

Construction et analyse numérique de schéma asymptotic preserving sur maillages non structurés. Application au transport linéaire et aux systèmes de Friedrichs

Emmanuel Franck

▶ **To cite this version:**

Emmanuel Franck. Construction et analyse numérique de schéma asymptotic preserving sur maillages non structurés. Application au transport linéaire et aux systèmes de Friedrichs. Analyse numérique [math.NA]. Université Pierre et Marie Curie - Paris VI, 2012. Français. NNT : . tel-00735956v3

HAL Id: tel-00735956

https://theses.hal.science/tel-00735956v3

Submitted on 23 Oct 2012

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.

Université Pierre et Marie Curie – Paris 6 CEA

Construction et analyse numérique de schémas « asymptotic preserving » sur maillages non structurés. Application à l'équation de transport et aux systèmes de Friedrichs.

THÈSE DE DOCTORAT

présentée et soutenue publiquement le 17 octobre 2012

pour l'obtention du

Doctorat de l'université Pierre et Marie Curie – Paris 6

Spécialité Mathématiques Appliquées

par

Emmanuel Franck

Composition du jury

Rapporteurs :	Christophe Berthon François Bouchut
Examinateurs :	Christophe Buet *Co-directeur de thèse* Bruno Després *Directeur de thèse* Edwige Godlewski Mohammed Lemou Nicolas Seguin

Après avis favorables des rapporteurs : Christophe Berthon et François Bouchut

École Doctorale, Sciences Mathématiques de Paris Centre **Laboratoire Jacques-Louis Lions**



Remerciements

Tout d'abord je souhaite remercier vivement mes directeurs de thèse Bruno Després et Christophe Buet qui ont accepté d'encadrer ce travail et l'ont accompagné de leurs conseils et remarques avisés, de leur enthousiasme et bonne humeur communicatifs ainsi que de leur patience, notamment devant mes premières rédactions (je sais que cela n'a pas été toujours d'un grand repos pour eux). Travailler avec eux pendant ces trois années fut très enrichissant d'un point de vue scientifique et a rendu cette expérience plus facile et agréable. J'espère que l'on pourra recommencer à travailler ensemble un jour.

Merci aussi à Eric Sonnendrücker de me donner la possibilité de continuer la recherche à Munich avec lui.

Je suis très reconnaissant envers Christophe Berthon et François Bouchut d'avoir accepté de rapporter sur ce travail ainsi qu'à Nicolas Seguin, Edwige Godlewski et Mohammed Lemou pour avoir accepté de faire partie du jury.

Je remercie Bruno Scheurer, Hervé Joudren du CEA pour leur suivi et leurs remarques pertinentes pendant ces trois ans ainsi que Stéphanie Perrot et Isabelle Visotto pour leur disponibilité.

J'aimerais aussi remercier Philippe Hoch, Gérald Samba, Emmanuel Labourasse et Xavier Blanc avec qui j'ai eu le grand plaisir de travailler ainsi que Clément Cances, Frédéric Coquel, Cedric Enaux, Stephane Jaouen pour nos échanges enrichissants. J'espère que ces échanges pourront continuer.

Durant ces trois années au CEA et un peu à Jussieu, j'ai croisé un certains nombre de doctorants, post-docs et stagiaires qui ont rendu ces années très agréables par leur bonne humeur, leur soutien. J'ai beaucoup aimé échanger avec eux même si parfois les débats ont un peu trop duré (voire beaucoup trop). Je remercie donc Alexandra, Gaêl, Gilles, François, Marc, Bertrand (qui fut un excellent co-bureau), Guillaume, Sebastien, Jean-Yves, Nicolas, Stephane, Camille, Jordan, Alexandre, Jean-Batiste, Matthieu, Julien, Khaled, Maxime, Morgan, Nafi. J'espère qu'on se re-croisera pour de nouvelles collaborations (scientifiques cette fois). Je voudrais aussi dire merci à tous les stagiaires (notamment a ceux qui ont partagé cette fin de thèse Méhdi, Gustave, Augustin, Christelle, Paul, Adrien, Jerome) qui sont passés au CEA durant ces trois ans et souhaiter bonne chance aux trois futurs doctorants : Thomas, Antoine et Emmanuelle (dans le fond cela reste un bon moment).

Evidemment ces remerciements ne seraient pas complet sans penser à mes amis "extérieurs" qui, pendant mes études, m'ont toujours soutenu et détendu (sans forcement le savoir) par leur présence malgré mes digressions pas toujours très compréhensibles. Ils peuvent continuer en venant me rendre visite à Munich, ils seront acceuillis les bras ouverts. Je me suis dis au début que je me lancerais pas dans une longue énumération, puis après tout c'est l'endroit ou je fais ce que je veux et le seul qui est lisible pour tous. Je commencerais par ceux qui sont là depuis longtemps, presque toujours en faite, David, Jerome, Deborah, Milena, Polo, Judie, Tina (que je félicite pour son récent doctorat), Julien, Rebecca et sa clique (Gwen, Lolo, Alex, Matthias, Raphaele). Il y aussi ceux qui ont partagé ma vie universitaire sur la butte aux cailles, Marine, Valentine, Flore, Vanessa, Jeremy, Johannes, Bik-

tou, Vince, Leila, Etienne, Antoine, Antoine, Than, Domen, Hadrien, Maeva, Seb, Olivier, Armand, Margaux, Maria, Charlotte, Guillaume, Amadou. Et puis il y a ceux qui j'ai partagé de nombreuses vacances, Julien, Will, Denis, Maxime, Marion, Jeanne, Leslie, Mathilde, Pierre, Nicolas, Christophe, Ghislain, ceux de l'agrégation de philosophie Paul, Jeanne, Germain, Rachel, Isabelle, Raph et ceux du vendredi thésard Pierre-Marie, Julien, Charlotte, Alex, Audray, Claire, Laetitia, et tous les autres (si j'ai oublié des gens je m'en excuse et je les remercie). Sans eux toutes ces années n'auraient pas eu le même charme. Vraiment merci à tous. Pour finir je souhaiterais adresser un merci tout particulier à Lucie qui a été toujours présente malgré les moments de doutes pendant cette période.

J'aimerais évidemment remercier mes parents qui se sont intéressés à ce travail (ce n'était pas gagné), m'ont soutenu à chaque instant et ont eu le courage de relire ce manuscrit, ainsi que ma famille pour son soutien et sa présence. Enfin j'aimerais dédier ce travail à mon grand père Claude qui m'a soutenu dans cette voie et qui hélas n'est plus là pour le lire.

Table des matières

I Introduction et contexte **3**

Introduction **5**

1 Contexte physique et mathématique **9**

1.1 Contexte physique 9

1.1.1 Fusion par Confinement Inertiel 9

1.1.2 Équations de transport 10

1.2 Modèles mathématiques pour l'approximation des problèmes de transport 13

1.2.1 Approximations P_n 14

1.2.2 Approximations S_n (méthodes des ordonnées discrètes) 16

1.2.3 Modèle M_1 16

1.2.4 Extension multi-groupes 17

1.2.5 Autres modèles 17

1.3 Contexte numérique 19

1.3.1 Schémas « asymptotic preserving » : introduction et applications 20

1.3.2 Schémas « asymptotic preserving » « centré » ou « décalé » 23

1.3.3 Schémas « asymptotic preserving » basés sur le schéma upwind 25

1.3.4 Maillages non structurés 39

1.3.5 Extension des schémas « asymptotic preserving » sur maillages non structurés 40

1.3.6 Schémas de diffusion sur maillages non structurés 43

1.4 Objectifs de l'étude 45

II Schémas « asymptotic preserving » pour des modèles linéaires **47**

2 Schémas « asymptotic preserving » aux noeuds pour l'équation de la chaleur hyperbolique **49**

2.1 Schémas « asymptotic preserving » avec les flux localisés aux noeuds 49

2.1.1 Rappels sur le schéma GLACE pour l'équation des ondes 50
2.1.2 Construction du schéma « asymptotic preserving » aux noeuds 51
2.1.3 Étude de l'inversibilité de la matrice locale 55
2.1.4 Variantes du schéma aux noeuds 63
2.2 Étude théorique des schémas aux noeuds 64
2.2.1 Résultats et hypothèses géométriques 64
2.2.2 Convergence du schéma limite de diffusion 65
2.2.3 Étude de stabilité des schémas « asymptotic preserving » 68
2.2.4 Convergence du schéma JL-(b) vers le schéma de diffusion 70
2.2.5 Dérivation des schémas JL-(a) et JL-(b) en dimension un 75
2.2.6 Conclusion sur les schémas « asymptotic preserving » 78
2.3 Discrétisation temporelle et conditions limites 79
2.3.1 Discrétisation temporelle 79
2.3.2 Conditions limites 82
2.4 Modes parasites et stabilisation 82
2.5 Résultats numériques 84
2.5.1 Résultats numériques pour le schéma de diffusion 84
2.5.2 Résultats numériques en régime de transport 87
2.5.3 Résultats numériques en régime asymptotique 89
2.6 Synthèse et perspectives 95

3 Schémas linéaires aux arêtes pour l'équation de la chaleur hyperbolique 97
3.1 Schémas de diffusion Breil-Maire et MPFA 97
3.2 Schéma aux arêtes pour l'équation de la chaleur hyperbolique 102
3.3 Résultats numériques 106
3.4 Synthèse 108

4 Étude des systèmes de Friedrichs avec termes sources raides 113
4.1 Étude théorique des systèmes de Friedrichs avec termes sources raides 113
4.1.1 Systèmes de Friedrichs : définitions 113
4.1.2 Limite de diffusion 115
4.1.3 Preuve rigoureuse d'existence de la limite de diffusion 117
4.2 Décomposition « diffusive - non diffusive » 122
4.2.1 Décomposition « diffusive - non diffusive » 122
4.2.2 Discrétisation de la décomposition 124
4.3 Applications aux modèles P_n 126
4.3.1 Modèles P_n 126

4.3.2 Analyse de la méthode en dimension un 128
4.3.3 Décomposition « diffusive - non diffusive » : application au modèle P_3 en 2D 130
4.4 Applications aux modèles S_n 132
4.4.1 Modèles S_n 132
4.4.2 Décomposition « diffusive - non diffusive » : application au modèle S_2 en 2D 134
4.5 Discrétisation en temps 136
4.5.1 Discrétisation implicite 136
4.5.2 Discrétisation semi-implicite 138
4.6 Résultats numériques 139
4.6.1 Résultats numériques pour le modèle S_2 139
4.6.2 Résultats numériques pour le modèle P_3 142
4.7 Synthèse 144

III Schémas « asymptotic preserving » pour des modèles non linéaires 147

5 Modèle P_1 et couplage avec la matière 149
5.1 Modèle P_1 couplé avec la matière et limite asymptotique 149
5.2 Schémas « asympotic preserving » en 1D pour le couplage avec la matière 151
5.2.1 Schémas de Jin-Levermore et Gosse-Toscani 151
5.2.2 Schéma « asymptotic preserving » et principe du maximum 154
5.3 Extension en dimension deux 157
5.3.1 Schéma semi-discret et limite asymptotique 158
5.3.2 Schéma semi-implicite pour le modèle P_1 couplé à la matière 159
5.3.3 Schéma semi-implicite par « splitting » pour le modèle P_1 couplé avec la matière 161
5.4 Résultats numériques 161
5.5 Synthèse 164

6 Discrétisation du modèle M_1 167
6.1 Rappels sur le modèle M_1 168
6.2 Rappels et « formulation comme un système de la dynamique des gaz » 169
6.3 Discrétisation du système Lagrangien associé au modèle M_1 171
6.3.1 Schéma pour le système Lagrangien et limite de diffusion 171
6.3.2 Entropie et principe du maximum pour le schéma Lagrangien 175
6.4 Étape de projection 177
6.4.1 Rappels sur le schéma aux noeuds d'advection 177
6.4.2 Projection et limite de diffusion 178

6.4.3 Principe du maximum pour la phase de projection . . . 179
6.4.4 Projection d'ordre 2, méthode MUSCL . . . 180
6.4.5 Autre schéma d'advection . . . 180
6.5 Schéma Lagrange+Projection à un pas . . . 181
6.5.1 Schéma Lagrange+Projection et limite de diffusion . . . 181
6.5.2 Schéma limite de diffusion : interprétation et remarques . . . 184
6.5.3 Condition d'entropie pour le schéma total . . . 184
6.6 Discrétisation semi-implicite en temps . . . 191
6.7 Résultats numériques . . . 192
6.7.1 Résultats numériques pour les schémas limites . . . 192
6.7.2 Résultats numériques pour le modèle M_1 . . . 195
6.8 Autres tenseurs d'Eddington . . . 197
6.9 Synthèse . . . 199

IV Perspectives **201**

7 Schémas positifs pour les modèles S_n d'ordonnées discrètes **203**
7.1 Schémas positifs, « asymptotic preserving » pour les modèles S_n en dimension un 203
7.1.1 Modèles S_n en dimension un . . . 204
7.1.2 Schéma sur grille non uniforme pour l'équation de la chaleur hyperbolique 204
7.1.3 Schéma « flux pair-impair » basé sur le schéma de Gosse-Toscani . . . 206
7.1.4 Résultats numériques . . . 210
7.2 Extension sur maillages non structurés . . . 213
7.2.1 Extension de la méthode « pair - impair » sur maillages de Delaunay . . . 213
7.2.2 Schéma « pair-impair » sur maillages de Delaunay et propriétés . . . 216
7.2.3 Correction du schéma sur maillage non structurés . . . 218
7.3 Synthèse et perspectives . . . 218

Conclusion et Perspectives **221**

Bibliographie **225**

Résumé/Abstract **231**

Glossaire

Calcul matriciel et vectoriel

Les espaces vectoriels et matriciels utilisés sont :

- $M^{p\times q}$: matrices de p lignes et q colonnes.
- M^p : matrices de p lignes et p colonnes.
- $\mathbb{R}^p$ vecteurs de taille p.

Les vecteurs sont notés en gras. Par exemple $\mathbf{U}$ un vecteur de $\mathbb{R}^n$ et U_i la ième composante de $\mathbf{U}$.
La transposition, la trace matricielle, le produit tensoriel, les normes matricielle et vectorielle et le produit scalaire sont respectivement donnés par :

- $A^t \in M^{q\times p}$ de composantes $A^t_{ij} = A_{ji}$,
- $\mathrm{Tr}(A) = \sum_{i=1}^{p} A_{ii}$,
- $\mathbf{x} \otimes \mathbf{y}$ avec $\mathbf{x} \in \mathbb{R}^p$, $\mathbf{y} \in \mathbb{R}^p$ de composantes $(\mathbf{x} \otimes \mathbf{y})_{ij} = x_i y_j$,
- $\parallel A \parallel = \sigma(A) = \sqrt{\lambda_{max}(A^t A)}$ avec $\lambda_{max}(A^t A)$ la valeur propre maximale de $A^t A$,
- $\parallel \mathbf{x} \parallel = \sqrt{\sum_i x_i^2}$,
- $(\mathbf{x}, \mathbf{y}) = \sum_i x_i y_i$.

Pour finir on note les matrices issues d'un produit tensoriel sous la forme $\widehat{\alpha} = \mathbf{x} \otimes \mathbf{y}$.

Calcul différentiel

Soit $f(\mathbf{x}, t)$ une fonction de $\mathbb{R}^p \times \mathbb{R}^+$, on note les dérivées partielles

$$\frac{\partial f(\mathbf{x}, t)}{\partial_{x_i}} = \partial_{x_i} f$$

Le vecteur gradient pour $f : D \in \mathbb{R}^2 \times \mathbb{R}^+ \longrightarrow \mathbb{R}$ est défini par

$$\nabla f = \begin{pmatrix} \partial_{x_1} f \\ \partial_{x_2} f \end{pmatrix}$$

et l'opérateur différentiel de divergence est défini par $\mathrm{div}(f) = \partial_{x_1} f + \partial_{x_2} f$.

Espaces fonctionnels et normes

Les espaces fonctionnels sont

- $L^p(D)$ l'espace des fonctions p-intégrables munies de la norme $\parallel f \parallel_{L^p(D)} = \left(\int_D f(\mathbf{x})^p d\mathbf{x}\right)^{\frac{1}{p}}$.
- $H^p_{\mathbf{x}}(D)$ un espace de Sobolev muni de la norme

$$\parallel f \parallel_{H^p_{\mathbf{x}}} = \sum_{a,b}^{a+b\leq p} \parallel \partial_{x^a, y^b} f \parallel_{L^2(D)},$$

avec $\alpha = a + b$.

Les normes $L^p(D)$ discrètes sont définies par

$$\parallel f \parallel_{L^p(D)} = \left(\sum_j \mid \Omega_j \mid f_j^p \right)^{\frac{1}{p}}.$$

Principales fonctions et paramètres

On introduit $\mathbf{x}$ le vecteur de position, $\mathbf{\Omega}$ le vecteur de direction, ν la fréquence d'un photon et e l'énergie associée à un neutron.

- p. 7 $\phi(t, \mathbf{x}, \mathbf{\Omega}, e)$ le flux de neutrons.
- p. 7 σ_a la section efficace (resp. opacité) d'absorption.
- p. 7 σ^* la section efficace de collision.
- p. 8 $I(t, \mathbf{x}, \mathbf{\Omega}, \nu)$ l'intensité radiative.
- p. 8 σ l'opacité de « scattering » (ou collision).
- p. 8 c la vitesse de la lumière, h la constante de Planck.
- p. 8 $p(\mathbf{\Omega}, \mathbf{\Omega}')$ la fonction de répartition angulaire.
- p. 9 $E(t, \mathbf{x})$ le moment d'ordre 0 de la fonction de distribution (flux neutronique, intensité radiative).
- p. 9 $\mathbf{F}(t, \mathbf{x})$ le moment d'ordre 1 de la fonction de distribution (flux neutronique, intensité radiative).
- p. 9 $\widehat{P}(t, \mathbf{x})$ le moment d'ordre 2 de la fonction de distribution (flux neutronique, intensité radiative).
- p. 9 $T(t, \mathbf{x})$ la température de la matière.
- p. 9 $T_r(t, \mathbf{x})$ la température radiative.
- p. 10 Y_l^m harmoniques sphériques.
- p. 12 $\mathbf{f}$ le flux radiatif normalisé.
- p. 12 χ le tenseur d'Eddington.
- p. 13 $\mathbf{\Omega}_i$ une vitesse de quadrature discrète pour la méthode des ordonnées discrètes.
- p. 13 $f(\mathbf{\Omega}_i)$ la fonction de distribution discrète associée à $\mathbf{\Omega}_i$.
- p. 13 w_i le poids de quadrature associé à $\mathbf{\Omega}_i$.

On donne la page associée à la première occurrence de la fonction ou du paramètre. **Principaux systèmes d'équations** :

- (1.1)-(1.2) Le système de l'hydrodynamique radiative.
- (1.3) L'équation de transport des neutrons.
- (1.6) L'équation de transfert radiatif.
- (1.4) L'équation de diffusion des neutrons.
- (1.7) L'équation de diffusion à l'équilibre des photons.
- (1.8) L'équation de diffusion hors équilibre des photons.
- (1.17) Le modèle M_1.
- (1.16) Les modèles S_n.
- (1.62) L'équation de la chaleur hyperbolique.
- (4.1) Un système de Friedrichs avec un terme source raide.
- (5.1) Le modèle P_1 couplé avec la matière.
- (4.52)-(4.53)-(4.55) Le modèle P_3.
- (4.59)-(4.60) Le modèle S_2.

Première partie

Introduction et contexte

Introduction

Contexte général

La motivation physique est celle de la Fusion par Confinement Inertiel (FCI). L'objectif de la FCI [DW94] est d'obtenir la fusion d'un mélange de Deutérium et Tritium dans une capsule en chauffant et en comprimant le mélange à l'aide d'un laser. Le milieu de la réaction est modélisé par les équations de l'hydrodynamique radiative qui décrit l'interaction entre le gaz ionisé (ou plasma) et le rayonnement.
On se focalisera uniquement dans ce mémoire sur les problèmes de transport qui interviennent dans les modèles d'interaction entre la matière (gaz, solide etc) et le rayonnement (photons). La dynamique des particules est décrite par une équation cinétique (échelle mesoscopique) dépendante de la vitesse, de la position voire, de l'énergie des particules.
Cette étude porte plus précisément sur l'approximation des problèmes de transport sur des maillages généraux, issus de la discrétisation sur maillages mobiles des modèles hydrodynamiques.

Après avoir utilisé une approximation en vitesse classique (approximations P_n [Bru02], méthode des ordonnées discrètes [Chan60]- [JL91]), on étudiera une famille de discrétisation spatiale adaptée à ce type de problème, basée sur la méthode des volumes finis. En effet, les approximations angulaires des équations cinétiques linéaires donnent des systèmes hyperboliques avec des termes sources raides qui induisent une limite asymptotique de diffusion. La raideur du terme source et l'existence de la limite de diffusion génèrent des contraintes, sur les discrétisations spatiales et temporelles. Afin de pallier ces contraintes, des méthode nommées « asymptotic preserving » ont été introduites depuis quelques années [JL96]- [GT01]- [BLeFT11]- [DDSV09]- [Jin11].
À notre connaissance ce type de méthode n'a pas été pour le moment, étendu en dimension deux sur maillages non structurés. Le but de cette étude est donc de construire et d'étudier des méthodes de types volumes finis en dimension deux pour des systèmes hyperboliques admettant une limite de diffusion. On se concentrera sur des modèles issus de la discrétisation angulaire de l'équation de transport linéaire.

Plan du mémoire

Chapitre 1

Le premier chapitre est dédié au contexte physique, mathématique et numérique de cette étude. On commencera par introduire le modèle de l'hydrodynamique radiative ainsi que les équations cinétiques décrivant un partie des phénomènes présents en FCI. Une fois ces modèles présentés on proposera un état de l'art sur la discrétisation en vitesse et sur les modèles simplifiés utilisés pour approcher les solutions des équations de transport. La seconde partie du chapitre sera consacrée à la présentation des méthodes « asymptotic preserving » en dimension un. On partira de l'équation de la chaleur

hyperbolique

$$\begin{cases} \partial_t E + \dfrac{1}{\varepsilon}\partial_x F = 0, \\ \\ \partial_t F + \dfrac{1}{\varepsilon}\partial_x E + \dfrac{\sigma}{\varepsilon^2} F = 0. \end{cases} \tag{1}$$

avec $\varepsilon > 0$. Dans le cadre de l'approximation de l'équation (1) on présentera des méthodes classiques et « asymptotic preserving » pour lesquelles on donnera des résultats de stabilité et de consistance. Enfin on proposera une extension en dimension deux d'une de ces méthodes à l'aide du formalisme classique des volumes finis. Ce dernier résultat nous permettra d'exhiber des problèmes de consistance inhérents aux schémas de diffusion sur maillages non structurés.

Chapitre 2

Le second chapitre est consacré à la construction de schémas « asymptotic preserving » pour l'équation de la chaleur hyperbolique sur maillages non structurés :

$$\begin{cases} \partial_t E + \dfrac{1}{\varepsilon} \operatorname{div} \mathbf{F} = 0, \\ \\ \partial_t \mathbf{F} + \dfrac{1}{\varepsilon}\nabla E + \dfrac{\sigma}{\varepsilon^2} \mathbf{F} = \mathbf{0}. \end{cases} \tag{2}$$

Afin d'obtenir ces nouvelles méthodes, on proposera de coupler un formalisme alternatif des volumes finis dans lequel les flux sont localisés aux noeuds [CDDL09] avec la méthode de Jin-Levermore [JL96]. Cette méthode construite afin d'obtenir des discrétisations « asymptotic preserving » est basée sur l'introduction des termes sources dans les flux upwind à l'aide de développement de Taylor. À partir de ce nouveau formalisme on obtiendra deux schémas qui peuvent être considérés comme l'extension des schémas de Jin-Levermore [JL96] et Gosse-Toscani [GT01] en dimension deux. Le schéma limite obtenu est un nouveau schéma pour lequel on donnera une preuve de convergence sous conditions sur les maillages. Enfin, on étudiera différentes discrétisations temporelles permettant de s'affranchir de la contrainte sur le pas de temps issue de la raideur du système (2) et on présentera quelques résultats numériques.

Chapitre 3

Le contenu du chapitre trois est issu d'un travail réalisé lors du CEMRACS 2010 avec G. Samba, P. Hoch et P. Navarro [FHSN11]. Il s'agit de construire des schémas « aysmptotic preserving » pour l'équation de la chaleur hyperbolique (2) basés sur deux schémas de diffusion [AE06]- [MABO07]. Ces schémas, contrairement aux schémas aux noeuds introduits dans le chapitre précédent, n'admettent pas de modes parasites (dégradation de la qualité de la solution numérique) pour certains types de données. Dans ce chapitre on rappellera les deux schémas de diffusion cible MPFA et Breil-Maire avant de construire un schéma hyperbolique compatible qui sera corrigé à l'aide de la méthode de Jin-Levermore. Enfin on donnera quelques résultats qui montrent le caractère « asymptotic preserving » de la méthode.

Chapitre 4

Afin d'obtenir des approximations précises de l'équation de transport aussi bien dans le régime de diffusion que dans le régime de transport ($\sigma \approx 1$, $\varepsilon = 1$), on proposera d'étendre les schémas

« asymptotic preserving » au cas des systèmes linéaires généraux du type

$$\partial_t \mathbf{U} + \frac{1}{\varepsilon} A_1 \partial_x \mathbf{U} + \frac{1}{\varepsilon} A_2 \partial_y \mathbf{U} = -\frac{\sigma}{\varepsilon^2} R\mathbf{U}. \tag{3}$$

Après avoir caractérisé la limite de diffusion de ce type de système, on proposera une décomposition proche de la décomposition « micro - macro » [LM07] qui consiste à séparer le modèle (3) entre un système proportionnel à l'équation de la chaleur hyperbolique et un système qui n'intervient pas dans le régime de diffusion. Afin d'obtenir une méthode « asymptotic preserving » on proposera de résoudre le premier système avec un des schémas précédemment introduits et le second système avec un schéma classique. Enfin, on montrera que les modèles P_n et S_n satisfont les hypothèses nécessaires pour pouvoir utiliser la décomposition « diffusive - non diffusive » . Des résultats numériques seront donnés pour les modèles P_3 et S_2.

Chapitre 5

Après avoir traité la discrétisation spatiale pour des systèmes hyperboliques linéaires issus du transfert radiatif, on proposera de discrétiser des modèles non linéaires. Le premier, qui sera étudié dans ce chapitre, correspond au modèle P_1 (1.62) couplé avec une énergie sur la matière [Tur03]. Ce couplage induit des termes sources non linéaires et raides modifiant ainsi la limite de diffusion. À travers cette étude on montrera que les schémas « asymptotic preserving » construits pour le modèle sans couplage permettent de capturer cette nouvelle limite de diffusion. Afin de traiter la raideur des termes sources, on proposera une discrétisation implicite adaptée.

Chapitre 6

Le second modèle non linéaire étudié appelé M_1 est un système hyperbolique décrivant l'évolution des deux premiers moments de la solution de l'équation de transport.

$$\begin{cases} \partial_t E + \dfrac{1}{\epsilon} \operatorname{div} \mathbf{F} = 0, \\ \\ \partial_t \mathbf{F} + \dfrac{1}{\epsilon} \nabla \widehat{P} = -\dfrac{\sigma}{\epsilon^2} \mathbf{F}, \end{cases} \tag{4}$$

avec le tenseur de pression

$$\widehat{P} = \frac{1}{2} \left((1 - \chi(\mathbf{f})) \widehat{I}_d + (3\chi(\mathbf{f}) - 1) \frac{\mathbf{f} \otimes \mathbf{f}}{||\mathbf{f}||} \right) E,$$

$\mathbf{f} = \mathbf{F}/E$, et le tenseur d'Eddington $\chi(\mathbf{f}) = \dfrac{3 + 4||\mathbf{f}||^2}{5 + 2\sqrt{4 - 3||\mathbf{f}||^2}}$. Afin d'utiliser le formalisme aux noeuds, on formulera le modèle M_1 comme un système de la dynamique des gaz. Une fois cette formulation introduite on proposera une discrétisation basée sur un schéma Lagrange+projection aux noeuds couplé avec la méthode de Jin-Levermore. Le schéma ainsi obtenu est entropique, préserve le principe du maximum et la limite asymptotique associés au modèle. Dans un premier temps, les caractères entropique et « asymptotic preserving » seront étudiés séparément pour la phase Lagrange et pour la phase de projection. Enfin ces résultats seront introduits pour le schéma total.

Chapitre 7

Dans ce dernier chapitre, on proposera une réflexion sur la construction de discrétisation positive et « asymptotic preserving » pour les modèles S_n. En effet la décomposition présentée dans le chapitre quatre ne permet pas de conserver la positivité des variables. À cette fin on introduira un nouveau schéma 1D sur grilles non uniformes. Cette nouvelle discrétisation est basée sur le schéma de Gosse-Toscani [GT01] couplé avec une formulation « pair - impair » des modèles considérés. Après une rapide étude théorique du schéma obtenu on donnera quelques pistes permettant d'étendre ces méthodes en dimension deux sur maillages non structurés.

Chapitre 1

Contexte physique et mathématique

1.1 Contexte physique

1.1.1 Fusion par Confinement Inertiel

Ce travail est motivé par la modélisation et la simulation de la Fusion par Confinement Inertiel (FCI). Ce processus de fusion nucléaire détaillé dans [DW94]- [AMTV04] consiste à comprimer une capsule de Deutérium-Tritium gazeux entourée d'une phase cryogénique à l'aide d'un laser puissant. Cette compression permet d'atteindre les seuils de température, pression et densité nécessaires à l'amorce d'une réaction de fusion produisant ainsi de l'énergie, un isotope d'hélium (ou particule α) et des neutrons. L'évolution du plasma (issu de la ionisation du gaz de Deutérium-Tritium) induite par la compression de la capsule, peut être modélisée par un modèle d'Euler multi-matériaux (voir [BD04]-[McL08]) à deux températures (utilisées pour décrire le comportement couplé des espèces ionique et électronique) avec des termes sources induits par le rayonnement et la réaction de fusion. Il est donné par

$$\begin{cases} D_t\rho + \rho \operatorname{div} \mathbf{u} = 0, \\ \rho D_t \mathbf{u} + \nabla p = \mathbf{F}, \\ \rho D_t E_e + p_e \operatorname{div} \mathbf{u} - \operatorname{div}(\chi_e \nabla T_e) + W_{ei} = Q + W_{\alpha_e} + W_s + S, \\ \rho D_t E_i + p_i \operatorname{div} \mathbf{u} + \operatorname{div}(\chi_i \nabla T_i) - W_{ei} = W_{\alpha_i}. \end{cases} \tag{1.1}$$

L'opérateur $D_t = \partial_t + \mathbf{u}.\nabla$ correspond à la dérivée matérielle. Les inconnues du système sont : la densité ρ du plasma, la vitesse $\mathbf{u}$ du plasma, les pressions électronique p_e et ionique p_i dont la somme donne la pression totale p. Les quantités T_e, T_i, E_e, E_i représentent les températures et énergies électroniques et ioniques. La relaxation entre les températures associées aux deux espèces est décrite par le terme source W_{ei}. Le rayonnement induit deux termes sources : le flux radiatif noté $\mathbf{F}$ et le terme d'échange d'énergie entre le rayonnement et le plasma noté Q. Pour conclure on introduit les termes $W_{\alpha_{e,i}}$ et W_s qui sont issus des particules « non thermiques » et S la puissance laser absorbée.

L'utilisation d'un modèle hydrodynamique deux-températures peut se justifier car les populations électronique et ionique sont distribuées en grande partie selon la distribution d'équilibre appelée Maxwellienne (il s'agit des particules « thermiques »). La partie restante des populations ionique et électronique n'est pas à l'équilibre thermodynamique local (particules dite « non thermiques »). Il s'agit des particules émises par la réaction et qui sont à des niveaux supérieurs d'énergie. Ces particules vont échanger de l'énergie avec le plasma à travers les termes sources $W_{\alpha_{e,i}}$, W_s jusqu'à atteindre l'équilibre thermodynamique local. La dynamique des particules dites « non thermiques » est modélisée par une équation de type Vlasov collisionnelle sur la fonction de distribution $f(t, \mathbf{x}, \mathbf{v})$

$$\partial_t f + \mathbf{v}.\nabla_{\mathbf{x}} f + \frac{\mathbf{E} + \frac{\mathbf{v}}{c} \wedge \mathbf{B}}{m}.\nabla_{\mathbf{v}} f = Q(f, f)$$

avec $Q(f,f)$ un opérateur de collision et $\mathbf{E}$, $\mathbf{B}$ les champs électrique et magnétique. Pour sa part l'interaction entre le rayonnement de photons et la matière (transfert radiatif) donne les termes sources Q et $\mathbf{F}$, calculés à partir de l'intensité radiative $I = I(t, \mathbf{x}, \mathbf{\Omega}, \nu)$, solution de l'équation de transfert radiatif :

$$\begin{cases} \frac{\rho}{c} D_t \left(\frac{I}{\rho}\right) + \mathbf{\Omega}.\nabla_{\mathbf{x}} I + \sigma_a I - \sigma_e B(T_e) + \\ (\mathbf{\Omega} \otimes \mathbf{\Omega} : \nabla \mathbf{u})(I - \partial_\nu(\nu I)) + \frac{1}{4\pi} \int_0^\infty \int_{S^2} \sigma p(\mathbf{\Omega}, \mathbf{\Omega}') \left(I(t, \mathbf{x}, \mathbf{\Omega}', \nu) - I\right) d\mathbf{\Omega}' d\nu, \end{cases} \tag{1.2}$$

avec le flux radiatif

$$\mathbf{F} = \frac{1}{4\pi} \int_0^\infty \int_{S^2} \mathbf{\Omega}(\sigma_a + \sigma) I d\mathbf{\Omega} d\nu$$

et le terme d'échange

$$Q = \frac{1}{4\pi} \int_0^\infty \int_{S^2} (\sigma_a I - \sigma_e B(T_e)) d\mathbf{\Omega} d\nu.$$

Dans ce cadre, ν représente la fréquence, $\mathbf{\Omega}$ représente la direction du photon, la fonction de répartition angulaire est définie par $p(\mathbf{\Omega}, \mathbf{\Omega}') = \frac{3}{4}(1 + (\mathbf{\Omega}, \mathbf{\Omega}'))$ et σ, σ_a, σ_e représentent les opacités de « scattering » , d'absorption et d'émission. De la même façon les neutrons émis par la réaction sont modélisés par une équation cinétique de type Boltzmann.

Pour résoudre les équations hydrodynamiques (1.1) on peut utiliser des méthodes Euleriennes (les équations sont écrites dans le référentiel terrestre, puis discrétisées), des méthodes Lagrangiennes [Des010]- [DM03]- [Maz07]- [Mai11], utilisant le référentiel du plasma où des méthodes de types ALE [BEHN10] qui passent arbitrairement de la formulation Lagrangienne à la formulation Eulérienne.
La résolution des équations écrites en coordonnée Eulérienne permet d'utiliser les méthodes numériques sur des maillages structurés (grille Cartésienne), contrairement aux deux autres formulations. En effet, en coordonnées Lagrangienne ou dans le formalisme ALE on se place dans le référentiel du plasma, par conséquent le maillage se déplace et se déforme à la vitesse de la matière. Les méthodes Lagrangienne et ALE génèrent des maillages non structurés et nécessitent donc des méthodes numériques valides sur des maillages non structurés pour résoudre les autres modèles intervenant en FCI.

1.1.2 Équations de transport

Comme il a été décrit dans la section précédente, la Fusion par Confinement Inertiel fait intervenir des modèles régissant la dynamique des neutrons, photons ou électrons. Pour décrire la dynamique d'une population de particules, on utilise des modèles dits cinétiques qui modélisent l'évolution de la fonction de distribution représentant la probabilité de présence d'une particule à une position donnée et une vitesse donnée. Dans le cadre de ce travail on s'intéresse à l'approximation d'équations cinétiques sans terme de champ moyen (terme modélisant l'interaction avec une force extérieure), avec des opérateurs de collision linéaires (terme décrivant les interactions binaires entre deux particules), comme les équations de la neutronique et de la photonique.

Équation de transport en neutronique

Afin d'établir l'équation de transport de la neutronique on commence par introduire les variables qui interviennent. Soit la position $\mathbf{x} \in \mathbb{R}^3$, la vitesse $\mathbf{v} \in \mathbb{R}^3$, la direction $\mathbf{\Omega} = \frac{\mathbf{v}}{\|\mathbf{v}\|} \in \mathbb{R}^3$ et l'énergie du neutron e. Le transport de neutrons est décrit par une équation de Boltzmann linéaire sur la densité

$f(t, \mathbf{x}, \mathbf{v}, e,)$. Afin d'étudier une quantité plus facilement observable, on introduit le flux de neutrons $\phi(t, \mathbf{x}, \mathbf{\Omega}, e) =\| \mathbf{v} \| f(t, \mathbf{x}, \mathbf{v}, e)$, qui satisfait l'équation de transport suivante :

$$(E_n) \begin{cases} \frac{1}{||\mathbf{v}||}\partial_t \phi(t, \mathbf{x}, \mathbf{\Omega}, e) + \mathbf{\Omega}.\nabla_{\mathbf{x}}\phi(t, \mathbf{x}, \mathbf{\Omega}, e) + \sigma_a \phi(t, \mathbf{x}, \mathbf{\Omega}, e) = \\ + \int_{e_{min}}^{e_{max}} \int_{\mathbb{R}^3} \sigma^*(\mathbf{x}, \mathbf{\Omega}, \mathbf{\Omega}', e, e') \left(\phi(t, \mathbf{x}, \mathbf{\Omega}', e') - \phi(t, \mathbf{x}, \mathbf{\Omega}, e)\right) d\mathbf{\Omega}' de' \end{cases} \tag{1.3}$$

avec la section efficace d'absorption σ_a qui représente le taux de neutrons capturés par les noyaux de la matière et σ^* la somme de la section efficace de collision et de celle de fission. Les quantités e_{min} et e_{max} représentent les valeurs maximale et minimale de l'énergie des neutrons. Dans le cas où le milieu est isotrope, σ^* ne dépend que de l'espace et de l'énergie.

Définition 1.1.1. *Le premier et second moments de $\phi(t, \mathbf{x}, \mathbf{\Omega}, e)$ sont définis par*

- *Le flux scalaire $E(t, \mathbf{x}) =< \phi >$,*
- *Le courant total $\mathbf{F}(t, \mathbf{x}) =< \mathbf{\Omega}\phi >$.*

avec

$$< . >= \int_{\mathbb{R}^3} \int_{e_{min}}^{e_{max}} .d\mathbf{\Omega} de.$$

Dans certains régimes liés à l'évolution de la matière, l'équation de transport peut être approchée par une équation de diffusion. Cette approximation de diffusion correspond au régime où les sections efficaces liés à la matière sont importantes.

Définition 1.1.2 (Limite de diffusion en neutronique)**.** *: Si $t >> 1$ et $\sigma^* >> 1$, l'équation des neutrons (1.3) tend vers l'équation*

$$\frac{1}{||\mathbf{v}_m||}\partial_t E(t, \mathbf{x}) - \text{div}\left(\frac{1}{3\sigma^*}\nabla E(t, \mathbf{x})\right) + \sigma_a E(t, \mathbf{x}) = 0. \tag{1.4}$$

avec $||\mathbf{v}_m||$ la vitesse moyenne des neutrons.

En effet, lorsque les sections efficaces sont importantes les phénomènes d'absorption et de collision génèrent des trajectoires Browniennes pour les neutrons, l'équation de diffusion limite étant le résultat macroscopique de nombreuses trajectoires Browniennes. Les équations de transport et de diffusion intervenant en neutronique ont été très étudiées dans la littérature notamment dans le cadre du calcul de criticité des réacteurs à fission nucléaire (voir J. Planchard [Pl95], P. Reuss [Reu03] ou le cours de G. Allaire et F. Golse [AG11]).

Équation de transport en photonique

On peut définir une équation cinétique régissant la propagation des photons et leurs interactions avec la matière. Les photons allant tous à la vitesse de la lumière, on introduit la direction $\mathbf{\Omega} = \frac{\mathbf{v}}{c}$. La population de photons est représentée par l'intensité radiative :

$$I(t, \mathbf{x}, \mathbf{\Omega}, \nu) = ch\nu f(t, \mathbf{x}, \mathbf{\Omega}, \nu) \geq 0,$$

où c est la vitesse de la lumière, h la constante de Planck et ν la fréquence du photon. En l'absence de matière les photons se déplacent à la vitesse de la lumière et ne sont soumis à aucune force ni

n'interagissent entre eux, par conséquent la distribution des photons satisfait une équation de transport classique sans terme source :

$$\frac{1}{c}\partial_t I(t,\mathbf{x},\mathbf{\Omega},\nu) + \mathbf{\Omega}.\nabla_{\mathbf{x}} I(t,\mathbf{x},\mathbf{\Omega},\nu) = 0.$$

Lorsque les photons interagissent avec la matière, plusieurs phénomènes interviennent :

- Le « scattering » : il s'agit du changement de direction et parfois de fréquence des photons du à une « collision » avec les électrons du milieu. Différents opérateurs de « scattering » sont décrits dans [AG11].
- L'absorption /émission : un photon peut être absorbé en excitant un électron d'un état d'énergie vers un autre si la fréquence de ce photon est proportionnelle (de proportion h) à la différence entre les deux niveaux d'énergies. Réciproquement un électron peut se désexciter en émettant un photon.

Afin de simplifier les mécanismes d'absorption et d'émission on suppose le milieu à l'équilibre thermodynamique local, il existe donc une température électronique T en tout point $\mathbf{x}$ et tout temps t. Enfin on introduit la fonction de Planck (Planckienne) qui correspond à l'intensité radiative d'un corps noir de température T :

$$B(T) = \frac{2h\nu^3}{c^2(e^{\frac{h\nu}{kT}} - 1)}, \tag{1.5}$$

où k est la constante de Boltzmann. L'énergie associée à la Planckienne suit la loi de Stefan-Boltzmann

$$E_B = \frac{4\pi}{c}\int_{S^2}\int_0^\infty B(T) d\mathbf{\Omega} d\nu = aT^4,$$

avec $a = \frac{4\sigma_r}{c}$, σ_r la constante de Stefan-Boltzmann. Les phénomènes de « scattering » et d'absorption sont guidés par les opacités, dépendantes de la matière. On nomme $\sigma(\rho,T,\nu)$ l'opacité de « scattering » $\sigma_a(\rho,T,\nu)$ l'opacité d'absorption et ρ la densité de la matière. En simplifiant le modèle décrivant la dynamique de la matière, on obtient l'équation du transfert radiatif (en faite un système) :

$$(ETR)\begin{cases} \frac{1}{c}\partial_t I(t,\mathbf{x},\mathbf{\Omega},\nu) + \mathbf{\Omega}.\nabla_{\mathbf{x}} I(t,\mathbf{x},\mathbf{\Omega},\nu) = \sigma_a(\rho,T,\nu)\left(B(T) - I(t,\mathbf{x},\mathbf{\Omega},\nu)\right) \\ +\frac{1}{4\pi}\int_0^\infty\int_{S^2}\sigma(\rho,T,\nu)p(\mathbf{\Omega},\mathbf{\Omega}')\left(I(\mathbf{x},\mathbf{\Omega}',\nu,t) - I(t,\mathbf{x},\mathbf{\Omega},\nu)\right)d\mathbf{\Omega}' d\nu, \\ \\ \frac{1}{c}\partial_t \epsilon\left(T(\mathbf{x},t)\right) = \frac{1}{4\pi}\int_0^\infty\int_{S^2}\sigma_a(\rho,T,\nu)\left(I(t,\mathbf{x},\mathbf{\Omega},\nu) - B(T)\right)d\mathbf{\Omega} d\nu. \end{cases} \tag{1.6}$$

avec $\epsilon(T)$ l'énergie du milieu, par exemple dans le cas d'un gaz parfait $\epsilon(T) = \rho C_v T$ et $p(\mathbf{\Omega},\mathbf{\Omega}')$ une fonction de répartition angulaire qui définit l'anisotropie du « scattering » . Elle sera supposée égale à 1 par la suite. Dans certains régimes l'équation du transfert radiatif tend vers une équation de diffusion sur le premier moment de $I(t,\mathbf{x},\mathbf{\Omega},\nu)$. Afin de définir cette limite on commence par définir les moments de l'intensité

Définition 1.1.3. *Les trois premiers moments de $I(t,\mathbf{x},\mathbf{\Omega},\nu)$ sont définis par*

- *L'énergie $E(t,\mathbf{x}) =< I >$, (la température radiative $T_r = (E/a)^{1/4}$),*
- *Le flux $\mathbf{F}(t,\mathbf{x}) =< \mathbf{\Omega} I >$,*
- *La pression $\widehat{P}(t,\mathbf{x}) =< \mathbf{\Omega}\otimes\mathbf{\Omega} I >$.*

avec

$$< . >= \int_0^\infty \frac{4\pi}{c}\int_{S^2} . d\mathbf{\Omega} d\nu$$

Comme dans le cas de la neutronique, l'équation de transfert radiatif peut être approchée par des équations de diffusion dans certains régimes de la matière [Tur03]- [BD06]- [BCT08].

Définition 1.1.4 (Limite à l'équilibre). *Si $t >> 1$ et $\sigma_a(\rho, T, \nu) >> \sigma(\rho, T, \nu)$, l'équation de transfert radiative tend vers l'équation de diffusion*

$$\partial_t \left(\epsilon(T) + aT^4\right) - \operatorname{div}\left(\frac{acT^3}{3\sigma_a}\nabla T\right) = 0. \tag{1.7}$$

Dans ce cas, le phénomène d'absorption et d'émission dominant permet à l'énergie radiative d'atteindre l'état d'équilibre $E = aT^4$ (équivalent à $T_r = T$). L'apparition d'une équation de diffusion s'explique, pour sa part, par le fait que les particules sont absorbées puis ré-émises dans une direction différente et aléatoire. Par conséquent, on peut considérer que les dynamiques des particules sont décrites par des mouvements Browniens dont l'équation de diffusion est le résultat macroscopique.

Définition 1.1.5 (Limite hors équilibre). *Si $t >> 1$ et $\sigma(\rho, T, \nu) >> \sigma_a(\rho, T, \nu)$, l'équation de transfert radiative tend vers l'équation*

$$\begin{cases} \partial_t E - \operatorname{div}\left(\dfrac{c}{3\sigma}\nabla E\right) = c\sigma_a\left(aT^4 - E\right), \\ \\ \partial_t \epsilon\left(T\right) = c\sigma_a\left(E - aT^4\right). \end{cases} \tag{1.8}$$

L'interprétation est assez simple puisque le milieu est fortement collisionnel, le photon suit une trajectoire Brownienne. La diffusion peut être vue comme le résultat macroscopique d'un grand nombre de photons suivant un mouvement Brownien.

Les équations de transport précédemment introduites sont des modèles précis mais elles sont délicates à utiliser en pratique. En effet, les méthodes déterministes directes sont trop coûteuses pour ces équations et les méthodes probabilistes de type Monte Carlo posent des problèmes dans le cadre du couplage avec les instabilités hydrodynamiques en raison du bruit statistique. Afin de résoudre ce problème, il est courant d'utiliser des modèles simplifiés approchant les solutions des équations de transport et moins coûteux en temps de calcul.

1.2 Modèles mathématiques pour l'approximation des problèmes de transport

L'équation de transport linéaire, comme les autres équations cinétiques, est très coûteuse à résoudre en raison du grand nombre de variables (six où sept en dimension trois). Afin de pallier ce problème il existe deux solutions on résout l'équation de transport à l'aide d'une méthode probabiliste de type Monte-Carlo : cela consiste à simuler de manière aléatoire le transport d'un grand nombre de particules. La vitesse de convergence de ce type de méthode est homogène à $\frac{1}{\sqrt{N}}$ où N est le nombre de particules. Cette méthode, bien que très efficace, entraîne des difficultés lorsqu'elle est couplée avec les équations de l'hydrodynamique notamment en présence d'instabilité. L'autre solution consiste à résoudre des modèles simplifiés ne dépendant que des variables d'espaces. Nous détaillerons la construction de ces modèles à partir de l'équation de transfert radiatif, mais le principe reste le même pour les équations de transport sans terme de champ moyen, avec un opérateur de collision linéaire. Ces modèles simplifiés sont construits pour approcher les solutions du problème de transport tout en conservant un maximum de propriétés liées à l'équation :

- positivité de $I(t, \mathbf{x}, \mathbf{\Omega}, \nu)$ (resp. de $\phi(t, \mathbf{x}, \mathbf{\Omega}, e)$), de $E(t, \mathbf{x})$,

- invariance par rotation,
- limite de diffusion,
- $\mathbf{f} = \frac{||\mathbf{F}||}{cE} \leq 1$, (pour le transfert radiatif).

1.2.1 Approximations P_n

On commence par introduire les approximations P_n définies dans [Bru05]- [Bru02]- [Bru05]- [HLMc10]. Elles sont obtenues en développant l'équation de transport sur la base des harmoniques sphériques et en tronquant le développement au moment n. Voici un bref rappel de la construction de ces approximations : les harmoniques sphériques sont les vecteurs propres du Laplacien défini sur la sphère, par conséquent elles forment une base Hilbertienne orthonormée des fonctions de carré intégrable sur la sphère. Soit $Y_l^m(\mathbf{\Omega}) = Y_l^m(\mu, \phi)$ une harmonique sphérique avec $\mathbf{\Omega} = (\sin\theta\cos\psi, \sin\theta\sin\psi, \cos\theta)$, $\mu = \cos\theta$, θ et ψ les coordonnées sphériques. Puisque les harmoniques sphériques forment une base de l'espace $L^2(S^2)$ on peut décomposer l'intensité radiative sous la forme suivante

$$I(t, \mathbf{x}, \mu, \psi, \nu) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} I_l^m(t, \mathbf{x}) Y_l^m(\mu, \psi), \tag{1.9}$$

avec

$$I_l^m(t, \mathbf{x}) = \int_{\mathbb{R}^+} \int_{4\pi} Y_l^m(\mu, \psi) I(t, \mathbf{x}, \mu, \psi) d\mathbf{\Omega} d\nu.$$

En injectant une troncature à l'ordre n du développement (1.9) dans l'équation de transport et en intégrant par rapport à chaque harmonique sphérique, on obtient les approximations P_n en 3D :

$$\left\{ \begin{array}{l} \frac{1}{c}\partial_t I_l^m + \frac{1}{2}\partial_x \left(-C_{l-1}^{m-1} I_{l-1}^{m-1} + D_{l+1}^{m-1} I_{l+1}^{m-1} + E_{l-1}^{m+1} I_{l-1}^{m+1} - F_{l+1}^{m+1} I_{l+1}^{m+1}\right) \\ +\frac{1}{2} i\partial_y \left(-C_{l-1}^{m-1} I_{l-1}^{m-1} + D_{l+1}^{m-1} I_{l+1}^{m-1} + E_{l-1}^{m+1} I_{l-1}^{m+1} - F_{l+1}^{m+1} I_{l+1}^{m+1}\right) \\ +\partial_z \left(A_{l-1}^m I_{l-1}^m + B_{l+1}^m I_{l+1}^m\right) + \sigma\left(I_0^0 \delta_{l0}\delta_{m0} - I_l^m\right) = 0, \end{array} \right. \tag{1.10}$$

pour $0 \leq l \leq n$ et $-l \leq m \leq l$, avec δ le produit de Kronecker.
Les calculs sont basés sur les remarques suivantes :

- Il existe des formules de récurrences entre les harmoniques sphériques $Y_l^m(\mathbf{\Omega})$ et $Y_l^{m+1}(\mathbf{\Omega})$, $Y_l^{m-1}(\mathbf{\Omega})$,
- $\int_{\mathbb{R}^+} \int_{4\pi} Y_l^m(\mathbf{\Omega}) I(t, \mathbf{x}, \mathbf{\Omega}, \mu) d\mathbf{\Omega} d\nu = I_l^m(t, \mathbf{x})$,
- $\int_{\mathbb{R}^+} \int_{4\pi} Y_l^m(\mathbf{\Omega}) \int_{4\pi} \frac{\sigma}{4\pi} I(t, \mathbf{x}, \mathbf{\Omega}', \nu) d\mathbf{\Omega}' d\mathbf{\Omega} d\nu = \sigma I_0^0(t, \mathbf{x}) \delta_{0l}\delta_{0m}$,
- $\int_{\mathbb{R}^+} \int_{4\pi} Y_l^m(\mathbf{\Omega}) \mathbf{\Omega}.\nabla I(t, \mathbf{x}, \mathbf{\Omega}, \nu) = \nabla^* . \left(\int_{\mathbb{R}^+} \int_{4\pi} Y_l^m(\mathbf{\Omega}) \mathbf{\Omega}^* I(t, \mathbf{x}, \mathbf{\Omega}, \nu) \right)$,

avec $\mathbf{\Omega} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, $\mathbf{\Omega}^* = (\sin\theta e^{i\phi}, \sin\theta e^{i\phi}, \cos\theta)$
et $\nabla^* = (\frac{1}{2}(\partial_x - i\partial_y), \frac{1}{2}(\partial_x + i\partial_y), \partial_z)$.
Afin de diminuer le nombre de variables on utilise la relation $\overline{I}_l^m(t, \mathbf{x}) = (-1)^m I_l^{-m}(t, \mathbf{x})$. Pour obtenir les modèles P_n en dimension deux, on supprime la dépendance en y afin de découpler la partie réelle et la partie imaginaire des équations, ce qui donne

$$\left\{ \begin{array}{l} \frac{1}{c}\partial_t I_l^m + \frac{1}{2}\partial_x \left(-C_{l-1}^{m-1} I_{l-1}^{m-1} + D_{l+1}^{m-1} I_{l+1}^{m-1} + E_{l-1}^{m+1} I_{l-1}^{m+1} - F_{l+1}^{m+1} I_{l+1}^{m+1}\right) \\ +\partial_z \left(A_{l-1}^m I_{l-1}^m + B_{l+1}^m I_{l+1}^m\right) - \sigma I_l^m = 0, \end{array} \right. \tag{1.11}$$

pour $m \neq 0$ et

$$\begin{cases} \frac{1}{c}\partial_t I_l^0 + \frac{1}{2}\partial_x \left(E_{l-1}^1 I_{l-1}^1 + F_{l+1}^1 I_{l+1}^1\right) \\ +\partial_z \left(A_{l-1}^0 I_{l-1}^0 + B_{l+1}^0 I_{l+1}^0\right) + \sigma \left(I_0^0 \delta_{l0} - I_l^0\right) = 0, \end{cases} \tag{1.12}$$

pour $m = 0$. Les coefficients associés aux modèles P_n sont définis par

$$A_l^m = \sqrt{\frac{(l-m+1)(l+m+1)}{(2l+3)(2l+1)}} \quad B_l^m = \sqrt{\frac{(l-m)(l+m)}{(2l+1)(2l-1)}}, \tag{1.13}$$

$$C_l^m = \sqrt{\frac{(l+m+1)(l+m+2)}{(2l+3)(2l+1)}} \quad D_l^m = \sqrt{\frac{(l-m)(l-m-1)}{(2l+1)(2l-1)}}, \tag{1.14}$$

$$E_l^m = \sqrt{\frac{(l-m+1)(l-m+2)}{(2l+3)(2l+1)}} \quad F_l^m = \sqrt{\frac{(l+m)(l+m-1)}{(2l+1)(2l-1)}}, \tag{1.15}$$

avec $A_{l-1}^m = B_l^m$, $C_l^m = F_{l+1}^{m+1}$ et $D_l^m = E_{l-1}^{m+1}$. Le système formé par les équations (1.11)-(1.12) n'est pas symétrique. En effet, la simplification proposée précédemment pour diminuer le nombre de variables rend le système non symétrique. Cependant, à l'aide d'un changement de variable trivial on peut obtenir un système symétrique. Soit $\tilde{I}_l^m$ les nouvelles inconnues définies par :

- $\tilde{I}_l^m = I_l^m$
- $\tilde{I}_l^m = -\sqrt{2} I_l^m$

Le système associé aux nouvelles inconnues $\tilde{I}_l^m$ est symétrique. Voici un tableau indiquant le nombre d'équations selon l'ordre de troncature n et la dimension d'espace. À partir de là il existe deux

dimension/ n	1	3	15	n
un	2	4	16	$n+1$
deux	3	10	136	$\frac{1}{2}n^2 + \frac{3}{2}n + 1$
trois	6	20	272	$n^2 + 3n + 2$

TABLE 1.1 – Nombre d'inconnues pour les modèles P_n en fonction de l'ordre de troncature et de la dimension

solutions pour classer les variables.
La première solution est de classer par l croissant, de l'harmonique la moins oscillante à l'harmonique la plus oscillante ce qui revient à classer les moments de l'équation de transport par ordre croissant (en effet $l = 0$ correspond au moment d'ordre 0, $l = 1$ au moment d'ordre 1). Ce classement a l'avantage de simplifier le passage du système P_n au système P_{n+1} car il suffit d'ajouter le nouveau moment. La seconde solution est de classer par m croissant, puisque les modes azimutaux m sont découplés des autres modes, ce classement permet d'obtenir une Jacobienne par rapport à z diagonale par bloc ou chaque bloc possède un spectre indépendant.
Les modèles P_n peuvent être également obtenus en minimisant l'entropie mathématique de l'équation de transport qu'est la norme L^2 sous contrainte de conservation des moments. Finalement, les modèles P_n peuvent être formalisés sous la forme de système linéaire hyperbolique symétrisable, appelé système de Friedrichs. Ces modèles sont invariants par rotation, ne préservent pas la positivité du premier moment $E(t, \mathbf{x})$ et de la reconstruction de $I(t, \mathbf{x}, \mathbf{\Omega}, \nu)$ à l'aide des moments et admettent la même limite de diffusion que l'équation de transport. Les démonstrations sur les limites asymptotiques des modèles seront apportées aux fur et à mesure de cette étude. D'autres solutions sont possibles, par

exemple Michael L. Hall utilise directement des harmoniques sphériques réelles contrairement à T. Brunner. Il est possible aussi d'utiliser à la place des harmoniques sphériques, un développement sur une base de tenseur symétrique bien choisi (voir [GC010]).

Remarque 1.2.1. *La base des harmoniques sphériques paraît adaptée à la discrétisation de l'équation de transport car les harmoniques sphériques forment les vecteurs propres de l'opérateur de collision* $\frac{1}{4\pi}\int_0^\infty \int_{S^2} p(\mathbf{\Omega},\mathbf{\Omega}')\left(I(t,\mathbf{x},\mathbf{\Omega}',\nu)-I(t,\mathbf{x},\mathbf{\Omega},\nu)\right)d\mathbf{\Omega}'d\nu$.

1.2.2 Approximations S_n (méthodes des ordonnées discrètes)

Les modèles S_n sont issus de la méthode des ordonnées discrètes [Chan60]- [JL91]- [BCLM02]. La stratégie consiste à discrétiser l'équation de transport mono-groupe (indépendante de la fréquence) à l'aide de vitesses discrètes et de résoudre le modèle formé par les équations qui dérivent de cette discrétisation. Dans le cas de l'opérateur de collision qui nous intéresse, la discrétisation en vitesse se fait à l'aide d'une formule de quadrature qui permet de discrétiser uniformément l'opérateur intégral. On choisit en général une discrétisation symétrique par rapport aux plans. Les systèmes construits ont l'avantage d'être diagonaux et de préserver la positivité de chaque inconnue mais ne sont pas invariants par rotation et génèrent ce qu'on appelle des effets de raies dégradant la solution numérique. Ces effets viennent de la nature discrète de l'approximation angulaire. Un exemple est donné dans [Bru02]. On pose $I_i(t,\mathbf{x}) = I_i(t,\mathbf{x},\mathbf{\Omega}_i)$ la fonction de distribution ponctuelle associée à la vitesse de quadrature $\mathbf{\Omega}_i$ et w_i le poids de quadrature associé. Les modèles S_n sont définis par

$$\frac{1}{c}\partial_t I_i + \mathbf{\Omega}_i.\nabla I_i = -\sigma\left(I_i - \sum_j w_j I_j\right). \tag{1.16}$$

Ces systèmes sont hyperboliques et symétriques (systèmes de Friedrichs). La formule de quadrature utilisée est construite de façon symétrique, ce qui se traduit par

$$\sum_i w_i = 1, \quad \sum_i w_i \mathbf{\Omega}_i = \mathbf{0}.$$

Contrairement aux modèles P_n, les modèles S_n ne donnent pas naturellement le coefficient de diffusion de l'équation de transport. En effet l'équation limite de diffusion est donnée par

$$\partial_t E - \operatorname{div}\left(\frac{\sum_i w_i \mathbf{\Omega}_i \otimes \mathbf{\Omega}_i}{\sigma}\nabla E\right) = 0,$$

avec $E = \sum_i w_i I_i$. Pour obtenir le coefficient de diffusion associé à l'équation de transport, la formule de quadrature est construite de façon à satisfaire

$$\sum_i w_i \mathbf{\Omega}_i \otimes \mathbf{\Omega}_i = \frac{1}{3}\widehat{I}_d.$$

1.2.3 Modèle M_1

Le modèle M_1 est un modèle hyperbolique non linéaire issu de l'équation de transfert radiatif (il n'est pas valide en neutronique). Il est obtenu en minimisant une entropie sous contrainte de la conservation des moments mais contrairement aux modèles P_n il s'agit de l'entropie physique associée à l'équation de transfert radiatif. Cette construction est détaillée dans [DF99].

Définition 1.2.1. *Le modèle M_1 est défini par*

$$\begin{cases} \dfrac{1}{c}\partial_t E + \operatorname{div} \mathbf{F} = 0, \\ \\ \dfrac{1}{c}\partial_t \mathbf{F} + \nabla \hat{P} = -\sigma \mathbf{F}, \end{cases} \tag{1.17}$$

avec le tenseur de pression suivant

$$\hat{P} = \frac{1}{2}\left((1-\chi(\mathbf{f}))\widehat{I}_d + (3\chi(\mathbf{f})-1)\frac{\mathbf{f}\otimes\mathbf{f}}{\| \mathbf{f} \|^2}\right) E,$$

le flux normalisé $\mathbf{f} = \frac{||\mathbf{F}||}{E}$ *et le tenseur d'Eddington* $\chi(\mathbf{f}) = \dfrac{3+4\mathbf{f}^2}{5+2\sqrt{4-3\mathbf{f}^2}}$.

Ce système conserve les principales propriétés liées à l'équation de transfert radiatif. Il conserve la positivité de E, la limitation du flux $||\mathbf{F}|| < cE$, et il est invariant par rotation. Il préserve aussi les régimes limites de l'équation de transport : le régime de diffusion ($\sigma >> 1$) et le régime libre (appelé aussi régime de transport pur), c'est à dire qu'il transporte exactement les solutions du type $\frac{||\mathbf{F}||}{cE} = 1$ quand $\sigma = 0$. En minimisant l'entropie des photons sous la contrainte de conserver les moments d'ordre supérieur on peut construire des systèmes plus grands appelés M_n. Cependant, dans ce cas il n'existe pas d'écriture analytique de tenseur d'Eddington (voir [MF11] pour le cas du M_2).

1.2.4 Extension multi-groupes

Une extension nommée multi-groupes existe pour le modèle M_1 et est aussi possible pour les modèles P_n ou S_n. En effet, les opacités (resp. les sections efficaces) dans la réalité sont dépendantes de la fréquence du photon (resp l'énergie des neutron). Afin d'obtenir une approximation en fréquence du problème de transport on définit des groupes de fréquence ou d'énergie et on dérive les modèles simplifiés de type P_n ou M_1 pour chaque groupe. Pour le transfert radiatif cela revient à définir des moments partiels à l'aide de la formule d'intégration

$$< . >_i = \int_{\nu_i}^{\nu_{i+1}} \int_{S^2} . d\mathbf{\Omega} d\nu,$$

avec $[\nu_i, \nu_{i+1}]$ un groupe de fréquence. On obtient ainsi un modèle simplifié pour chaque groupe qui sont couplés par des termes sources. Dans le cas du modèle M_1 ces extensions sont détaillées dans [Tur03], cependant le tenseur d'Eddington associé à chaque groupe n'admet pas une expression analytique. Dans cette étude nous ne traiterons pas du cas multi-groupes.

1.2.5 Autres modèles

Diffusion à flux limité

Les méthodes de flux limités consistent à résoudre l'équation de diffusion limite hors équilibre et à insérer dans cette équation un limiteur permettant de mieux approcher les régimes de transport notamment en récupérant la propriété $||\mathbf{f}|| = \frac{||\mathbf{F}||}{cE} < 1$. On résout donc

$$\partial_t E - \operatorname{div}\left(\frac{c\lambda}{3\sigma}\nabla E\right) = 0,$$

avec λ un limiteur dépendant d'un quantité R proportionnelle au gradient et qui vaut 1 dans le régime de diffusion. Par exemple le limiteur de Kershaw

$$\lambda = \frac{2}{3 + \sqrt{9 + 4R^2}}.$$

P_n fermé positivement

Les modèles PP_n ont été introduits récemment afin d'obtenir la positivité de la reconstruction de l'intensité $I(t, \mathbf{x}, \mathbf{\Omega}, \nu)$ (ou du flux neutronique) à partir de ses moments (voir [HMc11]). Le principe est simple, la construction des approximations P_n peut être obtenue en minimisant la norme L^2 sous la contrainte de conservation des moments. Pour obtenir la positivité on ajoute la contrainte de positivité sur l'intensité $I(t, \mathbf{x}, \mathbf{\Omega}, \nu)$ dans la minimisation de la fonctionnelle. De façon générale on ne peut imposer cette contrainte sur tout le domaine spatial, la méthode consiste donc à résoudre ce problème de minimisation avec la contrainte de positivité pour un nombre fini de points. Les auteurs proposent d'utiliser l'ensemble des points de quadrature associés aux modèles S_n. Lorsque les modèles P_n classiques sont positifs, les modèles PP_n dérivent exactement sur le modèle P_n. D'après les auteurs le régime de diffusion limite paraît difficile à étudier.

Modèles D_n

Il s'agit d'une correction des approximations P_n. Ces modèles ont été introduits dans [SFL11] pour le traitement des milieux semi-transparents. En effet les modèles P_n pour n grand sont très efficaces en milieu transparent mais très coûteux en temps de calcul. Quant aux approximations P_n pour n petit, ils n'approximent pas correctement l'équation de transport en milieu transparent. Les modèles D_n, d'après les auteurs, améliorent les résultats des modèles P_n pour n faible et sont très proches dans le cas des n grands.

P_n filtrés

Ces méthodes ont pour but d'obtenir des solutions positives et de diminuer le comportement ondulatoire des approximations P_n (voir [McH10]- [McHL10]), pour cela on ajoute un terme conservatif lors de la minimisation de la fonctionnelle. Ce terme force les coefficients à décroître quand l augmente afin de limiter l'aspect ondulatoire des modèles. Dans un certain sens l'ajout de ce terme revient à ajouter une viscosité artificielle dans le développement sur la base des harmoniques sphériques.

Modèle de moments partiels

La méthode des moments partiels peut s'appliquer au modèle M_1 comme aux approximations P_n. Les intégrations sur tout l'ensemble des vitesses peuvent poser problème. En effet, si on chauffe un milieu à gauche et à droite de la même façon, on crée un équilibre artificiel lorsque les deux ondes rentrent en contact parce qu'il n'existe qu'une seule manière d'obtenir le rayonnement à partir des moments E et $\mathbf{F}$, pour le modèle P_1 par exemple, [Tur03]. Le principe de construction est le même que pour les modèles multi-groupes, il suffit de définir des moments partiels en intégrant sur une portion de l'ensemble des vitesses la solution de l'équation de transport afin d'obtenir plusieurs modèles portant sur chaque moment partiel.

Conclusion

Tous ces modèles simplifiés offrent des alternatives aux méthodes de Monte Carlo pour la résolution des problèmes de neutronique ou de photonique. À présent on souhaite discrétiser ces modèles avec

des méthodes qui préservent le plus de propriétés physiques, notamment la limite asymptotique qui assure la validité du modèle pour tout type d'opacité.

1.3 Contexte numérique

Dans cette section, on propose de présenter quelques difficultés numériques issues de la discrétisation des systèmes hyperboliques avec termes sources raides admettant une limite de diffusion. Après un l'état de l'art sur les applications des schémas « asymptotic preserving » , on décrira le problème de consistance des schémas classiques en 1D ainsi que les méthodes « asymptotic preserving » construites pour être valides dans tous les régimes. Pour conclure on introduira les difficultés associées à l'extension sur maillages non structurés.

La résolution, sur maillages non structurés, des EDP précédemment citées, peut se faire à l'aide de plusieurs méthodes numériques : les volumes finis, les éléments finis ou encore les méthodes de Galerkin discontinues. Dans le cadre de cette thèse on se concentre sur la méthode des volumes finis [EGH00] qui possède l'avantage d'être moins coûteuse que les autres méthodes et qui est la plus naturelle pour la discrétisation des systèmes hyperboliques. On commence donc par un rappel du principe de cette méthode. Soit une EDP en dimension deux :

$$\partial_t \mathbf{U} + \partial_x f(\mathbf{U}, \nabla \mathbf{U}) + \partial_y g(\mathbf{U}, \nabla \mathbf{U}) = \mathbf{0}, \quad \mathbf{U} : D \in \mathbb{R}^2 \longrightarrow \mathbb{R}^n,$$

avec $f(\mathbf{U}, \nabla \mathbf{U})$, $g(\mathbf{U}, \nabla \mathbf{U})$ les flux. La première étape consiste à intégrer sur une maille l'équation afin d'utiliser le théorème de flux-divergence par conséquent

$$\frac{d}{dt} \int_{\Omega_j} \mathbf{U}(t,x,y) dxdy + \int_{\Omega_j} \partial_x f(\mathbf{U}, \nabla \mathbf{U}) + \partial_y g(\mathbf{U}, \nabla \mathbf{U}) dxdy = \mathbf{0},$$

est équivalent à

$$\frac{d}{dt} \int_{\Omega_j} \mathbf{U}(t,x,y) dxdy + \int_{\partial\Omega_j} f(\mathbf{U}, \nabla \mathbf{U}) n_j^x + g(\mathbf{U}, \nabla \mathbf{U}) n_j^y d\omega = \mathbf{0}.$$

avec $d\omega$ la mesure de surface. La méthode des volumes finis repose sur l'estimation de la quantité définie sur le bord de la maille qui peut être considérée comme l'approximation de la quantité d'informations échangées. On note $\mathbf{n}_{jk} = (n_{jk}^x, n_{jk}^y)$ la normale sortante associée à l'arête $\partial\Omega_{jk}$, l_{jk} la longueur de l'arête. En outre on définit notre inconnue $\mathbf{U}_j = \int_{\Omega_j} \mathbf{U}(t,x,y) dxdy$ la moyenne de la solution dans la maille, ce qui donne le schéma

$$|\Omega_j| \frac{\mathbf{U}_j^{n+1} - \mathbf{U}_j^n}{\Delta t} + \sum_k l_{jk} \left(f_{jk}^* n_j^x + g_{jk}^* n_j^y \right) = \mathbf{0}. \tag{1.18}$$

Afin de compléter ce dernier, il est nécessaire de construire le flux numérique $\left(f_{jk}^* n_j^x + g_{jk}^* n_j^y \right)$ qui approche correctement $f(\mathbf{U}, \nabla \mathbf{U}) n^x + g(\mathbf{U}, \nabla \mathbf{U}) n^y$. L'essentiel du travail porte sur la construction de ce flux numérique.
À présent on propose de rappeler quelques définitions et résultats standards utiles pour l'analyse des méthodes de différences finies et de volumes finis en 1D. On s'est inspiré directement du cours de [AG11] avec la réserve d'usage (remarque 2.2.5 du cours [AG11]). On se place dans le cadre d'une grille uniforme de pas Δx. Le pas de temps est noté Δt. Un schéma numérique pour l'EDP

$$F(\mathbf{U}) = \mathbf{0}, \text{ avec } \mathbf{U}(t,x) \in \mathbb{R}^+ \times \mathbb{R},$$

est donné pour tous j indice de maille et n indice de temps par

$$F_{\Delta t,\Delta x}\left(\left\{\mathbf{U}_{j+k}^{n+m}\right\}\right)=0, \tag{1.19}$$

avec $m^- \leq m \leq m^+$, $k^- \leq k \leq k^+$ où $m^\pm$, $k^\pm$ des entiers qui donnent le stencil du schéma.

