))\partial_t\varphi\,dx\,dt+\int_{\mathbb{R}} U(C_0(u_h^0,v))\varphi^0\,dx$$
$$+\int_0^{+\infty}\int_{\mathbb{R}} q(u_h,v)\partial_x\varphi\,dx\,dt-\int_0^{+\infty}\int_{\mathbb{R}}\partial_v q(u_h,v)\varphi\partial_x v\,dx\,dt\Big)\geq 0.$$

Hence we treat each one of the following error terms:

$$E_1=\int_0^{+\infty}\int_{\mathbb{R}}\Big(U(C_0(u_h,v))-U(C_0(u_h,v_h))\Big)\Delta_t\varphi_h\,dx\,dt,$$
$$E_2=\int_0^{+\infty}\int_{\mathbb{R}} U(C_0(u_h,v))\Big(\partial_t\varphi-\Delta_t\varphi_h\Big)dx\,dt,$$
$$E_3=\int_{\mathbb{R}}\Big(U(C_0(u_h^0,v)-U(C_0(u_h^0,v_h))\Big)\varphi_h^0\,dx,$$
$$E_4=\int_{\mathbb{R}} U(C_0(u_h^0,v))\Big(\varphi^0-\varphi_h^0\Big)dx,$$
$$E_5=\int_0^{+\infty}\int_{\mathbb{R}}\Big(q(u_h,v)-q(u_h,v_h)\Big)\Delta_x\varphi_h\,dx\,dt,$$

$$E_6=\int_0^{+\infty}\int_{\mathbb{R}} q(u_h,v)\Big(\partial_x\varphi-\Delta_x\varphi_h\Big)dx\,dt,$$
$$E_7=\int_0^{+\infty}\int_{\mathbb{R}}\Big(\partial_v q(u_h,v)-\partial_v q(u_h,v_h)\Big)\varphi_h\Delta_x v_h\,dx\,dt,$$
$$E_8=\int_0^{+\infty}\int_{\mathbb{R}}\partial_v q(u_h,v)\Big(\varphi-\varphi_h\Big)\Delta_x v_h\,dx\,dt,$$
$$E_9=\int_0^{+\infty}\int_{\mathbb{R}}\partial_v q(u_h,v)\varphi\Big(\partial_x v-\Delta_x v_h\Big)dx\,dt.$$

Using Lipschitz character for the comparison terms and L^∞ estimate for their multipliers, each one of these terms tends to 0 as Δx tends to 0.

The Proposition 4.9 ensures then μ is the unique entropy measure-valued solution with initial data u^0 and thus is a Dirac mass concentrated on the unique entropy solution to (4.1)-(4.2)-(4.3) u. □

4.5 The case of the state coupling. Total variation estimate

The case where $\gamma_-=\gamma_+=Id$ is particular, in the sense that we are there able to get then a total variation estimate and the convergence of the scheme can be directly deduced from this compactness property.

Theorem 4.10 (Convergence for the state coupling)**.** *Assume $\gamma_-=\gamma_+=Id$, then under the CFL restriction (4.18), the scheme is total variation diminishing.*

$$TV(w_h(t^{n+1},\cdot))\leq TV(w_h(t^n,\cdot)). \tag{4.42}$$

As a consequence, the numerical solution $(u_h)_{h>0}$ converges to the unique entropy solution of

(4.1)-(4.2)-(4.3), *as* Δx *tends to 0.*

Remark 4.2. Let remark this total variation estimate does not suppose any smoothness assumption on the function v.

Proof. In this particular case where $\gamma_- = \gamma_+ = Id$, the reconstruction function is such that $C_0(u,v) = w$ and thus $w^n_{j-1/2+} = w^n_{j+1/2-} = w^n_j$. Let consider the Riemann problem at interface $x_{j+1/2}$ (with $v_{j+1/2}$ constant) and denote $w(\cdot, w^n_j, w^n_{j+1})$ its selfsimilar solution that satisfies under

$$TV(w(\cdot, w^n_j, w^n_{j+1})) = |w^n_{j+1} - w^n_j|.$$

Under the CFL restriction $\Delta t/\Delta x \max h'(w,v) \leq 1/2 - \epsilon$, one gets for the solution $\tilde{w}_h(t^{n+1-}, \cdot)$ obtained before the projection step after time Δt

$$\begin{aligned} TV(\tilde{w}_h(t^{n+1-}, \cdot)) &= \sum_j TV_{[x_j + \epsilon\Delta x/2, x_{j+1} - \epsilon\Delta x/2]}(\tilde{w}_h(t^{n+1-}, \cdot)) \\ &\quad + \sum_j TV_{]x_j - \epsilon\Delta x/2, x_j + \epsilon\Delta x/2[}(\tilde{w}_h(t^{n+1-}, \cdot)) \\ &= \sum_j TV(w(\cdot, u^n_j, w^n_{j+1})) \\ &\quad + \sum_j TV_{]x_j - \epsilon\Delta x/2, x_j + \epsilon\Delta x/2[}(w_h(t^n, \cdot)) \end{aligned}$$

but the second term is zero because $w^n_{j-1/2+} = w^n_{j+1/2-}$ so that we get

$$TV(\tilde{w}_h(t^{n+1-}, \cdot)) = \sum_j |w^n_{j+1} - w^n_j| = TV(w_h(t^n, \cdot)).$$

Let denote $\mathbb{P}_{\Delta x}$ the projection step:

$$w_h(t^{n+1}, \cdot) = \mathbb{P}_{\Delta x}(\tilde{w}_h(t^{n+1-}, \cdot)),$$

that means, for $x \in (x_{j-1/2}, x_{j+1/2})$

$$w_h(t^{n+1}, x) = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} \tilde{w}(t^{n+1-}, y)\, dy.$$

The operator $\mathbb{P}_{\Delta x}$ is total variation diminishing, we thus get

$$\begin{aligned} TV(w_h(t^{n+1}, \cdot)) &\leq TV(\tilde{w}_h(t^{n+1-}, \cdot)) \\ &\leq TV(w_h(t^n, \cdot)). \end{aligned}$$

Using a classical compactness argument, we thus get the convergence of (a subsequence) □

4.6 Numerical experiments

All computations in this section are realized over the spatial domain $[-1,1]$ and the interface function v is chosen as an erf function with given "thickness" η according to the following

$$v(x) = \frac{\mathrm{erf}(x/\eta)+1}{2}. \tag{4.43}$$

Coupling condition For the first test (Fig. 4.2), we consider the coupling condition $w(t,0^-) = w(t,0^+)^2$ with initial data $w^0 \in \mathbb{R}^+$ so that $C_0(\cdot,v)$ is an admissible change of variable from $\mathbb{R}^+$ onto $\mathbb{R}^+$. The coupled fluxes are $f_-(w) = f_+(w) = w^2/2$ and w^0 is chosen as the constant function that equals 2. The numerical solution is computed with $N = 500$ points and $\eta = 2.10^{-2}$. Observe that as the time evolves ($t = 0.3$), an intermediate state $w_+ = \sqrt{2}$ appears at the right of the interface so that the expected coupling solution is satisfied and a rarefaction wave for the right-flux consequently develops that connects w_+ to the initial value on the right. In the right figure that presents the u-component, this rarefaction wave is alone because u becomes constant inside the interface.

For this specific example, the characteristics for both left and right advection problems are going from the left to the right and thus the numerical solution is in accordance with this property. For more complex coupling problems, resonance may occurs at the interface. We will consider such a situation in later examples.

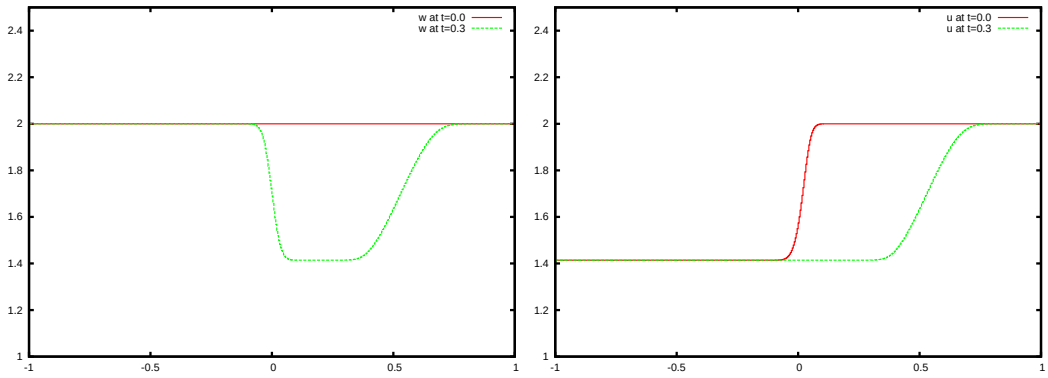


Figure 4.2: Non-trivial coupling condition : w (left) and u (right)

Convergence illustration This numerical test consists in illustrating the convergence rate of the scheme towards the expected solution. We consider the following fluxes $f_-(w) = w^2/2$ and $f_+(w) = (w-1)^2/2$. The initial data is discontinuous at the interface $x = 0$ with Riemann data $(w_\ell, w_r) = (-1.0, 1.5)$ and the expected coupling condition writes $w(t,0^-) = w(t,0^+)$. The solution to the corresponding thick coupling problem consists in two rarefaction waves "sticked" to the interface and with traces at the left (respectively at the right) of the interface that corresponds to the sonic point for the left (resp. right) flux (see Fig.4.3). Obviously the expected coupling condition is then not satisfied in a strong sense.

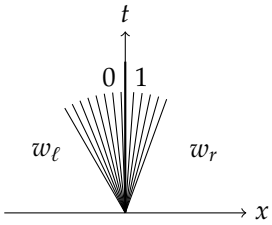


Figure 4.3: Structure of the solution for the thick coupling

Two different values for η are considered : $\eta = 0.01$ (Fig. 4.4) and $\eta = 0.001$ (Fig. 4.5). We observe in each case the convergence of the numerical solution to the expected double-rarefaction. Naturally, the more η is small, the more N needs to be high in order to have a sufficiently "smooth" numerical representation of the interface and a correct solution.

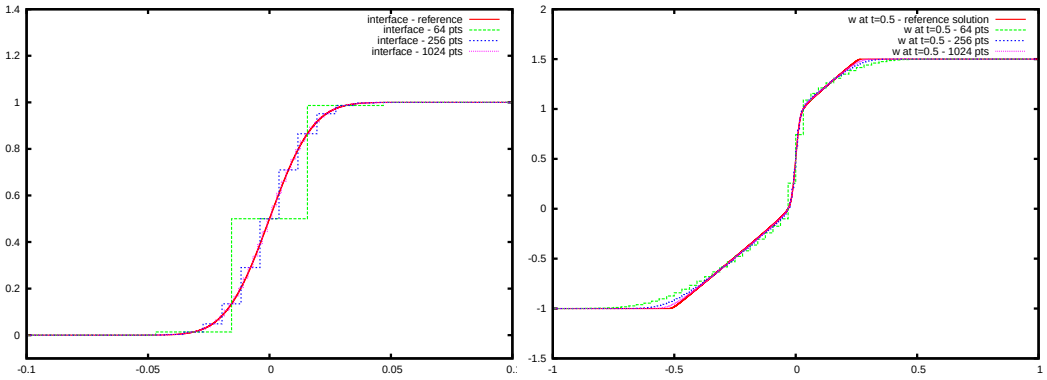


Figure 4.4: Interface profile (left) - Numerical solution (right) - $\eta = 0.01$

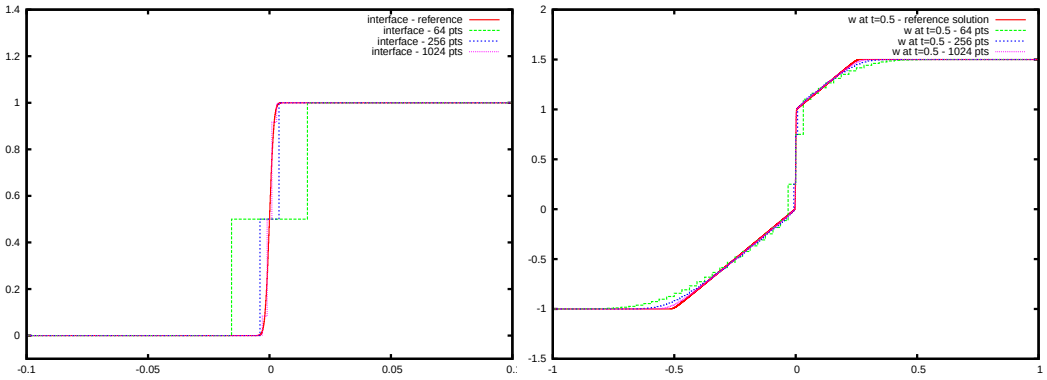


Figure 4.5: Interface profile (left) - Numerical solution (right) - $\eta = 0.001$

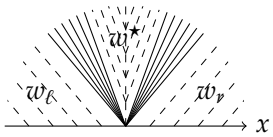
The Table 4.1 presents the L^1-errors comparing to a reference solution obtained with $N = 2^{16} = 65536$, for respectively $\eta = 0.01$ (left) and $\eta = 0.001$ (right).

N	err_{L^1} ($\times 10^{-2}$)	order
64	6.002	/
128	3.448	0.80
256	2.018	0.77
512	1.169	0.79
1024	0.665	0.81
2048	0.370	0.85
4096	0.199	0.89

N	err_{L^1} ($\times 10^{-2}$)	order
64	6.482	/
128	3.798	0.77
256	2.272	0.74
512	1.348	0.75
1024	0.749	0.85
2048	0.415	0.85
4096	0.223	0.89

Table 4.1: L^1-error for $\eta = 0.01$ (left) and $\eta = 0.001$ (right)

Non-uniqueness for the coupling problem and sensitiveness to the interface profile Next, we consider a coupling problem for which the uniqueness in the limiting solution fails. The flux functions are $f_-(w) = w^2/2$ and $f_+(w) = (w+1)^2/2$, the initial data is $(w_\ell, w_r) = (-1.0, 1.5)$ and the coupling condition writes $w(t,0^-) = w(t,0^+)$. Due to the outgoing characteristics at the interface, possible solutions for this coupling problem consists in two rarefaction waves with a constant intermediate state $w^\star$ that belongs to $[w_\ell, w_r]$.



$w_\star \in [w_\ell, w_r]$

Figure 4.6: Possible solutions for the thick coupling

The interface profile is slightly modified in order to add a kind of asymmetry inside, we thus give us ζ small enough and consider

$$v(x) = \frac{\mathrm{erf}(x/\eta + \zeta) + 1}{2}. \tag{4.44}$$

In Figure 4.7 and Figure 4.8, we use respectively 100 and 1000 points for the computations, $\eta = 5.10^{-3}$ and three different values for the parameter ζ are chosen: -0.5, 0.0 and 0.5. We observe a different intermediate state $w^\star$ that depends on ζ.

In the next Figure 4.9, with 5000 points, we fix $\zeta = 0.5$ and observe the convergence of the numerical solution as the thickness of the interface η tends to 0. We observe the intermediate state $w^\star$ does not depends on η but only on ζ (or here equivalently of $v(0)$).

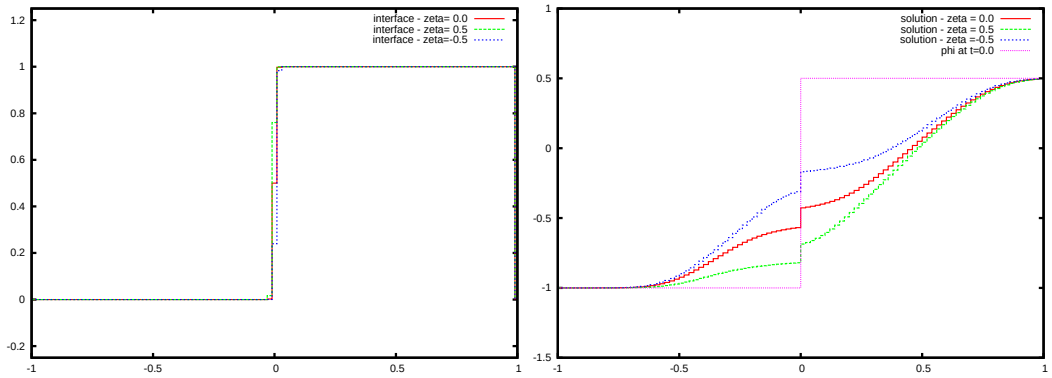


Figure 4.7: Three different interfaces (left) and corresponding solutions (right) - $N = 100$

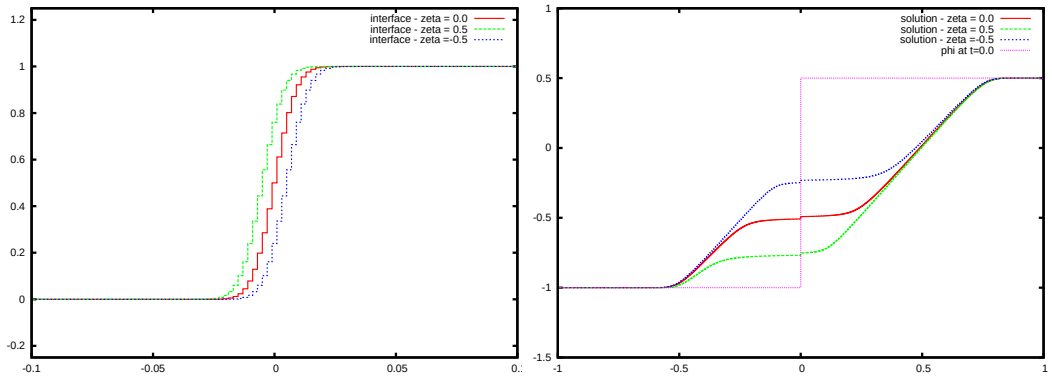


Figure 4.8: Three different interfaces (left) and corresponding solutions (right) - $N = 1000$

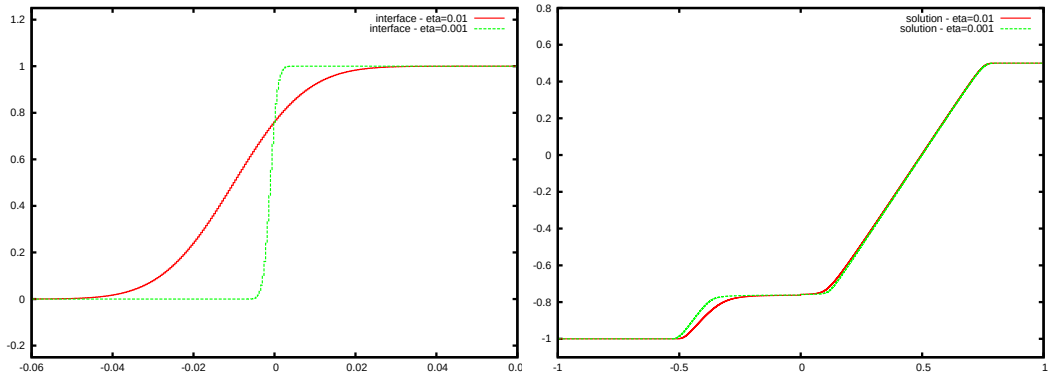
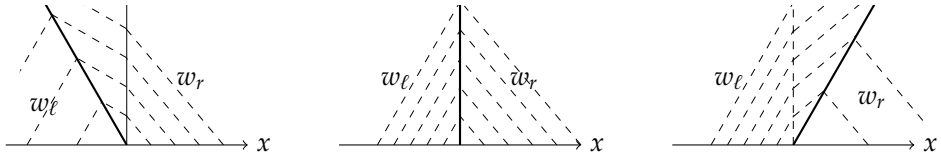


Figure 4.9: Thickness of the interface and convergence.

Shock-shock solutions Finally, this test provides an illustration for another non-uniqueness phenomenon, that occurs for example in the following coupling problem. The flux functions are $f_-(w) = w^2/2$ and $f_+(w) = (w+1)^2/2$ and the initial data is $(w_\ell, w_r) = (1.0, -2.0)$. The parameter are $\eta = 5.10^{-3}$ and three different values for the parameter ζ are chosen: -0.5, 0.0 and 0.5. The solutions are represented on the Figure 4.11. In this situation, we get either a shock for the left-flux if $\zeta < 0$ (Fig. 4.10a), or a shock for the right-flux if $\zeta > 0$ (Fig. 4.10c) or finally a steady shock if $\zeta = 0$ (Fig. 4.10b). It is noticeable that the latter is very sensitive to the value of the initial data. There is here a perfect symmetry of the problem in the sense that $f'_-(w_L) + f'_+(w_R) = 0$. For more general situations, the standing wave does not appear with $\zeta = 0$.



(a) A left-shock *(b) A steady shock* *(c) A right-shock*

Figure 4.10: Possible solutions for the thick coupling

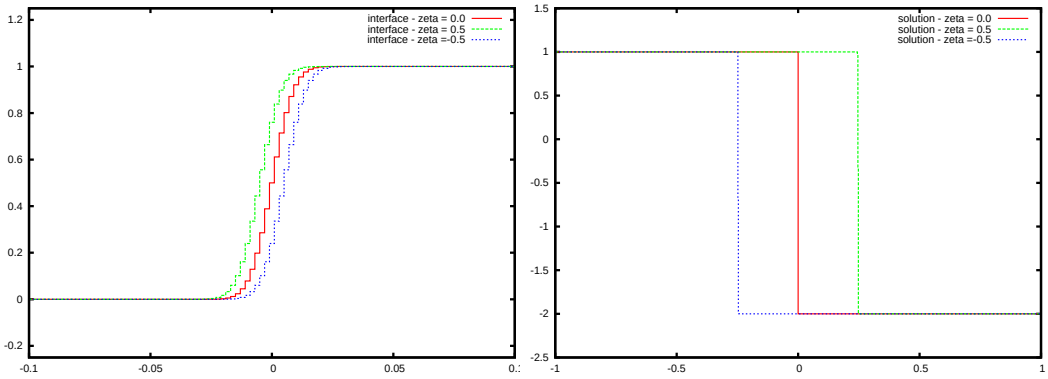


Figure 4.11: Three different interfaces (left) and corresponding solutions (right) - $N = 1000$

5 A MULTIDIMENSIONAL FINITE VOLUME FRAMEWORK

Dans les chapitres précédents, nous nous sommes intéressés au couplage d'équations hyperboliques non-linéaires. Ce couplage a été abordé à travers un système EDP augmenté, évitant la description géometrique directe des interfaces. Dans ce contexte, nous avons obtenu unicité pour le problème de Cauchy, au moins dans le cas d'une variable d'espace mono-dimensionnelle et du couplage de deux lois de conservation scalaires à une interface donnée. Dans ce chapitre, nous étendons cette approche au cas d'équations posées en plusieurs variables d'espace. Le formalisme de système EDP augmenté se révèle être assez flexible dans la mesure où il permet le couplage de lois de conservation avec possible recouvrement. L'existence et l'unicité de la solution du problème de Cauchy couplé est alors démontrée (pour des données initiales dans L^∞). La contribution majeure de ce chapitre réside dans l'élaboration et l'analyse d'un schéma de volumes finis équilibre qui concerne des triangulations générales et dont la convergence forte est établie en étendant le théorème de Coquel, Cockburn et LeFloch (valable pour une seule loi de conservation sans couplage).

Le travail présenté dans ce chapitre a été réalisé en collaboration avec Frederic Coquel et Philippe G. LeFloch et fait l'objet de la publication en préparation [35].

Coupling nonlinear hyperbolic equations (IV). A multidimensional finite volume framework

Abstract

This is a continuation of a series of papers devoted to the coupling of nonlinear hyperbolic equations. Such a coupling is performed in terms of an augmented PDE's system avoiding the direct geometric description of the interfaces. This strategy allowed us to develop an effective regularization technique based on thick interfaces. In this setting, uniqueness was obtained for the Cauchy problem, at least in the case of one space variable and the coupling of two scalar conservation laws at a given interface. In the present paper, we extend this framework and are able to encompass equations in several space dimensions. Our formulation for augmented PDE's systems turns out to be fairly flexible in that it allows the coupling of several distinct conservation laws with possible covering. The existence and uniqueness of the solution of the coupled Cauchy problem (for initial data in L^∞) follows easily. Our main contribution in this paper is the design and analysis of a well-balanced finite volume method which applies to general triangulations and whose strong convergence is established by extending Coquel, Cockburn, and LeFloch's theorem (for a single conservation laws without coupling).

5.1 Introduction

The present work deals with the coupling of distinct hyperbolic equations formulated in a partition of the physical domain. The main motivation stems from the study of complex systems resulting from the combination of elementary components modeled by different equations. Indeed, each of these components may be subject to phenomena involving fairly different time scales and space scales. Tackling this multiscale issue with accuracy and efficiency requires to consider distinct physical settings in the description of each component so as to end up with a sharp description of the whole operating system. Large scale powerplant provide a typical major example. Describing the time behaviour of the whole system therefore requires to exchange of transient informations at each of the physical boundary separating two discinct hyperbolic equations. These transient data are refered hereafter to as *coupling conditions*.

This problem seems to be rather new in the applied mathematic community: the pioneering contributions have been given by Godlewski and Raviart in a first work [76] in 2004. Devoted to the scalar setting in one space variable, these authors model the coupling problem in terms of two initial boundary value problems (IBVP) to be equipped with coupled boundary conditions at a given infinitely thin interface. These boundary conditions are stated in such a way that in most cases they resume to the continuity of the main unknown. The proposed continuity condition is verified, roughly speaking, as long as no wave from the left and right problems interact at the interface ; otherwise the interface is said to be resonant.

A subsequent serie of works by Ambroso *et al.* has proved that many other continuity conditions built on general nonlinear transform of the main unknown were also admissible candidates. An additional information coming from the physics is therefore mandatory in order to single out the relevant continuity condition or say transmission condition at a given interface. Various transmission conditions have been investigated in several physical settings ranging from the gaz dynamic equations in lagrangian coordinates [7] to euleriann ones, for polytopic gazes [12] to multiphase flows [6, 8].

Typically in the setting of two Euler equations with distinct pressure laws, one may wish to impose the continuity of the density, velocity, and pressure, or by contrast, the continuity of the main unknown $(\rho, \rho u, \rho E)$. Let us emphasize that the particular choice dictates the definition of the time and space constant solutions. In the previous examples, the constant solutions are either with constant density, velocity, and pressure, or by contrast with constant density, momentum, and total energy. Notice that in both cases, the proposed coupling conditions are non-conservative in the usual sense: the total mass of density, momentum, and total energy do vary with time. In [11], suitable relaxation procedure are proved to handle in the proposed general framework a fully conservative coupling. In this sense, the framework introduced in [76] and further developped in [7] and the subsequent papers generalizes the conservative coupling of two equations analyzed in the past decade [**?**] to a general setting motivated by the coupling of complex computing plateforms.

The above mentionned resonance phenomena likely to take place at thin interfaces, brings a main difficulty in the mathematical analysis in the coupled initial boundary value problems. Coupled solutions can be shown to exist under general conditions but resonance generally comes at the expense of uniqueness. We refer the reader to the work by Boutin *et al.* [30] in the scalar setting. Let us also quote the work [30] for a distinct behaviour due to the fact that interface is characteristic and not resonant. A selection criterion for discontinuous solutions at the interface is therefore required. Let us recall that in the fully conservative coupling setting several disctinct entropy criterion have been proposed, each of these criteria selects a distinct solution but in agreement with the physical context (see [36] for a short review and also [16, 82, 121]).

In the setting of general transmission conditions we have to deal with, a macroscopic selection principle like entropy inequalities does not seem to be available. The detailed description of microscopic mechanisms coming from suitable regularizing procedures seems to be needed. In the pioneering works [31] and [4], we have introduced an alternative modeling of the coupling problem for two hyperbolic equations in one space variable. This alternative relies on the introduction of an augmented PDE formulation that avoids the geometric description of the interfaces. The proposed formalism considers an additional unknown, the so-called *color function* taking values in the range [0,1]. Extreme values 0 and 1 are devoted to restore the left and right problems to be coupled, while intermediate values may be understood as a shift from one problem to the other.

The interest in this augmented formulation comes from its very capability to support various regularization mechanisms. A first mechanism has been built from a suitable viscous perturbation introduced in [32, 33] in the scalar setting. Dafermos self-similar

approach has allowed to study existence and uniqueness of the coupled Riemann solutions in the limit of a vanishing viscosity. These works have been extended to the case of systems in [32]. Existence of solutions is established there under fairly general assumptions. In [33], the precised analysis of the internal structure of resonant interfaces has allowed to characterize the cimplete set of admissible Riemann solutions. Despite of the viscous mechanisms a failure of uniqueness may be observed for resonant infinitely thin interfaces.

The origin of multiple self-similar solutions may be found in the mathematical interpretation of the Riemann problem proposed by Dafermos [58]. Riemann solutions may be indeed understood as the long time asymptotics of the solutions of the Cauchy problem. Failure of uniqueness for thin interfaces just reflects the property that distinct regularizations of thin interfaces may give rise to different solutions and thus with a distinct long time behaviour. This observation has motivated a second regularization procedure based on *thick interfaces*.

Thickened interfaces are readily obtained in the augmented PDE framework when considering a given regularization of the discontinuous color function considered in the thin regime. This natural approach has been introduced [34] within the frame of two conservation laws in a single space variable. Existence and uniqueness for the coupled Cauchy problem is proved in the general setting of initial data with bounded sup-norm. One of the main ingredients relies on the derivation of a well-balanced property voume method. The well-balanced property means that the constant solutions privileged by a given transmission condition are exactly preserved, and this whatever is the precise definition of the regularized form of the color function. This consistency property is of central importance.

In the present paper, we show how to extend the proposed formalism to coupled problems in several space dimensions. The coupling of several distinct hyperbolic equations can be addressed with a possible covering. Roughly speaking, a vector-valued color function can be conveniently involved so that each component represents a given equation. Each component again takes values in the interval $[0,1]$ and the precise definition of smooth-version of the color function models transition from an equation to the others (possibly more than one). We prove existence and uniqueness of the coupled Cauchy problem with initial data in L^∞ under fairly general assumptions for the transmission conditions and the equations under consideration. Again we propose a constructive proof based on a robust and flexible finite volume framework based on general triangulations. By construction, the proposed method is well-balanced. The strategy for achieving the well-balanced property has strong connection of the subcell reconstruction approach introduced in a work by Jin and Glimm [**?**] and analyzed by Bouchut in other settings in his monograph [28].

The format of the paper is as follows. In a first section, we show how to extend the two existing coupling framework discussed above to the coupling of two distinct hyperbolic equations but in several space dimensions. We then show how to extend the augmented PDE formalism to encompass the case of several hyperbolic equations with possible covering. In a second section we introduce the well-balanced finite volume method we intend to analyze. The required well-balanced property is easingly achieved when considering

two distinct meshes. The first one, the so-called primal mesh, is devoted to describe the main coupled unknown. The second mesh, referred as to the dual one, and inferred to the primal mesh, is devoted to represent the discrete color function. A comprehensive derivation of this dual mesh is in particular proposed. We will then prove for validity an uniform sup-norm estimate. Due to the subcell reconstruction procedure inherent to the present approach (but also to the general trangulation under consideration) uniform BV-estimates seems to be out of reach. Instead, we propose to use the DiPerna framework for entropy measure-valued solutions to establish the *a posteriori* strong convergence of the method. Some numerical illustrations involving problem with covering highlight the interest of the proposed coupling strategy.

5.2 Notation and objectives

5.2.1 First considerations

In this section, we introduce the coupling problematic when investigating the simplest setting of two hyperbolic equations to be coupled at a given interface (it suffices to think of an hyperplane at this stage, say $\{x_1 = 0\}$). In this aim, we will extend two distinct coupling strategies that have been developped in a single space variable. The first procedure consists in understanding the coupling problem as two initial boundary value problems (IBVP) with time dependant boundary conditions prescribing the evolution of traces of the coupled solutions on both sides of the hyperplane $\{x_1 = 0\}$. In deep contrast, the second strategy, recently introduced in [32–34], is grounded on augmented PDE systems, allowing to handle the coupling problem as an initial data problem written over the entire space $\mathbb{R}^d$. This new framework brings mathematical and numerical advantages, briefly pointed out at the end of this section and that will motivate a natural extension to much more general coupling issues.

The coupling problem modeled as a collection of interacting IBVP

Let us first consider an hyperplane of $\mathbb{R}^d$ with unit normal vector $\nu \in \mathbb{R}^d$, we denote $\mathcal{H} = \{x \in \mathbb{R}^d / x.\nu = 0\}$, partitioning $\mathbb{R}^d$ into two half-domains $\mathcal{D}_- = \{x \in \mathbb{R}^d / x.\nu < 0\}$ and $\mathcal{D}_+ = \{x \in \mathbb{R}^d / x.\nu > 0\}$. In each of these open subdomains, a distinct conservation law is prescribed:

$$\partial_t w + \sum_{i=1}^{d} \partial_{x_i} a_i^{\pm}(w) = 0, \qquad w(t,x) \in \mathbb{R}, \quad t > 0, \quad x \in \mathcal{D}_{\pm}, \tag{5.1}$$

where the flux-functions $A^{\pm} : \mathbb{R} \to \mathbb{R}^d$, with components $(a_i^{\pm})_{i=1,\dots,d}$, are assumed to be twice differentiable for definiteness. An initial data $w(0,x) = w_0(x)$ supplements this formulation, but obviously, some extra-condition, the so-called *coupling condition*, must be prescribed at the interface $\mathcal{H}$. For simplicity, we restrict ourselves in this introductory section to piecewise smooth solutions w with bounded left and right traces at the interface $\mathcal{H}$:

$$w(t, y^{\pm}) \equiv \lim_{z \to 0^+} w(t, y \pm z\nu), \quad y \in \mathcal{H}. \tag{5.2}$$

Then, it sounds natural that the coupling condition we seek should relate these traces

$$\mathfrak{C}(w(t,y^-),w(t,y^+))=0,\quad t>0,\ y\in\mathcal{H}, \tag{5.3}$$

for some nonlinear mapping $\mathfrak{C}$ to be specified. The implicit function theorem is assumed to apply so as to recast (5.3) in the more tractable form:

$$w(t,y^-)=\mathfrak{c}(w(t,y^+)),\quad t>0,\ y\in\mathcal{H}, \tag{5.4}$$

for some function $\mathfrak{c}$ mapping $\mathbb{R}$ onto $\mathbb{R}$. Assuming from now on $\mathfrak{c}$ to be strictly monotone, it turns more convenient to reexpress this above coupling condition in terms of a pair of nonlinear monotone functions, say θ_- and θ_+ with $\mathfrak{c}=\theta_-^{-1}\circ\theta_+$:

$$\theta_-(w(t,y^-))=\theta_+(w(t,y^+)),\quad t>0,\ y\in\mathcal{H}. \tag{5.5}$$

without loss of generality, θ_- and θ_+ are assumed to be strictly increasing and to map $\mathbb{R}$ onto $\mathbb{R}$. Their respective inverses are denoted by γ_- and γ_+. On the basis of this pair of functions, we introduce the following usefull change of unknown:

$$u(t,x)\equiv\begin{cases}\theta_-(w(t,x)), & t>0,\ x\in\mathcal{D}_-,\\ \theta_+(w(t,x)), & t>0,\ x\in\mathcal{D}_+,\end{cases} \tag{5.6}$$

so that the coupling condition (5.5) resumes to:

$$u(t,y^-)=u(t,y^+),\quad y\in\mathcal{H}, \tag{5.7}$$

namely to a continuity condition for the new unknown u. After the works [76] and [73], a coupling condition in the form (5.5) or equivalently (5.7) is referred to as a "u-transmission" condition.

It is worth underlining that the last formulation (5.7) obviously defines what are the constant solutions of the coupling problem (5.1)-(5.5), i.e. time independant functions $w(x)$ which solve (5.1) and (5.7). Such functions clearly obey:

$$u(w(x))=u^\star,\quad x\in\mathbb{R}^d\setminus\mathcal{H}, \tag{5.8}$$

for some real number $u^\star\in\mathbb{R}$.

This seemingly obvious remark actually just open the gate to the mathematical study of perturbed solutions built from the trivial solution (5.8). We refer the reader to the work [32] devoted to the existence of self-similar coupled solutions for systems.

Observe that the coupling condition (5.5) plays the role of a pair of transient boundary conditions for the interface $\mathcal{H}$. In other words, the coupling framework we address merely takes the form of two nonlinear hyperbolic IBVP linked via the transient boundary condition (5.5). With this respect, it becomes clear that the coupling conditions (5.5) is actually expressed in a strong sense, since it is formulated without reference to the signature

of the wave speeds at the interface $\mathcal{H}$. It is nevertheless well-known that the sign of the wave velocities at a boundary directly affects the condition to be prescribed. Hence, the coupling condition (5.5) or its equivalent form (5.7) must be given a weak formulation.

In that aim, we adopt in this paragraph a straightforward extension of a series of works devoted to coupled problem in one space dimension, initiated by Godlewski and Raviart [73], [76] [1]. In these works, weak form of the coupling condition (5.5) was derived in the spirit of the admissible set of boundary condition due to Dubois and LeFloch [65] and based on the notion of Riemann problem. Such a notion here readily extends since the coupling condition expressed in (5.5) just links the traces of the coupled solution w in the normal direction ν and thus essentially concerns the quasi-one dimensional form of (5.5) written for plane wave solution in the ν-direction. Thus it turns natural to consider the coupled problem in one space variable (up to some convenient shift in the space variable z)

$$\partial_t w + \partial_z A_\nu^\pm(w) = 0, \quad t > 0, \ \pm z > 0, \tag{5.9}$$

where we have set

$$A_\nu^\pm(w) \equiv A^\pm(w) \cdot \nu. \tag{5.10}$$

In order to state the weak form of the boundary condition $\theta_-(w(t,y^-)) = \theta_+(w(t,y^+))$, $y \in \mathcal{H}$, we first recall the Dubois-LeFloch framework for say the right IBVP:

$$\partial_t w + \partial_z A_\nu^+(w) = 0, \quad t > 0, \ z > 0, \tag{5.11}$$

$$w(t,0^+) = b, \quad t > 0, \tag{5.12}$$

for some given prescribed real number b. Following Dubois and LeFloch, a convenient weak formulation of (5.12) may be expressed in term of Riemann solutions for (5.11), which we denote $\mathcal{W}(\cdot; w_L, w_R)$, for given left and right states w_L, w_R. This weak form reads:

$$w(t,0^+) \in \mathcal{O}_\nu^+(b) \equiv \Big\{ \mathcal{W}(0^+; b, w), w \in \mathbb{R} \Big\}. \tag{5.13}$$

Observe that the analogous of (5.13) for the left IBVP built from A_ν^- would read

$$w(t,0^-) \in \mathcal{O}_\nu^-(b) \equiv \Big\{ \mathcal{W}(0^-; w, b), w \in \mathbb{R} \Big\}. \tag{5.14}$$

These considerations naturally yield the following weak formulation of the coupled boundary conditions (5.5) at any given point $y \in \mathcal{H}$:

$$\begin{array}{l} w(t,y^+) \in \mathcal{O}_\nu^-(\theta_+^{-1} \circ \theta_-(w(t,y^-))), \\ w(t,y^-) \in \mathcal{O}_\nu^+(\theta_-^{-1} \circ \theta_+(w(t,y^+))), \end{array} \quad t > 0. \tag{5.15}$$

This simple setting of two coupled equations at a given hyperplane can be easily extended to more general interfaces resulting from a partition of $\mathbb{R}^d$ into two non-overlapping open

[1] A joint research program on multiphase flows between CEA (French center of nuclear research) and University Pierre et Marie Curie-Paris6 (Laboratoire Jacques-Louis Lions) in the frame of the Neptune project [78] has produced several studies in the subject. See for instance [7, 8, 49] and the references therein

sets $\mathcal{D}_+$ and $\mathcal{D}_-$ such that $\overline{\mathcal{D}_-} \cup \overline{\mathcal{D}_+} = \mathbb{R}^d$, separated by a smooth boundary $\partial D = \overline{\mathcal{D}_-} \cap \overline{\mathcal{D}_+}$. Smoothness allows to define without ambiguity an unit normal vector $\nu(y)$ for all $y \in \partial D$ so that left and right traces at ∂D for piecewise smooth solutions of the coupled problem (5.5) may be defined as follows:

$$w(t, y^{\pm}) \equiv \lim_{z \to 0^+} w(t, y \pm z\nu(y)), \quad y \in \partial\mathcal{D}. \tag{5.16}$$

The expected coupling condition just takes the weak form (5.15).

The coupling problem as an augmented PDE system

As already emphasized, an alternative coupling framework has been introduced in [32]. Instead of dealing with two IBVP to be coupled at a given interface in terms of boundary conditions, this new setting proposes to treat the coupling problem as a single initial value problem set over the entire space $\mathbb{R}^d$, thanks to an augmented PDE formulation. The latter has been introduced in [32–34] for problems in a single space dimension. In order to encompass the setting of several space variables, we perform hereafter a comprehensive derivation of the proposed framework.

The derivation starts from the characteristic functions of the two open sets $\mathcal{D}_-$ and $\mathcal{D}_+$, we denote by:

$$v_- = \chi_{\mathcal{D}_-}, \quad v_+ = \chi_{\mathcal{D}_+}. \tag{5.17}$$

It heavily makes use of the change of unknown u introduced in (5.6), we rephraze as:

$$u(t,x) = \begin{cases} \theta_-(w(t,x)), & \text{if } v_-(x) = 1, \\ \theta_+(w(t,x)), & \text{if } v_-(x) = 0, \text{ i.e. if } v_+(x) = 1, \end{cases} \qquad t > 0,\ x \notin \partial\mathcal{D}. \tag{5.18}$$

Equipped with these notations, it is convenient to recast the two distinct hyperbolic equations respectively in $\mathcal{D}_-$ and $\mathcal{D}_+$ in terms of u:

$$\gamma'_{\pm}(u)\partial_t u + \sum_{i=1}^{d} \gamma'_{\pm}(u) a_i^{\pm\prime}(\gamma_{\pm}(u))\partial_{x_i} u = 0, \quad t > 0,\ x \in \mathcal{D}_{\pm}, \tag{5.19}$$

restricting ourselves to smooth solutions in a first step. Recall that γ_+ (resp. γ_-) denotes the inverse function of θ_+ (resp. θ_-). Let us further proceed rewritting the above two equations in terms of a single equation expressed for x in $R^d \setminus \partial D$:

$$\begin{aligned} &\Big(v_-\gamma'_-(u) + v_+\gamma'_+(u)\Big)\partial_t u \\ &\qquad + \sum_{i=1}^{d} \Big(v_-\gamma'_-(u)a_i^{-\prime}(\gamma_-(u)) + v_+\gamma'_+(u)a_i^{+\prime}(\gamma_+(u))\Big)\partial_{x_i} u = 0, \end{aligned} \qquad t > 0. \tag{5.20}$$

At this stage, it must be noticed that the two characteristic functions v_- and v_+ in the above equation may be replaced by a single function say v, setting for instance $v_-(x) = 1 - v(x)$ and $v_+(x) = v(x)$ for $x \in \mathbb{R}^d \setminus \partial D$ with $v = \chi_{\mathcal{D}_+}$. In the sequel, such a function v will be refered

to as a *color function*. For the moment v is nothing but a step function taking values in $\{0,1\}$ but it is important to conceive v as a function taking values in the interval $[0,1]$ so that the value 0 restores the equation set in $\mathcal{D}_-$ while the value 1 restores the equation set in $\mathcal{D}_+$. Intermediate values of v then may be thought as modelling a smooth shift from one problem to the other. Keeping this in mind we now recast (5.20) in the form of an augmented PDE system with unknow u and v, for $t>0$ and $x \in R^d \setminus \partial D$:

$$\begin{cases} \Big((1-v)\gamma'_-(u)+v\gamma'_+(u)\Big)\partial_t u \\ \qquad +\Big((1-v)\gamma'_-(u)\nabla A^-(\gamma_-(u))+v\gamma'_+(u)\nabla A^+(\gamma_+(u))\Big)\cdot\nabla_x u = 0, \\ \partial_t v = 0. \end{cases} \tag{5.21}$$

We stress that the 1-dimensional form of these equations written for plane wave solutions in the direction ν writes for $t>0$ and $x \in R^d \setminus \partial D$ (or $\pm z>0$):

$$\begin{cases} \Big((1-v)\gamma'_-(u)+v\gamma'_+(u)\Big)\partial_t u \\ \qquad +\Big((1-v)\gamma'_-(u)\nabla A^-(\gamma_-(u))\cdot\nu+v\gamma'_+(u)\nabla A^+(\gamma_+(u))\cdot\nu\Big)\partial_z u = 0, \\ \partial_t v = 0. \end{cases} \tag{5.22}$$

This system is easily seen to be hyperbolic if (and only if) the following quantity is not zero

$$(1-v)\gamma'_-(u)\nabla A^-(\gamma_-(u))\cdot\nu+v\gamma'_+(u)\nabla A^+(\gamma_+(u))\cdot\nu \neq 0. \tag{5.23}$$

For such states, the standing wave associated with the additional unknown v can be seen to admit u as a Riemann invariant. In other words, as long as the non-degeneracy condition (5.23) is valid, u stays continuous at the jumps of the color function v, namely accross the coupling boundary ∂D at which the value of v shifts from 0 to 1. In other words and whenever (5.23) is valid, the coupling condition (5.7) is verified in the strong sense accross the standing wave

$$u(t,y^-) = u(t,y^+), \quad y \in \partial D. \tag{5.24}$$

Violation of the condition (5.23) at a point of jump for v, namely at the interface $\partial\mathcal{D}$, expresses that waves from the left and right propagate with opposite sign at the interface ; the first order system (5.22) is then only weakly hyperbolic. This is the so-called resonance phenomena and we refer the reader to the recent work by Goatin-LeFloch [71]. As far as the coupling issue is concerned, the continuity condition (5.24) is no longer satisfied and the weak form (5.15) of the coupling condition must be addressed. Turning considering the augmented formulation (5.22), resonance phenomena has been studied in depth in [32] in the scalar setting thanks to a self-similar viscous perturbation . The Riemann solutions for (5.22) defined in the limit of a vanishing viscosity are proved to verify (5.15) when resonance takes place. To sum up, weak solutions of the augmented equations (5.22) and thus their multi-dimensional form (5.21) naturally encode the weak form of the coupling condition.

Here and for the sake of generality, we extend the particular form of the augmented

equation to the general setting addressed in our previous work [32] devoted to systems in one space variable. We thus introduce coupling functions $C_0 : \mathbb{R}\times[0,1] \to \mathbb{R}$ and $C_i : \mathbb{R}\times[0,1] \to \mathbb{R}$ with $i \in \{1,\dots,d\}$ satisfying the following consistency properties:

$$\begin{array}{ll} \lim_{v\to 0} C_0(u,v) = \gamma_-(u), & \lim_{v\to 1} C_0(u,v) = \gamma_+(u), \\ \lim_{v\to 0} C_i(u,v) = a_i^-(\gamma_-(u)), & \lim_{v\to 1} C_i(u,v) = a_i^+(\gamma_+(u)), \end{array} \tag{5.25}$$

so as to consider in place of (5.21) the more general augmented equations:

$$\begin{array}{ll} \partial_u C_0(u,v)\partial_t u + \sum_{i=1}^d \partial_u C_i(u,v)\partial_{x_i} u = 0, & t>0,\ x\in\mathbb{R}^d, \\ \partial_t v = 0, & \end{array} \tag{5.26}$$

which equivalently recasts as:

$$\begin{array}{ll} \partial_t C_0(u,v) + \sum_{i=1}^d \partial_{x_i} C_i(u,v) - \sum_{i=1}^d \partial_v C_i(u,v)\partial_{x_i} v = 0, & t>0,\ x\in\mathbb{R}^d. \\ \partial_t v = 0. & \end{array} \tag{5.27}$$

Remark 5.1. A particular example of coupling functions verifying the above conditions is obviously given by

$$\begin{aligned} C_0(u,v) &= (1-v)\gamma_-(u) + v\gamma_+(u), \\ C_i(u,v) &= (1-v)a_i^-(\gamma_-(u)) + v a_i^+(\gamma_+(u)),\ 1\le i\le d. \end{aligned} \tag{5.28}$$

In the sequel, the coupling functions C_0 and $(C_i)_{1\le i\le d}$ will assumed to be smooth with:

$$C_0, (C_i)_{1\le i\le d} \in C^2(\mathbb{R}\times[0,1]), \tag{5.29}$$

while C_0 will in addition obey:

$$\partial_u C_0(u,v) > 0, \quad u\in\mathbb{R},\ v\in[0,1]. \tag{5.30}$$

This last assumption is nothing but a non-degeneracy condition for the time arrow in (5.26).

