



HAL
open science

Analyse et approximation numérique de systèmes hyperboliques de lois de conservation avec termes sources. Application aux équations d'Euler et à un modèle simplifié d'écoulements diphasiques.

Laurent Gosse

► **To cite this version:**

Laurent Gosse. Analyse et approximation numérique de systèmes hyperboliques de lois de conservation avec termes sources. Application aux équations d'Euler et à un modèle simplifié d'écoulements diphasiques.. Analyse numérique [math.NA]. Université Paris Dauphine - Paris IX, 1997. Français. NNT : 97 PA09 0040 . tel-00773175

HAL Id: tel-00773175

<https://theses.hal.science/tel-00773175>

Submitted on 11 Jan 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

pour l'obtention du titre de
DOCTEUR EN SCIENCES
spécialité : Mathématiques Appliquées
arrêté du 23 novembre 1988

Sujet :

ANALYSE ET APPROXIMATION NUMÉRIQUE DE SYSTÈMES HYPERBOLIQUES
DE LOIS DE CONSERVATION AVEC TERMES SOURCES.
APPLICATION AUX ÉQUATIONS D'EULER
ET À UN MODÈLE SIMPLIFIÉ D'ÉCOULEMENTS DIPHASIQUES.

Candidat : Laurent GOSSE

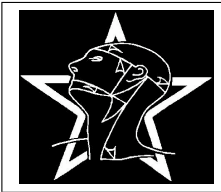
Directeur de recherche : Patrick LE TALLEC

Jury :

Président :	Monsieur Grégoire ALLAIRE
Rapporteur :	Monsieur Alain-Yves LEROUX
Rapporteur :	Monsieur Benoit PERTHAME
Examineur :	Monsieur Guy CHAVENT
Examineur :	Monsieur Patrick LE TALLEC
Examineur :	Monsieur Imad TOUMI

septembre 1997

A Stéphanie,
à mes parents.



Remerciements

Je tiens à remercier Patrick LeTallec de la confiance qu'il m'a témoignée en acceptant de diriger cette thèse, ainsi que pour ses relectures minutieuses du manuscrit.

L'impulsion décisive de cette recherche est due à Alain-Yves LeRoux qui m'a introduit au concept de schéma «équilibré». Je le remercie de sa générosité, de sa constante disponibilité, ainsi que d'avoir accepté de rapporter sur ce travail.

Grégoire Allaire a accepté d'être président de mon jury: je l'en remercie cordialement.

Le soutien financier a été obtenu par l'entremise d'Imad Toumi et Patrick Raymond, chefs successifs du Laboratoire d'Études Thermiques des Réacteurs au Commissariat à l'Énergie Atomique de Saclay. Je leur suis aussi gré de m'avoir confié un sujet de recherche fertile et porteur.

Je remercie Benoit Perthame pour avoir accepté d'être rapporteur de cette thèse, témoignant ainsi de son intérêt pour ce travail. J'apprécie par ailleurs le temps qu'il m'a consacré en discussions tout au long de ces trois années.

Enfin, je remercie Guy Chavent d'avoir accepté spontanément d'être membre du jury, ainsi que pour le soutien et la disponibilité dont il a fait montre durant mon séjour au sein du laboratoire.

Ce travail n'aurait pas abouti sans l'aide de l'équipe du Laboratoire d'Analyse Numérique de l'Université Paris VI, notamment d'Edwige Godlewski. Sa gentillesse et sa disponibilité constantes m'ont été d'un grand secours. La compétence et la sympathie de Frédéric Coquel me furent aussi très précieuses, ainsi que les conseils de Yann Brenier. Je n'oublie pas non plus les chantres des problèmes non-conservatifs, à savoir Florian DeVuyst et Xavier Louis.

J'aimerais maintenant témoigner ma sympathie aux membres du L.E.T.R., en particulier à Sebastien Clerc avec qui j'ai eu le plaisir de travailler pendant deux ans. Je voudrais remercier le spécialiste du *petit lemme d'algèbre linéaire* de s'être investi avec moi dans l'aventure des schémas non-conservatifs. Mes premiers pas dans le domaine de l'approximation numérique des problèmes de la Mécanique des Fluides (et aussi dans l'Internet!) furent guidés par Alain Zelmanse. Je le remercie tout particulièrement de l'amitié qu'il m'a témoignée durant la période de mon Service National. Les ennuis divers et variés générés par les mystères d'Unix furent le plus souvent résolus grâce à l'aide des administrateurs du réseau: Patrick Avice, Laurent Kettela et Emmanuel Mante.

Finalement, je profite de cette occasion pour remercier Eugène Camalet pour m'avoir «coaché» dans l'équipe enseignante de l'Université d'Évry Val d'Éssonne où j'ai eu l'occasion d'acquérir une expérience d'enseignement universitaire pendant deux ans dans d'excellentes conditions. Je remercie aussi Francis Hirsch, directeur du Département de Mathématiques de m'avoir fait immédiatement confiance.

Introduction

Beaucoup d'efforts ont été investis depuis les années 1960 dans le but de mieux comprendre les problèmes posés par les équations de la mécanique des fluides. Ceux-ci ont permis d'obtenir dans certains cas des justifications mathématiques rigoureuses des systèmes d'équations issues de la modélisation physique. D'autre part, grâce au développement à grande échelle des moyens informatiques, le calcul numérique dans ce domaine a connu un essor considérable et plusieurs problèmes commencent à être correctement maîtrisés. Une contribution décisive dans ce secteur a été l'introduction des méthodes de **volumes finis** reconnues pour leur robustesse et leur caractère conservatif.

Typiquement, les problèmes issus de la modélisation des écoulements de fluides compressibles se présentent sous la forme de systèmes d'équations aux dérivées partielles de nature hyperbolique. Chacune d'entre elles exprime la conservation d'une quantité isolée, ou bien son évolution lorsqu'elle est soumise à l'action de phénomènes extérieurs. Ces derniers peuvent être de natures extrêmement diverses: on peut citer l'action de la gravité, le travail des forces de frottement, l'effet d'une géométrie particulière dans laquelle est confiné l'écoulement, les changements de phase, l'apport calorifique dû à un chauffage, etc ... Alors que les mécanismes convectifs s'expriment dans les équations sous la forme de termes de flux, c'est-à-dire des gradients de fonctions des quantités conservées, les phénomènes extérieurs sont modélisés le plus souvent par des **termes sources**, i.e. des termes d'ordre zéro ne comportant pas de dérivées.

Au niveau de l'approximation de la solution de ce type de systèmes monodimensionnels, un outil fondamental est le **problème de Riemann** tant pour une investigation théorique que pour construire un algorithme numérique efficace. Celui-ci consiste à résoudre le système d'équations pour une classe particulière de données initiales, à savoir deux états constants séparés par une discontinuité. L'intérêt d'effectuer cette restriction est de pouvoir obtenir une expression analytique de la solution dans des cas représentatifs (système de l'élasticité, équations d'Euler pour un gaz parfait polytropique par exemple). C'est pratiquement le seul cas de figure en mécanique des fluides où de tels calculs peuvent être menés à bien. Cette approche a permis d'obtenir des résultats significatifs comme la théorie de Lax [57], le théorème d'existence de Glimm ensuite étendu par Liu [65]. Presque simultanément, a été introduite l'une des plus importantes contributions dans le domaine numérique, à savoir le schéma de Godunov [36] consistant à résoudre de façon itérative une suite de problèmes de Riemann dont les solutions élémentaires sont ensuite moyennées localement.

Il est cependant à noter que cette technique perd de son efficacité lorsque l'on traite des systèmes **non-homogènes**, ce qui est précisément le contexte mathématique qui convient aux problèmes munis de termes sources. En effet, la particularité des solutions de Riemann homogènes est leur caractère auto-similaire, c'est-à-dire leur invariance par les dilatations dans le plan espace-temps. C'est donc une façon de ramener un système aux dérivées partielles à un système différentiel ordinaire pour lequel on dispose d'outils puissants de résolution. Malheureusement, cette propriété essentielle disparaît en présence de termes au second membre des équations.

Des auteurs se sont néanmoins engagés dans cette voie en introduisant une extension connue sous le nom de **problème de Riemann généralisé**: on peut citer par exemple Ben-Artzi [6], Glimm *et al.* [34], Le Floch [59], Liu [65], Van Leer [84]. Mais les difficultés persistent du fait du manque de **bonnes propriétés** du modèle et compromettent sévèrement le cadre de travail ainsi obtenu. Les calculs deviennent beaucoup plus pénibles, la structure de la solution est complexe et peu utilisable en vue de construire un schéma numérique efficace et peu coûteux. Malgré tout, des résultats théoriques et numériques furent obtenus grâce à cette approche, particulièrement pour des problèmes issus de l'aéronautique.

Une autre voie a été ensuite explorée: celle-ci repose sur la décomposition d'opérateurs. Concrètement, on procède de la façon suivante: après avoir résolu la solution élémentaire de problèmes de Riemann locaux, on fait agir séparément le système différentiel lié aux termes sources. Il paraît naturel d'espérer la convergence d'une telle approche itérative lorsque l'on répète indéfiniment ces deux étapes élémentaires. Le principal avantage en est de remplacer le calcul d'une solution à la structure complexe par la succession de calculs d'approximations plus simples. Cette approche a le mérite d'être un moyen efficace et peu onéreux pour obtenir une solution approchée du problème non-homogène.

Outre les ennuis quantitatifs provenant des restrictions de stabilité de ces algorithmes, on constate l'apparition d'inconvénients qualitatifs pour chacune de ces deux approches. En effet, au niveau des solutions transitoires,

il est possible de mettre en évidence des anomalies graves: dans [62], LeVeque explique comment les vitesses de propagation des discontinuités peuvent être bouleversées par l'incidence du traitement des termes de réaction dans un contexte d'écoulement d'espèces chimiques réactives. Plus grave, on assiste dans une large variété de cas à la convergence en temps vers des profils stationnaires incorrects comme l'ont exposé (entre autres) Cargo et LeRoux.

Il paraît donc intéressant (du moins l'auteur l'espère-t-il!) d'envisager l'étude de ce type de problèmes, tant au niveau théorique qu'au niveau numérique. L'objectif est double: d'abord, il semble nécessaire d'acquérir une bonne connaissance de la classe de solutions que l'on désire obtenir. Malheureusement, les équations de la mécanique des fluides sont connues pour développer rapidement des singularités rendant ardue la compréhension mathématique de ce type de problèmes (voir par exemple [76]) exception faite des lois scalaires et de certaines classes de systèmes. Ensuite, et c'est sans doute la motivation principale, un nouvel angle d'attaque sur le plan numérique est souhaitable, si possible susceptible de corriger quelques uns des défauts précédemment relevés.

Le décor étant planté, nous donnons maintenant le plan de cette thèse.

Partie I Celle-ci est consacrée à l'étude théorique de quelques problèmes simples où l'influence de termes sources est intéressante à détailler. Le chapitre 1 introduit les éléments mathématiques de base pour ce type d'équations et donne un aperçu des complications induites par un second membre très simple dans le problème de Riemann pour l'équation de Burgers. Dans les chapitres 2 et 3, on établit l'existence et l'unicité de la solution L^∞ du problème de Cauchy pour une équation scalaire en présence d'une large classe de termes sources par la méthode de la viscosité évanescence et à l'aide des entropies de Kruzkov. On insiste sur les différences avec le cas homogène. L'argument clef est d'assurer une estimation uniforme en amplitude sur la suite d'approximations régulières afin de pouvoir appliquer la théorie de la compacité par compensation de Murat et Tartar. Cette voie laisse entrevoir une extension possible au cas de systèmes simples en suivant les idées de DiPerna. On a choisi de traiter un système très simplifié issu de la modélisation des écoulements diphasiques étudié auparavant par Toumi [81]. Il apparaît alors difficile de garantir encore une borne uniforme sur la taille des approximations visqueuses en présence d'un terme source car celui-ci interfère avec les flux convectifs. Il existe néanmoins quelques cas où il est possible d'établir un résultat de convergence en faisant des hypothèses sur la loi de pression.

Partie II On traite ici davantage de problèmes liés aux approximations numériques implicites des lois de conservation scalaires avec second membre. On se place donc dans le cadre théorique de la partie précédente. Au chapitre 1, on considère les discrétisations implicites les plus simples possibles pour ce type de problèmes (les θ -méthodes) et on établit une condition suffisante de type CFL sous laquelle le schéma numérique admet à chaque itération une unique solution. Cette restriction sur le maillage peut s'interpréter en termes simples: les chocs ne doivent pas être trop importants et la solution ne doit pas s'approcher trop près des équilibres répulsifs du terme source. Au chapitre 2, on rappelle les méthodes de quadrature plus sophistiquées, pour lesquelles on constate l'apparition d'oscillations parasites bien en-dessous de la limite prescrite par les résultats du chapitre précédent. En s'inspirant du travail de Shu et Osher [78] on complète cette approche en mettant en évidence les étapes explicites inhérentes aux méthodes implicites d'ordre élevé. Ceci entraîne dans la majorité des cas des restrictions supplémentaires sur le maillage numérique.

Partie III A la lecture des conclusions de la partie précédente, il semble que les techniques classiques de traitement numérique des termes sources ne soient pas optimales. On va donc chercher à construire un nouveau moyen d'approcher efficacement les solutions de ce type d'équations. Au chapitre 1, dans un travail commun avec A.-Y. LeRoux, on se restreint au cas scalaire pour lequel on construit un solveur de Riemann intégrant l'action du terme source et maintenant la propriété d'auto-similarité. Ceci est rendu possible grâce à une reformulation du second membre en produits non-conservatifs. Dans le contexte simple des lois scalaires, il est possible de lever l'ambiguïté liée à cette écriture et un schéma numérique s'en déduit naturellement. On montre en outre un résultat de convergence vers la solution entropique. Dans cette approche, le terme source n'a aucune influence sur le maillage qui n'est prescrit que par la condition CFL classique. Malheureusement, l'extension au cas des systèmes n'est pas limpide: c'est pourquoi on présente au chapitre 2 une recette basée sur des produits simplifiés de Vol'Pert dans le contexte particulier du système diphasique traité en Partie I. On observe encore une excellente robustesse pour cette généralisation. La voie la plus naturelle pour étendre le

formalisme non-conservatif du chapitre 1 au cas des systèmes hyperboliques généraux est exposée au chapitre 3. Elle est fondée sur la théorie des produits non-conservatifs proposée par DalMaso, Le Floch et Murat. On introduit donc une régularisation générique capable de fournir un solveur présentant les mêmes qualités que dans le cas scalaire. Ceci est exploité grâce aux travaux de Toumi et DeVuyst sur les linéarisations de type Roe dans le cas non-conservatif. On observe un gain de précision très net sur les profils stationnaires pour des cas-tests difficiles. D'autres expérimentations numériques ont été menées à bien par S. Clerc qui a collaboré à ce travail.

Partie IV Malgré les bons résultats de l'approche linéarisée, on déplore toujours les problèmes liés aux restrictions de stabilité numérique. Pour pallier à cela, on décide au chapitre 1 d'unifier les techniques développées aux chapitres 2 et 3 de la partie III. On introduit alors un schéma à décomposition de flux de type VanLeer [83] où les termes sources sont traités grâce aux relations de saut généralisées induites par la formulation non-conservative. On obtient de ce fait un algorithme robuste et fiable capable de fournir des profils stationnaires de bonne qualité. Les tests numériques couvrent une large plage d'applications: écoulements de gaz parfait dans une tuyère, en symétrie cylindrique, écoulements diphasiques avec pertes de charge, effets de recondensation, etc ... On présente aussi quelques résultats sur les systèmes relaxés pour lesquels on dispose d'une théorie satisfaisante [14]. Le chapitre 2 est consacré à une incursion dans le domaine des systèmes non-strictement hyperboliques. Le problème générique est proposé par l'aéronautique et consiste en un écoulement transsonique dans une tuyère divergente. On effectue une analyse théorique du système stationnaire afin de tester les schémas non-conservatifs dans ce cas. Enfin, le chapitre 3 est consacré aux applications bidimensionnelles: celles-ci s'effectuent sur la version 2D du modèle diphasique simplifié en présence d'effets combinés de pertes de charge singulières et de recondensation.

This thesis is devoted to the study of nonlinear hyperbolic systems of conservation laws with source terms allowed to become stiff. Applications come for the largest part from physical problems; for example Burgers or Buckley-Leverett equations and Euler-type systems endowed with pressure laws for a perfect gas or a simplified two-phase mixture.

The first part deals with theoretical aspects of this subject. We focus on the one-dimensional case and consider the scalar equation and a special case of a reactive two-phase flow system. The main issue one has to manage with when establishing the strong convergence of a sequence of approximations by means of the compensated compactness theory in such cases is to find invariant regions in the phases space. For the scalar problem, it is sufficient that the source term preserves an uniform bound on the L^∞ norm of the sequence. But it is less obvious for a system since invariant zones are rather complicated and may be perturbed due to the form of some forcing terms.

The second part is concerned with stability criterions for numerical schemes involving sophisticated implicit time quadrature formulas. We first present an analysis of a half-discretization (*method of lines*) and derive sufficient conditions for the differential system to be well-posed. This brings quite natural restrictions to ensure the stability of a one-step implicit scheme: the time-step must be small in highly compressive regimes and near repulsive points of the source term. A total-variation analysis is then carried out to derive second order schemes free from oscillations around shocks for the largest possible time-steps. Numerical experiments are shown on fully nonlinear scalar problems.

The third part is a direct sequel in order to fix some of the difficulties encountered before. The main objective is to build a new discretization which may handle an arbitrary stiff right-hand-side without any influence (except the usual CFL condition) on the time-step. Following the ideas of J. Greenberg and A.-Y. LeRoux, we first derive a Godunov-type scalar scheme and get a convergence result with the help of BV estimates towards the entropy solution. Since the main feature is to handle the sources as non-conservative products, we decide to use the formalism proposed by G. DalMaso, P.G. LeFloch and F. Murat to extend these ideas to inhomogeneous systems. Following I. Touni, we introduce Roe-type matrices along regularizing paths to obtain approximations showing nice properties. In fact, this is what is called *well-balanced* schemes, i.e. schemes which preserve theoretical first integrals of the motion. Numerical tests are based on the Euler system with geometrical source terms.

The fourth (and last) part concerns the introduction of this non-conservative formulation in a flux-splitting type scheme. The aim is to build a very robust, well-balanced and easy-to-implement scheme. The generalized jump relations coming from the sources are detailed and the scheme is tested in a very wide range of problems: nozzle flows, two-phase with chemistry and damping, relaxation systems ... A chapter is also devoted to resonant cases: following Majda, we investigate the stationary problem to validate our results. This work culminates with two-dimensional two-phase flow approximations which involve rather complicated non-conservative relations treated by a convergent iterative process.

Table des matières

I	Aspects théoriques de problèmes hyperboliques avec termes sources	7
1	Solutions faibles des lois scalaires	11
1.1	Lignes caractéristiques et explosion en temps fini	12
1.1.1	Généralités	12
1.1.2	Exemple: l'équation de Burgers amortie	12
1.2	Relations de Rankine-Hugoniot et perte d'unicité	13
1.2.1	Notion de solution faible	13
1.2.2	Multiplicité des solutions faibles	14
1.3	Limite visqueuse et critère d'entropie	15
1.3.1	Origine parabolique	15
1.3.2	Notion de solution entropique	15
1.4	Le problème de Riemann	16
1.4.1	Chocs admissibles et ondes de détente	17
1.4.2	Résolution explicite	17
2	Méthode de viscosité et existence	19
2.1	Problème parabolisé	20
2.1.1	Rappels généraux	20
2.1.2	Lemme du point fixe	21
2.1.3	Principe du maximum	22
2.2	Compacité par compensation	23
2.2.1	Présentation de la théorie	24
2.2.2	Emergence de la solution globale L^∞	25
2.3	Résultats d'entropie	28
2.3.1	Estimations d'entropie	28
2.3.2	Viscosité et entropie dans le cas non-homogène	29
2.4	Conclusion	31
3	Entropies de Kruzkov et unicité	33
3.1	L'inégalité fondamentale	34
3.1.1	Entropies au sens faible	34
3.1.2	Formulation différentielle	35
3.2	Le théorème de continuité	38
3.2.1	Formulation intégrale et continuité L^1_{loc}	38
3.2.2	Le résultat final	41
4	Existence de solutions pour le modèle diphasique équilibré	43
4.1	Formulation lagrangienne	44
4.1.1	Changement de variables	44
4.1.2	Propriétés élémentaires	45
4.2	Le problème de Cauchy	46
4.2.1	Expression des entropies	46
4.2.2	Estimation en amplitude	47
4.2.3	Compacité et existence de solutions	50

II	Méthodes de quadrature implicites adaptées aux lois scalaires	53
5	Stabilité linéaire des méthodes implicites à un pas - Cas des solutions régulières	57
5.1	Généralités	58
5.2	Analyse de Fourier et stabilité dans le cas linéaire	59
5.2.1	Écriture du θ -schéma	60
5.2.2	Stabilité linéaire au sens de Von Neumann	61
5.3	Stabilité dans le cas non-linéaire	62
5.3.1	Un cadre théorique adapté aux problèmes raides	62
5.3.2	L'estimation fondamentale	64
5.3.3	Domaine invariant et estimation globale	65
5.3.4	Conséquences numériques	66
5.4	Application au transport des fractions massiques	67
5.4.1	Choix du flux numérique	67
5.4.2	Inversibilité de la partie implicite	68
5.4.3	Exemples numériques	69
6	Stabilité des méthodes de Runge-Kutta implicites - Cas des ondes discontinues	71
6.1	Stabilité linéaire des routines de quadrature	72
6.1.1	Exemple: la décroissance exponentielle	72
6.1.2	La fonction de stabilité	72
6.1.3	Méthodes implicites de Runge-Kutta (cas linéaire)	73
6.1.4	Méthodes de Runge-Kutta semi-implicites (cas non-linéaire)	74
6.1.5	Exemples numériques et premier bilan	75
6.2	Stabilité en variation totale et méthode des lignes	78
6.2.1	Cadre théorique: les lois scalaires 1D	78
6.2.2	Variation totale des approximations de type Runge-Kutta	78
6.2.3	Première application: les θ -méthodes	82
6.2.4	Les méthodes A-stables d'ordre 2 à 2 pas	83
6.2.5	Stabilité en variation totale et méthodes d'ordre 2	84
6.2.6	Nouveaux essais numériques	85
6.3	Conclusion: méthodes de quadrature adaptées	86
III	Introduction d'un solveur hyperbolique vraiment non-homogène	91
7	Solveur scalaire θ-équilibre et convergence	95
7.1	Construction et convergence du schéma	96
7.1.1	Généralités	96
7.1.2	Problème auxiliaire et schéma de base	96
7.1.3	Principe du maximum discret	97
7.1.4	Analyse et convergence du schéma	99
7.1.5	Résultats numériques	103
7.2	Solveur de Riemann θ -équilibre général	105
7.2.1	Limite d'un problème régularisé	105
7.2.2	Écriture du schéma θ -équilibre général	107
7.2.3	Perte de bijectivité: l'asymptote horizontale	107
8	Première extension aux systèmes réactifs	109
8.1	Schémas à décomposition de flux	110
8.1.1	Construction traditionnelle	110
8.1.2	Fonctions de flux particulières	110
8.1.3	Exemple: le système d'Euler	111
8.1.4	Traitement des termes sources	112
8.2	Décomposition de flux vraiment non-homogène	112
8.2.1	Formulation du problème étudié	112

8.2.2	Réécriture des termes sources du système diphasique	114
8.2.3	Choix des relations de Rankine-Hugoniot stationnaires	114
8.2.4	Décomposition de flux modifiée pour le système diphasique	116
8.2.5	Principe du maximum pour la concentration massique	117
8.3	Essais numériques: termes de recondensation et de frottement pariétal	118
8.3.1	Ecriture précise du problème étudié	118
8.3.2	Relations de saut approchées	118
8.3.3	Exemples de cas-tests pour le système diphasique	119
9	Problèmes non-homogènes et formalisme non-conservatif - Approche linéarisée	123
9.1	Formulation faible du solveur non-conservatif	124
9.1.1	Définition d'un chemin \mathbb{J} équilibre $_{\mathcal{L}}$ scalaire	124
9.1.2	Le problème de Riemann	128
9.1.3	Extension aux systèmes hyperboliques	128
9.2	Formulation d'un solveur approché de type Roe	129
9.2.1	Structure de la linéarisée non-conservative	130
9.2.2	Moyenne \mathbb{J} équilibre $_{\mathcal{L}}$ de certains termes sources	130
9.2.3	Ecriture du schéma non-conservatif	132
9.3	Premier cas-test: tuyère quasi-1D	133
9.3.1	Mise en équations du phénomène	133
9.3.2	Relations de saut et quantités invariantes	133
9.3.3	Matrice de Roe associée $A_{\Phi}(V_L, V_R)$	135
9.3.4	Expression des flux numériques	135
9.3.5	Essai numérique: la tuyère supersonique divergente	136
9.4	Second cas-test: pertes de charge singulières	139
9.4.1	Modélisation proposée	139
9.4.2	Relations de saut et quantités invariantes	139
9.4.3	Matrice de Roe associée $A_{\Phi}(V_L, V_R)$	140
9.4.4	Résultats numériques	140
9.5	Conclusion	143
IV	Développement d'un schéma à décomposition de flux \mathbb{J}équilibre$_{\mathcal{L}}$	147
10	Méthodes de décomposition de flux et produits non-conservatifs	151
10.1	Schémas \mathbb{J} équilibre $_{\mathcal{L}}$ et décomposition de flux	152
10.1.1	Introduction des termes non-conservatifs	152
10.1.2	Ecriture du schéma proposé	153
10.1.3	Interprétation en problèmes de Riemann élémentaires	153
10.1.4	Propriétés du schéma ainsi obtenu	154
10.2	Les écoulements quasi-monodimensionnels	156
10.2.1	Expression des courbes d'équilibre	156
10.2.2	Ecriture du schéma pour ce type de systèmes	158
10.2.3	Cas-tests numériques	159
10.3	Le système diphasique équilibré	165
10.3.1	Expression des solutions stationnaires	165
10.3.2	Ecriture du schéma pour ce système	166
10.3.3	Cas-tests numériques	167
10.4	Le cas des termes de relaxation	171
10.4.1	Introduction à la théorie des systèmes relaxés	171
10.4.2	Application: le modèle diphasique équilibré	172
10.4.3	Formulation non-conservative associée	175
10.4.4	Ecriture et comportement du schéma \mathbb{J} équilibre $_{\mathcal{L}}$ en régime relaxé	176
10.5	Conclusion	177

11 Un mot sur les cas de résonance non-linéaire	179
11.1 Exemple théorique: calcul d'un permanent transsonique	180
11.1.1 Généralités	180
11.1.2 Méthode de résolution	180
11.1.3 Construction d'un permanent transsonique dans un profil divergent	181
11.2 Experiences numériques en régime résonant	185
11.2.1 Solution exacte du cas-test étudié	185
11.2.2 Comportement du schéma de type Roe	186
11.2.3 Comportement du schéma à décomposition de flux	189
11.3 Conclusion	189
12 Extension à des problèmes bidimensionnels	193
12.1 Introduction aux problèmes hyperboliques bidimensionnels	194
12.1.1 Le cas des lois scalaires	194
12.1.2 Extension aux systèmes quasi-linéaires	195
12.2 Traitement non-conservatif des sources dans un contexte 2D	196
12.2.1 Introduction générale de la méthode	196
12.2.2 Un exemple issu de la modélisation diphasique	197
12.2.3 Relations de saut non-conservatives	198
12.2.4 Ecriture du schéma proposé	200
12.3 Validation par des cas-tests numériques	201
12.3.1 Cas-test pseudo-1D	201
12.3.2 Frottement uniforme dans un canal	202
12.3.3 Pertes de charge dans un canal pourvu de deux demi-grilles	206
12.3.4 Effets combinés de recondensation et de perte de charge	210
A Le problème de Riemann pour le modèle diphasique équilibré	219
A.1 Hyperbolicité du système	220
A.2 Expression des courbes d'ondes	220
B Méthodes multipas stables en variation totale	223
B.1 Généralités	224
B.2 Représentation et stabilité des méthodes à 2 pas	225
C Rappels sur la définition des produits non-conservatifs	227
C.1 Interprétation des produits non-conservatifs comme mesures bornées	228
C.1.1 L'espace BV	228
C.1.2 Mesure de Borel associée	228
C.1.3 Deux propriétés importantes	229
C.2 Stabilité faible des produits non-conservatifs	230
C.2.1 Résultats forts de stabilité	230
C.2.2 Exemple: la fonction de Heaviside	230
C.2.3 Condition suffisante de convergence en loi	230
C.3 Application aux systèmes hyperboliques écrits sous forme non-conservative	231
C.3.1 Notion de solution faible	231
C.3.2 Conditions de saut généralisées	231
C.3.3 Structure du problème de Riemann	232
C.3.4 Matrices linéarisées de type Roe	232
D Une alternative pour les termes sources linéaires en ρ	233
D.1 Une formulation équivalente homogène	234
D.2 Application numérique	235
D.2.1 Linéarisée de type Roe	235
D.2.2 Etude d'un cas-test	236

E	Etude de l'approximation des relations de saut non-conservatives	239
E.1	Analyse du système différentiel 2×2	240
E.1.1	Convergence de la suite d'approximations	240
E.1.2	Estimation de l'erreur à la limite	242
E.2	Analyse du problème scalaire sur une variété	242
E.2.1	Convergence de la suite d'approximations	243
E.2.2	Estimation de l'erreur à la limite	244

Partie I

Aspects théoriques de problèmes hyperboliques avec termes sources

Résumé

L'objet de cette première partie est l'étude de certains aspects théoriques des problèmes hyperboliques munis de seconds membres non-linéaires. On commence par rappeler les phénomènes d'explosion en temps fini pour ce type d'équations. On se restreint aux cas où la non-linéarité du terme source est suffisamment faible pour maintenir une solution bornée en chaque instant puisque la solution et sa dérivée spatiale sont à priori susceptibles de devenir infinies. La résolution complète d'un problème de Riemann non-homogène est traitée dans le contexte des solutions faibles et les complications liées au second membre sont soulignées. Les chapitres suivants sont consacrés à la construction de la solution d'une loi scalaire monodimensionnelle munie d'un terme source très général par la méthode de viscosité évanescence. Bien que mettant en oeuvre des techniques devenues classiques (la compacité par compensation de L. Tartar [80] et les inégalités d'entropie de S.N. Kruzkov [53]), cette démonstration ne semble pas exister telle quelle dans la littérature, notamment en ce qui concerne les estimations du terme source. Le point central est d'imposer au second membre de respecter une borne uniforme en amplitude sur la suite d'approximations; cette restriction très proche des équations différentielles ordinaires est suffisante pour passer à la limite dans les termes non-linéaires. On obtient aussi une relation de dépendance continue modifiée de la solution par rapport aux données initiales: ceci montre qu'il n'est désormais généralement plus possible d'introduire un formalisme de semi-groupe de contractions comme dans le cas homogène. On poursuit par l'étude d'un système à trois équations issu de la modélisation des écoulements diphasiques sans écarts de vitesse. Sa formulation en variables de Lagrange révèle une structure particulière lui octroyant un grand nombre d'entropies. On décide d'exploiter cela grâce à la méthode due à DiPerna [25] pour le p -système. Malheureusement, ceci n'est possible que lorsque l'on peut trouver une région invariante pour le système régularisé, ce qui amène inexorablement des restrictions sur les lois de pression susceptibles d'être choisies. De plus, il n'est pas immédiat d'assurer que la présence de termes sources ne détruit pas cette zone privilégiée de l'espace des états. Cette dernière étape est très semblable au cas scalaire, mais le nombre d'équations complique les expressions mises en jeu; en effet, au vu des termes de flux, les régions globalement invariantes ne sont pas de forme simple (rectangles, parallélépipèdes, sphères ...) et celles-ci sont susceptibles d'être altérées par le second membre. Il est néanmoins possible de conclure pour certaines classes de systèmes. Pour terminer, le problème de Riemann pour ce système est étudié dans le cas homogène.

Chapitre 1

Solutions faibles des lois scalaires

Bien que les problèmes hyperboliques réalistes se présentent quasiment toujours sous la forme de *systèmes d'équations* (Euler, p -système, élasticité non-linéaire ...), il est commode d'introduire leur théorie en s'intéressant tout d'abord à des lois scalaires monodimensionnelles. En effet, leur structure mathématique encore suffisamment simple contient déjà beaucoup de traits communs avec des problèmes plus complexes. C'est ce que nous allons présenter ici.

1.1 Lignes caractéristiques et explosion en temps fini

1.1.1 Généralités

L'équation canonique s'écrit sous la forme suivante:

$$\begin{cases} \partial_t u + \partial_x f(u) = g(u) \\ u(x, 0) = u_0(x) \\ (x, t) \in (\mathbb{R} \times \mathbb{R}^+) \end{cases} \quad (1.1)$$

Il s'agit d'une *équation aux dérivées partielles* (EDP), et plus précisément d'un *problème de Cauchy*, c'est-à-dire que l'on désire connaître la valeur de $u(x, t)$ pour chaque instant t , et en chaque point x de l'axe réel à partir de la connaissance de la donnée initiale u_0 en chacun de ces points. L'équation (1.1) est appelée une *loi de conservation* puisqu'elle signifie que la variation de l'inconnue u entre t et $t + \varepsilon$ est égale à la différence du flux $f(u)$ de cette quantité entre x et $x + \eta$ modifié par l'action de forces extérieures modélisées par le *terme source* $g(u)$. On dit que ce problème admet une solution forte, ou classique, si pour tout couple (x, t) , on peut trouver une fonction $u \in C^1(\mathbb{R} \times \mathbb{R}^+)$ vérifiant la condition (1.1) point par point. C'est le cas de figure le plus immédiat auquel on pourrait s'attendre. Malheureusement, il est maintenant bien connu que tel n'est pas le cas: la dérivée spatiale $\partial_x u$ devient infinie au bout d'un temps fini dans une grande partie des cas. C'est la naissance d'une solution-choc discontinue. C'est ce phénomène qui rend difficile l'étude de ces équations puisqu'il est tout d'abord nécessaire de définir *en quel sens* il faut parler de solution pour (1.1), ce qui revient à dire: *comment définir la dérivée d'une fonction discontinue ?* Il faut tout de même souligner que quelques auteurs se ont intéressés aux solutions fortes de problèmes hyperboliques, comme par exemple [64]. Néanmoins, nous ne retiendrons ici que le modèle général pour lequel ce type de solutions n'existent pas dans la majorité des cas, ceci quelque soit la régularité de u_0 à l'origine des temps.

1.1.2 Exemple: l'équation de Burgers amortie

Voyons sur un cas simple le phénomène d'explosion typique des problèmes hyperboliques non-linéaires.

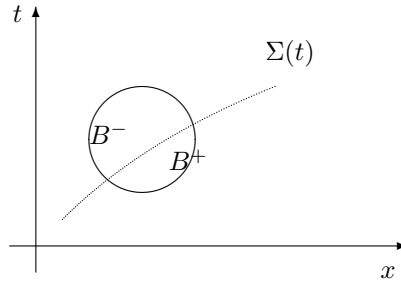
$$\begin{cases} \partial_t u + u \partial_x u = \alpha u \\ u_0(x) = \begin{cases} 1 & \text{pour } x \leq 0 \\ 1 - x & \text{pour } x \in]0, 1[\\ 0 & \text{pour } x \geq 1 \end{cases} \end{cases} \quad (1.2)$$

Comme souvent dans les problèmes de propagation d'ondes, il est utile d'introduire les *lignes caractéristiques* de l'équation (1.2). Ce sont les courbes $t \mapsto X(t)$ solutions de la condition différentielle:

$$\left. \begin{array}{l} \dot{X}(t) = u(X(t), t) \\ X(0) = x_0 \in \mathbb{R} \end{array} \right\} \text{ d'où } \frac{du}{dt}(X(t), t) = (\partial_t u + \dot{X} \partial_x u)(X(t), t) = \alpha u(X(t), t) \quad (1.3)$$

Par conséquent, il est maintenant clair que l'inconnue *voyage le long des courbes* $X(t)$, t appartenant à l'intervalle maximal pour lequel la définition (1.3) garde un sens. L'expression de $X(t)$ se trouve facilement par quadrature: $X(t) = \frac{u(x_0)}{\alpha} (e^{\alpha t} - 1) + x_0$. Remarquons déjà une différence importante avec le cas homogène $g \equiv 0$: les caractéristiques $t \mapsto X(t)$ *ne sont plus des droites du plan* (x, t) . D'autre part, la solution $u(X(t), t)$ n'est plus constante le long de ces directions de propagation: elle varie selon l'équation différentielle ordinaire $\dot{u} = g(u)$. Pour $x_0 < 1$, $t \mapsto X(t)$ est une fonction strictement croissante, alors que pour $x_0 \geq 1$, elle est constamment égale à x_0 . Ces deux familles de courbes vont donc se couper dans le plan (x, t) , et la solution $u(X(t), t)$ va devenir multi-valuée¹ en ces points d'intersection $((x_0, t) \mapsto (X(t), t))$ n'est plus un difféomorphisme de $\mathbb{R} \times \mathbb{R}^+$

1. C'est là le mécanisme d'explosion *géométrique* défini dans [2]

Figure 1.1: Relations de saut à travers Σ

pour t assez grand). Pour l'exemple choisi (1.2), le temps critique t_c où deux caractéristiques issues de $x_0 = 0$ et $x_0 = 1$ se coupent en $X = 1$ vaut: $t_c = \frac{1}{\alpha} \ln |1 + \alpha|$. En ce point, la dérivée $\partial_x u$ est bien sûr incompatible avec la définition classique de limite du taux de variation. Il devient donc nécessaire d'introduire une solution dans un sens *faible*. Il convient tout de même de remarquer que ce phénomène est avant tout d'origine convective, et qu'un terme source dissipatif peut parfois arranger les choses: si $\alpha < 0$, $|\alpha|$ assez grand, alors l'amortissement de u le long des caractéristiques sera tel que celles-ci vont devenir constantes et n'atteindront pas l'abscisse $X = 1$. L'apparition de discontinuités sera donc inhibée. A l'opposé de ce point de vue, la présence d'une *forte* non-linéarité au second membre (par ex. $g(u) = u^4$ en dimension 1 d'espace) peut introduire un type d'explosion *différentielle* et c'est un problème intéressant de déterminer quel mécanisme se déclenche en premier (voir [2] pp22/24).

1.2 Relations de Rankine-Hugoniot et perte d'unicité

1.2.1 Notion de solution faible

Pour continuer l'étude de (1.1), il devient nécessaire d'élargir le concept de dérivée d'une fonction au besoin discontinue. Pour cela, il est commode de se plonger dans l'espace des distributions \mathcal{D}' de Laurent Schwartz dans lequel il est possible de définir des dérivées au sens faible (voir [10], par ex.).

Définition I.1 La distribution $u \in \mathcal{D}'(\mathbb{R} \times \mathbb{R}^+)$ est dite solution de (1.1) au sens faible si elle vérifie, pour toute $\varphi \in C_0^\infty(\mathbb{R} \times \mathbb{R}^+)$, la condition:

$$\int_{\mathbb{R} \times \mathbb{R}^+} [u \partial_t \varphi + f(u) \partial_x \varphi + g(u) \varphi](x, t) dx dt + \int_{\mathbb{R}} u_0(x) \cdot \varphi(x, 0) dx = 0 \quad (1.4)$$

Cette définition étend donc la notion de solution classique pour des fonctions discontinues. Si d'aventure, u est suffisamment régulière, il est immédiat que la formule d'intégration par parties donne:

$$\forall \varphi \in C_0^\infty(\mathbb{R} \times \mathbb{R}^+), \int_{\mathbb{R} \times \mathbb{R}^+} [\partial_t u + \partial_x f(u) - g(u)] \varphi(x, t) dx dt = 0$$

Ce qui implique donc (1.1) au sens classique. Il reste maintenant à éclaircir ce qui se passe le long d'une ligne de discontinuité si la pertinence de ces singularités est admise par (1.4).

Soit $t \mapsto \Sigma(t)$ une courbe du plan (x, t) le long de laquelle u est discontinue, i.e.

$$\lim_{\substack{\epsilon \rightarrow 0 \\ \epsilon > 0}} \underbrace{u(\Sigma(t) - \epsilon, t)}_{u^-(t)} - \underbrace{u(\Sigma(t) + \epsilon, t)}_{u^+(t)} \neq 0$$

Alors sa pente $\sigma(t) \stackrel{def}{=} \dot{\Sigma}(t)$ vérifie le résultat suivant:

Théorème I.1 (Relations de Rankine-Hugoniot)

Le long d'une courbe de discontinuité, on a la condition:

$$\sigma(t) [u^+(t) - u^-(t)] = [f(u^+(t)) - f(u^-(t))] \quad (1.5)$$

Démonstration

Soit M un point sur Σ , et B une boule centrée sur M . La condition (1.4) donne:

$$\int_B [u \partial_t \varphi + f(u) \partial_x \varphi + g(u) \varphi](x, t) dx dt = 0$$

Ecrivons que $B = B^+ \cup B^-$ où B^+ et B^- sont disjointes car séparées par Σ (cf Fig 1.1). Alors l'égalité précédente est vérifiée sur chaque B^\pm . Néanmoins, sur chaque ouvert B^\pm , la solution u est régulière, donc la formule de Green permet d'écrire:

$$\int_{B^-} [\partial_t u + \partial_x f(u) - g(u)] \varphi(x, t) dx dt + \int_{B^+} [\partial_t u + \partial_x f(u) - g(u)] \varphi(x, t) dx dt + \int_{\Sigma \cap B} \varphi < U^-(t), d\vec{n} > - \int_{\Sigma \cap B} \varphi < U^+(t), d\vec{n} > = 0$$

où $\vec{n} = \begin{pmatrix} -\sigma(t) \\ 1 \end{pmatrix}$ désigne la normale orientée de B^- et $U^\pm(t) = \begin{pmatrix} u \\ f(u) \end{pmatrix}((\Sigma(t), t) \pm \vec{n})$. D'autre part, comme u est solution classique loin de la discontinuité, $g(u) \in L^\infty$ et que $\Sigma \cap B$ est de mesure nulle, on déduit que seuls comptent les deux derniers termes, c'est-à-dire:

$$\int_{\Sigma \cap B} \varphi < U^+(t) - U^-(t), d\vec{n} > = 0$$

Par définition de \vec{n} , cette relation est équivalente à (1.5). \square

Remarque Cette relation est à interpréter comme une équation différentielle, puisque les paliers situés de part et d'autre de la discontinuité varient en fonction du temps du fait de la présence du second membre $g(u)$.

1.2.2 Multiplicité des solutions faibles

La condition de saut (1.5) pourrait être un moyen simple de caractériser les solutions discontinues de (1.1). En effet, dans les régions régulières de u , on impose l'EDP au sens habituel, et on vérifie (1.5) à chaque point de discontinuité. Malheureusement, ceci est encore insuffisant pour définir de façon satisfaisante une solution faible. En effet, en acceptant des sauts via (1.4), on a introduit une très forte non-unicité dans notre problème. Pour s'en convaincre, reprenons l'EDP (1.2) avec cette fois la donnée initiale particulière $u_0 \in L^\infty(\mathbb{R})$:

$$\begin{cases} u_L \text{ pour } x < 0 \\ u_R \text{ pour } x > 0 \end{cases} \quad (1.6)$$

Alors, une solution faible de (1.4) est:

$$u(x, t) = \begin{cases} u_L(t) = u_L e^{at} \text{ pour } x < \Sigma(t) \\ u_R(t) = u_R e^{at} \text{ pour } x > \Sigma(t) \end{cases} \quad \text{avec } \sigma(t) = \frac{u_L(t) + u_R(t)}{2} \quad (1.7)$$

Mais u_a , $a \geq \max(u_L, -u_R)$ ainsi définie vérifie aussi (1.4):

$$u_a(x, t) = \begin{cases} u_L(t) = u_L e^{at} \text{ pour } x < \Sigma_1(t) \\ -a e^{at} \text{ pour } 0 > x > \Sigma_1(t) \\ a e^{at} \text{ pour } 0 < x < \Sigma_2(t) \\ u_R(t) = u_R e^{at} \text{ pour } x > \Sigma_2(t) \end{cases} \quad \text{avec } \begin{cases} \sigma_1(t) = \frac{u_L(t) - a e^{at}}{2} \\ \sigma_2(t) = \frac{a e^{at} + u_R(t)}{2} \end{cases}$$

Donc nous avons là un continuum de solutions faibles ... D'autre part, il est à noter que ces solutions sont sensibles aux manipulations de l'EDP (1.1) via la condition (1.5). En effet, multiplions (1.2) par u , il vient:

$$u \partial_t u + u^2 \partial_x u = \alpha u^2 \Rightarrow \partial_t \left(\frac{u^2}{2} \right) + \partial_x \left(\frac{u^3}{3} \right) = \alpha u^2 \quad (1.8)$$

Et les discontinuités relatives à (1.8), (1.6) vérifient cette fois:

$$\sigma(t) [u^+(t)^2 - u^-(t)^2] = \frac{2}{3} [u^+(t)^3 - u^-(t)^3]$$

On obtient donc des solutions différentes de celle fournie par (1.7).

1.3 Limite visqueuse et critère d'entropie

1.3.1 Origine parabolique

Un moyen de faire un choix avisé parmi cette infinité de solutions possibles est de revenir à l'origine physique du phénomène idéalisé décrit par (1.1). Considérons par exemple les équations d'Euler décrivant le mouvement d'un gaz via les variables de densité ρ , de quantité de mouvement (ρu) , et d'énergie totale (ρe) . Ce système est hyperbolique [35] et s'écrit:

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0 \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p) = 0 \\ \partial_t(\rho e) + \partial_x(u(\rho e + p)) = 0 \end{cases} \quad (1.9)$$

La quantité $p(\rho, \rho u, \rho e)$ est la pression. C'est la *loi de fermeture* du système. En amont de ce modèle idéalisé, on trouve le système de Navier-Stokes, qui comprend des termes au second ordre, dits *termes visqueux*.

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0 \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p) = (\lambda + \mu)\partial_{xx}u \\ \partial_t(\rho e) + \partial_x(u(\rho e + p)) = \partial_x(u(\lambda + \mu)\partial_x u) + k\partial_{xx}T \end{cases} \quad (1.10)$$

Ici, λ et μ sont les coefficients de viscosité de Lamé et k est le coefficient de conductivité thermique (T représente la température). Ce problème est parabolique [79] et admet des solutions continues. Alors, le système d'Euler peut être vu comme la limite de (1.10) pour lequel on a choisi des termes au second ordre infiniment faibles. Le système (1.9) est alors la *limite de viscosité* de (1.10).

Par analogie, nous définirons les solutions admissibles de (1.1) comme limites des solutions approchées u^ϵ du problème parabolisé suivant:

$$\begin{cases} \partial_t u^\epsilon + \partial_x f(u^\epsilon) = g(u^\epsilon) + \epsilon \partial_{xx}(u^\epsilon) \\ u^\epsilon(x, 0) = u_0^\epsilon(x) \\ (x, t) \in (\mathbb{R} \times \mathbb{R}^+) \end{cases} \quad (1.11)$$

Comme cette caractérisation est assez peu pratique, il est nécessaire d'introduire un autre critère d'admissibilité équivalent.

1.3.2 Notion de solution entropique

Définition I.2 On appelle entropie une fonction $\eta \in C^2(\mathbb{R})$ strictement convexe telle qu'il existe $q \in C^1(\mathbb{R})$ appelé flux d'entropie vérifiant: $q' = \eta' \cdot f'$

Voici donc comment reconnaître la limite $\epsilon \rightarrow 0$ des solutions des EDP parabolisées (1.11):

Proposition I.1 Supposons acquise l'existence d'un couple entropique (η, q) avec η strictement convexe. Alors si la suite $(u^\epsilon)_{\epsilon \rightarrow 0}$ est dans $L^\infty([0, T] \times \mathbb{R})$, on a

$$u^\epsilon \xrightarrow{L^{p < \infty}} u \Rightarrow \partial_t \eta(u) + \partial_x q(u) \leq \eta'(u) \cdot g(u) \quad (\mathcal{D}') \quad (1.12)$$

i.e. $\forall \varphi \geq 0 \in C_0^\infty(\mathbb{R} \times \mathbb{R}^+)$, $\int_{\mathbb{R} \times \mathbb{R}^+} \eta(u) \partial_t \varphi + q(u) \partial_x \varphi + \eta'(u) \cdot g(u) \varphi dx dt \geq 0$

Démonstration

Comme $u^\epsilon \in L^2(0, T; H^1(\mathbb{R}))$, on peut multiplier (1.11) par $\eta'(u^\epsilon)$:

$$\begin{aligned} \partial_t \eta(u^\epsilon) + \partial_x q(u^\epsilon) &= \eta'(u^\epsilon) \cdot g(u^\epsilon) + \epsilon \partial_{xx} \eta(u^\epsilon) - \eta''(u^\epsilon) |\sqrt{\epsilon} \partial_x u^\epsilon|^2 \\ &\stackrel{\text{Convexité de } \eta}{\leq} \eta'(u^\epsilon) \cdot g(u^\epsilon) + \epsilon \partial_{xx} \eta(u^\epsilon) \end{aligned}$$

Comme $(u^\epsilon)_\epsilon \in L^\infty$ et $u^\epsilon \xrightarrow{L^p} u$, on a $u^\epsilon \xrightarrow{pp^t} u$ à l'extraction d'une sous-suite près. Il vient donc:

$$\begin{cases} u^\epsilon \xrightarrow{\mathcal{D}'} u \\ \partial_t(u^\epsilon) \xrightarrow{\mathcal{D}'} \partial_t u \\ \epsilon \partial_{xx}(u^\epsilon) \xrightarrow{\mathcal{D}'} 0 \end{cases}$$

D'autre part, comme f et g sont régulières, donc fortement continues, q et $\eta'.g$ le sont aussi:

$$\begin{cases} \partial_x q(u^\epsilon) \xrightarrow{\mathcal{D}'} \partial_x q(u) \\ \partial_t \eta(u^\epsilon) \xrightarrow{\mathcal{D}'} \partial_t \eta(u) \\ \eta'(u^\epsilon).g(u^\epsilon) \xrightarrow{\mathcal{D}'} \eta'(u).g(u) \\ \epsilon \partial_{xx} \eta(u^\epsilon) \xrightarrow{\mathcal{D}'} 0 \end{cases}$$

A la limite, on obtient

$$\partial_t \eta(u) + \partial_x q(u) \leq \eta'(u).g(u) \quad (\mathcal{D}')$$

□

On peut démontrer que s'il est possible d'exhiber un couple (η, q) convexe pour une EDP hyperbolique scalaire, alors l'inégalité entropique est vérifiée pour tout couple convenable.

Théorème I.2 *Si u est solution $L^\infty([0, T] \times \mathbb{R})$ du problème (1.1) pour une entropie donnée, et si $u \mapsto f(u)$ est une fonction strictement convexe, alors u vérifie l'inégalité $\partial_t \eta(u) + \partial_x q(u) \leq \eta'(u).g(u)$ pour tout couple entropique (η, q) .*

Démonstration

Ce théorème est démontré rigoureusement dans [35]; la preuve est simplement esquissée ici. Celle-ci part de la condition de saut (1.5). Soit un couple entropique (η, q) . L'inégalité (1.12) entraîne une autre condition de saut:

$$\sigma(t) [\eta(u^+)(t) - \eta(u^-)(t)] \geq [q(u^+)(t) - q(u^-)(t)]$$

La réunion de celle-ci avec (1.5) implique:

$$\underbrace{\frac{f(u^+)(t) - f(u^-)(t)}{u^+(t) - u^-(t)} [\eta(u^+)(t) - \eta(u^-)(t)] - [q(u^+)(t) - q(u^-)(t)]}_{E_\eta(u^+(t), u^-(t))} \geq 0$$

Il est alors possible de montrer que $u \mapsto E_\eta(u, v)$, $E_\eta(v, v) = 0$, est une fonction (strictement) décroissante si η est une fonction (strictement) convexe. Donc, sous cette hypothèse, la condition d'entropie $E_\eta(u^+(t), u^-(t)) \geq 0$ pour une discontinuité $(u^+(t), u^-(t))$ s'écrit:

$$u^+(t) \leq u^-(t)$$

□

Par ailleurs, il peut être utile de préciser la représentation générale des couples entropiques.

Proposition I.2 *Les fonctions d'entropie $\eta(u)$ et $q(u)$ peuvent se définir comme*

$$\begin{cases} \eta(u) = \int^u H[f(\xi)].d\xi \\ q(u) = \int^{f(u)} H(\xi).d\xi \end{cases} \quad (1.13)$$

avec la condition de convexité

$$\eta''(u) = f'(u).H'[f(u)] \geq C > 0$$

Démonstration

La preuve est directe

$$\begin{cases} \eta'(u) = H[f(u)] \\ q'(u) = f'(u)H[f(u)] = \eta'(u).f'(u) \end{cases}$$

□

1.4 Le problème de Riemann

C'est un problème fondamental associé aux équations hyperboliques *homogènes* [35, 79]. Il est à la base des premiers schémas numériques pour ce type d'EDP [36]. Il se présente sous la forme d'un problème de Cauchy dont la donnée initiale est simplement une discontinuité en $x = 0$ (1.6).

1.4.1 Chocs admissibles et ondes de détente

On a vu que les discontinuités obtenues à la limite visqueuse de (1.11) sont, dans le cas d'une fonction de flux convexe, celles le long desquelles la solution décroît. Comme le flux associé à (1.2) s'écrit sous la forme: $u \mapsto \frac{u^2}{2}$, on déduit que la solution entropique de (1.2), (1.6) est celle fournie par l'expression (1.7) seulement dans le cas où $u_L > u_R$. Ceci entraîne aussi que $f'(u_L)(t) > f'(u_R)(t)$: on dit alors que le choc est *compressif* et c'est une condition nécessaire de stabilité [57, 79].

Dans le cas contraire $u_L < u_R$, la solution de (1.2) (1.6) est donc une onde *expansive* ou onde de détente. C'est-à-dire que u est C^1 -régulière et s'exprime simplement par le biais de la méthode des caractéristiques. C'est-à-dire que l'on cherche la forme des $X(t)$ associées à l'équation, et on construit la solution en résolvant l'équation différentielle ordinaire associée au terme source le long de ces courbes.

1.4.2 Résolution explicite

Tout ceci peut se résumer au sein d'un énoncé précis:

Proposition I.3 *En accord avec la condition (1.12), la solution du problème de Riemann (1.2), (1.6) s'écrit selon:*

- *Choc entropique ($u_L > u_R$)*

$$u(x, t) = \begin{cases} u_L(t) = u_L e^{\alpha t} \\ u_R(t) = u_R e^{\alpha t} \end{cases} \quad \text{avec } \Sigma(t) = \frac{u_L + u_R}{2\alpha} (e^{\alpha t} - 1) \quad (1.14)$$

- *Onde de détente ($u_L < u_R$)*

$$u(x, t) = \begin{cases} u_L(t) \text{ pour } x < X_L(t) \\ \frac{\alpha x}{1 - e^{-\alpha t}} \text{ pour } x \in [X_L(t), X_R(t)] \\ u_R(t) \text{ pour } x > X_R(t) \end{cases} \quad \text{avec } \begin{cases} X_L(t) = u_L \frac{e^{\alpha t} - 1}{\alpha} \\ X_R(t) = u_R \frac{e^{\alpha t} - 1}{\alpha} \end{cases} \quad (1.15)$$

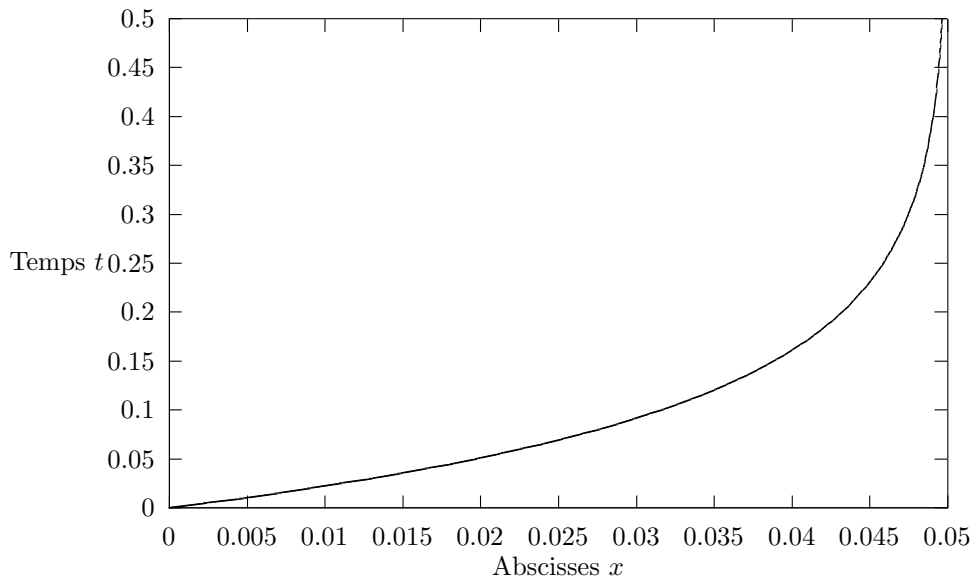


Figure 1.2: Trajectoire d'un choc amorti dans le plan (x, t)

Démonstration

Dans le premier cas (choc), les paliers $t \mapsto u_L(t)$ et $t \mapsto u_R(t)$ sont soumis à l'EDO $\dot{u} = \alpha u$. Quant à la condition

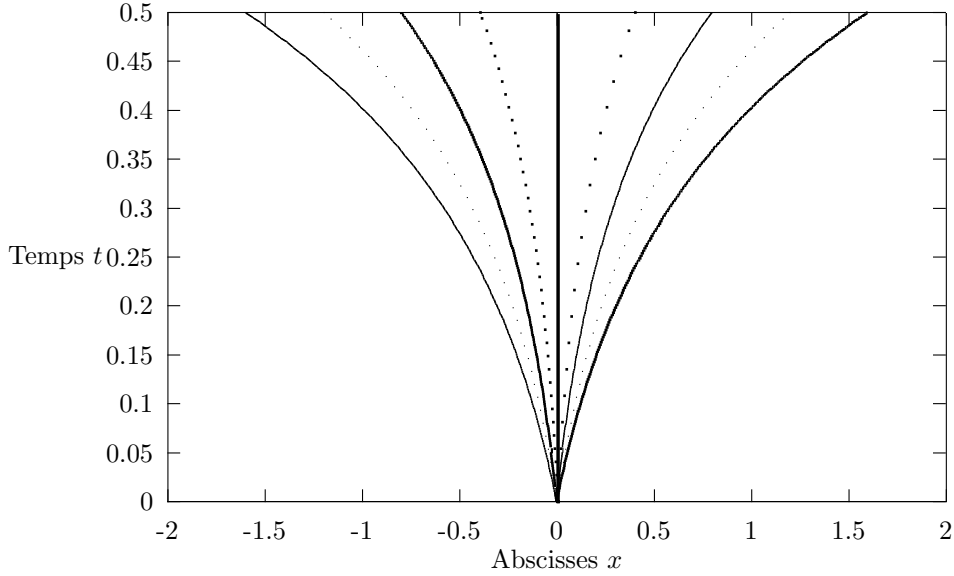


Figure 1.3: Caractéristiques dans le cône d'une détente sonique

(1.5) qui régit la trajectoire de la discontinuité dans le plan (x, t) , elle s'intègre aussi simplement:

$$\sigma(t) = \frac{u_L + u_R}{2} e^{\alpha t} \Rightarrow \Sigma(t) = \frac{u_L + u_R}{2\alpha} (e^{\alpha t} - 1)$$

puisque la donnée initiale (1.6) impose $\Sigma(0) = 0$.

Dans le cas d'une onde régulière, les lignes caractéristiques $t \mapsto X(t)$ issues de $(0, 0)$ s'écrivent $X(t) = \xi \frac{e^{\alpha t} - 1}{\alpha}$ où $\xi \in [u_L, u_R]$. Le long de celles-ci, la solution est une courbe intégrale de l'EDO $\dot{u} = \alpha u$: elle varie donc en $e^{\alpha t}$. On définit donc une quantité conservée par:

$$w(\xi) = w\left(\frac{\alpha x}{e^{\alpha t} - 1}\right) = \frac{u(x, t)}{e^{\alpha t}}$$

Il vient donc: $u(x, t) = e^{\alpha t} w(\xi)$. L'EDP (1.2) se réécrit donc sous la forme:

$$\alpha e^{\alpha t} w(\xi) + e^{2\alpha t} \frac{-\alpha^2}{(e^{\alpha t} - 1)^2} w'(\xi) + e^{\alpha t} w'(\xi) w(\xi) \frac{\alpha}{e^{\alpha t} - 1} = \alpha e^{\alpha t} w(\xi)$$

C'est-à-dire:

$$e^{2\alpha t} \frac{-\alpha^2}{(e^{\alpha t} - 1)} w'(\xi) + \alpha e^{\alpha t} w(\xi) w'(\xi) = 0 \Rightarrow w'(\xi) [-\xi + w(\xi)] = 0$$

Si on exclut la solution constante, on finit par trouver:

$$w(\xi) = \xi \Leftrightarrow u(x, t) = \frac{\alpha x}{1 - e^{-\alpha t}}$$

Finalement, on a bien trouvé les expressions (1.14), (1.15). \square

Remarquons que la propriété d'*autosimilarité* du cas homogène est perdue. En effet, la solution $u(x, t)$ ne peut plus se mettre sous la forme $w(x/t)$ et surtout, elle n'est plus *constante* le long de ses lignes de propagation. Il peut être utile de représenter graphiquement ces solutions. Les figures (1.2) et (1.3) montrent respectivement une courbe $t \mapsto \Sigma(t)$ pour $\alpha = -10$ et un cône de détente sonique pour $\alpha = 4$.

Chapitre 2

Méthode de viscosité et existence

Cette première partie est consacrée à la démonstration théorique de l'existence et de l'unicité de la solution $L^\infty([0, T] \times \mathbb{R})$ du problème générique scalaire:

$$\begin{cases} \partial_t u + \partial_x f(u) = g(u) \\ u(x, 0) = u_0(x) \in L^2 \cap L^\infty(\mathbb{R}) \\ f \in C^2(\mathbb{R}), f'' > 0 \\ g \in C^1(\mathbb{R}), g(0) = 0; \exists M/|x| > M \Rightarrow x.g(x) < 0 \end{cases} \quad (2.1)$$

La démarche adoptée dans ce travail suit les idées introduites par L. Tartar [80]:

-Existence d'une solution parabolique dans $L^2(0, T; H^1(\mathbb{R}))$: Ceci résulte de la théorie générale des EDP paraboliques (cf [10], par exemple)

-Estimation uniforme L^∞ de cette solution régulière: C'est une estimation importante à savoir qu'une donnée initiale bornée induit une solution uniformément bornée. On se place donc dans un cadre excluant les *explosions différentielles* [2].

-Mise en place des propriétés de compacité dans $*\sigma(L^1, L^\infty)$: L'estimation précédente nous donne un résultat de convergence faible insuffisant pour traiter les termes non-linéaires. Néanmoins, certaines de ces fonctionnelles ont de bonnes propriétés vis-à-vis de ce type de convergence.

-Estimation de la phrase entropique non-homogène et compacité de la suite d'approximations régulières: Pour s'assurer de la continuité faible de la fonction de flux, il faut au préalable vérifier une propriété de compacité de l'équation d'entropie dans un espace dual de H^1 .

-Passage à la limite de viscosité et conclusion

On a décidé de poser le problème en dimension $d = 1$ d'espace, ce qui signifie que l'on se borne à *une seule fonction de flux* $f \in C^2(\mathbb{R})$. En effet, dans le cas de $d \geq 2$ fonctions de flux, il convient de faire appel à une représentation cinétique pour obtenir les estimations suffisantes à l'application de cette théorie [63].

2.1 Problème parabolisé

2.1.1 Rappels généraux

On note Ω un ouvert non nécessairement borné du demi-espace $(\mathbb{R} \times \mathbb{R}^+)$ ainsi que $W(0, T)$ l'espace défini comme suit:

$$W(0, T) = \{v \in L^2(0, T; H^1(\mathbb{R})); \partial_t v \in L^2(0, T; H^{-1}(\mathbb{R})); T \in \mathbb{R}^+\}$$

On a ensuite le résultat classique d'existence parabolique suivant [10, 35]:

Théorème I.3 (*Lions*)

Il existe une unique fonction $u \in W(0, T)$ solution de:

$$\begin{cases} \partial_t u - \epsilon \partial_{xx} u + \lambda u = g \in L^2(0, T; H^{-1}(\mathbb{R})) \\ u(x, 0) = u_0(x) \in L^2(\mathbb{R}) \end{cases}$$

pour $(\lambda, \epsilon) \in (\mathbb{R}_*^+)^2$; elle est caractérisée par: $\forall v \in H^1(\mathbb{R}), \forall t \in [0, T]$

$$\langle \partial_t u(\cdot, t), v \rangle_{H^{-1}, H^1} + \epsilon \langle \partial_x u(\cdot, t), \partial_x v \rangle_{L^2} + \lambda \langle u(\cdot, t), v \rangle_{L^2} = \int_{\mathbb{R}} (g_0 v + g_1 \partial_x v) dx$$

avec $g = g_0 + \partial_x g_1$ et $g_i \in L^2([0, T] \times \mathbb{R})$.

On poursuit avec une série de quatre lemmes techniques.

Lemme I.1 Pour $f \in C^1(\mathbb{R}), f(0) = 0, f' \in L^\infty(\mathbb{R})$ et $u \in H^1(\mathbb{R})$, on a:

$$f \circ u \in H^1(\mathbb{R}) \text{ et } \partial_x (f \circ u) = (f' \circ u) \partial_x u$$

Lemme I.2 L'application $v \mapsto |v|$ est continue de $H^1(\mathbb{R})$ dans $H^1(\mathbb{R})$ et $\forall |v| \in H^1(\mathbb{R}), \partial_x |v| = \text{sgn}(v) \partial_x v$.

Lemme I.3 Soit $v \in W(0, T)$, alors $v_+ \in L^2(0, T; H^1(\mathbb{R}))$ et

$$\forall t_1, t_2 \in [0, T], \int_{t_1}^{t_2} \langle \partial_t v(\cdot, s), v_+(\cdot, s) \rangle ds = \|v_+(\cdot, t_2)\|_{L^2}^2 - \|v_+(\cdot, t_1)\|_{L^2}^2$$

Lemme I.4 (*Gronwall "faible"*) Si $f \in C^1([a, b])$ est telle que:

$$\forall x \in]a, b[, |f'(x)| \leq K|f(x)|$$

alors $f \equiv 0$, ou f ne s'annule en aucun point de $[a, b]$.

2.1.2 Lemme du point fixe

On établit maintenant un premier résultat d'existence pour l'EDP régularisée sous la condition que le flux f est globalement lipschitzien. La technique employée repose sur le théorème du point fixe. On s'attachera ensuite à relaxer cette restriction au moyen d'une majoration uniforme des valeurs prises par la solution ainsi construite.

Théorème I.4 (*Picard*)

Soit $F : E \rightarrow E$ (E métrique complet) telle qu'il existe $K < 1$ vérifiant pour tout couple $(x, y) \in E^2$, $d(F(x), F(y)) \leq Kd(x, y)$, alors il existe un unique $z \in E$ tel que $F(z) = z$.

On peut alors énoncer ceci:

Proposition I.4 Si f et g vérifient:

$$\begin{cases} (f, g) \in C^1(\mathbb{R}^2); \forall (x, y) \in \mathbb{R}^2, |f(x) - f(y)| \leq C|x - y| \\ f(0) = g(0) = 0; \forall (x, y) \in \mathbb{R}^2, |g(x) - g(y)| \leq C'|x - y| \end{cases}$$

alors il existe une unique fonction $u \in W(0, T)$ solution de:

$$\begin{cases} \partial_t u - \epsilon \partial_{xx} u = g(u) - \partial_x f(u) \\ u(x, 0) = u_0(x) \in L^2(\mathbb{R}) \end{cases} \quad (2.2)$$

Démonstration

On effectue le changement d'inconnue $v = u \cdot e^{-\lambda t}$ pour faire apparaître un terme linéaire. Donc u solution de (2.2) équivaut à v solution de:

$$\begin{cases} \partial_t v - \epsilon \partial_{xx} v + \lambda v = -e^{-\lambda t} [\partial_x f(v e^{\lambda t}) - g(v e^{\lambda t})] \\ v_0 = u_0 \in L^2(\mathbb{R}) \end{cases}$$

On considère alors le problème parabolique *linéaire*:

$$\begin{cases} \partial_t w - \epsilon \partial_{xx} w + \lambda w = -e^{-\lambda t} [\partial_x f(v e^{\lambda t}) - g(v e^{\lambda t})] \\ w_0 = v_0 \in L^2(\mathbb{R}) \end{cases}$$

Le membre de droite est alors dans $L^2(\mathbb{R}^+; H^{-1}(\mathbb{R}))$ par définition de l'espace dual $H^{-1}(\mathbb{R})$ puisque:

$$\begin{cases} e^{-\lambda t} |\partial_x f(v e^{\lambda t})| = e^{-\lambda t} |f(v e^{\lambda t}) - f(0)| \leq C|v| \in L^2(\mathbb{R}^+ \times \mathbb{R}) \\ e^{-\lambda t} |g(v e^{\lambda t})| = e^{-\lambda t} |g(v e^{\lambda t}) - g(0)| \leq C'|v| \in L^2(\mathbb{R}^+ \times \mathbb{R}) \end{cases}$$

Ce qui implique:

$$\begin{cases} e^{-\lambda t} \partial_x f(v e^{\lambda t}) \in L^2(\mathbb{R}^+; H^{-1}(\mathbb{R})) \\ e^{-\lambda t} g(v e^{\lambda t}) \in L^2(\mathbb{R}^+; H^{-1}(\mathbb{R})) \end{cases}$$

On déduit du théorème I.3 qu'il existe une unique fonction $w \in W(\mathbb{R}^+)$ solution de l'EDP linéaire ci-dessus posée par rapport à v .

Il reste à établir qu'il lui correspond un unique alter-ego non-linéaire v .

Pour cela, il faut montrer que l'opérateur

$$F_\lambda : L^2(\mathbb{R}^+ \times \mathbb{R}) \rightarrow W(\mathbb{R}^+)$$

$$v \mapsto F_\lambda(v) = w$$

est strictement contractant dans $L^2(\mathbb{R}^+ \times \mathbb{R})$, donc admet un unique point fixe.

$$\|w_1 - w_2\|_{L^2} \stackrel{?}{\leq} K \|v_1 - v_2\|_{L^2} \quad (K < 1)$$

On pose $w = w_1 - w_2$; comme le problème est linéaire, on obtient pour toute fonction-test $z \in H^1(\mathbb{R})$:

$$\begin{aligned} & \langle \partial_t w(\cdot, t), z \rangle_{H^{-1}, H^1} + \epsilon \langle \partial_x w(\cdot, t), \partial_x z \rangle_{L^2} + \lambda \langle w(\cdot, t), z \rangle_{L^2} \\ &= \mathbf{e}^{-\lambda t} \left\{ \int_{\mathbb{R}} [f(\mathbf{e}^{\lambda t} v_1) - f(\mathbf{e}^{\lambda t} v_2)](\cdot, t) \partial_x z + [g(\mathbf{e}^{\lambda t} v_1) - g(\mathbf{e}^{\lambda t} v_2)](\cdot, t) z \, dx \right\} \end{aligned}$$

On pose $z = w(\cdot, t)$, et on obtient en intégrant sur $[0, T]$ et en utilisant l'inégalité de Cauchy-Schwarz:

$$\begin{aligned} & \frac{1}{2} \|w(\cdot, t)\|_{L^2}^2 + \int_0^t (\epsilon \|\partial_x w(\cdot, s)\|_{L^2}^2 + \lambda \|w(\cdot, s)\|_{L^2}^2) \, ds \\ & \leq C \int_0^t \|v_1 - v_2\|_{L^2} \cdot \|\partial_x w(\cdot, s)\|_{L^2} \, ds + C' \int_0^t \|v_1 - v_2\|_{L^2} \cdot \|w(\cdot, s)\|_{L^2} \, ds \end{aligned}$$

D'après l'inégalité $xy \leq \alpha x^2 + \frac{1}{4\alpha} y^2$, il vient que $Cxy \leq \epsilon x^2 + \frac{C^2}{4\epsilon} y^2$. Ensuite, par celle de Minkowski, on obtient:

$$\frac{1}{2} \|w(\cdot, t)\|_{L^2}^2 + \int_0^t \lambda \|w(\cdot, s)\|_{L^2}^2 \, ds \leq \frac{C^2}{4\epsilon} \int_0^t \|v_1 - v_2\|_{L^2}^2 \, ds + \frac{C'}{2} \int_0^t (\|v_1 - v_2\|_{L^2}^2 + \|w(\cdot, s)\|_{L^2}^2) \, ds$$

D'où on déduit, pour $t \rightarrow +\infty$:

$$\int_0^{+\infty} \|w(\cdot, s)\|_{L^2}^2 \, ds \leq K \int_0^{+\infty} \|v_1 - v_2\|_{L^2}^2 \, ds$$

Pour λ assez grand, $K = \frac{C^2 + C'}{\lambda - \frac{C^2}{4\epsilon}} < 1$ et F_λ est strictement contractant dans $L^2(\mathbb{R}^+ \times \mathbb{R})$.

Donc il existe un unique $v \in W(\mathbb{R}^+)$ et donc un unique $u = \mathbf{e}^{\lambda t} v \in W(0, T)$ solutions du problème non-linéaire correspondant. \square

2.1.3 Principe du maximum

Nous cherchons maintenant à mettre en évidence une majoration uniforme de la norme infinie de la solution de l'EDP régularisée. Cette estimation, bien qu'insuffisante pour passer à la limite visqueuse par les moyens habituels, se révélera par la suite un ingrédient indispensable pour appliquer les résultats de compacité par compensation de F. Murat et L. Tartar. D'autre part, un tel résultat permettra aussi de s'affranchir de la restriction faite précédemment sur la fonction de flux et le terme réactionnel (f et g globalement Lipschitziennes). On l'énonce sous la forme suivante:

Proposition I.5 *Supposons que $(f, g) \in C^1(\mathbb{R})^2$; $f(0) = g(0) = 0$ et faisons l'hypothèse:*

$$\exists X \in \mathbb{R} / |x| > X \Rightarrow x.g(x) < 0 \tag{2.3}$$

alors le problème visqueux:

$$\begin{cases} \partial_t u - \epsilon \partial_{xx} u = g(u) - \partial_x f(u) \\ u(x, 0) = u_0(x) \in L^2 \cap L^\infty(\mathbb{R}) \end{cases}$$

admet une unique solution $u \in W(0, T) \cap L^\infty(\mathbb{R} \times \mathbb{R}^+)$ vérifiant:

$$\|u\|_{L^\infty(\mathbb{R} \times \mathbb{R}^+)} \leq \max(\|u_0\|_{L^\infty(\mathbb{R})}, X) \equiv M$$

Démonstration

Pour commencer, on tronque les fonctions f et g par $\varphi \in C_0^\infty(\mathbb{R})$ vérifiant:

$$\varphi(x) = \begin{cases} 1 & \text{pour } x \leq M \\ 0 & \text{pour } x \geq M + 1 \end{cases}$$

On pose alors $\tilde{f} = f \cdot \varphi$, $\tilde{g} = g \cdot \varphi$, et on applique le résultat d'existence précédent:

$$\exists ! u \in W(0, T) \text{ vérifiant } \begin{cases} \partial_t u - \epsilon \partial_{xx} u = \tilde{g}(u) - \partial_x \tilde{f}(u) \\ u(x, 0) = u_0(x) \in L^2 \cap L^\infty(\mathbb{R}) \end{cases}$$

On pose $v = u - M$; il faut donc montrer que $v_+ \equiv 0$.

Commençons par remarquer que:

$$\begin{cases} |v_+| \leq |u| \text{ donc } v_+ \in L^2([0, T] \times \mathbb{R}) \\ \partial_x v(\cdot, t) = \partial_x u(\cdot, t) \text{ donc } v \in H^1(\mathbb{R}) \\ x \mapsto |x| \text{ est continu dans } H^1(\mathbb{R}) \text{ donc } v_+(\cdot, t) \in H^1(\mathbb{R}) \end{cases}$$

Par construction, v est solution du problème tronqué:

$$\partial_t v - \epsilon \partial_{xx} v = \tilde{g}(u) - \partial_x \tilde{f}(u) = \tilde{g}(u) - \tilde{f}'(u) \partial_x u$$

Ce problème *linéaire* admet la formulation variationnelle: $\forall z \in H^1(\mathbb{R})$,

$$\langle \partial_t v(\cdot, t), z \rangle_{H^{-1}, H^1} + \epsilon \langle \partial_x v(\cdot, t), \partial_x z \rangle_{H^1} = \langle \tilde{g}(u)(\cdot, t), z \rangle_{H^1} - \langle \tilde{f}'(u) \partial_x u(\cdot, t), z \rangle_{H^1}$$

Par ailleurs, comme $u(\cdot, t) \in H^1(\mathbb{R})$, $g \in C^1(\mathbb{R})$, et $g(0) = 0$, on déduit que $g(u)(\cdot, t) \in H^1(\mathbb{R})$. L'expression est donc bien justifiée. On a aussi:

$$\begin{cases} \partial_x v(\cdot, t) \cdot \partial_x v_+(\cdot, t) = |\partial_x v_+|^2(\cdot, t) \\ \partial_x u(\cdot, t) \cdot v_+(\cdot, t) = \partial_x v(\cdot, t) \cdot v_+(\cdot, t) = (\partial_x v_+) \cdot v_+(\cdot, t) \end{cases}$$

On remarque alors qu'en posant $z = v_+(\cdot, t) \in H^1(\mathbb{R})$, il vient:

$$\int_{\mathbb{R}} \tilde{g}(u)(\cdot, t) \cdot v_+(\cdot, t) \cdot dx = \underbrace{\int_{u > M} \tilde{g}(u)(\cdot, t) \cdot v_+(\cdot, t) \cdot dx}_{u > M \geq X \Rightarrow \tilde{g}(u) \leq 0} + \underbrace{\int_{u \leq M} \tilde{g}(u)(\cdot, t) \cdot v_+(\cdot, t) \cdot dx}_{u \leq M \Rightarrow v \leq 0 \Rightarrow f \equiv 0} \leq 0$$

Par le lemme I.3, la formulation variationnelle devient:

$$\begin{aligned} & \frac{1}{2} (\|v_+(\cdot, t)\|_{L^2}^2 - \underbrace{\|v_+(\cdot, 0)\|_{L^2}^2}_{=0}) + \epsilon \int_0^t \|\partial_x v_+(\cdot, s)\|_{L^2}^2 \cdot ds \\ & \leq C \int_0^t \|\partial_x v_+(\cdot, s)\|_{L^2} \|v_+(\cdot, s)\|_{L^2} \cdot ds \end{aligned}$$

Par l'inégalité $Cxy \leq \epsilon x^2 + \frac{C^2}{4\epsilon} y^2$, on obtient:

$$\|v_+(\cdot, t)\|_{L^2}^2 \leq \frac{C^2}{2\epsilon} \int_0^t \|v_+(\cdot, s)\|_{L^2}^2 \cdot ds$$

Ce qui nous met en position d'utiliser le lemme de Gronwall pour conclure que v_+ est identiquement nulle. De même, pour montrer que $(-u - M)_+ = (u + M)_- \equiv w_-$ est nulle, on procède identiquement sachant que:

$$- \int_{\mathbb{R}} \tilde{g}(u) \cdot w_- \cdot dx = \underbrace{\int_{u < -M} -\tilde{g}(u)(\cdot, t) \cdot w_-(\cdot, t) \cdot dx}_{u < -M \leq -X \Rightarrow \tilde{g}(u) > 0} - \underbrace{\int_{u \geq -M} \tilde{g}(u)(\cdot, t) \cdot w_-(\cdot, t) \cdot dx}_{u \geq M \Rightarrow w \geq 0 \Rightarrow f \equiv 0} \leq 0$$

Finalement, $\|u\|_{L^\infty} \leq M$ en presque tout t , et u est l'unique solution puisqu'alors:

$$\begin{cases} \tilde{f}(u) = f(u) \\ \tilde{g}(u) = g(u) \end{cases}$$

□

2.2 Compacité par compensation

Cette théorie, construite à partir de celle des fonctions généralisées de L.C. Young, permet de mettre en lumière la continuité faible de la fonction de flux à partir de la majoration L^∞ de la solution régularisée. Bien sûr, cette information insuffisante (pas d'injection compacte) est compensée par les estimations d'entropie. Dans ce cas, celles-ci assurent que la solution trouvée est la limite quand $\epsilon \rightarrow 0$ des solutions approchées u^ϵ .

2.2.1 Présentation de la théorie

On rappelle ici les principaux résultats qui nous seront utiles. Ils sont tous issus du livre de B. Dacorogna [20] ou de l'article de L. Tartar [80] auxquels on renvoie pour plus de détails.

Nous en sommes à $\|u^\epsilon\|_{L^\infty} \leq M$, ce qui implique que la suite $(u^\epsilon)_{\epsilon \rightarrow 0}$ est compacte pour la topologie faible $^*\sigma(L^1, L^\infty)$ puisque L^1 n'est pas réflexif [10].

On désigne par $\mathcal{K}[\cdot]$ les ensembles compacts et par $\mathcal{B}[\cdot]$ les bornés.

Définition I.3 Soit Ω un ouvert régulier de \mathbb{R}^n , on considère les informations suivantes:

$$\begin{cases} u^\epsilon \xrightarrow{*L^\infty} u : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m \\ Au^\epsilon = \sum_{j=1}^m \sum_{k=1}^n \alpha_{ijk} \partial_{x_k} u_j^\epsilon \quad (i = 1 \dots q) \\ f(u^\epsilon) \xrightarrow{*L^\infty} l \end{cases}$$

On introduit alors:

$$B(\xi) : \mathbb{R}^m \rightarrow \mathbb{R}^q \\ \lambda \mapsto \left(\sum_{j=1}^m \sum_{k=1}^n \alpha_{ijk} \lambda_j \xi_k \right)_{i=1 \dots q}$$

On pose:

$$V = \{(\lambda, \xi) \in (\mathbb{R}^m \times \mathbb{R}^n) \text{ tels que } B(\xi)\lambda = 0\}$$

Et on définit l'espace:

$$\Lambda = \{\lambda \in \mathbb{R}^m \text{ tels que } \exists \xi \in \mathbb{R}^n \text{ vérifiant } (\lambda, \xi) \in V\}$$

Le résultat le plus utile dans le cadre de notre problème est concentré dans cet énoncé:

Théorème I.5 Soit $M \in \mathbf{M}_m(\mathbb{R})$, $M^t = M$, $Q(x) = \langle Mx, x \rangle_{\mathbb{R}^m}$ et supposons:

$$\begin{cases} u^\epsilon \xrightarrow{W^{-L^2}} u \\ Au^\epsilon = \sum_{j=1}^m \sum_{k=1}^n \alpha_{ijk} \partial_{x_k} u_j^\epsilon \text{ compact dans } H_{loc}^{-1}(\Omega) \text{ pour chaque } i=1 \dots q \\ Q(u^\epsilon) \xrightarrow{D'} l \end{cases}$$

Alors, si pour tout $\lambda \in \Lambda$, $Q(\lambda) = 0$, on a $Q(u) = l$. Autrement dit, Q ainsi définie est faiblement continue pour $(u^\epsilon)_\epsilon$ si Λ est inclus dans le noyau de M .

Ceci est sous-tendu par le résultat abstrait ci-dessous caractérisant les limites faibles dans $^*\sigma(L^1, L^\infty)$ en termes de mesures de Young.

Théorème I.6 (Tartar)

Posons K un compact de \mathbb{R}^m .

• Soit $u_s : \Omega \rightarrow \mathbb{R}^m$ telle que $u_s(x) \in K$, alors il existe une sous-suite encore notée $(u_s)_s$ et une famille de mesures de probabilités $(\nu_x)_{x \in \Omega}$ vérifiant pour toute fonction $f : \mathbb{R}^m \rightarrow \mathbb{R}$ continue:

$$\begin{cases} \text{supp}(\nu_x) \subset \overline{K} \\ f(u_s) \xrightarrow{*L^\infty} \bar{f} : x \mapsto \int_{\mathbb{R}^m} f(\lambda) . d\nu_x(\lambda) \end{cases}$$

• Réciproquement, si $(\nu_x)_{x \in \Omega}$ est une famille à un paramètre de mesures de probabilités à support dans \overline{K} , il existe une suite $(u_s)_s$ telle que $u_s(x) \in K$ et pour toute fonction f continue $K \rightarrow \mathbb{R}$:

$$f(u_s) \xrightarrow{*L^\infty} \bar{f} : x \mapsto \int_{\mathbb{R}^m} f(\lambda) . d\nu_x(\lambda)$$

On donne ensuite un critère établissant la convergence forte.

Théorème I.7 (Tartar)

Si $u_s \xrightarrow{*L^\infty} u$, alors $u_s \xrightarrow{L^{p < \infty}} u$ si et seulement si pour tout $x \in \Omega$, $\nu_x(\lambda) = \delta_{u(x)}(\lambda)$, la mesure de Dirac concentrée au point $u(x)$.

On termine par un lemme technique très utile en pratique.

Lemme I.5 (Murat) Soit Ω un ouvert de \mathbb{R}^n :

$$\mathcal{K}(H_{loc}^{-1}(\Omega)) \supset (\mathcal{K}(H^{-1}(\Omega)) + \mathcal{B}(\mathcal{M}(\Omega))) \cap \mathcal{B}(W^{-1, \infty}(\Omega))$$

2.2.2 Emergence de la solution globale L^∞

On donne ici le résultat le plus important de cette étape: les conditions sous lesquelles les fonctionnelles non-linéaires de l'EDP de convection-réaction sont continues pour la topologie faible-*. Nous commencerons par appliquer les théorèmes précédents pour préciser le comportement de $u^\epsilon \mapsto \int_{\mathbb{R}} f(u^\epsilon) \phi dx$ lorsque $\epsilon \rightarrow 0$. Ensuite, nous verrons que cela entraîne une convergence forte de la suite $(u^\epsilon)_{\epsilon \rightarrow 0}$ (compacité) en utilisant l'hypothèse de stricte convexité faite sur la fonction de flux.

Proposition I.6 Soit Ω le demi-espace borné en temps $[0, T] \times \mathbb{R}$.

On suppose en outre l'existence de couples (η, q) entropiques vérifiant:

$$q' = \eta' \cdot f' \text{ avec } \eta'' \geq 0$$

Alors la donnée de:

$$\begin{cases} u^\epsilon \overset{*}{\rightharpoonup} u \text{ et } u^\epsilon \text{ bornée dans } L^\infty(\Omega) \\ f \in C^2(\mathbb{R}) \text{ telle que } f'' > 0 \text{ et } f(0) = 0 \\ \forall \eta \text{ convexe : } \partial_t \eta(u^\epsilon) + \partial_x q(u^\epsilon) \in \mathcal{K}(H_{loc}^{-1}(\Omega)) \end{cases}$$

entraîne:

$$\begin{cases} f(u^\epsilon) \overset{*}{\rightharpoonup} f(u) \\ u^\epsilon \xrightarrow{L^p} u \text{ (} p < +\infty \text{)} \end{cases}$$

Démonstration

La (longue) démonstration sera scindée en deux parties pour plus de clarté.

Continuité *faible du flux

On pose $U^\epsilon = (u^\epsilon, f(u^\epsilon), \eta(u^\epsilon), q(u^\epsilon)) \in (L^\infty(\Omega))^4$.

Comme U^ϵ est uniformément borné dans L^∞ , on a la convergence faible

$$U^\epsilon \overset{*}{\rightharpoonup} U \stackrel{def}{=} (u, v, w, z)$$

On désire montrer au moins que $v = f(u)$. Pour cela, on a par hypothèse:

$$\begin{cases} \partial_t u^\epsilon + \partial_x v^\epsilon \in \mathcal{K}(H_{loc}^{-1}(\Omega)) \\ \partial_t w^\epsilon + \partial_x z^\epsilon \in \mathcal{K}(H_{loc}^{-1}(\Omega)) \end{cases}$$

puisque la première ligne correspond au cas particulier $\eta(x) = x$. Dans les notations de la définition I.3, l'estimation sur les dérivées s'interprète grace à l'opérateur $\lambda \mapsto B(\xi) \cdot \lambda$. Dans notre cas, ce dernier s'écrit sous la forme:

$$(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \mapsto B(\xi_1, \xi_2) \cdot (\lambda_1, \lambda_2, \lambda_3, \lambda_4) = \begin{cases} \xi_1 \cdot \lambda_1 + \lambda_2 \cdot \xi_2 \\ \xi_1 \cdot \lambda_3 + \lambda_4 \cdot \xi_2 \end{cases}$$

Ces relations impliquent donc que l'ensemble Λ se définit comme:

$$\Lambda = \{ \lambda \in \mathbb{R}^4 \text{ tels que } \begin{pmatrix} \lambda_1 & \lambda_2 \\ \lambda_3 & \lambda_4 \end{pmatrix} \cdot \vec{\xi} = \vec{0} \text{ pour } \vec{\xi} \in \Omega \subset (\mathbb{R}^+ \times \mathbb{R})^* \}$$

Il est à noter que les $(\xi_i)_{1 \leq i \leq d+1}$ représentent les opérateurs de dérivées partielles. Donc Λ est l'ensemble des $(\lambda_i)_{1 \leq i \leq 4}$ tels que le déterminant de la matrice 2×2 soit nul.

$$\Lambda = \{ \lambda \in \mathbb{R}^4 \text{ tels que } \lambda_1 \cdot \lambda_4 - \lambda_2 \cdot \lambda_3 = 0 \}$$

Posons:

$$Q(U) = \langle MU, U \rangle \text{ avec } M = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

D'après le théorème I.5, ceci implique que:

$$Q(U^\epsilon) = u^\epsilon \cdot z^\epsilon - v^\epsilon \cdot w^\epsilon \overset{*}{\rightharpoonup} Q(U) = u \cdot z - v \cdot w$$

C'est-à-dire que:

$$u^\epsilon \cdot q(u^\epsilon) - f(u^\epsilon) \cdot \eta(u^\epsilon) \xrightarrow{*-\underline{L}^\infty} u \cdot z - v \cdot w$$

On introduit alors la représentation suggérée par le théorème I.6. Comme $u^\epsilon \mapsto U^\epsilon = (u^\epsilon, f(u^\epsilon), \eta(u^\epsilon), q(u^\epsilon))$ est une application continue, il existe une famille de mesures paramétrées $(\nu_{x,t})_{x \in \mathbb{R}, t \in \mathbb{R}^+}$ telles que:

$$\begin{cases} u(x, t) = \int_{\mathbb{R}} \lambda \cdot d\nu_{x,t}(\lambda) \equiv \langle \nu_{x,t}, \lambda \rangle_{\mathcal{M}(\Omega), C_0^1(\Omega)} \\ v(x, t) = \langle \nu_{x,t}, f(\lambda) \rangle_{\mathcal{M}(\Omega), C_0^1(\Omega)} \\ w(x, t) = \langle \nu_{x,t}, \eta(\lambda) \rangle_{\mathcal{M}(\Omega), C_0^1(\Omega)} \\ z(x, t) = \langle \nu_{x,t}, q(\lambda) \rangle_{\mathcal{M}(\Omega), C_0^1(\Omega)} \end{cases}$$

Ce qui amène:

$$\langle \nu, U \rangle = (u, v, w, z)$$

Notons que λ représente ici les valeurs possibles pour $u(x, t)$, lorsque (x, t) parcourt Ω . On comprend donc que l'étalement du support de $\nu_{x,t}$ est directement lié aux oscillations subies par la limite $\epsilon \rightarrow 0$ de la suite $u^\epsilon(x, t)$. La limite faible $u(x, t) = \langle \nu_{x,t}, \lambda \rangle$ est alors l'expression de l'espérance de la variable aléatoire réelle de densité $\nu_{x,t}(\lambda)$. La convergence forte correspond alors à la concentration de cette densité en un point.

Par construction de $\nu_{x,t}$, on a:

$$Q(U^\epsilon) \xrightarrow{*-\underline{L}^\infty} \bar{Q} = \langle \nu, Q(U) \rangle = \langle \nu, \lambda q(\lambda) - f(\lambda) \eta(\lambda) \rangle$$

On a aussi $Q(U^\epsilon) \xrightarrow{*-\underline{L}^\infty} Q(U) = Q(\langle \nu, U \rangle)$, puisque Q est faiblement continue pour $(U^\epsilon)_{\epsilon \rightarrow 0}$. On obtient alors la propriété fondamentale de $Q(U)$:

$$\langle \nu, Q(U) \rangle = Q(\langle \nu, U \rangle)$$

Ceci s'écrit:

$$\langle \nu, \lambda \cdot q(\lambda) - f(\lambda) \cdot \eta(\lambda) \rangle = \langle \nu, \lambda \rangle \langle \nu, q(\lambda) \rangle - \langle \nu, f(\lambda) \rangle \langle \nu, \eta(\lambda) \rangle$$

D'où on déduit:

$$\langle \nu, \lambda \cdot q(\lambda) - f(\lambda) \cdot \eta(\lambda) \rangle = u \langle \nu, q(\lambda) \rangle - v \langle \nu, \eta(\lambda) \rangle$$

Par linéarité de $\langle \nu, \cdot \rangle$, on obtient:

$$\langle \nu, (\lambda - u)q(\lambda) - (f(\lambda) - v)\eta(\lambda) \rangle = 0 \quad (2.4)$$

On choisit maintenant les entropies faibles de Kruzkov: $\eta(\lambda) = |\lambda - u|$

Du fait de la relation $q'(\lambda) = \eta'(\lambda) \cdot f'(\lambda)$, on a: $q(\lambda) = \text{sgn}(\lambda - u)(f(\lambda) - f(u))$. Ceci implique:

$$(\lambda - u)q(\lambda) - (f(\lambda) - v)\eta(\lambda) = (v - f(u))|\lambda - u|$$

Et l'expression (2.4) devient alors:

$$\langle \nu, |\lambda - u| \rangle (v - f(u)) = 0$$

On obtient finalement l'alternative suivante pour tout $(x, t) \in \Omega$:

$$\begin{cases} - \text{ou bien } \langle \nu, |\lambda - u| \rangle \neq 0 \text{ et } v = f(u) \\ - \text{ou bien } \langle \nu, |\lambda - u| \rangle = 0 \end{cases}$$

La seconde possibilité s'écrit encore:

$$\int_{\mathbb{R}} |\lambda - u(x, t)| d\nu_{x,t}(\lambda) \equiv 0$$

Ceci implique:

$$\nu_{x,t} = \delta_{u(x,t)}$$

Convergence forte et second membre

Pour mettre en évidence la convergence forte, on va montrer que la convexité de la fonction de flux implique que ν est réduite à une masse de Dirac en établissant que son support est contenu dans un intervalle où f est affine. Pour cela, on suppose que $\nu_{x,t}$ est une mesure de probabilité admettant une densité par rapport à la mesure de Lebesgue. Comme dans ce cas $\nu = f(u)$, la relation (2.4) devient:

$$\langle \nu, (\lambda - u)q(\lambda) - (f(\lambda) - f(u))\eta(\lambda) \rangle = 0$$

La densité de $\nu_{x,t}$ s'exprime grâce à une fonction $\mathbb{R} \rightarrow \mathbb{R}$, on note l'enveloppe convexe de son support (qui est borné puisque les u^ϵ vérifient un principe du maximum):

$$\overline{\text{co}}(\text{supp}\{\nu_{x,t}(\lambda)\}) = [a, b]$$

On introduit les mesures suivantes à support dans $[a, b]$:

$$\begin{cases} r'(\lambda) = (\lambda - u)\nu \\ s'(\lambda) = (f(\lambda) - f(u))\nu \end{cases}$$

La relation (2.4) est donc équivalente à:

$$\langle r'(\lambda), q(\lambda) \rangle - \langle s'(\lambda), \eta(\lambda) \rangle = 0 \quad (2.5)$$

En vue d'intégrer par parties cette expression, on introduit les fonctions suivantes:

$$\begin{cases} r(x) = \int_a^x (\lambda - u)d\nu(\lambda) \\ s(x) = \int_a^x (f(\lambda) - f(u))d\nu(\lambda) \end{cases}$$

Le point important est d'assurer que $r(a) = r(b) = s(a) = s(b) = 0$. On a $r(a) = s(a) = 0$ puisque l'on a supposé que ν était à densité. D'autre part, on a:

$$r(b) = \int_a^b (\lambda - u)d\nu(\lambda) = \int_{\mathbb{R}} \lambda d\nu(\lambda) - u \underbrace{|\nu|}_{=1} = \langle \nu, \lambda \rangle - u = 0$$

Pour la même raison, on a $s(b) = 0$. Une intégration par parties dans (2.5) nous donne donc:

$$- \langle r(\lambda), \underbrace{q'(\lambda)}_{\eta'(\lambda)f'(\lambda)} \rangle + \langle s(\lambda), \eta'(\lambda) \rangle = 0$$

Ceci se réécrit sous la forme:

$$\langle (s - r.f')(\lambda), \eta'(\lambda) \rangle = 0$$

Ceci étant valable pour toute fonction croissante η' . Par linéarité, cette relation s'étend à toute fonction régulière, d'où on déduit:

$$s - r.f' \equiv 0$$

D'autre part, on a par construction $r'(\lambda).[f(\lambda) - f(u)] - s'(\lambda).[\lambda - u] = 0$ au sens des mesures. D'où:

$$\frac{d}{d\lambda}(r(\lambda).[f(\lambda) - f(u)] - s(\lambda).[\lambda - u]) = 0$$

Comme $r(a) = s(a) = 0$ et que $s(\lambda) = r(\lambda)f'(\lambda)$, il vient:

$$r(\lambda)[(f(\lambda) - f(u)) - f'(\lambda)(\lambda - u)] = 0$$

Comme par construction de r , on a $r(\lambda) \neq 0$ si $\lambda \in]a, b[$ (sinon le théorème de Rolle assurerait que la fonction *linéaire* $r'(\lambda)$ s'annule en deux points de $]a, b[$), on déduit de cette expression que:

$$\forall \lambda \in]a, b[, f''(\lambda) = 0$$

Pour conclure, comme on a supposé f strictement convexe, il n'existe aucun intervalle non trivial sur lequel f est affine. Donc le support de ν est réduit à un point pour tout (x, t) . On termine en écrivant:

$$\nu_{x,t}(\lambda) = \delta_{u(x,t)}(\lambda) \text{ d'où } u^\epsilon(x, t) \xrightarrow{\epsilon \rightarrow 0} u(x, t) \text{ dans } L^p \text{ (} p < +\infty \text{)}$$

et:

$$(f(u^\epsilon), \eta(u^\epsilon), q(u^\epsilon)) \xrightarrow{*} (f(u), \eta(u), q(u)) \text{ dans } L^\infty \text{ *-faible}$$

□

Une conséquence immédiate de ce résultat fort est la justification de l'écriture de la solution faible de l'EDP hyperbolique en tant que limite visqueuse. En effet, on a supposé $g \in C^1(\mathbb{R})$, donc continue pour la topologie forte de $L^p(\Omega)$ ($p < \infty$). Ceci implique la convergence au sens des distributions:

$$\langle g(u^\epsilon), \varphi \rangle_{\mathcal{D}', \mathcal{D}} \xrightarrow{\epsilon \rightarrow 0} \langle g(u), \varphi \rangle_{\mathcal{D}', \mathcal{D}}$$

Donc $u(x, t) = \lim_{\epsilon \rightarrow 0} u^\epsilon(x, t)$ est solution faible du problème scalaire non-homogène.

Proposition I.7 Soit Ω le demi-espace borné en temps $[0, T] \times \mathbb{R}$.

On suppose acquise la donnée de couples entropiques (η, q) . Alors, si on a:

$$\begin{cases} f \in C^2(\mathbb{R}) \text{ telle que } f'' > 0 \text{ et } f(0) = 0 \\ g \in C^1(\mathbb{R}) \text{ telle que } g(0) = 0; \exists M/|x| > M \Rightarrow x.g(x) < 0 \\ \forall \eta \text{ convexe} : \partial_t \eta(u^\epsilon) + \partial_x q(u^\epsilon) \in \mathcal{K}(H_{loc}^{-1}(\Omega)) \end{cases}$$

l'équation scalaire non-homogène admet une solution faible $L^\infty(\mathbb{R} \times [0, T])$ au sens des distributions:

$$\forall \varphi \in C_0^\infty(\mathbb{R} \times \mathbb{R}^+), \int_{\Omega} u \partial_t \varphi + f(u) \partial_x \varphi + g(u) \varphi dx dt = 0$$

et $(f(u^\epsilon), \eta(u^\epsilon), q(u^\epsilon)) \xrightarrow{*} (f(u), \eta(u), q(u))$ dans L^∞ *-faible.

2.3 Résultats d'entropie

Il reste encore à montrer que les estimations prescrites sur l'équation entropique ont un sens pour les solutions régulières du problème parabolisé. En effet, c'est avec la majoration uniforme sur $\|u^\epsilon\|_\infty$ l'ingrédient indispensable pour appliquer la méthode de compacité par compensation.

2.3.1 Estimations d'entropie

Proposition I.8 Soit le problème parabolisé

$$\begin{cases} \partial_t u^\epsilon + \partial_x f(u^\epsilon) = g(u^\epsilon) + \epsilon \partial_{xx} u^\epsilon \\ u^\epsilon(x, 0) = u_0(x) \in L^2 \cap L^\infty(\mathbb{R}) \end{cases}$$

alors si $\sqrt{\epsilon} \cdot \partial_x u^\epsilon \in L^2(\Omega)$, on a l'estimation d'entropie

$$\partial_t \eta(u^\epsilon) + \partial_x q(u^\epsilon) \in \mathcal{K}(H_{loc}^{-1}(\Omega))$$

Démonstration

Comme u^ϵ est régulière, on peut multiplier par $\eta'(u^\epsilon)$ afin d'obtenir:

$$\partial_t \eta(u^\epsilon) + \partial_x q(u^\epsilon) = \eta'(u^\epsilon) \cdot g(u^\epsilon) + \epsilon \partial_{xx} \eta(u^\epsilon) - \eta''(u^\epsilon) |\sqrt{\epsilon} \cdot \partial_x u^\epsilon|^2$$

D'après le lemme de Murat, il nous suffit de montrer

$$\partial_t \eta(u^\epsilon) + \partial_x q(u^\epsilon) \in (\mathcal{K}(H^{-1}) + \mathcal{B}(\mathcal{M})) \cap \mathcal{B}(W^{-1, \infty})$$

Avec $\mathcal{B}(\mathcal{M})$ signifiant borné dans l'espace des mesures sur Ω .

Tout d'abord, $\partial_t \eta(u^\epsilon) + \partial_x q(u^\epsilon)$ est borné dans $W^{-1,\infty}(\Omega)$ puisque $\|u^\epsilon\|_\infty \leq M$.

D'autre part, on a:

$$\begin{cases} \sqrt{\epsilon} \cdot \partial_x u^\epsilon \in L^2(\Omega) \\ \eta \in C^2(\mathbb{R}) \text{ et } u^\epsilon \in L^\infty(\Omega) \end{cases} \quad \text{d'où } \eta''(u^\epsilon) |\sqrt{\epsilon} \cdot \partial_x u^\epsilon|^2 \text{ borné dans } \mathcal{M}(\Omega) \supset L^1(\Omega)$$

Il reste à montrer que les deux autres termes sont compacts dans $H^{-1}(\Omega)$.

$$\begin{cases} g(u^\epsilon) \in L^\infty(0, T; H^1(\mathbb{R})) \subset L^2(\Omega) \\ \eta'(u^\epsilon) \in L^\infty(\Omega) \end{cases} \quad \text{d'où } \eta'(u^\epsilon)g(u^\epsilon) \text{ borné dans } L^2(\Omega)$$

Ceci implique que $[\eta'(u^\epsilon)g(u^\epsilon)]_\epsilon$ appartient à un compact pour la topologie faible de $L^2(\Omega)$. Par conséquent, il existe $\overline{\eta'g} \in L^2(\Omega)$ tel que:

$$\sup_{\psi \in L^2, \|\psi\|_{L^2} \leq 1} \int_{\Omega} [\eta'(u^\epsilon)g(u^\epsilon) - \overline{\eta'g}] \psi \cdot dx \cdot dt \xrightarrow{\epsilon \rightarrow 0} 0$$

Ceci implique clairement que:

$$\sup_{\varphi \in H_0^1, \|\varphi\|_{H_0^1} \leq 1} \int_{\Omega} [\eta'(u^\epsilon)g(u^\epsilon) - \overline{\eta'g}] \varphi \cdot dx \cdot dt \xrightarrow{\epsilon \rightarrow 0} 0$$

Montrons maintenant que $\|\epsilon \partial_{xx} \eta(u^\epsilon)\|_{H^{-1}} \xrightarrow{\epsilon \rightarrow 0} 0$: la norme de dualité s'écrit

$$\|\epsilon \partial_{xx} \eta(u^\epsilon)\|_{H^{-1}} = \sup_{\varphi \in H_0^1, \|\varphi\|_{H_0^1} \leq 1} \{ \langle \epsilon \partial_{xx} \eta(u^\epsilon), \varphi \rangle_{H^{-1}, H_0^1} = \int_{\Omega} \epsilon \partial_{xx} \eta(u^\epsilon) \cdot \varphi \cdot dx \cdot dt \}$$

Donc

$$\|\epsilon \partial_{xx} \eta(u^\epsilon)\|_{H^{-1}} = \sup_{\varphi \in H_0^1, \|\varphi\|_{H_0^1} \leq 1} \{ \langle \epsilon \partial_x \eta(u^\epsilon), \partial_x \varphi \rangle_{L^2} \}$$

Par l'inégalité de Schwarz

$$\langle \epsilon \partial_x \eta(u^\epsilon), \partial_x \varphi \rangle_{L^2} \leq \sqrt{\epsilon} \|\sqrt{\epsilon} \cdot \partial_x \eta(u^\epsilon)\|_{L^2} \cdot \|\partial_x \varphi\|_{L^2}$$

Et donc $\|\epsilon \partial_{xx} \eta(u^\epsilon)\|_{H^{-1}}$ tend vers 0 quand $\epsilon \rightarrow 0$ puisque l'on a

$$\|\epsilon \partial_{xx} \eta(u^\epsilon)\|_{H^{-1}} \leq \sqrt{\epsilon} \|\eta'(u^\epsilon)\|_{L^\infty} \|\sqrt{\epsilon} \cdot \partial_x u^\epsilon\|_{L^2} \|\varphi\|_{H_0^1}$$

On a donc la conclusion du lemme

$$\partial_t \eta(u^\epsilon) + \partial_x q(u^\epsilon) \in \mathcal{K}(H_{loc}^{-1}(\Omega))$$

□

Remarque: Par convexité de η , la convergence de $\epsilon \partial_{xx} \eta(u^\epsilon) + \eta'(u^\epsilon) \cdot g(u^\epsilon)$ vers $\overline{\eta'g}$ dans $H^{-1}(\Omega)$ et la convergence faible-* dans $L^\infty(\Omega)$ de $\eta(u^\epsilon)$ et $q(u^\epsilon)$ impliquent l'inégalité d'entropie à la limite:

$$- \langle \eta(u), \partial_t \varphi \rangle_{\mathcal{D}', \mathcal{D}} - \langle q(u), \partial_x \varphi \rangle_{\mathcal{D}', \mathcal{D}} \leq \langle \overline{\eta'g}, \varphi \rangle_{\mathcal{D}', \mathcal{D}} \quad \text{pour toute } \varphi \in \mathcal{D}'(\Omega)$$

Ceci donne au bout du compte:

$$\partial_t \eta(u) + \partial_x q(u) \leq \eta'(u) \cdot g(u) \quad (\mathcal{D}')$$

2.3.2 Viscosité et entropie dans le cas non-homogène

On démontre maintenant la dernière hypothèse sur $u^\epsilon(x, t)$ et on donne la caractérisation de la solution limite des problèmes parabolisés en termes d'inégalité d'entropie.