Définition 1.3.1. *Un schéma numérique de type (1.19) est dit consistant au sens des différences finies si, pour toute solution $\mathbf{U}(t,x)$ régulière, l'erreur de consistance*

$$F_{\Delta t,\Delta x}\left(\{\mathbf{U}(t+m\Delta t, x+k\Delta x\}\right)$$

tend vers zéro. Le schéma est dit consistant à l'ordre p en espace et à l'ordre q en temps si l'erreur de consistance est homogène à $O(\Delta x^p+\Delta t^q)$.

Dans [AG11], les auteurs précisent que cette définition est un légèrement ambiguë. En effet si on multiplie l'erreur de troncature par Δx^p, on peut rendre toute discrétisation consistante. Afin de lever cette ambiguïté, on suppose que la formule $F_{\Delta t,\Delta x}\left(\{\mathbf{U}(t+m\Delta t, x+k\Delta x\}\right)$ est écrite de manière à ce que l'erreur de troncature ne soit pas nulle lorsque $\mathbf{U}(t,\mathbf{x})$ n'est pas solution de l'équation $F(\mathbf{U})=\mathbf{0}$.

Définition 1.3.2. *Soit $\mathbf{U}_h^n=\left\{\mathbf{U}_j^n,...\mathbf{U}_N^n\right\}\in\mathbb{R}^N$ la solution numérique au temps $t_n=n\Delta t$. Le schéma est stable dans la norme $||.||_E$ s'il existe une constante $K>0$ indépendante de Δt et Δx telle que*

$$||\mathbf{U}_h^n||_E \leq K||\mathbf{U}_h^0||_E.$$

Théorème 1.1 (Théorème de Lax)**.** *Soit $\mathbf{U}_h^n=\left\{\mathbf{U}_{j=1}^n,...\mathbf{U}_{j=N}^n\right\}$ la solution numérique au temps $t_n=n\Delta t$ et $\mathbf{U}(t,x)$ la solution exacte telle que $\mathbf{U}_j^0=\mathbf{U}(t=0,x_j=j\Delta x)$. Si un schéma numérique linéaire à deux niveaux pour le système d'EDP linéaire*

$$F(\mathbf{U})=\mathbf{0},$$

est stable pour la norme $||.||_E$ et consistant aux sens des différences finies avec l'erreur $O(\Delta x^p+\Delta t^q)$ alors le schéma converge vers la solution exacte avec l'estimation

$$\sup_{t_n\leq T}||\mathbf{U}_h^n-\mathbf{U}(t_n)||_E \leq C_T\left(\Delta x^p+\Delta t^q\right),$$

Remarque 1.3.1. *Un schéma à deux niveaux est un schéma qui utilise une discrétisation temporelle dépendante uniquement de t_{n+1} et t_n.*

La notion de consistance au sens des différences finies ainsi que le théorème de Lax sont valables pour des schémas volumes finis linéaires sur grille uniforme. En effet, dans ces cas les méthodes de différences finies et de volumes finis se confondent.

1.3.1 Schémas « asymptotic preserving » : introduction et applications

Les problèmes de transport énoncés dans le chapitre précédent peuvent être interprétés comme des EDP dépendantes d'un paramètre $\varepsilon>0$, (qui est issu d'une mise à l'échelle d'un paramètre physique comme l'opacité, le temps ou la vitesse de la lumière pour les problèmes de rayonnement). La présence de ce paramètre induit une limite asymptotique pour l'EDP considérée lorsqu'il tend vers zéro. Pour traiter ce type de problème S. Jin [Jin99]- [Jin11] a introduit la notion de schéma « asymptotic preserving » . On propose donc de rappeler deux définitions associées à cette notion.

Définition 1.3.3. *Un schéma pour le système hyperbolique (P_h^ε sur la figure (1.1)) est dit « asymptotic preserving » lorsqu'il tend vers un schéma limite (P_h^0) consistant avec le modèle limite (P^0) quand ε tend vers zéro avec h (typiquement h correspond au pas de maillage Δx ou de temps Δt) le paramètre de discrétisation fixé.*

Définition 1.3.4. *Un schéma est dit uniformément « asymptotic preserving » si l'estimation d'erreur et la condition CFL associées sont indépendantes de ε.*

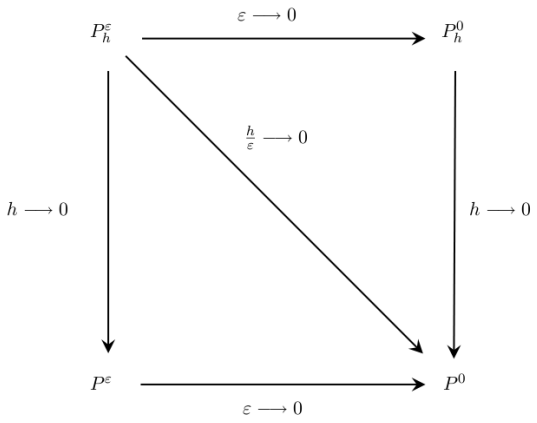


FIGURE 1.1 – Définition des schémas « asymptotic preserving » .

Plusieurs domaines de la physique contiennent des modèles dont la discrétisation nécessite la construction de schémas « asymptotic preserving » . Le premier considéré est celui du transfert radiatif. En effet, lorsque l'opacité de la matière est grande, l'équation de transport comme les modèles simplifiés tendent vers une équation de diffusion. Pour chacun de ces modèles un certain nombre de schémas ont été développés. L. Gosse, G. Toscani dans [GT01] et S. Jin, C. D. Levermore dans [JL96] ont développés des schémas « asymptotic preserving » basés sur l'introduction des termes sources dans les flux upwind pour l'équation de la chaleur hyperbolique :

$$\begin{cases} \partial_t E + \dfrac{1}{\varepsilon}\partial_x F = 0, \\ \\ \partial_t F + \dfrac{1}{\varepsilon}\partial_x E + \dfrac{\sigma}{\varepsilon} F = 0. \end{cases} \tag{1.20}$$

L'idée de S. Jin et D. Levermore consiste à incorporer les états stationnaires du modèle dans les flux upwind afin de ne pas ignorer les variations du gradient. Pour l'équation de la chaleur hyperbolique (1.20) cela consiste à incorporer, à l'aide d'un développement de Taylor, l'état stationnaire

$$\partial_x E = -\frac{\sigma}{\varepsilon} F.$$

La méthode de L. Gosse et G. Toscani propose de concentrer les termes sources aux interfaces, ce qui à pour conséquence de modifier le système car les masses de Dirac induisent une onde stationnaire dans le problème de Riemann. La solution du problème de Riemann est obtenue à l'aide de la solution stationnaire de l'équation de la chaleur hyperbolique. Cette méthode revient aussi à introduire les termes sources dans les flux. Ce type de méthode a été initialement utilisée par Greenberg et Leroux

[GL96] afin de construire des méthodes numériques « well - balanced » . M. Lemou et L. Mieussens dans [LM07] proposent une discrétisation de l'équation de transport linéaire basée sur une décomposition « micro-macro » couplée avec une discrétisation « décalée » en espace. Les auteurs commencent par décomposer la fonction de distribution de l'équation de transport entre un terme lié à la variable de diffusion E (partie macro) et un terme dépendant de ε (partie micro) :

$$f = aE + \varepsilon g$$

tel que le premier moment de g soit nul. En injectant cette décomposition dans l'équation de transport et en utilisant une projection sur l'espace des fonctions de distribution isotrope (indépendante de la vitesse) on obtient un système sur E et g. Afin d'obtenir le schéma, les auteurs discrétisent les termes de transport avec un schéma upwind et les autres termes avec un schéma centré sur la grille « décalée » . L. Gosse a aussi adapté ces méthodes à la résolution de l'équation de transport linéaire et au modèle BGK en utilisant le même principe de construction [GT04]- [Gos10] que pour l'équation de la chaleur hyperbolique. Dans [BCLM02] les auteurs proposent un schéma « asymptotic preserving » pour les approximations S_n avec un « scattering » de type Laplace-Beltrami. La discrétisation du modèle M_1 (1.17) à aussi été abordée à travers deux stratégies. La première consiste à utiliser une méthode de relaxation introduite dans [JX95], couplée au schéma de Gosse-Toscani (voir [BD06]- [BC07]). La seconde consiste à utiliser un schéma HLLC, modifié à l'aide des termes sources [BCT08]- [B010]. Plus récemment, il a été introduit, par exemple dans [CS10], une extension des schémas de L. Gosse aux modèles P_n dans laquelle les solutions stationnaires sont calculées de façon approchée. Des modèles similaires au modèle P_1 et à l'équation de transport existent en biologie. Par conséquent, les schémas précédents peuvent s'appliquer également à ces modèles, comme dans [Gos10]. On trouve ce type de problème également en mécanique des fluides, le problème le plus classique étant le modèle d'Euler avec friction. Ce modèle a été étudié par de N. Séguin, C. Coquel et leurs co-auteurs, [CGRS08]- [CCGRS010] ou par C. Enaux dans [En07]. Les équations de Saint-Venant avec friction ont elles aussi été traitées par C. Berthon, R. Turpault dans [BMT11]. Les schémas « asymptotic preserving » interviennent également en physique des plasmas. On peut citer notamment les travaux de P. Degond et co-auteurs [DDSV09]- [DGM07]- [DLV08]. Depuis peu, des schémas « asymptotic preserving » sont utilisés pour capturer les régimes fluides des équations cinétiques. Par exemple il y a les travaux de S. Jin et F. Filbet pour les équations de Boltzmann ou Landau [FJ10]- [JY11], ou les travaux de M. Lemou et ces co-auteurs pour les équations de Boltzmann ou de Vlasov [CL11]- [Lemou10].

Afin de mieux comprendre les schémas « asymptotic preserving » on propose d'étudier différentes discrétisations pour l'équation de la chaleur hyperbolique (1.20) introduite afin de pallier la vitesse de propagation infinie de l'équation de la chaleur. Ce système peut être considéré comme un système d'onde amortie. Nous supposerons que la solution est régulière et proche de l'état d'équilibre, c'est à dire que (E, F) solution de (1.20) satisfait

$$\|\partial_x^{m_1}\partial_t^{m_2}E\|_\infty \leq C_{m_1,m_2} \text{ et } \|\partial_x^{m_1}\partial_t^{m_2}F\|_\infty \leq \varepsilon C_{m_1,m_2},$$

avec $C_{m_1,m_2} > 0$ une constante, tel que $(m_1, m_2) \in \mathbb{N}^2$. Ces estimations peuvent être obtenues à partir de l'étude du problème de Cauchy associé à l'équation (1.20). L'estimation sur F est équivalente au fait que $\tilde{F} = \frac{F}{\varepsilon}$ satisfait $\|\partial_x^{m_1}\partial_t^{m_2}\tilde{F}\|_\infty \leq C_{m_1,m_2}$. Comme il a été rappelé plus haut, ce modèle peut être approximé par le régime de diffusion suivant quand ε tend vers 0 :

$$\partial_t E - \partial_x\left(\frac{1}{\sigma}\partial_x E\right) = 0. \tag{1.21}$$

Pour simplifier les notations nous supposons dans ce chapitre σ constant en espace sauf dans les sections sur le cas variable.

1.3.2 Schémas « asymptotic preserving » « centré » ou « décalé »

On commence par étudier le comportement des schémas dits « centré » ou « décalé » pour l'équation de la chaleur hyperbolique. On introduit le pas d'espace Δx, le pas de temps Δt, on note $x_j = j\Delta x$, $t_n = n\Delta t$ avec j et t des entiers.

Définition 1.3.5. *Le schéma explicite « centré » pour le modèle (1.20) est défini par*

$$\begin{cases} \dfrac{E_j^{n+1} - E_j^n}{\Delta t} + \dfrac{1}{\varepsilon}\dfrac{F_{j+1}^n - F_{j-1}^n}{2\Delta x} = 0, \\ \dfrac{F_j^{n+1} - F_j^n}{\Delta t} + \dfrac{1}{\varepsilon}\dfrac{E_{j+1}^n - E_{j-1}^n}{2\Delta x} = -\dfrac{\sigma}{\varepsilon^2}F_j^n. \end{cases} \tag{1.22}$$

On commence par étudier la consistance du schéma « centré » pour l'équation de la chaleur hyperbolique.

Lemme 1.2. *L'erreur de consistance du schéma (1.22) pour la première équation est* $O\left(\Delta x^2 + \Delta t\right)$. *L'erreur de consistance pour la seconde equation est* $O\left(\frac{\Delta x^2}{\varepsilon} + \Delta t\right)$.

Démonstration. L'erreur de consistance est étudiée en utilisant des développements de Taylor. L'erreur de consistance cE pour la première équation est défini par

$$cE_j^n = \frac{E(x_j, t_{n+1}) - E(x_j, t_n)}{\Delta t} + \frac{F(x_{j+1}, t_n) - F(x_{j-1}, t_n)}{2\varepsilon\Delta x}$$

$$= (\partial_t E(x_j, t_n) + O(\Delta t)) + \left(\frac{\partial_x F(x_j, t_n)}{\varepsilon} + \frac{\Delta x^2}{\varepsilon} O(\partial_{xx} F(x_j, t_n))\right) = O\left(\Delta x^2 + \Delta t\right).$$

Pour obtenir ce résultat on a utilisé $\| \partial_t^{m_1} \partial_x^{m_2} F \| \leq \varepsilon$. L'erreur de consistance cF_j^n pour la deuxième équation est

$$cF_j^n = \frac{F(x_j, t_{n+1}) - F(x_j, t_n)}{\Delta t} + \frac{E(x_{j+1}, t_n) - E(x_{j-1}, t_n)}{2\varepsilon\Delta x} + \frac{\sigma}{\varepsilon^2} F(x_j, t_n)$$

$$= (\partial_t F(x_j, t_n) + O(\Delta t)) + \left(\frac{\partial_x E(x_j, t_n)}{\varepsilon} + \frac{\Delta x^2}{\varepsilon} O(\partial_{xx} E(x_j, t_n))\right) + \frac{\sigma}{\varepsilon^2} F(x_j, t_n).$$

On utilise le fait que $\partial_t F(x_j, t_n) = O(\varepsilon)$. Par conséquent

$$cF^n = O\left(\frac{\Delta x^2}{\varepsilon} + \Delta t\right).$$

□

Le schéma « centré » est consistant pour la première équation qui donne l'équation de diffusion limite. Le schéma capture donc correctement la limite de diffusion. Cependant ce schéma souffre de plusieurs défauts. D'abord le schéma est inconditionnellement instable [Al05], ne préserve pas le

principe du maximum $E+F>0$, $E-F>0$ associé à l'équation et provoque des oscillations en présence de solutions discontinues. En outre, le schéma limite donné par la formule

$$\frac{E_j^{n+1}-E_j^n}{\Delta t}-\frac{E_{j+2}^n-2E_j^n+E_{j-2}^n}{4\Delta x^2}=0,$$

admet des modes parasites qui génèrent aussi des oscillations (voir [BCLM02]). Dans le cadre du couplage avec l'hydrodynamique, on ne souhaite pas utiliser de schéma « centré » à cause de son mauvais comportement dans les régimes de transport et à cause du caractère oscillant du schéma de diffusion limite. Une seconde solution est d'utiliser un schéma dit « décalé » . Ce type de schéma consiste à localiser une partie des inconnues aux centres $x_j=j\Delta x$ des mailles Ω_j et une autre partie des inconnues aux interfaces $x_{j+\frac{1}{2}}=(j+\frac{1}{2})\Delta x$. La discrétisation temporelle est définie par $t_n=n\Delta t$.

Définition 1.3.6. *Le schéma « décalé » est défini par*

$$\begin{cases} \dfrac{E_j^{n+1}-E_j^n}{\Delta t}+\dfrac{1}{\varepsilon}\dfrac{F_{j+\frac{1}{2}}^n-F_{j-\frac{1}{2}}^n}{\Delta x}=0, \\ \dfrac{F_{j+\frac{1}{2}}^{n+1}-F_{j+\frac{1}{2}}^n}{\Delta t}+\dfrac{1}{\varepsilon}\dfrac{E_{j+1}^n-E_j^n}{\Delta x}=-\dfrac{\sigma}{\varepsilon^2}F_{j+\frac{1}{2}}^n. \end{cases} \tag{1.23}$$

Lemme 1.3. *L'erreur de consistance du schéma (1.23) pour la première équation est $O(\Delta x+\Delta t)$. L'erreur de consistance pour la seconde équation est $O\left(\frac{\Delta x}{\varepsilon}+\Delta t\right)$.*

Démonstration. De la même façon que pour des différences finies classiques, l'erreur de consistance est étudiée en utilisant des développements de Taylor. L'erreur de consistance cE pour la première équation est

$$cE_j^n=\frac{E(x_j,t_{n+1})-E(x_j,t_n)}{\Delta t}+\frac{F(x_{j+\frac{1}{2}},t_n)-F(x_{j-\frac{1}{2}},t_n)}{\varepsilon\Delta x}$$
$$=(\partial_t E(x_j,t_n)+O(\Delta t))+\left(\frac{\partial_x F(x_j,t_n)}{\varepsilon}+\frac{\Delta x}{\varepsilon}O(\partial_{xx}F(x_j,t_n))\right)=O(\Delta x+\Delta t).$$

L'erreur de consistance $cF_{j+\frac{1}{2}}^n$ pour la deuxième équation est

$$cF_{j+\frac{1}{2}}^n=\frac{F(x_{j+\frac{1}{2}},t_{n+1})-F(x_{j+\frac{1}{2}},t_n)}{\Delta t}+\frac{E(x_{j+1},t_n)-E(x_j,t_n)}{\varepsilon\Delta x}+\frac{\sigma}{\varepsilon^2}F(x_{j+\frac{1}{2}},t_n)$$
$$=\Big(\partial_t F(x_{j+\frac{1}{2}},t_n)+O(\Delta t)\Big)+\left(\frac{\partial_x E(x_{j+\frac{1}{2}},t_n)}{\varepsilon}+\frac{\Delta x}{\varepsilon}O(\partial_{xx}E(x_{j+\frac{1}{2}},t_n))\right)+\frac{\sigma}{\varepsilon^2}F(x_{j+\frac{1}{2}},t_n).$$

On utilise le fait que $\partial_{xx}F(x_{j+\frac{1}{2}},t_n)=O(\varepsilon)$. Par conséquent

$$cF_j^n=O\left(\frac{\Delta x}{\varepsilon}+\Delta t\right).$$

□

Contrairement au schéma « centré » le schéma limite associé au schéma « décalé » n'exhibe pas de modes parasites, on dérive en effet sur le schéma classique à trois points. Les discrétisations « décalé », ou « centré » avec une discrétisation implicite sont donc « asymptotic preserving » au sens de la définition (1.3.3). Cependant les schémas « décalé » et « centré » souffrent de défauts de stabilité et

n'ont pas un bon comportement en régime de transport (notamment en présence de discontinuité). Dans le cadre de l'hydrodynamique radiative une bonne approximation du régime de transport ($\sigma \to 0$, $\varepsilon \approx 1$) est aussi importante qu'une bonne approximation du régime de diffusion. En outre les équations de l'hydrodynamique étant approximées, dans notre cas, avec une discrétisation où les inconnues sont localisées aux centres des mailles, on souhaiterait obtenir des schémas « asymptotic preserving » , « cell-centered » basés sur le schéma upwind.

1.3.3 Schémas « asymptotic preserving » basés sur le schéma upwind

Les schémas « décalés » (1.23) et « centrés » (1.22) sont valides dans le régime de diffusion, mais ont des mauvaises propriétés dans le régime de transport ($\sigma \to 0$, $\varepsilon \approx 1$). Par conséquent on souhaite étudier les schémas « décentrés » qui sont connus pour avoir de bonnes propriétés en régime de transport. Pour les schémas de type Godunov comme les schémas Rusanov, upwind, HLL, HLLC [Tor28]), la situation est différente. Ces schémas ont un bon comportement en régime de transport (principe du maximum préservé, stabilité, bon comportement en présence de discontinuité), mais l'erreur de consistance et la condition CFL de stabilité sont dépendantes de ε. Ceci a pour conséquence que ces schémas ne convergent pas lorsque ε tend vers 0 et ont un pas de temps contraint par ε lorsqu'on utilise des discrétisations purement explicites en temps. Cependant, on peut les modifier afin d'obtenir des schémas convergents et stables indépendamment de ε en régime de diffusion tout en préservant les propriétés du schéma dans le régime de transport. Pour l'approximation en temps l'utilisation d'un schéma inconditionnellement stable (implicite en général) suffit à régler le problème. Par contre, obtenir un schéma semi-implicite avec une CFL indépendante de ε et obtenir la convergence pour tous les régimes, nécessite un travail supplémentaire.

Maintenant on propose de rappeler et d'étudier le schéma upwind ainsi que deux schémas « asymptotic preserving » basés sur ce même schéma.

Le schéma upwind

Le schéma upwind est un schéma de volumes finis qui calcule les flux aux interfaces $x_{j+\frac{1}{2}}$ en diagonalisant le système et en écrivant un schéma « décentré » pour chaque équation d'advection obtenue. On introduit le pas d'espace Δx, le pas de temps Δt, on note $x_j = j\Delta x$, $t_n = n\Delta t$ avec j et t des entiers. Le schéma s'écrit

$$\left\{ \begin{array}{l} \dfrac{E_j^{n+1} - E_j^n}{\Delta t} + \dfrac{F_{j+\frac{1}{2}}^n - F_{j-\frac{1}{2}}^n}{\varepsilon \Delta x} = 0, \\ \dfrac{F_j^{n+1} - F_j^n}{\Delta t} + \dfrac{E_{j+\frac{1}{2}}^n - E_{j-\frac{1}{2}}^n}{\varepsilon \Delta x} + \dfrac{\sigma}{\varepsilon^2} F_j^n = 0. \end{array} \right. \tag{1.24}$$

Les flux sont donnés par la solution du système linéaire suivante, système obtenu en décentrant les deux invariants de Riemann associés à l'équation des ondes :

$$\left\{ \begin{array}{l} F_j^n + E_j^n = F_{j+\frac{1}{2}}^n + E_{j+\frac{1}{2}}^n, \\ F_{j+1}^n - E_{j+1}^n = F_{j+\frac{1}{2}}^n - E_{j+\frac{1}{2}}^n, \end{array} \right. \iff \left\{ \begin{array}{l} F_{j+\frac{1}{2}}^n = \frac{1}{2}\left(F_j^n + F_{j+1}^n + E_j^n - E_{j+1}^n\right), \\ E_{j+\frac{1}{2}}^n = \frac{1}{2}\left(E_j^n + E_{j+1}^n + F_j^n - F_{j+1}^n\right). \end{array} \right. \tag{1.25}$$

En incorporant les flux dans (1.24) on obtient la forme explicite du schéma

$$\left\{ \begin{array}{l} \dfrac{E_j^{n+1} - E_j^n}{\Delta t} + \dfrac{F_{j+1}^n - F_{j-1}^n}{2\varepsilon \Delta x} - \dfrac{E_{j+1}^n - 2E_j^n + E_{j-1}^n}{2\varepsilon \Delta x} = 0, \\ \dfrac{F_j^{n+1} - F_j^n}{\Delta t} + \dfrac{E_{j+1}^n - E_{j-1}^n}{2\varepsilon \Delta x} - \dfrac{F_{j+1}^n - 2F_j^n + F_{j-1}^n}{2\varepsilon \Delta x} + \dfrac{\sigma}{\varepsilon^2} F_j = 0. \end{array} \right. \tag{1.26}$$

Pour chaque équation le troisième terme est appelé viscosité numérique ou diffusion numérique. Ce terme stabilise la discrétisation « centrée » des dérivées premières. Maintenant que le schéma est introduit on va pouvoir l'étudier.

Consistance du schéma upwind

Lemme 1.4. *L'erreur de consistance du schéma upwind (1.26) est homogène à* $O\left(\frac{\Delta x}{\varepsilon}+\Delta t\right)$..

Démonstration. De la même façon que pour des différences finies classiques l'erreur de consistance est étudiée en utilisant des développements de Taylor. On note $x_j = j\Delta x$ et $t_n = n\Delta t$. L'erreur de consistance cE pour la première équation est

$$cE_j^n = \frac{E(x_j,t_{n+1})-E(x_j,t_n)}{\Delta t} + \frac{F(x_{j+1},t_n)-F(x_{j-1},t_n)}{2\varepsilon\Delta x} - \frac{E(x_{j+1},t_n)-2E(x_j,t_n)+E(x_{j-1},t_n)}{2\varepsilon\Delta x}$$

$$= (\partial_t E(x_j,t_n)+O(\Delta t)) + \left(\frac{\partial_x F(x_j,t_n)}{\varepsilon}+O(\Delta x^2)\right) + O\left(\frac{\Delta x}{\varepsilon}\right)\partial_{xx}E(x_j,t_n) = O\left(\frac{\Delta x}{\varepsilon}+\Delta t\right).$$

L'erreur de consistance cF pour la deuxième équation est

$$cF_j^n = \frac{F(x_j,t_{n+1})-F(x_j,t_n)}{\Delta t} + \frac{E(x_{j+1},t_n)-E(x_{j-1},t_n)}{2\varepsilon\Delta x}$$

$$-\frac{F(x_{j+1},t_n)-2F(x_j,t_n)+F(x_j,t_n)}{2\varepsilon\Delta x} + \frac{\sigma}{\varepsilon^2}F(x_j,t_n)$$

$$= (\partial_t F(x_j,t_n)+O(\Delta t)) + \left(\frac{\partial_x E(x_j,t_n)}{\varepsilon}+O(\frac{\Delta x^2}{\varepsilon})\right) + O\left(\frac{\Delta x}{\varepsilon}\right)\partial_{xx}F(x_j,t_n) + \frac{\sigma}{\varepsilon^2}F(x_j,t_n).$$

On utilise le fait que par hypothèse $\partial_{xx}F(x_j,t_n) = O(\varepsilon)$. Par conséquent

$$cF_j^n = O\left(\frac{\Delta x^2}{\varepsilon}+\Delta x+\Delta t\right).$$

□

L'erreur de convergence associée au schéma upwind (1.26) admet donc une erreur de consistance importante lorsque le maillage est grossier ($\frac{\Delta x}{\varepsilon} >> 1$).

Stabilité du schéma upwind

Lemme 1.5. *Le schéma upwind (1.26) satisfait le principe du maximum pour les invariants de Riemann* $u = E+F$ *et* $v = E-F$ *sous la CFL*

$$\frac{\Delta t}{\varepsilon\Delta x} + \frac{\sigma\Delta t}{2\varepsilon^2} \leq 1. \tag{1.27}$$

Démonstration. On commence par définir les invariants de Riemann $u = E+F$ et $v = E-F$, le système (1.26) se réécrit sous la forme suivante

$$\begin{cases} \dfrac{u_j^{n+1}-u_j^n}{\Delta t} + \dfrac{u_j^n-u_{j-1}^n}{\varepsilon\Delta x} + \dfrac{\sigma}{2\varepsilon^2}\left(u_j^n-v_j^n\right) = 0, \\ \dfrac{v_j^{n+1}-v_j^n}{\Delta t} - \dfrac{v_{j+1}^n-v_j^n}{\varepsilon\Delta x} + \dfrac{\sigma}{2\varepsilon^2}\left(-u_j^n+v_j^n\right) = 0, \end{cases} \tag{1.28}$$

par conséquent

$$\begin{cases} u_j^{n+1} = \left(1 - \dfrac{\Delta t}{\varepsilon \Delta x} - \dfrac{\sigma \Delta t}{2\varepsilon^2}\right) u_j^n + \dfrac{\Delta t}{\varepsilon \Delta x} u_{j-1}^n + \dfrac{\sigma \Delta t}{2\varepsilon^2} v_j^n, \\ v_j^{n+1} = \left(1 - \dfrac{\Delta t}{\varepsilon \Delta x} - \dfrac{\sigma \Delta t}{2\varepsilon^2}\right) v_j^n + \dfrac{\Delta t}{\varepsilon \Delta x} v_{j+1}^n + \dfrac{\sigma \Delta t}{2\varepsilon^2} u_j^n. \end{cases} \tag{1.29}$$

Les termes extra-diagonaux sont positifs, les termes diagonaux sont aussi positifs sous la condition $1 - \frac{\Delta t}{\varepsilon \Delta x} - \frac{\sigma \Delta t}{2\varepsilon^2} \geq 0$. Pour finir on remarque que u_j^{n+1} et v_j^{n+1} sont calculés à partir de combinaison convexe des quantités au temps n. Par conséquent le schéma satisfait donc la principe du maximum sous la condition CFL (1.27). □

Remarque 1.3.2. *Le principe du maximum implique la stabilité en norme L^p pour $1 \leq p \leq \infty$ et pour un domaine borné.*

Stabilité du schéma upwind semi-implicite

On souhaite utiliser des schémas semi-implicites qui ne nécessitent qu'une inversion locale afin d'obtenir une CFL moins contraignante. Pour le modèle (1.20) la discrétisation semi-implicite consiste uniquement à impliciter le terme source.

Lemme 1.6. *Le schéma upwind semi-implicite satisfait le principe du maximum sur les invariants de Riemann $u = E + F$, $v = E - F$ sous la CFL*

$$\frac{\Delta t}{\varepsilon \Delta x} \leq 1. \tag{1.30}$$

Démonstration. On part du schéma (1.51) avec une discrétisation implicite du terme source. On obtient

$$\begin{cases} -\dfrac{\sigma \Delta t}{2\varepsilon^2} v_j^{n+1} + \left(1 + \dfrac{\sigma \Delta t}{2\varepsilon^2}\right) u_j^{n+1} = \dfrac{\Delta t}{\Delta x \varepsilon} u_{j-1} + (1 - \dfrac{\Delta t}{\varepsilon \Delta x}) u_j, \\ -\dfrac{\sigma \Delta t}{2\varepsilon^2} u_j^{n+1} + \left(1 + \dfrac{\sigma \Delta t}{2\varepsilon^2}\right) v_j^{n+1} = \dfrac{\Delta t}{\Delta x \varepsilon} v_{j-1} + (1 - \dfrac{\Delta t}{\varepsilon \Delta x}) v_j. \end{cases} \tag{1.31}$$

On commence par inverser le système local par maille. La matrice, qui est une M-matrice, est donnée par

$$\begin{pmatrix} 1 + \frac{\sigma \Delta t}{2\varepsilon^2} & -\frac{\sigma \Delta t}{2\varepsilon^2} \\ -\frac{\sigma \Delta t}{2\varepsilon^2} & 1 + \frac{\sigma \Delta t}{2\varepsilon^2} \end{pmatrix}. \tag{1.32}$$

Un rapide calcul montre que les quatre coefficients issus de la résolution du système linéaire sont positifs sous la condition CFL (1.30) et que leur somme fait 1. On obtient donc une combinaison convexe et par conséquent le principe du maximum est préservé. □

On remarque que cette manipulation permet d'obtenir une CFL moins restrictive, notamment lorsque σ est grand. Par le théorème de Lax on sait que le schéma upwind converge avec un erreur de l'ordre de $\frac{\Delta x}{\varepsilon}$. En effet, contrairement aux schémas « centré » et « décalé » , le schéma upwind admet une viscosité numérique (ou diffusion numérique) qui domine par rapport à la diffusion physique (limite asymptotique de diffusion) lorsque $\frac{\Delta x}{\varepsilon} >> 1$. Pour obtenir une erreur satisfaisante, il faut choisir $\Delta x << \varepsilon$ (par exemple $\Delta x = \varepsilon^2$). Par conséquent le schéma est inutilisable en pratique. Le problème est le même pour les schémas HLL et Rusanov [Tor28]. Pour régler le problème de la discrétisation en temps il existe deux solutions : utiliser un schéma purement implicite ou trouver une discrétisation semi-implicite capable de lever la condition CFL en ε. On constate donc que le principal problème vient de la discrétisation en espace. C'est l'objet des schémas que nous allons étudier à présent.

Schéma de Jin-Levermore

Principe de construction

Le schéma de Jin-Levermore [JL96] est basé sur une modification des flux du schéma standard (1.24) qui permet au schéma (1.24) de devenir « asymptotic preserving » au sens de la définition (1.3.3). Sa construction est basée sur la constatation qu'une méthode classique de volumes finis, constante par maille, ignore les variations qui peuvent être fortes du gradient. L'idée de Jin-Levermore est d'incorporer dans le flux les états stationnaires du système afin de pouvoir capter les fortes variations de $\partial_x E$ et le régime de diffusion limite. On commence par écrire l'approximation de Taylor à l'ordre un

$$\begin{cases} E(x_j) \simeq E(x_{j+\frac{1}{2}}) - \frac{\Delta x}{2}\partial_x E(x_{j+\frac{1}{2}}), \\ E(x_{j+1}) \simeq E(x_{j+\frac{1}{2}}) + \frac{\Delta x}{2}\partial_x E(x_{j+\frac{1}{2}}). \end{cases} \tag{1.33}$$

En utilisant l'état stationnaire $\partial_x E = -\frac{\sigma}{\varepsilon}F$ avec la version discrète des équations (1.33) et en combinant cela aux flux (1.25) du schéma upwind on obtient

$$\begin{cases} F_j^n + E_j^n = F_{j+\frac{1}{2}}^n + E_{j+\frac{1}{2}}^n + \dfrac{\sigma\Delta x}{2\varepsilon}F_{j+\frac{1}{2}}^n, \\ F_{j+1}^n - E_{j+1}^n = F_{j+\frac{1}{2}}^n - E_{j+\frac{1}{2}}^n + \dfrac{\sigma\Delta x}{2\varepsilon}F_{j+\frac{1}{2}}^n, \end{cases} \iff \begin{cases} F_{j+\frac{1}{2}}^n = \dfrac{M}{2}\left(F_j^n + F_{j+1}^n + E_j^n - E_{j+1}^n\right), \\ E_{j+\frac{1}{2}}^n = \dfrac{1}{2}\left(E_j^n + E_{j+1}^n + F_j^n - F_{j+1}^n\right), \end{cases} \tag{1.34}$$

où le coefficient M est donné par

$$M = \frac{2\varepsilon}{2\varepsilon + \sigma\Delta x}.$$

Le schéma se réécrit sous la forme

$$\begin{cases} \dfrac{E_j^{n+1} - E_j^n}{\Delta t} + M\dfrac{F_{j+1}^n - F_{j-1}^n}{2\varepsilon\Delta x} - M\dfrac{E_{j+1}^n - 2E_j^n + E_{j-1}^n}{2\varepsilon\Delta x} = 0, \\ \dfrac{F_j^{n+1} - F_j^n}{\Delta t} + \dfrac{E_{j+1}^n - E_{j-1}^n}{2\varepsilon\Delta x} - \dfrac{F_{j+1}^n - 2F_j^n + F_{j-1}^n}{2\varepsilon\Delta x} + \dfrac{\sigma}{\varepsilon^2}F_j = 0. \end{cases} \tag{1.35}$$

Consistance du schéma de Jin-Levermore

Lemme 1.7. *L'erreur de consistance associée à la deuxième équation du schéma de Jin-Levermore (1.35) est $O\left(\frac{\Delta x}{\varepsilon} + \Delta t\right)$, et celle associée à la première équation est $O\left(\Delta x^2 + \varepsilon\Delta x + \Delta t\right)$.*

Démonstration. À présent on passe à l'étude de l'erreur de consistance du schéma (1.35) en utilisant les mêmes outils que pour la démonstration précédente. On voit que la seconde équation de (1.35) n'est pas modifiée par conséquent l'erreur de consistance reste en $O\left(\frac{\Delta x}{\varepsilon} + \Delta t\right)$. On définit l'erreur cE_j^n associée à la première équation du schéma (1.35) est

$$\begin{aligned} cE_j^n &= \frac{E(x_j, t_{n+1}) - E(x_j, t_n)}{\Delta t} + M\frac{F(x_{j+1}, t_n) - F(x_{j-1}, t_n)}{2\varepsilon\Delta x} - M\frac{E(x_{j+1}, t_n) - 2E(x_j, t_n) + E(x_{j-1}, t_n)}{2\varepsilon\Delta x} \\ &= \partial_t E(x_j, t_n) + O(\Delta t) + M\frac{2\Delta x\partial_x F(x_j, t_n) + O(\varepsilon\Delta x^3)}{2\varepsilon\Delta x} - M\frac{\Delta x^2\partial_{xx}E(x_j, t_n) + O(\Delta x^4)}{2\varepsilon\Delta x} \\ &= \left(\partial_t E(x_j, t_n) + \frac{M}{\varepsilon}\partial_x F(x_j, t_n) - M\frac{\Delta x}{2\varepsilon}\partial_{xx}E(x_j, t_n)\right) + O\left(\Delta x^2 + \frac{M\Delta x^3}{\varepsilon} + \Delta t\right). \end{aligned}$$

On note que $M\frac{\Delta x^3}{\varepsilon} = O\left(\Delta x^2\right)$. Puisque (E, F) est la solution exacte de (1.20) alors

$$\partial_{xx}E = -\frac{\sigma}{\varepsilon}\partial_x F - \varepsilon\partial_{xt}F = -\frac{\sigma}{\varepsilon}\partial_x F + O(\varepsilon^2).$$

le terme dans la première parenthèse s'écrit aussi

$$(\cdots) = \partial_t E(x_j, t_n) + \frac{M}{\varepsilon}\partial_x F(x_j, t_n) + M\frac{\Delta x}{2\varepsilon}\frac{\sigma}{\varepsilon}\partial_x F(x_j, t_n) + O\left(\frac{M\varepsilon^2\Delta x}{2\varepsilon}\right)$$

$$= \frac{1}{\varepsilon}\left(-1 + M + M\frac{\Delta x}{2}\frac{\sigma}{\varepsilon}\right)\partial_x F(x_j, t_n) + O\left(M\frac{\varepsilon^2\Delta x}{2\varepsilon}\right).$$

Par définition de M le terme entre parenthèse est $-1 + M + \frac{M\Delta x\sigma}{2\varepsilon} = 0$, donc $cE_j^n = O(\Delta x^2 + \varepsilon\Delta x + \Delta t)$. □

Stabilité du schéma de Jin-Levermore

Lemme 1.8. *Le schéma de Jin-Levermore (1.24)-(1.34) ne satisfait pas le principe du maximum pour les invariants de Riemann $u = E + F$ et $v = E - F$.*

Démonstration. On commence par analyser la stabilité au sens du principe du maximum. On somme les deux équations, cela donne

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{\left(u_j^n - aF_{j+\frac{1}{2}}^n\right) - \left(u_{j-1}^n - aF_{j-\frac{1}{2}}^n\right)}{\varepsilon\Delta x} + \frac{\sigma}{2\varepsilon^2}\left(u_j^n - v_j^n\right) = 0,$$

avec $a = \frac{\Delta x\sigma}{2\varepsilon}$. Puisque $F_{j+\frac{1}{2}}^n = \dfrac{u_j^n - v_{j+1}^n}{2(1+a)}$, on obtient

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \left(1 - \frac{a}{2(1+a)}\right)\frac{u_j^n - u_{j-1}^n}{\varepsilon\Delta x} + \frac{a}{2(1+a)\varepsilon\Delta x}\left(v_{j+1}^n - v_j^n\right) + \frac{\sigma}{2\varepsilon^2}\left(u_j^n - v_j^n\right) = 0$$

on isole les termes en $n+1$,

$$u_j^{n+1} = \left(1 - \frac{(2+a)\Delta t}{2(1+a)\varepsilon\Delta x} - \frac{\sigma\Delta t}{\varepsilon^2}\right)u_j^n + \frac{\sigma\Delta t}{4(1+a)\varepsilon^2}u_{j-1}^n + \left(\frac{a\Delta t}{2(1+a)\varepsilon\Delta x} + \frac{\sigma\Delta t}{2\varepsilon^2}\right)v_j^n - \frac{a\Delta t}{\varepsilon\Delta x}v_{j+1}^n.$$

Sous une certaine CFL on peut garantir la positivité du coefficient diagonal, mais par contre il n'est pas possible de garantir la positivité de tous les coefficients extra-diagonaux, en particulier $-\frac{a\Delta t}{\varepsilon\Delta x} < 0$. Pour cette raison le schéma ne satisfait pas le principe du maximum. □

Remarque 1.3.3. *Sous condition de stabilité en norme L^2 le schéma de Jin-Levermore converge et est « asymptotic preserving » au sens de la définition (1.3.3).*

Remarque 1.3.4. *L'erreur de consistance du schéma de Jin-Levermore $O(\Delta x^2 + \varepsilon\Delta x + \Delta t)$ montre que pour ε fixé, le schéma est d'ordre deux sur maillage grossier $\Delta x >> \varepsilon$ et d'ordre un pour un maillage fin $\Delta x << \varepsilon$. Par conséquent on s'attend à obtenir l'ordre deux lorsque ε est plus petit que Δx. Lorsque le pas de maillage s'approche de ε, on s'attend à ce que le schéma retombe sur l'ordre 1.*

Schéma pour un σ variable

Afin de comparer les différents schémas proposés on introduit les schémas pour les paramètres σ variables. En effet certaines méthodes vont donner des schémas non conservatifs dans ce cas. Lorsque σ est variable on obtient

$$\begin{cases} \dfrac{E_j^{n+1}-E_j^n}{\Delta t}+\dfrac{F_{j+\frac{1}{2}}^n-F_{j+\frac{1}{2}}^n}{\varepsilon\Delta x}=0, \\ \dfrac{F_j^{n+1}-F_j^n}{\Delta t}+\dfrac{E_{j+\frac{1}{2}}^n-E_{j+\frac{1}{2}}^n}{\varepsilon\Delta x}+\dfrac{\sigma_j}{\varepsilon^2}F_j^n=0. \end{cases} \tag{1.36}$$

avec

$$\begin{cases} E_{j+\frac{1}{2}}^n=\left(\frac{E_j^n+E_{j+1}^n}{2}+\frac{F_j^n-F_{j+1}^n}{2}\right) \\ F_{j+\frac{1}{2}}^n=M_{j+\frac{1}{2}}\left(\frac{F_j^n+F_{j+1}^n}{2}+\frac{E_j^n-E_{j+1}^n}{2}\right) \end{cases}$$

et $M_{j+\frac{1}{2}}=\frac{2\varepsilon}{2\varepsilon+\sigma_{j+\frac{1}{2}}\Delta x}$.

Schéma de Gosse-Toscani et de Jin-Levermore modifié

Maintenant on propose d'introduire un second schéma « asymptotic preserving » basé sur le schéma upwind pour l'équation de la chaleur hyperbolique (1.20). Le principe de construction du schéma de Gosse-Toscani consiste à utiliser un schéma de Godunov (upwind) où on localise les termes sources aux interfaces afin de capter les effets des termes sources. Une nouvelle onde stationnaire apparaît dans le problème de Riemann, les états associés à cette nouvelle onde sont calculés à partir des états stationnaires du système. On part de la version diagonalisée de l'équation de la chaleur hyperbolique (1.20) en posant $u=E+F$ et $v=E-F$. On obtient

$$\begin{cases} \partial_t u+\dfrac{1}{\varepsilon}\partial_x u=\dfrac{\sigma}{2\varepsilon^2}(v-u), \\ \partial_t v-\dfrac{1}{\varepsilon}\partial_x v=\dfrac{\sigma}{2\varepsilon^2}(u-v). \end{cases} \tag{1.37}$$

On commence donc par localiser les termes sources aux interfaces

$$\begin{cases} \partial_t u+\dfrac{1}{\varepsilon}\partial_x u=\dfrac{\sigma}{2\varepsilon^2}\displaystyle\sum_j \Delta x(v-u)\delta(x-x_{j-\frac{1}{2}}), \\ \partial_t v-\dfrac{1}{\varepsilon}\partial_x v=\dfrac{\sigma}{2\varepsilon^2}\displaystyle\sum_j \Delta x(u-v)\delta(x-x_{j-\frac{1}{2}}). \end{cases} \tag{1.38}$$

Le problème de Riemann associé à cette formulation admet une nouvelle onde stationnaire.

On peut calculer explicitement les solutions stationnaires du système d'EDP le long de cette onde. Pour cela on part du système

$$\begin{cases} \partial_x u^*=\dfrac{\Delta x\sigma}{2\varepsilon}(v^*-u^*), \\ -\partial_x v^*=\dfrac{\Delta x\sigma}{2\varepsilon}(u^*-v^*), \end{cases}$$

avec $u^*(0)=u_L$ et $v^*(\Delta x)=v_R$. On obtient les solutions

$$\begin{cases} u^*=u_L+\dfrac{\Delta x\sigma}{2\varepsilon+\sigma\Delta x}(v_R-u_L), \\ v^*=v_R-\dfrac{\Delta x\sigma}{2\varepsilon+\sigma\Delta x}(v_R-u_L). \end{cases} \tag{1.39}$$

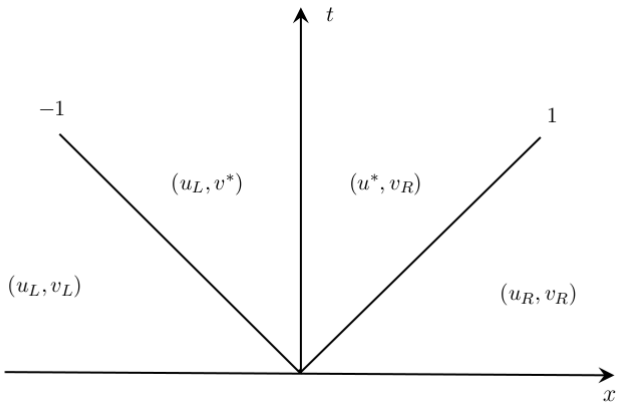


FIGURE 1.2 – Solution du nouveau problème de Riemann

Ces états intermédiaires (u^*, v^*) permettent de construire la solution du problème de Riemann donnée par la figure (1.2). Maintenant on écrit le schéma de Godunov, on intègre sur la maille j et entre les temps t_n et t_{n+1}, pour obtenir

$$\begin{cases} u_j^{n+1} = u_j^n - \dfrac{\Delta t}{\varepsilon \Delta x}(u^n_{j+\frac{1}{2}} - u^n_{j-\frac{1}{2}}), \\ \\ v_j^{n+1} = v_j^n + \dfrac{\Delta t}{\varepsilon \Delta x}(v^n_{j+\frac{1}{2}} - v^n_{j-\frac{1}{2}}). \end{cases} \tag{1.40}$$

Le problème de Riemann dans la maille j, décrit par les figures (1.3)-(1.2), donne les flux

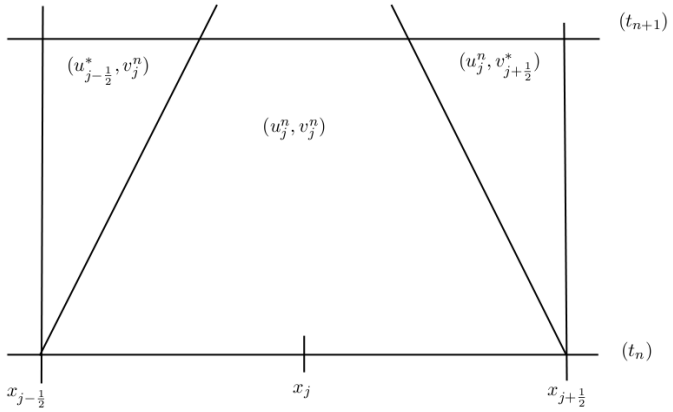


FIGURE 1.3 – Problème de Riemann dans la maille j.

$$\begin{aligned} u_{j-\frac{1}{2}} = u^*(x_{j-\frac{1}{2}}) = u_{j-1} + \frac{\Delta x \sigma}{2\varepsilon + \sigma \Delta x}(v_j - u_{j-1}),\ u_{j+\frac{1}{2}} = u_j, \\ v_{j+\frac{1}{2}} = v^*(x_{j+\frac{1}{2}}) = v_{j+1} - \frac{\Delta x \sigma}{2\varepsilon + \sigma \Delta x}(v_{j+1} - u_j),\ v_{j-\frac{1}{2}} = v_j. \end{aligned} \tag{1.41}$$

Une fois les flux obtenus on peut reformuler le schéma (1.40)-(1.41). On part de la première équation

$$u_j^{n+1} = u_j - \frac{\Delta t}{\varepsilon \Delta x}\left(u_j - \frac{\Delta x \sigma}{2\varepsilon + \sigma \Delta x}(v_j - u_{j-1}) - u_{j-1}\right). \tag{1.42}$$

On commence par introduire le facteur $M = \frac{2\varepsilon}{2\varepsilon + \sigma \Delta x}$, en factorisant on obtient

$$u_j^{n+1} = u_j - \frac{\Delta t}{\varepsilon \Delta x}\left(u_j - M u_{j-1} - \frac{\Delta x \sigma}{2\varepsilon} M v_j\right). \tag{1.43}$$

En utilisant la relation $(1 - M) = \frac{\Delta x \sigma}{2\varepsilon} M$, l'équation (1.43) peut se réécrire sous la forme

$$u_j^{n+1} = u_j - \frac{\Delta t}{\varepsilon \Delta x}\left(M(u_j - u_{j-1}) + \frac{\Delta x \sigma}{2\varepsilon} M(u_j - v_j)\right). \tag{1.44}$$

Enfin on agit de la même façon pour l'équation sur v on obtient le schéma

$$\begin{cases} \dfrac{u_j^{n+1} - u_j^n}{\Delta t} + M\dfrac{u_j^n - u_{j-1}^n}{\varepsilon \Delta x} - M\dfrac{\sigma}{2\varepsilon^2}\left(v_j^n - u_j^n\right) = 0, \\ \dfrac{v_j^{n+1} - v_j^n}{\Delta t} - M\dfrac{v_{j+1}^n - v_j^n}{\varepsilon \Delta x} - M\dfrac{\sigma}{2\varepsilon^2}\left(u_j^n - v_j^n\right) = 0, \end{cases} \tag{1.45}$$

Enfin, il suffit de repasser aux variables (E, F) de l'équation de la chaleur hyperbolique pour obtenir le schéma de Gosse-Toscani

$$\begin{cases} \dfrac{E_j^{n+1} - E_j^n}{\Delta t} + M\dfrac{F_{j+1}^n - F_{j-1}^n}{2\varepsilon \Delta x} - M\dfrac{E_{j+1}^n - 2E_j^n + E_{j-1}^n}{2\varepsilon \Delta x} = 0, \\ \dfrac{F_j^{n+1} - F_j^n}{\Delta t} + M\dfrac{E_{j+1}^n - E_{j-1}^n}{2\varepsilon \Delta x} - M\dfrac{F_{j+1}^n - 2F_j^n + F_{j-1}^n}{2\varepsilon \Delta x} + M\dfrac{\sigma}{\varepsilon^2} F_j^n = 0. \end{cases} \tag{1.46}$$

Le schéma de Gosse-Toscani peut donc être réécrit comme le schéma upwind multiplié par un coefficient M.

Lien avec le schéma de Jin-Levermore

Le schéma de Gosse-Toscani peut être obtenu à partir du schéma de Jin-Levermore avec une discrétisation du terme source différente. On introduit le schéma de Jin-Levermore modifié avec les mêmes flux (1.34) et une discrétisation du terme source aux interfaces (à l'aide des flux) et non aux centres des mailles.

$$\begin{cases} \dfrac{E_j^{n+1} - E_j^n}{\Delta t} + \dfrac{F_{j+\frac{1}{2}}^n - F_{j-\frac{1}{2}}^n}{\varepsilon \Delta x} = 0, \\ \dfrac{F_j^{n+1} - F_j^n}{\Delta t} + \dfrac{E_{j+\frac{1}{2}}^n - E_{j-\frac{1}{2}}^n}{\varepsilon \Delta x} + \dfrac{\sigma}{2\varepsilon^2}\left(F_{j+\frac{1}{2}}^n + F_{j-\frac{1}{2}}^n\right) = 0. \end{cases} \tag{1.47}$$

Lemme 1.9. *Le schéma de Jin-Levermore modifié est équivalent au schéma de Gosse-Toscani [GT01].*

Démonstration. On commence par incorporer les flux (1.34) dans le schéma (1.47). On a déjà montré que la première équation de (1.47) peut être réécrite comme la première équation de (1.46), reste la seconde équation. On part de

$$\frac{E_{j+\frac{1}{2}}^n - E_{j-\frac{1}{2}}^n}{\varepsilon \Delta x} + \frac{\sigma}{2\varepsilon^2}\left(F_{j+\frac{1}{2}}^n + F_{j-\frac{1}{2}}^n\right) = \frac{E_{j+1}^n - E_{j-1}^n}{2\varepsilon \Delta x} - \frac{F_{j+1}^n - 2F_j^n + F_{j-1}^n}{2\varepsilon \Delta x}$$

$$+\frac{M\sigma}{\varepsilon^2 4}\left(\left(F_j^n+F_{j+1}^n+E_j^n-E_{j+1}^n\right)+\left(F_{j-1}^n+F_j^n+E_{j-1}^n-E_j^n\right)\right)$$
$$=\left(\frac{1}{2\varepsilon\Delta x}-\frac{M\sigma}{\varepsilon^2 4}\right)\left(\left(E_{j+1}^n-E_{j-1}^n\right)-\left(F_{j+1}^n-2F_j^n+F_{j-1}^n\right)\right)+\frac{M\sigma}{\varepsilon^2}F_j^n.$$

On vérifie que

$$\frac{1}{2\varepsilon\Delta x}-\frac{M\sigma}{\varepsilon^2 4}=\frac{M}{2\varepsilon\Delta x}.$$

On obtient donc la relation suivante

$$\frac{E_{j+\frac{1}{2}}^n-E_{j-\frac{1}{2}}^n}{\varepsilon\Delta x}+\frac{\sigma}{2\varepsilon^2}\left(F_{j+\frac{1}{2}}^n+F_{j-\frac{1}{2}}^n\right)=M\left(\frac{E_{j+1}^n-E_{j-1}^n}{2\varepsilon\Delta x}-\frac{F_{j+1}^n-2F_j^n+F_{j-1}^n}{2\varepsilon\Delta x}+\frac{\sigma}{\varepsilon^2}F_j^n\right).$$

Par conséquent, le schéma de Jin-Levermore modifié admet l'écriture

$$\begin{cases}\dfrac{E_j^{n+1}-E_j^n}{\Delta t}+M\dfrac{F_{j+1}^n-F_{j-1}^n}{2\varepsilon\Delta x}-M\dfrac{E_{j+1}^n-2E_j^n+E_{j-1}^n}{2\varepsilon\Delta x}=0,\\ \dfrac{F_j^{n+1}-F_j.}{\Delta t}+M\dfrac{E_{j+1}^n-E_{j-1}^n}{2\varepsilon\Delta x}-M\dfrac{F_{j+1}^n-2F_j^n+F_{j-1}^n}{2\varepsilon\Delta x}+M\dfrac{\sigma}{\varepsilon^2}F_j^n=0.\end{cases}\tag{1.48}$$

qui correspond au schéma de Gosse-Toscani. □

Consistance du schéma de Gosse-Toscani

Lemme 1.10. *L'erreur de consistance du schéma de Gosse-Toscani (1.46) associée à la première équation est*

$$O\left(\Delta x^2+\varepsilon\Delta x+\Delta t\right),$$

L'erreur de consistance du schéma de Gosse-Toscani (1.46) associée à la seconde équation est

$$O(\Delta x+\Delta t).$$

Démonstration. L'erreur de consistance associée à la première équation a déjà été établie dans le cadre de l'étude du schéma de Jin-Levermore (1.35). Reste donc à étudier celle de la deuxième équation. On a

$$cF_j^n=\frac{F(x_j,t_{n+1})-F(x_j,t_n)}{\Delta t}+M\left(\frac{E(x_{j+1},t_n)-E(x_{j-1},t_n)}{2\varepsilon\Delta x}\right.$$
$$\left.-\frac{F(x_{j+1},t_n)-2F(x_j,t_n)+F(x_j,t_n)}{2\varepsilon\Delta x}+\frac{\sigma}{\varepsilon^2}F(x_j,t_n)\right)$$
$$=\partial_t F(x_j,t_n)+O(\Delta t)+M\left(\frac{2\Delta x\partial_x E(x_j,t_n)+O(\Delta x^3)}{2\varepsilon\Delta x}+\frac{\Delta x}{\varepsilon}\partial_{xx}F(x_j,t_n)+\frac{\sigma}{\varepsilon^2}F(x_j,t_n)\right).$$

On sait que $\partial_{xx}F=-\varepsilon\partial_{tx}E$ et $\|\partial_x^{m_1}\partial_t^{m_2}E\|_\infty\leq C_{m_1,m_2}$, par conséquent $\partial_{xx}F=O(\varepsilon)$ et

$$cF_j^n=(1-M)\partial_t F(x_j,t_n)+O\left(\Delta t+\frac{M\Delta x^2}{\varepsilon}+M\Delta x\right).$$

Le coefficient M par définition satisfait plusieurs propriétés : la première : $M\leq 1$, la seconde

$$\frac{M\Delta x}{\varepsilon}=\frac{2\varepsilon\Delta x}{2(2\varepsilon+\sigma\Delta x)\varepsilon}=\frac{2\Delta x}{2(2\varepsilon+\sigma\Delta x)}\leq\frac{1}{\sigma}\Longrightarrow O\left(\frac{M\Delta x^2}{\varepsilon}\right)=O(\Delta x),$$

et la troisième

$$1-M=\frac{\sigma\Delta x}{2\varepsilon+\sigma\Delta x}\Longrightarrow(1-M)O\left(\partial_t F(x_j,t_n)\right)=O\left(\frac{\sigma}{2+\frac{\sigma\Delta x}{\varepsilon}}\Delta x\right).$$

car $\partial_t F(x_j,t_n)=O(\varepsilon)$. Par conséquent

$$(1-M)\partial_t F(x_j,t_n)=O(\Delta x).$$

Ceci termine la preuve. □

Stabilité du schéma de Gosse-Toscani explicite

Lemme 1.11. *Le schéma de Gosse-Toscani (1.46) satisfait le principe du maximum sous la condition CFL*

$$M\frac{\Delta t}{\varepsilon\Delta x}\leq 1. \tag{1.49}$$

Démonstration. On commence par définir les invariants de Riemann $u=E+F$ et $v=E-F$, le système (1.26) se réécrit sous la forme suivante

$$\begin{cases}\dfrac{u_j^{n+1}-u_j^n}{\Delta t}+M\dfrac{u_j^n-u_{j-1}^n}{\varepsilon\Delta x}+M\dfrac{\sigma}{2\varepsilon^2}\left(u_j^n-v_j^n\right)=0,\\ \dfrac{v_j^{n+1}-v_j^n}{\Delta t}-M\dfrac{v_{j+1}^n-v_j^n}{\varepsilon\Delta x}+M\dfrac{\sigma}{2\varepsilon^2}\left(-u_j^n+v_j^n\right)=0,\end{cases} \tag{1.50}$$

par conséquent

$$\begin{cases}u_j^{n+1}=\left(1-M\dfrac{\Delta t}{\varepsilon\Delta x}-M\dfrac{\sigma\Delta t}{2\varepsilon^2}\right)u_j^n+M\dfrac{\Delta t}{\varepsilon\Delta x}u_{j-1}^n+M\dfrac{\sigma\Delta t}{2\varepsilon^2}v_j^n,\\ v_j^{n+1}=\left(1-M\dfrac{\Delta t}{\varepsilon\Delta x}-M\dfrac{\sigma\Delta t}{2\varepsilon^2}\right)v_j^n+M\dfrac{\Delta t}{\varepsilon\Delta x}v_{j+1}^n+M\dfrac{\sigma\Delta t}{2\varepsilon^2}u_j^n.\end{cases} \tag{1.51}$$

Les termes extra-diagonaux sont positifs, les termes diagonaux sont aussi positifs sous la condition $1-M\frac{\Delta t}{\varepsilon\Delta x}-M\frac{\sigma\Delta t}{2\varepsilon^2}\geq 0$. Pour finir on remarque que u_j^{n+1} et v_j^{n+1} sont calculés à partir de combinaison convexe des quantités au temps n. La CFL est donnée par $1-M\frac{\Delta t}{\varepsilon\Delta x}-M\frac{\sigma\Delta t}{2\varepsilon^2}\geq 0$. En remarquant que $\frac{M\sigma}{2\varepsilon}=\frac{1-M}{\Delta x}$, on obtient la condition CFL (1.49). □

Théorème 1.12. *Le schéma de Gosse-Toscani (1.46) est convergent avec l'erreur $O(\Delta x+\Delta t)$ en norme L^∞ sous la condition CFL (1.49).*

Par conséquent le schéma est uniformément « asymptotic preserving » (définition 1.3.4).

Ce résultat est donné par le théorème de Lax.

On remarque que le schéma de Gosse-Toscani est stable sous une condition CFL moins contraignante que le schéma upwind. En effet, le schéma de Gosse-Toscani modifie la viscosité numérique en modifiant le « scaling » du système de $\frac{1}{\varepsilon}$, $\frac{1}{\varepsilon^2}$ par 1 et $\frac{1}{\varepsilon}$. Par conséquent il est intéressant d'étudier l'effet de cette modification sur la discrétisation semi-implicite.

Stabilité du schéma de Gosse-Toscani semi-implicite

Lemme 1.13. *Le schéma de Gosse-Toscani semi-implicite satisfait le principe du maximum sur les invariants de Riemann sous la CFL*

$$\Delta t \leq \frac{\Delta x^2}{2\sigma} + \varepsilon \Delta x. \tag{1.52}$$

Démonstration. Le principe de preuve est exactement le même que pour le schéma upwind semi-implicite. On obtient la CFL

$$\frac{M\Delta t}{\varepsilon \Delta x} \leq 1.$$

Ce qui permet d'obtenir immédiatement le résultat (1.52). □

La condition CFL (1.52) paraît optimale. En effet celle-ci est la somme de la condition CFL hyperbolique et de la condition CFL parabolique.

Schéma de Gosse-Toscani avec σ variable

En modifiant la définition des flux du schéma de Gosse-Toscani (1.41) et des états intermédiaires (1.39) pour un coefficient σ variable, on obtient le schéma suivant sur les variables du modèle diagonalisé.

$$\begin{cases} \dfrac{u_j^{n+1} - u_j^n}{\Delta t} + M_{j-\frac{1}{2}} \dfrac{u_j^n - u_{j-1}^n}{\Delta x \varepsilon} = M_{j-\frac{1}{2}} \dfrac{\sigma_{j-\frac{1}{2}}}{\varepsilon^2}(v_j^n - u_j^n), \\ \dfrac{v_j^{n+1} - v_j^n}{\Delta t} - M_{j+\frac{1}{2}} \dfrac{v_{j+1}^n - v_j^n}{\Delta x \varepsilon} = M_{j+\frac{1}{2}} \dfrac{\sigma_{j+\frac{1}{2}}}{\varepsilon^2}(u_j^n - v_j^n), \end{cases} \tag{1.53}$$

avec $M_{j+\frac{1}{2}} = \frac{2\varepsilon}{2\varepsilon + \Delta x \sigma_{j+\frac{1}{2}}}$. On obtient, en repassant aux variables de l'équation de la chaleur hyperbolique, un schéma « asymptotic preserving » conservatif pour la quantité E et qui préserve le principe du maximum.

$$\begin{cases} \dfrac{E_j^{n+1} - E_j^n}{\Delta t} + \dfrac{F_{j+\frac{1}{2}}^n - F_{j-\frac{1}{2}}^n}{\varepsilon \Delta x} = 0, \\ \dfrac{F_j^{n+1} - F_j^n}{\Delta t} + \dfrac{E_{j+\frac{1}{2}}^n - E_{j-\frac{1}{2}}^n}{\varepsilon \Delta x} = \dfrac{M_{j+\frac{1}{2}} - M_{j-\frac{1}{2}}}{\varepsilon \Delta x} E_j^n \\ -\dfrac{1}{2}\left(M_{j+\frac{1}{2}} \dfrac{\sigma_{j+\frac{1}{2}}}{\varepsilon^2} + M_{j-\frac{1}{2}} \dfrac{\sigma_{j-\frac{1}{2}}}{\varepsilon^2} \right) F_j^n. \end{cases} \tag{1.54}$$

avec les flux

$$F_{j+\frac{1}{2}}^n = \frac{M_{j+\frac{1}{2}}}{2}\left(F_j^n + F_{j+1}^n + E_j^n - E_{j+1}^n\right), \; E_{j+\frac{1}{2}}^n = \frac{M_{j+\frac{1}{2}}}{2}\left(E_j^n + E_{j+1}^n + F_j^n - F_{j+1}^n\right).$$

Schéma de Berthon-Turpault

Récemment C. Berthon et R. Turpault ont construit un schéma « asymptotic preserving » pour des systèmes hyperboliques généraux [BLeFT11]- [BT10] du type

$$\partial_t \mathbf{U} + \frac{1}{\varepsilon} \partial_x F(\mathbf{U}) = -\sigma \frac{R(\mathbf{U})}{\varepsilon^2}. \tag{1.55}$$

Il est basé sur le solveur approché de Riemann appelé HLL. Il s'agit d'un solveur de Riemann à 2 ondes (quelque soit le nombre d'ondes du système total). Les deux ondes ont pour vitesse les vitesses

minimale et maximale du modèle étudié. On part du modèle (1.55) avec $\mathbf{U} = (u, v)$, $F(\mathbf{U}) = (u, -v)$, $R(\mathbf{U}) = (\frac{1}{2}(v-u), \frac{1}{2}(u-v))$. Le schéma HLL (ou upwind car dans ce cas ils sont identiques) s'écrit

$$\begin{cases} \dfrac{u_j^{n+1} - u_j^n}{\Delta t} + \dfrac{u_j^n - u_{j-1}^n}{\Delta x \varepsilon} + \dfrac{\sigma}{2\varepsilon^2}\left(u_j^n - v_j^n\right) = 0, \\ \dfrac{v_j^{n+1} - v_j^n}{\Delta t} - \dfrac{v_{j+1}^n - v_j^n}{\Delta x \varepsilon} + \dfrac{\sigma}{2\varepsilon^2}\left(-u_j^n + v_j^n\right) = 0. \end{cases}$$

Le solveur de Riemann approché HLL s'écrit

$$\mathbf{U}_{sol}(\frac{x}{t}, \mathbf{U}_L, \mathbf{U}_R) = \begin{cases} \mathbf{U}_L, \, x < -t, \\ \mathbf{U}^*, \, -t < x < t, \\ \mathbf{U}_R, \, x > t, \end{cases}$$

avec

$$\mathbf{U}^* = \frac{1}{2}\left(\mathbf{U}_L + \mathbf{U}_R\right) - \frac{1}{2}\left(F(\mathbf{U}_R) - F(\mathbf{U}_L)\right),$$

ce qui donne le schéma

$$\begin{cases} \mathbf{U}_j^{n+1} = \mathbf{U}_j - \dfrac{\Delta t}{\varepsilon \Delta x}\left(\mathbf{F}_{j+\frac{1}{2}} - \mathbf{F}_{j-\frac{1}{2}}\right), \\ \mathbf{F}_{j+\frac{1}{2}} = \dfrac{1}{2}\left(F(\mathbf{U}_j) + F(\mathbf{U}_{j+1})\right) - \dfrac{1}{2}\left(\mathbf{U}_{j+1} - \mathbf{U}_j\right). \end{cases} \tag{1.56}$$

L'idée est de modifier le schéma HLL en ajoutant une onde stationnaire (comme pour le schéma de Gosse-Toscani) afin d'incorporer les termes sources dans les flux. On obtient le nouveau problème de Riemann

$$\mathbf{U}_{sol}(\frac{x}{t}, \mathbf{U}_L, \mathbf{U}_R) = \begin{cases} \mathbf{U}_L, \, x < -t, \\ \mathbf{U}_L^*, \, -t < x < 0, \\ \mathbf{U}_R^*, \, 0 < x < t, \\ \mathbf{U}_R, \, x > t. \end{cases}$$

Les états intermédiaires étant donnés par une combinaison convexe entre l'état stationnaire du schéma HLL et le terme source modifié

$$\mathbf{U}_{L,R}^* = \widehat{\alpha}\mathbf{U}^* + \left((\widehat{I}_d - \widehat{\alpha})(\mathbf{U}_{L,R} - (\widehat{I}_d + \widehat{\sigma})^{-1} R(\mathbf{U}_{L,R}))\right),$$

où $\widehat{\alpha} = \left(\widehat{I}_d + \frac{\sigma \Delta x}{2\varepsilon}(\widehat{I}_d + \widehat{\sigma})\right)^{-1}$ et $\widehat{\sigma}$ est un paramètre de la méthode qui permet de récupérer le coefficient de diffusion de l'équation limite (en effet le schéma est défini pour des systèmes hyperboliques avec des limites de diffusion différentes, par conséquent il faut parfois corriger le coefficient de viscosité). Dans ce cas on choisit $\widehat{\sigma} = 0\widehat{I}_d$, ce qui donne

$$\begin{aligned} \mathbf{U}_j^{n+1} = \mathbf{U}_j^n - \frac{\Delta t}{\varepsilon \Delta x}\left(\alpha_{j+\frac{1}{2}} \mathbf{F}_{j+\frac{1}{2}}^n - \alpha_{j-\frac{1}{2}} \mathbf{F}_{j-\frac{1}{2}}^n\right) + \frac{\Delta t}{\varepsilon \Delta x}\left(\alpha_{j+\frac{1}{2}} - \alpha_{j-\frac{1}{2}}\right) F(\mathbf{U}_j^n) \\ -\Delta t \left(\alpha_{j+\frac{1}{2}} \frac{\sigma_{j+\frac{1}{2}}}{2\varepsilon^2} + \alpha_{j-\frac{1}{2}} \frac{\sigma_{j-\frac{1}{2}}}{2\varepsilon^2}\right) R(\mathbf{U}_j^n), \end{aligned}$$

avec $\alpha_{j+\frac{1}{2}} = \frac{2\varepsilon}{2\varepsilon + \Delta x \sigma_{j+\frac{1}{2}}}$. Pour un σ constant, on voit que ce schéma correspond au schéma de Gosse-Toscani. Par contre, dans le cas d'un σ variable il reste les termes non conservatifs

$$\frac{\Delta t}{\varepsilon \Delta x}(\alpha_{j+\frac{1}{2}} - \alpha_{j-\frac{1}{2}}) F(\mathbf{U}_j).$$

Le schéma final pour l'équation de la chaleur hyperbolique n'est pas conservatif pour la quantité E. Avec un σ constant le schéma dérive sur le schéma de Gosse-Toscani, donc les preuves de stabilité et de consistance sont les mêmes.

Analyse par équations modifiés

Afin de comprendre le comportement des schémas « asymptotic preserving », présentés précédemment, on propose de reprendre de façon formelle l'analyse par équations modifiées introduites dans [JL96]. Il s'agit d'étudier la limite de diffusion de l'équation associée au schéma considéré.