The central interest of the augmented formulation (5.26) over more classical coupling approaches built from a collection of IBVP stems from the fact it can be supplemented with a variety of regularizing mechanisms at the coupling interfaces. These regularization mechanisms are intended to handle the resonance phenomena which is likely to take place at the interfaces. A first regularization procedure relies on introduction of suitable viscous mechanisms. Suche mechanisms yield a non trivial internal structue to resonant interfaces which proves to be usefull in the selection of discontinuous solutions. It turns thant discontinuous solutions may not be unique for thin interfaces. The augmented formulation (5.26) actually allows for another regularization mechanism based on thick

interfaces. Roughly speaking, the color function which is naturally discontinuous in the setting of thin interfaces is given a regularization in the thick regime. Such a regularization technique has been analyzed in one space variable, existence and uniqueness of a solution for the Cauchy problem has been established. In the next section, we show how to extend this regularization procedure to several space dimensions.

5.2.2 Extension to multi-dimensional, multi-component coupling problems

The geometric setting

We are in a position to present the general coupling framework we intend to analyze in this paper. The proposed extension treats the coupling of $(L+1)$, $L \geq 1$, distinct conservations laws in several space dimension, with possible covering. The coupling modelling via augmented PDE equations relies on a partition of the space $\mathbb{R}^d$ in a finite number of non-overlapping, non-empty and open sets $(\mathcal{D}^l)_{0\leq l\leq L}$:

$$\bigcup_{l=0}^{L} \overline{\mathcal{D}^l} = \mathbb{R}^d. \tag{5.31}$$

The set of boundaries $\mathcal{B}$ are given by

$$\mathcal{B} \equiv \bigcup_{k\neq l} \overline{\mathcal{D}_k} \cap \overline{\mathcal{D}_l}. \tag{5.32}$$

An interface $\mathcal{H}_{kl}$ is by definition the part of the boundary of $\mathcal{D}_k$ which is only shared with $\mathcal{D}_l$ (see also Fig. 5.1 for an example with $N = 2$ and $L = 3$):

$$\mathcal{H}_{kl} \equiv (\overline{\mathcal{D}_k} \cap \overline{\mathcal{D}_l}) \setminus \bigcup_{i\neq k,l} \overline{\mathcal{D}_i}. \tag{5.33}$$

These interfaces $\mathcal{H}_{kl}$ are supposed to be smooth enough so that they admit an unit normal vector everywhere. We suppose the set of boundaries $\mathcal{B}$ to be of d-dimensional Lebesgue measure zero, and more precisely the remaining set $\mathcal{B} \setminus (\cup_{k\neq l}\mathcal{H}_{kl})$ has only components of Hausdorff dimension less than or equal to $(d-2)$ (see for example the four points underlined in Figure 5.1).

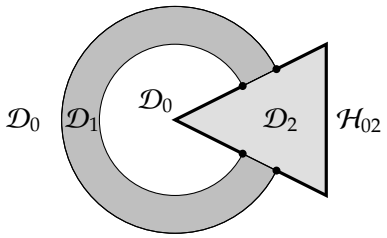


Figure 5.1: Boundaries (in bold-face $\mathcal{H}_{02}$, circle points being excluded)

In each domain $\mathcal{D}_l$, the unknown w is governed by a specific conservation law with

flux-function $A^l \equiv (a_i^l)_{1\le i\le d} : w \in \mathbb{R} \mapsto A^l(w) \in \mathbb{R}^d$:

$$\partial_t w + \sum_{i=1}^{d} \partial_{x_i} a_i^l(w) = 0, \qquad w(t,x) \in \mathbb{R}, \quad t > 0, \quad x \in \mathcal{D}_l. \tag{5.34}$$

Following the description introduced in the previous section, we start focusing the discussion on the definition of constant states (5.6)-(5.7)-(5.8) for the global problem set on the whole space $\mathbb{R}^d$. These solutions are recovered through a certain change of variable in each subdomain $\mathcal{D}_l$, for $l = 0, \dots, L$,

$$u(t,x) \equiv \theta_l(w(t,x)), \quad t > 0,\ x \in \mathcal{D}_l, \tag{5.35}$$

so that the stationnary solutions $w(x)$ for the coupled problem (5.34) are the given real constants $u^\star$ in the u variable:

$$u(w(x)) = u^\star, \quad x \in \mathbb{R}^d \setminus \mathcal{B}. \tag{5.36}$$

The coupling functions θ_l are supposed to map increasingly $\mathbb{R}$ onto itself and we denote once again γ_l their respective inverses.

$$\gamma_l \equiv \theta_l^{-1}, \quad l = 0, \dots, L. \tag{5.37}$$

Let observe importantely that a different outlook where the coupling functions would be associated to the interfaces $\mathcal{H}_{kl}$ rather than to the domains could only be local in space and would not allow an easy matching of local constant solutions so as to define a global constant solution. On the other hand, we take advantage here of the resulting local formulation at each interface in terms of the traces of w, say $w(t,y^k)$ and $w(t,y^l)$ respectively on the $\mathcal{D}_k$- and on the $\mathcal{D}_l$-side of $\mathcal{H}_{kl}$ relatively to its normal vector $\nu_{kl}(y)$. It writes

$$\theta^k(w(t,y^k)) = \theta^l(w(t,y^l)), \quad t > 0,\ y \in \mathcal{H}_{kl}. \tag{5.38}$$

The following augmented PDE formulation is based upon the introduction of a vector-valued *color function* that merges the geometric description of the coupling problem. In this multi-domain setting, this function is based on the set of characteristic functions of each domain:

$$v_0 = \chi_{\mathcal{D}_0},\ v_1 = \chi_{\mathcal{D}_1},\ \dots,\ v_L = \chi_{\mathcal{D}_L}, \tag{5.39}$$

so that the change of variable (5.35) may be also rewritten

$$u(t,x) = \theta_l(w(t,x)), \quad x \in \mathbb{R}^d \setminus \mathcal{B} \text{ s.t. } v_l(x) = 1. \tag{5.40}$$

Obviously, the $(L+1)$ domains being a partition of the whole space, only L of these functions

are usefull to get the description, say $v_1,\dots,v_L$, so that v_0 is finally recovered through

$$v_0(x) = 1 - \sum_{l=1}^{L} v_l(x), \quad x \in \mathbb{R}^d \setminus \mathcal{B}. \tag{5.41}$$

Augmented PDE formulation and thick interfaces

In the sequel we thus make use of the vector-valued *color function* $v = (v_1,\dots,v_L)$ that takes values at this stage in the discrete set $\{0\} \cup \{e_1\} \cup \dots \cup \{e_L\}$ where the e_l stand for the l-th canonical vector of $\mathbb{R}^L$. This color function is ultimately intended to receive a regularized formulation taking values in the convex hull of the above discrete set $\mathbb{B}_+^L = \left\{v = (v_1,\dots,v_L) \in \mathbb{R}^L \middle/ v_l \geq 0, \quad \sum_{l=1}^L v_l \leq 1\right\}$. The problem (5.34) is then understood under the augmented form

$$\begin{aligned} \partial_u C_0(u,v)\partial_t u + \sum_{i=1}^{d} \partial_u C_i(u,v)\partial_{x_i} u = 0, \qquad & t > 0,\ x \in \mathbb{R}^d, \\ \partial_t v = 0, & \end{aligned} \tag{5.42}$$

where the coupling functions C_0 and C_i are assumed to restore the formulation of (5.34) in terms of u in each open set $\mathcal{D}_l$, that is:

$$\begin{aligned} &\lim_{v\to 0} C_0(u,v) = \gamma_0(u), && \lim_{v\to e_l} C_0(u,v) = \gamma_l(u), \\ &\lim_{v\to 0} C_i(u,v) = a_i^0(\gamma_0(u)), && \lim_{v\to e_l} C_i(u,v) = a_i^l(\gamma_l(u)), \quad 1 \leq i \leq d. \end{aligned} \tag{5.43}$$

The following smoothness and monotonicity assumptions are moreover required

$$C_0, C_i \in C^2(\mathbb{R} \times \mathbb{B}_+^L), \tag{5.44}$$

$$\partial_u C_0(u,v) > 0, \qquad u \in \mathbb{R}, \quad v \in \mathbb{B}_+^L. \tag{5.45}$$

This last property ensures the validity of the change of variable $u \mapsto C_0(u,v)$ for any given fixed v, and the non-degenerate nature of the time-arrow in the augmented equations (5.42).

In this context the augmented system writes in the main unknown u as follows:

$$\begin{aligned} \partial_t C_0(u,v) + \sum_{i=1}^{d} \partial_{x_i} C_i(u,v) - \sum_{i=1}^{d}\sum_{l=1}^{L} \partial_{v_l} C_i(u,v)\partial_{x_i} v_l = 0, \\ \partial_t v = 0. \end{aligned} \tag{5.46}$$

In the sequel, it will be usefull to consider the same system written in the variable $w = C_0(u,v)$ (denoted by $w(u,v)$, and with inverse $u(w,v)$ for each fixed v). Equipped with such a change of unknown, (5.46) becomes

$$\begin{aligned} \partial_t w + \sum_{i=1}^{d} \partial_{x_i} f_i(w,v) - \sum_{i=1}^{d}\sum_{l=1}^{L} \ell_i^l(w,v)\partial_{x_i} v_l = 0, \\ \partial_t v = 0, \end{aligned} \tag{5.47}$$

where $f_i(w,v) := C_i(u(w,v),v)$ and $\ell_i^l(w,v) := \partial_{v_l} C_{i|u}(u(w,v),v)$ with $i \in \{1,\dots,d\}$ and $l \in \{1,\dots,L\}$ (i.e. $\ell = \nabla_v C$). Hereafter and to shorten the notations, we will adopt the following condensed form:

$$\begin{aligned} \partial_t w + \nabla \cdot f(w,v) - \ell(w,v) : \nabla v = 0, \\ \partial_t v = 0. \end{aligned} \tag{5.48}$$

Notation rules:
The first one is the tensor product of two vectors of different sizes: let be $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$, $a \otimes b$ denotes the $n \times m$ matrix with components $(a_i b_j)_{i,j}$.
The second one is derived from the canonical scalar product of matrices with compatible dimension. Let m and n be two integers and $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times m}$ two real matrices with size $m \times n$ and $n \times m$, then the real number $A : B$ is given by

$$A : B = \mathrm{tr}(AB) = \sum_{1\le i\le m,\ 1\le j\le n} A_{ij} B_{ji}. \tag{5.49}$$

Equivalently, this quantity is the sum of the vectorial scalar products of each row of A with the corresponding column of B (or the reverse). Observe the usefull identity $A : B = B : A$.
Moreover a noticeable algebraic property is the following

$$A : (a \otimes b) = (Aa) \cdot b. \tag{5.50}$$

For example, focusing on the contracted product $\ell(w,v) : \nabla v$, the matrices under consideration are $\ell = (\ell_i^l)_{1\le i\le d,\ 1\le l\le L}$ of size $d \times L$ and $\nabla v = (\partial_{x_i} v_l)_{1\le l\le L,\ 1\le i\le d}$ of size $L \times d$. Equivalently $\ell(w,v) : \nabla v$ is thus the sum of the scalar product of $(\ell_i^l)_{1\le i\le d}$ with $\nabla v_l = (\partial_{x_i} v_l)_{1\le i\le d}$ for $l = 1,\dots,L$. This quantity appears naturally when differentiating a vector-valued function with respect to a vectorial parameter, here: $\nabla_x C(u,v) = \nabla_u C \cdot \nabla_x u + \nabla_v C : \nabla_x v$.

Entropy like inequalities

As already emphasized, we promote in this work a regularization mechanism based in thick interfaces that are modelled by any suitable regularized version of the discontinuous vector-valued color function v introduced in (5.39)-(5.41). For definiteness, we shall consider color functions v in $W^{2,\infty}(\mathbb{R}^+ \times \mathbb{R}^d, \mathbb{B}_+^L)$. Obviously, it suffices to choose the initial data v_0 in $W^{2,\infty}(\mathbb{R}^d, \mathbb{B}_+^L)$ so as to inherit from the required smoothness in the v solution of the augmented equations (5.48). In turn and arguing about this smoothness property, the equations under consideration merely resume to a non homogeneous scalar equation in the main unkwnown w:

$$\partial_t w + \nabla \cdot f(w, v(x)) = \ell(w, v(x)) : \nabla v(x), \tag{5.51}$$

where the right hand side just plays the role of a classical source term ; namely this term does not contribute to the definition of the possible discontinuities of w. At a point of jump,

(5.51) just resumes to the classical Rankine-Hugoniot condition

$$-\sigma(w^+ - w^-) + \sum_{i=1}^{d} (f_i(w^+, v) - f_i(w^-, v)) = 0. \tag{5.52}$$

A selection criterion of the admissible weak solutions w is of course needed. In that aim, it is convenient to recast the balance law (5.51) expressed in terms of the following non homogeneous formulation promoting the main variable u:

$$\partial_t C_0(u,v) + \sum_{i=1}^{d} \partial_u C_i(u,v) \partial_{x_i} u = 0, \tag{5.53}$$

which we write for smooth solutions. For such solutions, natural additionnal non trivial equations are built from any given (strictly) convex function $\varpi \mapsto \mathcal{U}(\varpi)$ when multiplying (5.53) by $\mathcal{U}'(C_0(u,v))$ to infer:

$$\partial_t \mathcal{U}(C_0(u,v)) + \sum_{i=1}^{d} \partial_u Q_i(u,v) \partial_{x_i} u = 0, \tag{5.54}$$

where

$$Q_i(u,v) = \int^u \mathcal{U}'(C_0(\theta,v)) \partial_\theta C_i(\theta,v) d\theta, \quad 1 \le i \le d. \tag{5.55}$$

We thus get from (5.54) the equivalent form for smooth solutions u:

$$\partial_t \mathcal{U}(C_0(u,v)) + \sum_{i=1}^{d} \partial_{x_i} Q_i(u,v) = \sum_{i=1}^{d} \sum_{l=1}^{L} \partial_{v_l} Q_i(u,v) \partial_{x_i} v_l. \tag{5.56}$$

Observe that the above right hand side is nothing but a classical source term since we again emphasize that the color function v is smooth. As a consequence, the weak form of (5.56) for discontinuous solutions u reads:

$$\partial_t \mathcal{U}(C_0(u,v)) + \sum_{i=1}^{d} \partial_{x_i} Q_i(u,v) \le \sum_{i=1}^{d} \sum_{l=1}^{L} \partial_{v_l} Q_i(u,v) \partial_{x_i} v_l, \tag{5.57}$$

which naturally plays the role of a (non-homogenous) entropy inequality for selecting the relevant weak solutions. Hereafter, we shall make use of the inequalities (5.57) for all convex entropy $\mathcal{U}$. These will be alternatively invoked (essentially when the color function is locally constant) in the w variable:

$$\partial_t \mathcal{U}(w) + \sum_{i=1}^{d} \partial_{x_i} \mathcal{F}_i(w,v) - \sum_{i=1}^{d} \sum_{l=1}^{L} \mathcal{L}_i^l(w,v) \partial_{x_i} v_l \le 0, \tag{5.58}$$

with

$$\mathcal{F}_i(w,v) = Q_i(u(w,v),v), \quad \mathcal{L}_i(w,v) = \partial_v Q_{i|_u}(u(w,v),v), 1 \le i \le d. \tag{5.59}$$

To shorten the notations equation (5.58) are given the following condensed form:

$$\partial_t \mathcal{U}(w) + \nabla \cdot \mathcal{F}(w,v) - \mathcal{L}(w,v) : \nabla v \leq 0. \tag{5.60}$$

The non homogeneous scalar conservation law (5.51) supplemented with all the entropy inequalities (5.58) naturally falls within the frame of the usual Kruzkov theory for weak solutions since again the color function v belongs to $W^{2,\infty}(\mathbb{R}^d, \mathbb{B}_+^L)$:

$$v_0 \in W^{2,\infty}(\mathbb{R}^d, \mathbb{B}_+^L). \tag{5.61}$$

Therefore, the celebrated Kruzkov uniqueness theorem for scalar conservation law with smooth non homogeneities applies and asserts the uniqueness of the entropy weak solution of the Cauchy problem (5.51)-(5.58) with initial data $w_0 \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$.

Hereafter, we shall prove existence and uniqueness of a solution to the coupled problem (5.51)-(5.58) thanks to a well-balanced multidimensional finite volume method formulated on general triangulation. Here, the well-balanced property means that the solutions in the u variable will be kept constant in time and space as soon as the initial data u_0 is chosen constant whatever is the precise definition of the smoothly varying in space color function v. This well-balanced property is obviously a constancy property of primary importance.

5.3 Well-balanced finite volume scheme

Before stating our main result, we introduce the required notations and motivate the formulation of the finite volume method under consideration. To easily meet the well-balancing issues we have put forward along the lines of the Introduction, the finite volume framework we develop actually makes use of two families of triangulations. The first triangulation, denoted by $\mathcal{T}_h$, is made of general polyhedra and will be referred to as the primal mesh. Then a closely related triangulation is of concern, the so-called dual mesh $\mathcal{T}_h^\star$, whose polyhedra are derived from the edges of the primal one. As we shall see, dual meshes may not uniquely inferred from $\mathcal{T}_h$ and it will turn that a given choice essentially affects the closed-form of expression of the CFL restriction in the resulting time explicit finite volume method.

Equipped with these primal and dual meshes, approximate solutions u_h and v_h of the Cauchy problem (5.46) with initial data (u_0, v_0), are classically sought under the form of piecewise constant functions. But in opposition with the usual approach, constant values for u_h and v_h will not be colocalized : u_h (respectively v_h) will assume constant values in each polyhedron of the primal mesh (respectively in each polyhedron of the dual mesh).

To facilitate the derivation of the proposed well-balanced scheme, we shall take advantage of the smoothness of the color function v. Such a smoothness indeed gives room in the precise definition of the discrete approximation v_h : it may range from a local averaged form to a point-wise evaluation. Here and for convenience, we shall promote without real loss of generality an average value of v along each edge of the primal mesh. This choice indeed allows to bypass in a first step the precise derivation of a dual mesh from the edges

of the primal one : a convex sequence of real numbers will in turn provide a sufficient souvenir of the dual mesh. On the ground of this observation, we shall give a first brief but sustained mathematical presentation of the finite volume method under consideration. We shall then be in a position to state the main result of this paper. At last, we shall close this section with a comprehensive construction of the proposed finite volume approximation when deriving dual meshes from the primal one.

5.3.1 Terminology and assumptions

The primal mesh, $\mathcal{T}_h$, is a general (locally finite) triangulation of $\mathbb{R}^d$ made of non-overlapping, non-empty, and open polyhedra : $\cup_{K\in\mathcal{T}_h}\overline{K} = \mathbb{R}^d$. We assume that for every pair of distinct polyhedra $K, K' \in \mathcal{T}_h$ the set $K \cap K'$ is either an edge e of both K and K' or a set with Hausdorf dimension less than or equal to $d-2$. The set of edges of a polyhedron K is denoted by ∂K; and for each $e \in \partial K$, $\nu_{K,e} \in \mathbb{R}^d$ represents the outward unit normal vector to the edge e. The volume of K and the $(d-1)$-measure of e are denoted $|K|$ and $|e|$, respectively. Given an edge e in K, K_e denotes the unique polyhedron in $\mathcal{T}_h$ that shares the same edge e with K. We set

$$h = \sup_{K\in\mathcal{T}_h} h_K,$$

where h_K is the exterior perimeter of the polyhedron K, and assume that the triangulation $\mathcal{T}_h$ satisfies the following non degeneracy condition

$$\sup_K \frac{h_K p_K}{|K|} \leq C, \tag{5.62}$$

for a uniform constant $C > 0$. Here, p_K denotes the perimeter of K defined by

$$p_K = \sum_{e\in\partial K} |e|. \tag{5.63}$$

It is unnecessary, at this stage, to provide a comprehensive derivation of the *dual mesh* $\mathcal{T}_h^\star$ that one could define from the edges e in the primal mesh $\mathcal{T}_h$. Let us just recall that, by design, a dual mesh is made of non-overlapping, non-empty, and open polyhedra denoted by $K^\star(e)$ with $\cup_{e\in\mathcal{T}_h} K^\star(e) = \mathbb{R}^d$. By construction, both sets $K^\star(e) \cap K$ and $K^\star(e) \cap K_e$ are non-empty for all pair (K, K_e) of adjacent polyhedra parametrized by the edges e in $\mathcal{T}_h$. Note that the set $K^\star(e) \cap K$ is a subcell of K. Then, the only information about $\mathcal{T}_h^\star$ that is required in this section is a given convex sequence of real numbers prescribed in each polyhedron K in $\mathcal{T}_h$; we denote by $\{\alpha_{K,e}\}_{\{e,e\in\partial K\}}$, that satisfies for any given K in $\mathcal{T}_h$:

$$0 < \alpha_{K,e} < 1, \quad e \in \partial K; \quad \text{with} \quad \sum_{e\in\partial K} \alpha_{K,e} = 1. \tag{5.64}$$

We will see later that the coefficient $\alpha_{K,e}$ is nothing but the ratio of the volume of $K^\star(e) \cap K$ to the volume of K, where $K^\star(e)$ stands for the dual polyhedron of K attached to a given

edge e in ∂K :

$$\alpha_{K,e} \equiv \frac{|K^\star(e) \cap K|}{|K|}, \quad e \in \partial K. \tag{5.65}$$

At last, the time increment, denoted by τ, will be assumed to satisfy

$$\frac{\tau}{h} \leq C$$

and the primal mesh to be constrained by

$$C_1 \leq \frac{|e|}{h} \leq C_2$$

for uniform constants C, C_1, C_2. Observe that the latter is probably not an optimal condition for the present study but it ensures that all dimensions in both meshes will be of order h. A key property for the forthcoming CFL condition, is that under these assumptions the aera $|K^\star(e) \cap K|$ is not smaller than $O(h^2)$: there exists a uniform positive constant c such that

$$ch^2 \leq |K^\star(e) \cap K|. \tag{5.66}$$

We use the notation $t^n = n\tau$. As already underlined, we will seek at each time level t^n approximate solutions u_h and v_h of the Cauchy problem (5.46) with initial data (u_0, v_0), under the form of piecewise constant functions with :

$$\begin{aligned} &u_h(x,t^n) = u_K^n, \quad x \in K, \quad K \in \mathcal{T}_h, \\ &v_h(x,t^n) \equiv v_h(x) = v_e, \quad x \in K^\star(e), \quad e \in \mathcal{T}_h. \end{aligned} \tag{5.67}$$

Here and since the solution v in the Cauchy problem (5.46) does not depend on time, it seems natural to set $v_h(x,t^n) \equiv v(x) = {v_h}^0(x) \in \mathbb{R}^L$ for all time level t^n, for some discrete approximation ${v_h}^0$ of the smooth function v_0. As already emphasized, we promote the formula :

$$v_h(x) = v_e \equiv \frac{1}{|e|} \int_e v_0(y)dy, \quad x \in K^\star(e), \quad e \in \mathcal{T}_h, \tag{5.68}$$

while the discrete version of the possibly discontinuous initial data u_0 is chosen according to the usual full averaging procedure over each polyhedron K :

$$u_h^0(x) = u_K^0 \equiv \frac{1}{|K|} \int_K u_0(y)dy, \quad x \in K, \quad K \in \mathcal{T}_h. \tag{5.69}$$

Remark 5.2. In view of the smoothness of the data v_0, any other consistent definition for the constant value v_e in $K^\star(e)$ would have been relevant. The interest in the particular choice (5.68) stems from the following Green formula, valid for each of the polygonal domain K :

$$X\Big(\sum_{e \in \partial K} v_{el} \nu_{K,e} |e|\Big) = \int_K \nabla \cdot (v_l(x)\, X)dx = X\Big(\int_K \nabla v_l(x)dx\Big),$$

where X denotes any given fixed vector in $\mathbb{R}^d$ and v_{el} (respectively v_l) the l-th component

of the vector $v_e \in \mathbb{R}^L$ (resp. v). Hence the proposed average value in (5.68) comes with the identity :

$$\int_K \nabla v_l(x)dx = \sum_{e\in\partial K} v_{el}\nu_{K,e}|e|.$$

Under a tensorial notation, we thus get

$$\int_K \nabla v(x)dx = \sum_{e\in\partial K} v_e \otimes \nu_{K,e}|e|. \tag{5.70}$$

The evolution in time of the discrete solution u_h will rely on a family of so-called *numerical flux-functions,* associated with each edge e of any given polyhedron K in $\mathcal{T}_h$. Besides other properties, these numerical flux functions must meet some consistency property with the exact equation for governing u in (5.48), namely :

$$\partial_t w(u,v) + \nabla\cdot f(w(u,v),v) - \ell(w(u,v),v) : \nabla v = 0, \quad x \in K, \quad t \in (t^n, t^{n+1}). \tag{5.71}$$

Let us observe that in the neighbourhood $K^\star(e)$ of each edge e, precisely where v_h achieves a constant value v_e, the above equation boils down to the scalar equation in the unknown $w \equiv w(u, v_e)$:

$$\partial_t w + \nabla\cdot f(w, v_e) = 0, \quad x \in K^\star(e)\cap K, \quad t \in (t^n, t^{n+1}). \tag{5.72}$$

This in turn leads us to define the required numerical flux function at each edge e in $\mathcal{T}_h$ as a locally Lipschitz continuous two-point flux-function $g_{e,K}(.,.;v_e) : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ that satisfies the consistency property :

$$g_{e,K}(w,w;v_e) = f(w,v_e)\cdot\nu_{K,e}, \tag{5.73}$$

the conservation property :

$$g_{e,K}(w,w_e;v_e) = -g_{e,K_e}(w_e,w;v_e), \tag{5.74}$$

for all real numbers w and w_e, and the monotonicity property

$$\frac{\partial g(w,w_e;v_e)}{\partial w} \geq 0, \quad \frac{\partial g(w,w_e;v_e)}{\partial w_e} \leq 0. \tag{5.75}$$

In addition, we assume that the numerical flux depend (locally) Lipschitz continously upon the variable v_e.

Let us simply observe that the prominent 3-point monotone schemes of the scalar framework readily bring numerical flux functions which obey (5.73)–(5.75). Note that the main results in this paper are easily extended to the setting of E-schemes in the sense of Osher [117]. Here, the dependance in the parameter v_e has been kept in the numerical flux-function $g_{e,K}(.,.;v_e)$ for convenience in the forthcoming analysis.

Remark 5.3. Since the function $g(.,.;.) : \mathbb{R}\times\mathbb{R}\times\mathbb{R} \to \mathbb{R}$ locally Lipschitz continuous in its three arguments, for all compact $\mathcal{K} \subset \mathbb{R}^3$, there exists some positive constant $C_\mathcal{K}$ such that

for all triple $(w^{(1)}, w_e^{(1)}, v_e^{(1)})$ and $(w^{(2)}, w_e^{(2)}, v_e^{(2)})$ in $\mathcal{K}$, the following estimate holds true :

$$\begin{aligned} |g_{e,K}(w^{(2)}, w_e^{(2)}; v_e^{(2)}) - g_{e,K}(w^{(1)}, w_e^{(1)}; v_e^{(1)})| \leq \\ C_{\mathcal{K}}\Big(|w^{(2)} - w^{(1)}| + |w_e^{(2)} - w_e^{(1)}| + |v_e^{(2)} - v_e^{(1)}|\Big). \end{aligned} \tag{5.76}$$

5.3.2 Well-balanced scheme

We are now in a position to define the finite volume approximation of (5.71).

Let us assume that the approximate solution $u_h(.,t^n)$ is known at time t^n, this one is then evolved to the next time level t^{n+1} into two steps. These are given the following condensed form :

Subcell reconstruction.

Define at time t^n in each polyhedron K of $\mathcal{T}_h$ and for each edge e in ∂K, the subcell values :

$$w_{K,e}^n = C_0(u_K^n, v_e), \quad e \in \partial K, \tag{5.77}$$

and their $\{\alpha_{K,e}\}_{\{e,e\in\partial K\}}$-weighted average over K :

$$w_K^n = \sum_{e\in\partial K} \alpha_{K,e} w_{K,e}^n. \tag{5.78}$$

Evolution in time.

Update at time t^{n+1} the discrete solution u_h when defining in each polyhedron K, u_K^{n+1} as the unique solution of :

$$\sum_{e\in\partial K} \alpha_{K,e}\, C_0(u_K^{n+1}, v_e) = w_K^{n+1}, \tag{5.79}$$

where the update w_K^{n+1} is given by the following finite volume scheme :

$$w_K^{n+1} = w_K^n - \frac{\tau}{|K|}\sum_{e\in\partial K} g_{e,K}(w_{K,e}^n, w_{K_e,e}^n; v_e)|e| + \frac{\tau}{|K|}\sum_{e\in\partial K} f(w_{K,e}^n, v_e)\cdot \nu_{K,e}|e|. \tag{5.80}$$

The finite volume method we propose is a time explicit scheme and for stability purposes, we shall impose the following CFL restriction, for all polyhedron K in $\mathcal{T}_h$ and edge e in ∂K :

$$\frac{\tau}{|K|}\,\frac{|e|}{\alpha_{K,e}} \sup_{u\in[m,M]} \Big|\frac{\partial f(w(u,v_e),v_e)}{\partial w}\Big| \leq 1, \tag{5.81}$$

where $m = \inf_{x\in\mathbb{R}^d} u_0(x)$ and $M = \sup_{x\in\mathbb{R}^d} u_0(x)$. This completes the description of the discrete method.

Let remark that the quantity $|K|\alpha_{K,e}/|e|$ satisfies

$$|K|\frac{\alpha_{K,e}}{|e|} = \frac{|K^\star(e)\cap K|}{|e|} \geq \frac{c}{C_2}h,$$

so that the CFL condition can not imply the degeneracy of the time step τ, that decreases at most as $O(h)$. We will see in Section 5.4 how to build suitable primal and dual meshes.

Several comments are in order. First observe that the constitutive assumptions (5.44)–(5.45) on the coupling function $C_0(.,.)$ immediately yields existence and uniqueness of a solution to the nonlinear equation (5.77) so that the finite volume method (5.77)–(5.80) is well defined. Then, formulae (5.77) and (5.79) obviously express the same identity but respectively at time t^n and time t^{n+1} so that in practice, both are mostly redundant : the finite volume method essentially reduces to (5.79)–(5.80). As they stand, they nevertheless ease the description of the method.

Next, it worth observing that the consistency condition (5.73) allows in (5.80) to recast the flux balance $\sum_{e\in\partial K} f(w^n_{K,e},v_e)\nu_{K,e}|e|$ as $\sum_{e\in\partial K} g_{e,K}(w^n_{K,e},w^n_{K,e};v_e)|e|$. Here we stress that at each edge e in ∂K, both the numerical flux-function $g_{e,K}(w^n_{K,e},w^n_{K_e,e};v_e)$ and its counterpart $f(w^n_{K,e},v_e)\cdot\nu_{K,e}$ are evaluated thanks to the subcell values $w^n_{K,e}$ (5.77) and not to their averaged form w^n_K in (5.78). The motivation is twofold. In a first hand, the two flux balances involved in (5.80), namely $\sum_{e\in\partial K} g_{e,K}|e|$ and $\sum_{e\in\partial K} f(w^n_{K,e},v_e)\cdot\nu_{K,e}|e|$, make the proposed formula to be a consistent finite volume approximation of the exact equation (5.71) for governing u : namely, the first one will be seen hereafter to be consistent with $\nabla\cdot f(w,v)$ while the second one actually provides a consistent approximation of the source term $\ell(w,v):\nabla v$. In a second hand, the resulting discretization of the source term is seen to be well-balanced. This is the matter of the next statement :

Proposition 5.1 (Well-balanced property)**.** *Let the initial data u_0 in the Cauchy problem (5.71) be any given constant in space function, say for some given real number $u^\star$:*

$$u_0(x)=u^\star,\quad x\in\mathbb{R}^d.$$

Then and whatever is the precise closed-form of expression of the color function v in (5.71), the discrete solution u_h of (5.77)–(5.80) stays constant with :

$$u_h(x,t^n)=u_0(x)=u^\star,\quad x\in\mathbb{R}^d,\ \textit{for all time level } t^n. \tag{5.82}$$

In other words, the finite volume method (5.77)–(5.80) is well-balanced with respect to all the natural equilibria of (5.71).

Proof. The discrete initial data (5.69) clearly reads $u^0_h(x)=u^\star$ for all x in $\mathbb{R}^d$ so that at the first subcell reconstruction step, we get $w^0_{K,e}=C_0(u^\star,v_e)=w^0_{K_e,e}$ for any given edge e of an arbitrary polyhedron K in $\mathcal{T}_h$. Consequently, the numerical flux $g_{e,K}(w^0_{K,e},w^0_{K_e,e};v_e)$ at a given e boils down to $f(w^0_{K,e},v_e)\cdot\nu_{K,e}$ in view of the consistency condition (5.73). Namely the two flux balances in the updating formula (5.80) cancel out and we end up with $w^1_K=w^0_K=\sum_{e\in\partial K}\alpha_{K,e}C_0(u^\star,v_e)$ thanks to the definition (5.78). Arguing about uniqueness, we thus get when solving (5.79) $u^1_K=u^\star$ for any given polyhedron K of $\mathcal{T}_h$: namely $u_h(x,t^1)=u^\star$ for all x in $\mathbb{R}^d$. An immediate recursion extends the result to the subsequent time levels. □

To conclude this paragraph, it is worth illustrating that the last flux-balance entering the finite volume approximation (5.80) actually provides a consistent approximation of the source term $\ell(w,v) : \nabla v$. In that aim and for the sake of simplicity, it is convenient to temporarily adopt the following easy coupling framework (see indeed Remark 5.1) :

$$\begin{aligned} C_0(u,v) &= (1-v)\gamma_-(u) + v\gamma_+(u), \\ C_i(u,v) &= (1-v)a^-(\gamma_-(u)) + va^+(\gamma_+(u)), \quad 1 \le i \le d, \end{aligned}$$

so that $f(w,v)$ and $\ell(w,v)$ in (5.71) respectively read :

$$f(w(u,v),v) = (1-v)A^-(\gamma_-(u)) + vA^+(\gamma_+(u)),$$

and

$$\ell(w(u,v),v) = \Big(A^+(\gamma_-(u)) - A^-(\gamma_-(u))\Big).$$

It can be then readily computed :

$$\begin{aligned} \sum_{e\in\partial K} f(w^n_{K,e}, v_e)\cdot \nu_{K,e}|e| &= \sum_{e\in\partial K} \Big((1-v_e)A^-(\gamma_-(u^n_K)) + v_e A^+(\gamma_-(u^n_K))\Big)\cdot \nu_{K,e}|e| \\ &= \Big(A^+(\gamma_-(u^n_K)) - A^-(\gamma_-(u^n_K))\Big)\cdot \sum_{e\in\partial K} v_e|e|\nu_{K,e}, \\ &\quad + A^-(\gamma_-(u^n_K))\cdot\Big(\sum_{e\in\partial K}|e|\nu_{K,e}\Big) \end{aligned} \tag{5.83}$$

which is nothing else a consistent discretization of $\ell(w(u,v)) : \nabla v$, in view of the representation formula (5.70) for ∇v and the easy identity $\displaystyle\sum_{e\in\partial K}|e|\nu_{K,e} = 0$.

These straightforward calculations allows to bridge the finite volume formula (5.80) to the governing equation (5.71) for $w(u,v)$, expressed over K, namely where v_h does achieve distinct values. The gap in between (5.71) and its reduced version (5.72) (*i.e.* with $x \in K^\star(e)\cap K$) will be definitely closed when revisiting the finite volume approximation (5.77)–(5.80) within the frame of primal-dual meshes in Section 5.4.

5.3.3 Main convergence result

The main result of this paper is stated in the following theorem.

Theorem 5.2 (Convergence of the well-balanced finite volume scheme)**.** *Consider the Cauchy problem (5.46)–(5.57) with initial data $u_0 \in L^\infty(\mathbb{R}^d)$ and $v_0 \in W^{2,\infty}(\mathbb{R}^d)$ under the constitutive assumptions (5.44)–(5.45). Let u_h be the sequence of approximate solutions defined by the finite volume method (5.68)–(5.69), (5.77)–(5.80) with numerical flux-functions satisfying the conditions (5.73)–(5.75).*

Then under the CFL restriction (5.81), the sequence u_h is uniformly bounded in $L^\infty(\mathbb{R}_+\times\mathbb{R}^d)$ in term of the sup-norm of the initial data and converges when $h\to 0$ in the L^p_{loc} norm strongly, $1 \le p < \infty$, to the unique entropy solution u to the problem (5.46)–(5.57) : namely for all time $T > 0$

and for all compact $\mathcal{K}$ in $\mathbb{R}^d$

$$\|u - u_h\|_{L^p((0,T)\times\mathcal{K})} \leq o(h), \quad \textit{for} \quad 1 \leq p < \infty, \tag{5.84}$$

where $o(h)$ is a function tending to zero with h.

5.4 Finite volume approximations with primal-dual meshes

5.4.1 A convex combination

One of our objectives in this section is explaining how the coefficients $\alpha_{K,e}$ should be determined. Arguing about the formula-definitions (5.77)–(5.78) at time t^n and the consistency condition (5.73), the next statement is easily inferred.

Lemma 5.3 (Edge values and convex combination)**.** *For any given polyhedron K of $\mathcal{T}_h$ and edge e in ∂K, let us define the following subcell updates :*

$$w_{K,e}^{n+1,-} = w_{K,e}^n - \frac{|e|}{\alpha_{K,e}}\frac{\tau}{|K|}\Big(g_{e,K}(w_{K,e}^n, w_{K_e,e}^n; v_e) - g_{e,K}(w_{K,e}^n, w_{K,e}^n; v_e)\Big). \tag{5.85}$$

Then w_K^{n+1} in (5.80) can be equivalently recovered from the following $\{\alpha_{K,e}\}_{\{e,e\in\partial K\}}$-weighted averaging procedure :

$$w_K^{n+1} = \sum_{e\in\partial K} \alpha_{K,e} w_{K,e}^{n+1,-}. \tag{5.86}$$

Let us observe that the finite volume formula (5.85) for $w_{K,e}^{n+1,-}$ is nothing but a consistent approximation of the one dimensional conservation law :

$$\partial_t w + \nabla \cdot f(w, v_e) = 0. \tag{5.87}$$

The reason for quoting the update $w_{K,e}^{n+1,-}$ as a subcell update, and thus of multidimensional nature, will be grounded in this paragraph and actually stays at the core of the re-interpretation of the finite volume formula (5.80) in the frame of primal-dual meshes.

To further proceed, let us underline that the identity (5.86) just expresses that w_K^{n+1} actually writes as a convex decomposition of the subcell updates $w_{K,e}^{n+1,-}$. When understood in their quasi-one dimensional form (5.85), the latter can be recognized as extensions to

the present non-homogenous setting of partial updates entering similar convex decompositions that have proved well suited in the analysis of homogeneous multidimensional finite volume methods (see for instance Coquel-LeFloch [54], [55]). Indeed, the interest in such a convex decomposition primary stems from the fact that many of the basic stability properties satisfied by the scheme (5.85) in one space variable are right away inherited in several space dimensions thanks to convexity under some CFL restriction. Observe that the relevant CFL condition coming with (5.85) just writes:

$$\frac{\tau}{|K|}\frac{|e|}{\alpha_{K,e}}\left|\frac{g_{e,K}(w^n_{K,e},w^n_{K_e,e};v_e)-g_{e,K}(w^n_{K,e},w^n_{K,e};v_e)}{w^n_{K_e,e}-w^n_{K,e}}\right|\leq 1, \tag{5.88}$$

and hence the CFL restriction (5.81).

At last and arguing about the definition (5.85), the subcell reconstruction step (5.78) at time t^{n+1} and the update formula (5.79), we infer the seemingly trivial equalities:

$$\sum_{e\in\partial K}\alpha_{K,e}w^{n+1}_{K,e}=w^{n+1}_K=\sum_{e\in\partial K}\alpha_{K,e}w^{n+1,-}_{K,e}. \tag{5.89}$$

In other words, all the steps involved in the method are locally conservative: this natural property will play a central role in the forthcoming analysis.

5.4.2 Interpretation of the proposed well-balanced scheme

The derivation of a dual mesh $\mathcal{T}_h^\star$ from the edge of the primal one $\mathcal{T}_h$ may be performed as follow. For any given (open) polyhedron K, the idea is to pick an internal node x_K in K which choice is left arbitrary at this stage. Such a procedure is given below a systematic definition independent of the mesh refinement h. Equipped with the node x_K, we define for any given edge e in K the convex hull of e and x_K. The interior of this convex hull, we denote by $\mathcal{E}(x_K,e)$, yields a non-empty open polyhedron made of $(d+1)$ edges. Observe that the following properties are met by construction : for any given pair of edges e,e' in ∂K with K an arbitrary polyhedron in $\mathcal{T}_h$

$$\mathcal{E}(x_K,e)\cap K=\mathcal{E}(x_K,e),\quad \mathcal{E}(x_K,e)\cap\mathcal{E}(x_K,e')=\emptyset, \tag{5.90}$$

while

$$\sum_{e\in\partial K}\mathcal{E}(x_K,e)=K. \tag{5.91}$$

Then, the required definition of the polyhedron $K^\star(e)$ of the dual mesh $\mathcal{T}_h^\star$, attached to a given edge e in $\mathcal{T}_h$ with adjacent polyhedron K and K_e, follows from :

$$K^\star(e)=\mathcal{E}(x_K,e)\cup\mathcal{E}(x_{K_e},e). \tag{5.92}$$

We refer the reader to Figure 5.2 for an illustration.

The constructive procedure for defining the internal node x_K independently of h relies on the set of vertices ϑ of the polyhedron K, together with a convex sequence of real numbers

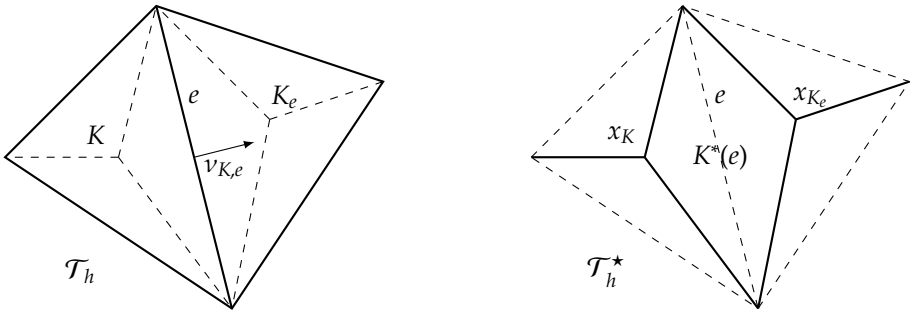


Figure 5.2: Primal and dual meshes, edges and vertices.

$\{\beta_{K,\vartheta}\}_{\{\vartheta,\vartheta\in K\}}$ satisfying :

$$0 < \beta_{K,\vartheta} < 1, \quad \vartheta \in K; \qquad \sum_{\vartheta\in K} \beta_{K,\vartheta} = 1.$$

The required internal node x_K in K is then defined by its co-ordinates in $\mathbb{R}^d$:

$$x_K = \sum_{\vartheta\in K} \beta_{K,\vartheta}\, x_\vartheta, \tag{5.93}$$

where x_ϑ stands for the co-ordinates of the vertex ϑ.
In fact, this kind of construction ensures the correct behaviour of the primal and dual meshes with the definition of the $\alpha_{K,e}$ and with the previous non-degeneracy assumptions (5.65)-(5.66), the CFL condition (5.81) is then only modified according to the choice of the function v and its discrete representation.

To further proceed in the comprehensive derivation of the finite volume framework, a few additional notations are in order. For any given K in $\mathcal{T}_h$ and e in ∂K, an edge of a dual polyhedron $K^\star(e) \in {\mathcal{T}_h}^\star$ or of the subcell $K^\star(e)\cap K$ of K will be indifferently denoted by $e^\star$. Observe that with little abuse in the notations, an edge e of some cell K of the primal mesh $\mathcal{T}_h$ is also a dual edge of the subcell $K^\star(e)\cap K$: see indeed Figure 5.2. At last $\nu_{K^\star(e),e^\star} \in \mathbb{R}^d$ stands for the outward unit vector normal to the edge $e^\star$.