Proposition I.9 *Si $\|u^\epsilon(x, t)\|_{L^\infty(\Omega)}$ est bornée, il en est de même pour $\|\sqrt{\epsilon} \cdot \partial_x u^\epsilon\|_{L^2(\Omega)}$.*

Démonstration

On considère l'EDP visqueuse que l'on multiplie par $\eta'(u^\epsilon)$. Comme u^ϵ est régulière, il vient:

$$\eta'(u^\epsilon) \cdot \partial_t u^\epsilon + \underbrace{\eta'(u^\epsilon) \cdot f'(u^\epsilon)}_{q'(u^\epsilon)} \cdot \partial_x u^\epsilon = \eta'(u^\epsilon) \cdot g(u^\epsilon) + \epsilon \underbrace{\eta'(u^\epsilon) \cdot \partial_{xx} u^\epsilon}_{\partial_{xx} \eta(u^\epsilon) - \eta''(u^\epsilon) (\partial_x u^\epsilon)^2}$$

Et on obtient donc:

$$\partial_t \eta(u^\epsilon) + \partial_x q(u^\epsilon) = \eta'(u^\epsilon) \cdot g(u^\epsilon) + \epsilon \partial_{xx} \eta(u^\epsilon) - \eta''(u^\epsilon) |\sqrt{\epsilon} \cdot \partial_x u^\epsilon|^2$$

On intègre sur la bande $\Omega = \mathbb{R} \times [0, T]$; comme les fonctions de $H^1(\mathbb{R})$ admettent un représentant continu nul à l'infini ([10] p130), il vient:

$$\begin{aligned} \int_{\mathbb{R}} \eta(u^\epsilon(s, T)) - \eta(u_0(s)) \cdot ds + \int_{\Omega} \epsilon \eta''(u^\epsilon) |\partial_x u^\epsilon|^2 \cdot dx \cdot dt + \underbrace{\int_0^T [q(u^\epsilon)]_{-\infty}^{+\infty}}_{\equiv 0} \\ = \int_0^T \epsilon [\partial_x \eta(u^\epsilon)]_{-\infty}^{+\infty} + \int_{\Omega} \eta'(u^\epsilon) \cdot g(u^\epsilon) \cdot dx \cdot dt \end{aligned}$$

Par la convexité de $x \mapsto \eta(x)$, le signe de $\int_{\Omega} \epsilon \eta''(u^\epsilon) |\partial_x u^\epsilon|^2 \cdot dx \cdot dt$ est connu. On écrit donc dans un premier temps que:

$$\int_{\mathbb{R}} [\eta(u^\epsilon(s, T)) - \eta(u_0(s))] \cdot ds \leq \int_{\Omega} \eta'(u^\epsilon) \cdot g(u^\epsilon) \cdot dx \cdot dt + \int_0^T \epsilon [\partial_x \eta(u^\epsilon)]_{-\infty}^{+\infty}$$

Pour évaluer le terme de droite, on commence par choisir une entropie particulière $\eta(u) = u^2$. Ce choix est motivé par le fait que l'on a posé le problème dans $L^2 \cap L^\infty(\mathbb{R})$. On pourrait le considérer dans tous les L^p avec $p = 2k$ afin de pouvoir interpréter $\int_{\mathbb{R}} \eta(u^\epsilon)(s, t) ds$ comme une norme fonctionnelle classique. Ceci nous donne:

$$\|u^\epsilon(\cdot, T)\|_{L^2(\mathbb{R})}^2 - \|u_0\|_{L^2(\mathbb{R})}^2 \leq 2 \int_{\Omega} |u^\epsilon| \cdot |g(u^\epsilon)| \cdot dx \cdot dt + \underbrace{\int_0^T \epsilon [u^\epsilon \partial_x u^\epsilon]_{-\infty}^{+\infty}}_{\equiv 0}$$

Par régularité Lipschitz de $x \mapsto g(x)$, et puisque u^ϵ est uniformément bornée en amplitude, on a:

$$|u^\epsilon| \cdot |g(u^\epsilon)| \leq |u^\epsilon| \cdot |g(u^\epsilon) - g(0)| \leq C |u^\epsilon|^2$$

Dans cette expression, $C = \sup_{|u| \leq \max(M, \|u_0\|_\infty)} |g'(u)|$. On a alors:

$$\|u^\epsilon(\cdot, T)\|_{L^2(\mathbb{R})}^2 \leq 2C \int_0^T \|u^\epsilon(\cdot, t)\|_{L^2(\mathbb{R})}^2 dt + \|u_0\|_{L^2(\mathbb{R})}^2$$

Par le lemme de Gronwall, il est clair que $\|u^\epsilon(\cdot, T)\|_{L^2(\mathbb{R})} \leq e^{CT} \|u_0\|_{L^2(\mathbb{R})}$. Par conséquent, il vient que:

$$\int_{\Omega} \epsilon |\partial_x u^\epsilon|^2 \cdot dx \cdot dt \leq 2 \int_{\Omega} |u^\epsilon| \cdot |g(u^\epsilon)| \cdot dx \cdot dt + \|u_0(s)\|_{L^2(\mathbb{R})}^2$$

Le résultat sur la norme L^2 de u^ϵ s'applique encore et il vient:

$$\int_{\Omega} \epsilon |\partial_x u^\epsilon|^2 \cdot dx \cdot dt \leq 2C \int_0^T e^{2Ct} \|u_0\|_{L^2(\mathbb{R})}^2 dt + \|u_0\|_{L^2(\mathbb{R})}^2$$

On a finalement que:

$$\int_{\Omega} \epsilon |\partial_x u^\epsilon|^2 \cdot dx \cdot dt \leq \|u_0\|_{L^2(\mathbb{R})}^2 e^{2CT}$$

□

On peut donc observer le défaut de compacité dans H^1 de la suite des solutions visqueuses: la norme L^2 du gradient explose lorsque $\epsilon \rightarrow 0$ et on ne peut utiliser le théorème d'injection de Rellich-Kondrakov pour traiter les termes non-linéaires. Ceci est la conséquence de l'apparition de solutions-chocs discontinues.

2.4 Conclusion

Dans ce chapitre, on a montré l'existence globale d'une solution entropique de l'EDP hyperbolique scalaire de convection-réaction modulo une restriction sur le terme source $g(\cdot)$. On peut regrouper tout cela dans un théorème.

Théorème I.8 *Soit (η, q) un couple entropique convexe. Alors le problème*

$$\begin{cases} \partial_t u + \partial_x f(u) = g(u) & (\mathcal{D}') \\ \partial_t \eta(u) + \partial_x q(u) \leq \eta'(u) \cdot g(u) & (\mathcal{D}') \\ u(x, 0) = u_0(x) \in L^2 \cap L^\infty(\mathbb{R}) \\ f \in C^2(\mathbb{R}), f'' > 0 \\ g \in C^1(\mathbb{R}), g(0) = 0; \exists M/|x| > M \Rightarrow x \cdot g(x) < 0 \end{cases}$$

admet une solution $u \in L^2 \cap L^\infty(\mathbb{R} \times [0, T])$. Elle se construit comme limite du problème parabolique associé lorsque le coefficient de viscosité $\epsilon \rightarrow 0$.

Chapitre 3

Entropies de Kruzkov et unicité

Dans ce chapitre, nous établirons, à l'aide des entropies faibles originellement introduites par Kruzkov [53], un résultat de continuité de la solution entropique du problème non-homogène par rapport à sa donnée initiale. On sait par les théorèmes relatifs à l'EDP de la chaleur que l'équation régularisée se résout dans $H^1(\mathbb{R})$ de manière unique [10]. Comme les éventuelles solutions entropiques furent construites sur la base de ce problème parabolique, on peut espérer que l'unicité soit conservée à la limite. Néanmoins, comme on a perdu beaucoup de régularité, il est prévisible que la démonstration n'en soit pas très aisée et fasse grand usage du critère d'entropie. En effet, c'est dans cette inégalité qu'est contenue l'information subsistant des solutions visqueuses.

3.1 L'inégalité fondamentale

3.1.1 Entropies au sens faible

Alors que l'on a à notre disposition la totalité des fonctions η convexes de $C^2(\mathbb{R})$, on ne fera usage que de la famille à un paramètre des entropies de Kruzkov déjà rencontrées dans le cadre de la compacité par compensation au chapitre précédent. Il y avait été occulté leur défaut de régularité ; il faut en fait les considérer au sens faible, i.e. dans le cadre de la théorie des distributions.

$$\eta(u) = |u - k| ; q(u) = \operatorname{sgn}(u - k)(f(u) - f(k))$$

Définition I.4 Une solution $u \in L^\infty(\mathbb{R} \times \mathbb{R}^+)$ est entropique au sens faible si elle vérifie pour toute fonction $\varphi \in C_0^\infty(\mathbb{R} \times \mathbb{R}^+)$ positive et pour tout $k \in \mathbb{R}$:

$$\int_{\mathbb{R} \times \mathbb{R}^+} |u - k| \partial_t \varphi + \operatorname{sgn}(u - k)(f(u) - f(k)) \partial_x \varphi + \operatorname{sgn}(u - k) g(u) \varphi dx dt \geq 0$$

Lemme I.6 Une solution entropique est aussi entropique au sens faible.

Démonstration

On introduit une régularisation $N \in C^\infty(\mathbb{R})$ de $x \mapsto |x - k|$

$$\begin{cases} N(x) = |x| \text{ pour } |x| \geq 1 \\ N'' \geq 0 \end{cases}$$

ainsi que la fonction N_ε telle que:

$$N_\varepsilon(x) = \varepsilon N\left(\frac{x - k}{\varepsilon}\right)$$

On note que $\lim_{\varepsilon \rightarrow 0} N_\varepsilon(x) = |x - k|$ est convexe en tant que limite de fonctions convexes. On a alors, au sens habituel:

$$\partial_t N_\varepsilon(u) + \partial_x Q_\varepsilon(u) \leq N'_\varepsilon(u) \cdot g(u) \quad (\mathcal{D}')$$

Q_ε s'écrit grace à la condition habituelle des couples entropiques $q = \int \eta' \cdot f'$:

$$Q_\varepsilon(u) = \int_k^u N'_\varepsilon(v) f'(v) dv$$

On a les limites ponctuelles suivantes:

$$Q_\varepsilon(u) \xrightarrow{\varepsilon \rightarrow 0} \begin{cases} f(u) - f(k) \text{ si } u \geq k \\ f(k) - f(u) \text{ si } u \leq k \end{cases}$$

Comme la convergence presque-partout implique la convergence au sens des distributions, on peut passer à la limite $\varepsilon \rightarrow 0$ et conclure. \square

3.1.2 Formulation différentielle

Résultats préliminaires

On donne ici une définition utile et des lemmes techniques qui nous serviront dans le paragraphe à venir. On ne précise pas les démonstrations, elles sont disponibles dans [35].

Définition I.5 Une fonction tronqueuse $\Gamma \in C_0^\infty(\mathbb{R}^d)$ est telle que:

$$\left\{ \begin{array}{l} \Gamma \geq 0 \text{ et } \text{supp}(\Gamma) \subset B(0,1)_{\mathbb{R}^d} \\ \int_{\mathbb{R}^d} \Gamma(x).dx = 1 \\ \Gamma(-x) = \Gamma(x) \end{array} \right.$$

On définit alors une approximation de la masse de Dirac:

$$\Gamma_\varepsilon(x) = \frac{1}{\varepsilon^d} \Gamma\left(\frac{x}{\varepsilon}\right)$$

Lemme I.7 Supposons que pour tout $(u, v) \in [-R, R]^2$, la fonction f vérifie la condition de Lipschitz:

$$|f(u) - f(v)| \leq M|u - v|$$

Alors on a pour tout $(u, v, u^*, v^*) \in [-R, R]^4$:

$$|\text{sgn}(u - v)(f(u) - f(v)) - \text{sgn}(u^* - v^*)(f(u^*) - f(v^*))| \leq M(|u - u^*| + |v - v^*|)$$

On poursuit avec un résultat de densité.

Lemme I.8 (Lebesgue)

Soit $u \in L^1_{loc}(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n)$

$$\mathcal{A} = \left\{ x \in \mathbb{R}^n \text{ tels que } \lim_{|Y| \rightarrow 0} \frac{1}{|Y|} \int_Y |u(x+y) - u(x)|.dy = 0 \right\}$$

L'ensemble \mathcal{A} est appelé l'ensemble de Lebesgue de u et est dense dans \mathbb{R}^n .

Le lemme fondamental

On est maintenant en position d'énoncer l'inégalité à l'origine de l'unicité des solutions entropiques. On l'énonce en premier lieu sous sa forme différentielle.

Lemme I.9 Soient u et v deux solutions entropiques dans $L^\infty(\mathbb{R} \times [0, T])$, alors:

$$\partial_t |u - v| + \partial_x (\text{sgn}(u - v)(f(u) - f(v))) \leq \text{sgn}(u - v)(g(u) - g(v)) \quad (\mathcal{D}') \quad (3.1)$$

Démonstration

Pour toutes $\varphi(x, s)$ et $\psi(y, t)$ positives dans $C_0^\infty(\mathbb{R} \times \mathbb{R}^+)$ et pour tout couple $(k, l) \in \mathbb{R}^2$, on a par définition des solutions entropiques:

$$\left\{ \begin{array}{l} \int_{\mathbb{R} \times \mathbb{R}^+} |u - k| \partial_s \varphi + \text{sgn}(u - k)(f(u) - f(k)) \partial_x \varphi + \text{sgn}(u - k)g(u)\varphi \, dx.ds \geq 0 \\ \int_{\mathbb{R} \times \mathbb{R}^+} |v - l| \partial_t \psi + \text{sgn}(v - l)(f(v) - f(l)) \partial_x \psi + \text{sgn}(v - l)g(v)\psi \, dy.dt \geq 0 \end{array} \right. \quad (3.2)$$

On se donne $\Phi(x, y, s, t) \in C_0^\infty(\mathbb{R}^2 \times (\mathbb{R}^+)^2)$ positive:

- on pose $\varphi(x, s) = \Phi(\cdot, y, \cdot, t)$, $k = v(y, t)$ à (y, t) fixés et on écrit la première inégalité de (3.2).
- on pose $\psi(y, t) = \Phi(x, \cdot, s, \cdot)$, $l = u(x, s)$ à (x, s) fixés et on écrit la deuxième inégalité de (3.2).
- on somme les deux inégalités ainsi obtenues et on intègre sur $\mathbb{R}^2 \times (\mathbb{R}^+)^2$. On obtient alors:

$$\int_{(\mathbb{R} \times \mathbb{R}^+)^2} |u(x, s) - v(y, t)| (\partial_s \varphi(x, s) + \partial_t \psi(y, t)) + \text{sgn}(u(x, s) - v(y, t))(f(u) - f(v)) (\partial_x \varphi(x, s) + \partial_y \psi(y, t)) + \text{sgn}(u(x, s) - v(y, t))(g(u)\varphi(x, s) - g(v)\psi(y, t)) \, dx.ds.dy.dt \geq 0$$

On impose maintenant à $\Phi(x, y, s, t)$ d'être de la forme suivante:

$$\Phi(x, y, s, t) = \phi\left(\frac{x+y}{2}, \frac{s+t}{2}\right) \cdot \Gamma_\epsilon\left(\frac{x-y}{2}, \frac{s-t}{2}\right) \text{ avec } 0 \leq \phi(\xi, \tau) \in C_0^\infty(\mathbb{R} \times \mathbb{R}^+)$$

On remarque alors que pour tout $(x, y, s, t) \in \mathbb{R}^2 \times (\mathbb{R}^+)^2$:

$$\begin{cases} \partial_s \varphi = \partial_s \Phi = (\partial_\tau \phi \cdot \Gamma_\epsilon + \phi \cdot \partial_\tau \Gamma_\epsilon)/2 \\ \partial_t \psi = \partial_t \Phi = (\partial_\tau \phi \cdot \Gamma_\epsilon - \phi \cdot \partial_\tau \Gamma_\epsilon)/2 \end{cases} \text{ d'où } \partial_s \varphi + \partial_t \psi = \partial_\tau \phi \cdot \Gamma_\epsilon$$

On obtiendra de même:

$$\partial_x \varphi + \partial_y \psi = \partial_\xi \phi \cdot \Gamma_\epsilon$$

Il devient donc naturel d'effectuer le changement de variables:

$$\begin{aligned} \mathbb{R}^2 \times (\mathbb{R}^+)^2 &\rightarrow \Omega = \{(X, Y, S, T) \in \mathbb{R}^4 \text{ tels que } S+T \geq 0; S-T \geq 0\} \\ (x, y, s, t) &\mapsto X = \frac{x+y}{2}; Y = \frac{x-y}{2}; S = \frac{s+t}{2}; T = \frac{s-t}{2} \end{aligned}$$

On définit alors $G : \Omega \rightarrow \mathbb{R}$ par:

$$\begin{aligned} G(X, Y, S, T) &= |u-v| \partial_\tau \phi(X, S) + \operatorname{sgn}(u-v)(f(u) - f(v)) \partial_x \phi(X, S) \\ &\quad + \operatorname{sgn}(u-v)(g(u) - g(v)) \phi(X, S) \end{aligned}$$

avec $u \equiv u(X+Y, S+T)$ et $v \equiv v(X-Y, S-T)$.

Il s'agit donc d'étudier la limite $\varepsilon \rightarrow 0$ de:

$$J_\varepsilon = \int_\Omega G(X, Y, S, T) \Gamma_\epsilon(Y, T) dX.dY.dS.dT$$

Comme le jacobien du changement de variables vaut $1/4$, on a $J_\varepsilon \geq 0$ en tant que somme de deux quantités positives. D'après sa définition, $\Gamma_\epsilon \rightarrow \delta_0$, mais on ne peut malheureusement pas passer à la limite directement avec la théorie des distributions car G n'est pas régulière.

Néanmoins, le problème est clairement formulé: il sera résolu dès qu'il sera établi que

$$J_\varepsilon \xrightarrow{\varepsilon \rightarrow 0} J_0 = \int_{\mathbb{R} \times \mathbb{R}^+} G(X, 0, S, 0) dX.dS \geq 0$$

On pose $\chi(S, T) = \mathbf{1}_{\{S-T \geq 0; S+T \geq 0\}}$

$$J_\varepsilon = \int_{\mathbb{R} \times \mathbb{R}^+} G(X, Y, S, T) \chi(S, T) \Gamma_\epsilon(Y, T) dX.dY.dS.dT$$

Comme par définition, $\|\Gamma_\epsilon\|_{L^1} = 1$, le théorème de Fubini donne:

$$\begin{aligned} J_0 &= \int_{\mathbb{R}^2} \Gamma_\epsilon(Y, T) dY.dT \times \int_{\mathbb{R}^2} G(X, 0, S, 0) \chi(S, 0) dX.dS \\ &= \int_{\mathbb{R}^4} G(X, 0, S, 0) \chi(S, 0) \Gamma_\epsilon(Y, T) dX.dS.dY.dT \end{aligned}$$

On note par la suite

$$\begin{cases} K = \operatorname{supp}(\phi) \subset \mathbb{R}^2 \\ C_\varepsilon = \operatorname{supp}(\Gamma_\epsilon) \subset \varepsilon \cdot B(0, 1)_{\mathbb{R}^2} \end{cases}$$

Et on obtient:

$$|J_\varepsilon - J_0| \leq \int_K \left[\int_{C_\varepsilon} |G(X, Y, S, T) \chi(S, T) - G(X, 0, S, 0) \chi(S, 0)| \Gamma_\epsilon(Y, T) dY.dT \right] dX.dS$$

On remarque maintenant que pour $(a, b, c, d) \in \mathbb{R}^4$, on a:

$$|ab - cd| \leq |b(a-c) + c(b-d)| \leq |b||a-c| + |c||b-d|$$

Donc si $|b| \leq 1$, il vient:

$$|ab - cd| \leq |a-c| + |c||b-d|$$

C'est-à-dire que $|J_\varepsilon - J_0| \leq A_\varepsilon + B_\varepsilon$ avec

$$\begin{cases} A_\varepsilon = \int_K \left[\int_{C_\varepsilon} |G(X, Y, S, T) - G(X, 0, S, 0)| \Gamma_\varepsilon(Y, T) dY.dT \right] dX.dS \\ B_\varepsilon = \int_K |G(X, 0, S, 0)| \left[\int_{C_\varepsilon} |\chi(S, T) - \chi(S, 0)| \Gamma_\varepsilon(Y, T) dY.dT \right] dX.dS \end{cases}$$

puisque $|\chi(S, T)| \leq 1$ en tant que fonction indicatrice.

D'autre part, d'après le lemme I.7 et les hypothèses de régularité sur f et g , on a les inégalités suivantes:

$$\begin{cases} \|u - v\| - \|u^* - v^*\| \leq \|u - u^*\| + \|v - v^*\| \\ |\operatorname{sgn}(u - v)(f(u) - f(v)) - \operatorname{sgn}(u^* - v^*)(f(u^*) - f(v^*))| \leq M(\|u - u^*\| + \|v - v^*\|) \\ |\operatorname{sgn}(u - v)(g(u) - g(v)) - \operatorname{sgn}(u^* - v^*)(g(u^*) - g(v^*))| \leq N(\|u - u^*\| + \|v - v^*\|) \end{cases}$$

Ces dernières ont pour conséquence:

$$\begin{aligned} |G(X, Y, S, T) - G(X, 0, S, 0)| &\leq (|u(X + Y, S + T) - u(X, S)| + |v(X - Y, S - T) - v(X, S)|) \partial_\tau \phi(X, S) \\ &\quad + M(|u(X + Y, S + T) - u(X, S)| + |v(X - Y, S - T) - v(X, S)|) \partial_x \phi(X, S) \\ &\quad + N(|u(X + Y, S + T) - u(X, S)| + |v(X - Y, S - T) - v(X, S)|) \phi(X, S) \end{aligned}$$

Et comme $\phi \in C_0^\infty(\mathbb{R} \times \mathbb{R}^+)$, il existe une constante $C(f, g, \phi, \|u\|_\infty, \|v\|_\infty)$ telle que:

$$|G(X, Y, S, T) - G(X, 0, S, 0)| \leq C(|u(X + Y, S + T) - u(X, S)| + |v(X - Y, S - T) - v(X, S)|)$$

D'après la définition I.5, il existe une constante $C'(\Gamma)$ telle que $|\Gamma_\varepsilon| \leq \frac{C'}{\varepsilon^2}$. Ceci implique:

$$A_\varepsilon \leq \frac{C.C'}{\varepsilon^2} \int_K dX.dS \left[\int_{C_\varepsilon} (|u(X + Y, S + T) - u(X, S)| + |v(X - Y, S - T) - v(X, S)|) dY.dT \right]$$

On est donc en position d'application du lemme de Lebesgue puisque $|C_\varepsilon| = C''\varepsilon^2$:

$$U_\varepsilon(X, S) = \frac{C'''}{|C_\varepsilon|} \int_{C_\varepsilon} |u(X + Y, S + T) - u(X, S)| dY.dT \xrightarrow{\varepsilon \rightarrow 0} 0$$

D'autre part, $U_\varepsilon \in L^1_{loc}(\mathbb{R} \times \mathbb{R}^+)$ à partir d'un certain rang puisque $\lim_{\varepsilon \rightarrow 0} U_\varepsilon = 0$. Par le théorème de convergence dominée, on déduit que

$$\lim_{\varepsilon \rightarrow 0} \int_K U_\varepsilon(X, S) dX.dS = 0$$

Il en est bien sûr de même pour

$$V_\varepsilon = \frac{C.C'}{\varepsilon^2} \int_{C_\varepsilon} |v(X + Y, S + T) - v(X, S)| dY.dT$$

Et finalement

$$\lim_{\varepsilon \rightarrow 0} A_\varepsilon = 0$$

Le cas de B_ε est bien plus simple. Comme $\|\Gamma_\varepsilon\|_\infty \leq \|\Gamma\|_\infty/\varepsilon^2$, on a:

$$\begin{aligned} \int_{C_\varepsilon} |\chi(S, T) - \chi(S, 0)| \Gamma_\varepsilon(Y, T) dY.dT &\leq \frac{\|\Gamma\|_\infty}{\varepsilon^2} \int_{C_\varepsilon} |\chi(S, T) - \chi(S, 0)| dY.dT \\ &\leq \frac{\|\Gamma\|_\infty}{\varepsilon} \int_{-\varepsilon}^\varepsilon |\chi(S, T) - \chi(S, 0)| dT \end{aligned}$$

Comme $G(X, 0, S, 0)$ et K sont bornés, cette majoration entraîne que:

$$B_\varepsilon \leq \frac{\|G\|_\infty \cdot |K| \cdot \|\Gamma\|_\infty}{\varepsilon} \int_{\mathbb{R}} dS \left[\int_{-\varepsilon}^\varepsilon |\chi(S, T) - \chi(S, 0)| dY.dT \right]$$

Géométriquement, cette intégrale correspond à la surface des triangles rectangles isocèles de bases $|S| = \varepsilon$ et de hauteur $T = \varepsilon$.

$$B_\varepsilon \leq \frac{\|G\|_\infty \cdot |K| \cdot \|\Gamma\|_\infty \varepsilon^2}{\varepsilon \cdot 2} \xrightarrow{\varepsilon \rightarrow 0} 0$$

Et finalement, on obtient le résultat:

$$|J_\varepsilon - J_0| \leq A_\varepsilon + B_\varepsilon \xrightarrow{\varepsilon \rightarrow 0} 0$$

Ce qui signifie qu'on a montré que l'inégalité est conservée à la limite quand $\varepsilon \rightarrow 0$.

$$\int_{\mathbb{R} \times \mathbb{R}^+} [|u - v| \partial_t \phi + \operatorname{sgn}(u - v)(f(u) - f(v)) \partial_x \phi + \operatorname{sgn}(u - v)(g(u) - g(v)) \phi](x, t) dx dt \geq 0$$

□
Comme on s'est borné à la théorie L^∞ via la compacité par compensation, le critère d'entropie a surtout servi à montrer que la suite des phrases entropiques était compacte dans H_{loc}^{-1} . Ici, nous voyons qu'il permet aussi de mesurer l'éloignement (en un certain sens) de deux solutions entropiques du problème non-homogène. Ceci entraîne que celles-ci sont astreintes à un comportement ordonné. Il est impossible que des données initiales proches donnent des trajectoires très divergentes dans l'espace des phases.

3.2 Le théorème de continuité

Comme cela a déjà été suggéré par la remarque précédente, il se trouve que l'on peut effectivement montrer un résultat de continuité par rapport aux données initiales pour les solutions de ces équations d'évolution. Il suffit pour cela d'intégrer l'inégalité précédente dans le cône de propagation de u et v .

3.2.1 Formulation intégrale et continuité L_{loc}^1

Proposition I.10 Soient u, v deux solutions entropiques issues de $u_0, v_0 \in L^2 \cap L^\infty(\mathbb{R})$ respectivement. Alors si on pose

$$\begin{cases} M = \max\{|f'(\xi)|, |\xi| \leq \max(\|u\|_\infty, \|v\|_\infty)\} \\ N = \max\{|g'(\xi)|, |\xi| \leq \max(\|u\|_\infty, \|v\|_\infty)\} \end{cases} \quad (3.3)$$

on a: $\forall R > 0, \forall T > 0$

$$\int_{|x| \leq R} |u(x, T) - v(x, T)| dx \leq e^{NT} \int_{|x| \leq R+MT} |u_0(x) - v_0(x)| dx$$

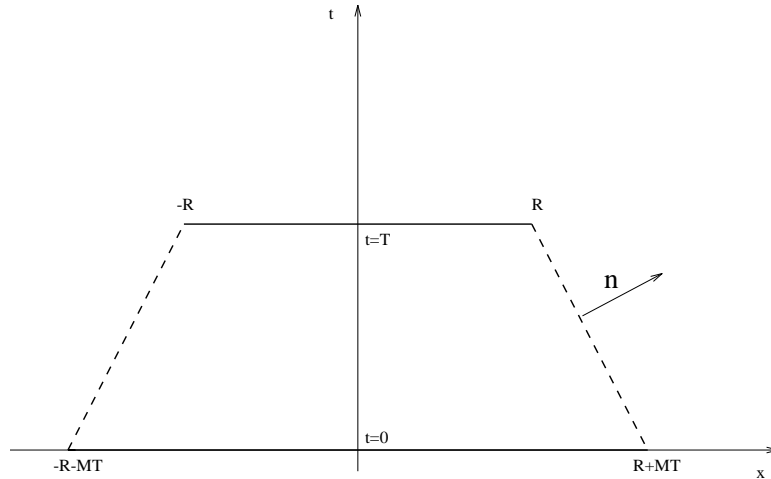


Figure 3.1: Cône de dépendance

Démonstration

L'idée est d'intégrer l'inégalité (3.1) dans le cône de dépendance (voir fig. 3.1) défini par:

$$D_{R,T} = \{(x, t) \in \mathbb{R} \times [0, T], |x| \leq R + M(T - t)\}$$

Pour cela, on commence par écrire que:

$$\begin{aligned} \partial_t |u - v| + \partial_x (\operatorname{sgn}(u - v)(f(u) - f(v))) &\leq \operatorname{sgn}(u - v)(g(u) - g(v)) \\ &\leq |(g(u) - g(v))| \\ &\leq N|u - v| \end{aligned}$$

En remarquant que:

$$\partial_t |u - v| - N|u - v| = \mathbf{e}^{-Nt} \partial_t [|u - v| \mathbf{e}^{-Nt}]$$

l'inégalité (3.1) devient:

$$\partial_t [|u - v| \mathbf{e}^{-Nt}] + \mathbf{e}^{-Nt} [\partial_x (\operatorname{sgn}(u - v)(f(u) - f(v)))] \leq 0$$

Pour toute fonction-test $\phi(x, t) \in C_0^\infty(\mathbb{R} \times \mathbb{R}^+)$, on a:

$$\int_{\mathbb{R} \times \mathbb{R}^+} |u - v| \mathbf{e}^{-Nt} \partial_t \phi + (\operatorname{sgn}(u - v)(f(u) - f(v))) \mathbf{e}^{-Nt} \partial_x \phi \, dx \, dt \geq 0 \quad (3.4)$$

On pose maintenant formellement $\phi \equiv \mathbf{1}_{D_{R,T}}$ qui n'appartient bien sûr pas à $C_0^\infty(\mathbb{R} \times \mathbb{R}^+)$. Si l'on introduit la fonction de Heaviside $x \mapsto Y(x) = \mathbf{1}_{x \geq 0}$, on a:

$$\phi(x, t) = Y(t) \cdot (1 - Y(t - T)) \cdot Y(x + R + M(T - t)) \cdot (1 - Y(x - R - M(T - t)))$$

Il faut donc calculer les dérivées partielles de $\phi(x, t)$ au sens des distributions puisque celle-ci est discontinue. Après élimination des termes inutiles, on trouve:

$$\begin{aligned} \partial_t \phi(x, t) &= \delta_{t=0} \cdot Y(x + R + MT) \cdot (1 - Y(x - R - MT)) \\ &\quad - \delta_{t=T} \cdot Y(x + R) \cdot (1 - Y(x - R)) \\ &\quad - M \delta_{x=-R-M(T-t)} \cdot Y(t) \cdot (1 - Y(t - T)) \\ &\quad - M \delta_{x=R+M(T-t)} \cdot Y(t) \cdot (1 - Y(t - T)) \\ \partial_x \phi(x, t) &= \delta_{x=-R-M(T-t)} \cdot Y(t) \cdot (1 - Y(t - T)) \\ &\quad - \delta_{x=R+M(T-t)} \cdot Y(t) \cdot (1 - Y(t - T)) \end{aligned}$$

Si l'on calcule formellement l'inégalité (3.4) avec cette fonction-test, il vient:

$$\begin{aligned} &\int_{-R-MT}^{R+MT} |u_0 - v_0|(s) \, ds - \mathbf{e}^{-NT} \int_{-R}^R |u - v|(s, T) \, ds \\ &+ \int_0^T -M|u - v|(-R - M(T - t), t) \mathbf{e}^{-Nt} + (\operatorname{sgn}(u - v)(f(u) - f(v)))(-R - M(T - t), t) \mathbf{e}^{-Nt} \, dt \\ &+ \int_0^T -M|u - v|(R + M(T - t), t) \mathbf{e}^{-Nt} - (\operatorname{sgn}(u - v)(f(u) - f(v)))(R + M(T - t), t) \mathbf{e}^{-Nt} \, dt \geq 0 \end{aligned} \quad (3.5)$$

Au vu de cette expression, on aura la conclusion du théorème si on établit que les deux derniers termes sont négatifs. Mais ceci est immédiat compte tenu de (3.3); en effet, on a:

$$|\operatorname{sgn}(u - v)(f(u) - f(v))| \leq M|u - v| \text{ puisque } |f(u) - f(v)| \leq M|u - v|$$

Donc la troisième intégrale de (3.5) est négative. Pour la quatrième, il suffit de remarquer que:

$$-\operatorname{sgn}(u - v)(f(u) - f(v)) \leq |\operatorname{sgn}(u - v)(f(u) - f(v))|$$

Il est donc clair que le problème est d'établir rigoureusement l'inégalité (3.4) avec $\phi \equiv \mathbf{1}_{D_{R,T}} \notin C_0^\infty(\mathbb{R} \times \mathbb{R}^+)$. Il reste donc à en prendre une régularisation et essayer de conserver l'inégalité par passage à la limite. Soit Γ_ϵ la masse de Dirac régularisée, on définit la fonction de Heaviside approchée Y_ϵ comme:

$$Y_\epsilon(t) = \int_{-\infty}^t \Gamma_\epsilon(s) \, ds$$

C'est-à-dire que $Y_\varepsilon(t)$ vaut approximativement:

- 0 si $t \leq -\varepsilon$
- $(t + \varepsilon)/2\varepsilon$ si $-\varepsilon \leq t \leq \varepsilon$
- 1 si $t \geq \varepsilon$

Comme pour la démonstration de l'inégalité (3.1), il s'agit de passer à la limite en utilisant des fonctions-test particulières. On impose donc à $\phi(x, t)$ d'être de la forme:

$$\phi(x, t) = [Y_\varepsilon(t - \delta) - Y_\varepsilon(t - T)][1 - Y_\theta(|x| - R - M(T - t))] \geq 0$$

C'est-à-dire:

$$\phi(x, t) = \begin{cases} 1 & \text{si } |x| \leq R + M(T - t) - \theta \text{ et } \varepsilon + \delta \leq t \leq T - \varepsilon \\ 0 & \text{si } |x| \geq R + M(T - t) - \theta \text{ ou } t \leq \delta - \varepsilon \text{ ou } t \geq T + \varepsilon \end{cases}$$

On suppose $\varepsilon < \delta$ et on intègre l'inégalité (3.1) avec $\phi(x, t)$ ainsi définie. Auparavant, on calcule ses dérivées partielles pour $x \neq 0$:

$$\begin{cases} \partial_t \phi(x, t) = [\Gamma_\varepsilon(t - \delta) - \Gamma_\varepsilon(t - T)][1 - Y_\theta(|x| - R - M(T - t))] \\ \quad - M \cdot [Y_\varepsilon(t - \delta) - Y_\varepsilon(t - T)][\Gamma_\theta(|x| - R - M(T - t))] \\ \partial_x \phi(x, t) = \frac{-x}{|x|} [Y_\varepsilon(t - \delta) - Y_\varepsilon(t - T)][\Gamma_\theta(|x| - R - M(T - t))] \end{cases}$$

On écrit de nouveau l'expression (3.1):

$$\int_{\mathbb{R} \times \mathbb{R}^+} |u - v| \partial_t \phi + \operatorname{sgn}(u - v)(f(u) - f(v)) \partial_x \phi + \operatorname{sgn}(u - v)(g(u) - g(v)) \phi \, dx \, dt \geq 0$$

Ceci entraîne l'expression (3.4):

$$\int_{\mathbb{R} \times \mathbb{R}^+} |u - v| e^{-Nt} \partial_t \phi + (\operatorname{sgn}(u - v)(f(u) - f(v))) e^{-Nt} \partial_x \phi \, dx \, dt \geq 0$$

On remplace dans cette expression les valeurs de $\partial_t \phi(x, t)$ et $\partial_x \phi(x, t)$ par celles données précédemment. Après avoir regroupé les termes, il vient:

$$\begin{aligned} 0 \leq & \int_{\mathbb{R} \times \mathbb{R}^+} |u - v| e^{-Nt} \cdot [\Gamma_\varepsilon(t - \delta) - \Gamma_\varepsilon(t - T)][1 - Y_\theta(|x| - R - M(T - t))] \, dx \, dt \\ - & \int_{\mathbb{R} \times \mathbb{R}^+} \underbrace{\left[M \cdot |u - v| + \frac{-x}{|x|} \operatorname{sgn}(u - v)(f(u) - f(v)) \right]}_{=A} e^{-Nt} \cdot \underbrace{[Y_\varepsilon(t - \delta) - Y_\varepsilon(t - T)]}_{=B} \cdot \underbrace{[\Gamma_\theta(|x| - R - M(T - t))]}_{=C} \, dx \, dt \end{aligned}$$

Tout comme précédemment, on va montrer que la seconde intégrale peut être enlevée de l'expression ci-dessus:

- la définition de M assure que $A \geq 0$.

- on calcule $Y_\varepsilon(t - \delta) - Y_\varepsilon(t - T) = \int_{-\infty}^{t-\delta} \Gamma_\varepsilon(s) ds - \int_{-\infty}^{t-T} \Gamma_\varepsilon(s) ds$. Comme $T > \delta$ et $\Gamma_\varepsilon(x) \geq 0$ pour tout x , on a $B = \int_{t-T}^{t-\delta} \Gamma_\varepsilon(s) ds \geq 0$.

- par définition de Γ_ε , le terme C est positif.

Par conséquent, la seconde intégrale est positive. Il est donc suffisant de traiter avec:

$$\int_{\mathbb{R} \times \mathbb{R}^+} \underbrace{|u - v| e^{-Nt}}_{\in L^\infty(\mathbb{R} \times \mathbb{R}^+)} \underbrace{(\Gamma_\varepsilon(t - \delta) - \Gamma_\varepsilon(t - T))(1 - Y_\theta(|x| - R - M(T - t)))}_{C^\infty \text{ à support compact}} \, dx \, dt \geq 0$$

On fait $\theta \rightarrow 0$ dans un premier temps, ce qui entraîne $Y_\theta \rightarrow \mathbf{1}_{|x| - R - M(T - t) \leq 0}$ presque-partout. Le théorème de Lebesgue permet de conclure à la convergence des intégrales puisque $|1 - Y_\theta(\cdot)| \leq 1$.

Posons:

$$D_{R, T}^\varepsilon = \{(x, t) \in \mathbb{R} \times [0, T], |x| \leq R + M(T - t); \delta - \varepsilon < t \leq T + \varepsilon\}$$

Alors on obtient à convergence que:

$$\int_{D_{R, T}^\varepsilon} |u - v| e^{-Nt} \underbrace{(\Gamma_\varepsilon(t - \delta) - \Gamma_\varepsilon(t - T))}_{\text{indépendant de } x} \, dx \, dt \geq 0$$

On peut réécrire ceci sous la forme suivante en appliquant le théorème de Fubini:

$$\int_{\mathbb{R}^+} \left[(\Gamma_\epsilon(t - \delta) - \Gamma_\epsilon(t - T)) \left[\int_{-R-M(T-t)}^{R+M(T-t)} |u - v| e^{-Nt} dx \right] \right] dt \geq 0$$

On définit alors une fonction $w : t \mapsto \int_{-R-MT}^{R+MT} |u - v| e^{-Nt} dx$. L'intégrale ci-dessus est alors la différence de deux produits de convolution. C'est-à-dire:

$$w_\epsilon(\delta) - w_\epsilon(T) \geq 0 \text{ avec } w_\epsilon \equiv w * \Gamma_\epsilon$$

Comme $w \in L^1_{loc}(\mathbb{R}^+)$, $w_\epsilon \xrightarrow{\epsilon \rightarrow 0} w$ en chaque point de Lebesgue de w . Ces points forment un ensemble dense dans \mathbb{R}^+ , donc pour presque tous les δ et T , on obtient:

$$w_\epsilon(\delta) - w_\epsilon(T) \xrightarrow{\epsilon \rightarrow 0} \int_{|x| \leq R+M(t-\delta)} |u(x, \delta) - v(x, \delta)| e^{-N\delta} dx - \int_{|x| \leq R} |u(x, T) - v(x, T)| e^{-NT} dx$$

Il reste à passer à la limite $\delta \rightarrow 0$. Pour cela, on écrit que:

$$\begin{aligned} \int_{|x| \leq R+M(t-\delta)} |u(x, \delta) - v(x, \delta)| e^{-N\delta} dx &\leq \int_{|x| \leq R+M(t-\delta)} |u(x, \delta) - u_0(x)| e^{-N\delta} dx \\ &\quad + \int_{|x| \leq R+M(t-\delta)} |u_0(x) - v_0(x)| e^{-N\delta} dx \\ &\quad + \int_{|x| \leq R+M(t-\delta)} |v_0(x) - v(x, \delta)| e^{-N\delta} dx \end{aligned}$$

Les premiers et troisièmes termes tendent vers 0 quand $\delta \rightarrow 0$ par le théorème de Lebesgue. Il reste au bout du compte:

$$\int_{|x| \leq R} |u(x, T) - v(x, T)| dx \leq e^{NT} \int_{|x| \leq R+MT} |u_0(x) - v_0(x)| dx$$

□

Cet important résultat de continuité a donc en particulier pour conséquence d'établir l'unicité de la solution entropique de notre problème hyperbolique non-homogène.

Proposition I.11 *Plaçons-nous dans les conditions d'applications du théorème d'existence pour le problème donné par*

$$\begin{cases} \partial_t u + \partial_x f(u) = g(u) \quad (\mathcal{D}') \\ \partial_t \eta(u) + \partial_x q(u) \leq \eta'(u) \cdot g(u) \quad (\mathcal{D}') \\ u(x, 0) = u_0(x) \in L^2 \cap L^\infty(\mathbb{R}) \end{cases}$$

Alors, celui-ci admet une unique solution entropique $u(x, t) \in L^\infty(\mathbb{R} \times [0, T])$.

3.2.2 Le résultat final

On énonce ici le théorème concluant ces deux chapitres consacrés à la construction théorique de la solution entropique d'une EDP hyperbolique de convection-réaction. C'est l'aboutissement de notre approche L^∞ .

Théorème I.9 *Soit le problème scalaire générique:*

$$\begin{cases} \partial_t u + \partial_x f(u) = g(u) \quad (\mathcal{D}') \\ \partial_t \eta(u) + \partial_x q(u) \leq \eta'(u) \cdot g(u) \quad (\mathcal{D}') \\ u(x, 0) = u_0(x) \in L^2 \cap L^\infty(\mathbb{R}) \\ f \in C^2(\mathbb{R}), f'' > 0 \\ g \in C^1(\mathbb{R}), g(0) = 0; \exists M/|x| > M \Rightarrow x \cdot g(x) < 0 \end{cases}$$

Alors sa solution entropique $u(x, t)$ est définie de manière unique dans $L^\infty(\mathbb{R} \times [0, T])$. De plus, elle est continue en norme $L^1_{loc}(\mathbb{R} \times \mathbb{R}^+)$ par rapport à sa donnée initiale.

Soient u, v deux solutions entropiques issues de $u_0, v_0 \in L^2 \cap L^\infty(\mathbb{R})$ respectivement.

Alors si on pose

$$\begin{cases} M = \max\{|f'(\xi)|, |\xi| \leq \max(\|u\|_\infty, \|v\|_\infty)\} \\ N = \max\{|g'(\xi)|, |\xi| \leq \max(\|u\|_\infty, \|v\|_\infty)\} \end{cases}$$

on a: $\forall R > 0, \forall t \in [0, T]$

$$\int_{|x| \leq R} |u(x, t) - v(x, t)| dx \leq e^{Nt} \int_{|x| \leq R+Mt} |u_0(x) - v_0(x)| dx$$

Ceci regroupe tout ce qui a pu être démontré dans ces deux premiers chapitres: l'existence et la dépendance continue des solutions d'une EDP scalaire hyperbolique réactive. L'extrême lourdeur des démonstrations qui conduisent à ces conclusions trouve son origine dans la pauvreté de ces fonctions en termes de régularité (au contraire des problèmes paraboliques, on n'a *jamais* de solutions fortes).

Remarques

- La méthode de compacité par compensation n'est pas l'unique moyen d'obtenir ce résultat, et on peut lui reprocher de laisser dans l'ombre les propriétés de régularité BV [23] de la solution faible. On peut consulter [35, 53] qui exploitent la compacité de l'injection $BV \hookrightarrow L^1$. Voir aussi [79] pour la convergence d'un schéma aux différences finies.
- On peut remarquer que la présence de l'exponentielle venant du terme source implique que la solution de l'équation n'est plus donnée par un semi-groupe de contractions.
- Dans [26], R. DiPerna a introduit le formalisme des solutions à valeur mesure pour les lois de conservation scalaires homogènes.
- Le comportement en grand temps de $u(x, t)$ est précisé par exemple dans [31]. (voir [57] pour des résultats dans le cas homogène)
- Le problème aux limites a été étudié dans [5] dans le cadre des fonctions BV . Il a été récemment étendu aux fonctions L^∞ [70].
- Pour finir, un $g(u)$ admettant plusieurs équilibres stables peut imposer une solution de type *couches-limites*. On peut consulter [22] comme point de départ dans cette direction.

Chapitre 4

Existence de solutions pour le modèle diphase équilibré

On considère maintenant le système des équations d'Euler isentropiques pour un mélange de deux espèces réactives. Ce modèle suffisamment général a déjà été abordé dans plusieurs publications; citons [8, 24, 55, 81]. Il peut s'écrire sous une forme quasi-conservative:

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0 \\ \partial_t(\rho c) + \partial_x(\rho u c) = \rho g(c) \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p) = 0 \end{cases} \quad (4.1)$$

où ρ, c, u désignent respectivement la densité globale du mélange considéré, la concentration d'une des phases mises en jeu dans l'écoulement, et la vitesse volumétrique commune aux deux espèces. La loi de fermeture est la pression p dépendant en général du couple (ρ, c) . La première équation exprime la conservation de la masse globale, la seconde décrit l'évolution d'une des deux fractions massiques soumise à un processus réactionnel exprimé par le terme source et au transport du mélange à la vitesse volumétrique u . Enfin la dernière régit la variation de quantité de mouvement en fonction de la somme des forces appliquées au système, celles-ci s'exprimant par la variation de pression. On ne met pas de terme au second membre de l'équation sur la quantité de mouvement: ce type de difficulté ayant déjà été abordé par exemple dans [14] pour le p -système. Ce problème est alors non-linéaire, strictement hyperbolique mais non-homogène. On désire en étudier le problème de Cauchy par l'intermédiaire de sa formulation en variables de Lagrange. On remarque alors que les entropies qui lui sont associées ont une structure particulière très proche de celles du p -système classique [57]. On adopte alors la méthode de viscosité évanescence. Dans ce cadre, il est nécessaire d'obtenir une estimation uniforme en amplitude sur la suite d'approximations régulières. Les résultats de Chueh, Conley et Smoller [79] ainsi que le théorème de DiPerna [25] permettent de conclure à l'existence de solutions pour certaines lois de pression particulières.

4.1 Formulation lagrangienne

4.1.1 Changement de variables

Vu le grand nombre de non-linéarités contenues dans le système considéré, on décide de le réécrire dans un nouveau système de coordonnées, afin de se ramener à l'expression d'un problème moins difficile. Soit le système différentiel

$$\begin{cases} \dot{x} = u(x, t) \\ x(0) = \xi \end{cases}$$

On note $t \mapsto x(\xi, t)$ la solution issue de ξ . Alors (ξ, t) sont les coordonnées de Lagrange associées au champ $u(x, t)$. Posons $J(\xi, t) = \frac{\partial x(\xi, t)}{\partial \xi}$, on trouve que $J \neq 0$ et

$$\frac{\partial J(\xi, t)}{\partial t} = \frac{\partial}{\partial \xi} \cdot \underbrace{\frac{\partial x(\xi, t)}{\partial t}}_{u(x(\xi, t), t)} = \underbrace{\frac{\partial x(\xi, t)}{\partial \xi}}_{J(\xi, t)} \cdot \frac{\partial u(x(\xi, t), t)}{\partial x}$$

D'où il vient:

$$\frac{\partial J}{\partial t}(\xi, t) = J(\xi, t) \cdot \frac{\partial u}{\partial x}(x(\xi, t), t)$$

On exprime maintenant une fonction $f(x, t)$ donnée en variables habituelles (Euler) en coordonnées de Lagrange: $f_L(\xi, t) \stackrel{\text{def}}{=} f(x(\xi, t), t)$. On remarque que l'on a:

$$\partial_t(f_L \cdot J) = J \cdot \underbrace{\partial_t(f_L)}_{\partial_t(f) + u \cdot \partial_x(f)} + f_L \cdot \underbrace{\partial_t J}_{J \cdot \partial_x(u)} = J[\partial_t(f) + \partial_x(f \cdot u)]$$

Le système (4.1) devient donc:

$$\begin{cases} \partial_t(\rho_L \cdot J) = J[\partial_t(\rho) + \partial_x(\rho u)] = 0 \\ \partial_t((\rho c)_L \cdot J) = (\rho \cdot g(c))_L \cdot J \\ \partial_t((\rho u)_L \cdot J) + J \partial_x(p_L) = 0 \end{cases}$$

On note que $\partial_{\xi} \cdot = \partial_{\xi}(x) \partial_x \cdot = J \partial_x \cdot$. Le système précédent se réécrit:

$$\begin{cases} \partial_t(\rho_L \cdot J) = 0 \\ (\partial_t c_L)(\rho_L \cdot J) + c_L \cdot \partial_t(\rho_L \cdot J) = (\rho \cdot g(c))_L \cdot J \\ \partial_t((\rho u)_L \cdot J) + \partial_{\xi}(p_L) = 0 \end{cases}$$

La première équation implique $\rho_L \cdot J = \rho_0(\xi)$, ce qui donne $J = \rho_0 \cdot \tau$ avec $\tau = (\rho_L)^{-1}$ désignant le volume spécifique. On en déduit:

$$\underbrace{\partial_t J}_{\rho_0 \cdot \partial_t \tau} = J \cdot \partial_x u = \partial_{\xi} u$$

Cela nous donne la première loi (linéaire) du système Lagrangien (on supprime les "L")

$$\rho_0 \cdot \partial_t \tau - \partial_{\xi} u = 0$$

Comme $\rho \cdot J \equiv \rho_0$, les deux dernières donnent

$$\begin{cases} \rho_0 \cdot \partial_t c = \rho_0 g(c) \\ \rho_0 \cdot \partial_t u + \partial_{\xi} p = 0 \end{cases}$$

Si on introduit la notation $\tilde{\tau}$ pour la pression exprimée en fonction de τ et la variable d'impulsion $dm = \rho_0 d\xi$, on trouve finalement:

$$\begin{cases} \partial_t \tau - \partial_m u = 0 \\ \partial_t u + \partial_m \tilde{p}(\tau, c) = 0 \\ \partial_t c = g(c) \end{cases} \quad (4.2)$$

C'est avec cette écriture que nous allons entamer l'étude du système (4.1). Bien sûr, il y a une restriction: pour définir τ , il est nécessaire de supposer que $\rho \geq \alpha > 0$. C'est-à-dire que le vide n'apparaît jamais.

4.1.2 Propriétés élémentaires

Nous donnons dans ce paragraphe les premières caractéristiques du système (4.2). Par commodité, nous pouvons le réécrire avec des notations plus habituelles sous la forme:

$$\begin{cases} \partial_t v - \partial_x u = 0 \\ \partial_t u + \partial_x p(v, c) = 0 \\ \partial_t c = g(c) \end{cases} \quad (4.3)$$

On fait les hypothèses suivantes sur la non-linéarité $p(v, c)$: $p_v < 0$ et $p_{vv} > 0$.

Si on pose

$$U = \begin{pmatrix} v \\ u \\ c \end{pmatrix} \quad F(U) = \begin{pmatrix} -u \\ p(v, c) \\ 0 \end{pmatrix} \quad S(U) = \begin{pmatrix} 0 \\ 0 \\ g(c) \end{pmatrix}$$

le problème (4.3) s'écrit alors

$$U_t + F(U)_x = S(U)$$

Sa stricte hyperbolicité est assurée par nos hypothèses puisque

$$dF(U) = \begin{pmatrix} 0 & -1 & 0 \\ p_v & 0 & p_c \\ 0 & 0 & 0 \end{pmatrix} \quad \text{ce qui donne le spectre } \Lambda = \{-\sqrt{-p_v(v, c)}; 0; \sqrt{-p_v(v, c)} \stackrel{\text{def}}{=} \lambda\}$$

Les vecteurs propres à droite associés à $0, -\lambda, \lambda$ sont:

$$\vec{R}_0(U) = \begin{pmatrix} p_c \\ 0 \\ \lambda^2 \end{pmatrix} \quad \vec{R}_{-1}(U) = \begin{pmatrix} 1 \\ \lambda \\ 0 \end{pmatrix} \quad \vec{R}_1(U) = \begin{pmatrix} -1 \\ \lambda \\ 0 \end{pmatrix} \quad (4.4)$$

Ceci implique que deux champs du système (4.3) sont **vraiment non-linéaires** (VNL). Le dernier est évidemment **linéairement dégénéré** (LD) puisqu'il est stationnaire. Il reste à exprimer les vecteurs propres à gauche:

$$\vec{L}_0(U) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \vec{L}_{-1}(U) = \begin{pmatrix} \lambda^2 \\ \lambda \\ -p_c \end{pmatrix} \quad \vec{L}_1(U) = \begin{pmatrix} -\lambda^2 \\ \lambda \\ p_c \end{pmatrix} \quad (4.5)$$

Quant aux invariants de Riemann, i.e. les fonctions $w(U)$ vérifiant $\langle \nabla w(U), \vec{R}_i \rangle_{\mathbb{R}^3} = 0$, on sait qu'il y en a deux par vecteur propre [79]; leur expression est:

$$\begin{cases} \vec{R}_0(U) : w = u \text{ ou } w = p(v, c) \\ \vec{R}_{-1}(U) : w = c \text{ ou } w = u - \int^v \lambda(s, c) ds \\ \vec{R}_1(U) : w = c \text{ ou } w = u + \int^v \lambda(s, c) ds \end{cases}$$

4.2 Le problème de Cauchy

4.2.1 Expression des entropies

Dans ce paragraphe, nous allons mettre en évidence la structure particulière des entropies du système (4.3). Rappelons que leur existence en grand nombre est fondamentale pour l'étude des solutions du p -système:

$$\begin{cases} \partial_t v - \partial_x u = 0 \\ \partial_t u + \partial_x v^{-\gamma} = 0 \end{cases} \quad (4.6)$$

par la méthode de compacité par compensation [25, 76]. On peut consulter [57, 79] pour une expression des entropies de (4.6). En fait, il se trouve que le système (4.3) en possède autant. Pour voir cela, on considère sa version homogène:

$$\begin{cases} \partial_t v - \partial_x u = 0 \\ \partial_t u + \partial_x p(v, c) = 0 \\ \partial_t c = 0 \end{cases} \quad (4.7)$$

On remarque que si l'on impose comme donnée initiale à ce dernier une concentration uniforme $c_0(x) \equiv C_0 \in [0, 1]$, on résout en fait le problème suivant:

$$\begin{cases} \partial_t v - \partial_x u = 0 \\ \partial_t u + \partial_x p(v, C_0) = 0 \end{cases} \quad (4.8)$$

L'idée est donc d'exprimer les entropies de (4.8) en fonction de la pression $p(v, C_0)$. Sous les hypothèses classiques $p_v < 0$, $p_{vv} > 0$, ce changement de variables est bijectif $\mathbb{R}_*^+ \rightarrow \mathbb{R}_*^+$ et on espère que les entropies du système complet (4.3) s'exprimeront en remplaçant $p(v, C_0)$ par $p(v, c)$.

Lemme I.10 *Tout couple entropique du p -système (4.8) $(\eta, q) \in C^2(\mathbb{R}_*^+ \times \mathbb{R})$ ne dépendant que de p et u définit un couple $(\tilde{\eta}, \tilde{q}) \in C^2(\mathbb{R}_*^+ \times \mathbb{R} \times [0, 1])$ pour le problème (4.7).*

Démonstration

Supposons les hypothèses habituelles [79] sur $v \mapsto p(v, c)$ pour (4.7): $p_v < 0$ et $p_{vv} > 0$. Alors, pour tout $C_0 \in [0, 1]$, l'application:

$$V : U := (v, u) \mapsto (p(v, C_0), u)$$

définit un difféomorphisme de $\mathbb{R}_*^+ \times \mathbb{R}$ sur lui-même. On peut donc considérer $(\bar{\eta}, \bar{q})$ tel que:

$$\begin{cases} \eta = \bar{\eta} \circ V \\ q = \bar{q} \circ V \end{cases}$$

avec:

$$\begin{cases} \nabla \eta = \frac{\partial V}{\partial U} \nabla \bar{\eta} \\ \nabla q = \frac{\partial V}{\partial U} \nabla \bar{q} \end{cases} \quad \text{et} \quad \left| \frac{\partial V}{\partial U} \right| = \begin{vmatrix} p_v & 0 \\ 0 & 1 \end{vmatrix} < 0$$

Si on note $f(U) = (-u, p(v, C_0))$, le flux d'entropie vérifie: $\nabla\eta \cdot \nabla f = \nabla q$. Ceci équivaut à:

$$\begin{cases} q_v = \eta_u \cdot p_v \\ q_u = -\eta_v \end{cases}$$

Par identification, il vient:

$$\begin{cases} p_v \cdot \bar{q}_v = \bar{\eta}_u \cdot p_v \\ \bar{q}_u = -\bar{\eta}_v \cdot p_v \end{cases} \quad \text{c'est-à-dire} \quad \begin{cases} \bar{q}_v = \bar{\eta}_u \\ \bar{q}_u = -\bar{\eta}_v \cdot p_v \end{cases}$$

On définit alors:

$$\begin{cases} \tilde{q}(v, u, c) = \bar{q}(p(v, c), u) \stackrel{\text{def}}{=} \bar{q} \circ W(v, u, c) \\ \tilde{\eta}(v, u, c) = \bar{\eta}(p(v, c), u) \stackrel{\text{def}}{=} \bar{\eta} \circ W(v, u, c) \end{cases}$$

où on a posé:

$$W : \mathbb{R}_*^+ \times \mathbb{R} \times [0, 1] \rightarrow \mathbb{R}_*^+ \times \mathbb{R} \\ (v, u, c) \mapsto (p(v, c), u)$$

Alors on a:

$$\nabla \tilde{q} = \begin{pmatrix} \bar{q}_v \cdot p_v \\ \bar{q}_u \\ \bar{q}_v \cdot p_c \end{pmatrix} \quad \text{et} \quad \nabla \tilde{\eta} = \begin{pmatrix} \bar{\eta}_u \cdot p_v \\ \bar{\eta}_u \\ \bar{\eta}_v \cdot p_c \end{pmatrix}$$

D'après les relations précédentes sur $\nabla\eta$ et ∇q , il vient:

$$\nabla \tilde{q} = \begin{pmatrix} \bar{\eta}_u \cdot p_v \\ -\bar{\eta}_v \cdot p_v \\ \bar{\eta}_u \cdot p_c \end{pmatrix} \quad \text{ce qui donne} \quad \nabla \tilde{q} = \begin{pmatrix} \tilde{\eta}_u \cdot p_v \\ -\tilde{\eta}_v \\ \tilde{\eta}_u \cdot p_c \end{pmatrix}$$

Autrement dit, on a $\nabla \tilde{q} = \nabla \tilde{\eta} \cdot A$ où A désigne la jacobienne des flux du système (4.3):

$$A = \begin{pmatrix} 0 & -1 & 0 \\ p_v & 0 & p_c \\ 0 & 0 & 0 \end{pmatrix}$$

En conclusion, on a construit un couple entropique $(\tilde{\eta}, \tilde{q})$ pour (4.3) à partir de (η, q) . \square

Cette formulation permet de retrouver immédiatement l'entropie écrite dans [81] à partir de celle donnée dans [79] (p.399):

$$\eta(v, u) = -\frac{u^2}{2} + \int^v p(s)ds \quad \text{donne} \quad \tilde{\eta}(v, u, c) = -\frac{u^2}{2} + \int^v p(s, c)ds$$

A la lumière de ce résultat, il semble donc possible de suivre la démarche de [25] pour l'étude des solutions de (4.3).

4.2.2 Estimation en amplitude

Avant toute autre chose, il est nécessaire de mettre en évidence une borne uniforme sur la suite d'approximations visqueuses de (4.3), i.e. sur les solutions U^ϵ du système:

$$\begin{cases} \partial_t v^\epsilon - \partial_x u^\epsilon = \epsilon \partial_{xx} v^\epsilon \\ \partial_t u^\epsilon + \partial_x p(v^\epsilon, c^\epsilon) = \epsilon \partial_{xx} u^\epsilon \\ \partial_t c^\epsilon = g(v^\epsilon, c^\epsilon) + \epsilon \partial_{xx} c^\epsilon \end{cases} \quad (4.9)$$

Remarquons déjà que l'expression de c^ϵ est donnée par le noyau de la chaleur [10, 15]. Le moyen généralement utilisé pour dériver une estimation L^∞ pour (4.9) est d'utiliser la notion de région positivement invariante introduite par K. Chueh, C. Conley et J. Smoller [79]. Bien sûr ceci ne nous donnera au mieux qu'une convergence dans la topologie **-faible* $\sigma(L^1, L^\infty)$ insuffisante pour passer à la limite dans les termes non-linéaires.

Définition I.6 *On appelle région positivement invariante un fermé Σ de l'espace des phases tel que si une solution U^ϵ a sa donnée initiale et sa condition aux limites à valeurs dans Σ , alors elle est totalement contenue dans Σ .*

Généralement, Σ est recherchée comme intersection de "demi-espaces".

$$\Sigma = \cap_{i=1}^3 \{U \in \mathbb{R}_*^+ \times \mathbb{R} \times [0, 1] \text{ tels que } G_i(U) \leq 0\}$$

avec les $G_i \in C^2$ telles que $\nabla G_i \neq \vec{0}$ dans la zone considérée.

Définition I.7 On dit que la fonction réelle $U \mapsto G(U)$ est quasi-convexe en U si:

$$\forall \xi \in \mathbb{R}^3, \langle \xi, \nabla G(U) \rangle_{\mathbb{R}^3} = 0 \Rightarrow \langle \xi, D^2 G(U) \cdot \xi \rangle_{\mathbb{R}^3} \geq 0$$

On énonce maintenant le critère caractérisant Σ .

Théorème I.10 (Chueh, Conley, Smoller)

Soit $D(v)$ une matrice définie positive, on considère le problème:

$$\begin{cases} \partial_t v = M(v) \partial_x v + \epsilon D(v) \partial_{xx} v + f(v) \\ v(0) = v_0 \in L^2(\mathbb{R}) \text{ et } \lim_{|x| \rightarrow \infty} v(x) = 0 \end{cases}$$

et la région $\Sigma = \cap_{i=1}^3 \{U \text{ tels que } G_i(U) \leq 0\}$.

Supposons que $\forall v_0 \in \partial \Sigma$ (i.e. $\exists i / G_i(v_0) = 0$), les conditions suivantes soient vérifiées:

- $\nabla G_i(v_0)$ soit vecteur propre à gauche pour $M(v_0)$ et $D(v_0)$.
- G_i est quasi-convexe en v_0 .
- $\langle \nabla G_i(v_0), f(v_0) \rangle_{\mathbb{R}^3} \leq 0$

Alors, si v dépend continûment de f , Σ est invariante pour tout $\epsilon > 0$.

D'après cet énoncé, il est vital que la jacobienne des flux de (4.3) admette des vecteurs propres à gauche (4.5) pouvant s'exprimer comme gradients de fonctions quasi-convexes. On va donc chercher pour quelles lois de pression cette propriété est vérifiée. Il faut aussi remarquer que ce théorème n'est pas l'unique moyen d'obtenir un principe du maximum pour les approximations visqueuses; on peut consulter [46] à ce sujet. Malheureusement, les résultats récents qui y sont exposés ne sont d'aucune utilité ici. Pour finir, la condition de dépendance continue de la solution en fonction du terme source a été étudiée dans [76] (tome2 - page 12).

Lemme I.11 Il est suffisant que la pression $p(v, c)$ soit de la forme:

$$p(v, c) = f(v + K(c)) \text{ où } \begin{cases} f' < 0, f'' > 0 \\ K \in C^2([0, 1], \mathbb{R}^+), K' > 0, K'' \leq 0 \end{cases} \quad (4.10)$$

pour qu'une zone positivement invariante existe pour (4.9).

Démonstration

D'après l'expression des vecteurs propres à gauche donnée par (4.5), on constate que \vec{L}_0 ne pose aucun problème. Par contre, les difficultés vont venir de $\vec{L}_{\pm 1}$, notamment du terme en p_c (sauf si celui-ci est nul, puisqu'alors les vecteurs propres à gauche sont des gradients d'invariants de Riemann indépendants de la concentration c). On décide de travailler avec \vec{L}_{-1} que l'on réécrit sous cette forme:

$$\vec{L}_{-1}(U) = \begin{pmatrix} -y \cdot p_v \\ y \cdot \sqrt{-p_v} \\ -y \cdot p_c \end{pmatrix} \text{ où } y(v, u, c) \in C^1 \text{ et } y \neq 0$$

La question est de savoir pour quel type de pression a-t-on: $\text{rot}(\vec{L}_{-1}) = \vec{0}$. Pour cela, on calcule le rotationnel de \vec{L}_{-1} :

$$\begin{vmatrix} \partial_v & -y \cdot p_v \\ \partial_u & y \cdot \sqrt{-p_v} \\ \partial_c & -y \cdot p_c \end{vmatrix} = \begin{pmatrix} -p_v \cdot y_u - (y \cdot \sqrt{-p_v})_v \\ (y \cdot \sqrt{-p_v})_c + p_c \cdot y_u \\ -p_{vc} \cdot y - p_v \cdot y_c + p_{cv} \cdot y + p_c \cdot y_v \end{pmatrix} = \vec{0}$$

La dernière ligne donne: $p_c \cdot y_v = p_v \cdot y_c$, ce qui est vérifié entre autres par $y = a(u) \cdot p(v, c)$. On reporte cette expression dans les deux premières équations:

$$a'(u) \cdot p \cdot \nabla p + a(u) \nabla(p \cdot \sqrt{-p_v}) = \vec{0}$$

Pour éliminer la dépendance en $u \mapsto a(u)$, on suppose $p_c \neq 0$, on passe les termes en $a(u)$ au membre de droite et on divise les deux expressions:

$$\frac{p_v}{p_c} = -\frac{p_v \cdot \sqrt{-p_v} - (p_{vv}/2\sqrt{-p_v})}{p_c \cdot \sqrt{-p_v} - (p_{vc}/2\sqrt{-p_v})}$$

En effectuant le produit en croix, il vient:

$$\frac{p_v}{p_c} = \frac{p_{vv}}{p_{vc}} \text{ c'est-à-dire } \frac{\partial}{\partial v} \left(\frac{p_v}{p_c} \right) = 0$$

Ce qui nous donne: $p_c = k(c)p_v$. Introduisons maintenant une fonction $K : [0, 1] \rightarrow \mathbb{R}^+$ strictement croissante telle que: $K'(c) = k(c)$ et $K(0) = 0$. Alors on obtient: $\tilde{p}_c - \tilde{p}_v = 0$ où $\tilde{p}(v, c) = p(v, K(c))$. On a donc une équation d'advection si on imagine que c joue le rôle du temps et v celui de l'espace. La solution est donc donnée par: $\tilde{p}(v, c) = \tilde{p}_0(v + c)$. où la "donnée initiale" \tilde{p}_0 est l'expression de la pression en régime monophasique $c = 0$. Par conséquent, il est clair que l'expression:

$$p(v, c) = f(v + K(c))$$

remplit les critères demandés: $p_c(v, c) = k(c)f'(v + K(c)) = k(c)p_v(v, c)$.

Il reste à vérifier la quasi-convexité des fonctions $U \mapsto G_i(U)$ induites par cette loi de pression. Cela va bien sûr nous donner une condition sur la dérivée seconde de $c \mapsto K(c)$. Introduisons:

$$G_{\pm 1}(v, u, c) = \pm u - \int^{v+K(c)} \lambda(s) ds; \nabla G_{\pm 1} = \begin{pmatrix} -\lambda \\ \pm 1 \\ -k(c)\lambda \end{pmatrix} \text{ avec } \lambda = \sqrt{-f'}$$

Alors on a:

$$\lambda \cdot \nabla G_{-1} = -\vec{L}_{-1} = \begin{pmatrix} f' \\ -\sqrt{-f'} \\ k(c)f' \end{pmatrix} \text{ et } \lambda \cdot \nabla G_1 = \vec{L}_1 = \begin{pmatrix} f' \\ \sqrt{-f'} \\ k(c)f' \end{pmatrix}$$

Les hypothèses sur les dérivées de f nous assurent que $-\lambda_v = \frac{f''}{2\sqrt{-f'}} > 0$. Donc la matrice:

$$D^2 G_{\pm 1}(U) = \begin{pmatrix} -\lambda_v & 0 & -k(c)\lambda_v \\ 0 & 0 & 0 \\ -k(c)\lambda_v & 0 & -K''\lambda - k(c)^2\lambda_v \end{pmatrix}$$

sera positive: $\forall \vec{X} = (x, y, z)$, on a $\langle \vec{X}, D^2 G_{\pm 1}(U) \cdot \vec{X} \rangle_{\mathbb{R}^3} \geq 0$. On trouve en fait que:

$$\langle \vec{X}, D^2 G_{\pm 1}(U) \cdot \vec{X} \rangle_{\mathbb{R}^3} = (-\lambda_v)(x + k(c)z)^2 + (-K''(c)\lambda)z^2$$

On conclut que les conditions de quasi-convexité seront vérifiées si $K' \geq 0$ et $K'' \leq 0$. \square

Lemme I.12 Si l'on suppose en plus des hypothèses (4.10):

$$g(c) \geq 0, g(1) = 0 \tag{4.11}$$

Alors il existe une zone invariante pour le système visqueux (4.9):

$$\Sigma = \left\{ U = (v, u, c) \text{ tels que } c \in [0, 1]; \left[\begin{array}{l} u \leq C_1 + \int^{v+K(c)} \sqrt{-f'(s)} ds \\ u \geq C_2 - \int^{v+K(c)} \sqrt{-f'(s)} ds \end{array} \right] \right\} \tag{4.12}$$

Démonstration

Tout d'abord, remarquons que la matrice de diffusion de (4.9) est l'identité, ce qui signifie que tout vecteur de \mathbb{R}^3 est propre. Soit $U \in \partial\Sigma$, 3 cas sont possibles:

- $i = 0$, i.e. $\vec{L}_0 = \nabla G_0(U)$. De la forme de \vec{L}_0 , il vient que $G_0(U) = c(c-1)$ convient. Par conséquent, $c = 0$ ou $c = 1$ et le produit scalaire $\langle \nabla G_0(U), S(U) \rangle$ est négatif en $c = 1$ et en $c = 0$ si $g(0) \geq 0$ et $g(1) \leq 0$.

- $i = \pm 1$: il est nécessaire de réécrire l'expression de $\vec{L}_{\pm 1}$ en tenant compte de (4.10):

$$\vec{L}_{\pm 1} = \begin{pmatrix} f'(v + K(c)) \\ \pm \sqrt{-f'(v + K(c))} \\ k(c)f'(v + K(c)) \end{pmatrix}$$

D'où $G_{\pm 1}(U) = \pm u - \int^{v+K(c)} \sqrt{-f'(s)} ds$ et (4.11) assure la condition de quasi-convexité. Lorsqu'on se place sur la frontière, on a $c \in [0, 1]$ et $u = \pm \int^{v+K(c)} \sqrt{-f'(s)} ds$. Il faut alors vérifier que:

$$\langle \nabla G_{\pm 1}(U), S(U) \rangle = g(c) \underbrace{k(c)f'(v+K(c))}_{\leq 0} \leq 0$$

Pour cela, il est suffisant que $g(c) \geq 0$ (sauf peut-être en $c = 1$ d'après précédemment, donc $g(1) = 0$) puisque $k(c)$ est supposée positive. \square

Ce résultat suscite plusieurs commentaires. En effet, on constate que même si on admet la restriction sur la pression (4.10), on ne pourra avoir une zone invariante que pour des termes de signe constant au second membre. Les conditions sévères sur les termes sources viennent du fait que les rectangles du plan $(v+K(c), u)$ ne sont pas invariants par (4.9). Par conséquent, la troisième hypothèse du théorème précédent mêle les composantes du vecteur $S(U)$ puisque les vecteurs de la base canonique ne sont pas propres (voir les exemples dans [79]). Dans [14], les auteurs s'intéressent à un terme de relaxation sur l'équation de la vitesse (du type $(f(v) - u)$). De la même façon, ils obtiennent une condition $((p'(v) - f'(v)^2) \geq 0)$ pour préserver la zone invariante habituelle du p -système.

4.2.3 Compacité et existence de solutions

Ce paragraphe contient simplement un corollaire du résultat de compacité de R. DiPerna [25] sur la suite d'approximations visqueuses du p -système homogène.

Théorème I.11 *Soit $(v_0, u_0, c_0) \in L^\infty \cap L^2(\mathbb{R})$, avec $v_0 \geq \alpha > 0$ et $c_0 \in [0, 1]$. Alors sous les hypothèses (4.10), (4.11), il existe un triplet de solutions $L^\infty([0, T] \times \mathbb{R})$ du système (4.3). Celles-ci sont obtenues comme limites $\epsilon \rightarrow 0$ des solutions visqueuses de (4.14).*

Démonstration

Remarquons que le changement de variables $U = (v, u, c) \mapsto \tilde{U} = (w = v + K(c), u, c)$ est un difféomorphisme et qu'il permet d'affaiblir le couplage entre v et c dans (4.3). On obtient donc:

$$\begin{cases} \partial_t w - \partial_x u = K(c)_t = k(c)g(c) \\ \partial_t u + \partial_x p(w) = 0 \\ \partial_t c = g(c) \end{cases} \quad (4.13)$$

On régularise la partie " p -système" que l'on sépare de l'équation différentielle:

$$\begin{cases} \partial_t w^\epsilon - \partial_x u^\epsilon = k(c)g(c) + \epsilon \partial_{xx} w^\epsilon \\ \partial_t u^\epsilon + \partial_x p(w^\epsilon) = \epsilon \partial_{xx} u^\epsilon \end{cases} \quad \text{et } \partial_t c = g(c) \quad (4.14)$$

L'expression de la concentration est donc:

$$c(x, t) = c_0(x) + \int_0^t g(c)(x, s). ds$$

D'autre part, les hypothèses (4.10), (4.11) assurent l'existence d'une zone invariante pour le système découplé $(G_{\pm 1}(w, u) = \pm u - \int^w \sqrt{-p'(s)} ds)$. Par conséquent, si $g(c)$ converge suffisamment vite vers zéro lorsque $t \rightarrow +\infty$, alors w reste bornée et le théorème de compacité de DiPerna [25] assure que $(w^\epsilon, u^\epsilon) \rightarrow (w, u)$ fortement dans tous les $L^{p < \infty}$ et presque-partout à l'extraction d'une sous-suite près. \square

Remarques:

- On a pu s'apercevoir que la principale difficulté pour traiter le système (4.3) par cette méthode est d'obtenir une estimation uniforme en norme L^∞ . Pour cela, le théorème de la zone invariante semble être le seul moyen de procéder dans le cas général. Néanmoins, comme la jacobienne des flux n'est généralement pas diagonale, les composantes de la base canonique ne sont pas vecteurs propres à gauche. Ceci oblige ces derniers à dériver d'un potentiel ainsi qu'à vérifier une propriété de signe à travers une combinaison avec les termes sources.

- Dans [81], I. Toumi met en évidence une zone invariante (non-convexe) pour le problème de Riemann appliqué au système (4.1) avec la loi de pression diphasique eau-vapeur:

$$p(\rho, c) = \left(\frac{1.6\rho c}{1.6 - \rho(1 - c)} \right)^2$$

Signalons que celle-ci n'est pas de la forme prescrite par (4.10) *uniquement* à cause du terme en v/c puisqu'elle s'exprime sous la forme:

$$p(v, c) = \left(\frac{v}{c} + \frac{c-1}{1.6c} \right)^{-2}$$

- Ce travail est très proche de [7] où les auteurs considèrent les équations d'Euler homogènes dans le cas d'une loi de pression telle que *la vitesse du son lagrangienne soit une fonction de la pression*. C'est-à-dire:

$$\rho^2 \partial_\rho(p) = a(p)(\rho, S)$$

où S désigne l'entropie physique du système de la dynamique des gaz. Si l'on se place dans ce contexte, la restriction faite sur la pression s'exprime sous la forme:

$$\partial_\rho(p) = -\partial_v(p) \frac{1}{\rho^2} \text{ c'est-à-dire } -\partial_v(p) = a(p)$$

Suivant les conventions de signe pour $\partial_v(p)$, cette équation différentielle s'intègre sous la forme:

$$\partial_v \ln(A(p)) = 1 \text{ et donc } A(p) = \mathbf{e}^{v+k(S)}$$

Comme le rôle joué par S dans le système d'Euler est semblable à celui de c dans (4.1), on conclut aisément à l'équivalence des deux points de vue. Il faut préciser aussi que ces résultats sont des cas particuliers de la théorie des systèmes *riches* [76].

- Précisons pour terminer que le passage en variables lagrangiennes est bijectif pour des fonctions discontinues [86]. Par conséquent, cette approche est bien justifiée dans le contexte des solutions faibles L^∞ des systèmes hyperboliques.

Partie II

Méthodes de quadrature implicites adaptées aux lois scalaires

Résumé

Il s'agit dans cette partie de mettre en évidence des critères de stabilité pour les méthodes numériques implicites dans le cadre de la résolution de lois scalaires 1D. Cette étude se fait suivant deux directions:

- définir soigneusement la solution du système différentiel issu de la méthode des lignes
- prévenir l'apparition d'oscillations parasites au voisinages des discontinuités.

Le premier point est d'abord étudié dans le cas linéaire où la transformation de Fourier permet de conclure à la bonne tenue des schémas implicites pour des solutions à priori régulières. Dans le cas général, on montre en utilisant un formalisme introduit par Dahlquist [43] que la méthode des lignes n'est bien posée que dans certaines conditions. En conséquence, les schémas numériques subissent des restrictions de maillage en présence de situations difficiles. En accord avec l'intuition, celles-ci correspondent à des zones de compression ou situées à proximité des équilibres instables du terme source. Néanmoins, en testant certaines routines de quadrature sophistiquées de type Runge-Kutta sur des problèmes d'advection dissipatifs, on constate que ce critère n'est pas suffisant pour garantir une approximation de bonne qualité si des ondes discontinues sont présentes. En effet, de fortes oscillations apparaissent au voisinage de celles-ci bien en-dessous du seuil d'inversibilité de l'opérateur discret. Une analyse en variation totale permet alors d'explicitier des conditions suffisantes de stabilité non-linéaires prolongeant les précédentes restrictions pour la plupart des méthodes de quadrature numériques. Ceci permet de proposer des routines robustes et non-oscillantes qui sont testées dans des situations de difficulté progressive: loi linéaire homogène, équation de Burgers amortie, puis équation de Buckley-Leverett munie d'un terme de réaction à plusieurs équilibres.

Chapitre 5

Stabilité linéaire des méthodes implicites à un pas - Cas des solutions régulières

Pour introduire l'analyse numérique des lois hyperboliques scalaires avec second membre, on s'intéresse aux approximations possibles de ce genre de problèmes. Les méthodes retenues sont implicites par souci de stabilité linéaire, même en régime fortement réactionnel. D'autre part, cette approche peut être aussi motivée par le besoin de se libérer de la restriction CFL pour des calculs de régimes stationnaires. Ce chapitre est donc dédié à l'étude numérique du problème scalaire 1D donné par l'équation générique

$$\begin{cases} \partial_t u + \partial_x f(u) = g(u) \\ u(x, 0) = u_0(x) \in L^2 \cap L^\infty(\mathbb{R}) \end{cases} \quad (5.1)$$

On se place naturellement dans les hypothèses du théorème d'existence et d'unicité de la partie précédente. On se restreindra tout d'abord aux méthodes implicites à un pas couramment appelées " θ -méthodes" par analogie avec la terminologie des équations différentielles ordinaires. L'objectif en est de construire des algorithmes précis, stables, et respectant une grande partie des propriétés de l'équation considérée.

5.1 Généralités

On commence par introduire les notations qui seront utilisées tout au long de ce chapitre. Tout d'abord, on désigne par Δx et Δt les pas d'espace et de temps du maillage. Comme la fonction que l'on cherche à approcher est discontinue mais localement sommable, on convient de discrétiser la donnée initiale u_0 sous la forme:

$$u_j^0 = \frac{1}{\Delta x} \int_{(j-\frac{1}{2})\Delta x}^{(j+\frac{1}{2})\Delta x} u_0(x) dx$$

Dans [35] (par exemple), la plupart des schémas numériques explicites couramment utilisés sont étudiés. Ils permettent de définir pour tout couple $(j, n) \in \mathbb{Z} \times \mathbb{N}$ une approximation u_j^n de la solution continue $u(\cdot, n\Delta t)$ du problème (5.1). On désigne alors par \mathbf{U}^n le vecteur des approximations $(u_j^n)_{j \in \mathbb{Z}}$. Ces algorithmes, pour rester stables, ont l'inconvénient d'être limités par la célèbre condition de Courant-Friedrichs-Lewy:

$$\sup_{j,n} |f'(u_j^n)| \cdot \frac{\Delta t}{\Delta x} \leq 1$$

D'autre part, la présence d'un second membre dans l'équation (5.1) impose des restrictions supplémentaires sur le pas de temps si on le traite de façon explicite. En réponse à quoi, des approximations semi-implicites furent étudiées dans ce contexte (citons [50] ou [68], par ex.). L'approche que nous choisissons ici concerne donc les discrétisations *totale*ment implicites du problème (5.1), et tout particulièrement les θ -méthodes qui s'écrivent sous la forme: (avec $\theta \in [0, 1]$ et λ désignant le rapport $\frac{\Delta t}{\Delta x}$)

$$\begin{aligned} u_j^{n+1} = u_j^n - \theta \cdot [\lambda \cdot [F^I(u_j^{n+1}, u_{j+1}^{n+1}) - F^I(u_{j-1}^{n+1}, u_j^{n+1})] - \Delta t G^I(u_{j-1}^{n+1}, u_j^{n+1}, u_{j+1}^{n+1})] \\ - (1 - \theta) \cdot [\lambda \cdot [F^E(u_j^n, u_{j+1}^n) - F^E(u_{j-1}^n, u_j^n)] - \Delta t G^E(u_{j-1}^n, u_j^n, u_{j+1}^n)] \end{aligned}$$

Les fonctions F^I et F^E d'une part, G^I et G^E d'autre part sont respectivement les flux numériques (implicites et explicites) et les moyennes des termes sources (implicites et explicites). Comme on se limite ici à des schémas d'ordre 1 en espace (i.e. réellement à 3 points), on peut simplifier $F^I = F^E = F$ et $G^I = G^E = G$. On définit donc u_j^n pour $(j, n) \in \mathbb{Z} \times \mathbb{N}$ comme la solution éventuelle du schéma suivant:

$$\begin{aligned} u_j^{n+1} + \theta \cdot [\lambda \cdot [F(u_j^{n+1}, u_{j+1}^{n+1}) - F(u_{j-1}^{n+1}, u_j^{n+1})] - \Delta t G(u_{j-1}^{n+1}, u_j^{n+1}, u_{j+1}^{n+1})] \\ = u_j^n - (1 - \theta) \cdot [\lambda \cdot [F(u_j^n, u_{j+1}^n) - F(u_{j-1}^n, u_j^n)] - \Delta t G(u_{j-1}^n, u_j^n, u_{j+1}^n)] \end{aligned} \quad (5.2)$$

Afin d'étudier précisément le comportement du schéma (5.2), on est amené à faire des hypothèses sur les fonctions F et G . On suppose en particulier F et G différentiables et on impose la condition de *consistance* sur la fonction de flux:

$$\forall u \in \mathbb{R}, F(u, u) = f(u) \quad (5.3)$$

Quant à l'hypothèse de *monotonie* [35] qui se traduit par:

$$\forall (u, v) \in \mathbb{R}^2 \begin{cases} F_u(u, v) \geq 0 \text{ et } F_v(u, v) \leq 0 \\ \lambda \cdot (|F_u(u, v)| + |F_v(u, v)|) \leq 1 \end{cases}$$

elle contient, dans sa deuxième assertion, la condition CFL. Il apparaît donc nécessaire de la relaxer.

Définition II.1 *Un flux numérique consistant sera dit monotone s'il vérifie:*

$$\forall (u, v) \in \mathbb{R}^2, F_u(u, v) \geq 0 \text{ et } F_v(u, v) \leq 0 \quad (5.4)$$

La discrétisation du second membre peut être centrée, auquel cas:

$$G(u_{j-1}^n, u_j^n, u_{j+1}^n) \stackrel{\text{def}}{=} g(u_j^n)$$

Pour $\theta = 0$, \mathbf{U}^{n+1} se déduit directement de \mathbf{U}^n , l'algorithme est alors explicite. Dans le cas contraire, \mathbf{U}^{n+1} est *implicitement* défini par un système d'équations qu'il faut inverser numériquement. Ce surcoût de calcul doit alors être compensé par le gain en stabilité de la méthode. La démarche usuelle pour obtenir le vecteur \mathbf{U}^{n+1} est alors d'itérer un algorithme de type Newton [16] sur ce système, ou même tout simplement de le linéariser autour de \mathbf{U}^n (on parle alors de schéma implicite linéarisé). Pour ce faire, on désigne par $A_{\Delta x}$ l'opérateur défini par:

$$A_{\Delta x} : \ell^2 \cap \ell^\infty(\mathbb{Z}) \rightarrow \ell^2 \cap \ell^\infty(\mathbb{Z})$$

$$\mathbf{U}^n = (u_j^n)_j \mapsto A_{\Delta x}(\mathbf{U}^n) = \left[\frac{1}{\Delta x} \cdot [F(u_j^n, u_{j+1}^n) - F(u_{j-1}^n, u_j^n)] - g(u_j^n) \right]_j$$

Et la méthode se réécrit sous la forme

$$\mathbf{U}^{n+1} + \theta[\Delta t \cdot A_{\Delta x}(\mathbf{U}^{n+1})] = \mathbf{U}^n - (1 - \theta)[\Delta t \cdot A_{\Delta x}(\mathbf{U}^n)] \quad (5.5)$$

C'est-à-dire plus explicitement:

$$\mathbf{U}^{n+1} = \mathbf{U}^n - (Id + \theta[\Delta t \cdot dA_{\Delta x}(\mathbf{U}^n)])^{-1} \cdot [\Delta t \cdot A_{\Delta x}(\mathbf{U}^n)]$$

où on a supposé: $A_{\Delta x}(\mathbf{U}^{n+1}) \simeq A_{\Delta x}(\mathbf{U}^n) + dA_{\Delta x}(\mathbf{U}^n) \cdot (\mathbf{U}^{n+1} - \mathbf{U}^n)$ avec $dA_{\Delta x}(\mathbf{U}^n)$ la différentielle de $A_{\Delta x}$ en \mathbf{U}^n . Une condition nécessaire à l'existence d'un \mathbf{U}^{n+1} est donc l'inversibilité pour tout \mathbf{U}^n de la matrice $(Id + \theta[\Delta t \cdot dA_{\Delta x}(\mathbf{U}^n)])$, le système étant en général non-linéaire. Néanmoins, on peut d'ores-et-déjà noter que cette opération ne peut à priori se faire (théorème d'inversion *locale*) que dans un petit voisinage de \mathbf{U}^n , ce qui limite la variation possible $(\mathbf{U}^{n+1} - \mathbf{U}^n)$, et conséquemment le pas de temps Δt .

Remarque: La θ -discrétisation peut être introduite au moyen de l'équation différentielle ordinaire issue de la méthode des lignes [68]. Cette approche consiste en un formalisme visant à découpler nettement les discrétisations temporelles et spatiales. On introduit de nouveau l'opérateur $A_{\Delta x}$, et on s'attache à résoudre le mieux possible le système différentiel ordinaire:

$$\begin{aligned} \dot{U} + A_{\Delta x}(U) &= 0 \\ U_j(0) &= (u_j^0) \\ t \mapsto U(t) &= (u_j(t))_{j \in \mathbb{Z}} \end{aligned} \quad (5.6)$$

Si on décide d'intégrer numériquement ce problème avec une θ -méthode [43], on retrouve directement la discrétisation écrite plus haut. L'ordre de précision (loin des chocs) est au moins le minimum entre celui de l'intégration temporelle et celui de l'opérateur spatial discrétisé $A_{\Delta x}$.

5.2 Analyse de Fourier et stabilité dans le cas linéaire

L'intérêt principal d'une méthode d'approximation implicite réside dans ses bonnes propriétés de stabilité linéaire. Cela signifie que si la solution analytique d'un problème continu est bornée en amplitude, alors il en sera de même pour la solution approchée. C'est cette caractéristique que nous allons commencer par étudier maintenant, puisqu'elle se révèle importante pour démontrer par exemple la convergence des schémas aux différences finies.

5.2.1 Ecriture du θ -schéma

L'équation considérée ici est un cas particulier de (5.1) totalement linéaire, donc nécessairement dissipative:

$$\begin{cases} \partial_t u + a \cdot \partial_x u = -\alpha \cdot u \text{ avec } \alpha \geq 0 \\ u(x, 0) = u_0(x) \in L^2 \cap L^\infty(\mathbb{R}) \end{cases} \quad (5.7)$$

On utilise habituellement la transformation de Fourier dans $L^2(\mathbb{R})$ (théorème de Plancherel [75]) pour traiter ce genre de problème. En effet, on peut faire correspondre au vecteur \mathbf{U}^n une fonction en escalier $U^n \in L^2(\mathbb{R})$ telle que $\int_{(j-\frac{1}{2})\Delta x}^{(j+\frac{1}{2})\Delta x} U^n(x) \cdot dx = \Delta x \cdot u_j^n$. Si le schéma est linéaire, on peut écrire:

$$(Id + \theta[\Delta t \cdot A_{\Delta x}]) \cdot \mathbf{U}^{n+1} = (Id - (1 - \theta)[\Delta t \cdot A_{\Delta x}]) \cdot \mathbf{U}^n$$

Comme le schéma est à trois points, on introduit les coefficients $(c_k, d_k)_{k=-1,0,1}$ tels que l'on aît:

$$\sum_{k=-1}^1 c_k U^{n+1}(x + k \cdot \Delta x) = \sum_{k=-1}^1 d_k U^n(x + k \cdot \Delta x)$$

La transformation de Plancherel donne, pour $\xi = \omega \cdot \Delta x$:

$$\underbrace{\left(\sum_{k=-1}^1 e^{ik\xi} \cdot c_k \right)}_{l_1(\xi)} \hat{U}^{n+1}(\omega) = \underbrace{\left(\sum_{k=-1}^1 e^{ik\xi} \cdot d_k \right)}_{l_2(\xi)} \hat{U}^n(\omega) \quad (5.8)$$

Comme $U \mapsto \hat{U}$ est un isomorphisme de $L^2(\mathbb{R})$, il est équivalent de s'assurer de la définition des suites $(\hat{U}^n)_n$ ou $(U^n)_n$. Il faut maintenant préciser la forme de la fonction de flux numérique. Comme le problème (5.7) est linéaire et que l'approximation est d'ordre 1, celle-ci est donc une forme bilinéaire $F(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$. Le gradient de flux approché sera donc:

$$F(u_j, u_{j+1}) - F(u_{j-1}, u_j) = \sum_{k=-1}^1 a_k \cdot u_{j+k}$$

Ceci implique:

$$u_j^{n+1} + \lambda \cdot \theta \left[\sum_{k=-1}^1 a_k \cdot u_{j+k}^{n+1} + \alpha \Delta x \cdot u_j^{n+1} \right] = u_j^n - \lambda \cdot (1 - \theta) \left[\sum_{k=-1}^1 a_k \cdot u_{j+k}^n + \alpha \Delta x \cdot u_j^n \right]$$

On obtient donc les valeurs de $(c_k, d_k)_{k=-1,0,1}$:

$$\begin{cases} c_{\pm 1} = \lambda \theta a_{\pm 1} \\ c_0 = 1 + \lambda \theta a_0 + \alpha \Delta t \end{cases} \quad \text{et} \quad \begin{cases} d_{\pm 1} = -\lambda(1 - \theta)a_{\pm 1} \\ d_0 = 1 - \lambda(1 - \theta)a_0 - \alpha \Delta t \end{cases}$$

Les valeurs des $(a_k)_{k=-1,0,1}$ peuvent être précisées grâce aux hypothèses faites sur le flux F . Si $\alpha = 0$, il est clair que les données initiales constantes ($u_j^0 \equiv cste$) doivent rester invariantes pour toute valeur de θ . Ceci implique:

$$\sum_{k=-1}^1 a_k = 0$$

Le flux correspondant $F(x, y) = -a_{-1} \cdot x + a_1 \cdot y$ sera consistant si $F(x, x) = a \cdot x$, ce qui se traduit par:

$$a_1 - a_{-1} = a$$

Et il sera monotone si:

$$a_{-1} \leq 0 \text{ et } a_1 \leq 0$$

Ceci implique donc que:

$$a_0 \geq 0$$

5.2.2 Stabilité linéaire au sens de Von Neumann

Sous ces hypothèses, on peut avancer un résultat quant à la stabilité ℓ^2 de la méthode.

Proposition II.1 *Pour un flux numérique consistant (5.3) et monotone (5.4), la suite d'approximations $(\mathbf{U}^n)_n$ issue d'une θ -discrétisation pour le problème dissipatif:*

$$\begin{cases} \partial_t u + a \cdot \partial_x u = -\alpha \cdot u \text{ avec } \alpha \geq 0 \\ u(x, 0) = u_0(x) \in L^2 \cap L^\infty(\mathbb{R}) \end{cases}$$

est uniquement définie et uniformément bornée dans $\ell^2(\mathbb{Z})$ pour $\theta \geq \frac{1}{2}$.

Démonstration

Pour cette équation, la méthode s'écrit:

$$\mathbf{U}^{n+1} = (Id + \theta[\Delta t \cdot A_{\Delta x}])^{-1} \cdot (Id - (1 - \theta)[\Delta t \cdot A_{\Delta x}]) \cdot \mathbf{U}^n$$

Puisque $A_{\Delta x}$ est un opérateur linéaire, $dA_{\Delta x} = A_{\Delta x}$ et $(Id + \theta \cdot \Delta t \cdot A_{\Delta x})$ sera inversible si et seulement si 0 n'appartient pas à son spectre. D'après [60], on peut s'en assurer en montrant que le module de $l_1(\xi)$ dans (5.8) est non nul. On pose: $\lambda' = \lambda \cdot \theta$, $\alpha' = \alpha \cdot \theta$

$$\begin{aligned} l_1(\xi) &= (1 + \alpha' \cdot \Delta t) + \lambda' \sum_{k=-1}^1 a_k \cdot \mathbf{e}^{ik\xi} \\ &= (1 + \alpha' \cdot \Delta t) + \lambda' a_0 \cdot (1 - \cos(\xi)) + i \cdot \lambda' a \cdot \sin(\xi) \end{aligned}$$

Et on obtient:

$$|l_1(\xi)|^2 = [1 + \underbrace{(\alpha' \cdot \Delta t + \lambda' a_0 \cdot (1 - \cos(\xi)))}_{\theta \cdot A \geq 0}]^2 + \underbrace{[\lambda' a \cdot \sin(\xi)]^2}_{\theta \cdot B} \geq 1 > 0$$

Ceci assure l'existence et l'unicité de \mathbf{U}^n pour tout $t = n \cdot \Delta t$ puisque la partie implicite du schéma est alors inversible. Pour obtenir une propriété de dissipation, il faut examiner le rapport $|l_2(\xi)/l_1(\xi)|$ qui correspond au facteur d'amplification de la méthode numérique. On note: $\lambda'' = \lambda \cdot (1 - \theta)$, $\alpha'' = \alpha \cdot (1 - \theta)$

$$\begin{aligned} l_2(\xi) &= 1 - \alpha'' \cdot \Delta t - \lambda'' \sum_{k=-1}^1 a_k \cdot \mathbf{e}^{ik\xi} \\ &= 1 - \alpha'' \cdot \Delta t - \lambda'' a_0 \cdot (1 - \cos(\xi)) - i \cdot \lambda'' a \cdot \sin(\xi) \end{aligned}$$

Et on obtient:

$$|l_2(\xi)|^2 = [1 - (\alpha'' \cdot \Delta t + \lambda'' a_0 \cdot (1 - \cos(\xi)))]^2 + [\lambda'' a \cdot \sin(\xi)]^2$$

Et donc:

$$\left| \frac{l_2(\xi)}{l_1(\xi)} \right|^2 = \frac{[1 - (\alpha'' \cdot \Delta t + \lambda'' a_0 \cdot (1 - \cos(\xi)))]^2 + [\lambda'' a \cdot \sin(\xi)]^2}{[1 + (\alpha' \cdot \Delta t + \lambda' a_0 \cdot (1 - \cos(\xi)))]^2 + [\lambda' a \cdot \sin(\xi)]^2}$$

Ce qui donne:

$$\left| \frac{l_2(\xi)}{l_1(\xi)} \right|^2 = \frac{[1 - (1 - \theta) \cdot A]^2 + [(1 - \theta) \cdot B]^2}{[1 + \theta \cdot A]^2 + [\theta \cdot B]^2} \text{ avec } A \geq 0$$

D'où on déduit:

$$\theta \geq \frac{1}{2} \text{ implique } \left| \frac{l_2(\xi)}{l_1(\xi)} \right| \leq 1$$

Et on conclut que:

$$\|\hat{\mathbf{U}}^{n+1}\|_{L^2} \leq \|\hat{\mathbf{U}}^n\|_{L^2}$$

□

La restriction $|l_2(\xi)/l_1(\xi)| \leq 1$ est connue sous le nom de condition de Von Neumann. Elle assure la stabilité du schéma pour $t \rightarrow +\infty$. Dans le cas d'une approximation explicite ($\theta = 0$), il faudrait imposer $|\lambda \cdot a| \leq 1$ et $|1 - \alpha \cdot \Delta t| \leq 1$ pour maintenir le résultat (voir [43, 60]).

5.3 Stabilité dans le cas non-linéaire

L'idée qui motive ce paragraphe est l'étude du système différentiel issu de la semi-discrétisation en espace (5.6) d'une équation de convection-réaction non-linéaire. Il s'agit donc de comprendre le comportement de ce type de schéma (5.5) sur un problème plus général, c'est-à-dire:

$$\begin{cases} \partial_t u + \partial_x f(u) = g(u) \\ f(0) = g(0) = 0 \text{ et } x.g(x) \leq 0 \text{ pour } |x| \geq M \\ u(x, 0) = u_0(x) \in L^2 \cap L^\infty(\mathbb{R}) \text{ à support compact} \end{cases} \quad (5.9)$$

La restriction sur le support de u_0 s'explique par le fait que l'on ne désire pas s'attaquer à un système différentiel infini. Il est donc nécessaire de considérer des approximations de $u(x, t)$ par des vecteurs de \mathbb{R}^N où N est le nombre de points de discrétisation en espace. La largeur de cet intervalle est directement liée au cône de lumière pour (5.9): elle dépend donc du temps. On travaille donc dans un compact de $\mathbb{R} \times \mathbb{R}^+$. On n'approfondira pas les conditions aux limites, la plus naturelle pour ce type de problèmes étant bien sûr celle de Neumann.

5.3.1 Un cadre théorique adapté aux problèmes raides

L'objet de ces rappels est de mettre en place un cadre de travail suffisamment peu restrictif pour pouvoir traiter efficacement le système différentiel issu de la méthode des lignes. Les théorèmes que l'on énonce ici proviennent du livre [19].

Théorème II.1 *Soit le problème différentiel:*

$$\begin{cases} \dot{U}(t) = \mathcal{L}(t, U(t)) \\ U(0) = U_0 \in E \subset \mathbb{R}^N \\ t \in I_0 = [0, T] \text{ où } T \in \mathbb{R}^+ \end{cases} \quad (5.10)$$

Si E est stable pour (5.10), si \mathcal{L} est continue, et s'il existe une fonction $l(t) \in L^1(I_0)$ telle que:

$$\forall t \in I_0, \forall (U, V) \in E^2, \langle \mathcal{L}(t, U) - \mathcal{L}(t, V), U - V \rangle_{\mathbb{R}^N} \leq l(t) |U - V|^2 \quad (5.11)$$

Alors le problème (5.10) admet une unique solution dans $E \times I_0$.

L'assertion « E stable pour (5.10)» signifie bien sûr:

$$U_0 \in E \Rightarrow \forall t \in I_0, U(t) \in E$$

Ce résultat nous assure de l'existence d'une unique solution au système semi-discrét et il est nécessaire de lui adjoindre un pendant numérique pour un schéma du type (5.5): à quelle condition ce dernier est-il inversible dans E ?

Théorème II.2 *Supposons qu'il existe une constante \bar{l} telle que:*

$$\forall t \in I_0, \forall (U, V) \in E^2, \langle \mathcal{L}(t, U) - \mathcal{L}(t, V), U - V \rangle_{\mathbb{R}^N} \leq \bar{l} |U - V|^2 \quad (5.12)$$

Alors, sous la condition $\theta \Delta t \bar{l} < 1$, le θ -schéma pour l'équation (5.10) admet une solution et une seule dans $E \times I_0$.

Remarque Lorsque $l(t) \leq 0$, la propriété (5.11) est parfois appelée *condition de contractivité*. Elle a été étudiée par (entre autres) G. Dahlquist qui cherchait à généraliser la théorie des routines A -stables à des problèmes non-linéaires. Une méthode numérique respectant ce type de propriété est dite B -stable [43].

Il est maintenant nécessaire d'indiquer la façon la plus naturelle d'estimer la constante \bar{l} . Cela peut se faire grâce à l'introduction de la *norme logarithmique* d'une matrice. Cette idée est due à G. Dahlquist mais les théorèmes cités se trouvent dans [42] (pp61/63).

Définition II.2 *Soit $M \in \mathbf{M}_N(\mathbb{R})$, la norme logarithmique de cette matrice est donnée par:*

$$\mu(M) = \lim_{h \rightarrow 0} \frac{\|Id + hM\| - 1}{h} \quad (5.13)$$

où $\|\cdot\|$ est une norme matricielle subordonnée.

Il existe une formulation plus claire de cette quantité lorsque l'on considère les normes matricielles classiques [16].

Théorème II.3 *Pour la norme Euclidienne $\|\cdot\|_2$, $\mu(M)_2$ est le plus petit réel tel que l'on ait:*

$$\forall U \in \mathbb{R}^N, \langle MU, U \rangle_{\mathbb{R}^N} \leq \mu(M)_2 \langle U, U \rangle_{\mathbb{R}^N}$$

D'autre part, on a:

$$\mu(M)_2 = \max_{1 \leq j \leq N} \left\{ \lambda_j \text{ valeur propre de } \frac{1}{2}(M + M^T) \right\}$$

On termine par trois lemmes techniques qui nous seront utiles:

Lemme II.1 *Soient deux solutions $U(t)$ et $V(t)$ issues du système (5.10) avec pour données initiales respectives U_0, V_0 , alors sous la condition (5.11), on a:*

$$|U(t) - V(t)| \leq e^{\sup_{s \in I_0} l(s) \cdot t} |U_0 - V_0| \quad (5.14)$$

Démonstration

voir [43] page 191/192

Cette propriété est respectée par les approximations issues des méthodes de quadrature B -stables. C'est par exemple le cas pour la méthode d'Euler rétrograde correspondant au choix $\theta = 1$.