Schéma upwind

L'équation modifiée peut s'interpréter comme l'EDP consistante au second ordre avec le schéma numérique. Pour le schéma upwind on obtient

$$\begin{cases} \partial_t E + \dfrac{1}{\varepsilon}\partial_x F - \dfrac{\Delta x}{2\varepsilon}\partial_{xx} E = 0, \\ \partial_t F + \dfrac{1}{\varepsilon}\partial_x E - \dfrac{\Delta x}{2\varepsilon}\partial_{xx} F = -\dfrac{\sigma}{\varepsilon^2} F. \end{cases} \tag{1.57}$$

Afin d'obtenir la limite de diffusion associée à l'équation (1.57), on remarque que $F = O(\varepsilon)$ et on incorpore la relation $\varepsilon\partial_x E + O(\varepsilon^2) = -\sigma F$ dans la première équation de (1.57). On obtient l'équation limite suivante

$$\partial_t E - \partial_x \frac{1}{\sigma}\partial_x E - \frac{\Delta x}{2\varepsilon}\partial_{xx} E = O(\varepsilon). \tag{1.58}$$

Lorsque la grille 1D est dite fine ($\frac{\Delta x}{\varepsilon} << 1$), l'équation limite (1.58) est consistante avec l'équation limite de diffusion associée à l'équation de la chaleur hyperbolique (1.20). Par contre, lorsque la grille est dite grossière ($\frac{\Delta x}{\varepsilon} >> 1$) la diffusion numérique devient dominante par rapport à la diffusion physique et l'équation (1.58) n'est plus consistante avec l'équation de diffusion limite $\partial_t E - \partial_x \frac{1}{\sigma}\partial_x E = 0$. Cette analyse montre que le schéma upwind ne capture pas la bonne limite de diffusion sur grille grossière.

Schéma de Gosse-Toscani

Maintenant on propose d'étudier l'équation limite associée au schéma de Gosse-Toscani (l'analyse reste valable pour les schémas de Jin-Levermore et Berthon-Turpault) :

$$\begin{cases} \partial_t E + \dfrac{M}{\varepsilon}\partial_x F - \dfrac{M\Delta x}{2\varepsilon}\partial_{xx} E = 0, \\ \partial_t F + \dfrac{M}{\varepsilon}\partial_x E - \dfrac{M\Delta x}{2\varepsilon}\partial_{xx} F = -M\dfrac{\sigma}{\varepsilon^2} F, \end{cases} \tag{1.59}$$

avec $M = \frac{2\varepsilon}{2\varepsilon + \sigma\Delta x}$.
De la même façon que précédemment, on remarque que $F = O(\varepsilon)$ et on incorpore la relation $M\varepsilon\partial_x E + O(M\varepsilon^2) = -M\sigma F$ dans la première équation de (1.59). On obtient donc l'équation limite suivante

$$\partial_t E - \frac{1}{\sigma}\partial_{xx} E - \frac{M\Delta x}{2\varepsilon}\partial_{xx} E = O(\varepsilon). \tag{1.60}$$

Puisque $M\frac{\Delta x\sigma}{2\varepsilon} = 1 - M$, l'équation limite associée au système (1.59) peut se réécrire sous la forme

$$\partial_t E - \frac{M}{\sigma}\partial_{xx} E - \frac{1-M}{\sigma}\partial_{xx} E = O(\varepsilon). \tag{1.61}$$

Plusieurs remarques :

- Lorsque la grille est grossière ($\frac{\Delta x}{\varepsilon} >> 1$), le coefficient M tend vers zéro, la limite de diffusion est donnée par la viscosité du schéma numérique.
- Lorsque la grille est fine ($\frac{\Delta x}{\varepsilon} << 1$), le coefficient M tend vers 1, la limite de diffusion est donnée par la relation de compatibilité $\varepsilon \partial_x E = -\sigma F$, couplé à l'équation $\partial_t E + \frac{1}{\varepsilon}\partial_x F = 0$.
- L'équation de diffusion limite est formée d'une combinaison linéaire des deux régimes limites précédents. La combinaison linéaire des coefficients de diffusion vaut toujours $\frac{1}{\sigma}$. Par conséquent la limite de diffusion est correctement capturée par le schéma de Gosse-Toscani pour toutes valeurs de ε et Δx.

Exemples numériques

Afin de montrer l'intérêt pratique des schémas « asymptotic preserving », on propose de comparer numériquement le schéma upwind et le schéma de Gosse-Toscani en terme de qualité des solutions et de temps de calcul. La condition initiale est donnée par $E(x, t = 0) = G(x)$ avec $G(x)$ une Gaussienne et $F(x, t = 0) = 0$. Les paramètres sont donnés par $\sigma = 1$ et $\varepsilon = 0.001$. On propose donc de comparer les solutions numériques ainsi que le temps de calcul donné par le schéma Gosse-Toscani avec 50 mailles, 500 mailles et le schéma upwind avec 500, 1000, 10000 mailles. La discrétisation en temps est explicite. Le pas de temps est donné par la moitié du pas de temps limite de stabilité. Les erreurs de convergence sont calculées par rapport à la solution exacte de diffusion. Les résultats explicités dans

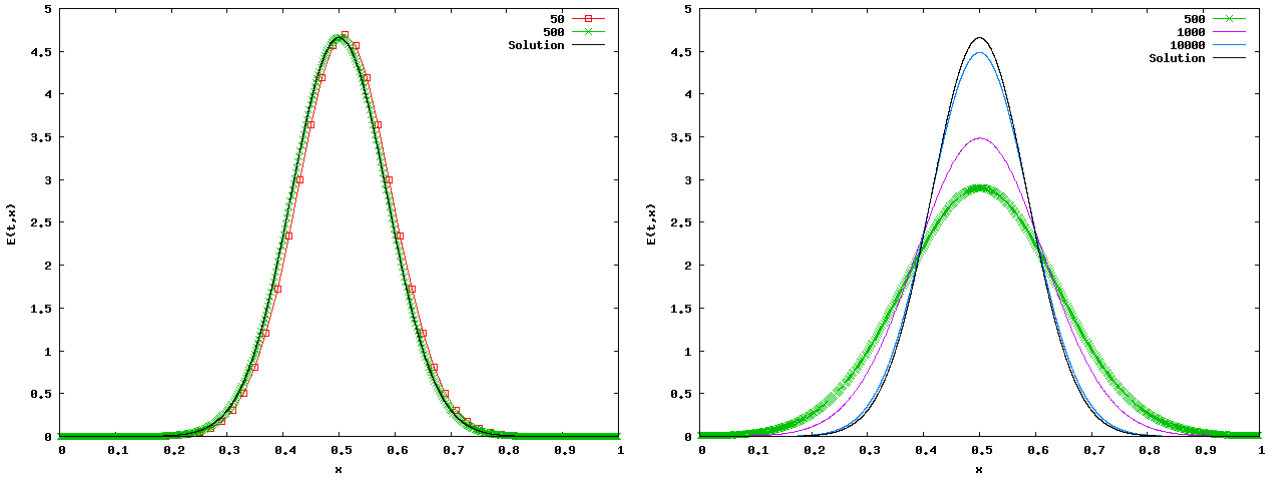


FIGURE 1.4 – À gauche : solutions numériques du schéma de Gosse Toscani avec 50 et 500 mailles, à droite : solutions numériques du schéma upwind avec 500, 1000 et 10000 mailles

le tableau (1.2) et par la figure (1.4) montrent que le schéma « asymptotic preserving » est nettement plus précis que le schéma upwind. En effet, les résultats explicités dans le tableau (1.2) montre que l'erreur associée au schéma « asymptotic preserving » avec 50 mailles est plus basse que celle associée au schéma upwind avec 10000 mailles. En outre on peux faire deux remarques sur le temps de calcul :

- Le temps de calcul, associé à la méthode « asymptotic preserving » est inférieur à celui associé au schéma upwind à cause de la condition CFL plus restrictive du schéma upwind pour $\varepsilon << \Delta x$ (lignes 2 et 3 du tableau (1.2)).
- Le temps de calcul nécessaire pour obtenir avec le schéma upwind, une erreur numérique comparable à celle obtenue avec un schéma « asymptotic preserving » est très important (ligne 1 et 5 du tableau (1.2)).

Schémas	Erreur L^1	Erreur L^2	Temps réel	Temps utilisateur
Schéma AP, 50 mailles	0.0065	0.0110	0m0.054s	0m0.157s
Schéma AP, 500 mailles	0.0001	0.00018	0m15.22s	1m1.680s
Schéma upwind, 500 mailles	0.445	0.647	0m24.317s	1m36.80s
Schéma upwind, 1000 mailles	0.279	0.113	2m9.530s	8m36.727s
Schéma upwind, 10000 mailles	0.0366	0.059	1485m4.26s	5140m56.11s

TABLE 1.2 – Tableau récapitulant l'erreur et le temps de calculs associés au schéma upwind et au schéma de Gosse-Toscani.

Ces remarques justifient largement l'utilisation d'un schéma « asymptotic preserving » par rapport au schéma upwind inutilisable en pratique.

1.3.4 Maillages non structurés

De nombreuses recherches étudient la résolution des équations de l'hydrodynamique (1.1) dans le cadre de la FCI. Les équations de l'hydrodynamique compressible non visqueuse sont résolues à l'aide de méthodes Eulérienne, Lagrangienne ou ALE. Cette dernière consiste à résoudre en coordonnées Lagrangienne les équations d'Euler et à projeter sur une grille régularisée lorsque le maillage est trop déformé ou le pas de temps trop petit. Le couplage avec les deux dernières méthodes nécessite de construire des méthodes numériques pour les problèmes de transport sur des maillages polygonaux vraiment déformés (exemple de maillage issu d'un calcul d'ALE (figure (3.1)). Par conséquent, la

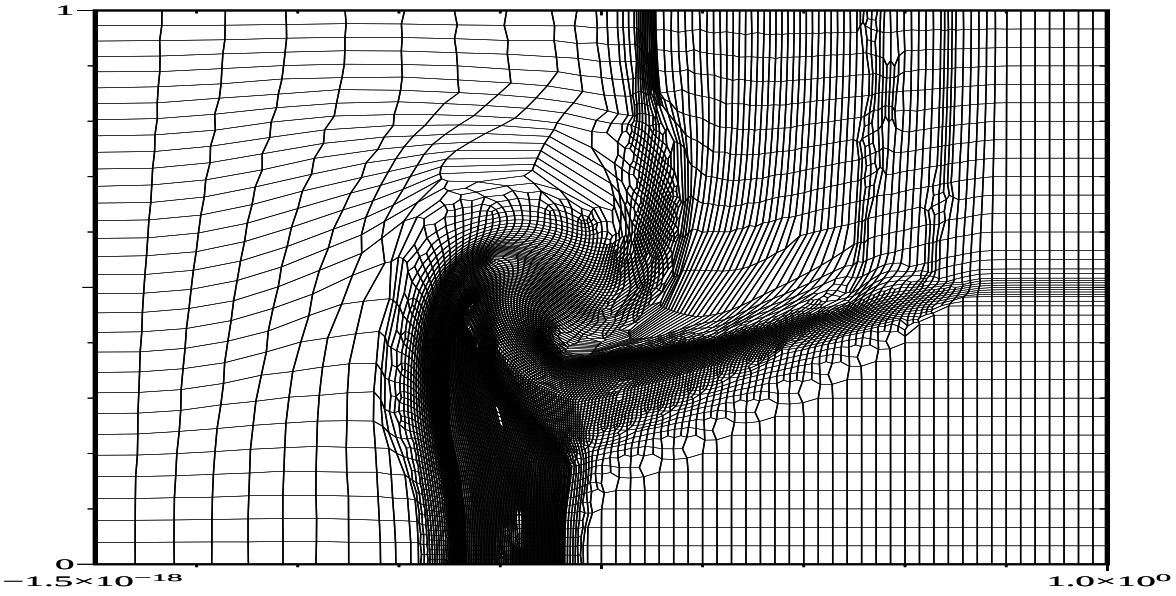


FIGURE 1.5 – Maillage polygonal 2D.

construction de schéma convergent pour des modèles simplifiés en transport sur des **maillages poly-**

gonaux, quadrangulaires et triangulaires très déformés est un enjeu important. Les maillages académiques aléatoires où de Kershaw sont, à ce titre, des bons exemples. En effet, le maillage aléatoire permet de vérifier si le schéma résiste à de fortes déformations locales, puisque des mailles peuvent être presque non convexes. Le maillage de Kershaw permet de vérifier si le schéma résiste à de fortes déformations globales. Le maillage aléatoire pose des problèmes essentiellement pour la consistance car la structure des mailles peut être très mauvaise ce qui empêche la construction d'un gradient consistant. Celui de Kershaw, quant à lui, est un bon exemple pour tester la robustesse de principe du maximum discret, car les fortes déformations rendent très difficile l'écriture du schéma sous forme de combinaison convexe qui permet d'obtenir le principe du maximum.

1.3.5 Extension des schémas « asymptotic preserving » sur maillages non structurés

Nous proposons d'appliquer la méthode de Jin-Levermore afin de construire un schéma « asymptotic preserving » de type volumes finis en dimension deux. Pour cela nous utiliserons la formulation classique des volumes finis qui consiste à localiser les flux aux centres des arêtes. Ensuite nous montrerons les limites de cette approche.
On commence par quelques notations, soit un maillage non structuré en dimension deux. Le maillage est défini par un nombre fini de sommets $\mathbf{x}_r$ et de mailles Ω_j. On note $\mathbf{x}_j$ un point arbitrairement choisi dans Ω_j. Par simplicité, on nomme ce point le centre de la maille.

L'arête commune entre les mailles j et k est notée $\partial\Omega_{jk}$ et le milieu de $\partial\Omega_{jk}$ est noté $\mathbf{x}_{jk}$. L'aire de la maille Ω_j est notée $|\Omega_j|$. Le pas de maillage est donné par $h = \max_j \operatorname{diam}(\Omega_j)$. Dans le cadre de cette formulation on note la normale $\mathbf{n}_{jk}$ de la maille j vers la maille k et l_{jk} la longueur de l'arête, associée à l'arête $\partial\Omega_{jk}$ (figure 1.5). Dans la suite de ce document j et k dénoteront les cellules, et r les noeuds.

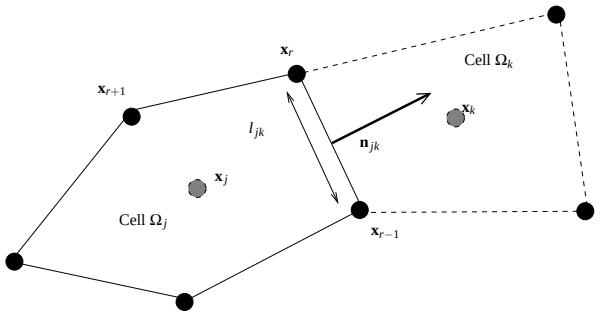


FIGURE 1.6 – Notation pour la formulation aux arêtes. l'interface entre la maille Ω_j et la maille Ω_k à une normale $\mathbf{n}_{jk} = -\mathbf{n}_{kj}$ et une longueur $l_{jk} = l_{kj}$. Les noeuds sont notés $\ldots, \mathbf{x}_{r-1}, \mathbf{x}_r, \mathbf{x}_{r+1}, \ldots$.

Nous considérons l'équation de la chaleur hyperbolique en dimension deux.

$$\begin{cases} \partial_t E + \dfrac{1}{\varepsilon} \operatorname{div} \mathbf{F} = 0, \\ \partial_t \mathbf{F} + \dfrac{1}{\varepsilon} \nabla E = -\dfrac{\sigma}{\varepsilon^2} \mathbf{F}. \end{cases} \tag{1.62}$$

On part du schéma upwind pour la résolution de cette équation

$$\left\{\begin{array}{l} |\Omega_j|\,\partial_t E_j + \dfrac{1}{\varepsilon}\displaystyle\sum_k l_{jk}(\mathbf{F}_{jk},\mathbf{n}_{jk}) = 0, \\ |\Omega_j|\,\partial_t \mathbf{F}_j + \dfrac{1}{\varepsilon}\displaystyle\sum_k l_{jk} E_{jk}\mathbf{n}_{jk} = -\,|\Omega_j|\,\dfrac{\sigma}{\varepsilon^2}\mathbf{F}_j, \end{array}\right. \tag{1.63}$$

où les flux associés à l'arête $\partial\Omega_{jk}$ commune aux mailles j et k sont définis par

$$\left\{\begin{array}{l} E_{jk} = \dfrac{1}{2}(E_j + E_k) + \dfrac{1}{2}(\mathbf{F}_j - \mathbf{F}_k, \mathbf{n}_{jk}), \\ (\mathbf{F}_{jk},\mathbf{n}_{jk}) = \dfrac{1}{2}(\mathbf{F}_j + \mathbf{F}_k, \mathbf{n}_{jk}) + \dfrac{1}{2}(E_j - E_k). \end{array}\right.$$

Ces flux peuvent être réécrits sous la forme suivante

$$\left\{\begin{array}{l} E_{jk} - E_j + (\mathbf{F}_{jk} - \mathbf{F}_j, \mathbf{n}_{jk}) = 0, \\ E_{jk} - E_k - (\mathbf{F}_{jk} - \mathbf{F}_k, \mathbf{n}_{jk}) = 0, \end{array}\right. \tag{1.64}$$

qui correspond au décentrement des invariants de Riemann du système 1D dans la direction normale $\mathbf{n}_{jk}$. La méthode de Jin-Levermore consiste à incorporer les états stationnaires dans les flux. On considère l'état stationnaire :

$$\nabla E = -\frac{\sigma}{\varepsilon}\mathbf{F}.$$

En utilisant un développement de Taylor et l'état stationnaire précédent, on obtient

$$\left\{\begin{array}{l} E(\mathbf{x}_{jk}) - E(\mathbf{x}_j) = \dfrac{\sigma}{\varepsilon}(\mathbf{F}(\mathbf{x}_{jk}), \mathbf{x}_j - \mathbf{x}_{jk}) + O(h^2), \\ \\ E(\mathbf{x}_{jk}) - E(\mathbf{x}_k) = \dfrac{\sigma}{\varepsilon}(\mathbf{F}(\mathbf{x}_{jk}), \mathbf{x}_k - \mathbf{x}_{jk}) + O(h^2), \end{array}\right. \tag{1.65}$$

avec $\mathbf{x}_j$ (resp. $\mathbf{x}_k$) le centre de la maille j et (resp. k), $\mathbf{x}_{jk}$ le milieu et h le pas de maillage.

L'idée à présent est de combiner les deux quantités (1.64) et (1.65) pour obtenir un nouveau système déterminant l'expression des flux :

$$\left\{\begin{array}{l} E_{jk} - E_j + (\mathbf{F}_{jk} - \mathbf{F}_j, \mathbf{n}_{jk}) = \dfrac{\sigma}{\varepsilon}(\mathbf{F}_{jk}, \mathbf{x}_j - \mathbf{x}_{jk}), \\ \\ E_{jk} - E_k - (\mathbf{F}_{jk} - \mathbf{F}_k, \mathbf{n}_{jk}) = \dfrac{\sigma}{\varepsilon}(\mathbf{F}_{jk}, \mathbf{x}_k - \mathbf{x}_{jk}). \end{array}\right. \tag{1.66}$$

Ce système linéaire est sous-déterminé, on a deux équations pour trois inconnues : la quantité scalaire E_{jk} et le vecteur $\mathbf{F}_{jk}$. Par conséquent le système (1.66) admet une infinité de solutions contrairement au cas de la dimension un.

Il est donc naturel, mais restrictif, de faire l'hypothèse que le maillage satisfait la condition de Delaunay [EGH00] qui implique que $(\mathbf{x}_{jk} - \mathbf{x}_j) = d_{jk}\mathbf{n}_{jk}$ et $(\mathbf{x}_{jk} - \mathbf{x}_k) = -d_{kj}\mathbf{n}_{jk}$, avec $d_{jk} = d(\mathbf{x}_j, \mathbf{x}_{jk}) > 0$ et $d_{kj} = d(\mathbf{x}_k, \mathbf{x}_{jk}) > 0$. Le système linéaire devient

$$\left\{\begin{array}{l} E_{jk} - E_j + (\mathbf{F}_{jk} - \mathbf{F}_j, \mathbf{n}_{jk}) = -\dfrac{\sigma}{\varepsilon} d_{jk}(\mathbf{F}_{jk}, \mathbf{n}_{jk}), \\ \\ E_{jk} - E_k - (\mathbf{F}_{jk} - \mathbf{F}_k, \mathbf{n}_{jk}) = \dfrac{\sigma}{\varepsilon} d_{kj}(\mathbf{F}_{jk}, \mathbf{n}_{jk}). \end{array}\right. \tag{1.67}$$

On obtient bien un système avec une unique solution donnée par

$$\begin{cases} (\mathbf{F}_{jk}, \mathbf{n}_{jk}) = \dfrac{(\mathbf{F}_j + \mathbf{F}_k, \mathbf{n}_{jk}) + (E_j - E_k)}{2 + (\frac{\sigma}{\varepsilon})(d_{jk} + d_{kj})}, \\ \\ E_{jk} = \dfrac{((\mathbf{F}_j, \mathbf{n}_{jk}) + E_j)(1 + d_{kj}(\frac{\sigma}{\varepsilon})) - ((\mathbf{F}_k, \mathbf{n}_{jk}) - E_k)(1 + d_{jk}(\frac{\sigma}{\varepsilon}))}{2 + (\frac{\sigma}{\varepsilon})(d_{jk} + d_{kj})}. \end{cases} \tag{1.68}$$

On obtient l'extension du schéma de Jin-Levermore en 2D défini par (1.63)-(1.68). Maintenant on exprime le schéma limite.

Proposition 1.14. *On suppose que le maillage satisfait les conditions de Delaunay. La limite asymptotique du schéma (1.63) avec les flux (1.68) est le schéma TPFA (« Two Points Flux Approximation ») défini par*

$$|\Omega_j| \, \partial_t E_j - \frac{1}{\sigma} \sum_k l_{jk} \frac{E_k - E_j}{d_{jk} + d_{kj}} = 0, \quad \text{avec } d_{jk} + d_{kj} = d(\mathbf{x}_j, \mathbf{x}_k). \tag{1.69}$$

Démonstration. En multipliant la première équation ε^2 par

$$\varepsilon^2 \, |\Omega_j| \, \partial_t \mathbf{F}_j + \varepsilon \sum_k l_{jk} E_{jk} \mathbf{n}_{jk} = - \, |\Omega_j| \, \sigma \mathbf{F}_j. \tag{1.70}$$

En factorisant le dénominateur par $\frac{1}{\varepsilon}$ on obtient la formulation suivante

$$\varepsilon E_{jk} = \frac{((\mathbf{F}_j, \mathbf{n}_{jk}) + E_j)(\varepsilon^2 + d_{kj}(\sigma\varepsilon)) - ((\mathbf{F}_k, \mathbf{n}_{jk}) - E_k)(\varepsilon^2 + d_{jk}(\sigma\varepsilon))}{(2\varepsilon + \sigma(d_{jk} + d_{kj}))}. \tag{1.71}$$

Quand ε tend vers 0, on voit que le terme de droite de (1.71) tend aussi vers zéro. Par conséquent en incorporant la relation (1.71) dans (1.70) on voit que $\mathbf{F}_j = \mathbf{0}$. Pour conclure il suffit d'utiliser ce résultat dans la définition (1.68) du flux $(\mathbf{F}_{jk}, \mathbf{n}_{jk})$ pour obtenir le schéma limite (1.69). Puisque le maillage satisfait la condition de Delaunay on sait que $(\mathbf{x}_j - \mathbf{x}_k) \perp \partial\Omega_{jk}$ et $d_{jk} + d_{kj} = d(\mathbf{x}_j, \mathbf{x}_k)$. Cela donne le schéma TPFA :

$$|\Omega_j| \partial_t E_j - \sum_k l_{jk} \frac{E_k - E_j}{d(\mathbf{x}_j, \mathbf{x}_k)} = 0. \tag{1.72}$$

□

Remarque 1.3.5. *Le schéma TPFA est aussi nommé FV4 (« Finite Volumes 4 ») pour des maillages triangulaires où FV5 (« Finite Volumes 5 ») pour des maillages quadrangulaires.*

Le schéma TPFA (« Two Points Flux Approximation ») (1.69) satisfait les propriétés suivantes :

- il est linéaire,
- il préserve le principe du maximum,
- il a un stencil à 5 points sur maillages quadrangulaires
- il est symétrique,
- S'il converge, il est d'ordre 2,
- il est coercif,

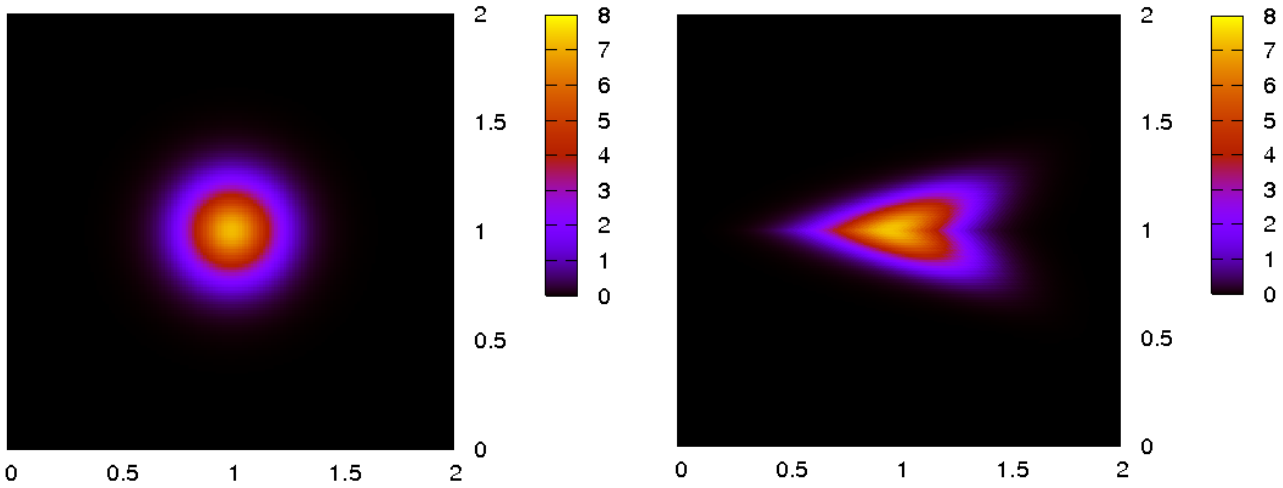


FIGURE 1.7 – A gauche la solution sur maillage Cartésien qui correspond à la solution. A droite la solution sur maillage de Kershaw

cependant il n'est pas consistant et ne converge pas sur des maillages déformés [EGH00]. En effet, lorsque la droite reliant les centres de mailles n'est pas alignée avec la normale (ceci est vérifié lorsque le maillage ne satisfait pas la condition de Delaunay) la quantité $\frac{E_k - E_j}{d(\mathbf{x}_j, \mathbf{x}_k)}$ n'approxime pas correctement le gradient dans la direction normale. On propose un cas test pour expliciter les problèmes géométriques liées au schéma TPFA. La condition initiale est donnée par une Gaussienne. Le temps final est $t_f = 0.1$. L'équation de la chaleur est résolue sur maillage Cartésien et sur maillage de Kershaw.

On constate sur la figure (1.7), que la qualité de la solution numérique donnée par le schéma TPFA est fortement dégradée par la déformation du maillage.

En dimension un, les méthodes « asymptotic preserving » et « cell-centered » consistent à modifier le coefficient de la viscosité du schéma pour le système hyperbolique afin que la viscosité domine dans le régime de diffusion et donne le schéma limite. En appliquant la même procédure en dimension deux on retrouve le schéma limite TPFA non consistant avec l'opérateur Laplacien sur maillages non structurés. Contrairement au 1D où la viscosité numérique est consistante, où il faut seulement modifier le coefficient de viscosité pour qu'elle domine avec le bon coefficient de diffusion lorsque ε tend vers zéro, en 2D, le travail est donc double : il faut construire une discrétisation hyperbolique telle que l'erreur de consistance est indépendante de ε et modifier la discrétisation géométrique de la diffusion numérique pour la rendre consistante avec un opérateur Laplacien. Depuis plusieurs années beaucoup de recherches ont été entreprises afin de construire des schémas de diffusion sur maillages non structurés. Ces schémas utilisent aussi une formulation aux arêtes mais différente de la discrétisation classique. Afin de mieux comprendre l'enjeu de la discrétisation de l'opérateur Laplacien sur maillages non structurés, nous proposons un rapide l'état de l'art sur les schémas de diffusion sur maillages non structurés.

1.3.6 Schémas de diffusion sur maillages non structurés

La construction de schémas de diffusion isotrope ou anisotrope sur maillages non structurés est depuis plusieurs années un thème de recherche très riche. C'est pourquoi dans cette section nous

proposons de faire un rappel sur certains schémas de diffusion valides sur maillages généraux pour l'équation

$$\partial_t E - \operatorname{div}(\nabla E) = 0.$$

En effet la construction de schémas « asymptotic preserving » sur maillages non structurés nécessite de comprendre la construction des schémas de diffusion convergent sur maillages généraux.

Schémas MPFA et de Breil-Maire

Le schéma MPFA fut un des premiers schémas de diffusion introduits sur maillages non structurés dans [ABBM96]- [AE06]. Comme pour le schéma TPFA on cherche à discrétiser un gradient dans la direction normale. L'idée consiste à introduire des inconnues fictives aux milieux des arêtes. À partir de ces inconnues on construit un gradient dans chaque triangle composé du centre de maille et des deux centres d'arêtes voisins. On impose ensuite la continuité des flux, ce qui donne un système linéaire à résoudre où les gradients dans la direction normale localisés aux centres des arêtes autour d'un noeud, sont les inconnues. Les inconnues fictives disparaissent lors de la résolution du système. Ce schéma sera détaillé dans un prochain chapitre. Ce schéma est consistant et convergent sous certaines conditions [AM08], mais les matrices locales ne sont pas toujours inversibles et le schéma est non symétrique. Pour résoudre ces problèmes J. Breil et P.H Maire ont introduit dans [BM06] une variante qui consiste au final à utiliser une formule du gradient consistante seulement si la maille admet une certaine régularité. Cette variante devient symétrique et les matrices locales sont toujours inversibles, par contre on perd la consistance sur des maillages où localement les mailles sont très déformées. Toutes ces variantes re-dérivent sur le schéma TPFA sur maillage Cartésien et ne préservent pas le principe du maximum.

Schémas SUSHI-Dioptre

Les schémas SUSHI et Dioptre ont été introduits par T. Gallouët, R. Herbin, R. Eymard dans [EGH09]. Comme pour les schémas précédents on définit les inconnues principales aux centres des mailles et d'autres inconnues aux milieux des arêtes. Afin d'éliminer ces inconnues localisées aux arêtes on utilise des formules barycentriques (ll existe aussi une formulation hybride du schéma SUSHI). Soit α une arête, on obtient donc

$$E_\alpha = \sum_k \gamma_k^\alpha E_k,$$

avec k un numéro de maille. A partir de là on peut définir naturellement un gradient discret dans la maille à l'aide de la formule

$$\nabla E_j = \frac{1}{|\,\Omega_j\,|} \sum_\alpha |\,\alpha\,|\,(E_\alpha - E_j)\mathbf{n}_{j,\alpha},$$

avec $|\,\alpha\,|$ la longueur de l'arête et $\mathbf{n}_{j,\alpha}$ la normale à l'arête α. Cette approche nécessite d'ajouter un stabilisateur du gradient discret, défini par l'erreur de consistance du gradient, on obtient les gradients

$$\nabla E_{j,\alpha} = \nabla E_j + R_{j,\alpha}\mathbf{n}_{j,\alpha},$$

avec $R_{j,\alpha} = \frac{\sqrt{d}}{d(\mathbf{x}_\alpha,\mathbf{x}_j)}\left(E_\alpha - E_j - (\nabla E_j, (\mathbf{x}_\alpha - \mathbf{x}_j))\right)$. À partir de ces gradients discrets et d'une formulation faible (ou variationnelle), on construit le schéma final [EGH09]. Le schéma obtenu est symétrique, possède un stencil large, dérive sur le schéma à TPFA (1.69) dans le cas d'un maillage Cartésien, ne préserve pas le principe du maximum et converge sur maillages non structurés. Une variante de ce schéma a été introduite afin de réduire le stencil au maximum, il s'agit du schéma Dioptre. Plutôt

que de définir les inconnues aux milieux des arêtes et de les exprimer à l'aide d'une combinaison barycentrique, on définit ces inconnues aux points formés par la moyenne harmonique des deux centres de mailles voisines. Par conséquent, l'inconnue localisée sur l'arête s'exprime par la moyenne harmonique des inconnues localisées aux centres de mailles voisines. Les reconstructions des gradients ne font intervenir que les mailles qui contiennent l'arête, ce qui réduit le stencil du schéma.

Schéma DDFV

Le schéma DDFV, contrairement aux schémas précédemment introduits, n'est pas un schéma « cell-centered » . En effet, une partie des inconnues est localisée aux centres des mailles et l'autre partie aux centres des mailles duales. Une cellule dite diamant est donc définie par les centres de maille du maillage physique et du dual. A partir de ces inconnues on reconstruit assez facilement un gradient dans la cellule diamant. Pour compléter le schéma il suffit d'écrire la discrétisation de la divergence pour chaque inconnue ou la discrétisation de la formulation variationnelle. Le schéma préserve la coercivité et converge sur maillages non structurés. L'utilisation de variables localisées aux noeuds et aux centres de mailles rend plus difficile le couplage avec des schémas hydrodynamiques de type Godunov où les inconnues sont situées uniquement aux centres des mailles. Ces schémas ont beaucoup été étudiés notamment dans [Her00]- [Her07].

Il existe encore un certain nombre de travaux sur la discrétisation des problèmes de diffusion sur maillages non structurés. On peut notamment citer les méthodes mimétiques [LSSV07]- [LMS04]- [BLS05]- [KLS04] ou encore les travaux de C. Le-Potier et ces co-auteurs sur des corrections non linéaire des schémas de diffusion permettant de récupérer la coercivité ou le principe du maximum [DlP09]- [LePo09]- [LePo09].

1.4 Objectifs de l'étude

A travers ce mémoire on souhaite étendre les schémas dit « asymptotic preserving » au cas des maillages non structurés et construire des schémas numériques adaptés pour les différents systèmes hyperboliques avec termes sources raides issus principalement des équations de transport linéaires. On souhaite donc construire des méthodes numériques avec à l'esprit plusieurs objectifs :

- Le schéma doit être « cell centered » (centré aux mailles) pour être compatible avec les codes hydrodynamiques Lagrangiennes.
- Le schéma doit être consistant dans tous les régimes, notamment le régime de diffusion.
- Si le modèle admet des propriétés tel que un principe du maximum ou une propriété d'entropie on souhaite les conserver.
- Les schémas doivent être valides pour des paramètres de relaxation variables (opacités pour le transfert radiatif, sections efficaces pour la neutronique).
- On souhaite construire des schémas semi-implicites avec des conditions CFL indépendantes de ε ou des schémas implicites dont le conditionnement est acceptable pour toute valeur de ε.
- Les schémas 2D doivent pouvoir se généraliser facilement au 3D.

Deuxième partie

Schémas « asymptotic preserving » pour des modèles linéaires

Chapitre 2

Schémas « asymptotic preserving » aux noeuds pour l'équation de la chaleur hyperbolique

Comme cela a été montré dans la section précédente, l'extension naturelle des schémas « asymptotic preserving » en dimension deux n'est pas valide en raison du défaut de consistance d'un schéma de diffusion limite. Pour obtenir des schémas consistants dans tous les régimes on propose deux solutions. La première consiste à partir d'un schéma de diffusion convergent et d'utiliser les mêmes principes de construction afin d'obtenir une discrétisation du système hyperbolique compatible avec le schéma de diffusion. À partir du nouveau schéma pour le système hyperbolique on obtient un schéma « asymptotic preserving » en modifiant les flux à l'aide de la méthode de Jin-Levermore [JL96]. Cette façon de procéder a été étudiée lors d'un projet CEMRACS dont les résultats seront énoncés dans un prochain chapitre. La deuxième solution consiste à utiliser un schéma pour le modèle hyperbolique utilisant une formulation différente. C'est cette solution que nous détaillons dans un premier temps. Dans cette section, on proposera d'étudier un formalisme volumes finis aux noeuds introduit pour l'hydrodynamique Lagrangienne dans [Maz07]- [CDDL09]. Après avoir rappelé le formalisme aux noeuds, on introduira l'extension en dimension deux des schémas de Jin-Levermore et Gosse-Toscani valides sur maillages non structurés. On obtiendra un nouveau schéma de diffusion pour lequel on peut donner une preuve de convergence. Enfin on donnera des résultats théoriques et numériques.

2.1 Schémas « asymptotic preserving » avec les flux localisés aux noeuds

Le formalisme aux noeuds pour la méthode des volumes finis à été introduit dans le cadre de l'hydrodynamique Lagrangienne par B. Després et C. Mazeran dans [Maz07] nommée depuis GLACE pour les équations d'Euler [Des09], [CDDL09]. Afin d'obtenir une vitesse pour chaque noeud, ce qui permet de déplacer naturellement le maillage, les auteurs ont proposés de localiser les flux aux noeuds du maillage contrairement au formalisme classique qui localise les flux aux milieux des arêtes. Afin de construire un schéma « asymptotic preserving » aux noeuds pour l'équation de la chaleur hyperbolique, nous utilisons l'analogie entre l'équation de la chaleur hyperbolique et les équations d'Euler linéarisées en coordonnée Eulerienne ou Lagrangienne. Une fois le schéma aux noeuds obtenu pour l'équation des ondes, on propose de le modifier à l'aide de la méthode de Jin-Levermore.

2.1.1 Rappels sur le schéma GLACE pour l'équation des ondes

On commence par quelques notations. Par convention on suppose que les noeuds sont numérotés dans le sens trigonométrique $...\mathbf{x}_{r-1}, \mathbf{x}_r, \mathbf{x}_{r+1}...$ avec les coordonnées $\mathbf{x}_r = (x_r, y_r)$. Comme pour la formulation classique on définit une longueur et une normale associées à chaque noeud

$$l_{jr} = \frac{1}{2}||\mathbf{x}_{r+1} - \mathbf{x}_{r-1}||_2 \text{ et } \mathbf{n}_{jr} = \frac{1}{2l_{jr}} \begin{pmatrix} -y_{r-1} + y_{r+1} \\ x_{r-1} - x_{r+1} \end{pmatrix}. \tag{2.1}$$

Les normales et les longueurs définies aux noeuds satisfont les propriétés suivantes

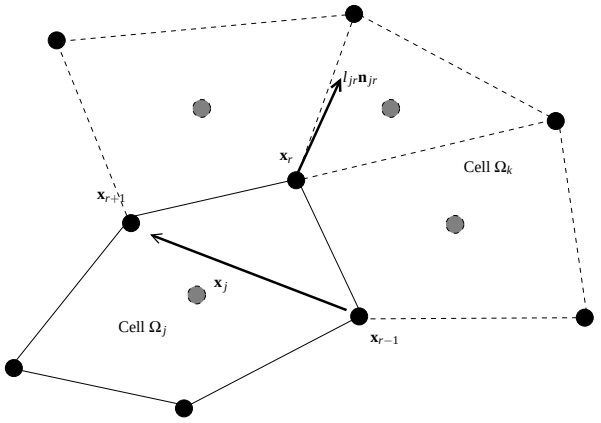


FIGURE 2.1 – Notation pour la formulation aux noeuds. La longueur associée au noeud l_{jr} et la normale associée au noeud $\mathbf{n}_{jr}$ sont définies par (2.1). On remarque $l_{jr}\mathbf{n}_{jr}$ est égal à la moitié du vecteur qui commence au noeud $\mathbf{x}_{r-1}$ et finit au noeud $\mathbf{x}_{r+1}$. Pour cette formulation le centre de la maille est également choisi arbitrairement.

$$\sum_j l_{jr}\mathbf{n}_{jr} = \mathbf{0}, \ \sum_r l_{jr}\mathbf{n}_{jr} = \mathbf{0}. \tag{2.2}$$

Enfin, on remarque $\nabla_{\mathbf{x}_r}\Omega_j = l_{jr}\mathbf{n}_{jr}$. Cette propriété est à la base de la construction du schéma. On peut se référer à [CDDL09]- [Des09].

Définition 2.1.1. *Le schéma GLACE pour l'équation des ondes (modèle (1.62) avec $\sigma = 0$) est défini par*

$$\begin{cases} |\Omega_j|\,\partial_t E_j + \sum_r l_{jr}(\mathbf{F}_r, \mathbf{n}_{jr}) = 0, \\ |\Omega_j|\,\partial_t \mathbf{F}_j + \sum_r l_{jr}E_{jr}\mathbf{n}_{jr} = \mathbf{0}. \end{cases} \tag{2.3}$$

avec les flux

$$\begin{cases} E_{jr} = E_j + (\mathbf{F}_j - \mathbf{F}_r, \mathbf{n}_{jr}), \\ \sum_j l_{jr}E_{jr}\mathbf{n}_{jr} = \mathbf{0}. \end{cases} \tag{2.4}$$

Le schéma GLACE satisfait plusieurs propriétés, conséquences des propriétés géométriques (2.2)

qui sont nécessaires pour obtenir un schéma conservatif :

$$\begin{cases} \sum_j l_{jr}(\mathbf{F}_r, \mathbf{n}_{jr}) = 0, \\ \sum_j l_{jr}(\mathbf{F}_r, \mathbf{n}_{jr}) E_{jr} = 0. \end{cases} \tag{2.5}$$

Dans [Maz07] l'auteur donne un résultat de convergence que nous rappelons.

Définition 2.1.2 (Résultats de convergence du schéma GLACE pour l'équation des ondes)**.** *Soit* $\mathbf{V}_h(t)$ *le vecteur de la solution du schéma discret en espace,* $\mathbf{V}(t)$ *le vecteur de la solution exacte. On note* h *le pas du maillage supposé régulier. On suppose que la donnée initiale appartient à* $H^3(\mathbb{R}^2)$*. Le schéma GLACE (2.3)-(2.4) pour l'équation des ondes converge avec l'ordre* $h^{\frac{1}{2}}$ *et l'estimation*

$$||\mathbf{V}_h(t) - \mathbf{V}(t)||_{L^2(\mathbb{R}^2)} \leq C\sqrt{\Big((1+t)||\nabla \mathbf{V}(0)||^2_{L^2(\mathbb{R}^2)} + t||\mathbf{V}(0)||^2_{H^3(\mathbb{R}^2)}\Big)} h^{\frac{1}{2}}.$$

2.1.2 Construction du schéma « asymptotic preserving » aux noeuds

On commence par écrire le schéma GLACE pour l'équation de la chaleur hyperbolique.

$$\begin{cases} |\Omega_j|\, \partial_t E_j + \dfrac{1}{\varepsilon} \sum_r l_{jr}(\mathbf{F}_r, \mathbf{n}_{jr}) = 0 \\ |\Omega_j|\, \partial_t \mathbf{F}_j + \dfrac{1}{\varepsilon} \sum_r l_{jr} E_{jr} \mathbf{n}_{jr} = -\,|\Omega_j|\, \dfrac{\sigma}{\varepsilon^2} \mathbf{F}_j, \end{cases} \tag{2.6}$$

avec les flux définis par

$$\begin{cases} E_{jr} = E_j + (\mathbf{F}_j - \mathbf{F}_r, \mathbf{n}_{jr}), \\ \sum_j l_{jr} E_{jr} \mathbf{n}_{jr} = \mathbf{0}. \end{cases} \tag{2.7}$$

La première formule de flux (2.7) correspond à l'invariant de Riemann écrit dans la direction normale. La seconde formule de flux (2.7) correspond à l'annulation de la somme des forces autour d'un noeud (point de vue mécanique). Afin d'obtenir l'expression du flux $\mathbf{F}_r$ on injecte $E_{jr} = E_j + (\mathbf{F}_j - \mathbf{F}_r, \mathbf{n}_{jr})$ dans l'égalité $\sum_j l_{jr} E_{jr} \mathbf{n}_{jr} = \mathbf{0}$, ce qui donne

$$\sum_j l_{jr} \left(E_j + (\mathbf{F}_j - \mathbf{F}_r, \mathbf{n}_{jr})\right) \mathbf{n}_{jr} = \mathbf{0}.$$

Par conséquent le flux $\mathbf{F}_r$ est solution du système linéaire :

$$\left(\sum_j l_{jr} \mathbf{n}_{jr} \otimes \mathbf{n}_{jr}\right) \mathbf{F}_r = \sum_j l_{jr} E_j \mathbf{n}_{jr} + \sum_j l_{jr} \mathbf{n}_{jr} \otimes \mathbf{n}_{jr} \mathbf{F}_j. \tag{2.8}$$

Remarque 2.1.1. *C. Mazeran [Maz07] a montré que la matrice du système (2.8) était toujours inversible sur les maillages non dégénérés.*

Afin d'obtenir un schéma « asymptotic preserving » on propose d'utiliser la méthode de Jin-Levermore. Pour cela on incorpore l'état stationnaire $\nabla E = -\frac{\sigma}{\varepsilon}\mathbf{F}$ dans les flux. On part de l'approximation de Taylor à l'ordre un

$$E(\mathbf{x}_j) \approx E(\mathbf{x}_r) + (\mathbf{x}_j - \mathbf{x}_r, \nabla E(\mathbf{x}_r)).$$

En utilisant l'état stationnaire de la seconde équation de (2.6), on obtient

$$E(\mathbf{x}_j) \approx E(\mathbf{x}_r) - \frac{\sigma}{\varepsilon}(\mathbf{x}_j - \mathbf{x}_r, \mathbf{F}(\mathbf{x}_r)). \tag{2.9}$$

On écrit maintenant l'équivalent discret de l'équation (2.9) à l'aide des flux et des valeurs centrées aux mailles puis on couple cette expression avec l'invariant de Riemann (2.7), ce qui permet d'obtenir

$$\begin{cases} E_{jr} = E_j + (\mathbf{F}_j - \mathbf{F}_r, \mathbf{n}_{jr}) - \dfrac{\sigma}{\varepsilon}(\mathbf{F}_r, \mathbf{x}_r - \mathbf{x}_j), \\ \displaystyle\sum_j l_{jr} E_{jr} \mathbf{n}_{jr} = \mathbf{0}. \end{cases} \tag{2.10}$$

Les flux (2.10) couplés au schéma (2.6) donne le schéma final que l'on nomme JL-(a).

Définition 2.1.3. *Le schéma JL-(a) s'écrit*

$$\begin{cases} |\Omega_j| \, \partial_t E_j + \dfrac{1}{\varepsilon} \displaystyle\sum_r l_{jr}(\mathbf{F}_r, \mathbf{n}_{jr}) = 0, \\ |\Omega_j| \, \partial_t \mathbf{F}_j + \dfrac{1}{\varepsilon} \displaystyle\sum_r l_{jr} E_{jr} \mathbf{n}_{jr} = - |\Omega_j| \, \dfrac{\sigma}{\varepsilon^2} \mathbf{F}_j, \end{cases} \tag{2.11}$$

avec les flux

$$\begin{cases} E_{jr} = E_j + (\mathbf{F}_j - \mathbf{F}_r, \mathbf{n}_{jr}) - \dfrac{\sigma}{\varepsilon}(\mathbf{F}_r, \mathbf{x}_r - \mathbf{x}_j), \\ \left(\displaystyle\sum_j l_{jr} \mathbf{n}_{jr} \otimes \mathbf{n}_{jr} + \frac{\sigma}{\varepsilon} \sum_j l_{jr} \mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j) \right) \mathbf{F}_r = \displaystyle\sum_j l_{jr} E_j \mathbf{n}_{jr} + \sum_j l_{jr} \mathbf{n}_{jr} \otimes \mathbf{n}_{jr} \mathbf{F}_j. \end{cases} \tag{2.12}$$

On commence par étudier l'asymptotique du schéma (2.11)-(2.12). Classiquement l'étude asymptotique de ce genre de problème utilise des développements de Hilbert qui consistent à décomposer chaque variable comme une série de puissances en ε. Un exemple d'utilisation de développement de Hilbert est donné dans [BCLM02]. Pour appliquer un développement de Hilbert dans le cas discret, on fixe le maillage (ce qui revient à étudier la limite asymptotique sur maillage grossier $\frac{h}{\varepsilon} >> 1$ avec h le pas de maillage).

Proposition 2.1. *Si la matrice de (2.12) est inversible, la limite asymptotique du schéma JL-(a) est*

$$\begin{cases} |\Omega_j| \, \partial_t E_j + \displaystyle\sum_r l_{jr}(\mathbf{F}_r, \mathbf{n}_{jr}) = 0, \\ \sigma \left(\displaystyle\sum_j l_{jr} \mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j) \right) \mathbf{F}_r = \displaystyle\sum_j l_{jr} E_j \mathbf{n}_{jr}. \end{cases} \tag{2.13}$$

Démonstration. On écrit le développement de Hilbert

$$\begin{aligned} E_j &= E_{j,0} + \varepsilon E_{j,1} + \varepsilon^2 E_{j,2} + o(\varepsilon^2), \\ E_{jr} &= E_{jr,0} + \varepsilon E_{jr,1} + \varepsilon^2 E_{jr,2} + o(\varepsilon^2). \end{aligned}$$

Le même développement est posé pour toutes les variables. En incorporant ces développements dans la première équation (2.11) on a

$$|\Omega_j|\partial_t E_{j,0} + \frac{1}{\varepsilon}\sum_r l_{jr}(\mathbf{F}_{r,1}, \mathbf{n}_{jr}) + \sum_r l_{jr}(\mathbf{F}_{r,0}, \mathbf{n}_{jr}) = O(\varepsilon).$$

Afin d'obtenir la limite de diffusion, on isole les termes associés aux puissances de $\frac{1}{\varepsilon}$. Le terme proportionnel à $\frac{1}{\varepsilon}$ pour la première équation de (2.11) est

$$\sum_r l_{jr}(\mathbf{F}_{r,0}, \mathbf{n}_{jr}) = 0. \tag{2.14}$$

le terme proportionnel à $O(1)$ est

$$|\Omega_j|\partial_t E_{j,0} + \sum_r l_{jr}(\mathbf{F}_{r,1}, \mathbf{n}_{jr}) = 0. \tag{2.15}$$

En utilisant la même méthode pour la deuxième équation de (2.11) on obtient trois termes

$$\frac{1}{\varepsilon^2}: \qquad \mathbf{F}_{j,0} = \mathbf{0}, \tag{2.16}$$

$$\frac{1}{\varepsilon}: \qquad \sum_r l_{jr} E_{jr,0}\mathbf{n}_{jr} = -\sigma|\Omega_j|\mathbf{F}_{j,1}, \tag{2.17}$$

$$\frac{1}{\varepsilon^0}: \qquad |\Omega_j|\partial_t \mathbf{F}_{j,0} + \sum_r l_{jr} E_{jr,1}\mathbf{n}_{jr} = -\sigma|\Omega_j|\mathbf{F}_{j,2}. \tag{2.18}$$

De la même façon on obtient pour le solveur nodal (2.10), le developpement de Hilbert donne

$$\frac{1}{\varepsilon}: \qquad \sigma\left(\sum_j l_{jr}\mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)\right)\mathbf{F}_{r,0} = \mathbf{0}, \tag{2.19}$$

et

$$\frac{1}{\varepsilon^0}: \qquad \left(\sum_j l_{jr}\mathbf{n}_{jr} \otimes \mathbf{n}_{jr}\right)\mathbf{F}_{r,0} + \sigma\left(\sum_j l_{jr}\mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)\right)\mathbf{F}_{r,1} \tag{2.20}$$

$$= \sum_j l_{jr} E_{j,0}\mathbf{n}_{jr} + \sum_j l_{jr}\mathbf{n}_{jr} \otimes \mathbf{n}_{jr}\mathbf{F}_{j,0}.$$

En supposant que la matrice du système (2.19) est inversible, on montre $\mathbf{F}_{r,0} = \mathbf{0}$. Puisque les équations (2.16) et (2.20) sont vérifiées, on obtient

$$\sigma\left(\sum_j l_{jr}\mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)\right)\mathbf{F}_{r,1} = \sum_j l_{jr} E_{j,0}\mathbf{n}_{jr}. \tag{2.21}$$

Cette égalité couplée à l'égalité (2.15) donnent le résultat final (2.13). □

Le schéma (2.11)-(2.12) peut s'interpréter comme l'extension en dimension deux et aux noeuds du schéma de Jin-Levermore. Il serait maintenant intéressant d'étendre aussi le schéma de Gosse-Toscani. Le principe de construction du schéma de Gosse-Toscani utilisant la localisation des termes sources aux interfaces et un problème de Riemann parait difficilement utilisable en dimension deux. On propose donc de partir du schéma JL-(a), avec une discrétisation du terme source localisé aux interfaces comme en dimension un (1.47). Cette méthode permet d'obtenir un schéma qui peut être vu comme l'extension du schéma de Gosse-Toscani en dimension deux. Pour cela on utilise la relation

$$\sum_r l_{jr}\mathbf{n}_{jr}\otimes\mathbf{x}_r = |\Omega_j|\,\widehat{I}_d.$$

Cette formule est à la base de l'étude de consistance menée sur GLACE dans [Des09]. On peut ainsi définir une nouvelle discrétisation du terme source qui reste homogène à la discrétisation classique.

Définition 2.1.4. *Le schéma JL-(b) est défini par*

$$\begin{cases} |\Omega_j|\,\partial_t E_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r l_{jr}(\mathbf{F}_r,\mathbf{n}_{jr}) = 0, \\ |\Omega_j|\,\partial_t \mathbf{F}_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r l_{jr}E_{jr}\mathbf{n}_{jr} = -\dfrac{\sigma}{\varepsilon^2}\displaystyle\sum_r l_{jr}\mathbf{n}_{jr}\otimes(\mathbf{x}_r-\mathbf{x}_j)\mathbf{F}_r. \end{cases} \tag{2.22}$$

avec les flux

$$\begin{cases} E_{jr} = E_j + (\mathbf{F}_j-\mathbf{F}_r,\mathbf{n}_{jr}) - \dfrac{\sigma}{\varepsilon}(\mathbf{F}_r,\mathbf{x}_r-\mathbf{x}_j), \\ \left(\displaystyle\sum_j l_{jr}\mathbf{n}_{jr}\otimes\mathbf{n}_{jr} + \frac{\sigma}{\varepsilon}\sum_j l_{jr}\mathbf{n}_{jr}\otimes(\mathbf{x}_r-\mathbf{x}_j)\right)\mathbf{F}_r = \displaystyle\sum_j l_{jr}E_j\mathbf{n}_{jr} + \sum_j l_{jr}\mathbf{n}_{jr}\otimes\mathbf{n}_{jr}\mathbf{F}_j. \end{cases} \tag{2.23}$$

Proposition 2.2. *Si la matrice nodale des flux (2.23) est inversible, le schéma JL-(b) (2.22)-(2.23) admet le schéma (2.13) comme limite asymptotique.*

Démonstration. Le schéma (2.22)-(2.23) peut se réécrire sous la forme suivante

$$\begin{cases} |\Omega_j|\,\partial_t E_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r l_{jr}(\mathbf{F}_r,\mathbf{n}_{jr}) = 0, \\ |\Omega_j|\,\partial_t \mathbf{F}_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r l_{jr}E^*_{jr}\mathbf{n}_{jr} = \mathbf{0}, \end{cases} \tag{2.24}$$

avec les flux

$$\begin{cases} E^*_{jr} = E_j + (\mathbf{F}_j-\mathbf{F}_r,\mathbf{n}_{jr}), \\ \left(\displaystyle\sum_j l_{jr}\mathbf{n}_{jr}\otimes\mathbf{n}_{jr} + \frac{\sigma}{\varepsilon}\sum_j l_{jr}\mathbf{n}_{jr}\otimes(\mathbf{x}_r-\mathbf{x}_j)\right)\mathbf{F}_r = \displaystyle\sum_j l_{jr}E_j\mathbf{n}_{jr} + \sum_j l_{jr}\mathbf{n}_{jr}\otimes\mathbf{n}_{jr}\mathbf{F}_j. \end{cases} \tag{2.25}$$

En utilisant les développements de Hilbert on obtient

$$\frac{1}{\varepsilon}: \qquad \sum_r l_{jr}(\mathbf{F}_{r,0},\mathbf{n}_{jr}) = 0, \tag{2.26}$$

$$\frac{1}{\varepsilon^0}: \qquad |\Omega_j|\partial_t E_{j,0} + \sum_r l_{jr}(\mathbf{F}_{r,0}, \mathbf{n}_{jr}) = 0. \tag{2.27}$$

Pour la seconde équation on obtient

$$\frac{1}{\varepsilon}: \qquad \sum_r l_{jr} E^*_{jr,0} \mathbf{n}_{jr} = \mathbf{0}, \tag{2.28}$$

$$\frac{1}{\varepsilon^0}: \qquad |\Omega_j|\partial_t \mathbf{F}_{j,0} + \sum_r l_{jr} E^*_{jr,1} \mathbf{n}_{jr} = \mathbf{0}. \tag{2.29}$$

Maintenant on utilise la définition du nouveau flux E^*_{jr}, ce qui donne

$$\sum_r l_{jr} E^*_{jr,0} \mathbf{n}_{jr} = \sum_r l_{jr} E_{j,0} \mathbf{n}_{jr} + \sum_r l_{jr} \mathbf{n}_{jr} \otimes \mathbf{n}_{jr} (\mathbf{F}_{j,0} - \mathbf{F}_{r,0}) = \mathbf{0}. \tag{2.30}$$

On montre que le second terme de (2.30) est nul en utilisant la propriété des schémas aux noeuds $\sum_r l_{jr}\mathbf{n}_{jr} = \mathbf{0}$. Par conséquent l'équation (2.30) se simplifie et donne

$$\sum_r l_{jr} \mathbf{n}_{jr} \otimes \mathbf{n}_{jr} (\mathbf{F}_{j,0} - \mathbf{F}_{r,0}) = \mathbf{0}. \tag{2.31}$$

Comme pour le premier schéma on utilise le développement de Hilbert dans le solveur nodal et sous condition de positivité de la matrice $\sum_j l_{jr}\mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)$, on obtient les deux relations suivantes

$$\left(\sum_j l_{jr} \mathbf{n}_{jr} \otimes \mathbf{n}_{jr} \right) \mathbf{F}_{r,0} + \sigma \left(\sum_j l_{jr} \mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j) \right) \mathbf{F}_{r,1} = \sum_j l_{jr} E_{j,0} \mathbf{n}_{jr} + \sum_j l_{jr} \mathbf{n}_{jr} \otimes \mathbf{n}_{jr} \mathbf{F}_{j,0} \tag{2.32}$$

et $\mathbf{F}_{r,0} = \mathbf{0}$.

Puisque $\mathbf{F}_{r,0} = \mathbf{0}$, alors la relation (2.31) est équivalente à $\sum_r l_{jr}\mathbf{n}_{jr} \otimes \mathbf{n}_{jr}\mathbf{F}_{j,0} = \mathbf{0}$. Il est trivial de montrer que la matrice $\sum_r l_{jr}\mathbf{n}_{jr} \otimes \mathbf{n}_{jr}$ est symétrique, définie positive pour un maillage non dégénéré, on peut se référer à la démonstration de C. Mazeran dans [Maz07], alors $\mathbf{F}_{j,0} = \mathbf{0}$. Ce résultat associé avec les relations (2.32) et (2.27) clôt la démonstration. □

Les précédentes démonstrations nécessitent que la matrice $\sum_j l_{jr}\mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)$ soit inversible. Nous allons donc étudier son inversibilité afin de vérifier si les schémas aux noeuds (2.11)-(2.12) et (2.22)-(2.23) sont correctement définis.

2.1.3 Étude de l'inversibilité de la matrice locale

Dans cette section nous donnons un critère sous optimal de positivité de la matrice locale du schéma de diffusion. Puis nous étudions ce critère dans le cas des triangles afin d'extraire une condition géométrique. La matrice locale issue des schémas JL-(a) (2.11)-(2.12) et JL-(b) (2.22)-(2.23) est la suivante

$$\sum_j l_{jr} \mathbf{n}_{jr} \otimes \mathbf{n}_{jr} + \frac{\sigma}{\varepsilon} \sum_j l_{jr} \mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j). \tag{2.33}$$

Son inversibilité est nécessaire à la bonne définition des schémas aux noeuds. Elle est composée de deux matrices. Une première symétrique et définie positive sur tous maillages non dégénérés (on peut trouver la preuve dans [Maz07]) :

$$\sum_j l_{jr}\mathbf{n}_{jr}\otimes\mathbf{n}_{jr}. \tag{2.34}$$

Une deuxième :

$$A_r=\sum_j l_{jr}\mathbf{n}_{jr}\otimes(\mathbf{x}_r-\mathbf{x}_j), \tag{2.35}$$

pour laquelle il n'existe pas de résultat d' inversibilité. On remarque que, sauf cas particulier, A_r n'est pas symétrique. Le but de cette section est d'étudier l'inversibilité et le caractère défini positif de cette matrice. On propose d'utiliser la définition suivante.

Définition 2.1.5. *Une matrice est définie positive si elle vérifie*

$$(\mathbf{y},A_r\mathbf{y})>0\quad\forall\mathbf{y}\in\mathbb{R}^2,\ \mathbf{y}\neq\mathbf{0}.$$

Remarque 2.1.2. *Pour les maillages triangulaires une possibilité existe afin de garantir que $A_r=A_r^t>0$. En effet, en supposant que l'orthocentre est à l'intérieur du triangle il suffit de choisir $\mathbf{x}_j$ égal à l'orthocentre. Dans ce cas $\mathbf{x}_r-\mathbf{x}_j=d_{jr}\mathbf{n}_{jr}$ avec $d_{jr}>0$. Par conséquent*

$$A_r=\sum_j l_{jr}d_{jr}\mathbf{n}_{jr}\otimes\mathbf{n}_{jr}=A_r^t.$$

Comme l'a prouvé C. Mazeran cette matrice est inversible car les normales engendrent $\mathbb{R}^2$ lorsque la somme sur les mailles j est suffisamment importante. Cependant, pour que l'orthocentre soit dans le triangle il faut que les angles soit inférieurs à $\pi/2$. Par conséquent cette solution ne nous permet d'obtenir des matrices inversibles que dans le cas de maillages triangulaires réguliers.

On a pour le moment échoué à obtenir un résultat aussi général que celui de C. Mazeran pour la matrice A_r. Cependant, on montre que si le maillage n'est pas trop déformé la matrice est définie positive. Pour cela on part de la constatation que $\mathrm{Tr}(A_r)=2\mid V_r\mid$ où $\mid V_r\mid$ est la mesure du volume de contrôle V_r (figure (2.2)) associé à un noeud $\mathbf{x}_r$.

Définition 2.1.6. *Le volume de contrôle V_r est le polygone défini par les points*

$$\ldots,\mathbf{x}_{j-\frac{1}{2}},\mathbf{x}_j,\mathbf{x}_{j+\frac{1}{2}},\ldots.$$

où $\mathbf{x}_j$ est le centre de la maille, $\mathbf{x}_{j+\frac{1}{2}}$ le centre des arêtes qui contiennent le noeud $\mathbf{x}_r$. Un exemple est donné avec la figure (2.2).

L'idée de l'étude est de comparer les matrices A_r et $\mid V_r\mid\widehat{I}_d$. Plus précisément il s'agit d'écrire A_r comme une perturbation de la matrice $\mid V_r\mid\widehat{I}_d$, puis de montrer que si le maillage n'est pas trop déformé, cette perturbation est suffisamment petite pour que A_r reste définie positive.

Proposition 2.3. *La matrice A_r satisfait :*

$$A_r=\mid V_r\mid\widehat{I}_d+\frac{1}{2}\sum_j\left(\mathbf{w}^{\perp}_{j+\frac{1}{2}}\otimes\mathbf{w}_{j+\frac{1}{2}}-\mathbf{v}^{\perp}_{j+\frac{1}{2}}\otimes\mathbf{v}_{j+\frac{1}{2}}\right)=\mid V_r\mid\widehat{I}_d+P, \tag{2.36}$$

avec $\mathbf{w}_{j+\frac{1}{2}}=(\mathbf{x}_{j+1}-\mathbf{x}_{j+\frac{1}{2}})$, $\mathbf{v}_{j+\frac{1}{2}}=(\mathbf{x}_{j+\frac{1}{2}}-\mathbf{x}_j)$ et

$$\mathrm{Tr}(A_r)=2\mid V_r\mid. \tag{2.37}$$

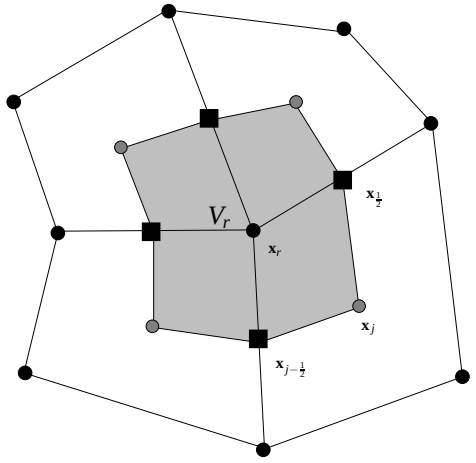


FIGURE 2.2 – Définition du volume de contrôle V_r autour du noeud $\mathbf{x}_r$. Le volume de contrôle est le polygone défini par les centres de maille $\mathbf{x}_j$ et les milieux des arêtes $\mathbf{x}_{j+\frac{1}{2}}$ autour du noeud.

Démonstration. Pour le polygone défini V_r par les points $\mathbf{x}_j$ et $\mathbf{x}_{j+\frac{1}{2}}$ autour du noeud $\mathbf{x}_r$, on peut écrire l'identité suivante qui est issue du théorème de Stokes.

$$\mid V_r \mid \widehat{I}_d = \sum_j (\mathbf{x}_{j+\frac{1}{2}} - \mathbf{x}_j)^\perp \otimes \left(\frac{1}{2}(\mathbf{x}_{j+\frac{1}{2}} + \mathbf{x}_j) - \mathbf{x}_r \right) + (\mathbf{x}_j - \mathbf{x}_{j-\frac{1}{2}})^\perp \otimes \left(\frac{1}{2}(\mathbf{x}_{j-\frac{1}{2}} + \mathbf{x}_j) - \mathbf{x}_r \right) \quad (2.38)$$

$$= \sum_j (\mathbf{x}_{j+\frac{1}{2}} - \mathbf{x}_j)^\perp \otimes \left(\frac{1}{2}(\mathbf{x}_{j+\frac{1}{2}} - \mathbf{x}_j) + \mathbf{x}_j - \mathbf{x}_r \right) + (\mathbf{x}_j - \mathbf{x}_{j-\frac{1}{2}})^\perp \otimes \left(\frac{1}{2}(\mathbf{x}_{j-\frac{1}{2}} - \mathbf{x}_j) + \mathbf{x}_j - \mathbf{x}_r \right). \quad (2.39)$$

Puisque $l_{jr}\mathbf{n}_{jr} = -(\mathbf{x}_{j+\frac{1}{2}} - \mathbf{x}_j)^\perp - (\mathbf{x}_j - \mathbf{x}_{j-\frac{1}{2}})^\perp$ on a

$$\mid V_r \mid \widehat{I}_d = A_r + \frac{1}{2} \sum_j (\mathbf{x}_{j+\frac{1}{2}} - \mathbf{x}_j)^\perp \otimes (\mathbf{x}_{j+\frac{1}{2}} - \mathbf{x}_j) - \sum_j (\mathbf{x}_j - \mathbf{x}_{j-\frac{1}{2}})^\perp \otimes (\mathbf{x}_j - \mathbf{x}_{j-\frac{1}{2}}), \quad (2.40)$$

qui se réécrit en utilisant la définition de P donnée par $A_r = \mid V_r \mid \widehat{I}_d + P$. Le second point est évident puisque pour tout vecteur $\mathbf{v}$ on a $\mathrm{Tr}(\mathbf{v}^\perp \otimes \mathbf{v}) = 0$. □

En introduisant le volume de contrôle V_{jr} associé à un noeud r et une maille j définie par $(\mathbf{x}_r, \mathbf{x}_j, \mathbf{x}_{j+\frac{1}{2}}, \mathbf{x}_{j+1})$ on peut écrire la matrice A_r sous la forme suivante :

$$A_r = \sum_j \left(|V_{jr}| \widehat{I}_d + \frac{1}{2} \left(\mathbf{w}_{j+\frac{1}{2}}^\perp \otimes \mathbf{w}_{j+\frac{1}{2}} - \mathbf{v}_{j+\frac{1}{2}}^\perp \otimes \mathbf{v}_{j+\frac{1}{2}} \right) \right) = \sum_j \left(|V_{jr}| \widehat{I}_d + P_j \right), \quad (2.41)$$

où si $\widetilde{V}_{jr}$ est défini par $(\mathbf{x}_r, \mathbf{x}_{j-\frac{1}{2}}, \mathbf{x}_j, \mathbf{x}_{j+\frac{1}{2}})$

$$A_r = \sum_j \left(|\widetilde{V}_{jr}| \widehat{I}_d + \frac{1}{2} \left(\mathbf{w}_{j-\frac{1}{2}}^\perp \otimes \mathbf{w}_{j-\frac{1}{2}} - \mathbf{v}_{j+\frac{1}{2}}^\perp \otimes \mathbf{v}_{j+\frac{1}{2}} \right) \right) = \sum_j \left(|\widetilde{V}_{jr}| \widehat{I}_d + P_j \right). \quad (2.42)$$

En utilisant les deux décompositions précédentes on peut obtenir des conditions sur les matrices P_j telle que la matrice A_r est définie positive.

Proposition 2.4. *Si pour toute maille j, l'une des inégalités suivantes est satisfaite*

$$|V_{jr}| > \frac{1}{2}||\mathbf{x}_{j+1} - \mathbf{x}_j||||\mathbf{x}_{j+\frac{1}{2}} - \frac{1}{2}(\mathbf{x}_{j+1} + \mathbf{x}_j)||, \tag{2.43}$$

$$|\widetilde{V}_{jr}| > \frac{1}{2}||\mathbf{x}_{j+\frac{1}{2}} - \mathbf{x}_{j-\frac{1}{2}}||||\mathbf{x}_j - \frac{1}{2}(\mathbf{x}_{j+\frac{1}{2}} + \mathbf{x}_{j-\frac{1}{2}})||, \tag{2.44}$$

alors la matrice A_r est définie positive.

Démonstration. Soit $\mathbf{x}$ un vecteur de $\mathbb{R}^2$. Pour la décomposition (2.41) (idem (2.42)) on obtient que

$$(\mathbf{x}, A_r\mathbf{x}) = \sum_j |V_{jr}|||\mathbf{x}||^2 + \sum_j (\mathbf{x}, P_j^s\mathbf{x}) \geq \sum_j \left(|V_{jr}| - \rho(P_j^s)\right)||\mathbf{x}||^2,$$

où $P_j^s = \frac{1}{2}(P_j + P_j^t)$ est la partie symétrique de P_j et $\rho(P_j^s)$ le rayon spectral de la matrice P_j^s. On constate donc immédiatement que si pour tout j $|V_{jr}| - \rho(P_j^s) > 0$, la matrice A_r est alors définie positive. Reste maintenant à étudier ce rayon spectral. On pose $C = \mathbf{w}^\perp \otimes \mathbf{w} - \mathbf{v}^\perp \otimes \mathbf{v}$ avec $\mathbf{v} = (a, b) \in \mathbb{R}^2$ et $\mathbf{w} = (\tilde{a}, \tilde{b}) \in \mathbb{R}^2$ alors

$$C^s = \begin{pmatrix} ab - \tilde{a}\tilde{b} & \dfrac{-a^2 + \tilde{a}^2 + b^2 - \tilde{b}^2}{2} \\ \dfrac{-a^2 + \tilde{a}^2 + b^2 - \tilde{b}^2}{2} & \tilde{a}\tilde{b} - ab \end{pmatrix}.$$

On définit $x = (a - \tilde{a})$, $y = (b - \tilde{b})$, $\alpha = \frac{1}{2}(b + \tilde{b})$ et $\beta = \frac{1}{2}(a + \tilde{a})$. Immédiatement on voit que $\mathrm{Tr}(C^s) = 0$. En utilisant la factorisation $ab - \tilde{a}\tilde{b} = x\alpha + y\beta$, on montre que

$$\mathrm{Det}(C^s) = -(x^2 + y^2)(\alpha^2 + \beta^2) = -||\mathbf{v} - \mathbf{w}||^2\, ||\frac{1}{2}(\mathbf{v} + \mathbf{w})||^2.$$

Puisque C^s est une matrice de taille deux à trace nulle on sait que :

$$\rho(C^s) = \lambda_{max} = \frac{1}{2}\left(\mathrm{Tr}(C^s) + \sqrt{\mathrm{Tr}(C^s)^2 - 4\,\mathrm{Det}(C^s)}\right) = ||\mathbf{w} - \mathbf{v}||\, ||\frac{1}{2}(\mathbf{v} + \mathbf{w})||.$$

quand on remplace $\mathbf{v}$ et $\mathbf{w}$ par les vecteurs définis dans la décomposition (2.41) de A_r, cela nous donne les critères de positivité (2.43). Le principe reste le même pour obtenir le critère (2.44) à partir de la décomposition (2.42). □

Remarque 2.1.3. *Pour des maillages composés de triangles équilatéraux ou de carrés, il est facile de vérifier que la matrice $A_r > 0$ car $\mathbf{x}_{j+\frac{1}{2}} = \frac{1}{2}(\mathbf{x}_{j+1} + \mathbf{x}_j)$, par conséquent l'inégalité (2.43) est vérifiée.*

Le critère obtenu est clairement sous optimal. En effet, dès le début de la preuve on voit que la minoration, n'est pas optimale. En outre on donne un critère local sur chacun des morceaux des décompositions (2.41) et (2.42). On peut penser que même si chacun des morceaux de la matrice A_r associé à chaque sous-maille ne respecte pas les précédents critères (2.41)-(2.42), la matrice pourrait être malgré tout, définie positive. Cela sera vérifié numériquement ultérieurement.