Equipped with this notations, we are in a position to re-interpret the quasi-one dimensional update $w_{K,e}^{n+1,-}$ introduced in (5.85) in term of an update in the subcell $K^\star(e)\cap K$ of K, thanks to following simple but key identity:

$$\sum_{e^\star\in K^\star(e)\cap K} |e^\star|\nu_{K^\star(e),e^\star} = 0, \quad \text{i.e.} \quad |e|\nu_{K,e} = -\sum_{e^\star\in K^\star(e)\cap K,\ e^\star\neq e} |e^\star|\nu_{K^\star(e),e^\star}.$$

It is then straightforward to recast $w_{K,e}^{n+1,-}$ according to:

$$\begin{aligned} w_{K,e}^{n+1,-} &= w_{K,e}^n - \frac{\tau}{\alpha_{K,e}|K|} g_{e,K}(w_{K,e}^n, w_{K_e,e}^n; v_e)|e| + \frac{\tau}{\alpha_{K,e}|K|} f(w_{K_e}^n)\cdot \nu_{K,e}|e|, \\ &= w_{K,e}^n - \frac{\tau}{|K^\star(e)\cap K|}\Big(g_{e,K}(w_{K,e}^n, w_{K_e,e}^n; v_e)|e| \\ &\qquad + \sum_{e^\star \in K^\star(e)\cap K,\ e^\star \neq e} f(w_{K,e^\star}^n, v_e)\cdot \nu_{K^\star(e),e^\star}|e^\star|\Big), \end{aligned} \tag{5.94}$$

where we have used the geometric interpretation (5.65) of $\alpha_{K,e}$. Introducing the numerical flux formula:

$$g_{e^\star,K^\star(e)} = \begin{cases} g_{e,K}(w_{K,e}^n, w_{K_e,e}^n; v_e), & \text{if } e^\star = e; \\ f(w_{K,e^\star}^n, v_e)\cdot \nu_{K^\star(e),e^\star}, & \text{otherwise,} \end{cases} \tag{5.95}$$

$w_{K,e}^{n+1,-}$ thus writes :

$$w_{K,e}^{n+1,-} = w_{K,e}^n - \frac{\tau}{|K^\star(e)\cap K|} \sum_{e^\star \in K^\star(e)\cap K} g_{e^\star,K^\star(e)}|e^\star|, \tag{5.96}$$

and hence under the expected form of an update in the subcell $K^\star(e)\cap K$ for some multi-dimensional balance law to be elucidated. In that aim, it suffices to clarify the origin of the formula-definition $g_{e^\star,K^\star(e)} = f(w_{K,e^\star}^n, v_e)\cdot \nu_{K^\star(e),e^\star}$ for edges $e^\star$ distinct from e. For such an edge $e^\star$, it is worth introducing the adjacent subcell $K^\star(e')\cap K$ to $K^\star(e)\cap K$ in K : *i.e.* with e' in ∂K such that $K^\star(e')\cap K^\star(e) = e^\star$. Note that $e^\star$ is of course distinct from e'. Let us then successively rewrite the left (respectively the right) numerical flux at $e^\star$, say $g_{e^\star,K^\star(e)}$ (respectively $g_{e^\star,K^\star(e')}$), as follows :

$$f(w(u_K^n, v_e), v_e)\cdot \nu_{K^\star(e),e^\star}, \quad \text{respectively : } -f(w(u_K^n, v_e{}'), v_e{}')\cdot \nu_{K^\star(e),e^\star},$$

since by definition (5.77) $w_{K,e}^n = w(u_K^n, v_e)$ and $w_{K,e'} = w(u_K^n, v_e{}')$, and then equivalently :

$$\begin{aligned} &\Big(f(w(u,v),v)\cdot \nu_{K^\star(e),e^\star}\Big)(\omega(0^-)), \\ &\qquad \text{respectively : } -\Big(f(w(u,v),v)\cdot \nu_{K^\star(e),e^\star}\Big)(\omega(0^+)), \end{aligned} \tag{5.97}$$

where $\omega(0^\mp)$ stands for the left and right traces at $\xi = 0$ of the self-similar function $\omega : \xi \in \mathbb{R}_\xi \to (u(\xi), v(\xi)) \in \mathbb{R}\times\mathbb{R}^L$ given by :

$$\omega(\xi) = \begin{cases} (u_K^n, v_e), & \xi < 0, \\ (u_K^n, v_e{}'), & \xi > 0. \end{cases} \tag{5.98}$$

It now suffices to recall (see indeed Section 5.2)that the Riemann solution of the Cauchy problem :

$$\begin{aligned} &\partial_t w + \partial_x\Big(f(w(u,v),v)\cdot \nu_{K^\star(e),e^\star}\Big)(u,v) - \partial_v\Big(f(w(u,v),v)\cdot \nu_{K^\star(e),e^\star}\Big) : \nabla v = 0, \\ &\qquad\qquad \partial_t v = 0 \end{aligned} \tag{5.99}$$

with initial data $((u_K^n, v_e),\ x<0,\ \ (u_K^n, v_e{}'),\ x>0)$ just consists in a standing wave separating (u_K^n, v_e) from $(u_K^n, v_e{}')$, and thus coincides with $\omega(\xi)$ in (5.98). It is therefore clear that the fluxes in (5.97) actually results from the Godunov method applied to the augmented system (5.99) at the edge $e^\star$. In other term, the finite volume formula (5.95)–(5.96) in each subcell $K^\star(e)\cap K$ may be understood as an approximation of the balance law for governing u in (5.71) :

$$\partial_t w(u,v) + \nabla\cdot f(w(u,v),v) - \ell(w(u,v),v) : \nabla v = 0, \quad x\in K,\ \ t\in(t^n,t^{n+1}).$$

This interpretation closes the gap in between the governing equation (5.71) for u and its reduced form (5.72) expressed in w :

$$\partial_t w + \nabla\cdot f(w,v_e) = 0, \quad x\in K^\star(e)\cap K,\ \ t\in(t^n,t^{n+1}).$$

5.5 Sup-norm estimates

Throughout the upcoming sections, the assumptions of Theorem 5.2 are tacitly assumed to be valid. Their formulations are thus skipped over in any forthcoming statements. The main result of this section ensures that the sequence of approximate solutions u_h stays uniformly bounded in $L^\infty(\mathbb{R}_+\times\mathbb{R}^d)$ as a consequence of the following maximum principle :

Proposition 5.4 (Maximum principle)**.** *The finite volume method satisfies the local maximum principle in the variable u :*

$$\min\Big(u_K^n,\ \min_{e\in\partial K} u_{K_e}^n\Big) \le u_K^{n+1} \le \max\Big(u_K^n, \max_{e\in\partial K} u_{K_e}^n\Big) \tag{5.100}$$

in each polyhedron K in $\mathcal{T}_h$ and at all time level t^n.

Since the boundedness assumption on v_0 (5.61) immediately implies an uniform sup-norm estimate for v_h given by (5.68), we easily infer from the maximum principle (5.100) an additional uniform sup-norm estimate but for $w_h = C_0(u_h, v_h)$ arguing about the smoothness properties (5.44) achieved by C_0 :

$$\|w_h\|_{L^\infty(\mathbb{R}^+\times\mathbb{R}^d)} \le O(1). \tag{5.101}$$

Besides the monotonicity assumption (5.75) met by the numerical flux functions, we stress that the preservation of conservativity (5.89) in the subcell reconstruction procedure plays a central role in the validity of the reported maximum principle, as highlighted in the pending proof. The latter will be carried out using a recursion procedure based on subsequent partitions of the set of edges e in K. To fix the notations and up to some relabelling, $\{e_1,...,e_{J_K}\}$ represents the full set of edges $e\in\partial K$ so that here the index J_K is given by $\#\{e, e\in\partial K\}$. Subsets of the form $\{e_1,...,e_J\}$, with increasing index $J\in\{1,...,J_K\}$, will be of concern as follows. Being given J with $1\le J\le K$, let us attach to the subset $\{e_1,...,e_J\}$ the

solution $u^{n+1-}_{K,\{e_1,\ldots,e_J\}}$ of the following nonlinear equation :

$$\sum_{1\le j\le J} \alpha_{K,e_j} C_0(u^{n+1-}_{K,\{e_1,\ldots,e_J\}}, v_{e_j}) = \sum_{1\le j\le J} \alpha_{K,e_j} w^{n+1-}_{K,e_j}, \tag{5.102}$$

where the subcell updates w^{n+1-}_{K,e_j} have been defined in (5.85), Lemma 5.3. Again, the constitutive assumptions (5.44)–(5.45) ensure existence and uniqueness of a solution to (5.102).

Arguing about the conservation property (5.89) verified at the subcell reconstruction step, it is worth observing that $u^{n+1-}_{K,\{e_1,\ldots,e_{J_K}\}}$ can be identified with the final update u^{n+1}_K at time t^{n+1} in the finite volume approximation (5.77)–(5.80). Therefore, the recursion under consideration naturally ends up as soon as the index J reaches the value J_K. In order to initiate the recursion and propagate it, we need the following statement concerned with the values $u^{n+1-}_{K,\{e_J\}}$, $1 \le J \le J_K$, solutions of :

$$C_0(u^{n+1-}_{K,\{e_J\}}, v_{e_J}) = w^{n+1-}_{K,e_J}. \tag{5.103}$$

Lemma 5.5 (Local maximum principle)**.** *The next maximum principle holds true at any given edge e_J in ∂K :*

$$\min(u^n_K, u^n_{K_{e_J}}) \le u^{n+1-}_{K,\{e_J\}} \le \max(u^n_K, u^n_{K_{e_J}}), \quad 1 \le J \le J_K. \tag{5.104}$$

Then the maximum principle property propagates to subsequent sets $\{e_1,\ldots,e_J\}$ with increasing index $J \in \{1,\ldots,J_K\}$ as we now state :

Lemma 5.6. *The solution $u^{n+1-}_{K,\{e_1,\ldots,e_J\}}$ to (5.102) with $J \in \{1,\ldots,J_K\}$, obeys the following maximum principle :*

$$\min\Big(u^n_K, \min_{1\le j\le J}(u^n_{K_{e_j}})\Big) \le u^{n+1-}_{K,\{e_1,\ldots,e_J\}} \le \max\Big(u^n_K, \max_{1\le j\le J}(u^n_{K_{e_j}})\Big). \tag{5.105}$$

The proposed lower and upper bounds for $u^{n+1-}_{K,\{e_1,\ldots,e_{J_K}\}}$, *i.e.* estimate (5.105) with $J = J_K$, just reads the expected local maximum principle (5.100) for u^{n+1}_K, since again u^{n+1}_K coincides with $u^{n+1-}_{K,\{e_1,\ldots,e_{J_K}\}}$ by construction.

Let us now prove the two technical Lemmas 5.5 and 5.6.

Proof of Lemma 5.5. To alleviate the notation we skip the index J . Let us first point out the next estimate, valid under the CFL restriction (5.81) for any given edge e in ∂K :

$$\min(w^n_{K,e}, w^n_{K_e,e}) \le w^{n+1,-}_{K,e} \le \max(w^n_{K,e}, w^n_{K_e,e}) \tag{5.106}$$

as a well-known consequence of the monotonicity assumptions (5.75) satisfied by the numerical flux function $g_{e,K}(.,.;v_e)$. We then recall that the subcell reconstruction step (5.77) builds $w^n_{K,e} = C_0(u^n_K, v_e)$ while the identity $w^{n+1-}_{K,e} = C_0(u^{n+1-}_{K,\{e\}}, v_e)$ holds from the definition (5.103). We can thus recast (5.106) as follows :

$$\min(C_0(u^n_K, v_e), C_0(u^n_{K_e}, v_e)) \le C_0(u^{n+1-}_{K,\{e\}}, v_e) \le \max(C_0(u^n_K, v_e), C_0(u^n_{K_e}, v_e)), \tag{5.107}$$

from which we immediately deduce the required estimate (5.105), namely

$$\min(u_K^n, u_{K_e}^n) \le u_{K,\{e\}}^{n+1-} \le \max(u_K^n, u_{K_e}^n), \quad e \in \partial K$$

since the function C_0 is by assumption (5.45) strictly increasing in its first argument. □

Proof of Lemma 5.6. Let us first observe that the proposed lower-upper bounds (5.105) with $J = 1$ are precisely the matter of Lemma 5.5. Then assuming the validity of the expected maximum principle at rank J, $1 \le J < J_K$, this one is proved to hold true at the next rank $(J+1)$ starting from the formula-definition (5.102):

$$\begin{aligned}\sum_{1\le j\le (J+1)} \alpha_{K,e_j} C_0(u_{K,\{e_1,\dots,e_{(J+1)}\}}^{n+1-}, v_{e_j}) &= \sum_{1\le j\le J} \alpha_{K,e_j} w_{K,e_j}^{n+1-} + \alpha_{K,e_{(J+1)}} w_{K,e_{(J+1)}}^{n+1-},\\ &= \sum_{1\le j\le J} \alpha_{K,e_j} C_0(u_{K,\{e_1,\dots,e_J\}}^{n+1-}, v_{e_j})\\ &+ \alpha_{K,e_{(J+1)}} C_0(u_{K,e_{(J+1)}}^{n+1-}, v_{e_{(J+1)}}).\end{aligned} \tag{5.108}$$

Let us recast the above identity as follows:

$$\begin{aligned}&\sum_{1\le j\le J} \alpha_{K,e_j} C_0(u_{K,\{e_1,\dots,e_{(J+1)}\}}^{n+1-}, v_{e_j}) - \sum_{1\le j\le J} \alpha_{K,e_j} C_0(u_{K,\{e_1,\dots,e_J\}}^{n+1-}, v_{e_j}) =\\ &- \alpha_{K,e_{(J+1)}} \Big(C_0(u_{K,\{e_1,\dots,e_{(J+1)}\}}^{n+1-}, v_{e_{(J+1)}}) - C_0(u_{K,e_{(J+1)}}^{n+1-}, v_{e_{(J+1)}}) \Big).\end{aligned} \tag{5.109}$$

To condense the notations, we introduce the two functions $u \mapsto \Psi_J(u) \equiv \sum_{1\le j\le J} \alpha_{K,e_j} C_0(u, v_{e_j})$ and $u \mapsto \psi_{(J+1)}(u) = \alpha_{K,e_{(J+1)}} C_0(u, v_{e_{(J+1)}})$ so as to infer from (5.109) :

$$\begin{aligned}&\Big(\Psi_J(u_{K,\{e_1,\dots,e_{(J+1)}\}}^{n+1-}) - \Psi_J(u_{K,\{e_1,\dots,e_J\}}^{n+1-})\Big)\times\\ &\qquad \Big(\psi_{(J+1)}(u_{K,\{e_1,\dots,e_{(J+1)}\}}^{n+1-}) - \psi_{(J+1)}(u_{K,e_{(J+1)}}^{n+1-})\Big) \le 0,\end{aligned} \tag{5.110}$$

since by assumption (5.64) $\alpha_{K,e_{(J+1)}} > 0$. But the monotonicity hypothesis (5.45) on C_0 together with again assumption (5.64) imply that both functions $u \to \Psi_J(u)$ and $u \to \psi_{(J+1)}(u)$ strictly increases with u so that (5.110) yields :

$$\min(u_{K,\{e_1,\dots,e_J\}}^{n+1-}, u_{K,e_{(J+1)}}^{n+1-}) \le u_{K,\{e_1,\dots,e_{(J+1)}\}}^{n+1-} \le \max(u_{K,\{e_1,\dots,e_J\}}^{n+1-}, u_{K,e_{(J+1)}}^{n+1-}). \tag{5.111}$$

This concludes the proof since Lemma 5.5 ensures the estimate :

$$\min(u_K^n, u_{K_{e_{(J+1)}}}^n) \le u_{K,e_{(J+1)}}^{n+1-} \le \max(u_K^n, u_{K_{e_{(J+1)}}}^n). \tag{5.112}$$

□

5.6 Entropy inequalities

Proposition 5.4 asserts uniform sup-norm boundedness for the sequence u_h which in the absence of an *a priori* strong compactness argument, leads us to study the structure of the Young measure μ associated with $\{u_h\}_{h>0}$. Recall that such a Young measure represents all the composite weak-star limits $a(u_h)$ of u_h with continuous functions $a \in C^0(\mathbb{R})$, namely with continuous functions of *a single* variable :

$$a(u_h) \quad \longrightarrow \quad < \mu, a > \equiv \int_{\mathbb{R}} a(\lambda) d\mu(\lambda), \quad w \star - L^\infty. \tag{5.113}$$

We propose to establish that the measure μ under consideration reduces to a Dirac measure, and hence to prove the *a posteriori* strong convergence of u_h, invoking the celebrated DiPerna's uniqueness Theorem [64] within the frame of entropy measure-valued solutions.

In this section we derive the required discrete entropy inequalities together with the *a priori* estimates that are needed to handle the passage to the limit in the sense of measure valued solutions. In this respect, the main issue is to assess the relevance of the Young measure μ in such a limit. Indeed and by contrast with (5.113), discrete entropy inequalities generically involve numerical flux functions, that are continuous functions but of (at least) *two* arguments : the sequence $u_h(.)$ itself and its shift $\Delta_h u_h \equiv u_h(. + h)$. Nonlinear superposition of possible discrete oscillations in u_h and its shift $\Delta_h u_h$ may prevent the usual Young measure μ to represent the composite weak-star limit of $G(u_h, \Delta_h u_h)$. Counterexamples have been proposed by Coquel-LeFloch in [55]. Some weak control over possible discrete oscillations is therefore mandatory in order to justify the applicability of μ in the limiting form of discrete entropy inequalities.

The requisite weak estimate classically corresponds to some estimate of the discrete entropy dissipation rate in the finite volume approximation. The derivation of several specific estimates with distinctive features have been the matter of a large body of literature since the pioneering work by Coquel and LeFloch [54]. The reader is referred to the Introduction where several subsequent contributions are quoted. In the present work, the type of estimate we derive is in the spirit of the one introduced by Cockburn, Coquel, and LeFloch [50]. This estimate is rather weak in the sense that it does not allow actually to pass weakly to the limit in arbitrary numerical entropy-flux functions but nevertheless it turns sufficient to handle the limiting form of convenient discrete entropy inequalities. The main interest in such an estimate stems from the simplicity of its derivation.

We first focus on the derivation of the convenient discrete entropy inequalities under interest and we then address the derivation of the required weak estimate. The passage to the limit in the discrete inequalities is the subject of the next section.

After Crandall and Majda [56], assumptions (5.73)–(5.75) on the numerical flux functions $g_{e,K}$ are known to yield a full set of discrete entropy inequalities in the scalar conservation law setting. Here and in the light of Section 5.2, the scalar conservation laws of concern have to be found locally at each edge e in $\mathcal{T}_h$, and take the generic form :

$$\partial_t w + \nabla \cdot f(w, v) = 0, \quad \text{for a given fixed } v \in \mathbb{R}. \tag{5.114}$$

Associated entropy pairs are then given by (5.55-5.59) Section 5.2. Inequalities in the next statement are naturally built from the subcell updates $w_{K,e}^{n+1,-}$ (5.85) of Lemma 5.3 and in this regard may be understood as subcell entropy inequalities.

5.6.1 Discrete entropy estimates

Lemma 5.7 (Entropy inequalities per cell)**.** *Let $(\mathcal{U},\mathcal{F}) : \mathbb{R} \to \mathbb{R}\times\mathbb{R}^d$ be any given convex entropy pair for the scalar conservation law (5.114), where e denotes any edge in ∂K for an arbitrarily given K in $\mathcal{T}_h$. Then there exists a numerical entropy flux function $G_{e,K} : \mathbb{R}^2 \to \mathbb{R}$ that satisfies the consistency property :*

$$G_{e,K}(w,w;v_e) = \mathcal{F}(w,v_e)\cdot\nu_{K,e}, \tag{5.115}$$

the conservation property:

$$G_{e,K}(w,w_e;v_e) = -G_{e,K_e}(w_e,w;v_e), \tag{5.116}$$

for all real numbers w and w_e, so that the following discrete entropy inequality holds true:

$$\mathcal{U}(w_{K,e}^{n+1,-}) - \mathcal{U}(w_{K,e}^{n}) + \frac{1}{\alpha_{K,e}}\frac{\tau|e|}{|K|}\Big(G_{e,K}(w_{K,e}^n, w_{K_e,e}^n; v_e) - \mathcal{F}(w_{K,e}^n, v_e)\cdot\nu_{K,e}\Big) \le 0. \tag{5.117}$$

We refer the reader to [74] for a proof of this classical result. As already claimed, the pending weak estimate will not allow to pass weakly to the limit in arbitrary numerical entropy flux-functions. We thus propose to merge inequalities (5.117) in such a way that solely exact entropy fluxes $\mathcal{F}(w,v_e)\cdot\nu_{K,e}$ enter the resulting weak form. This is the matter of the next preliminary Lemma.

Lemma 5.8. *Let ϕ be any given non-negative test function in $\mathcal{D}(\mathbb{R}_+^*\times\mathbb{R}^d)$. Define for any given edge e in $\mathcal{T}_h$, the average:*

$$\phi_e^n = \frac{1}{\tau|e|}\int_{t^n}^{t^{n+1}}\int_e \phi(x,t)dxdt. \tag{5.118}$$

Then the following discrete weak inequality holds true :

$$\begin{aligned}\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^{n+1,-}) - \mathcal{U}(w_{K,e}^n)\Big)\phi_e^n|K| \\ - \tau\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\mathcal{F}(w_{K,e}^n,v_e)\cdot\nu_{K,e}\phi_e^n|e| \le 0.\end{aligned} \tag{5.119}$$

The proof is postponed at the end of this section. We shall easily infer from the discrete in space weak inequality (5.119) the next companion continuous in space inequality :

Proposition 5.9. *The finite volume approximation (5.77)–(5.80) obeys at each time level t^n the following discrete in time weak formulation of the entropy inequality :*

$$\begin{aligned}&\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^{n+1,-}) - \mathcal{U}(w_{K,e}^n)\Big)\phi_e^n|K| \\ &- \iint_{]t^n,t^{n+1}[\times\mathbb{R}^d} Q(u_h^n,v(x))\cdot\nabla\phi(x,t) + \phi(x,t)\partial_v Q(u_h^n,v(x)) : \nabla v(x)dxdt \\ &\qquad \le O(h)\tau\,\|\phi\|_{W^{1,\infty}(]t^n,t^{n+1}[\times\mathbb{R}^d)}|supp(\phi)|.\end{aligned} \tag{5.120}$$

Let us emphasize that the proof, given at the end of this section, essentially makes use of the uniform sup-norm estimate (5.100) for the sequence u_h together with the smoothness assumption (5.61) made on the color function v.

Clearly, the Young measure μ can tackle the weak limit of the space derivatives involved in inequality (5.120) extended to any time interval $(0,T)$, $T>0$. Such a claim then naturally rises the question of passing weakly to the limit in the discrete time derivative. In this regard, the latter is conveniently decomposed into:

$$\begin{aligned}\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^{n+1,-})-\mathcal{U}(w_{K,e}^{n})\Big)\phi_e^n|K| \\ &=\sum_{K\in\mathcal{T}_h}\Big(\mathcal{U}(w_K^{n+1})-\mathcal{U}(w_K^n)\Big)\phi_K^n|K| \\ &-\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_K^{n+1})-\mathcal{U}(w_{K,e}^{n+1,-})\Big)\phi_e^n|K| \\ &-\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^{n})-\mathcal{U}(w_K^n)\Big)\phi_e^n|K|,\end{aligned} \tag{5.121}$$

where we have set

$$\phi_K^n=\sum_{e\in\partial K}\alpha_{K,e}\phi_e^n. \tag{5.122}$$

The last two error terms entering the right hand side of (5.121) are devoted to sum up

$$\begin{aligned}\sum_{n\geq 0}\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_K^{n+1})-\mathcal{U}(w_{K,e}^{n+1,-})\Big)\phi_e^n|K| \\ &+\sum_{n\geq 0}\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^{n})-\mathcal{U}(w_K^n)\Big)\phi_e^n|K|,\end{aligned} \tag{5.123}$$

with other error terms in the right hand side of the discrete entropy inequalities (5.120). The former must therefore be proved to go to zero with h. Sharp estimates of these error terms are precisely the subject of :

Lemma 5.10. *For any given polyhedron K in $\mathcal{T}_h$, we have:*

$$\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^{n})-\mathcal{U}(w_K^n)\Big)\phi_e^n\leq O(h^2)\|\phi\|_{W^{1,\infty}(]t^n,t^{n+1}[\times K)}, \tag{5.124}$$

while

$$\begin{aligned}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_K^{n+1})-\mathcal{U}(w_{K,e}^{n+1,-})\Big)\phi_e^n\leq -\sigma_{\mathcal{U}}\Big(\sum_{e\in\partial K}\alpha_{K,e}|w_K^{n+1}-w_{K,e}^{n+1,-}|^2\Big)\phi_K^n \\ &+O(h)\Big(\sum_{e\in\partial K}\alpha_{K,e}|w_K^{n+1}-w_{K,e}^{n+1,-}|\Big)\|\nabla\phi\|_{L^\infty(]t^n,t^{n+1}[\times K)}.\end{aligned} \tag{5.125}$$

where $\sigma_{\mathcal{U}}$ denotes some convexity-like modulus of $\mathcal{U}$: $\mathcal{U}''(u)\geq\sigma_{\mathcal{U}}>0$, for all $u\in(m,M)$ where the bounds m,M were introduced in (5.81) in agreement with the maximum principle (5.100).

We now turn proving the statements of this section. Let us first start with the preliminary lemma 5.8.

Proof of Lemma 5.8. Let e be any given edge in $\mathcal{T}_h$ and K, K_e the associated pair of adjacent polyhedra. Multiplying the subcell entropy inequality (5.117) valid for K by $\alpha_{K,e}|K|$ and the companion inequality for K_e by $\alpha_{K,e}|K_e|$, we get by summation:

$$\alpha_{K,e}|K|\Big(\mathcal{U}(w_{K,e}^{n+1,-}) - \mathcal{U}(w_{K,e}^{n})\Big) + \alpha_{K_e,e}|K_e|\Big(\mathcal{U}(w_{K_e,e}^{n+1,-}) - \mathcal{U}(w_{K_e,e}^{n})\Big) - \tau\Big(\mathcal{F}(w_{K,e}^{n}, v_e)\cdot\nu_{K,e} + \mathcal{F}(w_{K_e,e}^{n}, v_e)\cdot\nu_{K_e,e}\Big)|e| \le 0,$$

thanks to the conservation property (5.116) verified by the numerical entropy fluxes. Multiplying the above inequality by the discrete test function ϕ_e^n (5.118), then summing over the edges e in ∂K and the polyhedra K in $\mathcal{T}_h$ yield

$$\begin{aligned}
&\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^{n+1,-}) - \mathcal{U}(w_{K,e}^{n})\Big)\phi_e^n|K| \\
&+ \sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K_e,e}\Big(\mathcal{U}(w_{K_e,e}^{n+1,-}) - \mathcal{U}(w_{K_e,e}^{n})\Big)\phi_e^n|K_e| \\
&- \tau\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\Big(\mathcal{F}(w_{K,e}^{n}, v_e)\cdot\nu_{K,e} + \mathcal{F}(w_{K_e,e}^{n}, v_e)\cdot\nu_{K_e,e}\Big)\phi_e^n|e| \le 0.
\end{aligned}$$

To conclude the proof, it suffices to notice the following two identities :

$$\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^{n+1,-}) - \mathcal{U}(w_{K,e}^{n})\Big)\phi_e^n|K| = \sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K_e,e}\Big(\mathcal{U}(w_{K_e,e}^{n+1,-}) - \mathcal{U}(w_{K_e,e}^{n})\Big)\phi_e^n|K_e|,$$

and

$$\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\mathcal{F}(w_{K_e,e}^{n}, v_e)\cdot\nu_{K_e,e}\phi_e^n|e| = \sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\mathcal{F}(w_{K,e}^{n}, v_e)\cdot\nu_{K,e}\phi_e^n|e|.$$

□

Proof of Proposition 5.9. Let us start from the discrete inequality (5.119) of Lemma 5.8, focusing our attention on the flux balance:

$$\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\mathcal{F}(w_{K,e}^{n}, v_e)\cdot\nu_{K,e}\phi_e^n|e|. \tag{5.126}$$

Our purpose is to shift the mathematical expressions under consideration from the w to the u variable. Hence let us write $\mathcal{F}(w_{K,e}^{n}, v_e) = \mathcal{F}(w(u_K^n, v_e), v_e) = \mathcal{Q}(u_K^n, v_e)$ with $\mathcal{Q}(u,v)$ the exact entropy flux introduced in (5.55), which we repeat component-wise for convenience to the reader:

$$\mathcal{Q}_i(u,v) = \int^u \mathcal{U}'(C_0(\theta,v))\partial_\theta C_i(\theta,v)d\theta, \quad 1\le i\le d. \tag{5.127}$$

We then recast the flux balance as

$$\begin{aligned}\sum_{e\in\partial K}\mathcal{F}(w^n_{K,e},v_e)\cdot\nu_{K,e}\phi^n_e|e| = Q(u^n_K,v_K)\cdot\sum_{e\in\partial K}\phi^n_e|e|\nu_{K,e}\\ +\sum_{e\in\partial K}\Big(Q(u^n_K,v_e)-Q(u^n_K,v_K)\Big)\cdot\nu_{K,e}\phi^n_e|e|,\end{aligned}\tag{5.128}$$

where we have introduced the following $\{\alpha_{K,e}\}_{e,e\in\partial K}$-average of the v_e:

$$v_K=\sum_{e\in\partial K}\alpha_{K,e}v_e.\tag{5.129}$$

But in view of a representation formula for $\nabla\phi$ similar to the one (5.70) derived for ∇v, the proposed average form (5.118) for ϕ^n_e yields:

$$\sum_{e\in\partial K}\phi^n_e|e|\nu_{K,e}=\frac{1}{\tau}\int_{t^n}^{t^{n+1}}\Big(\sum_{e\in\partial K}\int_e\phi(x,t)\nu_{K,e}dx\Big)dt=\frac{1}{\tau}\int_{t^n}^{t^{n+1}}\int_K\nabla\phi(x,t)dxdt,\tag{5.130}$$

so that we get from (5.128):

$$\begin{aligned}\sum_{e\in\partial K}\mathcal{F}(w^n_{K,e},v_e)\cdot\nu_{K,e}\phi^n_e|e| = \frac{1}{\tau}\int_{t^n}^{t^{n+1}}\int_K Q(u^n_K,v_K)\cdot\nabla\phi(x,t)dxdt\\ +\sum_{e\in\partial K}\Big(Q(u^n_K,v_e)-Q(u^n_K,v_K)\Big)\cdot\nu_{K,e}\phi^n_e|e|.\end{aligned}\tag{5.131}$$

The treatment of the last remaining discrete term relies on the following identity:

$$Q(u^n_K,v_e)-Q(u^n_K,v_K)=\int_0^1\partial_v Q(u^n_K,v_K+s(v_e-v_K))ds\;(v_e-v_K)\tag{5.132}$$

which leads us to rewrite (5.131), using the tensor notation and its property(5.50), according to:

$$\begin{aligned}&\sum_{e\in\partial K}\mathcal{F}(w^n_{K,e},v_e)\cdot\nu_{K,e}\phi^n_e|e|-\frac{1}{\tau}\int_{t^n}^{t^{n+1}}\int_K Q(u^n_K,v_K)\cdot\nabla\phi(x,t)dxdt=\\ &+\partial_v Q(u^n_K,v_K):\Big(\sum_{e\in\partial K}\phi^n_e(v_e-v_K)\otimes\nu_{K,e}|e|\Big)\\ &+\sum_{e\in\partial K}\phi^n_e\Big(\int_0^1\Big(\partial_v Q(u^n_K,v_K+s(v_e-v_K))-\partial_v Q(u^n_K,v_K)\Big)ds\Big):\Big((v_e-v_K)\otimes\nu_{K,e}\Big)|e|.\end{aligned}\tag{5.133}$$

The matrix $(v_e-v_K)\otimes\nu_{K,e}|e|$ with size $L\times d$ appears as a discrete representation for the continuous function ∇v. The first term in the above right hand side is conveniently rewritten

as:

$$\begin{aligned}\partial_v \boldsymbol{Q}(u_K^n, v_K) : \Big(\sum_{e\in\partial K} \phi_e^n (v_e - v_K)\otimes \nu_{K,e}|e| \Big) \\ = \phi_K^n\, \partial_v \boldsymbol{Q}(u_K^n, v_K) : \Big(\sum_{e\in\partial K} (v_e - v_K)\otimes \nu_{K,e}|e| \Big) \\ + \Big(\sum_{e\in\partial K} (\phi_e^n - \phi_K^n) \partial_v \boldsymbol{Q}(u_K^n, v_K) : \big((v_e - v_K)\otimes \nu_{K,e} \big)|e| \Big),\end{aligned} \tag{5.134}$$

where the discrete flux function ϕ_K^n results from the following natural averaging:

$$\phi_K^n = \sum_{e\in\partial K} \alpha_{K,e} \phi_e^n. \tag{5.135}$$

In the one hand and owing to the identity

$$\sum_{e\in\partial K} (v_e - v_K)\otimes \nu_{K,e}|e| = \sum_{e\in\partial K} v_e \otimes \nu_{K,e}|e| \tag{5.136}$$

we get

$$\begin{aligned}\partial_v \boldsymbol{Q}(u_K^n, v_K) : \Big(\sum_{e\in\partial K} (v_e - v_K)\otimes \nu_{K,e}|e| \Big) = \\ \partial_v \boldsymbol{Q}(u_K^n, v_K) : \left(\frac{1}{\tau} \int_{t^n}^{t^{n+1}} \int_K \nabla v(x) dt dx \right),\end{aligned} \tag{5.137}$$

again thanks to the representation formula (5.70) for ∇v. In the second hand, the last error term in (5.134) is given the following successive bounds :

$$\begin{aligned}\Big| \sum_{e\in\partial K} (\phi_e^n - \phi_K^n) \partial_v \boldsymbol{Q}(u_K^n, v_K) : \big((v_e - v_K)\otimes \nu_{K,e} \big)|e| \Big| \\ \leq O(1) \sup_{e\in\partial K} |(\phi_e^n - \phi_K^n)(v_e - v_K)|) p_K, \\ \leq O(h_K^2) \|\nabla \phi\|_{L^\infty(]t^n, t^{n+1}[\times K)} p_K, \\ \leq O(h_K)\ \|\nabla \phi\|_{L^\infty(]t^n, t^{n+1}[\times K)} |K|.\end{aligned} \tag{5.138}$$

Here we have successively used the uniform sup-norm estimate (5.100) verified by u_h, the definition (5.63) of the perimeter p_K of K, the easy estimate

$$|v_e - v_K| \leq \sum_{e'\in\partial K} \alpha_{K,e} |v_e - v_{e'}| \leq O(h_K) \tag{5.139}$$

from the definition (5.129) of v_K and the smoothness (5.61) of the color function v, a similar estimate $|\phi_e^n - \phi_k^n| \leq O(h_K)$ and finally the non degeneracy assumption (5.62) on the triangulation $\mathcal{T}_h$. Involving (5.137)-(5.138), the identity (5.134) yields the following estimate

$$\begin{aligned}\Big| \partial_v \boldsymbol{Q}(u_K^n, v_K) : \Big(\sum_{e\in\partial K} \phi_e^n (v_e - v_K)\otimes \nu_{K,e}|e| \Big) - \frac{1}{\tau} \int_{t^n}^{t^{n+1}} \int_K \phi_K^n \partial_v \boldsymbol{Q}(u_K^n, v_K) : \nabla v(x) dt dx \Big| \\ \leq O(h_K)\ \|\nabla \phi\|_{L^\infty(]t^n, t^{n+1}[\times K)} |K|.\end{aligned} \tag{5.140}$$

Turning considering the last error term in the flux balance (5.133), we propose the following bounds:

$$\begin{aligned}\Big|\sum_{e\in\partial K}\phi_e^n\int_0^1(\partial_v\mathcal{Q}(u_K^n,v_K+s(v_e-v_K))-\partial_v\mathcal{Q}(u_K^n,v_K))ds:((v_e-v_K)\otimes\nu_{K,e})|e|\Big|\\ \leq O(1)\sup_{e\in\partial K}|v_e-v_K|^2\Big(p_K\|\phi\|_{L^\infty(]t^n,t^{n+1}[\times K)}\Big)\\ \leq O(h_K)\|\phi\|_{L^\infty(]t^n,t^{n+1}[\times K)}|K|,\end{aligned}\tag{5.141}$$

arguing successively about the smoothness of the exact entropy flux function $\mathcal{Q}$, the uniform sup-norm estimate (5.100) verified by the sequence u_h, the estimate (5.139) satisfied by $|v_e-v_K|$ and finally the non degeneracy assumption (5.62) on the triangulation $\mathcal{T}_h$.
To summarize we have proved the estimate for the flux balance on a single cell:

$$\begin{aligned}\Big|\frac{1}{\tau}\int_{t^n}^{t^{n+1}}\int_K\Big(\mathcal{Q}(u_K^n,v_K)\cdot\nabla\phi(x,t)+\phi_K^n\partial_v\mathcal{Q}(u_K^n,v_K):\nabla v(x)dtdx\Big)\\ -\sum_{e\in\partial K}\mathcal{F}(w_{K,e}^n,v_e)\cdot\nu_{K,e}\phi_e^n|e|\Big|\\ \leq O(h)\|\phi\|_{W^{1,\infty}(]t^n,t^{n+1}[\times K)}|K|.\end{aligned}\tag{5.142}$$

Taking care of the discrete weak entropy inequality (5.119) we recall here

$$\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^{n+1,-})-\mathcal{U}(w_{K,e}^n)\Big)\phi_e^n|K|-\tau\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\mathcal{F}(w_{K,e}^n,v_e)\cdot\nu_{K,e}\phi_e^n|e|\leq 0,$$

the sum of (5.142) over all cells K on the triangulation $\mathcal{T}_h$ gives

$$\begin{aligned}&\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^{n+1,-})-\mathcal{U}(w_{K,e}^n)\Big)\phi_e^n|K|\\ &\quad-\int_{t^n}^{t^{n+1}}\Big(\sum_{K\in\mathcal{T}_h}\int_K\mathcal{Q}(u_K^n,v_K)\cdot\nabla\phi+\phi\partial_v\mathcal{Q}(u_K^n,v_K):\nabla v dx\Big)dt\\ &\quad\leq O(h)\tau\sum_{K\in\mathcal{T}_h}\|\phi\|_{W^{1,\infty}(]t^n,t^{n+1}[\times K)}|K|\\ &\quad\leq O(h)\tau\|\phi\|_{W^{1,\infty}(]t^n,t^{n+1}[\times\mathbb{R}^d)}|supp(\phi)|.\end{aligned}$$

□

Proof of Lemma 5.10. Let us first establish the estimate (5.124). In that aim, consider the following decomposition involving again the $\{\alpha_{K,e}\}_{\{e,e\in\partial K\}}$-average ϕ_K^n of the ϕ_e^n (5.122):

$$\begin{aligned}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^n)-\mathcal{U}(w_K^n)\Big)\phi_e^n=\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^n)-\mathcal{U}(w_K^n)\Big)(\phi_e^n-\phi_K^n)\\ +\ \phi_K^n\Big(\sum_{e\in\partial K}\alpha_{K,e}\mathcal{U}(w_{K,e}^n)-\mathcal{U}(w_K^n)\Big),\end{aligned}\tag{5.143}$$

from which we deduce the following bound:

$$\begin{aligned}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w^n_{K,e})-\mathcal{U}(w^n_K)\Big)\phi^n_e&\\ \leq O(h_K)\|\nabla\phi\|_{L^\infty(]t^n,t^{n+1}[\times K)}\sup_{e\in\partial K}|w^n_{K,e}-w^n_K|&\\ +\ O(1)\ \|\phi\|_{L^\infty(]t^n,t^{n+1}[\times K)}\Big(\sum_{e\in\partial K}\alpha_{K,e}\mathcal{U}(w^n_{K,e})-\mathcal{U}(w^n_K)\Big),&\end{aligned}\tag{5.144}$$

in view of the sup-norm estimate (5.100) verified by u_h, the easy estimate $|\phi^n_e-\phi^n_K|\leq O(h_K)$ and the convexity of the entropy $\mathcal{U}(w)$. The first error term in (5.144) is given the following bound:

$$\begin{aligned}|w^n_{K,e}-w^n_K|&\leq\sum_{e'\in\partial K}\alpha_{K,e'}\big|C_0(u^n_K,v_{e'})-C_0(u^n_K,v_e)\big|\\ &\leq O(1)\sup_{e'\in\partial K}|v_{e'}-v_e|\leq O(h_K),\end{aligned}\tag{5.145}$$

while the second one may be handled as follows:

$$\begin{aligned}\sum_{e\in\partial K}\alpha_{K,e}\mathcal{U}(w^n_{K,e})-\mathcal{U}(w^n_K)&=\mathcal{U}'(w^n_K)\Big(\sum_{e\in\partial K}\alpha_{K,e}w^n_{K,e}-w^n_K\Big)\\ &+\sum_{e\in\partial K}\alpha_{K,e}\int_0^1\mathcal{U}''(w^n_{K,e}+s(w^n_K-w^n_{K,e}))ds(w^n_{K,e}-w^n_K)^2\\ &\leq O(1)\sup_{e\in\partial K}|w^n_{K,e}-w^n_K|^2\ \leq\ O(h_K^2),\end{aligned}\tag{5.146}$$

in view of the formula-definition (5.78) $w^n_K=\sum_{e\in\partial K}\alpha_{K,e}w^n_{K,e}$ and the estimate (5.145). Gathering bounds (5.145) and (5.146) yield the expected estimate (5.124) in Lemma 5.10.

Let us turn deriving the companion estimate (5.125). Starting from the decomposition:

$$\begin{aligned}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w^{n+1}_K)-\mathcal{U}(w^{n+1,-}_{K,e})\Big)\phi^n_e=\ &\phi^n_K\Big(\mathcal{U}(w^{n+1}_K)-\sum_{e\in\partial K}\alpha_{K,e}\mathcal{U}(w^{n+1,-}_{K,e})\Big)\\ &+\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w^{n+1}_K)-\mathcal{U}(w^{n+1,-}_{K,e})\Big)\Big(\phi^n_e-\phi^n_K\Big),\end{aligned}\tag{5.147}$$

we observe in the one hand:

$$\begin{aligned}\Big|\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w^{n+1}_K)-\mathcal{U}(w^{n+1,-}_{K,e})\Big)\Big(\phi^n_e-\phi^n_K\Big)\Big|&\\ \leq O(1)\sum_{e\in\partial K}\alpha_{K,e}|\phi^n_e-\phi^n_K\|w^{n+1,-}_{K,e}-w^{n+1}_K|&\\ \leq O(h_K)\Big(\sum_{e\in\partial K}\alpha_{K,e}|w^{n+1,-}_{K,e}-w^{n+1}_K|\Big)\|\nabla\phi\|_{L^\infty(]t^n,t^{n+1}[\times K)},&\end{aligned}\tag{5.148}$$

while in the second hand, we write the Taylor development

$$\begin{aligned}\sum_{e\in\partial K}\alpha_{K,e}\mathcal{U}(w_{K,e}^{n+1,-})-\mathcal{U}(w_K^{n+1})&=\mathcal{U}'(w_K^{n+1})\Big(\sum_{e\in\partial K}\alpha_{K,e}w_{K,e}^{n+1,-}-w_K^{n+1}\Big)\\&+\sum_{e\in\partial K}\alpha_{K,e}\int_0^1\mathcal{U}''(w_{K,e}^{n+1,-}+s(w_K^{n+1}-w_{K,e}^{n+1,-}))ds(w_{K,e}^{n+1,-}-w_K^{n+1})^2.\end{aligned}\tag{5.149}$$

Finally, in view of the convex decomposition (5.86) stating $w_K^{n+1}=\sum_{e\in\partial K}\alpha_{K,e}w_{K,e}^{n+1,-}$

$$\mathcal{U}(w_K^{n+1})-\sum_{e\in\partial K}\alpha_{K,e}\mathcal{U}(w_{K,e}^{n+1,-})\leq-\sigma_{\mathcal{U}}\sum_{e\in\partial K}\alpha_{K,e}|w_{K,e}^{n+1,-}-w_K^{n+1}|^2\tag{5.150}$$

where $\sigma_{\mathcal{U}}$ denotes the convexity like-modulus of $\mathcal{U}$ introduced in Lemma 5.10. This concludes the proof. □

5.6.2 Entropy dissipation rate and a posteriori strong convergence

The proposed estimates obtained in Lemma 5.10 deserve a few comments. Plugging first estimate (5.124) in (5.123) will be easily seen to yield the following upper-bound

$$\begin{aligned}\sum_{n\geq0}\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^n)-\mathcal{U}(w_K^n)\Big)\phi_e^n|K|&\\\leq O(h)\|\phi\|_{W^{1,\infty}(\mathbb{R}_+\times\mathbb{R}^d)}|supp(\phi)|&\end{aligned}$$

that obviously suffices to conclude. By contrast and turning considering (5.125), a crude upper-bound based on the sup-norm estimate (5.101), say

$$\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_K^{n+1})-\mathcal{U}(w_{K,e}^{n+1,-})\Big)\phi_e^n\leq O(h)\|\phi\|_{W^{1,\infty}(]t^n,t^{n+1}[\times K)}$$

would result in the useless estimate

$$\begin{aligned}\sum_{n\geq0}\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_K^{n+1})-\mathcal{U}(w_{K,e}^{n+1,-})\Big)\phi_e^n|K|&\\\leq O(1)\|\phi\|_{W^{1,\infty}(\mathbb{R}_+\times\mathbb{R}^d)}|supp(\phi)|.&\end{aligned}$$

Proving that the error term of concern in (5.123) actually vanishes with h requires therefore in turn a sharper control in (5.125) of the oscillations of the $w_{K,e}^{n+1,-}$ around their mean value w_K^{n+1}. Such a control over these discrete oscillations results from a sharp evaluation of the discrete entropy rate of dissipation. Its detailed form is the matter of the next statement:

Proposition 5.11. *Let $T>0$ be any given fixed time and let $N_T\in\mathbb{N}$ be the floor of T/τ we denote $[T/\tau]$. Then the finite volume approximation (5.77)–(5.80) obeys the following estimate on the discrete oscillations:*

$$\sum_{n=0}^{N_T}\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}|w_K^{n+1}-w_{K,e}^{n+1,-}|^2\psi_K|K|\leq O(1),\tag{5.151}$$

where ψ_K writes

$$\psi_K = \sum_{e\in\partial K} \alpha_{K,e}\psi_e, \qquad \psi_e = \frac{1}{|e|}\int_e \psi(x)dx, \tag{5.152}$$

for any given (time independent) non negative test function $\psi \in \mathcal{D}(\mathbb{R}^d)$.

Equipped with (5.151) we shall be in a position to infer the following entropy dissipation rate:

Corollary 5.12. *The sequence u_h verify the entropy like inequality*

$$\iint_{\mathbb{R}^+\times\mathbb{R}^d} \mathcal{U}(C_0(u_h,v))\partial_t\phi(x,t) + \mathcal{Q}(u_h,v)\cdot\nabla\phi + \phi\partial_v\mathcal{Q}(u_h,v):\nabla v dxdt \geq O(h^{1/2}), \tag{5.153}$$

for any given (smooth) convex entropy pair $(\mathcal{U},\mathcal{Q}) : \mathbb{R}\to\mathbb{R}\times\mathbb{R}^d$ introduced in (5.57) and (5.55).

Equipped with the above inequality valid for any given entropy pairs $(\mathcal{U},\mathcal{Q})$, we easily infer that the Young measure $\mu = \mu_{t,x}$ associated with the sequence $(u_h)_{h>0}$ is an entropy satisfying measure valued solution.

$$\partial_t\langle\mu,\mathcal{U}(C_0(\cdot,v))\rangle + \nabla_x\langle\mu,\mathcal{Q}(\cdot,v)\rangle - \langle\mu,\partial_v\mathcal{Q}(\cdot,v)\rangle:\nabla v \leq 0, \quad \mathcal{D}'. \tag{5.154}$$

Involving an extended version of the DiPerna's uniqueness theorem (see [34] for a proof), we deduce that the entropy measure-valued solution $\mu_{t,x}$ reduces to a Dirac measure $\delta_{u(t,x)}$ concentrated on a function $u(t,x)$ since the initial data μ_0 coincide with the Dirac measure δ_{u_0} where u_0 stands for the initial data in the Cauchy problem (**??**). Proving that the inital data u_0 is correctly handled amounts to show that for every compact subset $\mathcal{K}$ of $\mathbb{R}$ we have

$$\lim_{t\to 0^+}\int_0^t\int_{\mathcal{K}}\langle\mu_{s,x},|id-u_0(x)|\rangle\, dxds = 0. \tag{5.155}$$

This issue can be easily inferred from the previous analysis. This step is left to the reader. Theorem 5.2 is then established.