Lemme II.2 (*Gerschgorin-Hadamard*)

Si λ est valeur propre d'une matrice $M \in \mathbf{M}_N(\mathbb{R})$, alors on a:

$$\lambda \in \bigcup_{i=1}^N D_i \text{ où } D_i = \left\{ z \in \mathbb{C} / |z - M_{ii}| \leq \sum_{j \neq i} |M_{ij}| \right\} \quad (5.15)$$

Démonstration

voir [56] pp55/57

Lemme II.3 *Soit $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ différentiable, alors:*

$$\forall (\vec{a}, \vec{b}) \in (\mathbb{R}^2)^2, \exists \alpha \in [0, 1] \text{ tel que } F(\vec{b}) - F(\vec{a}) = \nabla F(\vec{c}) \cdot (\vec{b} - \vec{a}) \text{ avec } \vec{c} = (1 - \alpha)\vec{a} + \alpha\vec{b}$$

Démonstration

On sait que pour une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ continue sur $[a, b]$, dérivable sur $]a, b[$, on a:

$$\exists \alpha \in [0, 1] \text{ tel que } f(b) - f(a) = f'(c)(b - a) \text{ avec } c = (1 - \alpha)a + \alpha b$$

Pour $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ on considère le segment $(1 - \alpha)\vec{a} + \alpha\vec{b}$, alors on peut définir:

$$g : [0, 1] \rightarrow \mathbb{R} \\ x \mapsto F((1 - x)\vec{a} + x\vec{b})$$

La fonction g a autant de régularité que F et donc:

$$\exists \alpha \in [0, 1] \text{ tel que } g(0) - g(1) = g'(\alpha) = \frac{\partial F}{\partial x} \left(x(\vec{b} - \vec{a}) + \vec{a} \right)_{x=\alpha}$$

Comme $g(0) = F(\vec{a})$ et $g(1) = F(\vec{b})$, il vient:

$$F(\vec{b}) - F(\vec{a}) = \nabla F(\vec{c})(\vec{b} - \vec{a})$$

□

5.3.2 L'estimation fondamentale

On désire appliquer ces idées à l'approximation implicite du problème (5.9). Comme on a l'existence et l'unicité de la solution entropique L^∞ de (5.9) sur des bandes $\mathbb{R} \times [0, T]$, il est suffisant de considérer des solutions du système semi-discrète (5.10) sur des intervalles de temps bornés. D'autre part, l'opérateur approché intervenant dans la méthode des lignes induit un système différentiel *autonome*. Donc la fonction $t \mapsto l(t)$ peut être une constante. Les résultats qui vont suivre visent à dériver une constante \bar{l} suffisamment fine dans le cas d'un flux numérique consistant (5.3) et monotone (5.4). Ceci en vue d'estimer un domaine de définition du θ -schéma (5.5).

Proposition II.2 Soit \mathcal{L} l'application définie par:

$$\begin{aligned} \mathcal{L} : \mathbb{R}^N &\rightarrow \mathbb{R}^N \\ U &\mapsto \mathcal{L}(U) \stackrel{def}{=} -A_{\Delta x}(U) \end{aligned}$$

Alors sous les conditions de consistance (5.3) et de monotonie (5.4) pour la fonction de flux numérique, on a l'estimation:

$$\forall (U, V) \in (\mathbb{R}^N)^2, \exists l(U, V) \in \mathbb{R} \text{ telle que } \langle \mathcal{L}(U) - \mathcal{L}(V), U - V \rangle_{\mathbb{R}^N} \leq l(U, V) |U - V|^2$$

Ce résultat préliminaire ne nous donne qu'une majoration ponctuelle dépendant des états U et V . On l'étendra ensuite par compacité puisque les solutions entropiques de (5.9) sont bornées uniformément en amplitude.

Démonstration

Soit $(U, V) \in (\mathbb{R}^N)^2$, la j^{eme} composante du vecteur $\mathcal{L}(U) - \mathcal{L}(V)$ s'écrit:

$$[\mathcal{L}(U) - \mathcal{L}(V)]_j = \frac{-1}{\Delta x} [(F(u_j, u_{j+1}) - F(v_j, v_{j+1})) - (F(u_{j-1}, u_j) - F(v_{j-1}, v_j))] + (g(u_j) - g(v_j))$$

Comme $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ et $g : \mathbb{R} \rightarrow \mathbb{R}$ sont régulières, il existe des $(\alpha_i)_{i=1,2,3} \in [0, 1]$ et:

$$\begin{cases} \mathbb{R}^2 \ni \xi_{j+\frac{1}{2}} = (1 - \alpha_1)(u_j, v_j)^T + \alpha_1(u_{j+1}, v_{j+1})^T \\ \mathbb{R}^2 \ni \xi_{j-\frac{1}{2}} = (1 - \alpha_2)(u_{j-1}, v_{j-1})^T + \alpha_2(u_j, v_j)^T \\ \mathbb{R} \ni \zeta_j = (1 - \alpha_3)u_j + \alpha_3v_j \end{cases}$$

tels que l'on aît:

$$[\mathcal{L}(U) - \mathcal{L}(V)]_j = \frac{-1}{\Delta x} \left[\nabla F(\xi_{j+\frac{1}{2}}) \cdot \begin{pmatrix} u_j - v_j \\ u_{j+1} - v_{j+1} \end{pmatrix} - \nabla F(\xi_{j-\frac{1}{2}}) \cdot \begin{pmatrix} u_{j-1} - v_{j-1} \\ u_j - v_j \end{pmatrix} \right] + g'(\zeta_j)(u_j - v_j)$$

Au bout du compte, il existe une matrice tridiagonale $L(U, V)$ dans $\mathbf{M}_N(\mathbb{R})$ telle que l'on aît la linéarisation:

$$\mathcal{L}(U) - \mathcal{L}(V) = L(U, V)(U - V)$$

On décrit la forme générale de $L(U, V)$:

$$\begin{pmatrix} \ddots & & & & \\ & \frac{-1}{\Delta x} F_u(\xi_{j-\frac{1}{2}}) & & & \\ & g'(\zeta_j) - \frac{1}{\Delta x} F_u(\xi_{j+\frac{1}{2}}) + \frac{1}{\Delta x} F_u(\xi_{j-\frac{1}{2}}) & & & \\ & & \frac{1}{\Delta x} F_u(\xi_{j+\frac{1}{2}}) & & \\ & & & \ddots & \end{pmatrix} \begin{pmatrix} 0 \\ \frac{-1}{\Delta x} F_v(\xi_{j+\frac{1}{2}}) \\ 0 \\ \frac{-1}{\Delta x} F_v(\xi_{j-\frac{1}{2}}) \\ 0 \end{pmatrix}$$

Il faut noter que $L(U, V)$ n'est pas en général de la forme $d\mathcal{L}(W)$ pour $W \in \mathbb{R}^N$.

Posons $L = L(U, V)$ et $W = U - V$; pour évaluer le produit scalaire $\langle LW, W \rangle_{\mathbb{R}^N}$, on va introduire le formalisme des normes logarithmiques (5.13):

$$\langle LW, W \rangle_{\mathbb{R}^N} \leq \mu(L)_2 \langle W, W \rangle_{\mathbb{R}^N}$$

où $\mu(L)_2$ est la plus grande valeur propre de la symétrisée de L . Le résultat du lemme de Gerschgorin-Hadamard va justement nous servir à évaluer cette quantité (qui est réelle puisque $\frac{1}{2}(L + L^T)$ est symétrique).

Soit λ la plus grande valeur propre de $\frac{1}{2}(L + L^T)$, alors (5.15) implique que λ appartient à l'un des disques de Gerschgorin:

$$\exists j, 1 \leq j \leq N, \text{ tel que } \lambda - L_{jj} \leq |\lambda - L_{jj}| \leq \frac{1}{2} \left[\sum_{k \neq j} |L_{kj} + L_{jk}| \right]$$

D'après la construction de L et l'hypothèse de monotonie de F (5.4), on voit que les éléments hors de la diagonales sont tous positifs. Par conséquent, il vient:

$$\lambda = \mu(L)_2 \leq \max_{1 \leq j \leq N} \left[L_{jj} + \frac{1}{2} \sum_{k \neq j} (L_{kj} + L_{jk}) \right]$$

On conclut que pour tout couple (U, V) dans \mathbb{R}^N , il existe une constante $l(U, V)$ telle que:

$$\langle \mathcal{L}(U) - \mathcal{L}(V), U - V \rangle_{\mathbb{R}^N} \leq l(U, V) |U - V|^2$$

Celle-ci s'écrit:

$$l(U, V) = \max_{1 \leq j \leq N} \left(g'(\zeta_j) - \frac{1}{2\Delta x} [(F_u + F_v)(\xi_{j+\frac{1}{2}}) - (F_u + F_v)(\xi_{j-\frac{1}{2}})] \right)$$

□

5.3.3 Domaine invariant et estimation globale

Ce résultat fournit une condition suffisante pour définir une approximation discrète de la solution continue $u(x, t)$ du problème (5.9). Il reste à étendre cette majoration ponctuelle en une majoration uniforme. Pour cela, il est naturel de mettre en oeuvre des arguments de nature topologique tels que la compacité. En effet, on a un contrôle de la norme L^∞ de la solution entropique de l'équation (5.9). On a aussi une hypothèse sur le terme source. Par conséquent, si on pose m la borne supérieure des normes L^∞ des données initiales pour (5.9), il vient que le domaine convexe suivant¹ est stable par (5.10):

$$E = \{U \in \mathbb{R}^N \text{ tels que } \|U\|_\infty \leq \max(m, M)\} \quad (5.16)$$

Lemme II.4 *Si le flux numérique F est monotone (5.4) et si l'application $g \in C^1(\mathbb{R})$ vérifie:*

$$x.g(x) \leq 0 \text{ pour } |x| \geq M$$

alors le domaine convexe E (5.16) est invariant pour le système différentiel $\dot{U} = \mathcal{L}(U)$.

Démonstration

Il s'agit de montrer que la région $E = \Sigma^+ \cap \Sigma^-$ est invariante avec:

$$\begin{cases} \Sigma^+ = \cap_{i=1}^N \{U = (u_j)_{j=1, \dots, N} \in \mathbb{R}^N \text{ tels que } u_j - \max(m, M) \leq 0\} \\ \Sigma^- = \cap_{i=1}^N \{U = (u_j)_{j=1, \dots, N} \in \mathbb{R}^N \text{ tels que } u_j + \max(m, M) \geq 0\} \end{cases}$$

Choisissons une donnée initiale U^0 pour $\dot{U} = \mathcal{L}(U)$ dont une composante u_{j_0} touche une arête de Σ^+ c'est-à-dire:

$$U^0 = (u_j)_{1 \leq j \leq N} \text{ avec } \begin{cases} u_{j_0} = \max(m, M) \\ u_j \leq \max(m, M) \text{ pour } j \neq j_0 \end{cases}$$

Du fait de la régularité de la solution de $\dot{U} = \mathcal{L}(U)$, la région Σ^+ sera invariante si on a:

$$\mathcal{L}(U^0)_{j_0} \leq 0$$

En tenant compte de la définition de \mathcal{L} , il faut établir que:

$$g(\max(m, M)) - \frac{1}{\Delta x} [F(\max(m, M), u_{j_0+1}) - F(u_{j_0-1}, \max(m, M))] \leq 0 \quad (5.17)$$

1. Celui-ci n'est bien sûr pas unique: il dépend directement de la forme des fonctions f et g

En appliquant le lemme II.3, il existe $\xi = \theta.(M, u_{j_0+1})^T + (1 - \theta).(u_{j_0-1}, M)^T$ tel que:

$$[F(\max(m, M), u_{j_0+1}) - F(u_{j_0-1}, \max(m, M))] = \underbrace{[F_u(\xi)(\max(m, M) - u_{j_0-1})]}_{\geq 0} + \underbrace{[F_v(\xi)(u_{j_0+1} - \max(m, M))]}_{\geq 0}$$

Du fait de l'hypothèse sur g , on a $g(\max(m, M)) \leq 0$ et l'inégalité (5.17) est établie. On montre de la même façon la propriété pour Σ^- . Il est alors clair que la région $E = \Sigma^+ \cap \Sigma^-$ est invariante. La convexité est une conséquence de l'inégalité triangulaire. Soient $(U, V) \in E^2$, il vient:

$$\|\theta U + (1 - \theta)V\|_\infty \leq \theta\|U\|_\infty + (1 - \theta)\|V\|_\infty \leq \max(m, M)$$

□

Ceci est exactement l'estimation uniforme faite sur la suite d'approximations visqueuses dans la théorie de l'existence² pour le problème (5.9).

Proposition II.3 *Sous les hypothèses (5.3) et (5.4), il existe une constante réelle \bar{l} telle que:*

$$\forall (U, V) \in E^2, \langle \mathcal{L}(U) - \mathcal{L}(V), U - V \rangle_{\mathbb{R}^N} \leq \bar{l}|U - V|^2 \quad (5.18)$$

Démonstration

L'ensemble E (5.16) est compact en tant que fermé borné de \mathbb{R}^N . Par conséquent, toute application continue définie sur E y atteint ses bornes. Donc on peut écrire:

$$\bar{l} = \max_{(U, V) \in E^2} l(U, V)$$

où $l(U, V)$ est définie comme précédemment. □

5.3.4 Conséquences numériques

On va maintenant expliciter les conditions suffisant à affirmer l'existence et l'unicité de la solution à chaque pas de temps du θ -schéma (5.5).

Théorème II.4 *Sous la condition:*

$$1 - \theta\Delta t \max_{\substack{U = (u_j) \in E \\ 1 \leq j \leq N}} \left(g'(u_j) - \frac{1}{2\Delta x} [(F_u + F_v)(u_j, u_{j+1}) - (F_u + F_v)(u_{j-1}, u_j)] \right) > 0 \quad (5.19)$$

le schéma (5.5) admet une solution et une seule.

Démonstration

C'est une application directe du théorème II.2. □

On peut aller un peu plus loin en utilisant le lemme II.1. Supposons qu'à partir d'un certain rang \mathbf{U}^n , on ait $\nu = \theta\bar{l}\Delta t < 0$ (ce n'est pas déraisonnable puisque cela est vrai si la suite d'approximations ne comporte plus de chocs et est dans la zone d'attraction d'un équilibre stable du terme source), alors on voit que dans le cas $\theta = 1$ (Euler rétrograde):

$$|\mathbf{U}^{n+1} - \mathbf{U}^n| < e^{\nu\Delta t} |\mathbf{U}^n - \mathbf{U}^{n-1}|$$

Autrement dit, la variation en temps commence à décroître. Si la suite $(\mathbf{U}^n)_n$ reste dans une zone où la constante \bar{l} reste négative, alors elle est de Cauchy. Comme \mathbb{R}^N est complet, elle est convergente et en particulier uniformément bornée.

². Voir la proposition I.5 dans la première partie de ce manuscrit

Remarques • Pour une équation scalaire avec un flux régulier, la condition d'entropie pour un choc de vitesse σ impose que celui-ci soit **compressif**:

$$f'(u_R) < \sigma < f'(u_L)$$

Dans ce cas, si on admet que les termes de flux numériques présents dans l'expression (5.19) sont proches de $-\partial_x f'(u)/2$, alors cet effet de compression est une contrainte pour le pas de temps puisqu'il tend à augmenter la constante \bar{l} . En revanche, si on est en présence d'une onde expansive, alors ce terme convectif permet de vérifier plus facilement (5.19). Ceci explique un peu pourquoi les solutions-chocs peuvent être moins faciles à gérer en implicite. A noter qu'une variante de cette condition a été établie dans un cadre assez éloigné pour des systèmes linéaires multidimensionnels [17].

- Le paramètre déterminant est bien sûr le signe de la dérivée du terme source, c'est-à-dire le régime (attractif ou répulsif) dans lequel l'approximation \mathbf{U}^n se trouve. Ceci est cohérent avec l'analyse en stabilité linéaire (inconditionnelle dès que $\theta \geq \frac{1}{2}$) qui considère toujours un terme dissipatif au second membre.
- On voit clairement la différence avec le cas homogène où la solution entropique est construite grâce à un semi-groupe de contractions [53]:

$$\|u(\cdot, t) - v(\cdot, t)\|_{L^1_{loc}} \leq \|u_0 - v_0\|_{L^1_{loc}}$$

L'approche semi-discrétisée nous assure donc que \bar{l} est de la forme $-\partial_x f'(u)/2$ et les méthodes implicites sont assurées d'une bonne stabilité linéaire sur ce type de problèmes (avec tout de même la restriction en présence de chocs forts).

5.4 Application au transport des fractions massiques

Le point de départ est un système 3×3 de lois de conservations intervenant dans la théorie des écoulements diphasiques réactifs [81]:

$$\begin{cases} \partial_t(\rho) + \partial_x(\rho u) = 0 \\ \partial_t(\rho c) + \partial_x(\rho u c) = K \cdot \rho g(c, \rho) \\ \partial_t(\rho u^2) + \partial_x(\rho u^2 + P) = 0 \end{cases}$$

Si l'on suppose que $\frac{\partial P}{\partial c}(\rho, c) \equiv 0$, l'équation de transport-réaction sur la concentration peut s'écrire sans ambiguïté de façon non-conservative. On obtient une équation scalaire semi-linéaire du type que nous venons d'étudier. Dès que l'on a figé ρ , elle s'écrit:

$$\begin{cases} \partial_t c + \tilde{u} \cdot \partial_x c = K \cdot g(c) \text{ avec } c \in [0, 1], K > 0 \\ g(c) = c(1-c)(\mu - c) \text{ et } \mu \in]0, 1[\end{cases} \quad (5.20)$$

On va donc s'attacher à résoudre cette EDP le plus correctement possible au moyen des θ -méthodes. Le point important étant bien sûr de bien converger vers l'unique régime permanent stable $c \equiv \mu$.

L'existence et l'unicité de $c(x, t)$ se démontre grâce au changement de variable $c(x, t) = c(X(t), t)$ avec $X(t) = x_0 + \tilde{u} \cdot t$ qui nous ramène à une équation différentielle ordinaire. (méthode des caractéristiques)

5.4.1 Choix du flux numérique

Le choix du flux numérique ne doit pas se faire à la légère. Il en existe beaucoup (voir [35]), dont par exemple celui de Lax-Friedrichs qui correspond à:

$$F^{LF}(x, y) = \frac{\tilde{u}(x+y)}{2} + \frac{x-y}{2\lambda}$$

Malheureusement, bien que consistant, ce flux ne vérifie pas pour un Δt quelconque l'hypothèse faite tout au long de ce chapitre, à savoir:

$$\partial_x F^{LF}(x, y) \geq 0 \text{ et } \partial_y F^{LF}(x, y) \leq 0$$

D'autre part, un théorème présent dans [9] nous indique que ce schéma implicite cesse d'être TVD pour le problème homogène dès que $|\tilde{u}| \frac{\Delta t}{\Delta x} > 1$. Ce qui va entraîner des oscillations parasites dans la solution numérique et une catastrophe au vu de la dynamique induite par le terme de réaction.

Une autre possibilité est le flux de Godunov se ramenant dans notre cas au schéma décentré:

$$F^G(x, y) = \tilde{u} \cdot (x \cdot \mathbf{1}_{\tilde{u} \geq 0} + y \cdot \mathbf{1}_{\tilde{u} \leq 0})$$

Dans ce cas, il est simple de s'assurer que l'hypothèse faite sur le gradient du flux est vérifiée pour tout Δt . On pose donc: $\mathbf{C}^n = (c_j^n)$ l'approximation discrète de la moyenne des $c(x, t^n = n.\Delta t)$ sur les intervalles $[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$. On est donc en position d'écrire un θ -schéma pour (5.20). Supposons que $\tilde{u} > 0$, il vient:

$$\begin{aligned} \mathbf{C}^{n+1} + \theta\Delta t.A_{\Delta x}(\mathbf{C}^{n+1}) &= \mathbf{C}^n - (1 - \theta)\Delta t.A_{\Delta x}(\mathbf{C}^n) \\ \text{avec } A_{\Delta x}(\mathbf{C}^n) &= \left[\frac{\tilde{u}}{\Delta x}(c_j^n - c_{j-1}^n) - c_j^n(1 - c_j^n)(\mu - c_j^n) \right]_j \end{aligned} \quad (5.21)$$

5.4.2 Inversibilité de la partie implicite

On suppose à partir de maintenant que $c_0(x) \in [0, 1]$ et $\tilde{u} > 0$. On désire appliquer nos résultats de stabilité non-linéaire pour s'assurer de la cohérence d'une éventuelle approximation implicite. Le premier critère à vérifier est la condition (5.19):

$$\forall c \in [0, 1], 1 - \theta K \Delta t . g'(c) > 0$$

Lemme II.5 Une condition suffisante d'inversibilité du θ -schéma pour le problème (5.20) est donnée par:

$$K.\Delta t < \min \left[\frac{1}{\theta\mu}, \frac{1}{\theta(1-\mu)} \right] \quad (5.22)$$

Démonstration

Posons $f(x) \stackrel{def}{=} 1 - \theta\Delta t.g'(x)$. On cherche les racines du polynôme $f(x) = 0$:

$$3x^2 - 2(1 + \mu)x + \left(\mu - \frac{1}{K.\theta\Delta t}\right) = 0$$

Le discriminant est un polynôme du second degré en μ :

$$\Delta_f = 4\left(\mu^2 - \mu + \left(1 + \frac{3}{K.\theta\Delta t}\right)\right) \stackrel{def}{=} 4P(\mu)$$

On voit que $P(0) > 0$ et que le discriminant de P : $\Delta_P = 1 - 4\left(1 + \frac{3}{K.\theta\Delta t}\right) < 0$. Donc $\Delta_f > 0$ et l'équation $f(x) = 0$ admet deux racines réelles:

$$r_{1/2} = \frac{(1 + \mu) \pm \sqrt{(\mu^2 - \mu + (1 + \frac{3}{K.\theta\Delta t}))}}{3}$$

Il est naturel d'exiger que $r_2 < 0$ et $r_1 > 1$, i.e.

$$(1 + \mu) + \sqrt{(\mu^2 - \mu + (1 + \frac{3}{K.\theta\Delta t}))} > 3$$

$$(1 + \mu) - \sqrt{(\mu^2 - \mu + (1 + \frac{3}{K.\theta\Delta t}))} < 0$$

Ce qui donne:

$$K.\Delta t < \min \left(\frac{1}{\theta\mu}, \frac{1}{\theta(1-\mu)} \right)$$

□

Sous cette restriction de pas de temps, le théorème II.4 nous permet d'affirmer que la θ -méthode est bien définie pour le problème (5.20).

Conséquences: Forts de ces résultats, on peut donc remarquer que:

-Si le problème considéré est totalement linéaire et si la solution approchée est régulière, sa résolution numérique tombe sous le coup des théories A/L -stables [43] et ne présente donc pas d'ennuis particuliers.

-En régime raide (fortement réactionnel, par ex.), l'usage d'un *grand* pas de temps est susceptible d'occasionner la perte de certains équilibres stables si le terme source ou la fonction de flux numérique sont non-linéaires: voir à ce sujet [88] pour une étude des *spurious solutions* obtenues avec des méthodes de quadrature explicites.

5.4.3 Exemples numériques

On présente ici les comportements incorrects de la méthode d'Euler rétrograde ($\theta = 1$) pour le cas suivant:

$$\begin{cases} \partial_t c + \tilde{u} \cdot \partial_x c = K \cdot g(c) \\ g(c) = c(1-c)(\mu-c), \mu = 0.2 \\ c_0(x) = C_0 \text{ avec } C_0 = 0.95, C_0 = 0.01 \end{cases}$$

Comme la donnée initiale est constante, la solution ne va pas développer de discontinuités du fait de la forme de l'équation. Néanmoins, il est possible de mettre en évidence des comportements incorrects en exploitant la non-linéarité du terme source. On a vu que la condition d'inversibilité (5.19) ne se manifeste dans un problème de transport que du fait de la présence du terme source via la relation (5.22). Dans le cas présent, celle-ci impose que $K\Delta t < \min(5, 1.25)$ puisque l'on a fixé $\theta = 1$. On a donc effectué deux tests avec des données initiales constantes en espace mais situées près ou dans une zone répulsive ($g' > 0$) du second membre. On a choisi les valeurs 0.95 puis 0.01: les résultats sont présentés sous la forme d'une évolution en temps dans les figures 5.1 et 5.2.

Les traits pointillés représentent la convergence correcte vers l'unique équilibre stable μ obtenue avec un pas de temps petit. Le trait plein représente les états obtenus pour $K\Delta t = 6$: la solution numérique commence par sortir de l'intervalle physiquement réaliste $[0, 1]$ puis se stabilise autour de la racine instable la plus proche. On est donc en présence d'un comportement incorrect puisque le schéma ne respecte pas la nature des équilibres de l'équation (5.20). Si on augmente encore le pas de temps, on obtient des oscillations autour des équilibres instables comme on le voit sur la figure 5.3 obtenue avec la même équation et une valeur de $K\Delta t = 14$.

D'autres numerical artefacts ont d'ores-et-déjà été recensés pour ce type d'équations, notamment en présence de discontinuités. On peut citer [62] où des ondes progressant à des vitesses incorrectes furent mises en évidence pour un pas de temps en-deçà de la limite de stabilité linéaire. Il semble donc nécessaire de procéder à un examen plus attentif du comportement des discrétisations implicites en présence de solutions présentant des zones de forte variation, ce qui est somme toute un contexte pertinent lorsque l'on traite des problèmes hyperboliques.

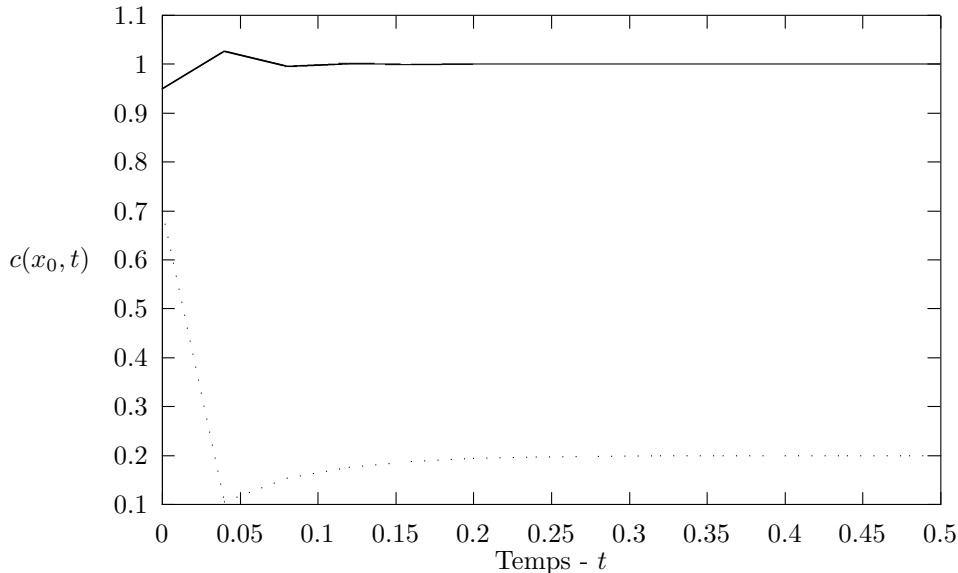


Figure 5.1: Convergence vers un équilibre instable avec $C_0 = 0.95$

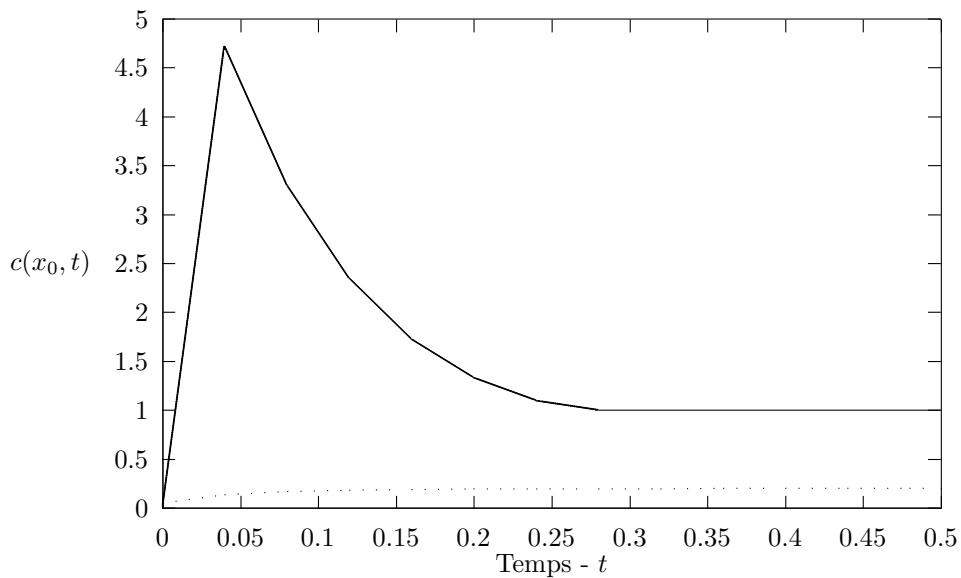


Figure 5.2: Convergence vers un équilibre instable avec $C_0 = 0.01$

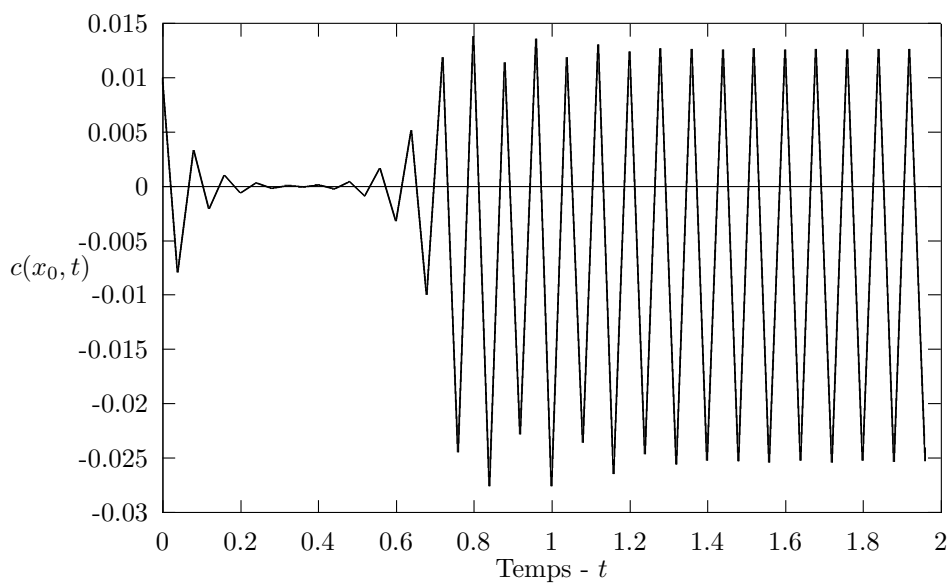


Figure 5.3: Solution oscillante issue de $C_0 = 0.01$

Chapitre 6

Stabilité des méthodes de Runge-Kutta implicites - Cas des ondes discontinues

Il a été vu au chapitre précédent que les méthodes implicites à un pas présentaient des avantages certains pour l'approximation des lois de conservation scalaires non-homogènes sous réserve que le régime ne soit pas trop raide. Néanmoins, ces routines ne sont au mieux que d'ordre 2. Nous allons donc maintenant détailler la construction de méthodes plus sophistiquées et linéairement stables pour les EDP hyperboliques. Mais il se révélera vite nécessaire de procéder à une analyse en stabilité du type *variation totale* puisque l'on verra que les méthodes d'ordre élevé ont tendance à générer des oscillations.

6.1 Stabilité linéaire des routines de quadrature

L'objectif est d'intégrer efficacement le problème différentiel générique $\dot{x} = f(x)$ où f est supposée suffisamment régulière. On s'apercevra vite que la précision en temps (ordre élevé) et la robustesse (stabilité linéaire) sont des objectifs difficiles à conjuguer, même pour des méthodes implicites. La référence permanente tout au long de cette section sera l'ouvrage de E. Hairer et G. Wanner [43].

6.1.1 Exemple: la décroissance exponentielle

On considère l'équation scalaire linéaire très simple:

$$\begin{aligned} \dot{x} &= ax \text{ avec } a \in \mathbb{R}_*^- \text{ et } x(0) = x_0 \\ \text{La solution est donnée par } x(t) &= x_0 e^{at} \xrightarrow{t \rightarrow +\infty} 0 \end{aligned} \quad (6.1)$$

Supposons que l'on effectue une intégration numérique explicite de cette équation. On désigne par x^n une approximation au temps t^n de la solution continue $x(n\Delta t)$ de (6.1):

$$x^{n+1} = x^n + \Delta t a x^n = (1 + a\Delta t)x^n \quad (6.2)$$

Ce qui implique que: $|x^n| \xrightarrow{n \rightarrow +\infty} +\infty$ si $|1 + a\Delta t| > 1$.

Ceci se traduit par des oscillations de plus en plus fortes de la solution numérique autour de zéro alors qu'elle devrait s'y écraser très vite. Cet exemple introduit le concept de *raideur numérique*: il existe une couche-limite dans laquelle $x(t)$ varie très fort alors qu'elle est quasiment stationnaire partout ailleurs. Néanmoins, pour rester qualitativement cohérent avec la réalité, il est nécessaire d'imposer un pas de temps Δt adapté à cette zone de forte variation.

6.1.2 La fonction de stabilité

Pour traiter cette difficulté, il est commode d'introduire une fonction $\mathcal{R} : \mathbb{C} \rightarrow \mathbb{C}$ telle que, pour une méthode de quadrature donnée du problème (6.1), on ait:

$$x^{n+1} = \mathcal{R}(a\Delta t)x^n$$

Cette fonction s'appelle la *fonction de stabilité* de la méthode.

Par exemple, la méthode (6.2) donne: $\mathcal{R} : z \mapsto 1 + z$.

Le résultat suivant est immédiat mais son contenu est essentiel:

Proposition II.4 Une méthode de quadrature numérique est stable sur le problème (6.1) à Δt fixé si et seulement si:

$$|\mathcal{R}(a\Delta t)| \leq 1$$

On introduit des notions pour décrire le comportement d'une routine sur (6.2).

Définition II.3 • On appelle *domaine de stabilité absolue* d'une méthode l'ensemble

$$D = \{z \in \mathbb{C} \text{ tels que } |\mathcal{R}(z)| \leq 1\}$$

- Une méthode est dite *A-stable* si on a: $\{Re(z) \leq 0\} \subset D$.
- Elle sera *L-stable* si en plus, $\mathcal{R}(z) \rightarrow 0$ lorsque $Re(z) \rightarrow -\infty$.

Toute la difficulté est bien sûr de construire des méthodes d'ordre élevé, peu coûteuses et L-stables. Ce n'est pas toujours possible:

Théorème II.5 (Dahlquist)

Une méthode multipas linéaire A-stable est d'ordre inférieur ou égal à 2 et implicite.

Exemple La méthode d'Euler rétrograde: $x^{n+1} = x^n + \Delta t a x^{n+1}$

$$x^{n+1} = \frac{x^n}{1 - a\Delta t} \tag{6.3}$$

La fonction de stabilité associée est donc $\mathcal{R}(z) = \frac{1}{1-z}$. C'est donc une méthode L -stable.

6.1.3 Méthodes implicites de Runge-Kutta (cas linéaire)

Tout comme leurs homologues explicites, elles se représentent usuellement à l'aide d'un tableau, mais dont la diagonale est cette fois non-nulle. Soit le problème linéaire:

$$\dot{X} = MX \text{ avec } X : \mathbb{R}^+ \rightarrow \mathbb{R}^N \text{ et } M \in \mathbf{M}_N(\mathbb{R}) \tag{6.4}$$

Une telle méthode à s étages s'écrit:

$$\begin{cases} K_i = M(X^n + \Delta t \sum_{j=1}^s a_{ij} K_j) \quad (i = 1, \dots, s) \\ X^{n+1} = X^n + \Delta t \sum_{i=1}^s b_i K_i \end{cases} \tag{6.5}$$

Les coefficients se notent dans un tableau (les c_i s'utilisent dans des problèmes non-autonomes):

$$\begin{array}{c|ccc} c_i & & \left(\begin{array}{ccc} & & \\ & a_{ij} & \\ & & \end{array} \right)_{1 \leq i, j \leq s} & \\ \hline & & & b_j \end{array} \tag{6.6}$$

Naturellement, plus ce tableau est plein, plus la méthode est implicite, donc robuste et plus elle est coûteuse. Il est donc légitime, au vu des problèmes aux dérivées partielles complexes que l'on désire traiter, de se limiter aux méthodes "simples" i.e. celles dont seuls les coefficients $(a_{ij})_{j \leq i}$ sont non-nuls. On peut aussi exiger que les valeurs sur la diagonales soient identiques: $a_{ii} \equiv \gamma$. Ceci afin de n'avoir qu'une seule matrice à inverser (SDIRK: Single Diagonally Implicit Runge-Kutta) ce qui est très intéressant si on utilise une routine d'inversion directe du type LU , par exemple.

$$\begin{array}{c|ccc} c_i & \left(\begin{array}{ccc} \gamma & 0 & 0 \\ & \ddots & 0 \\ a_{ij} & & \gamma \end{array} \right)_{1 \leq i, j \leq s} & \\ \hline & & & a_{sj} \end{array} \tag{6.7}$$

Proposition II.5 La fonction de stabilité d'une telle méthode à s étages est de la forme:

$$\mathcal{R}(z) = \frac{P(z)}{(1 - \gamma z)^s} \tag{6.8}$$

où P est un polynôme de degré inférieur ou égal à s

Il est donc clair que la L -stabilité exigera que $\deg(P) < s$ et $\gamma > 0$.

Exemple Méthode SDIRK L -stable d'ordre 4 à 5 étages [43]:

$$\begin{array}{c|ccccc} \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ \frac{3}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 \\ \frac{11}{20} & \frac{17}{50} & -\frac{1}{25} & \frac{1}{4} & 0 & 0 \\ \frac{1}{2} & \frac{371}{1360} & -\frac{137}{2720} & \frac{15}{544} & \frac{1}{4} & 0 \\ 1 & \frac{25}{24} & -\frac{49}{48} & \frac{125}{16} & -\frac{85}{12} & \frac{1}{4} \\ \hline & \frac{25}{24} & -\frac{49}{48} & \frac{125}{16} & -\frac{85}{12} & \frac{1}{4} \end{array} \tag{6.9}$$

6.1.4 Méthodes de Runge-Kutta semi-implicites (cas non-linéaire)

Les routines de Runge-Kutta implicites sont adaptées à la résolution de problèmes différentiels linéaires. Leur mise en oeuvre dans le cas non-linéaire $\dot{X} = G(X)$ implique l'utilisation d'un algorithme d'inversion de type Newton. Néanmoins, le fait d'en tronquer les itérations successives ne permet a priori plus d'en garantir l'ordre de précision théorique. Pour pallier à cet inconvénient, H.H. Rosenbrock [74] a introduit des méthodes de quadrature semi-implicites (SIRK) capables de maintenir un ordre prescrit en n'imposant qu'une seule itération de Newton.

Dans une méthode de Runge-Kutta implicite, si $(Id - \gamma\Delta t\mathcal{J})$ est inversible, on calcule les K_i en résolvant:

$$K_i = G \left(X^n + \Delta t \sum_{j=1}^i a_{ij} K_j \right)$$

Ceci s'inverse grâce à l'algorithme de Newton:

$$\begin{cases} K_i^{(0)} = 0 \\ K_i^{(l+1)} = K_i^{(l)} - (Id - \gamma\Delta t\mathcal{J})^{-1} \left[K_i^{(l)} - G \left(X^n + \Delta t \sum_{j=1}^{i-1} a_{ij} K_j + \Delta t a_{ii} K_i^{(l)} \right) \right] \\ \text{avec } \mathcal{J} = \left. \frac{\partial G(X)}{\partial X} \right|_{X=X^n} \end{cases} \quad (6.10)$$

Si on tronque à la première itération dans (6.10), il vient:

$$\begin{cases} K_i = (Id - \gamma\Delta t\mathcal{J})^{-1} G \left(X^n + \Delta t \sum_{j=1}^{i-1} a_{ij} K_j \right) \\ X^{n+1} = X^n + \Delta t \sum_{i=1}^s b_i K_i \end{cases} \quad (6.11)$$

Ces méthodes se notent au moyen d'un tableau similaire à (6.6) ou (6.7). On va maintenant en construire une à 2 étages, d'ordre 2 et L -stable. D'après les formules issues de développements de Taylor [43, 90], les conditions d'ordre 2 s'écrivent:

$$\begin{cases} b_1 + b_2 = 1 \\ b_1 a_{11} + b_2 (a_{21} + a_{22}) = \frac{1}{2} \end{cases} \quad (6.12)$$

D'autre part, la fonction de stabilité s'écrit:

$$\mathcal{R}(z) = \frac{1 + (b_1 + b_2 - (a_{11} + a_{22}))z + (a_{11}a_{22} - b_1a_{22} + b_2(a_{21} - a_{11}))z^2}{(1 - a_{11}z)(1 - a_{22}z)}$$

Posons $a_{11} = a_{22} = \gamma$, un critère de L -stabilité est d'annuler le terme en z^2 dans le numérateur de $\mathcal{R}(z)$:

$$\begin{cases} \gamma^2 - (b_1 + b_2)\gamma + b_2 a_{21} = 0 \\ b_1 + b_2 = 1 \\ (b_1 + b_2)\gamma + b_2 a_{21} = \frac{1}{2} \end{cases} \quad \text{d'où } \gamma + b_2 a_{21} = \frac{1}{2}$$

On obtient finalement $\gamma^2 - 2\gamma + \frac{1}{2} = 0$ dont la racine positive est donnée par $\gamma = 1 - \frac{1}{\sqrt{2}}$. D'où le tableau:

$$\begin{array}{c|cc} & \gamma & 0 \\ \hline & -\gamma & \gamma \\ \hline & 0 & 1 \end{array} \quad (6.13)$$

Remarques • Les méthodes SIRK n'ont généralement pas les mêmes propriétés (précision, stabilité) que les SDIRK à mêmes coefficients sauf à l'ordre 1 ou 2 [43]. Ceci signifie que pour ce degré de précision, il est inutile de converger l'algorithme de Newton. L'approche linéarisée du chapitre précédent est ainsi justifiée.

• La construction (6.13) est justifiée puisque la méthode de Gauss (ordre 2 - 1 étage - $a_{11} = \frac{1}{2}$ $b_1 = 1$) n'est que A -stable.

• Il peut exister des algorithmes explicites et stables [72], mais leur précision est souvent catastrophique.

6.1.5 Exemples numériques et premier bilan

On va tester la méthode SDIRK d'ordre 4 (6.9) sur une équation d'advection très simple:

$$\begin{cases} \partial_t u + \partial_x u = 0 \\ u_0(x) = \mathbf{1}_{x \leq 0.3} \\ x \in [0, 1] \end{cases}$$

La discrétisation du terme de flux se fera de manière décentrée et le système différentiel s'écrit:

$$\frac{d\vec{U}}{dt} = M\vec{U} \text{ avec } M = \frac{1}{\Delta x} \begin{pmatrix} -1 & 0 & 0 & \cdots & 0 \\ 1 & -1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & -1 \end{pmatrix}$$

Ce système est linéaire et ses valeurs propres sont toutes négatives, donc la condition d'inversibilité (5.19) n'entraîne aucune restriction sur Δt . On choisit d'intégrer ce système avec les paramètres $\Delta x = 0.01$, $\Delta t = 0.1$ et on itère jusqu'à $T = 0.3$. Le résultat est présenté dans la figure 6.1.

On effectue ensuite un test sur l'équation non-homogène:

$$\begin{cases} \partial_t u + \partial_x u = -10u \\ u_0(x) = \mathbf{1}_{x \leq 0.3} \\ x \in [0, 1] \end{cases}$$

La discrétisation du terme de flux se fera de manière décentrée, celle du terme dissipatif de façon centrée et le système différentiel obtenu s'écrit:

$$\frac{d\vec{U}}{dt} = M\vec{U} \text{ avec } M = \begin{pmatrix} \frac{-1}{\Delta x} - 10 & 0 & 0 & \cdots & 0 \\ \frac{1}{\Delta x} & \frac{-1}{\Delta x} - 10 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{\Delta x} & \frac{-1}{\Delta x} - 10 \end{pmatrix}$$

On choisit $\Delta x = 0.01$, $\Delta t = 0.07$, et on itère jusqu'à $T = 0.3$: voir la figure 6.2 pour la solution.

On propose pour finir un dernier test:

$$\begin{cases} \partial_t u + \partial_x u = -50u \\ u_0(x) = \mathbf{1}_{x \leq 0.3} \\ x \in [0, 1] \end{cases}$$

La discrétisation spatiale se fera comme précédemment. Le système différentiel sera donc:

$$\frac{d\vec{U}}{dt} = M\vec{U} \text{ avec } M = \begin{pmatrix} \frac{-1}{\Delta x} - 50 & 0 & 0 & \cdots & 0 \\ \frac{1}{\Delta x} & \frac{-1}{\Delta x} - 50 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{1}{\Delta x} & \frac{-1}{\Delta x} - 50 \end{pmatrix}$$

Les routines de quadrature seront cette fois Euler implicite (6.3) et SDIRK4 (6.9). On choisit $\Delta x = 0.01$, $\Delta t = 0.1$, et on itère jusqu'à $T = 0.5$: les figures 6.4, 6.3 montrent les solutions.

On remarque donc une différence *essentielle* entre une routine très simple d'ordre 1 et une autre plus sophistiquée: l'une préserve la monotonie de la solution théorique et l'autre pas. On voit donc qu'une analyse en stabilité linéaire est insuffisante lorsque l'on traite des solutions discontinues. Il est nécessaire d'étudier aussi la variation totale.

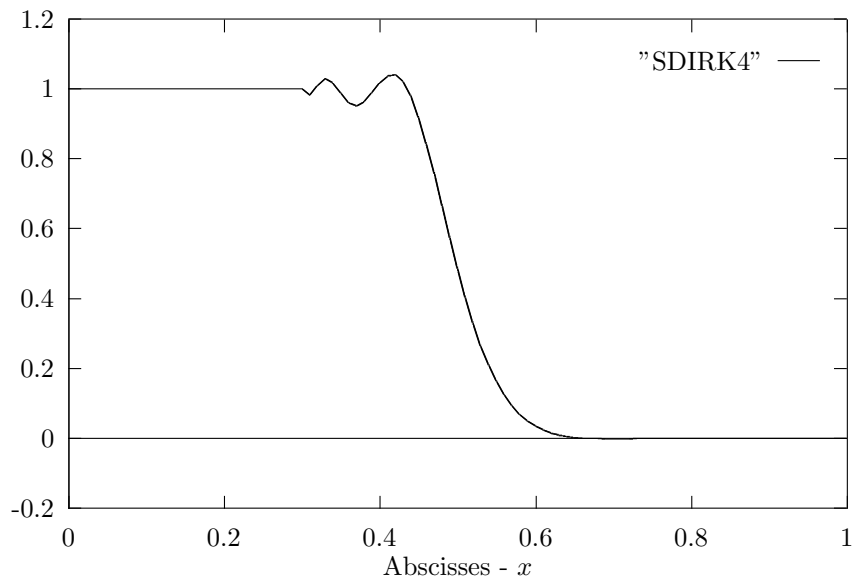


Figure 6.1: Oscillations CFL=10 - Cas homogène

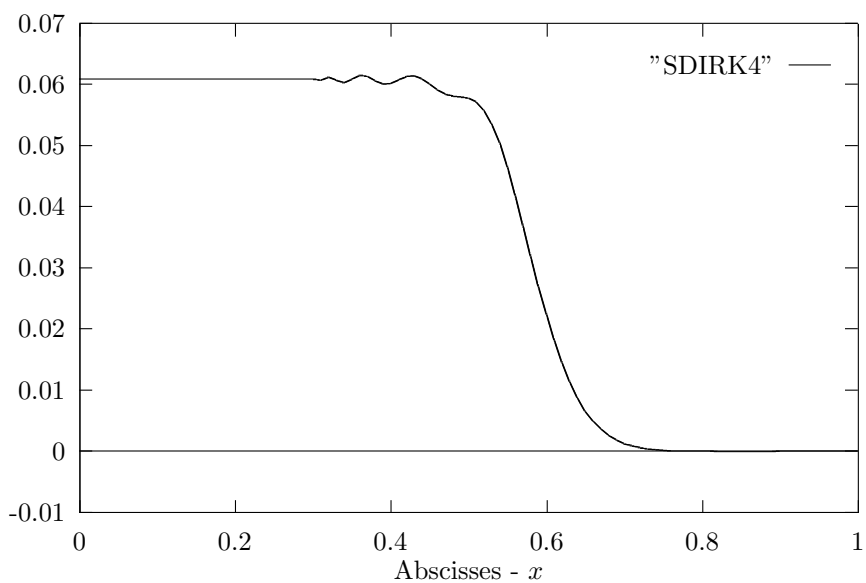
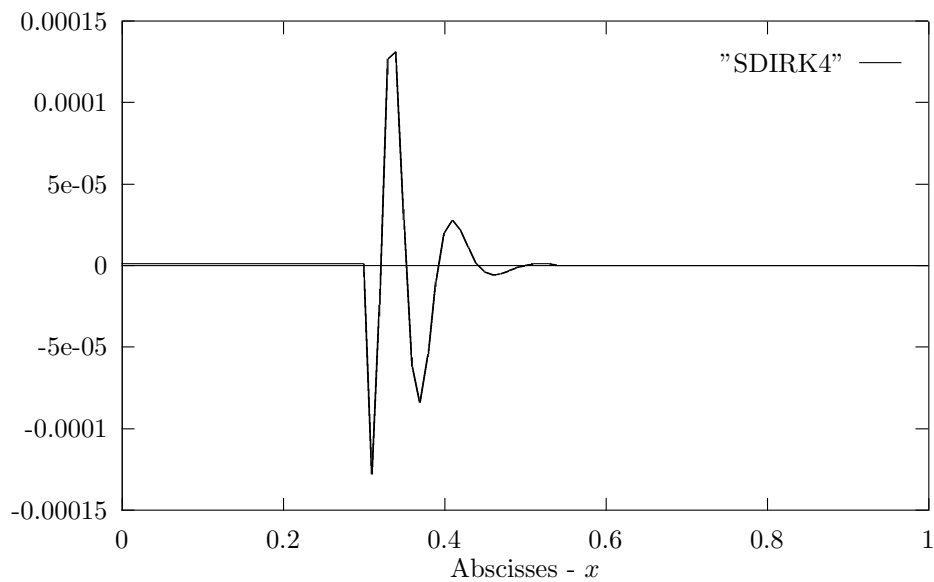
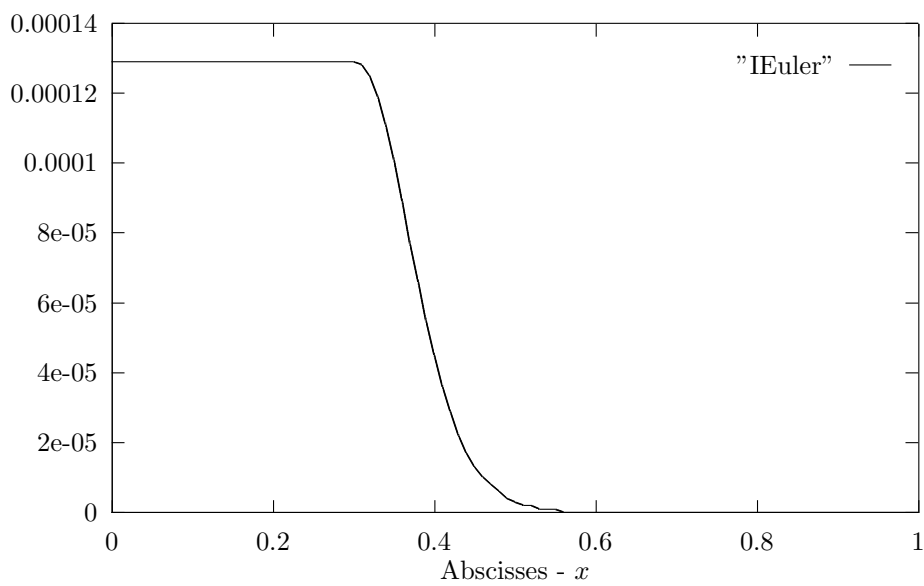


Figure 6.2: Oscillations CFL=7 - Second membre dissipatif: $-10u$

Figure 6.3: Méthode d'ordre 4 - Second membre dissipatif: $-50u$ Figure 6.4: Méthode d'Euler - Second membre dissipatif: $-50u$

6.2 Stabilité en variation totale et méthode des lignes

Il s'agit dans un premier temps de rappeler la méthode des lignes pour les problèmes hyperboliques non-homogènes. Celle-ci consiste à d'abord effectuer une discrétisation en espace au moyen d'une fonction de flux numérique et d'une approximation du terme de réaction. On obtient alors un système différentiel que l'on traitera avec une méthode de quadrature telle que celles décrites précédemment.

6.2.1 Cadre théorique: les lois scalaires 1D

On considère le problème scalaire:

$$\begin{cases} \partial_t u + \partial_x f(u) = g(u) \\ u_0(x) \in BV(\mathbb{R}) \end{cases} \quad (6.14)$$

où f et g sont suffisamment régulières. Introduisons une fonction de flux numérique $(u, v) \mapsto F(u, v)$ consistante et une discrétisation centrée du second membre $u \mapsto g(u)$. On obtient le système différentiel suivant:

$$\begin{cases} \dot{u}_j(t) + \frac{1}{\Delta x} [F(u_j(t), u_{j+1}(t)) - F(u_{j-1}(t), u_j(t))] = g(u_j(t)) \quad (j \in \mathbb{Z}) \\ u_j(0) = \frac{1}{\Delta x} \int_{(j-\frac{1}{2})\Delta x}^{(j+\frac{1}{2})\Delta x} u_0(s) ds \end{cases} \quad (6.15)$$

Tout d'abord, le théorème de Cauchy-Lipschitz assure l'existence d'une solution locale de (6.15) dès que F et G sont localement lipschitziennes. Par la suite, G. Mehlman [68] a montré le caractère global de celle-ci ainsi que la convergence vers une solution faible de (6.14) lorsque $\Delta x \rightarrow 0$ à l'aide d'estimations L^∞ et BV . Cette approche est donc *bien posée et convergente*.

En général, le système (6.15) est non-linéaire et les routines semi-implicites (6.11) sont les plus adaptées. Une approche implicite est souvent destinée à converger vers un *permanent* i.e. un état stationnaire. Néanmoins, on a vu précédemment que pour des équations scalaires très simples, ce type de méthodes pouvait générer des oscillations et faire exploser la suite d'approximations.

6.2.2 Variation totale des approximations de type Runge-Kutta

Généralités

L'idée est ici de généraliser la démarche introduite par C.W. Shu et S. Osher [78] à des méthodes de quadrature implicites afin de trouver un critère capable de contrôler la variation totale d'une approximation issue d'un algorithme du type (6.3), (6.5), (6.11).

Pour un problème linéaire $\dot{X} = MX$, une telle méthode à s étages s'écrit:

$$\begin{cases} K_i = M(X^n + \Delta t \sum_{j=1}^i a_{ij} K_j) \quad (i = 1, \dots, s) \\ X^{n+1} = X^n + \Delta t \sum_{i=1}^s b_i K_i \end{cases}$$

On remarque que si on pose formellement:

$$\begin{cases} X^{(0)} = X^n \\ X^{(i)} = M^{-1}(K^i) \\ X^{n+1} = X^{(s+1)} \end{cases}$$

il vient la formulation linéaire (voir aussi [1, 19]):

$$\begin{cases} X^{(i)} = X^{(0)} + \Delta t \sum_{j=1}^i a_{ij} M X^{(j)} \quad 1 \leq i \leq s \text{ étages implicites} \\ X^{(s+1)} = X^{(0)} + \Delta t \sum_{j=1}^s b_j M X^{(j)} \text{ dernière étape explicite} \end{cases}$$

D'où l'écriture modifiée pour la méthode (6.5):

$$\begin{cases} X^{(0)} = X^n \\ X^{(i)} = X^{(0)} + \Delta t \sum_{j=1}^i c_{ij} M X^{(j)} \quad (1 \leq i \leq s+1) \\ X^{n+1} = X^{(s+1)} \end{cases} \quad (6.16)$$

La matrice $C = (c_{ij})_{1 \leq i, j \leq s+1}$ est de la forme:

$$C = \begin{pmatrix} a_{11} & 0 & \cdots & \cdots & 0 \\ a_{21} & a_{22} & \ddots & 0 & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ a_{s1} & a_{s2} & \cdots & a_{ss} & 0 \\ b_1 & b_2 & \cdots & b_s & 0 \end{pmatrix} \quad (6.17)$$

Supposons que l'on s'intéresse à un système différentiel issu de la méthode des lignes, alors les routines les plus adaptées à un tel problème non-linéaire sont semi-implicites (6.11) et les étapes de (6.16) seront de la forme:

$$\begin{cases} \mathbf{U}^{(0)} = \mathbf{U}^n \\ u_j^{(i)} = u_j^{(0)} - \sum_{k=1}^i c_{ik} \left[\frac{\Delta t}{\Delta x} [F(u_j^{(k)}, u_{j+1}^{(k)}) - F(u_{j-1}^{(k)}, u_j^{(k)})] - \Delta t g(u_j^{(k)}) \right] \quad (1 \leq i \leq s+1) \\ \mathbf{U}^{n+1} = \mathbf{U}^{(s+1)} \end{cases} \quad (6.18)$$

L'objectif est alors d'estimer le seuil d'apparition d'oscillations parasites dans l'approximation issue de (6.18) en fonction de la condition CFL classique qui contrôle la stabilité des discrétisations explicites de (6.15).

La méthode d'Euler rétrograde

Dans ce paragraphe, \mathcal{L} désignera l'opérateur discrétisé en espace intervenant dans l'écriture (6.15) ou (6.18). C'est-à-dire qu'avec les notations du chapitre précédent, on a $\mathcal{L}(\mathbf{U}^n) = -A_{\Delta x}(\mathbf{U}^n)$ et \bar{l} est la constante introduite au chapitre précédent à la proposition II.3. On établit tout d'abord un résultat préliminaire concernant la méthode d'Euler rétrograde:

$$u_j^{n+1} = u_j^n - \left[\frac{\Delta t}{\Delta x} [F(u_{j+1}^{n+1}, u_j^{n+1}) - F(u_j^{n+1}, u_{j-1}^{n+1})] - \Delta t g(u_j^{n+1}) \right] \quad (6.19)$$

Lemme II.6 *Si le flux numérique $(u, v) \mapsto F(u, v)$ est consistant (5.3), monotone (5.4) et si la condition d'inversibilité (5.19) est vérifiée, alors la variation totale de $\mathbf{U}^{n+1} = (u_j^{n+1})$ est bornée et vérifie:*

$$(1 - \Delta t \bar{l}) . TV(\mathbf{U}^{n+1}) \leq TV(\mathbf{U}^n)$$

Ceci signifie en particulier que le schéma (6.19) est inconditionnellement borné sur un problème linéaire homogène ou dissipatif du type: $\partial_t u + a \partial_x u = -|b|u$ puisqu'alors, la condition (5.19) est vérifiée pour tout Δt .

Démonstration

On a:

$$\begin{cases} u_j^{n+1} = u_j^n - \left[\frac{\Delta t}{\Delta x} [F(u_j^{n+1}, u_{j+1}^{n+1}) - F(u_{j-1}^{n+1}, u_j^{n+1})] - \Delta t g(u_j^{n+1}) \right] \\ u_{j-1}^{n+1} = u_{j-1}^n - \left[\frac{\Delta t}{\Delta x} [F(u_{j-1}^{n+1}, u_j^{n+1}) - F(u_{j-2}^{n+1}, u_{j-1}^{n+1})] - \Delta t g(u_{j-1}^{n+1}) \right] \end{cases}$$

On effectue la différence:

$$(u_j^{n+1} - u_{j-1}^{n+1}) + \frac{\Delta t}{\Delta x} [F(u_j^{n+1}, u_{j+1}^{n+1}) - 2F(u_{j-1}^{n+1}, u_j^{n+1}) + F(u_{j-2}^{n+1}, u_{j-1}^{n+1})] - \Delta t [g(u_j^{n+1}) - g(u_{j-1}^{n+1})] = (u_j^n - u_{j-1}^n)$$

Comme $F : \mathbb{R}^2 \rightarrow \mathbb{R}$, par le lemme II.3, il existe $\xi_{j+\frac{1}{2}} \in [u_j^{n+1}, u_{j-1}^{n+1}] \times [u_{j+1}^{n+1}, u_j^{n+1}]$, $\xi_{j-\frac{1}{2}} \in [u_{j-2}^{n+1}, u_{j-1}^{n+1}] \times [u_{j-1}^{n+1}, u_j^{n+1}]$, $\zeta_{j+\frac{1}{2}} \in [u_j^{n+1}, u_{j-1}^{n+1}]$ tels que:

$$\begin{cases} F(u_j^{n+1}, u_{j+1}^{n+1}) = F(u_{j-1}^{n+1}, u_j^{n+1}) + \nabla F(\xi_{j+\frac{1}{2}}) \cdot \begin{pmatrix} u_j^{n+1} - u_{j-1}^{n+1} \\ u_{j+1}^{n+1} - u_j^{n+1} \end{pmatrix} \\ F(u_{j-2}^{n+1}, u_{j-1}^{n+1}) = F(u_{j-1}^{n+1}, u_j^{n+1}) + \nabla F(\xi_{j-\frac{1}{2}}) \cdot \begin{pmatrix} u_{j-2}^{n+1} - u_{j-1}^{n+1} \\ u_{j-1}^{n+1} - u_j^{n+1} \end{pmatrix} \\ g(u_j^{n+1}) - g(u_{j-1}^{n+1}) = g'(\zeta_{j+\frac{1}{2}})(u_j^{n+1} - u_{j-1}^{n+1}) \end{cases}$$

D'où l'expression:

$$\begin{aligned} (u_j^{n+1} - u_{j-1}^{n+1}) + \frac{\Delta t}{\Delta x} [F_u(\xi_{j+\frac{1}{2}})(u_j^{n+1} - u_{j-1}^{n+1}) + F_v(\xi_{j+\frac{1}{2}})(u_{j+1}^{n+1} - u_j^{n+1}) \\ - F_u(\xi_{j-\frac{1}{2}})(u_{j-1}^{n+1} - u_{j-2}^{n+1}) - F_v(\xi_{j-\frac{1}{2}})(u_j^{n+1} - u_{j-1}^{n+1})] \\ - \Delta t g'(\zeta_{j+\frac{1}{2}})(u_j^{n+1} - u_{j-1}^{n+1}) = (u_j^n - u_{j-1}^n) \end{aligned}$$

Sous les conditions (5.4) et (5.19), on a:

$$\begin{aligned} (u_j^{n+1} - u_{j-1}^{n+1}) \left[1 + \frac{\Delta t}{\Delta x} (F_u(\xi_{j+\frac{1}{2}}) - F_v(\xi_{j-\frac{1}{2}})) - \Delta t g'(\zeta_{j+\frac{1}{2}}) \right] \\ + (u_{j+1}^{n+1} - u_j^{n+1}) \underbrace{\frac{\Delta t}{\Delta x} F_v(\xi_{j+\frac{1}{2}})}_{\leq 0} - (u_{j-1}^{n+1} - u_{j-2}^{n+1}) \underbrace{\frac{\Delta t}{\Delta x} F_u(\xi_{j-\frac{1}{2}})}_{\geq 0} = (u_j^n - u_{j-1}^n) \end{aligned}$$

On prend alors le module de cette égalité. Comme $|a + b - c| \geq |a| - |b| - |c|$, il vient:

$$\begin{aligned} |u_j^{n+1} - u_{j-1}^{n+1}| \left[1 + \frac{\Delta t}{\Delta x} (F_u(\xi_{j+\frac{1}{2}}) - F_v(\xi_{j-\frac{1}{2}})) - \Delta t g'(\zeta_{j+\frac{1}{2}}) \right] \\ - |u_{j+1}^{n+1} - u_j^{n+1}| \left| \frac{\Delta t}{\Delta x} F_v(\xi_{j+\frac{1}{2}}) \right| - |u_{j-1}^{n+1} - u_{j-2}^{n+1}| \left| \frac{\Delta t}{\Delta x} F_u(\xi_{j-\frac{1}{2}}) \right| \leq |u_j^n - u_{j-1}^n| \end{aligned}$$

On remarque que l'hypothèse de monotonie sur F (5.4) assure que:

$$\begin{cases} -|F_v(\xi_{j+\frac{1}{2}})| = -(-F_v(\xi_{j+\frac{1}{2}})) = F_v(\xi_{j+\frac{1}{2}}) \\ -|F_u(\xi_{j-\frac{1}{2}})| = -F_u(\xi_{j-\frac{1}{2}}) \end{cases}$$

Donc l'expression précédente devient:

$$\begin{aligned} |u_j^{n+1} - u_{j-1}^{n+1}| \left[1 + \frac{\Delta t}{\Delta x} (F_u(\xi_{j+\frac{1}{2}}) - F_v(\xi_{j-\frac{1}{2}})) - \Delta t g'(\zeta_{j+\frac{1}{2}}) \right] \\ + |u_{j+1}^{n+1} - u_j^{n+1}| \left(\frac{\Delta t}{\Delta x} F_v(\xi_{j+\frac{1}{2}}) \right) - |u_{j-1}^{n+1} - u_{j-2}^{n+1}| \left(\frac{\Delta t}{\Delta x} F_u(\xi_{j-\frac{1}{2}}) \right) \leq |u_j^n - u_{j-1}^n| \end{aligned}$$

Finalement, si on pose $\Gamma_{j+\frac{1}{2}} = 1 - \Delta t g'(\zeta_{j+\frac{1}{2}})$, il est clair que:

$$\forall j, \Gamma_{j+\frac{1}{2}} \geq 1 - \Delta t \bar{l}$$

Comme $TV(\mathbf{U}^{n+1}) = \sum_{j \in \mathbb{Z}} |u_j^{n+1} - u_{j-1}^{n+1}|$, on voit que:

$$TV(\mathbf{U}^{n+1}) \leq \frac{TV(\mathbf{U}^n)}{1 - \Delta t \bar{l}}$$

□

Pour une équation d'advection homogène, on a un flux décentré et donc $TV(\mathbf{U}^{n+1}) \leq TV(\mathbf{U}^n)$ quelque soit le pas de temps Δt .

Le résultat général

L'idée pour contrôler la variation totale des méthodes à s étages consiste donc à réécrire chacun d'entre eux sous la forme de plusieurs étapes explicites (le minimum!) et d'une étape du type Euler rétrograde.

Théorème II.6 Soit $(\alpha_{ij})_{0 \leq i, j \leq s+1}$ tels que $\alpha_{ij} \geq 0$ et $\sum_{j=0}^{i-1} \alpha_{ij} = 1$, alors la méthode de Runge-Kutta (6.18) peut s'écrire sous la forme:

$$\begin{cases} \mathbf{U}^{(0)} = \mathbf{U}^n \\ \mathbf{U}^{(i)} = \sum_{j=0}^{i-1} (\alpha_{ij} \mathbf{U}^{(j)} + \Delta t \beta_{ij} \mathcal{L}(\mathbf{U}^{(j)})) + \Delta t c_{ii} \mathcal{L}(\mathbf{U}^{(i)}) \quad (1 \leq i \leq s+1) \\ \mathbf{U}^{n+1} = \mathbf{U}^{(s+1)} \end{cases} \quad (6.20)$$

avec $\beta_{ij} = c_{ij} - \sum_{l=j}^{i-1} \alpha_{il}c_{lj}$. De plus, si les conditions (5.3), (5.4), (5.19) avec $\theta = c_{ii}$ sont satisfaites et si $\forall (i, j), \beta_{ij} \geq 0$, le schéma ci-dessus (6.20) sera non-oscillant sous la condition de type CFL

$$\sup_{j,n} |f'(u_j^n)| \cdot \frac{\Delta t}{\Delta x} \leq \min_{i,j} \left(\frac{\alpha_{ij}}{\beta_{ij}} \right) \quad (6.21)$$

Il faut tout de même noter que si le système différentiel est non-linéaire, la mise en place d'une méthode implicite nécessite un algorithme d'inversion de type Newton qui est susceptible d'imposer lui aussi des restrictions de pas de temps.

Démonstration

Introduisons une matrice $(\alpha_{ij})_{0 \leq i, j \leq s+1}$ telle que:

$$\alpha_{ij} \geq 0 \text{ et } \sum_{j=0}^{i-1} \alpha_{ij} = 1$$

Alors:

$$\begin{aligned} \mathbf{U}^{(i)} &= \mathbf{U}^{(0)} + \Delta t \sum_{j=1}^i c_{ij} \mathcal{L}(\mathbf{U}^{(j)}) \\ &= \sum_{j=0}^{i-1} \alpha_{ij} \mathbf{U}^{(j)} + \Delta t \sum_{j=1}^{i-1} c_{ij} \mathcal{L}(\mathbf{U}^{(j)}) + \Delta t c_{ii} \mathcal{L}(\mathbf{U}^{(i)}) \\ &= \alpha_{i0} \mathbf{U}^{(0)} + \sum_{j=1}^{i-1} \alpha_{ij} [\mathbf{U}^{(j)} - \Delta t \sum_{k=1}^j c_{jk} \mathcal{L}(\mathbf{U}^{(k)})] + \Delta t \sum_{j=1}^{i-1} c_{ij} \mathcal{L}(\mathbf{U}^{(j)}) + \Delta t c_{ii} \mathcal{L}(\mathbf{U}^{(i)}) \end{aligned}$$

puisque à chaque itération, $\mathbf{U}^{(0)} = \mathbf{U}^{(j)} - \Delta t \sum_{k=1}^j c_{jk} \mathcal{L}(\mathbf{U}^{(k)})$. Il vient donc:

$$\begin{aligned} \mathbf{U}^{(i)} &= \alpha_{i0} \mathbf{U}^{(0)} + \sum_{j=1}^{i-1} [\alpha_{ij} \mathbf{U}^{(j)}] + \Delta t \sum_{j=1}^{i-1} [c_{ij} \mathcal{L}(\mathbf{U}^{(j)}) - \sum_{k=1}^j \alpha_{ij} c_{jk} \mathcal{L}(\mathbf{U}^{(k)})] + \Delta t c_{ii} \mathcal{L}(\mathbf{U}^{(i)}) \\ &= \sum_{j=0}^i [\alpha_{ij} \mathbf{U}^{(j)}] + \Delta t \sum_{j=1}^{i-1} [(c_{ij} - \sum_{l=j}^{i-1} \alpha_{il} c_{lj}) \mathcal{L}(\mathbf{U}^{(j)})] + \Delta t c_{ii} \mathcal{L}(\mathbf{U}^{(i)}) \end{aligned}$$

puisque

$$\begin{cases} j = 1 & : & c_{i1} \mathcal{L}(\mathbf{U}^{(1)}) - \alpha_{i1} c_{11} \mathcal{L}(\mathbf{U}^{(1)}) \\ j = 2 & : & c_{i2} \mathcal{L}(\mathbf{U}^{(2)}) - \alpha_{i2} [c_{21} \mathcal{L}(\mathbf{U}^{(1)}) + c_{22} \mathcal{L}(\mathbf{U}^{(2)})] \\ j = 3 & : & c_{i3} \mathcal{L}(\mathbf{U}^{(3)}) - \alpha_{i3} [c_{31} \mathcal{L}(\mathbf{U}^{(1)}) + c_{32} \mathcal{L}(\mathbf{U}^{(2)}) + c_{33} \mathcal{L}(\mathbf{U}^{(3)})] \end{cases}$$

et cela donne, par exemple:

$$\mathcal{L}(\mathbf{U}^{(j=1)}) [c_{i1} - \underbrace{(\alpha_{i1} c_{11} + \alpha_{i2} c_{21} + \alpha_{i3} c_{31} + \dots)}_{\sum_{l=j=1}^{i-1} \alpha_{il} c_{lj=1}}]$$

Finalement, on a bien l'écriture annoncée:

$$\mathbf{U}^{(i)} = \sum_{j=0}^{i-1} \left(\alpha_{ij} \mathbf{U}^{(j)} + \Delta t \beta_{ij} \mathcal{L}(\mathbf{U}^{(j)}) \right) + \Delta t c_{ii} \mathcal{L}(\mathbf{U}^{(i)}) \quad (1 \leq i \leq s+1)$$

avec $\beta_{ij} = c_{ij} - \sum_{l=j}^{i-1} \alpha_{il} c_{lj}$. Le résultat de stabilité provient du fait que, dans ce cas, on a une moyenne pondérée à coefficients positifs de schémas explicites non-oscillants du fait de la restriction CFL (6.21) et une méthode de type Euler rétrograde dont le pas de temps est pondéré par le coefficient c_{ii} . Sous cette condition, on obtient en remplaçant Δt par $c_{ii} \Delta t$ dans la démonstration du lemme II.6:

$$(1 - c_{ii} \Delta t \bar{l}). TV(\mathbf{U}^{(i)}) \leq \sum_{j=0}^{i-1} \left[\alpha_{ij} TV \left(\mathbf{U}^{(j)} + \Delta t \frac{\beta_{ij}}{\alpha_{ij}} \mathcal{L}(\mathbf{U}^{(j)}) \right) \right]$$

Comme l'expression de droite est une combinaison convexe d'opérations explicites, elle est soumise à la condition CFL habituelle et sera non-oscillante sous la condition suffisante:

$$\forall (1 \leq i \leq s+1), (0 \leq j \leq i-1), \sup_{j,n} |f'(u_j^n)| \cdot \frac{\Delta t \beta_{ij}}{\Delta x \alpha_{ij}} \leq 1$$

Ceci correspond bien à la majoration (6.21). \square

6.2.3 Première application: les θ -méthodes

On les a déjà rencontrées au chapitre précédent. Ce sont des routines de type Runge-Kutta très simples qui s'écrivent sous la forme:

$$\left| \begin{array}{c} \theta \in [0, 1] \\ 1 \end{array} \right. \quad (6.22)$$

D'où: $\mathbf{U}^{n+1} = \mathbf{U}^n + \Delta t K_1$ avec $K_1 = \mathcal{L}(\mathbf{U}^n + \theta \Delta t K_1)$. Pour $\theta = 1$, on a la méthode d'Euler rétrograde, pour $\theta = \frac{1}{2}$, la méthode de Gauss d'ordre 2.

$$\left\{ \begin{array}{l} \mathbf{U}^{(0)} = \mathbf{U}^n \\ \mathbf{U}^{(1)} = \mathbf{U}^{(0)} + \theta \Delta t \mathcal{L}(\mathbf{U}^{(1)}) \\ \mathbf{U}^{n+1} = \mathbf{U}^{(2)} = \mathbf{U}^{(0)} + \Delta t \mathcal{L}(\mathbf{U}^{(1)}) \end{array} \right. \quad \text{c'est-à-dire } C = \begin{pmatrix} \theta & 0 \\ 1 & 0 \end{pmatrix}$$

On calcule donc l'unique β_{ij} non-nul:

$$\left\{ \begin{array}{l} c_{11} = \theta \\ \beta_{21} = c_{21} - \sum_{l=1}^1 \alpha_{2l} c_{1l} = 1 - \alpha_{21} c_{11} = 1 - \alpha_{21} \theta \end{array} \right.$$

Par conséquent, si on choisit $\alpha_{21} = 1$ (on n'a pas le choix puisque la dernière étape d'une méthode de Runge-Kutta est explicite), il vient:

$$\sup_{j,n} |f'(u_j^n)| \cdot \frac{\Delta t}{\Delta x} \leq \frac{1}{1 - \theta}$$

Ce qui signifie en particulier que la méthode d'Euler rétrograde est inconditionnellement stable en variation totale pour un problème hyperbolique linéaire. Celle de Gauss ne l'est que pour le double de la condition CFL d'origine.

Pour vérifier ces résultats, nous allons faire un test numérique sur une équation d'advection très simple:

$$\left\{ \begin{array}{l} \partial_t u + \partial_x u = 0 \\ u_0(x) = \mathbf{1}_{x \leq 0.3} \\ x \in [0, 1] \text{ et } t \in [0, 0.1] \end{array} \right.$$

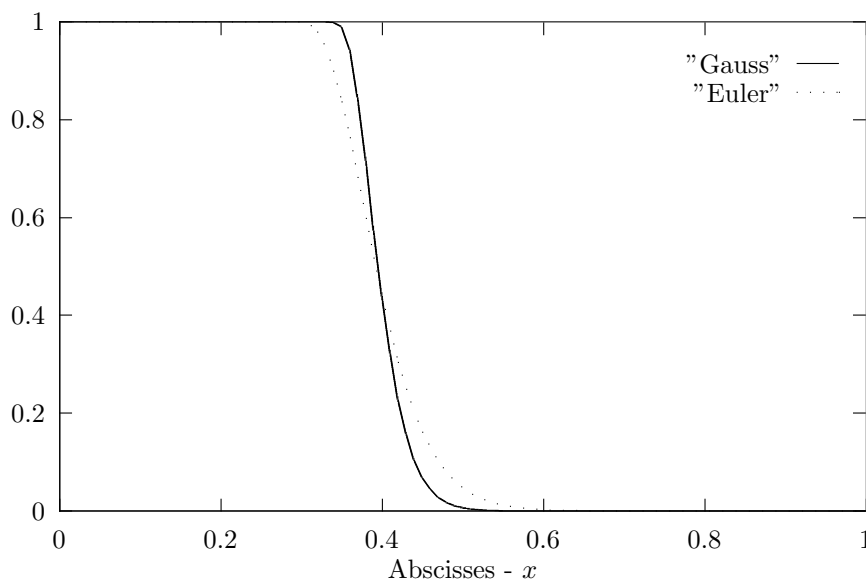
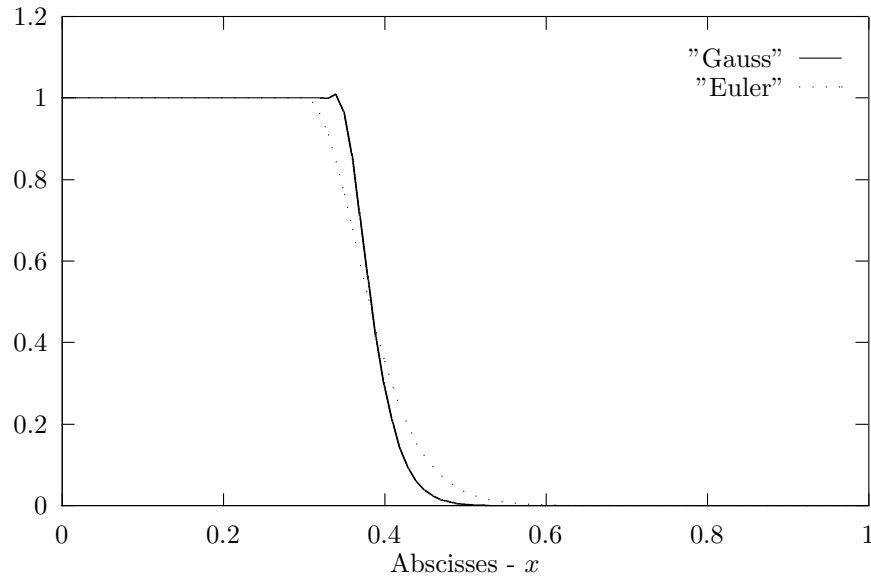


Figure 6.5: Comparaison $\theta = 1$ et $\theta = 0.5$ - CFL=1.9

Figure 6.6: Comparaison $\theta = 1$ et $\theta = 0.5$ - CFL=2.1

Sur les graphiques (6.5) et (6.6) figurent les résultats comparés de la méthode Euler rétrograde et de la méthode de Gauss. On a fixé $\Delta x = 0.01$ et on a surveillé le seuil d'apparition d'oscillations entre $\Delta t = 0.019$ (CFL=1.9) et $\Delta t = 0.021$ (CFL=2.1). Les résultats sont bien en accord avec les prévisions théoriques.

6.2.4 Les méthodes A-stables d'ordre 2 à 2 pas

La routine SIRK2 (6.13) ne tombe pas dans le champ d'application du résultat précédent à cause des restrictions de signe. Néanmoins, il est possible de constater numériquement l'apparitions d'oscillations parasites dans les solutions pour des CFL modérées (de l'ordre de 2 environ). Cela suggère que les méthodes de Runge-Kutta linéairement stables ne sont pas forcément bien adaptées aux problèmes hyperboliques. On va s'en assurer en étudiant les méthodes de quadrature proposées dans [90] pour des problèmes hypersoniques réactifs.

Pour ces méthodes, on a une matrice C de la forme:

$$C = \begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ b_1 & b_2 & 0 \end{pmatrix}$$

La A-stabilité se traduit par l'écriture des coefficients a_{ij} [90]:

$$\begin{cases} a_{11} = \frac{\frac{1}{2} - \gamma}{1 - \gamma} \\ a_{21} = (1 - \gamma) - \frac{\frac{1}{2} - \gamma}{1 - \gamma} \\ a_{22} = \gamma \end{cases} \quad \text{avec } \gamma \in [0, 1/2[$$

De plus, il faut imposer les conditions (6.12) pour garantir l'ordre 2. Ceci entraîne que:

$$\left(\frac{\frac{1}{2} - \gamma}{1 - \gamma}\right) b_1 + b_2 \left(1 - \frac{\frac{1}{2} - \gamma}{1 - \gamma}\right) = \frac{1}{2}$$

D'où:

$$\left(\frac{1}{2} - \gamma\right) b_1 + \frac{1}{2} b_2 = \frac{1 - \gamma}{2} \quad \text{donc } b_1 = b_2 = \frac{1}{2}$$

Pour contrôler la variation totale, on calcule les β_{ij} :

$$\begin{cases} \beta_{21} = c_{21} - \alpha_{21}c_{21} = a_{21} - a_{11} \\ \beta_{31} = c_{31} - \sum_{l=1}^2 \alpha_{3l}c_{l1} = b_1 - (\alpha_{31}a_{11} + \alpha_{32}a_{21}) \\ \beta_{32} = c_{32} - \alpha_{32}c_{22} = b_2 - \alpha_{32}a_{22} \end{cases}$$

avec toujours: $\alpha_{21} = 1$, $\alpha_{31} + \alpha_{32} = 1$ et $\alpha_{ij} \geq 0$.

On développe les expressions précédentes en posant $\alpha_{32} = \delta$:

$$\begin{cases} \beta_{21} = (1 - \gamma) - \frac{1-2\gamma}{1-\gamma} = \frac{\gamma^2}{1-\gamma} \\ \beta_{31} = \frac{1}{2} - \left[(1 - \delta) \frac{\frac{1}{2}-\gamma}{1-\gamma} + \delta \left((1 - \gamma) - \frac{\frac{1}{2}-\gamma}{1-\gamma} \right) \right] = \frac{1}{2} - \left[\delta(1 - \gamma) + (1 - 2\delta) \left(\frac{\frac{1}{2}-\gamma}{1-\gamma} \right) \right] \\ \beta_{32} = \frac{1}{2} - \delta\gamma \end{cases} \quad (6.23)$$

Pour obtenir une stabilité inconditionnelle grâce au théorème II.6, il faut maximiser les rapports $\frac{\alpha_{ij}}{\beta_{ij}}$, c'est-à-dire tenter d'annuler les β_{ij} . Pour cela, il faut que γ soit le plus faible possible, ce qui est absurde, car le second étage devient alors explicite.

Si on trace numériquement les surfaces de \mathbb{R}^3 définies par les équations (6.23) paramétrées par δ et γ , on voit que le meilleur choix correspond à:

$$\gamma \simeq 0.4, \delta \simeq 0.5 \text{ ce qui donne au mieux: } \sup_{j,n} |f'(u_j^n)| \cdot \frac{\Delta t}{\Delta x} \leq 2.5$$

6.2.5 Stabilité en variation totale et méthodes d'ordre 2

Après le résultat négatif précédent, il est naturel de se demander "combien" de méthodes de Runge-Kutta d'ordre 2 préservent la variation totale des approximations numériques. C'est l'objet de cette proposition.

Proposition II.6 *Il existe une famille à un paramètre réel de méthodes SDIRK/SIRK d'ordre 2 préservant la variation totale en espace.*

Démonstration

On garde les mêmes notations que dans le paragraphe précédent. Les conditions d'ordre sont données dans l'expression (6.12) et la forme des β_{ij} (qui doivent être positifs) est:

$$\begin{cases} \beta_{21} = a_{21} - a_{11} \\ \beta_{31} = b_1 - \alpha_{31}a_{11} - \alpha_{32}a_{21} = b_1 - (1 - \delta)a_{11} - \delta a_{21} \text{ où } \delta = \alpha_{32} \\ \beta_{32} = b_2 - \alpha_{32}a_{22} = (1 - b_1) - \delta a_{22} \end{cases}$$

Pour garantir une stabilité en variation totale pour des pas de temps arbitrairement grands, il faut annuler les β_{ij} . Par hypothèse, on a déjà que $\delta \in [0, 1]$ et $a_{21} = a_{11}$ nous assure que $\beta_{21} = 0$. Pour annuler β_{31} et β_{32} , il faut que:

$$\begin{cases} b_1 = (1 - \delta)a_{11} + \delta a_{21} = a_{21} \\ 1 - b_1 = \delta a_{22} \text{ donc } b_1 = 1 - \delta a_{22} \end{cases}$$

D'où: $a_{21} = b_1 = a_{11}$ et $a_{21} = 1 - \delta a_{22}$. La condition d'ordre impose: $a_{11}b_1 + (1 - b_1)(a_{21} + a_{22}) = \frac{1}{2}$, donc:

$$b_1(a_{11} - (a_{21} + a_{22})) + (a_{21} + a_{22}) = \frac{1}{2}$$

C'est-à-dire:

$$b_1 = \frac{\frac{1}{2} - (a_{21} + a_{22})}{a_{11} - (a_{21} + a_{22})}$$

Comme $a_{11} = a_{21}$, il vient:

$$b_1 = \frac{1}{2a_{22}}(1 - 2(a_{21} + a_{22}))$$

C'est cette dernière condition qui va donner l'expression de a_{22} :

$$1 - b_1 = \delta a_{22} \text{ d'où } 1 = \frac{1}{2a_{22}}(1 - 2(a_{21} + a_{22})) + \delta a_{22}$$

Ce qui donne:

$$2a_{22} = 1 - 2((1 - \delta a_{22}) + a_{22}) + 2\delta a_{22}^2$$

Finalement, a_{22} est racine du polynôme:

$$2\delta x^2 - 2(2 - \delta)x - 1 = 0$$

Son expression est:

$$x = \frac{(2 - \delta) \pm \sqrt{(2 - \delta)^2 + 2\delta}}{2\delta}$$

On choisit la racine négative et la méthode s'écrit:

$$\begin{cases} a_{11} = a_{21} = b_1 = 1 - \delta a_{22} \\ b_2 = 1 - b_1 = \delta a_{22} \\ a_{22} = \frac{(2 - \delta) - \sqrt{(2 - \delta)^2 + 2\delta}}{2\delta} \end{cases} \quad (6.24)$$

Il faut noter que ces méthodes ne peuvent pas avoir le même coefficient diagonal: $a_{11} = a_{22}$ donne $a_{22} = \frac{1}{1 + \delta}$. Ce qui n'est pas compatible avec l'expression donnée dans (6.24). \square

Remarque Ces méthodes ne sont pas A -stables puisque $a_{22} < 0$, mais elles sont telles que $\mathcal{R}(z) \rightarrow 0$ si $\text{Re}(z) \rightarrow -\infty$. Pour $\delta \simeq 0.5$, on a intégré le problème:

$$\begin{cases} \partial_t u + \partial_x u = -5000u \\ u_0(x) = \mathbf{1}_{x \leq 0.1} \end{cases}$$

de façon satisfaisante avec un maillage grossier.

6.2.6 Nouveaux essais numériques

On va maintenant tester la méthode (6.24) sur des problèmes scalaires de difficulté croissante. Le paramètre δ vaudra 0.5 puisque c'est la meilleure valeur d'après les expériences numériques. Le premier cas-test sera une équation d'advection:

$$\begin{cases} \partial_t u + \partial_x u = 0 \\ u_0(x) = \mathbf{1}_{x \leq 0.1} \end{cases}$$

On fixe $\Delta x = 0.001$, $\Delta t = 0.1$ (CFL=100) et on itère jusque $T = 0.2$. On constate sur la figure 6.7 une grande similitude entre la routine (6.24) d'ordre 2 et la méthode de référence (Euler rétrograde).

On décide maintenant d'approximer une équation d'advection dissipative:

$$\begin{cases} \partial_t u + \partial_x u = -50u \\ u_0(x) = \mathbf{1}_{x \leq 0.3} \end{cases}$$

On fixe $\Delta x = 0.01$, $\Delta t = 0.05$ (CFL=5) et on itère jusque $T = 0.3$. On voit que le palier est légèrement mieux approximé par la routine à l'ordre 2 (figure 6.8).

On présente maintenant deux problèmes non-linéaires. Le premier ne comporte qu'un terme de réaction à plusieurs équilibres:

$$\begin{cases} \partial_t u + \partial_x u = 15u(u - 1)(u - 0.2) \\ u_0(x) = \mathbf{1}_{x \leq 0.1} \end{cases}$$

On fixe $\Delta x = 0.01$, $\Delta t = 0.05$ (CFL=5) et on itère jusque $T = 0.3$. Remarquons que la donnée initiale (une discontinuité entre deux équilibres instables) est mal gérée par la méthode d'Euler. Cela corrobore les résultats du chapitre précédent (figure 6.9).

Le second se compose du flux de Buckley-Leverett (que l'on va traiter au moyen d'une linéarisation de Roe [35]) et du même terme de réaction un peu plus raide:

$$\begin{cases} \partial_t u + \partial_x \left(\frac{u^2}{u^2 + \frac{1}{2}(1-u)^2} \right) = 35u(u-1)(u-0.2) \\ u_0(x) = \mathbf{1}_{x \leq 0.1} \end{cases}$$

On fixe $\Delta x = 0.01$, $\Delta t = 0.05$ et on itère jusque $T = 3$ (voir figure 6.10). La CFL est supérieure à 2 sur ce problème puisque la méthode de Gauss oscille même en l'absence de terme source. Notons aussi que la méthode d'Euler se comporte bien mieux sans le second membre ou bien si on perturbe un peu les données initiales. La routine (6.24) continue à se comporter de façon plus convaincante sur ce genre de problèmes non-homogènes.

On termine par un test difficile pour la méthode (6.24) puisqu'il combine la même équation fortement non-linéaire avec un très grand nombre CFL. On fixe $u_0(x) = \mathbf{1}_{x \leq 0.3}$, $\Delta x = 0.0005$, $\Delta t = 0.05$ et on itère jusque $T = 30$. Les résultats sont grossiers du point de vue de la propagation des données, mais la solution reste monotone malgré un maillage très défavorable ... Celle-ci est présentée dans la figure 6.11.

6.3 Conclusion: méthodes de quadrature adaptées

Méthodes de Runge-Kutta

Grâce à cette analyse en variation totale pour des solutions de problèmes scalaires suffisamment généraux (qui complète les conditions d'existence et d'unicité discrètes du chapitre précédent), il ressort que la stabilité linéaire au sens de Dahlquist ne semble pas être un critère adapté à l'intégration en temps de ce type d'équations. On a vu en effet que les méthodes préservant la variation totale au second ordre ne sont ni A -stables, ni même $A(0)$ -stables puisque l'un des coefficients diagonaux est négatif.

Méthodes multipas et différences rétrogrades

Bien que ne faisant pas initialement partie de ce travail, celles-ci y trouvent naturellement leur place puisque pour évaluer les oscillations des méthodes de Runge-Kutta, on les a réécrites sous la forme de méthodes multipas successives. Par conséquent ces dernières seront stables en variation totale si elles ne comportent pas d'étapes explicites, i.e. si elles s'écrivent sous la forme de différences rétrogrades [19]. Bien que réputées peu précises, il en existe une d'ordre 2 et A -stable, elle s'écrit:

$$\frac{3}{2}\mathbf{U}^{n+1} - 2\mathbf{U}^n + \frac{1}{2}\mathbf{U}^{n-1} = \Delta t \mathcal{L}(\mathbf{U}^{n+1}) \quad (6.25)$$

Cette routine A -stable préserve aussi la variation totale pour tout Δt (voir l'annexe). Il faut aussi noter que (6.25) a été aussi remarquée par expérimentation directe sur un code industriel pour ses bonnes propriétés [3].

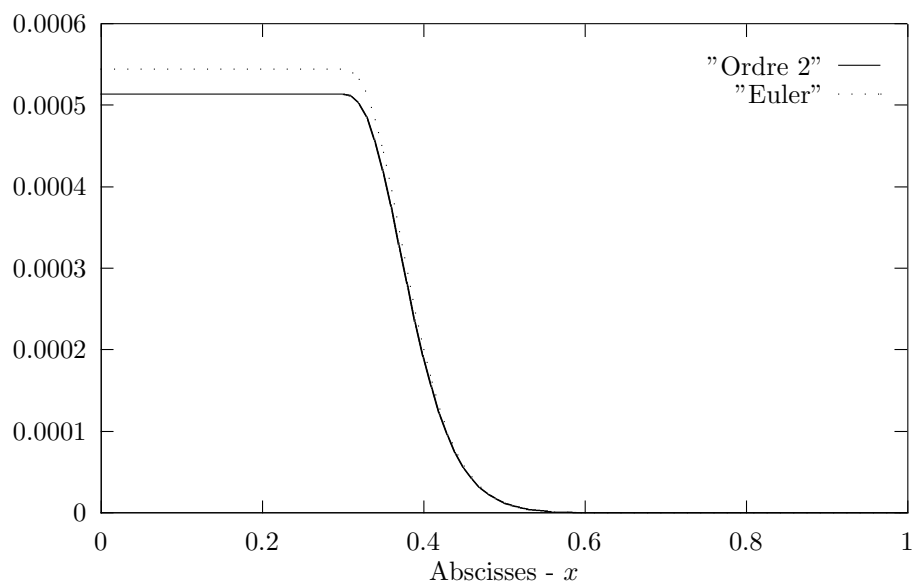


Figure 6.7: Comparaison Euler/ordre 2 - CFL=100

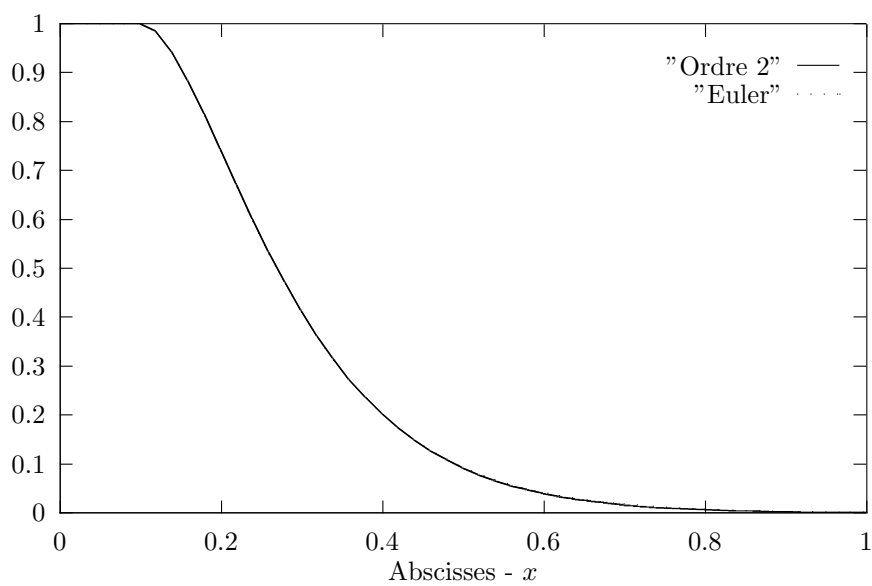


Figure 6.8: Flux d'advection - décroissance exponentielle

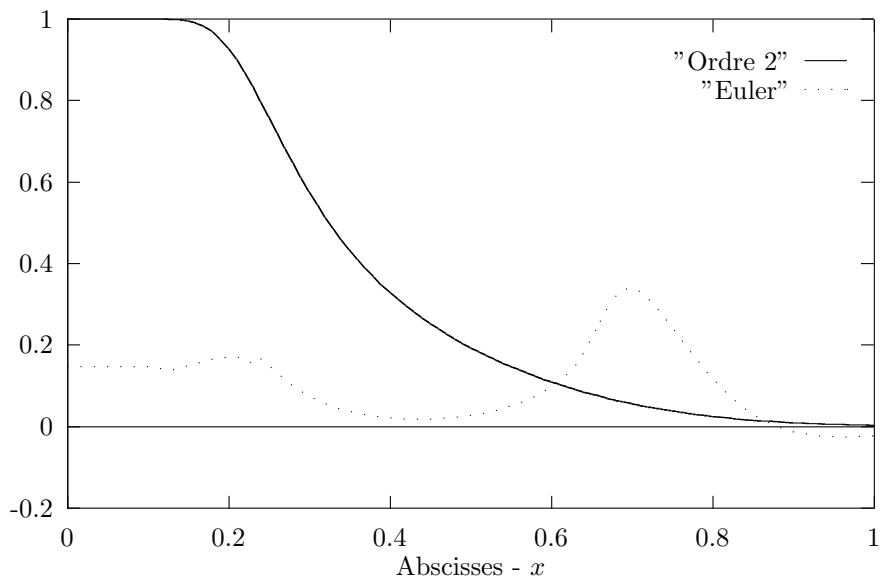


Figure 6.9: Flux d'advection - terme source de rappel

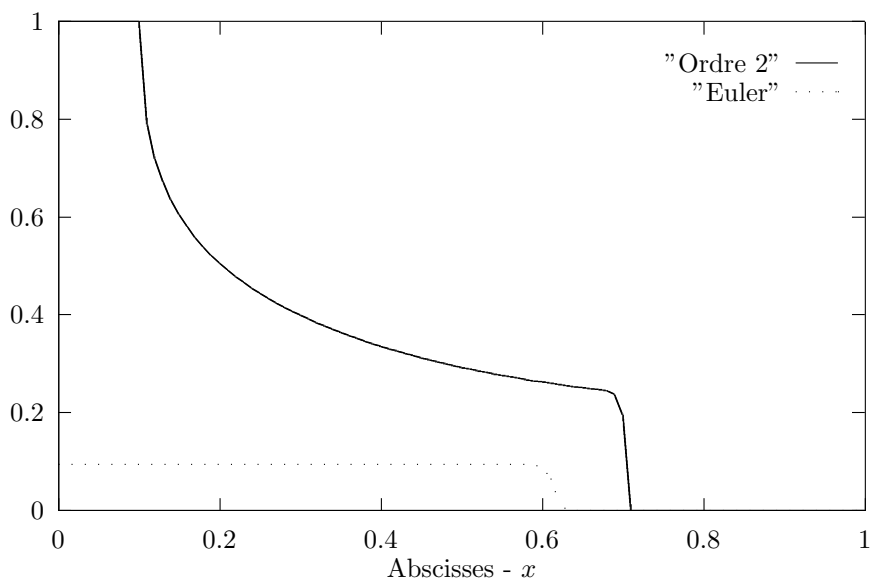


Figure 6.10: Flux de Buckley-Leverett - terme source de rappel

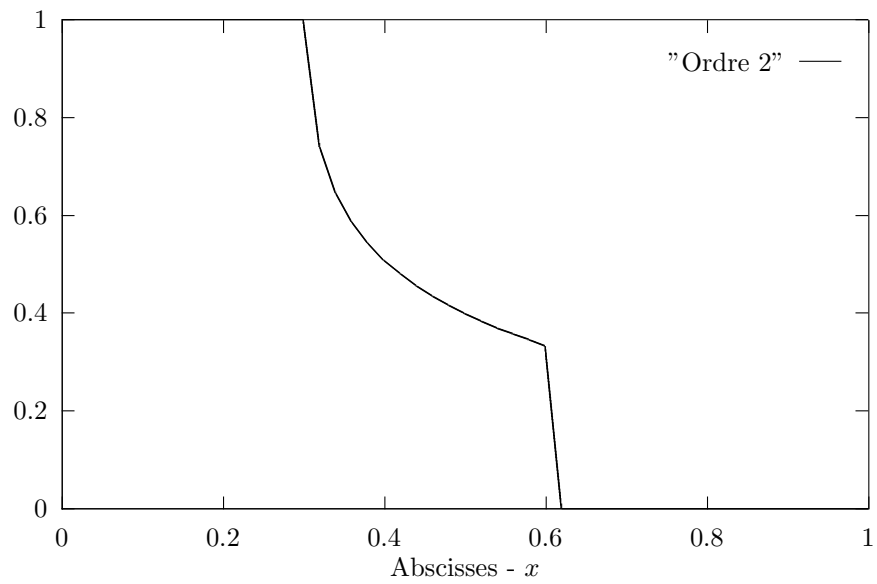


Figure 6.11: Flux de Buckley-Leverett - terme source de rappel - $CFL \geq 100$

Partie III

Introduction d'un solveur hyperbolique vraiment non-homogène

Résumé

Cette partie est consacrée à l'introduction d'une nouvelle classe de schémas numériques pour les systèmes hyperboliques avec termes sources. Les motivations sont très diverses mais dans le prolongement des travaux précédents:

- stabilité: on désire se débarrasser de toute l'influence que le second membre est susceptible d'avoir sur le pas de temps. Ceci peut être particulièrement intéressant pour les problèmes de relaxation.
- simplicité du solveur élémentaire: autant que possible, il est préférable d'éviter la mise en oeuvre d'un solveur de Riemann généralisé [6, 59] puisque l'on a pu constater sa complexité intrinsèque.
- qualité des approximations: il est connu que les schémas du type «pas fractionnaires» peuvent converger vers des solutions faibles incorrectes: oscillations parasites, non-respect des intégrales premières du mouvement, stationnaires irréalistes ... Il paraît donc souhaitable de pouvoir établir avec certitude des propriétés fortes de consistance avec le problème continu.

Pour cela, suivant les idées récentes exposées par J. Greenberg et A.-Y. LeRoux [41], on a décidé de s'orienter vers la classe des schémas «well-balanced» (équilibre). Ceux-ci sont tout d'abord introduits pour une loi scalaire et la convergence est établie indépendamment de la raideur réactionnelle. Cette construction élémentaire permet de déduire immédiatement une extension aux systèmes par le biais de la décomposition de flux de VanLeer [83]. Néanmoins, celle-ci respecte les deux premiers objectifs, mais le troisième peut être pris en défaut du fait d'une gestion hasardeuse d'un produit non-conservatif. Pour pallier à cela, on présente une extension plus rigoureuse fondée sur la théorie proposée par G. DalMaso, P.G. LeFloch et F. Murat. Un solveur général pour lequel le terme source est traité via des relations de saut généralisées s'en déduit. Ceci permet la construction d'un schéma linéarisé de type Roe grâce au formalisme proposé par I. Toumi [82]. Ce type de schéma est alors cohérent avec le troisième objectif et produit des permanents de bonne qualité. On rappelle aussi la technique de P. Cargo et A.-Y. LeRoux, conçue à l'origine pour des termes de gravité, que l'on généralise à une classe de seconds membres linéaires en ρ .

Plusieurs remarques sont possibles:

- on fait dorénavant une hypothèse de régularité BV en espace (définition des produits non-conservatifs oblige !) sur les inconnues. On retrouve ainsi le cadre classique introduit par J. Glimm, S.N. Kruzkov et T.P. Liu.
- ce type de schémas est au confluent de plusieurs travaux datant environ d'une décennie et motivés par l'aéronautique: voir [33, 34, 84]. Mais on en trouve aussi des traces dans des articles plus théoriques [29, 49]. L'interprétation non-conservative restant quant à elle originale.
- la régularisation des produits non-conservatifs s'effectue au moyen d'un profil asymptotique qui consiste le plus souvent en une onde progressive régulière. Celles-ci sont à l'étude depuis peu et seul le cas scalaire est clair. On renvoie à [67] pour les détails.
- par souci de lisibilité, les dérivées partielles seront dorénavant notées par des indices au pied des inconnues ($\partial_{\gamma} u \rightsquigarrow u_{\gamma}$). Quant aux représentations graphiques des solutions, on convient de *toujours* tracer en traits pleins celles issues des schémas «équilibre».

Enfin, il faut préciser que l'implémentation de ce solveur en 2D implicite a été menée à bien par S. Clerc [18] pour le système «shallow water» pour lequel une expression astucieuse des flux numériques est possible.

Chapitre 7

Solveur scalaire «équilibre» et convergence

1

On s'intéresse dans ce chapitre à l'approximation numérique d'une loi de conservation monodimensionnelle que l'on munit d'un terme de réaction non-linéaire pouvant devenir arbitrairement raide. Dans ce cadre, on présente un schéma nouveau, conçu pour respecter les états d'équilibre de l'équation originelle, et pour lequel on établit un résultat de convergence pour certains flux particuliers. Enfin, on indique une extension possible de cet algorithme à des problèmes plus généraux.

7.1 Construction et convergence du schéma

7.1.1 Généralités

Pour étudier les systèmes hyperboliques non-linéaires issus du formalisme des équations de la physique des fluides, il est naturel de se ramener à l'étude du problème de Cauchy pour la loi scalaire suivante:

$$\begin{cases} u_t + f(u)_x = g(u) \text{ avec } g(0) = 0 \text{ et } g(u)(\cdot, t) \text{ borné} \\ (x, t) \in \mathbb{R} \times \mathbb{R}^+ \\ u(x, 0) = u_0(x) \in BV(\mathbb{R}) \end{cases} \quad (7.1)$$

Sous des hypothèses de régularité C^1 pour la fonction de flux, Kruzkov [53] a démontré l'existence et l'unicité de la solution entropique $L^\infty([0, T]; BV(\mathbb{R}))$ de (7.1). Néanmoins, l'approximation numérique en reste encore délicate du fait de la présence d'un second membre non-identiquement nul qui tend à imposer parfois de sévères restrictions sur le maillage temporel. Par exemple, afin d'observer des états à l'équilibre cohérents avec (7.1), il sera nécessaire d'imposer sur le schéma de Godunov [36] implicite linéarisé:

$$u_j^{n+1} = u_j^n - \frac{\frac{\Delta t}{\Delta x} (F^G(u_j^n, u_{j+1}^n) - F^G(u_{j-1}^n, u_j^n)) + \Delta t g(u_j^n)}{1 - \Delta t \cdot g'(u_j^n)}$$

la restriction suivante:

$$\Delta t \leq \min \left\{ \frac{\Delta x}{\max_{(j,n)} |f'(u_j^n)|}, \frac{1}{\max_{(j,n)} |g'(u_j^n)|} \right\}$$

Nous proposons donc ici la construction d'un schéma numérique capable de fournir des équilibres réactionnels stables et réalistes pour (7.1) (schéma équilibre), sous la réserve que les zéros de g soient isolés et ceci sans restriction particulière sur les pas de discrétisation hormis l'habituelle condition de stabilité CFL. Il est commode pour cela d'introduire une équation auxiliaire homogène presque-partout à laquelle on adjoint des conditions aux limites exprimant l'action localisée du terme de réaction. Le problème de Riemann est ainsi résolu et permet la mise en oeuvre d'un schéma de Godunov, dont on donne ensuite des résultats de convergence dans un cas particulier où $f'(u)$ reste strictement positive. Par la suite, on indique une marche à suivre pour généraliser ce solveur de Riemann à une fonction de flux plus générale.