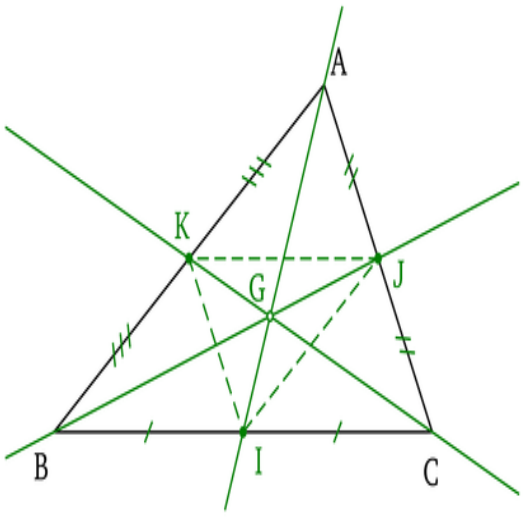


FIGURE 2.3 – Un triangle générique. Le barycentre est $G = \frac{1}{3}(A + B + C)$. Le milieu de l'arête BC est $I = \frac{1}{2}(B + C)$.

Inversibilité de la matrice nodale dans le cas triangulaire

Afin d'illustrer l'étude précédente, on propose dans le cas des triangles de passer des critères algébriques (2.43)-(2.44) à des critères géométriques. Le cas des triangles équilatéraux est traité dans la remarque (2.1.3). Pour obtenir ces conditions géométriques il suffit de caractériser les trois inégalités locales (2.44) (une par noeud) pour un triangle générique dépeint par la figure (2.3).

On définit trois angles. Le premier est θ_A l'angle entre le segment AG et KJ, le second θ_B est l'angle entre BG et KI et le dernier est θ_C l'angle entre CG et IJ (voir figure (2.3)).

Proposition 2.5. *Les trois inégalités locales (2.44) sont vérifiées si*

$$| \sin(\theta_{A,B,C}) | > \frac{1}{4}. \tag{2.45}$$

Démonstration. Par symétrie, il est suffisant d'étudier la condition de positivité (2.44) pour le sommet A par exemple, qui correspond à

$$\mathbf{x}_r = A, \quad \mathbf{x}_{j+\frac{1}{2}} = J = \frac{A+C}{2}, \quad \mathbf{x}_{j-\frac{1}{2}} = K = \frac{A+B}{2}, \quad \mathbf{x}_j = G = \frac{A+B+C}{3}.$$

Le quadrangle $\widetilde{V}_{jr}$ dans l'inégalité (2.44) correspond à $\widetilde{V}_{jr} = (A, K, G, J)$. On a

$$| \widetilde{V}_{jr} | = \frac{1}{2}||AG||||JK|| \, | \sin(\theta_A) | \, .$$

On voit que $||\mathbf{x}_{j+\frac{1}{2}} - \mathbf{x}_{j-\frac{1}{2}}|| = ||JK||$. Par conséquent on a

$$||\mathbf{x}_j - \frac{1}{2}(\mathbf{x}_{j+\frac{1}{2}} + \mathbf{x}_{j-\frac{1}{2}})|| = \left\| \frac{A+B+C}{3} - \frac{2A+B+C}{4} \right\|$$

$$= || - \frac{-2A + B + C}{12}|| = ||\frac{B + C}{2} - A||.$$

Comme G est le barycentre on sait que $||AG|| = \frac{2}{3}||AI||$. En outre $D = \frac{K + J}{2}$, ce qui donne

$$||AD|| = ||A - \frac{2A + B + C}{4}|| = \frac{1}{2}||AI||$$

$$\text{alors } ||AD|| = \frac{1}{2}||AI|| \text{ et } ||DG|| = \frac{1}{6}||AI||.$$

Par conséquent, l'inégalité (2.44) est équivalente à

$$\frac{1}{2}||AG||||JK|| \mid \sin(\theta_A) \mid > \frac{1}{2}||JK||||DG||.$$

En utilisant les relations $||AG|| = \frac{2}{3}||AI||$ et $||DG|| = \frac{1}{6}||AI||$ on obtient la condition $\mid \sin(\theta_A) \mid > \frac{1}{4}$. □

Afin de simplifier la suite on définit $\hat{A}$, $\hat{B}$ et $\hat{C}$ les angles associés aux sommets A,B et C.

Proposition 2.6. *On a la correspondance*

$$\mid \sin(\theta_A) \mid = \frac{1}{\sqrt{1 + \frac{\sin^2(\hat{B} - \hat{C})}{4 \sin^2 \hat{B} \sin^2 \hat{C}}}}. \tag{2.46}$$

Des formules similaires existent pour les deux autres angles.

Démonstration. Pour simplifier la preuve on choisit le sommet : $A = (x, y)$, $B = (0, 0)$, $C = (1, 0)$ et $D = (0.5, 0)$. On sait que $x = \cot(\hat{B})y$ et $1 - x = \cot(\hat{C})y$. Par conséquent

$$x = \frac{1}{\cot(\hat{B}) + \cot(\hat{C})} \text{ et } y = \frac{\cot(\hat{B})}{\cot(\hat{B}) + \cot(\hat{C})}.$$

À présent on peut remarquer que $x = \cot(\theta_A)y$, alors $\cot(\theta_A) = \frac{1}{2}(\cot(\hat{B}) - \cot(\hat{C}))$. Ceci est équivalent à

$$\cot(\theta_A) = \frac{\sin(\hat{C} - \hat{B})}{2 \sin(\hat{C}) \sin(\hat{B})}.$$

Pour conclure on utilise $1 + \cot^2(\theta_A) = \frac{1}{\sin^2(\theta_A)}$. □

On désire caractériser tous les angles de façon à satisfaire les conditions suffisantes de positivité (2.45) en utilisant la relation (2.46). On définit la fonction g par

$$g(\hat{B}, \hat{C}) = \frac{1}{\sqrt{1 + \frac{\sin^2(\hat{B} - \hat{C})}{4 \sin^2 \hat{B} \sin^2 \hat{C}}}}.$$

Puisque $\widehat{A} + \widehat{B} + \widehat{C} = \pi$, la condition (2.45) pour l'angle A se réécrit à l'aide de la relation (2.46) sous la forme

$$f(\hat{B}, \hat{C}) > \frac{1}{4}, \quad \text{avec } f(\hat{B}, \hat{C}) = \min\left(g(\hat{B}, \hat{C}), g(\pi - \hat{C} - \hat{B}, \hat{C}), g(\hat{B}, \pi - \hat{C} - \hat{B})\right).$$

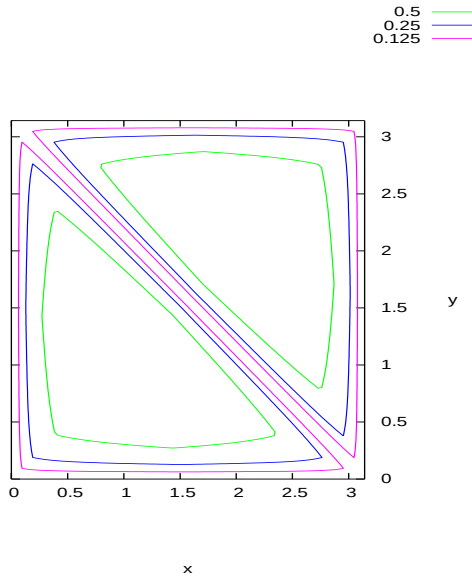


FIGURE 2.4 – Trace de la fonction f avec $x \in [0, \pi]$ et $y \in [0, \pi]$. La partie intéressante correspond à $x \geq 0$, $y \geq 0$ et $x + y \leq \pi$. Le domaine de positivité $f(x, y) > \frac{1}{4}$ est délimité par l'isoligne bleu $f(x, y) = \frac{1}{4}$, qui se trouve entre les deux autres.

On observe sur la figure (2.4) que si tous les angles sont strictement supérieurs à une valeur qui est proche de

$$\theta_{\text{ap}} = 0.2 \text{ rad}, \tag{2.47}$$

ce qui équivaut à peu près à 11 degrés, le critère de positivité (2.45) est alors respecté. Cette valeur correspond au coin gauche en bas de l'isoligne bleue, (figure 2.4). Finalement, dans le cadre d'un maillage triangulaire on obtient une condition géométrique assez peu restrictive. Maintenant il va s'agir de trouver une estimation plus précise de cette valeur. En effet l'inégalité (2.45) est équivalente à

$$\left| \frac{\sin(\hat{B} - \hat{C})}{\sin \hat{B} \sin \hat{C}} \right| < \sqrt{60} \Longleftrightarrow \left| \cot \hat{B} - \cot \hat{C} \right| < \sqrt{60}.$$

Par symétrie on obtient que $\left| \cot \hat{A} - \cot \hat{C} \right| < \sqrt{60}$ et $\left| \cot \hat{B} - \cot \hat{A} \right| < \sqrt{60}$.

$$\alpha = \cot \hat{A}, \ \beta = \cot \hat{B} \text{ et } \gamma = \cot \hat{C}. \tag{2.48}$$

Puisque $\hat{A}+\hat{B}+\hat{C} = \pi$ alors $\gamma = -\cot(\hat{A}+\hat{B}) = -\frac{\alpha\beta-1}{\alpha+\beta}$. Supposons que $\gamma \leq \beta \leq \alpha$. On obtient les deux inégalités suivantes $\alpha - \beta \leq \sqrt{60}$ et $\beta + \frac{\alpha\beta-1}{\alpha+\beta} \leq \sqrt{60}$. On considère le cas $\alpha = \beta$ and $\beta + \frac{\alpha\beta-1}{\alpha+\beta} = \sqrt{60}$. La solution est obtenue en résolvant l'équation quadratique $3\alpha^2 - 2\sqrt{60}\alpha - 1 = 0$. On obtient une valeur particulière $\alpha_+ = \beta_+ = \frac{\sqrt{60}+\sqrt{63}}{3}$: dans le cas $\gamma_- = \beta_+ - \sqrt{60} = \frac{-2\sqrt{60}+\sqrt{63}}{3}$. On définit l'angle θ_{opt} comme étant égal à

$$\cot \theta_{\text{opt}} = \frac{\sqrt{60} + \sqrt{63}}{3}. \tag{2.49}$$

Par construction $\cot(\pi - 2\theta_{\text{opt}}) = \gamma_-$.

Proposition 2.7. *Supposons que tous les angles satisfont*

$$\theta_{\text{opt}} < \hat{A}, \hat{B}, \hat{C} < \pi - 2\theta_{\text{opt}}. \tag{2.50}$$

avec θ_{opt} défini par (2.49) alors les trois inégalités (2.45) sont satisfaites.

À l'aide de l'équation (2.48) on montre que $\cot(\pi - 2\theta_{\text{opt}}) < \alpha, \beta, \gamma < \cot\theta_{\text{opt}}$. Donc

$$\frac{-2\sqrt{60} + \sqrt{63}}{3} < \alpha, \beta, \gamma < \frac{\sqrt{60} + \sqrt{63}}{3},$$

qui se transforme en $\max(|\alpha - \beta|, |\beta - \gamma|, |\gamma - \beta|) < \sqrt{60}$.

Remarque 2.1.4. *On note que*

$$\theta_{\text{opt}} \approx 0.189 \text{ rad.} \tag{2.51}$$

est une estimation plus précise que l'estimation graphique (2.47) donnée par l'examen de la figure 2.4.

Remarques sur l'inversibilité de la matrice nodale.

L'étude de l'inversibilité de la matrice nodale a permis d'exhiber des conditions sur la géométrie locale du maillage. À ce stade plusieurs remarques. D'abord, dans le cadre des triangles la condition de positivité suffisante est assez peu contraignante puisqu'elle admet tous les triangles dont les angles sont supérieurs à environ 11°. Les tests numériques effectués sur différents maillages montrent que même si les critères de positivité (2.43)-(2.44) ne sont pas respectés, la matrice reste définie positive. Numériquement nous observons que la matrice est définie positive sur tous les maillages utilisés. Afin de l'illustrer nous donnons les résultats suivants : soit λ_r la valeur propre minimale associée à la matrice A_r et donnée par

$$\lambda_r = \frac{1}{2}(\text{Tr}(A_r) - \sqrt{\text{Tr}(A_r)^2 - 4\,\text{Det}(A_r)}).$$

Le tableau (2.1) donne pour plusieurs maillages le $\min_r \lambda_r$ et $\max_r \lambda_r$ sur tous les noeuds du maillages. Pour ces résultats on prend 1600 mailles à chaque fois. Avec K un coefficient de déformation du maillage

Maillage	$\min \lambda_r$	$\max \lambda_r$
Maillage aléatoire quad.	0.001418	0.003128
Maillage de Kershaw, K=1	0.001288	0.003627
Maillage de Kershaw, K=1.9	0.0001523	0.00464155
Maillage aléatoire trig.	0.001163	0.004031

TABLE 2.1 – Valeurs minimale et maximale de λ_r pour plusieurs maillages.

de Kershaw. Le cas $K = 1$ correspond à la figure (2.7). Puisque la valeur propre minimale de A_r est toujours strictement positive sur ces maillages, la matrice est toujours définie positive sur ceux-ci.

La dernière remarque concerne la coercivité de la matrice nécessaire au théorème 2.11. Il s'agit de

$$||A_r^s|| \geq \alpha V_r, \tag{2.52}$$

où $A_r^s = \frac{A_r + A_r^t}{2}$ la partie symétrique A_r et $\alpha > 0$ une constante positive indépendante du maillage. Dans le cas de maillage triangulaire il suffit de modifier l'inégalité (2.50) par

$$\theta_{\text{opt}} < \theta_{\min} < \hat{A}, \hat{B}, \hat{C} < \pi - 2\theta_{\min} < \pi - 2\theta_{\text{opt}},$$

où $\theta_{\min}$ est un paramètre indépendant du pas de maillage h.

Cette étude, bien que donnant un résultat incomplet et sous optimal, nous donne une idée de l'ensemble des maillages admissibles dans le cas triangulaire. En outre les résultats numériques montrent que la méthode est bien plus robuste que ne le laisse présager ces résultats théoriques.

2.1.4 Variantes du schéma aux noeuds

Nous définissons un formalisme qui rend possible la construction de variantes pour les schémas JL-(a) et JL-(b). Ces variantes sont basées sur une formulation tensorielle des schémas aux noeuds introduite dans [Klu08]. Elle permet de proposer une discrétisation différente que ce soit pour la partie acoustique ou la partie liée au terme source dans les flux. Pour commencer on réécrit le schéma (2.11)-(2.12) sous la forme

$$\begin{cases} |\Omega_j| \, \partial_t E_j + \dfrac{1}{\varepsilon} \displaystyle\sum_r l_{jr}(\mathbf{F}_r, \mathbf{n}_{jr}) = 0, \\ |\Omega_j| \, \partial_t \mathbf{F}_j + \dfrac{1}{\varepsilon} \displaystyle\sum_r \mathbf{G}_{jr} = - |\Omega_j| \, \dfrac{\sigma}{\varepsilon^2} \mathbf{F}_j, \end{cases} \tag{2.53}$$

$$\begin{cases} \mathbf{G}_{jr} = l_{jr} E_j \mathbf{n}_{jr} + \widehat{\alpha}_{jr}(\mathbf{F}_j - \mathbf{F}_r) - \dfrac{\sigma}{\varepsilon} \widehat{\beta}_{jr} \mathbf{F}_r, \\ \left(\displaystyle\sum_j \widehat{\alpha}_{jr} + \dfrac{\sigma}{\varepsilon} \sum_j \widehat{\beta}_{jr} \right) \mathbf{F}_r = \displaystyle\sum_j l_{jr} E_j \mathbf{n}_{jr} + \sum_j \widehat{\alpha}_{jr} \mathbf{F}_j. \end{cases} \tag{2.54}$$

Le schéma (2.11)-(2.12) correspond au schéma (2.53)-(2.54) avec $\widehat{\beta}_{jr} = \widehat{\beta}_{jr}^I = l_{jr} \mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)$ et $\widehat{\alpha}_{jr} = \widehat{\alpha}_{jr}^I = l_{jr} \mathbf{n}_{jr} \otimes \mathbf{n}_{jr}$. Le système linéaire (2.54) est obtenu à l'aide de la propriété $\sum_j \mathbf{G}_{jr} = \mathbf{0}$.

Pour le tenseur associé au terme source on propose une deuxième solution $\widehat{\beta}_{jr}^{II} = \mid V_{jr} \mid \widehat{I}_d$ avec V_{jr} le volume de contrôle associé au noeud r et à la maille j. Par conséquent $\sum_j \widehat{\beta}_{jr} = \mid V_r \mid \widehat{I}_d$ est une matrice toujours inversible sur maillages non dégénérés. D'autres variantes seront définies par la suite. Pour la partie acoustique G. Kluth a montré que le schéma hydrodynamique EUCCLHYD (voir [MABO07]-[Mai11]) pouvait être incorporé dans ce formalisme. Nous proposons donc une seconde variante pour le tenseur acoustique qui permet de retomber sur le schéma EUCCLHYD. Le tenseur qui donne le schéma GLACE est (voir [Klu08])

$$\widehat{\alpha}_{jr}^I = l_{jr} \mathbf{n}_{jr} \otimes \mathbf{n}_{jr}. \tag{2.55}$$

Le tenseur qui donne le schéma EUCCLHYD est

$$\widehat{\alpha}_{jr}^{II} = \frac{1}{2 l_{jr}} \left(l_{j,r-1,r} \mathbf{n}_{j,r-1,r} \otimes \mathbf{n}_{j,r-1,r} + l_{j,r,r+1} \mathbf{n}_{j,r,r+1} \otimes \mathbf{n}_{j,r,r+1} \right), \tag{2.56}$$

avec $l_{j,r-1,r}$, $l_{j,r,r+1}$, $\mathbf{n}_{j,r-1,r}$ et $\mathbf{n}_{j,r,r+1}$ les longueurs et normales associées aux arêtes $[\mathbf{x}_{r-1}, \mathbf{x}_r]$ et $[\mathbf{x}_r, \mathbf{x}_{r+1}]$. Dans le cas linéaire les deux schémas EUCCLHYD et GLACE sont très proches, par contre

il peut exister des différences non négligeables dans le cas hydrodynamique. Dans le cadre de l'extension de ces schémas au modèle non linéaire M_1 équivalent à un système de la dynamique des gaz (chapitre 7), il est intéressant de pouvoir utiliser les deux variantes.

Remarque 2.1.5. *On peut introduire le même type de variante pour le schéma JL-(b) (2.22)-(2.23) en utilisant le schéma (2.53)-(2.54) avec le terme source* $-\frac{\sigma}{\varepsilon^2}\sum_r \widehat{\beta}_{jr}\mathbf{F}_r$.

Proposition 2.8. *Les schémas JL-(a) et JL-(b) ont la même limite asymptotique quelque soit le choix du tenseur acoustique* $\widehat{\alpha}_{jr}$. *Le schéma limite est entièrement déterminé par le tenseur* $\widehat{\beta}_{jr}$ *de la façon suivante*

$$\begin{cases} |\Omega_j|\,\partial_t E_j + \displaystyle\sum_r l_{jr}(\mathbf{F}_r, \mathbf{n}_{jr}) = 0, \\ \sigma\left(\displaystyle\sum_j \widehat{\beta}_{jr}\right)\mathbf{F}_r = \displaystyle\sum_j l_{jr} E_j \mathbf{n}_{jr}. \end{cases}$$

2.2 Étude théorique des schémas aux noeuds

On souhaite montrer la convergence du schéma JL-(b) dans tous les régimes. La première solution est de montrer que le schéma converge indépendement de ε (schémas uniformément « asymptotic preserving » au sens de la définition 1.3.4). La seconde consiste à démontrer successivement que le schéma JL-(b) converge lorsque ε tend vers zéro vers un schéma de diffusion convergent (schémas « asymptotic preserving » au sens de la définition 1.3.3, figure (1.1)). Dans cette section on étudiera des schémas continus en temps, afin de ce concentrer sur les problèmes de consistance. Le problème temporel sera traité intérieurement.

2.2.1 Résultats et hypothèses géométriques

On commence par rappeler des résultats d'estimations géométriques issues de [Des09]. On définit le pas de maillage $h = \max_j(\mathrm{diam}(\Omega_j))$. On commence par supposer que le maillage est régulier :

$$C_1 h \leq |\Omega_j| \leq C_2 h. \tag{2.57}$$

Par construction

$$||l_{jr}\mathbf{n}_{jr}||_2 \leq C\,\mathrm{diam}(\Omega_j) \leq C_3 h, \quad ||\widehat{\alpha}_{jr}||_2 \leq C_4 h$$

. En effet $\widehat{\alpha}_{jr} = \frac{l_{jr}\mathbf{n}_{jr}\otimes l_{jr}\mathbf{n}_{jr}}{l_{jr}}$, or la norme d'un produit tensoriel d'un vecteur avec lui même est le carré de la norme du vecteur. Un produit tensoriel de ce type admet un déterminant nul donc ρ le rayon spectral est égal à la trace. On peut immédiatement constater que la trace est égale au carré de la norme du vecteur. Donc

$$||\widehat{\alpha}_{jr}||_2 = ||l_{jr}\mathbf{n}_{jr}||_2,$$

d'où le résultat. Enfin on s'intéresse à la matrice A_r.

Proposition 2.9. *On suppose que le volume de contrôle* V_r *satisfait*

$$C_1 h^2 \leq |\ V_r\ | \leq C_2 h^2, \tag{2.58}$$

alors

$$||A_r||_2 \leq C h^2.$$

Démonstration. On part du résultat acquis dans le cas de l'étude de cette matrice (2.41) . On sait que

$$A_r = V_r \widehat{I}_d + \frac{1}{2}\sum_j (\mathbf{w}^{\perp}_{j+\frac{1}{2}} \otimes \mathbf{w}_{j+\frac{1}{2}} - \mathbf{v}^{\perp}_{j+\frac{1}{2}} \otimes \mathbf{v}_{j+\frac{1}{2}}),$$

avec $\mathbf{w}_{j+\frac{1}{2}} = (\mathbf{x}_{j+1} - \mathbf{x}_{j+\frac{1}{2}})$ et $\mathbf{v}_{j+\frac{1}{2}} = (\mathbf{x}_{j+\frac{1}{2}} - \mathbf{x}_j)$. En outre on a montré que

$$||\mathbf{w}^{\perp}_{j+\frac{1}{2}} \otimes \mathbf{w}_{j+\frac{1}{2}} - \mathbf{v}^{\perp}_{j+\frac{1}{2}} \otimes \mathbf{v}_{j+\frac{1}{2}}||_2 = ||\mathbf{w}_{j+\frac{1}{2}} - \mathbf{v}_{j+\frac{1}{2}}||\, ||\frac{1}{2}(\mathbf{v}_{j+\frac{1}{2}} + \mathbf{w}_{j+\frac{1}{2}})||.$$

Il est évident que $||\mathbf{w}_{j+\frac{1}{2}}|| \leq C_3 h$ et que $||\mathbf{v}_{j+\frac{1}{2}}|| \leq C_4 h$. On utilise deux fois l'inégalité triangulaire afin d'obtenir la majoration

$$||\mathbf{w}^{\perp}_{j+\frac{1}{2}} \otimes \mathbf{w}_{j+\frac{1}{2}} - \mathbf{v}^{\perp}_{j+\frac{1}{2}} \otimes \mathbf{v}_{j+\frac{1}{2}}|| \leq \frac{1}{2}(||\mathbf{v}_{j+\frac{1}{2}}|| + ||\mathbf{w}_{j+\frac{1}{2}}||)^2 \leq C_5 h^2.$$

Il suffit d'utiliser l'hypothèse sur le volume de contrôle pour conclure. □

On considérera par la suite que les quantités $||l_{jr}\mathbf{n}_{jr}||$ et $||\widehat{\alpha}_{jr}||$ sont homogènes à $O(h)$ et que $||A_r||$ est homogène à $O(h^2)$.

2.2.2 Convergence du schéma limite de diffusion

Les schémas construits ne sont « asymptotic preserving » que si le schéma limite est un schéma convergent pour la diffusion. Cette preuve utilise une approche différente des preuves classiques multi-D et est plus proche des méthodes de preuves pour les volumes finis 1D. On considère le schéma semi-discret

$$\left\{ \begin{array}{ll} \partial_t E_j - \dfrac{1}{|\Omega_j|}\displaystyle\sum_r l_{jr}\left(\mathbf{F}_r, \mathbf{n}_{jr}\right) = 0, & \\ A_r \mathbf{F}_r = \displaystyle\sum_j l_{jr} E_j \mathbf{n}_{jr}, & \text{avec } A_r = -\left(\sum_j l_{jr}\mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)\right). \end{array} \right. \tag{2.59}$$

Avant de commencer quelques remarques et hypothèses. Tout d'abord : pour la preuve le signe a été changé devant les opérateurs discrets par rapport à (2.13), cela nous permettra de montrer que $\mathbf{F}_r$ est une approximation de $\nabla E(\mathbf{x}_r)$ et non de $-\nabla E(\mathbf{x}_r)$, ce qui est plus agréable. On choisit $\sigma = 1$. Afin de simplifier la preuve et de ne pas prendre en compte les conditions limites le domaine $D = (\mathcal{T})$ de calcul est un tore. Enfin on suppose que la solution est régulière

$$E(t, \mathbf{x}) \in C^1\left([0,T], W^{3,\infty}(D)\right). \tag{2.60}$$

La technique de preuve est basée sur une étude de consistance des flux. En effet les méthodes de volumes sur maillages non uniformes ne sont pas consistantes au sens des différences finies. Puisqu'il n'existe pas un équivalent du théorème de Lax dans le cadre d'une étude de consistance de flux, on démontrera pour le schéma (2.59) que la consistance des flux et la stabilité impliquent la convergence. On commence par définir les erreurs de consistance suivantes

$$\left\{ \begin{array}{l} a_j(t) = \partial_t E(t, \mathbf{x}_j) - \dfrac{1}{|\Omega_j|}\displaystyle\sum_r l_{jr}\left(\nabla E(t, \mathbf{x}_r), \mathbf{n}_{jr}\right), \\ \mathbf{b}_r(t) = \dfrac{1}{V_r}\left(A_r \nabla E(t, \mathbf{x}_r) - \displaystyle\sum_j l_{jr} E(t, \mathbf{x}_j)\mathbf{n}_{jr}\right). \end{array} \right. \tag{2.61}$$

Lemme 2.10. *Il existe une constante $C > 0$ telle que les estimations suivantes soit vérifiées*

$$|a_j| \leq Ch \quad \textit{pour tout } j, \tag{2.62}$$

et

$$||\mathbf{b}_r|| \leq Ch, \quad \textit{pour tout } r. \tag{2.63}$$

Démonstration. Par construction on sait que

$$\partial_t E(t, \mathbf{x}_j) = \frac{\int_{\Omega_j} \partial_t E(t, \mathbf{x}_j) d\mathbf{x}}{|\Omega_j|} + O(h) = \frac{\int_{\Omega_j} \Delta E(t, \mathbf{x}) d\mathbf{x}}{|\Omega_j|} + O(h) = \frac{1}{|\Omega_j|} \int_{\partial\Omega_j} \partial_n E(t, \mathbf{x}) d\sigma + O(h).$$

Par définition de $l_{jr}\mathbf{n}_{jr}$ on a

$$\sum_r l_{jr} \left(\nabla E(t, \mathbf{x}_r), \mathbf{n}_{jr} \right) = \sum_k \int_{\partial\Omega_{jk}} \left(\frac{\nabla E\left(t, \mathbf{x}_{jk}^+\right) + \nabla E\left(t, \mathbf{x}_{jk}^-\right)}{2}, \mathbf{n}_j \right) d\sigma,$$

où les noeuds $\mathbf{x}_{jk}^+$ et $\mathbf{x}_{jk}^-$ sont les extrémités du segment $\partial\Omega_{jk}$. On en déduit que

$$a_j = O(h) + \frac{1}{|\Omega_j|} \sum_k \int_{\partial\Omega_{jk}} \left(\nabla E - \frac{\nabla E\left(t, \mathbf{x}_{jk}^+\right) + \nabla E\left(t, \mathbf{x}_{jk}^-\right)}{2}, \mathbf{n}_j \right) d\sigma. \tag{2.64}$$

Puisque la fonction dans l'intégrale de (2.64) est approximée par la formule des trapèzes, on en déduit que

$$\left| \int_{\partial\Omega_{jk}} \left(\nabla E - \frac{\nabla E\left(t, \mathbf{x}_{jk}^+\right) + \nabla E\left(t, \mathbf{x}_{jk}^-\right)}{2}, \mathbf{n}_j \right) d\sigma \right| \leq Ch^2 \left|\Sigma_{jk}\right| \leq Ch^3. \tag{2.65}$$

En divisant (2.65) par $|\Omega_j|$ et en utilisant les hypothèses de régularité sur la quantité E, on obtient que $a_j = O(h)$. Le second terme est défini par

$$\mathbf{b}_r = \frac{1}{|\ V_r\ |} \left(\left(\sum_j l_{jr} \mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j) \right) \nabla E(t, \mathbf{x}_r) + \sum_j l_{jr} E(t, \mathbf{x}_j) \mathbf{n}_{jr} \right)$$

$$= \frac{1}{|\ V_r\ |} \sum_j \left(l_{jr} \left(\mathbf{x}_r - \mathbf{x}_j, \nabla E(t, \mathbf{x}_r) \right) + E(t, \mathbf{x}_j) \right) \mathbf{n}_{jr}.$$

À l'aide d'un simple développement de Taylor on montre que

$$E(t, \mathbf{x}_j) + (\mathbf{x}_r - \mathbf{x}_j, \nabla E(t, \mathbf{x}_r)) = E(t, \mathbf{x}_r) + O_{jr}(h^2),$$

alors

$$\mathbf{b}_r = \frac{1}{|\ V_r\ |} \sum_j O_{jr}(h^2) l_{jr} \mathbf{n}_{jr} = O(h).$$

□

Maintenant que le résultat de consistance des flux (2.62)-(2.63) est établi on propose de démontrer la convergence du schéma en norme L^2. On définit les erreurs

$$e_j(t) = E_j(t) - E(t, \mathbf{x}_j) \text{ et } \mathbf{f}_r(t) = \mathbf{F}_r(t) - \nabla E(t, \mathbf{x}_r).$$

avec $E_j(t)$, $\mathbf{F}_j(t)$ la solution numérique en espace et $E(t, \mathbf{x}_j)$, $E(t, \mathbf{x}_r)$ la solution exacte. L'erreur e_j est définie de la façon suivante

$$e(t, \mathbf{x}) = e_j(t) \quad \forall \mathbf{x} \in \Omega_j.$$

l'erreur de la $\mathbf{f}_r$ est définie par

$$\mathbf{f}(t, \mathbf{x}) = \mathbf{f}_r(t) \quad \forall \mathbf{x} \in V_r.$$

Ces définitions permet d'introduire les erreurs numériques

$$E_r(t) = \|e(t)\|^2_{L^2(\Omega)} = \left(\sum_j |\Omega_j| \left(E_j(t) - E(t, \mathbf{x}_j) \right)^2 \right).$$

et

$$F_r(t) = \|\mathbf{f}\|^2_{L^2([0,t]\times\Omega)} = \left(\int_0^t \sum_r \mid V_r \mid |\mathbf{F}_r(t) - \nabla E(t, \mathbf{x}_r)|^2 \, dt \right).$$

Théorème 2.11. *Supposons $E(t, \mathbf{x}) \in C^1\left([0, T], W^{3,\infty}(\Omega)\right)$. Supposons qu'il existe une constante $\alpha > 0$ telle que pour tout r*

$$||A^s_r|| \geq \alpha \mid V_r \mid \qquad \text{où } A^s_r = \frac{A_r + A^t_r}{2} \text{ est la partie symétrique de } A_r. \tag{2.66}$$

Alors le schéma limite de diffusion est convergent pour tout temps $T > 0$. C'est à dire qu'il existe $C(T) > 0$ tel que

$$\|e(t)\|_{L^2(\Omega)} + \|\mathbf{f}\|_{L^2([0,T]\times\Omega)} \leq C(T)h, \qquad 0 < t \leq T. \tag{2.67}$$

Démonstration. Par construction on a

$$\begin{cases} \partial_t e(t, \mathbf{x}_j) - \dfrac{1}{\mid \Omega_j \mid} \displaystyle\sum_r l_{jr} \left(\mathbf{n}_{jr}, \mathbf{f}_r(t) \right) = a_j(t), \\ A_r \mathbf{f}_r(t) + \displaystyle\sum_j l_{jr} \mathbf{n}_{jr} e_j(t) = V_r \mathbf{b}_r(t). \end{cases} \tag{2.68}$$

On a l'identité

$$E'_r(t) = 2 \sum_j \mid \Omega_j \mid e_j(t) e'_j(t) = \sum_j e_j(t) \left(\left(\sum_r l_{jr} \left(\mathbf{n}_{jr}, \mathbf{f}_r(t) \right) \right) + |\Omega_j| \, a_j(t) \right)$$

$$= 2 \sum_j |\Omega_j| \, e_j(t) a_j(t) + 2 \sum_j e_j(t) \left(\sum_r l_{jr} \left(\mathbf{n}_{jr}, \mathbf{f}_r(t) \right) \right)$$

$$= 2 \left(e(t), a(t) \right)_{L^2(\Omega)} + 2 \sum_r \left(\sum_j l_{jr} \mathbf{n}_{jr} e_j(t), \mathbf{f}_r(t) \right)$$

$$= 2 \left(e(t), a(t) \right)_{L^2(\Omega)} + 2 \left(\mathbf{f}(t), \mathbf{b}(t) \right)_{L^2(\Omega)} - 2 \sum_r \left(A_r \mathbf{f}_r(t), \mathbf{f}_r(t) \right)$$

$$= 2\left(e(t), a(t)\right)_{L^2(\Omega)} + 2\left(\mathbf{f}(t), \mathbf{b}(t)\right)_{L^2(\Omega)} - 2\sum_r \left(A_r^s \mathbf{f}_r(t), \mathbf{f}_r(t)\right).$$

L' hypothèse sur la matrice locale permet de dire que $\sum_r \left(A_r^s \mathbf{f}_r(t), \mathbf{f}_r(t)\right) \geq \alpha \|\mathbf{f}(t)\|^2$. Donc

$$E_r'(t) \leq ||e(t)||^2 + ||a(t)||^2 + 2\left(\frac{\mu||\mathbf{f}(t)||^2 + \mu^{-1}||\mathbf{b}(t)||^2}{2}\right) - 2\alpha\|\mathbf{f}(t)\|^2, \qquad \forall \mu > 0.$$

En choisissant $\mu = \alpha$, on obtient

$$E_r'(t) \leq \frac{1}{2}E_r(t) - \alpha F_r(t) + ||a(t)||^2 + \frac{1}{\alpha}||\mathbf{b}(t)||^2 \leq \frac{1}{2}E_r(t) + Ch^2, \tag{2.69}$$

grâce aux erreurs de consistance (2.62)-(2.63) et au fait que le domaine D soit borné. Par construction $E_r(0) = O(h^2)$, donc en utilisant le lemme de Grönwall $E_r(t) \leq Cth^2$ qui prouve que $||a(t)|| \leq C'th$ pour toutes constantes C'. Enfin il suffit d'intégrer (2.69) sur intervalle de temps $[0, T]$ pour obtenir

$$E_r(T) + \alpha F_r(T) \leq E_r(0) + \int_0^T \left(\frac{1}{2}E_r(t) + Ch^2\right) dt \leq C''Th^2,$$

ce qui permet de déduire l'estimation de convergence finale (2.67). □

Le schéma limite converge sur maillages non structurés aux moins pour des solutions régulières. Ce schéma est un nouveau schéma, bien qu'il soit assez proche du schéma MPFA [AE06]- [AM08] ou du schéma à 9 points [Ker81].

2.2.3 Étude de stabilité des schémas « asymptotic preserving »

On étudie la stabilité L^2 des schémas JL-(a) (2.11)-(2.12) et JL -(b) (2.22)-(2.23) que nous avons définis précédemment. On suppose pour simplifier que le domaine est un tore. On considère ici les schémas semi-discrets en temps.

Proposition 2.12. *Le schéma semi-discret JL-(a) avec $\widehat{\beta}_{jr} = \mid V_{jr} \mid \widehat{I}_d$ et $\widehat{\alpha}_{jr}$ donné par (2.55) où (2.56) est stable en norme L^2.*

Démonstration. Le schéma semi-discret est donné par

$$\begin{cases} \mid \Omega_j \mid \partial_t E_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r l_{jr}(\mathbf{F}_r, \mathbf{n}_{jr}) = 0, \\ \mid \Omega_j \mid \partial_t \mathbf{F}_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r \mathbf{G}_{jr} = -\mid \Omega_j \mid \dfrac{\sigma}{\varepsilon^2}\mathbf{F}_j. \end{cases} \tag{2.70}$$

On commence par définir la norme L^2 associée au schéma (2.70)

$$E_r(t) = \frac{1}{2}\int_\Omega \left(\mid E_h(t) \mid^2 + (\mathbf{F}_h, \mathbf{F}_h)\right) d\mathbf{x}.$$

Si la norme L^2 décroît au cours du temps, le schéma semi-discret est stable en norme L^2. Par conséquent cela revient à prouver que $E_r^{'}(t) \leq 0$. On a

$$E_r^{'}(t) = \frac{1}{2}\int_\Omega \frac{d}{dt}\left(\mid E_h(t) \mid^2 + (\mathbf{F}_h(t), \mathbf{F}_h(t))\right) d\mathbf{x} = \int_\Omega E_h(t)\partial_t E_h(t) + (\mathbf{F}_h(t), \partial_t \mathbf{F}_h(t)) d\mathbf{x}$$

$$= \sum_j |\Omega_j| \left(E_j(t) \partial_t E_j(t) + (\mathbf{F}_j(t), \partial_t \mathbf{F}_j(t)) \right).$$

En utilisant la définition du schéma (2.70) on obtient

$$E_r'(t) = -\frac{1}{\varepsilon} \sum_j \sum_r l_{jr} E_j(\mathbf{F}_r, \mathbf{n}_{jr}) - \frac{1}{\varepsilon} \sum_j \sum_r (\mathbf{G}_{jr}, \mathbf{F}_j) - \frac{\sigma}{\varepsilon^2} \sum_j |\Omega_j| \, (\mathbf{F}_j, \mathbf{F}_j). \tag{2.71}$$

On développe le second terme de la précédente équation à l'aide de la formule de flux (2.23) afin d'obtenir

$$\sum_j \sum_r (\mathbf{G}_{jr}, \mathbf{F}_j) = \sum_j \sum_r l_{jr} E_j(\mathbf{F}_j, \mathbf{n}_{jr}) + \sum_j \sum_r (\widehat{\alpha}_{jr}(\mathbf{F}_j - \mathbf{F}_r), \mathbf{F}_j) - \frac{\sigma}{\varepsilon} \sum_j \sum_r (\widehat{\beta}_{jr} \mathbf{F}_r, \mathbf{F}_j). \tag{2.72}$$

Puisque $\sum_r l_{jr} \mathbf{n}_{jr} = \mathbf{0}$ (propriété fondamentale des schémas aux noeuds, qui permet d'obtenir un schéma conservatif) le premier terme de (2.72) est nul. En sommant sur r le terme (2.54) définissant le solveur nodal et en permutant les sommes, on montre que

$$\sum_j \sum_r l_{jr} E_j(\mathbf{F}_r, \mathbf{n}_{jr}) = \sum_j \sum_r ((\widehat{\alpha}_{jr} + \frac{\sigma}{\varepsilon} \widehat{\beta}_{jr}) \mathbf{F}_r, \mathbf{F}_r) - \sum_j \sum_r (\widehat{\alpha}_{jr} \mathbf{F}_j, \mathbf{F}_r). \tag{2.73}$$

Pour continuer il suffit d'incorporer les relations (2.72) et (2.73) dans l'équation (2.71) et de permuter les sommes dans $E_r'(t)$, ce qui donne

$$E_r'(t) = -\frac{1}{\varepsilon} \sum_j \sum_r (\widehat{\alpha}_{jr}(\mathbf{F}_j - \mathbf{F}_r), \mathbf{F}_j - \mathbf{F}_r) - \frac{\sigma}{\varepsilon^2} \sum_r \sum_j (\widehat{\beta}_{jr} \mathbf{F}_r, \mathbf{F}_r - \mathbf{F}_j) - \frac{\sigma}{\varepsilon^2} \sum_j |\Omega_j| \, (\mathbf{F}_j, \mathbf{F}_j). \tag{2.74}$$

Dans la suite de la démonstration on considère le tenseur $\widehat{\alpha}_{jr}^I$ (2.55), mais la démonstration reste identique avec le deuxième tenseur $\widehat{\alpha}_{jr}^{II}$ (2.56). En effet, c'est la propriété de symétrie du tenseur qui permet d'arriver au résultat. On obtient donc

$$E_r'(t) = -\frac{1}{\varepsilon} \sum_r \sum_j l_{jr} (\mathbf{n}_{jr}, (\mathbf{F}_j - \mathbf{F}_r))^2 - \frac{\sigma}{\varepsilon^2} \sum_r \sum_j (\widehat{\beta}_{jr} \mathbf{F}_r, \mathbf{F}_r - \mathbf{F}_j) - \frac{\sigma}{\varepsilon^2} \sum_j |\Omega_j| \, (\mathbf{F}_j, \mathbf{F}_j). \tag{2.75}$$

Pour conclure, il suffit d'utiliser le fait que $\widehat{\beta}_{jr} = |V_{jr}| \widehat{I}_d$ et que $\sum_r |V_{jr}| = |\Omega_j|$. Par conséquent on peut réécrire le second terme et le troisième terme de l'équation (2.75) sous la forme suivante

$$\frac{\sigma}{\varepsilon^2} \sum_j \sum_r \mid V_{jr} \mid ((\mathbf{F}_r, \mathbf{F}_r) + (\mathbf{F}_j, \mathbf{F}_j) - (\mathbf{F}_r, \mathbf{F}_j))$$

$$= \frac{\sigma}{2\varepsilon^2} \sum_j \sum_r \mid V_{jr} \mid (\|\mathbf{F}_r\|^2 + \|\mathbf{F}_j\|^2 + \|\mathbf{F}_r - \mathbf{F}_j\|^2) \geq 0.$$

Par conséquent $E_r'(t) \leq 0$. □

Proposition 2.13. *Le schéma semi-discret JL-(b) avec $\widehat{\beta}_{jr} =\mid V_{jr} \mid \widehat{I}_d$ ou $\widehat{\beta}_{jr} = l_{jr} \mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)$ et $\widehat{\alpha}_{jr}$ donné par (2.55) où (2.56) est stable en norme L^2.*

Dans le cas où $\widehat{\beta}_{jr} = l_{jr} \mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)$, on fait l'hypothèse que $A_r = \sum_j \widehat{\beta}_{jr}$ est positive. Des conditions suffisantes pour cela sont données dans la section 2.3.2.

Démonstration. La première partie de la preuve est exactement la même que pour le schéma JL-(a). On part de l'estimation

$$E_r^{'}(t) = -\frac{1}{\varepsilon}\sum_j\sum_r l_{jr}(\mathbf{n}_{jr},(\mathbf{F}_j-\mathbf{F}_r))^2 - \frac{\sigma}{\varepsilon^2}\sum_j\sum_r(\widehat{\beta}_{jr}\mathbf{F}_r,\mathbf{F}_r-\mathbf{F}_j) - \frac{\sigma}{\varepsilon^2}\sum_j|\Omega_j|\,(\mathbf{F}_r,\mathbf{F}_j). \quad (2.76)$$

Dans le cas de ces deux tenseurs on a $|\Omega_j| = \sum_r \widehat{\beta}_{jr}$. En effet si $\widehat{\beta}_{jr} = \widehat{I}_d \mid V_{jr} \mid$ c'est immédiat. Si $\widehat{\beta}_{jr} = l_{jr}\mathbf{n}_{jr}\otimes(\mathbf{x}_r-\mathbf{x}_j)$ on peut le montrer immédiatement en utilisant l'identité suivante

$$\sum_r l_{jr}\mathbf{n}_{jr}\otimes\mathbf{x}_r = \widehat{I}_d\,|\Omega_j|\,,$$

et $\sum_r l_{jr}\mathbf{n}_{jr} = \mathbf{0}$. Cette propriété nous permet d'obtenir

$$E_r^{'}(t) = -\frac{1}{\varepsilon}\sum_j\sum_r l_{jr}(\mathbf{n}_{jr},\mathbf{F}_j-\mathbf{F}_r))^2 - \frac{\sigma}{\varepsilon^2}\sum_j\sum_r(\widehat{\beta}_{jr}\mathbf{F}_r,\mathbf{F}_r). \quad (2.77)$$

Pour conclure il suffit de permuter les sommes. Puisque $\sum_j \widehat{\beta}_{jr}$ est positive alors $E_r^{'}(t)\leq 0$. □

Remarque 2.2.1. *La stabilité L^2 du schéma semi-discret implique naturellement la stabilité dans la même norme du schéma implicite.*

2.2.4 Convergence du schéma JL-(b) vers le schéma de diffusion

Dans cette section on souhaite montrer que pour un maillage donné, les solutions discrètes du schéma JL-(b) convergent vers les solutions du schéma de diffusion quand ε tend vers zéro. On suppose que le maillage satisfait les hypothèses (2.57)-(2.58). Afin de simplifier les notations on pose $B_j = \sum_r \widehat{\alpha}_{jr}$, $B_r = \sum_j \widehat{\alpha}_{jr}$, $A_r = \sum_j \widehat{\beta}_{jr}$. On part du schéma semi-discret JL-(b) sous la forme suivante

$$\begin{cases} \mid\Omega_j\mid \partial_t E_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r(l_{jr}\mathbf{n}_{jr},\mathbf{F}_r) = 0,\\ \mid\Omega_j\mid \partial_t \mathbf{F}_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r(l_{jr}E_j\mathbf{n}_{jr} + \widehat{\alpha}_{jr}(\mathbf{F}_j-\mathbf{F}_r)) = \mathbf{0},\\ \Big(B_r + \dfrac{\sigma}{\varepsilon}A_r\Big)\mathbf{F}_r - \displaystyle\sum_i l_{ir}E_i\mathbf{n}_{ir} - \sum_i \widehat{\alpha}_{ir}\mathbf{F}_i = \mathbf{0},\\ E_j(t=0) = E_j^0,\\ \mathbf{F}_j(t=0) = \mathbf{0}. \end{cases} \quad (2.78)$$

Ensuite on introduit la limite de diffusion. On définit les inconnues suivantes $(\tilde{E}_j,\tilde{\mathbf{F}}_j,\tilde{\mathbf{F}}_r)$ solution du système

$$\begin{cases} \mid\Omega_j\mid \partial_t\tilde{E}_j + \dfrac{1}{\sigma}\displaystyle\sum_r(l_{jr}\mathbf{n}_{jr}, A_r^{-1}\sum_i l_{ir}\tilde{E}_i\mathbf{n}_{ir}) = 0,\\ B_j\tilde{\mathbf{F}}_j - \displaystyle\sum_r \widehat{\alpha}_{jr}\tilde{\mathbf{F}}_r = \mathbf{0},\\ \dfrac{\sigma}{\varepsilon}A_r\tilde{\mathbf{F}}_r - \displaystyle\sum_i l_{ir}\tilde{E}_i\mathbf{n}_{ir} = \mathbf{0},\\ \tilde{E}_j(t=0) = E_j^0. \end{cases} \quad (2.79)$$

Avant d'introduire le résultat de convergence et la preuve, on définit les notations

$$\begin{cases} \mathbf{E}_c = (E_{j=1}(t), ...E_{j=n}(t)), \\ \mathbf{F}_c = (\mathbf{F}_{j=1}(t), ...\mathbf{F}_{j=n}(t)), \\ \mathbf{F}_n = (\mathbf{F}_{r=1}(t), ...\mathbf{F}_{r=nv}(t)), \\ \mathbf{e}_c = (e_{E,j=1}, ..., e_{E,j=n}, \mathbf{e}_{F,j=1}, ..., \mathbf{e}_{F,j=n}), \\ \mathbf{e}_n = (\mathbf{e}_{F,r=1}, ..., \mathbf{e}_{F,r=nv}), \end{cases}$$

avec $e_{E,j} = E_j(t) - \tilde{E}_j(t)$, $\mathbf{e}_{F,j} = \mathbf{F}_j(t) - \tilde{\mathbf{F}}_j(t)$, n le nombre de mailles, nv le nombre de noeuds. On introduit les mêmes définitions pour les quantités $\tilde{\mathbf{E}}_c$, $\tilde{\mathbf{F}}_c$, $\tilde{\mathbf{F}}_n$. On définit aussi la norme discrète associée à ces quantités. Par exemple pour la quantité $\mathbf{E}_c$ la norme L^2 discrète s'écrit

$$||\mathbf{E}_c||^2_{L^2} = \sum_j \mid \Omega_j \mid\mid E_j \mid^2 .$$

Pour la quantité $\mathbf{F}_n$ la norme L^2 s'écrit

$$||\mathbf{F}_n||^2_{L^2} = \sum_r \mid V_r \mid (\mathbf{F}_r, \mathbf{F}_r).$$

Une preuve de stabilité analogue à celle proposée pour le schéma JL-(b) permet de montrer que le schéma de diffusion correspondant à la première équation de (2.79) est stable en norme L^2. Par conséquent $||\tilde{\mathbf{E}}_c(t)||_{L^2} \leq ||\tilde{\mathbf{E}}_c(0)||_{L^2}$.

Lemme 2.14. *On suppose que*

$$B_j = \sum_r \alpha_{jr},\ A_r = \sum_j \beta_{jr}\ \textit{et}\ B_r + \frac{\sigma_r}{\varepsilon} A_r.$$

sont inversibles. Les quantités $\tilde{\mathbf{F}}_n$, $\tilde{\mathbf{F}}_c$ sont majorées par les estimations suivantes

$$||\tilde{\mathbf{F}}_n||_{L^2} \leq \frac{\varepsilon}{\sigma} h C_n ||\tilde{\mathbf{E}}_c||_{L^2}, \tag{2.80}$$

et

$$||\tilde{\mathbf{F}}_c||_{L^2} \leq \frac{\varepsilon}{\sigma} h C_{n_2} ||\tilde{\mathbf{E}}_c||_{L^2}, \tag{2.81}$$

Démonstration. On définit la norme L^2 discrète du vecteur formé par les quantités $\tilde{\mathbf{F}}_r$

$$||\tilde{\mathbf{F}}_n||^2_{L^2} = \sum_r \mid V_r \mid (\tilde{\mathbf{F}}_r, \tilde{\mathbf{F}}_r).$$

On injecte la définition de $\tilde{\mathbf{F}}_r$ donnée par (2.79) pour obtenir

$$||\tilde{\mathbf{F}}_n||^2_{L^2} = \frac{\varepsilon^2}{\sigma^2} \sum_r \mid V_r \mid ||A_r^{-1} \sum_i l_{jr} \tilde{E}_j \mathbf{n}_{jr}||_2^2.$$

On utilise les inégalités de norme matricielles, l'inégalité triangulaire et l'hypothèse (2.57) pour obtenir

$$||\tilde{\mathbf{F}}_n||^2_{L^2} \leq C_1 h^2 \frac{\varepsilon^2}{\sigma^2} \sum_r \mid V_r \mid ||A_r^{-1}||_2 \sum_j \mid \tilde{E}_j \mid^2 .$$

En définissant le ratio $\mu_{jr} = \frac{|V_r|}{|\Omega_j|}$ qui est borné par les hypothèses (2.57)-(2.58), on obtient donc la majoration suivante

$$||\tilde{\mathbf{F}}_n||_{L^2} = \sqrt{\sum_r \mid V_r \mid (\tilde{\mathbf{F}}_r, \tilde{\mathbf{F}}_r)} \leq \frac{\varepsilon}{\sigma} h^2 C_n \sqrt{\sum_j \mid \Omega_j \mid\mid \tilde{E}_j \mid^2}. \tag{2.82}$$

avec $C_n = C\sqrt{\max_r \left(||A_r^{-1}||_2\right)}$. On obtient (2.80). En utilisant $||\widehat{\alpha}_{jr}|| < Ch$ et les inégalités matricielles il est immédiat de montrer que

$$||\tilde{\mathbf{F}}_c||_{L^2} \leq C_{n_1} ||\tilde{\mathbf{F}}_n||_{L^2},$$

avec $C_{n_1} = \sqrt{\max_j(||B_j^{-1}||)}$. En utilisant la majoration de $||\tilde{\mathbf{F}}_n||_{L^2}$ on obtient (2.81) avec $C_{n_2} = C_n C_{n_1}$. □

Une fois ces estimations établies on propose de montrer que les solutions du système (2.79) associées au schéma de diffusion (2.79) sont solutions du système (2.78) associé au schéma JL-(b) avec des termes résiduels, l'estimation de ces termes permettant d'obtenir une estimation L^2 de l'erreur entre les solutions de (2.79) et (2.78).

Proposition 2.15. *La solution $(\tilde{E}_j, \tilde{\mathbf{F}}_j, \tilde{\mathbf{F}}_r)$ du système (2.79) est solution de l'équation suivante :*

$$\begin{cases} \mid \Omega_j \mid \partial_t \tilde{E}_j + \dfrac{1}{\varepsilon} \displaystyle\sum_r (l_{jr}\mathbf{n}_{jr}, \tilde{\mathbf{F}}_r) = 0, \\ \mid \Omega_j \mid \partial_t \tilde{\mathbf{F}}_j + \dfrac{1}{\varepsilon} \displaystyle\sum_r (l_{jr}\tilde{E}_j \mathbf{n}_{jr} + \widehat{\alpha}_{jr}(\tilde{\mathbf{F}}_j - \tilde{\mathbf{F}}_r)) = \mathbf{R}_j^c, \\ \left(B_r + \dfrac{\sigma}{\varepsilon} A_r\right) \tilde{\mathbf{F}}_r - \displaystyle\sum_i l_{ir}\tilde{E}_i \mathbf{n}_{ir} - \sum_i \widehat{\alpha}_{ir}\tilde{\mathbf{F}}_i = \mathbf{R}_r^n, \end{cases} \tag{2.83}$$

avec $\mathbf{R}^c = \left(\mathbf{R}^c_{j=1}, ..., \mathbf{R}^c_{j=n}\right)$ et $\mathbf{R}^n = (\mathbf{R}^n_{r=1}, ..., \mathbf{R}^n_{r=nv})$. Les termes résiduels sont majorés par

$$||\mathbf{R}^c||_{L^2} \leq \frac{\varepsilon}{\sigma} h^3 C_{r_1} ||\partial_t \tilde{\mathbf{E}}_c||_{L^2},$$

et

$$||\mathbf{R}^n||_{L^2} \leq Ch \left(||\tilde{\mathbf{F}}_n||_{L^2}\right).$$

Démonstration. On part donc du système

$$\begin{cases} \mid \Omega_j \mid \partial_t \tilde{E}_j + \dfrac{1}{\varepsilon} \displaystyle\sum_r (l_{jr}\mathbf{n}_{jr}, \tilde{\mathbf{F}}_r) = 0, \\ \mid \Omega_j \mid \partial_t \tilde{\mathbf{F}}_j + \dfrac{1}{\varepsilon} \displaystyle\sum_r (l_{jr}\tilde{E}_j \mathbf{n}_{jr} + \widehat{\alpha}_{jr}(\tilde{\mathbf{F}}_j - \tilde{\mathbf{F}}_r)) = \mathbf{R}_j^c, \\ \left(B_r + \dfrac{\sigma}{\varepsilon} A_r\right) \tilde{\mathbf{F}}_r - \displaystyle\sum_i l_{ir}\tilde{E}_i \mathbf{n}_{ir} - \sum_i \widehat{\alpha}_{ir}\tilde{\mathbf{F}}_i = \mathbf{R}_r^n. \end{cases} \tag{2.84}$$

Puisque $(\tilde{E}_j, \tilde{\mathbf{F}}_j, \tilde{\mathbf{F}}_r)$ sont solutions du système (2.79) et $\sum_r l_{jr}\mathbf{n}_{jr} = \mathbf{0}$ on obtient l'expression de terme résiduel

$$\begin{cases} \mathbf{R}_j^c = \mid \Omega_j \mid \partial_t \tilde{\mathbf{F}}_j, \\ \mathbf{R}_r^n = B_r \tilde{\mathbf{F}}_r - \displaystyle\sum_i \widehat{\alpha}_{ir}\tilde{\mathbf{F}}_i. \end{cases} \tag{2.85}$$

En utilisant le lemme précédent on montre que

$$||\tilde{\mathbf{R}}^c||_{L^2} \leq \frac{\varepsilon}{\sigma} h^3 C_{r_1} ||\partial_t \tilde{\mathbf{E}}_c||_{L^2},$$

avec $C_{r_1} = C_{n_1} C_{n_2}$. Maintenant on cherche à estimer le second terme résiduel $\mathbf{R}^n$.

$$||\mathbf{R}^n||_{L^2}^2 = \sum_r |V_r|(B_r \tilde{\mathbf{F}}_r - \sum_i \widehat{\alpha}_{ir} \tilde{\mathbf{F}}_i, B_r \tilde{\mathbf{F}}_r - \sum_i \widehat{\alpha}_{ir} \tilde{\mathbf{F}}_i) \leq$$

$$2 \left(\sum_r |V_r|(B_r \tilde{\mathbf{F}}_r, B_r \tilde{\mathbf{F}}_r) + \sum_r |V_r|(\sum_i \widehat{\alpha}_{ir} \tilde{\mathbf{F}}_i, \sum_i \widehat{\alpha}_{ir} \tilde{\mathbf{F}}_i) \right).$$

La majoration du premier terme est immédiate en utilisant la définition d'une norme matricielle subordonnée, des inégalités triangulaires, on obtient que

$$\sum_r |V_r|(B_r \tilde{\mathbf{F}}_r, B_r \tilde{\mathbf{F}}_r) \leq \max_r(||B_r||_2^2) ||\tilde{\mathbf{F}}_n||_{L^2}^2.$$

Sachant que $\max_r(||B_r||_2) < Ch$. L'inégalité de Cauchy-Schwarz donne

$$\sum_r |V_r|(\sum_i \widehat{\alpha}_{ir} \tilde{\mathbf{F}}_i, \sum_i \widehat{\alpha}_{ir} \tilde{\mathbf{F}}_i) \leq \sum_r \sum_i |V_r| ||\widehat{\alpha}_{ir}||^2 ||\tilde{\mathbf{F}}_i||^2.$$

On introduit le paramètre $\mu_{ir} = \frac{|V_r|}{|\Omega_i|}$ on obtient

$$\sum_r |V_r|(\sum_i \widehat{\alpha}_{ir} \tilde{\mathbf{F}}_i, \sum_i \widehat{\alpha}_{ir} \tilde{\mathbf{F}}_i) \leq h^2 \sum_i |\ \Omega_i\ | \ ||\tilde{\mathbf{F}}_i||^2 \left(\sum_r \mu_{ir} \right)$$

On sait que les μ_{jr} sont bornés par hypothèse sur le maillage donc

$$\sum_r |V_r|(\sum_i \widehat{\alpha}_{ir} \tilde{\mathbf{F}}_i, \sum_i \widehat{\alpha}_{ir} \tilde{\mathbf{F}}_i) \leq Ch^2 ||\tilde{\mathbf{F}}_c||_{L^2}^2.$$

Pour conclure il suffit d'utiliser la majoration sur $||\tilde{\mathbf{F}}_n||_{L^2}$ et celle sur $||\tilde{\mathbf{F}}_c||_{L^2}$ pour obtenir

$$||\mathbf{R}^n||_{L^2} \leq C_{n_3} h \left(||\tilde{\mathbf{F}}_n||_{L^2} \right) \tag{2.86}$$

□

Proposition 2.16. *On introduit $e_{E,j} = E_j - \tilde{E}_j$, $\mathbf{e}_{F,j} = \mathbf{F}_j - \tilde{\mathbf{F}}_j$, $\mathbf{e}_{F,r} = \mathbf{F}_r - \tilde{\mathbf{F}}_r$. Ces quantités satisfont*

$$\begin{cases} |\ \Omega_j\ | \ \partial_t e_{E,j} + \dfrac{1}{\varepsilon} \displaystyle\sum_r (l_{jr} \mathbf{n}_{jr}, \mathbf{e}_{F,r}) = 0, \\ |\ \Omega_j\ | \ \partial_t e_{F,j} + \dfrac{1}{\varepsilon} \displaystyle\sum_r (l_{jr} e_{E,j} \mathbf{n}_{jr} + \widehat{\alpha}_{jr}(\mathbf{e}_{F,j} - \mathbf{e}_{F,r})) = \mathbf{R}_j^c, \\ \left(B_r + \dfrac{\sigma}{\varepsilon} A_r \right) \mathbf{e}_{F,r} - \displaystyle\sum_i l_{ir} e_{E,i} \mathbf{n}_{ir} - \sum_i \widehat{\alpha}_{ir} \mathbf{e}_{F,j} = \mathbf{R}_r^n. \end{cases} \tag{2.87}$$

En outre les vecteurs $\mathbf{e}_c = (e_{E,j=1}, ..., e_{E,j=n}, \mathbf{e}_{F,j=1}, ..., \mathbf{e}_{F,j=n})$ et $\mathbf{e}_n = (\mathbf{e}_{F,r=1}, ..., \mathbf{e}_{F,r=nv})$ satisfont les estimations suivantes

$$||\mathbf{e}_c||_{L^2} \leq h \frac{\varepsilon}{\sigma} \left(C_0 ||\tilde{\mathbf{E}}_c(0)||_{L^2} + C_f h^2 T ||\partial_t \tilde{\mathbf{E}}_c||_{L^\infty([0:T]:L^2)} \right), \tag{2.88}$$

et

$$||\mathbf{e}_n||_{L^2} \leq \left(C_{f_2} ||\mathbf{e}_c||_{L^2} + C_{f_3} \frac{\varepsilon}{\sigma} h^2 ||\tilde{\mathbf{E}}_c||_{L^2} \right), \tag{2.89}$$

Démonstration. On pose $e_{E,j} = E_j - \tilde{E}_j$, $\mathbf{e}_{F,j} = \mathbf{F}_j - \tilde{\mathbf{F}}_j$, $\mathbf{e}_{F,r} = \mathbf{F}_r - \tilde{\mathbf{F}}_r$. En utilisant le lemme précédent on sait que les quantités qui viennent d'être définies satisfont

$$\begin{cases} \mid \Omega_j \mid \partial_t e_{E,j} + \dfrac{1}{\varepsilon}\sum_r (l_{jr}\mathbf{n}_{jr}, \mathbf{e}_{F,r}) = 0, \\ \mid \Omega_j \mid \partial_t e_{F,j} + \dfrac{1}{\varepsilon}\sum_r (l_{jr} e_{E,j}\mathbf{n}_{jr} + \widehat{\alpha}_{jr}(\mathbf{e}_{F,j} - \mathbf{e}_{F,r})) = R_j^c, \\ \left(B_r + \dfrac{\sigma}{\varepsilon}A_r\right)\mathbf{e}_{F,r} - \sum_i l_{ir} e_{E,i}\mathbf{n}_{ir} - \sum_i \widehat{\alpha}_{ir}\mathbf{e}_{F,i} = R_r^n. \end{cases} \tag{2.90}$$

On pose $\mathbf{e}^c = (e_{E,j=1}, ..., e_{E,j=n}, \mathbf{e}_{F,j=1}, ..., \mathbf{e}_{F,j=n})$ l'erreur entre la solution du schéma de diffusion et celle de JL-(b) et

$$E_r(t) = \int_\Omega (\mathbf{e}^c, \mathbf{e}^c) = \sum_j \mid \Omega_j \mid (\mid e_{E,j} \mid^2 + (\mathbf{e}_{F,j}, \mathbf{e}_{F,j})).$$

En reprenant mot pour mot la preuve de stabilité du schéma on sait que si on ignore les termes résiduels $\mathbf{R}^c$ alors $E_r^{'}(t) \leq 0$. Avec le terme résiduel on obtient

$$E_r^{'}(t) \leq 2\sum_j ||\mathbf{R}^c||_{L^2(D)} ||\mathbf{e}^c||_{L^2(D)}.$$

On utilise les majorations de la proposition 2.15 et du lemme 2.14 précédent pour obtenir

$$E_r^{'}(t) \leq C_f h^3 \frac{\varepsilon}{\sigma} ||\partial_t \tilde{\mathbf{E}}^c||_{L^2(D)} \sqrt{E(t)},$$

avec $C_f = C_{n_1} C_{n_3}$ et

$$||\partial_t \tilde{\mathbf{E}}^c||_{L^2(D)} = \sqrt{\sum_j \mid \Omega_j \mid (\partial_t \tilde{E}_j, \partial_t \tilde{E}_j)}.$$

En utilisant des arguments de type Gröndwall on obtient

$$\sqrt{E(T)} \leq \sqrt{E(0)} + T C_f h^3 \frac{\varepsilon}{\sigma} ||\partial_t \tilde{\mathbf{E}}_c||_{L^\infty([0,T]:L^2(D))}. \tag{2.91}$$

On sait que $e_{E,j}(0) = 0$ et $\mathbf{e}_{F,j}(0) = \tilde{\mathbf{F}}_j^0$. On utilise (2.80)-(2.81) pour obtenir

$$E_r(0) \leq C_0 \frac{h\varepsilon}{\sigma} ||\tilde{\mathbf{E}}_c(0)||_{L^2(D)}.$$

Pour conclure on sait que le schéma limite de diffusion est convergent et stable, par conséquent on sait que $||\partial_t \tilde{\mathbf{E}}^c||_{L^2}$ est borné. Maintenant on souhaite estimer la norme L^2 associée à l'erreur entre les flux $\mathbf{F}_r$ et $\tilde{\mathbf{F}}_r$. On introduit la norme L^2

$$\sum_r |V_r|(\mathbf{F}_r - \tilde{\mathbf{F}}_r, \mathbf{F}_r - \tilde{\mathbf{F}}_r)) = ||\mathbf{e}^n||_{L^2}^2.$$

On utilise la définition de $\mathbf{e}_{F,r}$ dans (2.90) afin d'obtenir la majoration suivante

$$\begin{aligned} \sum_r |V_r| ||\mathbf{e}_{F,r}||^2 \leq \sum_r \mid V_r \mid || \left(B_r + \frac{\sigma}{\varepsilon}A_r\right)^{-1} \left(\sum_j l_{jr} e_{E,j}\mathbf{n}_{jr} + \widehat{\alpha}_{jr}\mathbf{e}_{F,j}\right) ||^2 \\ + \sum_r |V_r| || \left(B_r + \frac{\sigma}{\varepsilon}A_r\right)^{-1} \mathbf{R}_r^n ||^2. \end{aligned} \tag{2.92}$$

En utilisant le fait que la matrice $(B_r + \frac{\sigma}{\varepsilon}A_r)$ soit inversible on montre que son inverse est majoré, par conséquent le second terme de (2.92) est majoré

$$\sum_r |\,|V_r|\,|\,|| \left(B_r + \frac{\sigma}{\varepsilon}A_r\right)^{-1} \mathbf{R}_r^n||^2 \leq \max_r \left(||C_r + \frac{\sigma}{\varepsilon}A_r)^{-1}||^2\right) ||\mathbf{R}^n||_{L^2}^2.$$

Toujours en prenant les mêmes arguments un rapide calcul montre que le premier terme de (2.92) est borné par

$$\max_r \left(||(B_r + \frac{\sigma}{\varepsilon}A_r)^{-1}||^2\right) \sum_r |\,V_r\,| \sum_j ||l_{jr}e_{E,j}\mathbf{n}_{jr} + \widehat{\alpha}_{jr}\mathbf{e}_{F,j}||^2$$

Puisque les quantités $||l_{jr}\mathbf{n}_{jr}||$ et $||\widehat{\alpha}_{jr}||$ sont bornées par Ch. On majore ainsi toutes ces quantités dans la somme précédente et on introduit le $\mu_{jr} = \frac{|V_r|}{|\Omega_j|}$. Par conséquent

$$\sum_r |V_r|| \left(B_r + \frac{\sigma}{\varepsilon}A_r\right)^{-1} \left(\sum_j l_{jr}e_{E,j}\mathbf{n}_{jr} + \widehat{\alpha}_{jr}\mathbf{e}_{F,j}\right) ||^2 \leq$$

$$\max_r \left(||(B_r + \frac{\sigma}{\varepsilon}A_r)^{-1}||^2\right) \sum_j |\Omega_j| \left(|e_{E,j}|^2 + ||\mathbf{e}_{F,j}||^2\right) \left(\sum_r \mu_{jr}\right)$$

Pour conclure il suffit de majorer les coefficients μ_{jr} et de reconnaître la norme L^2 de $\mathbf{e}^c$. Il suffit de sommer les estimations pour obtenir le résultat. □

Remarque 2.2.2. *Le schéma JL-(b) converge quand ε tend vers zéro vers le schéma de diffusion (2.13) lui même étant un schéma convergent pour la discrétisation de l'équation de la chaleur (théorème 2.11). Par conséquent, le schéma est « asymptotic preserving » au sens de la définition 1.3.3.*

2.2.5 Dérivation des schémas JL-(a) et JL-(b) en dimension un

Dans cette section on propose de démontrer que les schémas JL-(a) et JL-(b) dérivent en 1D sur des schémas ayant la même structure que les schémas de Jin-Levermore et Gosse-Toscani. Dans un premier temps on réécrit le schéma JL-(b) sous une forme qui est l'extension 2D du schéma (1.46). Puis on montre que en 1D le schéma obtenu est très proche du schéma (1.46) avant de conclure.