Proof of Proposition 5.11. Let us start from the discrete in time weak formulation (5.120) stated in Proposition 5.9 we repeat here for convenience

$$\begin{aligned}&\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^{n+1,-}) - \mathcal{U}(w_{K,e}^n)\Big)\phi_e^n|K| \\ &- \iint_{]t^n,t^{n+1}[\times\mathbb{R}^d} \mathcal{Q}(u_h^n,v(x))\cdot\nabla\phi(x,t) + \phi(x,t)\partial_v\mathcal{Q}(u_h^n,v(x)):\nabla v(x)dxdt \\ &\qquad \leq O(h)\tau\,\|\phi\|_{W^{1,\infty}(]t^n,t^{n+1}[\times\mathbb{R}^d)}|supp(\phi)|,\end{aligned} \tag{5.156}$$

in which we plug the decomposition (5.121)-(5.122). A discrete test function ψ_K given by

(5.152) for any given time independant test function $\psi \in \mathcal{D}(\mathbb{R}^d)$ is considered. We then get:

$$\begin{aligned}
&\sum_{K\in\mathcal{T}_h} \Big(\mathcal{U}(w_K^{n+1}) - \mathcal{U}(w_K^n)\Big)\psi_K|K| \\
&- \iint_{]t^n,t^{n+1}[\times\mathbb{R}^d} Q(u_h^n, v(x))\cdot\nabla\psi(x) + \psi(x)\partial_v Q(u_h^n, v(x)) : \nabla v(x)dxdt \\
&\qquad \le \sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K} \alpha_{K,e}\Big(\mathcal{U}(w_K^{n+1}) - \mathcal{U}(w_{K,e}^{n+1,-})\Big)\psi_e|K| \\
&\qquad + \sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K} \alpha_{K,e}\Big(\mathcal{U}(w_{K,e}^{n}) - \mathcal{U}(w_K^{n}))\Big)\psi_e|K| \\
&\qquad + O(h)\tau\|\psi\|_{W^{1,\infty}(\mathbb{R}^d)}|supp(\psi)|.
\end{aligned} \tag{5.157}$$

Invoquing estimates (5.124)-(5.125) then yields

$$\begin{aligned}
&\sum_{K\in\mathcal{T}_h} \Big(\mathcal{U}(w_K^{n+1}) - \mathcal{U}(w_{K,e}^{n+1,-})\Big)\psi_K|K| \\
&\quad + \sigma_{\mathcal{U}} \sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K} \alpha_{K,e}|w_K^{n+1} - w_{K,e}^{n+1,-}|^2\psi_K|K| \\
&\quad \le O(h)\tau\|\psi\|_{W^{1,\infty}(\mathbb{R}^d)}|supp(\psi)| \\
&\quad + O(h)\sum_{K\in\mathcal{T}_h} \|\nabla\psi\|_{L^\infty(K)}|K| \\
&\quad + O(h^2)\sum_{K\in\mathcal{T}_h} \|\psi\|_{W^{1,\infty}(K)}|K| \\
&\quad + \iint_{]t^n,t^{n+1}[\times\mathbb{R}^d} Q(u_h^n, v(x))\cdot\nabla\psi(x) + \psi(x)\partial_v Q(u_h^n, v(x)) : \nabla v(x)dxdt.
\end{aligned} \tag{5.158}$$

Observe that due to the estimate (5.100), the last contribution in the above right hand side can be given the following crude estimate $O(\tau)\|\psi\|_{W^{1,\infty}(\mathbb{R}^d)}$. Henceforth, we infer that

$$\begin{aligned}
&\sum_{K\in\mathcal{T}_h} \Big(\mathcal{U}(w_K^{n+1}) - \mathcal{U}(w_K^n)\Big)\psi_K|K| + \sigma_{\mathcal{U}} \sum_{K\in\mathcal{T}_h} \Big(\sum_{e\in\partial K} \alpha_{K,e}|w_K^{n+1} - w_{K,e}^{n+1,-}|^2\Big)\psi_K|K| \\
&\qquad \le O(h)\|\psi\|_{W^{1,\infty}(\mathbb{R}^d)}.
\end{aligned} \tag{5.159}$$

Summing over time indices $n \in [0, N_T]$ with $N_T = [T/\tau]$ for a given fixed time $T > 0$, we get

$$\begin{aligned}
&\int_{\mathbb{R}^d} \mathcal{U}(w_h(x,T))\psi_h(x)dx + \sigma_{\mathcal{U}} \sum_{n=0}^{N_T}\sum_{K\in\mathcal{T}_h} \Big(\sum_{e\in\partial K} \alpha_{K,e}|w_K^{n+1} - w_{K,e}^{n+1,-}|^2\Big)\psi_K|K| \\
&\qquad \le \int_{\mathbb{R}^d} \mathcal{U}(w_0(x))\psi_h(x)dx + O(1)T\|\psi\|_{W^{1,\infty}(\mathbb{R}^d)}.
\end{aligned} \tag{5.160}$$

which is the required result. □

Proof of Corollary 5.12. Let us again start from (5.120)-(5.121)-(5.122) to consider the following discrete in time weak formulation for the time dependant test function $\phi \in \mathcal{D}(\mathbb{R}_*^+ \times \mathbb{R}^d)$

and its discrete representation ϕ_K^n

$$\begin{aligned}
&\sum_{K\in\mathcal{T}_h}\Big(\mathcal{U}(w_K^{n+1})-\mathcal{U}(w_K^n)\Big)\phi_K^n|K| \\
&\quad-\iint_{]t^n,t^{n+1}[\times\mathbb{R}^d}\mathcal{Q}(u_h^n,v(x))\cdot\nabla\phi(x,t)+\phi(x,t)\partial_v\mathcal{Q}(u_h^n,v(x)):\nabla v(x)dxdt \\
&\quad\le O(h)\tau\|\phi\|_{W^{1,\infty}(]t^n,t^{n+1}[\times\mathbb{R}^d)}|supp(\phi)| \\
&\quad+\ O(h^2)\sum_{K\in\mathcal{T}_h}\|\phi\|_{W^{1,\infty}(]t^n,t^{n+1}[\times K)}|K| \\
&\quad+\ O(h)\sum_{K\in\mathcal{T}_h}\sum_{e\in\partial K}\alpha_{K,e}|w_K^{n+1}-w_{K,e}^{n+1,-}|\,\|\nabla\phi\|_{L^\infty(]t^n,t^{n+1}[\times K)}|K|.
\end{aligned}\tag{5.161}$$

where we have used estimates (5.124)-(5.125). Summing this inequality over time indices gives

$$\begin{aligned}
&-\sum_{n\ge0}\sum_{K\in\mathcal{T}_h}\mathcal{U}(w_K^{n+1})\frac{\phi_K^{n+1}-\phi_K^n}{\tau}\tau|K| \\
&-\iint_{\mathbb{R}^+\times\mathbb{R}^d}\mathcal{Q}(u_h^n,v(x))\cdot\nabla\phi(x,t)+\phi(x,t)\partial_v\mathcal{Q}(u_h^n,v(x)):\nabla v(x)dxdt \\
&\le O(h)\|\phi\|_{W^{1,\infty}(\mathbb{R}^+\times\mathbb{R}^d)} \\
&+O(1)\sum_{n\ge0}\sum_{K\in\mathcal{T}_h}\Big(\sum_{e\in\partial K}\alpha_{K,e}|w_K^{n+1}-w_{K,e}^{n+1,-}|\,\chi_\phi\|\nabla\phi\|_{L^\infty(]t^n,t^{n+1}[\times K)}|K|\tau\Big),
\end{aligned}\tag{5.162}$$

making use of the characteristic function χ_ϕ of the following compact subset of $\mathbb{R}^d$:

$$\bigcup_{0<t<T}supp(\phi(\cdot,t)),\tag{5.163}$$

where T is a finite time such that $supp(\phi(\cdot,t))=\emptyset$ for $t\ge T$. A Cauchy-Schwartz inequality then yields the following crude upper bound for the last term

$$\begin{aligned}
&\sum_{n\ge0}\sum_{K\in\mathcal{T}_h}\Big(\sum_{e\in\partial K}\alpha_{K,e}|w_K^{n+1}-w_{K,e}^{n+1,-}|\,\chi_\phi\Big)\|\nabla\phi\|_{L^\infty(]t^n,t^{n+1}[\times K)}|K|\tau \\
&\le\Big(\sum_{n\ge0}\sum_{K\in\mathcal{T}_h}\Big(\sum_{e\in\partial K}\alpha_{K,e}|w_K^{n+1}-w_{K,e}^{n+1,-}|\,\chi_\phi\Big)^2|K|\tau\Big)^{1/2} \\
&\qquad\times\Big(\sum_{n\ge0}\sum_{K\in\mathcal{T}_h}\|\nabla\phi\|^2_{L^\infty(]t^n,t^{n+1}[\times K)}|K|\tau\Big)^{1/2} \\
&\le O(1)\Big(\sum_{n\ge0}\sum_{K\in\mathcal{T}_h}\Big(\sum_{e\in\partial K}\alpha_{K,e}|w_K^{n+1}-w_{K,e}^{n+1,-}|^2\Big)\chi_\phi|K|\tau\Big)^{1/2}
\end{aligned}\tag{5.164}$$

as an easy consequence of the convexity property of the $\alpha_{K,e}-$average. The estimate (5.151) then yields with $\psi=\chi_\phi$

$$\sum_{n\ge0}\sum_{K\in\mathcal{T}_h}\Big(\sum_{e\in\partial K}\alpha_{K,e}|w_K^{n+1}-w_{K,e}^{n+1,-}|\,\chi_\phi\Big)\|\nabla\phi\|_{L^\infty(]t^n,t^{n+1}[\times K)}|K|\tau\le O(h^{1/2}).\tag{5.165}$$

Then routine arguments give the conclusion from (5.162). □

5.7 Numerical experiments

5.7.1 A two domains coupling problem

In this first test, we consider an heterogeneous medium located in the spatial domain $[-1,1]^2$ and constituted by an annular inclusion $\mathcal{D}_1$ centered at the origin $(0,0)$ with external radius $\sqrt{0.2}$ and with internal radius $\sqrt{0.1}$, and by its complement set $\mathcal{D}_0$. In these two domains, the following respective fluxes are considered in term of the unknown $w = w(t,x) \in \mathbb{R}$:

$$f_0(w) = w^2/2\begin{pmatrix}1\\1\end{pmatrix}, \quad f_1(w) = (w-0.9)^2/2\begin{pmatrix}1\\1\end{pmatrix}.$$

The smoothed color function v is represented on Figure 5.3b and consists in a smooth representation fo the characteristic function of $\mathcal{D}_1$. The coupling condition between $\mathcal{D}_0$ and $\mathcal{D}_1$ takes here the form

$$2w_-(t,x) = w_+(t,x), \quad x \in \partial\mathcal{D}_1,$$

where $w_\pm(t,x) = \lim_{\theta\to 0+} w(t, x \pm \theta \nu_x)$, with ν_x the incoming unit normal vector at $x \in \partial\mathcal{D}_1$.

The considered initial data, represented on Figure 5.3a, consists in the piecewise constant function

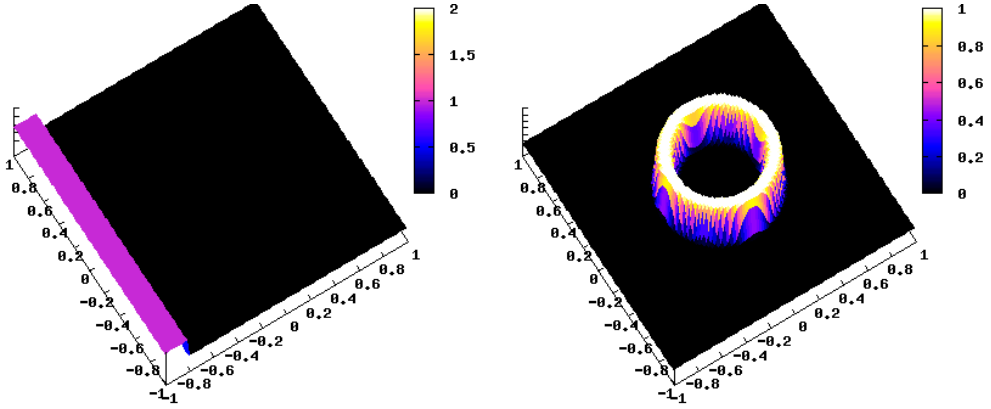
$$w_0(x,y) = \begin{cases} 1, & x < -0.8, \\ 0, & x \geq -0.8. \end{cases}$$

The computations are performed on a cartesian grid with 100×100 meshes. The CFL number is 0.5.

In a homogeneous domain with the sole flux f_0, such an initial data would develop a shock front moving with the uniform speed vector $0.5(1,1)^T$. In the present heterogeneous domain, this shock front has the same behaviour until it reaches the interface between both domains (see Figures 5.4a). The coupling condition at this interface is such that the value $w = 2$ appears then inside the domain $\mathcal{D}_+$. In this second domain, where the flux in consideration is f_1, we observe then a (non-straight) shock wave between the states $w = 2$ and $w = 0$ moving at the uniform speed given by the Rankine Hugoniot jump relation: $0.605(1,1)^T$ (see Figures 5.4c and 5.4e). Finally, the shock front go outside the whole domain $[-1,1]^2$ (see Figure 5.4g). On the right figures (5.4b, 5.4d, 5.4f, 5.4h), we present the u variable, that remains constant at each interface as expected.

5.7.2 A three domains coupling problem

In this second test, we consider three different domains as represented by the two components of v (see Figure **??**). The domain $\mathcal{D}_2$ is a triangular inclusion and the domain $\mathcal{D}_1$ is the complement of $\mathcal{D}_2$ relative to an annular inclusion. The fluxes in consideration are



(a) Initial data w_0 *(b) Color function v*

Figure 5.3: Initial data for the multidimensional test.

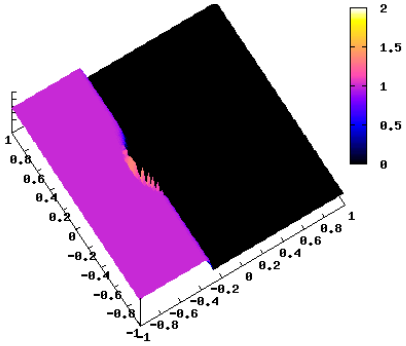
now the following

$$f_0(w) = w^2/2\begin{pmatrix}1\\0\end{pmatrix}, \quad f_1(w) = w^2/2\begin{pmatrix}0.5\\0\end{pmatrix}, \quad f_2(w) = w^2/2\begin{pmatrix}0\\1\end{pmatrix}, \tag{5.166}$$

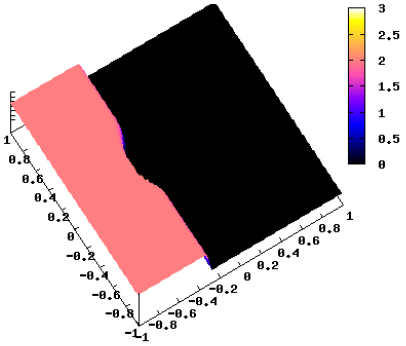
and the coupling relations are given through the change of unknown (5.35) where

$$\theta_0(w) = w, \quad \theta_1(w) = w/2, \quad \theta_2(w) = w/3. \tag{5.167}$$

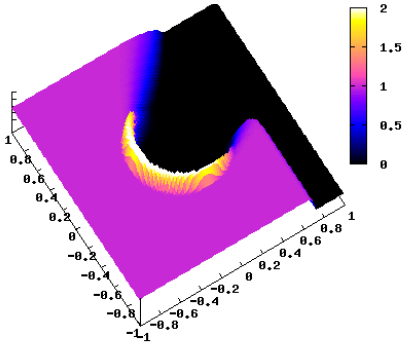
We consider the same initial data as previously, thus we expect a value $w = 2$ to appear in $\mathcal{D}_1$ and a value $w = 3$ in $\mathcal{D}_2$. The results are represented on Figures 5.6a to 5.6f. Once again, the limiting solution as the time grows satisfies the expected coupling relation.



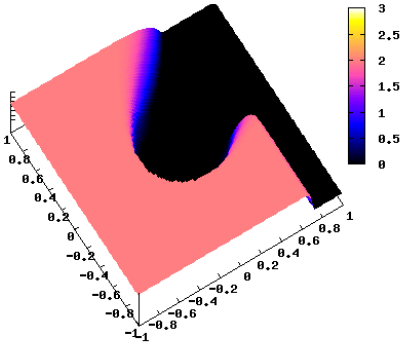
(a) Solution w at $t = 0.5$



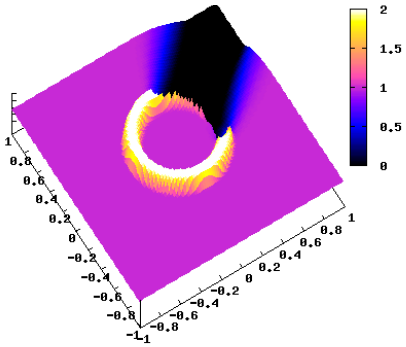
(b) Solution u at $t = 0.5$



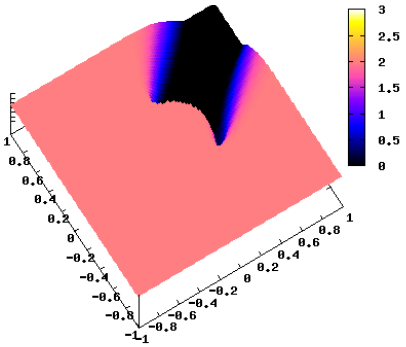
(c) Solution w at $t = 1.5$



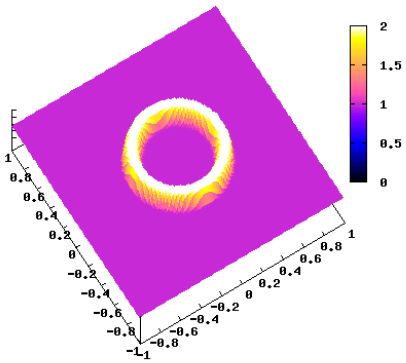
(d) Solution u at $t = 1.5$



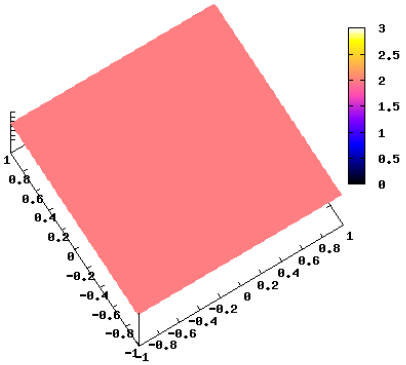
(e) Solution w at $t = 2.5$



(f) Solution u at $t = 2.5$

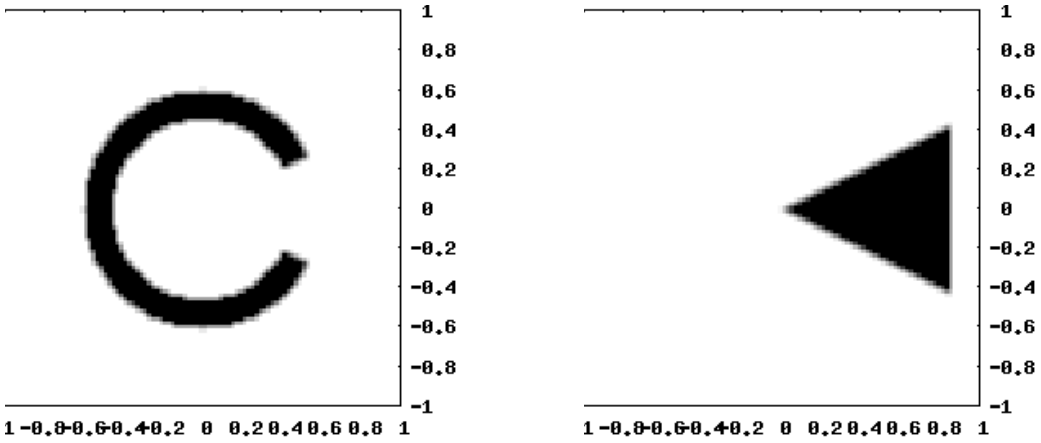


(g) Solution w at $t = 4.5$



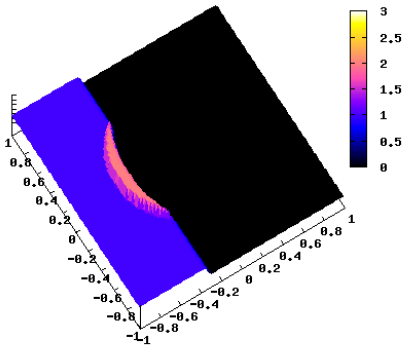
(h) Solution u at $t = 4.5$

Figure 5.4: Evolution of the solution for different times : w (left) and u (right).

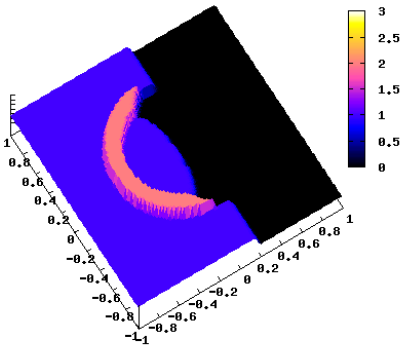


(a) Domain $\mathcal{D}_1$. *(b) Domain $\mathcal{D}_2$.*

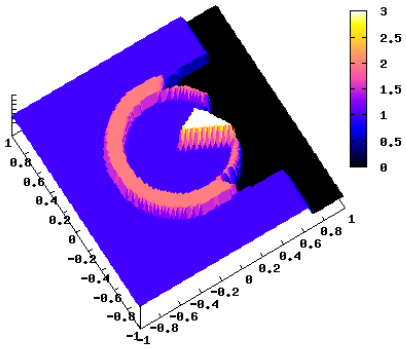
Figure 5.5: Geometry of the three domains.



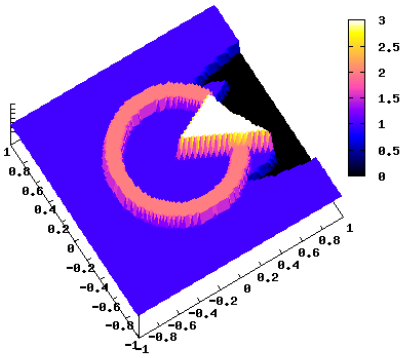
(a) Solution w at $t = 1.0$



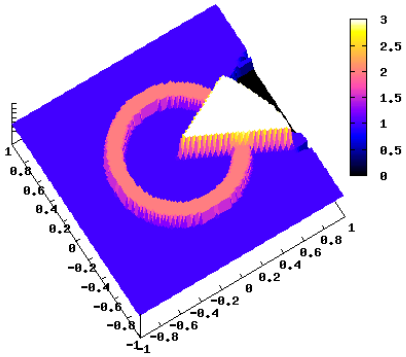
(b) Solution w at $t = 2.0$



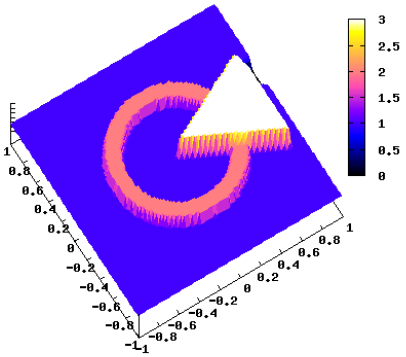
(c) Solution w at $t = 3.0$



(d) Solution w at $t = 4.0$



(e) Solution w at $t = 5.0$



(f) Solution w at $t = 6.0$

Figure 5.6: Three domains - evolution of the solution w.

Seconde partie

Solutions non classiques de lois de conservation hyperboliques

6 Convergent and conservative schemes for nonclassical solutions based on kinetic relations

Dans ce chapitre, nous proposons une nouvelle approche numérique permettant de calculer des solutions non classiques de lois de conservation hyperboliques. Les schémas aux différences finies présentés ici sont strictement conservatifs et conservent au niveau numérique le caractère discontinu des chocs non classiques, contrairement aux schémas de différences finies habituels. La difficulté principale consiste à réaliser, au niveau discret, une propriété de consistance avec une relation cinétique donnée, requise pour la sélection des chocs non classiques physiquement pertinents. Notre méthode s'appuie sur une reconstruction de la solution numérique, effectuée dans chaque maille de calcul pouvant contenir un choc non classique. Afin de valider cette approche, nous établissons plusieurs propriétés de consistance et de stabilité, et nous réalions des essais numériques approfondis. La convergence de l'algorithme vers les solutions physiquement valables satisfaisant à une relation cinétique donnée est illustrée numériquement pour plusieurs cas-tests, pour des flux concave-convexes comme pour des flux convexes-concaves.

Ce travail a été réalisé en collaboration avec Christophe Chalons, Frédéric Lagoutière et Philippe G. LeFloch et a été publié dans la revue *Interfaces and Free Boundaries* [29].

Convergent and conservative schemes for nonclassical solutions based on kinetic relations. I.

Benjamin Boutin, Christophe Chalons,
Frédéric Lagoutière and Philippe G. LeFloch

Abstract

We propose a new numerical approach to compute nonclassical solutions to hyperbolic conservation laws. The class of finite difference schemes presented here is fully conservative and keep nonclassical shock waves as sharp interfaces, contrary to standard finite difference schemes. The main difficulty is to achieve, at the discretization level, a consistency property with a prescribed kinetic relation which is known to be required for the selection of physically meaningful nonclassical shocks. Our method is based on a reconstruction technique performed in each computational cell that may contain a nonclassical shock. To validate this approach, we establish several consistency and stability properties, and we perform extensive numerical experiments. The convergence of the algorithm toward the physically meaningful solutions selected by a given kinetic relation is demonstrated numerically for several test cases, including concave-convex as well as convex-concave flux-functions.

6.1 Introduction

State of the art

We are interested here in the challenging issue of numerically computing *nonclassical* solutions (containing undercompressive shocks) to nonlinear hyperbolic conservation laws. Nonclassical solutions have the distinctive feature of being dynamically driven by small-scale effects such as diffusion, dispersion, and other high-order phenomena. Their selection requires an additional jump relation, called a kinetic relation, and introduced in the context of phase transition dynamics [1, 2, 47, 66, 101, 107, 122, 123, 130, 131], and extensively investigated by LeFloch and collaborators [102] in the context of general hyperbolic systems of conservation laws.

From pioneering work by Hayes and LeFloch [80] it is now recognized that standard finite difference schemes do not converge to nonclassical solutions selected by the prescribed kinetic function. In fact, kinetic functions can be associated not only with continuous models, but with the finite difference schemes themselves. Achieving a good agreement between the continuous and the numerical kinetic functions has been found to be very challenging.

It is worth noticing that the so-called nonclassical shocks and other phase transitions are naturally present in many models from continuum physics, especially in the modeling of real fluids governed by complex equation of state. This is the case, for instance, of models describing the dynamics of liquid-vapor phase changes in compressible fluids, or

of solid-solid phase transformations in materials such as memory alloys. For numerical work in this direction we refer to [46, 84, 115, 116, 133].

In the present paper, we attempt to enforce the validity of the kinetic relation at the numerical level, and we design a *fully conservative* scheme which combines the advantage of standard finite difference and Glimm-type (see below) approaches.

Setting for this paper

We restrict here attention to the scalar conservation law

$$\begin{aligned} &\partial_t u + \partial_x f(u) = 0, \qquad u(x,t) \in \mathbb{R},\ (x,t) \in \mathbb{R}\times\mathbb{R}^+, \\ &u(x,0) = u_0(x), \end{aligned} \tag{6.1}$$

and postpone the discussion of systems of conservation laws to the follow-up paper [29]. The above equation must be supplemented with an entropy inequality of the form

$$\partial_t U(u) + \partial_x F(u) \le 0. \tag{6.2}$$

Here, t denotes the time variable, x is the one dimensional space variable, $f : \mathbb{R} \to \mathbb{R}$ the flux function, and (U,F) is any strictly convex mathematical entropy pair, that is, $U : \mathbb{R} \to \mathbb{R}$ is strictly convex and $F : \mathbb{R} \to \mathbb{R}$ is given by $F' = U' f'$. Equations (6.1) and (6.2) are imposed in the distributional sense.

Importantly, the flux f is assumed to be *nonconvex*. This is a source of mathematical and numerical difficulties. From the mathematical standpoint, a single entropy inequality like (6.2) does not suffice to select a unique solution. This can be seen already at the level of the Riemann problem, corresponding to (6.1)-(6.2) when u_0 has the piecewise constant form

$$u_0(x) = \begin{cases} u_l & \text{if} \quad x < 0, \\ u_r & \text{if} \quad x > 0, \end{cases} \tag{6.3}$$

u_l and u_r being constant states. The Riemann problem admits (up to) a one-parameter family of solutions (see Chapter 2 in [102]). However, these solutions contain discontinuities violating the standard Lax shock inequalities, which are referred to as *nonclassical*. They are important from the physical standpoint, so still should be retained. This non-uniqueness can be fixed however, provided an additional algebraic condition, the so-called *kinetic relation,* is imposed on each nonclassical shock. Consider a shock connecting a left-hand state u_- to a right-hand state u_+ and propagating with the speed σ given by the usual Rankine-Hugoniot relation, that is,

$$u(x,t) = \begin{cases} u_- & \text{if} \quad x < \sigma t, \\ u_+ & \text{if} \quad x > \sigma t, \end{cases} \qquad \text{with} \quad \sigma = \sigma(u_-,u_+) = \frac{f(u_+) - f(u_-)}{u_+ - u_-}. \tag{6.4}$$

Then, the kinetic relation takes the form

$$u_+ = \varphi^\flat(u_-) \ \text{ for all nonclassical shocks,} \tag{6.5}$$

where $\varphi^\flat$ is the so-called *kinetic function*. Equivalently, denoting by $\varphi^{-\flat}$ the inverse of the kinetic function it may be preferably to write $u_- = \varphi^{-\flat}(u_+)$. The kinetic relation implies that the right-hand (respectively left-hand) state is no longer free (as in a classical shock wave) but depends explicitly on the left-hand (respectively right-hand) state. We refer the reader to the textbook [102] for the general theory of nonclassical solutions based on a kinetic relation.

Objectives in this paper

At the numerical level, several strategies exist in the literature in order to take into account the kinetic relation (6.5). We can distinguish between diffuse interface methods and sharp interface methods.

In the first approach, one assumes that the kinetic relation is derived from an augmented continuous model and, in order to take into account the internal structure of nonclassical discontinuities, one attempts to resolve the effects dues to (small) diffusive and dispersive terms that generate them. It is then possible to construct conservative schemes that mimic at the numerical level the effect of the regularized models. Due to the great sensitivity of nonclassical solutions with respect to small scales and numerical diffusion, it turns out that numerical results are satisfactory for shocks with moderate amplitude, but discrepancies between the exact and the numerical kinetic function arise with shocks with large amplitudes and in long-time computations. For this circle of ideas we refer the reader to [79, 80], and the follow-up papers [46, 47, 105].

In the second approach, small scale features are not explicitly taken into account. Instead, the kinetic relation is directly included, in a way or another, in the design of the numerical scheme. This is the case of the random choice and front tracking schemes. It should be mentioned here that the Glimm scheme and front tracking schemes do converge to exact solutions even in presence of nonconclassical shocks; see [101, 102, 107] for the theoretical aspects and Chalons and LeFloch [48] for a numerical study of the Glimm scheme. These schemes require the explicit knowledge of the underlying nonclassical Riemann solver, which may be expensive numerically, and this difficulty motivated the construction of the so-called transport-equilibrium scheme recently proposed in Chalons [44, 45].

More recently, Merckle and Rohde [116] developed a ghost-fluid type algorithm for a model of dynamics of phase transition. These schemes provide good numerical results, as nonclassical discontinuities are sharply and accurately computed. Although the convergence of the methods was demonstrated numerically, their main drawback in practice is similar to the Glimm-type schemes and the property of strict conservation of the conservative variable u fails.

Building on these previous works, our objective in this paper is to design a fully conservative, finite difference scheme for the approximation of nonclassical solutions to the hyperbolic conservation law (6.1). Our basic strategy relies on the discontinuous reconstruction technique proposed recently in Lagoutière [96, 97] which has been found to be particularly efficient to computing *classical* solutions of (6.1) with moderate numerical diffusion.

In our approach below, the kinetic function $\varphi^\flat$ is included explicit, in such a way that nonclassical shock are computed (essentially) exactly while classical shocks suffer a limited numerical diffusion. To validate our strategy we perform extensive numerical experiments and in particular draw the kinetic function associated with the scheme. As the mesh is refined, we observe that the approximate kinetic function converges toward the analytic kinetic function. The scheme also enjoys several fundamental stability properties of consistency with the conservative form of the equation. Interestingly enough, like the Glimm scheme, it also reproduces exactly single nonclassical discontinuities.

6.2 Nonclassical Riemann solver with kinetics

Assumption on the flux-function

We describe here the nonclassical Riemann solver introduced and investigated in LeFloch [102]. Note in passing that this solver was later extended in [107] to include also a nucleation criterion.

Consider the problem (6.1)-(6.2)-(6.5) for a given Riemann initial data (6.3). Throughout this paper we assume that the flux f is either *concave-convex* or *convex-concave*, that is, satisfies the conditions (for all $u \neq 0$)

$$uf''(u) > 0, \qquad f'''(0) \neq 0, \qquad \lim_{|u|\to+\infty} f'(u) = +\infty, \tag{6.6}$$

or

$$uf''(u) < 0, \qquad f'''(0) \neq 0, \qquad \lim_{|u|\to+\infty} f'(u) = -\infty, \tag{6.7}$$

respectively. The functions $f(u) = u^3 + u$ and $f(u) = -u^3 - u$ are prototypes of particular interest, used later in this paper for the validation of the proposed numerical strategy.

Let $\varphi^\natural : \mathbb{R} \to \mathbb{R}$ be the unique function defined by $\varphi^\natural(0) = 0$ and for all $u \neq 0$, $\varphi^\natural(u) \neq u$ is such that the line passing through the points $(u, f(u))$ and $(\varphi^\natural(u), f(\varphi^\natural(u)))$ is tangent to the graph of f at point $(\varphi^\natural(u), f(\varphi^\natural(u)))$:

$$f'(\varphi^\natural(u)) = \frac{f(u) - f(\varphi^\natural(u))}{u - \varphi^\natural(u)}.$$

This function is smooth, monotone decreasing and onto thanks to (6.6) or (6.7). We denote by $\varphi^{-\natural} : \mathbb{R} \to \mathbb{R}$ its inverse function.

Concave-convex flux functions

Let us assume that f obeys (6.6) and let $\varphi^\flat : \mathbb{R} \to \mathbb{R}$ be a kinetic function, that is (by definition) a monotone decreasing and Lipschitz continuous mapping such that

$$\begin{cases} \varphi_0^\flat(u) < \varphi^\flat(u) \leq \varphi^\natural(u) & \text{if} \quad u > 0, \\ \varphi^\natural(u) \leq \varphi^\flat(u) < \varphi_0^\flat(u) & \text{if} \quad u < 0. \end{cases} \tag{6.8}$$

From $\varphi^\flat$, we define the function $\varphi^\sharp : \mathbb{R} \to \mathbb{R}$ such that the line passing through the points $(u, f(u))$ and $(\varphi^\flat(u), f(\varphi^\flat(u)))$ with $u \neq 0$ also cuts the graph of the flux function f at point $(\varphi^\sharp(u), f(\varphi^\sharp(u)))$ with $\varphi^\sharp(u) \neq u$ and $\varphi^\sharp(u) \neq \varphi^\flat(u)$:

$$\frac{f(u) - f(\varphi^\flat(u))}{u - \varphi^\flat(u)} = \frac{f(u) - f(\varphi^\sharp(u))}{u - \varphi^\sharp(u)}.$$

The nonclassical Riemann solver associated with (6.1)-(6.2)-(6.5)-(6.3) is given as follows.
When $u_l > 0$:

(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.

When $u_l \leq 0$:

(1) If $u_r \leq u_l$, the solution is a rarefaction wave connecting u_l to u_r.

(2) If $u_r \in [u_l, \varphi^\sharp(u_l))$, the solution is a classical shock wave connecting u_l to u_r.

(3) If $u_r \in (\varphi^\sharp(u_l), \varphi^\flat(u_l))$, 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 $u_r \geq \varphi^\flat(u_l)$, 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.

Convex-concave flux functions

We next assume that f satisfies the condition (6.7). Let $\varphi^\flat : \mathbb{R} \to \mathbb{R}$ be a kinetic function, that is, a monotone decreasing and Lipschitz continuous map such that

$$\begin{cases} \varphi_0^\flat(u) < \varphi^\flat(u) \leq \varphi^{-\natural}(u) & \text{if} \quad u < 0, \\ \varphi^{-\natural}(u) \leq \varphi^\flat(u) < \varphi_0^\flat(u) & \text{if} \quad u > 0. \end{cases} \tag{6.9}$$

We then define $\rho(u, v) \in \mathbb{R}$ if $v \neq u$ and $v \neq \varphi^\natural(u)$ by

$$\frac{f(\rho(u,v)) - f(u)}{\rho(u,v) - u} = \frac{f(v) - f(u)}{v - u}$$

with $\rho(u, v) \neq u$ and $\rho(u, v) \neq v$, and extend the function ρ by continuity otherwise. Note that $\varphi^\sharp(u) = \rho(u, \varphi^\flat(u))$ where $\varphi^\sharp$ is defined as in the case of a concave-convex flux function.
The nonclassical Riemann solver associated with (6.1)-(6.2)-(6.5)-(6.3) is given as follows.
When $u_l > 0$:

(1) If $u_r \geq u_l$, the solution is a classical shock connecting u_l to u_r.

(2) If $u_r \in [0, u_l)$, the solution is a rarefaction wave connecting u_l to u_r.

(3) If $u_r \in (\varphi^\flat(u_l), 0)$, the solution contains a rarefaction wave connecting u_l to $\varphi^{-\flat}(u_r)$, followed by a nonclassical shock connecting $\varphi^{-\flat}(u_r)$ to u_r.

(4) If $u_r \leq \varphi^\flat(u_l)$, the solution contains:

(i) a classical shock connecting u_l to $\varphi^{-\flat}(u_r)$, followed by a nonclassical shock connecting $\varphi^{-\flat}(u_r)$ to u_r, if $u_l > \rho(\varphi^{-\flat}(u_r), u_r)$.

(ii) a classical shock connecting u_l to u_r, if $u_l \leq \rho(\varphi^{-\flat}(u_r), u_r)$.

When $u_l \leq 0$:

(1) If $u_r \leq u_l$, the solution is a classical shock connecting u_l to u_r.

(2) If $u_r \in (u_l, 0]$, the solution is a rarefaction wave connecting u_l to u_r.

(3) If $u_r \in (0, \varphi^{\flat}(u_l))$, the solution contains a rarefaction wave connecting u_l to $\varphi^{-\flat}(u_r)$, followed by a non classical shock connecting $\varphi^{-\flat}(u_r)$ to u_r.

(4) If $u_r \geq \varphi^{\flat}(u_l)$, the solution contains:

(i) a classical shock connecting u_l to $\varphi^{-\flat}(u_r)$, followed by a nonclassical shock connecting $\varphi^{-\flat}(u_r)$ to u_r, if $u_l < \rho(\varphi^{-\flat}(u_r), u_r)$.

(ii) a classical shock connecting u_l to u_r, if $u_l \geq \rho(\varphi^{-\flat}(u_r), u_r)$.

Observe that the convex-concave case can in principle be deduced from the concave-convex case, by replacing f by $-f$ and x by $-x$. Nevertheless, it is useful to keep the above two descriptions in mind, since there is an important difference between the Riemann solvers: the nonclassical shock always connects u_l to $\varphi^{\flat}(u_l)$ in the concave-convex case, and $\varphi^{-\flat}(u_r)$ to u_r in the convex-concave case. The numerical method we are going to describe must take this feature into account, and as we will explain it is necessary to take into account both $\varphi^{\flat}$ and $\varphi^{-\flat}$ in the design of the scheme.

6.3 Motivations and difficulties

Notation

Our aim is to design a scheme for the numerical approximation of the nonclassical solutions to (6.1)-(6.2)-(6.5). To this end, we consider the general class of finite volume methods. Introducing constant space and time lengths Δx and Δt for the space and time discretization, we can set $x_{j+1/2} = j\Delta x$, $j \in \mathbb{Z}$, and $t^n = n\Delta t$, $n \in \mathbb{N}$. The discretization consists, at each time t^n, of a piecewise constant function $x \mapsto u_\nu(x, t^n)$ which should be an approximation of the exact solution $u(x, t^n)$ on the cell $C_j = [x_{j-1/2}; x_{j+1/2})$:

$$u_\nu(x, t^n) = u_j^n, \qquad x \in C_j,\ j \in \mathbb{Z},\ n \in \mathbb{N}.$$

Here, ν refers to the ratio $\Delta t/\Delta x$. The initial data at the time $t = 0$ is denoted by u_0 and we define the sequence $(u_j^0)_{j\in\mathbb{Z}}$:

$$u_j^0 = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} u_0(x)dx, \qquad j \in \mathbb{Z}. \tag{6.10}$$

The starting point in the conception of our algorithm is a few conventional interpretation of the constant values u_j^n, $j \in \mathbb{Z}$. As suggested by the proposed initialization (6.10), u_j^n is usually, and rightly, seen as an approximate value of the average on cell C_j of the exact solution at time t^n. Integrating equation (6.1) over the slab $C_j \times [t^n, t^{n+1}]$ and using Green's

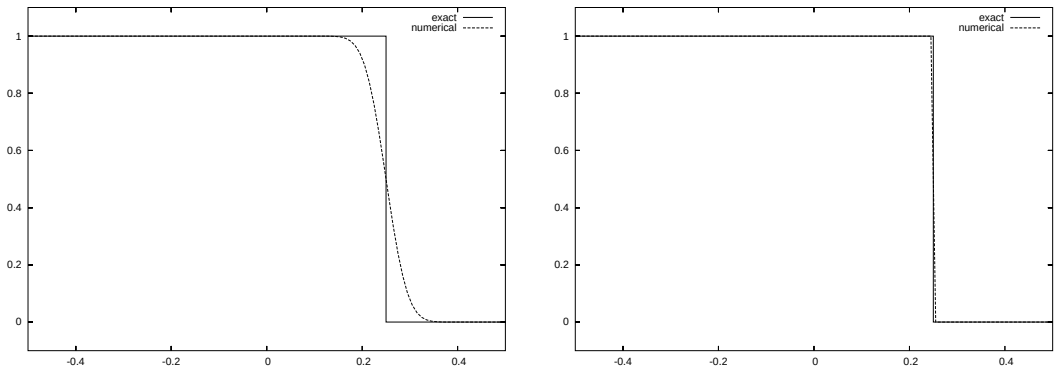
formula, it is thus natural to define $(u_j^{n+1})_j$ from $(u_j^n)_j$ and a conservative scheme of the following form

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x}(f_{j+1/2}^n - f_{j-1/2}^n), \text{ for all } j \in \mathbb{Z}, \tag{6.11}$$

where $f_{j+1/2}^n$ represents an approximate value of the flux that passes through the interface $x_{j+1/2}$ between times t^n and t^{n+1}.
Here, we shall also consider u_j^n as a given information, on cell C_j and at time t^n, on the structure of the exact Riemann solution associated with inital states $u_l = u_{j-1}^n$ and $u_r = u_{j+1}^n$ which will develop at the next times $t > t^n$. At this stage, it is quite easy to admit that the more this information will be precise (*i.e.* close to what will really happen), the more one will be in a good position to define the numerical fluxes $f_{j+1/2}^n$ and then predict the approximate values of the solution at time t^{n+1}.

Linear advection equation

As a first illustration, let us consider the linear advection with constant velocity $a > 0$, that is, the scalar conservation law with flux $f(u) = au$. In this case, the weak solution to the initial-value problem for (6.1) is unique, and is given explicitly as $u(t,x) = u_0(x-at)$. Hence, neither the entropy condition (6.2) nor the kinetic condition (6.5) are necessary. The basic scheme for approximating this solution is the so-called upwind scheme and corresponds to the choice $f_{j+1/2}^n = au_j^n$ for all $j \in \mathbb{Z}$. Remind that the CFL condition $a\Delta t/\Delta x \leq \alpha$ for a given $\alpha \leq 1$ is mandatory for the stability of the procedure. Figure 6.1a shows the corresponding numerical solution at time $t = 0.25$ for $a = 1$, $\alpha = 0.5$ and $u_l = 1$, $u_r = 0$ in (6.3). The mesh contains 100 points per unit interval. We observe that the numerical solution presents



(a) Upwind scheme. *(b) Reconstruction scheme.*

Figure 6.1: Linear advection.

a good agreement with the exact one but contains numerical diffusion. We propose the following interpretation. In some sense, the value u_j^n that we consider as an information on the Riemann solution associated with initial states $u_l = u_{j-1}^n$ and $u_r = u_{j+1}^n$ is sufficient to correctly approach this solution when defining $f_{j+1/2}^n = au_j^n$, but not enough to avoid the numerical diffusion. Note that the latter is expected but not hoped. In the present situation,

the fact is that we actually know what will happen in the future, namely a propagation of the Riemann initial states ($u_l = u^n_{j-1}$ and $u_r = u^n_{j+1}$) with speed a. In particular, no value different from u^n_{j-1} and u^n_{j+1} is created so that information given by u^n_j is clearly not optimal. In the process of calculation of the numerical flux $f^n_{j+1/2}$, we are thus tempted to add more details in cell C_j when replacing, as soon as possible, the constant state u^n_j with a discontinuity separating u^n_{j-1} on the left and u^n_{j+1} on the right, and located at point $\overline{x}_j \in C_j$. In the forthcoming developments, the left and right states of this reconstructed discontinuity will be noted $u^n_{j,l}$ and $u^n_{j,r}$, respectively. Hence, we have here

$$u^n_{j,l} = u^n_{j-1} \quad \text{and} \quad u^n_{j,r} = u^n_{j+1}. \tag{6.12}$$

See Figure 6.2 below. We claim that this provides a better information for calculating $f^n_{j+1/2}$ than the original one. Such a reconstruction is due to conserve u in order to be relevant, which defines $\overline{x}_j$ by the following constraint

$$(\overline{x}_j - x_{j-1/2})u^n_{j,l} + (x_{j+1/2} - \overline{x}_j)u^n_{j,r} = (x_{j+1/2} - x_{j-1/2})u^n_j$$

which equivalently recasts as

$$\overline{x}_j = x_{j-1/2} + \frac{u^n_{j,r} - u^n_j}{u^n_{j,r} - u^n_{j,l}}\Delta x. \tag{6.13}$$

Then, the reconstruction is possible provided we have $0 \le d^n_j \le 1$, with

$$d^n_j = \frac{u^n_{j,r} - u^n_j}{u^n_{j,r} - u^n_{j,l}}. \tag{6.14}$$

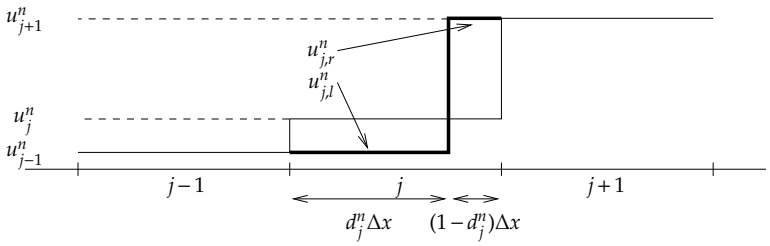


Figure 6.2: An example of discontinuous reconstruction with conservation property (the linear case).

Now, let us introduce $\Delta t_{j+1/2}$ the time needed by the reconstructed discontinuity to reach the interface $x_{j+1/2}$ (recall that $a > 0$). We clearly have

$$\Delta t_{j+1/2} = \frac{1 - d^n_j}{a}\Delta x.$$

In this case, the flux that passes through $x_{j+1/2}$ between times t^n and $t^{n+1} = t^n + \Delta t$ equals

$f(u^n_{j,r})$ until $t^n + \Delta t_{j+1/2}$, and $f(u^n_{j,l})$ after (if $\Delta t_{j+1/2} < \Delta t$). Therefore, we propose to set now

$$\Delta t f^n_{j+1/2} = \min(\Delta t_{j+1/2}, \Delta t) f(u^n_{j,r}) + \max(\Delta t - \Delta t_{j+1/2}, 0) f(u^n_{j,l}).$$

On Figure 6.1b, we have plotted the numerical solution given by this new numerical flux, leading to the so-called reconstruction scheme. The parameters of the simulation are the same than those of Figure 6.1a. We see that the more precise informations we have brought on each cell C_j for calculating the numerical fluxes make the scheme less diffusive than the original one. This strategy has first been proposed and is more detailed in [96], [97] (see also [61], [95]). In particular, it is shown that the numerical solution presented on Figure 6.1b is *exact* in the sense that u^n_j equals the average of the exact solution on C_j, that is

$$u^n_j = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} u(x, t^n) dx, \qquad j \in \mathbb{Z},\ n \in \mathbb{N}. \tag{6.15}$$

The corresponding numerical discontinuity separating u_l and u_r in then diffused on one cell at most.

Godunov scheme with a nonclassical Riemann solver

As a second illustration, let us go back to the problem (6.1)-(6.2)-(6.5) with a general concave-convex (or convex-concave) flux function f with however, for the sake of clarity,

$$f'(u) \geq 0,\ u \in \mathbb{R}. \tag{6.16}$$

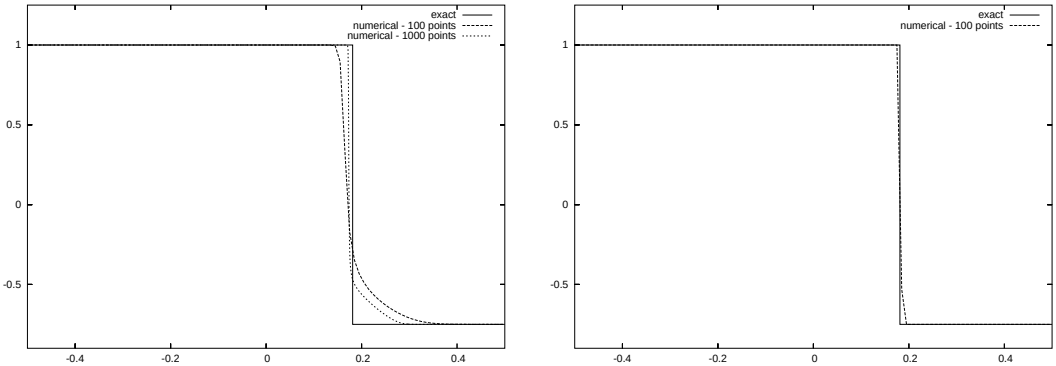
Here, we focus ourselves on a particular Riemann initial data (6.3) such that $u_r = \varphi^\flat(u_l)$. In other words, the kinetic criterion is imposed on the initial discontinuity. The exact solution then corresponds to the propagation of this discontinuity with speed $\sigma(u_l, u_r) > 0$ given by Rankine-Hugoniot relation:

$$\sigma(u_l, u_r) = \frac{f(u_r) - f(u_l)}{u_r - u_l}. \tag{6.17}$$

Figure 6.3a represents the numerical solution given by the upwind scheme $f^n_{j+1/2} = f(u^n_j)$ at time $t = 0.1$, for $f(u) = u^3 + u$ and $u_l = 1$. The kinetic function is taken to be $\varphi^\flat(u) = -0.75\,u$ so that $u_r = -0.75$.