7.1.2 Problème auxiliaire et schéma de base

La solution de (7.1) est essentiellement C^1 par morceaux. Dans les zones où $u(x, t)$ est suffisamment régulière et où $g(u(x, t)) \neq 0$, on écrit:

$$u_t + g(u) \left[\frac{f'(u)}{g(u)} u_x - 1 \right] = 0 \quad (7.2)$$

Soit $\phi(u)$ la fonction définie à une constante près telle que:

$$\phi'(u) = \frac{f'(u)}{g(u)}$$

Posons ensuite $a(x) = x$. Sous cette notation, l'équation (7.2) s'écrit alors:

$$u_t + g(u)[\phi(u) - a(x)]_x = 0 \quad (7.3)$$

Introduisons maintenant un maillage spatial $(x_j)_{j \in \mathbb{Z}}$ tel que $x_j = j\Delta x$ et une nouvelle fonction $a^{\Delta x}(x) \in BV(\mathbb{R})$ définie selon:

$$a^{\Delta x}(x) = x_j \text{ pour } x \in \left[\left(j - \frac{1}{2} \right) \Delta x, \left(j + \frac{1}{2} \right) \Delta x \right] \quad (7.4)$$

Alors, à l'intérieur des intervalles, $a^{\Delta x}(x)_x = 0$ et (7.3) se réduit à:

$$u_t + f(u)_x = 0 \quad (7.5)$$

En revanche, si x_0 est un point de discontinuité de $a^{\Delta x}$, on pose $w = \phi(u) - a^{\Delta x}(x)$, dont la variation se calcule selon:

$$w_t = \phi'(u)u_t = - \underbrace{\phi'(u)g(u)}_{f'(u)} \underbrace{[\phi(u) - a^{\Delta x}(x)]_x}_{w_x}$$

Ce qui donne:

$$w_t + f'(u)w_x = 0$$

Ceci exprime la convection passive de w le long des caractéristiques $\dot{x} = f'(u)$. Si on admet que $f'(u) \neq 0$, ceci peut se réécrire sous la forme:

$$\forall t > 0, \begin{cases} w(x_0, t) = w_{gauche} \text{ si } f'(u) > 0 \\ w(x_0, t) = w_{droite} \text{ si } f'(u) < 0 \end{cases} \quad (7.6)$$

On introduit un maillage temporel $(t_n)_{n \in \mathbb{N}}$ tel que $t_n = n\Delta t$. En vue de construire un schéma de type Godunov, on est conduit à résoudre dans chaque maille $(\Delta x \times \Delta t)$ le problème mixte suivant issu de (7.5):

$$\begin{cases} u_t^{\Delta x} + f(u^{\Delta x})_x = 0 \\ u^{\Delta x}(x, 0) = u_j^n \text{ pour } x \in](j - \frac{1}{2})\Delta x, (j + \frac{1}{2})\Delta x[\\ u^{\Delta x}((j - \frac{1}{2})\Delta x, t) = u_{j-\frac{1}{2}}^n \text{ pour } t \in [n\Delta t, (n+1)\Delta t[\end{cases} \quad (7.7)$$

dont la solution $u^{\Delta x}(x, t)$ s'exprime grâce à l'action d'un semi-groupe de contractions $t \mapsto \mathcal{S}(t)$ et est ensuite moyennée par maille. La valeur de la condition aux limites $u_{j-\frac{1}{2}}^n$ se détermine par la condition (7.6) appliquée au bord de la maille: $[w]_{x=x_{j-\frac{1}{2}}} = 0$. C'est-à-dire, si $f'(u) > 0$:

$$\phi(u_{j-\frac{1}{2}}^n) = \phi(u_{j-1}^n) + \Delta x \quad (7.8)$$

Sous réserve que ϕ soit un difféomorphisme local, le schéma de Godunov construit à partir de ce solveur définit alors une solution discrète $(u_j^n)_{(j,n)}$ pour tout couple $(j\Delta x, n\Delta t)$. Celle-ci s'obtient pour tout t^n par projection des solutions élémentaires sur l'espace des fonctions constantes par mailles. (cf Fig. 7.1)

$$(u_j^{n+1})_{j \in \mathbb{Z}} = P^{\Delta x} \circ \mathcal{S}(\Delta t)(u_j^n)_{j \in \mathbb{Z}}$$

Ce qui s'écrit:

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x} (f(u_j^n) - f(u_{j-\frac{1}{2}}^n)) \quad (7.9)$$

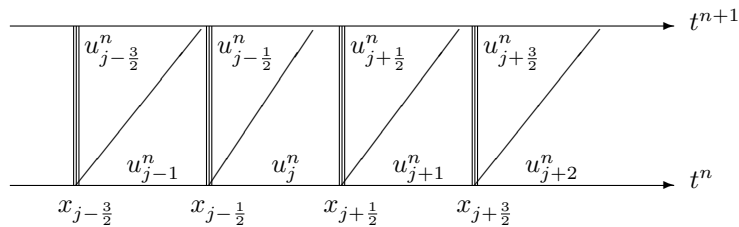


Figure 7.1: Problèmes mixtes de Riemann aux interfaces

7.1.3 Principe du maximum discrèt

L'objectif est d'établir une estimation uniforme en amplitude sur la suite d'approximations $(u_j^n)_{j,n}$ issues du schéma de type Godunov (7.9). Celle-ci est classique pour cette classe de schémas dans le cas homogène correspondant à $g(u) \equiv 0$ dans (7.1). Dans le cas contraire, il est possible de maintenir ce résultat en faisant des hypothèses sur le problème continu.

Lemme III.1 *Supposons que les fonctions f et g vérifient:*

$$\begin{cases} f, g \in C^1(\mathbb{R}) \\ f' \geq \alpha > 0 \\ \exists M \in \mathbb{R}^+ \text{ tel que } x.g(x) \leq 0 \text{ pour } |x| \geq M \end{cases} \quad (7.10)$$

Alors, sous la restriction CFL:

$$\max_{(j,n)} |f'(u_j^n)| \Delta t \leq \Delta x \quad (7.11)$$

on a l'estimation:

$$\sup_{j,n} |u_j^n| \leq \max(\|u_0\|_\infty, M) \quad (7.12)$$

Démonstration

La preuve se fait en deux temps: on montre que les valeurs aux interfaces $u_{j-\frac{1}{2}}^n$ sont bornées, puis on en déduit l'estimation (7.12) par récurrence.

On a supposé que $f' \geq \alpha > 0$, donc on a:

$$\alpha u_x \leq f(u)_x = g(u) \text{ pour } x \in [0, \Delta x]$$

Ceci implique que:

$$u_x \leq \frac{g(u)}{\alpha}$$

Posons $v = u - \max(\|u_0\|_\infty, M)$, alors on a $v_x = u_x$. Il s'agit de montrer que la partie positive $v_+ = \max(0, v)$ est identiquement nulle. Par l'inéquation précédente que l'on multiplie par v_+ , on obtient:

$$\left(\frac{(v_+)^2}{2} \right)_x = v_+ \cdot v_x \leq \frac{g(u) \cdot v_+}{\alpha}$$

On intègre sur $x \in [0, \Delta x]$ et il vient:

$$\left(\frac{(v_+)^2}{2} \right) (\Delta x) - \left(\frac{(v_+)^2}{2} \right) (0) = \int_0^{\Delta x} v_+ \cdot v_x(s) ds \leq \frac{g(u) \cdot v_+}{\alpha} (s) ds$$

L'intégrale de droite est clairement négative puisque $v_+ \geq 0$ n'a lieu que si $u \geq \max(\|u_0\|_\infty, M)$, mais ceci implique que $g(u) \leq 0$. Par conséquent, on obtient:

$$\left(\frac{(v_+)^2}{2} \right) (\Delta x) \leq \left(\frac{(v_+)^2}{2} \right) (0) = 0$$

On montre de la même manière que $w_- \stackrel{\text{def}}{=} (u + \max(\|u_0\|_\infty, M))_-$ est identiquement nulle. Comme la relation de saut (7.8) peut être obtenue en intégrant l'équation différentielle $f(u)_x = g(u)$ entre $x = 0$ et $x = \Delta x$, il vient que les valeurs à l'interface $u_{j-\frac{1}{2}}^n$ sont uniformément bornées.

La seconde étape se fait par récurrence: supposons qu'à l'instant $t^n = n\Delta t$, on ait $\sup_j |u_j^n| \leq \max(\|u_0\|_\infty, M)$, le schéma (7.9) donne:

$$\begin{aligned} |u_j^{n+1}| &= \left| u_j^n - \frac{\Delta t}{\Delta x} (f(u_j^n) - f(u_{j-\frac{1}{2}}^n)) \right| \\ &= \left| u_j^n - \frac{\Delta t}{\Delta x} f'(\xi_j) (u_j^n - u_{j-\frac{1}{2}}^n) \right| \\ &\leq \left(1 - \frac{\Delta t}{\Delta x} f'(\xi_j) \right) |u_j^n| + \frac{\Delta t}{\Delta x} f'(\xi_j) |u_{j-\frac{1}{2}}^n| \\ &\leq \max(\|u_0\|_\infty, M) \end{aligned}$$

On a utilisé successivement le théorème de la moyenne, la condition CFL et la positivité de f' . On a donc le résultat (7.12). \square

7.1.4 Analyse et convergence du schéma

On se restreint dans ce paragraphe au cas où les hypothèses (7.10) sont vérifiées et on renvoie à [5] pour l'analyse théorique du problème de Riemann avec conditions aux limites et du semi-groupe $t \mapsto \mathcal{S}(t)$ qui vont intervenir naturellement ici. Introduisons alors une fonction $u^h : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ telle que:

$$u^h(\cdot, t) = \mathcal{S}(t - n\Delta t) \circ [P^{\Delta x} \circ \mathcal{S}(\Delta t)]^n \circ P^{\Delta x}(u_0) \text{ où } \begin{cases} h = \Delta x \\ n \text{ est la partie entière de } t/\Delta t \end{cases} \quad (7.13)$$

Lemme III.2 *Si $u_0 \in L^1 \cap BV(\mathbb{R})$ est tel que le support de $g(u_0)$ est compact dans \mathbb{R} , et sous la condition CFL (7.11), la suite $(u^h)_h$ issue de (7.8), (7.9), (7.13) est relativement compacte dans $L^1([0, T] \times \mathbb{R})$ pour $T \in \mathbb{R}^+$.*

Démonstration

La définition de la variation totale en espace pour $u^h(x, t)$ s'écrit:

$$TV(u^h(\cdot, t)) = \lim_{\epsilon \rightarrow 0} \int_{\mathbb{R}} \frac{|u^h(x + \epsilon, t) - u^h(x, t)|}{\epsilon} dx$$

Ceci entraîne que pour la fonction définie par (7.13):

$$TV(u^h(\cdot, n\Delta t)) = \sum_{j \in \mathbb{Z}} |u_{j+1}^n - u_j^n|$$

La première estimation cherchée concerne la variation totale en espace à l'instant $(n+1)\Delta t$:

$$u_{j+1}^{n+1} - u_j^{n+1} = u_{j+1}^n - u_j^n - \lambda(f(u_{j+1}^n) - f(u_{j+\frac{1}{2}}^n)) + \lambda(f(u_j^n) - f(u_{j-\frac{1}{2}}^n))$$

où λ désigne comme de coutume le rapport $\frac{\Delta t}{\Delta x}$. Ceci peut se réécrire sous la forme suivante:

$$\begin{aligned} u_{j+1}^{n+1} - u_j^{n+1} &= u_{j+1}^n - u_j^n - \lambda(f(u_{j+1}^n) - f(u_j^n)) \\ &\quad + \lambda(f(u_j^n) - f(u_{j-1}^n)) \\ &\quad + \lambda(f(u_{j+\frac{1}{2}}^n) - f(u_j^n)) \\ &\quad - \lambda(f(u_{j-\frac{1}{2}}^n) - f(u_{j-1}^n)) \end{aligned}$$

On introduit alors la suite ξ_j^n de points définis par la relation:

$$f(u_j^n) - f(u_{j-1}^n) = f'(\xi_j^n)(u_j^n - u_{j-1}^n)$$

On désigne par Q la fonction telle que $f = Q \circ \phi$. Par construction de ϕ , on a $Q' = g \circ \phi^{-1}$. L'expression précédente devient alors:

$$\begin{aligned} u_{j+1}^{n+1} - u_j^{n+1} &= (u_{j+1}^n - u_j^n)(1 - \lambda f'(\xi_{j+1}^n)) \\ &\quad + \lambda f'(\xi_j^n)(u_j^n - u_{j-1}^n) \\ &\quad + \lambda(Q \circ \phi(u_{j+\frac{1}{2}}^n) - Q \circ \phi(u_j^n)) \\ &\quad - \lambda(Q \circ \phi(u_{j-\frac{1}{2}}^n) - Q \circ \phi(u_{j-1}^n)) \end{aligned}$$

En tenant compte de la relation:

$$Q \circ \phi(u_{j+\frac{1}{2}}^n) = Q[\phi(u_j^n) + \Delta x]$$

on obtient par le théorème de la moyenne:

$$\exists \theta_j^n \in [0, 1] \text{ tel que } Q \circ \phi(u_{j+\frac{1}{2}}^n) - Q \circ \phi(u_j^n) = Q_j^n \cdot \Delta x$$

avec:

$$Q_j^n = Q'[\phi(u_j^n) + \theta_j^n \cdot \Delta x] = g \circ \phi^{-1}[\phi(u_j^n) + \theta_j^n \cdot \Delta x]$$

Par la condition CFL (7.11) et le signe de f' , il vient en valeur absolue:

$$\begin{aligned} |u_{j+1}^{n+1} - u_j^{n+1}| &\leq |u_{j+1}^n - u_j^n|(1 - \lambda f'(\xi_{j+1}^n)) \\ &\quad + \lambda f'(\xi_j^n)|u_j^n - u_{j-1}^n| \\ &\quad + \Delta t |Q_j^n - Q_{j-1}^n| \end{aligned}$$

C'est-à-dire que:

$$TV(u^h(\cdot, (n+1)\Delta t)) \leq TV(u^h(\cdot, n\Delta t)) + \Delta t \sum_{j \in \mathbb{Z}} |Q_j^n - Q_{j-1}^n|$$

Comme $g(u_0)$ est à support compact et que la vitesse de propagation de l'information est finie, les Q_j^n sont nuls pour $|j|$ suffisamment grand². Ceci implique que la terme $\sum_{j \in \mathbb{Z}} |Q_j^n - Q_{j-1}^n|$ est borné. Il reste donc à effectuer la sommation de $n = 0$ à $(n+1)\Delta t$:

$$TV(u^h(\cdot, (n+1)\Delta t)) \leq TV(u_0) + \underbrace{\sum_{n=0}^{(n+1)\Delta t} \Delta t \sum_{j \in \mathbb{Z}} |Q_j^n - Q_{j-1}^n|}_{O(n+1)}$$

L'estimation en temps est plus rapide. Avec les mêmes arguments, le théorème de la moyenne fournit une suite de R_j correspondant aux indices j tels que $g(u_{j-1}^n) \neq 0$ permettant d'écrire:

$$u_j^{n+1} - u_j^n = -\lambda(f(u_j^n) - f(u_{j-\frac{1}{2}}^n)) = -\lambda R_j(\phi(u_j^n) - \phi(u_{j-\frac{1}{2}}^n))$$

Compte tenu de (7.8), si $g(u_{j-1}^n) \neq 0$, on obtient:

$$u_j^{n+1} - u_j^n = -\lambda R_j \phi_j(u_j^n - u_{j-1}^n) + \lambda R_j \Delta t$$

De même, si $g(u_{j-1}^n) = 0$, il suffit d'écrire:

$$u_j^{n+1} - u_j^n = -\lambda f'(\xi_j)(u_j^n - u_{j-1}^n)$$

Il existe donc une constante $C = \sup_j (f'(\xi_j), |R_j \phi_j|)$ telle que, si on effectue la somme sur $j \in \mathbb{Z}$

$$\int_{\mathbb{R}} |u^h(x, (n+1)\Delta t) - u^h(x, n\Delta t)| dx = \sum_{j \in \mathbb{Z}} \Delta x |u_j^{n+1} - u_j^n| \leq \Delta t \left[\underbrace{C \sum_{j \in \mathbb{Z}} |u_j^n - u_{j-1}^n|}_{=TV(u^h(\cdot, n\Delta t))} + \underbrace{\sum_{j/g(u_{j-1}^n) \neq 0} \Delta x |R_j|}_{\text{borné}} \right]$$

Il reste à voir que $u^h \in L^1([0, T] \times \mathbb{R})$. Pour cela, on écrit le schéma sous la forme suivante:

$$u_j^{n+1} = u_j^n - \lambda(f(u_j^n) - f(u_{j-1}^n)) + \lambda(f(u_{j-\frac{1}{2}}^n) - f(u_{j-1}^n))$$

En appliquant deux fois le théorème de la moyenne, on obtient:

$$|u_j^{n+1}| \leq |u_j^n| + C\lambda |u_j^n - u_{j-1}^n| + \Delta t |Q_{j-1}^n|$$

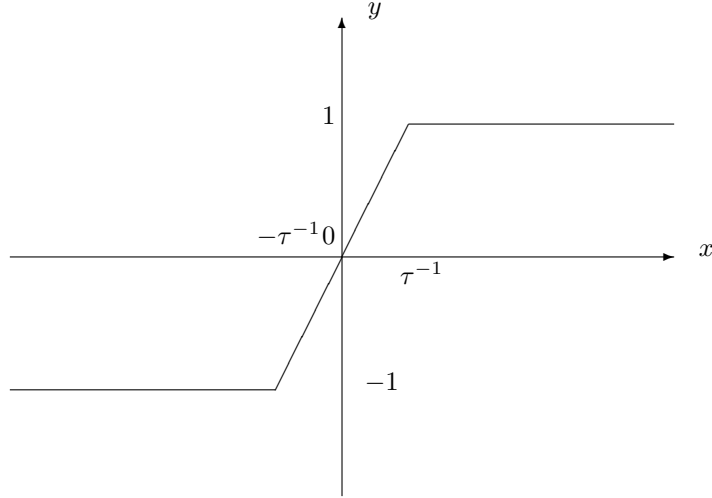
Et donc:

$$\sum_{j \in \mathbb{Z}} \Delta x |u_j^{n+1}| \leq \sum_{j \in \mathbb{Z}} \Delta x |u_j^n| + \Delta t \left[\underbrace{C \sum_{j \in \mathbb{Z}} |u_j^n - u_{j-1}^n|}_{TV(u^h(\cdot, n\Delta t))} + \underbrace{\sum_{j \in \mathbb{Z}} \Delta x |Q_j^n|}_{\text{borné}} \right]$$

Au bout du compte, on trouve que:

$$\|u^h(\cdot, t)\|_{L^1(\mathbb{R})} \leq \|u_0\|_{L^1(\mathbb{R})} + \sum_{n=0}^{t/\Delta t} \Delta t \left[C.TV(u^h(\cdot, n\Delta t)) + \sum_{j \in \mathbb{Z}} \Delta x |Q_j^n| \right]$$

2. On pourrait s'affranchir de cette hypothèse sur u_0 : pour cela, il faut majorer le terme $\sum_{j \in \mathbb{Z}} |Q_j^n - Q_{j-1}^n|$. Ceci peut se faire en montrant que $TV(\theta_j^n) \leq C \text{ste}.TV(u_j^n)$.

Figure 7.2: Régularisation de $u \mapsto \text{sgn}(u)$

Cette estimation entraîne que:

$$\begin{aligned} \|u^h\|_{L^1(\mathbb{R} \times \mathbb{R}^+)} &= \int_0^T \int_{\mathbb{R}} |u^h(x, t)| dx dt \\ &\leq T \sup_{t \in [0, T]} \int_{\mathbb{R}} |u^h(x, t)| dx \\ &\leq T \sup_{t \in [0, T]} \left[\|u_0\|_{L^1(\mathbb{R})} + \sum_{n=0}^{t/\Delta t} \Delta t \left(C.TV(u^h(\cdot, n\Delta t)) + \sum_{j \in \mathbb{Z}} \Delta x |Q_j^n| \right) \right] \end{aligned}$$

On est donc en position d'appliquer le théorème d'Ascoli pour avoir le résultat annoncé. \square

Nous examinons maintenant la nature de sa limite.

Théorème III.1 *Il existe une sous-suite de $(u^h)_h$ convergeant presque-partout vers la solution entropique au sens de Kruzkov de (7.1) lorsque $\Delta x \rightarrow 0$ sous la condition CFL (7.11).*

Démonstration

Pour montrer ce résultat, on commence par approcher la fonction discontinue $u \mapsto \text{sgn}(u - k)$ par S_τ croissante (voir figure 7.2) et définie par:

$$S_\tau(x) = \begin{cases} -1 & \text{pour } x < -\tau^{-1} \\ \tau x & \text{pour } |x| \leq \tau^{-1} \\ 1 & \text{pour } x > \tau^{-1} \end{cases}$$

Un couple entropique convexe (η, q) pour (7.1) est donné par:

$$\begin{cases} \eta(u) = \int^u S_\tau(x - k) dx \\ q(u) = \int^u S_\tau(x - k) f'(x) dx \\ \eta(k) = q(k) = 0 \end{cases}$$

Par l'intégration du schéma, la formulation faible de l'inégalité d'entropie avec conditions aux limites vérifiée dans chaque maille par $u^{\Delta x}$ solution de (7.7) s'écrit pour $0 \leq \psi \in C_0^\infty(\mathbb{R})$: (voir [41])

$$\begin{aligned} &\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \int_{t_n}^{t_{n+1}} (\eta(u^{\Delta x}) \psi_t + q(u^{\Delta x}) \psi_x)(x, t) dx dt + \\ &\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} (\eta(u^{\Delta x}) \psi)(x, t^n) dx - \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} (\eta(u^{\Delta x}) \psi)(x, t^{n+1} - 0) dx \\ &\geq \int_{t_n}^{t_{n+1}} q(u_j^n) \psi(x_{j+\frac{1}{2}}, t) dt - \int_{t_n}^{t_{n+1}} q(u_{j-\frac{1}{2}}^n) \psi(x_{j-\frac{1}{2}}, t) dt \end{aligned}$$

En effectuant la somme sur les indices j et n , on obtient:

$$\begin{aligned} &\sum_{j,n} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \int_{t_n}^{t_{n+1}} (\eta(u^{\Delta x}) \psi_t + q(u^{\Delta x}) \psi_x)(x, t) dx dt + \\ &\sum_{j,n} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} (\eta(u^{\Delta x})(x, t^n) - \eta(u^{\Delta x})(x, t^n - 0)) \psi(x, t^n) dx \\ &\geq \sum_{j,n} \int_{t_n}^{t_{n+1}} (q(u_j^n) - q(u_{j+\frac{1}{2}}^n)) \psi(x_{j+\frac{1}{2}}, t) dt \end{aligned}$$

Soit à déterminer le signe de $\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} (\eta(u^{\Delta x})(x, t^n) - \eta(u^{\Delta x})(x, t^n - 0)) \psi(x, t^n) dx$.

Pour cela, on pose $v^n(\cdot) = u^{\Delta x}(\cdot, t^n - 0)$ et $P^{\Delta x}$ l'opérateur de projection sur l'espace des fonctions constantes par mailles:

$$\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} [\eta(P^{\Delta x} v^n) - \eta(v^n)(x)] \psi(x, t^n) dx = \eta(P^{\Delta x} v^n) \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \psi(x, t^n) dx - \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \eta(v^n)(x) \psi(x, t^n) dx$$

Comme $\psi = P^{\Delta x} \psi + (\psi - P^{\Delta x} \psi)$, il vient:

$$\begin{aligned} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} [\eta(P^{\Delta x} v^n) - \eta(v^n)(x)] \psi(x, t^n) dx &= \eta(P^{\Delta x} v^n) \Delta x (P^{\Delta x} \psi) - [P^{\Delta x} \eta(v^n) \Delta x (P^{\Delta x} \psi)] \\ &\quad + \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \eta(P^{\Delta x} v^n) [(P^{\Delta x} \psi) - \psi(x, t^n)] dx \\ &= \Delta x (P^{\Delta x} \psi) [\eta(P^{\Delta x} v^n) - P^{\Delta x} \eta(v^n)(x)] \\ &\quad + \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \eta(P^{\Delta x} v^n) [(P^{\Delta x} \psi) - \psi(x, t^n)] dx \end{aligned}$$

Du fait de la convexité de η , on a par l'inégalité de Jensen [75]:

$$\eta(P^{\Delta x} v^n) = \eta \left(\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u^{\Delta x}(x, t^n - 0) \frac{dx}{\Delta x} \right) \leq \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \eta(u^{\Delta x}(x, t^n - 0)) \frac{dx}{\Delta x}$$

Donc:

$$\eta(P^{\Delta x} v^n) - P^{\Delta x} \eta(v^n)(x) \leq 0$$

Ce terme étant négatif, il peut être omis dans l'inégalité de départ. D'autre part, $u^{\Delta x}(x, t)$ est bornée dans $L^\infty(\mathbb{R} \times \mathbb{R}^+)$ en tant que solution d'un problème de Riemann élémentaire.

Au total, il reste à faire la somme pour $(j, n) \in \mathbb{Z} \times \mathbb{N}$:

$$\sum_{j,n} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \eta(P^{\Delta x} v^n) [P^{\Delta x} \psi - \psi(x, t^n)] dx \leq \|\eta(u^{\Delta x})\|_{L^\infty} \Delta x \underbrace{\sum_{j,n} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \max_{s \in [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]} |\psi_x(x, t^n)| dx}_{\simeq \|\psi_x\|_{L^1}}$$

On obtient alors l'inégalité suivante:

$$\sum_{j,n} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \int_{t_n}^{t_{n+1}} (\eta(u^{\Delta x}) \psi_t + q(u^{\Delta x}) \psi_x)(x, t) dx dt \geq \sum_{j,n} \int_{t_n}^{t_{n+1}} (q(u_j^n) - q(u_{j+\frac{1}{2}}^n)) \psi(x_{j+\frac{1}{2}}, t) dt - \Delta x \|\eta(u^{\Delta x})\|_{L^\infty} \|\psi_x\|_{L^1}$$

Par le lemme III.3, il existe $\xi \in [a, b]$ tel que

$$\frac{q(b) - q(a)}{\phi(b) - \phi(a)} = \frac{q'(\xi)}{\phi'(\xi)} = S_\tau(\xi - k) g(\xi)$$

Donc on peut définir une suite $(\xi_j^n)_{(j,n)}$ vérifiant:

$$q(u_j^n) - q(u_{j+\frac{1}{2}}^n) = \underbrace{(\phi(u_j^n) - \phi(u_{j+\frac{1}{2}}^n))}_{-\Delta x} S_\tau(\xi_j^n - k) g(\xi_j^n)$$

On introduit maintenant les fonctions suivantes:

$$\begin{cases} \xi^{\Delta x}(x, t) = \xi_j^n \text{ pour } x \in [(j - \frac{1}{2}) \Delta x, (j + \frac{1}{2}) \Delta x], t \in [n \Delta t, (n + 1) \Delta t] \\ \psi^{\Delta x}(x, t) = \psi(x_{j+\frac{1}{2}}, t) \text{ pour } x \in [(j - \frac{1}{2}) \Delta x, (j + \frac{1}{2}) \Delta x] \end{cases}$$

Alors, par construction, on a les estimations:

$$\begin{cases} \|\xi^{\Delta x} - u^{\Delta x}\|_{L^1(\mathbb{R} \times \mathbb{R}^+)} \leq C_1 \Delta x \\ \|\psi^{\Delta x} - \psi\|_{L^\infty(\mathbb{R} \times \mathbb{R}^+)} \leq C_2 \Delta x \end{cases}$$

Posons $r(x) = S_\tau(x - k)g(x)$ une fonction localement Lipschitzienne, on voit que:

$$\begin{aligned} \int_{t_n}^{t_{n+1}} (q(u_j^n) - q(u_{j+\frac{1}{2}}^n))\psi(x_{j+\frac{1}{2}}, t)dt &= - \int_{t_n}^{t_{n+1}} \Delta x \cdot r(\xi_j^n)\psi(x_{j+\frac{1}{2}}, t)dt \\ &= - \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \int_{t_n}^{t_{n+1}} r(\xi^{\Delta x})\psi^{\Delta x}(x, t)dxdt \end{aligned}$$

D'autre part, on a $\sum_{j,n} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \int_{t_n}^{t_{n+1}} (r(\xi^{\Delta x})\psi^{\Delta x}(x, t) - r(u)\psi)(x, t)dxdt \rightarrow 0$ lorsque $(\Delta t, \Delta x) \rightarrow 0$ par le théorème de Lebesgue puisque

$$\begin{cases} r(\xi^{\Delta x})\psi^{\Delta x} \rightarrow r(u)\psi \text{ presque-partout} \\ \|r(\xi^{\Delta x})\psi^{\Delta x}\|_{L^1} \leq C\|g(u)\psi\|_{L^1} \end{cases}$$

On obtient donc la consistance avec les entropies convexes régulières (η, q) .

$$\int_{\mathbb{R} \times \mathbb{R}^+} (\eta(u)\psi_t + q(u)\psi_x + r(u)\psi)(x, t)dxdt + \int_{\mathbb{R}} \eta(u_0)\psi(x, 0)dx \geq 0$$

Il reste à passer à la limite $\tau \rightarrow +\infty$ par le même procédé pour obtenir la formulation entropique faible de Kruzkov [53]:

$$\begin{cases} \eta(u) \rightarrow |u - k| \text{ presque-partout} \\ \|\eta(u)\psi_t\|_{L^1} \leq C \end{cases}$$

$$\int_{\mathbb{R} \times \mathbb{R}^+} (|u - k|\psi_t + \text{sgn}(u - k)(f(u) - f(k))\psi_x + \text{sgn}(u - k)g(u)\psi)(x, t)dxdt + \int_{\mathbb{R}} |u_0 - k|\psi(x, 0)dx \geq 0$$

□

Lemme III.3 Soient ϕ une fonction C^1 strictement monotone $[x, y] \rightarrow \mathbb{R}$ et $q \in C^1(\mathbb{R})$. Alors, pour tout couple $(a, b) \in [x, y]^2$ tel que $a \neq b$, on a:

$$\exists \xi \in [a, b] \text{ tel que } \frac{q(b) - q(a)}{\phi(b) - \phi(a)} = \frac{q'(\xi)}{\phi'(\xi)}$$

Démonstration

Posons $A = \phi(a)$, $B = \phi(b)$ et $\tilde{q} = q \circ \phi^{-1}$. Le théorème de la moyenne assure qu'il existe $C \in [A, B]$ tel que:

$$\tilde{q}(B) - \tilde{q}(A) = \tilde{q}'(C)(B - A)$$

Mais on remarque que: $\tilde{q}' = (q \circ \phi^{-1})' = (\phi^{-1})'(q' \circ \phi^{-1})$

D'autre part: $(\phi^{-1})' = \frac{1}{\phi' \circ \phi^{-1}}$

Finalement:

$$\frac{q \circ \phi^{-1}(\phi(b)) - q \circ \phi^{-1}(\phi(a))}{\phi(b) - \phi(a)} = \frac{q'}{\phi'}(\phi^{-1}(C))$$

Comme ϕ est inversible dans $[A, B]$, on a le résultat escompté. □

Remarque Tel qu'il est défini par (7.8), (7.9), ce schéma n'est pas consistant avec (7.1) au sens habituel. En effet, il suffit, pour s'en convaincre de chercher la solution de (7.7)-(7.8) pour des données gauche et droite identiques. Ceci n'a rien de surprenant puisque si le terme de flux est nul dans l'équation réactive (7.1), la solution varie tout de même selon l'équation différentielle $\dot{u} = g(u)$.

7.1.5 Résultats numériques

On présente dans cette section les tests pratiqués sur le schéma de Godunov issu du solveur de Riemann (7.8),(7.16) pour l'équation de Burgers non-visqueuse munie d'un terme source non-linéaire.

$$\begin{cases} u_t + \left(\frac{u^2}{2}\right)_x = Ku(1-u)(\mu - u) \\ u_0(x) = \begin{cases} u_g & \text{pour } x < 0.3 \\ u_d & \text{pour } x > 0.3 \end{cases} \end{cases} \quad (7.14)$$

Pour $\mu \in]0, 1[$, l'équilibre stable de l'équation est la solution constante $u(x, t) \equiv \mu$. Généralement, les schémas usuels implicites linéarisés cessent de respecter ce comportement lorsque le produit $K\Delta t$ devient trop important.³ C'est donc sur ce critère que nous allons tester le nouveau solveur. On impose juste au pas de temps (supposé uniforme) Δt la condition CFL (7.11). Le premier cas proposé est:

$$\left| \begin{array}{l} K = 3000, \mu = 0.2 \\ \Delta t = 0.04, \Delta x = 0.02 \\ u_0(x) = \begin{cases} u_g = 0.98 \\ u_d = 0.0 \end{cases} \end{array} \right.$$

Le palier initial gauche est situé à proximité d'un équilibre instable du terme source. Il faut donc s'attendre à obtenir une couche limite (initial layer) à l'issue de laquelle la solution annulera le second membre. Le palier gauche sera donc fixé en $\mu = 0.2$; quant à l'épaisseur de la couche de transition, elle sera de l'ordre de $1/3000 \simeq 3.10^{-4}$. L'intérêt du schéma est de s'affranchir de la restriction consistant à résoudre cette échelle temporelle (i.e. imposer $K.\Delta t \simeq 1$) tout en maintenant le caractère explicite de la discrétisation en temps. Hors de la couche-limite, la solution analytique est un choc entropique séparant les états 0.2 et 0. L'état 0 est par ailleurs un équilibre instable du terme de réaction, donc difficile à gérer au moyen d'un schéma implicite linéarisé. La validation consiste donc à observer que par ce schéma explicite, on obtient une onde entropique reliant les paliers corrects imposés par la forme du terme source. On itère jusqu'à $T = 0.4$: le résultat est présenté dans la figure 7.3.

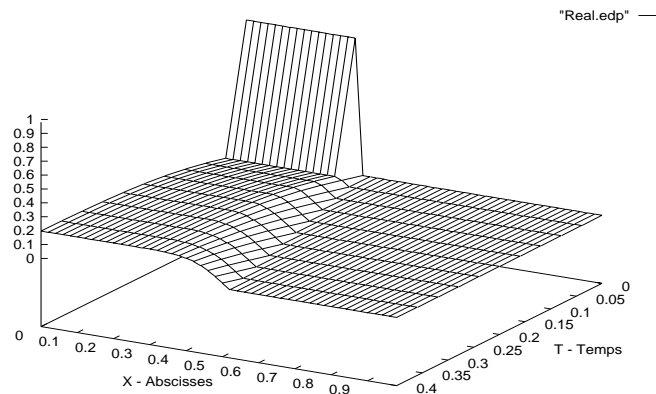


Figure 7.3: Convergence vers l'état stable $\mu = 0.2$

Le second cas-test proposé est:

$$\left| \begin{array}{l} K = 3000, \mu = 0.2 \\ \Delta t = 0.04, \Delta x = 0.02 \\ u_0(x) = \begin{cases} u_g = 0.0 \\ u_d = 0.2 \end{cases} \end{array} \right.$$

L'objectif de ce second test numérique est de montrer la capacité de ce nouveau schéma à préserver les différents états d'équilibre du terme de réaction. Pour cela, on considère un problème de Riemann avec comme donnée initiale deux paliers annulant le terme source. L'un correspond à un équilibre instable, l'autre à un équilibre stable. Le critère d'entropie pour l'équation de Burgers impose une onde de détente pour relier ceux-ci. On s'attend donc à obtenir une solution présentant une couche limite en espace de largeur directement liée à la raideur du terme source. Vu la forme de ce dernier, elle sera en $O(1/3000)$, échelle que l'on ne désire bien sûr pas résoudre puisque celle-ci se répercuterait directement sur le pas de temps via la condition CFL. On introduit donc un maillage grossier, et on observe l'évolution en temps de ce problème. On itère jusqu'à $T = 0.4$. Hors de la couche-limite, la solution analytique est constante et doit conserver les états d'équilibre initiaux 0 et 0.2. Les résultats issus du schéma sont présentés dans la figure 7.4.

3. Voir par exemple la partie II de ce manuscrit.

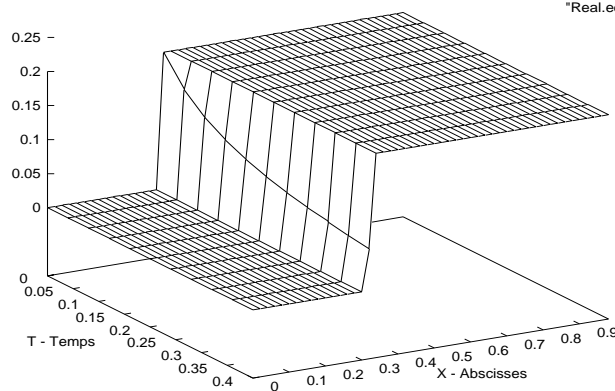


Figure 7.4: Préservation des états d'équilibre stables ou instables

7.2 Solveur de Riemann ;;équilibre; général

7.2.1 Limite d'un problème régularisé

Le schéma précédent peut s'étendre à des problèmes scalaires sans restriction sur le signe de f' qui peut être constant ou non. Il suffit pour cela d'adapter la condition de saut (7.6) qui se manifeste en bordure de chaque maille. Le solveur se construit en étudiant la limite $\epsilon \rightarrow 0$ dans un problème de Riemann régularisé:

$$\begin{cases} u_t + f(u)_x = a'_\epsilon(x)g(u) \\ u_0(x) = \begin{cases} U^G & \text{pour } x < -\epsilon \\ 0 & \text{pour } |x| \leq \epsilon \\ U^D & \text{pour } x > \epsilon \end{cases} \end{cases} \quad a_\epsilon(x) = \begin{cases} 0 & \text{pour } x < -\epsilon \\ \frac{\Delta x(x+\epsilon)}{2\epsilon} & \text{pour } |x| \leq \epsilon \\ \Delta x & \text{pour } x > \epsilon \end{cases} \quad (7.15)$$

Théorème III.2 *Supposons $f'(0) = g(0) = 0$, les traces $U^\pm = \lim_{\epsilon \rightarrow 0} u(\pm\epsilon, t), t > 0$ existent dans les cas suivants et vérifient les relations:*

$$\begin{cases} U^- = U^G; \phi(U^+) = \phi(U^G) + \Delta x \text{ pour } f'(U^G) \geq 0, f'(U^D) \geq 0 \\ \phi(U^-) = \phi(U^D) - \Delta x; U^+ = U^D \text{ pour } f'(U^D) \leq 0, f'(U^D) \leq 0 \\ U^+ = U^- = 0 \text{ pour } f'(U^G) \leq 0, f'(U^D) \geq 0 \end{cases} \quad (7.16)$$

Démonstration

Il s'agit de vérifier les conditions (7.16) dans chacun des cas:

$f'(U^G) \geq 0, f'(U^D) \geq 0$ La solution de (7.15) est telle que $u(-\epsilon, t) \equiv U^G$ pour $t \geq 0$ puisqu'aucune onde ne vient l'affecter. On étudie donc le problème réduit:

$$\begin{cases} u_t + f(u)_x = a'_\epsilon(x)g(u) \\ u_0(x) = \begin{cases} U^G & \text{pour } x < -\epsilon \\ 0 & \text{pour } |x| \geq -\epsilon \end{cases} \end{cases}$$

Pour commencer, un choc va propager vers la droite et atteindre l'abscisse $x = \epsilon$ au bout d'un temps $t = o(\epsilon)$. Par la suite, la trace U^+ est déterminée par la solution régulière $U(X(t), t)$ où $X(t)$ est la ligne caractéristique issue de $X_0 = -\epsilon$ et vérifiant $\dot{X}(t) = f'(U(X(t), t))$. Posons $V(t) = U(X(t), t)$ et soit T l'instant tel que $X(T) = \epsilon$ il vient:

$$2\epsilon = \int_0^T \dot{X}(s)ds = \int_0^T f'(V(s))ds = \int_{U^G}^{U^+} \frac{f'(V)}{g(V)a'_\epsilon} dV$$

puisque le long de $X(t)$, $V(t)$ varie selon $\dot{V} = g(V)a'_\epsilon$. Comme $a'_\epsilon = \frac{\Delta x}{2\epsilon}$, on obtient finalement, lorsque $\epsilon \rightarrow 0$:

$$\phi(U^+) - \phi(U^G) = \Delta x \text{ et } U^- = U^G$$

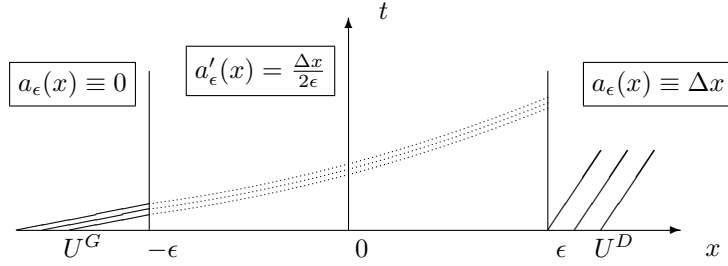


Figure 7.5: Problème de Riemann régularisé

$f'(U^G) \leq 0, f'(U^D) \leq 0$ Ce cas de figure est identique au précédent à ceci près que, cette fois, l'état droit est invariant $u(\epsilon, t) = U^D$, et c'est l'état gauche qui est déterminé par $U(X(t), t)$, $X(t)$ étant la ligne caractéristique issue de $X_0 = \epsilon$. Avec les mêmes notations, il vient:

$$-2\epsilon = \int_0^T f'(V(s))ds = \int_{U^D}^{U^-} \frac{f'(V)}{g(V)a'_\epsilon} dV$$

Là encore, les ϵ se simplifient et on obtient à la limite:

$$\phi(U^-) - \phi(U^D) = -\Delta x \text{ et } U^+ = U^D$$

$f'(U^G) \leq 0, f'(U^D) \geq 0$ C'est le cas de figure le plus simple, puisque la solution de (7.15) se compose de deux détente soniques restant constamment à l'extérieur de l'intervalle $[-\epsilon, \epsilon]$. Comme on a supposé que $f'(0) = 0$, on obtient donc des traces vérifiant pour tout $u(\pm\epsilon, t) = 0 \ t \geq 0$:

$$U^+ = U^- = 0$$

□

Il reste des cas non couverts par ce résultat, notamment ceux pour lesquels $f'(U^G) \geq 0, f'(U^D) \leq 0$. Ceux-ci doivent le plus souvent être étudiés pour chaque problème particulier. Dans cette éventualité, la solution de (7.15) est composée de deux chocs incidents, il s'agit donc de déterminer l'onde la plus forte et d'appliquer l'une des deux premières relations de (7.16).

Pour terminer, il reste un cas particulier pour lequel cette conclusion n'est pas valable, à savoir la présence d'un choc stationnaire dans l'intervalle $[-\epsilon, \epsilon]$ pour la solution de (7.15). Dans cette éventualité, on peut avancer un résultat en faisant des hypothèses supplémentaires.

Lemme III.4 *Supposons $f(u) = u^p$ avec $p \in 2\mathbb{N}$ et g une fonction impaire s'annulant en zéro. Alors les traces $U^\pm = \lim_{\epsilon \rightarrow 0} u(\pm\epsilon, t), t > 0$ existent dans le cas suivant et vérifient:*

$$\begin{cases} \phi(U^-) = \phi(U^G) + \frac{\Delta x}{2} \\ \phi(U^+) = \phi(U^D) - \frac{\Delta x}{2} \end{cases} \text{ pour } 0 \leq U^G = -U^D \leq 0 \quad (7.17)$$

Démonstration

Du fait de la forme de la fonction de flux, il est clair que les données (U^G, U^D) correspondent à un choc stationnaire pour le problème de Riemann et l'EDP homogène $u_t + f(u)_x = 0$. Dans le cas de l'équation (7.15), commençons par poser $t \mapsto X_i(t)$ ($i = 1, 2$) les caractéristiques vérifiant:

$$\begin{cases} \dot{X}_1(t) = f'(u(X_1(t), t)) \text{ et } X_1(0) = -\epsilon \\ \dot{X}_2(t) = f'(u(X_2(t), t)) \text{ et } X_2(0) = \epsilon \end{cases}$$

Le long de ces courbes, on introduit les quantités $V_i(t) = u(X_i(t), t)$ ($i = 1, 2$). Celles-ci vérifient:

$$\begin{cases} V_1(t) = U^G + \int_0^t g(V_1)(s)ds \\ V_2(t) = U^D + \int_0^t g(V_2)(s)ds \end{cases}$$

Comme $U^D = -U^G$ et que la fonction g est impaire, on a $V_1(t) = -V_2(t)$. D'après la parité de la fonction de flux, ceci implique $\dot{X}_1(t) = -\dot{X}_2(t)$ et donc $X_1(t) = -X_2(t)$. Par conséquent, après un temps $t = o(\epsilon)$, la solution de (7.15) est un choc stationnaire en $x = 0$. Les paliers (U^- , U^+) sont déterminés par les deux intégrales suivantes:

$$\begin{cases} \epsilon = \int_0^T f'(V_1(s)) ds = \int_{U^G}^{U^-} \frac{f'(V_1)}{g(V_1)a_\epsilon'} dV_1 \\ -\epsilon = \int_0^T f'(V_2(s)) ds = \int_{U^D}^{U^+} \frac{f'(V_2)}{g(V_2)a_\epsilon'} dV_2 \end{cases}$$

Comme précédemment, le changement de variables implique les relations de saut:

$$\begin{cases} \phi(U^-) = \phi(U^G) + \frac{\Delta x}{2} \\ \phi(U^+) = \phi(U^D) - \frac{\Delta x}{2} \end{cases}$$

□

7.2.2 Ecriture du schéma «équilibre» général

Le schéma «équilibre» général est un schéma de type Godunov pour le problème (7.1) construit sur la base du théorème III.2. Par conséquent, si on note $(u_j^n)_{j \in \mathbb{Z}}$ la fonction en escalier qui approche la solution de (7.1) au temps $t^n = n\Delta t$, on va calculer la suite $(u_j^{n+1})_{j \in \mathbb{Z}}$ en calculant à chaque interface $x_{j+\frac{1}{2}}$ du maillage spatial les traces $U_{j+\frac{1}{2}}^\pm$. Celles-ci correspondent aux valeurs en $x = [(j + \frac{1}{2})\Delta x]^\pm$ de la solution du problème de Riemann pour:

$$\begin{cases} u_t + f(u)_x = g(u) \cdot \Delta x \cdot \delta_{(j+\frac{1}{2})\Delta x} \\ u_0(x) = \begin{cases} u_j^n & \text{pour } x < (j + \frac{1}{2})\Delta x \\ u_{j+1}^n & \text{pour } x > (j + \frac{1}{2})\Delta x \end{cases} \end{cases}$$

Et donc, après l'étape de moyennisation par maille, on obtient le schéma pour (7.1):

$$\begin{cases} u_j^{n+1} = u_j^n - \lambda(f(u_{j+\frac{1}{2}}^-) - f(u_{j-\frac{1}{2}}^+)) \\ u_{j\pm\frac{1}{2}}^\pm \text{ déterminés à partir des relations (7.16)} \end{cases} \quad (7.18)$$

Cette expression généralise (7.9) au cas où f' n'est pas de signe constant. Malheureusement, toutes les éventualités possibles ne sont pas couvertes par les relations (7.16). Il est donc nécessaire de construire un solveur adapté à chaque cas si la solution cherchée est susceptible de comporter des zones fortement compressives, i.e. si f' peut changer de signe d'une maille à l'autre. Pour ces situations problématiques, une étude détaillée des problèmes de Riemann s'impose: on peut consulter [48, 49] comme points de départ dans cette direction.

7.2.3 Perte de bijectivité: l'asymptote horizontale

Jusqu'ici, nous ne nous sommes pas préoccupés de la validité des relations de saut du type $[\phi(u)] = \pm\Delta x$. En effet, ceci suppose que la fonction ϕ soit surjective pour définir à coup sûr le solveur de Riemann via une relation de saut du type (7.8). D'autre part, une condition suffisante d'injectivité est d'exiger que $u \mapsto \phi'(u)$ soit de signe constant à l'intérieur de ses intervalles de définition, la dérivée $u \mapsto f'(u)$ s'annulant éventuellement sur un ensemble de points disjoint de celui formé des zéros de $u \mapsto g(u)$. Néanmoins, il existe des cas où cette condition n'est pas réalisée: voyons ceci sur un exemple.

Soit le problème:

$$\begin{cases} u_t + \left(\frac{u^2}{2}\right)_x = -Ku(1-u)(\mu-u) \\ K \in \mathbb{R}_*^+ \text{ et } \mu \in]0, 1[\\ u_0(x) \in BV(\mathbb{R}) \end{cases} \quad (7.19)$$

Lemme III.5 L'équation (7.19) admet une fonction de saut donnée par:

$$\begin{aligned} \phi :]-\infty, \mu[\cup]\mu, 1[\cup]1, +\infty[&\rightarrow \mathbb{R} \\ u \mapsto \frac{1}{K(1-\mu)} \left[\ln \left| \frac{u-\mu}{1-u} \right| - \ln \left| \frac{u_0-\mu}{1-u_0} \right| \right] \end{aligned}$$

Ici, u_0 est une constante appartenant au même intervalle de définition que u .

Démonstration

Ceci s'obtient via une décomposition en éléments simples:

$$\frac{f'(u)}{g(u)} = \frac{u}{-Ku(1-u)(\mu-u)} = \frac{1}{-K(1-\mu)} \left[\frac{1}{\mu-u} - \frac{1}{1-u} \right]$$

On intègre de u_0 à u :

$$\phi(u) = \frac{1}{-K(1-\mu)} \int_{u_0}^u \frac{du}{\mu-u} - \frac{du}{1-u}$$

d'où il vient:

$$\phi(u) = \frac{1}{K(1-\mu)} \left[\ln \left| \frac{u-\mu}{1-u} \right| \right]_{u_0}^u$$

D'où le resultat. \square

Graphes correspondant au flux de Burgers muni de $g(u)=-u(1-u)(0.5-u)$

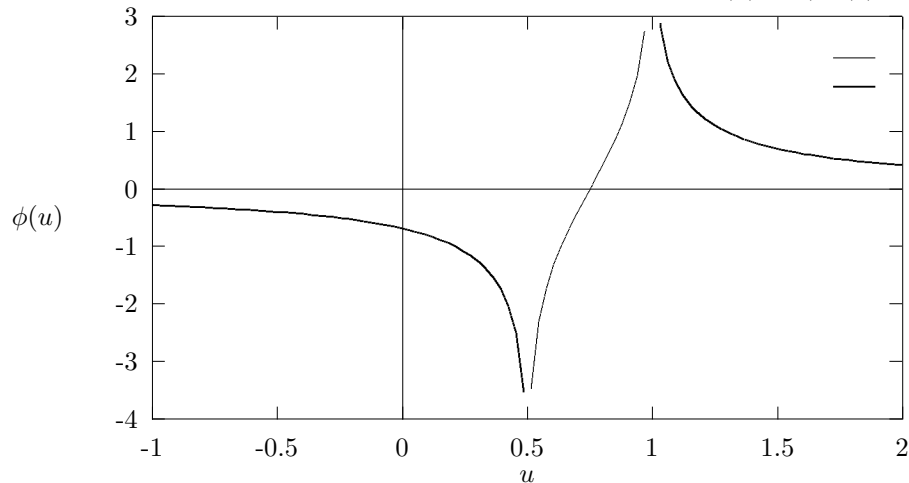


Figure 7.6: Fonction de saut du problème ($\mu = 0.5$)

Une remarque s'impose alors d'elle-même: cette fonction admet la droite $y \equiv 0$ pour asymptote horizontale. Ceci signifie qu'il sera impossible de réaliser le saut (7.8) pour des u suffisamment proches de 0. L'origine de ce problème réside dans le fait que $f'(u)$ et $g(u)$ admettent 0 comme zéro commun. Ce dernier disparaît donc dans la fraction rationnelle qui s'intègre pour donner $\phi(u)$. Ici, le problème est encore aggravé par le fait que $u \equiv 0$ est un équilibre stable pour (7.19). Dans les cas usuels, les zéros du terme source deviennent des pôles pour $\phi(u)$; ils constituent alors des asymptotes verticales (d'où la bijectivité sur \mathbb{R}) et empêchent la solution de «sauter» d'un intervalle de définition à l'autre. Ceci assure entre autres choses des propriétés telles que la positivité ou un principe du maximum pour ce type de schéma.

Il est donc clair qu'il va être ici nécessaire de forcer l'apparition d'une barrière de ce type au voisinage de $u = 0$. Par exemple, on peut imposer que $|u_j^n| \leq \varepsilon \Rightarrow u_{j+\frac{1}{2}}^n = 0$. Ce qui est cohérent avec le problème de Riemann: aux abords de $u = 0$, il n'y a plus assez de convection pour advecter la donnée à travers les interfaces du maillage (cf Fig. 7.6).

Chapitre 8

Première extension aux systèmes réactifs

L'objet de ce chapitre est de présenter une extension des idées précédentes pour des problèmes plus généraux se présentant sous la forme de *systèmes* hyperboliques munis de termes de réaction:

$$\begin{cases} U_t + F(U)_x = G(U) \text{ avec } U \in \mathbb{R}^N \\ U_0 \in [BV(\mathbb{R})]^N \end{cases} \quad (8.1)$$

Le principal obstacle à affronter est la complexité accrue des problèmes de Riemann élémentaires résolus par le schéma de Godunov. En effet, dans le cas (8.1), ces derniers sont constitués de $(N+1)$ états non nécessairement constants séparés par N ondes simples.

Il paraît donc difficile d'envisager l'action d'un terme source au moyen du saut d'une éventuelle fonction à travers les interfaces puisqu'il faudrait agir à partir des états *intermédiaires*. C'est-à-dire ceux utilisés dans le schéma de Godunov traditionnel pour (8.1). Il est alors clair qu'un tel algorithme ne peut être que coûteux et pénible à implanter. L'idée maîtresse est donc de *simplifier* au possible les problèmes de Riemann: ceci se fera par la décomposition du flux $F(U)$ originellement introduite par VanLeer [83]. Par ce biais, il est à nouveau possible de faire *sauter* les données à travers les interfaces, et un schéma numérique devient envisageable. Cette construction est appliquée à un système d'écoulements diphasiques en présence de termes sources modélisant de la recondensation et du frottement pariétal.

8.1 Schémas à décomposition de flux

8.1.1 Construction traditionnelle

Tout d'abord, il s'agit de rappeler, en suivant [36, 83] la mise en place de ce type de procédés d'approximation du problème (8.1) dans lequel on suppose, dans un premier temps $G \equiv 0$.

Formulation intégrale

Il est usuel de considérer dans un premier temps une approximation de $U(x, t)$, solution de (8.1) constante par morceaux. On définit alors

$$U(x, n\Delta t) = U_j^n \text{ pour } x \in \left[\left(j - \frac{1}{2} \right) \Delta x, \left(j + \frac{1}{2} \right) \Delta x \right]$$

Les notations classiques sont: $t^n = n\Delta t$, $x_j = j\Delta x$, $x_{j+\frac{1}{2}} = \left(j + \frac{1}{2} \right) \Delta x$.

En intégrant le système sur une maille $\left[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}} \right] \times [t^n, t^{n+1}]$, on obtient la formulation:

$$\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} U(s, t^{n+1}) - U(s, t^n) ds = \int_{t^n}^{t^{n+1}} F(U)(x_{j-\frac{1}{2}}, t) - F(U)(x_{j+\frac{1}{2}}, t) dt \quad (8.2)$$

Il paraît donc logique de fixer $U_j^n = \frac{1}{\Delta x} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} U(s, t^n) ds$, et de définir des *flux numériques*, i.e. des approximations du membre de droite de (8.2). Globalement, ces méthodes, dites conservatives, s'écrivent donc sous la forme:

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} + \frac{\mathcal{F}_{j+\frac{1}{2}} - \mathcal{F}_{j-\frac{1}{2}}}{\Delta x} = 0$$

Des résultats de convergence devenus classiques sont disponibles pour ce type d'approximations; c'est l'objet du théorème de Lax-Wendroff [58].

8.1.2 Fonctions de flux particulières

Un moyen commode de construire un schéma cohérent avec l'analyse théorique de (8.1) est d'interpréter (8.2) en termes de *problèmes de Riemann élémentaires* situés à chaque interface du maillage dual $(x_{j+\frac{1}{2}})_j$ (cf Fig. 8.1). Malheureusement, ceux-ci ont parfois des solutions complexes décourageant la mise en oeuvre de cette approche. La principale difficulté résidant dans le besoin de trouver l'état stationnaire situé à l'interface pour connaître les valeurs $(\mathcal{F}_{j+\frac{1}{2}})_j$.

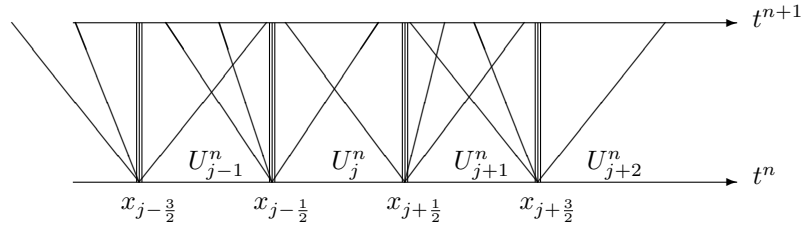


Figure 8.1: Problèmes de Riemann élémentaires

Il est tout de même des situations où les choses peuvent être sensiblement simplifiées. Par exemple, supposons-nous en régime supersonique: toutes les ondes propagent dans le même sens, et la valeur de $\mathcal{F}_{j+\frac{1}{2}}$ est simplement donnée par $F(U_{j/j+1}^n)$ suivant le sens de l'écoulement. Cette idée est la base de l'algorithme de décomposition de flux: *résoudre seulement des problèmes de Riemann pour lesquels la donnée ne circule que dans une seule direction*.

Bien sûr, il va falloir renoncer à travailler directement avec la fonction de flux réel $U \mapsto F(U)$. On introduit donc, suivant [83] les fonctions $F^\pm(U)$ telles que:

$$\forall U \in \Omega, \begin{cases} F(U) = F^+(U) + F^-(U) \\ \Lambda(dF^+(U)) \subset \mathbb{R}^+ \\ \Lambda(dF^-(U)) \subset \mathbb{R}^- \end{cases} \quad (8.3)$$

Ω désigne l'espace des états, un convexe inclus dans \mathbb{R}^N , et $\Lambda(\cdot)$, l'ensemble des valeurs propres (le spectre) d'une matrice. On définit alors la fonction de flux par l'expression:

$$\mathcal{F}(U_L, U_R) = F^+(U_L) + F^-(U_R) \quad (8.4)$$

Une remarque évidente est que ce flux est consistant par construction avec le flux réel puisque:

$$\mathcal{F}(U, U) = F^+(U) + F^-(U) = F(U)$$

Le schéma d'approximation (8.2) peut donc s'écrire grâce à ces fonctions sous la forme:

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} + \frac{F^+(U_j^n) - F^+(U_{j-1}^n) + F^-(U_{j+1}^n) - F^-(U_j^n)}{\Delta x} = 0 \quad (8.5)$$

Il est donc visible que cette expression peut s'interpréter géométriquement comme le montre la figure 8.2.

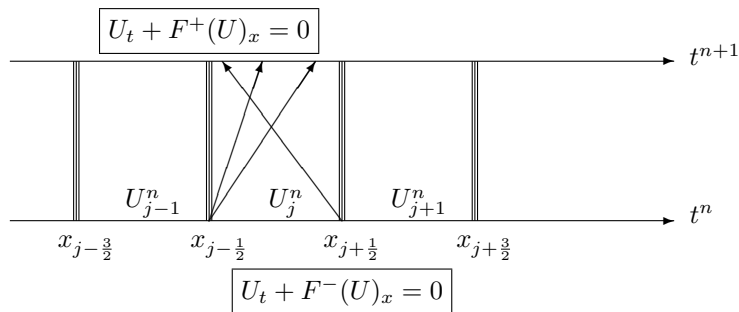


Figure 8.2: Schéma à décomposition de flux

8.1.3 Exemple: le système d'Euler

Dans le cas des équations décrivant la dynamique d'un gaz parfait polytropique en une dimension d'espace, la vitesse du son $a = \sqrt{\frac{\partial P}{\partial \rho}}$ joue un rôle de première importance. En effet, pour ce système, on a:

$$U = \begin{pmatrix} \rho \\ \rho u \\ \rho e \end{pmatrix} \text{ et } F(U) = \begin{pmatrix} \rho u \\ \rho u^2 + P \\ \rho u h \end{pmatrix}$$

où $\rho, u, e, h = e + \frac{P}{\rho}$ désignent la densité, la vitesse, l'énergie totale et l'enthalpie totale du gaz en chaque point $(x, t) \in \mathbb{R} \times \mathbb{R}^+$, alors que $P = (\gamma - 1) \left(\rho e - \frac{\rho u^2}{2} \right)$ est la loi de pression fermant le système.

La décomposition du flux pour ce problème se fait en cherchant une expression intéressante de la quantité de mouvement aux interfaces. On exprime ensuite les deux composantes restantes en fonction de la vitesse u (elle-même définie à partir de a via le nombre de Mach).

Proposition III.1 (*VanLeer*)

Soit $a = \sqrt{\frac{\gamma P}{\rho}}$ la vitesse du son, $M^+ = \frac{\rho}{4a}(u + a)^2$, $M^- = \frac{-\rho}{4a}(u - a)^2$ des quantités de mouvement auxiliaires vérifiant $M^+ + M^- = \rho u$, alors les flux F^\pm sont définis selon:

$$F^+(U) = \begin{cases} F(U) & \text{si } u \geq a \\ \begin{pmatrix} M^+ \\ M^+ \frac{(\gamma-1)u+2a}{\gamma} \\ M^+ \frac{((\gamma-1)u+2a)^2}{2(\gamma^2-1)} \end{pmatrix} & \text{si } |u| < a \\ \vec{0} & \text{si } u \leq -a \end{cases} \quad (8.6)$$

$$F^-(U) = \begin{cases} \vec{0} & \text{si } u \geq a \\ \begin{pmatrix} M^- \\ M^- \frac{(\gamma-1)u-2a}{\gamma} \\ M^- \frac{((\gamma-1)u-2a)^2}{2(\gamma^2-1)} \end{pmatrix} & \text{si } |u| < a \\ F(U) & \text{si } u \leq -a \end{cases}$$

On renvoie à [83] pour les propriétés des jacobiennes des flux $U \mapsto F^\pm(U)$.

8.1.4 Traitement des termes sources

Le vecteur des termes de réaction se traite conventionnellement de deux manières. Soit centré, auquel cas il est discrétisé sous la forme $G_j^n = G(U_j^n)$ et rajouté tel quel dans le membre de droite de l'expression (8.5). Soit décentré, c'est-à-dire que l'on exprime ses composantes en fonction des quantités définies aux interfaces intervenant dans $F^\pm(U)$ (telles que $M^\pm, u \dots$).

Dans tous les cas, son traitement est indépendant de la résolution des problèmes de Riemann, et son éventuelle action n'intervient pas dans la définition des flux numériques. Ceci n'est pas très conforme à l'analyse théorique des systèmes (8.1). En effet, si on retient l'idée consistant à chercher une solution approchée par la résolution successive de problèmes de Riemann élémentaires, alors il faudrait prendre en compte les termes d'ordre zéro à ce niveau. Ceci est néanmoins très difficile car de tels problèmes élémentaires non-homogènes ont une structure inhabituelle (voir [59] ou la partie I). La propriété d'autosimilarité habituelle est en particulier perdue.

8.2 Décomposition de flux vraiment non-homogène

L'objectif est maintenant d'apporter un palliatif à la construction précédente en s'inspirant des résultats annoncés dans le chapitre précédent. Cette construction originale est présentée par le biais d'un problème issu de la modélisation mathématique des écoulements à deux phases.

8.2.1 Formulation du problème étudié

Le système proposé s'écrit sous la forme suivante:

$$\begin{aligned} \rho_t + (\rho u)_x &= 0 \\ (\rho c)_t + (\rho u c)_x &= k_1(x) \rho g(c) \\ (\rho u)_t + (\rho u^2 + P)_x &= k_2(x) \rho u h(u) \end{aligned} \quad (8.7)$$

où ρ, c, u désignent respectivement la densité globale du mélange considéré, la concentration d'une des phases mises en jeu dans l'écoulement, et la vitesse volumétrique commune aux deux espèces. La loi de fermeture est la pression $P(\rho, c)$ supposée convexe par rapport à ρ (pas de transition de phases), et monotone par rapport à

c. Ces hypothèses garantissent que le problème de Riemann pour (8.7) est bien posé pour des discontinuités initiales quelconques dans $\Omega = \{(\rho, c, u) \in \mathbb{R}_*^+ \times [0, 1] \times \mathbb{R}\}$. On renvoie à [81] pour une étude détaillée de (8.7) avec toutefois des notations différentes et sans termes sources.

Quelques hypothèses sont nécessaires: on suppose $c \mapsto g(c)$ et $u \mapsto h(u)$ régulières, i.e. pour le moins Lipschitz-continues, et les fonctions $x \mapsto k_{1/2}(x)$ positives (on pourra au besoin faire des changements de signe dans g ou h). On suppose en outre qu'il existe γ tel que la vitesse du son $a = \sqrt{\frac{\partial P(\rho, c)}{\partial \rho}}$ soit égale à $a = \sqrt{\frac{\gamma P(\rho, c)}{\rho}}$. C'est par exemple le cas pour la loi de Dalton [55].

Les notations employées seront $U = (\rho, \rho c, \rho u)$ et $F(U) = (\rho u, \rho u c, \rho u^2 + P)$. La décomposition traditionnelle des flux du problème posé est donnée par:

Proposition III.2 *Soit $a = \sqrt{\frac{\gamma P(\rho, c)}{\rho}}$ la vitesse du son dans le mélange, $M^+ = \frac{\rho}{4a}(u + a)^2$, $M^- = \frac{\rho}{4a}(u - a)^2$ des quantités de mouvement auxiliaires vérifiant $M^+ + M^- = \rho u$ alors les flux F^\pm s'écrivent:*

$$F^+(U) = \begin{cases} F(U) & \text{si } u \geq a \\ \begin{pmatrix} M^+ \\ M^+ c \\ M^+ \frac{(\gamma-1)u+2a}{\gamma} \end{pmatrix} & \text{si } |u| < a \\ \vec{0} & \text{si } u \leq -a \end{cases} \quad (8.8)$$

$$F^-(U) = \begin{cases} \vec{0} & \text{si } u \geq a \\ \begin{pmatrix} M^- \\ M^- c \\ M^- \frac{(\gamma-1)u-2a}{\gamma} \end{pmatrix} & \text{si } |u| < a \\ F(U) & \text{si } u \leq -a \end{cases}$$

Démonstration

Dans le cas supersonique, $|u| \geq a$, les flux se résument au flux réel et l'analyse habituelle nous assure du signe des valeurs propres de $dF^\pm(U)$.

En régime subsonique, $|u| < a$, il faut vérifier que $\Lambda(dF^\pm(U)) \subset \mathbb{R}^\pm$. Notons $F^\pm(U)_i$ les composantes des vecteurs de flux décomposés et considérons le vecteur d'inconnues $\tilde{U} = (\rho, c(\rho, \rho c), \rho u)$. Les règles de dérivation sont:

$$\begin{cases} \frac{\partial}{\partial \rho} [M^+(\rho, c(\rho, \rho c))c] = c \left[M_\rho^+ + M_c^+ \frac{\partial c}{\partial \rho} \right] + M^+ \frac{\partial c}{\partial \rho} \\ \frac{\partial}{\partial \rho c} [M^+(\rho, c(\rho, \rho c))c] = c M_c^+ \frac{\partial c}{\partial \rho c} + M^+ \frac{\partial c}{\partial \rho c} \end{cases}$$

Les valeurs propres de dF^+ vont se trouver en calculant:

$$|dF^+(U) - \lambda Id_{\mathbb{R}^3}| = \begin{vmatrix} M_\rho^+ + M_c^+ \frac{\partial c}{\partial \rho} - \lambda & M_c^+ \frac{\partial c}{\partial \rho c} & M_{\rho u}^+ \\ c(M^+)_\rho + [M^+ + c(M^+)_c] \frac{\partial c}{\partial \rho} & [M^+ + c(M^+)_c] \frac{\partial c}{\partial \rho c} - \lambda & (M^+ c)_{\rho u} \\ (F^+(U)_3)_\rho + (F^+(U)_3)_c \frac{\partial c}{\partial \rho} & (F^+(U)_3)_c \frac{\partial c}{\partial \rho c} & (F^+(U)_3)_{\rho u} - \lambda \end{vmatrix} = 0$$

En additionnant la première colonne et la seconde multipliée par c et en remarquant que c est invariante par les dilatations de $(\rho, \rho c)$:

$$c(\rho, \rho c) = \frac{\rho c}{\rho} \text{ d'où on tire } \frac{\partial c}{\partial \rho} + c \frac{\partial c}{\partial \rho c} = 0$$

Il vient:

$$\begin{vmatrix} M_\rho^+ + M_c^+ \underbrace{\left[\frac{\partial c}{\partial \rho} + c \frac{\partial c}{\partial \rho c} \right]}_{=0} - \lambda & M_c^+ \frac{\partial c}{\partial \rho c} & M_{\rho u}^+ \\ c(M^+)_\rho + [M^+ + c(M^+)_c] \underbrace{\left[\frac{\partial c}{\partial \rho} + c \frac{\partial c}{\partial \rho c} \right]}_{=0} - c\lambda & [M^+ + c(M^+)_c] \frac{\partial c}{\partial \rho c} - \lambda & c(M^+)_\rho \\ (F^+(U)_3)_\rho + (F^+(U)_3)_c \underbrace{\left[\frac{\partial c}{\partial \rho} + c \frac{\partial c}{\partial \rho c} \right]}_{=0} & (F^+(U)_3)_c \frac{\partial c}{\partial \rho c} & (F^+(U)_3)_{\rho u} - \lambda \end{vmatrix} = 0$$

Finalement, en soustrayant "c fois" la première ligne à la seconde, et en développant par rapport à la seconde ligne, on finit par trouver, puisque $\frac{\partial c}{\partial \rho c} = \frac{1}{\rho}$:

$$|dF^+(U) - \lambda Id_{\mathbb{R}^3}| = \left(\frac{M^+}{\rho} - \lambda \right) \begin{vmatrix} M^+ - \lambda & M^+_{\rho u} \\ (F^+(U)_3)_\rho & (F^+(U)_3)_{\rho u} - \lambda \end{vmatrix} = 0$$

C'est-à-dire que toute la dépendance en $\frac{\partial}{\partial c}$ a été enlevée du déterminant pour se retrouver dans le facteur. Ceci donne une valeur propre évidemment positive par définition de M^+ , et la fin du calcul est identique [83]. Le résultat implique par ailleurs que les valeurs propres de $dF^-(U)$ sont négatives puisque le traitement des termes en $\frac{\partial}{\partial c}$ est exactement similaire. Il reste aussi la condition $\forall U \in \Omega, F^+(U) + F^-(U) = F(U)$ qui est immédiate. \square

8.2.2 Réécriture des termes sources du système diphasique

Les fonctions $x \mapsto k_{1/2}(x)$ représentent l'intensité avec laquelle les phénomènes extérieurs influent sur l'écoulement. Elles peuvent être très dissemblables d'une région de l'espace à l'autre, et ainsi induire plusieurs échelles de temps. C'est l'origine de la *raideur numérique*: il est alors difficile de choisir correctement un maillage susceptible de garantir une bonne approximation des événements. Des techniques ont été développées pour intégrer de tels systèmes *raides* [42, 43], néanmoins, lorsque l'on traite des solutions-chocs, *on ne résout pas un problème raide*¹. Il paraît donc cohérent d'essayer d'intégrer ces termes grâce à des techniques réellement adaptées aux systèmes hyperboliques. Pour cela, introduisons deux fonctions définies à des constantes près:

$$K_1(x) = \int_0^x k_1(s) ds \text{ et } K_2(x) = \int_0^x k_2(s) ds$$

Définissons $x \mapsto a(x) = x$, alors (8.7) se réécrit sous la forme:

$$\begin{aligned} \rho_t + (\rho u)_x &= 0 \\ (\rho c)_t + (\rho u c)_x - \rho g(c) K_1(a)_x &= 0 \\ (\rho u)_t + (\rho u^2 + P)_x - \rho u h(u) K_2(a)_x &= 0 \end{aligned}$$

Il n'y a pour l'instant aucune ambiguïté puisque les termes non-conservatifs sont bien définis ($x \mapsto a(x)$ est régulière). Introduisons maintenant la fonction en escalier $x \mapsto a^{\Delta x}(x)$ (7.4) du chapitre 1. Alors, nous obtenons au bout du compte un système vraiment non-conservatif:

$$\begin{aligned} \rho_t + (\rho u)_x &= 0 \\ (\rho c)_t + (\rho u c)_x - \rho g(c) K_1(a^{\Delta x})_x &= 0 \\ (\rho u)_t + (\rho u^2 + P)_x - \rho u h(u) K_2(a^{\Delta x})_x &= 0 \end{aligned} \tag{8.9}$$

Et c'est par ce dernier que nous allons intégrer les termes de réaction *dans la définition des flux numériques* en résolvant au mieux possible le problème de Riemann pour (8.9). La principale difficulté étant bien sûr de proposer des relations de saut cohérentes pour les parties non-conservatives, puisque ce sont elles qui vont traduire l'action des termes réactifs.

8.2.3 Choix des relations de Rankine-Hugoniot stationnaires

Tout comme dans le cas scalaire, introduisons une fonction $c \mapsto \phi(c)$ vérifiant $\phi'(c) = \frac{1}{g(c)}$, alors la seconde équation de (8.9) peut se réécrire:

$$(\rho \phi(c))_t + (\rho u \phi(c))_x - \rho K_1(a^{\Delta x})_x = 0$$

En dehors des points $(x_{j+\frac{1}{2}})$, (8.9) est strictement équivalent à (8.7) dans lequel on pose $g \equiv h \equiv 0$. Le traitement des termes sources peut donc se voir comme l'adjonction d'une ligne de saut stationnaire ($\sigma \equiv 0$) le long de laquelle les quantités vont subir des conditions de Rankine-Hugoniot liées aux termes non-conservatifs. Pour déterminer celles-ci, on introduit les solutions stationnaires des problèmes décomposés suivants:

$$\begin{aligned} \rho_t + (\rho u)_x &= 0 \\ (\rho \phi(c))_t + (\rho u \phi(c))_x - \rho [K_1(a^{\Delta x})] \delta_{j+\frac{1}{2}} &= 0 \\ (\rho u)_t + (\rho u^2 + P)_x - \rho u h(u) [K_2(a^{\Delta x})] \delta_{j+\frac{1}{2}} &= 0 \end{aligned}$$

1. G. Wanner - Congrès d'analyse numérique (Les Karellis 1994)

et:

$$\begin{aligned} \rho_t + (M^\pm)_x &= 0 \\ (\rho\phi(c))_t + (M^\pm\phi(c))_x - \rho[K_1(a^{\Delta x})]\delta_{j+\frac{1}{2}} &= 0 \\ (\rho u)_t + \left(M^\pm \frac{((\gamma-1)u \pm 2a)^2}{\gamma} \right)_x - M^\pm h(u)[K_2(a^{\Delta x})]\delta_{j+\frac{1}{2}} &= 0 \end{aligned}$$

Ces solutions stationnaires vérifient les problèmes singuliers:

$$\begin{aligned} (\rho u)_x &= 0 \\ (\rho u\phi(c))_x - \rho[K_1(a^{\Delta x})]\delta_{j+\frac{1}{2}} &= 0 \\ (\rho u^2 + P)_x - \rho u h(u)[K_2(a^{\Delta x})]\delta_{j+\frac{1}{2}} &= 0 \end{aligned} \quad (8.10)$$

ou bien:

$$\begin{aligned} (M^\pm)_x &= 0 \\ (M^\pm\phi(c))_x - \rho[K_1(a^{\Delta x})]\delta_{j+\frac{1}{2}} &= 0 \\ \left(M^\pm \frac{((\gamma-1)u \pm 2a)^2}{\gamma} \right)_x - M^\pm h(u)[K_2(a^{\Delta x})]\delta_{j+\frac{1}{2}} &= 0 \end{aligned} \quad (8.11)$$

Comme dans le cas scalaire, on les construit comme limite $\epsilon \rightarrow 0$ des problèmes régularisés à l'aide de la fonction $x \mapsto a_\epsilon(x)$ (7.15):

$$\begin{aligned} \rho u\phi(c)_x &= \rho K_1(a_\epsilon)_x \\ (\rho u^2 + P)_x &= \rho u h(u) K_2(a_\epsilon)_x \end{aligned}$$

et:

$$\begin{aligned} (M^\pm\phi(c))_x &= \rho K_1(a_\epsilon)_x \\ \left(\frac{((\gamma-1)u \pm 2a)^2}{\gamma} \right)_x &= h(u) K_2(a_\epsilon)_x \end{aligned}$$

Ces problèmes sont fortement couplés par la pression (ou la vitesse du son, ce qui revient au même!). Il est donc en général impossible de calculer une expression de la solution à la limite $\epsilon \rightarrow 0$. Ceci nous oblige à traiter les termes non-conservatifs de façon simplifiée à l'aide (par exemple) de produits de Vol'Pert [85] consistant à remplacer les inconnues ρ et $h(u)$ par des moyennes. Il est tout de même nécessaire de maintenir la propriété de consistance avec le système d'origine; on propose donc:

$$\bar{\rho} = \frac{M^\pm}{\bar{u}} \quad \text{où } \bar{u} = (u_L + u_R)/2$$

Avec ce choix arbitraire, on donne un sens aux produits ambigus dans (8.10), (8.11). On est donc en position d'écrire les relations de saut imposées le long des interfaces du maillage en intégrant entre $x = 0$ et $x = \Delta x$ les systèmes simplifiés suivants:

$$\begin{aligned} \rho u\phi(c)_x - \bar{\rho}[K_1(a^{\Delta x})]\delta_{j+\frac{1}{2}} &= 0 \\ (\rho u^2 + P)_x - \rho u h(u)[K_2(a^{\Delta x})]\delta_{j+\frac{1}{2}} &= 0 \end{aligned}$$

et:

$$\begin{aligned} (M^\pm\phi(c))_x &= \bar{\rho}[K_1(a^{\Delta x})]\delta_{j+\frac{1}{2}} \\ \left(\frac{((\gamma-1)u \pm 2a)^2}{\gamma} \right)_x &= \bar{h}(u)[K_2(a^{\Delta x})]\delta_{j+\frac{1}{2}} \end{aligned}$$

Le long des interfaces du maillage, ceci nous donne dans le cas supersonique:

$$\begin{aligned} [\rho u] &= 0 \text{ d'où } \rho u \equiv M \\ M[\phi(c)] &= \bar{\rho}[K_1(a^{\Delta x})] \\ [\rho u^2 + P] &= M\bar{h}(u)[K_2(a^{\Delta x})] \end{aligned} \quad (8.12)$$

Et en régime subsonique, il convient d'écrire:

$$\begin{aligned} [M^\pm] &= 0 \\ M^\pm[\phi(c)] &= \bar{\rho}[K_1(a^{\Delta x})] \\ M^\pm \left[\frac{((\gamma-1)u \pm 2a)^2}{\gamma} \right] &= M^\pm \bar{h}(u)[K_2(a^{\Delta x})] \end{aligned} \quad (8.13)$$

Si on admet la pertinence des produits de Vol'Pert, la consistance avec le système d'origine (8.9) est immédiate lorsque l'on fait la somme de ces deux expressions.

8.2.4 Décomposition de flux modifiée pour le système diphasique

Le problème (8.1) étant reformulé sous la forme

$$\begin{cases} U = (\rho, \rho c, \rho u) \text{ et } F(U) = (\rho u, \rho c u, \rho u^2 + P) \\ U_t + F(U)_x - G(U) a_x^{\Delta x} = 0 \\ U_0 \in [BV(\mathbb{R})]^N \end{cases} \quad (8.14)$$

le schéma de Godunov consiste maintenant à résoudre dans chaque maille le problème de Riemann pour (8.14). C'est-à-dire trouver une écriture explicite des états situés de part et d'autre des lignes de discontinuité $x = x_{j+\frac{1}{2}}$. Il devient alors évident que la présence des ondes propageant à la fois dans les directions positives et négatives induit des situations extrêmement complexes. On préfère alors introduire des flux conduisant à des problèmes de Riemann plus proches de ceux rencontrés dans le cas scalaire (i.e. possédant une solution composée uniquement d'ondes propageant dans un seul sens), d'où l'intérêt d'une décomposition de type VanLeer consistant à introduire les deux sous-systèmes:

$$\begin{cases} U_t + F^\pm(U)_x - G(U) a_x^{\Delta x} = 0 \\ F^\pm \text{ définies par la proposition III.2} \end{cases} \quad (8.15)$$

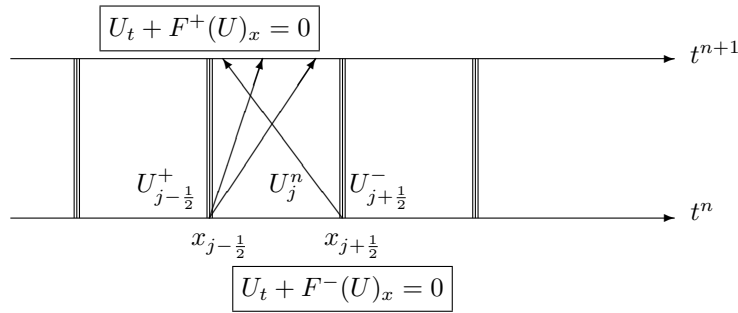


Figure 8.3: Décomposition de flux non-homogène

Le schéma numérique pour (8.15) se compose donc d'une concaténation de problèmes de Riemann aux limites associés à chaque sous-système. Ceux-ci sont bien posés puisque les lignes caractéristiques de ces derniers sont toujours dans le bon sens (i.e. rentrantes) suite à la définition des flux $U \mapsto F^\pm(U)$. On renvoie à [28] pour une analyse des problèmes de Riemann à conditions aux limites. On est donc conduit à résoudre dans chaque maille deux problèmes élémentaires d'un type similaire à ceux rencontrés pour les lois scalaires:

$$\begin{cases} U_t^{\Delta x} + F^+(U^{\Delta x})_x = 0 \\ U^{\Delta x}(x, 0) = U_j^n \text{ pour } x \in](j - \frac{1}{2}) \Delta x, (j) \Delta x[\\ U^{\Delta x}((j - \frac{1}{2}) \Delta x, t) = U_{j-\frac{1}{2}}^+ \text{ pour } t \in [n \Delta t, (n+1) \Delta t[\\ U_{j-\frac{1}{2}}^+ \text{ défini par les relations (8.12) (cas supersonique) ou (8.13) (cas subsonique) à partir de } U_{j-1}^n \end{cases} \quad (8.16)$$

$$\begin{cases} U_t^{\Delta x} + F^-(U^{\Delta x})_x = 0 \\ U^{\Delta x}(x, 0) = U_j^n \text{ pour } x \in](j) \Delta x, (j + \frac{1}{2}) \Delta x[\\ U^{\Delta x}((j + \frac{1}{2}) \Delta x, t) = U_{j+\frac{1}{2}}^- \text{ pour } t \in [n \Delta t, (n+1) \Delta t[\\ U_{j+\frac{1}{2}}^- \text{ défini par les mêmes relations de saut à partir de } U_{j+1}^n \end{cases}$$

En fin de compte, le schéma proposé pour le système (8.7) peut s'écrire sous la forme condensée:

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} + \frac{F^+(U_j^n) - F^+(U_{j-\frac{1}{2}}^+) + F^-(U_{j+\frac{1}{2}}^-) - F^-(U_j^n)}{\Delta x} = 0 \quad (8.17)$$

L'interprétation géométrique donnée dans la figure 8.2 est bien entendu toujours valable aux nouvelles conditions de saut près (voir la figure 8.3).

8.2.5 Principe du maximum pour la concentration massique

Dans le cas homogène $g \equiv 0$, il a été montré que ce type de solveur respectait des propriétés de positivité, ainsi qu'une borne uniforme sur la variable c [54]. On va voir que sous des conditions très naturelles sur le terme de réaction, il est possible d'établir un résultat analogue.

Proposition III.3 *Supposons $k_1(x) \geq 0$ et $k_2(x) \equiv 0$, alors sous les hypothèses suivantes:*

$$\begin{cases} g(0) \geq 0 \text{ et } g(1) \leq 0 \text{ et } g \in C^1([0, 1]; \mathbb{R}) \\ \forall(j, n), \frac{(u_j^n)^2 + (a_j^n)^2}{2a_j^n} \Delta t \leq \Delta x \text{ en régime subsonique} \\ \forall(j, n), |u_j^n| \Delta t \leq \Delta x \text{ en régime supersonique} \end{cases} \quad (8.18)$$

l'approximation $(c_j^n)_{j,n}$ vérifie:

$$\forall(j, n), c_j^n \in [0, 1] \quad (8.19)$$

Rappelons aussi que le schéma homogène est TVD sur la variable c en l'absence de terme source sur l'équation de la fraction massique [54].

Démonstration

Posons M_j^\pm la première composante du vecteur $F^\pm(U_j^n)$, et $\lambda = \frac{\Delta t}{\Delta x}$. Le schéma s'écrit pour les deux premières composantes:

$$\begin{cases} \rho_j^{n+1} = \rho_j^n - \lambda[M_j^+ - M_{j-\frac{1}{2}}^+ + M_{j+\frac{1}{2}}^- - M_j^-] \\ (\rho c)_j^{n+1} = (\rho c)_j^n - \lambda[(M^+ c)_j - (M^+ c)_{j-\frac{1}{2}} + (M^- c)_{j+\frac{1}{2}} - (M^- c)_j] \end{cases}$$

En reportant la première inégalité dans la seconde et en remarquant que M^\pm est invariant au travers des interfaces $x_{j\pm\frac{1}{2}}$, il vient:

$$c_j^{n+1} = \frac{c_j^n [\rho_j^n - \lambda(M_j^+ - M_j^-)] + \lambda M_{j-1}^+ c_{j-\frac{1}{2}}^+ - \lambda M_{j+1}^- c_{j+\frac{1}{2}}^-}{[\rho_j^n - \lambda(M_j^+ - M_j^-)] + \lambda M_{j-1}^+ - \lambda M_{j+1}^-} \quad (8.20)$$

Comme par définition, $M^+ \geq 0$ et $M^- \leq 0$, la seule condition à imposer pour que c_j^{n+1} appartienne à l'enveloppe convexe de $c_{j-\frac{1}{2}}^+, c_j^n, c_{j+\frac{1}{2}}^-$ est:

$$\rho_j^n \geq \lambda(M_j^+ - M_j^-)$$

On remarque ensuite que:

$$\begin{cases} M_j^+ - M_j^- = \rho_j \frac{u_j^2 + a_j^2}{2a_j} \text{ (Cas subsonique)} \\ M_j^+ - M_j^- = \rho_j |u_j| \text{ (Cas supersonique)} \end{cases}$$

Ce qui correspond bien à l'une ou l'autre des conditions imposées dans (8.18). Reste maintenant à s'assurer de la validité de $c_{j\pm\frac{1}{2}}^\pm$. Par des arguments similaires à ceux développés dans le cas scalaire (voir le lemme III.1), il est clair que l'intervalle $[0, 1]$ est invariant par les problèmes suivants du fait du signe de $k_1(x)$ et des hypothèses sur $c \mapsto g(c)$:

$$\begin{cases} M_{j-1}^+(c^+)_x = \bar{\rho} g(c^+) K_1(a_\epsilon)_x \text{ pour } x \in [0, \Delta x] \\ c^+(0) = c_{j-1}^n \\ M_{j+1}^-(c^-)_x = \bar{\rho} g(c^-) K_1(a_\epsilon)_x \text{ pour } x \in [-\Delta x, 0] \\ c^-(0) = c_{j+1}^n \end{cases}$$

Dans le cas où $k_2 \equiv 0$, il est possible de passer à la limite $\epsilon \rightarrow 0$ en introduisant une fonction ϕ telle que $\phi'(c) = \frac{1}{g(c)}$ et l'on obtient:

$$\begin{aligned} \phi(c^+(\Delta x)) - \phi(c_{j-1}) &= \frac{\bar{\rho}_{j-\frac{1}{2}}}{M_{j-1}^+} [K_1(a^{\Delta x})] \geq 0 \\ \phi(c^-(-\Delta x)) - \phi(c_{j+1}) &= \frac{\bar{\rho}_{j+\frac{1}{2}}}{M_{j+1}^-} [K_1(a^{\Delta x})] \geq 0 \end{aligned}$$

Par conséquent, $c_{j-\frac{1}{2}}^+ = c^+(\Delta x) \in [0, 1]$ et $c_{j+\frac{1}{2}}^- = c^-(-\Delta x) \in [0, 1]$. Ceci implique que si la concentration correspondant à la donnée initiale est confinée dans l'intervalle $[0, 1]$, alors l'étape des relations de saut provenant du terme source respectera cette propriété. Et donc, sous la condition de type CFL, l'égalité (8.20) assure que $c_j^{n+1} \in [0, 1]$ dès que la donnée initiale $(c_j^n)_{j,n}$ l'est. \square

8.3 Essais numériques: termes de recondensation et de frottement pariétal

L'objet de cette dernière partie est l'étude pratique du système (8.7) dans diverses configurations. L'expérimentation se fera essentiellement par le biais de *tubes à choc*, c'est-à-dire de *problèmes de Riemann non-homogènes*. Ce sera en particulier l'occasion de constater la perte d'auto-similarité de leurs solutions. Il pourra être aussi intéressant d'évaluer la stabilité de ce type de schéma.

8.3.1 Ecriture précise du problème étudié

Il s'agit donc du système (8.7) appliqué à la modélisation simplifiée d'un mélange eau/vapeur [81] pour lequel on choisit les lois suivantes:

$$\begin{cases} h(u) = -|u| \text{ et } g(c) = c(1-c)(\mu-c) \text{ où } \mu \in \mathbb{R} \\ P(\rho, c) = (\rho_{vapeur})^2 \text{ où } \rho_{vapeur} = \frac{1.6\rho c}{1.6 - \rho(1-c)} \end{cases} \quad (8.21)$$

Le terme $h(u) = -|u|$ modélise un effet de frottement pariétal sensé faire localement décroître la vitesse volumétrique. Quant à $g(c) = c(1-c)(\mu-c)$, il modélise un phénomène de recondensation vers un état privilégié. Lorsque μ est négatif, l'état stable est $c = 0$, et la pression est nulle: ceci correspond à un écoulement de liquide quasiment incompressible. Lorsque μ est supérieur à 1, l'état stable est $c = 1$, ce qui correspond à un écoulement de vapeur considérée comme un gaz parfait isentropique. Le plus souvent, $\mu \in]0, 1[$ est l'état stable et on a affaire à un écoulement diphasique particulier. On ne détaillera pas ici la façon de calculer la valeur spécifique μ .

La loi de pression (8.21) ne donne pas une vitesse du son de la forme $\sqrt{\frac{2P}{\rho}}$, il est donc judicieux d'en simplifier l'expression afin de récupérer cette propriété. On propose l'approximation:

$$P(\rho, c) = \rho^2 \sqrt{c}$$

Le problème résolu dans chaque maille sera donc:

$$\begin{aligned} \rho_t + (\rho u)_x &= 0 \\ (\rho c)_t + (\rho u c)_x - \rho c(1-c)(\mu-c)K_1(a^{\Delta x})_x &= 0 \\ (\rho u)_t + (\rho u^2 + \rho^2 \sqrt{c})_x + \rho u|u|K_2(a^{\Delta x})_x &= 0 \end{aligned} \quad (8.22)$$

Il sera donc utile d'introduire, comme à l'accoutumée, une fonction $c \mapsto \phi(c)$ vérifiant:

$$\phi'(c) = \frac{1}{c(1-c)(\mu-c)}$$

Ceci a pour conséquence de modifier l'écriture de l'équation de transport-réaction de la concentration massique:

$$(\rho\phi(c))_t + (\rho u\phi(c))_x - \rho K_1(a^{\Delta x})_x = 0$$

8.3.2 Relations de saut approchées

C'est le point faible de cette construction: trouver des quantités moyennes $\bar{\rho}$, $|\bar{u}|$ acceptables afin de pouvoir écrire des conditions de saut du type (8.12), (8.13). Théoriquement, celles-ci devraient s'obtenir (cf chapitre 1) via une régularisation affine par morceaux de la fonction $x \mapsto a^{\Delta x}(x)$. Malheureusement, ceci est très difficile dans le cas (8.22), même en ne prenant en compte qu'un seul terme source à la fois. On se replie donc derrière des produits de Vol'Pert [85] consistant simplement à choisir

$$\bar{u} = \frac{u_L + u_R}{2} \text{ et } \bar{\rho} = \begin{cases} M/\bar{u} & \text{dans le cas supersonique} \\ M^+/\bar{u} & \text{dans le cas subsonique et pour système associé à } F^+ \\ M^-/\bar{u} & \text{dans le cas subsonique et pour système associé à } F^- \end{cases}$$

Quant à la fonction ϕ introduite plus haut, elle s'exprime en forme fermée grâce encore à une décomposition en éléments simples.

Lemme III.6 Soient $\mu, c_0 \in]0, 1[$; le système (8.22) admet, pour la variable c , une fonction de saut de la forme:

$$\begin{aligned} \phi :]-\infty, 0[\cup]0, \mu[\cup]\mu, 1[\rightarrow \mathbb{R} \\ c \mapsto \phi(c) = \alpha_1 \ln \left(\frac{c}{c_0} \right) + \alpha_2 \ln \left(\frac{1-c}{1-c_0} \right) + \alpha_3 \ln \left(\frac{|\mu-c|}{|\mu-c_0|} \right) \\ \text{où } \alpha_1 = \frac{1}{\mu}; \alpha_2 = \frac{1}{1-\mu}; \alpha_3 = \frac{-1}{\mu(1-\mu)} \end{aligned} \quad (8.23)$$

De plus, cette fonction est bijective de $]0, \mu[\rightarrow \mathbb{R}$ et de $]\mu, 1[\rightarrow \mathbb{R}$.

La preuve de ce résultat est immédiate: on l'omet ici. Par contre, il faut noter qu'il est tout à fait possible d'étendre cette fonction de saut à des cas où $\mu \notin]0, 1[$; l'écriture du lemme étant pour le compte sensiblement alourdie.

Un point essentiel à la construction d'un tel schéma est de bien savoir ce que l'on désire connaître de l'autre côté des interfaces. En effet, lorsque l'on dispose d'une fonction de saut du type (8.23), il paraît clair que l'on cherche une composante de l'état $U_{j\pm\frac{1}{2}}^n$ satisfaisant à (8.12), (8.13). Il se peut néanmoins que ce ne soit pas toujours ce genre de quantité qu'il faille chercher. Par exemple, pour l'équation de (8.22) exprimant l'évolution de l'impulsion, la condition de saut s'écrit:

$$\begin{aligned} F^+(U_{j-\frac{1}{2}}^n)_3 - F^+(U_{j-1}^n)_3 &= M^+ |\bar{u}| [K_2(a^{\Delta x})] \\ F^-(U_{j+\frac{1}{2}}^n)_3 - F^-(U_{j+1}^n)_3 &= M^- |\bar{u}| [K_2(a^{\Delta x})] \end{aligned}$$

Au vu de l'expression de la troisième composante de $F^\pm(U)$, on réalise la difficulté à évaluer la quantité de mouvement ρu , alors qu'en fait, seule est intéressante la troisième composante des flux $F^\pm(U_{j\pm\frac{1}{2}}^n)_3$. Au bout du compte, le schéma utilisé pour approcher les solutions du système (8.22) s'écrit précisément:

$$\begin{aligned} U_j^+ &= U_j^n - \frac{2\Delta t}{\Delta x} \left(\begin{array}{c} M_j^+ - M_{j-1}^+ \\ M_j^+ c_j^n - M_{j-1}^+ c_{j-\frac{1}{2}}^+ \\ M_j^+ \frac{(\gamma-1)u_j^n + 2a_j^n}{\gamma} - M_{j-1}^+ \frac{(\gamma-1)u_{j-1}^n + 2a_{j-1}^n}{\gamma} + M_{j-1}^+ |u_j^n + u_{j-1}^n| [K_2]/2 \end{array} \right) \\ U_j^- &= U_j^n - \frac{2\Delta t}{\Delta x} \left(\begin{array}{c} M_{j+1}^- - M_j^- \\ M_{j+1}^- c_{j+\frac{1}{2}}^- - M_j^- c_j^n \\ M_{j+1}^- \frac{(\gamma-1)u_{j+1}^n + 2a_{j+1}^n}{\gamma} - M_j^- \frac{(\gamma-1)u_j^n + 2a_j^n}{\gamma} + M_{j+1}^- |u_{j+1}^n + u_j^n| [K_2]/2 \end{array} \right) \\ U_j^{n+1} &= \frac{1}{2}(U_j^+ + U_j^-) \end{aligned}$$

Avec les relations de saut:

$$\begin{cases} \phi(c_{j-\frac{1}{2}}^+) - \phi(c_{j-1}^n) = \frac{\bar{p}}{M_{j-1}^+} [K_1(j\Delta x) - K_1((j-1)\Delta x)] \\ \phi(c_{j+\frac{1}{2}}^-) - \phi(c_{j+1}^n) = \frac{\bar{p}}{M_{j+1}^-} [K_1(j\Delta x) - K_1((j+1)\Delta x)] \end{cases}$$

Ceci correspond bien à l'écriture du schéma (8.17) modulo les relations non-conservatives approchées.

8.3.3 Exemples de cas-tests pour le système diphasique

Les cas-tests sont tous issus de la thèse [81] où l'on peut avoir une idée de l'allure de la solution *homogène* de (8.22). Ces essais numériques se feront aussi pour des pas de temps Δt à la limite de la condition de stabilité C.F.L. (8.18). Le schéma-témoin choisi est de type Roe totalement implicite pour une plus grande robustesse. L'impression globale qui se dégage est un bon accord entre les deux types d'approximations, à la diffusion numérique près. Le plus grand avantage de cette nouvelle construction étant son moindre coût vu qu'elle reste explicite même en régime réactionnel très intense.

Les données initiales pour le système (8.22) correspondant à un tube à choc subsonique avec les paramètres:

$$\begin{cases} k_1(x) = 100e^{-100(x-\frac{1}{2})^2} \\ k_2(x) = 20e^{-100(x-\frac{1}{2})^2} \\ \mu = 0.1 \end{cases} \quad \text{et} \quad \begin{cases} P_L = 2.5, P_R = 0.5 \\ c_L = 0.2, c_R = 0.5 \\ u_L = u_R = 0 \end{cases} \quad (8.24)$$

Le maillage est tel que $\Delta x = 0.02$, $\Delta t = 0.01$, ce qui donne une C.F.L de 0.88, et on itère jusque $T = 0.3$. Les résultats du schéma aux interfaces ainsi que ceux d'un schéma implicite linéarisé construit à partir des résultats de [81] sont représentés dans les figures 8.4, 8.5, 8.6. On constate une bonne adéquation avec le schéma à pas fractionnaires et surtout l'absence complète de toute oscillation numérique au voisinage des discontinuités. La convergence en temps se fait de façon totalement monotone. Quant aux paliers à l'équilibre, ce sont bien entendu les valeurs attendues à savoir $c = 0.2$ à proximité de $x = 0$ et $c = 0.5$ à proximité de $x = 1$. La valeur $c = 0.1$ est atteinte au voisinage de $x = 0.5$.

Les données initiales du deuxième essai sont cette fois de type supersonique/subsonique et en régime plus sévère:

$$\begin{cases} k_1(x) = 5000e^{-100(x-\frac{1}{2})^2} \\ k_2(x) = 100e^{-100(x-\frac{1}{2})^2} \\ \mu = 0.5 \end{cases} \quad \text{et} \quad \begin{cases} P_L = 0.5, P_R = 0.5 \\ c_L = 0.2, c_R = 0.9 \\ u_L = 1.44, u_R = 0 \end{cases} \quad (8.25)$$

Le maillage est tel que $\Delta x = 0.02$, $\Delta t = 0.006$ et on itère jusque $T = 0.2$. Les résultats sont en figures 8.7,8.8,8.9. Le schéma est très robuste dans ce régime difficile où se mêlent une transition supersonique/subsonique ainsi qu'une forte raideur réactionnelle. Là encore, on observe que la structure de la solution obtenue par le nouveau schéma est bien en accord avec celle présentée dans [81] pour le cas homogène. D'autre part, il n'apparaît aucune oscillation parasite et les paliers de concentration sont bien conformes à ce qui est attendu au vu de l'expression du terme $g(c)$ (l'état stable $c = 0.5$ est respecté).

Ces résultats encourageants nous mettent en position de discuter les avantages et les limites de l'approche proposée dans ce chapitre:

Robustesse c'est effectivement le principal objectif de cette construction. On a réussi à fournir une bonne approximation *explicite* d'un système muni de termes de réaction très raides (celle-ci est validée par comparaison avec un schéma plus classique à pas fractionnaires où le second membre est traité de façon implicite en temps). La seule restriction sur le pas de temps est prescrite par la condition CFL qui permet en outre de garantir que la concentration massique est bien confinée dans un intervalle de valeurs réalistes pour le problème considéré.

Mise en oeuvre comme tous les schémas à décomposition de flux, celui-ci présente une alternative intéressante du point de vue de l'implémentation. D'autre part, les relations de saut associées aux termes sources ne sont pas trop complexes à mettre en place.

Equilibres délicats on touche là au point faible de cette construction à savoir la manière grossière avec laquelle on a traité les termes non-conservatifs. En effet, dans le cas où l'équilibre entre les termes convectifs et ceux de réaction est complexe (par exemple les paliers de vitesse et de pression de part et d'autre d'une grille de perte de charge modélisée sur une interface du maillage), il est vain d'espérer qu'une moyenne arithmétique permette d'obtenir des résultats précis. Il est donc nécessaire de raffiner la prise en compte des difficultés liées aux aspects non-conservatifs.