Proposition 2.17. *le schéma JL-(b) peut se réécrire sous la forme*

$$\begin{cases} |\,\Omega_j\,|\,\partial_t E_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r (l_{jr}\mathbf{n}_{jr}, M_r\tilde{\mathbf{F}}_r) = 0, \\ |\,\Omega_j\,|\,\partial_t \mathbf{F}_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r \widehat{\alpha}_{jr}M_r(\mathbf{F}_j - \tilde{\mathbf{F}}_r) = -\dfrac{1}{\varepsilon}\left(\displaystyle\sum_r \widehat{\alpha}_{jr}(\widehat{I}_d - M_r)\right)\mathbf{F}_j, \\ \left(\displaystyle\sum_j \widehat{\alpha}_{jr}\right)\tilde{\mathbf{F}}_r = \displaystyle\sum_j l_{jr}E_j\mathbf{n}_{jr} + \sum_j \widehat{\alpha}_{jr}\mathbf{F}_j, \end{cases} \tag{2.93}$$

avec

$$M_r = \left(\sum_j \widehat{\alpha}_{jr} + \frac{\sigma}{\varepsilon}\sum_j \widehat{\beta}_{jr}\right)^{-1} \left(\sum_j \widehat{\alpha}_{jr}\right),$$

et

$$\widehat{\alpha}_{jr} = l_{jr}\mathbf{n}_{jr} \otimes \mathbf{n}_{jr}, \quad \widehat{\beta}_{jr} = l_{jr}\mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j).$$

Démonstration. On part du schéma JL-(b) (2.22-(2.23) écrit sous la forme suivante.

$$\begin{cases} \mid \Omega_j \mid \partial_t E_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r (l_{jr}\mathbf{n}_{jr}, \mathbf{F}_r) = 0, \\ \mid \Omega_j \mid \partial_t \mathbf{F}_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r (l_{jr}E_j\mathbf{n}_{jr} + \widehat{\alpha}_{jr}(\mathbf{F}_j - \mathbf{F}_r)) = \mathbf{0}, \\ \left(\displaystyle\sum_j \widehat{\alpha}_{jr} + \frac{\sigma}{\varepsilon}\sum_j \widehat{\beta}_{jr}\right)\tilde{\mathbf{F}}_r = \displaystyle\sum_j l_{jr}E_j\mathbf{n}_{jr} + \sum_j \widehat{\alpha}_{jr}\mathbf{F}_j. \end{cases} \tag{2.94}$$

On pose $\tilde{\mathbf{F}}_r$ solution du solveur nodal classique

$$\left(\sum_j \widehat{\alpha}_{jr}\right)\tilde{\mathbf{F}}_r = \sum_j l_{jr}E_j\mathbf{n}_{jr} + \sum_j \widehat{\alpha}_{jr}\mathbf{F}_j.$$

On peut constater immédiatement que $\mathbf{F}_r = M_r\tilde{\mathbf{F}}_r$. Pour la première équation il suffit de remplacer $\mathbf{F}_r$ par $M_r\tilde{\mathbf{F}}_r$ pour obtenir le résultat. Pour la deuxième, par définition des normales et des longueurs associées aux noeuds on sait que

$$\sum_r l_{jr}E_j\mathbf{n}_{jr} = \mathbf{0}.$$

Pour conclure il suffit de remplacer $\mathbf{F}_r$ par $M_r\tilde{\mathbf{F}}_r$ dans la seconde équation et de remarquer que

$$\sum_r \widehat{\alpha}_{jr}(\mathbf{F}_j - M_r\tilde{\mathbf{F}}_r) = \sum_r \widehat{\alpha}_{jr}(\widehat{I}_d - M_r)\mathbf{F}_j + \sum_r \widehat{\alpha}_{jr}M_r(\mathbf{F}_j - \mathbf{F}_r).$$

Ceci clôt la preuve. □

On retrouve une écriture similaire au 1D. En effet, le schéma JL-(b) se réécrit comme le schéma aux noeuds classique où les flux sont corrigés à l'aide d'une matrice M_r.

Remarque 2.2.3. *La matrice M_r est une généralisation en dimension deux du coefficient M introduit dans les schémas 1D de Jin-Levermore et Gosse-Toscani. Cependant elle contient en plus des informations géométriques à travers le tenseur $\widehat{\beta}_{jr}$ qui permettent d'obtenir un gradient consistant dans le régime limite.*

Proposition 2.18. *Soit h le pas de maillage, sur maillage Cartésien, la matrice satisfait $M_r = M\widehat{I}_d$ avec $M = \frac{\sqrt{2}\varepsilon}{\sqrt{2}\varepsilon+\sigma h}$ et*

$$-\frac{1}{\varepsilon}\left(\sum_r \widehat{\alpha}_{jr}(\widehat{I}_d - M_r)\right) = -\frac{\sigma}{\varepsilon^2}h^2 M\widehat{I}_d.$$

Démonstration. Dans un premier temps un rapide calcul montre que

$$M_r = M\widehat{I}_d.$$

En effet dans le cas Cartésien, on montre que

$$M_r = \begin{pmatrix} \frac{1}{\sqrt{2}h+\frac{\sigma}{\varepsilon}h^2} & 0 \\ 0 & \frac{1}{\sqrt{2}h+\frac{\sigma}{\varepsilon}h^2} \end{pmatrix}\begin{pmatrix} \sqrt{2}h & 0 \\ 0 & \sqrt{2}h \end{pmatrix},$$

par factorisation on constate que

$$\frac{\sqrt{2}h}{\sqrt{2}h+\frac{\sigma}{\varepsilon^2}h^2}=\frac{\sqrt{2}\varepsilon}{\sqrt{2}\varepsilon+\sigma h}.$$

Il faut aussi montrer

$$-\frac{1}{\varepsilon}\left(\sum_r \widehat{\alpha}_{jr}(\widehat{I}_d-M_r)\right)=-\frac{\sigma}{\varepsilon^2}M\widehat{I}_d h^2.$$

Dans le cas Cartésien on montre que

$$-\frac{1}{\varepsilon}\left(\sum_r \widehat{\alpha}_{jr}(\widehat{I}_d-M_r)\right)=-\frac{1}{\varepsilon}\begin{pmatrix}\sqrt{2}h & 0\\ 0 & \sqrt{2}h\end{pmatrix}\begin{pmatrix}1-M & 0\\ 0 & 1-M\end{pmatrix}.$$

On commence par faire sortir le coefficient $\frac{\sigma}{\varepsilon^2}$ afin d'obtenir

$$-\frac{1}{\varepsilon}\left(\sum_r \widehat{\alpha}_{jr}(\widehat{I}_d-M_r)\right)=-\frac{\sigma}{\varepsilon^2}\begin{pmatrix}\frac{\sqrt{2}h\varepsilon}{\sigma}(1-M) & 0\\ 0 & \frac{\sqrt{2}h\varepsilon}{\sigma}(1-M)\end{pmatrix}.$$

En développant et en mettant au même dénominateur on obtient

$$\frac{\sqrt{2}h\varepsilon}{\sigma}(1-M)=Mh^2.$$

□

Sachant que le schéma au noeuds classique dérive en 1D sur le schéma upwind avec un coefficient de viscosité différent, le précédent résultat nous permet de montrer d'exhiber le schéma JL-(b) en 1D.

Proposition 2.19. *En dimension un, le schéma JL-(b) (2.22)-(2.23) s'écrit*

$$\begin{cases}\partial_t E_j+M\dfrac{F_{j+1}-F_{j-1}}{2\varepsilon\Delta x}-M\dfrac{E_{j+1}-E_j+E_{j-1}}{\sqrt{2}\varepsilon\Delta x}=0,\\ \\ \partial_t F_j+M\dfrac{E_{j+1}-E_{j-1}}{2\varepsilon\Delta x}-M\dfrac{F_{j+1}-2F_j+F_{j-1}}{\sqrt{2}\varepsilon\Delta x}+M\dfrac{\sigma}{\varepsilon^2}F_j=0,\ \forall j\in 0..n\end{cases}\tag{2.95}$$

avec $M=\frac{\sqrt{2}\varepsilon}{\sqrt{2}\varepsilon+\sigma h}$.

En utilisant le même type de manipulations algébriques on prouve que le schéma JL-(a) est très proche du schéma de Jin-Levermore en 1D (seul le coefficient de viscosité change).

Proposition 2.20. *Le schéma JL-(a) (2.11)-(2.12) est équivalent à*

$$\begin{cases}\mid\Omega_j\mid\partial_t E_j+\dfrac{1}{\varepsilon}\displaystyle\sum_r(l_{jr}\mathbf{n}_{jr},M_r\tilde{\mathbf{F}}_r)=0,\\ \\ \mid\Omega_j\mid\partial_t\mathbf{F}_j+\dfrac{1}{\varepsilon}\displaystyle\sum_r\widehat{\alpha}_{jr}(\mathbf{F}_j-\tilde{\mathbf{F}}_r)=-\dfrac{\sigma}{\varepsilon^2}\mid\Omega_j\mid\mathbf{F}_j,\\ \\ \left(\displaystyle\sum_j\widehat{\alpha}_{jr}\right)\tilde{\mathbf{F}}_r=\displaystyle\sum_j l_{jr}E_j\mathbf{n}_{jr}+\sum_j\widehat{\alpha}_{jr}\mathbf{F}_j,\end{cases}\tag{2.96}$$

avec

$$M_r = \left(\sum_j \widehat{\alpha}_{jr} + \frac{\sigma}{\varepsilon} \sum_j \widehat{\beta}_{jr} \right)^{-1} \left(\sum_j \widehat{\alpha}_{jr} \right)$$

et

$$\widehat{\alpha}_{jr} = l_{jr} \mathbf{n}_{jr} \otimes \mathbf{n}_{jr}, \quad \widehat{\beta}_{jr} = l_{jr} \mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j).$$

Démonstration. On part du schéma (2.11)-(2.12). Pour la première équation le calcul est trivial. Pour la seconde équation

$$\sum_j \widehat{\alpha}_{jr} (\mathbf{F}_j - \mathbf{F}_r) - \frac{\sigma}{\varepsilon} \sum_j \widehat{\beta}_{jr} \mathbf{F}_r = \sum_j \widehat{\alpha}_{jr} \mathbf{F}_j - \sum_j (\widehat{\alpha}_{jr} + \frac{\sigma}{\varepsilon} \widehat{\beta}_{jr}) M_r \tilde{\mathbf{F}}_r$$

Sur maillage Cartésien $\widehat{\alpha}_{jr} + \frac{\sigma}{\varepsilon} \widehat{\beta}_{jr} = \frac{h}{2\sqrt{2}} \widehat{I}_d + \frac{\sigma}{\varepsilon} \frac{h^2}{4} \widehat{I}_d$ et

$$M_r = \left(\sqrt{2} h \widehat{I}_d + \frac{\sigma}{\varepsilon} h^2 \widehat{I}_d \right)^{-1} (\sqrt{2} h \widehat{I}_d).$$

Par conséquent

$$\left(\widehat{\alpha}_{jr} + \frac{\sigma}{\varepsilon} \widehat{\beta}_{jr} \right) M_r = \widehat{\alpha}_{jr}.$$

On obtient la reformulation (2.96). □

En utilisant cette reformulation et les manipulations algébriques précédentes on obtient que

Proposition 2.21. *En 1D, le schéma JL-(a) (2.11)-(2.12) s'écrit*

$$\begin{cases} \partial_t E_j + M \dfrac{F_{j+1} - F_{j-1}}{2\varepsilon \Delta x} - M \dfrac{E_{j+1} - 2E_j + E_{j-1}}{\sqrt{2} \varepsilon \Delta x} = 0, \\ \partial_t F_j + \dfrac{E_{j+1} - E_{j-1}}{2\varepsilon \Delta x} - \dfrac{F_{j+1} - 2F_j + F_{j-1}}{\sqrt{2} \varepsilon \Delta x} + \dfrac{\sigma}{\varepsilon^2} F_j = 0, \ \forall j \in 0..n \end{cases} \tag{2.97}$$

avec $M = \frac{\sqrt{2}\varepsilon}{\sqrt{2}\varepsilon + \sigma h}$.

2.2.6 Conclusion sur les schémas « asymptotic preserving »

Le schéma limite des schémas JL-(a) et JL-(b) pour le tenseur $\widehat{\beta}_{jr} = l_{jr} \mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)$ étant consistant, on peut énoncer la propriété suivante

Proposition 2.22. *Le schéma JL-(a)*

$$\begin{cases} |\Omega_j| \, \partial_t E_j + \dfrac{1}{\varepsilon} \displaystyle\sum_r l_{jr} (\mathbf{F}_r, \mathbf{n}_{jr}) = 0, \\ |\Omega_j| \, \partial_t \mathbf{F}_j + \dfrac{1}{\varepsilon} \displaystyle\sum_r \mathbf{G}_{jr} = - |\Omega_j| \, \dfrac{\sigma}{\varepsilon^2} \mathbf{F}_j, \end{cases} \tag{2.98}$$

et le schéma JL-(b)

$$\begin{cases} |\Omega_j| \, \partial_t E_j + \dfrac{1}{\varepsilon} \displaystyle\sum_r l_{jr} (\mathbf{F}_r, \mathbf{n}_{jr}) = 0, \\ |\Omega_j| \, \partial_t \mathbf{F}_j + \dfrac{1}{\varepsilon} \displaystyle\sum_r \mathbf{G}_{jr} = - \dfrac{\sigma}{\varepsilon^2} \displaystyle\sum_r l_{jr} \mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j) \mathbf{F}_r, \end{cases} \tag{2.99}$$

avec les flux

$$\begin{cases} \mathbf{G}_{jr} = l_{jr}E_j\mathbf{n}_{jr} + \widehat{\alpha}_{jr}(\mathbf{F}_j - \mathbf{F}_r) - \dfrac{\sigma}{\varepsilon}\widehat{\beta}_{jr}\mathbf{F}_r, \\ \left(\displaystyle\sum_j \widehat{\alpha}_{jr} + \frac{\sigma}{\varepsilon}\sum_j \widehat{\beta}_{jr}\right)\mathbf{F}_r = \displaystyle\sum_j l_{jr}E_j\mathbf{n}_{jr} + \sum_j \widehat{\alpha}_{jr}\mathbf{F}_j, \end{cases} \tag{2.100}$$

et $\widehat{\beta}_{jr} = l_{jr}\mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)$ *sont « asymptotic preserving »au sens de la définition (1.3.3).*

Remarque 2.2.4. *Le théorème 2.11 et les propositions 2.15-2.16 montrent que le schéma JL-(b) est « asymptotic preserving » au sens de la définition (1.3.3). Pour montrer rigoureusement que le schéma JL-(a) « asymptotic preserving » au sens de la définition (1.3.3) il faudrait obtenir l'équivalent des propositions 2.15-2.16. Cependant les développements de Hilbert nous donnent un premier résultat formel.*

2.3 Discrétisation temporelle et conditions limites

2.3.1 Discrétisation temporelle

Trois discrétisations temporelles possibles : explicite, implicite et semi-implicite. Deux problèmes se posent. Les schémas sont ils stables sous une CFL indépendante de ε ? Selon la discrétisation temporelle restent ils « asymptotic preserving » ? Dans le cas des schémas implicites et explicites un rapide calcul en utilisant un développement de Hilbert montre formellement que les schémas restent « asymptotic preserving » . En ce qui concerne la stabilité des schémas, elle découle des résultats précédents de stabilité des schémas semi-discrets. En effet, la décroissance de la norme L^2 (propositions 2.12 - 2.13) revient à dire que si A est la matrice de discrétisation spatiale et $\mathbf{V}$ un vecteur alors

$$(\mathbf{V}, A\mathbf{V}) \leq 0.$$

A l'aide de cette propriété il est immédiat de montrer que le schéma implicite et explicite sont stables (sous CFL dans le cas explicite).
La norme L^2 discrète étant définie par

$$||\mathbf{V}||_{L^2} = \sum_j \mid \Omega_j \mid (\mathbf{V}_j, \mathbf{V}_j),$$

pour $\mathbf{V}$ le vecteur de variable du système et donc $\mathbf{V} = \{\mathbf{V}_{j=1}, ...\mathbf{V}_{j=n}\}$.

Cas implicite Soit M la matrice diagonale des volumes de mailles, le schéma implicite s'écrit

$$M\mathbf{V}^{n+1} = M\mathbf{V}^n + \Delta t A\mathbf{V}^{n+1}.$$

Puisque $(\mathbf{V}, A\mathbf{V}) \leq 0$ alors la matrice $\widehat{I}_d - \Delta t A$ est définie positive et donc inversible. On multiplie par $\mathbf{V}^{n+1}$ pour obtenir

$$(M\mathbf{V}^{n+1}, \mathbf{V}^{n+1}) = (M\mathbf{V}^n, \mathbf{V}^{n+1})_{L^2} + \Delta t(A\mathbf{V}^{n+1}, \mathbf{V}^{n+1})_{L^2}.$$

On utilise le fait que la matrice A soit négative ainsi que l'inégalité de Cauchy-Schwarz pour obtenir que

$$||\mathbf{V}^{n+1}||_{L^2} \leq ||\mathbf{V}^n||_{L^2}.$$

Cas explicite Maintenant on propose de montrer que $(\mathbf{V}, A\mathbf{V}) \leq 0$ implique que le schéma explicite en temps est stable en norme L^2 sous une condition CFL. On définit de la même façon le schéma explicite

$$M\mathbf{V}^{n+1} = M\mathbf{V}^n + \Delta t A\mathbf{V}^n.$$

En utilisant le même type de calcul on obtient la stabilité sous une condition CFL. En effet on obtient

$$||\mathbf{V}^{n+1}||^2 = (\sqrt{M}\mathbf{V}^{n+1}, \sqrt{M}\mathbf{V}^n) + \Delta t(\mathbf{V}^{n+1}, A\mathbf{V}^n),$$

$$||\mathbf{V}^{n+1}||^2 = (\sqrt{M}\mathbf{V}^{n+1}, \sqrt{M}\mathbf{V}^n) + \Delta t(\sqrt{M}\mathbf{V}^{n+1}, M^{-1}\sqrt{M}A\mathbf{V}^n),$$

On regroupe tout et on utilise l'inégalité de Cauchy-Schwarz pour obtenir

$$||\mathbf{V}^{n+1}|| \leq ||\widehat{I}_d + \Delta t M^{-1}A||_2 ||\mathbf{V}^{n+1}|| ||\mathbf{V}^n||.$$

Ce qui revient à dire que le schéma est stable si

$$||\widehat{I}_d + \Delta t M^{-1}A||_2 \leq 1. \tag{2.101}$$

Puisque A est négative on peut obtenir un pas de temps qui satisfait cette condition. Pour cela on utilise la définition suivante de la norme deux

$$||\widehat{I}_d + \Delta t M^{-1}A||_2 = \sup_{||\mathbf{x}||=1} \sqrt{\Big((\widehat{I}_d + \Delta t M^{-1}A)\mathbf{x}, (\widehat{I}_d + \Delta t M^{-1}A)\mathbf{x}\Big)}.$$

Par un rapide calcul on obtient

$$||\widehat{I}_d + \Delta t M^{-1}A||_2 = \sup_{||\mathbf{x}||=1} \sqrt{(1 + \Delta t(\mathbf{x}, (M^{-1}A + A^t M^{-1} + \Delta t A M^{-1}M^{-1}A^t)\mathbf{x}))}.$$

Puisque la matrice M^{-1} composée des volumes de mailles est strictement positive, la condition CFL (2.101) peut être respectée seulement si A est une matrice non positive. On souhaite maintenant exprimer cette CFL. On commence par remarquer que $(\mathbf{x}, AM^{-1}M^{-1}A^t\mathbf{x}) = ||AM^{-1}\mathbf{x}||^2$. La condition CFL revient à

$$\Delta t \left(\mathbf{x}, (M^{-1}A + A^t M^{-1})\mathbf{x}\right) + \Delta t^2 ||M^{-1}A\mathbf{x}||^2 \leq 0, \quad \forall \mathbf{x} \text{ tel que } ||\mathbf{x}|| = 1.$$

En remarquant que $\left(\mathbf{x}, (M^{-1}A + A^t M^{-1})\mathbf{x}\right) \leq 0$, on obtient la condition CFL suivante

$$\Delta t \leq \inf_{||\mathbf{x}||=1} \frac{|\left(\mathbf{x}, (AM^{-1} + M^{-1}A^t)\mathbf{x}\right)|}{||AS^{-1}\mathbf{x}||_2^2}.$$

Il est connu pour ce type de modèle que la CFL de la partie transport est contrainte par ε et que la CFL induite par le terme source dépend de $\frac{\varepsilon^2}{\sigma}$. Il est bien évident que lorsque ε tend vers 0 ou σ vers l'infini un schéma explicite est inutilisable. Le schéma implicite lève bien entendu cette restriction mais cet avantage est contrebalancé par un coût calcul à chaque itération nettement plus important (le conditionnement des méthodes itératives étant en général plus mauvais dans le régime de diffusion). Dans un certain nombre de cas physiques ε vaut $1/c$ avec c la vitesse de la lumière et σ peut être très grand. Pour cela une alternative étudiée dans [En07] consiste à impliciter juste le terme source. En effet, dans ce cas on obtient un schéma stable sous une condition qui dépend moins des grandes variations de σ et qui ne nécessite aucune inversion matricielle. Le fait que l'on conserve la propriété « asymptotic preserving » apparaît moins évident mais reste cependant vrai comme l'illustre le calcul

suivant. On pose $E(t_n, \mathbf{x}) = E^n$, $\mathbf{F}(t_n, \mathbf{x}) = \mathbf{F}^n$. On utilise le schéma semi-discret en temps et non discrétisé en espace.

$$\left\{ \begin{array}{l} \dfrac{E^{n+1} - E^n}{\Delta t} + \dfrac{1}{\varepsilon} \operatorname{div}(F^n) = 0, \\ \dfrac{\mathbf{F}^{n+1} - \mathbf{F}^n}{\Delta t} + \dfrac{1}{\varepsilon} \nabla E^n = -\dfrac{\sigma}{\varepsilon^2} \mathbf{F}^{n+1}, \end{array} \right. \tag{2.102}$$

la seconde équation montre que

$$\mathbf{F}^{n+1} = -\frac{\varepsilon}{\sigma} \nabla \mathbf{F}^n + O(\varepsilon^2).$$

Il suffit de combiner cette expression à un développement de Taylor en temps d'ordre 1 qui donne que $\mathbf{F}^{n+1} = \mathbf{F}^n + \Delta t \partial_t \mathbf{F}^n + O(\Delta t^2)$ pour obtenir l'équation limite

$$\frac{E^{n+1} - E^n}{\Delta t} - \operatorname{div}\left(\frac{1}{\sigma} \nabla E^n \right) + \frac{\Delta t}{\varepsilon} \partial_t \operatorname{div}(\mathbf{F}^n) + O(\Delta t^2 + \varepsilon^2) = 0$$

Enfin on remarque que $\partial_t \operatorname{div}(\mathbf{F}^n) = -\varepsilon(\partial_{tt} E^n + O(\Delta t))$. Il suffit d'injecter cette formule dans l'équation précédente et puisque E est régulier en temps, on obtient des dérivées bornées et donc la limite de diffusion de ce schéma semi-implicite est

$$\frac{E^{n+1} - E^n}{\Delta t} - \operatorname{div}\left(\frac{1}{\sigma} \nabla E^n \right) = O(\Delta t).$$

Ce schéma offre donc une troisième possibilité de discrétisation temporelle qui peut être utile pour un certain nombre cas test. Pour finir on introduit les discrétisations semi-implicites aux noeuds. La construction d'un schéma semi-implicite pour le schéma JL-(a) est immédiate. Pour le schéma JL-(b) on utilise la formulation suivante (2.93).

Le schéma « asymptotic preserving » semi-implicite JL-(a) est donné par

$$\left\{ \begin{array}{l} \mid \Omega_j \mid \dfrac{E_j^{n+1} - E_j^n}{\Delta t} + \dfrac{1}{\varepsilon} \displaystyle\sum_r (l_{jr} \mathbf{n}_{jr}, \mathbf{F}_r^n) = 0, \\ \mid \Omega_j \mid \dfrac{\mathbf{F}_j^{n+1} - \mathbf{F}_j^n}{\Delta t} + \dfrac{1}{\varepsilon} \displaystyle\sum_r \mathbf{G}_{jr}^n = -\dfrac{\sigma}{\varepsilon^2} \mid \Omega_j \mid \mathbf{F}_j^{n+1}, \\ \mathbf{G}_{jr}^n = l_{jr} E_j^n \mathbf{n}_{jr} + \widehat{\alpha}_{jr} (\mathbf{F}_j^n - \mathbf{F}_r^n) - \dfrac{\sigma}{\varepsilon} \widehat{\beta}_{jr} \mathbf{F}_r^n, \\ \left(\displaystyle\sum_j \widehat{\alpha}_{jr} + \dfrac{\sigma}{\varepsilon} \sum_j \widehat{\beta}_{jr} \right) \mathbf{F}_r^n = \displaystyle\sum_j l_{jr} \mathbf{n}_{jr} E_j^n + \sum_j \widehat{\alpha}_{jr} \mathbf{F}_j^n. \end{array} \right. \tag{2.103}$$

Le schéma « asymptotic preserving » semi-implicite JL-(b) est donné par

$$\left\{ \begin{array}{l} \mid \Omega_j \mid \dfrac{E_j^{n+1} - E_j^n}{\Delta t} + \dfrac{1}{\varepsilon} \displaystyle\sum_r (l_{jr} \mathbf{n}_{jr}, M_r \mathbf{F}_r) = 0, \\ \mid \Omega_j \mid \dfrac{\mathbf{F}_j^{n+1} - \mathbf{F}_j^n}{\Delta t} + \dfrac{1}{\varepsilon} \displaystyle\sum_r \widehat{\alpha}_{jr} M_r (\mathbf{F}_j - \mathbf{F}_r) = -\dfrac{1}{\varepsilon} \left(\displaystyle\sum_r \widehat{\alpha}_{jr} (\widehat{I}_d - M_r) \right) \mathbf{F}_j^{n+1}, \\ \left(\displaystyle\sum_j \widehat{\alpha}_{jr} \right) \tilde{\mathbf{F}}_r = \displaystyle\sum_j l_{jr} \mathbf{n}_{jr} E_j^n + \sum_j \widehat{\alpha}_{jr} \mathbf{F}_j^n. \end{array} \right. \tag{2.104}$$

avec

$$M_r = \left(\sum_j \widehat{\alpha}_{jr} + \sum_j \frac{\sigma}{\varepsilon} \widehat{\beta}_{jr} \right)^{-1} \left(\sum_j \widehat{\alpha}_{jr} \right), \quad \widehat{\beta}_{jr} = l_{jr} \mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j).$$

Les résultats numériques montrent que les schémas semi-implicites sont valides en régime de diffusion et que la condition CFL pour le schéma JL-(a) semi-implicite est de type transport. Pour le schéma JL-(b) semi-implicite les résultats numériques montre que la condition CFL est au pire parabolique.

2.3.2 Conditions limites

Nous détaillons dans ce paragraphe les conditions de Neumann (pour le schéma de diffusion (2.13)) et les conditions de flux nul (pour les schémas « asymptotic preserving »). La condition de Neumann pour l'équation de diffusion s'écrit

$$(\nabla E(t,\mathbf{x}),\mathbf{n}_{ext})=0.$$

La condition de type flux nul pour l'équation de la chaleur hyperbolique

$$(\mathbf{F}(t,\mathbf{x}),\mathbf{n}_{ext})=0,$$

où $\mathbf{n}_{ext}=(n^1_{ext},n^2_{ext})$ est la normale extérieure. Le résultat (2.63) montre que $\mathbf{F}_r$ est consistant avec le gradient $\nabla E(t,\mathbf{x}_r)$. Par conséquent la condition de Neumann pour l'équation de diffusion est équivalente $(\mathbf{F}_r,\mathbf{n}_{ext})=0$ avec $\mathbf{F}_r$ le flux de diffusion. Finalement cette condition se traite de la même façon que pour les schémas discrétisant l'équation de la chaleur hyperbolique et l'équation de diffusion. Cependant la matrice A_r n'est pas toujours inversible sur les bords. Pour résoudre cette difficulté on utilise la méthode décrite dans [Klu08]- [CDDL09]- [Maz07], dans laquelle la condition $(\mathbf{F},\mathbf{n})=0$ se réécrit de la façon suivante

$$\begin{cases} (\mathbf{F}_r,\mathbf{n}_{ext})=0 \\ (A_r\mathbf{F}_r,\mathbf{t}_{ext})=(\sum_j l_{jr}\mathbf{n}_{jr}E_j,\mathbf{t}_{ext}), \end{cases} \tag{2.105}$$

et $\mathbf{t}_{ext}=(-\mathbf{n}^2_{ext},\mathbf{n}^1_{ext})$ est le vecteur tangent à la frontière. Pour les coins du maillage on prend $\mathbf{F}_r=\mathbf{0}$. Ces résultats peuvent se retrouver en utilisant des arguments de mécanique comme des arguments de type mailles fantômes.

La condition de Dirichlet consiste à imposer une valeur de E aux bords. Pour traiter cette condition on utilise la condition dite de « pression imposée » définie [Klu08]- [CDDL09]- [Maz07]. Si on souhaite imposer une $E=E_{ext}$ sur le bord, il suffit de résoudre

$$\left(\sum_j l_{jr}\mathbf{n}_{jr}\otimes\mathbf{n}_{jr}+\frac{\sigma}{\varepsilon}\sum_j l_{jr}\mathbf{n}_{jr}\otimes(\mathbf{x}_r-\mathbf{x}_j)\right)\mathbf{F}_r=\sum_j l_{jr}(E_j-E_{ext})\mathbf{n}_{jr}+\sum_j l_{jr}\mathbf{n}_{jr}\otimes\mathbf{n}_{jr}\mathbf{F}_j.$$

2.4 Modes parasites et stabilisation

Il est connu que les schémas aux noeuds exhibent des modes parasites [Mai11], c'est à dire des comportements dégradant la qualité de la solution numérique du schéma. Ces modes parasites se présentent notamment sur maillage Cartésien et sont liés au fait que sur ce type de maillage les schémas aux noeuds admettent stencil en croix. En effet, supposons un maillage Cartésien où la longueur des carrés est h. Dans ce cas $l_{jr}=\frac{h}{\sqrt{2}}$, les vecteurs $\mathbf{n}_{jr}$ associés à la maille j sont

$$\begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix},\qquad \begin{pmatrix}\frac{-1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{pmatrix},\qquad \begin{pmatrix}\frac{1}{\sqrt{2}}\\ \frac{-1}{\sqrt{2}}\end{pmatrix},\qquad \begin{pmatrix}\frac{-1}{\sqrt{2}}\\ \frac{-1}{\sqrt{2}}\end{pmatrix}.$$

Après de rapides calculs on obtient que

$$A_r = \begin{pmatrix} h^2 & 0 \\ 0 & h^2 \end{pmatrix},$$

et que le schéma de diffusion (2.13) est équivalent à

$$\frac{E^{n+1}_{(i,j)} - E^n_{(i,j)}}{\Delta t} - \frac{E^n_{(i-1,j-1)} + E^n_{(i+1,j-1)} + E^n_{(i-1,j+1)} + E_{(i+1,j+1)^n} - 4E_{(i,j)^n}}{2h^2} = 0. \tag{2.106}$$

Pour des solutions régulières un stencil de ce type n'est pas gênant mais pour certaines données discontinues des modes parasites apparaissent. Par exemple, si on prend une masse de Dirac comme condition initiale, le schéma ne va pas converger. En effet, à cause de ce stencil en croix, un sous-maillage correspondant aux stencils du schéma TPFA ne sera pas utilisé par le schéma. Ceci donnera une solution en sablier [Mai11]. Le second cas se présente sur maillage triangulaire avec des conditions de Neumann, en effet, dans ce cas en temps long les solutions se dégradent et des modes parasites apparaissent. Le schéma peut converger donc vers un état stationnaire différent bien que proche de l'état stationnaire exact. Tant que le gradient discret $\mathbf{F}_r$ peut être nul sans que la solution discrète autour d'un noeud soit constante, le schéma admettra des modes parasites. On propose donc une méthode pour pallier ce défaut des schémas aux noeuds. Une voie possible est d'assurer le taux de production locale d'entropie est non nul lorsque la solution n'est pas constante. On rappelle que dans notre cas l'entropie est donnée par l'opposée de la norme L^2 et que le taux d'entropie discret est donné par $-\sum_r (\mathbf{u}_r, A_r \mathbf{u}_r)$.

Redéfinition des normales

Afin d'assurer que le taux de production d'entropie s'annule que pour des données constantes, on propose de redéfinir les normales et la longueur nodale. A cette fin on utilise

$$l_{jr}\mathbf{n}_{jr} = l_{j,r-\frac{1}{2}}\mathbf{n}_{j,r-\frac{1}{2}} + l_{j,r+\frac{1}{2}}\mathbf{n}_{j,r+\frac{1}{2}}$$

avec $l_{jr\pm\frac{1}{2}}$ les longueurs entre le noeud r et les milieux des arêtes et $\mathbf{n}_{j,r\pm\frac{1}{2}}$ les normales associées à ces arêtes.
L'idée consiste à modifier le produit $l_{jr}\mathbf{n}_{jr}$ en modifiant le milieu des arêtes. Cette modification préserve

$$\sum_j l_{jr}\mathbf{n}_{jr} = \mathbf{0},\ \sum_r l_{jr}\mathbf{n}_{jr} = \mathbf{0}.$$

Pour toute arête de taille d, on définit deux points qui sont le milieu de l'arête décalé de $\pm\alpha$ avec, par exemple $\alpha = \frac{d}{6}$. Au lieu de choisir le milieu de l'arête on choisit un de ces deux points mais le choix doit être le même pour les deux mailles qui ont cette arête en commun. Cette stratégie permet donc de redéfinir un couple $l_{jr}\mathbf{n}_{jr}$. Pour chaque couple de normale et de longueur on a quatre possibilités de définition car on peut utiliser $\pm\alpha$ pour chaque arête connectée à un noeud. Par conséquent autour d'un noeud on obtient un groupe de 16 produits longueur, normale possible pour un maillage quadrangulaire. Pour un maillage Cartésien la procédure est assez simple : Soit $P_{i,j}$ un noeud du maillage, (a,b) deux nombres appartenant à $\left\{\frac{1}{3}, \frac{2}{3}\right\}^2$, à la place du milieu des arêtes on utilise les points $P_{i,j} + a(P_{i+1,j} - P_{i,j})$ et $P_{i,j} + b(P_{i,j+1} - P_{i,j})$. Afin d'obtenir le schéma final on fait la moyenne des 4 choix possibles pour (a, b). Maintenant on obtient un nouveau taux de variation de la norme L^2

$$-\frac{1}{2}\sum_{c=1,4}\sum_r (A^c_r \mathbf{u}_r, \mathbf{u}_r).$$

Grâce à cette modification le gradient et le taux de production d'entropie ne sont nuls que si la solution discrète est constante. Pour obtenir un schéma symétrique sur maillage Cartésien, il faut choisir correctement les nouveaux points qui définissent les nouvelles normales. Cette méthode peut être aussi appliquée pour les schémas JL-(a) et JL-(b).

2.5 Résultats numériques

Dans cette section nous allons donner des résultats numériques pour le schéma de diffusion et pour les schémas « asymptotic preserving » dans différents régimes. On propose de tester ces schémas sur différents maillages introduits par les figures (2.5)-(2.6)-(2.7).

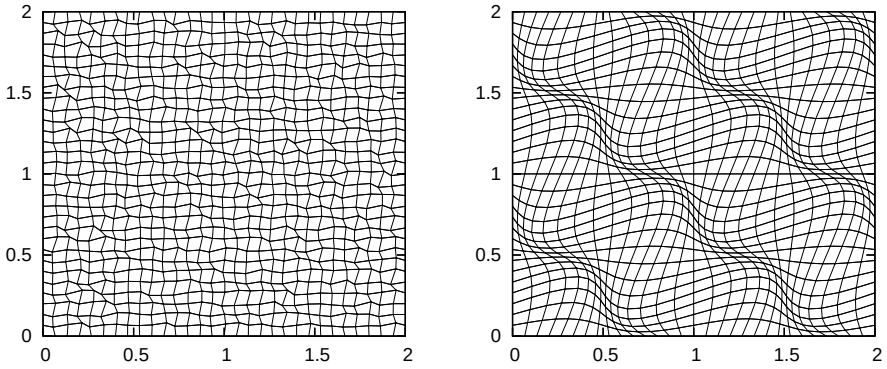


FIGURE 2.5 – Maillages quadrangulaires aléatoire et « smooth »

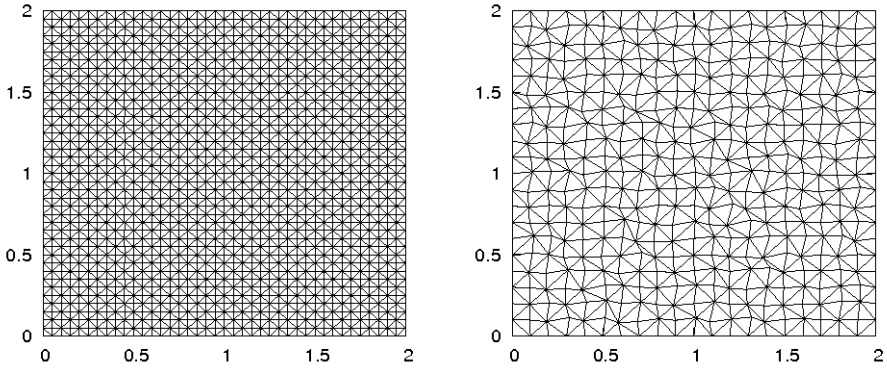


FIGURE 2.6 – Maillages triangulaires régulier et aléatoire

2.5.1 Résultats numériques pour le schéma de diffusion

Cas test : Solution fondamentale

Pour commencer nous étudions le schéma pour la diffusion dans le cas d'une solution régulière. La condition initiale est la solution fondamentale de l'équation de la chaleur à t=0.001. Le temps final est

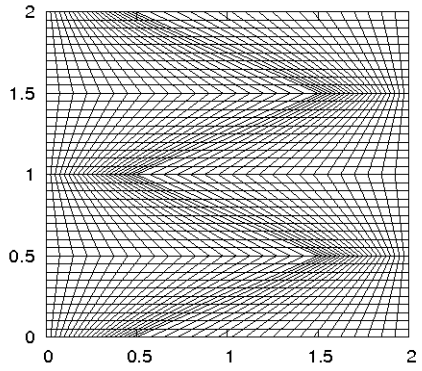


FIGURE 2.7 – Maillage de Kershaw, K=1

$T_f = 0.011$. Dans ce cas test on peut ignorer les conditions limites car la solution est nulle au voisinage du bord $\partial\Omega$.

On remarque sur la figure (2.8) que sur maillage Cartésien le schéma TPFA et le schéma de diffusion aux noeuds défini par (2.13) pour les deux variantes $\widehat{\beta}_{jr}$ (qui sont identiques sur maillage Cartésien) sont **convergents à l'ordre 2** en norme L^1 et L^2.

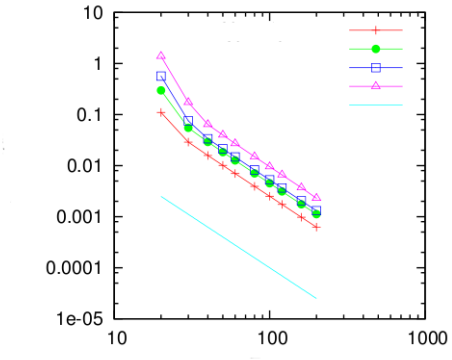


FIGURE 2.8 – Erreurs pour les schémas de diffusion TPFA (1.69) et aux noeuds (2.13) sur maillage **Cartésien**. L'abscisse correspond au nombre de mailles, l'ordonnée à l'erreur numérique liée au schéma. Les courbes composées de traits rouge et cercles vert correspondent à aux erreurs en norme L^1 et L^2 pour le schéma TPFA. Les courbes composées de carrés bleus et de triangles roses correspondent aux erreurs en norme L^1 et L^2 pour le schéma de diffusion aux noeuds. La courbe bleue correspond à la fonction h^2. On note la bonne convergence à l'ordre deux pour les deux schémas.

Comme prévu par la théorie on observe que le schéma TPFA ne converge pas sur maillages non structurés, figure (2.9). C'est pour cette raison que la méthode de Jin-Levermore (1.63)-(1.68) introduite dans le schéma de volumes finis standard avec des flux évalués aux arêtes, ne génère pas de schéma « asymptotic preserving » valide sur maillages non structurés. Cela justifie l'utilisation d'un schéma aux noeuds.

Pour le schéma de diffusion aux noeuds GLACE la convergence et l'ordre dépendent du maillage. Sur la figure (2.10) on note que le schéma (2.13) avec le tenseur $\widehat{\beta}^I_{jr}$ **converge à l'ordre 2** sur maillage aléatoire quadrangulaire. Par contre le schéma (2.13) avec le tenseur $\widehat{\beta}^{II}_{jr}$ ne converge pas sur maillage

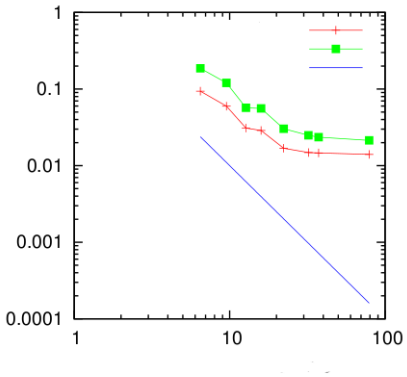


FIGURE 2.9 – Erreurs pour le schéma TPFA sur le maillage **aléatoire** de la figure (2.5). L'abscisse correspond au nombre de mailles, l'ordonnée à l'erreur numérique liée au schéma. Les courbes verte et rouge correspondent aux normes L^1 et L^2 de l'erreur numérique associée au schéma TPFA. La courbe bleue correspond à la fonction h^2. On note la non convergence du schéma TPFA

aléatoire.

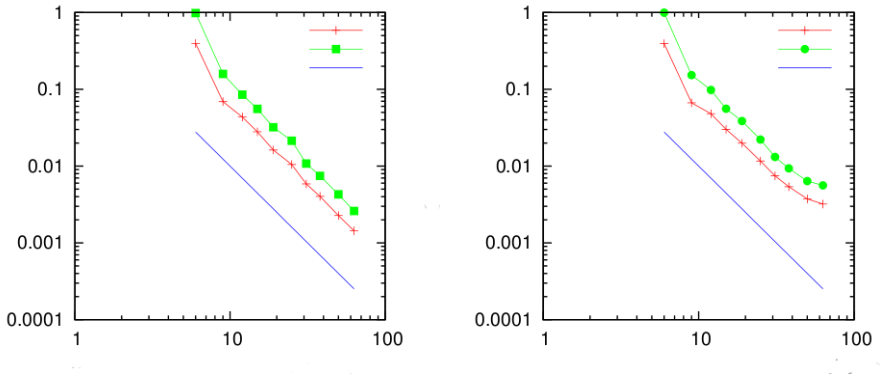


FIGURE 2.10 – Erreur pour le schéma aux noeuds avec la variante $\widehat{\beta}^I_{jr}$ sur la gauche et la variante $\widehat{\beta}^{II}_{jr}$ sur la droite pour un maillage **aléatoire** quadrangulaire. L'abscisse correspond au nombre de mailles, l'ordonnée à l'erreur numérique liée au schéma. Les courbes verte et rouge correspondent aux normes L^1 et L^2 de l'erreur numérique associée au schéma TPFA. La courbe bleue correspond à la fonction h^2. On note la non convergence pour la seconde variante.

Dans les tables (2.2) et (2.3) on donne l'erreur de convergence en norme L^1 entre la solution exacte et la solution numérique du schéma (2.13) pour $\widehat{\beta}_{jr} = l_{jr}\mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)$ sur différents maillages. Pour évaluer l'ordre nous utilisons la formule suivante : Soit h_1, h_2 les tailles caractéristiques des deux maillages (par exemple la longueur de la plus grande arête) et e_{h_1}, e_{h_2} les erreurs associées à ces deux maillages. L'ordre de convergence est calculé à partir de la formule

$$q = \frac{\log\left(e_{h_1}/e_{h_2}\right)}{\log\left(h_1/h_2\right)}.$$

Le schéma aux noeuds (2.13) est un schéma d'ordre deux sur maillage quadrangulaire. Par contre on remarque que sur maillage triangulaire on ne converge pas toujours à l'ordre 2.

Nb cellules\Maillages	Cartésien	Aléa. quad.	Smooth quad.	Kershaw
40-80	2.06	2.03	2.15	2.58
80-160	2.01	2.43	2.04	1.99
160-320	2.00	2.04	2.02	1.99

TABLE 2.2 – Ordre de convergence sur maillages quadrangulaires

Nb cellules\Maillages	Tri. régulier mod.	Tri. régulier	Tri aléa.
40-80	0.94	2.18	1.81
80-160	0.95	2.00	1.3
160-320	0.98	2.00	1.4

TABLE 2.3 – Ordre de convergence pour les maillages triangulaires. Le maillage triangulaire régulier est donné à gauche de la figure 2.6. Le maillage triangulaire régulier modifié est obtenu à partir d'un maillage Cartésien. On coupe toujours les quadrangles en deux dans la même direction.

Cas test : Solution périodique

Dans cette section nous regardons une solution périodique notamment afin de valider la discrétisation des conditions de Neumann. On prend $E(0,x,y) = \cos(\pi x)\cos(\pi y)$, $D =]0,2[\times]0,2[$. La solution de l'équation de la chaleur avec des conditions de Neumann est

$$E(t,x,y) = \exp(-2\pi^2 t)\cos(\pi x)\cos(\pi y).$$

Dans les tableaux (2.4)-(2.5) on présente les ordres de convergence du schéma aux noeuds avec la première variante tensorielle sur différents maillages. On constate que le schéma converge avec un ordre compris entre un et deux pour ce cas test.

Nb cellules\Maillages	Cartésien	Aléa. quad.	« Smooth » quad.	Kershaw
40-80	2.00	2.04	1.99	1.97
80-160	2.00	2.01	2.00	1.98
160-320	2.00	1.98	2.00	1.99

TABLE 2.4 – Ordre de convergence sur maillages quadrangulaires.

Nb cellules\Maillages	Tri. régulier mod.	Tri. régulier	Tri. aléa
40-80	1.00	2.00	1.28
80-160	1.02	2.00	1.31
160-320	1.00	2.00	1.33

TABLE 2.5 – Ordre de convergence sur maillages triangulaires.

2.5.2 Résultats numériques en régime de transport

Dans cette section on souhaite vérifier numériquement que la modification introduite dans le schéma pour l'équation de la chaleur hyperbolique n'altère pas la convergence du schéma initial en régime de

transport ($\varepsilon = O(1)$ et $\sigma = 0(1)$).

Cas test : solution fondamentale

Le premier cas test consiste à étudier la solution fondamentale de l'équation de la chaleur hyperbolique (1.62). La condition initiale est une masse de Dirac centrée au milieu du domaine. Le temps final est $T = 0.5$. La solution exacte est composée d'une masse de Dirac se propageant à la vitesse 1 et d'une fonction régulière entre la masse Dirac et sa position initiale. A $t = 0.5$ la partie régulière est négative, mais devient positive en temps long. La trace de la solution fondamentale sur maillage Cartésien avec et sans stabilisation des modes parasites est donnée par la figure (2.11).

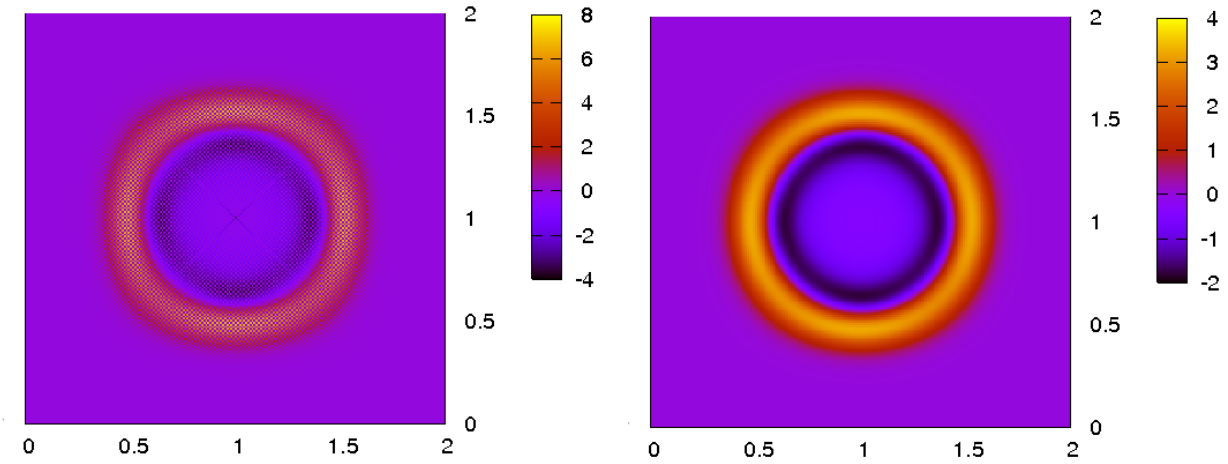


FIGURE 2.11 – Solution fondamentale de l'équation de la chaleur hyperbolique aux temps 0.5 sans et avec stabilisation. On trace la variable $E(t, \mathbf{x})$.

La solution numérique correspond approximativement à la solution exacte, la masse de Dirac se propage à la bonne vitesse, par contre, naturellement la masse de Dirac est étalée à cause de la diffusion numérique.

Cas test : solution périodique : le système (1.62) avec $\varepsilon = 1$ et $\sigma = 1$ est équivalent à l'équation scalaire suivante :

$$\begin{cases} \partial_{t^2} f + \partial_t f - \triangle f = 0, \\ f(t=0,x,y) = f_0(x,y), \qquad \partial_t f(t=0,x,y) = f_1(x,y). \end{cases} \tag{2.107}$$

Cette équation est nommée équation du télégraphe. On va utiliser l'équivalence entre notre système et cette équation pour construire une solution de référence. Soit la condition initiale suivante et le domaine $D =]0,2[\times]0,2[$

$$f_0(x,y) = \cos(\pi x)\cos(\pi y) \qquad f_1(x,y) = 0.$$

La solution est

$$f(t,x,y) = e^{-\frac{t}{2}} \left(\cos\left(\frac{\sqrt{8\pi^2 - 1}}{2} t \right) - \frac{1}{\sqrt{8\pi^2 - 1}} \sin\left(\frac{\sqrt{8\pi^2 - 1}}{2} t \right) \right) \cos(\pi x)\cos(\pi y).$$

L'équivalence entre les deux modèles se déduit du changement de variable $E = \partial_t f + f$ et $\mathbf{F} = -\nabla f$, on obtient donc une solution de référence pour l'équation de la chaleur hyperbolique. La condition limite est une condition de Neumann $(\mathbf{F}, \mathbf{n}) = 0$ décrite précédemment. On trace l'erreur pour différents choix de tenseur $\widehat{\alpha}_{jr}$ sur la figure (2.12) qui donne les courbes de convergence pour le schéma JL-(a) et pour le schéma de Jin-Levermore aux arêtes (1.63)-(1.68).

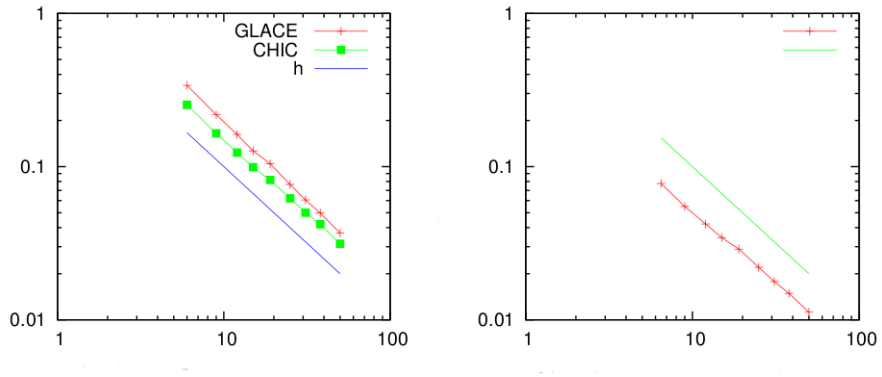


FIGURE 2.12 – Erreur en norme L^2 pour le schéma aux noeuds JL-(a) avec les différentes variantes du tenseur acoustique sur la gauche et pour le schéma aux arêtes sur la droite. Sur la figure de droite la courbe rouge correspond au l'erreur numérique et la courbe verte à la fonction $\frac{1}{h}$. Le maillage est un maillage aléatoire.

Pour toutes les variantes le schéma aux noeuds JL-(a) **converge avec un ordre autour de 1** sur les différents maillages.

Remarque 2.5.1. *On ne donne pas les résultats pour le schéma JL-(b) car il sont quasiment identiques à ceux donnés pour JL-(a).*

2.5.3 Résultats numériques en régime asymptotique

Cas test de diffusion

Pour les exemples présentés dans cette section il est important d'utiliser un schéma implicite où semi-implicite pour éviter la condition CFL dépendante de ε. Les résultats exposés sont calculés avec un schéma semi-implicite en temps. On considère que pour $\varepsilon = O(10^{-3})$ on est dans le régime asymptotique. On compare la solution numérique pour l'équation de la chaleur hyperbolique avec la solution exacte des deux cas test de diffusion. On observe que : **les schémas JL-(a) et JL-(b) convergent avec un ordre proche de 2 quand ε tend vers 0** (tableaux (2.6)-(2.7)-(2.8)-(2.9)-(2.10)-(2.11)).

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	1.90	1.99	2.00	2.00
80-160	1.82	1.97	2.00	2.00
160-320	1.72	1.95	2.00	2.00

TABLE 2.6 – Résultats de convergence pour le schéma JL-(b) sur maillage Cartésien

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	1.93	2.01	2.02	2.03
80-160	1.93	2.08	2.1	2.1
160-320	1.76	1.97	2.00	2.00

TABLE 2.7 – Résultats de convergence pour le schéma JL-(b) sur maillage Cartésien aléatoire

Nombre de cellules /ε	$\epsilon = 10^{-3}$	$\epsilon = 10^{-4}$	$\epsilon = 10^{-6}$	$\epsilon = 10^{-7}$
40-80	1.96	2.03	2.04	2.05
80-160	1.86	2.00	2.01	2.01
160-320	1.77	1.97	2.01	2.01

TABLE 2.8 – Résultats de convergence pour le schéma JL-(b) sur maillage « smooth »

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	2.4	2.5	2.55	2.56
80-160	1.74	1.82	1.83	1.84
160-320	1.80	1.96	1.99	2

TABLE 2.9 – Résultats de convergence pour le schéma JL-(b) sur maillage de Kershaw

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	2.26	2.14	2.15	2.15
80-160	1.99	2.00	2.00	2.00
160-320	1.99	2.00	2.00	1.99

TABLE 2.10 – Résultats de convergence pour le schéma JL-(b) sur maillage triangulaire régulier

Nombre de cellules /ε	$\epsilon = 10^{-3}$	$\epsilon = 10^{-4}$	$\epsilon = 10^{-6}$	$\epsilon = 10^{-7}$
40-80	1.80	1.81	1.83	1.84
80-160	1.32	1.30	1.29	1.29
160-320	1.43	1.43	1.43	1.43

TABLE 2.11 – Résultats de convergence pour le schéma JL-(b) sur maillage triangulaire aléatoire

Ces résultats de convergence méritent certaines remarques. L'ordre de convergence pour les valeurs $\varepsilon = 0.001$ et $\varepsilon = 0.0001$ décroît légèrement. Le phénomène est normal. En effet le résultat de convergence en dimension un [BDF11] montre que l'erreur de consistance est homogène à $(h^2 + \varepsilon h)$ avec h le pas de maillage. Par conséquent, sur maillage grossier ($h >> \varepsilon$) le schéma converge avec l'ordre deux. Si le maillage est fin ($h << \varepsilon$) le schéma converge avec l'ordre un. Quand $\varepsilon = 0.001$ et $\varepsilon = 0.0001$ le ratio $\frac{h}{\varepsilon}$ est proche de 1, il est donc normal de voir l'ordre se rapprocher de l'ordre 1. Il existe un autre argument existe pour justifier ce phénomène. L'erreur est calculée entre la solution numérique de l'équation de la chaleur hyperbolique et la solution exacte de diffusion. Or l'écart entre la solution exacte de l'équation de chaleur hyperbolique et la solution de diffusion est homogène à $O(\varepsilon)$. Il est

donc normal que l'ordre de convergence se dégrade lorsque l'erreur numérique s'approche de ε.

Il est également intéressant de regarder le comportement du schéma pour h fixé ε qui tend vers 0. On choisit un maillage aléatoire quadrangulaire avec 100 mailles dans chaque direction. Les résultats sont donnés dans le tableau (2.12). On remarque que quand ε tend vers 0, l'erreur commise par le

ε	0.01	0.001	0.0001	0.00001	diffusion
erreur L1	0.016401	0.00813369	0.00710269	0.00699941	0.00698787
erreur L2	0.0313253	0.01511494	0.0132377	0.0130463	0.0130249

TABLE 2.12 – Erreur entre la solution exacte de l'équation de diffusion et la solution numérique de l'équation de la chaleur hyperbolique pour un maillage donné de taille $\Delta x \approx \frac{1}{100}$. On utilise le schéma JL-(b). La dernière colonne donne l'erreur faite par le schéma de diffusion sur le même cas test.

schéma JL-(b) tend vers l'erreur commise par le schéma de diffusion.
Enfin on propose quelques remarques sur le coût de calcul et le conditionnement du schéma JL-(b) implicites lorsque ε est petit. On donne le nombre d'itérations de la méthode GMRES, associé au schéma JL-(b) (2.22)-(2.23) pour deux pas de temps : $\Delta t = \frac{1}{2}h$ et $\Delta t = \frac{1}{2}h^2$, pour deux maillages : Kershaw et Cartésien et pour différentes valeurs de ϵ. On utilise un maillage 80×80. On observe dans

Maillage, pas de temps ϵ	1	0.01	10^{-4}	10^{-6}	10^{-8}
Maillage de Kershaw, $\Delta t = \frac{1}{2}h$	9-12	195-310	1250-2470	1230-2325	1880-2010
Maillage de Kershaw, $\Delta t = \frac{1}{2}h^2$	2-3	4-12	17-39	19-39	19-39
Maillage Cartésien, $\Delta t = \frac{1}{2}h$	7-9	80-110	230-350	235-355	230-350
Maillage Cartésien, $\Delta t = \frac{1}{2}h^2$	2	3-10	7-18	8-19	8-19

TABLE 2.13 – Nombre d'itérations de la méthode GMRES appliquée au schéma JL-(b) pour différents pas de temps, maillages et valeur ε.

le tableau (2.13) que

- Le nombre d'itérations de la méthode GMRES décroît avec le pas de temps, ce qui n'est pas surprenant.
- Le nombre d'itérations de la méthode GMRES augmente avec la déformation du maillage. Cette remarque est à relier avec la précédente.
- Le nombre d'itérations de la méthode GMRES est indépendant de la valeur ε pour des petites valeurs ε.

Le schéma JL-(b) admet une CFL hyperbolique lorsque ε n'est pas trop petit et une CFL parabolique lorsque ε tend vers zéro. On observe le même type de comportement pour le conditionnement des matrices du schéma implicite. En effet à partir de $\varepsilon = 10^{-5}$, $\varepsilon = 10^{-6}$ le conditionnement ne se dégrade plus. Cela correspond à peu près au moment où la CFL parabolique domine pour le schéma semi-implicite.

Cas test avec un σ variable

Pour illustrer cette étude, nous proposons un cas test un peu plus proche de la physique liée au problème étudié pour l'équation de diffusion et l'équation de la chaleur hyperbolique. Le domaine est un carré $[0,1] \times [0,1]$ où $\sigma = 1$ possédant à l'intérieur deux autres carrés :

$$\left[\frac{3}{16}, \frac{7}{16}\right] \times \left[\frac{9}{16}, \frac{13}{16}\right] et \left[\frac{9}{16}, \frac{13}{16}\right] \times \left[\frac{3}{16}, \frac{7}{16}\right],$$

où $\sigma = 10000$. Le paramètre $\varepsilon = 1$ dans tout le domaine (on rappelle que le régime asymptotique est aussi atteint quand σ tend vers ∞). La condition initiale pour l'équation de diffusion est $E(t=0,\mathbf{x}) = 0.001 + 100\exp^{\left(-\left(\frac{||\mathbf{x}||}{0.1}\right)^2\right)}$ et les conditions initiales pour l'équation de la chaleur hyperbolique sont

- $E(t=0,\mathbf{x}) = 0.001 + 100\exp^{\left(-\left(\frac{||\mathbf{x}||}{0.1}\right)^2\right)}$,
- $\mathbf{F}(t=0,\mathbf{x}) = \mathbf{0}$.

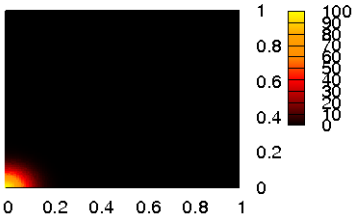


FIGURE 2.13 – Condition initiale pour $E(t,\mathbf{x})$.

On trace les résultats pour l'équation de diffusion.

Dans les deux carrés à l'intérieur du domaine le coefficient de diffusion $\frac{1}{\sigma}$ est très petit comparé aux coefficients dans le reste du domaine, par conséquent la diffusion est beaucoup plus lente dans les deux carrés. On observe ce phénomène sur la figure (2.14). L'équation de la chaleur hyperbolique est une équation d'onde amortie, par conséquent son comportement est très différent de celui de l'équation de diffusion.

Trois remarques : premièrement, dans les deux carrés l'équation de la chaleur hyperbolique se comporte comme l'équation de diffusion. L'équation de la chaleur hyperbolique peut développer des solutions négatives. On observe cette propriété sur la figure 2.15. Cette particularité, ainsi que son côté fortement ondulatoire en fait un assez mauvais candidat pour simuler la phase de transport de particules. Cependant on voit qu'un modèle comme la diffusion est encore plus mauvais en raison de sa vitesse de propagation infinie. L'approximation P_1 une avec limite asymptotique correctement capturée est donc une meilleure option que l'approximation diffusion pure.

Cas test de treillis

On regarde un cas test de [Bru02] issu de la neutronique pour le modèle P_1. Il s'agit de résoudre

$$\begin{cases} \partial_t E + \dfrac{1}{\varepsilon} \operatorname{div} \mathbf{F} = \sigma_a E + Q, \\ \partial_t \mathbf{F} + \dfrac{1}{\varepsilon} \nabla E = -\dfrac{\sigma}{\varepsilon^2} \mathbf{F}. \end{cases} \tag{2.108}$$

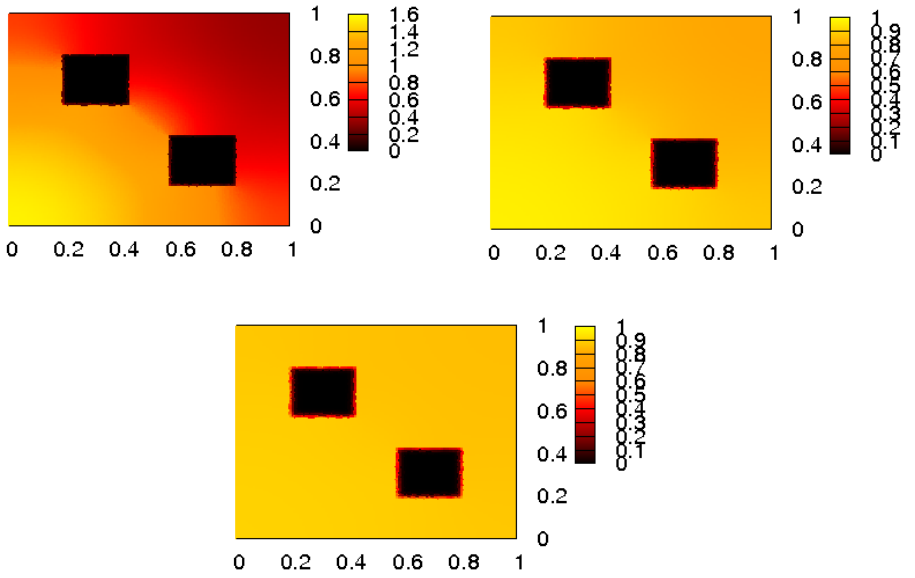


FIGURE 2.14 – Solution de l'équation de diffusion aux temps $t = 0.2$, $t = 0.5$, $t = 0.7$. Chaque carré est maillé de façon aléatoire. Le maillage compte 10000 mailles.

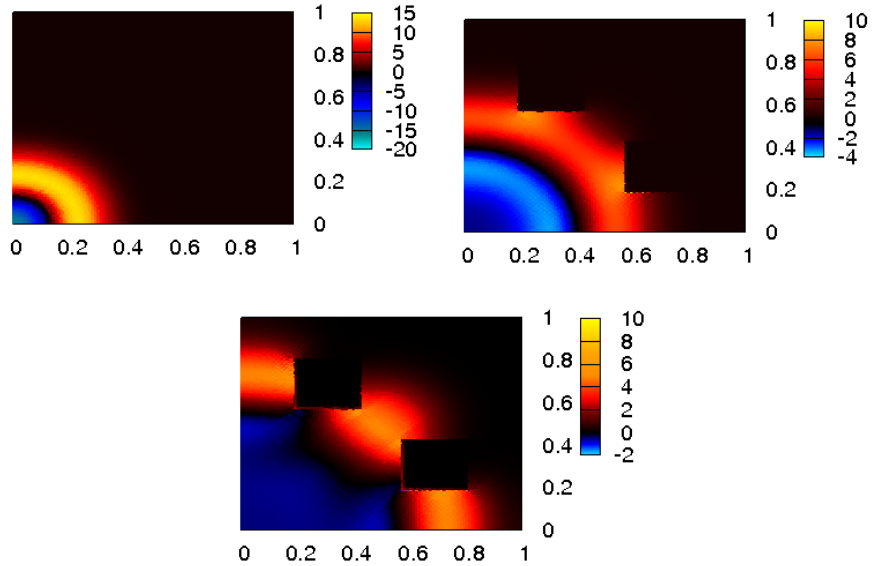


FIGURE 2.15 – Solution $E(t, \mathbf{x})$ de l'équation de la chaleur hyperbolique aux temps $t = 0.2$, $t = 0.5$, $t = 0.7$.

avec σ_a le coefficient d'absorption et Q une source. Le domaine est un carré de taille 7. Le maillage est un pavage de carrés représentés sur la figure (2.16).

Dans le domaine blanc et le carré central les paramètres sont définis par $\sigma = 3$ et $\sigma_a = 0$. Dans les autres carrés $\sigma = 0$ et $\sigma_a = 30$. Ce type de géométrie et de problème apparaissent pour des problèmes de criticités de réacteur nucléaire. Le coeur des réacteurs nucléaires présente cette structure de pavage,

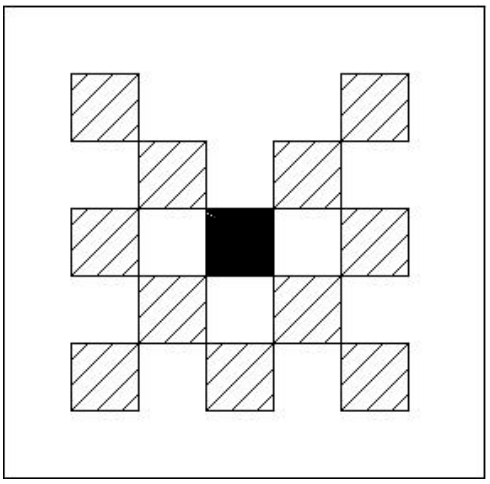


FIGURE 2.16 – Géométrie du problème

et les calculs de criticités nécessite de jouer sur les sections efficaces des pavés pour atteindre le régime critique. Afin de récupérer les bonnes vitesses du modèle P_1 et les bons coefficients on pose $\varepsilon = \sqrt{3}$. L'énergie et le flux sont initialement nuls et la source $Q = 1$. Le temps final est de 3.2 secondes. Le résultat (2.17) est comparable aux résultats obtenus dans [Bru02] avec de la diffusion numérique

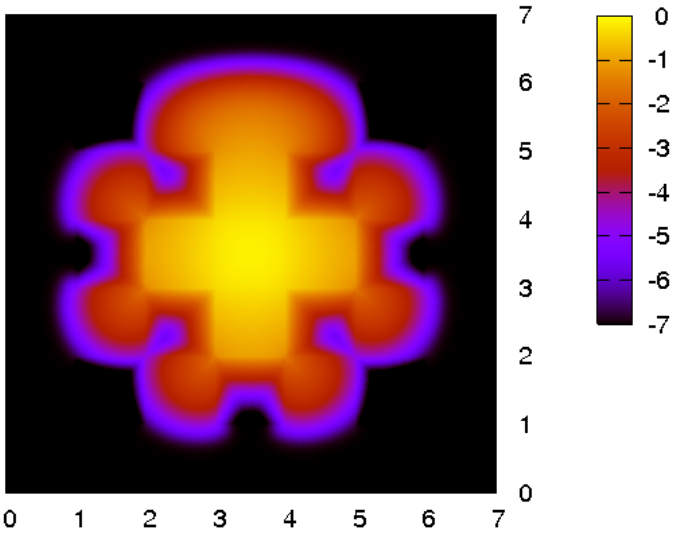


FIGURE 2.17 – Résultat du test sur un maillage 280*280 aléatoire sauf les frontières entre les domaines.

supplémentaire.

2.6 Synthèse et perspectives

Dans cette section, on a proposé deux schémas « asymptotic preserving » basés sur une formulation aux noeuds des méthodes de volumes finis pour l'équation de la chaleur hyperbolique. Le premier schéma JL-(a) (2.11)-(2.12) dérive en 1D sur le schéma de Jin-Levermore, quand au schéma JL-(b) (2.22)-(2.23) il dérive sur le schéma de Gosse-Toscani en 1D. Les deux schémas admettent un schéma de diffusion limite qui converge à l'ordre un d'un point de vue théorique et à l'ordre deux d'un point de vue numérique. En outre, on a montré que le schéma semi-discret JL-(b) (2.22)-(2.23) est stable en norme L^2 et qu'on pouvait obtenir un schéma semi-implicite avec une condition CFL indépendante de ε. Cependant ces schémas génèrent des modes parasites dans un certains nombre de cas. Une méthode de stabilisation a été proposée. Dans ce chapitre on montre que les schémas aux noeuds paraissent dans un premier temps plus adaptés pour la construction de schémas « asymptotic preserving » . Par ailleurs, ces schémas ont l'avantage d'être facilement extensibles en 3D et sur maillages curvilignes [BEHN10]. En effet les schémas aux noeuds peuvent être reformulés à l'aide de la quantité géométrique $\mathbf{C}_{jr} = \nabla_{\mathbf{x}_r}\Omega_j$, comme dans [CDDL09]. En 2D cette quantité est égale à $l_{jr}\mathbf{n}_{jr}$. Dans un premier temps on peut remarquer que la définition donnée précédemment est indépendante de la dimension ce qui permet de généraliser le schéma GLACE en dimension supérieure (voir [CDDL09]). Deuxièmement l'extension du schéma GLACE hydrodynamique à des maillages non conformes ou d'ordre 2 est faite uniquement en redéfinissant ces $\mathbf{C}_{jr}$. Par conséquent, ces extensions réalisées dans le cadre l'hydrodynamique devraient s'appliquer naturellement aux schémas aux noeuds pour l'équation de la chaleur hyperbolique et l'équation de diffusion. Dans la section suivante on propose une autre méthode afin d'obtenir une discrétisation « asymptotic preserving » de l'équation de la chaleur hyperbolique basée sur une formulation aux arêtes des méthodes de volumes finis.

Chapitre 3

Schémas linéaires aux arêtes pour l'équation de la chaleur hyperbolique

L'objectif de cette étude (issue d'une collaboration avec Gérald Samba, Philippe Hoch du CEA et Pierre Navarro de l'université de Strasbourg, effectuée lors du CEMRACS 2010), est de construire des schémas « asymptotic preserving » utilisant une formulation aux arêtes et admettant comme schémas limites les schémas de diffusion sur maillages non structurés polygonaux de type MPFA-O [AE06] où Breil-Maire [BM06]. L'approche introduite dans ce chapitre est différente de celle des chapitres précédent. On partait d'un schéma pour l'équation de la chaleur hyperbolique (schémas aux noeuds ou aux arêtes) puis on le couplait avec la méthode de Jin-Levermore. Dans le premier cas on obtenait un schéma limite aux noeuds de diffusion convergent et dans le second cas le schéma de diffusion TPFA (1.69) non consistant sur maillages déformés.