We observe a strong disagreement between the numerical solution and the exact one. Indeed, the former is made of a (classical) shock followed by a rarefaction wave while the latter is a single (nonclassical) shock from u_l to u_r. It is then clear that the usual upwind scheme (as many others actually) is not adapted for the computation of nonclassical solutions. The next result states that the upwind scheme always converges towards the classical solution of (6.1)-(6.2). This scheme is then adapted for the computation of classical solutions only.

Property 6.1. *Assume that $u_0 \in L^\infty(\mathbb{R})$ and f is a smooth function satisfying (6.16). Then, under*



(a) Upwind scheme. *(b) Reconstruction scheme.*

Figure 6.3: Propagating nonclassical shock.

the CFL condition

$$\frac{\Delta t}{\Delta x} \max_{u\in[\min_x u_0(x),\max_x u_0(x)]} |f'(u)| \le 1,$$

the upwind conservative scheme (6.11) with $f^n_{j+1/2} = f(u^n_j)$ *converges towards the unique classical solution of (6.1)-(6.2).*

To establish this property, we only need to observe that, under the assumption (6.16) (propagation is only in one direction), the upwind scheme is equivalent to the standard Godunov scheme associated with the *classical* Riemann solver of (6.1)-(6.2) Then, standard compactness and consistency arguments apply and allow us to conclude that the scheme converges towards the unique classical solution of (6.1)(6.2).

Obviously, the above property also holds if f is assumed to be decreasing and we define $f^n_{j+1/2} = f(u^n_{j+1})$.

6.4 A conservative scheme for nonclassical entropy solutions

Preliminaries

In view of the discussion in the previous section and in order to better evaluate the numerical fluxes $f^n_{j+1/2}$, let us try to get a better information than u^n_j on cell C_j. In the present instance of an isolated propagating discontinuity, it is expected that the Riemann solution associated with initial states u^n_{j-1} and u^n_{j+1} simply propagates the initial discontinuity. This is actually true if $u^n_{j-1} = u_l$ and $u^n_{j+1} = \varphi^\flat(u_l)$, or more generally if $u^n_{j+1} = \varphi^\flat(u^n_{j-1})$. So that here again, we propose to replace the constant state u^n_j with a discontinuity separating $u^n_{j,l}$ and $u^n_{j,r}$ and located at point $\overline{x}_j$ given by (6.13), as soon as possible *i.e.* when $0 \le d^n_j \le 1$. We take

$$u^n_{j,l} = \varphi^{-\flat}(u^n_{j+1}) \quad \text{and} \quad u^n_{j,r} = \varphi^\flat(u^n_{j-1}). \tag{6.18}$$

Note that this reconstruction is equivalent to (6.12) provided that $u^n_{j-1} = u_l$ and $u^n_{j+1} = \varphi^\flat(u_l)$, or more generally $u^n_{j+1} = \varphi^\flat(u^n_{j-1})$. Then, under the assumption (6.16), we again naturally set

$$\Delta t f^n_{j+1/2} = \min(\Delta t_{j+1/2}, \Delta t) f(u^n_{j,r}) + \max(\Delta t - \Delta t_{j+1/2}, 0) f(u^n_{j,l})$$

with now

$$\Delta t_{j+1/2} = \frac{1 - d^n_j}{\sigma(u^n_{j,l}, u^n_{j,r})} \Delta x. \tag{6.19}$$

Figure 6.3b highlights the benefit of such a reconstruction. The numerical solution now fully agrees with the exact one and is moreover free of numerical diffusion (the profile is composed of a single point). We will show below that it is exact in this case, in the sense that (6.15) is still valid as in the linear case.

The scheme

On the basis of the above motivations and illustrations, we follow the description of our algorithm by considering the general situation. Assuming as given a sequence $(u^n_j)_{j\in\mathbb{Z}}$ at time t^n, it is thus a question of defining its evolution towards the next time level t^{n+1}. More precisely, and in the context of a finite volume conservative scheme, we have to define the numerical fluxes $(f^n_{j+1/2})_{j\in\mathbb{Z}}$ coming in (6.11). For that, we still assume

$$\text{either } f'(u) \geq 0, \text{ for all } u \quad \text{or } f'(u) \leq 0, \text{ for all } u, \tag{6.20}$$

so that propagation makes in one direction only. According to the previous section, information in cell C_j is understood as an element of the inner structure of the Riemann problem associated with initial states u^n_{j-1} and u^n_{j+1}. This one will be used to compute either $f^n_{j+1/2}$ (if $f'(u) \geq 0$) or $f^n_{j-1/2}$ (if $f'(u) \leq 0$).

In section 6.2, it is stated that the Riemann problem associated with initial states u^n_{j-1} and u^n_{j+1} may contain a nonclassical shock between u^n_{j-1} and $\varphi^\flat(u^n_{j-1})$ if the function is concave-convex (and between $\varphi^{-\flat}(u^n_{j+1})$ and u^n_{j+1} if the function is convex-concave). Mind that these nonclassical waves are extremely difficult to properly capture from a numerical point of view and then deserve a particular attention (we have shown in the previous section that as many others, the upwind scheme does not suit). Instead of considering u^n_j as a sufficiently accurate information for the structure of the Riemann solution associated with initial states u^n_{j-1} and u^n_{j+1}, we propose to replace it (as soon as possible) with a discontinuity separating $u^n_{j,l} = \varphi^{-\flat}(u^n_{j+1})$ on the left and $u^n_{j,r} = \varphi^\flat(u^n_{j-1})$ on the right, and located at point $\overline{x}_j \in C_j$. In other words, we propose to introduce in the cell C_j the right (respectively left) state $\varphi^\flat(u^n_{j-1})$ (respectively $\varphi^{-\flat}(u^n_{j+1})$) of the nonclassical discontinuity which is expected to be present in the Riemann solution associated with u^n_{j-1} and u^n_{j+1} (depending on if f obeys (6.6) or (6.7)). As in the previous section, one requires the reconstructed discontinuity to satisfy the conservation property (6.13) and to be located inside C_j, that is $0 \leq d^n_j \leq 1$ with d^n_j given in (6.14). See Figure 6.4 for an illustration. Here, we let $u^n_{j,l} = u^n_{j,r} = u^n_j$ if d^n_j given in (6.14) does not belong to $[0,1]$.

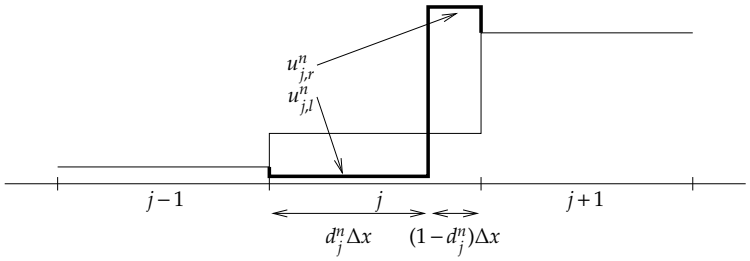


Figure 6.4: A general discontinuous reconstruction with conservation property (the general case).

Then, we naturally set for all $j \in \mathbb{Z}$:
(*i*) if f is non-decreasing

$$\Delta t f^n_{j+1/2} = \begin{cases} \min(\Delta t_{j+1/2}, \Delta t) f(u^n_{j,r}) + \max(\Delta t - \Delta t_{j+1/2}, 0) f(u^n_{j,l}) & \text{if} \quad 0 \le d^n_j \le 1, \\ \Delta t f(u^n_j) & \text{otherwise,} \end{cases} \tag{6.21}$$

with

$$\Delta t_{j+1/2} = \frac{1 - d^n_j}{\sigma(u^n_{j,l}, u^n_{j,r})} \Delta x. \tag{6.22}$$

(*ii*) if f is non-increasing:

$$\Delta t f^n_{j-1/2} = \begin{cases} \min(\Delta t_{j-1/2}, \Delta t) f(u^n_{j,l}) + \max(\Delta t - \Delta t_{j-1/2}, 0) f(u^n_{j,r}) & \text{if} \quad 0 \le d^n_j \le 1, \\ \Delta t f(u^n_j) & \text{otherwise,} \end{cases} \tag{6.23}$$

with

$$\Delta t_{j-1/2} = \frac{d^n_j}{-\sigma(u^n_{j,l}, u^n_{j,r})} \Delta x. \tag{6.24}$$

Note that contrary to the linear advection case (see the first illustration in the previous section), the local time step $\Delta t_{j+1/2}$ (respectively $\Delta t_{j-1/2}$) given by (6.22) (respectively (6.24)) is now only a prediction of the time needed by the reconstructed discontinuity to reach the interface $x_{j+1/2}$ (respectively $x_{j-1/2}$). The prediction is however exact in the case of an isolated nonclassical discontinuity (see the second illustration in the previous section) and more generally as soon as u^n_{j-1} and u^n_{j+1} verify $u^n_{j+1} = \varphi^\flat(u^n_{j-1})$.

Observe that the proposed scheme belongs to the class of 5-point schemes, since u^{n+1}_j depends on u^n_{j-2}, u^n_{j-1}, u^n_j, u^n_{j+1} and u^n_{j+2}.

Stability and consistency properties

We now state and prove important properties enjoyed by our algorithm.

Theorem 6.2. *Assume that the flux f is a smooth function satisfying the monotonicity condition (6.20) and either the concave-convex or concave-convex conditions (6.6) or (6.7) respectively. Then,*

under the CFL restriction

$$\frac{\Delta t}{\Delta x}\max_u |f'(u)| \le 1, \tag{6.25}$$

where the maximum is taken over all the u under consideration, the conservative scheme (6.11) with $f^n_{j+1/2}$ defined for all $j \in \mathbb{Z}$ by (6.21)-(6.23) is consistent with (6.1)-(6.2)-(6.5) in the following sense:
(i) Flux consistency: Assume that $u := u^n_{j-1} = u^n_j = u^n_{j+1}$, then $f^n_{j+1/2} = f(u)$ if $f' \ge 0$ and $f^n_{j-1/2} = f(u)$ if $f' \le 0$.
(ii) Classical solution (remaining in the region of convexity - or concavity - of f): Let us assume that u^n_{j-2}, u^n_{j-1}, u^n_j, u^n_{j+1} and u^n_{j+2} belong to the same region of convexity of f. Then the definition u^{n+1}_j given by the conservative scheme (6.11)-(6.21)-(6.23) coincides with the one given by the usual upwind conservative scheme. Then it obeys all the usual stability properties provided by this scheme. In particular, the strategy is convergent if the whole discrete solution belongs to the same region of convexity of f.
(iii) Isolated nonclassical shock: Let u_l and u_r be two initial states such that $u_r = \varphi^{\flat}(u_l)$. Assume that $u^0_j = u_l$ if $j \le 0$ and $u^0_j = u_r$ if $j \ge 1$. Then the conservative scheme (6.11)-(6.21)-(6.23) provides an exact numerical solution on each cell C_j in the sense that

$$u^n_j = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} u(x,t^n)dx, \qquad j \in \mathbb{Z},\ n \in \mathbb{N}, \tag{6.26}$$

where u denotes the exact Riemann solution of (6.1)-(6.2)-(6.3)-(6.5) given by $u(x,t) = u_l$ if $x < \sigma(u_l,u_r)t$ and $u(x,t) = u_r$ otherwise, and is convergent towards u. In particular, the numerical discontinuity is diffused on one cell at most.

Let us briefly comment this result. Property (*i*) shows that the proposed numerical flux function is consistent in the classical sense of finite volume methods. Properties (*ii*) and (*iii*) are more note-worthy and can be seen as two important stability/accuracy properties. Indeed, they state that the method is actually *convergent* if the solution either remains in the same region of convexity of f ((*ii*)), or more importantly consists in an isolated nonclassical discontinuity satisfying the prescribed kinetic relation ((*iii*)). Up to our knowledge, not one of the *conservative* schemes already existing in the literature verifies the latter property. To the authors mind, this explains the very good numerical results obtained in the next section.

Proof.
(i) If $u := u^n_{j-1} = u^n_j = u^n_{j+1}$ then

$$d^n_j = \frac{\varphi^{\flat}(u) - u}{\varphi^{\flat}(u) - \varphi^{-\flat}(u)}.$$

The property $0 \le d^n_j \le 1$ means $\min(\varphi^{-\flat}(u),\varphi^{\flat}(u)) \le u \le \max(\varphi^{-\flat}(u),\varphi^{\flat}(u))$ and cannot hold, since u and $\varphi^{\flat}(u)$ do not have the same sign for all u. Then we get $f^n_{j+1/2} = f(u)$ if $f' \ge 0$ and $f^n_{j-1/2} = f(u)$ if $f' \le 0$ by (6.21)-(6.23).
(ii) Arguments are quite similar. Assume without restriction that $f' \ge 0$ and recall that $0 \le d^n_{j-1} \le 1$ and $0 \le d^n_j \le 1$ respectively means that $\min(\varphi^{-\flat}(u^n_j),\varphi^{\flat}(u^n_{j-2})) \le u^n_{j-1} \le \max(\varphi^{-\flat}(u^n_j),\varphi^{\flat}(u^n_{j-2}))$ and $\min(\varphi^{-\flat}(u^n_{j+1}),\varphi^{\flat}(u^n_{j-1})) \le u^n_j \le \max(\varphi^{-\flat}(u^n_{j+1}),\varphi^{\flat}(u^n_{j-1}))$. These inequalities are not valid since by definition u and $\varphi^{\flat}(u)$ do not belong to the same region of convexity of f. By (6.21)-

(6.23), the numerical fluxes $f^n_{j\pm1/2}$ coincides with the usual upwind fluxes and the conclusion follows.

(iii) First, notice that there is no relevant reconstruction in the first iteration. Indeed, the property $0 \le d^n_j \le 1$ reads as follows if $j<0$ or $j>1$,

$$0 \le d^n_j \le 1 \iff \begin{cases} \min(\varphi^{-\flat}(u_l),\varphi^{\flat}(u_l)) \le u_l \le \max(\varphi^{-\flat}(u_l),\varphi^{\flat}(u_l)) & \text{if} \quad j<0, \\ \min(\varphi^{-\flat}(u_r),\varphi^{\flat}(u_r)) \le u_r \le \max(\varphi^{-\flat}(u_r),\varphi^{\flat}(u_r)) & \text{if} \quad j>1, \end{cases}$$

which again cannot hold (see (*i*) below), while if $j=0$ or $j=1$, the relation $u_r=\varphi^{\flat}(u_l)$ and definition (6.14) give

$$\begin{cases} d^n_j = \dfrac{u_r-u_l}{u_r-u_l} = 1 & \text{if} \quad j=0, \\ d^n_j = \dfrac{u_r-u_r}{u_r-u_l} = 0 & \text{if} \quad j=1, \end{cases}$$

so that the reconstructions exist but are trivial: $u_l=\varphi^{-\flat}(u_r)$ (respectively $u_r=\varphi^{\flat}(u_l)$) takes the whole cell associated with $j=0$ (respectively $j=1$).

Assume now without restriction that f is non-decreasing and let Δt be such that (6.25) holds. After one time step Δt, the exact solution given by $u(x,\Delta t)=u_l$ if $x<\sigma(u_l,u_r)\Delta t$ and $u(x,\Delta t)=u_r$ otherwise is such that

$$\frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} u(x,\Delta t)dx = \begin{cases} u_l & \text{if} \quad j\le 0, \\ u_r-\sigma(u_l,u_r)\frac{\Delta t}{\Delta x}(u_r-u_l) & \text{if} \quad j=1, \\ u_r & \text{if} \quad j>1. \end{cases} \tag{6.27}$$

But recall that $\sigma(u_l,u_r)$ is given by (6.17) so that we have

$$\frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} u(x,\Delta t)dx = \begin{cases} u_l-\frac{\Delta t}{\Delta x}(f(u_l)-f(u_l)) & \text{if} \quad j\le 0, \\ u_r-\frac{\Delta t}{\Delta x}(f(u_r)-f(u_l)) & \text{if} \quad j=1, \\ u_r-\frac{\Delta t}{\Delta x}(f(u_r)-f(u_r)) & \text{if} \quad j>1, \end{cases} \tag{6.28}$$

that is

$$u^1_j = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} u(x,\Delta t)dx, \qquad j\in\mathbb{Z}. \tag{6.29}$$

Equality (6.26) is then proved for the first iterate.

What happens now in the next time iteration ? At this stage, it is first clear (see the previous discussion just below) that only cell C_1 is going to be dealt with a reconstruction. Now, the main point of the proof lies in the fact that the reconstructed discontinuity in this cell actually joins the expected states $\varphi^{-\flat}(u^1_2)=\varphi^{-\flat}(u_r)=u_l$ and $\varphi^{\flat}(u^1_0)=\varphi^{\flat}(u_l)=u_r$ and is located exactly at point $x=\sigma(u_l,u_r)\Delta t$ by the conservation property (6.29). In other words, we have reconstructed the exact solution at time $t=\Delta t$. To get the required identity (6.26) for the second iterate, it is sufficient to recall that by Green's formula the conservative scheme (6.11) with $f^n_{j+1/2}$ defined for all $j\in\mathbb{Z}$ by (6.21)-(6.23) is equivalent for $n=2$ to average the evolution of this exact solution up to time $t^2=2\Delta t$. And the process is going on in a similar way for the next time iterations, which proves the result.

6.5 Numerical experiments

We consider $f(u) = u^3 + u$ thus f is concave-convex in the sense given in the second section. Concerning the entropy-entropy flux pair (U,F) used in (6.2), we set

$$U(u) = u^2, \quad F(u) = \frac{3}{2}u^4 + u^2.$$

Easy calculations give explicit definitions for $\varphi^\natural$ and $\varphi^{-\natural}$. We find $\varphi^\natural = -\frac{u}{2}$ and then $\varphi^{-\natural} = -2u$, $\varphi_0^\flat(u) = -u$. Concerning the definition of $\varphi^\sharp$, we have in this case $\varphi^\sharp(u) = -u - \varphi^\flat(u)$.

The choice of the kinetic function $\varphi^\flat$ must be in agreement with relations (6.8) with $\varphi^\natural$ and $\varphi_0^\flat$ just calculated. For instance, we set without restriction

$$\varphi^\flat(u) = -\beta u, \quad \beta \in [0.5, 1).$$

This kinetic relation can be realized by an augmented model based on nonlinear diffusion and dispersion terms; see [26]. In the following, we choose $\beta = 0.75$.

Test A: we want to insist numerically on the property (iii) of Theorem 6.2 that concerns the exact capture of any isolated nonclassical shocks. Thus let consider the following nonclassical shock as Riemann initial condition

$$u_0(x) = \begin{cases} 4, & x < 0, \\ \varphi^\flat(4) = -3, & x > 0, \end{cases}$$

The numerical solution (figure 6.5) is exact everywhere, except in the single cell containing the nonclassical shock (we exceptionally use a constant piecewise mode here to represent numerical solutions so that the concerned cell is easily recognizable). However the value in this cell coincides with the corresponding exact solution average (see (6.26)) and then allows after reconstruction to find the exact position of the discontinuity, thanks to the conservativity of scheme. And that is why the solution stays so precise when the time goes.

Test B: this test corresponds to the Riemann problem with initial data

$$u_0(x) = \begin{cases} 4, & x < 0, \\ -5, & x > 0, \end{cases}$$

which solution is a nonclassical shock followed by a rarefaction wave. The computations of Figure 6.6a are performed successively with $\Delta x = 0.01$ and $\Delta x = 0.002$. The nonclassical shock that appears is, as previously, localised in only one cell.
The Figure 6.6b represents the logarithm of the L1-norm between the exact and the numerical solution versus the logarithm of Δx. The numerical order of convergence is here of about 0.8374.

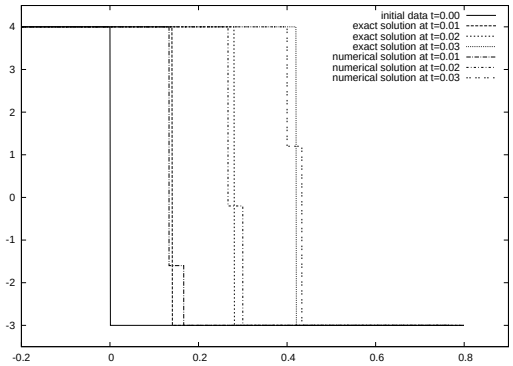
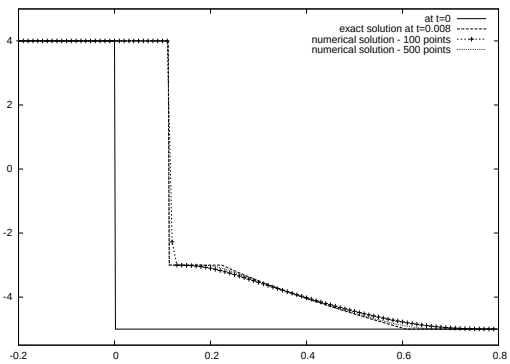
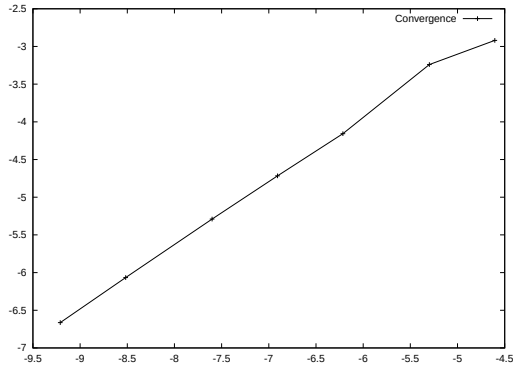


Figure 6.5: Test A: Nonclassical shock; 30 points.



(a) Nonclassical shock and rarefaction.



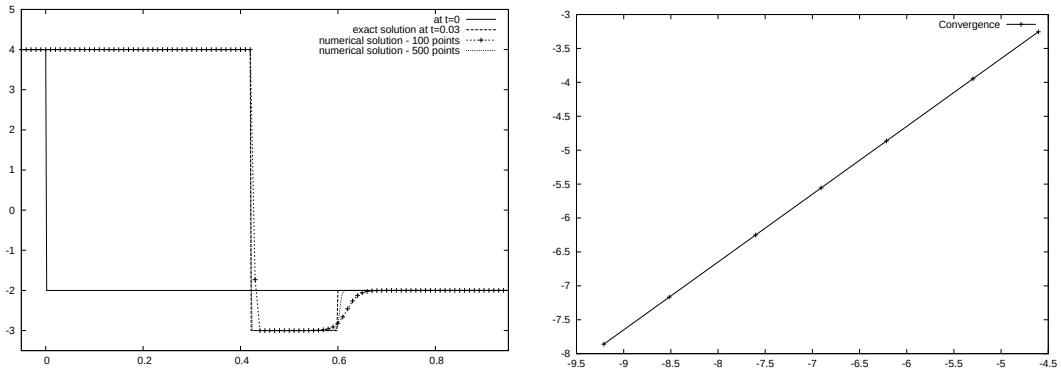
(b) L1 convergence ($log(E_{L^1})$ versus $log(\Delta x)$).

Figure 6.6: Test B.

Test C: (Figure 6.7) we choose another Riemann initial condition that develops a nonclassical shock followed by a classical shock.

$$u_0(x) = \begin{cases} 4, & x < 0, \\ -2, & x > 0. \end{cases}$$

We have the same observation as previously concerning the nonclassical shock computation that is well captured and arises in a tenuous spatial domain, however note that the classical shock diffuses: in fact our scheme is exactly the Godunov scheme in domains of same convexity of the flux f.
Once again, the L1-error figure ensures numerical convergence with numerical order of about 0.9999.



(a) Nonclassical and classical shocks. *(b) L1 convergence ($log(E_{L^1})$ versus $log(\Delta x)$).*

Figure 6.7: Test C.

Test D: (Figure 6.8) we take an initial data composed with two nonclassical shocks that interact.

$$u_0(x) = \begin{cases} 4 = \varphi^{-\flat}(-3), & x < 0.1 \\ -3, & 0.1 < x < 0.2 \\ 2.25 = \varphi^{\flat}(-3), & x > 0.2. \end{cases}$$

The computation is performed with $\Delta x = 0.05$ and plotted at four successive times $t = 0, 0.010, 0.017$ and 0.020. We observe that the two nonclassical shocks cancel each other, and generate a single classical shock, in accordance with theoretical results presented in [102].

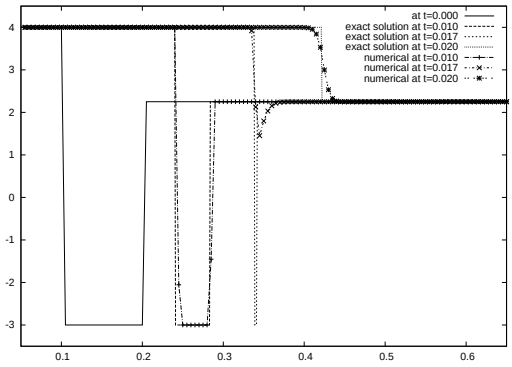


Figure 6.8: Test D: Interaction of two nonclassical shocks.

Test E: (Figure 6.9) let consider a periodic initial condition

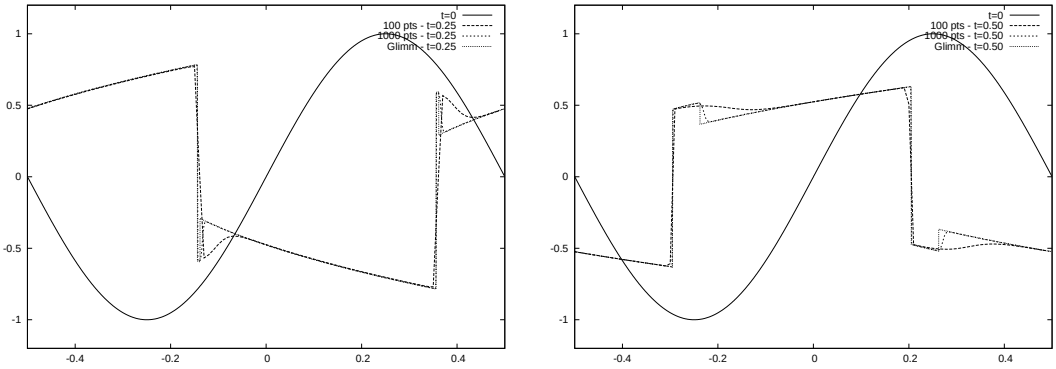
$$u_0(x) = \sin\left(\frac{x}{2\pi}\right),$$

with periodic boundary conditons $u(-0.5, t) = u(0.5, t)$. The exact solution is not known so we compare the reconstruction scheme solution with the numerical solution obtained by computing through Glimm random choice scheme (see [70]) with the nonclassical solver

described in Section 6.2 and using the van der Corput random sequence (a_n) defined by

$$a_n = \sum_{k=0}^{m} i_k 2^{-(k+1)},$$

where $n = \sum_{k=0}^{m} i_k 2^k$, $i_k \in \{0,1\}$, denotes the binary expansion of the integer n. The figure 6.9 represents the solutions at times $t = 0, 0.25$ and 0.5 for our scheme with $\Delta x = 0.01$, with $\Delta x = 0.0001$ and for the Glimm scheme with $\Delta x = 0.0001$ as a reference. Both methods seem to have the same good agreement. Roughly speaking, increasing parts of u_0 rarefact while decreasing parts compresse and develop in a first time classical shocks and then, when left and right states of those shocks change sign, nonclassical shocks (satisfying the expected kinetic relation) and new faster classical shocks on the right appear.



(a) Solutions at $t = 0.25$. *(b) Solutions at $t = 0.50$.*

Figure 6.9: Test E: Periodic initial data; reconstruction scheme vs Glimm scheme.

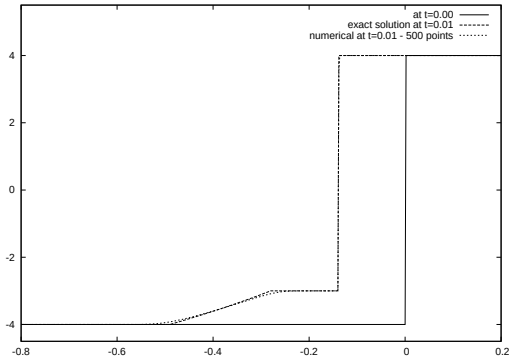
Test F: (Figure 6.10) as an illustration of the case of a convex-concave flux function, we compute two Riemann solutions with opposite flux $f(u) = -u^3 - u$ (so $f' < 0$ and the solutions move from right to left) and the same kinetic function $\varphi^{\flat}(u) = -0.75\, u$: the first one (Figure 6.10a) corresponds to an initial data

$$u_0(x) = \begin{cases} -4, & x < 0, \\ 4, & x > 0, \end{cases}$$

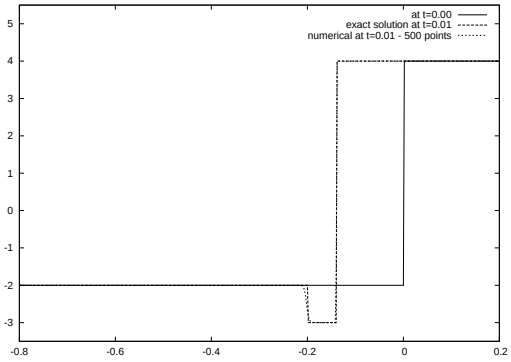
and develops a rarefaction wave and a nonclassical shock; the second one (Figure 6.10b) corresponds to the initial data

$$u_0(x) = \begin{cases} -2, & x < 0, \\ 4, & x > 0, \end{cases}$$

and the corresponding solution is a classical shock followed by a nonclassical shock.



(a) Rarefaction and nonclassical shock.



(b) Classical and nonclassical shocks.

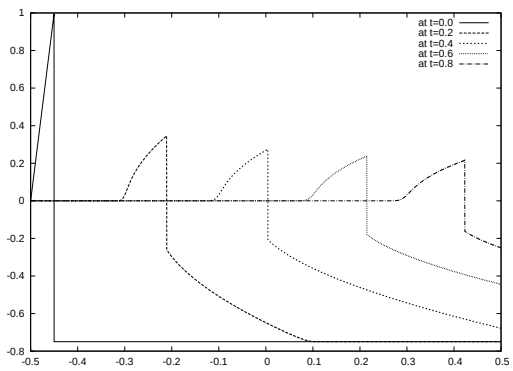
Figure 6.10: Test F - Two examples for the convex-concave case.

Test G: We now study how the kinetic relation $u_R = \varphi^\flat(u_L)$ is numerical performed through nonclassical shocks. On figure 6.11b, we plot points whose X-coordinates (respectively Y-coordinates) corresponds to the left (resp. right) traces around the reconstructed cell. The initial data allows to cover a large range of values:

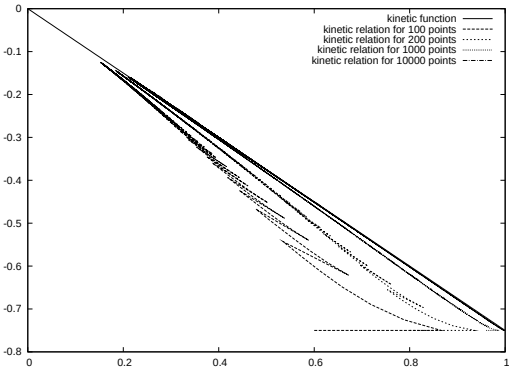
$$u_0(x) = \begin{cases} 0, & x < -0.5, \\ 1 + 20(x + 0.45), & -0.5 < x < -0.45, \\ -0.75, & x > -0.45. \end{cases}$$

The figure 6.11a represents the solution at different times with $\Delta x = 0.0002$.

We clearly observe the convergence of the numerical kinetic relation towards the prescribed one.



(a) Numerical solution.



(b) Numerical kinetic relation.

Figure 6.11: Test G.

Test H: We now study how the kinetic relation

In the course of designing the scheme proposed in the previous section we tried several variants. We report here one such scheme that is very similar to the proposed scheme, but

which *does not* converge to exact nonclassical solutions. This is due to the fact that small oscillations are generated in the scheme which "compete" with the dissipation mechanisms described by the prescribed kinetic function.
The variant is designed for the concave-convex flux $f(u) = u^3 + u$. The only difference with the scheme developed above is that it performs the reconstruction in C_j with $u^n_{j,l} = u^n_{j-1}$ (instead of $u^n_{j,l} = \varphi^{-\flat}(u^n_{j+1})$ and $u_{j,r} = \varphi^{\flat}(u^n_{j-1})$. This is equivalent in the case of a *pure* nonclassical shock (Test B) but different in the general case.
The Figure 6.12 presents the solution obtained for the same initial value as in Test E. Oscillations are generated because the reconstruction is not constrained enough in this version of the scheme. It is noticeable this behaviour persists with a change of CFL number.

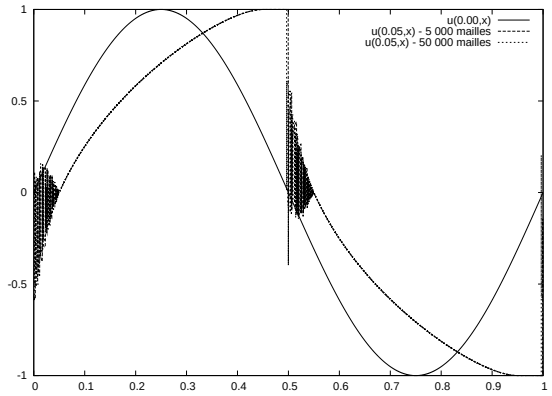


Figure 6.12: Test H: Another version of the scheme.

6.6 Concluding remarks

In this paper we have introduced a new numerical strategy for computing nonclassical solutions to nonlinear hyperbolic conservation laws. The method is based on a reconstruction technique performed in each computational cell which may exhibit a nonclassical shock. Importantly, the whole algorithm is *conservative* and propagates any admissible nonclassical discontinuity exactly. The convergence of the proposed method was demonstrated numerically for several test cases. This new approach brings a new perspective on the numerical approximation of nonclassical shocks and kinetic functions. The efficiency of the method is clearly demonstrated in the present paper. Among several open questions of interest we can mention the derivation of total variation bounds for the proposed scheme and generalization of our technique to hyperbolic systems of conservation laws, application to real materials undergoing phase transitions, as well as extension to higher-order schemes.

7 Extension au cas d'un flux concave-convexe non-monotone

Dans ce chapitre, nous proposons une version alternative au schéma mis au point dans le chapitre précédent, adaptée au cadre du calcul de solutions non classiques de lois de conservation scalaires pour un flux concave-convexe, l'hypothèse de monotonie étant levée.

7.1 Introduction

Dans le chapitre précédent, nous avons introduit un schéma conservatif pour le calcul des solutions non classiques de la loi de conservation scalaire

$$\partial_t u + \partial_x f(u) = 0. \tag{7.1}$$

Ce schéma prenait la forme conservative

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x}(f_{j+1/2}^n - f_{j-1/2}^n), \quad j \in \mathbb{Z}, \tag{7.2}$$

où la définition des flux $f_{j+1/2}^n$ était basée sur une étape de reconstruction de la solution numérique, assurant la consistance de la résolution avec la relation cinétique imposée aux changements de phase. Ce schéma était néanmoins adapté au cas spécifique $f'(u) > 0$, excluant de ce fait les flux de la forme $f(u) = u^3 - u$ qui présentent une zone de décroissance au voisinage du point d'inflexion. Nous allons ici nous concentrer sur un schéma analogue qui sera en mesure de traiter de tels flux.

7.2 Schéma numérique avec reconstructions délocalisées

7.2.1 Discontinuité non classique et états reconstruits

La première étape de la reconstruction consiste à substituer dans les cellules C_j et C_{j+1}, les valeurs respectives u_j, u_{j+1}, par une discontinuité connectant deux nouveaux états $u_{j+1/2,l}, u_{j+1/2,r}$. Cette reconstruction sera réalisée seulement si un changement de phase a lieu localement

$$u_j^n u_{j+1}^n < 0, \tag{7.3}$$

et sous d'autres contraintes ultérieurement précisées. La discontinuité sera positionnée en un point $\bar{x}_{j+1/2} = x_{j+1/2} + d_{j+1/2}\Delta x$ tel que représenté sur la figure 7.1, et afin de rendre cette étape de reconstruction conservative, on impose à la position $d_{j+1/2}$ de satisfaire la condition

$$u_j + u_{j+1} = (1 + d_{j+1/2})u_{j+1/2,l} + (1 - d_{j+1/2})u_{j+1/2,r}. \tag{7.4}$$

La reconstruction n'aura pas lieu si éventuellement $\bar{x}_{j+1/2}$ n'est pas dans l'ensemble $C_j \cup C_{j+1}$, c'est-à-dire si la condition

$$d_{j+1/2} \in (-1, 1), \tag{7.5}$$

n'est pas réalisée.

Remarque 7.1. La relation (7.4) semble assurer le caractère conservatif de l'étape de reconstruction, mais en fait un autre choix n'affecterait pas le caractère conservatif du schéma : l'écriture (7.2) suffit à l'assurer. Cette relation permettra en réalité d'assurer le transport exact des discontinuités non classiques, ceci via la définition des flux à venir, et constitue plutôt à ce titre une condition de consistance.

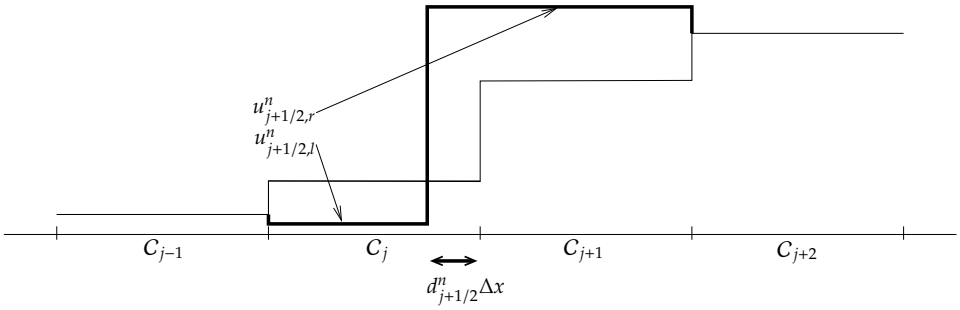


Figure 7.1: *Principe de reconstruction*

Pour récupérer la *consistance du schéma avec la fonction cinétique* employée dans le solveur non classique en considération, la discontinuité introduite vérifiera

$$u_{j+1/2,r} = \varphi^\flat(u_{j+1/2,l}). \tag{7.6}$$

Dans l'exemple mis en œuvre cette relation sera satisfaite au travers du choix suivant

$$u_{j+1/2,l} = u_{j-1}, \qquad u_{j+1/2,r} = \varphi^\flat(u_{j-1}). \tag{7.7}$$

L'assymétrie dans ce choix est motivée par le caractère sous-compressif des discontinuités non classiques, qui a pour conséquence que les caractéristiques entrent dans la discontinuité non classique à gauche, et en sortent à droite. L'information circule donc de la gauche vers la droite, quand bien même le choc non classique se déplace lui-même vers la gauche. En quelque sorte, le caractère monotone ou non-monotone du flux en considération n'a pas d'importance dans la définition des solutions non classiques, c'est son caractère concave-convexe qui joue un rôle dans le choix de la forme $u_+ = \varphi^\flat(u_-)$ pour la relation cinétique et du choix précédent pour l'étape de reconstruction du schéma numérique.

Remarque 7.2. Nous avons vu dans le chapitre précédent qu'un tel choix pour les états de reconstruction ne donnait pas de bons résultats pour une reconstruction par maille. Ici, en localisant les reconstructions au voisinage des interfaces, ce choix, nous le verrons, conviendra.

Remarque 7.3. Dans le cas d'un flux convexe-concave, pour lequel la fonction cinétique s'emploie sous la forme $u_- = \varphi^\flat(u_+)$, il serait plus naturel de favoriser l'état u_{j+2} afin de définir les états reconstruits, alors obtenus comme

$$u_{j+1/2,l} = \varphi^\flat(u_{j+2}), \qquad u_{j+1/2,r} = u_{j+2}. \tag{7.8}$$

7.2.2 Définition des flux

Nous raisonnons pour le moment sur une reconstruction isolée, dont nous supposerons qu'elle concerne la double-maille $C_j \cup C_{j+1}$, telle que représentée sur la figure 7.1. Soit σ la vitesse de la discontinuité non classique ainsi introduite, qui vérifie la condition de saut de

Rankine-Hugoniot

$$\sigma = \frac{f(u_{j+1/2,r}) - f(u_{j+1/2,l})}{u_{j+1/2,r} - u_{j+1/2,l}}. \tag{7.9}$$

Nous allons définir les flux numériques $f_{j-1/2}$, $f_{j+1/2}$ et $f_{j+3/2}$ respectivement aux interfaces $x_{j-1/2}$, $x_{j+1/2}$ et $x_{j+3/2}$, selon le comportement de la discontinuité non classique introduite. Notons auparavant $\Delta_l t_{j+1/2}$, $\delta t_{j+1/2}$, et $\Delta_r t_{j+1/2}$ les intervalles de temps nécessaires à cette discontinuité reconstruite pour atteindre, seule, respectivement l'interface gauche $x_{j-1/2}$, l'interface médiane $x_{j+1/2}$ et l'interface droite $x_{j+3/2}$ du patch de cellules. Ces quantités ne sont pas signées (elles sont éventuellement même infinies si $\sigma = 0$)

$$\begin{aligned} \delta t_{j+1/2} &= -d_{j+1/2}\Delta x/\sigma, \\ \Delta_r t_{j+1/2} &= (1 - d_{j+1/2})\Delta x/\sigma, \\ \Delta_l t_{j+1/2} &= -(1 + d_{j+1/2})\Delta x/\sigma. \end{aligned} \tag{7.10}$$

La condition de stabilité CFL suivante est imposée :

$$\frac{\Delta t}{\Delta x} \max_j \left(\max_{u \in I_j \cup N_j} |f'(u)| \right) \leq \frac{1}{2}, \tag{7.11}$$

où $I_j = [u_j, u_{j+1}]$ si $u_j \leq u_{j+1}$ (ou $[u_{j+1}, u_j]$ sinon) et $N_j = [u_j, \varphi^\flat(u_j)]$ si $u_j \leq \varphi^\flat(u_j)$ (ou $[\varphi^\flat(u_j), u_j]$ sinon).

Le flux $f_{j+1/2}$ est alors calculé en supposant que le choc non classique $(u_{j+1/2,l}, u_{j+1/2,r})$ évolue sans aucune interaction pendant un premier sous-pas de temps ($\Delta_r t_{j+1/2}$ ou $\Delta_l t_{j+1/2}$) correspondant à sa propagation dans $C_j \cup C_{j+1}$, puis après ce sous-pas de temps, s'il vient à quitter cet intervalle, en supposant que le choc a quitté ce domaine sans y avoir engendré aucune autre onde. En d'autres termes, on néglige toute éventuelle interaction du choc non classique avec des ondes provenant des points voisins $x_{j-1/2}$ et $x_{j+3/2}$. Naturellement, les quantités temporelles $\Delta_l t_{j+1/2}$, $\delta t_{j+1/2}$, et $\Delta_r t_{j+1/2}$ peuvent être non seulement négatives comme mentionné plus tôt mais même supérieures au pas de temps Δt. Le principe consistant à reconstruire certaines ondes et à en tenir compte de manière détaillée dans la définition des flux ne nous autorise pas à contraindre le pas de temps de calcul selon ces intervalles de temps intermédiaires. Les formules de flux s'écriront sous une forme faisant intervenir l'opérateur de projection $\pi := \pi(\alpha)$ de $\mathbb{R}$ dans $[0,1]$ défini par

$$\pi(\alpha) = \begin{cases} 0, & \alpha \leq 0, \\ \alpha, & 0 \leq \alpha \leq 1, \\ 1, & 1 \leq \alpha. \end{cases}$$

On distingue alors deux cas :

- Si $\sigma < 0$, alors on définit (voir Figure 7.2b)

$$\begin{aligned} f_{j-1/2} &= \pi\left(\frac{\Delta_l t_{j+1/2}}{\Delta t}\right) G(u_{j-1}, u_{j+1/2,l}) + \left(1 - \pi\left(\frac{\Delta_l t_{j+1/2}}{\Delta t}\right)\right) f(u_{j+1/2,r}), \\ f_{j+1/2} &= \pi\left(\frac{\delta t_{j+1/2}}{\Delta t}\right) f(u_{j+1/2,l}) + \left(1 - \pi\left(\frac{\delta t_{j+1/2}}{\Delta t}\right)\right) f(u_{j+1/2,r}), \\ f_{j+3/2} &= G(u_{j+1/2,r}, u_{j+2}). \end{aligned} \tag{7.12}$$

- Si $\sigma > 0$, alors on définit (voir Figure 7.2a)

$$\begin{aligned} f_{j-1/2} &= G(u_{j-1}, u_{j+1/2,l}), \\ f_{j+1/2} &= \pi\left(\frac{\delta t_{j+1/2}}{\Delta t}\right) f(u_{j+1/2,r}) + \left(1 - \pi\left(\frac{\delta t_{j+1/2}}{\Delta t}\right)\right) f(u_{j+1/2,l}), \\ f_{j+3/2} &= \pi\left(\frac{\Delta_r t_{j+1/2}}{\Delta t}\right) G(u_{j+1/2,r}, u_{j+2}) + \left(1 - \pi\left(\frac{\Delta_r t_{j+1/2}}{\Delta t}\right)\right) f(u_{j+1/2,l}). \end{aligned} \tag{7.13}$$

7.2.3 Séparation des reconstructions

La définition des flux précédemment introduite supposait que la reconstruction était isolée, de manière à bien définir un unique flux par interface. Nous allons préciser à présent par quels choix on se ramène à ce cas. Considérons le patch des 4 cellules voisines centré en $x_{j+1/2}$, que l'on notera dans la suite

$$O_{j+1/2} = C_{j-1} \cup C_j \cup C_{j+1} \cup C_{j+2}.$$

Soit m le nombre de changements de phase de la solution numérique u sur $O_{j+1/2}$, obtenu comme

$$m = \mathrm{Card}\{j-1 \le k \le j+1, u_k u_{k+1} < 0\}.$$

- Si $m = 0$, alors la situation est considérée comme classique en $x_{j+1/2}$ et le flux $f_{j+1/2}$ est simplement un flux numérique de notre choix, pour un solveur classique :

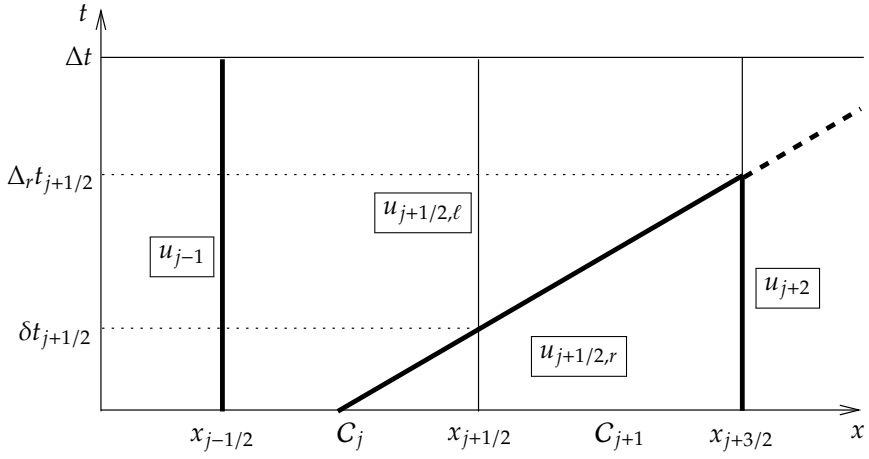
$$f_{j+1/2} = G(u_j, u_{j+1}),$$

  où G est un flux numérique classique à deux points consistant avec le flux exact f, que nous spécifierons plus tard pour les exemples.