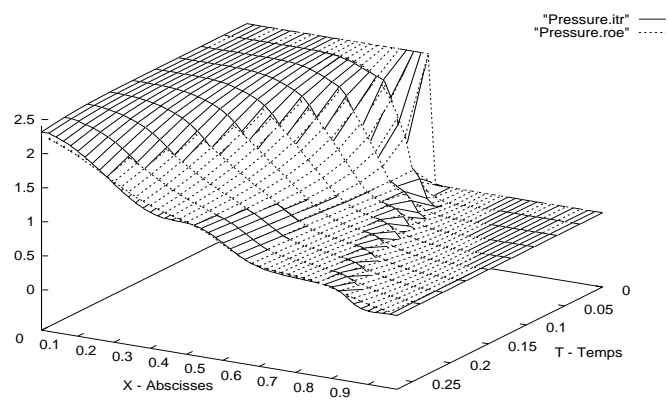


Figure 8.4: Pression (tube à choc subsonique)

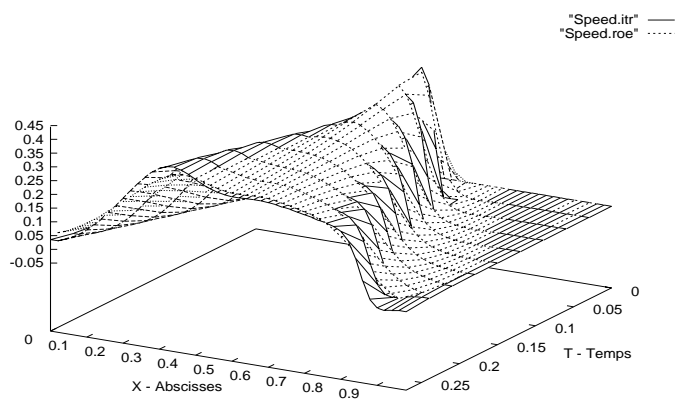


Figure 8.5: Vitesse (tube à choc subsonique)

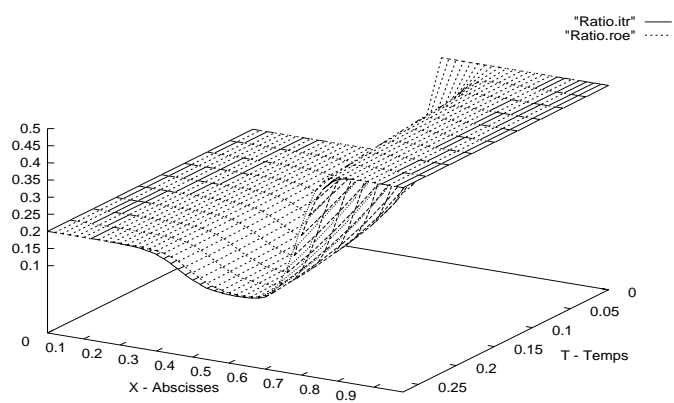


Figure 8.6: Concentration (tube à choc subsonique)

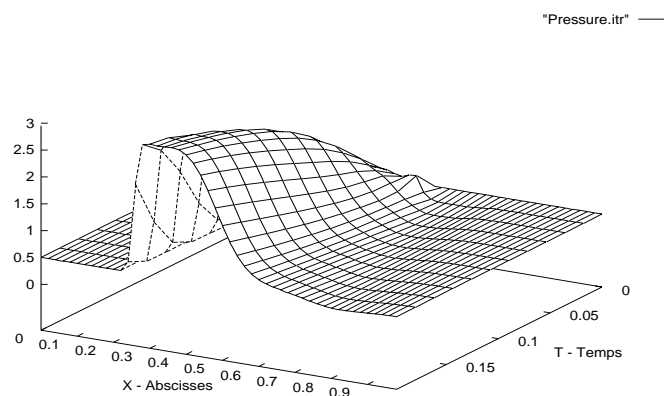


Figure 8.7: Pression (tube à choc supersonique - décomposition de flux)

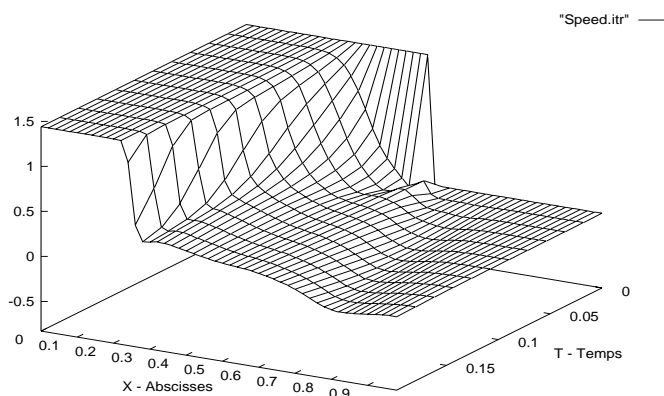


Figure 8.8: Vitesse (tube à choc supersonique - décomposition de flux)

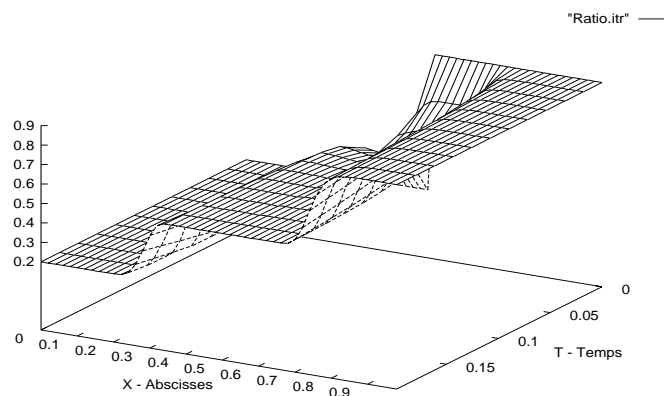


Figure 8.9: Concentration (tube à choc supersonique - décomposition de flux)

Chapitre 9

Problèmes non-homogènes et formalisme non-conservatif - Approche linéarisée

1

Ce chapitre est motivé par un autre type d'extension à des systèmes de lois de conservations non-homogènes. A la différence près que l'on va se concentrer, cette fois sur une résolution précise des termes non-conservatifs qui apparaissent dans les relations de saut à travers les interfaces. Il est pour cela commode de se plonger dans le formalisme des chemins introduit par [21], ce dernier pouvant se relier à la théorie des fonctions généralisées [12]. Le schéma scalaire peut alors se reformuler en des termes s'étendant sans difficulté à des problèmes plus complexes. Il est à noter que l'asymptote horizontale (chapitre 1) admet alors une interprétation géométrique claire. Les applications numériques se feront par le biais de la formulation faible du solveur de Roe [73, 82] qui permet de construire un schéma respectant par construction les états d'équilibres corrects du système considéré. Cette propriété n'est pas vraie pour des schémas conventionnels: on peut par exemple consulter [11] pour s'en convaincre. Les essais présentés ici ont pour cadre le système des équations d'Euler. Il est d'abord muni d'un terme modélisant une variation de section (cf [32]), comme par exemple une tuyère divergente (cf [27]). Le cas d'un écoulement dans un conduit pourvu de grilles d'aubes est aussi présenté, et la consistance avec les équilibres théoriques est alors soigneusement examinée.

9.1 Formulation faible du solveur non-conservatif

La motivation de cette partie est de mettre en place un cadre de travail adapté à des problèmes de termes sources devant conduire à un équilibre harmonieux entre termes sources et flux hyperboliques. Ce point a été négligé dans la construction du chapitre 2 puisque l'on s'est limité à des moyennes arithmétiques pour écrire les termes non-conservatifs. Il faut donc souligner que l'objectif n'est pas ici de bâtir des schémas très robustes en régime réactionnel raide. Il s'agit plutôt de s'attacher à respecter un comportement asymptotique correct, i.e. une bonne discrétisation de la solution stationnaire $U(x) \in BV(\mathbb{R})^N$, vérifiant $F(U)_x = G(U)$ [65].

9.1.1 Définition d'un chemin d'équilibre scalaire

Généralités

Se ramenant dans un premier temps à une loi scalaire (7.1,7.2) pour $u \in BV(\mathbb{R} \times \mathbb{R}^+)$ écrite sous la forme:

$$\begin{cases} u_t + f(u)_x - g(u)a_x^{\Delta x} = 0 \\ u_0 \in BV(\mathbb{R}) \end{cases}$$

où $x \mapsto a^{\Delta x}(x)$ est la fonction introduite au chapitre 1 (7.4), on cherche à expliciter une régularisation du produit ambigu $g(u)a_x^{\Delta x}$. D'après [21, 85], il est connu que ce dernier peut se définir comme une mesure borelienne bornée. Sa construction repose alors sur la donnée d'une famille de chemins lipschitziens Φ tracés dans l'espace des états et il convient de noter:

$$[g(u)a_x^{\Delta x}]_{\Phi}$$

la mesure associée à ce produit de distributions. Nous allons décomposer ce chemin Φ en deux sous-chemins associés à deux discontinuités de natures différentes: l'une de vitesse $\sigma \neq 0$ provient du terme de convection conservatif $f(u)_x$ et vérifie les relations classiques de Rankine-Hugoniot, l'autre, en revanche, est purement non-conservative et exprime l'action localisée du terme source $[g(u)a_x^{\Delta x}]_{\Phi}$. Le chemin Φ comportera donc deux sous-chemins suivant le sens de la propagation, i.e. suivant le signe de σ :

$$\begin{aligned} & \bullet \text{ Si } \sigma > 0 \\ \Phi(s; u_L, u_R) &= \begin{cases} \Phi_1(\gamma_1(s); u_L, u_M) \text{ avec } \phi(u_M) - \phi(u_L) = \Delta x \\ \Phi_2(\gamma_2(s); u_M, u_R) \text{ avec } [f(u)] = \sigma[u] \text{ et } [a^{\Delta x}] = 0 \end{cases} \\ & \bullet \text{ Si } \sigma < 0 \\ \Phi(s; u_L, u_R) &= \begin{cases} \Phi_2(\gamma_1(s); u_L, u_M) \text{ avec } [f(u)] = \sigma[u] \text{ et } [a^{\Delta x}] = 0 \\ \Phi_1(\gamma_2(s); u_M, u_R) \text{ avec } \phi(u_R) - \phi(u_M) = \Delta x \end{cases} \end{aligned} \tag{9.1}$$

$$\text{avec } \begin{cases} \gamma_1(s) = 2s \text{ pour } s \in [0, 1/2] \\ \gamma_2(s) = 2s - 1 \text{ pour } s \in [1/2, 1] \end{cases}$$

Néanmoins, et parce que l'on se borne à des problèmes écrits originellement sous forme conservative (i.e. de la forme $u_t + f(u)_x = g(u)$), il n'y a qu'une seule branche Φ_i de Φ qui soit non-triviale. En effet, lorsque l'on

dispose d'une fonction de flux, il vient:

$$\int_0^1 f'(\Phi(s))\Phi'(s)ds = f(\Phi(1)) - f(\Phi(0))$$

Par conséquent, seule la branche le long de laquelle $a_x^{\Delta x} \neq 0$ vaut la peine d'être explicitée. On fera donc dans toute la suite un abus de langage en désignant par *chemin* Φ la restriction $s \mapsto \Phi_j \circ \gamma_i(s)$ de (9.1) le long de laquelle $x \mapsto a^{\Delta x}(x)$ n'est pas constante. Les autres branches sont caduques puisqu'elles correspondent à des expressions sous forme de divergence, et peuvent être exprimées à l'aide de n'importe quelle paramétrisation (les lignes droites, par exemple).

Le système scalaire

Le premier jalon de ce travail est d'introduire le problème augmenté

$$\begin{cases} u_t + f(u)_x - g(u)a_x^{\Delta x} = 0 \\ a_t^{\Delta x} = 0 \end{cases} \quad (9.2)$$

où la fonction $x \mapsto a^{\Delta x}(x)$ joue maintenant le rôle d'une inconnue à part entière. On pose donc $V = (u, a^{\Delta x})$ le vecteur des inconnues. Le problème (9.2) est alors non-conservatif, non-strictement hyperbolique de valeurs propres 0 et $f'(u)$. Le champ associé à 0 est évidemment linéairement dégénéré: il induit une *discontinuité de contact stationnaire*. C'est par ce biais que l'on va pouvoir donner une interprétation des sauts à l'interface en termes de problèmes de Riemann augmentés. Pour l'instant, il est indispensable de donner un sens aux relations de saut pour (9.2): suivant le formalisme de [21], on écrit les relations de saut généralisées (voir la définition C.8)

$$\begin{aligned} \sigma[a^{\Delta x}] &= 0 \\ \sigma[u] &= [f(u)] - [g(u)a_x^{\Delta x}]_{\Phi} \end{aligned} \quad (9.3)$$

Ceci n'est d'aucun secours pour construire un profil microscopique donné par $s \mapsto \Phi(s; V_L, V_R)$. Néanmoins, et parce que nous travaillons sur des fonctions scalaires, il est possible d'exprimer (9.2) sous forme conservative. Pour cela, définissons deux fonctions de la façon suivante:

$$\psi(u) = \int^u \frac{dv}{g(v)} \quad \text{et} \quad \phi(u) = \int^u \frac{f'(v)}{g(v)} dv$$

Il est alors simple de voir que l'on obtient:

$$\begin{cases} (\psi(u))_t + (\phi(u) - a^{\Delta x})_x = 0 \\ a_t^{\Delta x} = 0 \end{cases} \quad (9.4)$$

Et ceci nous donne des conditions de Rankine-Hugoniot pour (9.2) qui ne sont pas sans rappeler celles obtenues au chapitre 1:

$$\begin{aligned} \sigma[a^{\Delta x}] &= 0 \\ \sigma[\psi(u)] &= [\phi(u)] - [a^{\Delta x}] \end{aligned} \quad (9.5)$$

En particulier, le long du saut stationnaire $\sigma = 0$, l'inconnue vérifie $[\phi(u)] = [a^{\Delta x}]$.

Théorème III.3 On suppose $a^{\Delta x} \in BV(\mathbb{R})$ et $u \mapsto \frac{g(u)}{f'(u)}$ localement lipschitzienne. Pour $u_L, a_L^{\Delta x}, a_R^{\Delta x}$ donnés, on construit les régularisations pour $x \in [0, \Delta x]$:

$$\begin{cases} \tilde{a}(x) = a_L^{\Delta x} + (a_R^{\Delta x} - a_L^{\Delta x})x/\Delta x \\ \tilde{u}(x) \text{ solution de } f(\tilde{u})_x = g(\tilde{u})\tilde{a}_x; \tilde{u}(0) = u_L \end{cases}$$

Alors, si $u_R = \tilde{u}(\Delta x)$, le chemin Φ :

$$[0, 1] \ni s \mapsto \Phi \left(s; \left(\begin{array}{c} u_L \\ a_L^{\Delta x} \end{array} \right), \left(\begin{array}{c} u_R \\ a_R^{\Delta x} \end{array} \right) \right) = \left(\begin{array}{c} \tilde{u}(s\Delta x) \\ \tilde{a}(s\Delta x) \end{array} \right) \quad (9.6)$$

vérifie les conditions d'admissibilité (C.3).

De plus, tout couple (V_L, V_R) tel que $u_R = \tilde{u}(\Delta x)$ vérifiant les relations généralisées (9.3) vérifie aussi les conditions de Rankine-Hugoniot (9.5).

C'est-à-dire que $x \mapsto a^{\Delta x}(x)$ est régularisée par une fonction affine par morceaux (le chemin des lignes droites). Ceci est cohérent avec la construction du chapitre 1.

Démonstration

On montre tout d'abord que tout couple (V_L, V_R) tel que $u_R = \tilde{u}(\Delta x)$ vérifiant (9.3) vérifie aussi les relations conservatives (9.5).

Supposons que u_L et u_R soient situés sur une courbe intégrale de $f(\tilde{u})_x = g(\tilde{u})\tilde{a}_x$, on pose:

$$u_R = \tilde{u}(\Delta x) \text{ avec } \begin{cases} f'(\tilde{u})\tilde{u}_x = g(\tilde{u})\tilde{a}_x \\ \tilde{u}(0) = u_L \end{cases}$$

L'équation peut se réécrire $\frac{f'(\tilde{u})}{g(\tilde{u})}\tilde{u}_x = \tilde{a}_x$, c'est-à-dire:

$$\int_0^{\Delta x} \phi(\tilde{u})_x dx = \int_0^{\Delta x} \tilde{a}_x dx$$

D'où la relation $[\phi(\tilde{u})] = [\tilde{a}]$. Comme on a $[a^{\Delta x}]_0^{\Delta x} = [\tilde{a}]_0^{\Delta x}$, il est clair que $[\phi(\tilde{u})] = [a^{\Delta x}]$ et on a le résultat. Il faut maintenant s'assurer de la régularité Lipschitz du chemin proposé. Soit une paire d'états $(V_L, V_R) \in \mathcal{V}$, \mathcal{V} borné dans \mathbb{R}^2 effectivement joignables par le chemin (9.6), alors il faut montrer qu'il existe une constante k telle que:

$$\left| \frac{\partial \Phi}{\partial s}(s; V_L, V_R) \right| \leq k |V_L - V_R|$$

avec la notation $\Delta(\cdot) = (\cdot)_R - (\cdot)_L$:

$$\frac{\partial \Phi}{\partial s}(s; V_L, V_R) = \begin{pmatrix} \Delta x \frac{\partial \tilde{u}}{\partial x} \\ \Delta x \frac{\partial \tilde{a}}{\partial x} \end{pmatrix} = \begin{pmatrix} \Delta x \frac{g(\tilde{u})(x)}{f'(\tilde{u})(x)} \frac{\Delta(\tilde{a})}{\Delta(\tilde{a})} \\ \Delta(\tilde{a}) \end{pmatrix}$$

Donc, comme $[\tilde{a}] = [a^{\Delta x}]$, on a:

$$\left| \frac{\partial \Phi}{\partial s}(s; V_L, V_R) \right| = \left| \frac{g(\tilde{u})(x)}{f'(\tilde{u})(x)} \frac{\Delta(a^{\Delta x})}{\Delta(a^{\Delta x})} \right| \leq \underbrace{\sup_{\mathcal{V} \text{ borné}} \left| \frac{g(\tilde{u})(x)}{f'(\tilde{u})(x)} \right|}_{\text{borné}} \cdot |\Delta(a^{\Delta x})| \leq k |\Delta(a^{\Delta x})|$$

et le résultat suit.

Soient deux paires d'états $(V_L^i, V_R^i)_{i=1,2} \in \mathcal{V}$, \mathcal{V} borné dans \mathbb{R}^2 effectivement joignables par le chemin (9.6), alors il faut montrer qu'il existe une constante K telle que:

$$\left| \frac{\partial \Phi}{\partial s}(s; V_L^1, V_R^1) - \frac{\partial \Phi}{\partial s}(s; V_L^2, V_R^2) \right| \leq K (|V_L^1 - V_L^2| + |V_R^1 - V_R^2|)$$

Pour cela, on décide d'établir la propriété pour la norme particulière $|\cdot|_1$. C'est-à-dire qu'on va devoir majorer le membre de gauche à l'aide de la quantité:

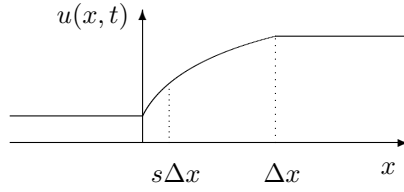
$$|V_L^1 - V_L^2|_1 + |V_R^1 - V_R^2|_1 = |u_R^1 - u_R^2| + |(a^{\Delta x})_R^1 - (a^{\Delta x})_R^2| + |u_L^1 - u_L^2| + |(a^{\Delta x})_L^1 - (a^{\Delta x})_L^2|$$

Pour commencer, on a:

$$\left| \frac{\partial \Phi}{\partial s}(s; V_L^1, V_R^1) - \frac{\partial \Phi}{\partial s}(s; V_L^2, V_R^2) \right|_1 = \left| \begin{pmatrix} \frac{g(\tilde{u}^1)(x)}{f'(\tilde{u}^1)(x)} \Delta(\tilde{a})^1 \\ \Delta(\tilde{a})^1 \end{pmatrix} - \begin{pmatrix} \frac{g(\tilde{u}^2)(x)}{f'(\tilde{u}^2)(x)} \Delta(\tilde{a})^2 \\ \Delta(\tilde{a})^2 \end{pmatrix} \right|_1$$

Commençons par calculer les composantes de ce vecteur:

$$\begin{aligned} \Delta(a^{\Delta x})^1 - \Delta(a^{\Delta x})^2 &= [(a^{\Delta x})_R^1 - (a^{\Delta x})_L^1] - [(a^{\Delta x})_R^2 - (a^{\Delta x})_L^2] \\ &= [(a^{\Delta x})_L^2 - (a^{\Delta x})_L^1] - [(a^{\Delta x})_R^2 - (a^{\Delta x})_R^1] \end{aligned}$$

Figure 9.1: Régularisation du profil de $u(x, t)$

Par conséquent, cette composante ne posera aucun problème. Il reste à remarquer que:

$$\frac{g(\tilde{u}^1)}{f'(\tilde{u}^1)} \Delta(a^{\Delta x})^1 - \frac{g(\tilde{u}^2)}{f'(\tilde{u}^2)} \Delta(a^{\Delta x})^2 = \frac{g(\tilde{u}^2)}{f'(\tilde{u}^2)} [\Delta(a^{\Delta x})^1 - \Delta(a^{\Delta x})^2] + \Delta(a^{\Delta x})^1 \left[\frac{g(\tilde{u}^1)}{f'(\tilde{u}^1)} - \frac{g(\tilde{u}^2)}{f'(\tilde{u}^2)} \right]$$

Par l'hypothèse de régularité sur $u \mapsto \frac{g(u)}{f'(u)}$ et par la théorie de Cauchy des équations différentielles, on a:

$$\begin{aligned} \forall x \in [0, \Delta x], \left| \frac{g(\tilde{u}^1)}{f'(\tilde{u}^1)}(x) - \frac{g(\tilde{u}^2)}{f'(\tilde{u}^2)}(x) \right| &\leq C |\tilde{u}^1(x) - \tilde{u}^2(x)| \\ &\leq CC' |\tilde{u}^1(0) - \tilde{u}^2(0)| \end{aligned}$$

Comme le domaine \mathcal{V} est borné dans l'espace des états, le terme $\frac{g(\tilde{u}^2)}{f'(\tilde{u}^2)}$ est borné et on a:

$$\left| \frac{g(\tilde{u}^1)}{f'(\tilde{u}^1)} \Delta(a^{\Delta x})^1 - \frac{g(\tilde{u}^2)}{f'(\tilde{u}^2)} \Delta(a^{\Delta x})^2 \right| \leq \sup_{\mathcal{V}} \left| \frac{g(\tilde{u}^2)}{f'(\tilde{u}^2)} \right| |\Delta(a^{\Delta x})^1 - \Delta(a^{\Delta x})^2| + CC' |\Delta(a^{\Delta x})^1| |u^1(0) - u^2(0)|$$

On déduit donc par définition de la norme $|\cdot|_1$ que si l'on pose $K = \max \left(1 + \sup_{\mathcal{V}} \left| \frac{g(\tilde{u}^2)}{f'(\tilde{u}^2)} \right|, CC' \sup_{\mathcal{V}} |\Delta(a^{\Delta x})^1| \right)$:

$$\left| \frac{\partial \Phi}{\partial s}(s; V_L^1, V_R^1) - \frac{\partial \Phi}{\partial s}(s; V_L^2, V_R^2) \right|_1 \leq K. [|\tilde{u}^1(0) - \tilde{u}^2(0)| + |\Delta(a^{\Delta x})^1 - \Delta(a^{\Delta x})^2|]$$

Finalement, comme $\tilde{u}^1(0) = u_L^1$ et $\tilde{u}^2(0) = u_L^2$, on obtient que:

$$\left| \frac{\partial \Phi}{\partial s}(s; V_L^1, V_R^1) - \frac{\partial \Phi}{\partial s}(s; V_L^2, V_R^2) \right|_1 \leq K. (|V_L^1 - V_L^2|_1 + |V_R^1 - V_R^2|_1)$$

Il reste à conclure que cette majoration est indépendante du choix de la norme puisque celles-ci sont toutes équivalentes en dimension finie. \square

Remarque Ce chemin n'est pas issu d'une régularisation visqueuse du problème (9.2) comme c'est souvent le cas². Il paraît en outre difficile d'en introduire une autre tout en préservant le résultat de convergence du chapitre 1.

Adoptant les notations habituelles, on pose

$$A(V) = \begin{pmatrix} f'(u) & -g(u) \\ 0 & 0 \end{pmatrix} \text{ et } \frac{\partial \Phi}{\partial s}(s; V_L, V_R) = \begin{pmatrix} \Delta x \frac{\partial \tilde{u}}{\partial x} \\ \Delta x \end{pmatrix}$$

Par conséquent, si $a^{\Delta x}$ désigne la fonction définie en (7.4) régularisée par les lignes droites \tilde{u} , les relations (9.3,9.5) impliquent le long du saut immobile:

$$0 = [f(u)] - \int_0^1 g(\Phi(s; V_L, V_R)) \Delta x \cdot ds \text{ donc } \Delta x = \frac{[f(u)]}{\int_0^1 g(\Phi(s; V_L, V_R)) ds} = [\phi(u)]$$

Si $u \mapsto \phi(u)$ est localement bijective, le lemme III.3 nous assure de l'existence d'un $\xi \in [u_L, u_R]$ tel que:

$$\Delta x = \frac{[\phi(u)]}{[f(u)]} [f(u)] = \underbrace{\frac{\phi'(\xi)}{f'(\xi)}}_{1/g(\xi)} [f(u)]$$

2. Dans [21], il est suggéré de régulariser les produits non-conservatifs grâce à une onde progressive issu d'un tenseur de diffusion. Dans le cas présent, il s'agit plutôt d'un profil asymptotique ($t \rightarrow +\infty$) du problème étudié.

On en déduit que ξ et Φ sont liés par la relation suivante:

$$g(\xi) = \int_0^1 g(\Phi(s; V_L, V_R)) ds = \frac{[f(u)]}{[\phi(u)]} \tag{9.7}$$

D'où le caractère totalement décentré de cette approche non-conservative.

Exemple: l'équation d'advection-réaction

Soit le problème simple dans lequel $a^{\Delta x}$ désigne la fonction définie en (7.4):

$$u_t + u_x = g(u) \text{ i.e. } \begin{cases} u_t + u_x - g(u)a^{\Delta x} = 0 \\ a_t^{\Delta x} = 0 \end{cases} \tag{9.8}$$

En ce cas, il est immédiat que la caractéristique issue de zéro s'écrit $\tau \mapsto X(\tau) = \tau$, et que $\dot{u} = g(u)$ le long de celle-ci. Le chemin (9.6) est donc donné par:

$$\Phi(s; V_L, V_R) = \begin{pmatrix} \tilde{u}(s\Delta x) \\ \tilde{a}(s\Delta x) \end{pmatrix} \text{ où } \tilde{u}_x = g(\tilde{u}); \tilde{u}(0) = u_L$$

9.1.2 Le problème de Riemann

L'écriture (9.2) suggère une interprétation de (7.1,7.2) via l'intersection de courbes d'ondes dans le plan de phases comme pour un problème de Riemann [système]. Il s'agit donc de connecter les états $V_L = (u_L, 0)$ et $V_R = (u_R, \Delta x)$ grâce à l'adjonction d'un état intermédiaire V_M situé sur la courbe intégrale de l'équation $f(\tilde{u})_x = g(\tilde{u})$ passant par V_L . (cf Fig. 9.2) Cette onde stationnaire est aussi introduite dans [49] sous le vocable *standing wave*.

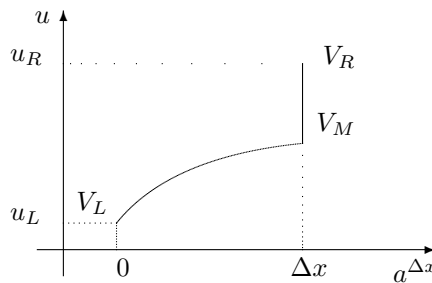


Figure 9.2: Problème de Riemann

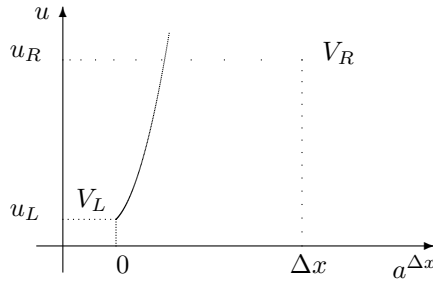
Le cas de [l'asymptote horizontale] (cf chapitre 1) se comprend alors aisément: il est impossible d'expliciter un V_M tel que $a^{\Delta x}$ avance de Δx via le chemin (9.6). On est alors dans une situation du type décrit par la figure 9.3. Comme il a déjà été vu, cela est dû à l'apparition d'un point u^* pour lequel $f'(u^*) = 0$. Ceci implique que l'équation générique $f(\tilde{u})_x = g(\tilde{u})$ devient caduque et le chemin Φ ne peut plus être défini comme précédemment³. On est alors en présence d'un phénomène de *resonance* qui se traduit (entre autres choses) par une très forte non-unicité des solutions du problème de Riemann [49].

9.1.3 Extension aux systèmes hyperboliques

Le plus grand avantage que présente le résultat (9.6) est sa capacité à être étendu à des problèmes plus complexes sans changer sa formulation. En effet, considérons le système:

$$\begin{cases} U_t + F(U)_x = G(U) \text{ avec } U \in \mathbb{R}^N \\ U_0 \in [BV(\mathbb{R})]^N \end{cases}$$

3. C'est un cas de figure où le profil asymptotique est discontinu [31, 67].

Figure 9.3: Situation lorsque $f'(u) = 0$

il paraît naturel d'introduire et de régulariser les mêmes termes non-conservatifs via une courbe intégrale du système stationnaire $F(U)_x = G(U)$. C'est-à-dire que l'on va obtenir un système (évidemment encore hyperbolique puisque l'on ne fait que rajouter un champ associé à la valeur propre 0) non-conservatif de dimension $N + 1$, agissant sur un vecteur d'inconnues augmenté $V = (U, a^{\Delta x})$, écrit sous la forme:

$$\begin{cases} U_t + F(U)_x - [G(U)a_x^{\Delta x}]_{\Phi} = 0 \\ a_t^{\Delta x} = 0 \end{cases}$$

avec $\Phi \left(s; \begin{pmatrix} U_L \\ a_L^{\Delta x} \end{pmatrix}, \begin{pmatrix} U_R \\ a_R^{\Delta x} \end{pmatrix} \right) = \begin{pmatrix} \tilde{U}(s\Delta x) \\ \tilde{a}(s\Delta x) \end{pmatrix}$ (9.9)

où \tilde{U} est une courbe intégrale de $F(\tilde{U})_x = G(\tilde{U})\tilde{a}_x$; $\tilde{U}(0) = U_L$
et $\tilde{a}(x) = a_L^{\Delta x} + (a_R^{\Delta x} - a_L^{\Delta x})x/\Delta x$

Bien évidemment, cette régularisation ne peut encore être envisagée que dans le cas où le système $F(U)_x = G(U)$ a un sens, i.e. pour des valeurs de U pour lesquelles $dF(U)$ est inversible.

9.2 Formulation d'un solveur approché de type Roe

On s'intéresse maintenant à la construction d'un schéma numérique exploitant la formulation non-conservative introduite précédemment. Pour cela, il est de mise d'envisager un solveur de type Godunov prenant en compte les termes de la forme $[G(U)a_x^{\Delta x}]_{\Phi}$ dans l'expression des flux numériques. Le chemin défini par Φ permet en théorie de résoudre le problème de Riemann pour le système:

$$\begin{cases} U_t + F(U)_x - G(U)a_x^{\Delta x} = 0 \\ a_t^{\Delta x} = 0 \end{cases} \quad (9.10)$$

avec la donnée initiale :

$$U(x, 0) = \begin{cases} U_L \in \mathbb{R}^N & \text{pour } x < 0 \\ U_R \in \mathbb{R}^N & \text{pour } x > 0 \end{cases} \quad a^{\Delta x}(x, 0) = \begin{cases} a_L^{\Delta x} & \text{pour } x < 0 \\ a_R^{\Delta x} & \text{pour } x > 0 \end{cases} \quad (9.11)$$

au moins dans les cas où la jacobienne des flux reste inversible. Dans la pratique, cela donne lieu à un système d'équations non-linéaires difficile à résoudre. Il est donc légitime de chercher un solveur approché de ce problème de Riemann.

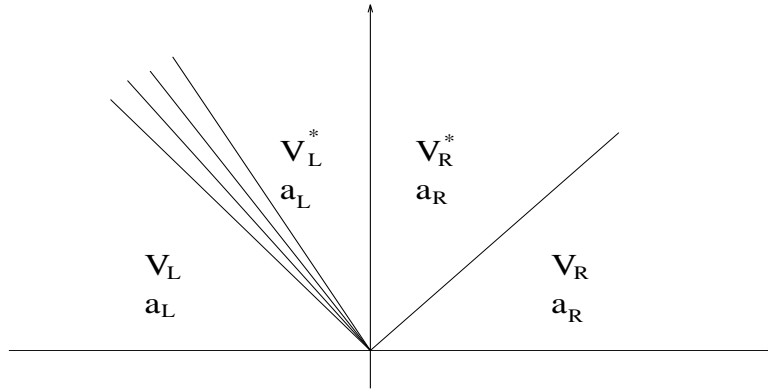


Figure 9.4: Problème de Riemann pour le système non-conservatif.

9.2.1 Structure de la linéarisée non-conservative

Dans ce chapitre, on va donc s'orienter vers un solveur linéarisé de type Roe. Il s'agit alors de trouver pour tout couple (V_L, V_R) de la forme (9.11) une matrice de la forme

$$A_{\Phi}(V_L, V_R) = \begin{pmatrix} A & -G_{\Phi} \\ 0 & 0 \end{pmatrix} (V_L, V_R) \in \mathbf{M}_{N+1}(\mathbb{R}) \quad (9.12)$$

telle que le problème de Riemann pour l'équation linéaire

$$\begin{cases} U_t + A(V_L, V_R)U_x - G_{\Phi}(V_L, V_R)a_x^{\Delta x} = 0 \\ a_t^{\Delta x} = 0 \end{cases}$$

et avec la même donnée initiale (9.11) approche (9.10). Ici, G_{Φ} désigne une moyenne du terme source $G(U)$ le long du chemin Φ plus adaptée que celle utilisée au chapitre précédent. Suivant [73] et [82], on impose les propriétés suivantes:

- le problème linéarisé est strictement hyperbolique pour tout couple (V_L, V_R) , i.e. $A_{\Phi}(V_L, V_R)$ est diagonalisable dans \mathbb{R} et ses valeurs propres sont distinctes.
- si $V_L = V_R$, alors $A(V_L, V_R) = dF(U)$ et $G_{\Phi}(V_L, V_R) = G(U)$.
- la condition de consistance (voir la définition IV.9 dans l'annexe A et [24]):

$$A_{\Phi}(V_L, V_R) \cdot (V_R - V_L) = A(U_R - U_L) - G_{\Phi}(a_R^{\Delta x} - a_L^{\Delta x}) = F(U_R) - F(U_L) - [G(U)a_x^{\Delta x}]_{\Phi} \quad (9.13)$$

Tout comme pour un problème homogène, c'est cette dernière exigence qui permet d'explicitier la matrice A_{Φ} . En effet, supposons que A soit une matrice de Roe classique pour le système homogène $U_t + F(U)_x = 0$, alors on a pour tout couple (U_L, U_R) :

$$F(U_R) - F(U_L) = A(U_L, U_R) \cdot (U_R - U_L)$$

Par conséquent, si on reporte cette propriété dans (9.13), il reste (puisque $a^{\Delta x}$ est régularisée par les lignes droites, le terme $(\tilde{a})_x = a_R^{\Delta x} - a_L^{\Delta x}$ est constant):

$$G_{\Phi}(a_R^{\Delta x} - a_L^{\Delta x}) = [G(U)a_x]_{\Phi} = \left(\int_0^1 G(\Phi(s; V_L, V_R)) ds \right) (a_R^{\Delta x} - a_L^{\Delta x}). \quad (9.14)$$

Ceci est une condition très contraignante puisque le calcul de l'intégrale sur le chemin Φ nécessite la résolution du problème de Riemann : il semble en effet qu'il faille connaître les bornes V_L^* et V_R^* qui sont les états de part et d'autre de la discontinuité stationnaire liée au terme source (voir la figure 9.4).

9.2.2 Moyenne d'équilibre de certains termes sources

L'objectif est ici de présenter une façon de calculer une moyenne approchée $G_{\Phi}(V_L, V_R)$ en s'affranchissant de la résolution du problème de Riemann non-conservatif. Pour cela, on exige seulement que (9.14) soit vérifié

si $U_L = U_L^*$ et $U_R = U_R^*$. Cela revient à dire que l'expression:

$$G_\Phi(V_L, V_R) = \int_0^1 G(\Phi(s; V_L, V_R)) ds$$

est réalisée seulement pour des états $(V_L, V_R) = \left(\begin{array}{c} U_L \\ a_L^{\Delta x} \end{array} \right), \left(\begin{array}{c} U_R \\ a_R^{\Delta x} \end{array} \right)$ vérifiant:

$$F(\tilde{U})_x = G(\tilde{U})\tilde{a}_x; \tilde{U}(0) = U_L \text{ et } \tilde{U}(\Delta x) = U_R \quad (9.15)$$

Pour ces états particuliers, on peut calculer:

$$\int_0^1 G(\Phi(s; V_L, V_R)) ds = \int_0^{\Delta x} G(\tilde{U})(x) \frac{dx}{\Delta x} = \frac{1}{2}[G(U_L) + G(U_R)] + o(\Delta x^2)$$

Et on définit donc la moyenne approchée:

$$G_\Phi(V_L, V_R) = \frac{1}{2}[G(U_L) + G(U_R)]$$

Afin d'obtenir une méthode de calcul pour un couple (V_L, V_R) quelconque, on aura besoin de faire deux hypothèses sur le vecteur de termes sources à savoir:

i) Pour tout $U \in \mathbb{R}^N$, $G(U)$ admet $(N - 1)$ composantes identiquement nulles. Donc on note:

$$G(U) = (0, 0, \dots, 0, G_N(U_1, U_2, \dots, U_N)) \text{ avec } F(U) = (F_1, F_2, \dots, F_N)(U_1, U_2, \dots, U_N)$$

ii) Comme la fonction de flux F est supposée localement inversible, il y a $(N - 1)$ quantités conservées: $F_1(U)(0), F_2(U)(0), \dots, F_{N-1}(U)(0)$ par l'équation $F(U)_x = G(U)\tilde{a}_x$ qui peut se réécrire selon:

$$F_N(U_1(0), U_2(0), \dots, U_N(x))_x = G_N \circ F^{-1}(F_1(U)(0), F_2(U)(0), \dots, F_N(U)(x))\tilde{a}_x \text{ i.e. } f(u)_x = g(u)\tilde{a}_x$$

où $u(x) = U_N(x)$. On suppose alors que la valeur $g(u)$ associée à la solution de ce problème *scalaire* peut s'écrire au moyen d'une fonction régulière φ sous la forme suivante:

$$g(u)(\Delta x) = g(u) \frac{\varphi(u, a^{\Delta x})(0)}{\varphi(u, a^{\Delta x})(\Delta x)}$$

Sous ces hypothèses, on propose de définir $G_\Phi(V_L, V_R)$ pour tout couple (V_L, V_R) en calculant l'intégrale:

$$\begin{aligned} \int_0^1 G(\Phi(s; V_L, V_R)) ds &= (0, 0, \dots, \int_0^{\Delta x} g(u)(x) \frac{dx}{\Delta x}) \\ &\simeq (0, 0, \dots, \frac{1}{2}[g(u)(0) + g(u)(\Delta x)]) \end{aligned}$$

Du fait de la forme particulière de $g(u)(x)$, il vient que:

$$G_\Phi(V_L, V_R) = \frac{G(U)(0)}{2} \left(1 + \frac{\varphi(U_N(0), a_L^{\Delta x})}{\varphi(U_N(\Delta x), a_R^{\Delta x})} \right)$$

Dans cette expression, il reste un degré de liberté puisque l'on ne veut pas calculer l'état $U_N(\Delta x)$. On choisit donc de fixer $U(\Delta x) = U_R$. Ce choix est justifié au moins dans le cas où la donnée initiale est proche d'un état d'équilibre. Ceci nous permet d'obtenir finalement l'expression approchée:

$$G_\Phi(V_L, V_R) = \frac{G(U_L)}{2} \left(1 + \frac{\varphi(U_L, a_L^{\Delta x})}{\varphi(U_R, a_R^{\Delta x})} \right) \quad (9.16)$$

On remarque que si (V_L, V_R) vérifient (9.15), alors cette moyenne redonne $G_\Phi(V_L, V_R) = \frac{1}{2}[G(U_L) + G(U_R)]$. Autrement dit, le solveur approché ainsi construit est exact dans le cas où la donnée initiale est une discrétisation d'une courbe à l'équilibre du type $F(U)_x = G(U)\tilde{a}_x$. Ceci est suffisant pour garantir la propriété suivante:

Proposition III.4 *La linéarisation préserve les équilibres à l'ordre 2: si V_L et V_R sont sur la même courbe d'équilibre i.e. $F(U_R) - F(U_L) - [G(U)a_x^{\Delta x}]_\Phi = 0$ alors $A(U_R - U_L) - G_\Phi(a_R^{\Delta x} - a_L^{\Delta x}) = o(\Delta x^2)$.*

Dans les sections suivantes, on calcule de cette manière des linéarisations G_Φ préservant les équilibres pour certains termes sources issus de systèmes physiquement réalistes.

9.2.3 Ecriture du schéma non-conservatif

Une fois la linéarisation du problème de Riemann effectuée, on peut écrire un schéma décentré d'ordre 1. Introduisons tout d'abord quelques notations : la matrice A étant diagonalisable dans \mathbb{R} , on appelle λ_i , ($i = 1, \dots, N$) ses valeurs propres. On note ensuite Π_i le projecteur sur le $i^{\text{ème}}$ vecteur propre de A et on pose

$$\Pi^+ = \sum_{i, \lambda_i > 0} \Pi_i \quad , \quad \Pi^- = \sum_{i, \lambda_i < 0} \Pi_i \quad .$$

Concrètement, si $A = \vec{r}\Lambda\vec{l}$ avec $\Lambda = \text{diag}(\lambda_i)$, alors Π^+ est une matrice $N \times N$: $\Pi^+ = \vec{r}E^+\vec{l}$,

$$E^+ = \text{diag}(\varepsilon_i^+) \quad , \quad \varepsilon_i^+ = \begin{cases} 1 & \text{si } \lambda_i > 0 \\ 0 & \text{sinon.} \end{cases}$$

Proposition III.5 *Le schéma s'écrit alors :*

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} [F_{j+1/2} - F_{j-1/2}] + \Delta t \left[\Pi_{j+1/2}^- (G_\Phi)_{j+1/2} + \Pi_{j-1/2}^+ (G_\Phi)_{j-1/2} \right] \quad (9.17)$$

avec le flux de Roe classique: $F_{j+1/2} = \frac{1}{2}[F(U_{j+1}^n) + F(U_j^n)] - \frac{1}{2}|A_{j+1/2}|[U_{j+1}^n - U_j^n]$ et $(G_\Phi)_{j+1/2} = G_\Phi(V_{j+1}^n, V_j^n)$.

Démonstration

Le schéma de Roe non-conservatif au sens de [24] (p. 46) pour le système (9.10) consiste à résoudre à chaque interface $x_{j+\frac{1}{2}}$ du maillage le problème de Riemann linéarisé:

$$\begin{cases} V_i + A_\Phi(V_L, V_R) = 0 \\ V_L = (U_L, a_L^{\Delta x})^T; V_R = (U_R, a_R^{\Delta x})^T \end{cases}$$

Le schéma s'écrit donc:

$$V_j^{n+1} = V_j^n - \frac{\Delta t}{\Delta x} [(A_\Phi(V_j^n, V_{j+1}^n)^-)(V_{j+1}^n - V_j^n) + (A_\Phi(V_{j-1}^n, V_j^n)^+)(V_j^n - V_{j-1}^n)] \quad (9.18)$$

avec les matrices définies par:

$$A_\Phi(V_L, V_R)^+ = \frac{1}{2}(A_\Phi(V_L, V_R) + |A_\Phi(V_L, V_R)|) \quad \text{et} \quad A_\Phi(V_L, V_R)^- = \frac{1}{2}(A_\Phi(V_L, V_R) - |A_\Phi(V_L, V_R)|)$$

On calcule donc l'expression de $|A_\Phi|$. Comme A_Φ est diagonalisable dans \mathbb{R} , on peut écrire $A_\Phi = \vec{R}\Lambda_\Phi\vec{L}$ ($\Lambda_\Phi = \Lambda \cup \{0\}$) où la matrice de passage \vec{R} est de la forme:

$$\vec{R} = \begin{pmatrix} \vec{r} & \vec{r}_0 \\ 0 & 1 \end{pmatrix} \quad \text{avec} \quad A_\Phi \begin{pmatrix} \vec{r}_0 \\ 1 \end{pmatrix} = \vec{0}$$

Par la structure de A_Φ , cela amène:

$$\vec{r}_0 = A^{-1}G_\Phi$$

D'autre part, comme $\vec{R}\vec{L} = Id$, \vec{L} est de la forme:

$$\vec{L} = \begin{pmatrix} \vec{l} & \vec{l}_0 \\ 0 & 1 \end{pmatrix} \quad \text{où} \quad \vec{l}_0 = -\vec{l}A^{-1}G_\Phi$$

Comme $A = \vec{r}\Lambda\vec{l}$, il vient que $\vec{l}_0 = -\Lambda^{-1}\vec{l}G_\Phi$.

D'autre part, pour tout $V \in \mathbb{R}^{N+1}$ de la forme $V = \begin{pmatrix} U \in \mathbb{R}^N \\ a^{\Delta x} \end{pmatrix}$, on a:

$$|A_\Phi|V = \vec{R}|\Lambda_\Phi|\vec{L}V = \vec{r}|\Lambda|(\vec{l}U - \Lambda^{-1}\vec{l}G_\Phi[a^{\Delta x}])$$

Finalement, comme $[a^{\Delta x}] = \Delta x$, il vient:

$$|A_\Phi|V = |A|U - \underbrace{\vec{r}|\Lambda|\Lambda^{-1}\vec{l}}_{\Pi^+ + \Pi^-} G_\Phi \Delta x$$

Et donc on a :

$$A_{\Phi}^{\pm}(V_L, V_R) = \begin{pmatrix} A^{\pm}(U_L, U_R) & -\Pi^{\pm}(U_L, U_R)G_{\Phi}(V_L, V_R) \\ 0 & 0 \end{pmatrix} \text{ avec } A^{\pm}(U_L, U_R) = (\Pi^{\pm}A)(U_L, U_R)$$

Le schéma (9.18) se réécrit donc sous la forme:

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} [(A(U_j^n, U_{j+1}^n)^-)(U_{j+1}^n - U_j^n) + (A(U_{j-1}^n, U_j^n)^+)(U_j^n - U_{j-1}^n)] \\ + \Delta t [\Pi^-(U_j^n, U_{j+1}^n)G_{\Phi}(V_j^n, V_{j+1}^n) + \Pi^+(U_{j-1}^n, U_j^n)G_{\Phi}(V_{j-1}^n, V_j^n)]$$

Du fait des relations de Roe, $A(U_j^n, U_{j+1}^n)(U_{j+1}^n - U_j^n) = F(U_{j+1}^n) - F(U_j^n)$ et on obtient:

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{2\Delta x} [(F(U_{j+1}^n) + F(U_j^n)) - (F(U_j^n) + F(U_{j-1}^n))] \\ + \frac{\Delta t}{2\Delta x} [|A(U_j^n, U_{j+1}^n)|(U_{j+1}^n - U_j^n) + |A(U_{j-1}^n, U_j^n)|(U_j^n - U_{j-1}^n)] \\ + \Delta t [\Pi^-(U_j^n, U_{j+1}^n)G_{\Phi}(V_j^n, V_{j+1}^n) + \Pi^+(U_{j-1}^n, U_j^n)G_{\Phi}(V_{j-1}^n, V_j^n)]$$

On vérifiera sans peine que ce schéma coïncide avec l'écriture (9.17). \square

Remarque

Il faut noter que cette expression linéarisée traite les termes sources de façon explicite. Par conséquent, ce schéma n'est d'ores-et-déjà pas un bon candidat pour les problèmes raides.

9.3 Premier cas-test: tuyère quasi-1D

9.3.1 Mise en équations du phénomène

Considérons un conduit possédant une symétrie axiale dans lequel s'effectue un écoulement de gaz parfait. On suppose en outre que la variation de section $x \mapsto S(x)$ est régulière et lentement variable ($S \in C^1([0, 1])$) de manière à pouvoir considérer un système d'équations monodimensionnelles (cf [32, 89]).

$$\begin{cases} (\rho S)_t + (\rho u S)_x = 0 \\ (\rho u S)_t + (S(\rho u^2 + p))_x = pS'(x) \\ (\rho e S)_t + (S u(\rho e + p))_x = 0 \\ x \in [0, 1] \end{cases} \quad (9.19)$$

Comme toujours, ρ , u , i désignent respectivement la densité, la vitesse, l'énergie interne du fluide; $e = i + \frac{p}{\rho}$, $p = (\gamma - 1)\rho i$ désignent quant à elles l'énergie totale et la pression, en fonction du rapport de chaleur massique γ . On note:

$$U = S.(\rho, \rho u, \rho e) \quad F(U) = S. \left(\rho u, \rho u^2 + p, \rho u \left(e + \frac{p}{\rho} \right) \right) \quad G(U) = (0, p, 0)$$

9.3.2 Relations de saut et quantités invariantes

La première étape est d'effectuer un changement de variables [32] dans (9.19): soit $R = \rho S$, $P = pS$, $h = i + \frac{u^2}{2} + \frac{P}{R}$

$$\begin{cases} R_t + (Ru)_x = 0 \\ (Ru)_t + (Ru^2 + P)_x - \frac{P}{S}S_x = 0 \\ (Re)_t + (Ruh)_x = 0 \\ S_t = 0 \end{cases} \quad (9.20)$$

$$V = (R, Ru, Re, S) \text{ et } \tilde{F}(V) = (Ru, Ru^2 + P, Ruh, 0)$$

C'est le système des équations d'Euler avec, en plus, un champ stationnaire associé à la nouvelle variable $S(x)$. Pour en expliciter les relations de saut, il faut d'abord donner un sens au produit $\frac{P}{S}S_x$. On introduit pour cela le chemin (9.6).

$$\Phi(s; V_L, V_R) = \begin{pmatrix} \tilde{U}(s\Delta x) \\ \tilde{S}(s\Delta x) \end{pmatrix}$$

où \tilde{U} est une courbe intégrale de $F(\tilde{U})_x = G(\tilde{U})\tilde{S}_x$; $\tilde{U}(0) = U_L$
et $\tilde{S}(x) = S_L + (S_R - S_L)x/\Delta x$

Les relations de Rankine-Hugoniot conservatives donnent déjà:

$$[Ru] = 0 \text{ et } [Ruh] = (Ru)[h] = 0$$

Il reste à trouver de quelle manière le saut de densité équilibre les pressions de part et d'autre du saut de section $[S]$. C'est-à-dire qu'il va falloir résoudre:

$$\begin{cases} (Ru)_x = 0 \\ \left[\frac{(Ru)^2}{R} + P \right]_x = \frac{PS_x}{S} \\ (Ru)h_x = 0 \end{cases} \quad (9.21)$$

Ce qui va permettre de mener à bien les calculs est de remarquer:

$$P = \frac{\gamma-1}{\gamma} \rho S \left(h - \frac{u^2}{2} \right) = \frac{\gamma-1}{\gamma} \left(Rh - \frac{(Ru)^2}{2R} \right) \text{ et donc } P = P(R)$$

Pour $P(R) \neq 0$, la seconde équation de (9.21) se réécrit donc:

$$(Ru)^2 \frac{-R_x}{R^2 P(R)} + \frac{P_x}{P} = \frac{S_x}{S}$$

Posons $a = \frac{Ru}{\sqrt{2h}}$ ($R > a$ puisque $R^2 > (Ru)^2/(2h)$ est assuré par $h > u^2/2$), il vient:

$$R^2 P(R) = \frac{\gamma-1}{\gamma} h R \left(R^2 - \frac{(Ru)^2}{2h} \right) = \frac{(\gamma-1)h}{\gamma} R(R+a)(R-a)$$

On obtient donc la condition au premier ordre en R :

$$\frac{2a^2\gamma}{\gamma-1} \frac{-R_x}{R(R+a)(R-a)} + \frac{P_x}{P} = \frac{S_x}{S}$$

Une décomposition en éléments simples fournit:

$$\frac{1}{x(x+a)(x-a)} = \frac{-1}{a^2x} + \frac{1}{2a^2(x-a)} + \frac{1}{2a^2(x+a)}$$

D'où:

$$\frac{\gamma}{\gamma-1} \left(\frac{2R_x}{R} - \frac{R_x}{(R-a)} - \frac{R_x}{(R+a)} \right) + \frac{P_x}{P} = \frac{S_x}{S}$$

Posons $\varphi(R) = \frac{R^2}{(R+a)(R-a)}$, il vient finalement:

$$\frac{\gamma}{\gamma-1} \ln \left(\frac{\varphi(R(x))}{\varphi(R_L)} \right) + \ln \left(\frac{P(x)}{P_L} \right) = \ln \left(\frac{S(x)}{S_L} \right)$$

Il reste à traiter le cas où la vitesse est nulle, i.e. $Ru = 0$. Par conséquent, la variation d'enthalpie totale peut être quelconque (non-unicité): on la choisit nulle, d'où $h_x = 0$. Il reste donc:

$$-\frac{(Ru)^2}{R^2} R_x + P(R)_x = \frac{P}{S} S_x \text{ et donc } \frac{P(R)}{P(R_L)} = \frac{S}{S_L}$$

Proposition III.6 *Le long du contact $\sigma = 0$, les relations suivantes sont vérifiées:*

$$\begin{aligned} &\bullet \text{Si } u_L \neq 0 : \\ &\left\{ \begin{array}{l} (Ru)_R = (Ru)_L \text{ (conservation du débit)} \\ h_R = h_L \text{ (conservation de l'enthalpie totale)} \\ \left(\frac{\varphi(R_R)}{\varphi(R_L)} \right)^{\frac{\gamma}{\gamma-1}} \frac{P_R}{S_R} = \frac{P_L}{S_L} \text{ (équilibre de la pression)} \end{array} \right. \end{aligned} \quad (9.22)$$

$$\begin{aligned} &\bullet \text{Si } u_L = 0 : \\ &\left\{ \begin{array}{l} (Ru)_R = (Ru)_L \text{ et } h_R = h_L \\ \frac{P_R}{P_L} = \frac{S_R}{S_L} \text{ (équilibre de la pression)} \end{array} \right. \end{aligned}$$

9.3.3 Matrice de Roe associée $A_{\Phi}(V_L, V_R)$

Les expressions (9.22) sont compliquées et il sera pénible de les mettre en application dans un schéma de Godunov. On va plutôt les exploiter grâce à la linéarisation de type Roe généralisée construite au paragraphe précédent. Soit A la linéarisée des équations d'Euler homogènes [73], elle est telle que:

$$A(U_L, U_R) \cdot (U_R - U_L) = F(U_R) - F(U_L)$$

Pour obtenir l'expression de $A_{\Phi}(V_L, V_R)$ associée à (9.20), il suffit donc de calculer la moyenne approchée $G_{\Phi}(V_L, V_R)$ du terme source par la formule (9.16). D'autre part, la proposition III.6 fournit une expression du rapport P/S de part et d'autre de chacune des interfaces $x_{j+\frac{1}{2}}$ du maillage. On déduit donc que pour le problème (9.19), (9.20), on a:

$$\begin{cases} G_{\Phi}(V_L, V_R) = \frac{-P(V_L)}{2S_L} \left(1 + \left(\frac{\varphi(R_R)}{\varphi(R_L)} \right)^{\frac{-\gamma}{\gamma-1}} \right) \\ \text{avec } \varphi(R) = \frac{R^2}{(R+a)(R-a)} \end{cases}$$

Voilà l'expression du terme supplémentaire présent dans $A_{\Phi}(V_L, V_R)$, apte à équilibrer l'action du terme de variation de section aux interfaces par la biais du système (9.20) homogène. Il reste maintenant à calculer la matrice de passage de A_{Φ} ainsi que son inverse pour construire les flux numériques.

9.3.4 Expression des flux numériques

Soient (V_L, V_R) donnés et A_{Φ} la linéarisée qui s'en déduit; on note \vec{R} la matrice de passage et \vec{L} son inverse. Pour les calculer, il est d'abord nécessaire d'avoir les valeurs propres de A_{Φ} . On introduit aussi les notations suivantes où $h = i + \frac{p}{\rho}$ désigne l'enthalpie spécifique:

$$\bar{u} = \frac{\sqrt{R_L}u_L + \sqrt{R_R}u_R}{\sqrt{R_L} + \sqrt{R_R}} \quad \text{et} \quad \bar{h} = \frac{\sqrt{R_L}h_L + \sqrt{R_R}h_R}{\sqrt{R_L} + \sqrt{R_R}}$$

$$A_{\Phi}(V_L, V_R) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ (\gamma-3)\frac{\bar{u}^2}{2} & (3-\gamma)\bar{u} & (\gamma-1) & G_{\Phi}(V_L, V_R) \\ (\gamma-3)\frac{\bar{u}^2}{2} - \bar{h}\bar{u} & \bar{h} - (\gamma-1)\bar{u}^2 & \gamma\bar{u} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

La diagonalisée est $\Lambda_{\Phi} = \{\bar{u} + a, \bar{u} - a, \bar{u}, 0\}$ où $a = \sqrt{(\gamma-1)(\bar{h} - \frac{\bar{u}^2}{2})}$ est la vitesse du son dans le milieu considéré. La matrice de passage R s'écrit:

$$\vec{R}(V_L, V_R) = \begin{pmatrix} 1 & 1 & 1 & G_{\Phi}(V_L, V_R) \\ \bar{u} + a & \bar{u} - a & \bar{u} & 0 \\ \bar{h} + \bar{u}a & \bar{h} - \bar{u}a & \frac{\bar{u}^2}{2} & \frac{G_{\Phi}(V_L, V_R)}{\gamma} \left(\bar{h} - (\gamma-1)\frac{\bar{u}^2}{2} \right) \\ 0 & 0 & 0 & \underbrace{\frac{\bar{u}^2}{2} \frac{(\gamma+1)}{\gamma} - \frac{(\gamma-1)\bar{h}}{\gamma}}_K \end{pmatrix} \quad (9.23)$$

Il est intéressant de remarquer que le dernier vecteur propre est associé à la valeur propre 0. Il n'a donc aucune incidence sur le schéma et son unique fonction est de permettre le calcul de la matrice inverse \vec{L} .

$$\vec{L}(V_L, V_R) = \frac{1}{2a^2} \begin{pmatrix} (\gamma-1)\frac{\bar{u}^2}{2} - \bar{u}a & a - (\gamma-1)\bar{u} & (\gamma-1) & \frac{G_{\Phi}(V_L, V_R)}{\gamma K} \left(\gamma\bar{u}a - (\gamma-1) \left(\bar{h} + \frac{\bar{u}^2}{2} \right) \right) \\ (\gamma-1)\frac{\bar{u}^2}{2} + \bar{u}a & -a - (\gamma-1)\bar{u} & (\gamma-1) & \frac{-G_{\Phi}(V_L, V_R)}{\gamma K} \left(\gamma\bar{u}a + (\gamma-1) \left(\bar{h} + \frac{\bar{u}^2}{2} \right) \right) \\ 2a^2 - (\gamma-1)\bar{u}^2 & 2\bar{u}(\gamma-1) & -2(\gamma-1) & \frac{2G_{\Phi}(V_L, V_R)}{\gamma K} \left(-\gamma a^2 + (\gamma-1) \left(\bar{h} + \frac{\bar{u}^2}{2} \right) \right) \\ 0 & 0 & 0 & \frac{2a^2}{K} \end{pmatrix} \quad (9.24)$$

9.3.5 Essai numérique: la tuyère supersonique divergente

Le test de validation est effectué sur une tuyère supersonique divergente de contour illustré en figure 9.5, proposée par [27, 87]. Le schéma-témoin est celui proposé par P. Glaister [32]. Ce dernier met en oeuvre un décentrement des termes sources selon la base propre associée à la matrice de Roe homogène. On constate une très bonne adéquation des courbes numériques, ce qui valide donc la présente approche non-conservative linéarisée.

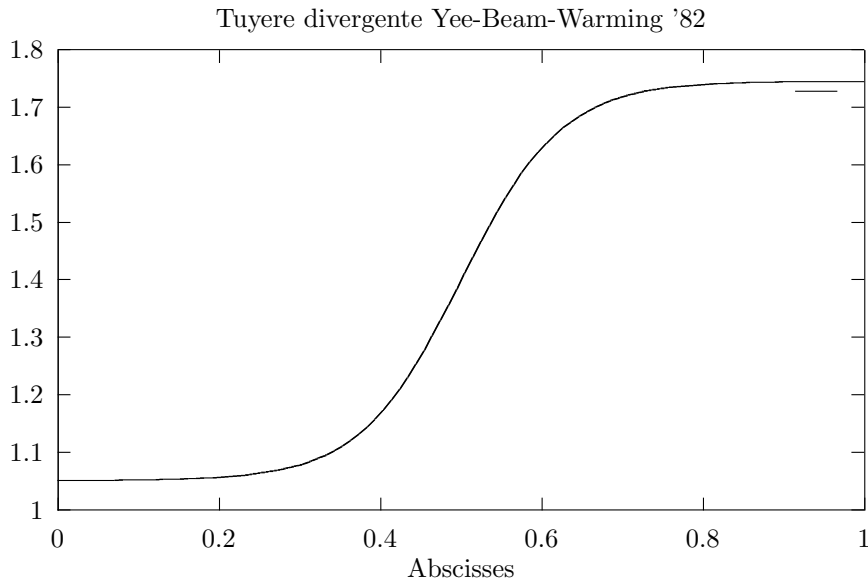


Figure 9.5: Variation de section pour la tuyère supersonique

L'équation du profil est: $S(x) = 1.398 + 0.347 \tanh(8x - 4)$, $x \in [0, 1]$. Les données initiales dans la tuyère sont celles d'un gaz au repos:

$$\rho_0 = 0.125, \quad u = 0, \quad e_0 = i_0 = 2.5$$

Par contre, on impose la condition supersonique à l'extrémité gauche:

$$\begin{cases} \rho_l = 0.502 \\ u_l = 1.299 \\ p_l = 0.381 \end{cases}$$

L'écoulement est fortement instationnaire, jusqu'à l'apparition d'un régime permanent vers l'instant $t \simeq 2$. Les deux schémas concordent et sont aussi en accord avec [27]. Les itérations en temps sont présentées dans les figures 9.6, 9.7, 9.8. On note à propos de cette dernière figure que les débits massiques issus des deux schémas sont bien constants au régime stationnaire. Les paramètres choisis pour les deux schémas sont identiques à savoir: $\Delta x = 0.02$, $\Delta t = 0.007$, ce qui donne une CFL maximale de 0.925. On présente aussi une coupe en $t = 2$ afin de comparer précisément les résultats fournis par le présent schéma et celui de P. Glaister. On constate sur les figures (9.9) et (9.10) que les courbes au permanent sont difficiles à distinguer.

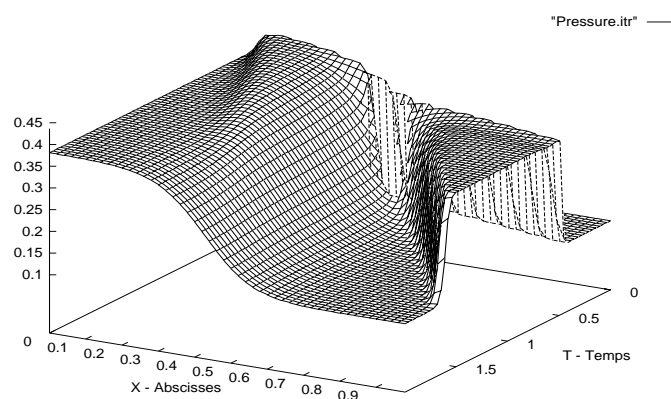


Figure 9.6: Pression (schéma non-conservatif)

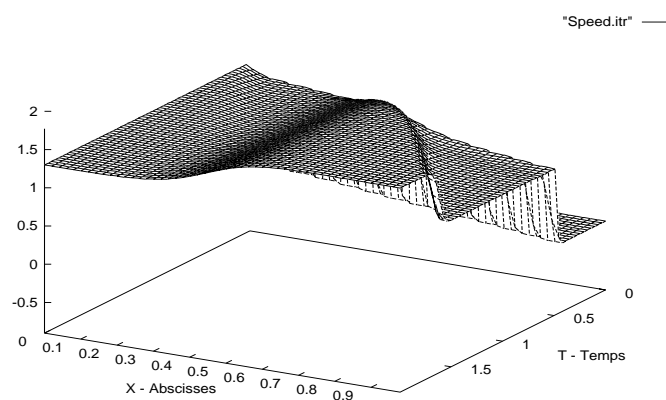


Figure 9.7: Vitesse (schéma non-conservatif)

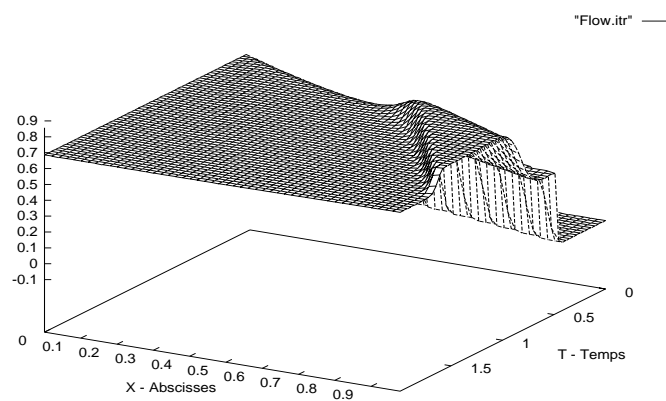


Figure 9.8: Débit (schéma non-conservatif)

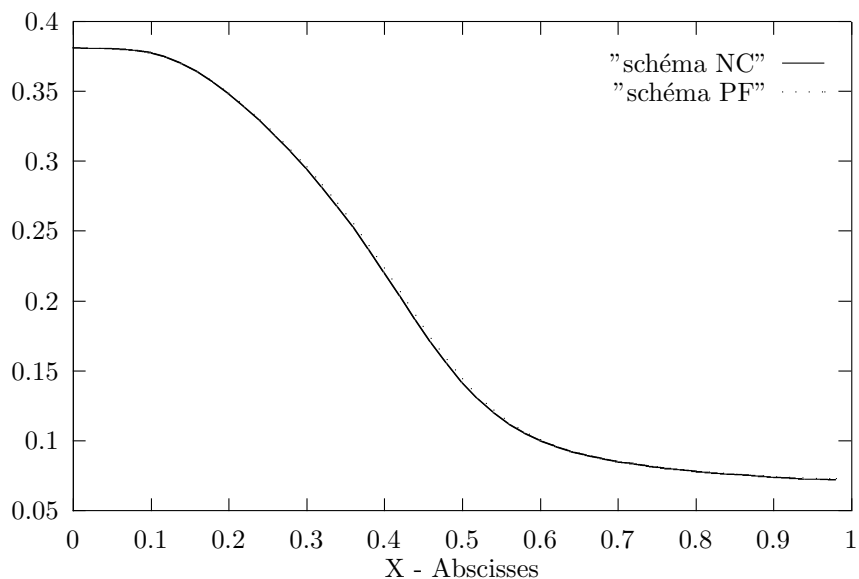


Figure 9.9: Comparaison des pressions stationnaires (schéma non-conservatif / pas-fractionnaires)

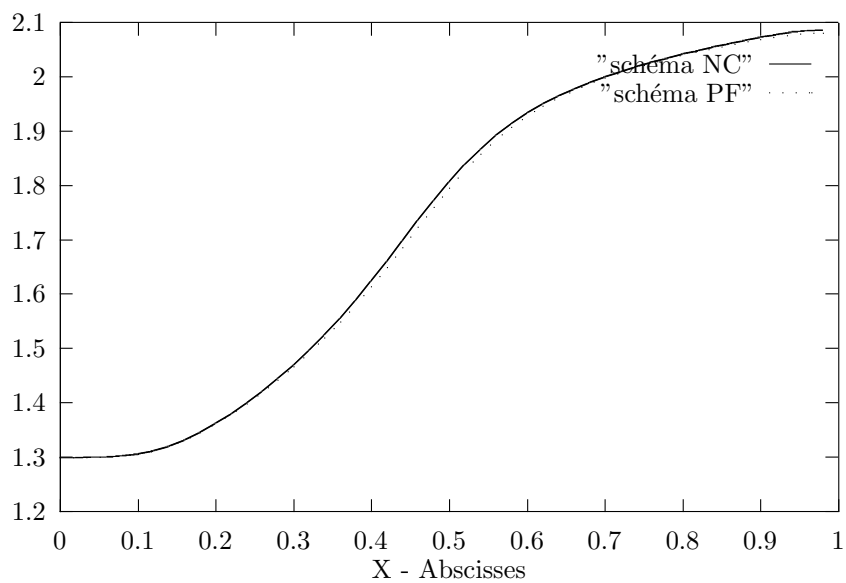


Figure 9.10: Comparaison des vitesses stationnaires (schéma non-conservatif / pas-fractionnaires)

9.4 Second cas-test: pertes de charge singulières

9.4.1 Modélisation proposée

Il s'agit maintenant d'examiner un écoulement de gaz parfait à travers un réseau de grilles à aubes. Le système utilisé est encore celui des équations d'Euler avec un terme source de type frottement :

$$\begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho u)_t + (\rho u^2 + p)_x = -k(x)\rho u|u| \\ (\rho e)_t + (u(\rho e + p))_x = 0 \\ x \in [0, 1] \end{cases} \quad (9.25)$$

Les notations employées sont identiques au cas précédent (9.19).

$$U = (\rho, \rho u, \rho e) \quad F(U) = (\rho u, \rho u^2 + p, u(\rho e + p)) \quad G(U) = (0, -\rho u|u|, 0)$$

9.4.2 Relations de saut et quantités invariantes

Introduisons maintenant une variable dite de raideur :

$$K(x) = \int_0^x k(r) dr \quad (9.26)$$

Le système non-homogène (9.25) se réécrit maintenant sous une forme homogène, mais non-conservative:

$$\begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho u)_t + (\rho u^2 + p)_x + \operatorname{sgn}(\rho u) \frac{(\rho u)^2}{\rho} K_x = 0 \\ (\rho e)_t + (u(\rho e + p))_x = 0 \\ K_t = 0 \end{cases} \quad (9.27)$$

Là encore, cette variable supplémentaire induit un champ linéairement dégénéré, matérialisé par un saut à l'interface $x = 0$. Dans le formalisme de [21], les conditions de Rankine-Hugoniot le long de cette discontinuité stationnaire s'écrivent:

$$\begin{aligned} [\rho u] &= 0 \\ (\rho u)[h] &= 0 \\ [\rho u^2 + p] + \operatorname{sgn}(\rho u) \left[\frac{(\rho u)^2}{\rho} K_x \right]_{\Phi} &= 0 \end{aligned}$$

Celles-ci imposent déjà au débit ρu et à l'enthalpie totale h de rester constants à travers les grilles. Pour expliciter la dernière condition, il est nécessaire d'introduire le chemin (9.6):

$$\begin{aligned} V &= (\rho, \rho u, \rho e, K) \\ \Phi(s; V_L, V_R) &= \begin{pmatrix} \tilde{U}(s\Delta x) \\ \tilde{K}(s\Delta x) \end{pmatrix} \\ \text{où } \tilde{U} &\text{ est une courbe intégrale de } F(\tilde{U})_x = G(\tilde{U})\tilde{K}_x; \tilde{U}(0) = U_L \\ \text{et } \tilde{K}(x) &= K_L + (K_R - K_L)x/\Delta x \end{aligned}$$

Il va encore être profitable de réécrire la pression en fonction de ρ et des deux autres quantités invariantes ρu et h :

$$p(\rho) = \frac{\gamma - 1}{\gamma} \rho \left(h - \frac{(\rho u)^2}{2\rho^2} \right)$$

Le système à résoudre est donc:

$$\begin{cases} (\rho u)_x = 0 \\ (\rho u)^2 \left[\frac{1}{\rho} + p \right]_x + \operatorname{sgn}(\rho u) \frac{(\rho u)^2}{\rho} K_x = 0 \\ (\rho u)h_x = 0 \end{cases} \quad (9.28)$$

La deuxième équation de (9.28) s'écrit maintenant:

$$(\rho u)^2 \left(\frac{-\rho_x}{\rho} \right) + \rho p_x + \operatorname{sgn}(\rho u) (\rho u)^2 K_x = 0$$

Définissons une fonction $\rho \mapsto P(\rho)$ telle que $P_\rho = \rho p_\rho$. L'équation au premier ordre peut alors s'intégrer simplement sous la forme:

$$\ln \left(\frac{\rho_L}{\rho(x)} \right) = - \left(\frac{P(x) - P_L}{(\rho u)^2} \right) - \operatorname{sgn}(\rho u) (K(x) - K_L)$$

Dans ce cas, on peut définir la fonction φ :

$$\varphi(\rho, K) = \mathbf{e}^{\left(\frac{P(\rho)}{(\rho u)^2} + \operatorname{sgn}(\rho u) K \right)}$$

et on obtient:

$$\frac{\varphi(\rho, K)}{\rho}(x) = \frac{\varphi(\rho, K)}{\rho}(0)$$

Proposition III.7 *Le long du contact $\sigma = 0$, les relations suivantes sont vérifiées:*

$$\begin{cases} (\rho u)_L = (\rho u)_R & (\text{conservation du débit}) \\ h_L = h_R & (\text{conservation de l'enthalpie totale}) \\ \frac{1}{\rho_R} = \frac{1}{\rho_L} \mathbf{e}^{-\left(\frac{P_R - P_L}{(\rho u)^2} + \operatorname{sgn}(\rho u) (K_R - K_L) \right)} & (\text{équilibre de la pression}) \end{cases} \quad (9.29)$$

9.4.3 Matrice de Roe associée $A_\Phi(V_L, V_R)$

Ici encore, c'est par le biais d'une linéarisée de Roe consistante avec (9.6) que nous allons exploiter les conditions (9.29). D'une façon rigoureusement identique au cas-test précédent, nous introduisons le terme $G_\Phi(V_L, V_R)$ dans la matrice A_Φ ; celui-ci se calcule à partir de la proposition précédente:

$$G_\Phi(V_L, V_R) = \frac{(\rho u)_L^2}{2\rho_L} \left(1 + \mathbf{e}^{-\left(\frac{P_R - P_L}{(\rho u)_L^2} + \operatorname{sgn}(\rho u) (K_R - K_L) \right)} \right)$$

On obtient au bout du compte la linéarisée de Roe des équations d'Euler avec un terme supplémentaire représentant une moyenne de la densité destinée à équilibrer les flux numériques en respectant les relations (9.29). Pour le calcul de ces flux, les expressions de \vec{R} et \vec{L} sont strictement identiques à celles présentées en (9.23) et (9.24).

9.4.4 Résultats numériques

Le cas-test comprend la résolution des systèmes (9.25), (9.27) respectivement par le schéma décentré proposé par P. Glaister [32] et le schéma NC associé à A_Φ . Le critère déterminant pour la validation sera de vérifier la consistance avec les relations (9.29). La donnée initiale est celle d'un écoulement uniforme subsonique:

$$\rho_0 = 0.125 \quad u_0 = 0.5 \quad i_0 = 2.5$$

Quant aux grilles, elles sont représentées par la fonction $k(x)$, i.e.

$$k(x) = \begin{cases} 10 & \text{pour } x = 0.25, x = 0.5, x = 0.75 \\ 0 & \text{ailleurs} \end{cases}$$

Les figures 9.11 et 9.12 représentent respectivement les itérations en temps du schéma «équilibre» pour la vitesse $u(x, t)$ et la pression $p(x, t)$ avec $(x, t) \in [0, 1] \times [0, 1.5]$. Le schéma employé est celui décrit dans la section précédente. On observe une bonne stabilité et une solution non-oscillante.

C'est aussi le cas pour le schéma décentré plus classique [32] utilisé en témoin; néanmoins, un examen attentif de son comportement va révéler certaines faiblesses. En effet, si on trace les trois quantités de (9.29), à savoir

le débit, l'enthalpie totale et la fonction de saut, on peut deviner l'emplacement des grilles sur la solution stationnaire issue du schéma classique alors que ces quantités devraient être constantes (voir les figures 9.13, 9.14 et 9.15). Ce qui n'est pas le cas pour le schéma non-conservatif linéarisé qui respecte bien les conditions de saut (9.29).

On a choisi les paramètres $\Delta x = 0.02$, $\Delta t = 0.012$ pour les deux schémas, ce qui donne une CFL maximale de 0.98. On présente en outre deux coupes en $t = 1.5$ qui permettent une comparaison précise entre la solution stationnaire issue du présent schéma et celle fournie par celui proposé par P. Glaister. On constate des écarts croissants avec le nombre de grilles sur la pression et des overshoots sur la vitesse: voir les figures (9.16) et (9.17). Il semble que la présence de ceux-ci soient liée à l'étalement excessif des discontinuités de pression à l'emplacement des grilles. On peut donc conclure à la meilleure qualité des approximations fournies par le schéma équilibré.

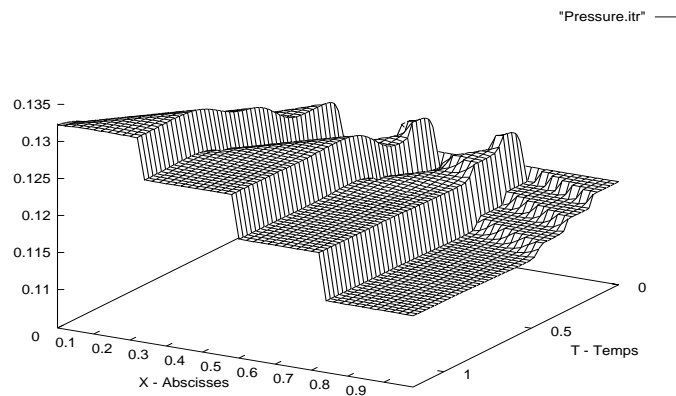


Figure 9.11: Pression (schéma non-conservatif)

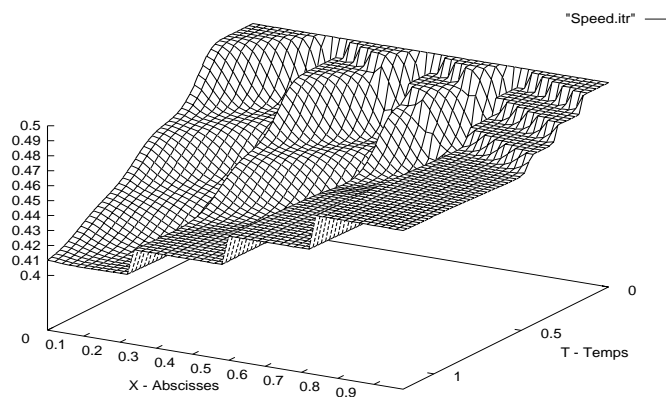


Figure 9.12: Vitesse (schéma non-conservatif)

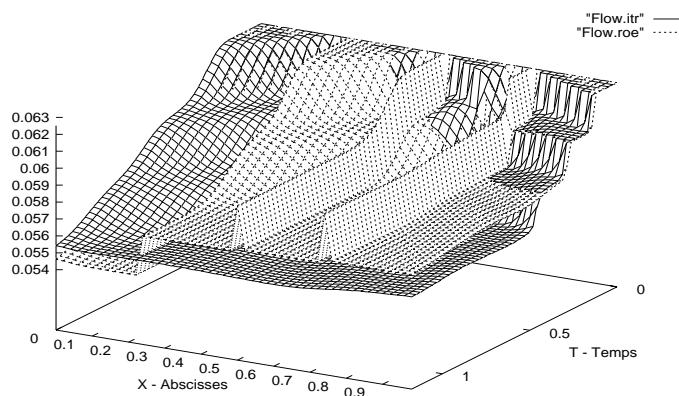


Figure 9.13: Débit (schéma non-conservatif / pas-fractionnaires)

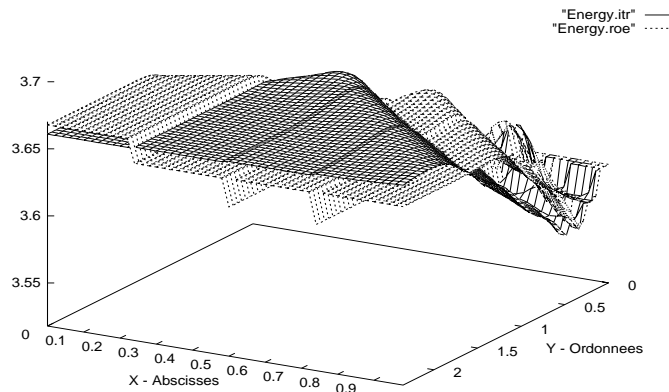


Figure 9.14: Enthalpie totale (schéma non-conservatif / pas-fractionnaires)

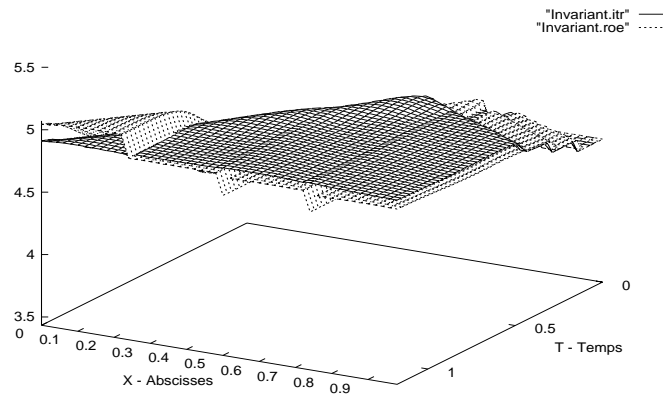


Figure 9.15: Quantité invariante (schéma non-conservatif / pas-fractionnaires)

9.5 Conclusion

On a construit dans ce chapitre un nouveau schéma de type Roe non-conservatif pour la résolution numérique des systèmes d'équations hyperboliques avec termes sources. Cette construction s'appuie sur une reformulation non-conservative du second membre que l'on traite en premier lieu grâce au formalisme de [21], puis au niveau numérique en suivant les idées de [24, 82]. Il est alors nécessaire d'effectuer une moyenne sophistiquée du terme source afin de respecter les états d'équilibres réalistes du système d'origine. On a présenté une façon d'en obtenir une sans avoir à résoudre le problème de Riemann non-conservatif sous-jacent. Malheureusement, cette construction est limitée au cas d'un seul terme source; voir cependant [18] pour un autre calcul adapté aux équations shallow-water avec terme de relief.

Apports spécifiques liés à ce nouveau schéma L'objectif de celui-ci est d'assurer que les régimes permanents numériques soient des approximations correctes des stationnaires théoriques. Ceci n'est pas une propriété immédiate puisque les schémas à pas fractionnaires appliquant un décentrement classique ne la vérifient pas. En effet, il n'y a aucune raison pour qu'une technique classique fournisse un régime stationnaire discret cohérent avec les profils théoriques. La première contribution pour remédier à cela est [11] dans laquelle une formulation homogène *conservative* équivalente au système de l'atmosphère avec gravité est mise en place et un schéma de

Godunov s'en déduit. Cette technique est malheureusement limitée à des cas particuliers et le schéma obtenu est coûteux. On présente donc ici un moyen de traiter des cas plus généraux (moins de restrictions sur le terme source) avec un schéma plus simple. Les résultats numériques obtenus sont encourageants et montrent en particulier le bien-fondé de la moyenne approchée (9.16).

Limites de cette approche Ce schéma est par construction limité dans deux directions pour le moins. La première est relative au fait que la moyenne du terme source est calculée de façon approchée, et que cette technique est en fait limitée aux problèmes «presque scalaires». En effet, pour l'intégration du système à N équations $F(U)_x = G(U)$, on suppose l'existence de $(N - 1)$ quantités conservées afin d'obtenir une solution par quadratures. D'autre part, l'expression du schéma (9.17) montre que les termes sources sont traités de façon explicite, ce qui ne manquera pas de créer des instabilités numériques en régime raide. Le remède habituel à ce genre de problème est d'impliciter le schéma: ceci a été testé dans [18] avec succès. Néanmoins, ceci amène des limitations de pas de temps similaires à celles étudiées pour les problèmes scalaires dans la partie II. Il paraît donc plus judicieux, au vu des résultats obtenus par l'approche «décomposition de flux», d'unir la formulation non-conservative présentée ici pour le calcul des états aux interfaces avec le cadre robuste offert par ce type de schéma non-linéaire.

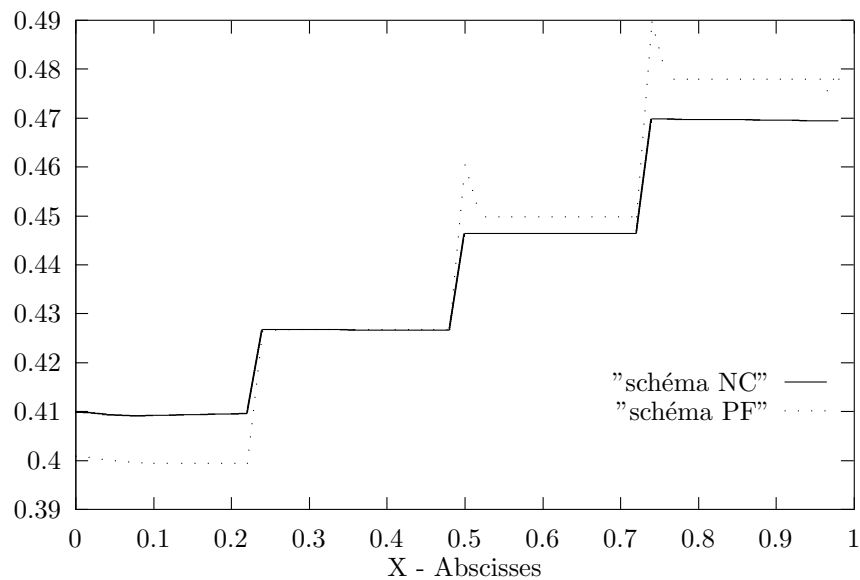


Figure 9.16: Comparaison des pressions stationnaires (schéma non-conservatif / pas-fractionnaires)

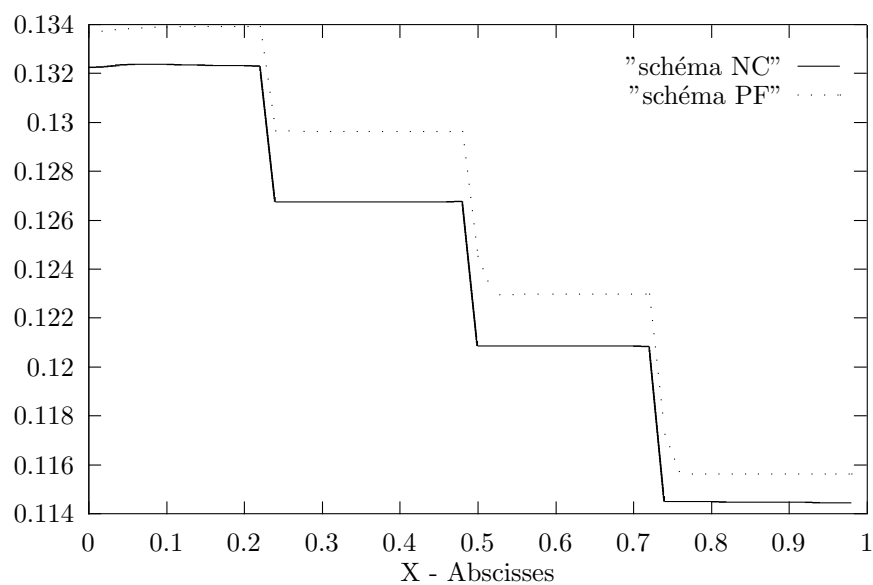


Figure 9.17: Comparaison des vitesses stationnaires (schéma non-conservatif / pas-fractionnaires)

Partie IV

Développement d'un schéma à décomposition de flux «équilibre»

Résumé

Dans cette dernière partie, on présente en détail l'application des idées précédemment exposées à une classe assez large de systèmes hyperboliques avec des termes sources variés allant des équations d'Euler quasi-monodimensionnelles aux problèmes diphasiques 1D ou 2D. Bien qu'il soit évident que le solveur le plus adapté aux objectifs d'une discrétisation "équilibre" soit celui de Godunov [36], il est malheureusement hors de portée d'en faire un usage courant du fait de sa complexité décourageante. On propose donc de s'en tenir à l'alternative d'une décomposition de flux de VanLeer [83] alliée aux formulations non-conservatives. Ceci pour plusieurs raisons:

- on préserve ainsi une interprétation en termes de solveurs de Riemann élémentaires, ce qui permet d'assurer des propriétés de positivité ou des principes du maximum.
- les termes sources sont traités de façon très robuste par les relations de saut généralisées.
- l'implémentation reste simple et un tel code fournit des solutions réalistes.

En conséquence, on présente un schéma "équilibre" sur la base de cette décomposition et on détaille le traitement original des termes sources. Ceci est appliqué à des écoulements de gaz parfait dans des géométries diverses et à des modèles diphasiques réactifs simples. Le cas des termes de relaxation est aussi traité; on vérifie à ce propos la cohérence des approximations avec les résultats prévus par la théorie de G.Q. Chen, C.D. Levermore et T.P. Liu [14]. Il existe toutefois une pierre dans ce jardin à savoir l'éventualité d'une situation de résonance lors de laquelle la construction proposée devient caduque. Dans ce contexte, on s'intéresse au problème modèle posé par un choc sonique dans une tuyère divergente: on construit la solution stationnaire exacte et on teste les schémas "en aveugle". Enfin, le dernier pan de ce travail consiste à proposer une extension bidimensionnelle de ces techniques. Il est bien évidemment hors d'atteinte d'explicitier des profils asymptotiques 2D, c'est pourquoi on met en oeuvre des "volumes finis". La construction d'un schéma scalaire est présentée ainsi qu'une extension possible pour des systèmes plus généraux. Ceci ouvre la voie des tests sur un modèle diphasique réactif qui clôt ce travail.

Chapitre 10

Méthodes de décomposition de flux et produits non-conservatifs

Encouragés par les bons résultats de l'approche linéarisée de type Roe non-conservatif pour des cas où l'équilibre correct entre flux convectifs et termes sources non-linéaires est délicat, nous allons tenter de poursuivre le développement de ce type de schémas. L'idée de base reste identique: réécrire les termes sources sous forme de produits non-conservatifs afin de les décentrer au mieux. Néanmoins, il s'agit de pallier à un inconvénient inhérent à l'approche linéarisée: le manque de robustesse en régime fortement réactionnel. Cela revient donc à chercher une meilleure généralisation du schéma scalaire telle que le pas de temps reste indépendant de la grandeur du terme source. Pour cela, il faut de nouveau le traiter au moyen d'une condition de saut (issu de la formulation du chemin régularisant), et par conséquent, préserver à toute force une interprétation en termes de problèmes de Riemann les plus simples possibles. Il est donc logique de revenir aux idées du chapitre 2 de la partie précédente en mettant en oeuvre une décomposition de flux plus sophistiquée qui s'appuie sur le formalisme du chemin régularisant.

10.1 Schémas à l'équilibre et décomposition de flux

Par le biais de ce type de schéma, on va introduire des problèmes de Riemann pour lesquels la convection ne s'effectue que dans un unique sens. Le produit non-conservatif traduisant l'action localisée du terme de réaction va donc amener une relation de saut en bordure des mailles $\Delta x \times \Delta t$ que l'on va pouvoir traiter à la main. Dans les cas où il sera possible d'obtenir les états à l'interface en forme fermée, on aura une approximation très robuste pour laquelle le pas de temps ne sera prescrit que par les termes convectifs.

10.1.1 Introduction des termes non-conservatifs

Soit le problème:

$$\begin{cases} U_t + F(U)_x = G(U) \text{ avec } U \in \mathbb{R}^N \\ U_0 \in [BV(\mathbb{R})]^N \end{cases} \quad (10.1)$$

Sans hypothèses particulières sur la fonction de flux $U \mapsto F(U)$ (hormis la régularité: [77]), il est possible de définir une décomposition de flux du type:

$$\forall U \in \mathbb{R}^N, F(U) = F^+(U) + F^-(U), \pm \Lambda(dF^\pm(U)) \subset \mathbb{R}^+$$

Au moyen de la fonction $a : \mathbb{R} \rightarrow \mathbb{R}$ définie par $a(x) = x$, le système (10.1) peut alors se réécrire sous la forme:

$$\begin{cases} U_t + F^+(U)_x + F^-(U)_x - G(U)a_x = 0 \\ a_t = 0 \\ U_0 \in [BV(\mathbb{R})]^N \end{cases} \quad (10.2)$$

Introduisons un maillage spatial $(x_j)_{j \in \mathbb{Z}}$ tel que $x_j = j\Delta x$: il reste maintenant à remplacer la fonction régulière $x \mapsto a(x) = x$ par $a^{\Delta x} \in BV(\mathbb{R})$ définie par:

$$a^{\Delta x}(x) = x_j \text{ pour } x \in \left[\left(j - \frac{1}{2} \right) \Delta x, \left(j + \frac{1}{2} \right) \Delta x \right] \quad (10.3)$$

On obtient le système vraiment non-conservatif:

$$\begin{cases} U_t + F^+(U)_x + F^-(U)_x - G(U)a_x^{\Delta x} = 0 \\ a_t^{\Delta x} = 0 \\ U_0 \in [BV(\mathbb{R})]^N \end{cases} \quad (10.4)$$

Posons $V = (U, a^{\Delta x})$, on voudrait régulariser le système (10.4) à l'aide d'une variante du chemin à l'équilibre découvert au chapitre 3 de la partie III: (on fera encore l'abus de langage habituel consistant à n'en expliciter qu'une seule branche)

$$\Phi \left(s; \begin{pmatrix} U_L \\ a_L^{\Delta x} \end{pmatrix}, \begin{pmatrix} U_R \\ a_R^{\Delta x} \end{pmatrix} \right) = \begin{pmatrix} \tilde{U}(s\Delta x) \\ \tilde{a}(s\Delta x) \end{pmatrix} \quad (10.5)$$

où \tilde{U} est une courbe intégrale de $F(\tilde{U})_x = G(\tilde{U})\tilde{a}_x$; $\tilde{U}(0) = U_L$
et $\tilde{a}(x) = a_L^{\Delta x} + (a_R^{\Delta x} - a_L^{\Delta x})x/\Delta x$

10.1.2 Ecriture du schéma proposé

A partir de (10.4), il est possible d'écrire un schéma en faisant intervenir les quantités:

$$\begin{cases} U^+ = U_j^n - \frac{2\Delta t}{\Delta x} (F^+(U_j^n) - F^+(U_{j-\frac{1}{2}}^+)) \\ U^- = U_j^n - \frac{2\Delta t}{\Delta x} (F^-(U_{j+\frac{1}{2}}^-) - F^-(U_j^n)) \end{cases} \quad (10.6)$$

où $U_{j-\frac{1}{2}}^+$ et $U_{j+\frac{1}{2}}^-$ sont obtenus par résolution de conditions de saut généralisées susceptibles de maintenir constants les discrétisations des courbes à l'équilibre de (10.1). Donc, sous la condition CFL:

$$\sup_{j,n} |\lambda^\pm(U_j^n)| \frac{\Delta t}{\Delta x} \leq \frac{1}{2} \quad (10.7)$$

on obtient une valeur pour U_j^{n+1} en faisant la moyenne des deux demi-maillages (10.6):

$$U_j^{n+1} = \frac{U^+ + U^-}{2} = U_j^n - \frac{\Delta t}{\Delta x} [F^+(U_j^n) - F^+(U_{j-\frac{1}{2}}^+) + F^-(U_{j+\frac{1}{2}}^-) - F^-(U_j^n)] \quad (10.8)$$

Cette expression est identique à (8.17) au chapitre 2 de la partie III. La différence vient de la manière avec laquelle on va maintenant calculer les valeurs aux interfaces $U_{j\pm\frac{1}{2}}^\pm$. Par construction, ces vecteurs sont calculés par intégration des courbes intégrales suivantes:

$$\begin{aligned} U_{j-\frac{1}{2}}^+ &= \tilde{U}(\Delta x) \text{ où } \begin{cases} F(\tilde{U})_x = G(\tilde{U})\tilde{a}_x \\ \tilde{U}(0) = U_{j-1}^n \end{cases} \\ U_{j+\frac{1}{2}}^- &= \tilde{U}(-\Delta x) \text{ où } \begin{cases} F(\tilde{U})_x = G(\tilde{U})\tilde{a}_x \\ \tilde{U}(0) = U_{j+1}^n \end{cases} \end{aligned} \quad (10.9)$$

Si ces relations peuvent être traitées explicitement en forme fermée (c'est ce qui se passe dans le cas scalaire au moyen de la fonction $u \mapsto \phi(u)$ qui permet d'écrire (10.9) sous la forme $\phi(u)u_x = \tilde{a}_x$), alors on obtient un schéma robuste par construction puisque le terme en $G(U)$ n'amène à priori aucune restriction sur le maillage.

10.1.3 Interprétation en problèmes de Riemann élémentaires

Il s'agit donc de formuler une régularisation pour le problème (10.4) au niveau des systèmes décomposés à l'aide des fonctions de flux F^\pm . A priori, lorsque $F^\pm \neq F$ (régime subsonique), la relation de saut stationnaire $[F^\pm(U)] = [G(U)a_x^{\Delta x}]_\Phi$ n'est pas correcte puisqu'elle ne préserve pas les équilibres discrets de (10.1). On propose donc de chercher des fonctions G^\pm donnant leur sens aux problèmes i,j :

$$\begin{cases} U_t + F^+(U)_x - [G^+(U)a_x^{\Delta x}]_\Phi = 0 \\ a_t^{\Delta x} = 0 \end{cases} \quad \text{et} \quad \begin{cases} U_t + F^-(U)_x - [G^-(U)a_x^{\Delta x}]_\Phi = 0 \\ a_t^{\Delta x} = 0 \end{cases} \quad (10.10)$$

et susceptibles de fournir un schéma respectueux des équilibres corrects de (10.1).

Théorème IV.1 *A F^\pm différentiables donnés, si $U \mapsto dF(U)^{-1}G(U)$ est localement Lipschitzienne, il existe un unique couple de fonctions régulières (G^+, G^-) telles que:*

$$\begin{cases} \forall U \in \mathbb{R}^N, G^+(U) + G^-(U) = G(U) \\ [F(U)] = [G(U)a_x^{\Delta x}]_\Phi \text{ implique } [F^+(U)] = [G^+(U)a_x^{\Delta x}]_\Phi \\ [F(U)] = [G(U)a_x^{\Delta x}]_\Phi \text{ implique } [F^-(U)] = [G^-(U)a_x^{\Delta x}]_\Phi \end{cases} \quad (10.11)$$

avec Φ désignant le chemin i,j équilibre i,j (10.5). Celles-ci sont données par:

$$\forall U \in \mathbb{R}^N, G^\pm(U) = dF^\pm(U).dF(U)^{-1}.G(U)$$

Démonstration

On considère la relation stationnaire:

$$[F(U)] = [G(U)a_x^{\Delta x}]_\Phi = \int_0^1 G(\Phi(s; V_L, V_R)) \frac{\partial \Phi}{\partial s}(s; V_L, V_R).ds$$

Comme $x \mapsto a^{\Delta x}(x)$ est régularisée par les lignes droites, on a :

$$[F(U)] = \int_0^1 G(U(s))\Delta x ds = \int_0^{\Delta x} G(U(x))\tilde{a}_x(x)dx$$

avec $x \mapsto U(x)$ la solution de $F(U)_x = G(U)\tilde{a}_x$; $U(0) = U_L$. C'est-à-dire que l'intégrale curviligne le long de Φ revient à suivre le flot de ce système différentiel.

On introduit maintenant le problème modifié $U_t + F^+(U)_x - [G^+(U)a_x^{\Delta x}]_{\Phi} = 0$ où G^+ est à priori différente de G mais capable de préserver les équilibres de la forme :

$$[F(U)] = [G(U)a_x^{\Delta x}]_{\Phi} = \int_0^{\Delta x} G(U(x))\tilde{a}_x(x)dx$$

En dérivant par rapport à Δx et en tenant compte des hypothèses, il vient :

$$F(U)_x = G(U)\tilde{a}_x \text{ d'où } F^+(U)_x = (dF^+ \cdot dF^{-1} \cdot G)(U)\tilde{a}_x \stackrel{def}{=} G^+(U)\tilde{a}_x$$

En intégrant cette équation différentielle entre $x = 0$ et $x = \Delta x$, on trouve :

$$[F^+(U)] = \int_0^{\Delta x} G^+(U(x))\tilde{a}_x(x)dx = \int_0^1 G^+(U(s))\Delta x ds = \int_0^1 G^+(\Phi(s; V_L, V_R)) \frac{\partial \Phi}{\partial s}(s; V_L, V_R) \cdot ds$$

De la même façon, on définit $G^-(U) = (dF^- \cdot dF^{-1} \cdot G)(U)$ et la première relation de (10.11) est immédiate du fait des propriétés de F^{\pm} . Les deux relations suivantes assurent la préservation des équilibres réalistes de (10.1) et sont assurées par construction. \square

Compte-tenu des relations de saut généralisées (C.8), on va donc résoudre puis moyenner dans la demi-maille $[x_{j-\frac{1}{2}}, x_j]$, la solution du problème de Riemann: (avec les notations du chapitre 2 - partie III)

$$\left\{ \begin{array}{l} U_t + F^+(U)_x - [G^+(U)a_x^{\Delta x}]_{\Phi} = 0 \\ a_t^{\Delta x} = 0 \\ U_0 = \begin{cases} U_{j-1}^n & \text{si } x < (j - \frac{1}{2})\Delta x \\ U_j^n & \text{si } x > (j - \frac{1}{2})\Delta x \end{cases} \end{array} \right. \quad \text{c'est-à-dire} \quad \left\{ \begin{array}{l} U_t + F^+(U)_x = 0 \\ U_0 = \begin{cases} U_{j-\frac{1}{2}}^+ & \text{si } x < (j - \frac{1}{2})\Delta x \\ U_j^- & \text{si } x > (j - \frac{1}{2})\Delta x \end{cases} \\ F^+(U_{j-\frac{1}{2}}^+) - F^+(U_{j-1}^n) = [G^+(U)a_x^{\Delta x}]_{\Phi} \end{array} \right. \quad (10.12)$$

Tandis que dans $[x_j, x_{j+\frac{1}{2}}]$:

$$\left\{ \begin{array}{l} U_t + F^-(U)_x - [G^-(U)a_x^{\Delta x}]_{\Phi} = 0 \\ a_t^{\Delta x} = 0 \\ U_0 = \begin{cases} U_j^n & \text{si } x < (j + \frac{1}{2})\Delta x \\ U_{j+1}^n & \text{si } x > (j + \frac{1}{2})\Delta x \end{cases} \end{array} \right. \quad \text{c'est-à-dire} \quad \left\{ \begin{array}{l} U_t + F^-(U)_x = 0 \\ U_0 = \begin{cases} U_j^- & \text{si } x < (j + \frac{1}{2})\Delta x \\ U_{j+\frac{1}{2}}^+ & \text{si } x > (j + \frac{1}{2})\Delta x \end{cases} \\ F^-(U_{j+\frac{1}{2}}^+) - F^-(U_{j+1}^n) = [G^-(U)a_x^{\Delta x}]_{\Phi} \end{array} \right. \quad (10.13)$$

Les problèmes (10.12), (10.13) sont le plus souvent non-strictement hyperboliques puisqu'ils comportent au moins un champ linéairement dégénéré à vitesse nulle. Sous réserve qu'il n'y ait pas d'interactions entre champs VNL et LD (points soniques), ils possèdent une solution dans la classe de Lax [48]. En toute rigueur, il est de mise de calculer ces états au moyen des relations $[F^+(U)] = [G^+(U)a_x^{\Delta x}]_{\Phi}$ et $[F^-(U)] = [G^-(U)a_x^{\Delta x}]_{\Phi}$; néanmoins, du fait de la complexité des systèmes différentiels provenant de l'expression des flux décomposés, on préfère se ramener aux relations mettant en jeu les flux réels.

10.1.4 Propriétés du schéma ainsi obtenu

Respect des équilibres corrects Le schéma (10.8) a été construit en gardant une interprétation en termes de moyennes de solutions exactes de demi-problèmes de Riemann. Par conséquent, il vérifie des propriétés de positivité sur la densité et la pression sous la restriction CFL (10.7). (voir [71] pour l'exemple du splitting de Lax-Friedrichs). On désire maintenant s'assurer que cette construction vérifie aussi une propriété de *stabilité des équilibres* similaire à (9.14) pour s'assurer d'un bon comportement sur des problèmes du type (9.25).

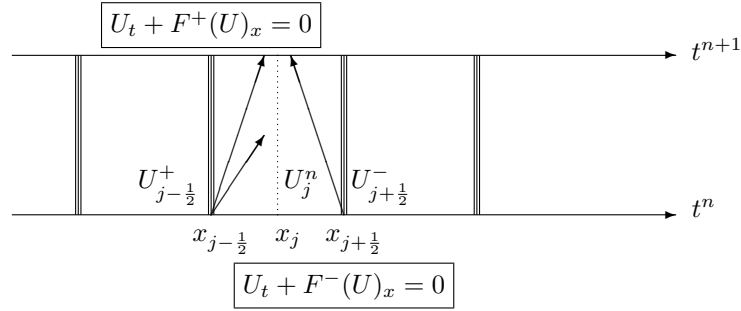


Figure 10.1: Demi-problèmes de Riemann non-conservatifs

Proposition IV.1 *Le schéma (10.8) est jéquilibré*: si les états initiaux $(U_j^n)_j$ sont tels que $F(U_{j+1}^n) - F(U_j^n) = [G(U)a_x^{\Delta x}]_{\Phi}$, alors, pour tout j , on a: $U_j^{n+1} = U_j^n$.

Démonstration

C'est immédiat du fait des relations (10.9): soient $U_{j=1,2,3}$ et $U_{3/2}^+, U_{5/2}^-$ les états aux interfaces, ils s'écrivent:

$$U_{3/2}^+ = \tilde{U}(\Delta x) \text{ où } \begin{cases} F(\tilde{U})_x = G(\tilde{U})\tilde{a}_x \\ \tilde{U}(0) = U_1^n \end{cases}$$

$$U_{5/2}^- = \tilde{U}(-\Delta x) \text{ où } \begin{cases} F(\tilde{U})_x = G(\tilde{U})\tilde{a}_x \\ \tilde{U}(0) = U_3^n \end{cases}$$

D'autre part, on a supposé que $F(U_{j+1}^n) - F(U_j^n) = [G(U)a_x^{\Delta x}]_{\Phi}$, donc si on intègre les équations (10.9) ci-dessus entre $x = 0$ et $x = \Delta x$, on obtient:

$$F(U_{3/2}^+) = F(U_2^n) = F(U_{5/2}^-)$$

Le fait d'introduire une donnée initiale sous la forme $F(U_{j+1}^n) - F(U_j^n) = [G(U)a_x^{\Delta x}]_{\Phi}$ implique que la fonction $U \mapsto dF(U)^{-1}.G(U)$ est Lipschitzienne au voisinage de tous les U_j^n . Par conséquent, la jacobienne des flux est partout inversible, ce qui implique que l'on a par le théorème d'inversion locale:

$$U_{3/2}^+ = U_2^n = U_{5/2}^-$$

En considérant l'expression du schéma donnée en (10.8), ceci entraîne que:

$$U_2^{n+1} = U_2^n - \frac{\Delta t}{\Delta x} \underbrace{[F^+(U_2^n) - F^+(U_{3/2}^+)]}_{=0} + \underbrace{[F^-(U_{5/2}^-) - F^-(U_2^n)]}_{=0}$$

Ce qui signifie que la discrétisation (U_j^n) est invariante par le schéma proposé. \square

Remarque C'est réellement l'extension la plus naturelle du schéma scalaire pour des problèmes où on ne peut plus calculer une fonction de saut $u \mapsto \phi(u)$ (partie III - chapitre 1) et en déduire directement un schéma de type Godunov.

Consistance Il peut aussi être intéressant de voir en quoi le schéma (10.8) est *consistant* avec le problème d'origine (10.1). Pour cela, il est nécessaire d'effectuer des développements limités des conditions de saut (10.9):

$$U_{j-1/2}^+ = V(\Delta x) = U_{j-1} + \Delta x dF(U_{j-1})^{-1}G(U_{j-1}) + o(\Delta x)$$

$$U_{j+1/2}^- = V(-\Delta x) = U_{j+1} - \Delta x dF(U_{j+1})^{-1}G(U_{j+1}) + o(\Delta x)$$

En reportant dans l'expression des flux numériques, il vient:

$$F^+(U_j) - F^+(U_{j-1/2}^+) = F^+(U_j) - F^+(U_{j-1} + \Delta x dF(U_{j-1})^{-1}G(U_{j-1}) + o(\Delta x))$$

$$F^-(U_{j+1/2}^-) - F^-(U_j) = F^-(U_{j+1} - \Delta x dF(U_{j+1})^{-1}G(U_{j+1}) + o(\Delta x)) - F^-(U_j)$$

On développe à l'ordre 1 en Δx , et il vient que:

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} [F^+(U_j^n) - F^+(U_{j-1}^n) + F^-(U_{j+1}^n) - F^-(U_j^n)] \\ + \Delta t \left[\underbrace{dF^+(U_{j-1}^n) dF(U_{j-1}^n)^{-1} G(U_{j-1}^n)}_{G^+(U_{j-1}^n)} + \underbrace{dF^-(U_{j+1}^n) dF(U_{j+1}^n)^{-1} G(U_{j+1}^n)}_{G^-(U_{j+1}^n)} \right] \\ + o(\Delta x)$$

avec $G^+ + G^- = G$ du fait de l'identité $dF^+ + dF^- = dF$ (voir l'expression (10.11)).