Ici la façon de procéder est différente : on part d'un schéma de diffusion existant et valable sur maillages non structurés, puis on construit un schéma pour l'équation de la chaleur hyperbolique sans terme source compatible avec le schéma de diffusion choisi. Enfin on couple ce nouveau schéma à l'approche de Jin-Levermore afin d'obtenir une discrétisation « asymptotic preserving » qui converge vers le schéma de diffusion cible lorsque ε tends vers 0. Ces schémas cibles ont l'avantage, contrairement au schéma de diffusion aux noeuds, de ne pas admettre de modes parasites et de coïncider avec le schéma TPFA pour un maillage Cartésien. Cette méthode pourrait s'adapter à d'autres schémas de diffusion comme le schéma SUSHI [EGH09].
La principale difficulté est d'obtenir des schémas convergents indépendemment de ε sur des maillages très déformés et polygonaux. Les figures (3.1) et (3.2) donnent des exemples de maillages issus de simulations d'hydrodynamique effectuées à l'aide d'une méthode de type ALE.

3.1 Schémas de diffusion Breil-Maire et MPFA

Le premier schéma cible est celui de Breil-Maire [BM06]. Ce schéma est symétrique, admet un stencil à 9 points sur les maillages quadrangulaires et est correctement défini pour tous les maillages. Par contre, il est non consistant sur maillage aléatoire. On commencera par des rappels sur le schéma de Breil-Maire, puis on proposera une modification de ce schéma afin de le rendre consistant, ce qui nous permettra de retrouver le schéma de diffusion MPFA-O. Cette manipulation permet de fusionner les deux schémas dans le même formalisme afin de construire un unique schéma « asymptotic preserving » pour l'équation de la chaleur hyperbolique.

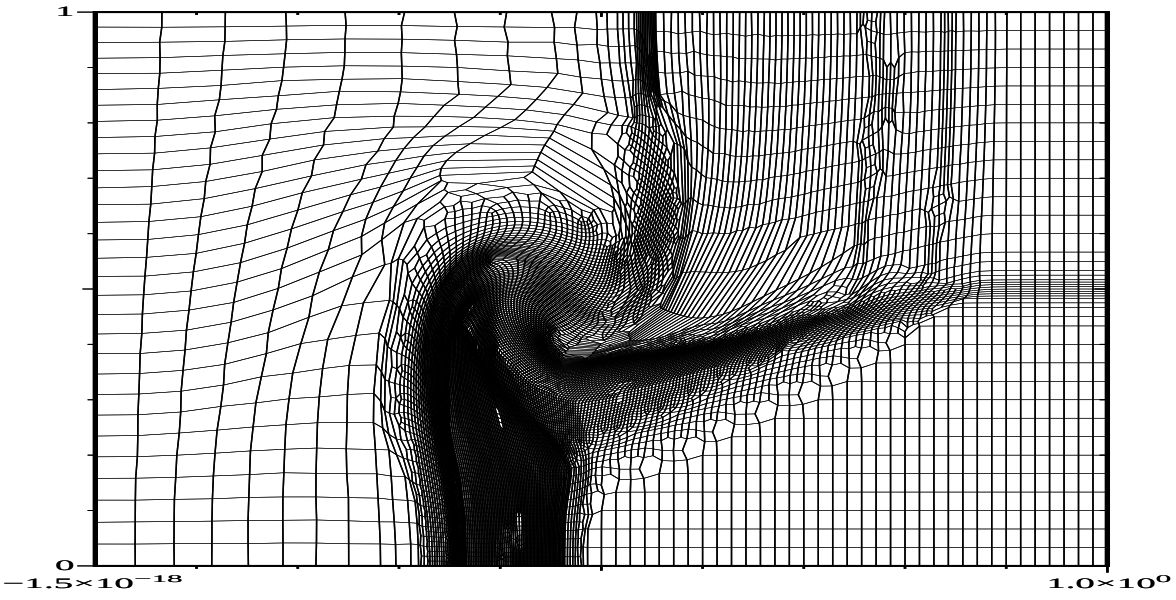


FIGURE 3.1 – Maillage polygonal, issu d'une simulation hydrodynamique avec une méthode ALE

Définition 3.1.1. *Le schéma de Breil-Maire [BM06] s'écrit :*

$$\begin{cases} \mid \Omega_j \mid \partial_t E_j(t) + \displaystyle\sum_r \frac{1}{2}(l_{j,r-1,r}\boldsymbol{\Phi}^j_{r-\frac{1}{2},r} + l_{j,r,r+1}\boldsymbol{\Phi}^j_{r,r+\frac{1}{2}}) = 0, \\ \\ \frac{1}{2}l_k(\boldsymbol{\Phi}^{k-1}_k + \boldsymbol{\Phi}^k_k) = \mathbf{0}, \\ \\ \begin{pmatrix} \boldsymbol{\Phi}^j_{r-\frac{1}{2},r} \\ \boldsymbol{\Phi}^j_{r,r+\frac{1}{2}} \end{pmatrix} = -\dfrac{1}{\sigma}\dfrac{1}{2\omega^j_r}T^j_r \begin{pmatrix} l_{j,r-1,r}(\overline{E}_{r-\frac{1}{2},r} - E_j) \\ l_{j,r,r+1}(\overline{E}_{r,r+\frac{1}{2}} - E_j), \end{pmatrix}, \end{cases} \tag{3.1}$$

avec

$$T^j_r = \begin{pmatrix} 1 & -\cos(\theta^j_r) \\ -\cos(\theta^j_r) & 1 \end{pmatrix}, \tag{3.2}$$

où ω^r_j est l'aire de la sous-maille associée aux noeuds r et à la maille j. On note $\mathbf{x}_{r+\frac{1}{2}}$ le milieu du segment $[\mathbf{x}_{r-1}, \mathbf{x}_r]$, et θ^j_r l'angle entre $\mathbf{x}_r$, $\mathbf{x}_{r+\frac{1}{2}}$ et $\mathbf{x}_{r-\frac{1}{2}}$. La quantité $l_{j,r-1,r}$ correspond à la longueur associée au segment $[\mathbf{x}_{r-1}, \mathbf{x}_r]$.

Comme pour le schéma TPFA (1.69), le schéma de Breil-Maire estime un gradient dans la direction normale, nommé $\boldsymbol{\Phi}$. Le gradient est d'abord construit dans chaque sous-maille (figure 3.3), avant d'imposer la continuité des flux. C'est la seconde équation du schéma (3.1) qui assure la continuité des flux normaux $\boldsymbol{\Phi}^k_{k-1}$ et $\boldsymbol{\Phi}^k_k$ de chaque côté de l'arête k.

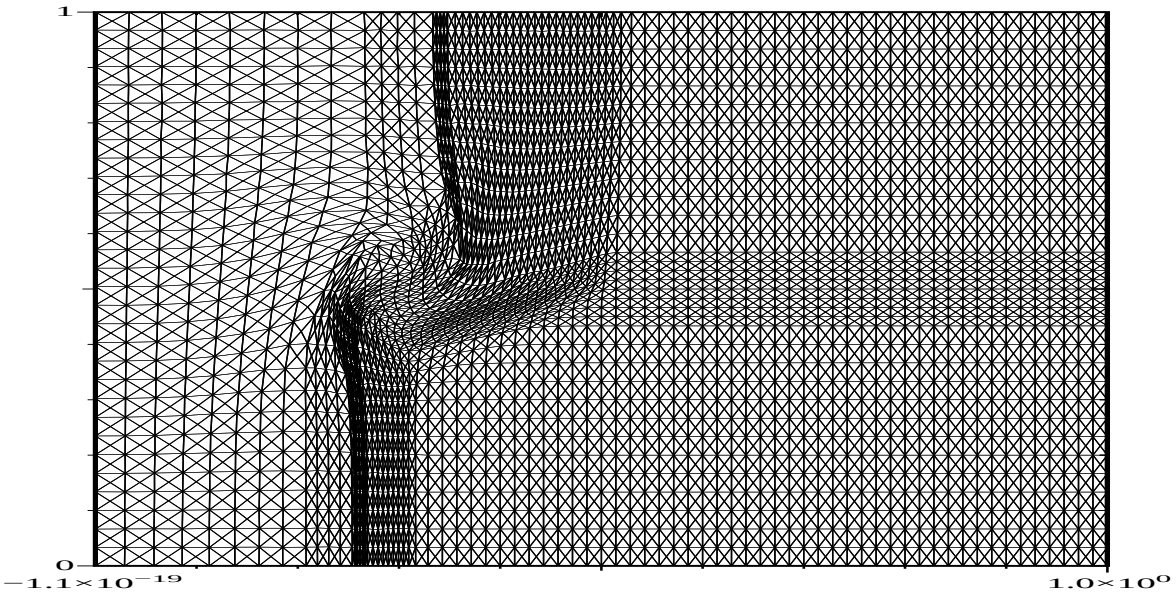


FIGURE 3.2 – Maillage triangulaire, issu d'une simulation hydrodynamique avec une méthode ALE

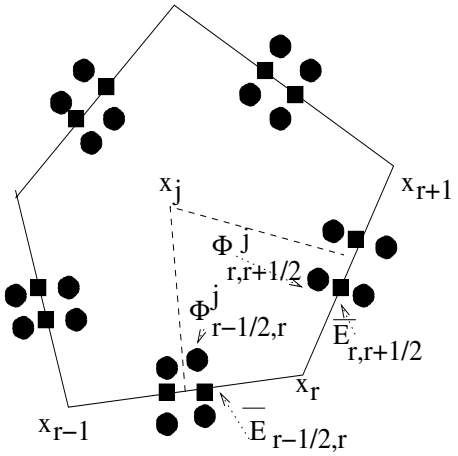


FIGURE 3.3 – Notations

La dernière équation du schéma de Breil-Maire estime les gradients normaux associés aux deux arêtes de la sous-maille associée à la maille j et au noeud r en utilisant la valeur aux centres de la maille E_j et les valeurs aux milieux des arêtes $\overline{E}_{r-\frac{1}{2},r}$ et $\overline{E}_{r,r+\frac{1}{2}}$ (ce qui correspond à A et C sur la figure 3.4). Cependant, la continuité des flux permet d'éliminer ces valeurs aux centres des arêtes afin d'obtenir un schéma où les inconnues sont localisées uniquement aux centres des mailles. La formule de reconstruction du gradient dans les sous-mailles (dernière formule du schéma (3.1)) n'est valable que si celles-ci sont des parallélogrammes. Afin de rendre consistant le schéma de Breil-Maire , il suffit

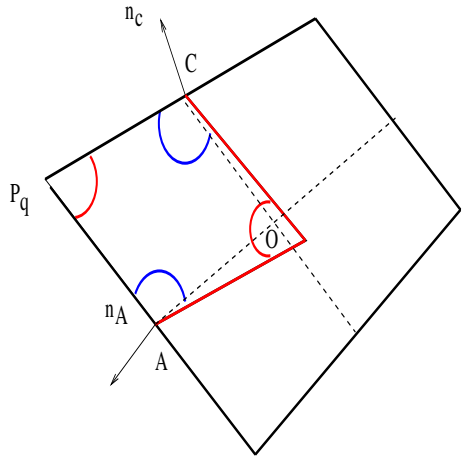


FIGURE 3.4 – Hypothèse du schéma de Breil-Maire

de retrouver la formule du gradient valide indépendamment de la géométrie de la sous-maille. On se place dans le cadre de la sous-maille décrite par la figure (3.4) et on introduit l'identité géométrique suivante

$$\begin{cases} \mathbf{n}_A = \dfrac{(\mathbf{n}_A, \mathbf{OC}^\perp)}{(\mathbf{OA}, \mathbf{OC}^\perp)}\mathbf{OA} + \dfrac{(\mathbf{n}_A, \mathbf{OA}^\perp)}{(\mathbf{OC}, \mathbf{OA}^\perp)}\mathbf{OC}, \\ \\ \mathbf{n}_C = \dfrac{(\mathbf{n}_C, \mathbf{OC}^\perp)}{(\mathbf{OA}, \mathbf{OC}^\perp)}\mathbf{OA} + \dfrac{(\mathbf{n}_C, \mathbf{OA}^\perp)}{(\mathbf{OC}, \mathbf{OA}^\perp)}\mathbf{OC}. \end{cases}$$

À partir de cette formule qui décompose la normale associée à une face en fonction des segments formant la sous-maille associée, on peut obtenir l'approximation suivante du gradient dans la direction normale

$$\begin{cases} \mathbf{\Phi}_k^k = -\dfrac{1}{\sigma}(\nabla E, \mathbf{n}_A) \simeq \dfrac{1}{\sigma}\dfrac{1}{\beta}((\mathbf{n}_A, \mathbf{OC}^\perp)(E(O) - E(A)) + (\mathbf{n}_A, \mathbf{OA}^\perp)(E(O) - E(C)), \\ \\ \mathbf{\Phi}_{k+1}^k = -\dfrac{1}{\sigma}(\nabla E, \mathbf{n}_C) \simeq \dfrac{1}{\sigma}\dfrac{1}{\beta}((\mathbf{n}_C, \mathbf{OC}^\perp)(E(O) - E(A)) + (\mathbf{n}_C, \mathbf{OA}^\perp)(E(O) - E(C)), \end{cases}$$

avec $\mathbf{OA}^\perp = \begin{pmatrix} y_A - y_O \\ -(x_A - x_O) \end{pmatrix}$, $\mathbf{OC}^\perp = \begin{pmatrix} -(y_C - y_O) \\ x_C - x_O \end{pmatrix}$ et $\beta = (\mathbf{OA}, \mathbf{OC}^\perp) = (\mathbf{OC}, \mathbf{OA}^\perp)$.
À partir de là on obtient une estimation du gradient utilisant la valeur associée au centre de la maille et les valeurs associées aux interfaces. On définit une nouvelle estimation du gradient pour le schéma de Breil-Maire.

Définition 3.1.2. *Le schéma de Breil-Maire modifié s'écrit*

$$\begin{cases} \mid \Omega_j \mid \partial_t E_j(t) + \sum_r \frac{1}{2}(l_{j,r-1,r}\mathbf{\Phi}^j_{r-\frac{1}{2},r} + l_{j,r,r+1}\mathbf{\Phi}^j_{r,r+\frac{1}{2}}) = 0, \\ \\ \frac{1}{2}l_k(\mathbf{\Phi}^{k-1}_k + \mathbf{\Phi}^k_k) = \mathbf{0}, \\ \\ \begin{pmatrix} \mathbf{\Phi}^j_{r-\frac{1}{2},r} \\ \mathbf{\Phi}^j_{r,r+\frac{1}{2}} \end{pmatrix} = -\frac{1}{\sigma}\frac{1}{2\omega^j_r}T^j_r \begin{pmatrix} l_{j,r-1,r}(\overline{E}_{r-\frac{1}{2},r} - E_j) \\ l_{j,r,r+1}(\overline{E}_{r,r+\frac{1}{2}} - E_j) \end{pmatrix}, \end{cases} \tag{3.3}$$

avec

$$T^j_r = \begin{pmatrix} 2w^j_r \dfrac{(\mathbf{n}_{jr-1,r}, (\mathbf{x}_{r+\frac{1}{2}} - \mathbf{x}_j)^\perp)}{\beta l_{jr-1,r}} & 2w^j_r \dfrac{(\mathbf{n}_{jr-1,r}, (\mathbf{x}_{r-\frac{1}{2}} - \mathbf{x}_j)^\perp)}{\beta l_{jr,r+1}} \\ 2w^j_r \dfrac{(\mathbf{n}_{jr,r+1}, (\mathbf{x}_{r+\frac{1}{2}} - \mathbf{x}_j)^\perp)}{\beta l_{jr-1,r}} & 2w^j_r \dfrac{(\mathbf{n}_{jr,r+1}, (\mathbf{x}_{r-\frac{1}{2}} - \mathbf{x}_j)^\perp)}{\beta l_{jr,r+1}} \end{pmatrix}, \tag{3.4}$$

et $\beta = ((\mathbf{x}_{r-\frac{1}{2}} - \mathbf{x}_j), (\mathbf{x}_{r+\frac{1}{2}} - \mathbf{x}_j)^\perp) = ((\mathbf{x}_{r-\frac{1}{2}} - \mathbf{x}_j)^\perp, (\mathbf{x}_{r+\frac{1}{2}} - \mathbf{x}_j))$.

Remarque 3.1.1.
- *Les matrices locales obtenues après avoir injecté les définitions des gradients dans la condition de continuité ne sont pas toujours inversibles. Un exemple est introduit dans [MV06].*
- *Les matrices locales* T^j_r *ne sont pas symétriques, ce qui implique la même propriété pour la matrice globale associée au schéma.*
- *Lorsque les sous-mailles sont des parallélogrammes ce schéma modifié est égal au schéma de Breil-Maire.*

D'un certain point de vue le schéma de Breil-Maire consiste à renoncer à la consistance de la formule du gradient dans certains cas afin de récupérer la symétrie du schéma et obtenir des matrices locales toujours inversibles. Avant d'introduire le schéma pour l'équation de la chaleur hyperbolique, on propose une remarque sous forme de proposition.

Proposition 3.1. *Le schéma de Breil-Maire modifié est égal au schéma MPFA-O défini dans [AE06].*

Démonstration. On utilisera les mêmes notations, en effet le schéma MPFA-O [AE06] utilise le même type de discrétisation. Le gradient associé au schéma MPFA-O est défini pour toute maille j de noeud r par

$$\nabla E_j = \frac{\overline{E}_{r,r+\frac{1}{2}} - E_j}{2V^j_r}(\mathbf{x}_{r-\frac{1}{2}} - \mathbf{x}_j)^\perp + \frac{\overline{E}_{r-\frac{1}{2},r} - E_j}{2V^j_r}(\mathbf{x}_{r+\frac{1}{2}} - \mathbf{x}_j)^\perp, \tag{3.5}$$

avec V^j_r l'aire du triangle formé par les noeuds $\mathbf{x}_{r-\frac{1}{2}}$, $\mathbf{x}_{r+\frac{1}{2}}$ et $\mathbf{x}_j$. Pour construire le schéma MPFA on part de ce gradient pour obtenir les gradients normaux associés aux demi-faces (en multipliant ∇E_j par la normale) puis on impose la continuité des gradients normaux. Pour montrer que le schéma de Breil-Maire modifié correspond au schéma MPFA il suffit de vérifier si les gradients normaux introduits avant d'utiliser la condition de compatibilité sont égaux. Pour cela, on développe le troisième terme de (3.3), l'aire w^j_r et la longueur disparaissent. En outre, on remarque que $\beta = 2V^j_r$, par conséquent on a

$$\begin{pmatrix} \mathbf{\Phi}^j_{r-\frac{1}{2},r} \\ \mathbf{\Phi}^j_{r,r+\frac{1}{2}} \end{pmatrix} = \tag{3.6}$$

$$-\frac{1}{\sigma}\begin{pmatrix} \dfrac{(\mathbf{n}_{jr-1,r},(\mathbf{x}_{r+\frac{1}{2}}-\mathbf{x}_j)^\perp)}{2V_r^j}(\overline{E}_{r-\frac{1}{2},r}-E_j)+\dfrac{(\mathbf{n}_{jr-1,r},(\mathbf{x}_{r-\frac{1}{2}}-\mathbf{x}_j)^\perp)}{2V_r^j}(\overline{E}_{r,r+\frac{1}{2}}-E_j), \\ \dfrac{(\mathbf{n}_{jr,r+1},(\mathbf{x}_{r+\frac{1}{2}}-\mathbf{x}_j)^\perp)}{2V_r^j}(\overline{E}_{r-\frac{1}{2},r}-E_j)+\dfrac{(\mathbf{n}_{jr,r+1},(\mathbf{x}_{r-\frac{1}{2}}-\mathbf{x}_j)^\perp)}{2V_r^j}(\overline{E}_{r,r+\frac{1}{2}}-E_j), \end{pmatrix}. \tag{3.7}$$

Il suffit d'écrire le gradient du schéma MPFA-O (3.5) dans les directions normales pour obtenir les formules (3.7) du schéma de Breil-Maire modifié. Par conséquent, les flux des deux schémas sont égaux ainsi que la condition de compatibilité. Les schémas sont identiques. □

Ce calcul montre que les schémas MPFA-O et Breil-Maire sont des variantes d'un même schéma, seule la matrice locale discrétisant le gradient diffère. Le schéma MPFA utilise une matrice qui permet de calculer un gradient consistant, ce qui donne un schéma non symétrique. En outre, la matrice locale issue des conditions de continuité n'est pas toujours inversible (un exemple est donné dans [MV06]). À présent on souhaite construire une famille de schémas « asymptotic preserving » dont deux variantes qui admettent comme limite asymptotique ces deux schémas de diffusion.

3.2 Schéma aux arêtes pour l'équation de la chaleur hyperbolique

Le but de cette section est de construire un schéma pour l'équation de la chaleur hyperbolique utilisant les mêmes principes de construction que les schémas de Breil-Maire et MPFA-O. Une fois ce schéma introduit on le modifie avec la méthode de Jin-Levermore afin d'obtenir un schéma « asymptotic preserving ». On utilise le même principe de construction du schéma de Breil-Maire afin d'obtenir

$$\begin{cases} \mid \Omega_j \mid \partial_t E_j + \dfrac{1}{2\varepsilon}\displaystyle\sum_r (l_{j,r-1,r}\mathbf{\Phi}^j_{r-\frac{1}{2},r} + l_{j,r,r+1}\mathbf{\Phi}^j_{r,r+\frac{1}{2}}) = 0, \\ \mid \Omega_j \mid \partial_t \mathbf{F}_j + \dfrac{1}{2\varepsilon}\displaystyle\sum_r (l_{j,r-1,r}\mathbf{n}_{j,r-1,r}\overline{E}_{r-\frac{1}{2},r} + l_{j,r,r+1}\mathbf{n}_{j,r,r+1}\overline{E}_{r,r+\frac{1}{2}}) = -\dfrac{\sigma}{\varepsilon^2} \mid \Omega_j \mid \mathbf{F}_j, \\ \dfrac{1}{2} l_k(\mathbf{\Phi}^{k-1}_k + \mathbf{\Phi}^k_k) = 0, \\ \overline{E}_{r-\frac{1}{2},r} - E_j + (\mathbf{\Phi}^j_{r-\frac{1}{2},r} - (\mathbf{F}_j, \mathbf{n}_{j,r-1,r})) = 0, \\ \overline{E}_{r,r+\frac{1}{2}} - E_j + (\mathbf{\Phi}^j_{r,r+\frac{1}{2}} - (\mathbf{F}_j, \mathbf{n}_{j,r,r+1})) = 0. \end{cases} \tag{3.8}$$

Les première et seconde équations correspondent à la discrétisation de l'équation de la chaleur hyperbolique. Pour la première équation le principe est le même que pour le schéma de Breil-Maire. On introduit deux flux $\mathbf{\Phi}$ par arêtes correspondant à la quantité $\mathbf{F}$ localisée à l'interface. Pour la seconde équation on introduit $\overline{E}_{r-\frac{1}{2},r}$ et $\overline{E}_{r-\frac{1}{2},r}$ qui contrairement au schéma de diffusion de Breil-Maire, ne sont pas des valeurs fictives aux interfaces mais des flux associés à la seconde équation. La troisième équation correspond à la condition de continuité. Les dernières équations correspondent aux invariants de Riemann de l'équation de la chaleur hyperbolique écrits dans les directions $\mathbf{n}_{jr-1,r}$, $\mathbf{n}_{jr,r+1}$. Ce schéma n'est pas capable de capter la limite de diffusion. Afin d'obtenir cette propriété nous procédons comme pour les schémas aux noeuds en utilisant la méthode de Jin-Levermore. On

part de la formule de gradient discret suivante

$$\begin{pmatrix} \overline{E}_{r-\frac{1}{2},r} - E_j \\ \overline{E}_{r,r+\frac{1}{2}} - E_j \end{pmatrix} = \begin{pmatrix} \dfrac{2\omega_r^j}{l_{j,r-1,r}} & 0 \\ 0 & \dfrac{2\omega_r^j}{l_{j,r,r+1}} \end{pmatrix} (T_r^j)^{-1} \begin{pmatrix} \nabla E^j_{r-\frac{1}{2},r} \\ \nabla E^j_{r,r+\frac{1}{2}} \end{pmatrix}. \tag{3.9}$$

introduite par les schémas de Breil-Maire (3.1) et MPFA-O (3.3). Les quantités $\nabla E^j_{r,r+\frac{1}{2}}$ et $\nabla E^j_{r,r+\frac{1}{2}}$ correspondent aux gradients de E aux centres des deux arêtes associées au noeud r de la maille j. Sachant que $\mathbf{\Phi}^j_{r-\frac{1}{2},r}$ et $\mathbf{\Phi}^j_{r,r+\frac{1}{2}}$ représentent la quantité $\mathbf{F}$ localisée aux centres des deux arêtes associées au noeud r de la maille j, on peut écrire l'équivalent discret de l'état stationnaire $\nabla E = -\frac{\sigma}{\varepsilon}\mathbf{F}$. On obtient $\nabla E^j_{r-\frac{1}{2},r} = -\frac{\sigma}{\varepsilon}\mathbf{\Phi}^j_{r-\frac{1}{2},r}$ et $\nabla E^j_{r,r+\frac{1}{2}} = -\frac{\sigma}{\varepsilon}\mathbf{\Phi}^j_{r,r+\frac{1}{2}}$. Une fois ces relations obtenues, on les injecte dans la définition du gradient, ce qui permet d'obtenir

$$\begin{pmatrix} \overline{E}_{r-\frac{1}{2},r} - E_j \\ \overline{E}_{r,r+\frac{1}{2}} - E_j \end{pmatrix} = - \begin{pmatrix} \dfrac{2\omega_r^j}{l_{jr-1,r}} & 0 \\ 0 & \dfrac{2\omega_r^j}{l_{jr,r+1}} \end{pmatrix} (T_r^j)^{-1} \frac{\sigma}{\varepsilon} \begin{pmatrix} \mathbf{\Phi}^j_{r-\frac{1}{2},r} \\ \mathbf{\Phi}^j_{r,r+\frac{1}{2}} \end{pmatrix}. \tag{3.10}$$

La relation (3.10) peut être interprétée comme une discrétisation consistante de l'état d'équilibre $\nabla E = -\frac{\sigma}{\varepsilon}\mathbf{F}$. À partir de là on souhaite injecter cette relation dans les flux du schéma de transport (3.8). À cette fin, dans les flux de (3.8) on remplace E_j par $E_j + (\overline{E}_{r-\frac{1}{2},r} - E_j)$ (resp $E_j + (\overline{E}_{r,r+\frac{1}{2}} - E_j)$), ce qui donne

$$\begin{cases} \overline{E}_{r-\frac{1}{2},r} - (E_j + (\overline{E}_{r-\frac{1}{2},r} - E_j)) + (\mathbf{\Phi}^j_{r-\frac{1}{2},r} - (\mathbf{F}_j, \mathbf{n}_{j,r-1,r})) = 0, \\ \overline{E}_{r,r+\frac{1}{2}} - (E_j + (\overline{E}_{r,r+\frac{1}{2}} - E_j)) + (\mathbf{\Phi}^j_{r,r+\frac{1}{2}} - (\mathbf{F}_j, \mathbf{n}_{j,r,r+1})) = 0. \end{cases}$$

Pour terminer il suffit de remplacer les quantités $(\overline{E}_{r-\frac{1}{2},r} - E_j)$ (resp $(\overline{E}_{r,r+\frac{1}{2}} - E_j)$) par la relation (3.10) issue de la méthode de Jin-Levermore.

Définition 3.2.1. *Les schémas MPFA-AP et Breil-Maire-AP pour l'équation de la chaleur hyperbolique sont définis par*

$$\begin{cases} \mid \Omega_j \mid \partial_t E_j(t) + \dfrac{1}{2\varepsilon}\displaystyle\sum_r (l_{j,r-1,r}\mathbf{\Phi}^j_{r-\frac{1}{2},r} + l_{j,r,r+1}\mathbf{\Phi}^j_{r,r+\frac{1}{2}}) = 0, \\ \mid \Omega_j \mid \partial_t \mathbf{F}_j(t) + \dfrac{1}{2\varepsilon}\displaystyle\sum_r (l_{j,r-1,r}\overline{E}_{r-\frac{1}{2},r}\mathbf{n}_{j,r-1,r} + l_{j,r,r+1}\overline{E}_{r,r+\frac{1}{2}}\mathbf{n}_{j,r,r+1}) = -\dfrac{\sigma}{\varepsilon^2} \mid \Omega_j \mid \mathbf{F}_j, \\ \dfrac{1}{2} l_k(\mathbf{\Phi}^{k-1}_k + \mathbf{\Phi}^k_k) = \mathbf{0}, \\ \begin{pmatrix} \mathbf{\Phi}^j_{r-\frac{1}{2},r} \\ \mathbf{\Phi}^j_{r,r+\frac{1}{2}} \end{pmatrix} = -\dfrac{1}{2\omega^j_r}\tilde{M}^j_r \begin{pmatrix} l_{j,r-1,r}(\overline{E}_{r-\frac{1}{2},r} - E_j) - l_{j,r-1,r}(\mathbf{F}_j, \mathbf{n}_{j,r-1,r}) \\ l_{j,r,r+1}(\overline{E}_{r,r+\frac{1}{2}} - E_j) - l_{j,r,r+1}(\mathbf{F}_j, \mathbf{n}_{j,r,r+1}) \end{pmatrix}, \end{cases} \tag{3.11}$$

avec

$$\tilde{M}^j_r = 2\omega^j_r \begin{pmatrix} l_{jr-1,r}(1 + (S^j_{r,\epsilon})^{xx}) & l_{jr-1,r}(S^j_{r,\epsilon})^{xy} \\ l_{jr,r+1}(S^j_{r,\epsilon})^{yx} & l_{jr,r+1}(1 + (S^j_{r,\epsilon})^{yy}) \end{pmatrix}^{-1}, \tag{3.12}$$

$$S^j_{r,\epsilon} = \begin{pmatrix} \dfrac{2\omega^j_r}{l_{jr-1,r}} & 0 \\ 0 & \dfrac{2\omega^j_r}{l_{jr,r+1}} \end{pmatrix} \frac{\sigma}{\epsilon}(T^j_r)^{-1}, \tag{3.13}$$

et T^j_r la matrice locale d'estimation du gradient associée aux schémas de Breil-Maire (3.1) ou de MPFA-O (3.3).

Maintenant on vérifie que le schéma proposé admet comme limite de diffusion les schémas de diffusion cible.

Proposition 3.2. *La limite asymptotique du schéma (3.11)-(3.12)-(3.13) est le schéma de Breil-Maire si on choisit comme matrice locale T^j_r (3.2) et le schéma MPFA-O si on choisit comme matrice locale T^j_r (3.4).*

Démonstration. Comme pour les autres schémas on utilise un développement de Hilbert

$$E_j = E^0_j + \varepsilon E^1_j + \varepsilon^2 E^2_j + o(\varepsilon^2),$$

$$E_{r\pm\frac{1}{2},r} = E^0_{r\pm\frac{1}{2},r} + \varepsilon E^1_{r\pm\frac{1}{2},r} + \varepsilon^2 E^2_{r\pm\frac{1}{2},r} + o(\varepsilon^2).$$

Le développement est le même pour toutes les variables. En incorporant le développement précédent dans le schéma on obtient pour la première équation :
le terme proportionnel à $\frac{1}{\varepsilon}$

$$\sum_r \frac{1}{2}(l_{j,r-1,r}\mathbf{\Phi}^{j,0}_{r-\frac{1}{2},r} + l_{j,r,r+1}\mathbf{\Phi}^{j,0}_{r,r+\frac{1}{2}}) = 0, \tag{3.14}$$

le terme proportionnel à $\frac{1}{\varepsilon^0}$

$$|\Omega_j|\partial_t E^0_j + \sum_r \frac{1}{2}(l_{j,r-1,r}\mathbf{\Phi}^{j,1}_{r-\frac{1}{2},r} + l_{j,r,r+1}\mathbf{\Phi}^{j,1}_{r,r+\frac{1}{2}}) = 0. \tag{3.15}$$

On utilise le même type de procédure pour la seconde équation.

$$\frac{1}{\varepsilon^2}: \qquad \mathbf{F}_j^0 = \mathbf{0}, \tag{3.16}$$

$$\frac{1}{\varepsilon^1}: \qquad \sum_r \frac{1}{2\varepsilon}(l_{j,r-1,r}\mathbf{n}_{j,r-1,r}\overline{E}^0_{r-\frac{1}{2},r} + l_{j,r,r+1}\mathbf{n}_{j,r,r+1}\overline{E}^0_{r,r+\frac{1}{2}}) = -\sigma \mid \Omega_j \mid \mathbf{F}_j^1, \tag{3.17}$$

$$\frac{1}{\varepsilon^0}: \qquad |\Omega_j|\partial_t \mathbf{F}_j^0 + \sum_r \frac{1}{2\varepsilon}(l_{j,r-1,r}\mathbf{n}_{j,r-1,r}\overline{E}^1_{r-\frac{1}{2},r} + l_{j,r,r+1}\mathbf{n}_{j,r,r+1}\overline{E}^1_{r,r+\frac{1}{2}}) = -\sigma \mid \Omega_j \mid \mathbf{F}_j^2. \tag{3.18}$$

L'étape suivante consiste à réécrire l'équation sur le flux $\mathbf{\Phi}$ (3.12)-(3.13), ce qui donne

$$\begin{pmatrix} l_{j,r-1,r}(1+(S^j_{r,\epsilon})^{xx}) & l_{j,r-1,r}(S^j_{r,\epsilon})^{xy} \\ l_{j,r,r+1}(S^j_{r,\epsilon})^{yx} & l_{j,r,r+1}(1+(S^j_{r,\epsilon})^{yy}) \end{pmatrix} \begin{pmatrix} \mathbf{\Phi}^j_{r-\frac{1}{2},r} \\ \mathbf{\Phi}^j_{r,r+\frac{1}{2}} \end{pmatrix}$$

$$= -\begin{pmatrix} l_{j,r-1,r}(\overline{E}_{r-\frac{1}{2},r} - E_j) - l_{j,r-1,r}(\mathbf{F}_j, \mathbf{n}_{j,r-1,r}) \\ l_{j,r,r+1}(\overline{E}_{r,r+\frac{1}{2}} - E_j) - l_{j,r,r+1}(\mathbf{F}_j, \mathbf{n}_{j,r,r+1}) \end{pmatrix}, \tag{3.19}$$

Pour finir on incorpore le résultat du développement de Hilbert dans l'équation (3.19)

$$\frac{1}{\varepsilon^1}: \qquad 2w_r^j\sigma T_r^{j,-1}\begin{pmatrix} \mathbf{\Phi}^{j,0}_{r-\frac{1}{2},r} \\ \mathbf{\Phi}^{j,0}_{r,r+\frac{1}{2}} \end{pmatrix} = \mathbf{0}, \tag{3.20}$$

$$\frac{1}{\varepsilon^0}: \qquad \begin{pmatrix} l_{j,r-1,r} & 0 \\ 0 & l_{j,r,r+1} \end{pmatrix}\begin{pmatrix} \mathbf{\Phi}^{j,0}_{r-\frac{1}{2},r} \\ \mathbf{\Phi}^{j,0}_{r,r+\frac{1}{2}} \end{pmatrix} + 2w_r^j\sigma T_r^{j,-1}\begin{pmatrix} \mathbf{\Phi}^{j,1}_{r-\frac{1}{2},r} \\ \mathbf{\Phi}^{j,1}_{r,r+\frac{1}{2}} \end{pmatrix}$$

$$= \begin{pmatrix} l_{j,r-1,r}(\overline{E}^0_{r-\frac{1}{2},r} - E_j^0) - l_{j,r-1,r}(\mathbf{F}_j^0, \mathbf{n}_{j,r-1,r}) \\ l_{j,r,r+1}(\overline{E}^0_{r,r+\frac{1}{2}} - E_j^0) - l_{j,r,r+1}(\mathbf{F}_j^0, \mathbf{n}_{j,r,r+1}) \end{pmatrix}. \tag{3.21}$$

L'équation (3.20) donne

$$\begin{pmatrix} \mathbf{\Phi}^{j,0}_{r-\frac{1}{2},r} \\ \mathbf{\Phi}^{j,0}_{r,r+\frac{1}{2}} \end{pmatrix} = \mathbf{0}. \tag{3.22}$$

Par conséquent, en utilisant (3.22), (3.16), (3.15) et (3.21) on obtient

$$\begin{cases} \mid \Omega_j \mid \partial_t E_j^0 + \dfrac{1}{2}\displaystyle\sum_r (l_{j,r-1,r}(\mathbf{\Phi}^j_{r-\frac{1}{2},r})^1 + l_{j,r,r+1}(\mathbf{\Phi}^j_{r,r+\frac{1}{2}})^1) = 0, \\ \begin{pmatrix} (\mathbf{\Phi}^j_{r-\frac{1}{2},r})^1 \\ (\mathbf{\Phi}^j_{r,r+\frac{1}{2}})^1 \end{pmatrix} = -\dfrac{1}{\sigma}\dfrac{1}{2\omega_r^j}T_r^j \begin{pmatrix} l_{j,r-1,r}((\overline{E}_{r-\frac{1}{2},r})^0 - (E_j)^0) \\ l_{j,r,r+1}((\overline{E}_{r,r+\frac{1}{2}})^0 - (E_j)^0) \end{pmatrix}. \end{cases}$$

Enfin il suffit d'opérer le développement dans la condition de continuité, d'incorporer (3.22) et de simplifier par ε pour obtenir

$$\frac{1}{2}L_k(\mathbf{\Phi}_k^{k-1,1} + \mathbf{\Phi}_k^{k,1}) = \mathbf{0}.$$

Ce résultat clôt la démonstration. □

3.3 Résultats numériques

Les résultats numériques ont été réalisés au CEMRACS dans un code différent de celui utilisé pour le reste du mémoire. On utilise d'autres cas tests que ceux proposés dans le chapitre précédent.

Test 1

Il s'agit d'un cas test analytique et stationnaire introduit dans [BM06].
Soit l'équation $-\triangle E_{stat} = Q$ avec le domaine $x \in [0,1]$, les conditions limites $\partial_x E_{stat} = 0$ et le terme source $Q = \left(\frac{cos(1)-1}{sin(1)}\right) cos(x) + sin(x)$. La solution ne dépend que de la variable x et est définie à une constante additive près. La solution stationnaire est d'intégrale nulle et elle est définie par

$$E_{stat}(x) = -x + \left(\frac{cos(1)-1}{sin(1)}\right) cos(x) + sin(x) + 0.5.$$

Le cas test consiste à résoudre l'équation de la chaleur en temps long

$$\frac{\partial E}{\partial t} - \triangle E = Q, \tag{3.23}$$

avec la donnée initiale $E(t=0,\mathbf{x}) = 0$ et la condition limite $(\nabla E, \mathbf{n}) = 0$ sur ∂D. En temps long la solution coïncide avec la solution stationnaire E_{stat} puisque l'intégrale de Q est nulle. Afin de valider les schémas « asymptotic preserving » on résout aussi le modèle

$$\begin{cases} \partial_t E_\varepsilon + \dfrac{1}{\varepsilon} \operatorname{div} \mathbf{F}_\varepsilon = Q, \\ \\ \partial_t \mathbf{F}_\varepsilon + \dfrac{1}{\varepsilon} \nabla E_\varepsilon = -\dfrac{1}{\varepsilon^2} \mathbf{F}_\varepsilon, \end{cases}$$

avec la condition limite $(\mathbf{F}_\varepsilon, \mathbf{n}) = 0$ sur ∂D et les données initiales $E_\varepsilon(t=0,\mathbf{x}) = 0$ et $\mathbf{F}_\varepsilon(t=0,\mathbf{x}) = \mathbf{0}$. Lorsque ε tend vers zéro, la solution E_ε converge vers E la solution de (3.23). On souhaite tester les schémas limites (les schémas MPFA-O et Breil-Maire) et les schémas « asymptotic preserving » associés en régime de diffusion. On compare les résultats sur différents maillages donnés par les schémas de diffusion et les schémas « asymptotic preserving » avec $\sigma = 1$ et ε petit.
On constate sur la gauche de la figure (3.5) que tous les schémas de diffusion (2.13), (3.1)-(3.2) et (3.3)-(3.4) convergent à l'ordre 2. Par contre, sur maillage aléatoire le schéma de Breil-Maire ne converge pas comme cela avait été observé dans [BM06], en effet les sous-mailles n'étant plus forcément des parallélogrammes le gradient calculé n'est plus consistant. Les schémas MPFA-O et de diffusion aux noeuds (noté GLACE ici) sont eux d'ordre deux. Sur la même figure on a tracé les résultats donnés par les schémas JL-(a) (2.11)-(2.12), P_1 Breil-Maire et P_1-MPFA pour $\varepsilon = 0.001$. Les ordres de convergence associés à ces schémas sont les mêmes que ceux obtenus pour les schémas de diffusion cibles.

Les méthodes hydrodynamiques Lagrangiennes ou les méthodes ALE actuelles utilisent les maillages polygonaux, par conséquent il est intéressant d'observer comment les schémas se comportent sur des maillages polygonaux issus de simulations hydrodynamiques. On propose de comparer la solution du cas test précédent sur le maillage donné par la figure (3.1), contenant quelques mailles polygonales issues d'un remaillage. Les résultats obtenus sont satisfaisants (fig 3.6), en effet les schémas arrivent à conserver la forme 1D de la solution et les erreurs absolues en norme L^1 sont respectivement de 3.07952×10^{-5} pour MPFA et 3.63853×10^{-5} pour le schéma de diffusion aux noeuds. Le schéma de Breil-Maire admet une erreur plus grande 6.3×10^{-4}, qui est probablement liée à de la non consistance du schéma dans la zone déformée. On compare aussi les schémas MPFA-O et Breil-Maire sur un

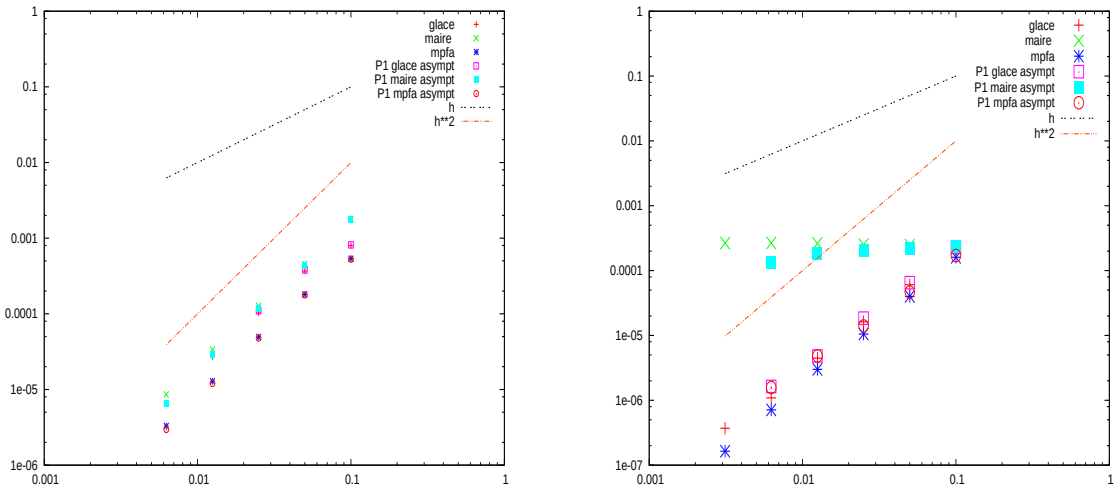


FIGURE 3.5 – Cas test de diffusion : convergence sur le maillage de Kershaw à gauche et sur maillage aléatoire à droite

maillage triangulaire issu de l'hydrodynamique (3.2). Les erreurs numériques associées aux schémas MPFA-O et Breil-Maire sont respectivement 2.7217×10^{-6} et 6.15042×10^{-6}. On trace la solution donnée par le schéma MPFA sur la figure (3.7).

Test 2

Dans un second cas test on étudie la solution fondamentale du modèle P_1 comme lors du chapitre précédent. Le domaine est défini par $[0,4] \times [0,4]$, avec 800 mailles par direction. La condition initiale est donnée par $E(t=0,\mathbf{x}) = \delta_{2,2}$ et $\mathbf{F}(t=0,\mathbf{x}) = \mathbf{0}$, de plus $\sigma = 1$ et $\varepsilon = 1$. La condition limite est donnée par $(\mathbf{F},\mathbf{n}) = 0$. La solution est composée d'une fonction de Dirac qui se propage à la vitesse 1 et d'une fonction régulière entre la position initiale de la fonction de Dirac et la fonction de Dirac qui se propage à la vitesse 1. En temps court la solution régulière est négative et devient positive en temps long. Sur la figure (3.8), on trace la solution donnée par le schéma P_1-MPFA et la solution exacte calculée à l'aide d'un logiciel de calcul formel. On remarque que bien que la singularité soit approximée, on retrouve la solution. Sur la figure (3.9) on trace la solution donnée par le schéma JL-(a) et le schéma P_1-MPFA sur maillage aléatoire. On remarque que contrairement au schéma JL-(a), le schéma P_1-MPFA n'admet pas de modes parasites. Par ailleurs, on retrouve le comportement de la solution exacte dans les résultats donnés par les deux schémas.

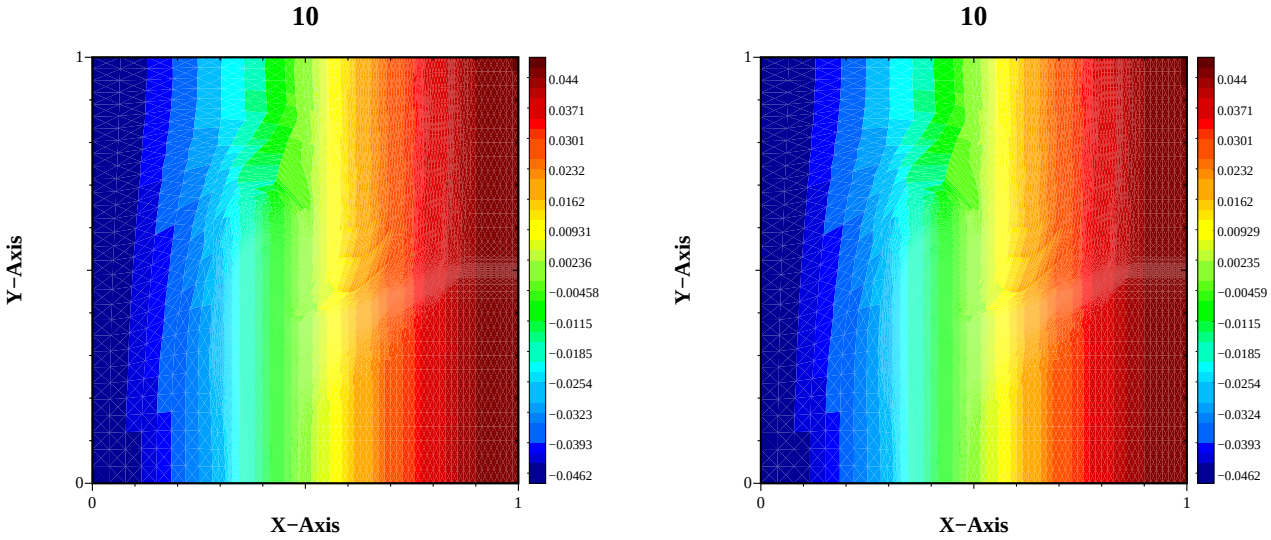


FIGURE 3.6 – Solution sur le maillage hybride polygones-quadrangles (Fig. 3.1) pour GLACE à gauche, pour MPFA à droite.

3.4 Synthèse

Dans ce chapitre nous avons proposé de construire des schémas « asymptotic preserving » ayant une limite de diffusion valable sur maillages non structurés. Les schémas ainsi construits ont l'avantage, contrairement aux schémas aux noeuds, de ne pas admettre des modes parasites. En outre, cette étude montre que l'on peut obtenir un schéma « asymptotic preserving » à partir d'un schéma de diffusion cible, en construisant un schéma de transport compatible géométriquement couplé avec la méthode de Jin-Levermore. Il suffit de coupler la discrétisation du gradient donnée par le schéma de diffusion cible avec la discrétisation de l'état stationnaire $\nabla E = -\frac{\sigma}{\varepsilon}\mathbf{F}$, puis d'injecter les relations obtenues dans les invariants de Riemann discret. On suppose que la méthode pourrait être utilisée avec d'autres schémas de diffusion comme les schémas Hybride ou SUSHI [EGH09]. En guise de conclusion on propose de rappeler les propriétés des schémas « asymptotic preserving » introduits dans ce chapitre et le précédent. À la suite de quoi, un choix se dégagera.

Schéma P_1-Breil-Maire
+ Schéma d'ordre 1 en régime de transport.
+ Schéma limite d'ordre 2.
+ Schéma limite valable sur un certain nombre de maillages non structurés.
+ Matrice locale toujours inversible.

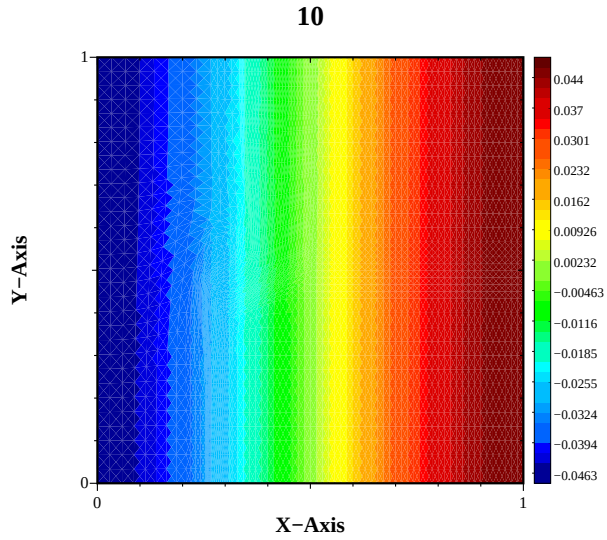


FIGURE 3.7 – Solution sur le maillage triangulaire (Fig. 3.2) pour le schéma MPFA.

+ Schéma limite symétrique.
+ Schéma TPFA sur maillage Cartésien.
+ Linéaire.
+ Stencil à 9 points sur maillage quadrangulaire.
- La condition CFL des schémas explicite et semi-implicite dépend de ε.
- Non consistant donc non convergent sur maillage aléatoire en limite de diffusion.
- Pas d'extensible naturellement en 3D.
- Pas de résultats de convergence pour le schéma quand $\varepsilon = O(1)$.
- Schéma non positif.

Schéma P_1-MPFA O
+ Schéma d'ordre 1 en régime de transport.
+ Schéma limite d'ordre 2.
+ Schéma limite valable sur maillages non structurés.
+ Schéma TPFA sur maillage Cartésien.
+ Linéaire.
+ Stencil à 9 points sur maillage quadrangulaire.
- La condition CFL des schémas explicite et semi-implicite dépend de ε.
- Non symétrique.
- Matrice locale pas toujours inversible.

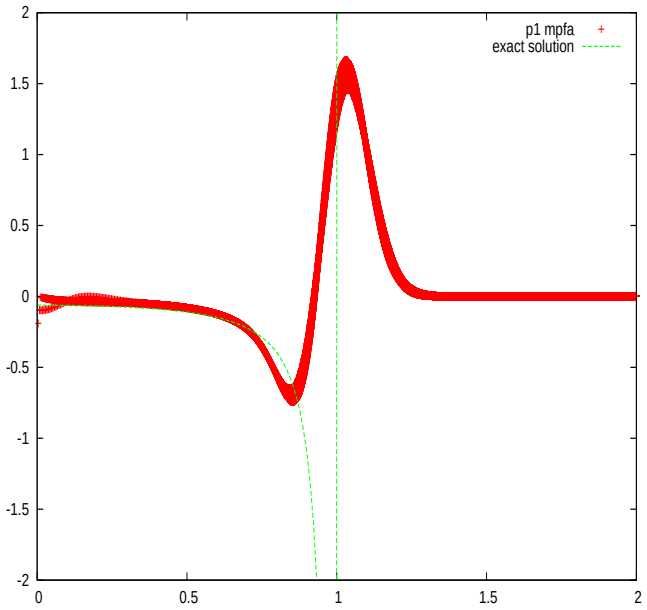


FIGURE 3.8 – Comparaison de la solution donnée par le schéma MPFA avec la solution exacte au temps $t = 1$.

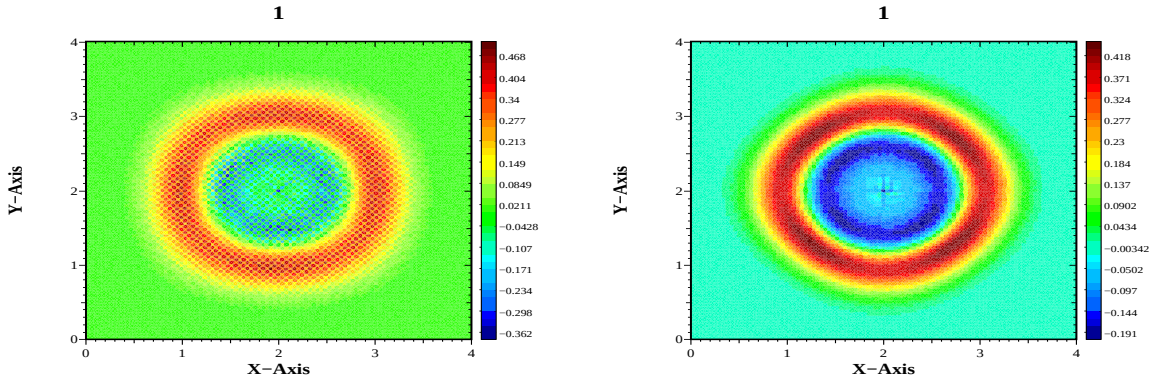


FIGURE 3.9 – Solution donnée par le schéma JL-(a) à gauche, solution donnée par le schéma MPFA à droite.

- Pas d'extensible naturellement en 3D.
- Schéma non positif.

Schémas JL-(a) et JL-(b)
+ Schéma d'ordre 1 en régime de transport.
+ Schéma limite d'ordre 2.
+ Schéma limite valable sur maillages non structurés.
+ Extensibilité naturelle en 3D (par l'introduction des quantités $\mathbf{C}_{jr}$ dans [CDDL09]).
+ Linéaire.
+ Stencil à 9 points sur maillage quadrangulaire.
+ Extensibilité naturelle à des maillages d'ordre 2. Même discrétisation géométrique que l'hydrodynamique.
+ La condition CFL associée au schéma semi-implicite JL-(b) est indépendante de ε.
- Mode parasite sur maillage Cartésien.
- Non symétrique.
- Matrice locale pas toujours inversible.
- Pas de principe du maximum en limite de diffusion.

Le schéma P_1 MPFA-O paraît donc être le plus complet, puisqu'il converge sur tous les maillages étudiés et n'exhibe pas de modes parasites. Cependant, le schéma JL-(b) semi-implicite admet une CFL indépendamment de ε et paraît plus naturel à étendre en 3D en suivant la démarche proposée dans [CDDL09], ce qui en fait une alternative intéressante.

Chapitre 4

Étude des systèmes de Friedrichs avec termes sources raides

Dans le premier chapitre, on a rappelé que les équations de transport, après discrétisation de type P_n ou S_n sont approchées par des grands systèmes hyperboliques linéaires qui admettent une limite de diffusion. En conséquence, il est essentiel d'étendre les schémas « asymptotic preserving » décrits précédemment à ces modèles linéaires afin d'approcher au mieux les équations de transport sur des maillages non structurés dans tous les régimes. Dans ce chapitre on cherche donc à discrétiser des systèmes hyperboliques symétriques de Friedrichs avec des termes sources raides (Les approximations P_n ou S_n entrent dans ce formalisme). A cette fin on propose une méthode de décomposition proche de la méthode micro-macro [LM07] pour une classe de systèmes de Friedrichs. L'objectif initial étant de résoudre des problèmes de transport, les exemples sont issus de l'ensemble des modèles simplifiés utilisés pour approcher les équations (1.3)-(1.6).

4.1 Étude théorique des systèmes de Friedrichs avec termes sources raides

On considère des systèmes linéaires généraux de type Friedrichs avec des termes sources raides. On introduira la limite asymptotique ainsi que les hypothèses nécessaires à l'obtention d'une équation limite de diffusion valide.

4.1.1 Systèmes de Friedrichs : définitions

Définissons le cadre nécessaire. Soit $D \subset \mathbb{R}^2$ est un domaine spatial régulier. La discrétisation des systèmes de Friedrichs sans termes sources raides et avec problèmes aux limites est traitée dans [EG06I]-[EG06II] pour des problèmes stationnaires (méthode de Galerkin discontinue) ou dans [CVV00] pour les problèmes instationnaires. On définit les espaces fonctionnels :

$$L^2(D) = \left\{ \mathbf{U} : D \to \mathbb{R}^n,\, ||\mathbf{U}||_{L^2(D)} = \left(\int_D (\mathbf{U}, \mathbf{U}) \right)^{\frac{1}{2}} dxdy < \infty \right\}$$

$$H^p(D) = \left\{ \mathbf{U} : D \to \mathbb{R}^n, ||\mathbf{U}||_{H^p(D)} = \sum_{a,b}^{a+b \leq p} ||\partial_{x^a,y^b}\mathbf{U}||_{L^2(D)} < \infty \right\}$$

$$X^p(D) = L^\infty([0,T] : H^p(D)) = \left\{ \exists C \text{ tel que } ||\mathbf{U}(t)||_{H^p(D)} < C \quad \forall t \in [0,T]\,, \quad ||\mathbf{U}||_{X^p(D)} = \operatorname{Inf} C \right\}$$

Le coefficient de relaxation σ est supposé strictement positif. Le cadre étant posé, on définit les systèmes de Friedrichs avec des termes sources raides.

Définition 4.1.1. *Les systèmes de Friedrichs avec termes sources raides, que nous considérons, sont de la forme :*

$$\partial_t \mathbf{U} + \frac{1}{\varepsilon} A_1 \partial_x \mathbf{U} + \frac{1}{\varepsilon} A_2 \partial_y \mathbf{U} = -\frac{\sigma}{\varepsilon^2} R\mathbf{U}, \tag{4.1}$$

avec $\mathbf{U}$ *une fonction de* $D \times \mathbb{R}^+$ *dans* $\mathbb{R}^n$, A_1, A_2, R *des matrices carrées symétriques à coefficients constants réels. On suppose* R *positive et non inversible (*$\operatorname{Ker} R \neq \{0\}$*).*

L'hypothèse de non inversibilité de la matrice R est essentielle à l'existence d'un régime limite non trivial. Les modèles simplifiés issus du transport (introduits dans le chapitre 1), les modèles ondulatoires (équation des ondes avec amortissement, équations de Maxwell) sont des systèmes de Friedrichs. Avant d'étudier le régime limite, on rappelle des résultats de stabilité pour les systèmes (4.1). Puisque chaque matrice est symétrique et réelle, la matrice $A_1 n_x + A_2 n_y$ avec $\mathbf{n} = (n_x, n_y) \in \mathbb{R}^2$ est diagonalisable à valeurs propres réelles et le système est donc hyperbolique.

Proposition 4.1. *Soit* $D = [0,1]^2$. *Ces systèmes munis de condition limite périodique sont stables dans la norme associée à l'espace* $H^p(D)$.

Démonstration. On commence par montrer la stabilité pour la norme de $L^2(D)$ en introduisant la dérivée de la norme $L^2(D)$:

$$\frac{1}{2}\frac{d}{dt}||\mathbf{U}||^2_{L^2(D)} = \int_D (\mathbf{U}, \partial_t \mathbf{U}) = -\frac{1}{\varepsilon}\int_D (A_1 \partial_x \mathbf{U} + A_2 \partial_y \mathbf{U} + \frac{\sigma}{\varepsilon} R\mathbf{U}, \mathbf{U}) dxdy.$$

La première intégrale s'écrit

$$\int_D (A_1 \partial_x \mathbf{U}, \mathbf{U}) = \sum_{i,j} \int_D (A_{1,i,j} \partial_x U_i U_j) dxdy. \tag{4.2}$$

Par intégration par parties on obtient

$$\int_D (A_1 \partial_x \mathbf{U}, \mathbf{U}) = -\sum_{i,j} \int_D \partial_x (A_{1,i,j} U_i U_j) dxdy - \sum_{i,j} \int_D (A_{1,i,j} U_i \partial_x U_j) dxdy. \tag{4.3}$$

Puisque le domaine est périodique $\sum_{i,j} \int_D \partial_x (A_{1,i,j} U_i U_j) dxdy = 0$ Après avoir sommé (4.2) et (4.3), on obtient $2\int_D (A_1 \partial_x \mathbf{U}, \mathbf{U}) dxdy = 0$ car la matrice A_1 est symétrique. En procédant de la même manière pour la matrice A_2, on constate qu'il ne reste plus que l'intégrale associée à la matrice R :

$$\frac{1}{2}\frac{d}{dt}||\mathbf{U}||^2_{L^2(D)} = -\int_D \frac{\sigma}{\varepsilon^2} (R\mathbf{U}, \mathbf{U}) dxdy.$$

La matrice R étant positive, la norme $L^2(D)$ décroît. Le système est donc L^2-stable. La fonction $\mathbf{V} = \partial_{x^a, y^b} \mathbf{U}$, avec $a + b \leq \alpha$ est aussi solution du modèle (4.1) et le modèle est donc stable en norme $H^p(D)$. □

4.1.2 Limite de diffusion

Dans cette sous-section, on se propose de déterminer la limite asymptotique d'un système de Friedrichs (4.1). On commencera par obtenir, d'un point de vue formel, la limite de diffusion associée aux systèmes de Friedrichs ainsi que les conditions nécessaires ou suffisantes à l'existence de cette limite asymptotique. Puis on déduira une preuve de convergence rigoureuse et des estimations d'erreurs entre les systèmes de Friedrichs et les limites asymptotiques. On suppose que σ est constant et strictement positif. On commence par introduire une hypothèse de structure. Soit $(\mathbf{E}_1, ..., \mathbf{E}_n)$ la base orthonormée des vecteurs propres de la matrice R.

Hypothèse (H_1) : *Soit* $(\mathbf{E}_1, ...\mathbf{E}_p)$ *les vecteurs propres de* R *formant la base du noyau de* R. *On suppose qu'il existe deux vecteurs propres particuliers* $\mathbf{E}_{i_1}$, $\mathbf{E}_{i_2}$ *linéairement indépendants associés à deux valeurs propres non nulles* $\lambda_{i_1} > 0$, $\lambda_{i_2} > 0$. *On suppose en outre que ces deux vecteurs propres satisfont l'hypothèse de structure*

$$\begin{cases} A_1\mathbf{E}_i = \gamma_i^1\mathbf{E}_{i_1} & \forall i \in \{1...p\}, \\ A_2\mathbf{E}_i = \gamma_i^2\mathbf{E}_{i_2} & \forall i \in \{1...p\}. \end{cases}$$

Remarque 4.1.1. *L'hypothèse* (H_1) *n'est pas nécessaire. Elle peut être relaxée par l'hypothèse* $A_k\mathbf{E}_i \in \operatorname{Ker} R^\perp$ *pour* $i \in \{1..p\}$. *Cependant l'expression de l'équation de diffusion devient, de ce fait, plus compliquée.*

On montrera par la suite que les modèles simplifiés en transport linéaire comme les modèles S_n et P_n satisfont cette hypothèse de structure. On peut prendre comme exemple l'équation de la chaleur hyperbolique (1.62). Dans ce cas, $p = 1$, $A_1\mathbf{E}_1 = \mathbf{E}_2$ et $A_2\mathbf{E}_1 = \mathbf{E}_3$ où $\mathbf{E}_1$ est le vecteur propre associé à la valeur propre nulle et $\mathbf{E}_2$, $\mathbf{E}_3$ des vecteurs propres associés à la valeur propre 1 de la matrice R du terme source.

Proposition 4.2. *Sous la condition de structure* (H_1), *le système (4.1) admet la limite de diffusion formelle*

$$\partial_t\mathbf{V} - \frac{1}{\lambda_{i_1}\sigma}K_1\partial_{xx}\mathbf{V} - \frac{1}{\lambda_{i_2}\sigma}K_2\partial_{yy}\mathbf{V} = \mathbf{0}, \tag{4.4}$$

avec $\mathbf{V} = ((\mathbf{U}, \mathbf{E}_1), ..., (\mathbf{U}, \mathbf{E}_p))$, $K_1 = \gamma^1 \otimes \gamma^1$, $K_2 = \gamma^2 \otimes \gamma^2$, *des matrices symétriques positives où les vecteurs* $\gamma^k = (\gamma_1^k, ..., \gamma_p^k)$ *sont définis par l'hypothèse* (H_1).

Démonstration. En utilisant un développement de Hilbert $\mathbf{U} = \mathbf{U}_0 + \varepsilon\mathbf{U}_1 + \varepsilon^2\mathbf{U}_2 + o(\varepsilon^2)$ dans (4.1), on obtient la hiérarchie d'équations suivantes :

$$\frac{1}{\varepsilon^2} : \qquad R\mathbf{U}_0 = \mathbf{0}, \tag{4.5}$$

$$\frac{1}{\varepsilon^1} : \qquad A_1\partial_x\mathbf{U}_0 + A_2\partial_y\mathbf{U}_0 = -\sigma R\mathbf{U}_1, \tag{4.6}$$

$$\frac{1}{\varepsilon^0} : \qquad \partial_t\mathbf{U}_0 + A_1\partial_x\mathbf{U}_1 + A_2\partial_y\mathbf{U}_1 = -\sigma R\mathbf{U}_2. \tag{4.7}$$

La première équation (4.5) montre que $\mathbf{U}_0$ appartient au noyau de R. Par conséquent,

$$\mathbf{U}_0 = \sum_{j=1}^{p}(\mathbf{U}, \mathbf{E}_j)\mathbf{E}_j. \tag{4.8}$$

La matrice R étant non inversible, l'équation (4.6) implique que $\mathbf{U}_1$ existe à un élément du noyau près sous la condition de compatibilité suivante

$$A_1\partial_x\mathbf{U}_0 + A_2\partial_y\mathbf{U}_0 \in (\mathrm{Ker}\, R)^\perp.$$

L'hypothèse (H_1) et l'équation (4.8) indiquent que les termes d'ordre 0 s'écrivent sous la forme

$$A_1\partial_x\mathbf{U}_0 = \left(\sum_j^p \partial_x(\mathbf{U}_0, \mathbf{E}_j)\gamma_j^1\right)\mathbf{E}_{i_1}, \quad A_2\partial_y\mathbf{U}_0 = \left(\sum_j^p \partial_y(\mathbf{U}_0, \mathbf{E}_j)\gamma_j^2\right)\mathbf{E}_{i_2},$$

avec γ_j^1, γ_j^2, définis par l'hypothèse de structure (H_1). Par définition des vecteurs propres et par linéarité on déduit la relation

$$R\left(\frac{A_1\partial_x\mathbf{U}_0}{\lambda_{i_1}} + \frac{A_2\partial_y\mathbf{U}_0}{\lambda_{i_2}}\right) = A_1\partial_x\mathbf{U}_0 + A_2\partial_y\mathbf{U}_0,$$

ce qui permet d'exprimer $\mathbf{U}_1$ sous la forme suivante

$$\mathbf{U}_1 = -\frac{1}{\sigma}\left(\frac{A_1\partial_x\mathbf{U}_0}{\lambda_{i_1}} + \frac{A_2\partial_y\mathbf{U}_0}{\lambda_{i_2}}\right) + \mathbf{z}, \tag{4.9}$$

avec $\mathbf{z} \in \mathrm{Ker}\, R$. En utilisant la relation (4.7), on montre de même que $\mathbf{U}_2$ existe à un élément du noyau près sous la condition de compatibilité suivante

$$\partial_t\mathbf{U}_0 + A_1\partial_x\mathbf{U}_1 + A_2\partial_y\mathbf{U}_1 \in (\mathrm{Ker}\, R)^\perp.$$

Puisque $\mathrm{Ker}(R) = \mathrm{Vect}(\mathbf{E}_1, ..., \mathbf{E}_p)$, la condition de compatibilité se réécrit sous la forme de p équations

$$\partial_t(\mathbf{U}_0, \mathbf{E}_i) + \partial_x(A_1\mathbf{U}_1, \mathbf{E}_i) + \partial_y(A_2\mathbf{U}_1, \mathbf{E}_i) = \mathbf{0} \quad i \in \{1..p\}\,. \tag{4.10}$$

On injecte la relation (4.9) sur $\mathbf{U}_1$ dans les p équations (4.10) afin d'obtenir les équations de diffusion

$$\begin{cases} \partial_t(\mathbf{U}, \mathbf{E}_i) - \dfrac{1}{\lambda_1\sigma}\partial_{xx}\left(A_1\mathbf{U}, A_1\mathbf{E}_i\right) - \dfrac{1}{\lambda_2\sigma}\partial_{yy}\left(A_2\mathbf{U}, A_2\mathbf{E}_i\right) \\ \dfrac{1}{\lambda_2\sigma}\partial_{xy}(A_1\mathbf{U}, A_2\mathbf{E}_i) - \dfrac{1}{\lambda_1\sigma}\partial_{yx}(A_2\mathbf{U}, A_1\mathbf{E}_i) + \mathbf{N} = \mathbf{0}, \text{ pour } i \in \{1..p\} \end{cases} \tag{4.11}$$

où le terme nommé $\mathbf{N}$, dépendant de $\mathbf{z}$ est donné par

$$\mathbf{N} = \partial_x(A_1\mathbf{z}, \mathbf{E}_i) + \partial_y(A_2\mathbf{z}, \mathbf{E}_i) = \partial_x(\mathbf{z}, A_1\mathbf{E}_i) + \partial_y(\mathbf{z}, A_2\mathbf{E}_i). \tag{4.12}$$

L'hypothèse (H_1) et l'orthogonalité de la base des vecteurs propres de R montrent que les termes $A_1\mathbf{E}_i$, $A_2\mathbf{E}_i$ sont orthogonaux à $\mathbf{z}$. Le terme résiduel $\mathbf{N}$ (4.12) est donc nul. Maintenant on propose de simplifier l'équation limite (4.11). On commence par étudier les termes croisés $(A_1\mathbf{U}_0, A_2\mathbf{E}_i)$ et $(A_2\mathbf{U}_0, A_1\mathbf{E}_i)$.

$$(A_1\mathbf{U}_0, A_2\mathbf{E}_i) = \left(A_1\left(\sum_{j=1}^p(\mathbf{U}_0, \mathbf{E}_j)\mathbf{E}_j\right), A_2\mathbf{E}_i\right) = \left(\left(\sum_{j=1}^p \gamma_j^1(\mathbf{U}_0, \mathbf{E}_j)\right)\mathbf{E}_{i_1}, \gamma_i^2\mathbf{E}_{i_2}\right) = 0, \tag{4.13}$$

$$(A_2\mathbf{U}_0, A_1\mathbf{E}_i) = \left(A_2\left(\sum_{j=1}^p(\mathbf{U}_0, \mathbf{E}_j)\mathbf{E}_j\right), A_1\mathbf{E}_i\right) = \left(\left(\sum_{j=1}^p \gamma_j^2(\mathbf{U}_0, \mathbf{E}_j)\right)\mathbf{E}_{i_2}, \gamma_i^1\mathbf{E}_{i_1}\right) = 0. \tag{4.14}$$

On remarque donc que l'hypothèse de structure et la symétrie des matrices permettent de montrer que les termes croisés sont nuls. Pour les autres termes on obtient

$$(A_k\mathbf{U}, A_k\mathbf{E}_i) = \sum_j (\mathbf{U}_0, \mathbf{E}_j)(A_k\mathbf{E}_j, A_k\mathbf{E}_i) = \sum_j (\mathbf{U}_0, \mathbf{E}_j)\gamma_j^k\gamma_i^k. \tag{4.15}$$

Par conséquent les équations (4.11) avec $\mathbf{U}_0 = \mathbf{U}$ sont équivalentes aux équations (4.1) avec $K_1 = \gamma^1 \otimes \gamma^1$, $K_2 = \gamma^2 \otimes \gamma^2$. Ces matrices sont symétriques par définition. De plus,

$$(\mathbf{X}, K_k\mathbf{X}) = (\gamma^k, \mathbf{X})^2 \geq 0 \text{ avec } \mathbf{X} \in \mathbb{R}^n,$$

donc la matrice K_k est positive. □

Remarque 4.1.2.
- *les matrices K_i étant positives, le système est dissipative.*
- *La dimension de l'équation de diffusion est égale à la multiplicité de la valeur propre nulle de la matrice R.*
- *L'hypothèse (H_1) est suffisante mais non nécessaire. L'hypothèse $A_k\mathbf{E}_i \in KerR^\perp$ pour $i \in \{1..p\}$ est plus optimal.*
- *Si $\lambda_{i_1} = \lambda_{i_2}$ la limite de diffusion est isotrope.*
- *Le principe de preuve est le même en 3D.*

Le calcul formel précédent, basé sur un développement de Hilbert montre que l'équation de diffusion limite provient des conditions de compatibilités assurant l'existence des termes d'ordre un et d'ordre deux dans la solution formelle du développement de Hilbert. Cependant ces calculs ne peuvent pas justifier à eux seuls la limite de diffusion. En revanche, ils permettent de construire une preuve rigoureuse de convergence des solutions du système de Friedrichs vers celles de l'équation de diffusion pour certaines données initiales ce qui est l'objet de la sous-section suivante.