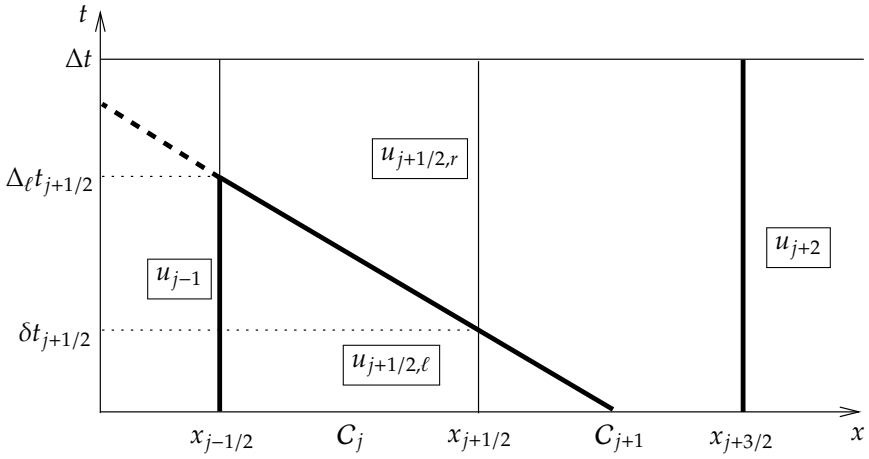
- Si $m = 1$, nous réaliserons donc une seule reconstruction au voisinage de $x_{j+1/2}$, soit dans la double-maille $C_{j-1} \cup C_j$, soit dans $C_j \cup C_{j+1}$, soit enfin dans $C_{j+1} \cup C_{j+2}$. Le flux $f_{j+1/2}$ sera donc défini selon le comportement de cette discontinuité non classique comme l'un des flux définis dans (7.12)-(7.13).

- Si enfin $m \ge 2$, alors nous utiliserons en $x_{j+1/2}$ le flux classique G et négligerons ainsi les interactions possibles entre les ondes non classiques présentes dans ce voisinage :

$$f_{j+1/2} = G(u_j, u_{j+1}).$$



(a) Définition des flux si $\sigma > 0$.



(b) Définition des flux si $\sigma < 0$.

Figure 7.2: *Méthodologie pour la définition des flux à partir de la solution recontruite.*

Par le principe précédent, chaque interface $x_{j+1/2}$ n'est ainsi associée qu'à une reconstruction au plus.

Remarque 7.4. On pourrait penser que, par ce choix de séparation des reconstruction, le schéma ne sera pas en mesure de gérer correctement les interactions entre discontinuités non classiques. Les tests numériques ultérieurs mettront cette idée à l'épreuve et nous constaterons que ces interactions seront correctement restituées.

7.3 Tests numériques pour un flux concave-convexe non-monotone

Dans les tests qui suivent, nous considérons la fonction flux $f(u) := u^3 - u$ et la fonction cinétique $\varphi^\flat(u) = -\beta u$ avec pour coefficient $\beta = 0.75$. Précisons de plus que le flux numérique

classique employé sera le flux d'Engquist-Osher :

$$G(u,v) = \int_u^v \left|f'(w)\right| dw. \tag{7.14}$$

Rappelons que pour une fonction flux telle que celle choisie, le flux numérique d'Osher se réduit à la définition suivante :

$$G(u,v) = \begin{cases} f(u), & f'(u) > 0 \ \& \ f'(v) > 0, \\ f(v), & f'(u) < 0 \ \& \ f'(v) < 0, \\ f(u_*), & f'(u) < 0 \ \& \ v < u_*, \\ f(u^*), & f'(u) < 0 \ \& \ v > u^*, \\ f(u) + f(v) - f(u_*), & u < u_* \ \& \ f'(v) < 0, \\ f(u) + f(v) - f(u^*), & u > u^* \ \& \ f'(v) < 0, \end{cases}$$

où $u_* = -1/\sqrt{3}$ et $u^* = 1/\sqrt{3}$ sont les deux points soniques de f.

Pour condition initiale, dans ces premiers tests, nous choisirons une donnée de type Riemann (u_l, u_r), à savoir

$$u_0(x) = \begin{cases} u_l, & x < 0, \\ u_r, & x \geq 0. \end{cases}$$

Quant aux conditions aux limites elles seront du type Neumann à chacun des deux bords du domaine.

7.3.1 Discontinuités non classiques pures

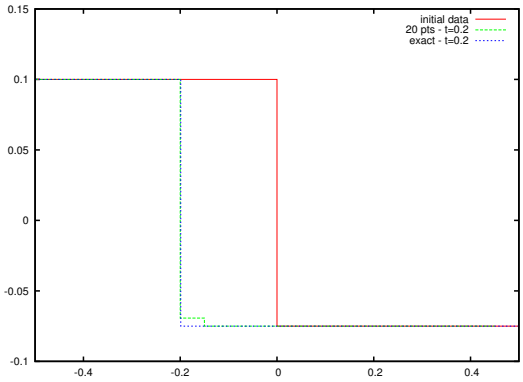
Ces trois premiers tests concernent des discontinuités non classiques isolées, avec donc des données de Riemann de la forme $(u_l, \varphi^\flat(u_l))$ qui seront respectivement les suivantes $(0.1, -0.075)$, $(2.0, -1.5)$ en enfin $(4/\sqrt{13}, -3/\sqrt{13}) \simeq (1.109, -0.832)$. Les solutions exactes pour chacune de ces données initiales sont respectivement des discontinuités non classiques de vitesse négative strictement, positive strictement et nulle. Dans chacun des cas, nous effectuons les calculs avec seulement 20 mailles sur $[-0.5, 0.5]$. La solution obtenue est exactement la projection sur ce maillage de la solution exacte (Figure 7.3), c'est-à-dire :

$$u_j^n = \frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} u(x, t^n) dx, \qquad j \in \mathbb{Z}, n \in \mathbb{N}, \tag{7.15}$$

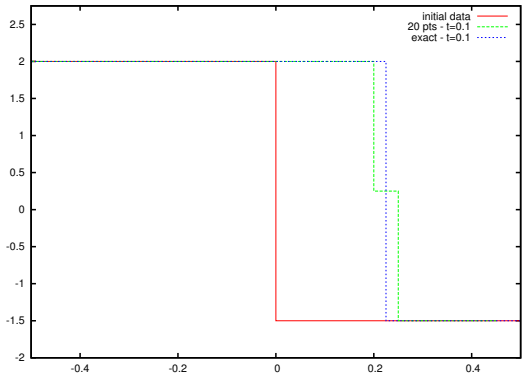
où u est la solution exacte. En particulier, la discontinuité n'est donc diffusée que sur une maille au plus.

7.3.2 Problèmes de Riemann

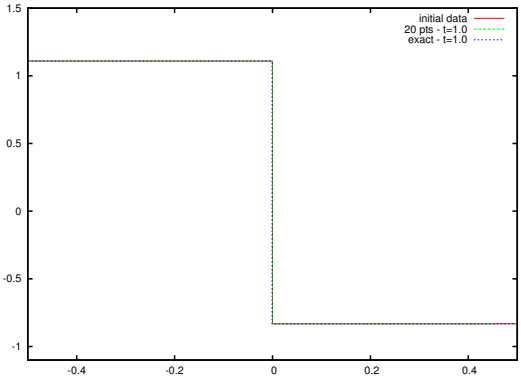
Nous considérons à présent des données de Riemann plus générales qui couvriront toutes les structures envisageables pour la solution exacte. La solution numérique est de nouveau calculée sur l'intervalle $[-0.5, 0.5]$, avec successivement 100 et 1000 points de calcul.



(a) Choc NC de vitesse négative



(b) Choc NC de vitesse positive



(c) Choc NC stationnaire

Figure 7.3: *Calcul de chocs non classiques (NC) isolés*

La figure 7.5 réunit l'ensemble des résultats obtenus. Ils concordent bien avec la solution exacte et leur convergence est illustrée sur la figure 7.4 sur laquelle est tracée en échelle logarithmique l'erreur L^1 relativement à la solution exacte, en fonction du pas du maillage Δx.

Test N_+R **:** $(u_l, u_r) = (1.5, -1.5)$. La solution est constituée d'un choc non classique entre les état 1.5 et $\varphi^\flat(1.5)$, suivi d'une raréfaction entre $\varphi^\flat(1.5)$ et -1.5.

Test N_+S **:** $(u_l, u_r) = (1.5, -0.6)$. La solution est constituée d'un choc non classique entre les états 1.5 et $\varphi^\flat(1.5)$, suivi d'un choc compressif entre $\varphi^\flat(1.5)$ et -0.6.

Test S_+ **:** $(u_l, u_r) = (1.5, -0.2)$. La solution est constituée exclusivement d'un choc classique (compressif) entre les états 1.5 et -0.2.

Test N_-R **:** $(u_l, u_r) = (0.5, -1.2)$. La solution est constituée d'un choc non classique entre les états 0.5 et $\varphi^\flat(0.5)$, de vitesse négative, suivi d'une raréfaction entre les états $\varphi^\flat(0.5)$ et -1.2.

Test N_-S **:** $(u_l, u_r) = (0.5, -0.2)$. La solution est constituée d'un choc non classique entre les états 0.5 et $\varphi^\flat(0.5)$, de vitesse négative, suivi d'un choc classique (compressif) entre les états $\varphi^\flat(0.5)$ et -0.2.

Test S_- **:** $(u_l, u_r) = (0.5, -0.1)$. La solution est constituée d'un unique choc classique (compressif) séparant les états 0.5 et -0.1.

Observons en particulier que dans les tests S_+ et S_-, la solution est entièrement classique, ceci en raison de l'inégalité $(u_r - \varphi^\sharp(u_l))(u_r - u_l) \leq 0$: même s'il y a changement de phase $u_l u_r < 0$, aucun choc non classique ne se développe. On observe dans ces deux cas que la solution numérique présente une couche limite au voisinage direct du choc. Elle est due à ce que les reconstructions sont effectuées dès qu'un changement de phase est présent, sans souci de l'apparition effective d'un choc non classique. Cependant cette couche limite n'a pas d'effet néfaste sur la convergence numérique car elle reste bornée et localisée sur quelques mailles seulement.

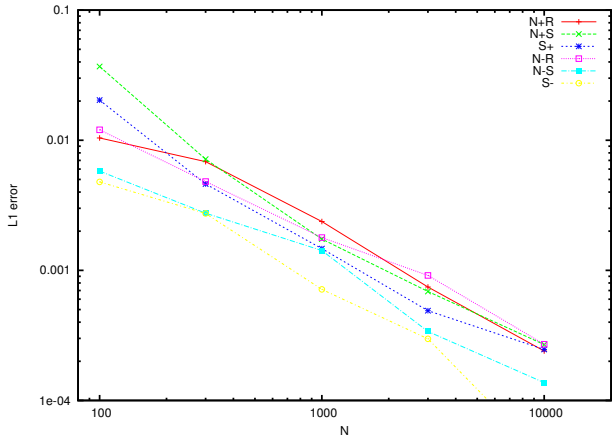
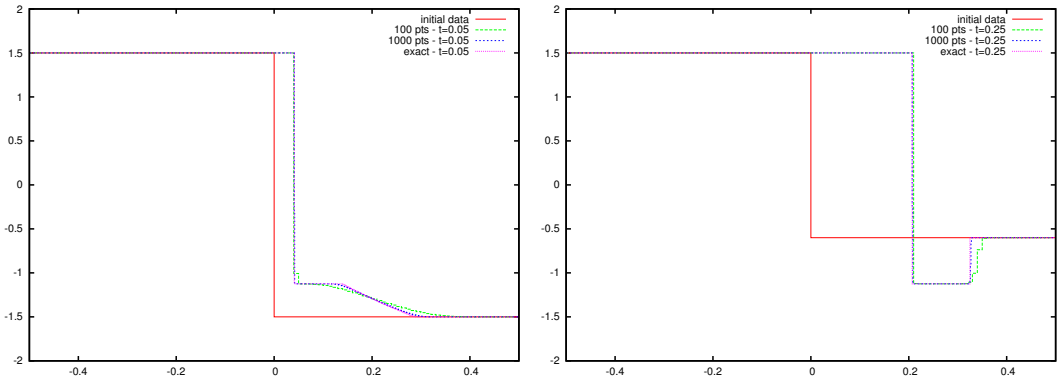
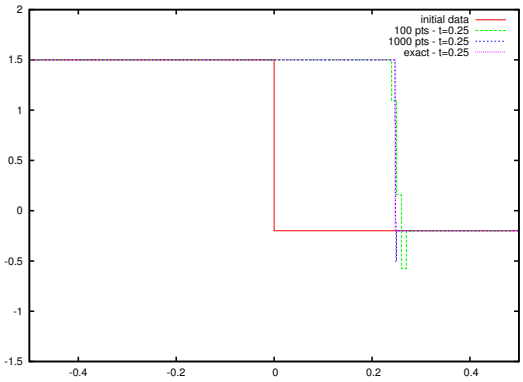


Figure 7.4: *Convergence en norme* L^1 *(tracé de* $\log(E_{L^1})$ *en fonction de* $\log(N)$*).*

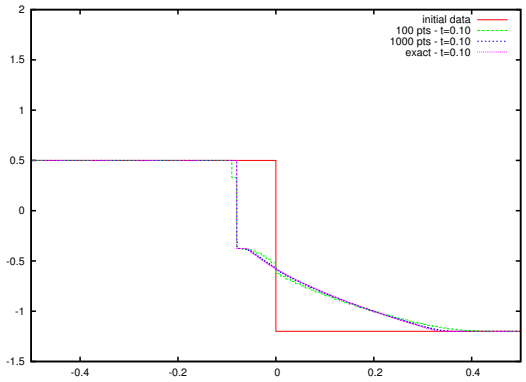


(a) Test N_+R

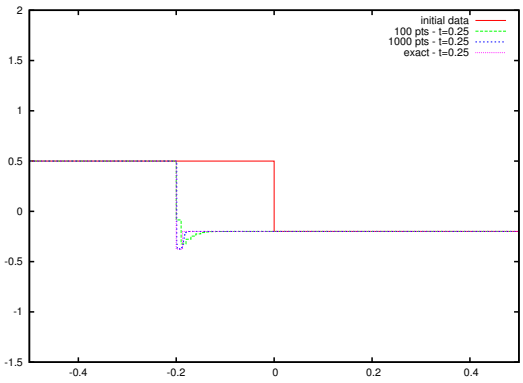
(b) Test N_+S



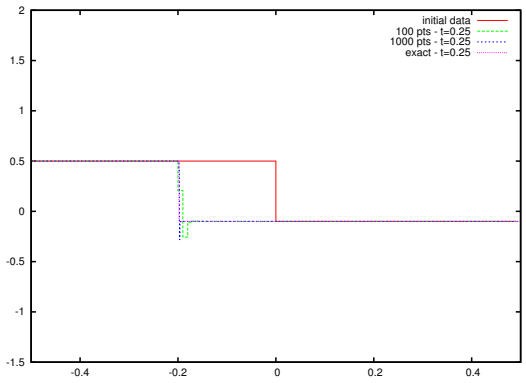
(c) Test S_+



(d) Test N_-R



(e) Test N_-S



(f) Test S_-

Figure 7.5: *Problèmes de Riemann non classiques.*

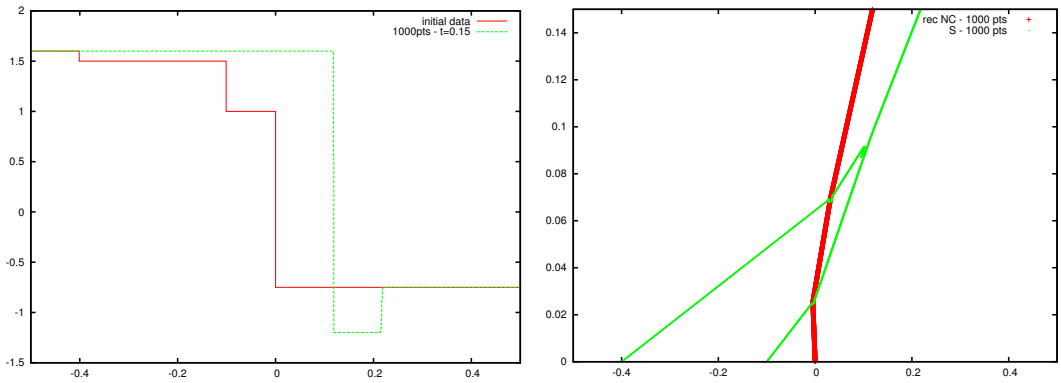
7.3.3 Problèmes de Cauchy et interaction d'ondes

Considérons à présent des problèmes plus généraux, de manière à éprouver le schéma dans des situations dans lesquelles des ondes viennent à interagir.

Test $S-N$ **:**
Ce test concerne ainsi l'interaction de chocs classiques avec un choc non classique. La donnée initiale est ainsi choisie constante par morceaux, prenant 4 valeurs successives.

$$u_0(x) = \begin{cases} 1.6 & x < -0.4, \\ 1.5 & -0.4 \le x < -0.1, \\ 1.0 & -0.1 \le x < 0, \\ -0.75 & x \ge 0. \end{cases}$$

La figure 7.6a représente la solution numérique pour 1000 points de calcul, à l'instant initial puis à l'instant $t = 0.15$. La figure 7.6b représente quant à elle, dans le plan (x,t) la position des discontinuités, distinguant les discontinuités classiques en trait fin, et les discontinuités non classiques en trait plus épais.



(a) Solution à l'instant $t = 0.15$ *(b) Structure dans le plan (x,t)*

Figure 7.6: *Interaction de chocs classiques et non classique.*

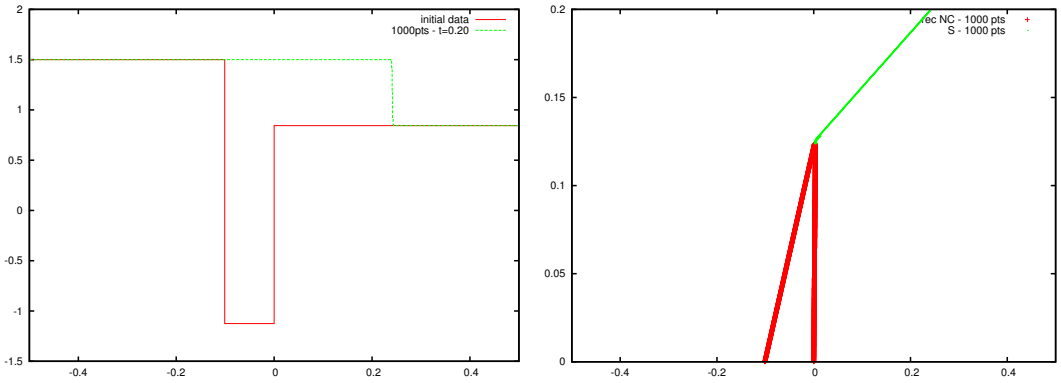
La position du choc non classique est simplement obtenue en repérant les mailles reconstruites. En revanche pour les discontinuités classiques, la position est obtenue numériquement en repérant dans le profil numérique de u, à chaque pas de temps, les mailles qui sont sujettes à un changement de convexité de u (c'est-à-dire les points d'inflexions x_j pour lesquels la quantité $u_{j-1} - 2u_j + u_{j+1}$ change de signe). Cette méthode permet pour des solutions de structure suffisamment simple de localiser rapidement les ondes de type choc.

On observe bien le caractère sous-compressif de l'onde non classique, qui laisse passer les chocs classiques vers la droite. Les états sont alors modifiés de manière à vérifier la relation cinétique. En revanche les chocs classiques, de caractère compressif, se combinent tout deux pour engendrer un unique choc classique.

Test $N-N$:
Ce test concerne l'interaction entre deux chocs non classiques. La donnée initiale est la suivante :

$$u_0(x) = \begin{cases} 1.5, & x < -0.1, \\ \varphi^\flat(1.5) = -1.125, & -0.1 \le x < 0, \\ \varphi^\flat(\varphi^\flat(1.5)) = 0.844, & x \ge 0. \end{cases}$$

La figure 7.7a représente la solution obtenue pour 1000 points de calculs à l'instant initial, puis à l'instant $t = 0.20$. La figure 7.7b représente, comme dans le test précédent la position des discontinuités dans le plan (x,t). Souvenons-nous que la méthode de reconstruction proposée est conCcue sans tenir compte des possibles interactions entre ondes non classiques au moment de la définition des flux. Ce test permet de s'assurer que cette caractéristique ne met pas en défaut la résolution d'un pareil cas. On observe bien une annihilation des deux chocs non classiques qui engendre un choc classique.



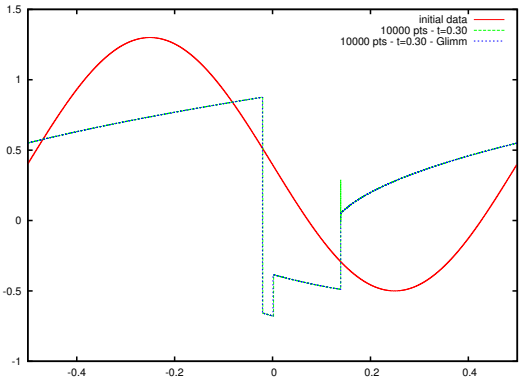
(a) Solution à l'instant $t = 0.20$ *(b) Structure dans le plan (x,t)*

Figure 7.7: *Interaction de deux chocs non classiques isolés.*

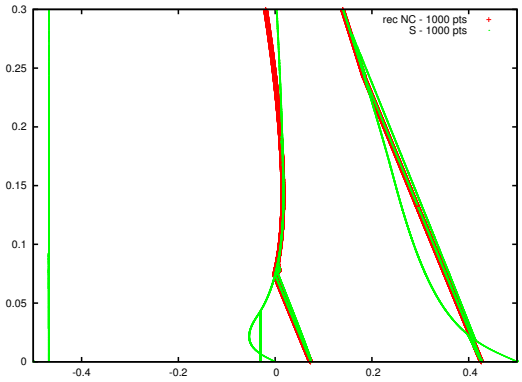
Test périodique :
À présent nous considérons un problème de Cauchy périodique, la donnée initiale est choisie sous la forme

$$u_0(x) = 0.4 + 0.9\sin(2\pi x), \qquad x \in [-0.5, 0.5],$$

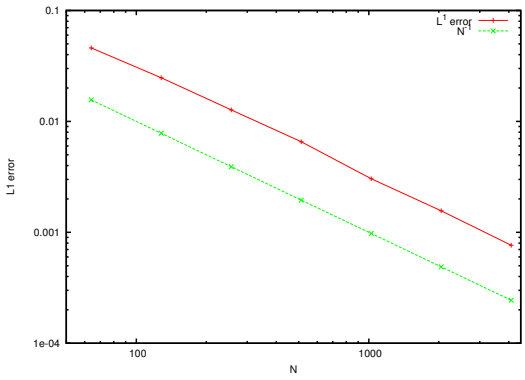
avec des conditions de bord périodiques $u(-0.5,t) = u(0.5,t),\ t \geq 0$. Le problème est résolu avec successivement 1000 et 10000 points de calculs (figure 7.8a). Une solution de référence est utilisée, qui nous est fournie par un schéma de Glimm, utilisant le solveur non classique exact avec la relation cinétique convenable. Nous renvoyons le lecteur au chapitre précédent pour des informations sur la méthode employée. Rappelons que la méthode de Glimm pour les solveurs non classiques en scalaire a été étudiée par Chalons et LeFloch dans [48] et la convergence en a été illustrée numériquement. Nous utilisons pour sa mise en œuvre la suit de van der Corput déjà mentionnée dans le chapitre précédent. Les deux solutions ainsi constituées concordent, comme l'atteste le diagramme de convergence présenté en figure 7.8c, représentant l'erreur en norme L^1 de la solution calculée à l'aide de notre schéma par rapport à une solution de référence obtenue par le schéma de Glimm pour $N = 2^{15}$ points, à l'instant final $t = 0.30$. La convergence numérique observée est d'ordre 1.



(a) Solution à l'instant $t = 0.30$



(b) Structure dans le plan (x,t)



(c) Convergence d'ordre 1 en norme L^1.

Figure 7.8: *Problème de Cauchy périodique avec des chocs non classiques.*

7.4 Conclusions et perspectives

7.4.1 Bilan positif

Dans les chapitre 6 et 7, nous avons proposé une approche numérique pour le calcul de solutions non classiques de lois de conservation scalaires. Rappelons-en les propriétés principales.

- Le schéma est conservatif et prend la forme

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

- Les flux numériques $f_{j+1/2}^n$ sont des flux classiques, choisis par l'utilisateur, dans les domaines où la solution numérique ne change pas de phase, et sont déduits de l'étape de reconstruction au niveau des changements de phase.
- Les discontinuités non classiques isolées sont parfaitement calculées.
- Le schéma est en mesure de restituer la relation cinétique prescrite pour des problèmes plus généraux, faisant éventuellement intervenir des interactions entre ondes de différents types (classiques et/ou non classiques).

7.4.2 Perspectives

La convergence de ces schémas, observée numériquement pourrait être étayée par un *résultat théorique*, les différents arguments habituels (estimation en variation totale, inégalités d'entropie numérique) sont délicats à mettre en œuvre au travers de l'étape de reconstruction. Une relecture du schéma en terme de schéma-équilibre (cf. Bouchut [28]) pourrait être une piste pour y parvenir.

Les heuristiques qui fondent ces schémas pourraient être réalisées par des choix différents de ceux qui ont été considérés. En ce qui concerne les "degrés de liberté", on peut relever par exemple le *choix des états de reconstruction*. Dans le chapitre 6, le choix était $u_{j,l}^n = \varphi^{-\flat}(u_{j+1}^n)$ et $u_{j,r}^n = \varphi^{\flat}(u_{j-1}^n)$, ne satisfaisant pas au sens propre la relation cinétique, mais restituant cependant correctement une discontinuité non classique exacte ($u_{j+1} = \varphi^{\flat}(u_{j-1}^n)$). Nous avons vu que le choix $u_{j,l}^n = u_{j-1}^n$ et $u_{j,r}^n = \varphi^{\flat}(u_{j-1}^n)$, pourtant également consistant avec la relation cinétique ne permettait pas de calculer les solutions. Au contraire, dans la version du chapitre 7, les états reconstruits ont été choisis précisément comme $u_{j+1/2,l}^n = u_{j-1}^n$ et $u_{j+1/2,r}^n = \varphi^{\flat}(u_{j-1}^n)$ satisfaisant alors la relation cinétique . Nous avons observé que ce choix convenait alors. Ces différentes observations méritent un éclaircissement que nous n'avons pas été en mesure de fournir à ce jour.

De la même manière, il y a un degré de liberté dans le *choix de définition des flux au voisinage d'une reconstruction*. Différents choix reviennent à négliger différentes interactions d'ondes entre les problèmes de Riemann classiques et le transport de l'onde non classique

notamment pour le calcul des flux. Il est probable que certains choix ne garantissent pas la stabilité du schéma, et conduisent éventuellement à une convergence vers une solution non classique différente.

7.4.3 Cas des systèmes de lois de conservation

Dans la série de papiers de LeFloch et Thanh [108; 109; 110], les auteurs étudient un modèle hyperbolique non-linéaire décrivant les transitions de phases, concernant l'elastodynamique non-linéaire

$$\begin{aligned} &\partial_t v - \partial_x \sigma(w) = 0, \\ &\partial_t w - \partial_x v = 0, \end{aligned} \tag{7.16}$$

ou le type p-système décrivant les fluides de Van des Waals

$$\begin{aligned} &\partial_t u + \partial_x p(v) = 0, \\ &\partial_t v - \partial_x u = 0. \end{aligned} \tag{7.17}$$

La loi d'état (σ ou p) est considérée de la forme concave-convexe telle qu'elle apparaît de manière standard dans ces situations et le modèle est additionné d'une relation cinétique pour caractériser les chocs sous-compressifs. L'étude proposée précédemment dans ce cadre scalaire prétend proposer une approche numérique pour parvenir au calcul des solutions non classiques pour ces systèmes. Des tests prospectifs ont été effectués sur la base du schéma de reconstruction proposé en scalaire, dans le cas de l'elastodynamique non-linéaire. Dans ce cadre système, les ondes non classiques sont de deux types, il y a en effet une famille de discontinuités non classiques par champ caractéristique et autant de relations cinétiques à prescrire. Les premiers tests mis en œuvre ont pour le moment échoué et soulignent la nécessité d'une révision du cas scalaire.

Annexes

TABLE DES FIGURES

Liste des travaux & Communications

Voici la liste des publications acceptées ou soumises dont il est question dans l'ensemble de ce manuscrit.

Articles :

❶ *Existence result for the coupling problem of two scalar conservation laws with Riemann initial data*,
Boutin B., Chalons C., and Raviart P.-A.,
Soumis. Repris dans le Chapitre 1 p.23.

❷ *Coupling nonlinear hyperbolic equations (I). Selfsimilar viscous approximation for thin interfaces*,
Boutin B., Coquel F., and LeFloch P.G.,
En préparation. Repris dans le Chapitre 2 p.65.

❸ *Coupling nonlinear hyperbolic equations (II). Resonant interfaces with internal structure*,
Boutin B., Coquel F., and LeFloch P.G.,
En préparation. Repris dans le Chapitre 3 p.113.

❹ *Coupling nonlinear hyperbolic equations (III). A regularization method via thick interfaces*,
Boutin B., Coquel F., and LeFloch P.G.,
En préparation. Repris dans le Chapitre 4 p.129.

❺ *Coupling nonlinear hyperbolic equations (IV). A multidimensional finite volume framework*,
Boutin B., Coquel F., and LeFloch P.G.,
En préparation. Repris dans le Chapitre 5 p.155.

❻ *Convergent and conservative schemes for nonclassical solutions based on kinetic relations. I*,
Boutin B., Chalons C.,Lagoutière F., and LeFloch P.G.,
Interf. and Free Bound., 10(3):399-421, (2008).
Repris dans le Chapitre 6 p.207.

Actes de congrès :

① *Existence result for the coupling problem of two scalar conservation laws with Riemann initial data*,
Boutin B., Chalons C., and Raviart P.-A.,
Proceedings of ENUMATH 2009.

② *Fluid-particles flows: a thin spray model with energy exchanges*,
Boudin L., Boutin B., Fornet B., Goudon T., Lafitte P., Lagoutière F., and Merlet B.,
ESAIM: Proceedings of CEMRACS 2008, (2009).
Inclus à la suite de ce manuscrit p.257.

③ *Dafermos regularization for interface coupling of conservation laws*,
Boutin B., Coquel F., and Godlewski E.,
Proceedings of HYP 2006, Eleventh International Conference on Hyperbolic Problems, *Hyperbolic problems. Theory, numerics and applications, Springer*, (2008).
Inclus à la suite de ce manuscrit p.273.

④ *Coupling two scalar conservation laws via Dafermos' self-similar regularization*,
Ambroso A., Boutin B., Coquel F., Godlewski E., and LeFloch P.G.,
Proceedings of ENUMATH 2007, 7th European Conference on Numerical Mathematics and Advanced Applications, *Numerical Mathematics and Advanced Applications, Springer*, (2008).

⑤ *A sharp interface and fully conservative scheme for computing nonclassical shocks*,
Boutin B., Chalons C., Lagoutière F., and LeFloch P.G.,
Proceedings of ENUMATH 2007, 7th European Conference on Numerical Mathematics and Advanced Applications, *Numerical Mathematics and Advanced Applications, Springer*, (2008).

Communications

mai 2009 Smai 2009, la Colle sur Loup, France

Couplage interfacial épais de deux lois de conservations scalaires

août 2008 Cemracs 2008, Luminy, France

Exposé de compte-rendu du projet Spraynergy

juin 2008 Hyp 2008, University Maryland, USA

Convergent and conservative schemes for nonclassical solutions based on kinetic relations

mai 2008 Canum 2008, Saint-Jean de Monts, France

Approximation autosemblable du problème de Riemann. Systèmes hyperboliques non-conservatifs avec résonance

sept. 2007 Enumath 2007, Karl-Franzens-Universität Graz, Österreich

The coupling of two scalar conservation laws by a Dafermos regularization

ANNEXES : PROCEEDINGS

Dans les pages qui suivent sont repris deux actes de congrès.

Fluid-particles flows: a thin spray model with energy exchanges,
BOUDIN L., BOUTIN B., FORNET B., GOUDON T., LAFITTE P., LAGOUTIÈRE F., and MERLET B., *ESAIM: Proceedings of CEMRACS 2008*, (2009).

Cet acte de congrès (pp.257–270) capitalise les recherches menées durant la session 2008 du CEMRACS 2008. Elles concernent l'élaboration d'un schéma préservant l'asymptotique (AP) d'un régime d'écoulement décrivant l'interaction entre une phase fluide (modèle fluide : équations d'Euler) et une phase particulaire (modèle cinétique de type Vlasov), prenant en compte les échanges d'énergies entre ces deux phases. Le proceedings présente les modèles en considération et un schéma dédié au modèle limite. Le schéma AP, quant à lui, doit permettre de se détacher de l'asymptotique exacte ici calculée, tout en la restituant correctement. Il fera l'objet de recherches futures.

Dafermos regularization for interface coupling of conservation laws,
BOUTIN B., COQUEL F., and GODLEWSKI E.,
Proceedings of HYP 2006, Eleventh International Conference on Hyperbolic Problems, *Hyperbolic problems. Theory, numerics and applications, Springer*, (2008).

Cet acte de congrès (pp.273–278) marque le point de départ de l'approche développée tout au long de la Première Partie de ce mémoire, à savoir l'utilisation d'un système EDP augmenté pour modéliser le couplage de deux équations hyperboliques, ceci via l'introduction d'une fonction couleur. Nous avons alors vu quelle richesse cette vision des choses apporte tant du point de vue mathématique que numérique.

ESAIM: PROCEEDINGS, Vol. ?, 2009, 1-10
Editors: Will be set by the publisher

FLUID-PARTICLES FLOWS: A THIN SPRAY MODEL WITH ENERGY EXCHANGES

Laurent Boudin[1, 2], Benjamin Boutin[1, 3], Bruno Fornet[4], Thierry Goudon[5, 6], Pauline Lafitte[5,6], Frédéric Lagoutière[5, 7] and Benoît Merlet[8]

Abstract. This paper is devoted to an asymptotic analysis of a fluid-particles coupled model, in the *bubbling* regime. On the theoretical point of view, we extend the analysis done in [4] for the case of an isentropic gas to the case of an ideal gas, thus adding the internal energy, or temperature, which is unknown. We formally derive the bubbling limit system in the same way as in [4] and propose a numerical scheme to solve this limit system.

The numerical resolution of the non-limit system, and the numerical analysis of the asymptotic properties of the scheme (e.g. the asymptotic preserving property), as performed in [4], is at study.

Résumé. Nous proposons ici une analyse asymptotique formelle d'un modèle de couplage entre une densité de particules et un fluide, dans la limite dite *bubbling*. Cette analyse est effectuée en suivant les pas de [4] où le fluide considéré est isentropique tandis qu'il est ici un gaz parfait (où donc l'énergie interne, ou la température, est une inconnue supplémentaire). Nous identifions le système limite et proposons un algorithme pour le résoudre de manière approchée.

La suite de ce travail, en cours, concerne l'écriture d'un algorithme de résolution du système non limite, et l'étude des propriétés asymptotiques dudit schéma.

1. Introduction

We are interested in a PDE system describing the interaction between a fluid and a set of droplets immersed in the fluid. This situation occurs in combustion theory [15], motivated for instance by the design of engines or propulsors [11]. We also mention the dynamics of sprays with many applications e.g. biomedical sprays [2], dispersion of pollutants [13], the optimization of fine water spray fire suppression systems... The fluid is described by the evolution of its density $\rho(t,x) \geq 0$, its velocity $u(t,x) \in \mathbb{R}^N$ ($N = 1$, 2 or 3) and its total energy $E(t,x) \geq 0$, which are functions of time $t \geq 0$ and position $x \in \mathbb{R}^N$. We define the internal

[1] UPMC Paris 06, Lab. J.-L. Lions, 175 rue du Chevaleret, BC 187, F-75013 Paris, France; e-mail: laurent.boudin@upmc.fr & boutin@ann.jussieu.fr

[2] INRIA Paris-Rocquencourt, REO Project team, BP 105, F-78153 Le Chesnay Cedex, France.

[3] CEA, DEN/DANS/DM2S/SFME/LETR, F-91191 Gif-sur-Yvette, France.

[4] DTIM/ONERA Centre de Toulouse, 2 avenue Edouard Belin, 31055 Toulouse, France. e-mail: bruno.fornet@onera.fr

[5] Project-Team SIMPAF, INRIA Lille Nord Europe Research Centre Park Plazza, 40 avenue Halley 59650 Villeneuve d'Ascq CEDEX, France; e-mail: thierry.goudon@inria.fr & pauline.lafitte-godillon@math.univ-lille1.fr & lagoutie@math.jussieu.fr

[6] Laboratoire Paul Painlevé, USTL-CNRS UMR 8524, Cité Scientifique, 59655 Villeneuve d'Ascq CEDEX, France.

[7] Université Paris Diderot-Paris 7, Lab. J.-L. Lions, 175 rue du Chevaleret, BC 187, F-75013 Paris, France.

[8] Université Paris Nord - Institut Galilée LAGA (Laboratoire d'Analyse, Géométrie et Applications) Avenue J.B. Clément 93430 Villetaneuse, France; e-mail: merlet@math.univ-paris13.fr

energy e, the pressure p, the temperature Θ by the relations

$$e = \frac{p}{(\gamma-1)\rho} \geq 0, \qquad p = R\rho\Theta, \qquad E = e + \frac{u^2}{2}$$

where R is the perfect gas constant and $\gamma > 1$ is the adiabatic constant. The disperse phase is described by its density distribution in phase space $f(t,x,v) \geq 0$, where the variable $v \in \mathbb{R}^N$ stands for the velocity of the particles. Macroscopic quantities can be defined as moments with respect to v; in what follows we need

$$\left\{\begin{array}{ll} \text{the macroscopic density:} & n(t,x) = \displaystyle\int_{\mathbb{R}^N} f(t,x,v)\ \mathrm{d}v, \\ \text{the bulk velocity:} & nV(t,x) = \displaystyle\int_{\mathbb{R}^N} v\ f(t,x,v)\ \mathrm{d}v, \\ \text{the temperature:} & n|V|^2(t,x) + Nn\Theta_p = \displaystyle\int_{\mathbb{R}^N} |v|^2\ f(t,x,v)\ \mathrm{d}v, \\ \text{the heat flux:} & q(t,x) = \displaystyle\int_{\mathbb{R}^N} v\ \frac{|v|^2}{2}\ f(t,x,v)\ \mathrm{d}v. \end{array}\right.$$

The evolution of the density f is governed by

$$\partial_t f + v\cdot\nabla_x f = \mathrm{div}_v\big((v-u)f + \Theta\nabla_v f\big) - \eta_p \nabla_x\Phi\cdot\nabla_v f. \tag{1}$$

The divergence term in the right-hand side accounts for both the friction force exerted by the fluid on the particles, which is supposed to be proportional to the relative velocity $(v-u)$, and the Brownian motion of the particles, which induces diffusion with respect to the velocity variable, depending on the surrounding temperature Θ. The second term in the right-hand side comes from an external force with potential Φ and η_p is a positive constant. The evolution of the fluid obeys the Euler system

$$\left\{\begin{array}{l} \partial_t\rho + \mathrm{div}_x(\rho u) = 0, \\ \partial_t(\rho u) + \mathrm{Div}_x(\rho u\otimes u) + \nabla_x p = \mathscr{F} - \eta_f\rho\nabla_x\Phi, \\ \partial_t(\rho E) + \mathrm{div}_x\big((\rho E + p)u\big) = \mathscr{E} - \eta_f\rho u\cdot\nabla_x\Phi, \end{array}\right. \tag{2}$$

where η_f is another positive coefficient that accounts for a possible difference of amplitude in the forces applied to the fluid or the disperse phase. Remark that assuming η_p and η_f positive means that the external force associated to the potential Φ acts on opposite directions on the particles and on the fluid. Bearing in mind the example of gravity we have $\nabla_x\Phi = g \in \mathbb{R}^N$ and we are dealing with a situation where particles are light compared to the fluid: gravity pushes the fluid downward, while buoyancy effects push the particles upward. We refer to [3], [4] for a discussion on this modeling issue.

In (2), the force $\mathscr{F}$ arising in the right-hand side of the momentum equation is given by the friction force exerted by the particles on the fluid and it reads

$$\mathscr{F}(t,x) = n(V-u)(t,x) = \int_{\mathbb{R}^N} (v-u)f(t,x,v)\ \mathrm{d}v. \tag{3}$$

The energy exchanges between the two phases split as follows:

$$\begin{array}{rcl} \mathscr{E}(t,x) & = & \Big(n(V-u)\cdot u + Nn(\Theta_p - \Theta)\Big)(t,x) + \mathscr{E}'(t,x) \\ & = & \displaystyle\int_{\mathbb{R}^N} (v-u)f(t,x,v)\ \mathrm{d}v\cdot u \\ & & \displaystyle + N(\Theta_p-\Theta)(t,x)\int_{\mathbb{R}^N} f(t,x,v)\ \mathrm{d}v + \mathscr{E}'(t,x). \end{array} \tag{4}$$

In this expression, the first term is nothing but the work of the friction force $\mathscr{F}$ while the second describes the heat transfer between the two phases; the last term, which will be specified later on, guarantees the total

energy conservation. As a matter of fact, we remark that the evolution of the internal energy is driven by

$$\partial_t(\rho e) + \mathrm{div}_x(\rho e u) + p\mathrm{div}_x u = Nn(\Theta_p - \Theta) + \mathscr{E}'(t,x).$$

The goal is to extend, at least formally, the analysis performed in [3] (see also [9,10]) and, having identified relevant asymptotic regimes, to design adapted Asymptotic Preserving schemes, in the spirit of [4,5,8]. The dissipative or relaxation properties of the system are crucial to this approach.

2. Dissipation properties

Let us start by considering the evolution of the macroscopic quantities

$$\begin{aligned}
&\partial_t n + \mathrm{div}_x(nV) = 0,\\
&\partial_t(nV) + \mathrm{Div}_x\Big(\int_{\mathbb{R}^N} v\otimes v f\ \mathrm{d}v\Big) = -n(V-u) + \eta_p n\nabla_x\Phi,\\
&\partial_t\Big(\int_{\mathbb{R}^N}\frac{v^2}{2} f\ \mathrm{d}v\Big) + \mathrm{div}_x q = -\int_{\mathbb{R}^N}(v-u)\cdot v\ f\ \mathrm{d}v + Nn\Theta + \eta_p nV\cdot\nabla_x\Phi,
\end{aligned} \tag{5}$$

where we use integration by parts for evaluating the right-hand sides. Observe that the total momentum is conserved (up to the gravity term) since

$$\partial_t(\rho u + nV) + \mathrm{Div}_x\Big(\rho u\otimes u + \int_{\mathbb{R}^N} v\otimes v f\ \mathrm{d}v\Big) + \nabla_x p = (\eta_p n - \eta_f\rho)\nabla_x\Phi. \tag{6}$$

Next, since we consider the mixture fluid/particles as a whole, the total energy should also be conserved, which will give the definition of $\mathscr{E}'$. We remark that

$$\begin{aligned}
\int_{\mathbb{R}^N}(v-u)\cdot v\ f\ \mathrm{d}v &= \int_{\mathbb{R}^N}|v-u|^2 f\ \mathrm{d}v + n(V-u)\cdot u\\
&= n(|V|^2 + N\Theta_p - 2V\cdot u + |u|^2) + n(V-u)\cdot u\\
&= n|V-u|^2 + Nn\Theta_p + n(V-u)\cdot u.
\end{aligned}$$

Therefore the kinetic energy of the particles obeys

$$\partial_t\Big(\int_{\mathbb{R}^N}\frac{|v|^2}{2} f\ \mathrm{d}v\Big) + \mathrm{div}_x q = -n|V-u|^2 + Nn(\Theta-\Theta_p) - n(V-u)\cdot u + \eta_p nV\cdot\nabla_x\Phi.$$

Accordingly, we set

$$\mathscr{E}'(t,x) = n|V-u|^2$$

so that the total energy is driven by

$$\partial_t\Big(\rho E + \int_{\mathbb{R}^N}\frac{|v|^2}{2} f\ \mathrm{d}v\Big) + \mathrm{div}_x\big((\rho E + p)u + q\big) = (\eta_p nV - \eta_f\rho u)\cdot\nabla_x\Phi. \tag{7}$$

Hence, the term $\mathscr{E}'$ appears as a source of internal energy, or a source of heat, for the fluid, produced by the friction with the particle, and proportional to the macroscopic kinetic energy defined by the relative velocity $V-u$.

Next, we consider the entropy $S(t,x)$ defined by the relation

$$S = -\frac{R}{\gamma-1}\ln\left(p\rho^{-\gamma}\right) = -\frac{R}{\gamma-1}\ln\Big(R\frac{\Theta}{\rho^{\gamma-1}}\Big).$$

We check that it satisfies

$$\begin{aligned}
\partial_t(\rho S) + \mathrm{div}_x(\rho S u) &= \frac{R\rho}{p}\left(n(V-u)\cdot u - n(V-u)\cdot u - Nn(\Theta_p-\Theta) - \mathscr{E}'\right)\\
&= -Nn\Big(\frac{\Theta_p}{\Theta} - 1\Big) - \frac{n}{\Theta}|V-u|^2.
\end{aligned} \tag{8}$$

Finally, we look at the entropy of the disperse phase

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} f\ln(f)\ \mathrm{d}v\ \mathrm{d}x &= -\int_{\mathbb{R}^N}\int_{\mathbb{R}^N}\Big((v-u)f\cdot\frac{\nabla_v f}{f}+\Theta\frac{|\nabla_v f|^2}{f}\Big)\ \mathrm{d}v\ \mathrm{d}x\\ &= N\int_{\mathbb{R}^N} n\ \mathrm{d}x-\int_{\mathbb{R}^N}\int_{\mathbb{R}^N}\Theta\frac{|\nabla_v f|^2}{f}\ \mathrm{d}v\ \mathrm{d}x.\end{aligned}$$

Hence, the total entropy satisfies

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\Big(\int_{\mathbb{R}^N}\rho S\ \mathrm{d}x+\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} f\ln(f)\ \mathrm{d}v\ \mathrm{d}x\Big)\\ &=-\int_{\mathbb{R}^N}\int_{\mathbb{R}^N}\Theta\frac{|\nabla_v f|^2}{f}\ \mathrm{d}v\ \mathrm{d}x+2N\int_{\mathbb{R}^N} n\ \mathrm{d}x\\ &\qquad -N\int_{\mathbb{R}^N} n\frac{\Theta_p}{\Theta}\ \mathrm{d}x-\int_{\mathbb{R}^N}\frac{n|V-u|^2}{\Theta}\ \mathrm{d}x\end{aligned}$$

The next argument is two-fold. On the one hand, we observe that

$$N\int_{\mathbb{R}^N} n\ \mathrm{d}x=-\int_{\mathbb{R}^N}\int_{\mathbb{R}^N}\frac{v-V}{\sqrt{\Theta}}\sqrt{f}\cdot\sqrt{\Theta}\frac{\nabla_v f}{\sqrt{f}}\ \mathrm{d}v\ \mathrm{d}x,$$

and on the other hand we have

$$\int_{\mathbb{R}^N}|v-V|^2 f\ \mathrm{d}v=n|V|^2+Nn\Theta_p-2nV\cdot V+n|V|^2=Nn\Theta_p.$$

It follows that

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\Big(\int_{\mathbb{R}^N}\rho S\ \mathrm{d}x+\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} f\ln(f)\ \mathrm{d}v\ \mathrm{d}x\Big)\\ &\quad=-\int_{\mathbb{R}^N}\int_{\mathbb{R}^N}\Big(\Theta\frac{|\nabla_v f|^2}{f}+2\frac{v-V}{\sqrt{\Theta}}\sqrt{f}\cdot\sqrt{\Theta}\frac{\nabla_v f}{\sqrt{f}}+\frac{|v-V|^2}{\Theta}f\Big)\ \mathrm{d}v\ \mathrm{d}x\\ &\qquad-\int_{\mathbb{R}^N} n\frac{|V-u|^2}{\Theta}\ \mathrm{d}x\\ &\quad=-\int_{\mathbb{R}^N}\int_{\mathbb{R}^N}\Big|\sqrt{\Theta}\frac{\nabla_v f}{\sqrt{f}}+\frac{v-V}{\sqrt{\Theta}}\sqrt{f}\Big|^2\ \mathrm{d}v\ \mathrm{d}x-\int_{\mathbb{R}^N} n\frac{|V-u|^2}{\Theta}\ \mathrm{d}x,\end{aligned}$$

which indicates that the total entropy of the system is dissipated.