Cette expression est à rapprocher de celle du schéma de Roe non-conservatif trouvée à la proposition III.5. Il est donc visible que la contribution du terme de réaction évalué en $U_{j\pm 1}^n$ est répartie approximativement suivant les lignes de propagation de l'information. On ne dispose malheureusement plus du théorème de la moyenne (comme dans le cas scalaire) pour préciser les états en lesquels le terme source est évalué.

10.2 Les écoulements quasi-monodimensionnels

10.2.1 Expression des courbes d'équilibre

On réécrit le problème considéré (9.20): il s'agit du système d'Euler pour un écoulement de gaz parfait dans un conduit à section $x \mapsto S(x)$ régulière et lentement variable. Il se présente sous la forme:

$$\begin{cases} R_t + (Ru)_x = 0 \\ (Ru)_t + (Ru^2 + P)_x - \frac{P}{S} S_x = 0 \\ (Re)_t + (Ruh)_x = 0 \\ S_t = 0 \end{cases} \quad (10.14)$$

avec les variables normalisées $(R, Ru, Re, P) = S(\rho, \rho u, \rho e, p = (\gamma - 1)\rho i = \frac{\gamma-1}{\gamma}\rho(h - \frac{u^2}{2}))$ où (i, h) désignent l'énergie interne et l'enthalpie totale. Les notations sont identiques à celles du chapitre précédent. Le terme $\frac{P}{S} S_x$ modélise l'action du contour sur le mouvement du fluide; cette formulation est très générale. On peut ainsi simuler des profils de tuyères, ou bien des écoulements dans des géométries possédant une symétrie cylindrique ou sphérique. La littérature sur ce sujet est très vaste, on peut citer [6, 32, 33, 34, 35]. On se propose donc d'effectuer des approximations de la solution de (10.14) dans différents contextes à l'aide du schéma introduit dans la section précédente. Pour cela, on mettra en oeuvre la décomposition de VanLeer [83] déjà rencontrée (voir la proposition III.1) et on traitera le terme source de façon non-conservative grâce au chemin régularisant (10.5). Pour cela, il est nécessaire de rappeler l'expression des courbes intégrales du système à l'équilibre établies à la proposition III.6.

Proposition IV.2 *En l'absence de tout point sonique, les courbes à l'équilibre de (10.14) issues d'un état (R_0, u_0, h_0, S_0) s'écrivent:*

$$\begin{cases} (Ru)(x) = (Ru)_0 \\ h(x) = h_0 \\ \left(\frac{\varphi(R(x))}{\varphi(R_0)} \right)^{\frac{\gamma}{\gamma-1}} \frac{P(R(x))}{P(R_0)} = \frac{S(x)}{S_0} \end{cases} \quad (10.15)$$

Soit $a = \frac{Ru_0}{\sqrt{2h_0}}$, la fonction φ est telle que:

$$\begin{aligned} \varphi :]a, +\infty[&\rightarrow \mathbb{R}^+ \\ R &\mapsto \frac{R^2}{(R+a)(R-a)} \end{aligned}$$

Il est donc possible d'écrire un schéma «équilibre» au moyen de la décomposition de flux, des courbes (10.15) et des relations (10.9). Néanmoins, en vue d'implanter un algorithme de Newton pour obtenir les valeurs $U_{j\pm\frac{1}{2}}^\pm$, il convient de réécrire les relations (10.15) sous une forme plus simple.

Lemme IV.1 *Sous les mêmes hypothèses, $R(x)$ vérifie:*

$$\psi(R(x)) = \left(\frac{S(x)}{S_0} \right)^{1-\gamma} \psi(R_0) \quad (10.16)$$

où ψ est la fonction définie par:

$$\begin{aligned} \psi &:]a, +\infty[\rightarrow \mathbb{R}^+ \\ R &\mapsto R^{1-\gamma} - \frac{a^2}{R^{1+\gamma}} \end{aligned}$$

Démonstration

La condition de saut sur R s'écrit:

$$\left(\frac{R^2}{R^2 - a^2} \right)^{\frac{\gamma}{\gamma-1}} P(R) = \frac{S}{S_0} \left(\frac{R_0^2}{R_0^2 - a^2} \right)^{\frac{\gamma}{\gamma-1}} P(R_0)$$

L'expression de la pression des gaz parfaits à Ru et h constants se réécrit sous la forme:

$$P(R) = \frac{\gamma-1}{\gamma} \left(Rh - \frac{(Ru)^2}{2R} \right) = \frac{\gamma-1}{\gamma} \frac{h}{R} (R^2 - a^2)$$

La relation

$$P(R) \cdot \varphi(R)^{\frac{\gamma}{\gamma-1}} = \frac{S}{S_0} \cdot P(R_0) \cdot \varphi(R_0)^{\frac{\gamma}{\gamma-1}}$$

devient donc:

$$\frac{\gamma-1}{\gamma} \cdot h \cdot \frac{R^{\frac{2\gamma}{\gamma-1}}}{R} \cdot \frac{R^2 - a^2}{(R^2 - a^2)^{\frac{\gamma}{\gamma-1}}} = \frac{S}{S_0} \cdot \frac{\gamma-1}{\gamma} \cdot h \cdot \frac{R_0^{\frac{2\gamma}{\gamma-1}}}{R_0} \cdot \frac{R_0^2 - a^2}{(R_0^2 - a^2)^{\frac{\gamma}{\gamma-1}}}$$

Les exposants se calculent selon:

$$\begin{cases} \frac{2\gamma}{\gamma-1} - 1 = \frac{\gamma+1}{\gamma-1} \\ 1 - \frac{\gamma}{\gamma-1} = \frac{-1}{\gamma-1} \end{cases}$$

On obtient donc:

$$\left(\frac{R^{\gamma+1}}{R^2 - a^2} \right)^{\frac{1}{\gamma-1}} = \frac{S}{S_0} \left(\frac{R_0^{\gamma+1}}{R_0^2 - a^2} \right)^{\frac{1}{\gamma-1}}$$

Il reste à élever à la puissance $\gamma-1 > 0$ et à prendre l'inverse ($S \neq 0$ et $R \neq a$). Il vient finalement:

$$R^{2-(\gamma+1)} - \frac{a^2}{R^{\gamma+1}} = \left(\frac{S}{S_0} \right)^{1-\gamma} \left(R_0^{2-(\gamma+1)} - \frac{a^2}{R_0^{\gamma+1}} \right)$$

D'où la définition de ψ et la relation (10.16). \square

Tout comme dans le premier chapitre de la partie III, l'équation (10.16) sous-entend que $R \mapsto \psi(R)$ est une bijection. Ce n'est bien sûr pas le cas au vu de la forme de ψ (voir figure 10.2). Néanmoins, on dispose d'un résultat un peu plus faible.

Lemme IV.2 Soit $R_{sonic} = a\sqrt{\frac{\gamma+1}{\gamma-1}}$ avec $a = \frac{Ru}{\sqrt{2h}}$, les restrictions de ψ :

$$]a, R_{sonic}[\rightarrow]0, \psi(R_{sonic})[\text{ et } [R_{sonic}, +\infty[\rightarrow [\psi(R_{sonic}), 0[$$

réalisent des bijections. De plus, l'état R_{sonic} (à h et Ru fixés) est un point sonique pour (10.14).

Démonstration

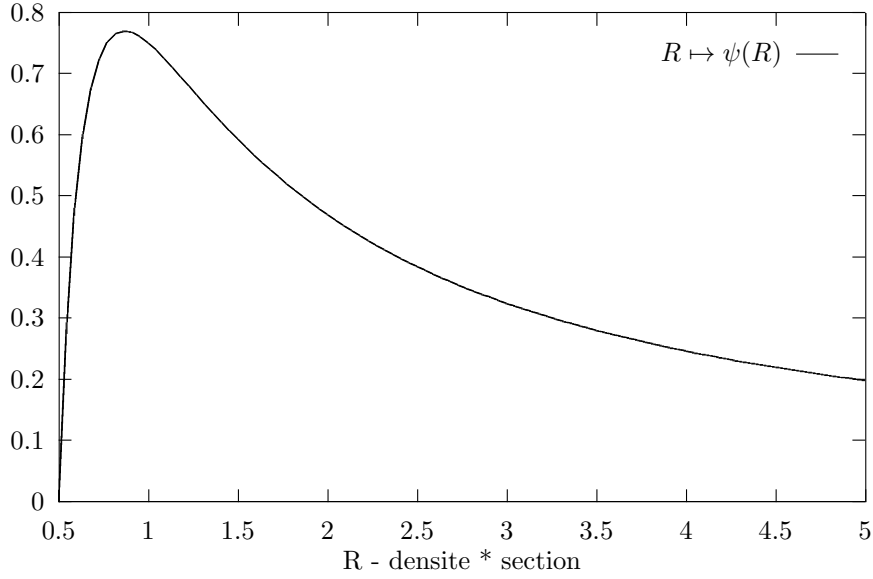
La dérivée de ψ est $\psi'(R) = (1-\gamma)R^{-\gamma} + \frac{(1+\gamma)a^2}{R^{2+\gamma}}$. On peut écrire:

$$\begin{aligned} \psi'(R) &= R^{-\gamma} \left[(1-\gamma) + (1+\gamma) \frac{a^2}{R^2} \right] \\ &= (\gamma+1)R^{-\gamma} \left[\frac{a^2}{R^2} - \frac{\gamma-1}{\gamma+1} \right] \\ &= (\gamma+1)R^{-\gamma} \left[\frac{a}{R} - \sqrt{\frac{\gamma-1}{\gamma+1}} \right] \cdot \left[\frac{a}{R} + \sqrt{\frac{\gamma-1}{\gamma+1}} \right] \end{aligned}$$

Donc $\psi' > 0$ si et seulement si $R < a\sqrt{\frac{\gamma+1}{\gamma-1}}$. Il reste donc à vérifier que l'état $R = a\sqrt{\frac{\gamma+1}{\gamma-1}}$ est un point sonique.

D'après l'expression $\psi'(R) = (1-\gamma)R^{-\gamma} + \frac{(1+\gamma)a^2}{R^{2+\gamma}}$, on remarque que ψ' s'annule en changeant de signe si:

$$\frac{\gamma-1}{\gamma+1} \frac{R^{2+\gamma}}{R^\gamma} = a^2 \text{ d'où } \frac{\gamma-1}{\gamma+1} R^2 = \frac{(Ru)^2}{2h} \text{ et donc } \frac{u^2}{2} = \frac{\gamma-1}{\gamma+1} h$$

Figure 10.2: Allure de la fonction ψ ($a = 0.5$, $\gamma = 2$)

D'autre part, pour un point sonique, $u^2 = \frac{\partial P}{\partial R} = \frac{\partial p}{\partial \rho}$. On introduit donc l'expression de la pression à enthalpie totale et débits constants:

$$P(R) = \frac{\gamma - 1}{\gamma} R \left(h - \frac{u^2}{2} \right) = \frac{\gamma - 1}{\gamma} \left(Rh - \frac{(Ru)^2}{2R} \right)$$

On obtient donc:

$$P'(R) = \frac{\gamma - 1}{\gamma} \left(h - \frac{u^2}{2} \right)$$

Donc la condition $u^2 = P'(R)$ donne:

$$(\gamma + 1) \frac{u^2}{2} = (\gamma - 1)h$$

Finalement, $R_{sonic} = a \sqrt{\frac{\gamma+1}{\gamma-1}}$ est un point sonique pour (10.14). D'après l'expression de ψ , on voit que $\psi(a) = 0$, $\lim_{R \rightarrow +\infty} \psi(R) = 0$, et $\psi(R_{sonic}) > 0$. Le graphe de ψ est donc de la forme illustrée par le figure 10.2. \square

Par conséquent, le point d'extremum pour ψ correspond à un état où la jacobienne des flux est singulière. Ceci implique que l'expression de Φ (10.5) dégénère et qu'il est impossible de franchir ce dernier de façon régulière (voir [34], par exemple, pour une discussion similaire portant sur le nombre de Mach). On est alors en position de résonance: ce point fait l'objet du chapitre suivant.

10.2.2 Ecriture du schéma pour ce type de systèmes

On décrit maintenant précisément l'algorithme numérique permettant de calculer les approximations du problème (10.14) au moyen du schéma à décomposition de flux modifié (10.8). Si on désigne par

$$U = (R, Ru, Re), F(U) = (Ru, Ru^2 + P, Ruh)$$

il suffit de poser F^\pm les flux décomposés de VanLeer et d'écrire que:

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} [F^+(U_j^n) - F^+(U_{j-\frac{1}{2}}^+) + F^-(U_{j+\frac{1}{2}}^-) - F^-(U_j^n)]$$

avec les états à l'interface $U_{j\pm\frac{1}{2}}^\pm$ qui se calculent au moyen de la relation (10.16).

Concrètement, on a pour l'état gauche $U_{j-\frac{1}{2}}^+$:

$$\psi(R_{j-\frac{1}{2}}^+) = \left(\frac{S(j\Delta x)}{S((j-1)\Delta x)} \right)^{1-\gamma} \psi(R_{j-1}) \text{ avec } (Ru)_{j-\frac{1}{2}}^+ = (Ru)_{j-1} \text{ et } h_{j-\frac{1}{2}}^+ = h_{j-1}$$

et pour l'état droit $U_{j+\frac{1}{2}}^-$:

$$\psi(R_{j+\frac{1}{2}}^-) = \left(\frac{S(j\Delta x)}{S((j+1)\Delta x)} \right)^{1-\gamma} \psi(R_{j+1}) \text{ avec } (Ru)_{j+\frac{1}{2}}^- = (Ru)_{j+1} \text{ et } h_{j+\frac{1}{2}}^- = h_{j+1}$$

Ces relations se font donc à Ru et h constants. Ceci suffit à définir complètement tous les états $U_{j\pm\frac{1}{2}}^\pm$ en l'absence de tout point sonique.

10.2.3 Cas-tests numériques

La tuyère supersonique Le premier cas-test correspond à la même configuration que dans le chapitre 3 de la partie III. La validation se fera encore avec le schéma à pas fractionnaires décentré proposé par P. Glaister [32]. Il s'agit de calculer un écoulement d'air ($\gamma = 1.4$) dans une tuyère divergente (voir figure 9.5). Les données initiales et la condition imposée à gauche sont:

$$\left. \begin{array}{l} \rho_0 = 0.125 \\ u_0 = 1 \\ i_0 = 2.5 \end{array} \right\} \text{ et } \left. \begin{array}{l} \rho_l = 0.502 \\ u_l = 1.299 \\ p_l = 0.381 \end{array} \right\}$$

Le régime est supersonique et fortement instationnaire au commencement, puis converge vers un régime permanent donné par [27] et recoupé par le schéma de référence [32]. On présente les itérations en temps sur la densité, la pression, la vitesse et le débit (figures 10.3, 10.4, 10.5, 10.6) qui s'accordent parfaitement avec les résultats du chapitre 3, ainsi que ceux de [27, 32]. On obtient donc bien le même état stationnaire convergé où le débit est constant en espace. Ceci est illustré par les deux coupes en $t = 1$ qui permettent de comparer précisément les résultats du présent schéma avec celui de P. Glaister. Les paramètres utilisés sont bien sûr identiques à savoir $\Delta x = 0.02$, $\Delta t = 0.006$, ce qui donne une CFL maximale de 0.9.

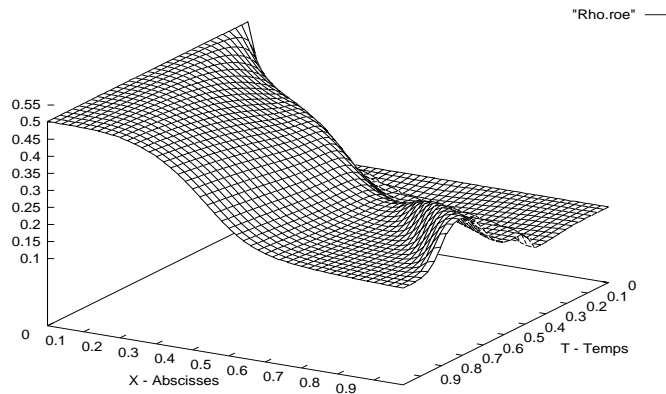


Figure 10.3: Densité (schéma à décomposition de flux)

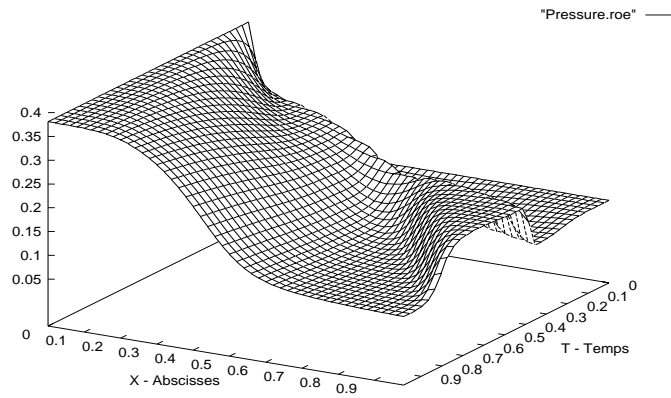


Figure 10.4: Pression (schéma à décomposition de flux)

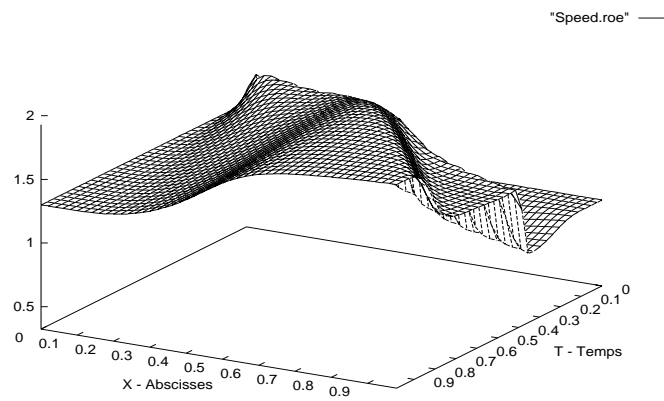


Figure 10.5: Vitesse (schéma à décomposition de flux)

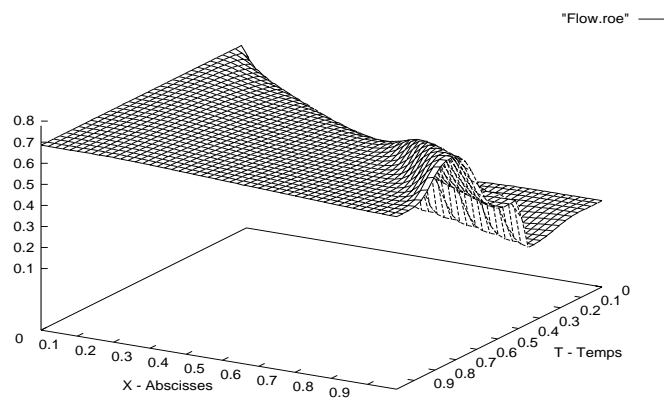


Figure 10.6: Débit (schéma à décomposition de flux)

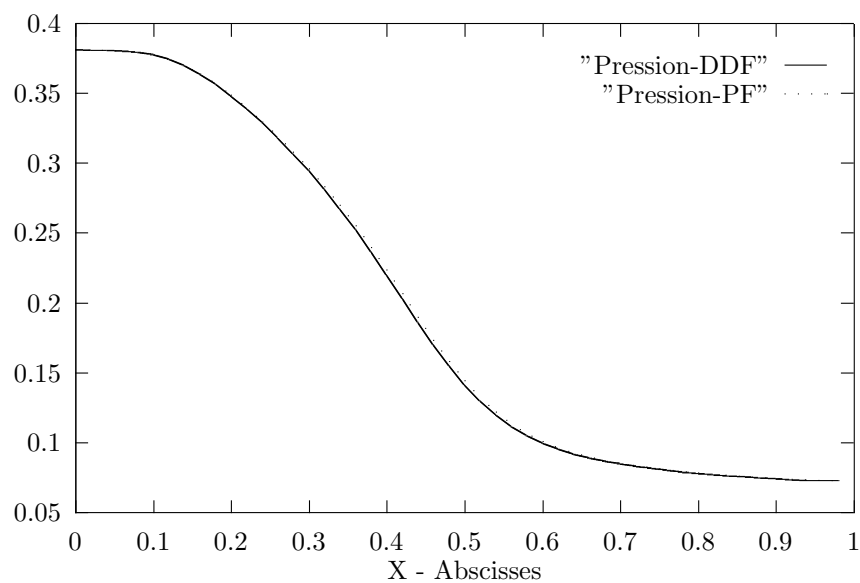


Figure 10.7: Pressions en $t = 1$ (schéma à décomposition de flux / pas fractionnaires)

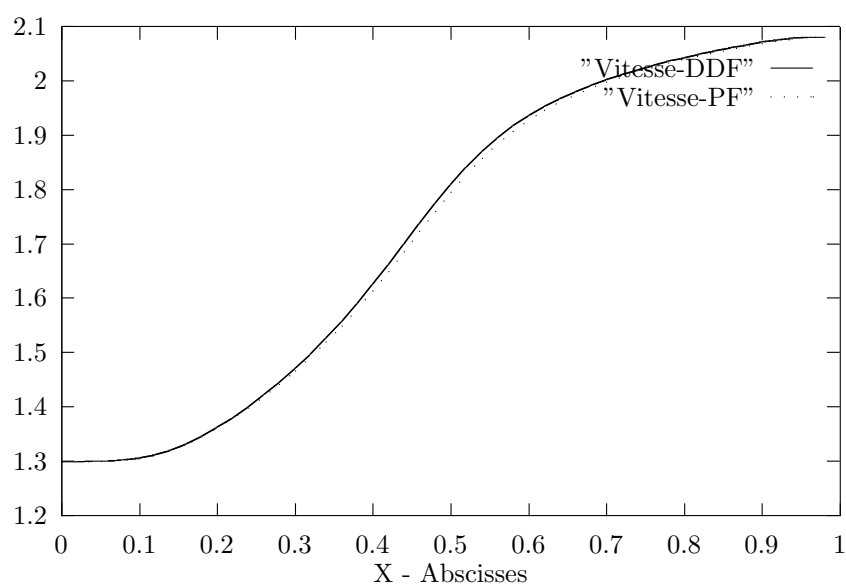


Figure 10.8: Vitesses en $t = 1$ (schéma à décomposition de flux / pas fractionnaires)

Le tube à choc cylindrique Le second cas-test reprend l'exemple du tube à choc avec symétrie cylindrique proposé dans [6]. Ceci correspond au cas où la variable $x \in [0, 1]$ désigne l'éloignement à l'axe du cylindre et la fonction $x \mapsto S(x)$ est définie par:

$$\begin{cases} S(x) = x \text{ pour } x > \Delta x \\ S(x) = \varepsilon \Delta x \text{ sinon} \end{cases}$$

C'est-à-dire qu'à une distance suffisante de l'axe $x = 0$ qui est une singularité, on résout le problème donné par:

$$\begin{cases} (\rho a^{\Delta x})_t + (\rho a^{\Delta x} u)_x = 0 \\ (\rho a^{\Delta x} u)_t + (\rho a^{\Delta x} u^2 + P)_x - P[\ln(a^{\Delta x})]_x = 0 \\ (\rho a^{\Delta x} e)_t + (\rho a^{\Delta x} u h)_x = 0 \\ a_t^{\Delta x} = 0 \end{cases}$$

où $a^{\Delta x}$ est la fonction définie en (10.3) et $P = (\gamma - 1)\rho a^{\Delta x} i$ est la loi de pression des gaz parfaits (i désigne l'énergie interne). Les données initiales proposées dans [6] sont celles d'un tube à choc de type Sod où les discontinuités convergent vers l'axe du cylindre:

$$\left| \begin{array}{l} \rho_L = 1 \\ u_L = 0 \\ i_L = 2.5 \end{array} \right. \quad \text{et} \quad \left| \begin{array}{l} \rho_R = 4 \\ u_R = 0 \\ i_R = 2.5 \end{array} \right.$$

Le domaine est monodimensionnel et $x \in [0, 1]$; le saut initial est en $x = 0.5$. Les ondes convergent vers le centre du cylindre où on a placé une condition de réflexion du type «état miroir». Le maillage est identique au cas précédent. En $t = 0.55$, on observe que le choc réfléchi a heurté la discontinuité de contact; comme le schéma est d'ordre 1 en espace, le choc convergent faible qui en résulte n'est pas visible. Les itérations en temps sont présentées aux figures 10.9, 10.10, 10.11. On peut observer le pic de pression en $x = 0$ dû à la condition-limite choisie; ceci est proche des résultats de [6]. On ne constate en outre aucune oscillation au voisinage des discontinuités. Trois coupes en $t = 0.55$ sont présentées aux figures 10.13, 10.14, 10.15. On y présente une comparaison entre le présent schéma à décomposition de flux (en trait plein fin), le schéma à pas fractionnaires de P. Glaister [32] (en trait pointillé), et le schéma de type Roe développé au chapitre 3 de la partie III (en trait plein gras). On constate un bon accord entre les deux approches «par interfaces» excepté sur la densité au voisinage de $x = 0$. Ceci est dû à la condition-limite «miroir» qui perturbe la relation de saut donnant l'état $U_{\frac{1}{2}}^+$; de ce fait, le débit met très longtemps à s'équilibrer (cf figure 10.12 pour une coupe en $t \simeq 20$). On constate néanmoins que les profils d'ondes sont corrects et les paliers sont proches des résultats à l'ordre 2 de [6]. Ceci n'est pas du tout le cas pour le schéma de P. Glaister qui donne une solution très différente, notamment pour la vitesse et la pression (voir la figure 10.15). Pour cette dernière quantité, la solution de référence présente un palier constant entre l'axe $x = 0$ et le choc divergent; le schéma à pas fractionnaires ne respecte pas du tout cela (voir figure 10.15).

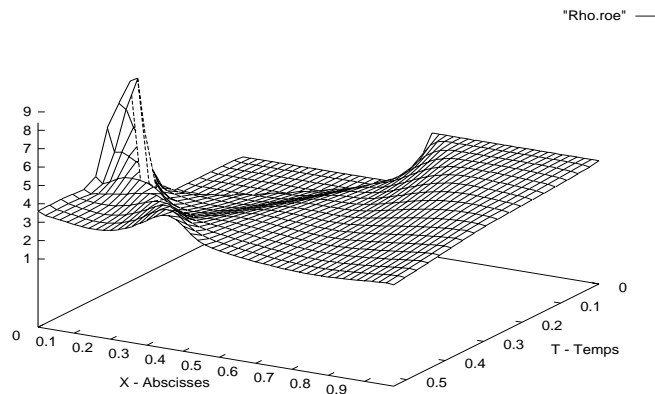


Figure 10.9: Densité (schéma à décomposition de flux)

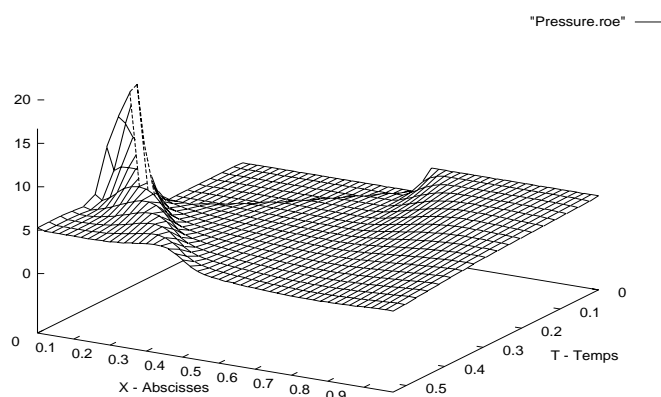


Figure 10.10: Pression (schéma à décomposition de flux)

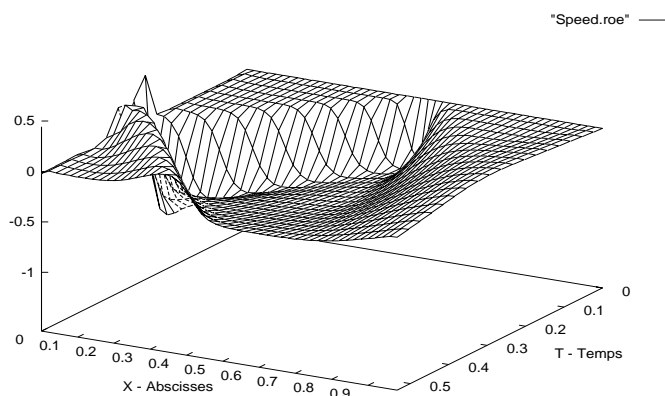


Figure 10.11: Vitesse (schéma à décomposition de flux)

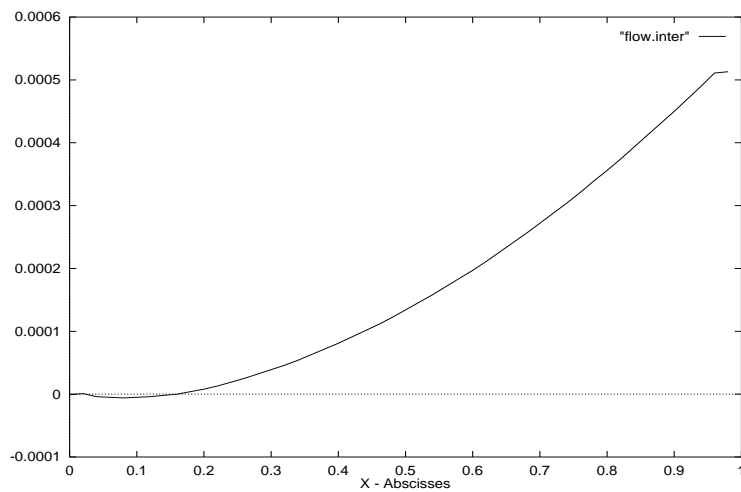


Figure 10.12: Débit en $t \simeq 20$ (schéma à décomposition de flux)

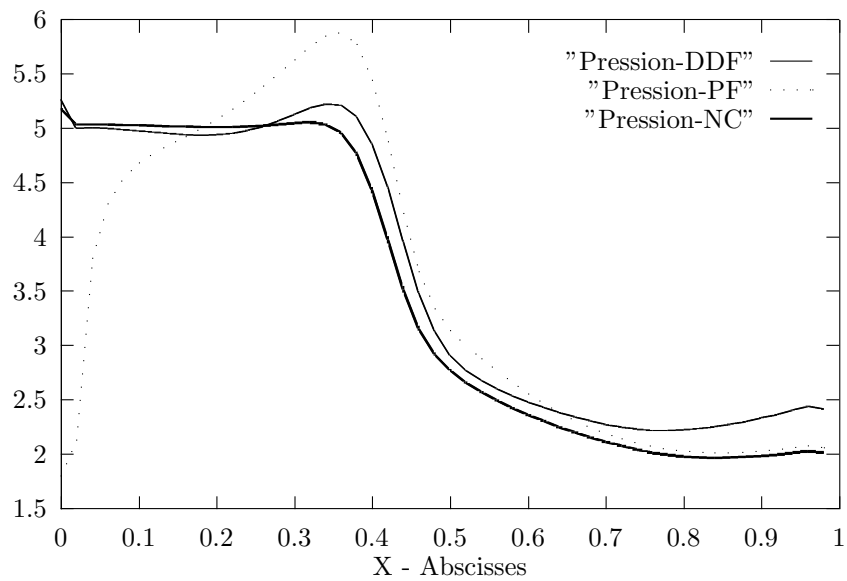


Figure 10.13: Densité en $t = 0.55$ pour les 3 schémas
DDF: décomposition de flux; PF: pas fractionnaires; NC: schéma de type Roe non-conservatif

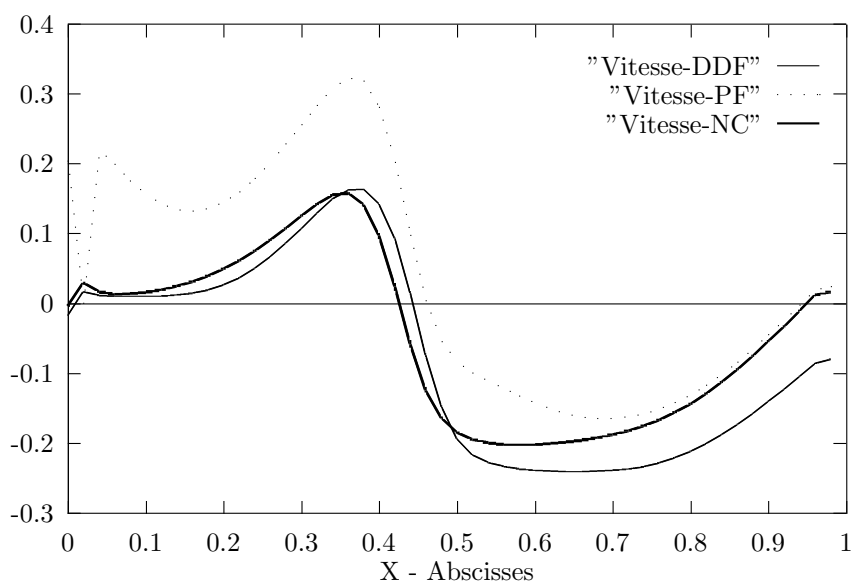


Figure 10.14: Vitesse en $t = 0.55$ pour les 3 schémas
DDF: décomposition de flux; PF: pas fractionnaires; NC: schéma de type Roe non-conservatif

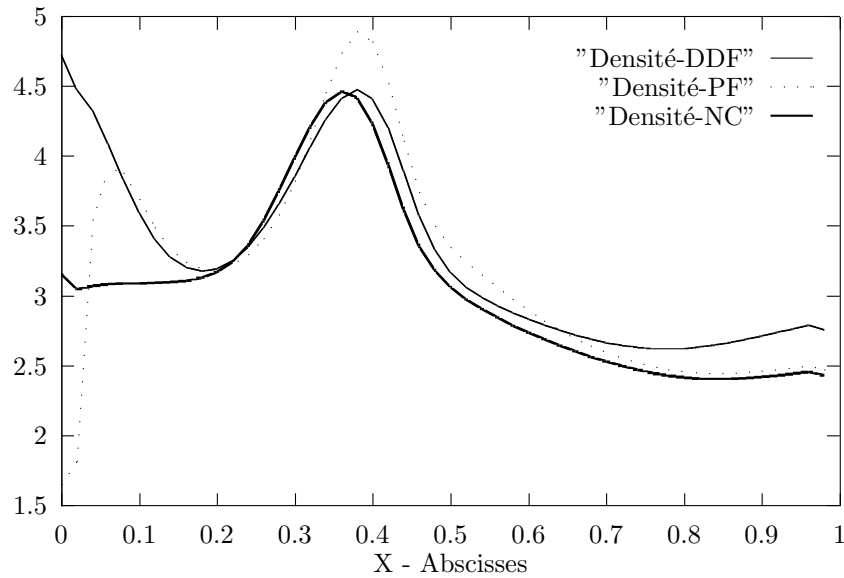


Figure 10.15: Pression en $t = 0.55$ pour les 3 schémas
DDF: décomposition de flux; PF: pas fractionnaires; NC: schéma de type Roe non-conservatif

10.3 Le système diphasique équilibré

10.3.1 Expression des solutions stationnaires

On considère maintenant le problème suivant:

$$\begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho c)_t + (\rho u c)_x = k_1(x) \cdot \rho g(c) \\ (\rho u)_t + (\rho u^2 + p)_x = -k_2(x) \cdot \rho u |u| \end{cases} \quad (10.17)$$

où $g(c) = c(1-c)(\mu - c)$ avec $\mu \in]0, 1[$. Ce modèle a déjà été abordé à maintes reprises, notamment au chapitre 2 de la partie III où on l'avait traité sans le recours du formalisme non-conservatif du chapitre suivant. On rappelle la signification des symboles: $\rho \in \mathbb{R}_*^+$, $u, c \in [0, 1]$, p, M, h désignent respectivement la densité, vitesse volumétrique commune aux deux phases, concentration massique, pression isentropique simplifiée ($p(\rho, c) = \rho^2 \sqrt{c}$), quantité de mouvement ($M = \rho u$) et enthalpie totale. Le solveur de Riemann homogène employé ici sera le splitting de VanLeer modifié pour les mélanges diphasiques (voir la proposition III.2). Les termes sources modélisent ici une recondensation vers un état privilégié stable donné par $c = \mu$ ainsi qu'un effet de frottement pariétal (k_1 uniforme) ou de perte de charge (k_1 variant fortement avec x). Comme précédemment, il est nécessaire de clarifier l'expression des courbes intégrales du système (10.17) stationnaire afin d'obtenir l'expression du chemin (10.5). On introduit aussi la fonction $x \mapsto a^{\Delta x}(x)$ et les primitives définies à des constantes près $x \mapsto K_{1/2}(x)$ des fonctions $x \mapsto k_{1/2}(x)$:

$$\begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho c)_t + (\rho u c)_x - [\rho g(c) K_1(a^{\Delta x})_x]_{\Phi} = 0 \\ (\rho u)_t + (\rho u^2 + p)_x + [\rho u |u| K_2(a^{\Delta x})_x]_{\Phi} = 0 \\ a_t^{\Delta x} = 0 \end{cases} \quad (10.18)$$

Le problème non-conservatif obtenu est bien sûr non-strictement hyperbolique, c'est pourquoi on se limitera à des expériences subsoniques afin d'éviter toute situation résonante. En vue d'implanter le schéma (10.8) pour le système (10.18), il est nécessaire d'en expliciter les courbes stationnaires afin de pouvoir calculer les états à l'interface. Malheureusement, celles-ci imposent la résolution d'un système différentiel algébrique non-linéairement couplé qui ne peut être envisagée ici. On propose donc une simplification.

Lemme IV.3 Pour des états non-soniques et lorsque $k_1 \equiv 0$, une expression du chemin régularisant (10.5) pour le système (10.18) est donnée par:

$$\Phi \left(s; \begin{pmatrix} \rho_L \\ \rho_L c_L \\ \rho_L u_L \\ a_L^{\Delta x} \end{pmatrix}, \begin{pmatrix} \rho_R \\ \rho_R c_L \\ \rho_L u_L \\ a_R^{\Delta x} \end{pmatrix} \right) = \begin{pmatrix} \rho(x) \\ \rho(x) c_L \\ \rho_L u_L \\ a_L^{\Delta x} + (a_R^{\Delta x} - a_L^{\Delta x}) s \end{pmatrix} \quad (10.19)$$

où $x = s\Delta x$, et $\rho(x)$ est solution de:

$$\ln \left(\frac{\rho(x)}{\rho_L} \right) - \frac{2\sqrt{c_L}}{3(\rho_L u_L)^2} (\rho(x)^3 - \rho_L^3) = \text{sgn}(\rho_L u_L) (K_2(a_R^{\Delta x}) - K_2(a_L^{\Delta x})) \quad (10.20)$$

Démonstration

Posons $M_L = \rho_L u_L$: le système stationnaire pour (10.17) donne:

$$\begin{cases} (\rho u)_x = 0 \text{ donc } (\rho u)(x) = M_L \\ (\rho u) \cdot c_x = 0 \text{ donc } c(x) = c_L \\ (\rho u^2 + p)_x = -\text{sgn}(M_L) \cdot \frac{M_L^2}{\rho} \cdot K_2(a^{\Delta x})_x \end{cases}$$

On insère dans la dernière équation l'expression de la pression:

$$-\frac{M_L^2}{\rho^2} \rho_x + 2\rho \rho_x \sqrt{c_L} = -\text{sgn}(M_L) \cdot \frac{M_L^2}{\rho} \cdot K_2(a^{\Delta x})_x$$

En multipliant par ρ et en intégrant en x , il vient:

$$-M_L^2 \ln(\rho)_x + \frac{2}{3} (\rho^3)_x \sqrt{c_L} = -\text{sgn}(M_L) \cdot M_L^2 K_2(a^{\Delta x})_x$$

On suppose $M_L \neq 0$ (sinon le terme de frottement est nul et $\rho_x = 0$) et on trouve:

$$[\ln(\rho)] - \frac{2\sqrt{c_L}}{3M_L^2} [\rho^3] = \text{sgn}(M_L) [K_2]$$

□

L'équation (10.20) ne permet en général pas d'obtenir la valeur de $\rho(\Delta x)$ en forme fermée, c'est pourquoi il est nécessaire de mettre en place un algorithme d'inversion de type Newton. A ce propos, on peut remarquer qu'il est intéressant d'itérer sur la variable $X = \rho^3$.

Lorsque $k_1(x) \neq 0$, utiliser la relation (10.20) telle quelle revient à négliger la variation de c dans l'équilibrage des quantités intervenant dans la dernière équation. Ceci est justifié dans trois cas au moins:

- La pression dépend très peu de la concentration
- k_1 est très faible, donc la variation de concentration est négligeable
- c_L est très proche d'un équilibre réactionnel stable

10.3.2 Ecriture du schéma pour ce système

Lorsque $k_{1/2} \neq 0$, la connaissance des courbes à l'équilibre de (10.17) passe par la résolution d'un système différentiel couplé:

$$\begin{cases} (\rho u)_x = 0 \\ (\rho u) c_x = k_1(x) \cdot \rho g(c) \\ (\rho u^2 + p(\rho, c))_x = -k_2(x) \cdot \rho |u| \end{cases}$$

Même pour la loi de pression simplifiée $p(\rho, c) = \rho^2 \sqrt{c}$, il n'est pas possible de le résoudre par quadratures. Néanmoins, on persistera à utiliser le lemme IV.3 dans d'autres configurations où $k_1 \neq 0$ en mettant en place un algorithme itératif du point fixe qui calcule une solution approchée de ce système à $o(\Delta x^2)$ près:

- On fixe les valeurs initiales $\rho_0 = \rho_L$, $c_0 = c_L$ et on calcule $\rho_1 = \rho(\Delta x)$ par la relation (10.20) grâce à un algorithme de Newton (variable $X = \rho^3$ et fonction $f_{c_0}(X) = \frac{2\sqrt{c_0}}{M_L^2} \cdot X - \ln(X)$).

- L'équation sur la concentration peut alors se réécrire $M_L \phi(c)_x = \rho_1 \cdot K_2(a^{\Delta x})_x$ où ρ_1 est la valeur calculée précédemment. La fonction $c \mapsto \phi(c)$ est celle explicitée au lemme III.6. Ceci nous définit une valeur $c_1 = c(\Delta x)$.
 - On recommence le processus avec comme valeurs initiales (ρ_1, c_1) pour trouver des valeurs plus fines (ρ_2, c_2) , voire au-delà ...

Le choix s'est porté sur cet algorithme très simple pour des raisons de facilité d'implémentation. En effet, il tient sur quelques lignes de code supplémentaires. D'autre part, sa vitesse de convergence est apparue satisfaisante dans les cas traités ici. L'impact de l'erreur commise sur la solution numérique semble faible puisque qu'il est difficile de distinguer des profils obtenus avec des tolérances allant de 10^{-2} à 10^{-6} . Bien entendu, ce n'est pas l'unique façon de programmer un schéma \llcorner interfaces \lrcorner : on peut calculer les états $U_{j\pm\frac{1}{2}}^\pm$ à l'aide d'un solveur différentiel de type Runge-Kutta-Fehlberg à pas adaptatif, ou bien avec une routine d'inversion de Newton dans \mathbb{R}^2 .

L'étude de cet algorithme ainsi qu'une estimation de l'erreur commise à convergence en itérant ces relations de saut simplifiées sont rejetées en annexe A. On est donc en position d'écrire le schéma (10.8) pour le système (10.18) dans le cas général. On pose pour cela:

$$U = (\rho, \rho c, \rho u), F(U) = (\rho u, \rho c u, \rho u^2 + p)$$

La décomposition de VanLeer adaptée aux mélanges diphasiques (voir la proposition III.2 et le lemme III.6) permet de définir les fonctions régulières $U \mapsto F^\pm(U)$ nécessaires pour le schéma \llcorner équilibre \lrcorner :

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} [F^+(U_j^n) - F^+(U_{j-\frac{1}{2}}^+) + F^-(U_{j+\frac{1}{2}}^-) - F^-(U_j^n)]$$

Il reste à expliciter les états $U_{j\pm\frac{1}{2}}^\pm$. A gauche, l'état $U_{j-\frac{1}{2}}^+$ vérifie les relations suivantes:

$$(\rho u)_{j-\frac{1}{2}}^+ = (\rho u)_{j-1}$$

et:

$$\begin{cases} \ln \left(\frac{\rho_{j-\frac{1}{2}}^+}{\rho_{j-1}} \right) - \frac{2\sqrt{c_{j-\frac{1}{2}}^+}}{3(\rho u)_{j-1}^2} [(\rho_{j-\frac{1}{2}}^+)^3 - \rho_{j-1}^3] = \text{sgn}((\rho u)_{j-1})(K_2(j\Delta x) - K_2((j-1)\Delta x)) \\ (\rho u)_{j-1} [\phi(c_{j-\frac{1}{2}}^+) - \phi(c_{j-1})] = \rho_{j-\frac{1}{2}}^+ (K_1(j\Delta x) - K_1((j-1)\Delta x)) \end{cases}$$

L'état droit $U_{j+\frac{1}{2}}^-$, quant à lui vérifie:

$$(\rho u)_{j+\frac{1}{2}}^- = (\rho u)_{j+1}$$

et:

$$\begin{cases} \ln \left(\frac{\rho_{j+\frac{1}{2}}^-}{\rho_{j+1}} \right) - \frac{2\sqrt{c_{j+\frac{1}{2}}^-}}{3(\rho u)_{j+1}^2} [(\rho_{j+\frac{1}{2}}^-)^3 - \rho_{j+1}^3] = \text{sgn}((\rho u)_{j+1})(K_2(j\Delta x) - K_2((j+1)\Delta x)) \\ (\rho u)_{j+1} [\phi(c_{j+\frac{1}{2}}^-) - \phi(c_{j+1})] = \rho_{j+\frac{1}{2}}^- (K_1(j\Delta x) - K_1((j+1)\Delta x)) \end{cases}$$

10.3.3 Cas-tests numériques

Il s'agit dans un premier temps de s'assurer de la pertinence des simplifications effectuées dans la gestion des termes non-conservatifs au moyen de l'algorithme du point fixe. Pour cela, on décide de comparer le schéma à décomposition de flux et un schéma de Roe [81] à pas fractionnaires dans lequel les termes sources sont traités de façon centrée. Le cas-test est un problème aux limites où on impose un débit et une concentration en entrée, et où on place une condition de Neumann en sortie. La solution se compose d'une onde de choc sur la vitesse ainsi que d'une discontinuité de contact sur la concentration. Cela donne:

$$\left| \begin{array}{l} \rho_0 = 2 \\ c_0 = 0.5 \\ u_0 = 0.5 \end{array} \right. \text{ et } \left| \begin{array}{l} c_l = 0.9 \\ (\rho u)_l = 1 \end{array} \right.$$

Les valeurs choisies sont $\mu = 0.5$, $k_1 \equiv 50$ et $k_2 \equiv 5$. Ce choix aura tendance à amortir les ondes présentes dans la solution. En effet, le terme de frottement dissipe de la vitesse, et la recondensation tend à créer une couche-limite de transition entre la valeur en entrée ($c = 0.9$) et la valeur à l'équilibre $c = \mu = 0.5$. D'autre part, on choisit les paramètres $\Delta x = 0.02$ (50 mailles), $\Delta t = 0.006$, ce qui donne une CFL maximale de 0.9. On itère jusque $T = 0.6$: les courbes comparées sont présentées aux figures (10.16) et (10.17). On observe une bonne adéquation des deux approches en régime peu raide puisque les profils de transition sont très proches. D'autre part, les valeurs de la vitesse sont identiques en amont du choc et très proches en aval sur toute la durée des itérations.

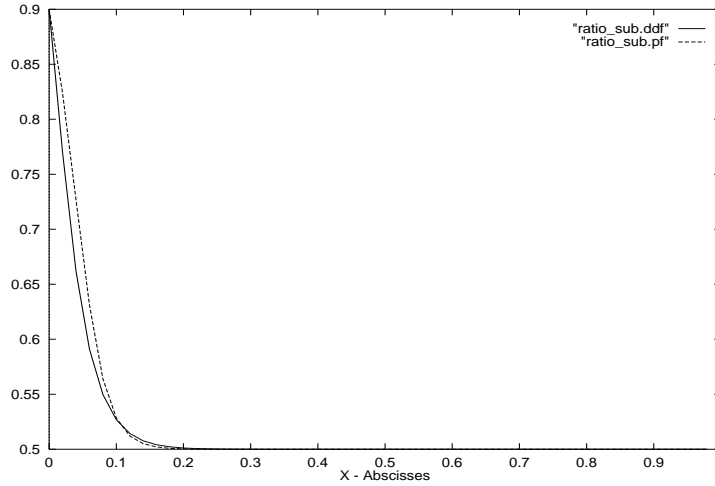


Figure 10.16: Concentration en $T = 0.6$ (schéma à décomposition de flux/pas fractionnaires)

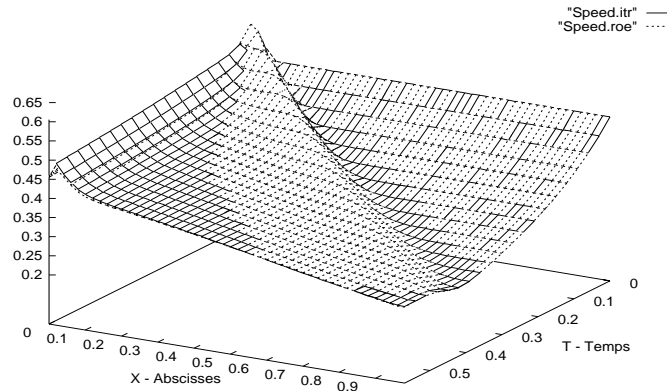


Figure 10.17: Vitesse pour $t \in [0, 0.6]$ (schéma à décomposition de flux/pas fractionnaires)

Dans un deuxième temps, on traite un problème de couche limite plus difficile, i.e. on veut contrôler la stabilité du schéma en régime réactionnel intense. En outre, on place quatre grilles à aubes (modélisée comme au chapitre 3 de la partie III) pour complexifier l'écoulement. Les données initiales et les conditions aux limites sont identiques à celles du problème précédent. Cela donne:

$$k_1 \equiv 10000 \text{ et } k_2(x) = \begin{cases} 50 & \text{si } x = 0.2, 0.4, 0.6, 0.8 \\ 0 & \text{sinon} \end{cases}$$

On utilise les mêmes paramètres que précédemment, ce qui donne une CFL maximale de 0.9. Les figures donnant les résultats sont numérotées selon 10.18 pour la concentration, 10.19 pour la vitesse, 10.20 pour la pression,

10.21 pour le débit. On constate là encore un débit presque invariant malgré la présence des grilles: on peut comparer la courbe 10.21 avec celle présentée à l'annexe B de la partie III: figure D.3 où les pertes de charges sont traitées de façon centrée. Cette particularité permet d'obtenir des solutions stationnaires monotones malgré la couche-limite de concentration. On ne note en effet aucun *overshoot* sur la vitesse et l'étalement des paliers de pression reste modéré. Signalons d'autre part que le palier de pression en sortie de couche limite est conservé si l'on varie (raisonnablement) l'étalement de celle-ci via la valeur de k_1 : voir en figure 10.22. Cela signifie qu'en régime raide, il n'est pas nécessaire de résoudre l'échelle de la zone de transition en espace.

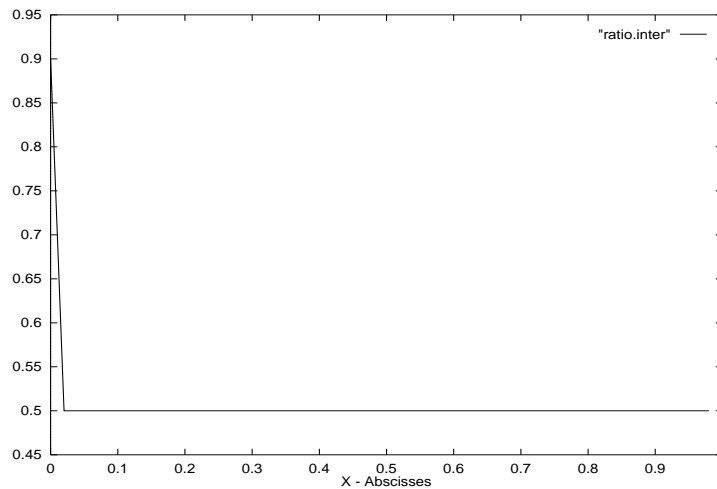


Figure 10.18: Concentration en $t = 2.5$ (schéma à décomposition de flux)

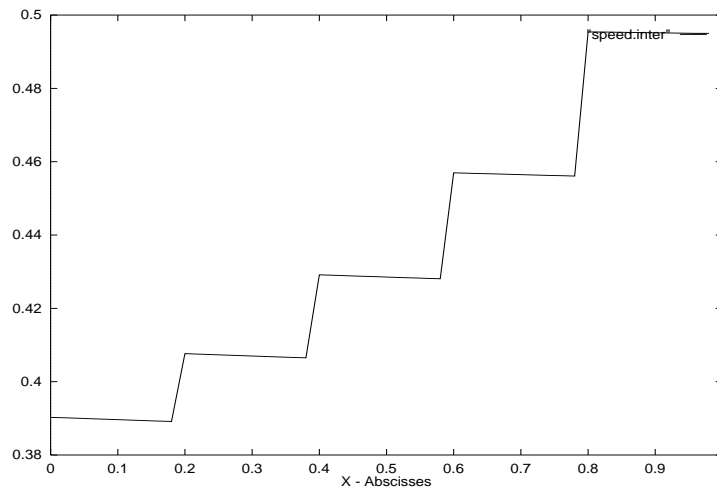


Figure 10.19: Vitesse en $t = 2.5$ (schéma à décomposition de flux)

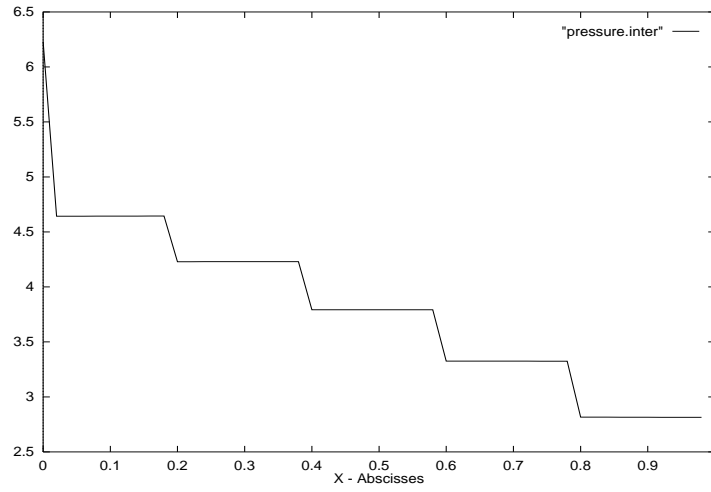


Figure 10.20: Pression en $t = 2.5$ (schéma à décomposition de flux)

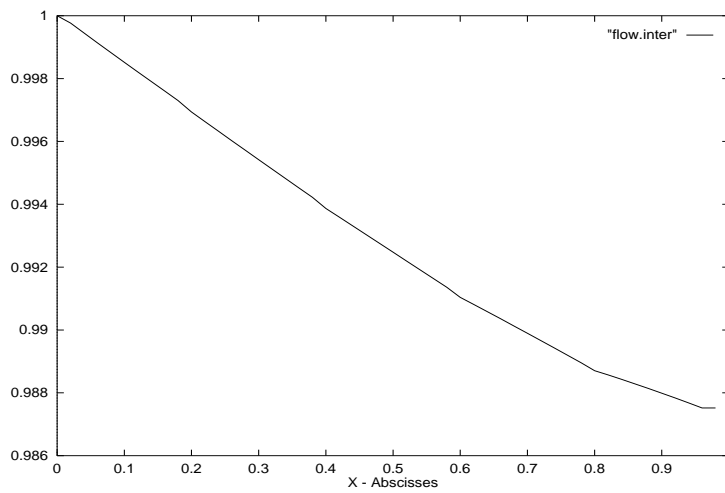


Figure 10.21: Débit en $t = 2.5$ (schéma à décomposition de flux)

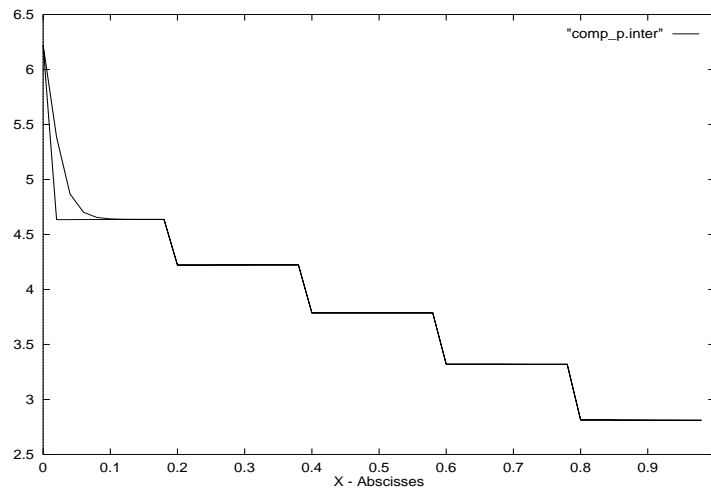


Figure 10.22: Comparaison des paliers de pression pour $k_1 = 100, 10000$ en $t = 2.5$

10.4 Le cas des termes de relaxation

Parmi la variété des problèmes hyperboliques non-homogènes de la forme (10.1), il en existe une sous-classe particulièrement digne d'intérêt: les systèmes relaxés. En effet, une théorie [14, 15, 65] a récemment été élaborée à propos de leur comportement asymptotique (i.e. lorsque le paramètre de relaxation tend vers 0). Ceci a entraîné la publication de plusieurs travaux dans le domaine numérique [8, 50, 51]: il est donc logique d'y consacrer un paragraphe.

10.4.1 Introduction à la théorie des systèmes relaxés

Un tel système se présente usuellement sous la forme suivante:

$$\begin{cases} U_t + F(U)_x = \frac{G(U)}{\epsilon} \text{ avec } U \in \mathbb{R}^N \\ U_0 \in [BV(\mathbb{R})]^N \end{cases} \quad (10.21)$$

Dans [65], T.P. Liu a considéré l'exemple:

$$\begin{cases} u_t + f(u, v)_x = 0 \\ v_t + g(u, v)_x = \frac{v_*(u) - v}{\tau} \equiv h(u, v) \end{cases} \quad (10.22)$$

Le terme $h(u, v)$ est une relaxation au sens où:

$$h(u, v_*(u)) = 0 \text{ et } \frac{\partial h}{\partial v}(u, v) < 0 \quad (10.23)$$

Lorsque la solution est proche de l'équilibre réactionnel, elle vérifie approximativement:

$$\begin{cases} u_t + f_*(u)_x = 0 \\ f_*(u) = f(u, v_*(u)) \end{cases} \quad (10.24)$$

La vitesse caractéristique pour (10.24) est $\lambda_*(u) = f'_*(u)$. Ce travail a donc consisté à vérifier la stabilité de la solution équilibre $\bar{U} = \begin{pmatrix} u \\ v_*(u) \end{pmatrix}$ de (10.24) lorsque τ est petit sous une condition appelée *sous-caractéristique*:

$$\lambda_1 < \lambda_* < \lambda_2 \text{ où } \Lambda \left(\frac{\partial(f, g)}{\partial(u, v)} \right) = \{\lambda_1, \lambda_2\} \quad (10.25)$$

Ceci parce que la correction au premier ordre en τ pour (10.24) s'écrit:

$$\begin{cases} u_t + f_*(u)_x = \tau(\beta(u)u_x)_x \\ f_*(u) = f(u, v_*(u)) \text{ et } \beta(u) = -(\lambda_* - \lambda_1)(\lambda_* - \lambda_2) \end{cases}$$

Ce problème sera parabolique, i.e. bien posé sous la condition (10.25). Il peut aussi survenir des cas où le comportement asymptotique soit purement diffusif [47]:

$$\begin{cases} v_t - u_x = 0 \\ u_t + p(v)_x = -\alpha u \\ \alpha < 0 \text{ et } p'(v) < 0, p''(v) > 0 \end{cases} \quad t \xrightarrow{+} \infty \begin{cases} v_t + \frac{p(v)_{xx}}{\alpha} = 0 \\ p(v)_x = -\alpha u \end{cases}$$

Toutes ces idées furent unifiées ultérieurement dans [14] pour décrire théoriquement le comportement de (10.21) à la limite $\epsilon \rightarrow 0$: notamment, sous quelle condition le système réduit du type (10.24) est-il hyperbolique ?

Il est alors nécessaire de préciser quelques notations (identiques à celles employées dans [14]): Soit $U \in \mathbb{R}^N$, le terme de relaxation $G(U)$ induit un projecteur \mathcal{Q} de rang $n < N$ et de matrice $n \times N$ tel que:

$$\forall U \in \mathbb{R}^N, \mathcal{Q}G(U) = 0$$

Il y a donc n quantités invariantes notées $u = \mathcal{Q}U$ par le champ de vecteurs $\dot{U} = G(U)$. On suppose en outre que chaque u détermine un unique \bar{U} à l'équilibre réactionnel:

$$\bar{U} = \mathcal{E}(u) \text{ avec } G(\mathcal{E}(u)) = 0 \text{ et } \mathcal{Q}\mathcal{E}(u) = u \quad (10.26)$$

L'image de \mathcal{E} dans \mathbb{R}^N est la variété à l'équilibre de G . Par conséquent, la forme canonique de G est:

$$G(U) = \mathcal{E}(QU) - U$$

Pour l'exemple (10.22), on a (sous réserve que v_* soit injective):

$$Q = (1 \ 0) \quad \mathcal{E}(u) = \begin{pmatrix} u \\ v_*(u) \end{pmatrix}$$

Les quantités invariantes par G peuvent être obtenues à la limite $\epsilon \rightarrow 0$ dans (10.21) et vérifient un système du premier ordre. Pour voir cela, soit $U \in \mathbb{R}^N$ solution de (10.21), il est possible de la développer sous la forme:

$$\begin{cases} U = U^0(u) + \epsilon U^1(u, u_x) + \dots \\ G(U^0) \equiv 0 \end{cases} \quad (10.27)$$

avec $QU = QU^0$ et $QU^k = 0$ pour $k \geq 1$. Comme $G(U^0) = 0$, il vient la condition à l'ordre 0 en ϵ :

$$U_t^0 + F(U^0)_x = 0$$

Comme chaque $u \in \mathbb{R}^n$ détermine un unique état à l'équilibre réactionnel, $U^0 = \mathcal{E}(u)$ et il vient en multipliant à gauche par Q :

$$(QU^0)_t + (QF(U^0))_x = 0$$

C'est-à-dire:

$$Q\mathcal{E}(u)_t + (QF(\mathcal{E}(u)))_x = 0$$

Par définition de \mathcal{E} , $Q\mathcal{E}(u) = u$. Il reste à définir une fonction de flux relaxée et il vient:

$$\begin{aligned} u_t + f(u)_x &= 0 \\ \text{où } f(u) &= QF(\mathcal{E}(u)) \end{aligned} \quad (10.28)$$

Ce calcul formel est justifié rigoureusement dans [14] où l'hyperbolicité du problème relaxé (10.28) et la stabilité de sa solution sont établies.

Théorème IV.2 (Chen, Levermore, Liu)

S'il existe un couple entropique (η, q) strictement convexe pour (10.21) compatibles avec $G(U)$ au sens suivant:

$$\forall U \in \mathbb{R}^N, \begin{cases} \nabla \eta(U) \cdot G(U) \leq 0 \\ \nabla \eta(U) \cdot G(U) = 0 \text{ implique } G(U) = 0 \end{cases}$$

alors si on note $u = QU \in \mathbb{R}^n$ la limite relaxée, celle-ci vérifie le système strictement hyperbolique:

$$\begin{cases} u_t + f(u)_x = 0 \text{ où } f(u) = QF(\mathcal{E}(u)) \\ u_0 = QU_0 \in [BV(\mathbb{R})]^n \\ \eta(\mathcal{E}(u))_t + q(\mathcal{E}(u))_x \leq 0 \end{cases}$$

De plus, la correction à l'ordre 1 en ϵ est strictement dissipative vis-à-vis de $\eta(\mathcal{E}(u))$.

Le critère de stabilité 2×2 (10.25) est un cas particulier de ce résultat très général. On renvoie à [14] pour des preuves détaillées.

10.4.2 Application: le modèle diphasique équilibré

On s'intéresse maintenant à un cas particulier du modèle diphasique proposé par [81] et muni de termes sources particuliers. Plus précisément, il s'agit du système:

$$\begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho c)_t + (\rho u c)_x = \frac{\rho}{\epsilon} g(\rho, c) \\ (\rho u)_t + (\rho u^2 + p(\rho, c))_x = 0 \end{cases} \quad (10.29)$$

Les notations sont identiques à celles employées dans (10.17). La seule hypothèse supplémentaire porte sur le terme de réaction $g(\rho, c)$ afin de pouvoir appliquer les résultats de la théorie relaxée. Pour cela, on suppose que celui-ci n'admet *qu'un seul équilibre*; il doit de plus être *stable*. En revanche, il peut tout à fait dépendre de ρ :

$$g(\rho, c) = \mu(\rho) - c \quad (10.30)$$

Si $\rho \mapsto \mu(\rho)$ est régulière et bijective sur \mathbb{R}_*^+ , ce terme est une relaxation au sens de [14]. La limite asymptotique $\epsilon \rightarrow 0$ est alors donnée par le système:

$$\begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho u)_t + (\rho u^2 + p(\rho, \mu(\rho)))_x = 0 \end{cases} \quad (10.31)$$

Un cas-test intéressant peut être de faire évoluer l'écoulement de fluide à travers une large variété de régimes différents par le biais de la forme choisie pour (10.30). Plus précisément, on se propose de déterminer $\mu(\rho)$ afin que la vitesse du son du système à l'équilibre réactionnel (10.31) soit différente de celle de (10.29).

Pour cela, on se donne la loi de pression approchée (cf chapitre 2 - partie III) $p(\rho, c) = \rho^2 \sqrt{c}$, et on impose à l'équilibre réactionnel la pression isentropique habituelle $p(\rho, \mu(\rho)) = \rho^\gamma$. Ceci implique que

$$\mu(\rho) = \rho^{2(\gamma-2)}$$

pour que la définition de la pression à l'équilibre soit cohérente. Finalement, on obtient le problème:

$$\begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho c)_t + (\rho u c)_x = \frac{\rho}{\epsilon} (\rho^{2(\gamma-2)} - c) \\ (\rho u)_t + (\rho u^2 + p(\rho, c))_x = 0 \end{cases} \quad (10.32)$$

Si on choisit une valeur pour γ inférieure à 2, le processus de relaxation couvrira une large plage de régimes d'écoulements. Il reste à s'assurer de la stabilité de la solution relaxée grâce au théorème IV.2.

Proposition IV.3 *Le système (10.29) admet une entropie strictement convexe compatible avec le terme de relaxation. Celle-ci s'écrit:*

$$\eta(\rho, \rho c, \rho u) = \frac{\rho u^2}{2} + \rho \left(\int_{\mu^{-1}(c)}^{\rho} \frac{p(s, c)}{s^2} ds + \int^{\mu^{-1}(c)} \frac{\bar{p}(s)}{s^2} ds \right) \quad (10.33)$$

avec $p(\rho, c) = \rho^2 \sqrt{c}$, $\bar{p}(\rho) = \rho^\gamma$ et $\gamma < 2$.

Démonstration

Il faut d'abord montrer que $\frac{\partial \eta}{\partial \rho c}(\rho, \rho c, \rho u)(\mu(\rho) - c) \leq 0$. Pour cela, on écrit:

$$\frac{\partial \eta}{\partial \rho c}(\rho, \rho c, \rho u) = \frac{1}{\rho} \frac{\partial \eta}{\partial c}(\rho, c, \rho u)$$

Donc on obtient:

$$\frac{\partial \eta}{\partial \rho c}(\rho, \rho c, \rho u) = \frac{\partial}{\partial c} \left[\int_{\mu^{-1}(c)}^{\rho} \frac{p(s, c)}{s^2} ds + \int^{\mu^{-1}(c)} \frac{\bar{p}(s)}{s^2} ds \right]$$

On effectue les dérivations à l'intérieur du signe somme et aux bornes de l'intervalle d'intégration:

$$\frac{\partial \eta}{\partial \rho c}(\rho, \rho c, \rho u) = \frac{\partial}{\partial c} \left[\int_{\mu^{-1}(c)}^{\rho} \frac{\partial p(s, c)}{\partial c} \frac{1}{s^2} ds - \frac{\partial \mu^{-1}(c)}{\partial c} \cdot \frac{p(\mu^{-1}(c), c)}{\mu^{-1}(c)^2} + \frac{\partial \mu^{-1}(c)}{\partial c} \cdot \frac{\bar{p}(\mu^{-1}(c))}{\mu^{-1}(c)^2} \right] \quad (10.34)$$

Remarquons maintenant que:

$$\begin{aligned} p(\mu^{-1}(c), c) &= (c^{1/2(\gamma-2)})^2 \cdot \sqrt{c} \\ &= c^{1/(\gamma-2) + 1/2} \\ &= (c^{1/2(\gamma-2)})^\gamma \\ &= \bar{p}(c^{1/2(\gamma-2)}) \\ &= \bar{p}(\mu^{-1}(c)) \end{aligned}$$

Par conséquent, les deux derniers termes de l'expression (10.34) s'éliminent. D'autre part, $p(s, c)/s^2 = \sqrt{c}$; on a donc:

$$\frac{\partial \eta}{\partial \rho c}(\rho, \rho c, \rho u) = \frac{1}{2\sqrt{c}}(\rho - \mu^{-1}(c))$$

Il faut donc vérifier que:

$$\frac{\partial \eta}{\partial \rho c}(\rho, \rho c, \rho u)(\mu(\rho) - c) = \frac{1}{2\sqrt{c}}(\rho - \mu^{-1}(c))(\mu(\rho) - c) \leq 0$$

Supposons pour cela que $\mu(\rho) - c \leq 0$, i.e. $\rho^{2(\gamma-2)} \leq c$. Comme $\gamma < 2$, ceci implique $\rho \geq c^{1/2(\gamma-2)}$ et donc $\rho - \mu^{-1}(c) \geq 0$. Pour des signes opposés, la propriété se montre de la même façon. On a donc bien la propriété recherchée à savoir:

$$\nabla \eta(\rho, \rho c, \rho u) \cdot G(\rho, \rho c, \rho u) \leq 0$$

De plus, la forme de η assure que $\mu(\rho) - c = 0$ si on a l'égalité dans l'expression ci-dessus.

Il reste donc à s'assurer que η est bien une entropie du système (10.29). Pour cela, on remarque que:

$$\eta(\rho, \rho c, \rho u) = \frac{\rho u^2}{2} + \rho \left(\sqrt{c}(\rho - \mu^{-1}(c)) + \int^{\mu^{-1}(c)} s^{\gamma-2} ds \right)$$

Le premier produit donne:

$$\sqrt{c}\mu^{-1}(c) = c^{\frac{1}{2} + \frac{1}{2(\gamma-2)}} = c^{\frac{\gamma-1}{2(\gamma-2)}}$$

De même il vient que:

$$\int^{\mu^{-1}(c)} s^{\gamma-2} ds = \frac{\mu^{-1}(c)^{\gamma-1}}{\gamma-1} = \frac{c^{\frac{\gamma-1}{2(\gamma-2)}}}{\gamma-1}$$

On obtient donc que:

$$\eta(\rho, \rho c, \rho u) = \frac{\rho u^2}{2} + \rho^2 \sqrt{c} + \left(\frac{1}{\gamma-1} - 1 \right) \rho c^{\frac{\gamma-1}{2(\gamma-2)}}$$

Dans [81] (voir aussi le chapitre 4 de la partie I), il est montré qu'une entropie strictement convexe pour (10.29) est donnée par:

$$\bar{\eta}(\rho, \rho c, \rho u) = \frac{\rho u^2}{2} + \rho \int^{\rho} \frac{p(s, c)}{s^2} ds = \frac{\rho u^2}{2} + \rho^2 \sqrt{c}$$

On peut alors écrire que:

$$\eta(\rho, \rho c, \rho u) = \bar{\eta}(\rho, \rho c, \rho u) + \frac{2-\gamma}{\gamma-1} \rho c^{\frac{\gamma-1}{2(\gamma-2)}}$$

Soit $dF(U)$ la jacobienne des flux du système (10.29), il faut montrer:

$$\nabla \eta(\rho, \rho c, \rho u) \cdot dF(U) = \nabla q(\rho, \rho c, \rho u) \text{ avec } dF(U) = \begin{pmatrix} 0 & 0 & 1 \\ -uc & u & c \\ -u^2 + p_\rho & p_{\rho c} & 2u \end{pmatrix}$$

Posons $\alpha = \frac{\gamma-1}{2(\gamma-2)}$, on calcule les dérivées partielles:

$$\frac{\partial \rho c^\alpha}{\partial \rho} = \frac{\partial (\rho c)^\alpha}{\partial \rho \rho^{\alpha-1}} = (1-\alpha)c^\alpha$$

ainsi que:

$$\frac{\partial \rho c^\alpha}{\partial \rho c} = \frac{\partial (\rho c)^\alpha}{\partial \rho c \rho^{\alpha-1}} = \alpha c^{\alpha-1}$$

Et donc, on obtient:

$$\nabla \eta(\rho, \rho c, \rho u) = \nabla \bar{\eta}(\rho, \rho c, \rho u) + \frac{2-\gamma}{\gamma-1} \begin{pmatrix} (1-\alpha)c^\alpha \\ \alpha c^{\alpha-1} \\ 0 \end{pmatrix}$$

Notons \bar{q} le flux de l'entropie $\bar{\eta}$ donnée plus haut, il vient:

$$\nabla\eta(\rho, \rho c, \rho u).dF(U) = \nabla\bar{q}(\rho, \rho c, \rho u) + \frac{2-\gamma}{\gamma-1} \begin{pmatrix} -\alpha u c^\alpha \\ \alpha u c^{\alpha-1} \\ c^\alpha \end{pmatrix}$$

Et donc, le flux d'entropie associé à η est donné par:

$$q(\rho, \rho c, \rho u) = \bar{q}(\rho, \rho c, \rho u) + \frac{2-\gamma}{\gamma-1}(\rho u c^\alpha)$$

Il reste à vérifier la convexité de η . Pour cela, on calcule la dérivée seconde de $(\rho, \rho c) \mapsto \rho c^\alpha$:

$$D^2(\rho c^\alpha) = \begin{pmatrix} (\alpha-1)\alpha \frac{c^\alpha}{\rho} & -(\alpha-1)\alpha \frac{c^{\alpha-1}}{\rho} \\ -(\alpha-1)\alpha \frac{c^{\alpha-1}}{\rho} & (\alpha-1)\alpha \frac{c^{\alpha-2}}{\rho} \end{pmatrix}$$

On vérifie donc que pour $\vec{v} = (v_1, v_2) \neq \vec{0}$:

$$\langle D^2(\rho c^\alpha).\vec{v}, \vec{v} \rangle = (\alpha-1)\alpha \frac{c^{\alpha-2}}{\rho} (c v_1 + v_2)^2 \geq 0$$

Il reste donc à déterminer le signe de $\alpha = \frac{\gamma-1}{2(\gamma-2)}$. Comme $1 < \gamma < 2$, $\alpha < 0$, et donc $(\alpha-1)\alpha > 0$. La somme d'une fonction strictement convexe et d'une fonction convexe est strictement convexe, donc on a le résultat annoncé. \square

Par conséquent, la solution relaxée du système (10.31) est stable et il est pertinent de tester la convergence des schémas numériques sur ce problème.

10.4.3 Formulation non-conservative associée

On commence par réécrire le système (10.32) sous la forme désormais familière:

$$\begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho c)_t + (\rho u c)_x - [\frac{\rho}{\epsilon}(\rho^{2(\gamma-2)} - c)a_x^{\Delta x}]_{\Phi} = 0 \\ (\rho u)_t + (\rho u^2 + p(\rho, c))_x = 0 \\ a_t^{\Delta x} = 0 \end{cases} \quad (10.35)$$

au moyen de la fonction (10.3). La régularisation se fera une fois encore grâce à la donnée du chemin \llbracket équilibre \rrbracket (10.5). Pour ce faire, posons:

$$V = (\rho, \rho c, \rho u, a^{\Delta x}); U = (\rho, \rho c, \rho u); F(U) = (\rho u, \rho u c, \rho u^2 + p); G(U) = (0, \rho(\rho^{2(\gamma-2)} - c), 0) \quad (10.36)$$

On a donc:

$$\mathcal{Q} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \mathcal{E}(\rho, \rho u) = (\rho, \rho \mu(\rho), \rho u)$$

Alors, le système (10.35) peut se réécrire sous la forme:

$$\begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho c)_t + (\rho u c)_x - [\rho(\rho^{2(\gamma-2)} - c)a_x^{\Delta x}]_{\Phi^\epsilon} = 0 \\ (\rho u)_t + (\rho u^2 + p(\rho, c))_x = 0 \\ a_t^{\Delta x} = 0 \end{cases}$$

Avec la régularisation suivante:

$$\Phi^\epsilon(s; V_L, V_R) = \begin{pmatrix} \tilde{U}(\frac{s\Delta x}{\epsilon}) \\ \tilde{a}(s\Delta x) \end{pmatrix} \quad (10.37)$$

où \tilde{U} est une courbe intégrale de $F(\tilde{U})_x = G(\tilde{U})\tilde{a}_x$; $\tilde{U}(0) = U_L$
et $\tilde{a}(x) = a_L^{\Delta x} + (a_R^{\Delta x} - a_L^{\Delta x})x/\Delta x$

Posons $\rho u = M$, l'expression de Φ^ϵ passe donc par la résolution d'une équation différentielle sur la variété $\frac{M^2}{\rho} + p(\rho, c) = cste$:

$$\begin{cases} M c_x = \rho(\rho^{2(\gamma-2)} - c) \\ \left(\frac{M^2}{\rho} + p(\rho, c) \right)_x = 0 \end{cases} \quad (10.38)$$

Comme le problème (10.38) est fortement couplé par le terme de pression, il ne peut être intégré en général. Dans le cadre d'un schéma aux interfaces, on décide de se ramener à un processus itératif sur une suite $(\rho^k(x), c^k(x))_{k \geq 0}$. Ce type d'approximation est étudié en détail en annexe A, on se contente ici de le décrire en vue d'une implémentation¹. Soit $(U_j^n)_j$ une approximation de $U(\cdot, n\Delta t)$,

- on fixe pour chaque $j \in \mathbb{Z}$ un état d'équilibre $\mu_j = \mu(\rho_j^n)$

- on fixe $\rho^0(x) \equiv \rho_L$, $c^0(x) \equiv c_L$ et on résout en forme fermée la première égalité de (10.38). On trouve: $c^1(x) = \mu_j(1 - e^{-(\rho_L x/M)}) + c_L e^{-(\rho_L x/M)}$.

- on projette $c^1(\Delta x/\epsilon)$ sur la variété localement définie par la seconde égalité de (10.38) au moyen d'un algorithme de Newton. Ceci nous donne une valeur pour $\rho_1(\Delta x/\epsilon)$.

- on itère jusqu'au point fixe:

$$\begin{cases} c^k(\Delta x/\epsilon) = \mu_j \left[1 - e^{-(\rho^{k-1}(\Delta x/\epsilon)\Delta x/\epsilon M)} \right] + c_L e^{-(\rho^{k-1}(\Delta x/\epsilon)\Delta x/\epsilon M)} \\ \frac{M^2}{\rho^k(\Delta x/\epsilon)} + p(\rho^k(\Delta x/\epsilon), c^k(\Delta x/\epsilon)) = \frac{M^2}{\rho_L} + p(\rho_L, c_L) \end{cases} \quad (10.39)$$

On remarque que les relations (10.39) sont inexploitable dès que $(\Delta x M) < 0$ puisque les concentrations approchées vont tendre vers l'infini lorsque $\epsilon \rightarrow 0$. On doit donc se restreindre à des problèmes supersoniques pour lesquels le présent schéma coïncide avec celui de Godunov.

10.4.4 Ecriture et comportement du schéma \llcorner équilibre \lrcorner en régime relaxé

Pour ce type de problèmes, le schéma (10.8) s'écrit de la manière suivante:

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} [F^+(U_j^n) - F^+(U_{j-\frac{1}{2}}^+) + F^-(U_{j+\frac{1}{2}}^-) - F^-(U_j^n)]$$

avec les états à l'interface $U_{j \pm \frac{1}{2}}^\pm$ définis de la façon suivante ($M_{j-1} = (\rho u)_{j-1}^n$):

$$\begin{cases} c_{j-\frac{1}{2}}^+ = \mu(\rho_j^n) \left[1 - e^{-(\rho_{j-\frac{1}{2}}^+ \Delta x/\epsilon M_{j-1})} \right] + c_{j-1}^n e^{-(\rho_{j-\frac{1}{2}}^+ \Delta x/\epsilon M_{j-1})} \\ \frac{M_{j-1}^2}{\rho_{j-\frac{1}{2}}^+} + p(\rho_{j-\frac{1}{2}}^+, c_{j-\frac{1}{2}}^+) = \frac{M_{j-1}^2}{\rho_j^n} + p(\rho_j^n, c_{j-1}^n) \end{cases}$$

Ainsi que ($M_{j+1} = (\rho u)_{j+1}^n$):

$$\begin{cases} c_{j+\frac{1}{2}}^- = \mu(\rho_j^n) \left[1 - e^{(\rho_{j+\frac{1}{2}}^- \Delta x/\epsilon M_{j+1})} \right] + c_{j+1}^n e^{(\rho_{j+\frac{1}{2}}^- \Delta x/\epsilon M_{j+1})} \\ \frac{M_{j+1}^2}{\rho_{j+\frac{1}{2}}^-} + p(\rho_{j+\frac{1}{2}}^-, c_{j+\frac{1}{2}}^-) = \frac{M_{j+1}^2}{\rho_j^n} + p(\rho_j^n, c_{j+1}^n) \end{cases}$$

Ceci permet de construire un schéma stable sous la condition CFL et pour tout $\epsilon > 0$ dans le cas d'écoulements supersoniques. Toute la question est maintenant de s'assurer que les approximations numériques issues d'un schéma du type (10.8) pour le problème (10.21) convergent vers la limite asymptotique correcte (10.28) à la limite $\epsilon \rightarrow 0$. A ce propos, on peut déjà consulter [50, 51] pour une étude des schémas plus classiques à splitting d'opérateurs.

1. Comme on l'a fait remarquer dans la section précédente, cet algorithme du point fixe n'est pas l'unique moyen de calculer les états aux interfaces. Son principal avantage réside dans sa facilité de programmation. Il peut être remplacé par un solveur différentiel sophistiqué, voire par un algorithme de Newton dans \mathbb{R}^2 .

Cas-test supersonique Le problème étudié est du type couche-limite supersonique: on propose les données initiales suivantes pour $x \in]0, 1[$:

$$\begin{cases} \rho_0 = 2 \\ c_0 = 0.5 + \frac{\sin(10\pi x)}{5(1+x)} \\ u_0 = 2 \end{cases} \quad \text{avec } \epsilon = 10^{-7} \text{ et } \gamma = 1.1$$

ainsi que la condition-limite à gauche ($x = 0$) supersonique:

$$\begin{cases} \rho_l = 1.5 \\ c_l = 0.8 \\ u_l = 4 \end{cases}$$

Pour différentes valeurs de γ , on constate une bonne consistance du schéma non-conservatif avec le système réduit: les figures 10.23, 10.24 et 10.25 représentent les pressions, vitesses, ainsi que la concentration massique en $T = 0.2$ pour $\gamma = 1.1$. La courbe en trait plein montre les résultats issus du schéma `''équilibre''`, celle en pointillés ceux d'un schéma à décomposition de flux classique pour le système relaxé correspondant à la limite $\epsilon \rightarrow 0$. Les paramètres choisis sont $\Delta x = 0.02$ et $\Delta t = 0.002$ pour les deux algorithmes. On constate une précision moyenne sur la densité, et une approximation à peine meilleure sur la vitesse. La diffusion numérique reste importante (ceci est aussi observé pour les schémas à pas fractionnaires [51]). Le point important est bien sûr d'obtenir un comportement en accord avec le régime prescrit par le système asymptotique, ce qui passe par une concentration stationnaire non-oscillante.

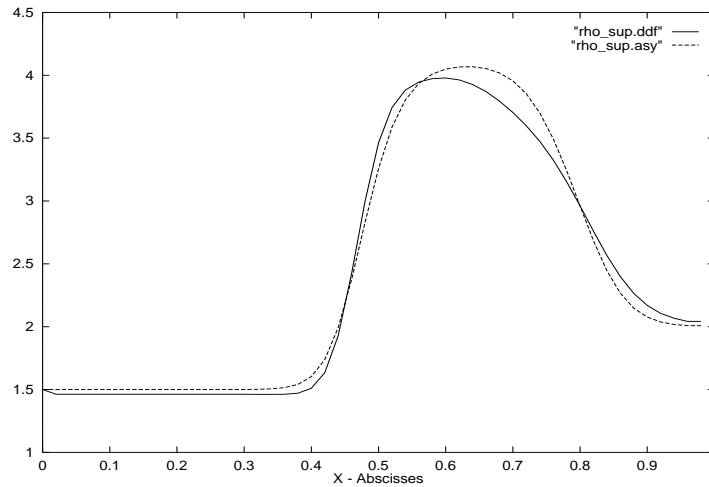


Figure 10.23: Densité ($\epsilon = 10^{-7}$) et densité asymptotique (schéma à décomposition de flux)

10.5 Conclusion

On a présenté dans ce chapitre un schéma nouveau, construit sur la base d'une décomposition de flux afin de pallier au manque de robustesse du schéma de type Roe non-conservatif de la partie III tout en maintenant ses bonnes propriétés en régime stationnaire. D'autre part, on a évité l'écueil d'un schéma de Godunov qui serait très difficile à mettre en oeuvre au vu des relations de saut non-conservatives à gérer.

Avantages Le schéma ainsi construit est capable de traiter des termes sources raides sans restriction particulière sur le pas de temps (sauf pour le cas de la relaxation) tout en maintenant une intégration en temps explicite. La validation s'est effectuée sur une large plage de problèmes physiquement réalistes (tuyères, tubes à choc cylindriques, diphasique équilibré) avec termes sources de natures différentes (géométrie, perte de charge, recondensation). De plus, les algorithmes qui découlent d'une approche par décomposition de flux présentent une alternative facile à programmer comparativement aux schémas de Roe ou de Godunov.

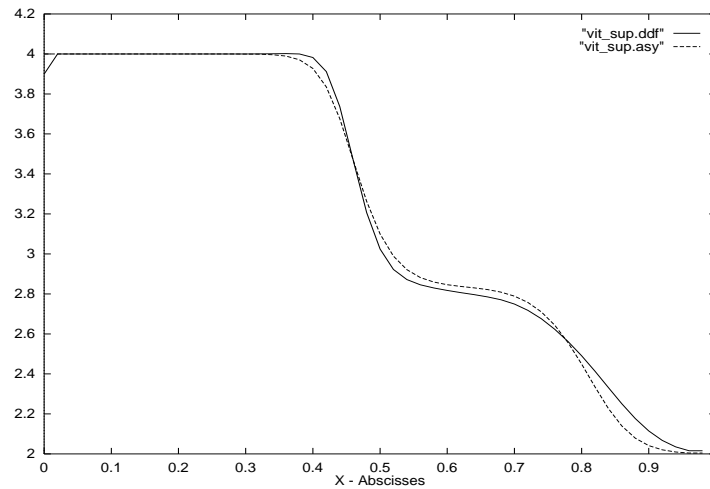


Figure 10.24: Vitesse ($\epsilon = 10^{-7}$) et vitesse asymptotique (schéma à décomposition de flux)

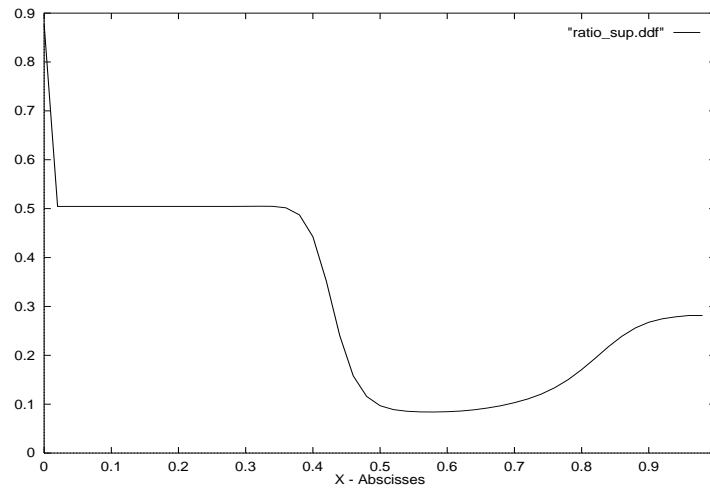


Figure 10.25: Concentration ($\epsilon = 10^{-7}$) (schéma à décomposition de flux)

Limites Cette approche souffre néanmoins de plusieurs limitations, notamment dans le cas de la relaxation raide en régime subsonique. Néanmoins, dans ce cas, il est facile de mettre en place un schéma à pas fractionnaires muni d'un solveur exact du terme source. Comme l'écoulement à l'équilibre est presque totalement imposé par le terme réactionnel, il n'y a pas de balance délicate à trouver avec les termes de flux. La principale difficulté reste bien sûr de trouver une expression des courbes stationnaires du problème. Dans le cas où il n'y a qu'un seul terme source, on peut se ramener à une équation différentielle scalaire. Dans l'autre alternative, il faut intégrer de façon approchée un système différentiel couplé. On a proposé de faire cela par le biais d'un algorithme itératif.

Chapitre 11

Un mot sur les cas de résonance non-linéaire

Dans le contexte des systèmes hyperboliques avec termes sources, l'idée qui est à la base des présents travaux peut s'exprimer simplement: ayant constaté la bonne fiabilité des solveurs de Riemann homogènes, on essaie à toute force de se ramener à une situation où il sera possible de les mettre en oeuvre avec succès. C'est-à-dire qu'il est nécessaire de reformuler les seconds membres comme des "termes de flux" supplémentaires. Voulant récupérer un système homogène, on est en droit de s'attendre à une complication des non-linéarités intervenant dans les opérateurs différentiels du problème modifié. Le premier écueil a été de formuler correctement ces termes se manifestant en général sous la forme de *produits non-conservatifs*: ceci a été résolu dans la majorité des cas par le chemin "équilibre" découvert à la partie III. Néanmoins, il reste à considérer la particularité du problème modifié à être *non-strictement hyperbolique*, c'est-à-dire que des champs VNL peuvent tout à fait interférer localement avec le champ stationnaire induit par le terme de réaction. On est alors dans une situation de *résonance* (cf partie III - Figs 9.2, 9.3) où la solution du problème de Riemann instationnaire devient très singulière (au mieux non-unique). Dans ces cas, le formalisme de [21] doit être abandonné puisqu'il n'existe pas de courbe régulière joignant les états gauche et droits. Il existe des références sur ce type de problèmes (au demeurant ouverts pour la plupart), notamment E. Isaacson et B. Temple [48, 49] ainsi que [45]. Dans ce chapitre, nous suivons la démarche introduite dans [30] qui se borne au calcul de solutions stationnaires. L'objet est bien sûr de contrôler le comportement des schémas introduits précédemment (linéarisation de type Roe ou bien décomposition de flux). Le problème des oscillations numériques derrière les chocs stationnaires [52] ne sera cependant pas abordé ici.

11.1 Exemple théorique: calcul d'un permanent transsonique

11.1.1 Généralités

On considère le système générique modélisant un écoulement de gaz parfait dans un conduit à section variable au moyen des équations d'Euler munies d'un terme de contour. Les notations sont identiques à celles des chapitres précédents (voir par exemple (9.19), (9.20) en partie III et (10.14) en partie IV).

$$\begin{cases} R_t + (Ru)_x = 0 \\ (Ru)_t + (Ru^2 + P)_x - \left[\frac{P}{S}S_x\right]_{\Phi} = 0 \\ (Re)_t + (Ruh)_x = 0 \\ S_t = 0 \end{cases} \quad (11.1)$$

La fonction $x \mapsto S(x)$ modélise le profil de la tuyère étudiée et les variables $R = \rho S$ et $P = pS = (\gamma-1)\rho S(e - \frac{u^2}{2})$ sont des facilités d'écritures largement employées. La construction proposée dans [30] exploite le fait que $S(\cdot)$ est à l'origine une fonction très régulière (sinon le système (11.1) n'a pas de sens puisque la recirculation du fluide ne peut pas être ignorée). C'est au niveau de l'approximation numérique que celle-ci devient *discontinue* pour les besoins d'un schéma de type Godunov (voir aussi un exemple similaire dans [11] avec cette fois une pression hydrostatique). Néanmoins, le point essentiel est d'observer le **caractère physiquement réaliste** des chocs résonants. Par conséquent, ceux-ci ne sont aucunement liés au caractère *artificiellement discontinu* de $S(\cdot)$, ils sont d'origine purement *convective* et leur profil infiniment mince n'est en rien affecté par un terme source au second membre. Ceci pour arriver à la conclusion fondamentale: **les chocs résonants vérifient les conditions de Rankine-Hugoniot du système homogène**

C'est ce point qui va permettre de construire théoriquement des régimes permanents contenant des chocs soniques stationnaires.

11.1.2 Méthode de résolution

D'après l'expression (10.15), les relations de saut pour le système non-conservatif (11.1) régularisé s'écrivent:

$$\begin{cases} (Ru)_L = (Ru)_R \text{ et } h_L = h_R \\ \left(\frac{\varphi(R_R)}{\varphi(R_L)}\right)^{\frac{\gamma}{\gamma-1}} \frac{P_R}{P_L} = \frac{S_R}{S_L} \end{cases} \quad (11.2)$$

avec $\varphi(R) = \frac{R^2}{(R+a)(R-a)}$ et $a = \frac{Ru}{\sqrt{2h}}$. On rappelle par ailleurs que $R > a$ équivaut à $h > u^2/2$. La dernière équation peut se réécrire sous la forme suivante:

$$S(x) = S(0) \cdot \left(\frac{\varphi(R)(0)}{\varphi(R)(x)} \right)^{\frac{\gamma}{\gamma-1}} \frac{P(0)}{P(x)} \text{ à } (Ru) \text{ et } h \text{ constants} \quad (11.3)$$

Ceci est l'expression d'une *courbe de Bernoulli* exprimant la conservation de la charge totale malgré la variation du profil dans lequel s'effectue l'écoulement. Par conséquent, à chaque point $(R(0), S(0)) \in (\mathbb{R}_*^+)^2$, il est possible d'associer une courbe régulière d'états connectables via la relation (11.3). Lorsque l'on est en présence d'un régime résonant, on a affaire à deux états $(R(0), S(0))$ et $(R(x_0), S(x_0))$ qu'il est impossible de joindre de cette façon. Ceci signifie, d'un point de vue physique qu'il existe un choc stationnaire compressif dans la solution séparant la région supersonique en amont de la région subsonique en aval. En ce cas, l'influence du terme source étant nulle dans la zone de transition infiniment mince, on décide de raccorder les deux portions de courbes par une discontinuité entropique vérifiant:

$$\begin{cases} (Ru)_L = (Ru)_R \text{ et } h_L = h_R \\ (Ru^2)_L + P_L = (Ru^2)_R + P_R \text{ et } S_L = S_R \\ P \text{ et } u \text{ décroissantes strictement} \end{cases} \quad (11.4)$$

Par conséquent, il est naturel de considérer la courbe d'états $(R(x), S(x))$ pouvant être joints à la *portion supersonique* de la courbe de Bernoulli (11.3) par un choc du type (11.4). On déduit donc que si celle-ci coupe la portion subsonique en un unique point, alors le régime transsonique discontinu est totalement caractérisé.

11.1.3 Construction d'un permanent transsonique dans un profil divergent

On se restreint à un profil divergent (i.e. $\frac{\partial S(x)}{\partial x} \geq 0$) pour un raison simple: d'après [30], il est tout à fait possible de construire de cette façon des chocs stationnaires entropiques non-physiques. Par exemple, la présence d'un choc dans la partie convergente d'une tuyère de Laval n'est pas à exclure; ce qui est incohérent avec les observations. Ce phénomène peut aussi être mis en évidence par expérimentations numériques [30].

Lemme IV.4 *A R supersonique donné et (Ru) , h constants, il existe un unique $\bar{R} > R$ qui puisse lui être joint par un choc stationnaire entropique vérifiant (11.4).*

De plus, si on note $R_{sonic}^2 \stackrel{def}{=} \frac{\gamma+1}{\gamma-1} \cdot \frac{(Ru)^2}{2h}$, alors on a: $\bar{R} = \frac{R_{sonic}^2}{R}$.

Démonstration

D'après les conditions (11.4), il faut vérifier la relation de saut:

$$\frac{(Ru)^2}{\bar{R}} + P(\bar{R}) = \frac{(Ru)^2}{R} + P(R) \text{ avec } P(R) = \frac{\gamma-1}{\gamma} \left(Rh - \frac{(Ru)^2}{2R} \right)$$

D'où il vient une équation du second degré en \bar{R} :

$$\frac{(\gamma-1)h}{\gamma} \bar{R}^2 - \left(\frac{(Ru)^2}{R} + P(R) \right) \bar{R} + \frac{(\gamma+1)(Ru)^2}{2\gamma} = 0$$

Du fait de l'expression de $P(R)$, le coefficient de \bar{R} s'écrit:

$$\frac{(Ru)^2}{R} + P(R) = \frac{(\gamma+1)(Ru)^2}{2\gamma R} + \frac{(\gamma-1)Rh}{\gamma}$$

Le discriminant $\Delta(R)$ est-il positif ?

$$\Delta(R) = \left(\frac{(\gamma+1)(Ru)^2}{2\gamma R} + \frac{(\gamma-1)Rh}{\gamma} \right)^2 - \frac{(\gamma-1)4h}{\gamma} \cdot \frac{(\gamma+1)(Ru)^2}{2\gamma R}$$

Finalement, il vient que:

$$\Delta(R) = \left(\frac{(\gamma+1)(Ru)^2}{2\gamma R} - \frac{(\gamma-1)Rh}{\gamma} \right)^2$$

Cette expression est évidemment positive, et ne s'annule que si R est un point sonique (voir le lemme IV.2). La condition de Lax impose que u décroisse à (Ru) constant: on choisit donc la racine positive de l'équation. Par conséquent, il vient directement que:

$$\bar{R} = \frac{\gamma + 1}{\gamma - 1} \cdot \frac{(Ru)^2}{2hR} \stackrel{\text{def}}{=} \frac{R_{sonic}^2}{R} \quad (11.5)$$

Ceci signifie en particulier que si R appartient à la zone supersonique, alors $\frac{\gamma+1}{\gamma-1} \cdot \frac{u^2}{2} > h$ et on a $\bar{R} > R$. \square

Soit l'état (S_0, R_0) supersonique, il existe une unique courbe de Bernoulli à Ru et h constants passant par ce point du fait de la théorie de Cauchy des équations différentielles. On peut associer de façon univoque à celle-ci une courbe d'états pouvant être lui connectés par un choc stationnaire. Le problème est d'établir que cette dernière coupe la courbe de Bernoulli issue de l'état (S_1, R_1) subsonique en un unique point.

Pour cela, posons $\mathcal{B}(S, R) = S \cdot \varphi(R)^{\frac{\gamma}{\gamma-1}} \cdot P(R)$; l'expression (11.3) devient:

$$\mathcal{B}(S(x), R(x)) = \mathcal{B}(S(0), R(0))$$

Donc, l'équation des courbes de Bernoulli peut se réécrire:

$$\mathcal{B}(S, R) = \mathcal{B}(S_0, R_0) \text{ et } \mathcal{B}(S, R) = \mathcal{B}(S_1, R_1)$$

Pour finir, on adopte la notation $\psi(R) = \varphi(R)^\gamma P(R)^{\gamma-1}$ déjà utilisée au chapitre précédent. Cette fonction ψ est strictement croissante pour les états supersoniques, et strictement décroissante pour les états subsoniques (voir lemme IV.2). Celle-ci nous donne:

$$\mathcal{B}(S, R) = S \psi(R)^{\frac{1}{\gamma-1}}$$

Au bout du compte, on obtient l'équation des courbes d'équilibres pour (11.1) issues de (S_0, R_0) et (S_1, R_1) :

$$S = \frac{\mathcal{B}(S_0, R_0)}{\psi(R)^{\frac{1}{\gamma-1}}} \text{ et } S = \frac{\mathcal{B}(S_1, R_1)}{\psi(R)^{\frac{1}{\gamma-1}}}$$

Le premier résultat est le suivant:

Lemme IV.5 *Etant donnés $(R_0, R_1) \in]a, R_{sonic}[\times]R_{sonic}, R_{sonic}^2/a[$ et $(S_0, S_1) \in (\mathbb{R}_*^+)^2$ tels que $\mathcal{B}(S_0, R_0) < \mathcal{B}(S_1, R_1)$, il existe un unique triplet $(S, R, \bar{R}) \in \mathbb{R}_*^+ \times]a, R_{sonic}[\times]R_{sonic}, R_{sonic}^2/a[$ tels que l'on ait:*

$$\begin{cases} \mathcal{B}(S, R) = \mathcal{B}(S_0, R_0) \\ \mathcal{B}(S, \bar{R}) = \mathcal{B}(S_1, R_1) \\ \bar{R} = R_{sonic}^2/R \end{cases}$$

Démonstration

Il s'agit de vérifier plusieurs choses. Tout d'abord, on s'assure qu'à l'abscisse $R = R_{sonic}$, la courbe passant par (S_1, R_1) subsonique est au-dessus de celle passant par (S_0, R_0) supersonique. Ensuite, on va montrer que la valeur $\mathcal{B}(S, \bar{R})$ est strictement croissante et non-bornée le long de la ligne de chocs stationnaires. Cela impliquera un unique point d'intersection entre celle-ci et la portion de courbe subsonique.

Au point sonique $R = R_{sonic}$, les équations des deux courbes de Bernoulli imposent:

$$S = \frac{S_0 [\psi(R_0)]^{1/(\gamma-1)}}{[\psi(R_{sonic})]^{1/(\gamma-1)}} \text{ et } \tilde{S} = \frac{S_1 [\psi(R_1)]^{1/(\gamma-1)}}{[\psi(R_{sonic})]^{1/(\gamma-1)}}$$

Par conséquent, l'hypothèse $\mathcal{B}(S_0, R_0) < \mathcal{B}(S_1, R_1)$ implique que $S < \tilde{S}$, donc la courbe passant par (S_1, R_1) est au-dessus de celle passant par (S_0, R_0) .

Il s'agit maintenant d'établir clairement l'existence d'un unique point d'intersection entre la courbe subsonique et la ligne de choc. Soit (S, \bar{R}) un point sur cette dernière, on a:

$$\mathcal{B}(S, \bar{R}) = S [\psi(\bar{R})]^{1/(\gamma-1)} \quad (11.6)$$

Pour obtenir une expression de $\mathcal{B}(S, \bar{R})$, on remarque que le choc stationnaire se fait à S constant. D'autre part, l'antécédent supersonique de \bar{R} (i.e. $R(\bar{R}) = R_{sonic}^2/\bar{R}$) appartient à la portion de courbe supersonique issue de (S_0, R_0) . On a donc:

$$\mathcal{B}(S_0, R_0) = \mathcal{B}(S, R(\bar{R}))$$

En développant les expressions précédentes, il vient:

$$S = S_0 \left[\frac{\psi(R_0)}{\psi(R(\bar{R}))} \right]^{1/(\gamma-1)} \quad (11.7)$$

En reportant (11.7) dans l'expression (11.6), on obtient:

$$\mathcal{B}(S, \bar{R}) = S_0 \left[\frac{\psi(R_0) \cdot \psi(\bar{R})}{\psi(R(\bar{R}))} \right]^{1/(\gamma-1)} \quad (11.8)$$

Il faut donc établir dans un premier temps que:

$$\forall \bar{R} > R_{sonic}, \frac{\partial \mathcal{B}(S, \bar{R})}{\partial \bar{R}} > 0$$

On développe les expressions intervenant dans (11.8) à savoir:

$$\begin{cases} \psi(\bar{R}) = \bar{R}^{1-\gamma} - \frac{(Ru)^2}{2h} \bar{R}^{-(1+\gamma)} = \left(\frac{1}{\bar{R}} \right)^{\gamma-1} \left(1 - \frac{\gamma-1}{\gamma+1} \cdot \frac{R_{sonic}^2}{\bar{R}^2} \right) \\ \psi(R(\bar{R})) = \left(\frac{\bar{R}}{R_{sonic}^2} \right)^{\gamma-1} \left(1 - \frac{\gamma-1}{\gamma+1} \cdot \frac{\bar{R}^2}{R_{sonic}^2} \right) \end{cases}$$

Posons $\alpha = \frac{\gamma-1}{\gamma+1}$ et $y = \frac{\bar{R}^2}{R_{sonic}^2} > 1$, alors la fonction $\frac{\psi(\bar{R})}{\psi(R(\bar{R}))}$ peut se réécrire:

$$\frac{\psi(\bar{R})}{\psi(R(\bar{R}))} = \left(\frac{R_{sonic}^2}{\bar{R}^2} \right)^{\gamma-1} \left(\frac{1 - \alpha R_{sonic}^2/\bar{R}^2}{1 - \alpha \bar{R}^2/R_{sonic}^2} \right) = \left(\frac{1}{y} \right)^{\gamma-1} \left(\frac{1 - \alpha/y}{1 - \alpha y} \right) \quad (11.9)$$

On introduit donc la fonction:

$$f(y) = \frac{1 - \alpha/y}{y^{\gamma-1} - \alpha y^\gamma}$$

On voit qu'on ne changera pas de signe en dérivant par rapport à y puisque $\frac{\partial}{\partial \bar{R}} = \frac{2\bar{R}}{R_{sonic}^2} \frac{\partial}{\partial y}$.