4.1.3 Preuve rigoureuse d'existence de la limite de diffusion

Dans cette sous-section, on apporte une preuve rigoureuse de convergence des solutions de systèmes de Friedrichs vers celles de l'équation de diffusion. On détaille la preuve en dimension deux. Le principe de démonstration est le même en dimension trois. Le coefficient de relaxation σ est supposé constant et strictement positif. Le principe de la preuve est décrit dans le cours de G. Allaire et F. Golse [AG11] pour une équation de transport linéaire. On commence par construire la solution formelle de Hilbert tronquée à l'ordre deux que l'on obtient à partir de la preuve formelle précédente qui est solution du système de Friedrichs (4.1) avec un terme source résiduel. Pour finir il suffit d'étudier la différence entre la solution du système de Friedrichs et la solution formelle de Hilbert tronquée à l'ordre deux et de montrer que le terme résiduel est borné par ε.

Théorème 4.3. *On suppose que l'hypothèse de structure (H_1) est vérifiée par les matrices A_1, A_2 et R de (4.16). Le vecteur $\mathbf{U}_\varepsilon$ est solution du système de Friedrichs*

$$\begin{cases} \partial_t\mathbf{U}_\varepsilon + \dfrac{1}{\varepsilon}A_1\partial_x\mathbf{U}_\varepsilon + \dfrac{1}{\varepsilon}A_2\partial_y\mathbf{U}_\varepsilon = -\dfrac{\sigma}{\varepsilon^2}R\mathbf{U}_\varepsilon, \\ (\mathbf{U}_\varepsilon(t=0), \mathbf{E}_i) = V_{in,i}, \ \textit{pour } i \leq p, \\ (\mathbf{U}_\varepsilon(t=0), \mathbf{E}_i) = 0, \ \textit{pour } i > p, \end{cases} \tag{4.16}$$

où $\mathbf{V}_{in}$ une fonction de $H^2(D)$. Soit $\widetilde{\mathbf{U}}_0$, $\widetilde{\mathbf{U}}_1$ et $\widetilde{\mathbf{U}}_2$ construits de la façon suivante. Le terme d'ordre zéro $\widetilde{\mathbf{U}}_0 = \sum_j (\widetilde{\mathbf{U}}_0, \mathbf{E}_j)\mathbf{E}_j$ est solution de

$$\begin{cases} \partial_t\mathbf{V} - \dfrac{1}{\lambda_{i_1}\sigma}K_1\partial_{xx}\mathbf{V} - \dfrac{1}{\lambda_{i_2}\sigma}K_2\partial_{yy}\mathbf{V} = \mathbf{0}, \\ (\widetilde{\mathbf{U}}_0, \mathbf{E}_i) = 0, \ i \in \{p+1...n\}, \end{cases} \tag{4.17}$$

avec $\mathbf{V} = ((\widetilde{\mathbf{U}}_0, \mathbf{E}_1), ..., (\widetilde{\mathbf{U}}_0, \mathbf{E}_p))$. *Les termes d'ordres supérieurs* $\widetilde{\mathbf{U}}_k$ ($k = 1$ *ou* $k = 2$) *sont définis par*

$$\begin{cases} (\widetilde{\mathbf{U}}_k, \mathbf{E}_i) = 0,\ i \in \{1...p\}, \\ (\widetilde{\mathbf{U}}_k, \mathbf{E}_i) = -\dfrac{1}{\lambda_i \sigma}(A_1 \partial_x \widetilde{\mathbf{U}}_{k-1} + A_2 \partial_y \widetilde{\mathbf{U}}_{k-1}, \mathbf{E}_i),\ i \in \{p+1...n\}. \end{cases} \tag{4.18}$$

Alors la solution formelle tronquée à l'ordre deux $\tilde{\mathbf{U}}_\varepsilon = \tilde{\mathbf{U}}_0 + \varepsilon \tilde{\mathbf{U}}_1 + \varepsilon^2 \tilde{\mathbf{U}}_2$ *satisfait l'estimation*

$$||\mathbf{U}_\varepsilon(T) - \widetilde{\mathbf{U}}_\varepsilon(T)||_{L^2(D)} \leq \varepsilon C_0 ||\widetilde{\mathbf{U}}_0(t=0)||_{H^2(D)} + TC_{mf}\varepsilon \left(||\widetilde{\mathbf{U}}_0||_{X^3} + ||\partial_t \widetilde{\mathbf{U}}_0||_{X^2} \right). \tag{4.19}$$

Pour terminer on obtient l'estimation suivante entre les solutions du système de Friedrichs (4.16) et celles de l'équation de diffusion (4.17)

$$||\mathbf{U}_\varepsilon(T) - \widetilde{\mathbf{U}}_0(T)||_{L^2(D)} \leq \varepsilon C \left(||\widetilde{\mathbf{U}}_0(t=0)||_{H^2(D)} + ||\widetilde{\mathbf{U}}_0(T)||_{H^2(D)} \right) + TC_f \varepsilon \left(||\widetilde{\mathbf{U}}_0||_{X^3} + ||\partial_t \widetilde{\mathbf{U}}_0||_{X^2} \right), \tag{4.20}$$

Démonstration. Afin d'obtenir ce résultat, on propose de montrer que les solutions de l'équation de chaleur sont solutions d'un système de Friedrichs avec un terme source supplémentaire homogène à ε. Une fois ce résultat obtenu on pourra en déduire une estimations entre les solutions de diffusion et les solutions du système de Friedrichs.

Étape 1

Dans cette première étape on souhaite montrer que la solution formelle tronquée à l'ordre deux $\tilde{\mathbf{U}}_\varepsilon = \widetilde{\mathbf{U}}_0 + \varepsilon \widetilde{\mathbf{U}}_1 + \varepsilon^2 \widetilde{\mathbf{U}}_2$ avec $\widetilde{\mathbf{U}}_0$, $\widetilde{\mathbf{U}}_1$, $\widetilde{\mathbf{U}}_2$ définis par (4.17)-(4.18) est solution du système

$$\partial_t \widetilde{\mathbf{U}}_\varepsilon + \frac{1}{\varepsilon} A_1 \partial_x \widetilde{\mathbf{U}}_\varepsilon + \frac{1}{\varepsilon} A_2 \partial_y \widetilde{\mathbf{U}}_\varepsilon + \frac{\sigma}{\varepsilon^2} R \widetilde{\mathbf{U}}_\varepsilon = \mathbf{S} \tag{4.21}$$

avec $\mathbf{S}$ défini par

$$\mathbf{S} = \varepsilon \partial_t \widetilde{\mathbf{U}}_1 + \varepsilon^2 \partial_t \widetilde{\mathbf{U}}_2 + \varepsilon(\partial_x A_1 \widetilde{\mathbf{U}}_2 + \partial_y A_2 \widetilde{\mathbf{U}}_2). \tag{4.22}$$

On commence par introduire la matrice Q, $Q = (\mathbf{E}_1, ...\mathbf{E}_n)$ formée des vecteurs propres de R et le vecteur $\mathbf{F}$ tels que

$$\mathbf{F} = Q^t \left(\partial_t \widetilde{\mathbf{U}}_\varepsilon + \frac{1}{\varepsilon} A_1 \partial_x \widetilde{\mathbf{U}}_\varepsilon + \frac{1}{\varepsilon} A_2 \partial_y \widetilde{\mathbf{U}}_\varepsilon + \frac{\sigma}{\varepsilon^2} R \widetilde{\mathbf{U}}_\varepsilon \right)$$

qui peut se réécrire sous la forme

$$\mathbf{F} = \begin{cases} \partial_t \begin{pmatrix} (\mathbf{E}_1, \widetilde{\mathbf{U}}_\varepsilon) \\ \vdots \\ (\mathbf{E}_p, \widetilde{\mathbf{U}}_\varepsilon) \end{pmatrix} + \frac{1}{\varepsilon} \begin{pmatrix} (\mathbf{E}_1, A_1 \partial_x \widetilde{\mathbf{U}}_\varepsilon) \\ \vdots \\ (\mathbf{E}_p, A_1 \partial_x \widetilde{\mathbf{U}}_\varepsilon) \end{pmatrix} + \frac{1}{\varepsilon} \begin{pmatrix} (\mathbf{E}_1, A_2 \partial_x \widetilde{\mathbf{U}}_\varepsilon) \\ \vdots \\ (\mathbf{E}_p, A_2 \partial_y \widetilde{\mathbf{U}}_\varepsilon) \end{pmatrix}, \\ \partial_t \begin{pmatrix} (\mathbf{E}_{p+1}, \widetilde{\mathbf{U}}_\varepsilon) \\ \vdots \\ (\mathbf{E}_n, \widetilde{\mathbf{U}}_\varepsilon) \end{pmatrix} + \frac{1}{\varepsilon} \begin{pmatrix} (\mathbf{E}_{p+1}, A_1 \partial_x \widetilde{\mathbf{U}}_\varepsilon) \\ \vdots \\ (\mathbf{E}_n, A_1 \partial_x \widetilde{\mathbf{U}}_\varepsilon) \end{pmatrix} + \frac{1}{\varepsilon} \begin{pmatrix} (\mathbf{E}_{p+1}, A_2 \partial_y \widetilde{\mathbf{U}}_\varepsilon) \\ \vdots \\ (\mathbf{E}_n, A_2 \partial_y \widetilde{\mathbf{U}}_\varepsilon) \end{pmatrix} + \frac{\sigma}{\varepsilon^2} \begin{pmatrix} \lambda_{p+1}(\mathbf{E}_{p+1}, \widetilde{\mathbf{U}}_\varepsilon) \\ \vdots \\ \lambda_n(\mathbf{E}_n, \widetilde{\mathbf{U}}_\varepsilon) \end{pmatrix} \end{cases} \tag{4.23}$$

On souhaite donc montrer que $\mathbf{F} = ((\mathbf{S}, \mathbf{E}_1), ..., (\mathbf{S}, \mathbf{E}_n)) = Q^t \mathbf{S}$. Comme les vecteurs propres forment une base de $\mathbb{R}^n$, on peut écrire chaque terme de la reconstruction de Hilbert sous la forme

$$\widetilde{\mathbf{U}}_k = \sum_{j=1}^{n} (\widetilde{\mathbf{U}}_k, \mathbf{E}_j) \mathbf{E}_j.$$

Afin de simplifier les notations on considère d'abord les p premières équations de (4.23) puis, dans un second temps, les $n-p$ suivantes.

Résultats pour les p premières relations de (4.23) :
On injecte $\widetilde{\mathbf{U}}_\varepsilon$ dans les p premières équations, et on étudie le terme $\partial_x(A_1\widetilde{\mathbf{U}}_0,\mathbf{E}_i)+\partial_y(A_2\tilde{\mathbf{U}}_0,\mathbf{E}_i)$ du système obtenu. En utilisant l'hypothèse (H_1) et la définition de $\widetilde{\mathbf{U}}_0$ qui montre que $\widetilde{\mathbf{U}}_0\in\operatorname{Ker}R$, on remarque que $(A_1\widetilde{\mathbf{U}}_0,\mathbf{E}_i)=(\widetilde{\mathbf{U}}_0,\mathbf{E}_{i_1})=0$ et $(A_2\widetilde{\mathbf{U}}_0,\mathbf{E}_i)=0$. À présent on étudie les termes d'ordre 1 : $(A_1\partial_x\widetilde{\mathbf{U}}_1,\mathbf{E}_i)$ et $(A_2\partial_x\widetilde{\mathbf{U}}_1,\mathbf{E}_i)$. En utilisant l'hypothèse (H_1), on a

$$(A_1\widetilde{\mathbf{U}}_1,\mathbf{E}_i)=\gamma_i^1(\widetilde{\mathbf{U}}_1,\mathbf{E}_{i_1})\quad\text{ et }(A_2\widetilde{\mathbf{U}}_1,\mathbf{E}_i)=\gamma_i^2(\widetilde{\mathbf{U}}_1,\mathbf{E}_{i_2})\quad,\forall i\in\{1..p\},\tag{4.24}$$

Par conséquent

$$\begin{cases}\partial_x(A_1\widetilde{\mathbf{U}}_1,\mathbf{E}_i)=\gamma_i^1\partial_x(\widetilde{\mathbf{U}}_1,\mathbf{E}_{i_1}) & \forall i\in\{1..p\},\\ \partial_x(A_2\widetilde{\mathbf{U}}_1,\mathbf{E}_i)=\gamma_i^2\partial_y(\widetilde{\mathbf{U}}_1,\mathbf{E}_{i_2}) & \forall i\in\{1..p\}.\end{cases}\tag{4.25}$$

En utilisant l'hypothèse (H_1) et la définition de $\widetilde{\mathbf{U}}_1$, on a

$$\begin{cases}\partial_x(A_1\widetilde{\mathbf{U}}_1,\mathbf{E}_i)=-\gamma_i^1\partial_x\dfrac{1}{\lambda_{i_1}\sigma}\left(A_1\partial_x\widetilde{\mathbf{U}}_0,A_1\dfrac{\mathbf{E}_i}{\gamma_i^1}\right)-\gamma_i^1\partial_x\dfrac{1}{\lambda_{i_1}\sigma}\left(A_2\partial_y\widetilde{\mathbf{U}}_0,A_1\dfrac{\mathbf{E}_i}{\gamma_i^1}\right) & \forall i\in\{1..p\},\\ \partial_y(A_2\tilde{\mathbf{U}}_1,\mathbf{E}_i)=-\gamma_i^2\partial_y\dfrac{1}{\lambda_{i_2}\sigma}\left(A_1\partial_x\widetilde{\mathbf{U}}_0,A_2\dfrac{\mathbf{E}_i}{\gamma_i^2}\right)-\gamma_i^2\partial_y\dfrac{1}{\lambda_{i_2}\sigma}\left(A_2\partial_y\widetilde{\mathbf{U}}_0,A_2\dfrac{\mathbf{E}_i}{\gamma_i^2}\right) & \forall i\in\{1..p\}.\end{cases}\tag{4.26}$$

On simplifie, et on obtient alors

$$\partial_x(A_1\widetilde{\mathbf{U}}_1,\mathbf{E}_i)=-\frac{1}{\lambda_{i_1}\sigma}\Big(\partial_{xx}(A_1\widetilde{\mathbf{U}}_0,A_0\mathbf{E}_i)+\partial_{yx}(A_2\widetilde{\mathbf{U}}_0,A_1\mathbf{E}_i)\Big)\quad\forall i\in\{1..p\},\tag{4.27}$$

et

$$\partial_y(A_2\widetilde{\mathbf{U}}_1,\mathbf{E}_i)=-\frac{1}{\lambda_{i_2}\sigma}\Big(\partial_{xy}(A_1\widetilde{\mathbf{U}}_0,A_2\mathbf{E}_i)+\partial_{yy}(A_2\widetilde{\mathbf{U}}_0,A_2\mathbf{E}_i)\Big)\quad\forall i\in\{1..p\}.\tag{4.28}$$

Enfin, on peut remarquer que les termes croisés $(A_1\widetilde{\mathbf{U}}_0,A_2\mathbf{E}_i)$ et $(A_2\widetilde{\mathbf{U}}_0,A_1\mathbf{E}_i)$ sont nuls en utilisant les calculs (4.13)-(4.14). Pour conclure, on utilise le même type de raisonnement que pour la relation (4.15). Par conséquent, les p premières relations de (4.23) sont égales à

$$F_i=\partial_t\mathbf{V}_i-\frac{1}{\lambda_{i_1}\sigma}\partial_{xx}(K_1\mathbf{V})_i-\frac{1}{\lambda_{i_2}\sigma}\partial_{yy}(K_2\mathbf{V})_i+\varepsilon\partial_t(\widetilde{\mathbf{U}}_1,\mathbf{E}_i)+\varepsilon^2(\widetilde{\mathbf{U}}_2,\mathbf{E}_i)+\varepsilon\partial_t(\partial_xA_1\widetilde{\mathbf{U}}_2+\partial_yA_2\widetilde{\mathbf{U}}_2,\mathbf{E}_i).$$

Puisque $\mathbf{V}$ satisfait l'équation de diffusion (4.17), on obtient alors

$$F_i=\varepsilon\partial_t(\widetilde{\mathbf{U}}_1,\mathbf{E}_i)+\varepsilon^2\partial_t(\widetilde{\mathbf{U}}_2,\mathbf{E}_i)+\varepsilon(\partial_xA_1\widetilde{\mathbf{U}}_2+\partial_yA_2\widetilde{\mathbf{U}}_2,\mathbf{E}_i),\ i\in\{1..p\}\,.\tag{4.29}$$

Résultats pour les $n-p$ dernières relations de (4.23) :
On incorpore la définition de $\widetilde{\mathbf{U}}_\varepsilon$ (4.17)-(4.18) dans ces relations. Par conséquent chaque équation de (4.23) pour $i\in\{p+1..n\}$ est équivalente à

$$F_i=\varepsilon\partial_t(\widetilde{\mathbf{U}}_1,\mathbf{E}_i)+\varepsilon^2\partial_t(\widetilde{\mathbf{U}}_2,\mathbf{E}_i)+\varepsilon(\partial_xA_1\widetilde{\mathbf{U}}_2+\partial_yA_2\widetilde{\mathbf{U}}_2,\mathbf{E}_i),\ i\in\{p+1...n\}\,.\tag{4.30}$$

Il suffit de multiplier par Q pour conclure, en remarquant que $Q\mathbf{F}=\mathbf{S}$. La solution formelle tronquée à l'ordre deux est donc solution du système (4.21) avec $\mathbf{S}$ défini par (4.22).
Étape 2

On souhaite maintenant estimer le terme résiduel du système (4.21). On commence par étudier la norme $L^2(D)$ du terme résiduel $\mathbf{S}$. En développant et en utilisant l'inégalité triangulaire, on déduit que

$$||\mathbf{S}(t)||_{L^2(D)} = \int_D (\mathbf{S},\mathbf{S})_{\mathbb{R}^n} \leq \varepsilon \left(||\partial_t \widetilde{\mathbf{U}}_1||_{L^2(D)} + \varepsilon ||\partial_t \widetilde{\mathbf{U}}_2||_{L^2(D)} + ||A_1 \partial_x \widetilde{\mathbf{U}}_2 + A_2 \partial_y \widetilde{\mathbf{U}}_2||_{L^2(D)} \right) \quad (4.31)$$

On étudie chaque terme de cette estimation. En utilisant les relations (4.18) on obtient

$$\partial_t \widetilde{\mathbf{U}}_1 = - \sum_{j=p+1}^{n} \frac{1}{\lambda_i \sigma} \left(A_1 \partial_{t,x} \widetilde{\mathbf{U}}_0 + A_2 \partial_{t,y} \widetilde{\mathbf{U}}_0, \mathbf{E}_j \right) \mathbf{E}_j$$

Puisque les vecteurs $\mathbf{E}_j$ sont orthonormés, en utilisant l'inégalité de Cauchy-Schwarz et l'inégalité triangulaire on obtient

$$||\partial_t \widetilde{\mathbf{U}}_1||_{L^2(D)} \leq C_1 ||\partial_t \widetilde{\mathbf{U}}_0||_{H^1(D)}$$

avec la constante C_1 dépendant des normes et des valeurs propres de A_1, A_2 ainsi que de σ. De la même façon on montre que

$$\widetilde{\mathbf{U}}_2 = - \sum_{j=p+1}^{n} \frac{1}{\lambda_i^2 \sigma^2} \left(A_1^2 \partial_{xx} \widetilde{\mathbf{U}}_0 + (A_1 A_2 + A_2 A_1) \partial_{xy} \widetilde{\mathbf{U}}_0 + A_2^2 \partial_{yy} \widetilde{\mathbf{U}}_0, \mathbf{E}_j \right) \mathbf{E}_j.$$

À partir de là en utilisant les inégalités sur les normes matricielles et vectorielles ainsi que l'inégalité de Cauchy-Schwarz on montre que

$$||\partial_t \widetilde{\mathbf{U}}_2||_{L^2(D)} \leq C_2 ||\partial_t \widetilde{\mathbf{U}}_0||_{H^2(D)}$$

et

$$||A_1 \partial_x \widetilde{\mathbf{U}}_2 + A_2 \partial_y \widetilde{\mathbf{U}}_2||_{L^2(D)} \leq C_3 ||\widetilde{\mathbf{U}}_0||_{H^3(D)}$$

On obtient donc la majoration avec la constante $C_{mf} = \max(C_1, \varepsilon C_2, C_3)$

$$||\mathbf{S}(t)||_{L^2(D)} \leq C_{mf} \varepsilon \left(||\partial_t \widetilde{\mathbf{U}}_0||_{H^2(D)} + ||\widetilde{\mathbf{U}}_0||_{H^3(D)} \right). \quad (4.32)$$

Pour finir il suffit de prendre sur $[0,T]$ le maximum des normes introduites dans (4.32).

$$||\mathbf{S}||_{L^\infty([0,T]:L^2(D))} \leq C_{mf} \varepsilon \left(||\partial_t \widetilde{\mathbf{U}}_0||_{X^2(D)} + ||\widetilde{\mathbf{U}}_0||_{X^3(D)} \right). \quad (4.33)$$

Le terme résiduel $\mathbf{S}$ étant majoré par ε, reste à démontrer que l'erreur entre le système de Friedrichs et le système (4.21) est homogène à ce terme résiduel pour assurer la convergence entre la solution $\mathbf{U}_\varepsilon$ du système (4.16) et la solution $\widetilde{\mathbf{U}}_\varepsilon$ du système (4.21).

Étape 3
On pose $\mathbf{e}_\varepsilon = \mathbf{U}_\varepsilon - \widetilde{\mathbf{U}}_\varepsilon$ avec $\mathbf{U}_\varepsilon$ la solution du système de Friedrichs et $\widetilde{\mathbf{U}}_\varepsilon$ est solution de (4.17)-(4.18). La quantité $\mathbf{e}_\varepsilon$ est solution du système

$$\begin{cases} \partial_t \mathbf{e}_\varepsilon + \dfrac{1}{\varepsilon} A_1 \partial_x \mathbf{e}_\varepsilon + \dfrac{1}{\varepsilon} A_2 \partial_y \mathbf{e}_\varepsilon + \dfrac{\sigma}{\varepsilon^2} R \mathbf{e}_\varepsilon = \mathbf{S}, \\ \\ \mathbf{e}_\varepsilon(t=0) = -\varepsilon \widetilde{\mathbf{U}}_1(t=0) - \varepsilon^2 \widetilde{\mathbf{U}}_2(t=0), \end{cases}$$

En multipliant par $\mathbf{e}_\varepsilon$ et en intégrant sur D, on obtient

$$\partial_t \frac{1}{2}\left(||\mathbf{e}_\varepsilon||^2_{L^2(D)}\right) + \frac{1}{\varepsilon}\int_D (\mathbf{e}_\varepsilon, A_1\partial_x \mathbf{e}_\varepsilon) + (\mathbf{e}_\varepsilon, A_2\partial_y \mathbf{e}_\varepsilon)dxdy = -\frac{\sigma}{\varepsilon^2}\int_D (\mathbf{e}_\varepsilon, R\mathbf{e}_\varepsilon)dxdy + \int_D (\mathbf{S}, \mathbf{e}_\varepsilon)dxdy.$$

On a

$$\int_D (\mathbf{e}_\varepsilon, A_1\partial_x \mathbf{e}_\varepsilon)dxdy = 0,\ \int_D (\mathbf{e}_\varepsilon, A_2\partial_y \mathbf{e}_\varepsilon)dxdy = 0 \text{ et } \int_D (\mathbf{e}_\varepsilon, R\mathbf{e}_\varepsilon)dxdy \leq 0.$$

En utilisant l'inégalité de Cauchy-Schwarz, on déduit que

$$\partial_t \frac{1}{2}\left(||\mathbf{e}_\varepsilon||^2_{L^2(D)}\right) \leq ||\mathbf{S}||_{L^2(D)}||\mathbf{e}_\varepsilon||_{L^2(D)},$$

On pose $E(t) = ||\mathbf{e}_\varepsilon(t)||^2_{L^2(D)}$ et on remarque que l'inégalité précédente est équivalente à l'équation différentielle

$$E'(t) \leq 2||\mathbf{S}||_{L^2(D)}\sqrt{E(t)}. \tag{4.34}$$

En intégrant sur $[0, T]$, on majore l'erreur au temps T de la façon suivante

$$\sqrt{E(T)} \leq \sqrt{E(0)} + T||\mathbf{S}||_{L^\infty([0,T]:L^2(D))}.$$

La seconde partie de l'estimation (4.35) est obtenue à partir du résultat de l'étape 2 :

$$||\mathbf{S}||_{L^\infty([0,T]:L^2(D))} \leq C_{mf}\varepsilon\left(||\widetilde{\mathbf{U}}_0||_{X^3} + ||\partial_t\widetilde{\mathbf{U}}_0||_{X^2}\right).$$

Pour conclure, on doit estimer $E(0)$. En utilisant les définitions de $\widetilde{\mathbf{U}}_1$ et $\widetilde{\mathbf{U}}_2$, on obtient

$$||\mathbf{e}_\varepsilon(t=0)||_{L^2(D)} \leq \varepsilon C_0\left(||\widetilde{\mathbf{U}}_0(t=0)||_{H^2(D)}\right)$$

et

$$\sqrt{E(T)} \leq \varepsilon C_0||\widetilde{\mathbf{U}}_0(t=0)||_{H^2(D)} + TC_{mf}\varepsilon\left(||\widetilde{\mathbf{U}}_0||_{X^3} + ||\partial_t\widetilde{\mathbf{U}}_0||_{X^2}\right).$$

Ces résultats donnent l'estimation suivante en temps T :

$$||\mathbf{e}_\varepsilon(T)||_{L^2(D)} = ||\mathbf{U}_\varepsilon(T) - \widetilde{\mathbf{U}}_\varepsilon(T)||_{L^2(D)} \leq \varepsilon C_0||\widetilde{\mathbf{U}}_0(t=0)||_{H^2(D)} + TC_{mf}\varepsilon\left(||\widetilde{\mathbf{U}}_0||_{X^3} + ||\partial_t\widetilde{\mathbf{U}}_0||_{X^2}\right). \tag{4.35}$$

Étape 4

Pour finir on utilise l'estimation (4.35) pour montrer la convergence des solutions du système de Friedrichs vers les solutions de l'équation de diffusion. L'étape précédente donne

$$||\mathbf{U}_\varepsilon(T) - \widetilde{\mathbf{U}}_\varepsilon(T)||_{L^2(D)} \leq \varepsilon C_0||\widetilde{\mathbf{U}}_0(t=0)||_{H^2(D)} + TC_{mf}\varepsilon\left(||\widetilde{\mathbf{U}}_0||_{X^3} + ||\partial_t\widetilde{\mathbf{U}}_0||_{X^2}\right).$$

On montre ainsi

$$||\mathbf{U}_\varepsilon(T) - \widetilde{\mathbf{U}_0}(T)||_{L^2(D)} \leq ||\mathbf{U}_\varepsilon(T) - \widetilde{\mathbf{U}}_\varepsilon(T)||_{L^2(D)} + \varepsilon||\widetilde{\mathbf{U}}_1(T) + \varepsilon\widetilde{\mathbf{U}}_2(T)||_{L^2(D)}.$$

On utilise le même type de majoration que précédemment pour obtenir

$$||\widetilde{\mathbf{U}}_1(T) + \varepsilon\widetilde{\mathbf{U}}_2(T)||_{L^2(D)} \leq \varepsilon\left(||\widetilde{\mathbf{U}}_0(T)||_{H^2}\right).$$

Ce qui donne le résultat final. □

Remarque 4.1.3.
- *Le résultat est donné pour un domaine périodique et pour des données initiales correctement choisies (la différence entre la condition initiale du système de Friedrichs et celle du système limite est homogène à ε). Le principe de démonstration reste valable si on introduit des conditions limites ou pour des données initiales éloignées l'une de l'autre. Cependant, une difficulté supplémentaire apparaît due a l'étude des couches limites [AG11].*
- *Le principe de la preuve est le même dans le cas d'un paramètre de relaxation σ variable. Cependant, il faut une régularité suffisante pour la fonction $\frac{1}{\sigma}$ et pour la solution de l'équation de diffusion afin d'utiliser des inégalités plus fines pour la majoration de $\mathbf{S}$.*

Remarque 4.1.4. *Dans [DM04] les auteurs démontrent un résultat de convergence similaire pour des systèmes de Friedrichs avec termes sources non constant raides et non raides. Le principe de preuve est différent, en effet les auteurs proposent d'utiliser des arguments de compacité. Le résultat est plus général mais ne permet d'obtenir une estimation d'erreur entre les solutions du système de Friedrichs et les solutions de diffusion. Pour obtenir ce résultat les auteurs font une hypothèse de structure équivalente à $A\mathbf{E}_i \in \operatorname{Ker} R^{\perp}$, $B\mathbf{E}_i \in \operatorname{Ker} R^{\perp}$ pour $i \in \{1...p\}$. Cependant il faut ajouter d'autres hypothèses afin d'assurer le caractère dissipatif de l'équation limite. L'hypothèse (H_1) que nous avons utilisé est certe plus restrictive mais permet d'assurer automatiquement l'existence de la limite de diffusion et le caractère dissipatif de l'équation limite.*

Maintenant que la limite de diffusion est correctement définie, on propose une méthode numérique « asymptotic preserving », basée sur une version légèrement plus restrictive de l'hypothèse (H_1).

4.2 Décomposition « diffusive - non diffusive »

4.2.1 Décomposition « diffusive - non diffusive »

On se place dans le cas d'une équation limite scalaire ($\dim \operatorname{Ker}(R) = 1$). On propose une méthode numérique qui consiste à décomposer le système de Friedrichs (4.1) entre un modèle du type équation de la chaleur hyperbolique et un autre système hyperbolique qui n'intervient pas dans la limite de diffusion. Cette méthode est comparable à la méthode « micro - macro » utilisée dans [LM07] afin de construire des schémas « asymptotic preserving » en dimension un. Elle consiste à isoler puis séparer les termes négligeables en régime de diffusion des termes associés à l'inconnue du système de diffusion limite. Une fois cette décomposition effectuée, il est plus aisé de trouver une discrétisation valide. On décrit la méthode en dimension deux, elle reste valable en dimension trois. On se restreint au cas où

$$\dim(\operatorname{Ker} R) = 1. \qquad (H_2)$$

Cette hypothèse est vérifiée par les modèles simplifiés utilisés pour approcher les problèmes de transport. En transport linéaire R est une Z-matrice avec termes diagonaux négatifs. En outre, les vecteurs propres associés à la valeur propre nulle ont des composantes positives. Puisque R est symétrique réelle il existe une décomposition de la forme $R = QDQ^t$ où D est une matrice diagonale et Q une matrice orthogonale formée des vecteurs propres de R. La matrice R étant symétrique positive, les éléments diagonaux de D (qui sont les valeurs propres) sont positifs ou nuls. On note λ_i les valeurs propres de R et par convention $\lambda_1 = 0$ avec $\mathbf{E}_1$ le vecteur propre associé. On note $\mathbf{E}_i$ un vecteur propre de R. On suppose aussi que $A_1\mathbf{E}_1$, $A_2\mathbf{E}_1$, sont deux vecteurs propres indépendants associés à la même valeur propre de R, en d'autres termes l'équation de diffusion limite est isotrope. On traduit cela par l'hypothèse suivante proche de l'hypothèse (H_1) :

$$A_1\mathbf{E}_1 = a\mathbf{E}_2,\ A_2\mathbf{E}_1 = a\mathbf{E}_3. \qquad (H_3)$$

On pose $\mathbf{V} = Q^t\mathbf{U}$, $\mathbf{V}$ satisfaisant

$$\partial_t \mathbf{V} + \frac{1}{\varepsilon}\left(A_1'\partial_x \mathbf{V} + A_2'\partial_y \mathbf{V}\right) = -\frac{\sigma}{\varepsilon^2} D\mathbf{V}, \tag{4.36}$$

avec $A_1^{'} = Q^t A_1 Q$, $A_2^{'} = Q^t A_2 Q$.

Lemme 4.4. *Sous les hypothèses de la Définition 4.1.1 et les Hypothèses (H_2) et (H_3) les matrices $A_1^{'}$, $A_2^{'}$ ont la structure suivante*

$$A_1^{'} = \begin{pmatrix} 0 & a & 0 & 0 & .. & 0 \\ a & .. & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. & .. \end{pmatrix} A_2^{'} = \begin{pmatrix} 0 & 0 & a & 0 & .. & 0 \\ 0 & .. & .. & .. & .. & .. \\ a & .. & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. & .. \\ .. & .. & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. & .. \end{pmatrix} \tag{4.37}$$

Démonstration. On détaille le calcul pour la matrice $A_1^{'}$. Celui-ci est immédiatement transposable à la matrice $A_2^{'}$. Cette structure dérive uniquement des hypothèses de symétrie et de l'hypothèse (H_3). En effet

$$A_{1,11}^{'} = (\mathbf{E}_1, A_1\mathbf{E}_1) = (\mathbf{E}_1, a\mathbf{E}_2) = 0,$$

car la base est orthonormée. Pour $A_{1,12}^{'} = (\mathbf{E}_1, A_1\mathbf{E}_2)$, on utilise la symétrie de A_1 afin d'obtenir $(a\mathbf{E}_2, \mathbf{E}_2) = a$. De la même façon, on montre que les autres coefficients ont la bonne forme. □

En utilisant le lemme précédent, on peut immédiatement réécrire le système (4.36) sous la forme suivante

$$\partial_t \mathbf{V} + \frac{1}{\varepsilon}\left(P_{1,x}\partial_x \mathbf{V} + P_{1,y}\partial_y \mathbf{V}\right) + \frac{1}{\varepsilon}\left(A_1^*\partial_x \mathbf{V} + A_2^*\partial_y \mathbf{V}\right) = -\frac{\sigma}{\varepsilon^2} D\mathbf{V}, \tag{4.38}$$

où les matrices $P_{1,x}$, $P_{1,y}$ sont des matrices carrées de taille n :

$$P_{1,x} = \begin{pmatrix} 0 & a & 0 & .. & .. & 0 \\ a & 0 & 0 & .. & .. & 0 \\ 0 & 0 & 0 & .. & .. & 0 \\ .. & .. & .. & .. & .. & .. \\ 0 & 0 & 0 & .. & .. & 0 \end{pmatrix} P_{1,y} = \begin{pmatrix} 0 & 0 & a & 0 & .. & 0 \\ 0 & 0 & 0 & 0 & .. & 0 \\ a & 0 & 0 & 0 & .. & 0 \\ 0 & 0 & 0 & 0 & .. & 0 \\ .. & .. & .. & .. & .. & .. \\ 0 & 0 & 0 & .. & .. & 0 \end{pmatrix}, \tag{4.39}$$

et $A_1^* = A_1^{'} - P_{1,x}$, $A_2^* = A_2^{'} - P_{1,y}$ sont des matrices symétriques. Pour finir il suffit de décomposer (4.38) en deux systèmes. Le premier système du type équation de la chaleur hyperbolique

$$\partial_t \mathbf{V} + \frac{1}{\varepsilon}\left(P_{1,x}\partial_x \mathbf{V} + P_{1,y}\partial_y \mathbf{V}\right) = -\frac{\sigma}{\varepsilon^2} D^{'}\mathbf{V} \tag{4.40}$$

avec $D_{11}^{'} = 0$, $D_{22}^{'} = D_{33}^{'} = 1$ et $D_{ii}^{'} = 0$ $i \geq 4$. Le second système est donné par

$$\partial_t \mathbf{V} + \frac{1}{\varepsilon}\left(A_1^*\partial_x \mathbf{V} + A_2^*\partial_y \mathbf{V}\right) = -\frac{\sigma}{\varepsilon^2} D^{''}\mathbf{V}, \tag{4.41}$$

avec $D_{11}^{''} = D_{22}^{''} = D_{33}^{''} = 0$ et $D_{ii}^{''} = 1$ $i \geq 4$. Une rapide analyse asymptotique montre que cette décomposition isole les inconnues en $O(1)$ et $O(\varepsilon)$ (premier système (4.40) de la décomposition) qui donnent l'équation de diffusion limite et les inconnues homogènes à $O(\varepsilon^2)$ (second système (4.41) de la décomposition) n'intervenant pas en régime de diffusion.

4.2.2 Discrétisation de la décomposition

Principe de discrétisation *: on propose de discrétiser le système homogène à l'équation de la chaleur hyperbolique issue de la décomposition avec un schéma « asymptotic preserving » comme les schémas aux noeuds JL-(a) (2.11)-(2.12), JL-(b) (2.22)-(2.23) ou comme les schémas aux arêtes P_1-MPFA et P_1-Breil-Maire (3.11)-(3.12)-(3.13) définis dans le Chapitre 3. La phase dite « non diffusive » est discrétisée en utilisant un schéma aux arêtes classiques, par exemple le schéma upwind, ou le schéma de Rusanov.*

Rappel : schéma upwind en dimension deux

On souhaite discrétiser le système suivant

$$\partial_t \mathbf{U} + \frac{1}{\varepsilon} A_1 \partial_x \mathbf{U} + \frac{1}{\varepsilon} A_2 \partial_y \mathbf{U} = \mathbf{0}, \quad \mathbf{U}(t=0,x,y) = \mathbf{U}_0. \tag{4.42}$$

On nomme $G\mathbf{n}$ la matrice $A_1 n^x + A_2 n^y$ avec $\mathbf{n} = (n^x, n^y)$ un vecteur de $\mathbb{R}^2$.

Définition 4.2.1 (Schéma upwind pour un système de Friedrichs)**.** *La discrétisation en espace est donnée par*

$$|\Omega_j|\, \partial_t \mathbf{U}_j + \sum_k l_{jk} \mathbf{U}_{jk} = \mathbf{0}, \tag{4.43}$$

avec les flux

$$\mathbf{U}_{jk} = (G\mathbf{n}_{jk})^+ \mathbf{U}_j + (G\mathbf{n}_{jk})^- \mathbf{U}_k,$$

pour les arêtes intérieures, $\mathbf{n}_{jk}$ étant la normale associée à l'arête $\partial\Omega_{jk}$.

Les matrices $(G\mathbf{n}_{jk})^{-,+}$ sont définies par $G\mathbf{n}_{jk}^+ = P^{-1}D^+P$, $G\mathbf{n}_{jk}^- = P^{-1}D^-P$. La matrice D^+ est définie par $D_{ii}^+ = \max(0,\lambda_i)$ et la matrice D^+ est définie par $D_{ii}^- = \min(0,\lambda_i)$ avec $\{\lambda_1..\lambda_n\}$ les valeurs propres de $G\mathbf{n}_{jk}$. Des preuves de stabilité et de convergence du schéma sont données dans [CVV00]. Cependant, ce schéma nécessite de connaître et de stocker les valeurs propres et vecteurs propres de la matrice $G\mathbf{n}_{jk}$. Sur des maillages non structurés et pour des modèles d'ordre élevé en discrétisation angulaire, le temps de calcul et la capacité de stockage nécessaire deviennent important. C'est pourquoi on propose une alternative : le schéma de Rusanov [Sah11].

Rappel : schéma de Rusanov

Le schéma de Rusanov est un schéma centré et stabilisé par un terme de viscosité numérique. Ce schéma nécessite seulement la connaissance d'une majoration de la valeur propre maximale de $G\mathbf{n}_{jk}$ dans chaque maille.

Définition 4.2.2. *Le schéma de Rusanov est défini par*

$$|\ \Omega_j\ |\ \partial_t \mathbf{U}_j + \sum_k l_{jk} \mathbf{U}_{jk} = \mathbf{0},$$

et par le flux numérique

$$\mathbf{U}_{jk} = G\mathbf{n}_{jk} \frac{(\mathbf{U}_j + \mathbf{U}_k)}{2} - \frac{1}{2} S_{jk} (\mathbf{U}_k - \mathbf{U}_j),$$

avec la vitesse de Rusanov locale $S_{jk} \geq \max(\lambda_{jk})$ et λ_{jk} les valeurs propres de $M = A_1 n_{jk}^x + A_2 n_{jk}^y$.

Ce schéma est stable et convergent [Sah11]. Maintenant que la décomposition « diffusive - non diffusive » ainsi que les différents schémas utilisés ont été introduits, on propose l'algorithme final permettant d'obtenir une discrétisation « asymptotic preserving » pour des systèmes de Friedrichs satisfaisant les hypothèses (H_2) et (H_3).

Algorithme final

Nous résumons ici l'algorithme de la méthode.

Algorithme 4.5. *L'algorithme de la décomposition « diffusive - non diffusive »*

- ***Étape 1*** *: Diagonalisation dans la base de R, du système*

$$\partial_t \mathbf{U} + \frac{1}{\varepsilon} A_1 \partial_x \mathbf{U} + \frac{1}{\varepsilon} A_2 \partial_y \mathbf{U} = -\frac{\sigma}{\varepsilon^2} R\mathbf{U}.$$

- ***Étape 2*** *: Décomposition du système diagonalisé*

$$\partial_t \mathbf{V} + \frac{1}{\varepsilon} A_1^{'} \partial_x \mathbf{V} + \frac{1}{\varepsilon} A_2^{'} \partial_y \mathbf{V} = -\frac{\sigma}{\varepsilon^2} D\mathbf{V},$$

  avec $\mathbf{V} = Q^t\mathbf{U}$*,* $A_1^{'} = Q^t A_1 Q$ *et* $A_2^{'} = Q^t A_2 Q$*. On obtient*

$$\partial_t \mathbf{V} + \frac{1}{\varepsilon} \left(P_{1,x} \partial_x \mathbf{V} + P_{1,y} \partial_y \mathbf{V}\right) + \frac{1}{\varepsilon} \left(A_1^* \partial_x \mathbf{V} + A_2^* \partial_y \mathbf{V}\right) = -\frac{\sigma}{\varepsilon^2} D\mathbf{V},$$

  avec $P_{1,x}$*,* $P_{1,y}$ *des matrices définies par (4.39) et* $A_1^* = A_1^{'} - P_{1,x}$*,* $A_1^* = A_1^{'} - P_{1,y}$*.*
- ***Étape 3*** *: Discrétisation du système homogène à l'équation de la chaleur hyperbolique*

$$\partial_t \mathbf{V} + \frac{1}{\varepsilon} \left(P_{1,x} \partial_x \mathbf{V} + P_{1,y} \partial_y \mathbf{V}\right) = -\frac{\sigma}{\varepsilon^2} D^{'} \mathbf{V},$$

  avec $D^{'}$ *définie par* $D_{22}^{'} = D_{33}^{'} = 1$ *et* $D_{ii}^{'} = 0$ $i \neq 2$ *et* $i \neq 3$ *avec un schéma « asymptotic preserving » du type JL-(a), JL-(b) ou* P_1*-MPFA.*
- ***Étape 4*** *: Discrétisation du second système*

$$\partial_t \mathbf{V} + \frac{1}{\varepsilon} \left(A_1^* \partial_x \mathbf{V} + A_2^* \partial_y \mathbf{V}\right) = -\frac{\sigma}{\varepsilon^2} D^{''} \mathbf{V},$$

  avec $D^{''}$ *définie par* $D_{11}^{''} = D_{22}^{''} = D_{33}^{''} = 0$ *et* $D_{ii}^{''} = 1$ $i \geq 4$*, avec un schéma classique comme le schéma de Rusanov ou le schéma upwind.*
- ***Étape 5*** *: Calcul des inconnues* $\mathbf{U}$ *à partir de* $\mathbf{V}$*.*

4.3 Applications aux modèles P_n

Dans cette section, on rappellera que les modèles P_n en dimension deux satisfont les hypothèses de la définition (4.1.1) ainsi que les hypothèses (H_2) et (H_3) nécessaires pour l'utilisation de notre algorithme. Enfin on proposera une analyse de la décomposition en dimension un.

4.3.1 Modèles P_n

La décomposition « diffusive - non diffusive » nécessite que le système hyperbolique satisfasse un certain nombre d'hypothèses. Afin d'appliquer cette méthode à la discrétisation des modèles P_n, on souhaite montrer que ces modèles satisfont les hypothèses nécessaires de la définition (4.1.1) et l'hypothèse (H_2) et (H_3). On rappelle que le transport de photons ou de neutrons est modélisé par

l'équation de transport avec un terme de collision :

$$\partial_t f(t,\mathbf{x},\mathbf{\Omega}) + \mathbf{\Omega}.\nabla f(t,\mathbf{x},\mathbf{\Omega}) = \sigma \int_{S^2} \Big(f(t,\mathbf{x},\mathbf{\Omega}') - f(t,\mathbf{x},\mathbf{\Omega}) \Big)\, d\mathbf{\Omega}',\ \mathbf{x} \in \mathbb{R}^2,\ \mathbf{\Omega} \in \mathbb{R}^2, t \in \mathbb{R}^+. \quad (4.44)$$

Comme nous l'avons rappelé dans la partie introductive, on ne résout pas cette équation directement mais on résout des modèles hyperboliques approchant cette équation comme les modèles P_n. On introduit par exemple, les matrices du modèle P_1 en dimension deux :

$$A_1 = \begin{pmatrix} 0 & \frac{1}{\sqrt{3}} & 0 \\ \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{3}} \\ 0 & 0 & 0 \\ \frac{1}{\sqrt{3}} & 0 & 0 \end{pmatrix}, \quad R = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Les modèles P_n multi-dimensionnels présentés, par exemple dans [Bru05] entrent dans le cadre des systèmes de Friedrichs avec la structure (4.1).

Propriétés Les modèles P_n (4.1) satisfont

- R est une matrice diagonale, 0 est valeur propre de multiplicité 1, et 1 valeur propre de multiplicité $n-1$ (hypothèse (H_2)).
- Les valeurs propres du système appartiennent à $]-1,1[$.
- En symétrisant le système on montre que les matrices A_1, A_2 satisfont l'hypothèse (H_3) et admettent la structure introduite dans le lemme 4.5
- Comme l'équation de transport (4.44), les modèles P_n (1.11)-(1.12) admettent la limite de diffusion suivante :

$$\partial_t U_1 - \operatorname{div}\left(\frac{1}{3\sigma}\nabla U_1\right) = 0,$$

avec U_1 la première inconnue des systèmes P_n correspondant au premier moment de $f(\mathbf{x},\mathbf{\Omega},t)$. Les trois premières propriétés sont démontrées dans [Bru05]- [Bru05], les hypothèses de la définition 4.1 sont donc satisfaites. On part des modèles P_n sous la forme (4.1). On applique le changement de variable introduit dans la section 1.2.1 ce qui permet d'obtenir exactement la structure des matrices suivantes

$$A_1 = \begin{pmatrix} 0 & a & 0 & 0 & 0 \\ a & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. \end{pmatrix},\ A_2 = \begin{pmatrix} 0 & 0 & a & 0 & 0 \\ 0 & .. & .. & .. & .. \\ a & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. \end{pmatrix}, \quad (4.45)$$

avec $a = \frac{1}{\sqrt{3}}$ si on numérote les inconnues ϕ_l^m par l croissant et

$$A_1 = \begin{pmatrix} 0 & a & 0 & 0 & 0 \\ a & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. \end{pmatrix},\ A_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & a \\ 0 & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. \\ 0 & .. & .. & .. & .. \\ a & .. & .. & .. & .. \end{pmatrix}, \quad (4.46)$$

si on numérote par m croissant. Dans les deux cas, on remarque que le système (4.1) satisfait l'hypothèse (H_3) et que les matrices A_1, A_2 admettent une structure qui permet utiliser la décomposition « diffusion - non diffusion » . Par conséquent, les approximations P_n peuvent être décomposées en un modèle homogène à l'équation de la chaleur hyperbolique et en un système qui n'intervient pas dans la limite de diffusion. Cette décomposition est assez naturelle pour les modèles P_n. En effet, lorsque

l'ordre des modèles P_n augmente, les harmoniques sphériques ajoutées dans le développement de la solution de l'équation de transport sont de plus en plus anisotropes. Le régime de diffusion étant isotrope, il est donc naturel que les harmoniques d'ordre supérieur n'interviennent pas en régime de diffusion. La décomposition consiste donc juste à séparer les modèles P_n entre les équations associées au modèle P_1 (correspondant aux deux premières harmoniques) qui est suffisant pour obtenir le régime de diffusion et les équations issues des harmoniques sphériques d'ordre supérieur qui sont ajoutées uniquement pour enrichir l'approximation du régime de transport ($\sigma << 1$).

Remarque 4.3.1. *On peut utiliser la décomposition « diffusive - non diffusive » pour des modèles P_n qui dérivent de l'équation de transport avec des opérateurs de collision comme*

$$Q(f) = \int_{S^2} p(\mathbf{\Omega}, \mathbf{\Omega}^{'}) \Big(f(t, \mathbf{x}, \mathbf{\Omega}^{'}) - f(t, \mathbf{x}, \mathbf{\Omega}) \Big) \, d\mathbf{\Omega}^{'}, \ ou \quad Q(f) = \triangle_{\theta} f(t, \mathbf{x}, \mathbf{\Omega}),$$

avec $\triangle_{\mathbf{\Omega}}$ l'opérateur de Laplace défini sur la sphère (Laplace-Beltrami) et $p(\mathbf{\Omega}, \mathbf{\Omega}^{'})$ une fonction de répartition angulaire.

4.3.2 Analyse de la méthode en dimension un

Afin de montrer la pertinence de la méthode de décomposition « diffusive - non diffusive » , on propose de valider cette méthode pour les discrétisations du modèle P_n à l'aide d'une étude de consistance en dimension un. Cette étude de consistance justifie l'utilisation d'un schéma classique pour la partie non diffusive. On propose donc d'utiliser le schéma de Gosse-Toscani (1.46) pour le système homogène à l'équation de la chaleur hyperbolique dans la décomposition et le schéma de Rusanov pour le second système. L'étude de consistance reste aussi valable si on utilise d'autres schémas pour chaque phase de la décomposition. Par exemple, on peut utiliser aussi le schéma de Jin-Levermore (1.34) pour la phase « diffusive » de la décomposition et un schéma upwind pour la phase « non diffusive » . On introduit le modèle en 1D

$$\partial_t \mathbf{U} + \frac{1}{\varepsilon} A \partial_x \mathbf{U} = -\frac{\sigma}{\varepsilon^2} D\mathbf{U},$$

avec la matrice A qui admet la décomposition $A = P_1 + A^{'}$ tel que P_1 soit la matrice définie par

$$P_{1,12} = P_{1,21} = \frac{1}{\sqrt{3}} \text{ et } P_{1,ij} = 0 \text{ pour } (i,j) \neq (1,2) \text{ et } (i,j) \neq (2,1)$$

et $A^{'}$ est une matrice qui satisfait

$$A^{'}_{1j} = A^{'}_{i1} = 0.$$

De plus D est une matrice diagonale définie par $D_{11} = 0$ et $D_{ii} = 1$ avec $i > 1$. Cette structure correspond à celle obtenue après la décomposition « diffusive - non diffusive » . On suppose ainsi que les données initiales satisfont $||(\partial_{t^a,x^b} U_1)(t=0)||_{L^\infty} \leq C_{a,b}$ et $||(\partial_{t^a,x^b} U_i)(t=0)||_{L^\infty} \leq \varepsilon C_{a,b}$ avec $i > 1$ et $(a,b) \in \mathbb{N}^2$. Par stabilité en norme L^∞ du système 1D ces estimations sont propagées en temps avec des constantes différentes. Maintenant on introduit le schéma final, après décomposition défini, par

$$\frac{\mathbf{U}_j^{n+1} - \mathbf{U}_j^n}{\Delta t} + \frac{A\mathbf{U}^n_{j+\frac{1}{2}} - A\mathbf{U}^n_{j-\frac{1}{2}}}{\varepsilon \Delta x} = -\frac{\sigma}{\varepsilon^2} D\mathbf{U}_j^n, \tag{4.47}$$

avec les flux

$$\begin{cases} (A\mathbf{U}_{j+\frac{1}{2}})^n_1 = M\left(\frac{U^n_{2,j}+U^n_{2,j+1}}{2\sqrt{3}} + \frac{U^n_{1,j}-U^n_{1,j+1}}{2\sqrt{3}}\right), \\ (A\mathbf{U}_{j+\frac{1}{2}})^n_2 = M\left(\frac{U^n_{1,j}+U^n_{1,j+1}}{2\sqrt{3}} + \frac{U^n_{2,j}-U^n_{2,j+1}}{2\sqrt{3}}\right) + \frac{(A^{'}\mathbf{U})^n_{2,j}+(A^{'}\mathbf{U})^n_{2,j+1}}{2} + S_{j+\frac{1}{2}} \frac{U^n_{2,j}-U^n_{2,j+1}}{2}, \\ (A\mathbf{U}_{j+\frac{1}{2}})^n_i = \left(\frac{(A^{'}\mathbf{U})^n_{i,j}+(A^{'}\mathbf{U})^n_{i,j+1}}{2} + S_{j+\frac{1}{2}} \frac{U^n_{i,j}-U^n_{i,j+1}}{2}\right) \text{ pour } i \geq 3, \end{cases} \tag{4.48}$$

avec U_i la ième composante du vecteur $\mathbf{U}$, $(A\mathbf{U})_i$ (resp. $(A^{'}\mathbf{U})_i$) la $i$ème ligne du vecteur $A\mathbf{U}$ (resp. $A^{'}\mathbf{U}$), $S_{j+\frac{1}{2}} = 1$ la vitesse locale de Rusanov et $M = \frac{2\sqrt{3}}{2\sqrt{3}+3\sigma\Delta x}$.

Lemme 4.6. *L'erreur de consistance de la première composante de l'équation vectorielle (4.47)*

$$O\left(\Delta x^2 + \varepsilon\Delta x + \Delta t\right). \tag{4.49}$$

L'erreur de consistance des autres composantes de l'équation vectorielle (4.47)

$$O(\Delta x + \Delta t). \tag{4.50}$$

Démonstration. La démonstration de consistance pour les deux premières composantes de (4.47)-(4.48) est identique à celle du Lemme (1.10) introduite dans le premier chapitre de cette étude, sauf pour le terme dans la seconde équation

$$\frac{(A^{'}\mathbf{U})^n_{2,j} + (A^{'}\mathbf{U})^n_{2,j+1}}{2} + S_{j+\frac{1}{2}}\frac{U^n_{2,j} - U^n_{2,j+1}}{2}.$$

Ceci montre que l'erreur de consistance pour la seconde équation est

$$CU^n_{2,j} = O(\Delta t + \Delta x) + \Delta x^2(A^{'}\partial_x\mathbf{U})_2(x_j, t_n) - \frac{\Delta x}{2\varepsilon}\partial_{xx}U_2(x_j, t_n).$$

Le premier terme vient de l'erreur de consistance du flux de Gosse-Toscani. Il suffit de constater que $U_2(x_j, t_n)$ et $(A^{'}\partial_x\mathbf{U})_2(x_j, t_n)$ sont homogènes à ε, par hypothèse et puisque $(A^{'}\mathbf{U})_2$ ne dépend pas de $U_1(x_j, t_n)$. Par conséquent l'erreur de consistance de la seconde équation est

$$CU^n_{2,j} = O(\Delta t + \Delta x).$$

On définit l'erreur de consistance associée aux autres équations $CU^n_{i,j}$ avec $i > 2$

$$CU^n_{i,j} = \frac{U_i(x_j, t_{n+1}) - U_i(x_j, t_n)}{\Delta t} + \frac{1}{\varepsilon}\frac{(A\mathbf{U})_{i,j+\frac{1}{2}} - (A\mathbf{U})_{i,j-\frac{1}{2}}}{\Delta x} = -\frac{\sigma}{\varepsilon^2}U_i(x_j, t_n),\ i > 2\ ,$$

avec

$$(A\mathbf{U})_{i,j+\frac{1}{2}} = \left(\frac{(A^{'}\mathbf{U})_i(x_j, t_n) + (A^{'}\mathbf{U})_i(x_{j+1}, t_n)}{2} + \frac{U_i(x_j, t_n) - U_i(x_{j+1}, t_n)}{2}\right).$$

Le même type de calcul que celui utilisé pour les autres démonstrations de consistance montre que

$$CU_{i,j} = \partial_t U_i(x_j, t_n) + O(\Delta t) + \frac{1}{\varepsilon}\partial_x(A^{'}\mathbf{U})_i(x_j, t_n) + \frac{\Delta x^2}{\varepsilon}\partial_{xx}(A^{'}\mathbf{U})_i(x_j, t_n) +$$

$$\frac{\Delta x}{\varepsilon}\partial_{xx}U_i(x_j, t_n) + \frac{\Delta x^3}{\varepsilon}\partial_{xx}U_i(x_j, t_n) + \frac{\sigma}{\varepsilon^2}U_i(x_j, t_n),$$

pour $i > 2$. Étant donnée la structure des matrices, en remarquant $A_i = A^{'}_i$ pour $i > 2$, on simplifie (4.51)

$$CU^n_{i,j} = O(\Delta t) + \frac{\Delta x^2}{\varepsilon}\partial_{xx}(A^{'}\mathbf{U})_i(x_j, t_n) - \frac{\Delta x}{\varepsilon}\partial_{xx}U_i(x_j, t_n) - \frac{\Delta x^3}{\varepsilon}\partial_{xx}U_i(x_j, t_n). \tag{4.51}$$

Par hypothèse, $\partial_{xx}U_i(x_j, t_n) = O(\varepsilon)$ pour $i > 2$. En outre, la structure de la matrice $A^{'}$ indique que $(A^{'}\mathbf{U})_i$ ne dépend que de U_i pour $i > 1$. Par conséquent, $\partial_{xx}(A^{'}\mathbf{U})_i = O(\varepsilon)$. On obtient donc l'erreur de consistance suivante

$$CU^n_{i,j} = O(\Delta t + \Delta x).$$

□

Cette démonstration justifie l'utilisation de la décomposition proposée afin de construire un schéma « asymptotic preserving » . En effet, les dérivées spatiales et les discrétisations des équations associées aux inconnues U_i avec $i > 2$ ne dépendent pas de la première inconnue. Par conséquent, l'analyse asymptotique montre que ces mêmes dérivées spatiales sont homogènes à ε, ce qui permet de compenser le coefficient de viscosité numérique homogène à $\frac{\Delta x}{\varepsilon}$.

Remarque 4.3.2. *Le schéma ainsi présenté est consistant uniformément en ε (voir (4.49)-(4.50)). Cependant, comme pour le schéma de Jin-Levermore (1.35) le schéma (4.47)-(4.48) ne préserve pas à priori le principe du maximum. On pourrait multiplier les flux du schéma utilisé pour discrétiser la partie « non diffusive » par un coefficient M bien choisi afin d'obtenir un schéma stable en norme L^∞. L'analyse précédente montre cependant que cette manipulation n'est pas nécessaire pour obtenir la consistance dans tous les régimes.*

4.3.3 Décomposition « diffusive - non diffusive » : application au modèle P_3 en 2D

On propose de valider la décomposition « diffusive - non diffusive » en 2D à l'aide du modèle P_3 décrit dans [Bru05]. Initialement, c'est un système symétrique de $\mathbb{R}^{16}$. Afin de réduire le coût de calcul, on propose la simplification décrite dans la section 1.2, ainsi que le changement de variable nécessaire pour obtenir un système symétrique. Le modèle P_3 ainsi obtenu, est défini par

$$\partial_t \mathbf{U} + \frac{1}{\varepsilon} A_1 \partial_x \mathbf{U} + \frac{1}{\varepsilon} A_2 \partial_y \mathbf{U} = -\frac{\sigma}{\varepsilon^2} R\mathbf{U}, \text{ avec } \mathbf{U} \in \mathbb{R}^{10}, \tag{4.52}$$

avec

$$A_1 = \begin{pmatrix}
0 & \sqrt{\frac{1}{3}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\sqrt{\frac{1}{3}} & 0 & \sqrt{\frac{4}{15}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \sqrt{\frac{4}{15}} & 0 & \sqrt{\frac{9}{35}} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \sqrt{\frac{9}{35}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \sqrt{\frac{1}{5}} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \sqrt{\frac{1}{5}} & 0 & \sqrt{\frac{8}{35}} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \sqrt{\frac{8}{35}} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{1}{7}} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{1}{7}} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}, \tag{4.53}$$

$$A_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & \sqrt{\frac{1}{3}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sqrt{\frac{1}{5}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\sqrt{\frac{1}{15}} & 0 & \sqrt{\frac{6}{35}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\sqrt{\frac{3}{35}} & 0 & 0 & 0 & 0 \\ \sqrt{\frac{1}{3}} & 0 & -\sqrt{\frac{1}{15}} & 0 & 0 & 0 & 0 & -\sqrt{\frac{1}{5}} & 0 & 0 \\ 0 & \sqrt{\frac{1}{5}} & 0 & -\sqrt{\frac{3}{35}} & 0 & 0 & 0 & 0 & -\sqrt{\frac{1}{7}} & 0 \\ 0 & 0 & \sqrt{\frac{6}{35}} & 0 & 0 & 0 & 0 & \sqrt{\frac{1}{70}} & 0 & 0 \\ 0 & 0 & 0 & 0 & -\sqrt{\frac{1}{5}} & 0 & \sqrt{\frac{1}{70}} & 0 & 0 & -\sqrt{\frac{3}{14}} \\ 0 & 0 & 0 & 0 & 0 & -\sqrt{\frac{1}{7}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\sqrt{\frac{3}{14}} & 0 & 0 \end{pmatrix}. \tag{4.54}$$

et la matrice associée au terme source directement diagonale

$$R = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.55}$$

Puisque R est déjà diagonale, il ne reste plus qu'à scinder ce modèle en un système du type équation de la chaleur hyperbolique qui correspond exactement au modèle P_1 et un système résiduel. La matrice $P_{1,x}$ de la décomposition « diffusive - non diffusive » est définie par $P_{1,x,12} = P_{1,x,21} = \frac{1}{\sqrt{3}}$ et les autres termes sont égaux à zéro.
La matrice $P_{1,y}$ de la décomposition « diffusive - non diffusive » est définie par $P_{1,y,14} = P_{1,y,41} = \frac{1}{\sqrt{3}}$ et les autres termes sont égaux à zéro.
Les matrices du système de la phase « non diffusive » sont définies par $A_1^{'} = A_1 - P_{1,x}$ et $A_2^{'} = A_2^{*} - P_{1,y}$. Afin d'éviter le calcul des valeurs, on propose d'utiliser un schéma de Rusanov pour la partie dite « non-diffusive » avec $S_{jk} = 1$ comme coefficient de viscosité numérique. Cette majoration de la valeur propre maximale est d'ailleurs valide pour tous les modèles P_n.

Définition 4.3.1. *Le schéma numérique pour le modèle P_3 est*

$$
\left\{
\begin{array}{l}
|\Omega_j|\,\partial_t \widetilde{U}_{1,j} + \dfrac{1}{\sqrt{3}\varepsilon} F_{\text{div}} = 0, \\
|\Omega_j|\,\partial_t \widetilde{U}_{2,j} + \dfrac{1}{\sqrt{3}\varepsilon} \mathbf{F}^x_{\text{grad}} + \dfrac{1}{\varepsilon}\displaystyle\sum_k l_{jk}\left(\frac{1}{2}(G\mathbf{n}_{jk}(\widetilde{\mathbf{U}}_j + \widetilde{\mathbf{U}}_k))_2 - \frac{1}{2}(\widetilde{U}_{2,k} - \widetilde{U}_{2,j})\right) = -S_j \widetilde{U}_{2,j}, \\
|\Omega_j|\,\partial_t \widetilde{U}_{3,j} + \dfrac{1}{\varepsilon}\displaystyle\sum_k l_{jk}\left(\frac{1}{2}(G\mathbf{n}_{jk}(\widetilde{\mathbf{U}}_j + \widetilde{\mathbf{U}}_k))_3 - \frac{1}{2}(\widetilde{U}_{3,k} - \widetilde{U}_{3,j})\right) = -\,|\Omega_j|\,\dfrac{\sigma}{\varepsilon^2}\widetilde{U}_{3,j}, \\
|\Omega_j|\,\partial_t \widetilde{U}_{4,j} + \dfrac{1}{\sqrt{3}\varepsilon} \mathbf{F}^y_{\text{grad}} + \dfrac{1}{\varepsilon}\displaystyle\sum_k l_{jk}\left(\frac{1}{2}(G\mathbf{n}_{jk}(\widetilde{\mathbf{U}}_j + \widetilde{\mathbf{U}}_k))_4 - \frac{1}{2}(\widetilde{U}_{4,k} - \widetilde{U}_{4,j})\right) = -S_j \widetilde{U}_{4,j}, \\
|\Omega_j|\,\partial_t \widetilde{U}_{5,j} + \dfrac{1}{\varepsilon}\displaystyle\sum_k l_{jk}\left(\frac{1}{2}(G\mathbf{n}_{jk}(\widetilde{\mathbf{U}}_j + \widetilde{\mathbf{U}}_k))_5 - \frac{1}{2}(\widetilde{U}_{5,k} - \widetilde{U}_{5,j})\right) = -\,|\Omega_j|\,\dfrac{\sigma}{\varepsilon^2}\widetilde{U}_{5,j}, \\
|\Omega_j|\,\partial_t \widetilde{U}_{6,j} + \dfrac{1}{\varepsilon}\displaystyle\sum_k l_{jk}\left(\frac{1}{2}(G\mathbf{n}_{jk}(\widetilde{\mathbf{U}}_j + \widetilde{\mathbf{U}}_k))_6 - \frac{1}{2}(\widetilde{U}_{6,k} - \widetilde{U}_{6,j})\right) = -\,|\Omega_j|\,\dfrac{\sigma}{\varepsilon^2}\widetilde{U}_{6,j}, \\
|\Omega_j|\,\partial_t \widetilde{U}_{7,j} + \dfrac{1}{\varepsilon}\displaystyle\sum_k l_{jk}\left(\frac{1}{2}(G\mathbf{n}_{jk}(\widetilde{\mathbf{U}}_j + \widetilde{\mathbf{U}}_k))_7 - \frac{1}{2}(\widetilde{U}_{7,k} - \widetilde{U}_{7,j})\right) = -\,|\Omega_j|\,\dfrac{\sigma}{\varepsilon^2}\widetilde{U}_{7,j}, \\
|\Omega_j|\,\partial_t \widetilde{U}_{8,j} + \dfrac{1}{\varepsilon}\displaystyle\sum_k l_{jk}\left(\frac{1}{2}(G\mathbf{n}_{jk}(\widetilde{\mathbf{U}}_j + \widetilde{\mathbf{U}}_k))_8 - \frac{1}{2}(\widetilde{U}_{8,k} - \widetilde{U}_{8,j})\right) = -\,|\Omega_j|\,\dfrac{\sigma}{\varepsilon^2}\widetilde{U}_{8,j}, \\
|\Omega_j|\,\partial_t \widetilde{U}_{9,j} + \dfrac{1}{\varepsilon}\displaystyle\sum_k l_{jk}\left(\frac{1}{2}(G\mathbf{n}_{jk}(\widetilde{\mathbf{U}}_j + \widetilde{\mathbf{U}}_k))_9 - \frac{1}{2}(\widetilde{U}_{9,k} - \widetilde{U}_{9,j})\right) = -\,|\Omega_j|\,\dfrac{\sigma}{\varepsilon^2}\widetilde{U}_{9,j}, \\
|\Omega_j|\,\partial_t \widetilde{U}_{10,j} + \dfrac{1}{\varepsilon}\displaystyle\sum_k l_{jk}\left(\frac{1}{2}(G\mathbf{n}_{jk}(\widetilde{\mathbf{U}}_j + \widetilde{\mathbf{U}}_k))_{10} - \frac{1}{2}(\widetilde{U}_{10,k} - \widetilde{U}_{10,j})\right) = -\,|\Omega_j|\,\dfrac{\sigma}{\varepsilon^2}\widetilde{U}_{10,j},
\end{array}
\right. \tag{4.56}
$$

où la matrice du schéma de Rusanov est définie par $G\mathbf{n}_{jk} = A_1 n^x_{jk} + A_2 n^y_{jk}$, les flux associés au schéma « asymptotic preserving » pour l'équation de la chaleur hyperbolique sont F_{div}, $\mathbf{F}_{grad}$ et la discrétisation du terme source associé au schéma « asymptotic preserving » est S_j.

Les schémas « asymptotic preserving » présentés dans les chapitres deux et trois sont construits pour l'équation de la chaleur hyperbolique. Les vitesses d'ondes et l'état stationnaire du système associés à la partie « diffusive » étant différents de ceux associés à l'équation de la chaleur hyperbolique, il faut dériver à nouveau le schéma « asymptotic preserving » pour le système associé à la partie « diffusive ». Dans ce cas cela revient à multiplier σ par dans $\sqrt{3}$ dans les flux « asymptotic preserving »

Un développement de Hilbert permet de montrer, comme dans le cas de l'équation de la chaleur hyperbolique, que ce schéma tend vers un schéma de diffusion limite (celui associé au schéma « asymptotic preserving » utilisé). Le développement de Hilbert étant du même type que les précédents, on ne le détaille pas. En effet, on montre rapidement que tous les termes d'ordre zéro dans le développement pour les moments d'ordre supérieur à 1 sont nuls.

4.4 Applications aux modèles S_n

Dans cette section, on propose de montrer que les modèles S_n en dimension deux satisfont les hypothèses de la définition (4.1.1) et les hypothèses (H_2)-(H_3) nécessaires à l'utilisation de notre algorithme.