Let us summarize the computation as follows.

Proposition 1. *Let (ρ,u,E,f) be a (smooth enough) solution of (1)–(4). Then, both the total momentum and the total energy are conserved while the total entropy is dissipated and we have*

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\Big(\int_{\mathbb{R}^N}\rho E\ \mathrm{d}x+\int_{\mathbb{R}^N}\int_{\mathbb{R}^N}\frac{v^2}{2}\ f\ \mathrm{d}v\ \mathrm{d}x\Big)=\int_{\mathbb{R}^N}(\eta_p nV-\eta_f\rho u)\cdot\nabla_x\Phi\ \mathrm{d}x,\\ &\frac{\mathrm{d}}{\mathrm{d}t}\Big(\int_{\mathbb{R}^N}\rho S\ \mathrm{d}x+\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} f\ln(f)\ \mathrm{d}v\ \mathrm{d}x\Big)\\ &=-\int_{\mathbb{R}^N}\int_{\mathbb{R}^N}\Big|\sqrt{\Theta}\frac{\nabla_v f}{\sqrt{f}}+\frac{v-V}{\sqrt{\Theta}}\sqrt{f}\Big|^2\ \mathrm{d}v\ \mathrm{d}x-\int_{\mathbb{R}^N}\int_{\mathbb{R}^N} f\frac{|V-u|^2}{\Theta}\ \mathrm{d}v\ \mathrm{d}x\le 0.\end{aligned}\tag{9}$$

Note that the dissipation terms vanish when

$$u=V\quad\text{and}\quad f(t,x,v)=\frac{n(t,x)}{\big(2\pi\Theta(t,x)\big)^{N/2}}\ \exp\Big(-\frac{|v-V(t,x)|^2}{2\Theta(t,x)}\Big).$$

3. Bubbling Regime

According to [3], the "Bubbling Regime" relies on the following scaling:

$$\left\{\begin{array}{l}\partial_t f^\epsilon + \dfrac{1}{\epsilon} v\cdot\nabla_x f^\epsilon = \dfrac{1}{\epsilon^2}\mathrm{div}_v\big((v-\epsilon u^\epsilon)f^\epsilon + \Theta^\epsilon\nabla_v f^\epsilon\big) - \dfrac{1}{\epsilon}\nabla_x\Phi\cdot\nabla_v f^\epsilon,\\ \partial_t\rho^\epsilon + \mathrm{div}_x(\rho^\epsilon u^\epsilon) = 0,\\ \partial_t(\rho^\epsilon u^\epsilon) + \mathrm{Div}_x(\rho^\epsilon u^\epsilon\otimes u^\epsilon) + \nabla_x p^\epsilon = \dfrac{1}{\epsilon}\displaystyle\int_{\mathbb{R}^N}(v-\epsilon u^\epsilon)f^\epsilon\ \mathrm{d}v - \dfrac{1}{1-\epsilon^2}\rho^\epsilon\nabla_x\Phi,\\ \partial_t(\rho^\epsilon E^\epsilon) + \mathrm{div}_x\big((\rho^\epsilon E^\epsilon + p^\epsilon)u^\epsilon\big)\\ \qquad = \dfrac{1}{\epsilon^2}\left(\displaystyle\int_{\mathbb{R}^N}(v-\epsilon u^\epsilon)f^\epsilon\ \mathrm{d}v\cdot\epsilon u^\epsilon + Nn^\epsilon(\Theta_p^\epsilon-\Theta^\epsilon) + n^\epsilon|V^\epsilon-\epsilon u^\epsilon|^2\right) - \dfrac{1}{1-\epsilon^2}\rho^\epsilon u^\epsilon\cdot\nabla_x\Phi,\end{array}\right. \tag{10}$$

where it is convenient to introduce the following notation

$$\begin{array}{ll} n^\epsilon = \displaystyle\int_{\mathbb{R}^N} f^\epsilon\ \mathrm{d}v, & J^\epsilon = \dfrac{1}{\epsilon}n^\epsilon V^\epsilon = \displaystyle\int_{\mathbb{R}^N}\frac{v}{\epsilon}\ f^\epsilon\ \mathrm{d}v,\\ n^\epsilon|V^\epsilon|^2 + Nn^\epsilon\Theta_p^\epsilon = \displaystyle\int_{\mathbb{R}^N} v^2\ f^\epsilon\ \mathrm{d}v, & q^\epsilon = \displaystyle\int_{\mathbb{R}^N}\frac{v}{\epsilon}\frac{v^2}{2}\ f^\epsilon\ \mathrm{d}v.\end{array}$$

In this context Proposition 1 recasts as

$$\begin{array}{l}\dfrac{\mathrm{d}}{\mathrm{d}t}\Big(\displaystyle\int_{\mathbb{R}^N}\rho^\epsilon E^\epsilon\ \mathrm{d}x + \int_{\mathbb{R}^N}\int_{\mathbb{R}^N}\frac{v^2}{2}\ f^\epsilon\ \mathrm{d}v\ \mathrm{d}x\Big) = \displaystyle\int_{\mathbb{R}^N}\left(\frac{n^\epsilon V^\epsilon}{\epsilon} - \frac{\rho^\epsilon u^\epsilon}{1-\epsilon^2}\right)\cdot\nabla_x\Phi\ \mathrm{d}x,\\ \dfrac{\mathrm{d}}{\mathrm{d}t}\Big(\displaystyle\int_{\mathbb{R}^N}\rho^\epsilon S^\epsilon\ \mathrm{d}x + \int_{\mathbb{R}^N}\int_{\mathbb{R}^N} f^\epsilon\ln(f^\epsilon)\ \mathrm{d}v\ \mathrm{d}x\Big)\\ \qquad + \dfrac{1}{\epsilon^2}\displaystyle\int_{\mathbb{R}^N}\int_{\mathbb{R}^N}\left[\left|\sqrt{\Theta^\epsilon}\frac{\nabla_v f^\epsilon}{\sqrt{f^\epsilon}} + \frac{v-V^\epsilon}{\sqrt{\Theta^\epsilon}}\sqrt{f^\epsilon}\right|^2 + f^\epsilon|V^\epsilon-\epsilon u^\epsilon|^2\right]\mathrm{d}v\ \mathrm{d}x \le 0,\end{array} \tag{11}$$

and we infer the following relaxation effects

$$\begin{array}{l} V^\epsilon \simeq \epsilon u^\epsilon\\ f^\epsilon(t,x,v) \simeq \dfrac{n^\epsilon(t,x)}{\big(2\pi\Theta^\epsilon(t,x)\big)^{N/2}}\ \exp\Big(-\dfrac{|v-V^\epsilon(t,x)|^2}{2\Theta^\epsilon(t,x)}\Big) \simeq n^\epsilon(t,x)\ M_{\Theta^\epsilon(t,x)}(v)\end{array}$$

where M_Θ stands for the centered Maxwellian with temperature Θ

$$M_\Theta(v) = (2\pi\Theta)^{-N/2}\ e^{-v^2/(2\Theta)}.$$

Accordingly, we obtain

$$\int_{\mathbb{R}^N} v^2 f^\epsilon\ \mathrm{d}v = n^\epsilon|V^\epsilon|^2 + Nn^\epsilon\Theta_p^\epsilon \simeq \int_{\mathbb{R}^N} v^2\ n^\epsilon M_{\Theta^\epsilon}\ \mathrm{d}v = Nn^\epsilon\Theta^\epsilon$$

with V^ϵ expected to be of order $\mathcal{O}(\epsilon)$. We deduce that

$$\Theta^\epsilon - \Theta_p^\epsilon \xrightarrow[\epsilon\to 0]{} 0$$

(an observation which can be seen also by considering the energy equation). Eventually, we guess that the asymptotic behavior is described by the evolution of the macroscopic quantities n, ρ, u, Θ only, and we assume that

$$(n^\epsilon, \rho^\epsilon, u^\epsilon, \Theta^\epsilon) \xrightarrow[\epsilon\to 0]{} (n, \rho, u, \Theta)$$

(in a strong enough sense...), and consequently

$$f^\epsilon(t,x,v) \to n(t,x)\ M_{\Theta(t,x)}(v). \tag{12}$$

It thus remains to determine the equations satisfied by these quantities.

To this end, we go back to the moments equations (5) which recast here in the following rescaled form

$$\left\{\begin{array}{l} \partial_t n^\epsilon + \mathrm{div}_x(J^\epsilon) = 0, \\ \epsilon^2 \partial_t J^\epsilon + \mathrm{Div}_x\Big(\displaystyle\int_{\mathbb{R}^N} v\otimes v f^\epsilon \,\mathrm{d}v\Big) = -J^\epsilon + n^\epsilon u^\epsilon + n^\epsilon \nabla_x \Phi, \\ \partial_t\Big(\displaystyle\int_{\mathbb{R}^N} \frac{v^2}{2} f^\epsilon \,\mathrm{d}v\Big) + \mathrm{div}_x q^\epsilon = \dfrac{1}{\epsilon^2}\Big(-\displaystyle\int_{\mathbb{R}^N}(v-\epsilon u^\epsilon)\cdot v\ f^\epsilon \,\mathrm{d}v + N n^\epsilon \Theta^\epsilon\Big) - \dfrac{1}{\epsilon}\nabla_x \cdot \displaystyle\int_{\mathbb{R}^N} \frac{|v|^2}{2}\nabla_v f^\epsilon \,\mathrm{d}v \\ \qquad\qquad = \dfrac{1}{\epsilon^2}\ n^\epsilon(-|V^\epsilon|^2 - N\Theta_p^\epsilon + \epsilon u^\epsilon\cdot V^\epsilon + N\Theta^\epsilon\Big) + \dfrac{n^\epsilon V^\epsilon}{\epsilon}\cdot\nabla_x\Phi \\ \qquad\qquad = N n^\epsilon\ \dfrac{\Theta^\epsilon - \Theta_p^\epsilon}{\epsilon^2} + J^\epsilon\cdot\Big(u^\epsilon - \dfrac{J^\epsilon}{n^\epsilon} + \nabla_x\Phi\Big). \end{array}\right. \tag{13}$$

We also suppose that

$$(J^\epsilon, q^\epsilon) \xrightarrow[\epsilon\to 0]{} (J, q)$$

holds and we wish to relate these limits to (n, ρ, u, Θ). The ansatz (12) allows to compute the limit of the second moment

$$\int_{\mathbb{R}^N} v\otimes v\ f^\epsilon \,\mathrm{d}v \xrightarrow[\epsilon\to 0]{} n\ \int_{\mathbb{R}^N} v\otimes v\ M_\Theta(v)\,\mathrm{d}v = n\Theta\ \mathbb{I}.$$

Therefore, letting ϵ go to 0 in the momentum equation leads to

$$0 + \nabla_x(n\Theta) = -J + n(u + \nabla_x\Phi). \tag{14}$$

Next, we consider the equation for the total energy

$$\partial_t\Big(\rho^\epsilon E^\epsilon + \int_{\mathbb{R}^N}\frac{|v|^2}{2} f^\epsilon \,\mathrm{d}v\Big) + \mathrm{div}_x\big((\rho^\epsilon E^\epsilon + p^\epsilon)u^\epsilon + q^\epsilon\big) = (J^\epsilon - \frac{\rho^\epsilon u^\epsilon}{1-\epsilon^2})\cdot\nabla_x\Phi. \tag{15}$$

As $\epsilon\to 0$ we get

$$\partial_t\Big(\rho E + \frac{N}{2} n\Theta\Big) + \mathrm{div}_x\big((\rho E + p)u + q\big) = (J - \rho u)\cdot\nabla_x\Phi. \tag{16}$$

We are left with the task of identifying the heat flux q. To this purpose, we observe that

$$\epsilon^2\partial_t q^\epsilon + \mathrm{Div}_x\Big(\int_{\mathbb{R}^N} v\otimes v\frac{|v|^2}{2} f^\epsilon \,\mathrm{d}v\Big) = -3q^\epsilon + \Big(\int_{\mathbb{R}^N}\Big(v\otimes v + \frac{|v|^2}{2}\mathbb{I}\Big) f^\epsilon \,\mathrm{d}v\Big)u^\epsilon + (N+2)\Theta^\epsilon J^\epsilon - \int_{\mathbb{R}^N} v\frac{|v|^2}{2}(\nabla_v f\cdot\nabla_x\Phi)\,\mathrm{d}v$$

Hence, when ϵ goes to 0 we find

$$\begin{aligned} 0 + \nabla_x\Big(\frac{n}{2N}\int_{\mathbb{R}^N}|v|^4 M_\Theta(v)\,\mathrm{d}v\Big) &= \nabla_x\Big(\frac{N+2}{2}\ n\Theta^2\Big) \\ &= \frac{N+2}{2}\Theta^2\nabla_x n + (N+2)n\Theta\nabla_x\Theta \\ &= -3q + \frac{N+2}{2} n\Theta(u+\nabla_x\Phi) + (N+2)\Theta\big(n(u+\nabla_x\Phi) - \nabla_x(n\Theta)\big) \\ &= -3q + \frac{3}{2}(N+2)n\Theta(u+\nabla_x\Phi) - (N+2)(n\Theta\nabla_x\Theta + \Theta^2\nabla_x n). \end{aligned}$$

since $\int_{\mathbb{R}^N}|v|^4 M_1(v)\,\mathrm{d}v = N(N+2)$. It defines the heat flux q by means of n, u and Θ:

$$q = \frac{N+2}{2} n\Theta(u+\nabla_x\Phi) - \frac{N+2}{2}\Theta^2\nabla_x n - \frac{2(N+2)}{3} n\Theta\nabla_x\Theta.$$

The conclusion of the computations states as follows.

Theorem 1. *Assuming strong enough convergence of the macroscopic quantities, the limit as ϵ goes to 0 is described by the following set of PDEs*

$$\left\{\begin{array}{l} \partial_t\rho + \mathrm{div}_x(\rho u) = 0, \\ \partial_t n + \mathrm{div}_x\big(nu - \nabla_x(n\Theta)\big) = -\mathrm{div}_x(n\nabla_x\Phi), \\ \partial_t(\rho u) + \mathrm{Div}_x(\rho u\otimes u + p) = -\nabla_x(n\Theta) + (n-\rho)\nabla_x\Phi, \\ \partial_t\Big(\rho E + \dfrac{N}{2}n\Theta\Big) + \mathrm{div}_x\Big((\rho E+p)u + \dfrac{N+2}{2}n\Theta u\Big) \\ \qquad = \dfrac{N+2}{2}\mathrm{div}_x(\Theta^2\nabla_x n) + \dfrac{2(N+2)}{3}\mathrm{div}_x(n\Theta\nabla_x\Theta) \\ \qquad\quad - \dfrac{N+2}{2}\mathrm{div}_x(n\Theta\nabla_x\Phi) + (n-\rho)u\nabla_x\Phi + n|\nabla_x\Phi|^2 - \nabla_x\Phi\cdot\nabla_x(n\Theta). \end{array}\right. \tag{17}$$

In System (17) the unknowns are the fluid density ρ, the particle (macroscopic) density n and the common velocity u and temperature Θ. Assuming a constant temperature and dropping the energy equation, we recover the equations derived in [3]. There appear some unusual terms in the equations which are reminiscent of the so-called Soret and Dufour effects, see [14]. The Soret effect relies on the conduction current $n\nabla_x\Theta$: the temperature gradient produces a flow of particles. Note that particles are also subject to diffusion with a diffusion coefficient proportional to the temperature. The Dufour effect relies on the effects of concentration gradients on the evolution of the temperature.

Another way of deriving the limit equations consists in considering fluctuations in the ansatz (12). This approach motivates the design of "Asymptotic Preserving" schemes, see [4, 5, 8]. Namely, we set

$$f^\epsilon(t,x,v) = n^\epsilon(t,x)M_{\Theta^\epsilon(t,x)}(v) + \epsilon r^\epsilon(t,x,v)$$

and we wish to identify the limit r of (r^ϵ) by using the equation

$$\epsilon\partial_t f^\epsilon + v\cdot\nabla_x f^\epsilon = \mathscr{L}_{\Theta^\epsilon} r^\epsilon - u^\epsilon\cdot\nabla_v f^\epsilon - \nabla_x\Phi\cdot\nabla_v f^\epsilon \tag{18}$$

where $\mathscr{L}_\Theta$ stands for the Fokker-Planck operator

$$\mathscr{L}_\Theta f = \mathrm{div}_v(vf + \Theta\nabla_v f)$$

(recall that $\mathscr{L}_\Theta(M_\Theta) = 0$ for every $\Theta\in\mathbb{R}$). Indeed, assuming that r^ϵ converges to r, we can express the mass and heat fluxes as follows

$$J^\epsilon = \int_{\mathbb{R}^N} vr^\epsilon\,\mathrm{d}v \xrightarrow[\epsilon\to 0]{} J = \int_{\mathbb{R}^N} vr\,\mathrm{d}v, \qquad q^\epsilon = \int_{\mathbb{R}^N} v\frac{|v|^2}{2}r^\epsilon\,\mathrm{d}v \xrightarrow[\epsilon\to 0]{} q = \int_{\mathbb{R}^N} v\frac{|v|^2}{2}r\,\mathrm{d}v.$$

As a matter of fact, the leading contribution is the last term of (18) reads

$$-(u^\epsilon + \nabla_x\Phi)\cdot\nabla_v(n^\epsilon M_{\Theta^\epsilon}) = +(u^\epsilon+\nabla_x\Phi)\cdot\frac{v}{\Theta^\epsilon}n^\epsilon M_{\Theta^\epsilon}$$

Therefore, r is characterized by the relation

$$\begin{array}{rcl} \mathscr{L}_\Theta r &=& v\cdot\nabla_x(nM_\Theta) - (u+\nabla_x\Phi)\cdot\dfrac{v}{\Theta}nM_\Theta \\ &=& vM_\Theta\cdot\Big(\nabla_x n - (u+\nabla_x\Phi)\dfrac{n}{\Theta} - \dfrac{N}{2}\dfrac{n}{\Theta}\nabla_x\Theta\Big) + v\dfrac{|v|^2}{2}M_\Theta\cdot\dfrac{n}{\Theta^2}\nabla_x\Theta. \end{array} \tag{19}$$

The solution splits as follows

$$\begin{array}{l} r = r^1 + r^2, \\ r^1(t,x,v) = \chi(t,x,v)\cdot\Big(\nabla_x n - (u+\nabla_x\Phi)\dfrac{n}{\Theta} - \dfrac{N}{2}\dfrac{n}{\Theta}\nabla_x\Theta\Big), \\ r^2(t,x,v) = \lambda(t,x,v)\cdot\dfrac{n}{\Theta^2}\nabla_x\Theta, \end{array}$$

where $\chi = (\chi_1, ..., \chi_N)$ and $\lambda = (\lambda_1, ..., \lambda_N)$ are solutions of the auxiliary problems (where t, x appear only as parameters through the temperature)

$$\mathscr{L}_{\Theta(t,x)}\chi_j = v_j M_{\Theta(t,x)}, \qquad \mathscr{L}_{\Theta(t,x)}\lambda_j = v_j \ \frac{|v|^2}{2}\ M_{\Theta(t,x)}.$$

We check readily that

$$\chi = -vM_\Theta, \qquad \lambda = -vM_\Theta \ \frac{1}{3}\Big(\frac{|v|^2}{2} + (N+2)\Theta\Big).$$

The associated mass fluxes are given by

$$\begin{aligned} J_\chi &= \int_{\mathbb{R}^N} v \otimes \chi \ \mathrm{d}v = -\int_{\mathbb{R}^N} v \otimes vM_\Theta \ \mathrm{d}v = -\Theta\mathbb{I}, \\ J_\lambda &= \int_{\mathbb{R}^N} v \otimes \lambda \ \mathrm{d}v = -\int_{\mathbb{R}^N} v \otimes v\frac{1}{3}\Big(\frac{v^2}{2} + (N+2)\Theta\Big)M_\Theta \ \mathrm{d}v \\ &= -\frac{1}{3}(N+2)\ (1/2+1)\ \Theta^2\mathbb{I} = -\frac{N+2}{2}\ \Theta^2\mathbb{I}, \end{aligned}$$

while for the heat flux we obtain

$$\begin{aligned} q_\chi &= \int_{\mathbb{R}^N} v\frac{v^2}{2} \otimes \chi \ \mathrm{d}v = -\frac{N+2}{2}\Theta^2\mathbb{I}, \\ q_\lambda &= \int_{\mathbb{R}^N} v\frac{v^2}{2} \otimes \lambda \ \mathrm{d}v = -\frac{1}{3}\Big(\frac{(N+4)(N+2)}{4} + \frac{(N+2)^2}{2}\Big)\Theta^3\mathbb{I}. \end{aligned}$$

Finally, we are led to the following asymptotic behavior

$$\begin{aligned} J^\epsilon \xrightarrow[\epsilon\to 0]{} J_\chi \cdot \Big(\nabla_x n - (u+\nabla_x\Phi)\frac{n}{\Theta} - \frac{N}{2}\frac{n}{\Theta}\nabla_x\Theta\Big) + J_\lambda \cdot \frac{n}{\Theta^2}\nabla_x\Theta, \\ = -\Theta\nabla_x n + n(u+\nabla_x\Phi) + \frac{N}{2}n\nabla_x\Theta - \frac{N+2}{2}n\nabla_x\Theta \\ = n(u+\nabla_x\Phi) - \nabla_x(n\Theta). \end{aligned}$$

Similarly, we obtain

$$\begin{aligned} q^\epsilon \xrightarrow[\epsilon\to 0]{} q_\chi \cdot \Big(\nabla_x n - (u+\nabla_x\Phi)\frac{n}{\Theta} - \frac{N}{2}\frac{n}{\Theta}\nabla_x\Theta\Big) + q_\lambda \cdot \frac{n}{\Theta^2}\nabla_x\Theta \\ = -\frac{N+2}{2}\Big(\Theta^2\nabla_x n - n\Theta(u+\nabla_x\Phi) - \frac{N}{2}n\Theta\nabla_x\Theta\Big) \\ -\frac{(N+4)(N+2)+2(N+2)^2}{12}n\Theta\nabla_x\Theta \\ = \frac{N+2}{2}n\Theta(u+\nabla_x\Phi) - \frac{N+2}{2}\Theta^2\nabla_x n - \frac{2(N+2)}{3}n\Theta\nabla_x\Theta. \end{aligned}$$

We summarize the result as follows.

Proposition 2. *The fluctuation r^ϵ converges to*

$$r = -vM_\Theta \cdot \Big(\nabla_x n - (u+\nabla_x\Phi)\frac{n}{\Theta} - \frac{N}{2}\frac{n}{\Theta}\nabla_x\Theta\Big) - vM_\Theta \ \frac{1}{3}\Big(\frac{v^2}{2} + (N+2)\Theta\Big) \cdot \frac{n}{\Theta^2}\nabla_x\Theta,$$

and the mass and heat fluxes have the following behavior

$$\begin{aligned} J &= n(u+\nabla_x\Phi) - \nabla_x(n\Theta), \\ q &= \frac{N+2}{2}n\Theta(u+\nabla_x\Phi) - \frac{N+2}{2}\Theta^2\nabla_x n - \frac{2(N+2)}{3}n\Theta\nabla_x\Theta. \end{aligned}$$

It allows to identify the limit in (13), and, of course, we recover in this way System (17).

4. Numerical scheme for the limit system

In the section we explore the limit system (17) from a numerical point of view. The space dimension is $N = 1$ in the following, and the potential is gravitational: $\nabla_x \Phi = g$; System (17) reads

$$\left\{\begin{array}{l} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t n + \partial_x\big(nu\big) = \partial^2_{x,x}\big(n\Theta\big) - g\partial_x(n), \\ \partial_t(\rho u) + \partial_x(\rho u \otimes u + p) = -\partial_x(n\Theta) + (n-\rho)g, \\ \partial_t\Big(\rho E + \dfrac{N}{2} n\Theta\Big) + \partial_x\Big((\rho E + p)u + \dfrac{3}{2} n\Theta u\Big) \\ \qquad = \dfrac{3}{2}\partial_x(\Theta^2 \partial_x n) + 2\partial_x(n\Theta\partial_x\Theta) - \dfrac{5}{2} g\partial_x(n\Theta) + (n-\rho)ug + ng^2. \end{array}\right.$$

The numerical scheme that we derived is based on the following form of the system: denoting $\tilde{p} = p + n\Theta$ and $\rho\tilde{E} = \rho E + n\Theta/2$, the system rewrites

$$\left\{\begin{array}{l} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t n + \partial_x\big(nu\big) = \partial^2_{x,x}\big(n\Theta\big) - g\partial_x(n), \\ \partial_t(\rho u) + \partial_x(\rho u \otimes u + \tilde{p}) = (n-\rho)g, \\ \partial_t\Big(\rho\tilde{E}\Big) + \partial_x\Big((\rho\tilde{E} + \tilde{p})u\Big) \\ \qquad = \dfrac{3}{2}\partial_x(\Theta^2 \partial_x n) + 2\partial_x(n\Theta\partial_x\Theta) - \dfrac{5}{2} g\partial_x(n\Theta) + (n-\rho)ug + ng^2. \end{array}\right.$$

Then the design of the scheme is based on a splitting strategy. Every time step is decomposed in 3 stages.

- The first stage solves the system without the right-hand side above:

$$\left\{\begin{array}{l} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t n + \partial_x\big(nu\big) = 0, \\ \partial_t(\rho u) + \partial_x(\rho u \otimes u + \tilde{p}) = 0, \\ \partial_t\Big(\rho\tilde{E}\Big) + \partial_x\Big((\rho\tilde{E} + \tilde{p})u\Big) = 0. \end{array}\right.$$

  Here we note the similarity with the classical Euler system. For this stage a standard Lagrange-remap scheme, as in [6], [7], is used. The stability of this stage relies on a hyperbolic-type condition on the time step: the time step of the whole time iteration is determined by this "hyperbolic" condition.
- The second stage solves the dissipative terms as well as the "transport" terms due to the gravity:

$$\left\{\begin{array}{l} \partial_t \rho = 0, \\ \partial_t n = \partial^2_{x,x}\big(n\Theta\big) - g\partial_x(n), \\ \partial_t(\rho u) = 0, \\ \partial_t\Big(\rho\tilde{E}\Big) = \dfrac{3}{2}\partial_x(\Theta^2 \partial_x n) + 2\partial_x(n\Theta\partial_x\Theta) - \dfrac{5}{2} g\partial_x(n\Theta). \end{array}\right.$$

  This stage imposes a parabolic-type stability condition on the time step and is sub-cycled in the time step.
- the last stage solves the gravity terms:

$$\left\{\begin{array}{l} \partial_t \rho = 0, \\ \partial_t n = 0, \\ \partial_t(\rho u) = (n-\rho)g, \\ \partial_t\Big(\rho\tilde{E}\Big) = (n-\rho)ug + ng^2. \end{array}\right.$$

  This is done numerically with an explicit Euler scheme.

5. Numerical results

The space domain is a bounded interval (namely $[0, 4]$ in the following) and the chosen boundary conditions are wall boundary conditions: concerning the fluid, $u = 0$ and $\partial_x p = 0$. The boundary condition to be imposed on Θ is to be derived from the microscopic equations. The wall condition on the microscopic density f reads $f(t, x, v) = f(t, x, -v)$ for every $v \in \mathbb{R}$, when x belongs to the boundary. From (19) we have that r obeys

$$\mathscr{L}_\Theta r = vM_\Theta\Big(\partial_x n - (u+g)\frac{n}{\Theta} - \frac{1}{2}\frac{n}{\Theta}\partial_x\Theta + \frac{v^2 n}{2\Theta^2}\partial_x\Theta\Big).$$

We remark that, imposing that r is an even function in the v variable on the boundary, the left-hand side is an even function and the right-hand side is an odd function. Thus one has

$$\partial_x n - (u+g)\frac{n}{\Theta} - \frac{1}{2}\frac{n}{\Theta}\partial_x\Theta + \frac{v^2 n}{2\Theta^2}\partial_x\Theta = 0,$$

for every $v \in \mathbb{R}$. Thus $\partial_x\Theta = 0$ and $\partial_x n = (u+g)n/\Theta = gn/\Theta$ on the boundary.

The first numerical result we propose is the case of an isentropic fluid with pressure law $p = \rho^\gamma$ ($\gamma \geq 1$, namely here $\gamma = 1.4$). In this case the energy equation for the fluid is dropped. This allows to compare the results with those obtained for the microscopic (kinetic) system (for $\epsilon > 0$ but small) in [4]. The test-case is one of those presented in this reference: this initial condition is at rest, with $u(0, x) = 0$, $\rho(0, x) = 1$ and $n(0, x) = 0.5$. The final time is $T = 2$. On figures 1,2,3 we observe the strong similarity between the two numerical solutions. This similarity is improved when refining the mesh and having ϵ decrease.

The second series of numerical results are obtained with the same initial conditions, but we want here to analyze the effect of temperature. We thus compare the results of the limit system for an isentropic gas and for an ideal gas. On this first and preliminary result, we see that the thermal effect is not negligible. Figure 7 shows that the temperature has strong variations in the space variable. The consequences concerning the fluid density are particularly important, inducing a high fluid density region on the left-hand side (which can be considered as the bottom, since gravity makes the fluid go leftwards) of the domain. There is a lot of energy on the left-hand side, because of the fluid "falling down" with a high velocity, as we can see in Figures 4, 6 and 7. On the right-hand side, on the contrary, since the particles are going rightwards slowly because of the buoyancy effect, there is very little energy.

References

[1] C. Baranger, G. Baudin, L. Boudin, B. Després, F. Lagoutière, E. Lapébie and T. Takahashi. Liquid jet generation and break-up. In *Numerical Methods for Hyperbolic and Kinetic Equations*, S. Cordier, Th. Goudon, M. Gutnic, E. Sonnendrücker Eds., IRMA Lectures in Mathematics and Theoretical Physics, vol. 7, EMS Publ. House, 2005.

[2] C. Baranger, L. Boudin, P.-E. Jabin and S. Mancini. A modeling of biospray for the upper airways. *ESAIM:Proc*, 14:41–47, 2005.

[3] J.-A. Carrillo and Th. Goudon. Stability and asymptotics analysis of a fluid-particles interaction model. *Comm. PDE*, 31:1349–1379, 2006.

[4] J.-A. Carrillo, Th. Goudon and P. Lafitte. Simulation of Fluid & Particles Flows: Asymptotic Preserving Schemes for Bubbling and Flowing Regimes. *J. Comput. Phys.*, 227(16):7929-7951, 2008.

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

[6] B. Després, Inégalité entropique pour un solveur conservatif du système de la dynamique des gaz en coordonnées de Lagrange. C. R. Acad. Sci. Paris Sér. I Math. 324 (1997) no 11, p. 1301–1306.

[7] G. Gallice, Schémas de type Godunov entropiques et positifs préservant les discontinuités de contact. C. R. Acad. Sci. Paris Sér. I Math. 331 (2000) no 2, p. 149–152.

[8] Th. Goudon, P. Lafitte. Splitting schemes for the simulation of non equilibrium radiative flows. *Preprint*, 2007.

[9] Th. Goudon, P.-E. Jabin and A. Vasseur. Hydrodynamic limit for the Vlasov-Navier-Stokes equations. I. Light particles regime. *Indiana Univ. Math. J.*, 53(6):1495–1515, 2004.

[10] Th. Goudon, P.-E. Jabin and A. Vasseur. Hydrodynamic limit for the Vlasov-Navier-Stokes equations. II. Fine particles regime. *Indiana Univ. Math. J.*, 53(6):1517–1536, 2004.

[11] M. Massot and Ph. Villedieu. Modélisation multi-fluide eulérienne pour la simulation de brouillards denses polydispersés. *C. R. Acad. Sci. Paris Sér. I Math.*, 332(9):869–874, 2001.

[12] J. Mathiaud. Etude de systèmes de type gaz-particules. PhD thesis, ENS Cachan, Sept. 2006.

[13] B. Sportisse. Modélisation et simulation de la pollution atmosphérique. Habilitation à Diriger les Recherches, Sciences de l'Univers, Université Pierre et Marie Curie, June 2007.
[14] R. F. Streater. The Soret and Dufour effects in statistical dynamics. Preprint 1998 (arXiv:math-ph/9910043v1).
[15] F.A. Williams. *Combustion theory*. Benjamin Cummings, 1985.

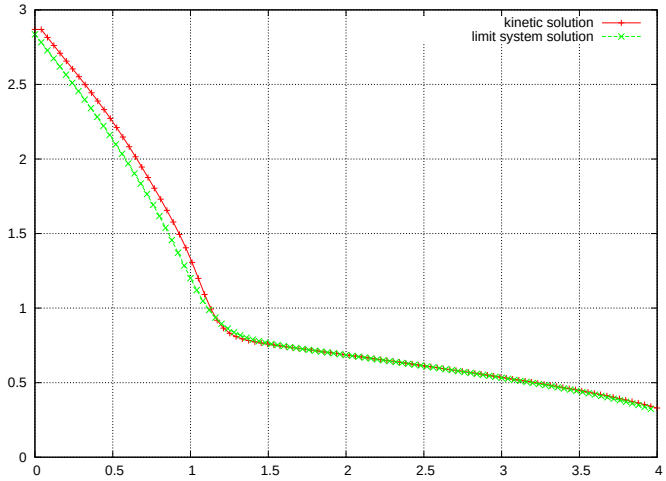


FIGURE 1. Fluid density: comparison of the limit solution and the solution with $\epsilon = 0.1$.

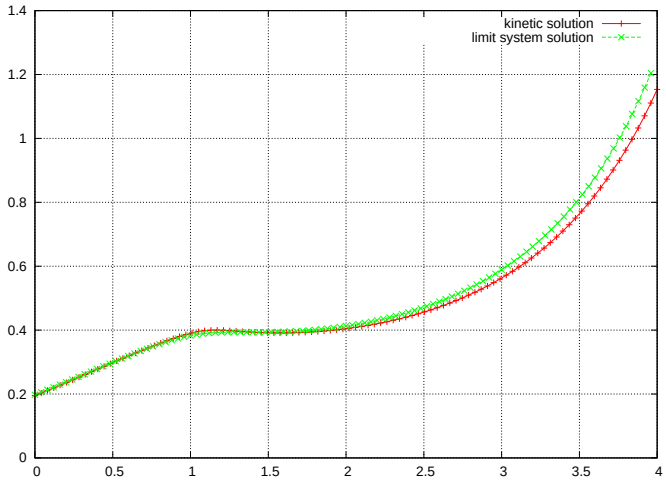


FIGURE 2. Particles macroscopic density: comparison of the limit solution and the solution with $\epsilon = 0.1$.

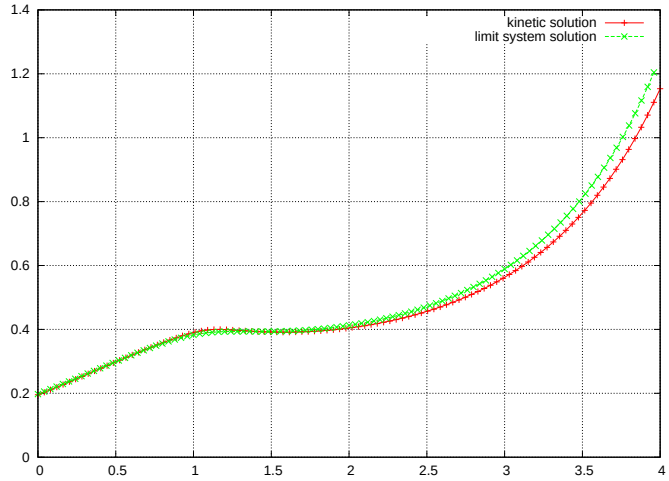


FIGURE 3. Fluid velocity: comparison of the limit solution and the solution with $\epsilon = 0.1$.

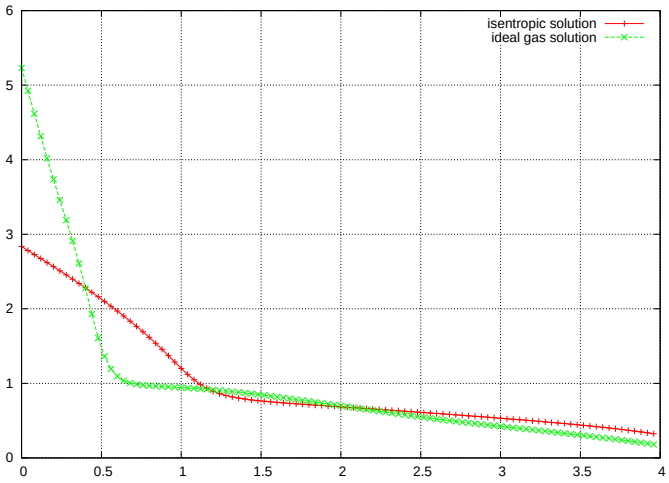


FIGURE 4. Fluid density: comparison of the isentropic and the ideal gas solutions.

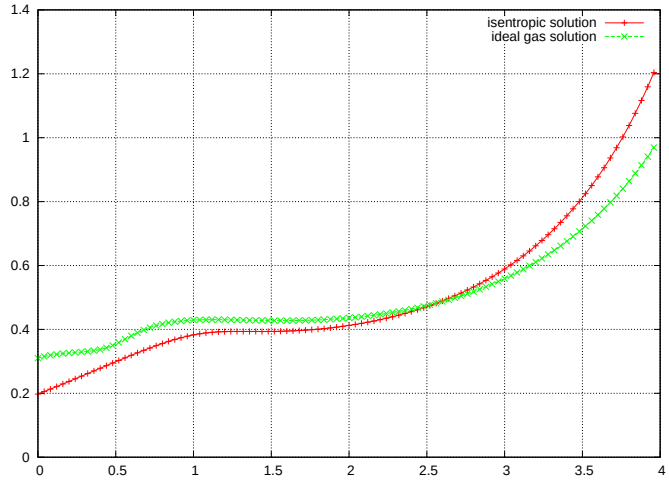


FIGURE 5. Particles macroscopic density: comparison of the isentropic and the ideal gas solutions.

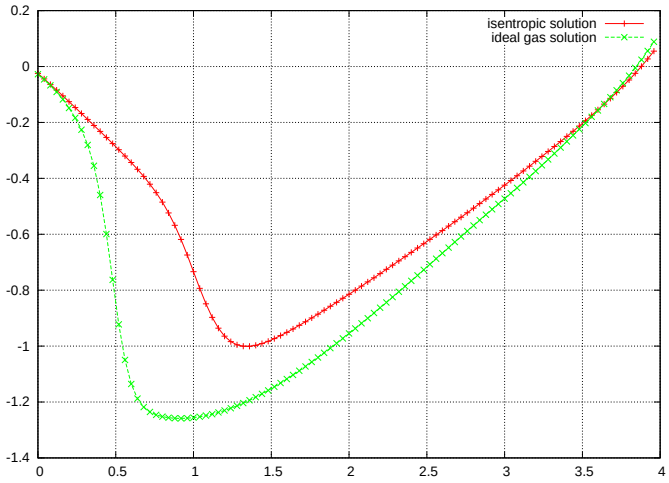


FIGURE 6. Fluid velocity: comparison of the isentropic and the ideal gas solutions.

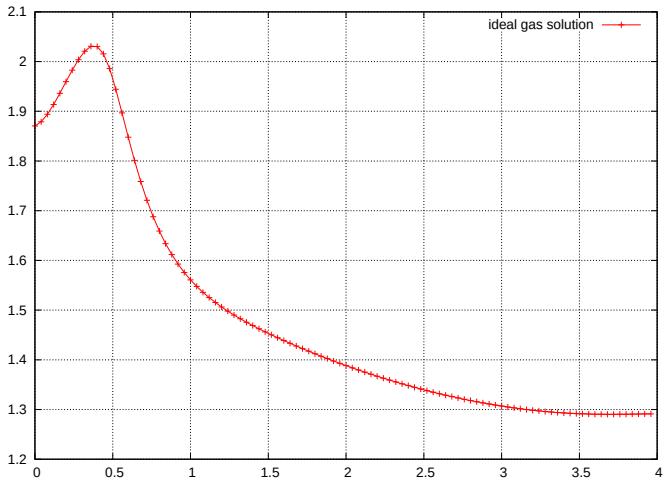


Figure 7. Temperature (ideal gas pressure law).

The Dafermos regularization for the interface coupling of conservation laws

B. Boutin[1,2], F. Coquel[1], E. Godlewski[1]

[1] Université Pierre et Marie Curie-Paris6, UMR 7598 LJLL, F-75005 France; CNRS, UMR 7598 LJLL, Paris, F-75005 France godlewski@ann.jussieu.fr
[2] CEA-Saclay, F-91191, Gif-sur-Yvette cedex, France

Summary. We study the coupling of two conservation laws with different fluxes at the interface $x = 0$. The coupling condition yields as far as possible the continuity of the solution at the interface and thus the coupled model is not conservative in general. This gives rise to interesting questions such as non-uniqueness of self-similar solutions which we have chosen to analyze via a viscous regularization. Introducing a color function, we rewrite the problem in a conservative form involving a source term which is a Dirac measure and in turn this leads to a nonconservative system which may be resonant. In this work we analyze the regularization of this system by a viscous term following Dafermos's approach. We prove the existence of a viscous solution to the Cauchy problem with Riemann data and study the convergence as the viscosity parameter goes to zero to the solution of the coupled Riemann problem.

1 Introduction

We are interested in the study of the coupling of two different conservation laws at a fixed interface

$$\partial_t u + \partial_x f_L(u) = 0,\ x < 0,\ \partial_t u + \partial_x f_R(u) = 0,\ x > 0, \quad t > 0, \tag{1}$$

where $f_\alpha, \alpha = L, R$, are two smooth fluxes. The function u satisfies some initial condition

$$u(x, 0) = u_0(x), \quad x \in \mathrm{R}, \tag{2}$$

and a *coupling condition* at the interface $x = 0$. When u_0 is a Riemann data, as in the following sections,

$$u_0(x) = \begin{cases} u_L,\ x < 0, \\ u_R,\ x > 0, \end{cases} \tag{3}$$

whre $u_L, u_R \in \mathrm{R}$ are two given constant states, we will say that u is solution of a *coupled Riemann problem*.

The theoretical study of a new coupling condition (CC) was initiated in the scalar case [7]. It results by expressing that two boundary value problems

should be well-posed, and resumes to impose as far as possible the continuity of the solution at the interface, we will speak of *state coupling*, without imposing the overall conservativity of the coupled model.

The coupling problem may be ill-posed because the velocities may change sign at the interface and non uniqueness of self-similar solutions of the coupled Riemann problem was observed in [7]. We will consider only solutions u which are entropy solution in each half-space. Contrary to the conservative approach, no *natural* global entropy criterium selecting a unique solution is associated to the nonconservative formulation, and using Dafermos's procedure [6] enables us to recover some of these solutions as self-similar zero-viscosity limits of a regularized system. In case of nonuniqueness, one can construct solutions of a coupled Riemann problem with arbitrary intermediate states but they are not allowed by this regularization procedure.

1.1 Conservative approach: flux coupling

Let us introduce a color function a and consider the following conservative system

$$\left\{ \begin{array}{l} \partial_t u + \partial_x f(u,a) = 0, \quad x \in \mathrm{R},\ t > 0, \\ \partial_t a = 0, \end{array} \right. \tag{4}$$

with

$$f(u,a) = a f_L(u) + (1-a) f_R(u) \tag{5}$$

and

$$a(x,0) = \left\{ \begin{array}{l} 1,\ x < 0 \\ 0,\ x > 0. \end{array} \right. \tag{6}$$

For the conservative system (4), Rankine-Hugoniot condition requires the continuity of the flux at $x = 0$, i.e.,

$$f_L(u(0-,t)) = f_R(u(0+,t)) \tag{7}$$

which we may call *flux coupling*. Note that system (4), (5) is resonant (see [8]) when $\partial_u f(u,a) = a f_L'(u) + (1-a) f_R'(u)$ vanishes, which for $a \in [0,1]$ may occur only if the f_α''s have not the same sign. Moreover, entropy conditions which select a unique solution can be naturally introduced in a number of situations (see [3], [1], [4]).

1.2 Non conservative approach: state coupling

If (7) is natural in a number of physical situations, it is not necessarily the case when the interface is artificially situated. Hence we write instead

$$\partial_t u + \partial_x f(u,a) = \mathcal{M}, \quad x \in \mathrm{R},\ t > 0,\ \partial_t a = 0, \tag{8}$$

with (5) where $\mathcal{M}$ is a Dirac measure concentrated at the interface which involves $[f(u,a)] = f_R(u(0+,t)) - f_L(u(0-,t)) = (f_R(u) - f_L(u))(0,t)$, if one

forces the continuity $u(0+,t) = u(0-,t)$ in the present case of *state coupling*. One notes that

$$f_R(u) - f_L(u) = -\partial_a(af_L(u) + (1-a)f_R(u))$$

and moreover since a is a Heaviside function, one gets the nonconservative formulation

$$\begin{cases} \partial_t u + (af_L'(u) + (1-a)f_R'(u))\partial_x u = 0, \\ \partial_t a = 0. \end{cases} \tag{9}$$

Note that the nonconservative product is well-defined when u is continuous at $x = 0$. Let us set

$$\lambda(u,a) \equiv af_L'(u) + (1-a)f_R'(u). \tag{10}$$

System (9) is strictly hyperbolic if $\lambda \neq 0$, it has two characteristic fields, one genuinely nonlinear (GNL) associated to the eigenvalue λ, the other linearly degenerate (LD) asssociated to 0. As (5), it is resonant when (10) vanishes. Again this corresponds to the fact that the f_α''s may change sign. In order to define the nonconservative product $\lambda(u,a)\partial_x u$ even when u is not continuous, following [10], we use the Dafermos regularization (see (13) below). We prove below the existence of self similar solutions of (13) and the convergence of an extracted subsequence to a solution of the coupled Riemann problem for (1) which satisfies our coupling condition (see ([5]). A closer study of the behavior of the limit profiles characterizes the possible boundary layers at the interface. Explicit conditions are obtained in the case of convex fluxes.