D'autre part, la restriction à \mathbb{R}^+ de la fonction $z \mapsto z^{1/(\gamma-1)}$ est strictement croissante pour $\gamma > 1$. Pour avoir le signe de $\partial \mathcal{B}(S, \bar{R})/\partial \bar{R}$, il est donc suffisant de montrer que $f'(y) > 0$ pour $y > 1$. Comme on a affaire à une fraction rationnelle en y , le signe de la dérivée est donné par celui du numérateur.

$$N(y) = (\alpha/y^2)(y^{\gamma-1} - \alpha y^\gamma) - ((\gamma-1)y^{\gamma-2} - \alpha \gamma y^{\gamma-1})(1 - \alpha/y)$$

C'est-à-dire:

$$sgn[N(y)] = sgn[\alpha \gamma y^2 - (\alpha^2(\gamma+1) + (\gamma-1)y + \alpha \gamma)]$$

Ceci est une simple équation du second degré en $y > 1$ dont on connaît explicitement les solutions. Pour cela, on en réécrit les coefficients:

$$\begin{cases} (\alpha^2(\gamma+1) + (\gamma-1)) = \left(\frac{\gamma-1}{\gamma+1} \right)^2 (\gamma+1) + \left(\frac{\gamma-1}{\gamma+1} \right) (\gamma-1) = \frac{2\gamma(\gamma-1)}{\gamma+1} \\ \alpha \gamma = \frac{\gamma(\gamma-1)}{\gamma+1} \end{cases}$$

Il est donc immédiat que le discriminant est nul et l'unique racine double est $y = 1$. D'autre part, comme $\alpha \gamma > 0$, le numérateur est toujours strictement positif dès que $y > 1$. La fonction $y \mapsto f(y)$ est donc strictement croissante. Ceci implique bien que:

$$\forall \bar{R} > R_{sonic}, \frac{\partial \mathcal{B}(S, \bar{R})}{\partial \bar{R}} > 0$$

On remarque ensuite que dans l'expression (11.9), le dénominateur $1 - \alpha y$ s'annule pour

$$y = \frac{\bar{R}^2}{R_{sonic}^2} = \frac{\gamma + 1}{\gamma - 1}$$

Du fait de la valeur de R_{sonic}^2 , ceci correspond à:

$$\frac{\bar{R}^2}{R_{sonic}^2} = \bar{R}^2 \cdot \left(\frac{\gamma + 1}{\gamma - 1} \cdot \frac{(Ru)^2}{2h} \right)^{-1} = \frac{\gamma + 1}{\gamma - 1}$$

Par conséquent, cela donne:

$$\bar{R} = \frac{R_{sonic}^2}{a}$$

Cela signifie donc que:

$$\mathcal{B}(S, \bar{R}) \rightarrow +\infty \text{ lorsque } \bar{R} \rightarrow R_{sonic}^2/a$$

Il y a donc nécessairement intersection entre la ligne de chocs et la portion subsonique de courbe de Bernoulli $\mathcal{B}(S, R) = \mathcal{B}(S_1, R_1) < +\infty$ dans l'intervalle $[R_{sonic}, R_{sonic}^2/a[$ puisqu'au point sonique, la constante de Bernoulli sur la ligne de chocs vaut $\mathcal{B}(S_0, R_0) < \mathcal{B}(S_1, R_1)$.

L'unicité est une conséquence de la stricte monotonie de la fonction $\bar{R} \mapsto \mathcal{B}(S, \bar{R})$ le long de la ligne de chocs. En effet, comme la courbe d'états subsonique est définie par $\mathcal{B}(S, R) = \mathcal{B}(S_1, R_1)$, on déduit que ces deux courbes ne se coupent qu'au plus en un point. \square

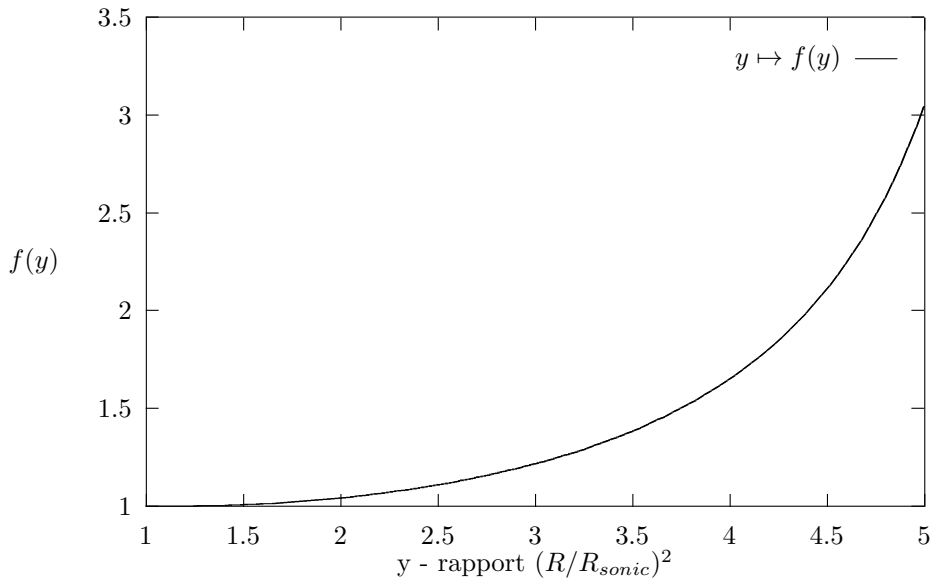


Figure 11.1: Allure de la fonction $y \mapsto f(y)$ pour $\gamma = 1.4$ et $a = 1/6$

Théorème IV.3 *Etant donnés (S_0, R_0) et (S_1, R_1) , tels que:*

$$\frac{\mathcal{B}(S_1, R_1)}{\mathcal{B}(S_0, R_0)} \in \left] 1, \left(\frac{\psi(R_1)}{\psi(R_{sonic}^2/R_1)} \right)^{1/(\gamma-1)} \right[\quad (11.10)$$

il existe un unique triplet $(S, R, \bar{R}) \in]S_0, S_1[\times]a, R_0[\times]R_{sonic}, R_1[$ tels que l'on aît:

$$\begin{cases} \mathcal{B}(S, R) = \mathcal{B}(S_0, R_0) \\ \mathcal{B}(S, \bar{R}) = \mathcal{B}(S_1, R_1) \\ \bar{R} = R_{sonic}^2/R \end{cases}$$

Démonstration

Il s'agit juste de s'assurer que sous l'hypothèse supplémentaire (11.10), le choc sonique se trouve bien dans le domaine. On remarque immédiatement que si on a:

$$\mathcal{B}(S_1, R_1) < \left(\frac{\psi(R_1)}{\psi(R_{sonic}^2/R_1)} \right)^{1/(\gamma-1)} \cdot \mathcal{B}(S_0, R_0)$$

Alors, cela implique:

$$S_1[\psi(R_1)]^{1/(\gamma-1)} < \left(\frac{\psi(R_1)}{\psi(R_{sonic}^2/R_1)} \right)^{1/(\gamma-1)} \cdot S_0[\psi(R_0)]^{1/(\gamma-1)}$$

Et donc:

$$S_1 \cdot [\psi(R_{sonic}^2/R_1)]^{1/(\gamma-1)} < S_0[\psi(R_0)]^{1/(\gamma-1)}$$

Reprenons l'expression (11.8) donnant la valeur de $\mathcal{B}(S, \bar{R})$:

$$\mathcal{B}(S, \bar{R}) = S_0 \left[\frac{\psi(R_0) \cdot \psi(\bar{R})}{\psi(R_{sonic}^2/\bar{R})} \right]^{1/(\gamma-1)}$$

Plaçons-nous au point de la ligne de chocs déterminé par $\bar{R} = R_1$.

La constante de Bernoulli correspondante vaut:

$$\begin{aligned} \mathcal{B}(S, R_1) &= S_0 \left[\frac{\psi(R_0) \cdot \psi(R_1)}{\psi(R_{sonic}^2/R_1)} \right]^{1/(\gamma-1)} \\ &> S_1 \left[\frac{\psi(R_{sonic}^2/R_1) \cdot \psi(R_1)}{\psi(R_{sonic}^2/R_1)} \right]^{1/(\gamma-1)} \\ &> \mathcal{B}(S_1, R_1) \end{aligned}$$

Alors ceci implique directement qu'au point de la ligne de choc correspondant à $\bar{R} = R_1$, la valeur de la constante de Bernoulli est strictement supérieure à $\mathcal{B}(S_1, R_1)$. On sait aussi que $\mathcal{B}(S, \bar{R})$ est strictement croissante avec \bar{R} le long de la ligne de chocs.

Par conséquent, l'intersection a lieu pour une valeur \bar{R} subsonique et strictement inférieure à R_1 . De ce fait, comme pour la zone subsonique, ψ croît quand R décroît (voir figure 10.2), on déduit qu'au point d'intersection correspond une valeur de S strictement inférieure à S_1 . \square

On conclut donc au caractère "bien posé" du problème stationnaire aux limites pour (11.1) sous les hypothèses (11.10). Le procédé de construction de la solution nous fournit un moyen fiable pour obtenir un profil permanent exact grâce auquel il est possible de tester les schémas "équilibre".

11.2 Expériences numériques en régime résonant

11.2.1 Solution exacte du cas-test étudié

Pour fixer les idées, on choisit d'appliquer ces résultats au cas de la tuyère divergente proposée dans [87]. L'équation en est:

$$x \mapsto S(x) = 1.398 + 0.347 \tanh(8x - 4) \quad (11.11)$$

Le profil est représenté en partie III à la figure 9.5. Ceci permet de construire le régime permanent pour (11.1) au moyen d'un graphique du type:

La solution stationnaire est donc composée de deux portions de courbes de Bernoulli régulières passant par deux points donnés (S_0, R_0) et (S_1, R_1) correspondant chacune à une zone où l'écoulement est supersonique (en amont) puis subsonique (en aval). On sait aussi que le raccord entre ces deux courbes disjointes ne peut se faire que par un choc sonique à S constant. Comme les quantités $\mathcal{B}(S, R)$ varient de façon strictement monotone le long de cette ligne de chocs, on obtient donc les valeurs de la densité de part et d'autre du choc ainsi que

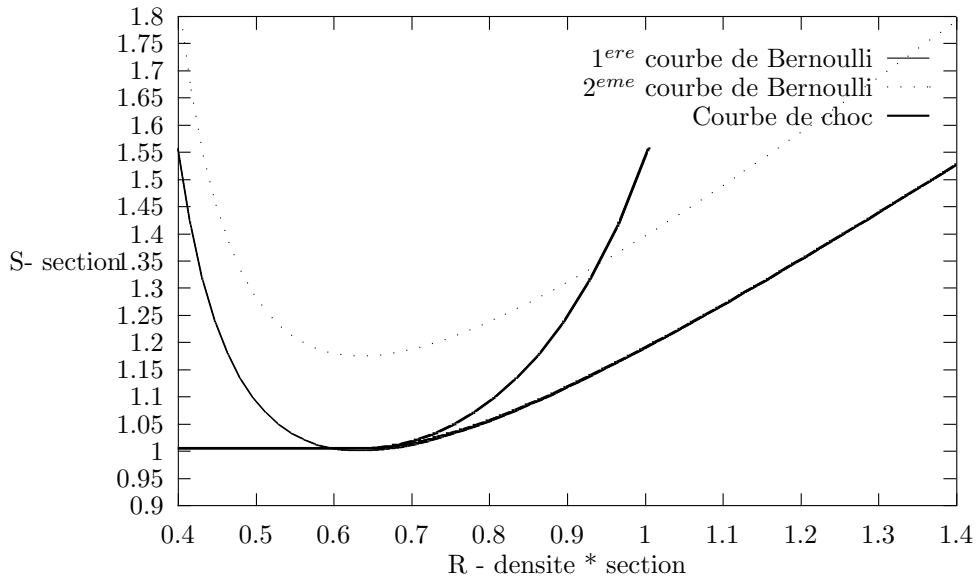


Figure 11.2: Régime permanent pour la tuyère divergente

l'endroit où il se produit si $x \mapsto S(x)$ est bijective. La figure 11.2 correspond à des valeurs gauche et droite proposées dans [6] et données par:

$$\left. \begin{array}{l} \rho_L = 0.502 \\ u_L = 1.299 \\ p_L = 0.381 \\ \rho_R = 0.776 \end{array} \right\} \text{ce qui donne} \left\{ \begin{array}{l} R_0 = 0.53 \\ R_1 = 1.36 \\ (Ru) = 0.69 \\ h = 3.49 \\ S_0 = 1.05 \\ S_1 = 1.75 \end{array} \right. \quad (11.12)$$

On est dans un cas de figure couvert par le théorème IV.3 puisque $]1, 1.98[\ni \mathcal{B}(S_1, R_1)/\mathcal{B}(S_0, R_0) \simeq 1.2$. L'intersection entre la ligne de chocs et la portion subsonique issue de (S_1, R_1) a lieu au point:

$$S = 1.345, R = 0.475, \bar{R} = 0.945, \rho = 0.353, \bar{\rho} = 0.702 \quad (11.13)$$

11.2.2 Comportement du schéma de type Roe

Le schéma obtenu au chapitre 3 de la partie précédente n'est pas conçu pour supporter les régimes résonants puisqu'en ces points singuliers, la définition du chemin régularisant le terme non-conservatif présent dans (11.1) devient caduque. On persiste néanmoins à tester $\ddot{\text{en}} \text{ aveugle}$ ce type de schéma sur le problème donné par (11.1), (11.11), (11.12). Le schéma témoin sera celui proposé par P. Glaister [32]. On présente un transitoire en ρ et u obtenu pour $T = 1.4$ aux figures 11.3 et 11.4: on peut observer que les solutions sont très proches pour $t \in [0, 1.4]$.

Enfin, le régime permanent est atteint pour $T = 14$: les figures 11.5 à 11.8 présentent la densité, la vitesse, le débit et l'enthalpie totale convergées. On constate que ces deux dernières variables ne sont pas constantes à l'endroit du choc stationnaire pour les deux algorithmes proposés. Néanmoins, bien que l'approximation non-conservative soit moins fiable sur ce cas-test défavorable, les écarts sur la densité et la vitesse sont très faibles et les quantités conservées (débit, enthalpie totale) sont mieux respectées par le schéma non-conservatif (les pics sont plus importants mais les valeurs sont à peu près maintenues en aval du choc sonique). Bien sûr, si on compare avec les valeurs exactes pour la densité, on constate que les valeurs extrémales sont excessives, mais les erreurs restent faibles. On peut donc conclure que les valeurs obtenues sont bien concordantes avec (11.13). Les paramètres utilisés sont $\Delta x = 0.02$, $\Delta t = 0.007$, pour une CFL maximale de 0.92.

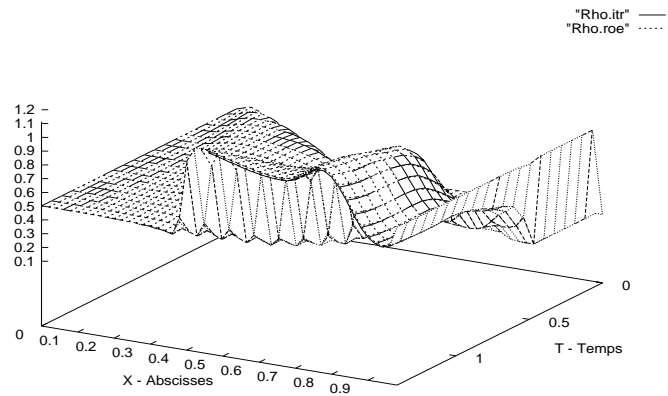


Figure 11.3: Densité transitoire (schéma non-conservatif linéarisé/pas fractionnaires)

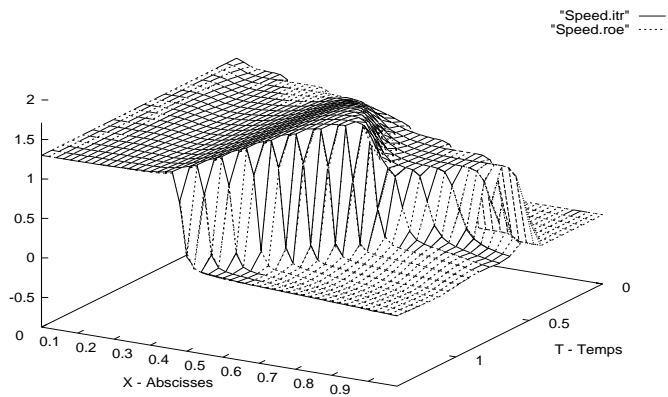


Figure 11.4: Vitesse transitoire (schéma non-conservatif linéarisé/pas fractionnaires)

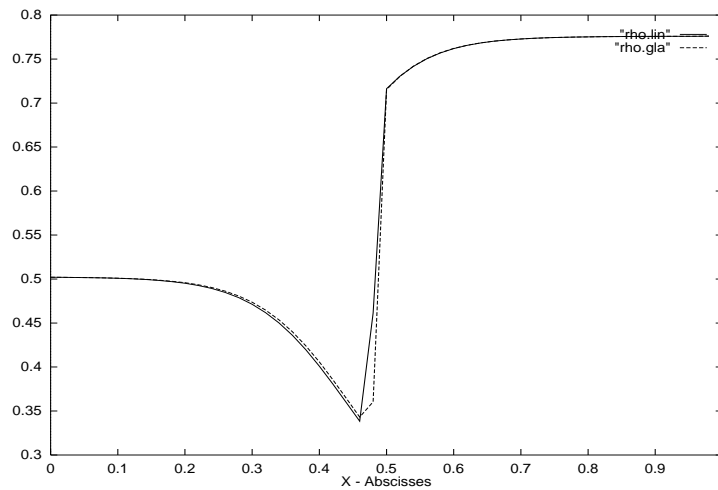


Figure 11.5: Densité stationnaire (schéma non-conservatif linéarisé/pas fractionnaires)

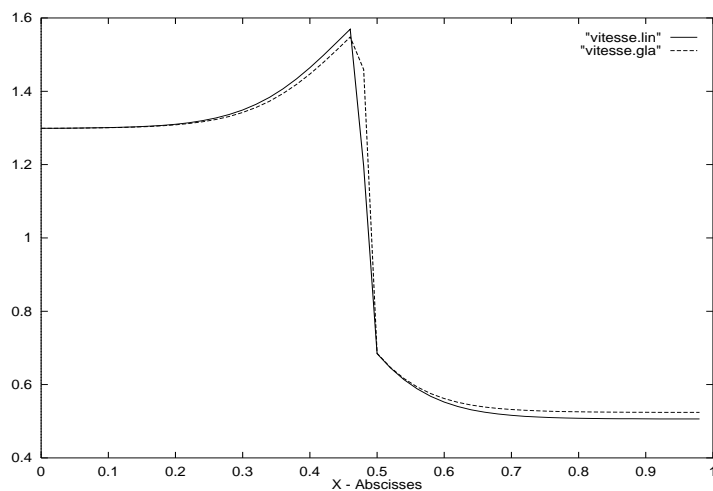


Figure 11.6: Vitesse stationnaire (schéma non-conservatif linéarisé/pas fractionnaires)

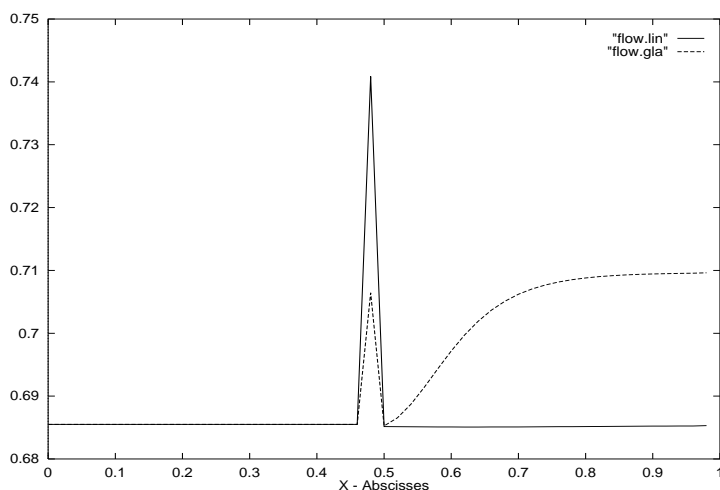


Figure 11.7: Débit stationnaire (schéma non-conservatif linéarisé/pas fractionnaires)

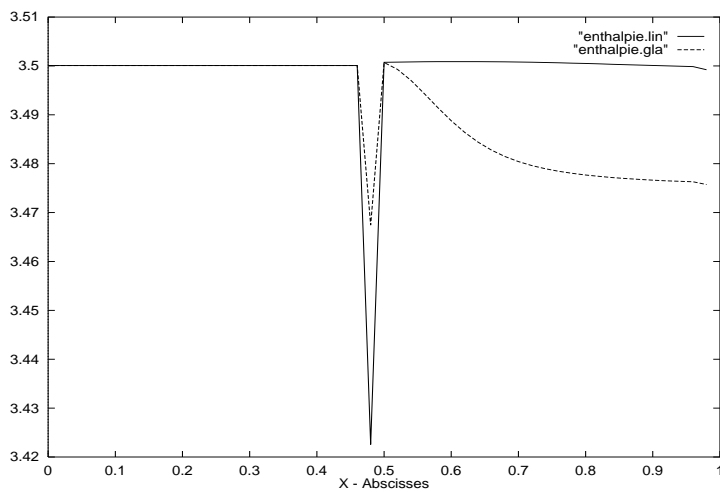


Figure 11.8: Enthalpie totale stationnaire (schéma non-conservatif linéarisé/pas fractionnaires)

11.2.3 Comportement du schéma à décomposition de flux

On rappelle l'écriture du schéma à décomposition de flux:

$$U_j^{n+1} = \frac{U^+ + U^-}{2} = U_j^n - \frac{\Delta t}{\Delta x} [F^+(U_j^n) - F^+(U_{j-\frac{1}{2}}^+) + F^-(U_{j+\frac{1}{2}}^-) - F^-(U_j^n)]$$

avec les relations servant à déterminer l'état gauche $U_{j-\frac{1}{2}}^+$:

$$\psi(R_{j-\frac{1}{2}}^+) = \left(\frac{S(j\Delta x)}{S((j-1)\Delta x)} \right)^{1-\gamma} \psi(R_{j-1}) \text{ avec } (Ru)_{j-\frac{1}{2}}^+ = (Ru)_{j-1} \text{ et } h_{j-\frac{1}{2}}^+ = h_{j-1}$$

et l'état droit $U_{j+\frac{1}{2}}^-$:

$$\psi(R_{j+\frac{1}{2}}^-) = \left(\frac{S(j\Delta x)}{S((j+1)\Delta x)} \right)^{1-\gamma} \psi(R_{j+1}) \text{ avec } (Ru)_{j+\frac{1}{2}}^- = (Ru)_{j+1} \text{ et } h_{j+\frac{1}{2}}^- = h_{j+1}$$

De par sa définition même, ce schéma développé au chapitre précédent ne peut être implémenté tel quel dans le cadre d'un pareil problème. En effet, comme les états à l'interface $U_{j\pm\frac{1}{2}}^\pm$ sont calculés à partir des courbes de Bernoulli, il est clair que certaines relations de saut généralisées vont être impossibles à réaliser. Celles-ci correspondent à une transition sonique dans laquelle la courbe à l'équilibre devient discontinue. Numériquement, cela se traduit par une dérive irrémédiable de l'algorithme de Newton qui essaie vainement de résoudre l'équation (10.16). Le remède que l'on a mis en oeuvre pour obtenir les résultats présentés ici s'inspire du procédé utilisé dans la construction du permanent théorique. Celui-ci consiste à forcer la présence d'un choc sonique vérifiant (11.4) en bordure des mailles où l'algorithme de Newton pose problème. C'est-à-dire:

$$\bar{R} = \frac{R_{sonic}^2}{R} \text{ à } (Ru), h \text{ constants}$$

Cette idée a aussi été utilisée dans [13] pour montrer la convergence d'un schéma de Glimm pour la restriction isentropique du système instationnaire (11.1).

On répète donc les mêmes tests numériques que dans le paragraphe précédent. On constate encore une dérive du schéma non-conservatif à ceci près: les pics dans les quantités conservées sont plus modérés que ceux du schéma de type Roe non-conservatif (voir les figures 11.13 et 11.7, 11.14 et 11.8) et donc les états de part et d'autre du choc stationnaire sont plus précis. Malheureusement, les paliers qui devraient rester constants ne sont pas maintenus en aval du choc sonique. D'autre part, en régime transitoire, on observe sur les itérations en temps (voir figures 11.9 et 11.10 un défaut de vitesse de propagation du choc sonique ainsi que des états érronés en aval de celui-ci lorsque l'on compare la solution du schéma à décomposition de flux avec celui de P. Glaister [32]. Ceci n'est pas sans rappeler les observations numériques reportées dans [24] à propos des discrétisations totalement non-conservatives pour le système d'Euler. Les chocs progressent à une vitesse incorrecte et l'état en aval est incorrect: la régularisation du produit non-conservatif est donc mauvaise. Le schéma témoin et les paramètres utilisés sont bien sûr identiques à savoir $\Delta x = 0.02$ et $\Delta t = 0.007$.

11.3 Conclusion

On a testé dans ce chapitre le comportement des schémas à l'équilibre sur un problème pathologique comprenant un stationnaire discontinu. Signalons que dans ce cas, la régularisation proposée pour les termes non-conservatifs devient caduque puisque la jacobienne des flux devient sigulière. On est en situation de résonance. Pour le problème modèle abordé ici, il est possible de construire la solution théorique stationnaire afin de valider précisément les schémas. On constate une nette supériorité du schéma de type Roe construit à la partie précédente. Néanmoins, la différence avec le schéma à décomposition de flux n'est pas flagrante en régime stationnaire. Le choc sonique issu de celui-ci est plus étalé du fait de sa plus grande diffusion numérique. Pour ce qui est du transitoire, on a choisi comme schéma témoin celui de P. Glaister [32] et on constate des écarts importants avec l'approche par décomposition de flux. Celle-ci comporte des termes non-conservatifs régularisés de façon approximative au voisinage des points soniques, ce qui se traduit par des vitesses de propagation érronées et des états en aval incorrects. Par contre, le schéma de type Roe semble fiable, malgré la perte de justification théorique.

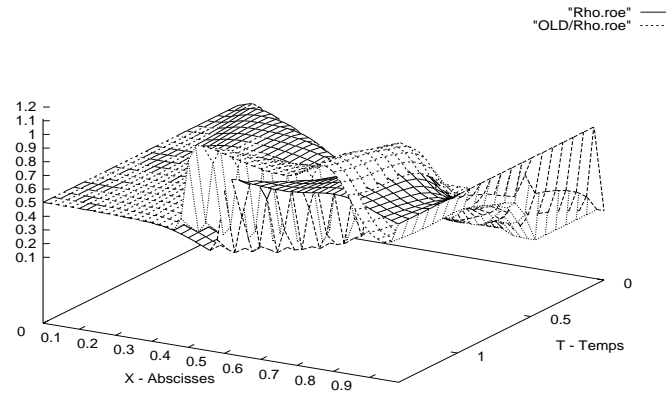


Figure 11.9: Densité transitoire (schéma non-conservatif/pas fractionnaires)

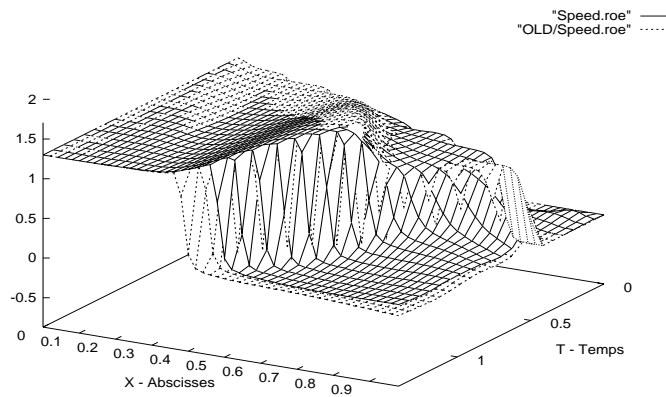


Figure 11.10: Vitesse transitoire (schéma non-conservatif/pas fractionnaires)

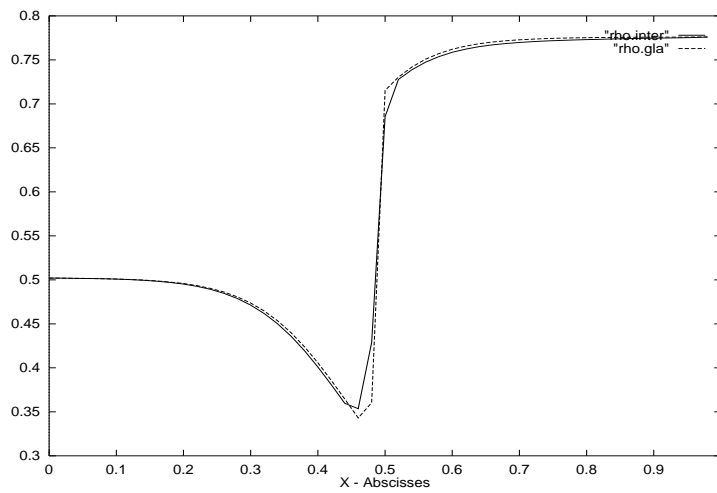


Figure 11.11: Densité stationnaire (schéma non-conservatif/pas fractionnaires)

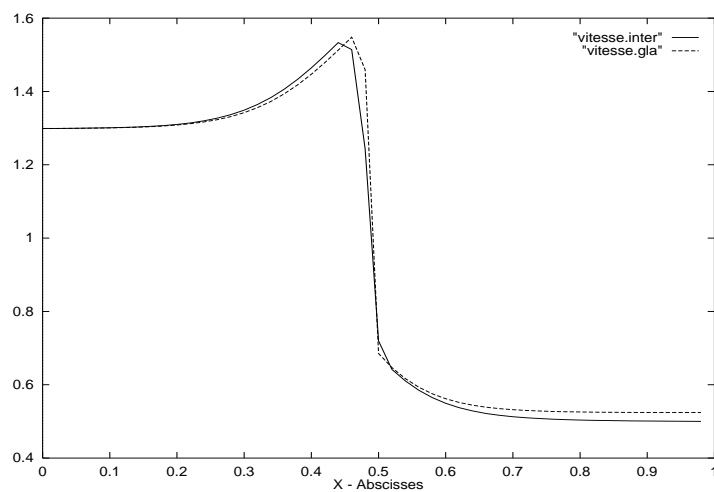


Figure 11.12: Vitesse stationnaire (schéma non-conservatif/pas fractionnaires)

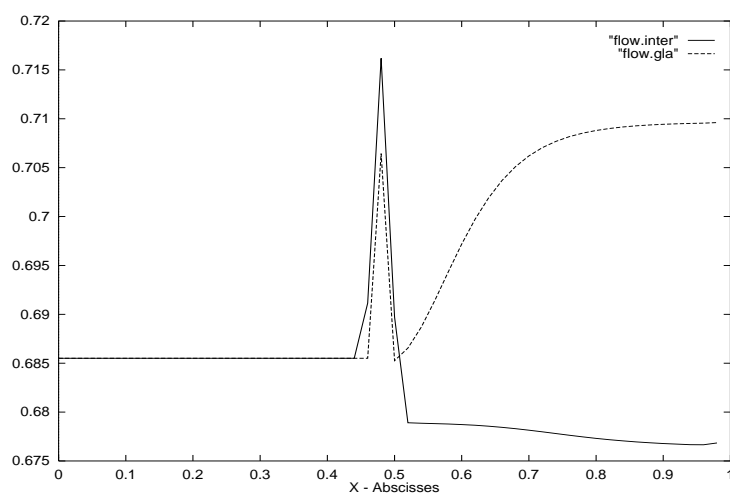


Figure 11.13: Débit stationnaire (schéma non-conservatif/pas fractionnaires)

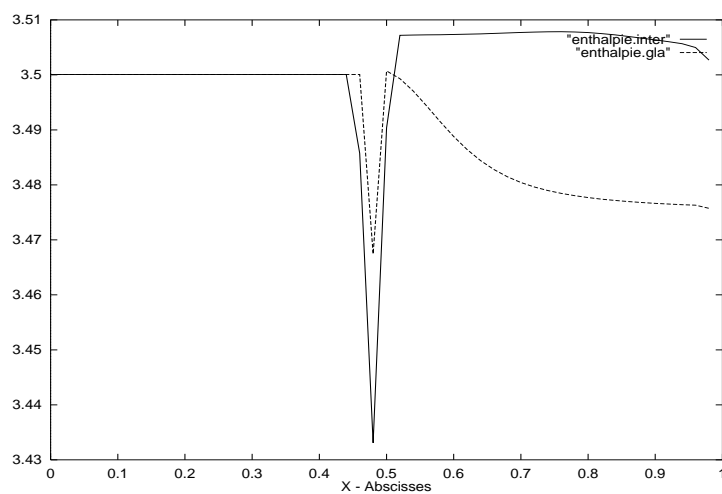


Figure 11.14: Enthalpie totale stationnaire (schéma non-conservatif/pas fractionnaires)

Chapitre 12

Extension à des problèmes bidimensionnels

Ce chapitre présente des applications bidimensionnelles pour la classe de schémas développée jusqu'ici. C'est en quelque sorte l'aboutissement de ce travail. On va tout d'abord introduire des éléments de théorie 2D pour les problèmes hyperboliques dans le cadre des lois scalaires pour lesquelles on dispose de la théorie de Kruzkov [53], puis pour les systèmes. Dans le cas de ceux-ci, les rappels suivent de près le livre de D. Serre [76], notamment pour la définition de l'hyperbolicité. Le principal problème auquel il faut faire face est l'adaptation des relations de saut non-conservatives liées à la formulation des termes sources. En effet, en 1D, le choix de la fonction $x \mapsto a^{\Delta x}(x)$ (7.4) est clair. Pour des équations de type divergence, la question est nettement plus ouverte et plusieurs choix sont possibles. C'est pourquoi on présentera pour finir un cas-test pour le modèle diphasique équilibré en dimension deux réunissant plusieurs termes sources en régime raide.

12.1 Introduction aux problèmes hyperboliques bidimensionnels

12.1.1 Le cas des lois scalaires

Généralités Soit le problème de Cauchy dans $\mathbb{R}^2 \times \mathbb{R}^+$ pour:

$$\begin{cases} u_t + f(u)_x + g(u)_y = h(u) \\ u(x, y, 0) = u_0(x, y) \text{ dans un espace fonctionnel adéquat} \end{cases} \quad (12.1)$$

La solution de cette équation est soumise au même processus d'explosion géométrique en temps fini que celui décrit à la partie I. Pour s'en convaincre, il suffit de choisir $f = g$ dans (12.1). On effectue ensuite le changement de variables $x' = x + y$, $y' = x - y$ et on obtient une équation 1D:

$$u_t + 2f(u)_{x'} = h(u)$$

Celle-ci est du type étudié précédemment; c'est donc en un sens faible qu'il est nécessaire de chercher des solutions pour (12.1). C'est-à-dire dans un espace où il est possible de donner un sens à la dérivée d'une fonction discontinue. Une *solution faible* pour (12.1) sera donc une distribution vérifiant pour toute fonction-test $\varphi \in C_0^\infty(\mathbb{R}^2 \times \mathbb{R}^+)$:

$$\int_{\mathbb{R}^2 \times \mathbb{R}^+} u \varphi_t + f(u) \varphi_x + g(u) \varphi_y + h(u) \varphi dx dy dt + \int_{\mathbb{R}^2} u_0 \varphi dx dy = 0 \quad (12.2)$$

Relations de Rankine-Hugoniot Soit Σ une surface régulière de discontinuité dans $\mathbb{R}^2 \times \mathbb{R}^+$ séparant deux portions de solution C^1 . On choisit un point $M(x, y, t) \in \Sigma$ et B une boule centrée en M séparée en deux composantes connexes B^\pm par Σ . On désigne par $\vec{n} = (n_t, n_x, n_y)$ la normale à Σ orientée dans le sens des t croissants. Si l'on applique la formule de Green à la solution faible (12.2), il vient:

$$\begin{aligned} 0 &= \int_B u \varphi_t + f(u) \varphi_x + g(u) \varphi_y + h(u) \varphi dx dy dt \\ &= \left(\int_{B^+} + \int_{B^-} \right) \varphi (h(u) - u_t - f(u)_x - g(u)_y) dx dy dt + \int_{B \cap \Sigma} \varphi (n_t [u] + n_x [f(u)] + n_y [g(u)]) ds(x, t) \end{aligned}$$

Les deux premiers termes s'annulent puisque u est supposée classique dans B^\pm . Si on pose $F(u) = \begin{pmatrix} f(u) \\ g(u) \end{pmatrix}$, le troisième terme exprime que:

$$\sigma(t)[u] = \langle \vec{\xi}, [F(u)] \rangle_{\mathbb{R}^2} \quad \text{avec } \sigma(t) = -\frac{n_t}{\sqrt{n_x^2 + n_y^2}} \text{ et } \vec{\xi} = \frac{1}{\sqrt{n_x^2 + n_y^2}} \begin{pmatrix} n_x \\ n_y \end{pmatrix} \in \mathbb{S}^1 \quad (12.3)$$

Notion de solution entropique Par analogie avec le cas monodimensionnel, on décide de construire rigoureusement une solution de (12.2) par une méthode de perturbation visqueuse évanescence. Pour cela, on introduit le problème (où Δ désigne le Laplacien dans \mathbb{R}^2):

$$u_t^\epsilon + f(u^\epsilon)_x + g(u^\epsilon)_y = h(u^\epsilon) + \epsilon \Delta u^\epsilon$$

Ce problème admet une solution régulière, il est donc pertinent de multiplier à gauche par $\eta'(u^\epsilon)$ où $\eta \in C^2(\mathbb{R})$ est une fonction strictement convexe. Par la règle de composition, on introduit q et r définies à des constantes additives près par $q' = \eta' \cdot f'$ et $r' = \eta' \cdot g'$ et on obtient:

$$\eta(u^\epsilon)_t + q(u^\epsilon)_x + r(u^\epsilon)_y = \eta'(u^\epsilon) \cdot h(u^\epsilon) + \epsilon \Delta \eta(u^\epsilon) - \epsilon \eta''(u^\epsilon) |\nabla(u^\epsilon)|^2$$

C'est-à-dire que sous réserve d'établir la convergence forte de la suite $(u^\epsilon)_{\epsilon \rightarrow 0}$, la limite ainsi obtenue vérifie:

$$\eta(u)_t + q(u)_x + r(u)_y \leq \eta'(u).h(u) \text{ au sens de } \mathcal{D}'(\mathbb{R}^2 \times \mathbb{R}^+) \quad (12.4)$$

Le théorème de Kruzkov L'inégalité (12.4) est essentielle à la théorie BV de l'équation (12.1) [53]. Elle permet en particulier d'assurer l'unicité à la limite en établissant une relation de dépendance continue de u par rapport à sa donnée initiale u_0 .

Théorème IV.4 (*Kruzkov*)

Soit $u_0 \in (L^\infty \cap BV)(\mathbb{R}^2)$ tel que $h(u)(\cdot, t)$ soit uniformément borné, alors le problème scalaire

$$\begin{cases} u_t + f(u)_x + g(u)_y = h(u) \\ \eta(u)_t + q(u)_x + r(u)_y \leq \eta'(u).h(u) \\ u(x, y, 0) = u_0(x, y) \end{cases} \quad (12.5)$$

admet une unique solution $u \in L^\infty([0, T]; BV(\mathbb{R}^2)) \cap C^0([0, T]; L^1_{loc}(\mathbb{R}^2))$.

La restriction sur $h(u)$ est nécessaire pour conserver un principe du maximum sur la suite d'approximations visqueuses $(u^\epsilon)_\epsilon$. Ceci est indispensable pour maintenir une vitesse de propagation finie tout en évitant une restriction du type f et g globalement Lipschitziennes.

12.1.2 Extension aux systèmes quasi-linéaires

Hyperbolicité dans le cas linéaire Soit $U(x, t)$ un vecteur de $\mathbb{R}^{N \geq 2}$, on considère le système:

$$\begin{cases} U_t + A.U_x + B.U_y = 0 \\ U_0 \in (L^2(\mathbb{R}^2))^N \end{cases} \quad (12.6)$$

Pour $N \geq 2$, la méthode des caractéristiques ne permet plus de construire une solution de (12.6) dans le cas général comme c'est le cas en 1D [18]. Néanmoins, ce problème peut se traiter par une transformation de Plancherel qui est une isométrie de L^2 [75]. On prévient l'explosion de la norme de $\hat{u}(t, \zeta)$ ($\zeta \in \mathbb{R}^2$ désignant les variables duales) sous la condition suivante:

Définition IV.1 Le problème (12.6) est bien posé dans L^2 s'il est hyperbolique, i.e.

$$\exists C > 0 \text{ telle que } \sup_{\xi \in \mathbb{S}^1} \|e^{-i(A\xi_1 + B\xi_2)}\| \leq C \quad (12.7)$$

Réciproquement, si (12.7) est vérifiée, il existe une unique solution de (12.6) dans $C^0(\mathbb{R}, L^2(\mathbb{R}^2))$.

Le problème linéaire homogène (12.7) est donc réversible.

Définition IV.2 Le problème (12.6) est strictement hyperbolique si les matrices $A\xi_1 + B\xi_2$, $\vec{\xi} \in \mathbb{S}^1$ sont diagonalisables dans \mathbb{R} et si les N valeurs propres sont différentes et dépendent continûment de $\vec{\xi}$.

La terminologie suggère qu'un problème strictement hyperbolique soit hyperbolique; c'est bien sûr le cas. On pourra se reporter à [76] pour plus de détails.

Retour au cas général On considère maintenant le problème non-linéaire:

$$\begin{cases} U_t + F(U)_x + G(U)_y = H(U) \\ U_0 \text{ dans un espace fonctionnel convenable} \end{cases} \quad (12.8)$$

Définition IV.3 On dit que le problème (12.8) est hyperbolique si pour tout \bar{U} dans l'espace des états admissibles (un convexe de \mathbb{R}^N), le problème linéarisé:

$$U_t + dF(\bar{U})U_x + dG(\bar{U})U_y = 0$$

est hyperbolique. Une condition suffisante est d'en garantir la stricte hyperbolicité.

Le choix de l'espace fonctionnel pour (12.8) n'est pas simple: en 1D, on sait que les solutions d'un tel système ne sont pas forcément à variation bornée. En effet, les champs linéairement dégénérés véhiculent les oscillations sans les amortir et il faut avoir recours à la théorie de la compacité par compensation puisque la seule estimation uniforme n'est que L^∞ . Ceci n'est valable que pour les systèmes 2×2 ou bien la classe de Temple. En 2D, le problème reste désespérément ouvert, puisque ni ce procédé, ni le schéma de Glimm ne permettent d'en construire des solutions (un simple problème de Riemann devient déjà très complexe!). Il subsiste qu'à la suite des phénomènes d'explosion du gradient, il est nécessaire de chercher des solutions de (12.8) très peu régulières. Celles-ci doivent néanmoins être bornées en amplitude afin de maintenir la condition physiquement réaliste d'une *vitesse de propagation finie*. On les suppose donc à priori comme des distributions de $L^\infty([0, T] \times \mathbb{R}^2)$

Définition IV.4 Soit $U_0 \in L^\infty(\mathbb{R}^2)^N$, $U \in L^\infty([0, T] \times \mathbb{R}^2)^N$ est une solution faible de (12.8) si pour toute fonction-test φ dans $C_0^\infty([0, T] \times \mathbb{R}^2)^N$, on a:

$$\int_{\mathbb{R}^2 \times \mathbb{R}^+} U \varphi_t + F(U) \varphi_x + G(U) \varphi_y + H(U) \varphi dx \cdot dy \cdot dt + \int_{\mathbb{R}^2} U_0 \varphi dx \cdot dy = 0$$

Par analogie avec les problèmes 1D, on introduit la condition d'entropie de Lax:

Définition IV.5 Une solution faible $U \in L^\infty([0, T] \times \mathbb{R}^2)^N$ est une solution entropique de (12.8) au sens de Lax si pour toute entropie η strictement convexe et si pour toute φ dans $C_0^\infty([0, T] \times \mathbb{R}^2)$, on a:

$$\int_{\mathbb{R}^2 \times \mathbb{R}^+} \eta(U) \varphi_t + q(U) \varphi_x + r(U) \varphi_y + \nabla \eta(U) H(U) \varphi dx \cdot dy \cdot dt + \int_{\mathbb{R}^2} \eta(U_0) \varphi dx \cdot dy \geq 0$$

i.e. $\eta(U) + q(U)_x + r(U)_y \leq \nabla \eta(U) \cdot H(U)$ au sens de $\mathcal{D}'(\mathbb{R}^2 \times \mathbb{R}^+)$. Les fonctions q et r vérifient $\nabla q = \nabla \eta \cdot dF$ et $\nabla r = \nabla \eta \cdot dG$ et sont les flux d'entropie.

Pour terminer, on énonce l'unique résultat de portée un tant soit peu générale, à savoir un théorème de 1963 concernant l'existence et l'unicité locales de solutions régulières de (12.8).

Théorème IV.5 (Garding-Leray)

Soit $\mathbb{R} \ni s > 2$ et $U_0 \in H^s(\mathbb{R}^2)$, alors il existe un temps $T > 0$ tel que le problème (12.8) admette une unique solution forte $U \in C^1([0, T] \times \mathbb{R}^2)^N$ si ce dernier possède une entropie strictement convexe.

De plus, $U \in C^0([0, T]; H^s(\mathbb{R}^2)) \cap C^1([0, T]; H^{s-1}(\mathbb{R}^2))$.

12.2 Traitement non-conservatif des sources dans un contexte 2D

12.2.1 Introduction générale de la méthode

Celle-ci se fera sur un exemple scalaire très simple à savoir la loi linéaire:

$$\begin{cases} u_t + au_x + bu_y = \alpha u \\ u_0 \in C^1(\mathbb{R}^2) \text{ par exemple} \end{cases} \quad (12.9)$$

Sa solution exacte est immédiate: $u(x, y, t) = u_0(x - at, y - bt) \mathbf{e}^{\alpha t}$; néanmoins, obtenir une bonne approximation numérique l'est beaucoup moins. D'une façon générale, on convient de se placer dans un contexte volumes finis afin d'exploiter au mieux le savoir-faire acquis pour les équations monodimensionnelles. On renvoie à [18] pour des détails sur ce type de méthode.

La tactique la plus directe pour attaquer (12.9) est de chercher une fonction régulière $K(x, y) \in C^1(\mathbb{R}^2)$ dont la divergence vaut α afin d'obtenir:

$$\begin{cases} (\ln |u|)_t + a(\ln |u| - K)_x + b(\ln |u| - K)_y = 0 \\ K_t = 0 \end{cases} \quad (12.10)$$

Un choix immédiat (et simpliste!) est donné par:

$$K(x, y) = \frac{\alpha}{2}(x + y)$$

Malheureusement, celui-ci conduit à des problèmes 1D grossiers dans le cas où a (ou bien b) est très proche de zéro. Ceci révèle qu'en 2D, il est nécessaire de faire dépendre K du champ de vitesses $\vec{V} = \begin{pmatrix} a \\ b \end{pmatrix}$. On décide donc de réécrire l'équation (12.9) sous la forme:

$$u_t + \langle \vec{V}, \nabla u \rangle_{\mathbb{R}^2} = \langle \vec{V}, \nabla K \rangle_{\mathbb{R}^2} u$$

Cette formulation au moyen du produit scalaire de \mathbb{R}^2 suggère l'usage de la norme euclidienne:

$$K(x, y) = \frac{\alpha}{\|\vec{V}\|^2} \left\langle \vec{V}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle_{\mathbb{R}^2} \quad (12.11)$$

Les problèmes 1D élémentaires intervenant dans un schéma volumes finis pour (12.9) sont donc de la forme:

$$\begin{cases} (\ln |u|)_t + a(\ln |u| - K)_x = 0 \\ K_t = 0 \end{cases} \quad \text{et} \quad \begin{cases} (\ln |u|)_t + b(\ln |u| - K)_y = 0 \\ K_t = 0 \end{cases} \quad (12.12)$$

12.2.2 Un exemple issu de la modélisation diphasique

On décide de s'intéresser au système suivant:

$$\begin{cases} \rho_t + (\rho u)_x + (\rho v)_y = 0 \\ (\rho c)_t + (\rho u c)_x + (\rho v c)_y = k_1(x, y) \rho g(c) \\ (\rho u)_t + (\rho u^2 + p)_x + (\rho u v)_y = -k_2(x, y) \rho u |u| \\ (\rho v)_t + (\rho v u)_x + (\rho v^2 + p)_y = 0 \end{cases} \quad (12.13)$$

Comme à l'accoutumée, les symboles $\rho, c, (u, v), p$ désignent la densité globale du mélange, la concentration massique d'une des phases, les composantes axiales et transverses du vecteur-vitesse volumétrique et la loi de pression commune aux deux phases. Dans le cadre d'un schéma à décomposition de flux, on convient d'utiliser l'expression simplifiée:

$$p(\rho, c) = \rho^2 \sqrt{c}$$

Quant aux termes sources, ils sont sensés modéliser un rappel vers un état d'équilibre privilégié ainsi qu'un effet de frottement pariétal et une perte de charge dans une grille à aubes. On pose donc:

$$\begin{aligned} U &= (\rho, \rho c, \rho u, \rho v) \\ F(U) &= (\rho u, \rho u c, \rho u^2 + p, \rho v u) \\ G(U) &= (\rho v, \rho v c, \rho u v, \rho v^2 + p) \\ H(x, y, U) &= (0, k_1(x, y) \rho g(c), -k_2(x, y) \rho u |u|, 0) \end{aligned}$$

Et on obtient l'écriture condensée:

$$U_t + F(U)_x + G(U)_y = H(x, y, U)$$

Pour construire un schéma \llcorner équilibre \lrcorner sur la base d'une décomposition de flux, il faut tout d'abord étendre les rappels du chapitre 2 de la partie III.

Proposition IV.4 *Le problème (12.13) peut se réécrire sous la forme suivante:*

$$U_t + F^+(U)_x + F^-(U)_x + G^+(U)_y + G^-(U)_y = H(x, y, U)$$

où les flux F^\pm, G^\pm sont différentiables et vérifient les propriétés:

$$\forall U \in (\mathbb{R}_*^+ \times [0, 1] \times \mathbb{R}^2), \Lambda(dF^\pm(U)) \subset \mathbb{R}^\pm \quad \text{et} \quad \Lambda(dG^\pm(U)) \subset \mathbb{R}^\pm$$

Ceux-ci sont définis par:

$$F^+(U) = \begin{cases} F(U) & \text{si } u \geq a \\ \begin{pmatrix} M^+ \\ M^+ c \\ M^+(\frac{u}{2} + a) \\ M^+ v \end{pmatrix} & \text{si } |u| < a \\ \vec{0} & \text{si } u \leq -a \end{cases} \quad \text{et} \quad F^-(U) = \begin{cases} \vec{0} & \text{si } u \geq a \\ \begin{pmatrix} M^- \\ M^- c \\ M^-(\frac{u}{2} - a) \\ M^- v \end{pmatrix} & \text{si } |u| < a \\ F(U) & \text{si } u \leq -a \end{cases} \quad (12.14)$$

$$G^+(U) = \begin{cases} G(U) \text{ si } u \geq a \\ \begin{pmatrix} N^+ \\ N^+c \\ N^+u \\ N^+(\frac{v}{2} + a) \\ \vec{0} \end{pmatrix} \text{ si } |u| < a \\ \vec{0} \text{ si } u \leq -a \end{cases} \quad \text{et } G^-(U) = \begin{cases} \vec{0} \text{ si } u \geq a \\ \begin{pmatrix} N^- \\ N^-c \\ N^-u \\ N^-(\frac{v}{2} - a) \\ G(U) \end{pmatrix} \text{ si } |u| < a \\ G(U) \text{ si } u \leq -a \end{cases}$$

$a = \sqrt{\frac{2p(\rho,c)}{\rho}}$ désigne la vitesse du son dans le mélange et les quantités M^\pm, N^\pm sont des décompositions du terme de quantité de mouvement:

$$\begin{aligned} M^+ &= \frac{\rho}{4a}(u+a)^2 & N^+ &= \frac{\rho}{4a}(v+a)^2 \\ M^- &= \frac{-\rho}{4a}(u-a)^2 & N^- &= \frac{-\rho}{4a}(v-a)^2 \end{aligned}$$

Démonstration

C'est une extension directe de la proposition III.2; voir le chapitre 2 - partie III. \square

Dans le contexte des volumes finis où l'on fixe $k_1 \equiv k_2 \equiv 0$, on est conduit à résoudre dans chaque maille $\Delta x \times \Delta y$ quatre problèmes de Riemann du type:

$$\begin{cases} \text{Selon } O\vec{x} \begin{cases} U_t^+ + F^+(U^+)_x = 0 \text{ pour } x \in [(i - \frac{1}{2})\Delta x, i\Delta x[\\ U_t^- + F^-(U^-)_x = 0 \text{ pour } x \in [i\Delta x, (i + \frac{1}{2})\Delta x[\\ \text{Selon } O\vec{y} \begin{cases} V_t^+ + G^+(V^+)_x = 0 \text{ pour } y \in [(j - \frac{1}{2})\Delta y, j\Delta y[\\ V_t^- + G^-(V^-)_x = 0 \text{ pour } y \in [j\Delta y, (j + \frac{1}{2})\Delta y[\end{cases} \end{cases} \end{cases}$$

Dans le cas général, il faut introduire l'action des termes sources par le biais de relations de saut aux interfaces du maillage.

12.2.3 Relations de saut non-conservatives

On commence en premier lieu par définir les fonctions de raideur:

$$K_i^{\Delta x}(x, y) = \int^x k_i(s, y) ds \text{ et } K_i^{\Delta y}(x, y) = \int^y k_i(x, s) ds \text{ pour } i = 1, 2 \quad (12.15)$$

Par analogie avec les lois scalaires, on décompose les termes sources de la façon suivante:

$$H(x, y, U) = \left(\frac{M^2}{M^2 + N^2} + \frac{N^2}{M^2 + N^2} \right) \cdot H(x, y, U)$$

avec les notations $M = \rho u$ et $N = \rho v$. On va donc résoudre les problèmes 1D suivants:

$$\begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho c)_t + (\rho uc)_x = \rho \cdot g(c) \frac{M^2}{M^2 + N^2} (K_1^{\Delta x})_x \\ (\rho u)_t + (\rho u^2 + p)_x = -\rho u |u| \frac{M^2}{M^2 + N^2} (K_2^{\Delta x})_x \\ (\rho v)_t + (\rho v u)_x = 0 \end{cases} \quad (12.16)$$

$$\begin{cases} \rho_t + (\rho v)_y = 0 \\ (\rho c)_t + (\rho vc)_y = \rho \cdot g(c) \frac{N^2}{M^2 + N^2} (K_1^{\Delta y})_y \\ (\rho u)_t + (\rho v u)_y = -\rho u |u| \frac{N^2}{M^2 + N^2} (K_2^{\Delta y})_y \\ (\rho v)_t + (\rho v^2 + p)_y = 0 \end{cases} \quad (12.17)$$

On régularise les termes non-conservatifs par le biais des courbes intégrales des systèmes stationnaires dans chacune des directions $O\vec{x}$, $O\vec{y}$. Ceci est vraiment une généralisation 2D du travail exposé au chapitre 1. D'une façon similaire, les systèmes différentiels couplés ne peuvent être traités en forme fermée; pour parer à cette déconvenue, on présente un résultat un peu plus faible.

Lemme IV.6 *Supposons que $(K_1^{\Delta y})_x \equiv 0$ et $(K_1^{\Delta y})_y \equiv 0$ dans les systèmes (12.16), (12.17), alors les relations de saut suivantes sont vérifiées:*

- pour (12.16): $[\rho u] = [c] = [v] = 0$ et

$$-M^2[\ln(\rho)] + [p_1] - \frac{v^2}{2}[\rho^2] + \frac{v^2}{M^2}[p_2] = -M|M|[(K_2^{\Delta x})] \quad (12.18)$$

- pour (12.17): $[\rho v] = [c] = [v] = 0$ et

$$v \left[\frac{1}{u} \right] - \frac{1}{v}[u] = \text{sgn}(u)[K_2^{\Delta y}] \quad (12.19)$$

où p_1, p_2 désignent respectivement des primitives à c constant de $\rho.p_\rho$ et $\rho^3.p_\rho$.

Comme précédemment, le cas $k_1(x, y) \neq 0$ sera traité par un algorithme itératif du type «point fixe».

Démonstration

Cas du système en $O\vec{x}$

$$\left(\frac{M^2}{\rho} + p(\rho, c_0) \right)_x = -\frac{M^2}{\rho^2}\rho_x + p_\rho\rho_x = -M|M|\frac{M^2}{\rho(M^2 + \rho^2 v_0^2)}(K_2^{\Delta x})_x$$

D'où il vient, si $M \neq 0$:

$$\left(-\frac{M^2}{\rho^2}\rho_x + p_\rho\rho_x \right) \left(\rho + \rho^3 \frac{v_0^2}{M^2} \right) = -M|M|(K_2^{\Delta x})_x$$

Et donc:

$$-\frac{M^2}{\rho}\rho_x + (\rho p_\rho)\rho_x - v_0^2 \left(\frac{\rho^2}{2} \right)_x + \frac{v_0^2}{M^2}(\rho^3 p_\rho)\rho_x = -M|M|(K_2^{\Delta x})_x$$

On pose: $p_1(\rho, c_0)$ telle que $(p_1)_\rho = \rho p_\rho$ et $p_2(\rho, c_0)$ telle que $(p_2)_\rho = \rho^3 p_\rho$. En intégrant par rapport à x , on obtient (12.18)

Cas du système en $O\vec{y}$ On obtient clairement $[\rho v] = [c] = 0$. Par ailleurs, la conservation de la charge $Q = \rho v^2 + p$ impose $[v] = [\rho] = 0$. On obtient donc:

$$v.u_y = -u|u|\frac{v^2}{u^2 + v^2}(K_2^{\Delta y})_y$$

En multipliant par $(1 + u^2/v^2)$ et par $1/u|u|$, il vient:

$$v \cdot \frac{-u_y}{u|u|} \left(1 + \frac{u^2}{v^2} \right) = \text{sgn}(u) \left[v \left(\frac{1}{u} \right)_y - \left(\frac{1}{v} \right) (u)_y \right] = (K_2^{\Delta y})_y$$

Comme $v \equiv v_0$, on peut intégrer en y et ainsi obtenir l'expression (12.19). \square

Il reste à prendre en compte les variations de concentration dues au terme réactionnel en $\rho.g(c)$ dans l'éventualité $k_1(x, y) \neq 0$. Comme cela va condamner toute tentative de résolution analytique, on met en place un algorithme itératif afin d'obtenir les valeurs à l'interface. Pour cela, on utilise encore la fonction $\phi(c)$ définie en partie III:

$$\begin{aligned} \phi : [0, 1] &\rightarrow \mathbb{R} \\ c &\mapsto \phi(c) \text{ avec } \phi'(c) = \frac{1}{g(c)} \end{aligned}$$

(voir par exemple le lemme III.6).

A chaque interface, on fera une approximation du type:

- calculer ρ^k (ou bien v^k) par le lemme IV.6 en supposant $c \equiv c^{k-1}(\Delta x)$ (ou bien $c \equiv c^{k-1}(\Delta y)$).
- calculer c^k en gelant ρ^k en itérant vers un point fixe jusqu'à satisfaire le critère d'arrêt choisi. On est donc en position d'écrire un schéma non-conservatif «équilibre» pour le système (12.13).

12.2.4 Ecriture du schéma proposé

Dans cette section, on commence par préciser les notations employées pour désigner la solution approchée du système (12.13). Les quantités Δx , Δy supposées constantes caractérisent le maillage spatial; le pas de temps est noté Δt . On se limite à un schéma explicite, par conséquent, ce dernier sera soumis à la condition de stabilité CFL. La donnée initiale est supposée uniformément bornée, donc localement sommable; on la discrétise donc sous la forme:

$$U_{i,j}^0 = \frac{1}{\Delta x \Delta y} \int_{(i-\frac{1}{2})\Delta x}^{(i+\frac{1}{2})\Delta x} \int_{(j-\frac{1}{2})\Delta y}^{(j+\frac{1}{2})\Delta y} U^0(x, y) dx dy$$

On introduit donc un schéma numérique $\bar{U}_{i,j}$ pour calculer à chaque instant $t^n = n\Delta t$ les valeurs $U_{i,j}^n$ approchant les moyennes locales de la solution $U(x, y, t^n)$ du problème continu. On espère donc que l'on a:

$$U_{i,j}^n \simeq \frac{1}{\Delta x \Delta y} \int_{(i-\frac{1}{2})\Delta x}^{(i+\frac{1}{2})\Delta x} \int_{(j-\frac{1}{2})\Delta y}^{(j+\frac{1}{2})\Delta y} U(x, y, n\Delta t) dx dy$$

Par analogie avec le schéma monodimensionnel, on propose d'écrire:

$$\begin{aligned} U_{i,j}^{n+1} = U_{i,j}^n & - \frac{\Delta t}{\Delta x} [F^+(U_{i,j}^n) - F^+(U_{i-\frac{1}{2},j}^+) + F^-(U_{i+\frac{1}{2},j}^-) - F^-(U_{i,j}^n)] \\ & - \frac{\Delta t}{\Delta y} [G^+(U_{i,j}^n) - G^+(U_{i,j-\frac{1}{2}}^+) + G^-(U_{i,j+\frac{1}{2}}^-) - G^-(U_{i,j}^n)] \end{aligned} \quad (12.20)$$

Les états à l'interface $U_{i\pm\frac{1}{2},j}^\pm$ et $U_{i,j\pm\frac{1}{2}}^\pm$ sont déterminés à l'aide des relations de saut (12.18) et (12.19) lorsque $k_1(x, y) \equiv 0$. Dans le cas où $k_1(x, y) \neq 0$, on met en place l'algorithme du point fixe décrit au chapitre 1 et étudié en annexe A. Concrètement, soit $(U_{i,j}^n)$ une approximation des moyennes locales autour des points $(i\Delta x, j\Delta y)$ au temps $n\Delta t$, on détermine une solution approchée en $(n+1)\Delta t$ grâce au schéma (12.20) et aux états à l'interface calculés de la façon suivante:

- $U_{i-\frac{1}{2},j}^+$ est tel que:

$$\left\{ \begin{array}{l} (\rho u)_{i-\frac{1}{2},j}^+ = (\rho u)_{i-1,j}^n \text{ et } v_{i-\frac{1}{2},j}^+ = v_{i-1,j}^n \\ (\rho u)_{i-1,j}^n [\phi(c_{i-\frac{1}{2},j}^+) - \phi(c_{i-1,j}^n)] = \rho_{i-\frac{1}{2},j}^+ [K_1(i\Delta x, j\Delta y) - K_1((i-1)\Delta x, j\Delta y)] \\ \ln \left(\frac{\rho_{i-\frac{1}{2},j}^+}{\rho_{i-1,j}^n} \right) - [p_1(\rho_{i-\frac{1}{2},j}^+, c_{i-\frac{1}{2},j}^+) - p_1(\rho_{i-1,j}^n, c_{i-1,j}^n)] / [(\rho u)_{i-1,j}^n]^2 \\ \quad + \frac{(v_{i-1,j}^n)^2}{2[(\rho u)_{i-1,j}^n]^2} [(\rho_{i-\frac{1}{2},j}^+)^2 - (\rho_{i-1,j}^n)^2] \\ \quad - \frac{(v_{i-1,j}^n)^4}{[(\rho u)_{i-1,j}^n]^4} [p_2(\rho_{i-\frac{1}{2},j}^+, c_{i-\frac{1}{2},j}^+) - p_2(\rho_{i-1,j}^n, c_{i-1,j}^n)] \\ = \text{sgn}(\rho u)_{i-1,j}^n [K_2(i\Delta x, j\Delta y) - K_2((i-1)\Delta x, j\Delta y)] \end{array} \right.$$

- $U_{i+\frac{1}{2},j}^-$ est tel que:

$$\left\{ \begin{array}{l} (\rho u)_{i+\frac{1}{2},j}^- = (\rho u)_{i+1,j}^n \text{ et } v_{i+\frac{1}{2},j}^- = v_{i+1,j}^n \\ (\rho u)_{i+1,j}^n [\phi(c_{i+\frac{1}{2},j}^-) - \phi(c_{i+1,j}^n)] = \rho_{i+\frac{1}{2},j}^- [K_1(i\Delta x, j\Delta y) - K_1((i+1)\Delta x, j\Delta y)] \\ \ln \left(\frac{\rho_{i+\frac{1}{2},j}^-}{\rho_{i+1,j}^n} \right) - [p_1(\rho_{i+\frac{1}{2},j}^-, c_{i+\frac{1}{2},j}^-) - p_1(\rho_{i+1,j}^n, c_{i+1,j}^n)] / [(\rho u)_{i+1,j}^n]^2 \\ \quad + \frac{(v_{i+1,j}^n)^2}{2[(\rho u)_{i+1,j}^n]^2} [(\rho_{i+\frac{1}{2},j}^-)^2 - (\rho_{i+1,j}^n)^2] \\ \quad - \frac{(v_{i+1,j}^n)^4}{[(\rho u)_{i+1,j}^n]^4} [p_2(\rho_{i+\frac{1}{2},j}^-, c_{i+\frac{1}{2},j}^-) - p_2(\rho_{i+1,j}^n, c_{i+1,j}^n)] \\ = \text{sgn}(\rho u)_{i+1,j}^n [K_2(i\Delta x, j\Delta y) - K_2((i+1)\Delta x, j\Delta y)] \end{array} \right.$$

- $U_{i,j-\frac{1}{2}}^+$ est tel que:

$$\left\{ \begin{array}{l} \rho_{i,j-\frac{1}{2}}^+ = \rho_{i,j-1}^n \text{ et } v_{i,j-\frac{1}{2}}^+ = v_{i,j-1}^n \\ v_{i,j-1}^n [\phi(c_{i,j-\frac{1}{2}}^+) - \phi(c_{i,j-1}^n)] = [K_1(i\Delta x, j\Delta y) - K_1(i\Delta x, (j-1)\Delta y)] \\ v_{i,j-1}^n \left(\frac{1}{u_{i,j-\frac{1}{2}}^+} - \frac{1}{u_{i,j-1}^n} \right) - \frac{1}{v_{i,j-1}^n} (u_{i,j-\frac{1}{2}}^+ - u_{i,j-1}^n) = \text{sgn}(u_{i,j-1}^n) [K_2(i\Delta x, j\Delta y) - K_2(i\Delta x, (j-1)\Delta y)] \end{array} \right.$$

- $U_{i,j+\frac{1}{2}}^-$ est tel que:

$$\begin{cases} \rho_{i,j+\frac{1}{2}}^- = \rho_{i,j+1}^n \text{ et } v_{i,j+\frac{1}{2}}^- = v_{i,j+1}^n \\ v_{i,j+1}^n [\phi(c_{i,j+\frac{1}{2}}^-) - \phi(c_{i,j+1}^n)] = [K_1(i\Delta x, j\Delta y) - K_1(i\Delta x, (j+1)\Delta y)] \\ v_{i,j+1}^n \left(\frac{1}{u_{i,j+\frac{1}{2}}^-} - \frac{1}{u_{i,j+1}^n} \right) - \frac{1}{v_{i,j+1}^n} (u_{i,j+\frac{1}{2}}^- - u_{i,j+1}^n) = \text{sgn}(u_{i,j+1}^n) [K_2(i\Delta x, j\Delta y) - K_2(i\Delta x, (j+1)\Delta y)] \end{cases}$$

12.3 Validation par des cas-tests numériques

Les tests numériques pour (12.13) sont des raffinements de la version monodimensionnelle présentée au chapitre 1 (voir les figures 10.16, 10.17, 10.18, 10.19, 10.20, 10.21). Il s'agit de tester la qualité des approximations obtenues par le schéma (12.20) par rapport à celles fournies par un algorithme plus classique du type `¶¶pas fractionnaires` mettant en oeuvre une discrétisation centrée des termes sources.

12.3.1 Cas-test pseudo-1D

Ce premier test est une validation simple de l'algorithme 2D proposé au sens où on va traiter un problème sans vitesse transverse. Cela permettra en outre de comparer les profils d'ondes obtenus avec ceux issus des systèmes monodimensionnels du chapitre 1. Ceci assurera aussi que la décomposition du terme source préserve si besoin est la symétrie axiale de la solution et ne génère pas de vitesses transverses parasites. On considère donc un canal maillé de façon régulière avec $x \in [0; 1]$ et $y \in [0; 0.3]$. On fixe des valeurs constantes pour $k_1(x, y)$, $k_2(x, y)$ et μ à savoir:

$$k_1(x, y) \equiv 10; k_2(x, y) \equiv 5; \mu \equiv 0.5$$

Le cas-test est du type `¶¶couche-limite` subsonique avec les données initiales et la condition imposée à gauche:

$$\begin{cases} \rho_0 = 2 \\ c_0 = 0.5 \\ u_0 = 0.5 \\ v_0 = 0 \end{cases} \quad \text{et} \quad \begin{cases} c_l = 0.9 \\ (\rho u)_l = 1 \\ v_l = 0 \end{cases}$$

On choisit les paramètres $\Delta x = \Delta y = 0.02$ (750 mailles) ainsi que $\Delta t = 0.005$. On itère jusque $T = 0.2$: les figures 12.1, 12.2, 12.3 représentent les courbes issues du présent schéma non-conservatif avec celles issues d'une approche classique `¶¶pas fractionnaires`. On constate bien une bonne adéquation sur ce problème simple.

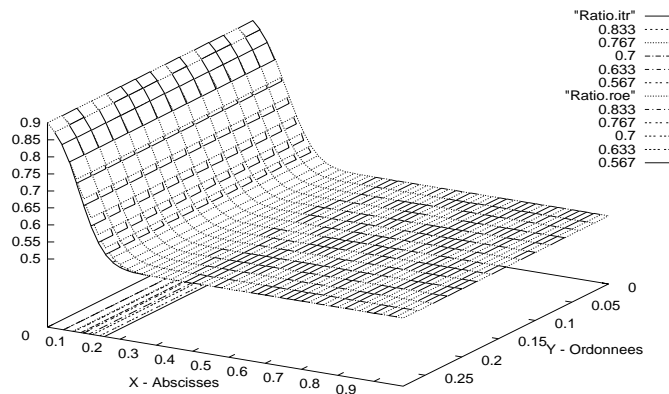


Figure 12.1: Concentration (schéma non-conservatif/pas fractionnaires)

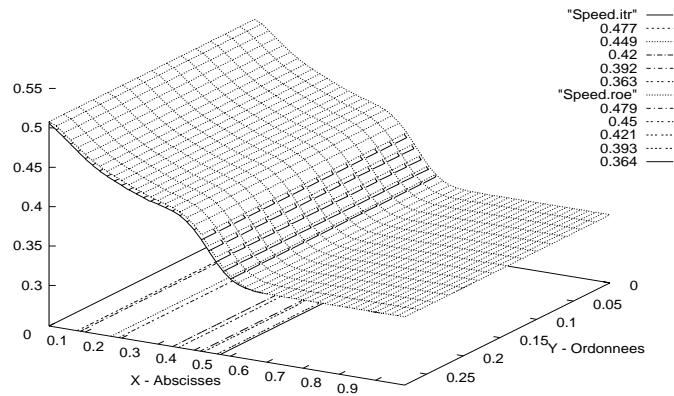


Figure 12.2: Vitesse (schéma non-conservatif/pas fractionnaires)

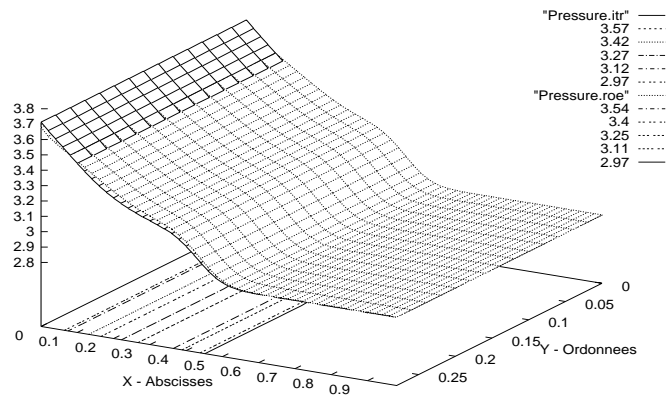


Figure 12.3: Pression (schéma non-conservatif/pas fractionnaires)

12.3.2 Frottement uniforme dans un canal

On considère maintenant un canal semblable dont la géométrie est décrite à la figure 12.4. Les parois sont matérialisées par des conditions aux limites de type μmiroir_i .

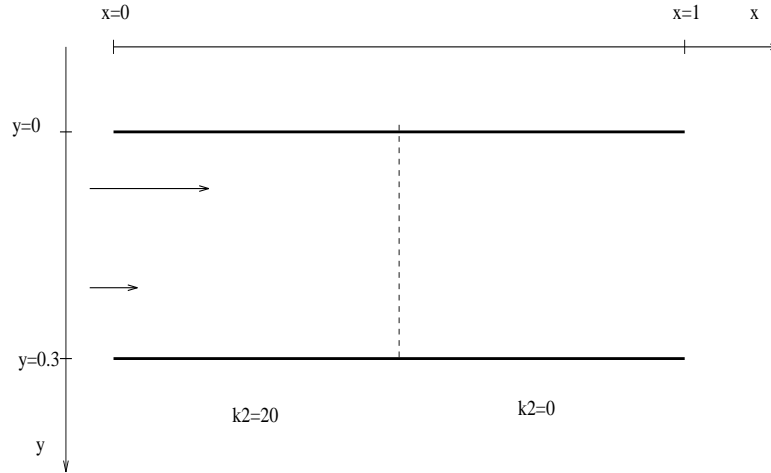


Figure 12.4: Géométrie du canal

Afin de modéliser un effet de frottement dû à des obstacles microscopiques, on introduit la fonction k_2 que l'on choisit de la forme:

$$k_2(x, y) = \bar{k}_2 \mathbf{1}_{x \leq 0.5} \text{ avec } \bar{k}_2 = 20$$

Il n'y a pas de reconcondensation (i.e. $k_1 \equiv 0$). La donnée initiale et la condition-limite à gauche simulent un écoulement cisailé¹. On a choisi:

$$\left\{ \begin{array}{l} \rho_0 = 2 \\ c_0 = 0.8 \mathbf{1}_{y < 0.15} + 0.2 \mathbf{1}_{y > 0.15} \\ u_0 = \mathbf{1}_{y < 0.15} + 0.2 \mathbf{1}_{y > 0.15} \\ v_0 = 0 \end{array} \right. \text{ et } \left\{ \begin{array}{l} c_l = 0.8 \mathbf{1}_{y < 0.15} + 0.2 \mathbf{1}_{y > 0.15} \\ u_l = \mathbf{1}_{y < 0.15} + 0.2 \mathbf{1}_{y > 0.15} \\ v_l = 0 \end{array} \right.$$

A maillage égal, on constate au bout d'un certain nombre d'itérations une dégradation spontanée de la suite d'approximations issues du schéma classique à pas fractionnaires sous la forme d'oscillations parasites sur la vitesse transverse. Tout d'abord, celles-ci restent localisées aux abords de $x = 0.5$, puis elles contaminent toute la solution. Quant au débit axial, il est discontinu, même à proximité de la paroi où le problème est quasiment monodimensionnel. Ceci corrobore les observations du chapitre 1. Les figures 12.5, 12.6, 12.7 12.8 présentent les valeurs stationnaires obtenues pour $t \simeq 10$ de la vitesse axiale, transverse, de la pression et de la concentration pour le schéma non-conservatif (12.20). On présente aussi à la figure 12.9 les historiques des résidus pour le présent schéma l'approche plus classique à pas fractionnaires où les termes sources sont traités de façon centrée. Les résidus sont calculés de la façon suivante: on désigne par $\mathbf{U}^n = (U_{i,j}^n)_{i,j}$ le vecteur des approximations à l'instant $t^n = n\Delta t$ où $U_{i,j}^n = (\rho_{i,j}^n, (\rho c)_{i,j}^n, (\rho u)_{i,j}^n, (\rho v)_{i,j}^n)$. Le résidu est alors la norme infinie du vecteur $\mathbf{U}^{n+1} - \mathbf{U}^n$. L'efficacité de l'approche par interfaces est alors claire au vu de la figure 12.9 puisque le permanent obtenu atteint la précision machine, ce qui n'est vraiment pas le cas de l'approche centrée classique. Les paramètres utilisés avec chacun des schémas sont: $\Delta x = 0.02$, $\Delta y = 0.02$, $\Delta t = 0.003$.

1. Ce cas-test a été suggéré par S. Clerc

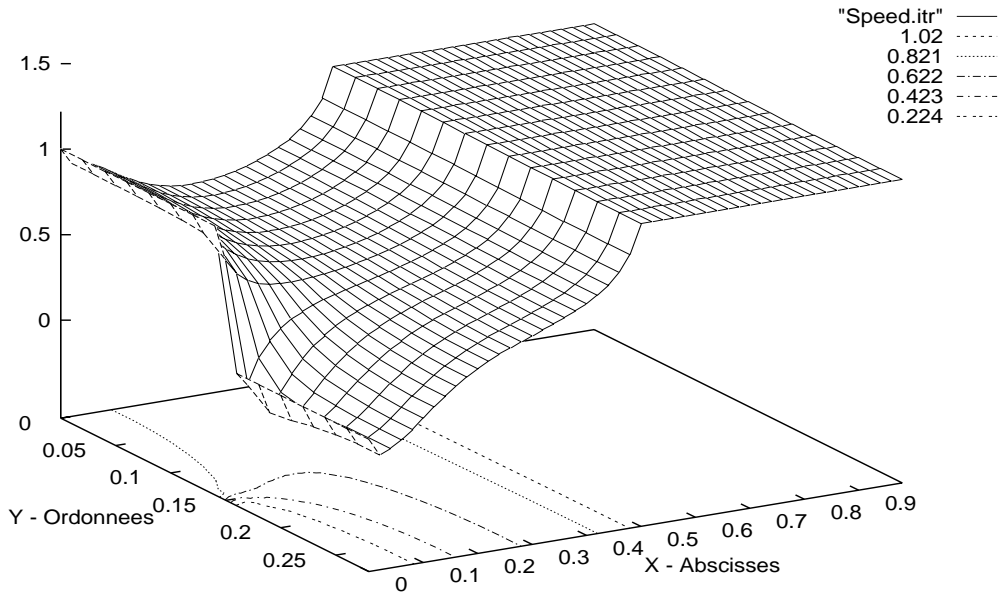


Figure 12.5: Vitesse axiale stationnaire (schéma non-conservatif)

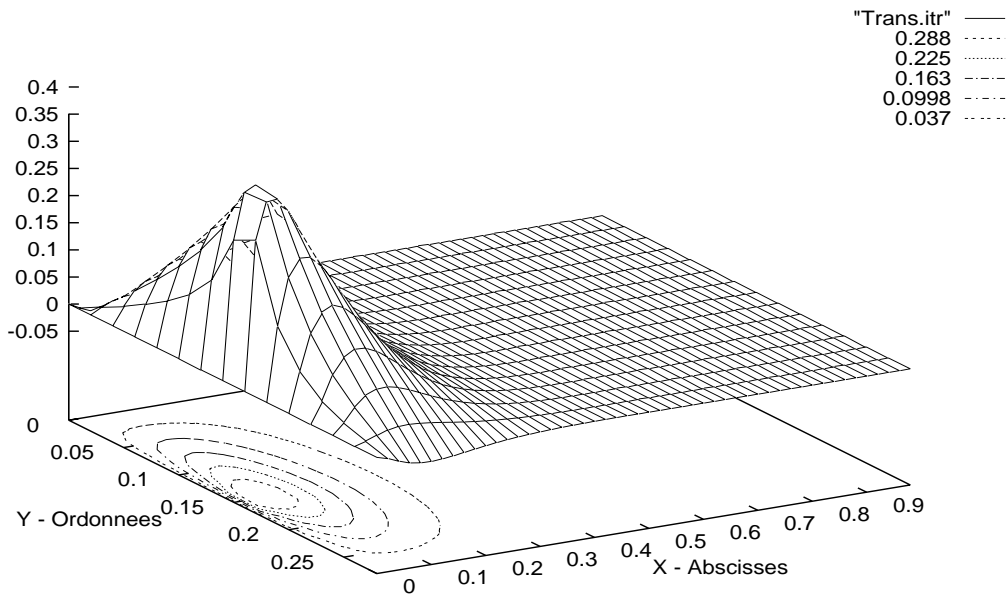


Figure 12.6: Vitesse transverse stationnaire (schéma non-conservatif)

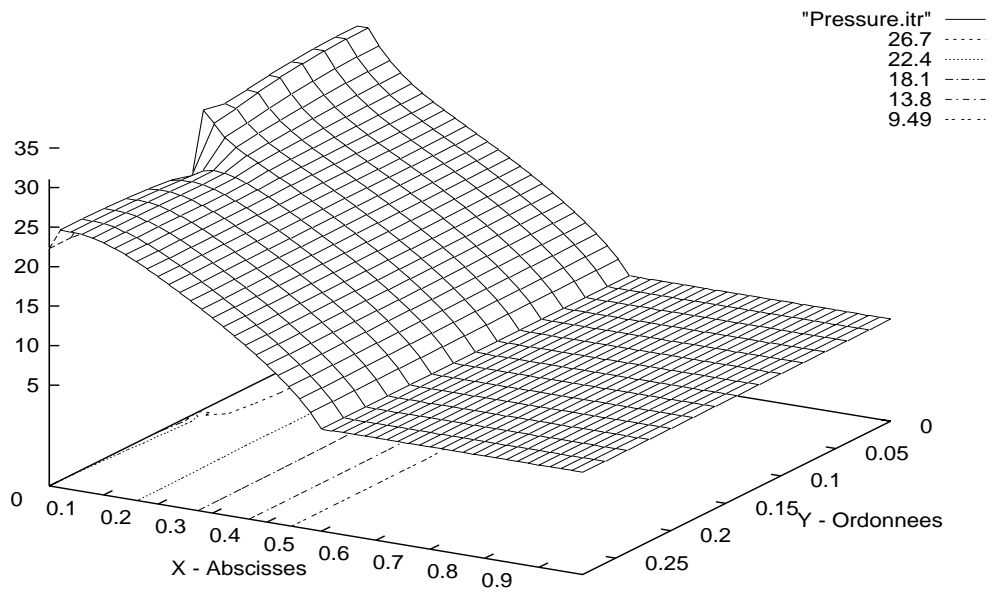


Figure 12.7: Pression stationnaire (schéma non-conservatif)

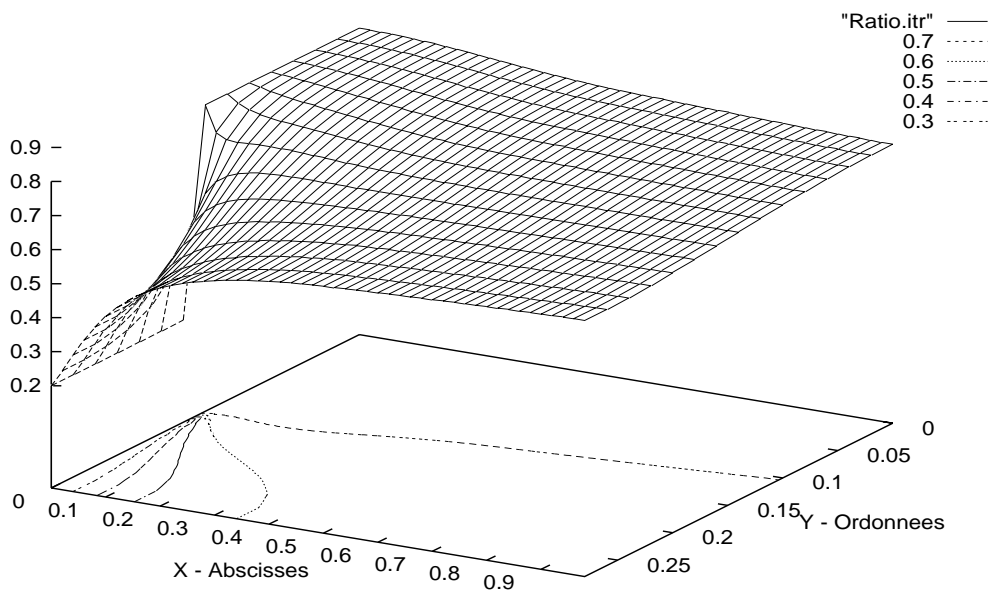


Figure 12.8: Concentration stationnaire (schéma non-conservatif)

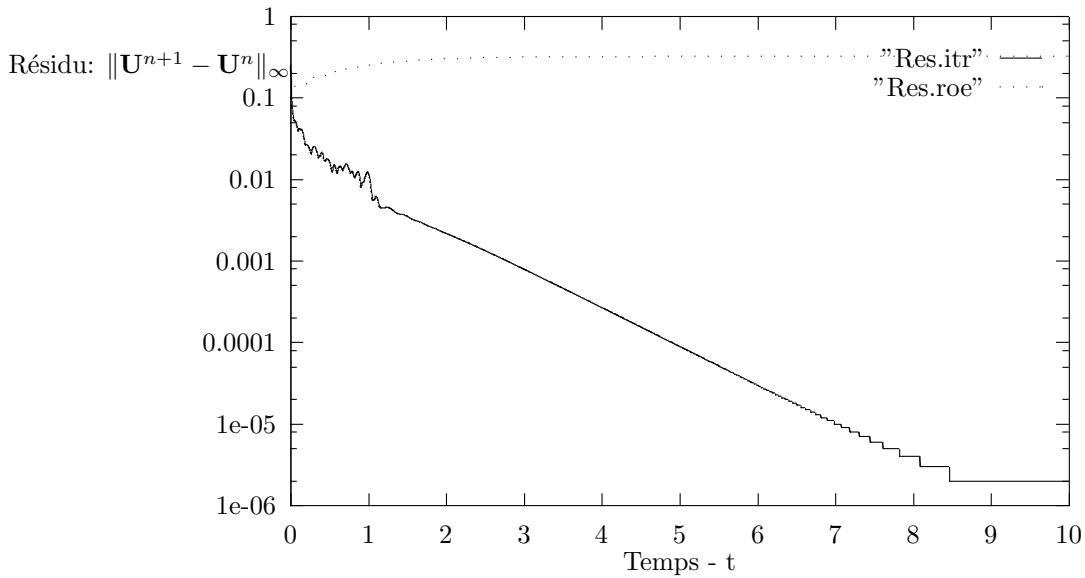


Figure 12.9: Historique des résidus (schéma non-conservatif/pas fractionnaires)

12.3.3 Pertes de charge dans un canal pourvu de deux demi-grilles

On considère maintenant un canal semblable dont la géométrie est décrite à la figure 12.10. Les parois sont matérialisées par des conditions aux limites de type «miroir». On maintient le terme réactionnel nul: $k_1 \equiv 0$.

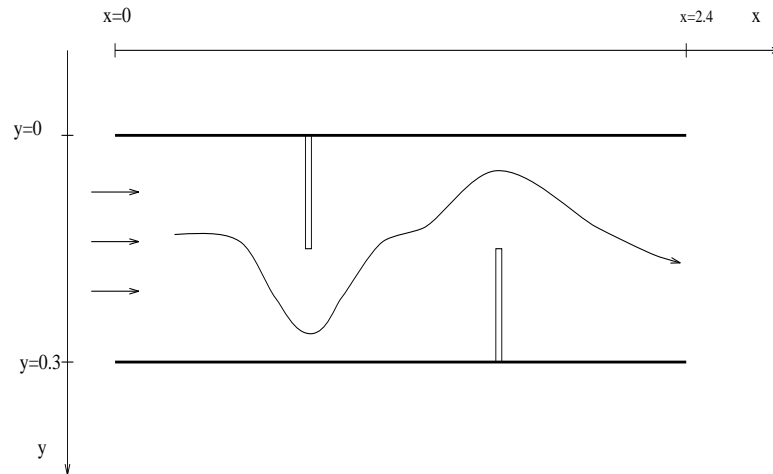


Figure 12.10: Géométrie du canal et allure des lignes de courant

Afin de simuler les effets de pertes de charges dus aux grilles de mélange, on définit la fonction k_2 sous la forme suivante:

$$k_2(x, y) = 20\mathbf{1}_{x=0.6; y \leq 0.15} + 50\mathbf{1}_{x=1.2; y \geq 0.15}$$

La donnée initiale et la condition-limite à gauche correspondent à un écoulement uniforme. Concrètement, on a choisi:

$$\left| \begin{array}{l} \rho_0 = 2 \\ c_0 = 0.5 \\ u_0 = 0.5 \\ v_0 = 0 \end{array} \right. \quad \text{et} \quad \left| \begin{array}{l} c_l = 0.5 \\ (\rho u)_l = 1 \\ v_l = 0 \end{array} \right.$$

Le domaine de calcul court maintenant sur l'intervalle $x \in [0; 2.4]$ en vue d'observer le retour à un écoulement monodimensionnel loin en aval des grilles. Le maillage est identique: $\Delta x = \Delta y = 0.02$, ce qui impose $\Delta t = 0.005$. On itère jusque $T = 10$ où un régime permanent est établi. Les figures 12.11, 12.12, 12.15, 12.14, 12.13 donnent respectivement la vitesse axiale, la vitesse transverse, le débit axial, le nombre de Mach et la pression stationnaires. On observe des solutions stationnaires non-oscillantes malgré les importantes discontinuités de pression. Le profil de vitesse axiale est réaliste, mais illustre les limites d'une simulation monodimensionnelle puisque les discontinuités sont modérées en comparaison de celles subies par la pression (ce qui n'est pas le cas pour un problème 1D!) même à proximité de la paroi. Le débit axial est continu y compris dans les pertes de charge singulières. Quant à la vitesse transverse, elle n'est affectée par les grilles que très localement: il n'y a aucune remontée amont des perturbations. Les paliers de pression sont stables et ne s'étalent pas sur plusieurs mailles (ceci n'est pas vérifié dans l'approche classique). Le nombre de Mach reste subsonique dans tout le domaine, ce qui est une condition nécessaire de bon fonctionnement du schéma non-conservatif (12.20). La figure 12.16 permet de comparer l'historique des résidus pour les deux approches numériques. Le schéma non-conservatif atteint la précision machine, alors que l'algorithme conventionnel stagne. Ceci semble être dû au fait que la pression est mal évaluée de part et d'autre des grilles de mélange, ce qui crée des oscillations sur le profil de vitesse axiale. Il faut alors raffiner le maillage spatial pour corriger cela.

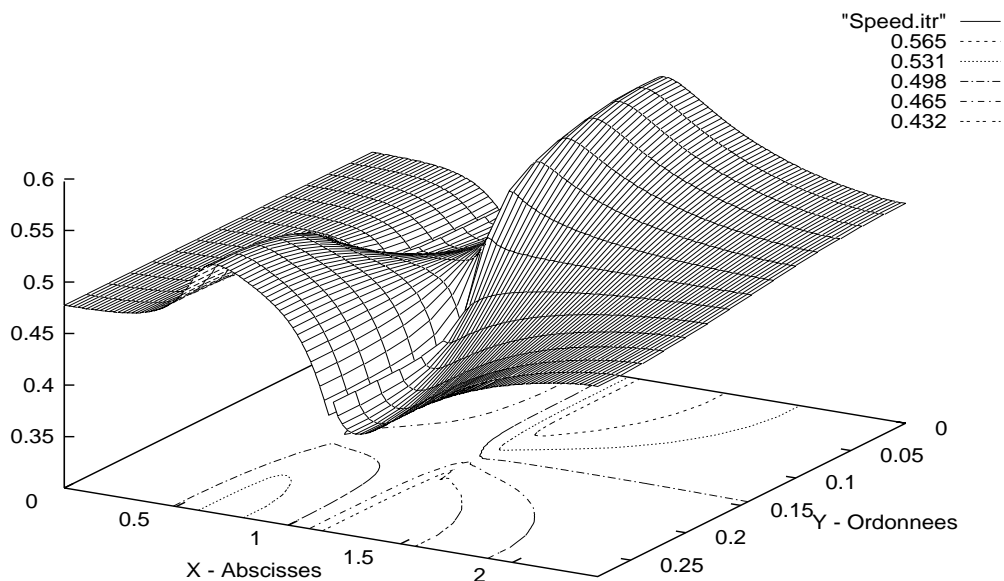


Figure 12.11: Vitesse axiale stationnaire; $T=10$ (schéma non-conservatif)

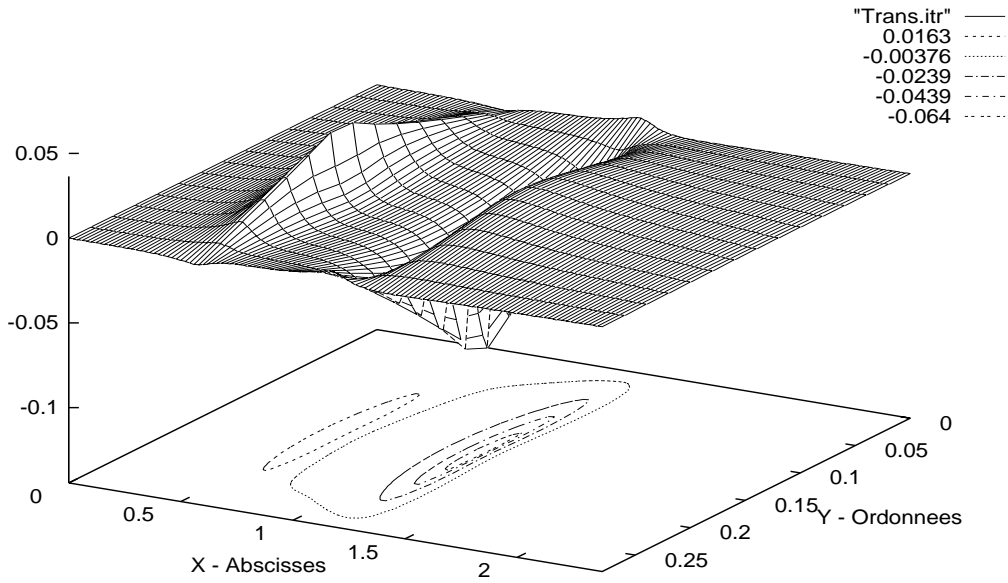


Figure 12.12: Vitesse transverse stationnaire; T=10 (schéma non-conservatif)

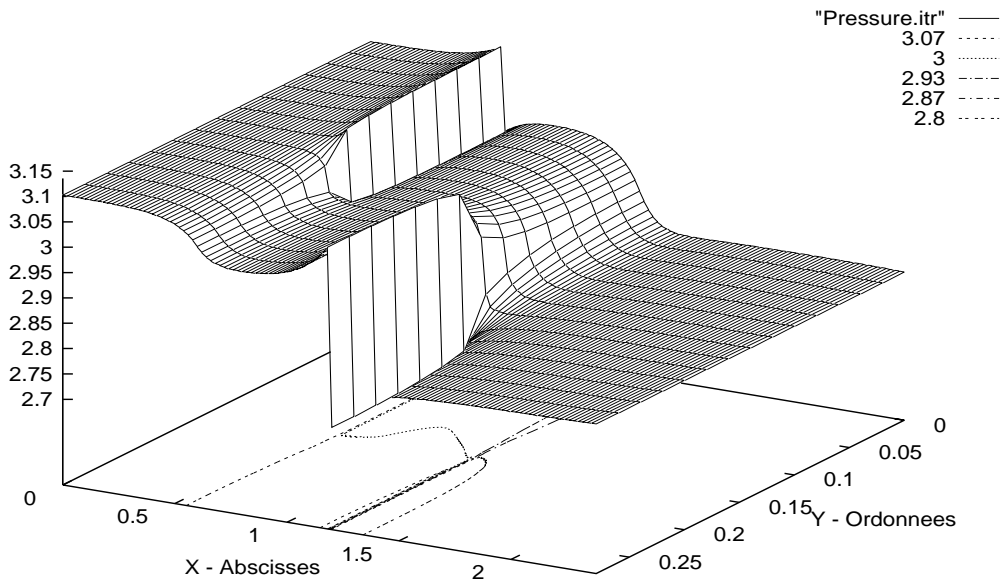


Figure 12.13: Pression stationnaire; T=10 (schéma non-conservatif)

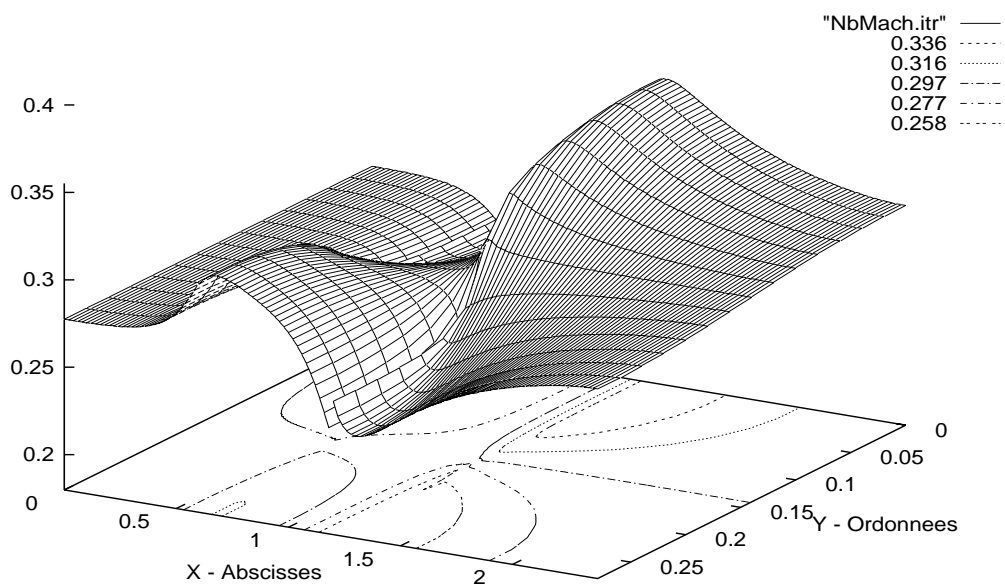


Figure 12.14: Nombre de Mach stationnaire; T=10 (schéma non-conservatif)

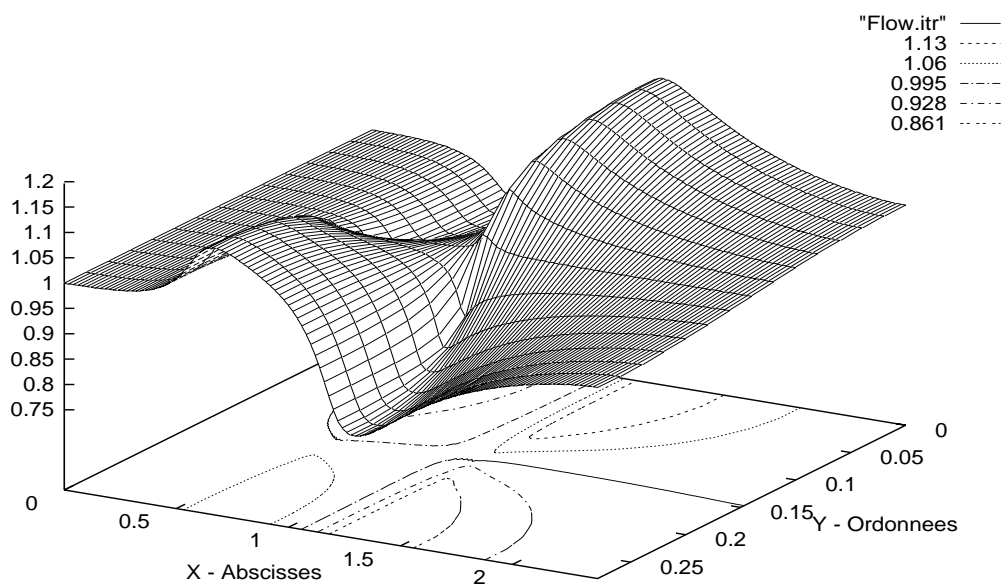


Figure 12.15: Débit axial stationnaire; T=10 (schéma non-conservatif)

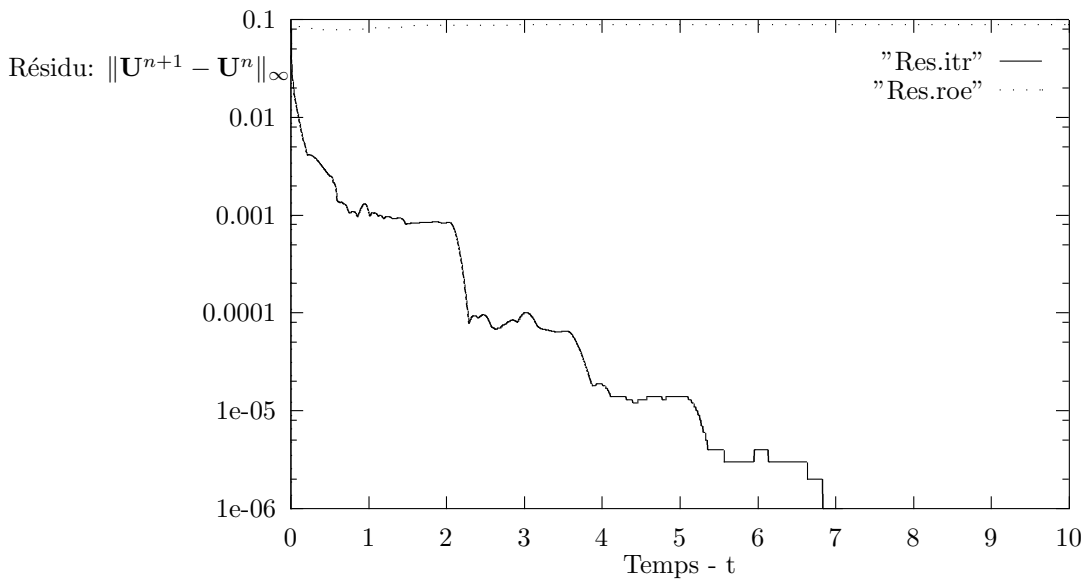


Figure 12.16: Historique des résidus (schéma non-conservatif/pas fractionnaires)

12.3.4 Effets combinés de recondensation et de perte de charge

Il reste à tester le comportement du schéma non-conservatif pour un problème combinant l'action des deux termes sources (i.e. $k_1 \neq 0$ et $k_2 \neq 0$). Pour cela, on considère un canal dont la géométrie est donnée par la figure 12.17. Les parois sont matérialisées par des conditions aux limites de type mirror .

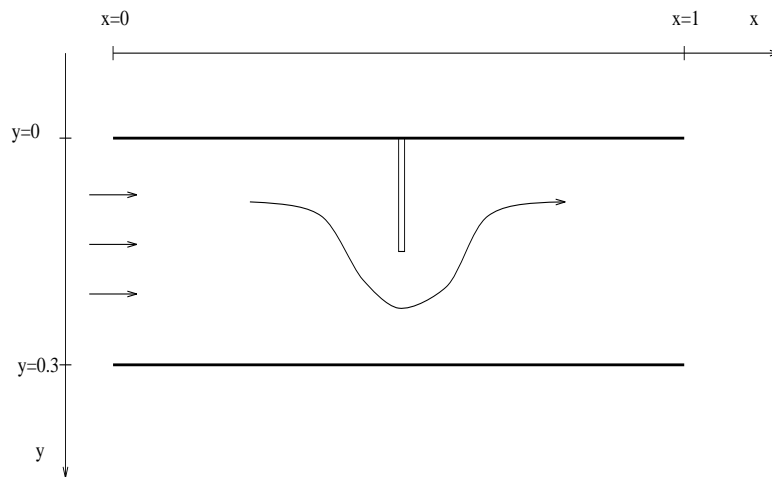


Figure 12.17: Géométrie du canal et allure des lignes de courant

On simule un écoulement diphasique cisailé soumis à un effet de recondensation inégal en espace et à une perte de charge singulière occasionnée par une grille de mélange. Concrètement, les données initiales et la condition-limite à gauche sont de la forme:

$$\left\{ \begin{array}{l} \rho_0 = 2 \\ c_0 = 0.8\mathbf{1}_{y < 0.15} + 0.2\mathbf{1}_{y > 0.15} \\ u_0 = 1 \\ v_0 = 0 \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{l} c_l = 0.8\mathbf{1}_{y < 0.15} + 0.2\mathbf{1}_{y > 0.15} \\ u_l = 1 \\ v_l = 0 \end{array} \right.$$

Quant aux coefficients des termes sources, ils sont donnés par ²:

$$k_1 \equiv 100; \mu = 0.3\mathbf{1}_{y < 0.15} + 0.1\mathbf{1}_{y > 0.15}; k_2 = 10\mathbf{1}_{x=0.5; y \leq 0.15}$$

Ce problème est motivé par les expérimentations industrielles réalisées sur le code FLICA-4 mettant en oeuvre une discrétisation conventionnelle des termes sources. On observe en régime suffisamment raide l'apparition de concentrations négatives ainsi qu'une remontée amont des perturbations de la vitesse transverse jusque l'entrée du domaine de calcul. Dans le cas présent, celui-ci court uniquement sur l'intervalle $x \in [0; 1]$. Le maillage est identique: $\Delta x = \Delta y = 0.02$, ce qui impose $\Delta t = 0.005$. On itère jusque $T = 2$ où un régime permanent est établi. Les figures 12.18, 12.19, 12.22, 12.21, 12.20, 12.23 donnent respectivement la vitesse axiale, la vitesse transverse, le débit axial, le nombre de Mach, la pression et la concentration massique stationnaires. On observe des solutions non-oscillantes et stables sur ce cas-test raide. La vitesse transverse est localement perturbée, que ce soit à cause de l'effet de recondensation inégal ou bien de l'effet de la grille de mélange. Les profils de pression et de vitesse axiale sont réalistes et ne présentent pas d'oscillations parasites. Le débit axial est continu à travers la perte de charge singulière. Quant au nombre de Mach, il reste subsonique dans tout le domaine. La figure 12.24 permet de comparer l'historique des résidus pour les deux approches numériques. On constate que l'algorithme du point fixe servant à résoudre le système différentiel couplé introduit une erreur qui empêche d'obtenir la convergence du résidu à précision machine. Néanmoins, il faut souligner que la différence fondamentale entre les deux schémas numériques provient du fait que l'approximation conventionnelle est totalement fautive. En effet, le schéma à pas fractionnaires calcule mal les paliers de part et d'autre de la perte de charge et ceux-ci interfèrent au bout d'un certain temps avec la couche-limite de concentration aux abords de l'entrée du canal. Ceci crée rapidement une onde supersonique (voir figure 12.25) qui s'empresse de dégrader toute la solution et fait exploser le schéma. Ceci n'est pas dû au traitement du terme de réaction puisque l'on constate sur la figure 12.26 que les profils de concentration massique issus des deux schémas sont quasiment identiques. Lorsque l'on raffine le maillage, on observe que l'on retarde l'apparition de la couche-limite supersonique, mais celle-ci finit toujours par arriver. Ceci amène à penser que les stationnaires issus d'un schéma classique à pas fractionnaires sont le plus souvent faux. Le raffinement de maillage ne fait que retarder l'apparition des erreurs; cette conclusion est partagée par [18].