1.3 More general state coupling

The above formalism enables us to treat more general cases where one forces the continuity of another variable (see [2] for a justification of transmitting other variables than the conservative ones in the state coupling). Indeed, consider some continuity condition associated to two monotone mappings Φ_α, simultaneously increasing (or decreasing)

$$\Phi_R(u(0+,t)) = \Phi_L(u(0-,t)). \tag{11}$$

Introducing a change of variables $\Phi_\alpha(u_\alpha(\varphi)) = \varphi$, one wants whenever possible $\varphi(0-) = \varphi(0+)$. In this aim, one defines the function $g(\varphi,a) = f_R(u_R(\varphi)) - f_L(u_L(\varphi))$ and introduces the equation

$$\begin{cases} \partial_t\Big(au_L(\varphi) + (1-a)u_R(\varphi)\Big) + \partial_x\Big(af_L(u_L(\varphi)) + (1-a)f_R(u_R(\varphi))\Big) \\ \qquad\qquad + g(\varphi,a)\partial_x a = 0. \end{cases}$$

Setting for simplicity (in view of (10), this notation is a little abusing)

$$\lambda(\varphi,a) = \frac{af_L'(u_L)u_L' + (1-a)f_R'(u_R)u_R'}{au_L' + (1-a)u_R'}(\varphi)$$

we get the nonconservative system

$$\begin{cases} \partial_t \varphi + \lambda(\varphi, a)\partial_x \varphi = 0 \\ \partial_t a = 0. \end{cases} \tag{12}$$

The two systems (9) and (12) coincide for the identity mapping, i.e., when $\Phi_R(u) = \Phi_L(u) = u$. Again, system (12) has two eigenvalues, 0 and $\lambda(\varphi, a)$. We will now focuss on (9) but a similar analysis can be developed for (12).

2 The Dafermos regularization

In order to approach solutions of the Riemann problem for the non conservative system (9), following [6] we introduce for $\varepsilon > 0$ a system with a viscous regularization

$$\begin{cases} \partial_t u_\varepsilon + (a_\varepsilon f_L'(u_\varepsilon) + (1 - a_\varepsilon) f_R'(u_\varepsilon))\partial_x u_\varepsilon = t\varepsilon \partial_{xx} u_\varepsilon, \\ \partial_t a_\varepsilon = t\varepsilon^2 \partial_{xx} a_\varepsilon, \end{cases} \tag{13}$$

with Riemann data $u_\varepsilon(x, 0) = u_0(x)$ given by (3) (where u_L, u_R are given) and (6) for $a_\varepsilon(x, 0)$. In (13), two facts are noteworthy: one is the presence of the time variable t in the right-hand side viscous term and the other, the power 2 for ε in the second equation.

The term with t can be seen to correspond to a classical viscous regularization in variable $\xi = x/t, T = \ln t$ (see [11] for details). Having scale invariant solutions $u(x, t) = \tilde{u}(x/t)$, it allows to study the approximation of self-similar solutions. Indeed, Dafermos's conjecture for a system of conservation laws $u_t + f(u)_x = 0$ says that any Riemann solution is the limit (as $\varepsilon \to 0$) of self-similar solutions to the Dafermos regularized system $u_t + f(u)_x = \epsilon t u_{xx}$ (this is partially proven in [13], see also [12], [9]).

The difference in the power of the small viscosity coefficient ε corresponds to the fact that the system has two characteristic fields, one is GNL, associated to λ and the first equation, and one associated to 0 (thus LD) and the second equation, which happens to be characteristic. If ε is the size of the boundary layer for a GNL field, ε^2 is typical of a characteristic LD field.

We look for self similar solutions of system (13) of the form $u_\varepsilon(x, t) = \tilde{u}_\varepsilon(x/t)$, $a_\varepsilon(x, t) = \tilde{a}_\varepsilon(x/t)$. Dropping the tilde for simplicity, they satisfy the ODE

$$\begin{cases} (-\xi + \lambda(u_\varepsilon, a_\varepsilon))d_\xi u_\varepsilon = \varepsilon d_{\xi\xi} u_\varepsilon \\ -\xi d_\xi a_\varepsilon = \varepsilon^2 d_{\xi\xi} a_\varepsilon \end{cases} \tag{14}$$

with boundary conditions

$$\begin{cases} \lim_{\xi \to -\infty} u_\varepsilon(\xi) = u_L, \ \lim_{\xi \to +\infty} u_\varepsilon(\xi) = u_R \\ \lim_{\xi \to -\infty} a_\varepsilon(\xi) = 1, \ \lim_{\xi \to +\infty} a_\varepsilon(\xi) = 0. \end{cases} \tag{15}$$

In order to deal with a compact interval $[-M, M]$ for $\xi = x/t$, we use the fact that the propagation speeds of the limit problem are finite and bounded. Hence we may set for $M > \max |\lambda(u, a)|$

$$\begin{cases} u_\varepsilon(-M) = u_L, u_\varepsilon(+M) = u_R \\ a_\varepsilon(-M) = 1, a_\varepsilon(+M) = 0. \end{cases} \tag{16}$$

One extends naturally $u_\varepsilon, a_\varepsilon$ respectively by u_L, u_R and $1, 0$ outside $[-M, M]$. Let us define the interval $B_{LR} = [\min(u_L, u_R), \max(u_L, u_R)]$.

Proposition 1. *Assume $(u, a) \to \lambda(u, a)$ defined by (10) is bounded on $B_{LR} \times [0, 1]$ and f'_L, f'_R are Lipschitz on B_{LR}. There exists a solution $(u_\varepsilon(\xi), a_\varepsilon(\xi))$ of (14), (15) and an extracted sequence $(u_{\varepsilon'}(\xi), a_{\varepsilon'}(\xi))$ such that $u_{\varepsilon'}(\xi)$ converges as $\varepsilon' \to 0$ to $u \in \mathrm{L}^1_{loc}$ where $\overline{u}(x, t) \equiv u(x/t)$ is a weak self-similar solution of (1), (3) satisfying the entropy condition in each half space.*

Proof. The second equation of (14) can be integrated

$$a_\varepsilon(\xi) = \frac{\int_\xi^M e^{-s^2/2\varepsilon^2} ds}{\int_{-M}^M e^{-s^2/2\varepsilon^2} ds},$$

the resulting expression can be written in terms of the *error function* and the whole sequence converges to (6). The first one is nonlinear and we use a fixed point argument. Define for fixed u a primitive of $\xi - \lambda(u(\xi), a_\varepsilon(\xi))$:

$$h(\xi; u, m) = \int_m^\xi (s - \lambda(u(s), a_\varepsilon(s))) ds \tag{17}$$

for some $m \in]-M, M[$ (which will be chosen in such a way that h remains non-negative). Note that h also depends on ε through a_ε. Then, freezing u_ε, one has to solve the auxiliary linear problem

$$\varepsilon d_{\xi\xi} v_\varepsilon = -h(\xi; u_\varepsilon, \alpha) d_\xi v_\varepsilon \tag{18}$$

which can be integrated twice

$$v_\varepsilon(\xi) = u_L + (u_R - u_L) \frac{\int_{-M}^\xi e^{-h(s)/\varepsilon} ds}{\int_{-M}^M e^{-h(s)/\varepsilon} ds} \tag{19}$$

(with shorthand notations for h). The right-hand side of (19) is in fact independent of m. Thus u_ε solution of (14) must be a fixed point of the mapping $T_\varepsilon : u_\varepsilon \to v_\varepsilon$, v_ε given by (19), defined on the set $\mathcal{C} = \{v \in \mathcal{C}^0([-M, M]), v(\xi) \in B_{LR}\}$. Schauder's theorem can be applied since T_ε is an operator on the closed convex set $\mathcal{C}$ of the Banach $\mathcal{C}^0([-M, M])$ which transforms any (bounded) sequence v_n of $\mathcal{C}$ in a relatively compact set of $\mathcal{C}$. Moreover u_ε is BV since from (19), it satisfies the implicit representation formula

$$u_\varepsilon(\xi) = u_L + (u_R - u_L) \frac{\int_{-M}^\xi e^{-h(s; u_\varepsilon(s))/\varepsilon} ds}{\int_{-M}^M e^{-h(s; u_\varepsilon(s))/\varepsilon} ds}. \tag{20}$$

The embedding of $BV(-M,M)$ in L^1 is compact and we can extract a subsequence which converges to some $u \in \mathrm{L}^1$. It is not difficult to prove that u satisfies the ODE

$$-\xi d_\xi u + d_\xi f_\alpha(u) = 0,\ \alpha = L \text{ in } x < 0,\ \alpha = R \text{ in } x > 0,$$

together with the corresponding entropy inequalities, for $\eta \in \mathcal{C}^0(\mathrm{R})$ strictly convex

$$-\xi d_\xi \eta(u) + d_\xi q_\alpha(u) \leq 0 \text{ in } \mathcal{D}'$$

where q_α is the associated entropy flux, satisfying $q'_\alpha = \eta' f'_\alpha$. This is proved first in the intervals $(-M,0)$, $(0,M)$ then in each half space (we refer to [5] for details). Thus $\overline{u}(x,t) = u(x/t)$ is a self-similar solution and a good candidate for a solution of the coupled Riemann problem if we can get more information on its behavior at $x = 0$. Note that the preceding proof does not guarantee uniqueness.

3 Results at the interface

We study the behavior of the limit solution and the existence of a possible boundary layer at the interface. For the *linear* case, $f'_L(.) = a_L, f'_R(.) = a_R$, we can study directly the limit u. It generally exibits one discontinuity between u_L and u_R, propagating at speed a_L (if $a_L \leq 0$ and $a_R < -a_L$) or a_R (if $a_R \geq 0$ and $a_R > -a_L$) then u is continuous at $x = 0$, or at speed 0, i.e., stationary (if $a_L \geq 0 \geq a_R$) in which case $u(0-) = u_L, u(0+) = u_R$. It may have two discontinuities propagating at speed a_L and a_R only if $a_L = -a_R < 0$, then the intermediate state is $(u_L + u_R)/2$. No solution with arbitrary intermediate state is thus approached.

The general case requires a zooming: we introduce a *stretched* variable $y = \xi/\varepsilon$ and define

$$\mathcal{U}_\varepsilon(y) = u_\varepsilon(\varepsilon y),\ \mathcal{A}_\varepsilon(y) = a_\varepsilon(\varepsilon y).$$

For what concerns $\mathcal{A}_\varepsilon$, one shows easily that it converges to the smooth function

$$\mathcal{A}(y) = (1 - \mathrm{erf}(y/\sqrt{2}))/2,$$

satisfying $\mathcal{A}(-\infty) = 1, \mathcal{A}(+\infty) = 0$. If we had put ε in the second equation of (13), it would have resulted in a trivial limit profile for a ($a \equiv 1/2$) while, as we see, ε^2 provides a non trivial smooth profile connecting $1 = a_L$ to $0 = a_R$. Now for $\mathcal{U}$, the characterization of the boundary layer, i.e., the position of $u(0-)$, $u(0+)$, is linked to the existence of non trivial profiles for the limit $\mathcal{U}$ of $\mathcal{U}_\varepsilon$ since we might expect a profile connecting $\mathcal{U}(-\infty) = u(0-), \mathcal{U}(+\infty) = u(0+)$. The first equation (14) gives by rescaling

$$(-\varepsilon y + \lambda(\mathcal{U}_\varepsilon, \mathcal{A}_\varepsilon)(y)) d_y \mathcal{U}_\varepsilon(y) = d_{yy} \mathcal{U}_\varepsilon(y).$$

Hence once we have proved the convergence of $\mathcal{U}_\varepsilon$, the limit $\mathcal{U}$ satisfies the ODE

$$\lambda(\mathcal{U}, \mathcal{A})(y) d_y \mathcal{U}(y) = d_{yy}\mathcal{U}(y).$$

The following result is not straightforward and requires a close study of the integral for $y < 0$ and $y > 0$ and some technical lemmas, we refer to [5] for details.

Proposition 2. *Assume the hypothesis of Proposition 1. The functions u,$\mathcal{U}$ are both simultaneously increasing or decreasing and satisfy*

$$f_L(\mathcal{U}(-\infty)) = f_L(u(0-)),\ f_R(\mathcal{U}(+\infty)) = f_R(u(0+))$$

together with inequalities for the entropy fluxes

$$q_L(\mathcal{U}(-\infty)) \geq q_L(u(0-)),\ q_R(\mathcal{U}(+\infty)) \leq q_R(u(0+))$$

Moreover, if f_L and f_R are assumed to be strictly convex, we have

$$\begin{cases} left: either\ \ \mathcal{U}(-\infty) = u(0-), & right: either\ \ \mathcal{U}(+\infty) = u(0+) \\ \quad or\ \ \mathcal{U}(-\infty) < u(0-), & \quad or\ \ \mathcal{U}(+\infty) > u(0+), \\ f_L'(\mathcal{U}(-\infty)) < 0 < f_L'(u(0-)), & f_R'(\mathcal{U}(+\infty)) > 0 > f_R'(u(0+)). \end{cases} \tag{21}$$

We can get even more information when f_α are both strictly convex. A deeper study of all possible cases (respecting the above constraints) gives 35 cases, according to the signs of $f_L'(u(0-), f_R'(u(0+))$. The solutions we obtain are indeed solution of the *coupled Riemann problem* satisfying the coupling condition. Moreover, as in the linear case, if a solution involves an intermediate state, this state is not arbitrary. This occurs either with two shocks, one with negative $\sigma_L(u_L, \hat{u})$, one with positive speed $\sigma_R(u_R, \hat{u})$, then the intermediate state solves

$$\sigma_R(u_R, \hat{u})(2f_R'(\hat{u}) - \sigma_R(u_R, \hat{u})) = \sigma_L(u_L, \hat{u})(2f_L'(\hat{u}) - \sigma_L(u_L, \hat{u})),$$

where $\sigma_\alpha(u, v) = (f_\alpha(u) - f_\alpha(v))/(u - v)$. Or with two rarefactions on each side of the interface and $f_L'(\hat{u})^2 = f_R'(\hat{u})^2$. Thus we emphasize that some solutions of the coupled Riemann problem we might have constructed directly as in [7] are not attained by this regularization procedure.

For quadratic fluxes, $f_L(u) = u^2/2$, $f_R(u) = (u-c)^2/2$, example which was treated in [7], one can represent them in the $(u_L, u_R)-$plane. For example in case $c > 0$, 8 regions of uniqueness, 2 of non uniqueness. Also, we mention that for $c < 0$, we may have up to 4 solutions when the velocities change sign, more precisely for initial states such that $\sigma_L(u_L, u_R) = (u_L + u_R)/2 < 0$, $\sigma_R(u_L, u_R) = (u_L + u_R)/2 - c > 0$, $f_L'(u_L) = u_L > 0$, $f_R'(u_R) = u_R - c < 0$, and such that $u_R - u_L - 2c < 0$ (the last inequality ensures $\sigma_L(u_L, \hat{u}) < 0$). Moreover some stationary discontinuities are not *stable*, in the sense that they are not obtained numerically.

The above study has given some insight on the existence of multiple solutions which can be attained by a regularization procedure. Some selecting criteria such as continuity of the shock speed might be considered, i.e. given u_L, and u_R varying, the possibility of selecting a solution which ensures the continuity of the shock speed.

This work falls within the scope of an ingoing joint research program on multiphase flows between CEA and University Pierre et Marie Curie-Paris6.

References

1. Adimurthi, Mishra, S., Gowda, G. D. V.: Optimal entropy solutions for conservation laws with discontinuous flux-functions. J. Hyperbolic Differ. Equ. **2**, No 4, 783–837 (2005)
2. Ambroso, A., Chalons, C., Coquel, F., Godlewski, E., Lagoutière, F., Raviart, P.-A., Seguin, N.: Coupling of general Lagrangian systems, submitted (2006)
3. Audusse, E., Perthame, B.: Uniqueness for a scalar conservation law with discontinuous flux via adapted entropies. Proc. Roy. Soc. Edinburgh A **1**35 (2006), no. 2, 253–265.
4. Bachman F., Vovelle, J.: Existence and uniqueness of entropy solution of scalar conservation laws with a flux function involving discontinuous coefficients. Comm. Partial Differential Equations **3**1 (2006), no. 1-3, 371–395.
5. Boutin, B.: Couplage de systèmes de lois de conservation scalaires par une régularisation à la Dafermos, Master dissertation, Paris (2005)
6. Dafermos, C. M.: Solution of the Riemann problem for a class of hyperbolic systems of conservation laws by the viscosity method. Arch. Ration. Mech. Anal., **52** 1–9 (1973)
7. Godlewski E., Raviart, P.-A., The numerical interface coupling of nonlinear hyperbolic systems of conservation laws: I. The scalar case, Numer. Math., **97**, 81–130 (2004)
8. Isaacson, E., Temple, B.: Nonlinear resonance in systems of conservation laws. Siam J. Appl. Math. **52**, no.5, 1260–1278(1992)
9. Joseph, K.T., LeFloch, P.G.: Boundary layers in weak solutions of hyperbolic conservation laws II. Self-similar vanishing diffusion limits. Comm. Pure Appl. Anal. **1**, no. 1, 51–76 (2002)
10. LeFloch, P.G., Tzavaras, A.: Existence theory for the Riemann problem for non-conservative hyperbolic systems. C. R. Acad. Sci. Paris Sér. I Math. **23**, no. 4, 347–352 (1996)
11. Lin, X.-B., Schecter, S.: Stability of selfsimilar solutions of the Dafermos regularization of a system of conservation laws, SIAM J. Math. Anal. **35**, no 4, 884–921 (2003)
12. Schecter, S., Szmolyan, P.: Composite waves in the Dafermos regularization. J. Dynam. Differential Equations **16** no. 3, 847–867(2004)
13. Tzavaras, A. E.: Wave interactions and variation estimates for self-similar zero-viscosity limits in systems of conservation laws. Arch. Rational Mech. Anal. **135**, no. 1, 1–60 (1996)

BIBLIOGRAPHIE

[1] Abeyaratne, R. et Knowles, J. K. (1991a). Implications of viscosity and strain-gradient effects for the kinetics of propagating phase boundaries in solids. *SIAM J. Appl. Math.*, 51(5):1205–1221.

[2] Abeyaratne, R. et Knowles, J. K. (1991b). Kinetic relations and the propagation of phase boundaries in solids. *Arch. Rational Mech. Anal.*, 114(2):119–154.

[3] Amadori, D., Baiti, P., LeFloch, P. G. et Piccoli, B. (1999). Nonclassical shocks and the Cauchy problem for nonconvex conservation laws. *J. Differential Equations*, 151(2):345–372.

[4] Ambroso, A., Boutin, B., Coquel, F., Godlewski, E. et LeFloch, P. G. (2008a). Coupling two scalar conservation laws via Dafermos' self-similar regularization. *ENUMATH2007 proceedings, Numerical Mathematics and Advanced Applications, Springer*.

[5] Ambroso, A., Chalons, C., Coquel, F. et Galié, T. (2009). Relaxation and numerical approximation of a two fluid two pressure diphasic model. to appear.

[6] Ambroso, A., Chalons, C., Coquel, F., Galié, T., Godlewski, E., Raviart, P.-A. et Seguin, N. (2008b). The drift-flux asymptotic limit of barotropic two-phase two-pressure models. *Commun. Math. Sci.*, 6(2):521–529.

[7] Ambroso, A., Chalons, C., Coquel, F., Godlewski, E., Lagoutière, F., Raviart, P.-A. et Seguin, N. (2006). Extension of interface coupling to general Lagrangian systems. *In Numerical mathematics and advanced applications*, pages 852–860. Springer, Berlin.

[8] Ambroso, A., Chalons, C., Coquel, F., Godlewski, E., Lagoutière, F., Raviart, P.-A. et Seguin, N. (2007a). The coupling of homogeneous models for two-phase flows. *Int. J. Finite Vol.*, 4(1):39.

[9] Ambroso, A., Chalons, C., Coquel, F., Godlewski, E., Lagoutière, F., Raviart, P.-A. et Seguin, N. (2007b). A relaxation method for the coupling of systems of conservation laws. *In Hyp2006 Proceedings*. Springer, Berlin.

[10] Ambroso, A., Chalons, C., Coquel, F., Godlewski, E., Lagoutière, F., Raviart, P.-A. et Seguin, N. (2008c). Coupling of general Lagrangian systems. *Math. Comp.*, 77(262):909–941.

[11] Ambroso, A., Chalons, C., Coquel, F., Godlewski, E., Lagoutière, F., Raviart, P.-A. et Seguin, N. (2008d). Relaxation methods and coupling procedures. *Internat. J. Numer. Methods Fluids*, 56(8):1123–1129.

[12] AMBROSO, A., CHALONS, C., COQUEL, F., GODLEWSKI, E. et RAVIART, P.-A. (2005). Couplage de deux systèmes de la dynamique des gaz. *In 17e congrès francais de mécanique*. Université de technologie de Troyes, Troyes.

[13] AUDUSSE, E. et PERTHAME, B. (2005). Uniqueness for scalar conservation laws with discontinuous flux via adapted entropies. *Proc. Roy. Soc. Edinburgh Sect. A*, 135(2):253–265.

[14] AZEVEDO, A. V. et MARCHESIN, D. (1995). Multiple viscous solutions for systems of conservation laws. *Trans. Amer. Math. Soc.*, 347(8):3061–3077.

[15] BACHMANN, F. (2005). *Équations hyperboliques scalaires à flux discontinu*. Thèse de doctorat, Université de Provence.

[16] BACHMANN, F. et VOVELLE, J. (2006). Existence and uniqueness of entropy solution of scalar conservation laws with a flux function involving discontinuous coefficients. *Comm. Partial Differential Equations*, 31(1-3):371–395.

[17] BAITI, P., LEFLOCH, P. G. et PICCOLI, B. (2000). BV stability via generalized characteristics for nonclassical solutions of conservation laws. *In International Conference on Differential Equations, Vol. 1, 2 (Berlin, 1999)*, pages 289–294. World Sci. Publ., River Edge, NJ.

[18] BAITI, P., LEFLOCH, P. G. et PICCOLI, B. (2001). Uniqueness of classical and nonclassical solutions for nonlinear hyperbolic systems. *J. Differential Equations*, 172(1):59–82.

[19] BAITI, P., LEFLOCH, P. G. et PICCOLI, B. (2004). Existence theory for nonclassical entropy solutions of scalar conservation laws. *Z. Angew. Math. Phys.*, 55(6):927–945.

[20] BANDA, M. K., HERTY, M. et KLAR, A. (2006a). Coupling conditions for gas networks governed by the isothermal Euler equations. *Netw. Heterog. Media*, 1(2):295–314 (electronic).

[21] BANDA, M. K., HERTY, M. et KLAR, A. (2006b). Gas flow in pipeline networks. *Netw. Heterog. Media*, 1(1):41–56 (electronic).

[22] BEDJAOUI, N. et LEFLOCH, P. G. (2001). Diffusive-dispersive traveling waves and kinetic relations. III. An hyperbolic model of elastodynamics. *Ann. Univ. Ferrara Sez. VII (N.S.)*, 47:117–144.

[23] BEDJAOUI, N. et LEFLOCH, P. G. (2002a). Diffusive-dispersive traveling waves and kinetic relations. I. Nonconvex hyperbolic conservation laws. *J. Differential Equations*, 178(2):574–607.

[24] BEDJAOUI, N. et LEFLOCH, P. G. (2002b). Diffusive-dispersive travelling waves and kinetic relations. II. A hyperbolic-elliptic model of phase-transition dynamics. *Proc. Roy. Soc. Edinburgh Sect. A*, 132(3):545–565.

[25] BEDJAOUI, N. et LEFLOCH, P. G. (2003). Diffusive-dispersive traveling waves and kinetic relations. IV. Compressible Euler equations. *Chinese Ann. Math. Ser. B*, 24(1):17–34.

[26] BEDJAOUI, N. et LEFLOCH, P. G. (2004). Diffusive-dispersive travelling waves and kinetic relations. V. Singular diffusion and nonlinear dispersion. *Proc. Roy. Soc. Edinburgh Sect. A*, 134(5):815–843.

[27] BENZONI-GAVAGE, S. et COLOMBO, R. M. (2003). An n-populations model for traffic flow. *European J. Appl. Math.*, 14(5):587–612.

[28] BOUCHUT, F. (2004). *Nonlinear stability of finite volume methods for hyperbolic conservation laws and well-balanced schemes for sources*. Birkhäuser.

[29] BOUTIN, B., CHALONS, C., LAGOUTIERE, F. et LEFLOCH, P. G. (2008a). A convergent and conservative schemes for nonclassical solutions based on kinetic relations. I. *Interfaces and Free Bound.*, 10(3):399–421.

[30] BOUTIN, B., CHALONS, C. et RAVIART, P.-A. Existence result for the coupling problem of two scalar conservation laws with Riemann initial data. Submitted.

[31] BOUTIN, B., COQUEL, F. et GODLEWSKI, E. (2008b). Dafermos regularization for interface coupling of conservation laws. *Hyp2006 proceedings, Springer*.

[32] BOUTIN, B., COQUEL, F. et LEFLOCH, P. G. Coupling nonlinear hyperbolic equations (I). Selfsimilar viscous approximation for thin interfaces. In preparation.

[33] BOUTIN, B., COQUEL, F. et LEFLOCH, P. G. Coupling nonlinear hyperbolic equations (II). Resonant interfaces with internal structure. In preparation.

[34] BOUTIN, B., COQUEL, F. et LEFLOCH, P. G. Coupling nonlinear hyperbolic equations (III). A regularization method via thick interfaces. In preparation.

[35] BOUTIN, B., COQUEL, F. et LEFLOCH, P. G. Coupling nonlinear hyperbolic equations (IV). A multidimensional finite volume framework. In preparation.

[36] BÜRGER, R. et KARLSEN, K. H. (2008). Conservation laws with discontinuous flux : a short introduction. *J. Engrg. Math.*, 60(3-4):241–247.

[37] BÜRGER, R., KARLSEN, K. H., RISEBRO, N. H. et TOWERS, J. D. (2004). Well-posedness in BV_t and convergence of a difference scheme for continuous sedimentation in ideal clarifier-thickener units. *Numer. Math.*, 97(1):25–65.

[38] CAETANO, F. (2006). *Sur certains problèmes de linéarisation et de couplage pour les systèmes hyperboliques non linéaires*. Thèse de doctorat, Paris VI.

[39] CANCÈS, C. (2008a). Nonlinear parabolic equations with spatial discontinuities. *NoDEA Nonlinear Differential Equations Appl.*, 15(4-5):427–456.

[40] CANCÈS, C. (2008b). Two-phase flows involving discontinuities on the capillary pressure. *In Finite volumes for complex applications V*, pages 249–256. ISTE, London.

[41] CANCÈS, C. (2008c). *Écoulements diphasiques en milieux poreux hétérogènes : modélisation et analyse*. Thèse de doctorat, Université de Provence.

[42] CHALONS, C. (2002). *Bilans d'entropie discrets dans l'approximation numérique des chocs non classiques. Application aux équations de NS multi-pression 2D et à quelques systèmes visco-capillaires*. Thèse de doctorat, École Polytechnique.

[43] CHALONS, C. (2006). Transport-equilibrium schemes for computing nonclassical shocks. *C. R. Math. Acad. Sci. Paris*, 342(8):623–626.

[44] CHALONS, C. (2007). Numerical approximation of a macroscopic model of pedestrian flows. *SIAM J. Sci. Comput.*, 29(2):539–555 (electronic).

[45] CHALONS, C. (2008). Transport-equilibrium schemes for computing nonclassical shocks. Scalar conservation laws. *Numer. Methods Partial Differential Equations*, 24(4): 1127–1147.

[46] CHALONS, C. et LEFLOCH, P. G. (2001a). A fully discrete scheme for diffusive-dispersive conservation laws. *Numer. Math.*, 89(3):493–509.

[47] CHALONS, C. et LEFLOCH, P. G. (2001b). High-order entropy-conservative schemes and kinetic relations for van der Waals fluids. *J. Comput. Phys.*, 168(1):184–206.

[48] CHALONS, C. et LEFLOCH, P. G. (2003). Computing undercompressive waves with the random choice scheme. Nonclassical shock waves. *Interfaces Free Bound.*, 5(2):129–158.

[49] CHALONS, C., RAVIART, P.-A. et SEGUIN, N. (2008). The interface coupling of the gas dynamics equations. *Quart. Appl. Math.*, 66(4):659–705.

[50] COCKBURN, B., COQUEL, F. et LEFLOCH, P. G. (1995). Convergence of the finite volume method for multidimensional conservation laws. *SIAM J. Numer. Anal.*, 32(3):687–705.

[51] COLOMBO, R. M. et GARAVELLO, M. (2006). A well posed Riemann problem for the *p*-system at a junction. *Netw. Heterog. Media*, 1(3):495–511 (electronic).

[52] COLOMBO, R. M. et ROSINI, M. D. (2005). Pedestrian flows and non-classical shocks. *Math. Methods Appl. Sci.*, 28(13):1553–1567.

[53] COQUEL, F., DIEHL, D., MERKLE, C. et ROHDE, C. (2005). Sharp and diffuse interface methods for phase transition problems in liquid-vapour flows. *In Numerical methods for hyperbolic and kinetic problems*, volume 7 de *IRMA Lect. Math. Theor. Phys.*, pages 239–270. Eur. Math. Soc., Zürich.

[54] COQUEL, F. et LEFLOCH, P. (1990). Convergence de schémas aux différences finies pour des lois de conservation à plusieurs dimensions d'espace. *C. R. Acad. Sci. Paris Sér. I Math.*, 310(6):455–460.

[55] COQUEL, F. et LEFLOCH, P. (1993). Convergence of finite difference schemes for conservation laws in several space dimensions : a general theory. *SIAM J. Numer. Anal.*, 30(3):675–700.

[56] CRANDALL, M. G. et MAJDA, A. (1980). Monotone difference approximations for scalar conservation laws. *Math. Comp.*, 34(149):1–21.

[57] DAFERMOS, C. M. (1973). Solution of the Riemann problem for a class of hyperbolic systems of conservation laws by the viscosity method. *Arch. Rational Mech. Anal.*, 52:1–9.

[58] DAFERMOS, C. M. (1973/74). Structure of solutions of the Riemann problem for hyperbolic systems of conservation laws. *Arch. Rational Mech. Anal.*, 53:203–217.

[59] DAFERMOS, C. M. et DIPERNA, R. J. (1976). The Riemann problem for certain classes of hyperbolic systems of conservation laws. *J. Differential Equations*, 20(1):90–114.

[60] DAL MASO, G., LEFLOCH, P. G. et MURAT, F. (1995). Definition and weak stability of nonconservative products. *J. Math. Pures Appl. (9)*, 74(6):483–548.

[61] DESPRÉS, B. et LAGOUTIÈRE, F. (2001). Contact discontinuity capturing schemes for linear advection and compressible gas dynamics. *J. Sci. Comput.*, 16(4):479–524 (2002).

[62] DIEHL, S. (1996). Scalar conservation laws with discontinuous flux function. I. The viscous profile condition. *Comm. Math. Phys.*, 176(1):23–44.

[63] DIPERNA, R. J. (1983). Convergence of approximate solutions to conservation laws. *Arch. Rational Mech. Anal.*, 82(1):27–70.

[64] DIPERNA, R. J. (1985). Measure-valued solutions to conservation laws. *Arch. Rational Mech. Anal.*, 88(3):223–270.

[65] DUBOIS, F. et LEFLOCH, P. G. (1988). Boundary conditions for nonlinear hyperbolic systems of conservation laws. *J. Differential Equations*, 71(1):93–122.

[66] FAN, H. T. et SLEMROD, M. (1993). The Riemann problem for systems of conservation laws of mixed type. *In Shock induced transitions and phase structures in general media*, volume 52 de *IMA Vol. Math. Appl.*, pages 61–91. Springer, New York.

[67] GALIÉ, T. (2009). *Couplage interfacial de modèles en dynamique des fluides. Application aux écoulements diphasiques*. Thèse de doctorat, Paris VI.

[68] GARAVELLO, M. et PICCOLI, B. (2006a). *Traffic flow in networks*. AIMS book series.

[69] GARAVELLO, M. et PICCOLI, B. (2006b). Traffic flow on a road network using the Aw-Rascle model. *Comm. Partial Differential Equations*, 31(1-3):243–275.

[70] GLIMM, J. (1965). Solutions in the large for nonlinear hyperbolic systems of equations. *Comm. Pure Appl. Math.*, 18:697–715.

[71] GOATIN, P. et LEFLOCH, P. G. (2004). The Riemann problem for a class of resonant hyperbolic systems of balance laws. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 21(6):881–902.

[72] GODLEWSKI, E. (2008). Coupling fluid models. Exploring some features of interfacial coupling. *In Finite volumes for complex applications V*, pages 87–102. ISTE, London.

[73] GODLEWSKI, E., LE THANH, K.-C. et RAVIART, P.-A. (2005). The numerical interface coupling of nonlinear hyperbolic systems of conservation laws. II. The case of systems. *M2AN Math. Model. Numer. Anal.*, 39(4):649–692.

[74] GODLEWSKI, E. et RAVIART, P.-A. (1996). *Numerical approximation of hyperbolic systems of conservation laws*, volume 118 de *Applied Mathematical Sciences*. Springer-Verlag, New York.

[75] GODLEWSKI, E. et RAVIART, P.-A. (2002). Coupling nonlinear hyperbolic systems : mathematical and numerical analysis. *In Finite volumes for complex applications, III (Porquerolles, 2002)*, pages 197–204. Hermes Sci. Publ., Paris.

[76] GODLEWSKI, E. et RAVIART, P.-A. (2004). The numerical interface coupling of nonlinear hyperbolic systems of conservation laws. I. The scalar case. *Numer. Math.*, 97(1):81–130.

[77] GODLEWSKI, E. et SEGUIN, N. (2006). The Riemann problem for a simple model of phase transition. *Commun. Math. Sci.*, 4(1):227–247.

[78] GUELFI, A., BESTION, D., BOUCKER, M., BOUDIER, P., FILLION, P., GRANDOTTO, M., HÉRARD, J.-M., HERVIEU, E. et PÉTURAUD, P. (2007). NEPTUNE : A New Software Platform for Advanced Nuclear Thermal Hydraulic. *Nuclear Science and Engineering*, 156:281–324.

[79] HAYES, B. T. et LEFLOCH, P. G. (1997). Non-classical shocks and kinetic relations : scalar conservation laws. *Arch. Rational Mech. Anal.*, 139(1):1–56.

[80] HAYES, B. T. et LEFLOCH, P. G. (1998). Nonclassical shocks and kinetic relations : finite difference schemes. *SIAM J. Numer. Anal.*, 35(6):2169–2194 (electronic).

[81] HAYES, B. T. et LEFLOCH, P. G. (2000). Nonclassical shocks and kinetic relations : strictly hyperbolic systems. *SIAM J. Math. Anal.*, 31(5):941–991 (electronic).

[82] HELLUY, P. et SEGUIN, N. (2006). Relaxation models of phase transition flows. *M2AN Math. Model. Numer. Anal.*, 40(2):331–352.

[83] HERTY, M. et RASCLE, M. (2006). Coupling conditions for a class of second-order models for traffic flow. *SIAM J. Math. Anal.*, 38(2):595–616 (electronic).

[84] HOU, T. Y., ROSAKIS, P. et LEFLOCH, P. (1999). A level-set approach to the computation of twinning and phase-transition dynamics. *J. Comput. Phys.*, 150(2):302–331.

[85] HURISSE., O. (2006). *Techniques de couplage de modèles hyperboliques en thermohydraulique diphasique*. Thèse de doctorat, Université de Provence.

[86] ISAACSON, E. et TEMPLE, B. (1992). Nonlinear resonance in systems of conservation laws. *SIAM J. Appl. Math.*, 52(5):1260–1278.

[87] ISAACSON, E. et TEMPLE, B. (1995). Convergence of the 2×2 Godunov method for a general resonant nonlinear balance law. *SIAM J. Appl. Math.*, 55(3):625–640.

[88] JOSEPH, K. T. et LEFLOCH, P. G. (1999). Boundary layers in weak solutions of hyperbolic conservation laws. *Arch. Ration. Mech. Anal.*, 147(1):47–88.

[89] JOSEPH, K. T. et LEFLOCH, P. G. (2002a). Boundary layers in weak solutions of hyperbolic conservation laws. II. Self-similar vanishing diffusion limits. *Commun. Pure Appl. Anal.*, 1(1):51–76.

[90] JOSEPH, K. T. et LEFLOCH, P. G. (2002b). Boundary layers in weak solutions of hyperbolic conservation laws. III. Vanishing relaxation limits. *Port. Math. (N.S.)*, 59(4):453–494.

[91] JOSEPH, K. T. et LEFLOCH, P. G. (2007). Singular limits for the Riemann problem : general diffusion, relaxation, and boundary conditions. *C. R. Math. Acad. Sci. Paris*, 344(1):59–64.

[92] KARLSEN, K. H., RASCLE, M. et TADMOR, E. (2007). On the existence and compactness of a two-dimensional resonant system of conservation laws. *Commun. Math. Sci.*, 5(2):253–265.

[93] KLINGENBERG, C. et RISEBRO, N. H. (1995). Convex conservation laws with discontinuous coefficients. Existence, uniqueness and asymptotic behavior. *Comm. Partial Differential Equations*, 20(11-12):1959–1990.

[94] KRUZKOV, S. N. (1970). First order quasilinear equations with several independent variables. *Mat. Sb. (N.S.)*, 81 (123):228–255.

[95] LAGOUTIÈRE, F. (2000). *Modélisation mathématique et résolution numérique de problèmes de fluides compressibles à plusieurs constituants*. Thèse de doctorat, Paris VI.

[96] LAGOUTIÈRE, F. (2008). Stability of reconstruction schemes for scalar hyperbolic conservation laws. *Commun. Math. Sci.*, 6(1):57–70.

[97] LAGOUTIÈRE, F. (2009). Non-dissipative entropic discontinuous reconstruction schemes for hyperbolic conservation laws. Submitted.

[98] LATTANZIO, C. et SERRE, D. (2001). Convergence of a relaxation scheme for hyperbolic systems of conservation laws. *Numer. Math.*, 88(1):121–134.

[99] LEFLOCH, P. et LIU, T.-P. (1993). Existence theory for nonlinear hyperbolic systems in nonconservative form. *Forum Math.*, 5(3):261–280.

[100] LEFLOCH, P. G. (1989). Shock waves for nonlinear hyperbolic systems in nonconservative form. *Institute for Math. and its Appl., Minneapolis*, Preprint # 593.

[101] LEFLOCH, P. G. (1993). Propagating phase boundaries : formulation of the problem and existence via the Glimm method. *Arch. Rational Mech. Anal.*, 123(2):153–197.

[102] LEFLOCH, P. G. (2002). *Hyperbolic systems of conservation laws*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel. The theory of classical and nonclassical shock waves.

[103] LEFLOCH, P. G., MERCIER, J. M. et ROHDE, C. (2002). Fully discrete, entropy conservative schemes of arbitrary order. *SIAM J. Numer. Anal.*, 40(5):1968–1992 (electronic).

[104] LEFLOCH, P. G. et MOHAMMADIAN, M. (2008). Why many theories of shock waves are necessary : kinetic functions, equivalent equations, and fourth-order models. *J. Comput. Phys.*, 227(8):4162–4189.

[105] LEFLOCH, P. G. et ROHDE, C. (2000). High-order schemes, entropy inequalities, and nonclassical shocks. *SIAM J. Numer. Anal.*, 37(6):2023–2060 (electronic).

[106] LEFLOCH, P. G. et ROHDE, C. (2001). Zero diffusion-dispersion limits for self-similar Riemann solutions to hyperbolic systems of conservation laws. *Indiana Univ. Math. J.*, 50(4):1707–1743.

[107] LEFLOCH, P. G. et SHEARER, M. (2004). Non-classical Riemann solvers with nucleation. *Proc. Roy. Soc. Edinburgh Sect. A*, 134(5):961–984.

[108] LEFLOCH, P. G. et THANH, M. D. (2000). Nonclassical Riemann solvers and kinetic relations. III. A nonconvex hyperbolic model for van der Waals fluids. *Electron. J. Differential Equations*, pages No. 72, 19 pp. (electronic).

[109] LEFLOCH, P. G. et THANH, M. D. (2001). Nonclassical Riemann solvers and kinetic relations. I. A nonconvex hyperbolic model of phase transitions. *Z. Angew. Math. Phys.*, 52(4):597–619.

[110] LEFLOCH, P. G. et THANH, M. D. (2002). Non-classical Riemann solvers and kinetic relations. II. An hyperbolic-elliptic model of phase-transition dynamics. *Proc. Roy. Soc. Edinburgh Sect. A*, 132(1):181–219.

[111] LEFLOCH, P. G. et THANH, M. D. (2003). Properties of Rankine-Hugoniot curves for van der Waals fluids. *Japan J. Indust. Appl. Math.*, 20(2):211–238.

[112] LEFLOCH, P. G. et TZAVARAS, A. E. (1996). Existence theory for the Riemann problem for non-conservative hyperbolic systems. *C. R. Acad. Sci. Paris Sér. I Math.*, 323(4):347–352.

[113] LIN, X.-B. et SCHECTER, S. (2003). Stability of self-similar solutions of the Dafermos regularization of a system of conservation laws. *SIAM J. Math. Anal.*, 35(4):884–921 (electronic).

[114] LIU, T. P. (1975). The Riemann problem for general systems of conservation laws. *J. Differential Equations*, 18:218–234.

[115] MERKLE, C. et ROHDE, C. (2006). Computation of dynamical phase transitions in solids. *Appl. Numer. Math.*, 56(10-11):1450–1463.

[116] MERKLE, C. et ROHDE, C. (2007). The sharp-interface approach for fluids with phase change : Riemann problems and ghost fluid techniques. *M2AN Math. Model. Numer. Anal.*, 41(6):1089–1123.

[117] OSHER, S. (1984). Riemann solvers, the entropy condition, and difference approximations. *SIAM J. Numer. Anal.*, 21(2):217–235.

[118] SCHECTER, S. (2004). Existence of Dafermos profiles for singular shocks. *J. Differential Equations*, 205(1):185–210.

[119] SCHECTER, S. (2006). Eigenvalues of self-similar solutions of the Dafermos regularization of a system of conservation laws via geometric singular perturbation theory. *J. Dynam. Differential Equations*, 18(1):53–101.

[120] SCHECTER, S. et SZMOLYAN, P. (2004). Composite waves in the Dafermos regularization. *J. Dynam. Differential Equations*, 16(3):847–867.

[121] SEGUIN, N. et VOVELLE, J. (2003). Analysis and approximation of a scalar conservation law with a flux function with discontinuous coefficients. *Math. Models Methods Appl. Sci.*, 13(2):221–257.

[122] SLEMROD, M. (1983). Admissibility criteria for propagating phase boundaries in a van der Waals fluid. *Arch. Rational Mech. Anal.*, 81(4):301–315.

[123] SLEMROD, M. (1989). The viscosity-capillarity approach to phase transitions. *In PDEs and continuum models of phase transitions (Nice, 1988)*, volume 344 de *Lecture Notes in Phys.*, pages 201–206. Springer, Berlin.

[124] SLEMROD, M. (1995). A comparison of two viscous regularizations of the Riemann problem for Burgers's equation. *SIAM J. Math. Anal.*, 26(6):1415–1424.

[125] SLEMROD, M. et TZAVARAS, A. E. (1989). A limiting viscosity approach for the Riemann problem in isentropic gas dynamics. *Indiana Univ. Math. J.*, 38(4):1047–1074.

[126] SZEPESSY, A. (1989a). An existence result for scalar conservation laws using measure valued solutions. *Comm. Partial Differential Equations*, 14(10):1329–1350.

[127] SZEPESSY, A. (1989b). Measure-valued solutions of scalar conservation laws with boundary conditions. *Arch. Rational Mech. Anal.*, 107(2):181–193.

[128] TARTAR, L. (1979). Compensated compactness and applications to partial differential equations. *In Nonlinear analysis and mechanics : Heriot-Watt Symposium, Vol. IV*, volume 39 de *Res. Notes in Math.*, pages 136–212. Pitman, Boston, Mass.

[129] TEMPLE, B. (1982). Global solution of the Cauchy problem for a class of 2×2 nonstrictly hyperbolic conservation laws. *Adv. in Appl. Math.*, 3(3):335–375.

[130] TRUSKINOVSKY, L. (1987). Dynamics of nonequilibrium phase boundaries in a heat conducting non-linearly elastic medium. *Prikl. Mat. Mekh.*, 51(6):1009–1019.

[131] TRUSKINOVSKY, L. (1993). Kinks versus shocks. *In Shock induced transitions and phase structures in general media*, volume 52 de *IMA Vol. Math. Appl.*, pages 185–229. Springer, New York.

[132] TZAVARAS, A. E. (1996). Wave interactions and variation estimates for self-similar zero-viscosity limits in systems of conservation laws. *Arch. Rational Mech. Anal.*, 135(1):1–60.

[133] ZHONG, X., HOU, T. Y. et LEFLOCH, P. G. (1996). Computational methods for propagating phase boundaries. *J. Comput. Phys.*, 124(1):192–216.

Étude mathématique et numérique d'équations hyperboliques non-linéaires : couplage de modèles et chocs non classiques.

Résumé

Cette thèse concerne l'étude mathématique et numérique d'équations aux dérivées partielles hyperboliques non-linéaires.

Une première partie traite d'une problématique émergente : le couplage d'équations hyperboliques. Les applications poursuivies relèvent du couplage mathématique de plateformes de calcul, en vue d'une simulation adaptative de phénomènes multi-échelles. Nous proposons et analysons un nouveau formalisme de couplage construit sur des systèmes EDP augmentés permettant de s'affranchir de la description géométrique des frontières. Ce nouveau formalisme permet de poser le problème en plusieurs variables d'espace en autorisant l'éventuel recouvrement des modèles à coupler. Ce formalisme autorise notamment à munir la procédure de couplage de mécanismes de régularisation visqueuse utiles à la sélection de solutions discontinues naturelles. Nous analysons alors les questions d'existence et d'unicité dans le cadre d'une régularisation parabolique autosemblable. L'existence est acquise sous des conditions très générales mais de multiples solutions sont susceptibles d'apparaître dès que le phénomène de résonance survient. Ensuite, nous montrons que notre formalisme de couplage à l'aide de modèles EDP augmentés autorise une autre stratégie de régularisation basée sur l'épaississement des interfaces. Nous établissons dans ce cadre l'existence et l'unicité des solutions au problème de Cauchy pour des données initiales L^∞. À cette fin, nous développons une technique de volumes finis sur des triangulations générales que nous analysons dans la classe des solutions à valeurs mesures entropiques de DiPerna.

La seconde partie est consacrée à la définition d'un schéma de volumes finis pour l'approximation des solutions non classiques d'une loi de conservation scalaire basée sur une relation cinétique. Ce schéma présente la particularité d'être *stricto sensu* conservatif contrairement à une approche à la Glimm qui ne l'est que statistiquement. Des illustrations numériques étayent le bien-fondé de notre approche.

Mots-clefs : équations hyperboliques – couplage d'équations – résonance – régularisation de Dafermos – solutions non classiques – méthodes de volumes finis.

Mathematical and numerical study of nonlinear hyperbolic equations: model coupling and nonclassical shocks.

Abstract

This thesis concerns the mathematical and numerical study of nonlinear hyperbolic partial differential equations.

A first part deals with an emergent problematic: the coupling of hyperbolic equations. The pursued applications are linked with the mathematical coupling of computing platforms, dedicated to an adaptative simulation of multi-scale phenomena. We propose and analyze a new coupling formalism based on extended PDE systems avoiding the geometric treatment of the interfaces. In addition, it allows to formulate the problem in a multidimensional setting, with possible covering of the coupled models. This formalism allows in particular to equip the coupling procedure with viscous regularization mechanisms, useful in the selection of natural discontinuous solutions. We analyze existence and uniqueness in the framework of a parabolic regularization *à la Dafermos*. Existence of a solution holds true under very general conditions but failure of uniqueness may naturally arise as soon as resonance occurs at the interfaces. Next, we highlight that our extended PDE framework gives rise to another regularization strategy based on thick interfaces. In this setting, we prove existence and uniqueness of the solutions of the Cauchy problem for initial data in L^∞. The main tool consists in the derivation of a flexible and robust finite volume method for general triangulation which is analyzed in the setting of entropy measure-valued solutions by DiPerna.

The second part is devoted to the definition of a finite volume scheme for the computing of nonclassical solutions of a scalar conservation law based on a kinetic relation. This scheme offers the feature to be *stricto sensu* conservative, in opposition to a Glimm approach that is only statistically conservative. The validity of our approach is illustrated through numerical examples.

Keywords: hyperbolic equations – coupling of equations – resonance – Dafermos regularization – nonclassical solutions – finite volume methods.