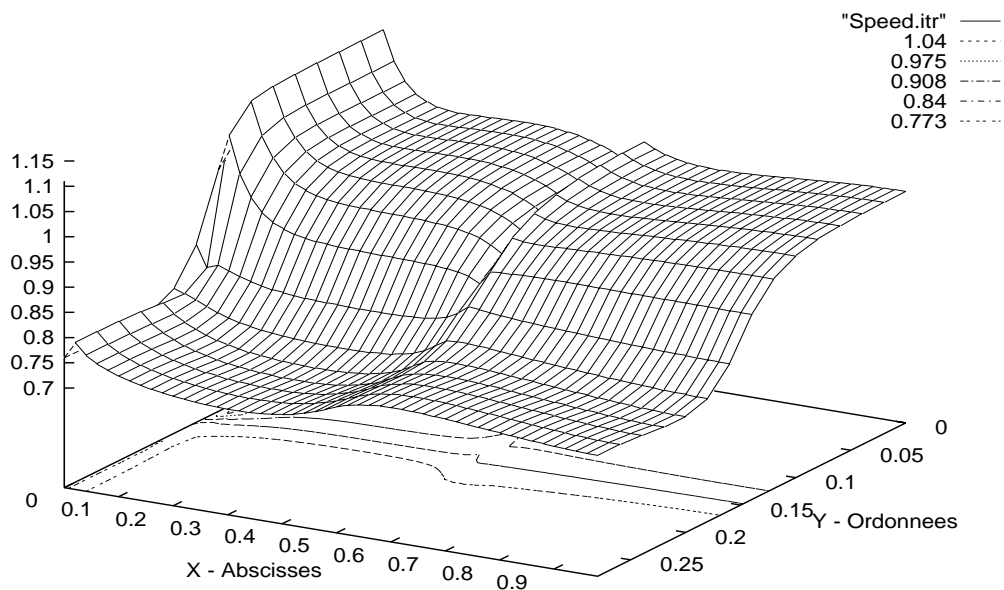


Figure 12.18: Vitesse axiale stationnaire; T=2 (schéma non-conservatif)

2. Ce cas-test a été proposé par I. Toumi

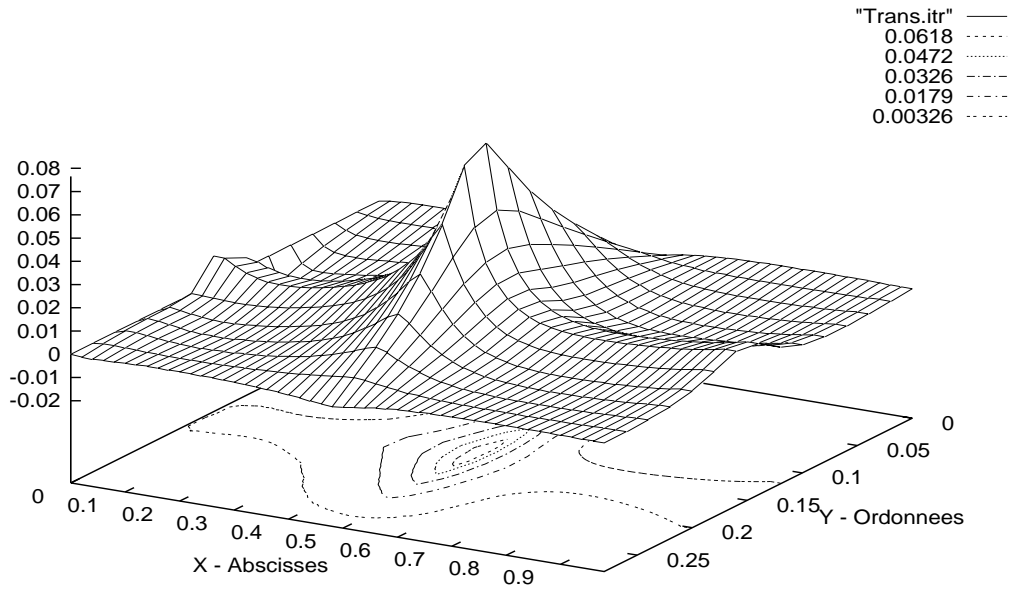


Figure 12.19: Vitesse transverse stationnaire; T=2 (schéma non-conservatif)

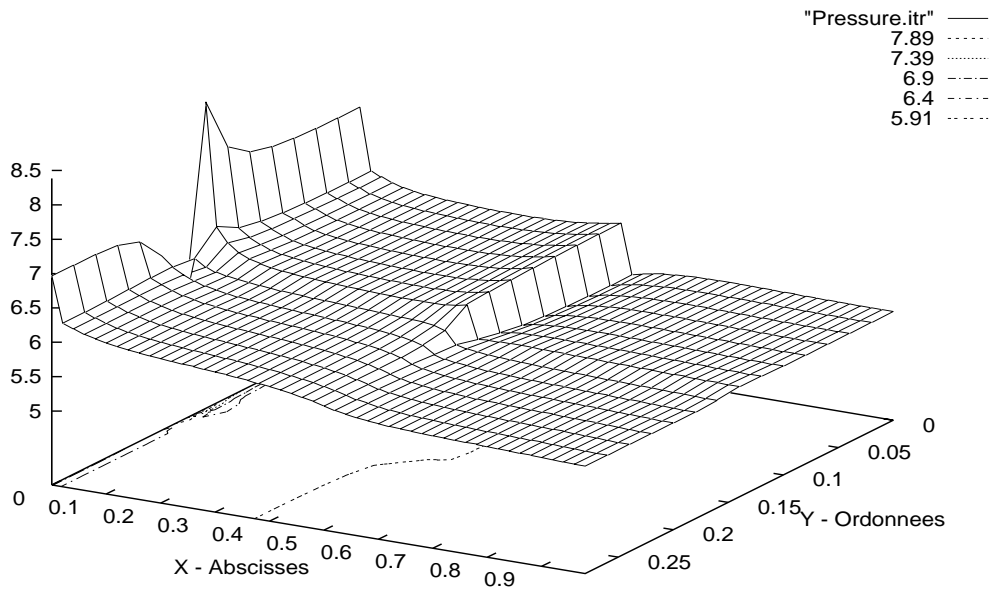


Figure 12.20: Pression stationnaire; T=2 (schéma non-conservatif)

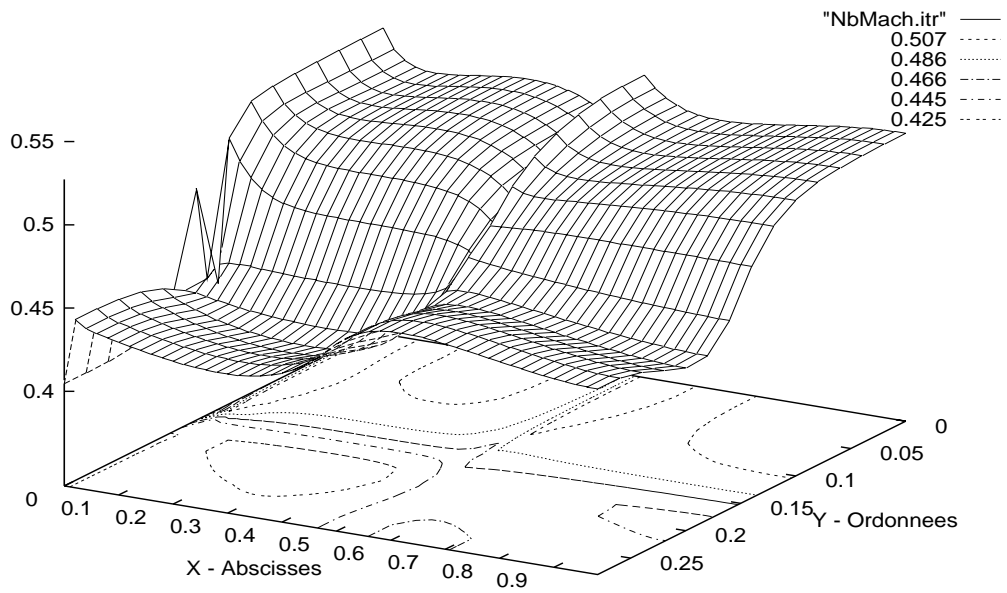


Figure 12.21: Nombre de Mach stationnaire; T=2 (schéma non-conservatif)

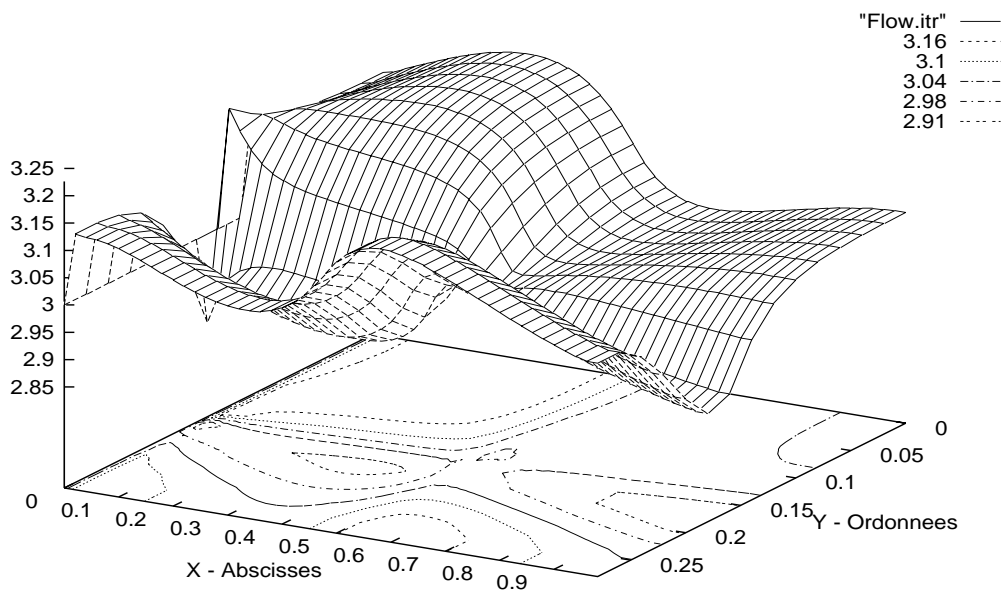


Figure 12.22: Débit axial stationnaire; T=2 (schéma non-conservatif)

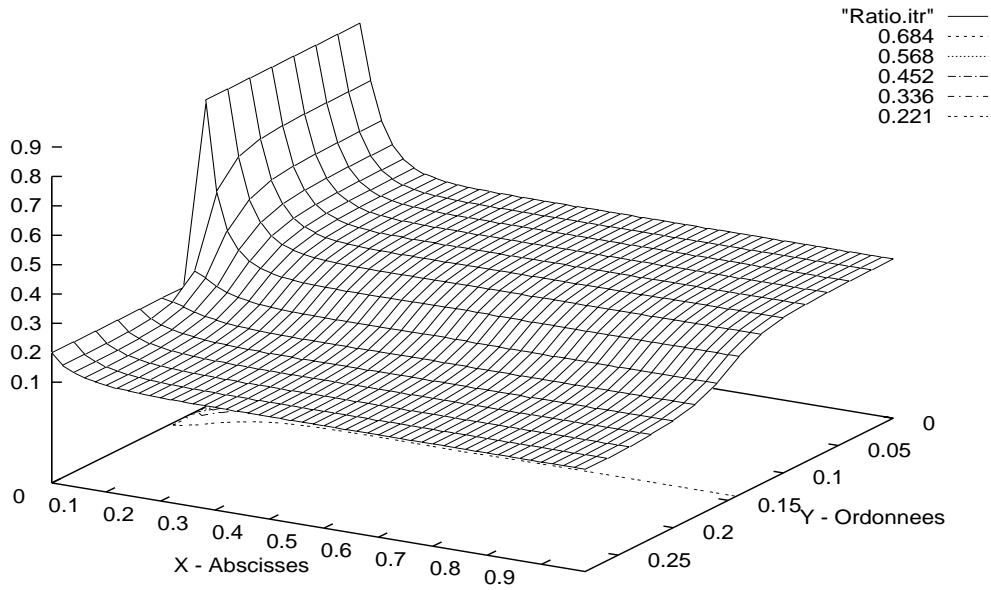


Figure 12.23: Concentration massique stationnaire; T=2 (schéma non-conservatif)

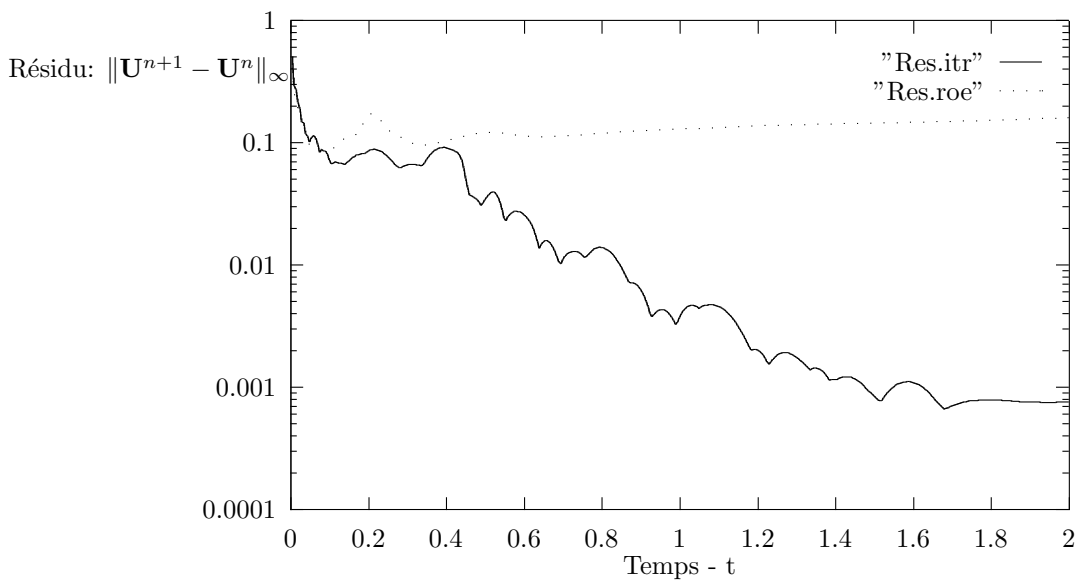


Figure 12.24: Historique des résidus (schéma non-conservatif/pas fractionnaires)

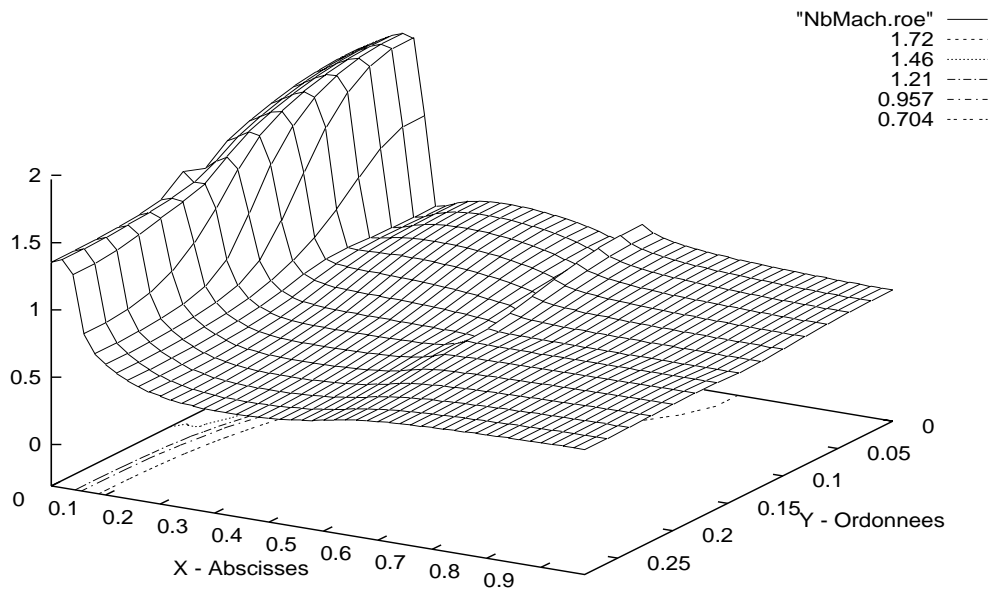


Figure 12.25: Nombre de Mach transitoire; $T=1.5$ (schéma à pas fractionnaires)

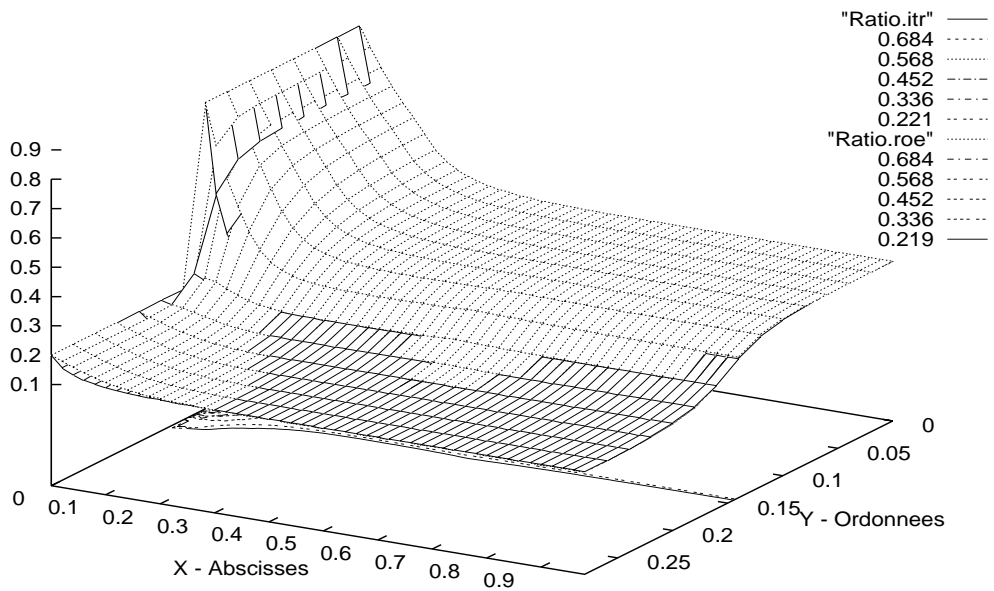


Figure 12.26: Comparaison des concentrations transitoires; $T=1.5$ (schéma non-conservatif / pas fractionnaires)

Conclusion générale

On s'est attaché dans cette thèse à comprendre le mieux possible les problèmes posés par l'adjonction de termes sources dans les systèmes hyperboliques de lois de conservation, tant du point de vue théorique que pour construire un schéma numérique efficace, robuste et capable de converger vers des régimes stationnaires réalistes.

Les travaux menés dans la Partie I montrent qu'il est possible d'envisager la construction de la solution entropique d'une loi scalaire réactive avec seulement une estimation en amplitude sur la suite d'approximations obtenue par perturbations visqueuses. Ceci est bien adapté au cas des problèmes avec termes sources, vu que le contrôle de la norme infinie peut s'exprimer de façon simple sur ces derniers. Le problème est beaucoup plus délicat lorsque l'on traite des systèmes puisqu'il est alors nécessaire de mettre en évidence des zones globalement invariantes dans l'espace des phases; les expressions des flux convectifs et du second membre se mêlent alors et rendent hypothétique l'existence de celles-ci. On est donc dans la majorité des cas incapable d'obtenir une estimation a priori sur la solution recherchée. Ceci semble condamner cette approche pour établir une théorie satisfaisante de ce type d'équations. Quant au problème de Riemann, sa structure autosimilaire est irrémédiablement perdue, ce qui le prive d'une grande partie de ses attraits.

A la suite des résultats obtenus sur les lois scalaires, il est apparu intéressant d'étudier les approximations numériques de celles-ci. Pour cela, on a considéré en Partie II les algorithmes implicites de façon à se libérer le plus possible d'éventuelles contraintes de pas de temps. En fait, on constate que pour s'assurer d'un bon comportement de ce type de schémas, il est encore nécessaire de respecter une condition d'inversibilité de type CFL. D'autre part, dans le cas où le second membre possède plusieurs équilibres, il est indispensable de prévenir l'apparition d'oscillations parasites dans la solution approchée. Une analyse en variation totale montre que des étapes explicites sont présentes dans la majorité des routines implicites. Il est donc nécessaire de moduler le pas de temps en fonction de celles-ci si on veut éviter une dérive sévère de la suite d'approximations au voisinage des équilibres instables du terme source.

Pour tenter de corriger cet état de fait, on a proposé à la Partie III un nouveau solveur de Riemann construit à partir d'une réécriture du terme source sous la forme d'un produit non-conservatif. L'intérêt de cette démarche est triple:

- on maintient une structure autosimilaire au niveau des problèmes de Riemann locaux.
- la stabilité du schéma de Godunov ainsi obtenu est uniquement prescrite par les termes convectifs.
- le schéma converge en temps vers un profil stationnaire correct: propriété *well-balanced*.

L'extension de cet algorithme à des systèmes hyperboliques généraux s'est effectuée par l'intermédiaire des travaux de DalMaso *et al.* et de Toumi. En effet, il a été possible de mettre en place une régularisation générique des produits non-conservatifs susceptibles d'apparaître dans une large variété de problèmes. Ceci nous a conduit à la formulation d'un schéma numérique construit sur la base d'une linéarisation de type Roe. Les résultats obtenus à l'aide de celle-ci furent très satisfaisants.

La principale faiblesse apparue lors de l'extension aux systèmes menée précédemment a été la perte de robustesse liée à la linéarisation. Pour corriger cela, on a présenté en Partie IV un schéma non-linéaire construit à partir de la décomposition de flux de VanLeer où les termes sources traités de façon non-conservative induisent des relations de saut généralisées à chaque interface du maillage. Cette approche permet de retrouver une grande partie de la stabilité observée dans le cas scalaire. Cette construction permet de résoudre une grande variété de problèmes physiquement réalistes; elle a aussi été testée dans le contexte des systèmes relaxés où une théorie récente prescrit les comportements asymptotiques des solutions. L'extension bidimensionnelle de ce type de schémas n'est pas directe: on a donc proposé une formulation conduisant à un algorithme simple à implémenter. Celui-ci a été testé dans le contexte des écoulements diphasiques en présence de frottements uniformes, de pertes de charges singulières et d'effets combinés de recondensation. Le comportement observé est très satisfaisant puisque l'on obtient au stationnaire des régimes de meilleure qualité que dans le cas d'une approche classique à pas fractionnaires.

Annexe A

Le problème de Riemann pour le modèle diphasique équilibré

Cette annexe est consacrée à l'étude du système (4.1) dans le contexte du *problème de Riemann* dans le cas homogène. Les résultats proviennent de la thèse de I. Toumi [81]: la motivation est ici de les exposer sous une forme plus simple. On insiste aussi sur une particularité liée au choix de la variable de concentration massique plutôt que le taux de vide volumique. Pour tout ce qui concerne l'influence des termes sources sur la structure de la solution, on renvoie au travail de P.G. LeFloch [59].

A.1 Hyperbolicité du système

Commençons par rappeler que la solution du problème de Riemann pour un système strictement hyperbolique existe et est unique dans la classe des fonctions composées de chocs, détente, ou discontinuités de contact, si toutefois les états initiaux sont suffisamment proches [57, 79].

Proposition IV.5 *Supposons $p_\rho(\rho, c) > 0$, $p_{\rho\rho}(\rho, c) > 0$, le système:*

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0 \\ \partial_t(\rho c) + \partial_x(\rho u c) = 0 \\ \partial_t(\rho u) + \partial_x(\rho u^2 + p) = 0 \end{cases} \quad (\text{A.1})$$

admet 3 vitesses caractéristiques réelles $\lambda_1 = u - a$, $\lambda_2 = u$, $\lambda_3 = u + a$ où $a = \sqrt{p_\rho(\rho, c)}$ désigne la vitesse du son dans le mélange. Les champs associés aux vecteurs propres:

$$\vec{R}_1 = \begin{pmatrix} \rho \\ 0 \\ a \end{pmatrix} \quad \vec{R}_2 = \begin{pmatrix} p_c \\ -p_\rho \\ 0 \end{pmatrix} \quad \vec{R}_3 = \begin{pmatrix} \rho \\ 0 \\ -a \end{pmatrix} \quad (\text{A.2})$$

sont vraiment non-linéaires (1^{er} et 3^{eme}) ou linéairement dégénéré (2nd).

Le problème (4.1) est donc strictement hyperbolique et le problème de Riemann pour (A.1) admet une unique solution pour des discontinuités initiales suffisamment modérées.

On peut d'autre part remarquer que la structure des entropies est identique à celle de (4.3). Par analogie avec la dynamique des gaz isentropiques en simple phase, on déduit que $\eta(\rho, \rho c, \rho u) = \frac{\rho u^2}{2} + \rho \int^\rho \frac{p(s, c)}{s^2} ds$ convient pour (A.1).

Démonstration

On passe en variables non-conservatives pour alléger l'écriture.

Soient $W = (\rho, c, u)$, le système (A.1) se réécrit $W_t + A(W).W_x = 0$ avec:

$$A(W) = \begin{pmatrix} u & 0 & \rho \\ 0 & u & 0 \\ \frac{p_\rho}{\rho} & \frac{p_c}{\rho} & u \end{pmatrix} \text{ de spectre } \Lambda = \{u \pm a, u\} \text{ avec } a = \sqrt{p_\rho}$$

Les vecteurs propres à droite \vec{R}_i sont tels que $(\frac{\partial F(W)}{\partial W} - \lambda_i Id).\vec{R}_i = \vec{0}$. On obtient donc les expressions (A.2). Il est alors immédiat que $\langle \nabla \lambda_2, \vec{R}_2 \rangle_{\mathbb{R}^3} = 0$ puisque \vec{R}_2 ne dépend pas de u . D'autre part, il vient:

$$\langle \nabla \lambda_{1/3}, \vec{R}_{1/3} \rangle_{\mathbb{R}^3} = \pm(\rho \cdot a_\rho + a) \neq 0$$

sauf si a est linéaire en ρ . \square

A.2 Expression des courbes d'ondes

Afin de déterminer les courbes de choc et de détente du problème (A.1), il est nécessaire d'en chercher les *invariants de Riemann*.

Lemme IV.7 *Les invariants de Riemann associés à (A.1) sont de la forme:*

$$\begin{cases} w = c \\ w = u \pm \int^\rho \frac{a(s, c)}{s} ds & \text{pour les champs 1 et 3} \\ w = p \\ w = u & \text{pour le champ 2} \end{cases} \quad (\text{A.3})$$

Démonstration

Par définition [79], on a:

$$\langle \nabla w, \vec{R}_{1/3} \rangle_{\mathbb{R}^3} = 0 \text{ d'où } \begin{cases} 0.w_c = 0 \\ w = w_\rho \pm \frac{a}{\rho} w_u = 0 \end{cases}$$

D'autre part, on sait déjà que $\lambda_2 = u$ est constante le long des courbes intégrales de \vec{R}_2 puisque le champ est LD Il reste a remarquer que:

$$\langle \nabla w, \vec{R}_2 \rangle_{\mathbb{R}^3} = 0 \text{ impose } w_\rho p_c - w_c p_\rho = 0$$

D'où (A.3). \square

On peut donc donner l'expression des courbes d'ondes puisque les chocs se trouvent en éliminant la propagation dans les relations de Rankine-Hugoniot et les détentees sont telles que les invariants de Riemann restent constants.

Proposition IV.6 *Les courbes d'ondes entropiques associées au problème de Riemann (U_L, U_R) pour (A.1) sont de la forme suivante:*

$$\begin{cases} C_{1/3} = \left\{ U \in \mathbb{R}^+ \times [0, 1] \times \mathbb{R} \text{ tels que } \begin{array}{l} (u - u_L)^2 = \frac{\rho - \rho_L}{\rho \cdot \rho_L} (p - p_L) \\ c = c_L \end{array} \right\} \\ p \text{ augmente dans les 1-chocs et diminue dans les 3-chocs} \\ D_{1/3} = \left\{ U \in \mathbb{R}^+ \times [0, 1] \times \mathbb{R} \text{ tels que } \begin{array}{l} (u - u_L) = \pm \int_{\rho_L}^{\rho} \frac{a(s, c_L)}{s} ds \\ c = c_L \end{array} \right\} \end{cases} \quad (\text{A.4})$$

Démonstration

C'est une réécriture d'un résultat de [81]. La différence vient du fait que l'on travaille avec une fraction massique c au lieu d'un taux de vide α . Les courbes de détente sont immédiates du fait de l'expression des invariants (A.3). Pour les 1/3-chocs de vitesse σ , on écrit les relations de Rankine-Hugoniot:

$$\begin{cases} \sigma[\rho] = [\rho u] \text{ donc } \sigma = \frac{[\rho u]}{[\rho]} \\ \sigma[\rho c] = [\rho u c] \\ \sigma[\rho u] = [\rho u^2 + p] \text{ donc } \frac{[\rho u]^2}{[\rho]} = [\rho u^2 + p] \end{cases}$$

Tout d'abord, la concentration c est constante puisque si on pose $v = u - \sigma \neq 0$, il vient:

$$[\rho v] = 0 \Rightarrow (\rho v) \cdot [c] = 0$$

Par conséquent, c est conservé le long des ondes de vitesse différente de u , c'est-à-dire que c ne varie qu'au contact.

Pour le reste, on pose $M = \rho u$ et il vient:

$$(M - M_L)^2 = (\rho - \rho_L)(u \cdot M - u_L \cdot M_L) + (\rho - \rho_L)(p - p_L)$$

On développe le membre de droite et on trouve:

$$\rho \cdot \rho_L (u - u_L)^2 = (\rho - \rho_L)(p - p_L)$$

Les conditions d'entropie sont celles de Lax [57] précisées dans [81]. \square

Une conséquence directe de cette proposition est:

- c est constante à travers les ondes vraiment-non-linéaires (en particulier les chocs)
- c n'est discontinue qu'aux contacts le long desquels u est constante. Et donc, on déduit que le terme $u \cdot c_x$ n'est jamais un produit de distributions puisque les deux quantités ne sont jamais discontinues aux mêmes endroits. Dans le cadre du problème de Riemann homogène pour (A.1), le profil transitoire de concentration est explicitement calculable au moyen de la fonction de Heaviside Y :

$$c(x, t) = c_L + (c_R - c_L) \cdot Y(x - \bar{u}t)$$

où c_L, c_R désignent les paliers initiaux de concentration et \bar{u} la valeur de la vitesse dans la zone délimitée par les deux ondes acoustiques. Ceci justifie aussi une écriture non-conservative de l'équation de transport sur la concentration lorsque la pression ne dépend que de la densité ρ (voir par exemple [24] pour une exploitation numérique de cette remarque).

Annexe B

Méthodes multiples stables en variation totale

On présente rapidement ici une extension des idées du chapitre précédent aux méthodes multipas. Après un récapitulatif des résultats classiques pour ce genre d'approximations, on donne l'expression de l'unique méthode d'ordre 2 A -stable préservant la variation totale pour des pas de temps arbitrairement grands. Les références sont les livres [19, 42].

B.1 Généralités

Les méthodes multipas se présentent comme une alternative *linéaire* aux autres méthodes de résolution d'équations différentielles ordinaires (comme les routines de Runge-Kutta, par exemple). Soit le problème générique:

$$\dot{y} = f(y) \text{ avec } y(0) = y_0$$

alors on peut écrire ce type d'algorithme en toute généralité:

$$\begin{cases} \alpha_k y_{n+k} + \alpha_{k-1} y_{n+k-1} + \dots + \alpha_0 y_n = \Delta t (\beta_k f_{n+k} + \beta_{k-1} f_{n+k-1} + \dots + \beta_0 f_n) \\ \alpha_k \neq 0 \text{ et } |\alpha_0| + |\beta_0| > 0 \\ y_i, f_i \text{ sont des approximations de } y(i\Delta t), f(y)(i\Delta t) \end{cases} \quad (\text{B.1})$$

Le schéma (B.1) est dit implicite si $\beta_k \neq 0$; il est explicite sinon. Pour étudier l'ordre de précision en temps et la stabilité, il est naturel d'introduire les polynômes suivants (prendre $f(y) = y$):

$$\begin{cases} \rho(z) = \alpha_k z^k + \alpha_{k-1} z^{k-1} + \dots + \alpha_0 \\ \sigma(z) = \beta_k z^k + \beta_{k-1} z^{k-1} + \dots + \beta_0 \end{cases}$$

On a alors le résultat:

Théorème IV.6 *La méthode (B.1) est d'ordre $p \geq 1$ si et seulement si:*

$$\sum_{j=0}^k \alpha_j = 0 \text{ et } \sum_{j=0}^k \alpha_j j^q = q \sum_{j=0}^k \beta_j j^{q-1} \text{ pour } q = 1 \dots p$$

La consistance correspond à imposer un ordre $p = 1$. Cela se traduit par:

$$\rho(1) = 0 \text{ et } \sigma(1) = \rho'(1)$$

Théorème IV.7 (Dahlquist)

Le schéma (B.1) est stable si le polynôme ρ vérifie:

- les racines de ρ sont de module inférieur à 1
- les racines de module 1 sont simples

Pour le cas particulier des différences rétrogrades ($\beta_0 = \beta_1 = \dots = \beta_{k-1} = 0$), on a:

Théorème IV.8 *Les méthodes du type:*

$$\sum_{j=1}^k \alpha_j y_{n+j} = \Delta t f_{n+k}$$

sont instables pour $k \geq 7$.

Pour étudier la A -stabilité de ces méthodes, on introduit l'équation $\dot{y} = \lambda y$ et il vient:

Définition IV.6 *Soit D le domaine du plan complexe défini par les $\mu = \lambda \Delta t \in \mathbb{C}$ tels que:*

- les racines de $\zeta(z) = \rho(z) - \mu \sigma(z)$ sont de module inférieur à 1
- les racines de module 1 sont simples

Alors la méthode (B.1) est dite A -stable si on a:

$$\{\mu \in \mathbb{C} \text{ tels que } \operatorname{Re}(\mu) \leq 0\} \subset D$$

Pour $\mu = 0$, il vient $\zeta(z) = \rho(z)$ et la stabilité usuelle correspond à $0 \in D$. Les méthodes aux différences rétrogrades d'ordre inférieur ou égal à 2 sont A -stables.

B.2 Représentation et stabilité des méthodes à 2 pas

Se restreignant aux méthodes (B.1) à deux pas, on a une représentation:

$$(1 + \xi)y_{n+2} - (1 - 2\xi)y_{n+1} + \xi y_n = \Delta t[\theta f_{n+2} + (1 - \theta + \phi)f_{n+1} - \phi f_n] \quad (\text{B.2})$$

On a alors les polynômes:

$$\begin{cases} \rho(z) = (1 + \xi)z^2 - (1 + 2\xi)z + \xi \\ \sigma(z) = \theta z^2 - (1 - \theta + \phi)z - \phi \end{cases}$$

Il est alors possible [3] d'établir que:

- (B.2) sera d'ordre 2 sous la condition $\phi = \xi - \theta + \frac{1}{2}$
- La A -stabilité correspond à:

$$\theta \geq \phi + \frac{1}{2} \text{ et } -\frac{1}{2} \leq \xi \leq \theta + \phi - \frac{1}{2}$$

Ceci est suffisant pour obtenir:

Proposition IV.7 *Il existe une unique méthode d'ordre 2 A -stable et non-oscillante pour tout Δt .*

Démonstration

D'après le chapitre précédent, il suffit d'imposer:

$$\beta_1 = \beta_0 = 0 \Leftrightarrow 1 - \theta + \phi = 0 = \phi$$

D'autre part, l'ordre 2 impose: $\phi = 0 = \xi - \theta + \frac{1}{2}$

On rassemble et il vient: $\theta = 1$ et $\xi = \frac{1}{2}$

Ceci donne (voir (6.25)):

$$\frac{3}{2}y_{n+2} - 2y_{n+1} + \frac{1}{2}y_n = \Delta t f_{n+2}$$

On vérifie immédiatement que:

$$\phi = \xi - \theta + \frac{1}{2} \Leftrightarrow 0 = \frac{1}{2} - 1 + \frac{1}{2}$$

D'où la A -stabilité. \square

Ceci est bien cohérent avec la théorie des différentiations rétrogrades qui donnent des méthodes assez peu précises mais très bien adaptées aux problèmes mal conditionnés.

Annexe C

Rappels sur la définition des produits non-conservatifs

La théorie exposée ici vise à donner un sens à l'expression $A(W)\frac{\partial W}{\partial x}$ lorsque $W \in [BV(\mathbb{R})]^{N \geq 2}$ et que $W \mapsto A(W) \in \mathbf{M}_N(\mathbb{R})$ est une application régulière. Dans le contexte des problèmes hyperboliques, il s'agit d'attribuer une valeur à un produit du type Dirac \times Heaviside $\delta_0(x)Y(x)$. Outre le formalisme développé par J.F. Colombeau et A.Y. LeRoux (l'algèbre des fonctions généralisées $\mathcal{G}(\Omega) \supset \mathcal{D}'(\Omega)$ [12]), G. DalMaso, P.G. LeFloch et F. Murat [21] ont réussi à définir rigoureusement de tels produits de distributions au moyen de la donnée d'une famille de chemins tracés dans l'espace des états Ω . Signalons que l'équivalence des deux points de vue pour des fonctions à variation bornée a été soulignée dans [44].

C.1 Interprétation des produits non-conservatifs comme mesures bornées

C.1.1 L'espace BV

Il est indispensable de rappeler la définition de l'espace BV des fonctions à variation bornée de $\mathbb{R} \rightarrow \mathbb{R}^N$ qui jouera un rôle de première importance dans la suite:

$$BV(\mathbb{R}, \mathbb{R}^N) = \left\{ W : \mathbb{R} \rightarrow \mathbb{R}^N / W_i \in L^1_{loc}(\mathbb{R}), \frac{\partial W_i}{\partial x} \in \mathcal{M}(\mathbb{R}) \right\} \quad (\text{C.1})$$

Où $L^1_{loc}(\mathbb{R})$ désigne l'espace des fonctions Lebesgue-intégrable sur tout compact de \mathbb{R} et \mathcal{M} l'espace des mesures de Borel (cf [75]). C'est-à-dire que l'on a:

$$\forall \varphi \in C^1_0(\mathbb{R}), \quad - \langle \varphi', W \rangle = \left\langle \varphi, \frac{\partial W}{\partial x} \right\rangle_{C^1_0 \times \mathcal{M}} = \int_{\mathbb{R}} \varphi \, d\nu$$

$\frac{\partial W}{\partial x}$ ainsi définie est la *dérivée au sens faible* de W . La mesure ν coïncide avec $\frac{\partial W}{\partial x}$ sur tous les boréliens de \mathbb{R} . Une semi-norme naturellement définie sur cet espace est la *variation totale*:

$$TV(W) = \left| \sup_{\varphi \in C^1_0(\mathbb{R}), \|\varphi\|_\infty \leq 1} \left\langle \varphi, \frac{\partial W}{\partial x} \right\rangle_{C^1_0 \times \mathcal{M}} \right| \quad (\text{C.2})$$

Rappelons aussi l'injection en dimension 1 d'espace: $BV(\mathbb{R}) \subset L^\infty(\mathbb{R})$. C'est donc un espace de Banach pour les deux normes $\|\cdot\|_{L^1} + TV(\cdot)$ et $\|\cdot\|_{L^\infty} + TV(\cdot)$. Il n'y a pas de résultat de densité des fonctions C^∞ à support compact puisque BV n'est pas inclus dans l'espace de Sobolev $W^{1,1}$.

C.1.2 Mesure de Borel associée

Le premier auteur à avoir interprété le produit ambigu $A(W)\frac{\partial W}{\partial x}$ est Vol'Pert [85]. Les résultats exposés dans la suite sont donc une extension de ses travaux.

Définition IV.7 On désigne par Φ une famille de chemins dans $\Omega \subset \mathbb{R}^N$, i.e. une famille d'applications régulières: $[0, 1] \times \Omega \times \Omega \rightarrow \Omega$ vérifiant les propriétés suivantes:

$$\begin{aligned} & \bullet \Phi(0; W_L, W_R) = W_L \\ & \bullet \Phi(1; W_L, W_R) = W_R \\ & \bullet \forall \mathcal{V} \text{ borné dans } \Omega, \exists k \text{ t.q. } \forall s \in [0, 1], \forall (U_L, U_R) \in \mathcal{V} \\ & \quad \left| \frac{\partial \Phi}{\partial s}(s; U_L, U_R) \right| \leq k |U_L - U_R| \\ & \bullet \forall \mathcal{V} \text{ borné dans } \Omega, \exists K \text{ t.q. } \forall s \in [0, 1], \forall (U_L^i, U_R^i)_{i=1,2} \in \mathcal{V} \\ & \quad \left| \frac{\partial \Phi}{\partial s}(s; U_L^1, U_R^1) - \frac{\partial \Phi}{\partial s}(s; U_L^2, U_R^2) \right| \leq K [|U_L^1 - U_L^2| + |U_R^1 - U_R^2|] \end{aligned} \quad (\text{C.3})$$

Enonçons de suite le résultat fondamental:

Théorème IV.9 (Dal Maso, LeFloch, Murat)

Soient $W \in BV([a, b], \mathbb{R}^N)$ et $A : \mathbb{R}^N \times]a, b[\rightarrow \mathbb{R}^N$ une fonction localement bornée, i.e.

$$\forall X \in \mathbb{R}^N \text{ borné, } \exists C > 0 \text{ telle que } \forall W \in X, \forall x \in]a, b[, |A(W, x)| \leq C$$

Alors il existe une unique mesure borelienne μ sur $]a, b[$ caractérisée par les deux propriétés:

- Si $x \mapsto W(x)$ est continue sur un ouvert $B \subset]a, b[$, alors

$$\mu(B) = \int_B A(W, x) \frac{\partial W}{\partial x} = \int_B A(W(x), x) \frac{\partial W(x)}{\partial x} dx$$

- Si $x_0 \in]a, b[$ est un point de discontinuité de $x \mapsto W(x)$, alors la mesure de Dirac en ce point est donné par

$$\mu(x_0) = \left[\int_0^1 A(\Phi(s; W(x_0^-), W(x_0^+)), x_0) \frac{\partial \Phi}{\partial s}(s; W(x_0^-), W(x_0^+)) ds \right] \delta(x_0)$$

La mesure de Borel μ ainsi introduite est appelée *produit non-conservatif* et se note habituellement sous la forme:

$$\left[A(W) \frac{\partial W}{\partial x} \right]_{\Phi} \quad (\text{C.4})$$

afin de faire clairement référence au chemin qui lui donne son sens.

Démonstration

Tout d'abord, on note que les intégrales sont bien définies:

$$\mu(B) = \int_B \underbrace{A(W(x), x)}_{\text{continue et bornée}} \underbrace{\frac{\partial W(x)}{\partial x}}_{\text{bornée}} dx \quad (\text{cas continu})$$

$$\mu(x_0) = \left[\int_0^1 A \left(\Phi \left(s; \underbrace{W(x_0^-), W(x_0^+)}_{\text{existent car } u \in BV} \right), x_0 \right) \underbrace{\frac{\partial \Phi}{\partial s}(s; W(x_0^-), W(x_0^+))}_{\text{Lipschitzienne en } W(x_0^\pm)} ds \right] \delta(x_0)$$

En un point de continuité x_0 , comme $\Phi(s; W(x_0^-), W(x_0^+)) = W(x_0)$, les deux formules coïcident. Supposons l'existence de $\mu \in \mathcal{M}$ vérifiant les deux intégrales, alors on peut l'écrire (cf [75]):

$$\mu(B) = \mu(\overset{\circ}{B}) + \sum_m \mu(x_m) \quad (\text{C.5})$$

avec $\overset{\circ}{B}$ ensemble des ouverts où $x \mapsto W(x)$ est continue et $(x_m)_m$ la collection (au plus dénombrable) de ses points de discontinuité. L'existence (et l'unicité) dans \mathcal{M} de μ est donc équivalente à la convergence uniforme de la série des modules dans (C.5). D'après la définition, comme A est localement bornée, $W \in BV \subset L^\infty$, et Φ Lipschitzienne, il vient que:

$$\sum_m |\mu(x_m)| \leq (\sup k) \|A\|_\infty \underbrace{\sum_m |W(x_m^-) - W(x_m^+)|}_{TV(W) < \infty}$$

□

L'exemple le plus simple est le *chemin des lignes droites* ou *chemin canonique* qui fait correspondre le produit non-conservatif ainsi défini avec le produit de superposition de Vol'Pert.

$$\Phi_S(s; W_L, W_R) = sW_L + (1-s)W_R$$

C.1.3 Deux propriétés importantes

Il s'agit de l'indépendance de (C.4) par rapport à la paramétrisation du chemin Φ , ainsi que la cohérence de la définition avec la théorie conservative dans le cas où $A(W) = dF(W)$.

Proposition IV.8 *Sous les hypothèses du théorème IV.9, la mesure $[A(W) \frac{\partial W}{\partial x}]_{\Phi}$ est indépendante de la paramétrisation de Φ au sens suivant:*

$$\forall \gamma \text{ difféomorphisme } [0, 1] \rightarrow [0, 1]$$

$$\tilde{\Phi}(s, W_L, W_R) \stackrel{\text{def}}{=} \Phi(\gamma(s), W_L, W_R) \Rightarrow [A(W) \frac{\partial W}{\partial x}]_{\tilde{\Phi}} = [A(W) \frac{\partial W}{\partial x}]_{\Phi}$$

Proposition IV.9 *Sous les mêmes hypothèses, si pour tout W , $A(W) = dF(W)$, alors la mesure $\left[A(W) \frac{\partial W}{\partial x}\right]_{\Phi}$ est indépendante de la donnée de Φ , et on a :*

$$\left[A(W) \frac{\partial W}{\partial x}\right]_{\Phi} = \frac{\partial F(W)}{\partial x}$$

C.2 Stabilité faible des produits non-conservatifs

C.2.1 Résultats forts de stabilité

Théorème IV.10 *(Dal Maso, LeFloch, Murat)*

• Soit $(W_n)_{n \in \mathbb{N}}$ une suite de fonctions de $BV(]a, b[, \mathbb{R}^N)$ convergeant fortement dans BV i.e. $W_n \xrightarrow{BV} W$, alors :

$$\left[A(W_n) \frac{\partial W_n}{\partial x}\right]_{\Phi} \xrightarrow{\mathcal{M}} \left[A(W) \frac{\partial W}{\partial x}\right]_{\Phi}$$

• Soit $(W_n)_{n \in \mathbb{N}}$ une suite de fonctions de $BV(]a, b[, \mathbb{R}^N)$ convergeant uniformément i.e. $W_n \xrightarrow{L^\infty} W$, alors :

$$\left[A(W_n) \frac{\partial W_n}{\partial x}\right]_{\Phi} \xrightarrow{weak} \left[A(W) \frac{\partial W}{\partial x}\right]_{\Phi}$$

i.e. $\int_{]a, b[} \varphi \left[A(W_n) \frac{\partial W_n}{\partial x}\right]_{\Phi} \rightarrow \int_{]a, b[} \varphi \left[A(W) \frac{\partial W}{\partial x}\right]_{\Phi}$ pour toute fonction test $\varphi \in C_0^1(]a, b[, \mathbb{R}^N)$.

Par exemple, la suite de fonctions à variation bornée $W_n(x) = \frac{1}{n} \sin(2\pi nx)$ tend uniformément vers 0, mais ne converge pas dans BV .

C.2.2 Exemple: la fonction de Heaviside

Soient W_L et W_R deux vecteurs distincts de \mathbb{R}^N , et Ψ une fonction vérifiant les hypothèses (C.3). On définit une suite de fonctions continues $(W_n)_n :]-1, 1[\rightarrow \mathbb{R}^N$ par :

$$W_n(x) = \begin{cases} W_L & \text{pour } x \in]-1, -1/n[\\ \Psi\left(\frac{nx+1}{2}\right) & \text{pour } x \in]-1/n, 1/n[\\ W_R & \text{pour } x \in]1/n, 1[\end{cases} \quad (\text{C.6})$$

On a clairement $W_n \xrightarrow{BR} W = W_L + H(x)(W_R - W_L)$ où $Y(x)$ désigne la fonction de Heaviside $Y(x) = \mathbf{1}_{x \geq 0}$. Il est alors possible de voir la fonction Ψ comme un chemin définissant le produit non-conservatif $A(W) \frac{\partial W}{\partial x}$. Néanmoins, ceci ne nous dit rien au sujet de la suite de produits $(A(W_n) \frac{\partial W_n}{\partial x})_n \dots$

C.2.3 Condition suffisante de convergence en loi

La suite précédente ne vérifie pas les hypothèses du théorème de stabilité. Néanmoins, il est possible de montrer "à la main", la convergence vague des produits non-conservatifs.

Proposition IV.10 *Sous les hypothèses (C.6), et si le chemin Φ est défini par $\Phi = \Psi$, alors on a le résultat de convergence dans la topologie faible $\sigma(C_0^1, \mathcal{M})$:*

$$A(W_n) \frac{\partial W_n}{\partial x} \xrightarrow{weak} \left[A(W) \frac{\partial W}{\partial x}\right]_{\Phi}$$

Si la définition du chemin est différente, le résultat de convergence est en général perdu.

Démonstration

Soit $\varphi \in C_0^1(]-1, 1[, \mathbb{R}^N)$, on a :

$$\begin{aligned} \int_{-1}^1 \varphi \left[A(W_n) \frac{\partial W_n}{\partial x}\right]_{\Phi} &= \int_{-1/n}^{1/n} \varphi(x) \left[A(W_n) \frac{\partial W_n}{\partial x}\right]_{\Phi}(x) dx \\ &= \int_{-1/n}^{1/n} \varphi(x) A(\Psi) \frac{\partial \Psi}{\partial x} \left(\frac{nx+1}{2}\right) dx \\ &= \int_0^1 \varphi\left(\frac{2x-1}{n}\right) A(\Psi(x)) \frac{\partial \Psi(x)}{\partial x} dx \end{aligned}$$

A la limite, le théorème de Lebesgue implique:

$$\int_{-1}^1 \varphi \left[A(W_n) \frac{\partial W_n}{\partial x} \right]_{\Phi} \rightarrow \varphi(0) \int_0^1 A(\Psi(s)) \Psi'(s) ds$$

Donc, si $\forall s \in [0, 1]$, on a $\Phi(s; U_L, U_R) = \Psi(s)$, on a:

$$\int_{-1}^1 \varphi \left[A(W) \frac{\partial W}{\partial x} \right]_{\Phi} = \lim_{n \rightarrow +\infty} \int_{-1}^1 \varphi \left[A(W_n) \frac{\partial W_n}{\partial x} \right]_{\Phi}$$

et le produit non-conservatif passe donc à la limite faible. \square

On peut bien sûr généraliser ce résultat à une fonction ayant m discontinuités régularisées par m profils visqueux différents (voir entre autres [24, 81] et le chapitre 3). Le chemin Φ est alors composé de m sous-chemins $(\Phi_i \circ \gamma_i)_{1 \leq i \leq m}$ de la forme:

$$\Phi(s; W_L, W_R) = \left(\begin{array}{c} \Phi_1(ms; W_L, W_1) \text{ pour } s \in [0, \frac{1}{m}] \\ \vdots \\ \underbrace{\Phi_i(ms - (i-1); W_{i-1}, W_i)}_{\gamma_i(s)} \text{ pour } s \in [\frac{i-1}{m}, \frac{i}{m}] \\ \vdots \\ \underbrace{\Phi_m((m-1)s + 1; W_{m-1}, W_R)}_{\gamma_m(s)} \text{ pour } s \in [1 - \frac{1}{m}, 1] \end{array} \right)$$

C.3 Application aux systèmes hyperboliques écrits sous forme non-conservative

C.3.1 Notion de solution faible

Ce formalisme va maintenant être appliqué aux systèmes hyperboliques non-conservatifs:

$$\begin{cases} \frac{\partial W}{\partial t} + A(W) \frac{\partial W}{\partial x} = 0 \\ W_0 \in BV(\mathbb{R})^N \end{cases} \quad (\text{C.7})$$

Définition IV.8 Soit Φ une famille de chemins dans Ω au sens de (C.3), un N -uplet W de fonctions de $L^\infty \cap BV([a, b] \times [0, T])$ est dit solution faible de (C.7) s'il vérifie l'égalité suivante

$$\frac{\partial W}{\partial t} + \left[A(W) \frac{\partial W}{\partial x} \right]_{\Phi} = 0$$

au sens des mesures boreliennes bornées de $\mathbb{R} \times \mathbb{R}^+$.

C.3.2 Conditions de saut généralisées

Ces solutions développent généralement des discontinuités au bout d'un temps fini. Il est donc nécessaire d'expliciter des relations de saut de part et d'autre de ces singularités. Soit $t \mapsto \Sigma(t)$ une courbe de $\mathbb{R} \times \mathbb{R}^+$ en dehors de laquelle W est C^1 -régulière, (x_0, t_0) un point de cette courbe, on désigne par W^\pm les limites gauche et droite en ce point:

$$W^\pm = \lim_{h \rightarrow 0, h \geq 0} W(x_0 \pm h, t_0)$$

Alors le théorème implique des relations de Rankine-Hugoniot généralisées pour une discontinuité de vitesse σ :

$$\int_0^1 (-\sigma \cdot Id + A(\Phi(s; W^-, W^+))) \frac{\partial \Phi}{\partial s}(s; W^-, W^+) ds = 0 \quad (\text{C.8})$$

C.3.3 Structure du problème de Riemann

Les auteurs de [21] retrouvent les résultats habituels de la théorie conservative à savoir que la solution du problème de Riemann pour (C.7) se compose de $(N + 1)$ états constants séparés par N ondes simples. Ces dernières sont d'une part des chocs ou des détente si le champ est vraiment non-linéaire (VNL), d'autre part des discontinuités de contact si le champ est linéairement dégénéré. La différence majeure vient du fait que tout ce qui concerne les sauts dépend explicitement de la donnée du chemin Φ .

C.3.4 Matrices linéarisées de type Roe

Dans [82], I. Touni propose une extension de la construction de Roe [73] à des systèmes écrits sous la forme (C.7).

Définition IV.9 Une matrice $A_\Phi(W_L, W_R) \in \mathbf{M}_N(\mathbb{R})$ est appelée linéarisée de Roe généralisée si elle est pourvue des propriétés suivantes:

- $A_\Phi(W_L, W_R)$ est diagonalisable et possède une base de N vecteurs propres associés à N valeurs propres réelles.
- $A_\Phi(W_L, W_R)$ est consistante avec (C.7) au sens suivant:

$$A_\Phi(W, W) = A(W)$$

- Pour tout couple (W_L, W_R) dans Ω , on a:

$$\int_0^1 A(\Phi(s; W_L, W_R)) \frac{\partial \Phi}{\partial s}(s; W_L, W_R) ds = A_\Phi(W_L, W_R)(W_R - W_L)$$

La dernière condition est bien sûr à rapprocher de l'expression des relations de Rankine-Hugoniot généralisées écrites précédemment.

Annexe D

Une alternative pour les termes sources linéaires en ρ

L'idée à partir de laquelle cette annexe est construite est très largement due à P. Cargo et A.Y. LeRoux [11]. Elle consiste à reformuler le vecteur de termes sources d'un système hyperbolique du type:

$$\begin{cases} U_t + F(U)_x = G(U) \text{ avec } U \in \mathbb{R}^N \\ U_0 \in [BV(\mathbb{R})]^N \end{cases} \quad (\text{D.1})$$

sous la forme de "faux produits non-conservatifs"; c'est-à-dire des termes du type $G(U) \cdot Q_x$, mais dont l'écriture est toujours pertinente. Dans le cas des équations d'Euler avec un second membre modélisant la gravité, il est même possible de dériver un système modifié homogène s'écrivant de façon totalement conservative [11]. Pour ce problème, la variable $Q(x, t)$ s'identifie à une pression hydrostatique. Néanmoins, cela arrive assez rarement, mais l'idée sous-jacente peut tout de même s'appliquer à un certain nombre de problèmes notamment *lorsque $G(U)$ ne comporte que des termes linéaires en ρ .*

D.1 Une formulation équivalente homogène

On décide d'exposer la méthode sur un exemple issu de la modélisation des écoulements diphasiques [81] au moyen du système:

$$\begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho c)_t + (\rho u c)_x = 0 \\ (\rho u)_t + (\rho u^2 + p(\rho, c))_x = -\rho h(u) \end{cases} \quad (\text{D.2})$$

Comme à l'accoutumée, ρ, u, c désignent respectivement la densité du mélange, la vitesse volumétrique et la concentration massique d'une des phases. Par ailleurs, on ne fera sur la pression $p(\rho, c)$ que les hypothèses explicitées à la proposition IV.5 (partie I) permettant de résoudre le problème de Riemann pour (D.2). On suppose en outre que la fonction $u \mapsto h(u)$ est au moins Lipschitz-continue. Le résultat fondamental est le suivant:

Lemme IV.8 (*Cargo, LeRoux*)

Il existe un potentiel scalaire $Q \in W_{loc}^{1,\infty}([0, T] \times \mathbb{R})$ vérifiant:

$$\begin{cases} Q_t = \rho u \text{ et } Q_x = -\rho \\ (\rho Q)_t + (\rho u Q)_x = 0 \end{cases} \quad (\text{D.3})$$

Démonstration

La première équation de (D.2) signifie que la divergence du vecteur $(\rho, \rho u)$ dans le demi-espace $\mathbb{R} \times \mathbb{R}^+$ est nulle. Conséquent, celui-ci est le rotationnel d'un potentiel scalaire noté Q vérifiant:

$$\rho = -Q_x \text{ et } \rho u = Q_t$$

On a donc: $Q(x, t) = \int_{-\infty}^x \rho(s, t) \cdot ds$ et d'après cette définition, on a la relation: $Q_t + u Q_x = 0$. En multipliant cette dernière par ρ et en ajoutant Q fois l'équation de conservation de la masse, on obtient la forme conservative annoncée dans (D.3). La régularité est une conséquence de la borne L^∞ sur ρ et ρu . \square

Il est donc opportun d'introduire un nouveau problème *apparamment* non-conservatif:

$$\begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho c)_t + (\rho u c)_x = 0 \\ (\rho u)_t + (\rho u^2 + p(\rho, c))_x + h(u) Q_x = 0 \\ (\rho Q)_t + (\rho u Q)_x = 0 \end{cases} \quad (\text{D.4})$$

Proposition IV.11 *Pour des solutions $(\rho, \rho u, Q, c) \in [BV([0, T] \times \mathbb{R})]^4$, le système suivant est équivalent à (D.4):*

$$\begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho u)_t + (\rho u^2 + p(\rho, c))_x + h(u) Q_x = 0 \\ Q_t + u Q_x = 0 \\ c_t + u c_x = 0 \end{cases} \quad (\text{D.5})$$

En particulier, les fonctions Q et c sont constantes à travers les ondes VNL et les produits $u \cdot c_x$, $u \cdot Q_x$ et $h(u) \cdot Q_x$ sont bien définis dans $\mathcal{D}'(\mathbb{R} \times \mathbb{R}^+)$.

Démonstration

Par analogie avec la proposition IV.5, on note $W = (\rho, u, Q, c)$ le vecteur des variables non-conservatives. On a alors un problème $W_t + A(W)W_x = 0$ non-strictement hyperbolique puisque:

$$A(W) = \begin{pmatrix} u & \rho & 0 & 0 \\ p_\rho/\rho & u & h(u) & p_c/\rho \\ 0 & 0 & u & 0 \\ 0 & 0 & 0 & u \end{pmatrix} \text{ de spectre } \Lambda = \{u \pm a, u\} \text{ avec } a = \sqrt{p_\rho}$$

Les vecteurs propres associés à $u \pm a$ s'écrivent:

$$\begin{pmatrix} \rho \\ -a \\ 0 \\ 0 \end{pmatrix} \text{ et } \begin{pmatrix} \rho \\ a \\ 0 \\ 0 \end{pmatrix}$$

Donc les champs associés sont VNL et les trois invariants de Riemann sont:

$$u \pm \int^{\rho} \frac{a}{s} . ds, Q, c$$

L'espace propre associé à la valeur propre double u est engendré par:

$$\begin{pmatrix} p_c \\ 0 \\ 0 \\ -p_\rho \end{pmatrix} \text{ et } \begin{pmatrix} \rho \cdot h(u) \\ 0 \\ -p_\rho \\ 0 \end{pmatrix}$$

Donc le champ associé est LD et u (mais pas p) est constant le long de cette discontinuité de contact. D'autre part, le long d'un choc de vitesse σ , les relations de Rankine-Hugoniot donnent:

$$\begin{cases} \sigma[\rho] = [\rho u] \\ \sigma[\rho c] = [\rho u c] \\ \sigma[\rho Q] = [\rho u Q] \\ \sigma[\rho u] = [\rho u^2 + p] + [h(u) \cdot Q_x]_\Phi \end{cases}$$

avec Φ un chemin régularisant du type [21] à préciser. Posons $v = u - \sigma \neq 0$ pour se placer dans le repère mobile du choc, il vient:

$$[\rho v] = 0 \text{ d'où } [c] = [Q] = 0$$

Au bout du compte, c et Q sont constantes à travers les ondes VNL, et u est constant au contact. Il n'y a donc aucune ambiguïté dans (D.5) et le choix du chemin Φ est libre. \square

Dans le cadre du problème de Riemann pour (D.5), il s'ensuit qu'il est possible de donner une formulation explicite pour $Q(x/t)$ et $c(x/t)$ au moyen de la fonction de Heaviside.

D.2 Application numérique

Bien que les systèmes (D.4), (D.5) possèdent d'intéressantes propriétés, notamment en ce qui concerne les termes non-conservatifs, on renonce à construire un schéma de Godunov comme il a été fait dans le cas conservatif de [11]. L'objet de cette section est de mettre en oeuvre un solveur linéarisé pour résoudre efficacement le système (D.2). Pour cela, on reprend les idées de [82] pour définir une matrice de type Roe dans un tel cadre "apparamment non-conservatif".

D.2.1 Linéarisée de type Roe

Pour exploiter au mieux les bonnes propriétés du système (D.5), il est d'abord nécessaire de définir un vecteur d'inconnues U et une fonction de flux $F(U)$ correctes. Le point important est que le couplage entre

la concentration et la densité dans l'expression de la pression oblige à garder la variable conservative ρc . On propose donc:

$$U = (\rho, \rho c, \rho u, Q) \text{ et } F(U) = (\rho u, \rho u c, \rho u^2 + p(\rho, c), 0)$$

Le système résolu par le schéma sera donc:

$$\begin{cases} \rho_t + (\rho u)_x = 0 \\ (\rho c)_t + (\rho u c)_x = 0 \\ (\rho u)_t + (\rho u^2 + p(\rho, c))_x + h(u)Q_x = 0 \\ Q_t + uQ_x = 0 \end{cases} \quad (\text{D.6})$$

D'après la section précédente, il est immédiat que le choix du chemin régularisant pour (D.6) est libre. On choisit donc le plus simple à savoir le chemin canonique "des lignes droites":

$$[0, 1] \ni s \mapsto \Phi_S(s; U_L, U_R) = U_L + s(U_R - U_L)$$

Soit $\bar{A}(U_L, U_R)$ une matrice de Roe pour (D.6), celle-ci a une structure du type:

$$\bar{A}(U_L, U_R) = \begin{pmatrix} & & 0 \\ & A(U_L, U_R) & \frac{0}{h(u)} \\ 0 & 0 & 0 \quad \bar{u} \end{pmatrix}$$

avec $A(U_L, U_R)$ une matrice de Roe (par exemple celle explicitée dans [81]) pour le problème (D.2) homogène. Le choix du chemin canonique Φ_S nous donne une expression directe des moyennes de Roe pour les "faux" termes non-conservatifs:

$$\begin{cases} \overline{h(u)} = \int_0^1 h(u_L + s(u_R - u_L)) ds = \frac{H(u_R) - H(u_L)}{u_R - u_L} \\ \bar{u} = \int_0^1 u_L + s(u_R - u_L) ds = \frac{1}{2}(u_L + u_R) \end{cases} \quad (\text{D.7})$$

avec $H(\cdot)$ désignant une primitive de $h(\cdot)$ (le cas $u_L = u_R$ est évident). On peut remarquer que l'équation sur Q étant traitée de façon non-conservative, un principe du maximum est assuré pour cette variable.

D.2.2 Etude d'un cas-test

On considère le système (D.4) posé sur l'intervalle $x \in [0, 1]$ avec un second membre du type $-\rho u|u|.k(x)$ où $k(x) = \mathbf{1}_{x=0.5}$. Celui-ci modélise l'action d'une grille de mélange disposée dans un conduit droit grossièrement maillé. Le test consiste à observer l'évolution vers un régime permanent à partir d'une donnée initiale correspondant à un écoulement uniforme subsonique. On a choisi:

$$\begin{cases} \rho_0 = 0.75, c_0 = 0.5, u_0 = 0.1 \\ p(\rho, c) = \left(\frac{1.6\rho c}{1.6 - \rho(1-c)} \right)^2 \end{cases}$$

Le maillage est $\Delta x = 0.02$ et $\Delta t = 0.01$, ce qui donne une CFL maximale de 0.9. Le schéma témoin est une approximation à pas fractionnaires sans décentrement pour le terme source. On a représenté la vitesse, la pression et le débit pour $x \in [0, 1]$ aux figures D.1, D.2, D.3. On constate clairement la meilleure qualité des approximations fournies par le présent schéma.

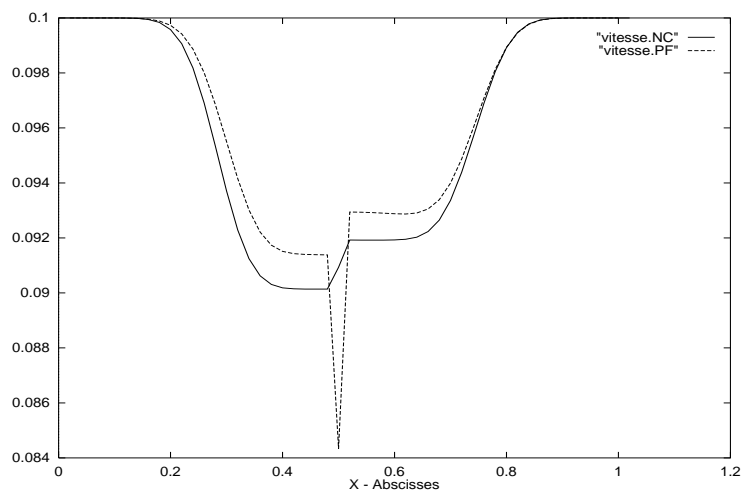


Figure D.1: Vitesse (schéma non-conservatif/pas fractionnaires)

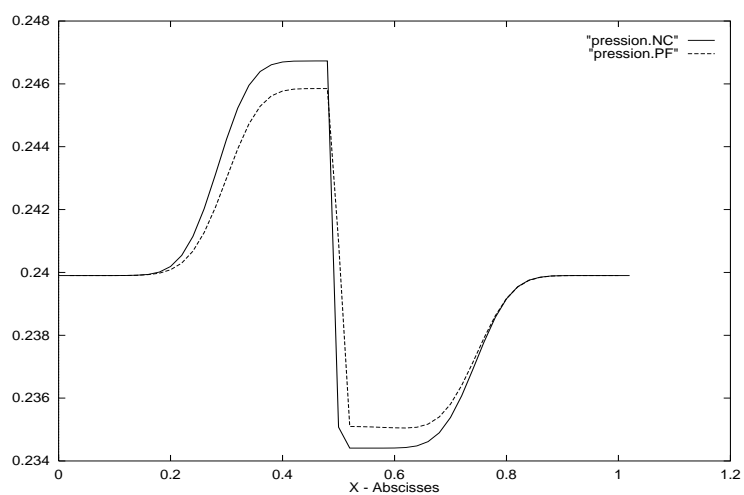


Figure D.2: Pression (schéma non-conservatif/pas fractionnaires)

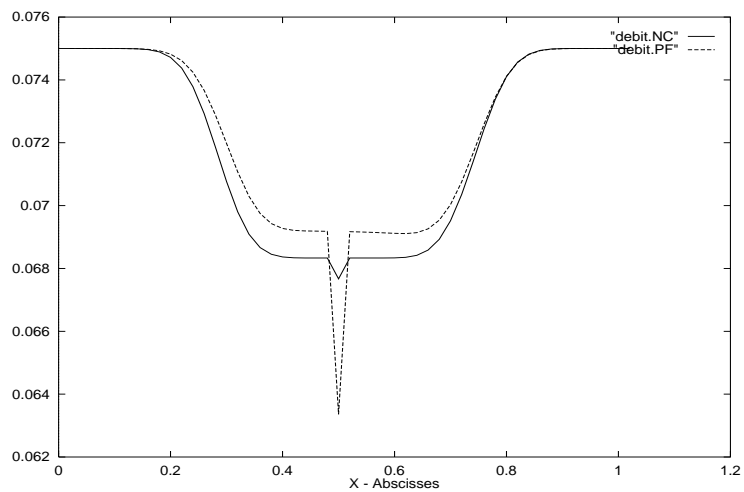


Figure D.3: Débit (schéma non-conservatif/pas fractionnaires)

Annexe E

Etude de l'approximation des relations de saut non-conservatives

On s'intéresse ici aux erreurs commises par les simplifications opérées dans la détermination des états à l'interface pour les schémas de la présente partie. Celles-ci proviennent de deux origines différentes:

- on remplace un système différentiel couplé par 2 équations découplées que l'on itère afin que l'éventuel point fixe vérifie le système d'origine.
- on remplace les variables de couplage par des constantes, plus précisément par les valeurs des itérées à l'extrémité de l'intervalle $[0, \Delta x]$.

Pour fixer les idées, on étudie les deux problèmes modèles:

$$\begin{cases} u_x = f(u, v) \\ v_x = g(u, v) \end{cases} \quad \text{et} \quad \begin{cases} u_x = f(u, v) \\ h(u, v) = 0 \end{cases} \quad \text{pour } x \in [0, \Delta x]$$

pour lesquels on fait les approximations:

$$\begin{cases} u_x^{n+1} = f(u^{n+1}, v^n(\Delta x)) \\ v_x^{n+1} = g(u^{n+1}(\Delta x), v^{n+1}) \end{cases} \quad \text{et} \quad \begin{cases} u_x^{n+1} = f(u^{n+1}, v^{n+1}(\Delta x)) \\ h(u^{n+1}, v^{n+1})(\Delta x) = 0 \end{cases}$$

Pour terminer, on fera un usage important d'un lemme classique dans la théorie des équations différentielles:

Lemme IV.9 (*Gronwall "fort"*)

Soit $\phi \in C^0([0, T])$ positive telle qu'il existe $(a, b) \in (\mathbb{R}_*^+)^2$ vérifiant:

$$\forall t \in [0, T], \phi(t) \leq a \int_0^t \phi(s) ds + bt$$

Alors on a la majoration:

$$\phi(t) \leq \frac{b}{a} (e^{at} - 1)$$

E.1 Analyse du système différentiel 2×2

Dans cette section, on s'intéresse au problème de Cauchy suivant pour $x \in [0, \Delta x]$:

$$\begin{cases} u_x = f(u, v) \\ v_x = g(u, v) \end{cases} \quad \text{avec} \quad \begin{cases} u(0) = u_0 \\ v(0) = v_0 \end{cases} \quad (\text{E.1})$$

On décide de l'approcher par la suite $(u^n, v^n)_{n \geq 0}$ définie par:

$$\begin{cases} u_x^{n+1} = f(u^{n+1}, v^n(\Delta x)) \\ v_x^{n+1} = g(u^{n+1}(\Delta x), v^{n+1}) \end{cases} \quad \text{avec} \quad \begin{cases} u^{n+1}(0) = u_0 \text{ et } u^0(x) \equiv u_0 \\ v^{n+1}(0) = v_0 \text{ et } v^0(x) \equiv v_0 \end{cases} \quad (\text{E.2})$$

E.1.1 Convergence de la suite d'approximations

Par analogie avec les résultats de la partie I, on fait des hypothèses sur les fonctions f et g à savoir:

-Régularité: généralement, ces fonctions sont polynômiales, il est donc justifié d'exiger que:

$$(f, g) \in C^1(\mathbb{R}^2, \mathbb{R})^2 \quad (\text{E.3})$$

-Borne uniforme en amplitude: il est commun, dans le cadre des problèmes hyperboliques, d'assurer une estimation L^∞ . Pour cela, il est suffisant que:

$$\begin{aligned} \forall v \in \mathbb{R}, u. f(u, v) < 0 \text{ si } |u| > M \\ \forall u \in \mathbb{R}, v. g(u, v) < 0 \text{ si } |v| > M \end{aligned} \quad (\text{E.4})$$

Ceci permet d'avoir un premier résultat.

Lemme IV.10 *Sous les hypothèses (E.3) et (E.4), la suite $(u^n, v^n)_{n \geq 0}$ définie par (E.2) est convergente dans $C^1([0, \Delta x], \mathbb{R})$ pour $\Delta x > 0$ suffisamment petit.*

Démonstration

Il s'agit de vérifier que cette suite est de Cauchy pour la norme de la convergence uniforme de la fonction et de sa dérivée. Comme C^1 est un Banach pour cette norme, on aura le résultat. On suppose pour simplifier que $|u_0| < M$ et $|v_0| < M$; d'après (E.4), la suite est confinée dans le compact $[-M, M]^2$. Il est donc naturel de définir:

$$|f_u|_\infty = \max_{(u,v) \in [-M,M]^2} \left| \frac{\partial f}{\partial u}(u,v) \right| \quad (\text{E.5})$$

Cette quantité existe puisque par régularité C^1 , la dérivée partielle est continue donc atteint son maximum sur tout compact. On définit de même $|f_v|_\infty, |g_u|_\infty, |g_v|_\infty$. On note les écarts:

$$E_u^{n+1}(x) = |u^{n+1}(x) - u^n(x)| \quad \text{et} \quad E_v^{n+1}(x) = |v^{n+1}(x) - v^n(x)|$$

Comme les fonctions f et g vont de $\mathbb{R}^2 \rightarrow \mathbb{R}$, la règle du point milieu s'applique:

$$f(u, v) - f(u_0, v_0) = f_u(\bar{u}, \bar{v}) \cdot (u - u_0) + f_v(\bar{u}, \bar{v}) \cdot (v - v_0)$$

Par conséquent, on majore le module par:

$$|f(u, v) - f(u_0, v_0)| \leq |f_u|_\infty \cdot |u - u_0| + |f_v|_\infty \cdot |v - v_0|$$

Par conséquent, en intégrant (E.2) entre 0 et x , il est immédiat que:

$$\begin{aligned} E_u^{n+1}(x) &\leq \int_0^x |f(u^{n+1}(s), v^n(\Delta x)) - f(u^n(s), v^{n-1}(\Delta x))| ds \\ &\leq |f_u|_\infty \int_0^x |u^{n+1}(s) - u^n(s)| ds + |f_v|_\infty \int_0^x |v^n(\Delta x) - v^{n-1}(\Delta x)| ds \\ &\leq |f_u|_\infty \int_0^x E_u^{n+1}(s) ds + |f_v|_\infty E_v^n(\Delta x) \cdot x \end{aligned}$$

On est alors en position d'appliquer le lemme de Gronwall:

$$E_u^{n+1}(x) \leq \frac{|f_v|_\infty}{|f_u|_\infty} (\mathbf{e}^{|f_u|_\infty x} - 1) E_v^n(\Delta x) \quad (\text{E.6})$$

De même:

$$\begin{aligned} E_v^{n+1}(x) &\leq \int_0^x |g(u^{n+1}(\Delta x), v^{n+1}(s)) - g(u^n(\Delta x), v^n(s))| ds \\ &\leq |g_u|_\infty \int_0^x |u^{n+1}(\Delta x) - u^n(\Delta x)| ds + |g_v|_\infty \int_0^x |v^{n+1}(s) - v^n(s)| ds \\ &\leq |g_v|_\infty \int_0^x E_v^{n+1}(s) ds + |g_u|_\infty E_u^{n+1}(\Delta x) \cdot x \end{aligned}$$

Ce qui donne:

$$E_v^{n+1}(x) \leq \frac{|g_u|_\infty}{|g_v|_\infty} (\mathbf{e}^{|g_v|_\infty x} - 1) E_u^{n+1}(\Delta x)$$

Comme (E.6) est une fonction croissante de x , il vient que:

$$E_v^{n+1}(x) \leq \frac{|f_v|_\infty |g_u|_\infty}{|f_u|_\infty |g_v|_\infty} (\mathbf{e}^{|f_u|_\infty \Delta x} - 1) (\mathbf{e}^{|g_v|_\infty x} - 1) E_v^n(\Delta x)$$

Par le même argument, il vient la majoration uniforme en x :

$$E_v^{n+1}(x) \leq \left[\frac{|f_v|_\infty |g_u|_\infty}{|f_u|_\infty |g_v|_\infty} (\mathbf{e}^{|f_u|_\infty \Delta x} - 1) (\mathbf{e}^{|g_v|_\infty \Delta x} - 1) \right]^n E_v^1(\Delta x)$$

La suite géométrique $(E_v^n)_n$ tend vers 0 si sa raison est de module strictement inférieur à 1. Cela impose:

$$\frac{|f_v|_\infty |g_u|_\infty}{|f_u|_\infty |g_v|_\infty} (\mathbf{e}^{|f_u|_\infty \Delta x} - 1) (\mathbf{e}^{|g_v|_\infty \Delta x} - 1) < 1 \quad (\text{E.7})$$

En développant en Δx , cela nous donne:

$$|f_v|_\infty |g_u|_\infty \Delta x^2 < 1$$

Sous cette condition, il vient que $\sup_{x \in [0, \Delta x]} E_v^n(x) \rightarrow 0$, d'où $\sup_{x \in [0, \Delta x]} E_u^n(x) \rightarrow 0$ par (E.6). Par régularité de f et g , ceci entraîne que la suite des dérivées $(u_x^n, v_x^n)_n$ est aussi de Cauchy. Finalement, comme C^1 est complet pour cette norme, on a le résultat annoncé. \square

E.1.2 Estimation de l'erreur à la limite

A convergence, il existe donc \bar{u}, \bar{v} dans $C^1([0, \Delta x], \mathbb{R})$ vérifiant le système différentiel:

$$\begin{cases} \bar{u}_x = f(\bar{u}, \bar{v}(\Delta x)) \\ \bar{v}_x = g(\bar{u}(\Delta x), \bar{v}) \end{cases} \quad \text{avec} \quad \begin{cases} \bar{u}(0) = u_0 \\ \bar{v}(0) = v_0 \end{cases} \quad (\text{E.8})$$

Comme (E.8) est différent de (E.1), il est clair que l'on introduit une erreur dans la détermination des états aux interfaces. Pour ne pas perdre la consistance du schéma global, il est donc nécessaire d'estimer celle-ci en fonction de Δx .

Proposition IV.12 *Sous les hypothèses (E.3), (E.4) et (E.7), on a l'estimation:*

$$|\bar{u}(\Delta x) - u(\Delta x)| + |\bar{v}(\Delta x) - v(\Delta x)| \leq o(\Delta x^2) \quad (\text{E.9})$$

Démonstration

En intégrant (E.8) entre 0 et Δx , il vient:

$$\begin{cases} |\bar{u}(\Delta x) - u(\Delta x)| \leq \int_0^{\Delta x} |f(\bar{u}(s), \bar{v}(\Delta x)) - f(u(s), v(s))| ds \\ |\bar{v}(\Delta x) - v(\Delta x)| \leq \int_0^{\Delta x} |g(\bar{u}(\Delta x), \bar{v}(s)) - g(u(s), v(s))| ds \end{cases}$$

J'appelle $e_u(\Delta x)$ et $e_v(\Delta x)$ les erreurs en u et v définies ci-dessus. Du fait de l'hypothèse (E.4) et de la majoration des dérivées du type (E.5), on a:

$$e_u(\Delta x) \leq |f_u|_\infty \int_0^{\Delta x} e_u(s) ds + |f_v|_\infty \int_0^{\Delta x} |\bar{v}(\Delta x) - v(s)| ds$$

On fait la majoration:

$$\begin{aligned} |\bar{v}(\Delta x) - v(s)| &\leq |\bar{v}(\Delta x) - \bar{v}(s)| + |\bar{v}(s) - v(s)| \\ &\leq \int_s^{\Delta x} |g(\bar{u}(\Delta x), \bar{v}(t))| dt + e_v(s) \\ &\leq \Delta x \max_{s \in [0, \Delta x]} |g(\bar{u}, \bar{v})(s)| + e_v(s) \end{aligned}$$

D'où:

$$\begin{cases} e_u(\Delta x) \leq |f_u|_\infty \int_0^{\Delta x} e_u(s) ds + |f_v|_\infty \Delta x \max_{s \in [0, \Delta x]} e_v(s) + |f_v|_\infty \max_{s \in [0, \Delta x]} |g(\bar{u}, \bar{v})(s)| \frac{\Delta x^2}{2} \\ e_v(\Delta x) \leq |g_v|_\infty \int_0^{\Delta x} e_v(s) ds + |g_u|_\infty \Delta x \max_{s \in [0, \Delta x]} e_u(s) + |g_u|_\infty \max_{s \in [0, \Delta x]} |f(\bar{u}, \bar{v})(s)| \frac{\Delta x^2}{2} \end{cases}$$

Comme $e_u(0) = e_v(0) = 0$, on déduit que:

$$\begin{cases} e_u(\Delta x) \leq |f_v|_\infty \Delta x \max_{s \in [0, \Delta x]} e_v(s) + o(\Delta x^2) \\ e_v(\Delta x) \leq |g_u|_\infty \Delta x \max_{s \in [0, \Delta x]} e_u(s) + o(\Delta x^2) \end{cases}$$

Si on admet que Δx est suffisamment petit pour que les fonctions $x \mapsto e_u(x)$ et $x \mapsto e_v(x)$ soient croissantes, il vient que:

$$\begin{cases} e_u(\Delta x) \leq |f_v|_\infty |g_u|_\infty \Delta x^2 e_u(\Delta x) + o(\Delta x^2) \\ e_v(\Delta x) \leq |f_v|_\infty |g_u|_\infty \Delta x^2 e_v(\Delta x) + o(\Delta x^2) \end{cases}$$

Et finalement, on a le résultat:

$$e_u(\Delta x) + e_v(\Delta x) \leq \frac{o(\Delta x^2)}{1 - |f_v|_\infty |g_u|_\infty \Delta x^2}$$

□

E.2 Analyse du problème scalaire sur une variété

Il s'agit maintenant d'étudier un problème du type:

$$\begin{cases} u_x = f(u, v) \\ h_{u_0, v_0}(u, v) = 0 \end{cases} \quad \text{avec} \quad \begin{cases} u(0) = u_0 \\ v(0) = v_0 \end{cases} \quad (\text{E.10})$$

La fonction h_{u_0, v_0} est supposée localement inversible pour lui appliquer le théorème des fonctions implicites et se ramener à:

$$\begin{cases} u_x = f(u, v) \\ v = g_{u_0, v_0}(u) \end{cases} \text{ avec } \begin{cases} u(0) = u_0 \\ v(0) = v_0 \end{cases} \quad (\text{E.11})$$

Le schéma itératif approche (E.11) sous la forme d'une suite $(u^n, v^n)_{n \geq 0}$ de fonctions $C^1([0, \Delta x], \mathbb{R})$ définie par:

$$\begin{cases} u_x^{n+1} = f(u^{n+1}, v^n(\Delta x)) \\ v^{n+1} = g_{u_0, v_0}(u^{n+1}) \end{cases} \text{ avec } \begin{cases} u^{n+1}(0) = u_0 \text{ et } u^0(x) \equiv u_0 \\ v^{n+1}(0) = v_0 \text{ et } v^0(x) \equiv v_0 \end{cases} \quad (\text{E.12})$$

E.2.1 Convergence de la suite d'approximations

Comme dans la section précédente, il faut faire des hypothèses sur les fonctions f et g_{u_0, v_0} (dorénavant notée simplement g):

-Régularité: on demande plus que la Lipschitz-continuité:

$$(f, g) \in C^1(\mathbb{R}^2)^2 \quad (\text{E.13})$$

-Estimation uniforme en amplitude: on veut s'assurer que les solutions des deux problèmes (E.11) et (E.12) sont confinées dans un compact de \mathbb{R}^2 :

$$\forall v \in \mathbb{R}, u.f(u, v) < 0 \text{ si } |u| > M \quad (\text{E.14})$$

Ceci assure que $u \in [-M, M]$ dès que $u_0 \in [-M, M]$ (si l'hypothèse sur u_0 n'est pas satisfaite, on remplace M par $|u_0|$ et on garde une majoration uniforme). Par conséquent, la régularité C^1 de g impose que $v \in [-N, N]$ et on a un premier résultat:

Lemme IV.11 *Sous les hypothèses (E.13) et (E.14), la suite $(u^n)_{n \geq 0}$ est convergente dans $C^1([0, \Delta x], \mathbb{R})$ pour $\Delta x > 0$ suffisamment petit.*

Ceci entraîne la convergence de $(v^n)_{n \geq 0}$ par régularité de g .

Démonstration

L'expression (E.12) entraîne que:

$$u_x^{n+1} = f(u^{n+1}, g(u^n)(\Delta x)) \quad (\text{E.15})$$

On intègre entre 0 et x et on pose $E^{n+1}(x) = |u^{n+1}(x) - u^n(x)|$ l'écart au point x . On a l'estimation:

$$E^{n+1}(x) \leq \int_0^x |f(u^{n+1}(s), g(u^n)(\Delta x)) - f(u^n(s), g(u^{n-1})(\Delta x))| ds$$

Comme précédemment, on introduit:

$$|f_u|_\infty = \max_{(u, v) \in [-M, M] \times [-N, N]} \left| \frac{\partial f}{\partial u}(u, v) \right| \quad (\text{E.16})$$

De même pour $|f_v|_\infty$ et $|g_u|_\infty$. Alors on a:

$$E^{n+1}(x) \leq \int_0^x E^{n+1}(s) ds + |g_u|_\infty |f_v|_\infty x E^n(\Delta x)$$

Par l'inégalité de Gronwall, il vient:

$$E^{n+1}(x) \leq \frac{|g_u|_\infty |f_v|_\infty}{|f_u|_\infty} (e^{|f_u|_\infty x} - 1) E^n(\Delta x)$$

Comme cette fonction est croissante en $x \in [0, \Delta x]$, on trouve:

$$E^{n+1}(x) \leq E^{n+1}(\Delta x) \leq \frac{|g_u|_\infty |f_v|_\infty}{|f_u|_\infty} (e^{|f_u|_\infty \Delta x} - 1) E^n(\Delta x)$$

Et donc la suite des écarts tend vers 0 si on a :

$$\frac{|g_u|_\infty |f_v|_\infty}{|f_u|_\infty} (\mathbf{e}^{|f_u|_\infty \Delta x} - 1) < 1 \quad (\text{E.17})$$

A l'ordre 1 en Δx , cela donne :

$$|g_u|_\infty |f_v|_\infty \Delta x < 1$$

Par la forme de (E.15), ceci entraîne l'accumulation de la suite des dérivées $(u_x^n)_{n \geq 0}$. La suite $(u^n)_{n \geq 0}$ est donc de Cauchy dans $C^1([0, \Delta x], \mathbb{R})$ sous la condition (E.17). Comme cet espace est un Banach pour cette norme, on a le résultat de convergence escompté. \square

E.2.2 Estimation de l'erreur à la limite

A la limite $n \rightarrow +\infty$, il existe donc \bar{u}, \bar{v} dans $C^1([0, \Delta x], \mathbb{R})$ vérifiant le système :

$$\begin{cases} \bar{u}_x = f(\bar{u}, \bar{v}(\Delta x)) \\ \bar{v} = g(\bar{u}(\Delta x)) \end{cases} \quad \text{avec} \quad \begin{cases} \bar{u}(0) = u_0 \\ \bar{v}(0) = v_0 \end{cases} \quad (\text{E.18})$$

Comme (E.18) est différent de (E.11), il est clair que l'on introduit une erreur dans la détermination des états aux interfaces. Pour ne pas perdre la consistance du schéma global, il est donc nécessaire d'estimer celle-ci en fonction de Δx .

Proposition IV.13 *Sous les hypothèses (E.13), (E.14) et (E.17), on a l'estimation :*

$$|\bar{u}(\Delta x) - u(\Delta x)| + |\bar{v}(\Delta x) - v(\Delta x)| \leq o(\Delta x^2) \quad (\text{E.19})$$

Démonstration

Pour évaluer l'erreur $e(x) = |\bar{u}(x) - u(x)|$, il suffit d'intégrer entre 0 et x l'expression (E.15) :

$$\begin{aligned} e(x) &\leq \int_0^x |f(\bar{u}(s), g(\bar{v}(\Delta x))) - f(u(s), g(u)(s))| ds \\ &\leq |f_u|_\infty \int_0^x e(s) ds + |f_v|_\infty |g_u|_\infty \int_0^x |\bar{u}(\Delta x) - \bar{u}(s)| + |\bar{u}(s) - u(s)| ds \\ &\leq (|f_u|_\infty + |f_v|_\infty |g_u|_\infty) \int_0^x e(s) ds + |f_v|_\infty |g_u|_\infty \int_0^x \int_s^x |f(\bar{u}(t), \bar{v}(\Delta x))| dt \end{aligned}$$

On obtient finalement que :

$$e(x) \leq (|f_u|_\infty + |f_v|_\infty |g_u|_\infty) \int_0^x e(s) ds + |f_v|_\infty |g_u|_\infty \max_{s \in [0, \Delta x]} |f(\bar{u}(s), \bar{v}(\Delta x))| \frac{x^2}{2}$$

Comme $e(0) = 0$, on déduit que $e(x) \leq o(x^2)$.

L'estimation sur \bar{v} est une conséquence immédiate de la régularité de g . \square

Bibliographie

- [1] Albrecht P. - The Runge-Kutta theory in a nutshell - SIAM J. Numer. Anal. Vol.33 No5 (1996) pp1712/1735
- [2] Alinhac S. - Blowup for nonlinear hyperbolic equations - P.N.L.D.E. Birkhäuser Verlag (1995)
- [3] Allaire G. Bec X. - Schéma d'ordre 2 en temps pour le code FLICA-4 - Note interne CEA DMT 96/395 SERMA/LETR 96/1943
- [4] Bardos C. - Introduction aux problèmes hyperboliques non-linéaires - Lecture notes in math. 1047 (Springer Verlag) pp2/74
- [5] Bardos C., LeRoux A.Y., Nedelec J.C. - First order quasilinear equations with boundary conditions - Comm. part. diff. eqns - No 4 (1979) pp 1017/1034
- [6] Ben-Artzi M. Falconvitz J. - An upwind 2nd order scheme for compressible duct flows - SIAM J. of Sci. Stat. Comp. Vol 7 No 3 (1986) pp744/768
- [7] Bereux F. Bonnetier E. LeFloch P. - Gas dynamics system: two special cases - Rapport interne CMAP No 325 (Ecole Polytechnique 1995)
- [8] Bereux F. Sainsaulieu L. - A Roe-type Riemann solver for hyperbolic systems with relaxation based on time-dependant wave decomposition - Mémoire d'habilitation à diriger des recherches (Ecole Polytechnique 1995)
- [9] Blunt M. Rubin B. - Implicit flux limiting schemes for petroleum reservoir simulation - J.C.P. 102 (1992) pp194/210
- [10] Brézis H. - Analyse fonctionnelle: théorie et applications - Maitrise de mathématiques Masson (1983)
- [11] Cargo P. LeRoux A.Y. - Un schéma équilibre adapté au modèle d'atmosphère avec termes de gravité - C.R. Acad. Sc. Série A t. 318 (1994) pp 73/76
- [12] Cauret J.J. Colombeau J.F. LeRoux A.Y. - Discontinuous generalized solutions of nonlinear nonconservative hyperbolic equations - J. Math. Anal. Applic. No139 (1989) pp552/573
- [13] Chen G.Q. Glimm J. - Global solutions to the compressible Euler equations with geometrical structure - Comm. Math. Phys. vol.180 No1 (1996) pp159/178
- [14] Chen G.Q. Levermore C.D. Liu T.P. - Hyperbolic conservation laws with stiff relaxation terms and entropy - Comm. Pure and Applied Math. vol.47 (1994) pp787/830
- [15] Chen G.Q. Liu T.P. - Zero relaxation and dissipation limits for hyperbolic systems of conservation laws - Comm. Pure and Applied Math. vol.46 (1993) pp755/781
- [16] Ciarlet P.G. - Introduction à l'analyse numérique matricielle et à l'optimisation - Maitrise de mathématiques Masson (1980)
- [17] Clerc S. - Décomposition de domaine pour un système hyperbolique: applications aux schémas implicites en mécanique des fluides - Note interne CEA DMT 96/284 SERMA/LETR 96/1915
- [18] Clerc S. - Etude de schémas décentrés implicites pour le calcul numérique en mécanique des fluides. Résolution par décomposition de domaines - Thèse de l'université Paris VI - 1997 (non publiée)
- [19] Crouzeix M. Mignot A.L. - Analyse numérique des équations différentielles - Maitrise de mathématiques Masson (2^{ème} édition 1992)
- [20] Dacorogna B. - Weak continuity and weak lower-continuity of non-linear functionals - Lecture notes in math. 922 (Springer Verlag)

- [21] Dal Maso G. LeFloch P. Murat F. - Definition and weak stability of a non-conservative product - J. Math. Pures et Appliquées - Vol. 74 No 6 (1995) pp483/548
- [22] DeSanti A.J. - Boundary and interior layer behavior of solutions of some singularly perturbed semilinear elliptic boundary value problems - J. Math. pures et appliquées - Vol. 65 (1986) pp 423/468
- [23] DeVore R.A. Lucier B.J. - On the size and smoothness of solutions to nonlinear hyperbolic conservation laws - SIAM J. of Math. Anal. Vol 27 No 3 (1996) pp684/707
- [24] DeVuyst F. - Schémas non-conservatifs et cinétiques pour la simulation numérique d'écoulements hypersoniques non-visqueux en déséquilibre thermo-chimique - Thèse de doctorat de l'Université Paris VI - 1994 (non publiée)
- [25] DiPerna R. - Convergence of approximate solutions to conservation laws - Arch. rational mech. and anal. 82 (1983) pp27/70
- [26] DiPerna R. - Measure-valued solutions to conservation laws - Arch. rational mech. and anal. 88 (1985) pp223/270
- [27] Dubois F. - Conditions aux limites fortement non-linéaires pour les équations d'Euler - Ecole CEA, EDF, INRIA sur les problèmes non-linéaires appliqués (P.L. Lions org. 29/9-2/12 1988)
- [28] Dubois F. LeFloch P. - Boundary conditions for nonlinear hyperbolic systems of conservation laws - J. of Diff. Eqs. vol.71 (1988) pp93/122
- [29] E Weinan - Homogenization of scalar conservation laws with oscillatory forcing terms - SIAM J. of Appl. Math. Vol.52 No4 (1992) pp959/972
- [30] Embid P. Goodman J. Majda A. - Multiple steady-states for 1D transonic flow - SIAM J. of Sci. Stat. Comp. Vol 5 No 1 (1984) pp21/41
- [31] Fan H. Hale J. - Large-time behavior in inhomogeneous conservation laws - Arch. rational mech. and anal. 125 (1993) pp 201/216
- [32] Glaister P. - Flux difference splitting for hyperbolic systems of conservation laws with source terms - Comp. Math. Applic. vol.26 No7 (1993) pp79/96
- [33] Glaz H.M. Liu T.P. - The asymptotic analysis of wave interactions and numerical calculations of transonic nozzle flows - Adv. in Applied Math. No 5 (1984) pp111/146
- [34] Glimm J. Marshall G. Plohr - A generalized Riemann problem for quasi one dimensional gas flows - Adv. in Applied Math. No 5 (1984) pp1/30
- [35] Godlewski E. Raviart P.A. - Hyperbolic systems of conservation laws (Ellipses-SMAI 1991)
- [36] Godunov S.K. - Finite difference schemes for numerical computation of solutions of the equations of fluid dynamics - Math. USSR Sbornik - Vol 47 (1959) pp 271/306
- [37] Gosse L. - Systèmes hyperboliques non homogènes: introduction de solveurs de Riemann à l'équilibre - Note interne CEA DMT 96/260 SERMA/LETR 96/1908
- [38] Gosse L. - Stabilité des méthodes de Runge-Kutta pour les EDP hyperboliques non-homogènes - Note interne CEA DMT 96/525 SERMA/LETR 96/1973
- [39] Gosse L. LeRoux A.Y. - Un schéma-équilibre adapté aux lois de conservation scalaires non-homogènes - C.R. Acad. Sc. Série A - t.323 (1996) pp543/546
- [40] Greenberg J. LeRoux A.Y. - A well balanced scheme for the numerical processing of source terms in hyperbolic equations - SIAM J. Numer. Anal. - Vol. 33 No 1 (1996) pp1/16
- [41] Greenberg J. LeRoux A.Y. Baraille R. Noussair A. - Analyse et approximations de lois de conservation avec terme source - C.R. Acad. Sc. Série A - t.321 (1995) pp1073/1076
- [42] Hairer E. Norsett S.P. Wanner G. - Solving ordinary differential equations I nonstiff problems - Springer series in computational mathematics (S.C.M. 8)
- [43] Hairer E. Wanner G. - Solving ordinary differential equations II stiff problems - Springer series in computational mathematics (S.C.M. 14)
- [44] Heibig A. Colombeau J.F. - Non-conservative products in bounded variation functions - SIAM J. Math. Anal. Vol. 23 No4 (1992) pp941/949
- [45] Holger H. Risebro N. - Riemann problems with a kink - Preprint (1996)

- [46] Holger H. Risebro N. Tveito A. - Maximum principles for a class of conservation laws - SIAM J. of Appl. Math. Vol.55 No3 (1995) pp651/661
- [47] Hsiao L. Liu T.P. - Convergence to nonlinear diffusion waves for solutions of a system of hyperbolic conservation laws with damping - Comm. Math. Phys. 143 (1992) pp599/605
- [48] Isaacson E. Temple B. - Nonlinear resonance for systems of conservation laws - SIAM J. Appl. Math. - Vol. 52 No 5 (1992) pp1260/1278
- [49] Isaacson E. Temple B. - Convergence of the 2x2 Godunov method for a general resonant nonlinear balance law - SIAM J. Appl. Math. - Vol. 55 No 3 (1995) pp625/640
- [50] Jin S. - Implicit schemes for hyperbolic conservation laws with stiff relaxation terms - J.C.P. Vol.122 No51 (1995)
- [51] Jin S. Levermore C.D. - Numerical schemes for hyperbolic conservation laws with stiff relaxation terms - J.C.P. No 126 (1996) pp449/467
- [52] Jin S. Liu J.G. - The effects of numerical viscosities (I) Slowly moving shocks - J.C.P. No 126 (1996) pp373/389
- [53] Kruzkov S.N. - First order quasilinear equations in several independent space variables - Math. USSR Sbornik - No 10 (1970) pp 217/243
- [54] Larroutourou B. - How to preserve the mass fractions positivity when computing multi-component flows - J.C.P. vol.95 (1991) pp59/84
- [55] Larroutourou B. Fezoui L. - On the equations of multi-component perfect or real gas inviscid flow - in Nonlinear hyperbolic problems, Lecture notes in math. vol.1402 (Springer Verlag 1989)
- [56] Lascaux P. Théodor R. - Analyse matricielle appliquée à l'art de l'ingénieur I Méthodes directes - Masson (2^{eme} édition 1993)
- [57] Lax P.D. - Hyperbolic systems of conservation laws and the mathematical theory of shock waves - Conference Board in Mathematical Sciences / National Science Foundation Vol.11
- [58] Lax P. Wendroff B. - Systems of conservation laws - Comm. Pure and Applied Math. No7 (1954) pp159/193
- [59] LeFloch P.G. - Thèse de doctorat de l'Ecole Polytechnique - 1987 (non publiée)
- [60] Lerat A. - Implicit time dependant methods for the solution of the Euler equations (Lecture 04 - VonKarman Institute - 1984)
- [61] LeRoux A.Y. - Convergence of the Godunov scheme for first order quasilinear equations - Proc. Japan Acad. - No 52 (1976) pp 488/491
- [62] LeVeque R.J. - Numerical methods for conservation laws: stiff source terms (Lecture 04 - VonKarman Institute - 1988)
- [63] Lions P.L. Perthame B. Tadmor E. - Formulation cinétique de lois de conservation scalaires multidimensionnelles - C.R. Acad. Sc. Série A - t.312 (1991) pp97/102
- [64] Ta-Tsien L. - Global classical solutions for hyperbolic systems (Coll. RMA 24 - Masson 1994)
- [65] Liu T.P. - Quasilinear hyperbolic systems - Comm. Math. Phys. 68 (1979) pp141/172
- [66] Liu T.P. - Hyperbolic conservation laws with relaxation - Comm. Math. Phys. 108 (1987) pp153/175
- [67] Mascia C. - Travelling wave solutions for a balance law - Preprint 1996
- [68] Mehlman G. - Thèse de doctorat de l'Ecole Polytechnique (Palaiseau France - 1991)
- [69] Murat F. - A survey on compensated compactness, in Contributions to modern calculus of variations - L. Cesari Eds. Research notes in math. (No 148 pp145/183 - 1987)
- [70] Otto F. - Initial-boundary value problem for a scalar conservation law - C.R. Acad. Sc. t.322 Série I (1996) pp729/734
- [71] Perthame B. Shu C.W. - On positivity preserving finite volume schemes for the Euler equations - Num. Math. No73 (1996) pp119/130
- [72] Richardson J.L. Ferrell R.C. - Unconditionally stable explicit algorithms for nonlinear fluid dynamics problems - J.C.P. 104 (1993) pp69/74
- [73] Roe P.L. - Approximate Riemann solvers, parameter vectors, and difference schemes - J.C.P. 43 (1981) pp357/372

-
- [74] Rosenbrock H.H. - Some general implicit processes for the numerical solutions of differential equations - Comput. J. No5 (1963) pp329/330
- [75] Rudin W. - Analyse réelle et complexe - Masson (1982)
- [76] Serre D. - Systèmes de lois de conservation - Coll. Fondations Diderot (1996)
- [77] Shu C.W. - A numerical method for systems of conservation laws of mixed type admitting hyperbolic flux splitting - J.C.P. 100 (1992) pp424/429
- [78] Shu C.W. Osher S. - Efficient implementation of essentially non-oscillatory shock-capturing schemes - J.C.P. 77 (1988) pp439/471
- [79] Smoller J. - Shock waves and reaction-diffusion equations (258 - Springer Verlag 1982)
- [80] Tartar L. - Compensated compactness and applications to partial differential equations, in Nonlinear analysis and mechanics: Heriot-Watt symposium vol. 4 - R.J. Knops Eds. Research notes in math. (No 39 pp 136/192 - Pitman 1979)
- [81] Toumi I. - Schémas de type Godunov pour les écoulements diphasiques multidimensionnels - Thèse de doctorat de l'Université Paris VI - 1989 (non publiée)
- [82] Toumi I. - A weak formulation of Roe's approximate Riemann solver - J.C.P. vol.102 No2 (1992) pp360/373
- [83] VanLeer B. - Flux-vector splitting for the Euler equations - Lectures Notes in Physics - vol.170 pp507/512 (Springer Verlag 1982)
- [84] VanLeer B. - On the relation between the upwind differencing schemes of Engquist-Oscher, Godunov and Roe - SIAM J. of Sci. Stat. Comp. Vol 5 No 1 (1984) pp1/20
- [85] Vol'pert A. - The space BV and quasi-linear equations - Math. USSR Sbornik - No2 (1967) pp225/267
- [86] Wagner D.H. - The transformation from Eulerian to Lagrangian coordinates for solutions with discontinuities - Nonlinear hyperbolic problems Proceedings St Etienne 1986 (Lecture notes in math. 1270 - Springer Verlag)
- [87] Yee H.C. Beam R.M. Warming R.F. - Boundary approximations for implicit schemes for one-dimensional inviscid equations of gas dynamics - AIAA Journal vol.20 No9 (1982) pp1203/1211
- [88] Yee H.C. Sweby P.K. and Griffiths D.F. - Dynamical approach study of spurious steady-state numerical solutions for nonlinear differential equations, Part I: the dynamics of time discretizations and its implications for algorithm development in computational fluid dynamics - J.C.P. 97 (1991) pp249/310
- [89] Zelmanse A. - Formulation cinétique et schémas de Boltzmann pour le calcul numérique en mécanique des fluides - Thèse de doctorat de l'Université Paris XIII - 1994 (non publiée)
- [90] Zhong X. - Additive semi-implicit Runge-Kutta methods for computing high-speed non equilibrium reactive flows - J.C.P. 128 (1996) pp19/31