4.4.1 Modèles S_n

Pour résoudre les équations de transport (équation de la neutronique, équation de transfert radiatif, etc) une alternative aux modèles P_n est la méthode des ordonnées discrètes qui permet de construire

les modèles S_n. La méthode consiste à discrétiser le terme de « scattering » à l'aide d'une formule de quadrature et à écrire l'équation de transport (4.44) pour chaque vitesse discrète. La fonction de distribution discrète $f_i = f(\mathbf{\Omega}_i)$ associée à la vitesse discrète $\mathbf{\Omega}_i$ satisfait l'équation suivante :

$$\partial_t f_i + \frac{1}{\varepsilon}\mathbf{\Omega}_i.\nabla f_i = -\frac{\sigma}{\varepsilon^2}(f_i - \sum_j f_j w_j). \tag{4.57}$$

avec les vitesses discrètes $\mathbf{\Omega}_i$ et les poids de quadratures w_i tels que

$$\left\{ \sum_i w_i = 1, \quad \sum_i w_i \mathbf{\Omega}_i = \mathbf{0}, \quad \sum_i w_i \mathbf{\Omega}_i \otimes \mathbf{\Omega}_i = D\widehat{I}_d \right.$$

avec $D = \frac{1}{3}$ si les vitesses appartiennent à S^2 et $D = \frac{1}{2}$ si les vitesses appartiennent à S^1. Le système défini par l'ensemble de ces équations admet pour limite de diffusion :

$$\partial_t E - \operatorname{div}\left(\frac{\sum_i w_i \mathbf{\Omega}_i \otimes \mathbf{\Omega}_i}{\sigma}\nabla E\right) = 0, \tag{4.58}$$

avec $E = \sum_i w_i f_i$. On souhaite donc montrer, comme pour les modèles P_n, que les modèles S_n rentrent dans le cadre introduit par la définition (4.1.1) et satisfont les hypothèses (H_2) et (H_3) nécessaire à l'écriture de la décomposition « diffusive - non diffusive » .

Proposition 4.7. *Les modèles S_n s'écrivent de la façon suivante :*

$$\partial_t \mathbf{U} + \frac{1}{\varepsilon}A_1\partial_x \mathbf{U} + \frac{1}{\varepsilon}A_2\partial_y \mathbf{U} = -\frac{\sigma}{\varepsilon^2}R\mathbf{U},$$

avec $U_i = \sqrt{w_i} f_i$ et $R = \sqrt{\mathbf{w}} \otimes \sqrt{\mathbf{w}} - \widehat{I}_d$. Le vecteur $\sqrt{\mathbf{w}}$ est composé de la racine des poids w_i.
Ces systèmes satisfont les propriétés suivantes

- $\dim \operatorname{Ker} R = 1$,
- *A_1 et A_2 sont diagonales,*
- *0 est valeur propre de R de multiplicité 1 et de vecteur propre $\mathbf{E}_1 = (\sqrt{w_1}, ..., \sqrt{w_n})$,*
- *1 est valeur propre de R de multiplicité $n-1$,*
- *La matrice R est diagonalisable,*
- *La matrice R est symétrique.*
- *$A_1\mathbf{E}_1 = a\mathbf{E}_2$, $A_2\mathbf{E}_1 = a\mathbf{E}_3$ avec $\mathbf{E}_2$, $\mathbf{E}_3$ des vecteurs propres de R associés à la valeur propre 1 de R.*

Démonstration. On commence par démontrer que 1 est valeur propre de multiplicité $n-1$ en partant du système. A cette fin on remarque que la matrice $\sqrt{\mathbf{w}}\otimes\sqrt{\mathbf{w}} - \widehat{I}_d$ correspond à la projection orthogonal dans l' hyperplan orthogonal à $\sqrt{\mathbf{w}}$. Le vecteur $\sqrt{\mathbf{w}}$ correspond au vecteur composé de la racine des poids de quadrature. Cette projection dans l'hyperplan orthogonal à $\sqrt{\mathbf{w}}$ est égale à $\widehat{I}_d$. Par conséquent 1 est valeur propre de multiplicité $n-1$ de la matrice R . La projection dans l'espace de dimension un généré par $\sqrt{\mathbf{w}}$ est identiquement nulle. La matrice R admet donc 0 comme valeur propre de vecteur propre $\mathbf{E}_1 = \sqrt{\mathbf{w}}$. Dans un second temps, on propose de montrer que la dernière propriété est vérifiée. La symétrie des vitesse discrètes $\sum_i w_i \mathbf{\Omega}_i = \mathbf{0}$ implique que $(A_1\sqrt{\mathbf{w}}, \sqrt{\mathbf{w}}) = 0$ et $(A_2\sqrt{\mathbf{w}}, \sqrt{\mathbf{w}}) = 0$. Par conséquent $A\mathbf{E}_1 \in \operatorname{Ker}(R)^\perp$ et $A\mathbf{E}_2 \in \operatorname{Ker}(R)^\perp$. En définissant les vecteurs propres $\mathbf{E}_2$, $\mathbf{E}_3$ de la matrice R associée à la valeurs propres 1 :

$$\mathbf{E}_2 = \frac{1}{\sqrt{\sum_i w_i \Omega_i^{x,2}}}\begin{pmatrix} \Omega_1^x \\ \vdots \\ \Omega_n^x \end{pmatrix}, \ \mathbf{E}_3 = \frac{1}{\sqrt{\sum_i \sqrt{w_i}\Omega_i^{y,2}}}\begin{pmatrix} \Omega_1^y \\ \vdots \\ \Omega_n^y \end{pmatrix}.$$

on obtient $A_1\mathbf{E}_1 = a\mathbf{E}_2$, $A_2\mathbf{E}_1 = a\mathbf{E}_3$ avec $a = \sqrt{\sum_i w_i\Omega_i^{x,2}} = \sqrt{\sum_i w_i\Omega_i^{y,2}}$. L'égalité $\sqrt{\sum_i w_i\Omega_i^{x,2}} = \sqrt{\sum_i w_i\Omega_i^{y,2}}$ vient de la construction de la formule de quadrature qui impose

$$\sum_i w_i \mathbf{\Omega}_i \otimes \mathbf{\Omega}_i = D\widehat{I}_d$$

avec $D = \frac{1}{3}$ ou $D = \frac{1}{2}$, afin d'obtenir la limite de diffusion de l'équation de transport. □

Remarque 4.4.1. *En dimension un, les modèles S_n (avec les points de quadrature de Gauss-Legendre) diagonalisés dans la base de la matrice R sont égaux aux modèles P_n.*

4.4.2 Décomposition « diffusive - non diffusive » : application au modèle S_2 en 2D

Le schéma est simple, il suffit d'utiliser un des schémas « asymptotic preserving » (par exemple les schémas JL-(a), JL-(b) où P_1-MPFA) pour la partie « diffusive » et un schéma aux arêtes, upwind ou Rusanov pour la partie restante. On propose de construire le schéma pour ce modèle S_2 en utilisant un schéma upwind pour la partie « non diffusive » . On commence par définir le modèle S_2

$$\partial_t \mathbf{U} + \frac{1}{\varepsilon} A_1 \partial_x \mathbf{U} + \frac{1}{\varepsilon} A_2 \partial_y \mathbf{U} = -\frac{\sigma}{\varepsilon^2} R\mathbf{U}, \tag{4.59}$$

avec

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad R = \begin{pmatrix} \frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{3}{4} \end{pmatrix}. \tag{4.60}$$

La limite de diffusion est donnée par

$$\partial_t E - \operatorname{div}\left(\frac{1}{2\sigma} E\right) = 0$$

et $E = \dfrac{1}{4}\displaystyle\sum_{i=1}^{4} U_i$. La matrice de passage Q composée des vecteurs propres de R, et D la matrice diagonale sont données par

$$Q = \begin{pmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{2} \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Soit $\mathbf{E}_1 = \frac{1}{2}(1,1,1,1)$ le vecteur propre normé associé à la valeur propre 0, et $\mathbf{E}_2 = \frac{1}{\sqrt{2}}(1,0,-1,0)$, $\mathbf{E}_3 = \frac{1}{\sqrt{2}}(0,1,0,-1)$, $\mathbf{E}_4 = \frac{1}{2}(1,-1,1-1)$ les vecteurs propres normés associés à la valeur propre 1. Il est trivial de vérifier que $A_1\mathbf{E}_1 = \frac{1}{\sqrt{2}}\mathbf{E}_2$, $A_2\mathbf{E}_1 = \frac{1}{\sqrt{2}}\mathbf{E}_3$.
A présent, on applique un changement de variable. On pose $\mathbf{V} = Q^t\mathbf{U}$. Le nouveau système est donc

$$\partial_t \mathbf{V} + \frac{1}{\varepsilon} A_1' \partial_x \mathbf{V} + \frac{1}{\varepsilon} A_2' \partial_y \mathbf{V} = -\frac{\sigma}{\varepsilon^2} D\mathbf{V}, \tag{4.61}$$

avec

$$A_1^{'} = Q^t A_1 Q = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 \end{pmatrix}, \quad A_2^{'} = Q^t A_2 Q = \begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \end{pmatrix},$$

et

$$\mathbf{V} = \begin{pmatrix} \frac{1}{2}(U_1 + U_2 + U_3 + U_4) \\ \frac{1}{\sqrt{2}}(U_1 - U_3) \\ \frac{1}{\sqrt{2}}(U_2 - U_4) \\ \frac{1}{2}(U_1 + U_3 - U_2 - U_4) \end{pmatrix}.$$

Ces matrices se décomposent de la façon suivante

$$A_1^{'} = P_{1,x} + A_1^*, \quad A_2^{'} = P_{1,y} + A_2^*.$$

La matrice $P_{1,x}$ est définie par $P_{1,x,12} = \frac{1}{\sqrt{2}}$, $P_{1,x,21} = \frac{1}{\sqrt{2}}$ et $P_{1,x,ij} = 0$ pour $(i,j) \neq (1,2)$ et $(i,j) \neq (2,1)$.
La matrice $P_{1,y}$ est définie par $P_{1,y,13} = \frac{1}{\sqrt{2}}$, $P_{1,y,31} = \frac{1}{\sqrt{2}}$ et $P_{1,y,ij} = 0$ pour $(i,j) \neq (1,3)$ et $(i,j) \neq (3,1)$.
A partir de là on peut introduire le schéma upwind à l'aide des parties positive et négative de $G^*\mathbf{n} = A_1^* n^x + A_2^* n^y$.

Proposition 4.8. *La matrice $G^*\mathbf{n}$ a pour valeur propre $(0, 0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ et pour matrice de passage*

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ n^x & 0 & n^y & n^x \\ -n^y & 0 & n^x & -n^y \\ -||\mathbf{n}|| & 0 & 0 & ||\mathbf{n}|| \end{pmatrix}.$$

Démonstration. La matrice $G^*\mathbf{n}$ s'écrit

$$G^*\mathbf{n} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & n^x \\ 0 & 0 & 0 & -n^y \\ 0 & n^x & -n^y & 0 \end{pmatrix}.$$

Son polynôme caractéristique se réduit à

$$\mathrm{Det}(G^*\mathbf{n} - \lambda \widehat{I}_d) = \lambda^2(\lambda^2 - \frac{1}{2})$$

Le spectre ρ de la matrice G^* est donné par $\rho(G^*\mathbf{n}) = (0, 0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. Pour conclure il suffit de calculer les vecteurs propres. □

Les parties positive et négative de $G\mathbf{n}$ nécessaires à la construction du schéma upwind sont données par

$$G^+\mathbf{n} = \frac{1}{2\sqrt{2}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & n^x n^x & -n^x n^y & n^x \\ 0 & -n^x n^y & n^y n^y & -n^y \\ 0 & n^x & -n^y & 1 \end{pmatrix}, \quad G^-\mathbf{n} = \frac{1}{2\sqrt{2}} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -n^x n^x & n^x n^y & n^x \\ 0 & n^x n^y & -n^y n^y & -n^y \\ 0 & n^x & -n^y & -1 \end{pmatrix}.$$

Une fois le schéma upwind construit pour la partie « non diffusive » de la décomposition, il suffit de choisir un schéma « asymptotic preserving » pour la partie « diffusive » . On obtient donc le schéma numérique suivant :

Définition 4.4.1 (Schéma numérique pour le modèle S_2)**.** *On applique le changement de variable*

$$\mathbf{V} = Q^t \mathbf{U}.$$

On applique ensuite le schéma numérique semi-discret :

$$\left\{ \begin{array}{l} |\Omega_j| \, \partial_t V_{1,j} + \dfrac{1}{\sqrt{2}\varepsilon} F_{\text{div}} = 0, \\ |\Omega_j| \, \partial_t V_{2,j} + \dfrac{1}{\varepsilon} \left(\dfrac{1}{\sqrt{2}} \mathbf{F}^x_{\text{grad}} + \displaystyle\sum_k l_{jk} (G\mathbf{n}^+_{jk} \mathbf{V}_j)_2 + \sum_k l_{jk} (G\mathbf{n}^-_{jk} \mathbf{V}_k)_2 \right) = -S_j V_{2,j}, \\ |\Omega_j| \, \partial_t V_{3,j} + \dfrac{1}{\varepsilon} \left(\dfrac{1}{\sqrt{2}} \mathbf{F}^y_{\text{grad}} + \displaystyle\sum_k l_{jk} (G\mathbf{n}^+_{jk} \mathbf{V}_j)_3 + \sum_k l_{jk} (G\mathbf{n}^-_{jk} \mathbf{V}_k)_3 \right) = -S_j V_{3,j}, \\ |\Omega_j| \, \partial_t V_{4,j} + \dfrac{1}{\varepsilon} \left(\displaystyle\sum_k l_{jk} (G\mathbf{n}^+_{jk} \mathbf{V}_j)_4 + \sum_k l_{jk} (G\mathbf{n}^-_{jk} \mathbf{V}_k)_4 \right) = - |\Omega_j| \, \dfrac{\sigma}{\varepsilon^2} V_{4,j}. \end{array} \right. \tag{4.62}$$

où les matrices $G\mathbf{n}^-_{jk}$, $G\mathbf{n}^+_{jk}$ représentent les parties négative et positive de $G\mathbf{n}_{jk}$. On note les flux numériques et la discrétisation du terme source associés au schéma « asymptotic preserving » pour l'équation de la chaleur hyperbolique : F_{div}, $\mathbf{F}_{\text{grad}}$ et S_j.
Pour finir

$$\mathbf{U} = Q\mathbf{V}.$$

Les schémas « asymptotic preserving » présentés dans le chapitres deux et trois sont construits pour l'équation de la chaleur hyperbolique. Les vitesses d'ondes et l'état stationnaire du système associés à la partie « diffusive » étant différents de ceux associés à l'équation de la chaleur hyperbolique, il faut dériver à nouveau le schéma « asymptotic preserving » pour le système associé à la partie « diffusive ». Dans ce cas cela revient à multiplier σ par dans $\sqrt{2}$ dans les flux « asymptotic preserving » .

4.5 Discrétisation en temps

La méthode de décomposition précédemment introduite permet d'obtenir des discrétisations spatiales valides en régime de diffusion. La stabilité du schéma complet obtenu dépend, elle, des conditions de stabilité des schémas utilisés pour discrétiser chaque système de la décomposition. Certains schémas « asymptotic preserving » pour la partie diffusive ainsi que les schémas classiques pour la partie non diffusive admettent des conditions CFL dépendantes de ε. Pour régler ce problème on peut utiliser une version implicite de ces schémas ou bien, comme dans le cas de l'équation de la chaleur hyperbolique, construire des schémas semi-implicites stables sous une condition CFL optimale. On commencera par étudier la discrétisation implicite de la décomposition « diffusive non-diffusive » . Par la suite on proposera une modification de la décomposition afin d'obtenir des schémas semi-implicites stables sous une condition CFL indépendante de ε.

4.5.1 Discrétisation implicite

Dans un premier temps, on propose d'étudier la stabilité de la discrétisation implicite en temps associée à la « diffusive - non diffusive » d'un système de Friedrichs. On résout le système suivant

$$\partial_t \mathbf{U} + \frac{1}{\varepsilon} A_1 \partial_x \mathbf{U} + \frac{1}{\varepsilon} A_2 \partial_y \mathbf{U} = -\frac{\sigma}{\varepsilon^2} R\mathbf{U}. \tag{4.63}$$

Afin d'obtenir une discrétisation « asymptotic preserving » on applique la décomposition « diffusive-non diffusive » Soit Q la matrice composée des vecteurs propres de R, $\mathbf{V} = Q^t\mathbf{U}$, $A_1^{'} = Q^t A_1 Q$ et $A_2^{'} = Q^t A_2 Q$. Le système associé à la décomposition « diffusive-non diffusive » est donné par

$$\partial_t \mathbf{V} + \frac{1}{\varepsilon}\left(P_{1,x}\partial_x \mathbf{V} + P_{1,y}\partial_y \mathbf{V}\right) + \frac{1}{\varepsilon}\left(A_1^*\partial_x \mathbf{V} + A_2^*\partial_y \mathbf{V}\right) = -\frac{\sigma}{\varepsilon^2} D\mathbf{V}, \tag{4.64}$$

avec $P_{1,x}$, $P_{1,y}$ des matrices définies par (4.39) et $A_1^* = A_1^{'} - P_{1,x}$, $A_1^* = A_1^{'} - P_{1,y}$. La partie diffusive est décrite par le système

$$\partial_t \mathbf{V} + \frac{1}{\varepsilon}\left(P_{1,x}\partial_x \mathbf{V} + P_{1,y}\partial_y \mathbf{V}\right) = -\frac{\sigma}{\varepsilon^2} D^{'}\mathbf{V} \tag{4.65}$$

avec $D^{'}$ définie par $D_{22}^{'} = D_{33}^{'} = 1$ et $D_{ii}^{'} = 0$ $i \neq 2$ et $i \neq 3$. On discrétise le système (4.65) avec un schéma « asymptotic preserving » stable en norme L^2. On pose $\mathbf{V}_h$ le vecteur des inconnues défini par $\mathbf{V}_h = (\mathbf{V}_{j=1}, ..., \mathbf{V}_{j=n})$, M la matrice diagonale composée des volumes de mailles. Le schéma implicite pour le système (4.65) est défini par

$$M\mathbf{V}_h^{n+1} = M\mathbf{V}_h^n + \Delta t P_h \mathbf{V}_h^{n+1},$$

avec P_h la matrice de discrétisation spatiale, telle que $(\mathbf{X}, P_h\mathbf{X}) \leq 0$, ou $\mathbf{X} \in \mathbb{R}^{n\times n_c}$ et n_c correspond au nombre de mailles. La stabilité en norme L^2 du schéma implicite dérive de l'inégalité $(\mathbf{X}, P_h\mathbf{X}) \leq 0$. De la même façon on définit le système associé à la partie « non diffusive »

$$\partial_t \mathbf{V} + \frac{1}{\varepsilon}\left(A_1^*\partial_x \mathbf{V} + A_2^*\partial_y \mathbf{V}\right) = -\frac{\sigma}{\varepsilon^2} D^{''}\mathbf{V}, \tag{4.66}$$

avec $D^{''}$ définie par $D_{11}^{''} = D_{22}^{''} = D_{33}^{''} = 0$ et $D_{ii}^{''} = 1$ $i \geq 4$. Le système (4.66) est discrétisé avec un schéma de volumes finis classique comme le schéma upwind ou le schéma de Rusanov stable en norme L^2, défini par

$$M\mathbf{V}_h^{n+1} = M\mathbf{V}_h^n + \Delta t A_h \mathbf{V}_h^{n+1},$$

avec $(\mathbf{X}, A_h\mathbf{X}) \leq 0$.

Proposition 4.9. *Soit $\mathbf{V}_h$ le vecteur des inconnues et M la matrice diagonale composée des volumes de mailles. Soit P_h la matrice de la discrétisation spatiale du schéma « asymptotic preserving » pour le système (4.65) telle que $(\mathbf{X}, P_h\mathbf{X}) \leq 0$. Soit A_h la matrice de la discrétisation spatiale du schéma pour le système (4.66) telle que $(\mathbf{X}, A_h\mathbf{X}) \leq 0$. Le schéma implicite*

$$M\mathbf{V}_h^{n+1} = M\mathbf{V}_h^n + \Delta t P_h \mathbf{V}_h^{n+1} + \Delta t A_h \mathbf{V}_h^{n+1},$$

pour le système (4.64), est stable en norme $L^2(D)$.

Démonstration. On part du schéma implicite

$$M\mathbf{V}_h^{n+1} = M\mathbf{V}_h^n + \Delta t P_h \mathbf{V}_h^{n+1} + \Delta t A_h \mathbf{V}_h^{n+1}.$$

On multiplie par $\mathbf{V}^{n+1}$ pour obtenir

$$(M\mathbf{V}_h^{n+1}, \mathbf{V}_h^{n+1}) = (M\mathbf{V}_h^n, \mathbf{V}_h^{n+1}) + \Delta t(P_h\mathbf{V}_h^{n+1}, \mathbf{V}_h^{n+1}) + \Delta t(A_h\mathbf{V}_h^{n+1}, \mathbf{V}_h^{n+1}).$$

Pour conclure on utilise les inégalités $(\mathbf{V}_h^{n+1}, P_h\mathbf{V}_h^{n+1}) \leq 0$, $(\mathbf{V}_h^{n+1}, A_h\mathbf{V}_h^{n+1}) \leq 0$ et l'inégalité de Cauchy-Schwarz pour obtenir que

$$||\mathbf{V}^{n+1}||_{L^2} \leq ||\mathbf{V}^n||_{L^2}.$$

avec $||\mathbf{V}^{n+1}||_{L^2} = \sum_j |\Omega_j|(\mathbf{V}_j^{n+1}, \mathbf{V}_j^{n+1})$. □

4.5.2 Discrétisation semi-implicite

Afin d'obtenir une discrétisation temporelle non contrainte par ε, on peut aussi utiliser un schéma semi-implicite. En s'inspirant du travail réalisé dans le chapitre deux, on propose de construire une discrétisation semi-implicite de la décomposition « diffusive - non diffusive » stable sous une condition CFL indépendante de ε. Dans le chapitre un on a montré que le schéma de Jin-Levermore en 1D admet une CFL explicite

$$\Delta t \left(\frac{1}{\varepsilon \Delta x} + \frac{\sigma}{\varepsilon^2} \right) \leq 1. \tag{4.67}$$

Le schéma de Gosse-Toscani, lui, est stable en norme L^∞ sous la condition CFL

$$\Delta t \left(\frac{M}{\varepsilon \Delta x} + \frac{M\sigma}{\varepsilon^2} \right) \leq 1,$$

avec $M = \dfrac{2\varepsilon}{2\varepsilon + \sigma \Delta x}$. Cette condition CFL est équivalente à $\Delta t \left(\dfrac{1}{\varepsilon \Delta x} \right) \leq 1$. Cela vient du fait que le facteur M déraidit le schéma. Le principe est le même en dimension deux. Si on discrétise le terme source dans le schéma JL-(a) (2.98)-(2.100) on ne peut obtenir une CFL indépendante de ε. Par contre le schéma de Gosse-Toscani en dimension un ainsi que la reformulation du schéma JL-(b) (2.104) avec une discrétisation du terme source implicite sont stables sous la condition CFL suivante

$$\Delta t \left(\frac{1}{\varepsilon \Delta x + \frac{\sigma}{\Delta x^2}} \right) \leq 1. \tag{4.68}$$

Afin d'obtenir une condition CFL de stabilité du schéma indépendante de ε en pratique pour un système de Friedrichs, on propose de modifier les schémas upwind ou de Rusanov discrétisant la phase « non diffusive » afin que la condition CFL de ces schémas soit du même ordre que (4.68). Pour cela, on propose de « déraidir » le schéma à l'aide du coefficient M du même type que ceux qui interviennent en dimension un, puis à impliciter le terme source. Puisque le facteur M est discrétisé aux interfaces, il est aussi intéressant de discrétiser le coefficient de relaxation $\dfrac{\sigma}{\varepsilon^2}$ aux interfaces. Le rôle de la modification, à l'aide du coefficient M des schémas classiques pour la partie « non diffusive » est simplement de déraidir les schémas afin d'obtenir une CFL optimale. Par contre, les modifications introduites afin d'obtenir les schémas « asymptotic preserving » présentés dans les chapitres deux et trois contiennent des informations géométriques qui permettent d'obtenir un schéma limite consistant. Afin d'expliquer cette modification, on propose un schéma modifié pour le modèle S_2.

Définition 4.5.1 (Schéma « asymptotic preserving » semi-implicite pour le modèle S_2)**.** *On applique le changement de variable*

$$\mathbf{V}^n = Q^t \mathbf{U}^n.$$

On applique ensuite le schéma

$$\left\{\begin{array}{l} |\Omega_j| \dfrac{V_{1,j}^{n+1}-V_{1,j}^n}{\Delta t} + \dfrac{1}{\sqrt{2}\varepsilon} F_{\rm div} = 0 \\ |\Omega_j| \dfrac{V_{2,j}^{n+1}-V_{2,j}^n}{\Delta t} + \dfrac{1}{\varepsilon}\left(\dfrac{1}{\sqrt{2}} \mathbf{F}_{\rm grad}^x + \sum_k l_{jk} M_{jk} (G\mathbf{n}_{jk}^+ \mathbf{V}_j)_2^n + \sum_k l_{jk} M_{jk} (G\mathbf{n}_{jk}^- \mathbf{V}_k)_2^n \right) = -S_j V_{2,j}^{n+1}, \\ |\Omega_j| \dfrac{V_{3,j}^{n+1}-V_{3,j}^n}{\Delta t} + \dfrac{1}{\varepsilon}\left(\dfrac{1}{\sqrt{2}} \mathbf{F}_{\rm grad}^y + \sum_k l_{jk} M_{jk} (G\mathbf{n}_{jk}^+ \mathbf{V}_j)_3^n + \sum_k l_{jk} M_{jk} (G\mathbf{n}_{jk}^- \mathbf{V}_k)_3^n \right) = -S_j V_{3,j}^{n+1}, \\ |\Omega_j| \dfrac{V_{4,j}^{n+1}-V_{4,j}^n}{\Delta t} + \dfrac{1}{\varepsilon}\left(\sum_k l_{jk} M_{jk} (G\mathbf{n}_{jk}^+ \mathbf{V}_j)_4^n + \sum_k l_{jk} M_{jk} (G\mathbf{n}_{jk}^- \mathbf{V}_k)_4^n \right) \\ = - |\Omega_j| \left(\frac{1}{N(j)} \sum_k M_{jk} \frac{\sigma_{jk}}{\varepsilon^2} \right) V_{4,j}^{n+1}. \end{array}\right. \tag{4.69}$$

On note les flux numériques et la discrétisation du terme source associés au schéma « asymptotic preserving » pour l'équation de la chaleur hyperbolique : $F_{\rm div}$, $\mathbf{F}_{\rm grad}$ *et* S_j. *Le facteur* M_{jk} *est défini par* $M_{jk} = \frac{2\sqrt{2}\varepsilon}{2\sqrt{2}\varepsilon + 2\sigma_{jk}(d_{jk}+d_{kj})}$.
Pour finir

$$\mathbf{U}^{n+1} = Q\mathbf{V}^{n+1}.$$

4.6 Résultats numériques

Dans cette dernière partie, on expose des résultats numériques pour les deux systèmes précédemment introduits, le modèle P_3 et le modèle S_2. La partie « diffusive » de la décomposition est discrétisée à l'aide des schémas JL-(a) (2.98)-(2.100) et JL-(b) (2.99)-(2.100). On peut aussi utiliser le schéma P_1-MPFA ou n'importe quel autre schéma « asymptotic preserving » valide pour l'équation de la chaleur hyperbolique.

4.6.1 Résultats numériques pour le modèle S_2

On résout le système S_2 défini par (4.59)-(4.60) à l'aide du schéma (4.62) ou les flux « asymptotic preserving » sont donnés par le schéma JL-(b) (2.99)-(2.100). On définit $E = \frac{1}{4}\sum_{i=1}^4 U_i$ avec U_i la $i$ème composante du vecteur $\mathbf{U}$.

Limite de diffusion

Soit $SF(t)$ la solution fondamentale de la chaleur avec un coefficient de diffusion de 1/2. Les conditions initiales sont données par $U_i(t=0,\mathbf{x}) = SF(0.01)$. La solution de diffusion est donnée par $U_i(t,\mathbf{x}) = SF(0.01+T_f)$ avec $T_f = 0.01$. A ce temps final la solution est négligeable au bord. On compare la solution de diffusion et la solution numérique du modèle S_2 pour la variable $E(t,\mathbf{x}) = \frac{1}{4}\sum_{i=1}^4 U_i(t,\mathbf{x})$.

Les tableaux (4.1)-(4.2)-(4.3)-(4.4)-(4.5)-(4.6) donnent les ordres de convergence du schéma en limite de diffusion pour plusieurs valeurs de ε et pour plusieurs maillages.
On remarque que l'ordre de convergence est proche de l'ordre de convergence du schéma limite. On remarque aussi que lorsque $\frac{\Delta x}{\varepsilon} = O(1)$ l'ordre de convergence descend lentement. Les explications possibles sont les mêmes que celles apportées dans la section résultats numériques du chapitre deux.

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	2.00	1.98	1.99	1.99
80-160	1.80	1.97	2.	2
160-320	1.69	1.97	2.01	2.01

TABLE 4.1 – Résultats de convergence pour le modèle S_2 sur maillage Cartésien

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	1.92	1.99	2.00	2.00
80-160	1.88	2.02	2.03	2.03
160-320	1.76	2.01	2.04	2.03

TABLE 4.2 – Résultats de convergence pour le modèle S_2 sur maillage Cartésien aléatoire

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	1.89	1.96	1.96	1.96
80-160	1.84	1.94	1.96	1.96
160-320	1.79	1.97	1.99	1.99

TABLE 4.3 – Résultats de convergence pour le modèle S_2 sur maillage « smooth »

Nombre de cellules/ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	1.89	1.96	1.96	1.96
80-160	1.84	1.94	1.96	1.96
160-320	1.79	1.97	2.00	1.99

TABLE 4.4 – Résultats de convergence pour le modèle S_2 sur maillage de Kershaw

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	1.98	2.02	2.02	2.02
80-160	1.91	1.99	2.00	2.00
160-320	1.83	2.01	2.01	2.01

TABLE 4.5 – Résultats de convergence pour le modèle S_2 sur maillage triangulaire régulier

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	1.65	1.65	1.65	1.65
80-160	1.39	1.38	1.38	1.38
160-320	1.26	1.25	1.25	1.25

TABLE 4.6 – Résultats de convergence pour le modèle S_2 sur maillage triangulaire aléatoire

Cas test de transport

Cas test 1

On initialise $U_1(\mathbf{x}, t = 0) = g(\mathbf{x}) = \chi_{[0.4,0.6]^2(x)}$ où χ est la fonction indicatrice. Les autres variables

sont initialisées à 0. On prend $\sigma = 0$ et $\varepsilon = 1$. La solution est alors donnée par $U_1(\mathbf{x}, t) = g(\mathbf{x} - \mathbf{a}t)$ avec $\mathbf{a} = (1, 0)$ et $U_i(t, \mathbf{x}) = 0$ pour $i > 1$. Comme la condition initiale est discontinue on s'attend à constater un ordre proche de 0.5 en norme L^1. On choisit comme temps final $T_f = 0.1$. Le tableau des

maillage ordre	40-80	80-160	160-320
maillage cartésien	0.45	0.48	0.51
maillage bruité quad.	0.47	0.48	0.50
maillage « smooth »	0.47	0.46	0.47
maillage triang régulier	0.48	0.48	0.48
maillage triang bruité	0.49	0.47	0.47
Kershaw	0.38	0.42	0.43

TABLE 4.7 – Ordre de convergence pour le cas test 1 de transport sur différents maillages.

ordres de convergence (4.7) montre que le schéma converge avec un ordre proche de l'ordre théorique de 0.5 pour une solution discontinue.

Remarque 4.6.1. *Le schéma ne préserve pas exactement la nullité des variables U_2, U_3 et U_4 en raison du changement de variable dû à la diagonalisation du système, qui lie les variables U_i entre elles.*

Cas test 2

Soit $G(\mathbf{x})$ une gaussienne. Les conditions initiales sont données par $U_i(\mathbf{x}, t = 0) = G(\mathbf{x})$ et les paramètres sont définis par $\sigma = 0$ et $\varepsilon = 1$. La solution correspond alors à l'advection de ces quatre Gaussienne aux quatre vitesses $(0, 1)$, $(0, -1)$, $(1, 0)$ et $(-1, 0)$. Le temps final est 0.2. On compare les solutions exacte et approchée pour la variable $E = \frac{1}{4}\sum_{i=1} U_i^4$. On constate que le schéma converge à

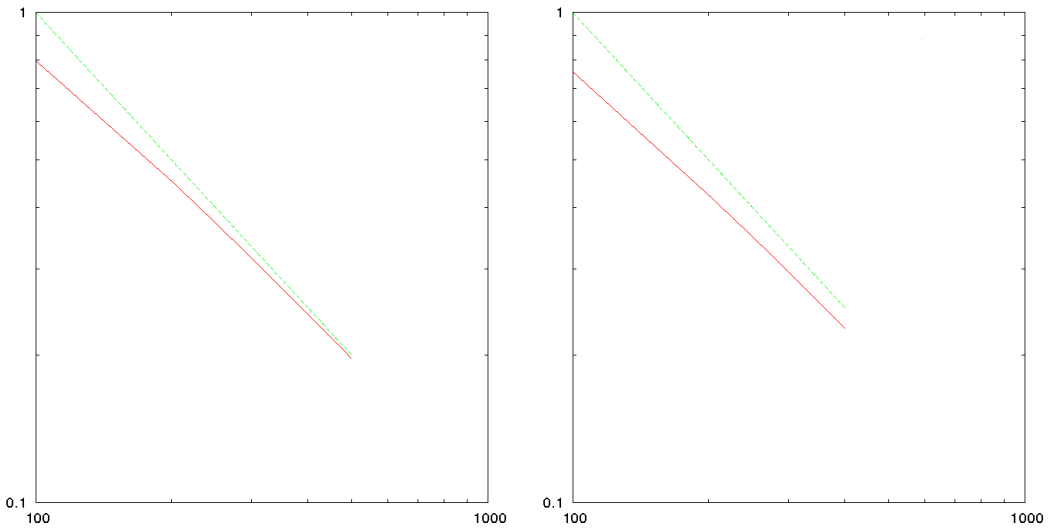


FIGURE 4.1 – En rouge la courbe de l'erreur numérique sur maillage Cartésien (à gauche) et pour maillage aléatoire (à droite). En vert la courbe de $1/h$.

l'ordre 1. Par contre on constate que le schéma est très diffusif pour ce cas test.

Cas test 3
On étudie la solution fondamentale du modèle S_2. Les données initiales sont $U_i(t=0,x) = \delta_{0.5,0.5}$ avec δ une masse de Dirac. Les paramètres sont définis par $\varepsilon = 1$ et $\sigma = 1$. La solution correspond à $U_i(t=0,x) = \delta_{0.5-v_{i,x}t,0.5-v_{i,y}t}$ avec $\mathbf{v}_1 = (0,1)$, $\mathbf{v}_2 = (0,-1)$, $\mathbf{v}_3 = (1,0)$ et $\mathbf{v}_1 = (-1,0)$. Le cas test est résolu sur maillage aléatoire. On remarque sur la figure (4.2) que les quatre fonctions de Dirac sont

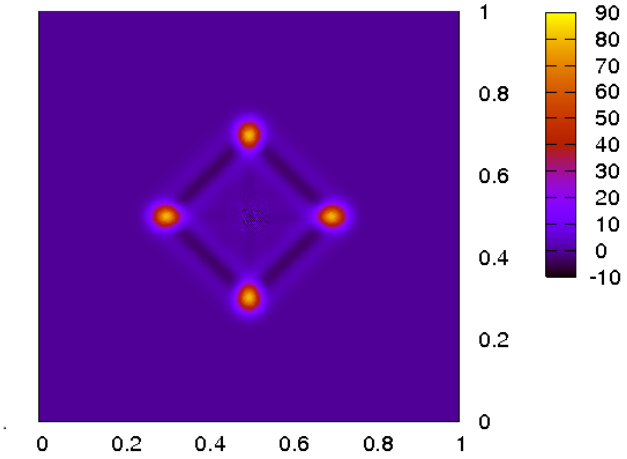


FIGURE 4.2 – Solution fondamentale. On trace la quantité $E(t,\mathbf{x}) = \sum_{i=1}^{4} U^i(t,\mathbf{x})$

bien advectées à la bonne vitesse. Cependant on remarque que la méthode est assez diffusive a cause du couplage numérique réalisé entre les équations par la diagonalisation du système.

4.6.2 Résultats numériques pour le modèle P_3

On résout le modèle P_3 défini par (4.52)-(4.53)-(4.54)-(4.55) à l'aide du schéma (4.69). Pour la partie « diffusive », on utilise le schéma JL-(b) (2.99)-(2.100).

Limite de diffusion

Soit $SF(t)$ la solution fondamentale de la chaleur avec un coefficient de diffusion de $\frac{1}{3}$. Les conditions initiales sont données par $U_1(\mathbf{x}, t=0) = SF(0.01)$ et $U_i(\mathbf{x}, t=0) = 0$ pour $i > 1$. La solution de diffusion est donnée par $U_1(\mathbf{x},t) = SF(0.01 + T_f)$ et $U_i(\mathbf{x},t) = 0$ pour $i > 1$. Le temps final est $T_f = 0.01$. À ce temps la solution est nulle au bord.

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	1.89	1.97	1.98	1.98
80-160	1.82	1.97	1.99	1.99
160-320	1.70	1.96	1.99	2.00

TABLE 4.8 – Résultats de convergence pour le modèle P_3 sur maillage Cartésien

Les tableaux de convergence (4.8)-(4.9)-(4.10)-(4.11)-(4.12)-(4.13) montrent que le schéma proposé converge avec les mêmes ordres de convergence que les schémas pour le modèle P_1. Les remarques

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	1.89	1.95	1.95	1.95
80-160	1.87	1.99	2.00	2.00
160-320	1.77	2.01	2.03	2.03

TABLE 4.9 – Résultats de convergence pour le modèle P_3 sur maillage Cartésien aléatoire

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	1.99	2.05	2.06	2.06
80-160	1.88	2.00	2.01	2.01
160-320	1.76	1.98	2.01	2.01

TABLE 4.10 – Résultats de convergence pour le modèle P_3 sur maillage "smooth"

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	1.86	1.88	x	1.92
80-160	1.85	1.94	x	1.95
160-320	1.83	1.97	x	1.99

TABLE 4.11 – Résultats de convergence pour le modèle P_3 sur maillage de Kershaw

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	1.98	2.01	2.02	2.02
80-160	1.92	2.00	2.00	2.00
160-320	1.81	1.99	2.00	2.00

TABLE 4.12 – Résultats de convergence pour le modèle P_3 sur maillage triangulaire régulier

Nombre de cellules /ε	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$\varepsilon = 10^{-7}$
40-80	1.72	1.74	1.73	1.73
80-160	1.42	1.42	1.43	1.45
160-320	1.29	1.27	1.27	1.27

TABLE 4.13 – Résultats de convergence pour le modèle P_3 sur maillage triangulaire aléatoire

exprimées précédemment lors de l'analyse des cas test de convergence en régime de diffusion pour l'équation de la chaleur hyperbolique où le modèle S_2 sont également valables pour ce cas test.

Solution fondamentale

Afin de valider le schéma dans le régime de transport, on propose le cas test suivant. Le domaine est défini par $D = [0, 2.5]^2$. La condition initiale est donnée par $U_1(0, \mathbf{x}) = \delta_{1.25,1.25}$, $U_i(0, \mathbf{x}) = 0$ pour $i \neq 0$. Pour finir $\varepsilon = 1$, $\sigma = 1$. La solution est donnée dans [HMc11]- [Bru02]. La solution fondamentale est composée des différentes ondes séparées par des fonctions régulières négatives en temps court et

positives en temps long. On résout ce cas test en utilisant le schéma JL-(b)-stabilisé (chapitre deux), présenté à la fin du second chapitre pour discrétiser l'équation de la chaleur hyperbolique et un schéma de Rusanov pour le système qui n'intervient pas dans la limite de diffusion. La vitesse maximale du modèle P_3 est approximativement 0.86. Le résultat numérique (4.3) montre que le schéma capture correctement la vitesse maximale ainsi que la dynamique de la solution. En outre on montre que la

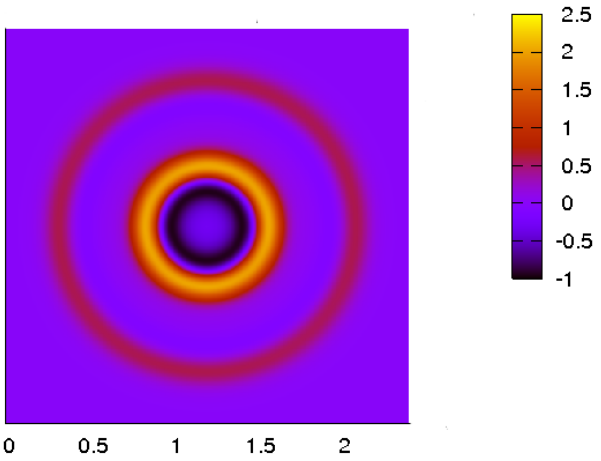


FIGURE 4.3 – On trace la première quantité du modèle P_3 en fonction de $\mathbf{x}$.

version stabilisée du schéma du schéma JL-(b) couplé avec la décomposition « diffusive - non diffusive » n'admet pas de mode parasite.

Cas test du Treillis

On reprend le cas test du Treillis, issu de la neutronique, détaillé pour le modèle P_1 dans le chapitre deux. On résout le modèle P_3 avec un terme source supplémentaire $-\sigma_a U_1 + Q$ sur la première équation, où σ_a est le coefficient d'absorption et Q est une source. Le domaine est un carré de taille 7. Le maillage est un pavage de carrés représentés sur la figure (2.16) introduite dans le chapitre deux. Dans le domaine blanc et le carré central (2.16) $\sigma = 1$ et $\sigma_a = 0$. Dans les autres carrés (2.16), les opacités sont définies par $\sigma = 0$ et $\sigma_a = 10$. Tous les moments sont initialement nuls et la source vaut $Q = 1$. Le temps final est de 3.2 secondes. La solution correspond aux solutions de référence donné dans [Bru02].

4.7 Synthèse

Dans le cas de modèles hyperboliques symétriques avec termes sources raides satisfaisant l'hypothèse (H_3) et admettant une limite de diffusion scalaire, la décomposition entre l'équation de la chaleur hyperbolique et un autre système qui n'intervient pas en limite de diffusion, permet d'obtenir un schéma « asymptotic preserving » . De plus, la difficulté liée à la consistance géométrique du schéma limite est entièrement reportée sur la discrétisation de l'équation de la chaleur hyperbolique. Dans le cadre de l'approximation de transport on montre que l'on peut utiliser cette décomposition pour des modèles simplifiés comme les systèmes P_n et S_n. Dans le cadre des systèmes P_n, cette décomposition est naturelle. En effet, en régime de diffusion, seuls les deux premiers moments forment le régime limite. Les autres moments sont négligeables par rapport aux deux premiers moments. Pour les modèles

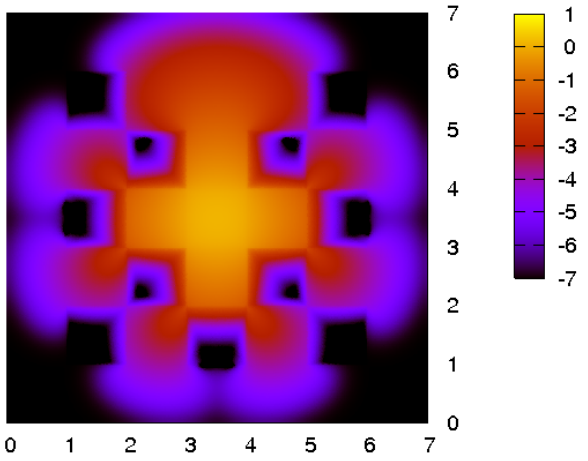


FIGURE 4.4 – Cas test Treillis pour le modèle P_3. On trace $\log_{10}(E)$.

S_n, il suffit de remarquer que leur forme diagonalisée est très proche de la structure des modèles P_n. Puisque cette décomposition découle du fait que les moments d'ordre supérieur ne sont là que pour enrichir l'approximation du régime de transport et n'interviennent pas dans le régime de diffusion, on peut supposer que la méthode reste utilisable pour les modèles dérivés des systèmes P_n (par exemples les modèles PP_n, FP_n introduits dans le chapitre 1). La décomposition « diffusive - non diffusive » permet donc d'obtenir des schémas consistants pour tous les régimes, pour les modèles simplifiés issus du transport. En outre, à l'aide d'une modification simple dans les schémas de Rusanov ou upwind utilisés pour le second système, on peut obtenir une condition CFL indépendante de ε. Cependant, cette méthode est incomplète notamment dans le cadre de la discrétisation des modèles S_n. En effet, ces modèles préservent la positivité des variables. Or la méthode numérique proposée ne préserve pas la positivité des variables pour l'instant. Par ailleurs, selon le schéma utilisé pour l'équation de la chaleur hyperbolique, la positivité peut ne pas être préservée même lorsque $\sigma = 0$.

Troisième partie

Schémas « asymptotic preserving » pour des modèles non linéaires

Chapitre 5

Modèle P_1 et couplage avec la matière

Lors de l'interaction entre la matière et un rayonnement, certaines particules sont absorbées puis réémises par le milieu dans une direction aléatoire, ce qui provoque un échange d'énergie. Ces mécanismes sont détaillés dans [Tur03]- [AG11]. Dans les équations du rayonnement, ces phénomènes se traduisent par un couplage raide et non linéaire entre le modèle de rayonnement et l'équation de la matière (exemple de couplage : l'équation de transfert radiatif (1.6)). La présence de ces termes sources induit une limite de diffusion différente de celle étudiée lors des chapitres précédents. On propose donc d'étendre les schémas « asymptotic preserving » (décrits précédemment) au cadre du couplage avec la matière. On commence par étudier le modèle P_1 couplé avec la matière, d'abord du point de vue continu puis numériquement en 1D. Enfin, on proposera une extension en 2D, couplée avec une discrétisation en temps qui permet de traiter la raideur et la non linéarité de la relaxation entre la matière et le rayonnement. Afin d'étudier ce mécanisme de façon simplifiée, on remplace les équations d'Euler par une équation sur la température qu'on introduira par la suite.

5.1 Modèle P_1 couplé avec la matière et limite asymptotique

Tout d'abord quelques définitions : soit σ le coefficient de « scattering » ou de dispersion, σ_a le coefficient d'absorption/émission des particules par la matière. On note T et ρ la température et la densité de la matière.

Définition 5.1.1. *Le modèle P_1 couplé à la matière considéré*

$$\begin{cases} \partial_t E + c \operatorname{div} \mathbf{F} = c\sigma_a(aT^4 - E), \\ \partial_t \mathbf{F} + \dfrac{c}{3}\nabla E = -c(\sigma + \sigma_a)\mathbf{F}, \\ \partial_t(\rho C_v T) = c\sigma_a(E - aT^4). \end{cases} \tag{5.1}$$

avec a la constante de Stefan-Boltzmann, C_v la capacité calorifique de la matière et c la vitesse de la lumière.

Il existe plusieurs limites asymptotiques pour le système (5.1). En temps long, quand le phénomène d'absorption/émission domine par rapport aux collisions (« scattering »), on parle de limite à l'équilibre car les températures de la matière et du rayonnement s'équilibrent. Lorsque le phénomène de « scattering » est dominant par rapport à l'absorption, on parle alors de limite hors équilibre. Afin d'étudier ces limites, on réécrit le système avec le « scaling » en ε. Le paramètre ε peut venir d'une mise à l'échelle du temps ou de la vitesse de la lumière.

Définition 5.1.2 (**Limite à l'équilibre**). *En utilisant le « scaling » $\bar{t} = \varepsilon t$, $\overline{\sigma}_a = \varepsilon\sigma_a$, $\sigma = 0$ on obtient le système suivant*

$$\begin{cases} \partial_{\bar{t}} E + \dfrac{1}{\varepsilon} \operatorname{div} \mathbf{F} = \dfrac{\overline{\sigma}_a}{\varepsilon^2}(aT^4 - E), \\ \partial_{\bar{t}} \mathbf{F} + \dfrac{1}{\varepsilon}\nabla E = -\dfrac{\overline{\sigma}_a}{\varepsilon^2}\mathbf{F}, \\ \partial_{\bar{t}}(\rho C_v T) = \dfrac{\overline{\sigma}_a}{\varepsilon^2}(E - aT^4). \end{cases} \tag{5.2}$$

La limite de diffusion de ce système est

$$\partial_{\bar{t}}(\rho C_v T + aT^4) - \operatorname{div}\left(\frac{a}{\overline{\sigma}_a}\nabla T^4\right) = 0. \tag{5.3}$$

On parle de limite à l'équilibre.

Afin de simplifier les notations on remplace les quantités $\bar{t}$ et $\overline{\sigma}_a$ par t et σ_a. Pour obtenir l'équivalence entre les deux modèles (5.1) et (5.2) on pose $\varepsilon = \frac{\sqrt{3}}{c}$, $\overline{\sigma}_a = \frac{3}{c}\sigma_a$ et $\overline{\mathbf{F}} = \frac{\sqrt{3}}{c}\mathbf{F}$.

Démonstration. La démonstration formelle se fait à l'aide d'un développement de Hilbert. On commence par introduire les termes homogènes à $\frac{1}{\varepsilon^2}$ et $\frac{1}{\varepsilon}$ dans le développement de Hilbert.

$$\frac{1}{\varepsilon^2} : \qquad E_0 = aT_0^4, \quad \mathbf{F}_0 = \mathbf{0}. \tag{5.4}$$

$$\frac{1}{\varepsilon} : \operatorname{div}\mathbf{F}_0 = \sigma_a(aT_1^4 - E_1), \quad \nabla E_0 = -\sigma_a \mathbf{F}_1, \quad \sigma_a(aT_1^4 - E_1) = 0 \tag{5.5}$$

Les termes homogènes à $\frac{1}{\varepsilon^0}$ sont :

- $\partial_t E_0 + \operatorname{div}\mathbf{F}_1 = \sigma_a(aT_2^4 - E_2)$.
- $\partial_t \mathbf{F}_0 + \nabla E_1 = -\sigma_a \mathbf{F}_2 \Longleftrightarrow \nabla E_1 = -\sigma_a \mathbf{F}_2$.
- $\partial_t(\rho C_v T_0) = \sigma_a(E_2 - aT_2^4)$.

Pour obtenir la limite de diffusion, on somme la première et la troisième équations, on incorpore $\nabla E_0 = -\sigma_a \mathbf{F}_1$ et on utilise $E_0 = aT_0^4$. □

Remarque 5.1.1. *Lorsque le rayonnement est distribué selon une Planckienne (1.5), la température du rayonnement est donnée par $T_r = (E/a)^{\frac{1}{4}}$. Par conséquent la limite à l'équilibre revient à $T_r = T$.*

Définition 5.1.3 (**Limite hors équilibre**). *En utilisant le « scaling » $\bar{t} = \varepsilon t$,$\overline{\sigma} = \varepsilon\sigma$ et $\overline{\sigma}_a = \frac{1}{\varepsilon}\sigma_a$ on obtient le système suivant*

$$\begin{cases} \partial_{\bar{t}} E + \dfrac{1}{\varepsilon} \operatorname{div} \mathbf{F} = \overline{\sigma}_a(aT^4 - E), \\ \partial_{\bar{t}} \mathbf{F} + \dfrac{1}{\varepsilon}\nabla E = -\dfrac{\overline{\sigma}}{\varepsilon^2}\mathbf{F} - \overline{\sigma}_a\mathbf{F}, \\ \partial_{\bar{t}}(\rho C_v T) = \overline{\sigma}_a(E - aT^4). \end{cases} \tag{5.6}$$

La limite de diffusion de ce système est

$$\begin{cases} \partial_{\bar{t}} E - \operatorname{div}\left(\dfrac{1}{\overline{\sigma}}\nabla E\right) = \overline{\sigma}_a(aT^4 - E), \\ \\ \partial_{\bar{t}}(\rho C_v T) = \overline{\sigma}_a(E - aT^4). \end{cases} \tag{5.7}$$

On parle de limite hors équilibre.

Afin de simplifier les notations on remplace les quantités $\bar{t}$, $\overline{\sigma}_a$ et $\overline{\sigma}$ par t, σ_a et σ.

Démonstration. On utilise un développement de Hilbert.

$$\frac{1}{\varepsilon^2} : \mathbf{F}_0 = \mathbf{0}. \tag{5.8}$$

$$\frac{1}{\varepsilon} : \operatorname{div}\mathbf{F}_0 = 0, \quad \nabla E_0 = -\sigma\mathbf{F}_1. \tag{5.9}$$

Les termes homogènes à $\frac{1}{\varepsilon^0}$: sont

- $\partial_t E_0 + \operatorname{div}\mathbf{F}_1 = \sigma_a(aT_0^4 - E_0)$.
- $\partial_t \mathbf{F}_0 + \nabla E_1 = -\sigma\mathbf{F}_2 - \sigma\mathbf{F}_0 \Longleftrightarrow \nabla E_1 = -\sigma\mathbf{F}_2$.
- $\partial_t(\rho C_v T_0) = \sigma_a(E_0 - aT_0^4)$.

Pour obtenir la limite de diffusion, on incorpore $\nabla E_0 = -\sigma_a\mathbf{F}_1$ dans l'équation $\partial_t E_0 + \operatorname{div}\mathbf{F}_1 = \sigma_a(aT_0^4 - E_0)$. □

La limite hors équilibre correspond à celle étudiée dans les chapitres précédents, en effet les termes sources d'échanges ne modifient pas la structure du système. On souhaite donc étendre les méthodes « asymptotic preserving » au cas de la limite à l'équilibre.

5.2 Schémas « asympotic preserving » en 1D pour le couplage avec la matière

Maintenant on propose d'étendre les schémas 1D de Jin-Levermore et Gosse-Toscani au modèle couplé (5.2). La construction de schéma « asymptotic preserving » pour le modèle P_1 dans le cas de la limite hors équilibre, nécessite de modifier les flux du schéma upwind classique. Dans cette étude on souhaite déterminer si la discrétisation du modèle couplé et par conséquent la modification des flux doit être repensée pour capturer la limite à l'équilibre. Pour comprendre ceci, on commence par étendre, sans modification les schémas « asymptotic preserving » , 1D au cas du couplage avec la matière.

5.2.1 Schémas de Jin-Levermore et Gosse-Toscani

On propose d'étudier le caractère « asymptotic preserving » du schéma de Jin-Levermore appliqué à l'équation de la chaleur hyperbolique couplé avec la matière

$$\begin{cases} \partial_t E + \dfrac{1}{\varepsilon}\partial_x F = \dfrac{\sigma_a}{\varepsilon^2}(aT^4 - E), \\ \\ \partial_t F + \dfrac{1}{\varepsilon}\partial_x E = -\dfrac{\sigma_a}{\varepsilon^2}F, \\ \\ \partial_t(\rho C_v T) = \dfrac{\sigma_a}{\varepsilon^2}(E - aT^4). \end{cases} \tag{5.10}$$

avec $E \in \mathbb{R}^+$, $F \in \mathbb{R}$ et $T \in \mathbb{R}^+$. Pour utiliser des estimations de stabilité L^∞ sur la solution continue et simplifier la preuve de consistance on choisit $\rho C_v = aT^3$, par conséquent le système (5.10) devient linéaire. Cette simplification permet malgré tout d'étudier les spécificités de la limite de diffusion à l'équilibre. Le schéma de Jin-Levermore couplé avec une discrétisation classique des termes sources liés à la matière s'écrit

$$\begin{cases} \dfrac{E_j^{n+1}-E_j^n}{\Delta t} + M\dfrac{F_{j+1}^n - F_{j-1}^n}{2\varepsilon\Delta x} - M\dfrac{E_{j+1}^n - 2E_j^n + E_{j-1}^n}{2\varepsilon\Delta x} = \dfrac{\sigma_a}{\varepsilon^2}(aT_j^{4,n} - E_j^n), \\ \dfrac{F_j^{n+1}-F_j^n}{\Delta t} + \dfrac{E_{j+1}^n - E_{j-1}^n}{2\varepsilon\Delta x} - \dfrac{F_{j+1}^n - 2F_j^n + F_{j-1}^n}{2\varepsilon\Delta x} = -\dfrac{\sigma_a}{\varepsilon^2}F_j^n, \\ \dfrac{(\rho C_v T)_j^{n+1} - (\rho C_v T)_j^n}{\Delta t} = \dfrac{\sigma_a}{\varepsilon^2}(E_j^n - aT_j^{4,n}). \end{cases} \tag{5.11}$$

avec $M = \frac{2\varepsilon}{2\varepsilon + \Delta x \sigma_a}$.
Afin d'étudier la consistance, on choisit des données initiales proches de l'équilibre du système : $||(\partial_{t^a}\partial_{x^b}E)(t=0,x)|| < C^1_{a,b}$, $||(\partial_{t^a}\partial_{x^b}T^4)(t=0,x)|| < C^2_{a,b}$, $||(\partial_{t^a}\partial_{x^b})F(t=0,x)|| < \varepsilon C^3_{a,b}$ et $||(\partial_{t^a}\partial_{t^b}(E-T^4))(t=0,x)|| < \varepsilon^2 C^4_{a,b}$. Puisque que le système couplé (5.10) avec $\rho C_v = a4T^3$ est linéaire et stable en norme L^∞, ces estimations sont propagées en temps avec de nouvelles constantes.

Lemme 5.1. *L'erreur de consistance associée à la deuxième équation du schéma de Jin-Levermore (5.11) pour le modèle (5.10) est $O\left(\frac{\Delta x}{\varepsilon} + \Delta t\right)$, celle associée à la première équation $O\left(\Delta x^2 + \varepsilon\Delta x + \Delta t\right)$ et celle de l'équation sur la température sont homogènes à $O(\Delta t)$.*

Démonstration. L'étude de consistance utilise les mêmes outils que pour les démonstrations du chapitre 1. L'erreur de consistance associée à la première équation de (5.11) est

$$cE_j^n = \frac{E(x_j,t_{n+1}) - E(x_j,t_n)}{\Delta t} + M\frac{F(x_{j+1},t_n) - F(x_{j-1},t_n)}{2\varepsilon\Delta x} -$$

$$M\frac{E(x_{j+1},t_n) - 2E(x_j,t_n) + E(x_{j-1},t_n)}{2\varepsilon\Delta x} - \frac{\sigma_a}{\varepsilon^2}(aT(x_j,t_n)^4 - E(x_j,t_n))$$

$$= \partial_t E(x_j,t_n) + O(\Delta t) + M\frac{\partial_x F(x_j,t_n) + O(\varepsilon\Delta x^2)}{\varepsilon} - M\frac{\Delta x\partial_{xx}E(x_j,t_n) + O(\Delta x^3)}{2\varepsilon} - \frac{\sigma_a}{\varepsilon^2}(aT(x_j,t_n)^4 - E(x_j,t_n))$$

L'erreur de consistance $\partial_x F(x_j,t_n) + O(\varepsilon\Delta x^2)$ vient du fait que $\partial_{xx}F = O(\varepsilon)$.

$$cE_j^n = \left(\partial_t E(x_j,t_n) + \frac{M}{\varepsilon}\partial_x F(x_j,t_n) - \frac{M\Delta x}{2\varepsilon}\partial_{xx}E(x_j,t_n)\right)$$

$$-\frac{\sigma_a}{\varepsilon^2}(aT(x_j,t_n)^4 - E(x_j,t_n)) + O\left(M\Delta x^2 + \frac{M\Delta x^3}{2\varepsilon} + \Delta t\right).$$

On réécrit la précédente équation

$$cE_j^n = \left(\partial_t E(x_j,t_n) + \frac{1}{\varepsilon}\partial_x F(x_j,t_n) - \frac{1-M}{\varepsilon}\partial_x F(x_j,t_n) -\right.$$

$$\left.\frac{M\Delta x}{2\varepsilon}\partial_{xx}E(x_j,t_n) - \frac{\sigma_a}{\varepsilon^2}(aT(x_j,t_n)^4 - E(x_j,t_n))\right) + O\left(M\Delta x^2 + \frac{M\Delta x^3}{2\varepsilon} + \Delta t\right).$$

Maintenant on utilise $1 - M = \frac{M\sigma_a \Delta x}{2\varepsilon}$, la relation sur les quantités continues

$$\partial_{xx} E = -\frac{\sigma_a}{\varepsilon}\partial_x F - \varepsilon \partial_{xt} F = -\frac{\sigma_a}{\varepsilon}\partial_x F + O(\varepsilon^2).$$

et la première équation continue. On obtient donc

$$cE_j^n = \left(-\frac{M\Delta x \sigma_a}{2\varepsilon^2} + \frac{M\Delta x \sigma_a}{2\varepsilon^2}\right)\partial_x F + O\left(\frac{\varepsilon^2 M}{2\varepsilon \Delta x}\right) + O\left(M\Delta x^2 + \frac{M\Delta x^3}{2\varepsilon} + \Delta t\right).$$

Enfin on remarque $M\frac{\Delta x^3}{2\varepsilon} = O(\Delta x^2)$. En outre $M \leq 1$, par conséquent $\frac{\varepsilon^2 M}{2\varepsilon \Delta x} = O(\varepsilon \Delta x)$ et $O(M\Delta x^2) = O(\Delta x \varepsilon)$. L'erreur de consistance cF pour la deuxième équation est

$$cF_j^n = \frac{F(x_j, t_{n+1}) - F(x_j, t_n)}{\Delta t} + \frac{E(x_{j+1}, t_n) - E(x_{j-1}, t_n)}{2\varepsilon \Delta x}$$

$$-\frac{F(x_{j+1}, t_n) - 2F(x_j, t_n) + F(x_j, t_n)}{2\varepsilon \Delta x} + \frac{\sigma_a}{\varepsilon^2} F(x_j, t_n)$$

$$= (\partial_t F(x_j, t_n) + O(\Delta t)) + \left(\frac{\partial_x E(x_j, t_n)}{\varepsilon} + O(\frac{\Delta x^2}{\varepsilon})\right) + O\left(\frac{\Delta x}{\varepsilon}\right)\partial_{xx} F(x_j, t_n) + \frac{\sigma_a}{\varepsilon^2} F(x_j, t_n)$$

On utilise le fait que $\partial_{xx} F(x_j, t_n) = O(\varepsilon)$. Par conséquent

$$cF_j^n = O\left(\frac{\Delta x^2}{\varepsilon} + \Delta x + \Delta t\right).$$

Pour finir il est trivial de montrer que

$$cT_j^n = O(\Delta t).$$

□

On remarque que le schéma de Jin-Levermore pour l'équation de la chaleur hyperbolique couplé à la matière est consistant avec les mêmes estimations d'erreurs que pour l'équation sans couplage (section 1.3.3). Il capture donc correctement la limite de diffusion et est « asymptotic preserving » au sens de la définition 1.3.3. Si on utilise le schéma de Gosse-Toscani, le schéma est aussi consistant puisque seule la discrétisation de la seconde équation est modifiée. On introduit le schéma de Gosse-Toscani avec couplage

$$\begin{cases} \dfrac{E_j^{n+1} - E_j^n}{\Delta t} + M\dfrac{F_{j+1}^n - F_{j-1}^n}{\varepsilon \Delta x} - M\dfrac{E_{j+1}^n - 2E_j^n + E_{j-1}^n}{\varepsilon \Delta x} = \dfrac{\sigma_a}{\varepsilon^2}(aT_j^{4,n} - E_j^n), \\ \dfrac{F_j^{n+1} - F_j^n}{\Delta t} + M\dfrac{E_{j+1}^n - E_{j-1}^n}{\varepsilon \Delta x} - M\dfrac{F_{j+1}^n - 2F_j^n + F_{j-1}^n}{\varepsilon \Delta x} = -M\dfrac{\sigma_a}{\varepsilon^2} F_j^n, \\ \dfrac{(\rho C_v T)_j^{n+1} - (\rho C_v T)_j}{\Delta t} = \dfrac{\sigma_a}{\varepsilon^2}(E_j^n - aT_j^{4,n}). \end{cases} \tag{5.12}$$

Lemme 5.2. *On choisit $\rho C_v = aT^3$ pour le modèle (5.10). L'erreur de consistance associée à la première équation du schéma de Gosse-Toscani (5.12) pour le modèle (5.10) en 1D est $O\left(\Delta x^2 + \varepsilon \Delta x + \Delta t\right)$ et celle associée à la seconde équation est $(\Delta x + \Delta t)$.*

Démonstration. La consistance de la première équation est donnée par l'étude de consistance du schéma de Jin-Levermore avec couplage (démonstration lemme 5.1).
La consistance de la seconde équation est donnée par l'étude de consistance du schéma de Gosse-Toscani sans couplage (démonstration lemme 1.10).
La consistance de la troisième équation est donnée par l'étude de consistance du schéma de Jin-Levermore avec couplage (démonstration lemme 5.1).

□

Lemme 5.3. *On suppose ρC_v constant. Le schéma de Gosse-Toscani (5.12) pour le modèle (5.10) ne préserve pas le principe du maximum pour les quantités $E-F$, $E+F$, et T^4.*

Démonstration. On pose $u = E + F$ et $v = E - F$. Pour ces variables, le système s'écrit

$$\begin{cases} \dfrac{u_j^{n+1}-u_j^n}{\Delta t} + M\dfrac{u_j^n - u_{j-1}^n}{\varepsilon\Delta x} + M\dfrac{\sigma_a}{2\varepsilon^2}\left(u_j^n - v_j^n\right) = \dfrac{\sigma_a}{2\varepsilon^2}\left(2aT_j^{4,n} - (u_j^n + v_j^n)\right), \\ \dfrac{v_j^{n+1}-v_j^n}{\Delta t} - M\dfrac{v_{j+1}^n - v_j^n}{\varepsilon\Delta x} + M\dfrac{\sigma}{2\varepsilon^2}\left(-u_j^n + v_j^n\right) = \dfrac{\sigma_a}{2\varepsilon^2}\left(2aT_j^{4,n} - (u_j^n + v_j^n)\right), \\ \dfrac{(\rho C_v T)_j^{n+1} - (\rho C_v T)_j^n}{\Delta t} = \dfrac{\sigma_a}{2\varepsilon^2}\left(u_j^n + v_j^n - 2aT_j^{4,n}\right). \end{cases}$$

par conséquent

$$\begin{cases} u_j^{n+1} = \left(1 - M\dfrac{\Delta t}{\varepsilon\Delta x} - (1+M)\dfrac{2\sigma_a\Delta t}{\varepsilon^2}\right)u_j^n + M\dfrac{\Delta t}{\varepsilon\Delta x}u_{j-1}^n + \dfrac{\sigma_a\Delta t}{\varepsilon^2}aT_j^{4,n} + (M-1)\left(\dfrac{\sigma_a\Delta t}{2\varepsilon^2}\right)v_j^n, \\ v_j^{n+1} = \left(1 - M\dfrac{\Delta t}{\varepsilon\Delta x} - (1+M)\dfrac{2\sigma_a\Delta t}{\varepsilon^2}\right)v_j^n + M\dfrac{\Delta t}{\varepsilon\Delta x}v_{j+1}^n + \dfrac{\sigma_a\Delta t}{\varepsilon^2}aT_j^{4,n} + (M-1)\left(\dfrac{\sigma_a\Delta t}{2\varepsilon^2}\right)u_j^n, \\ T_j^{n+1} = \left(1 - \dfrac{\sigma_a}{\rho C_v\varepsilon^2}aT_j^{3,n}\right)T_j^n + M\dfrac{\sigma_a}{2\varepsilon^2}u_j^n + M\dfrac{\sigma_a}{2\varepsilon^2}v_j^n. \end{cases} \tag{5.13}$$

Le schéma n'est pas stable car $M - 1 \leq 0$ et donc des coefficients extra-diagonaux sont négatifs. □

Ces résultats montrent que les schémas de Jin-Levermore (5.11) et Gosse-Toscani (5.12) sont consistants avec les mêmes erreurs de consistance que pour le modèle sans couplage. Par conséquent le schéma de Jin-Levermore (5.11) avec une discrétisation classique des termes de couplage est « asymptotic preserving » au sens de la définition (1.3.3). Le schéma de Gosse-Toscani (5.12) avec une discrétisation classique des termes de couplage est « asymptotic preserving » au sens de la définition (1.3.4). Par contre, les deux schémas ne préservent pas le principe du maximum, il faut introduire des modifications pour récupérer cette propriété.

5.2.2 Schéma « asymptotic preserving » et principe du maximum

On propose une modification possible du schéma de Gosse-Toscani pour le problème (5.2) afin d'obtenir un schéma « asymptotic preserving », préservant le principe du maximum. Afin d'obtenir un schéma positif on propose de multiplier les termes sources de couplage par le facteur M et d'introduire une viscosité numérique dans l'équation sur la température qui permettra de conserver la consistance

du schéma.

$$\begin{cases} \dfrac{E_j^{n+1}-E_j^n}{\Delta t}+M\dfrac{F_{j+1}^n-F_{j-1}^n}{2\varepsilon\Delta x}-M\dfrac{E_{j+1}^n-2E_j^n+E_{j-1}^n}{2\varepsilon\Delta x}=M\dfrac{\sigma_a}{\varepsilon^2}(aT_j^{4,n}-E_j^n), \\ \dfrac{F_j^{n+1}-F_j^n}{\Delta t}+M\dfrac{E_{j+1}^n-E_{j-1}^n}{2\varepsilon\Delta x}-M\dfrac{F_{j+1}^n-2F_j^n+F_{j-1}^n}{2\varepsilon\Delta x}=-M\dfrac{\sigma_a}{\varepsilon^2}F_j^n, \\ \dfrac{(\rho C_vT)_j^{n+1}-(\rho C_vT)_j^n}{\Delta t}-M\dfrac{T_{j+1}^{4,n}-2T_j^{4,n}+T_{j-1}^{4,n}}{2\varepsilon\Delta x}=M\dfrac{\sigma_a}{\varepsilon^2}(E_j^n-aT_j^{4,n}), \end{cases} \tag{5.14}$$

avec $M=\frac{2\varepsilon}{2\varepsilon+2\sigma_a\Delta x}$.

Remarque 5.2.1. *Le coefficient M est corrigé en utilisant le même genre de manipulation que dans [BT10]. Cependant il est probable que cette manipulation ne suffise plus pour obtenir un schéma positif et « asymptotic preserving » pour des opacités variables. Le schéma (5.14) est un exemple de schéma « asymptotic preserving » positif, il existe d'autres possibilités qu'il serait utile d'étudier.*

On commence par supposer que $Cv=aT^3$, $||(\partial_{t^a}\partial_{x^b}E)(t=0,x)||<C_{a,b}$, $||(\partial_{t^a}\partial_{x^b}T^4)(t=0,x)||<C_{a,b}$, $||(\partial_{t^a}\partial_{x^b}F)(t=0,x)||<\varepsilon C_{a,b}$ et $||(\partial_{t^a}\partial_{t^b}(E-T^4))(t=0,x)||<\varepsilon^2C_{a,b}$. Par stabilité L^∞ des systèmes hyperboliques linéaires en 1D, ces estimations sont propagées en temps avec de nouvelles constantes.

Lemme 5.4. *Les erreurs de consistance associées aux seconde et troisième équations de (5.14) pour le modèle (5.10) en 1D avec $C_v=a4T^3$ sont $O(\Delta x+\Delta t)$, et celle associée à la première équation $O\left(\Delta x^2+\varepsilon\Delta x+\Delta t\right)$*

Démonstration. Afin de démontrer la consistance du schéma (5.14), on propose le changement de variable suivant : $w=\frac{E+aT^4}{2}$ et $z=\frac{E-aT^4}{2}$. Le modèle continu (5.10) devient

$$\begin{cases} \partial_t w+\dfrac{1}{2\varepsilon}\partial_x F=0, \\ \partial_t F+\dfrac{1}{\varepsilon}\partial_x(w+z)=-\dfrac{\sigma_a}{\varepsilon^2}F, \\ \partial_t z+\dfrac{1}{2\varepsilon}\partial_x F=-\dfrac{2\sigma_a}{\varepsilon^2}z. \end{cases} \tag{5.15}$$

Les estimations sur les conditions initiales deviennent $||(\partial_{t^a}\partial_{x^b}w)(t=0,x)||<C_{a,b}$, $||(\partial_{t^a}\partial_{x^b}F)(t=0,x)||<\varepsilon C_{a,b}$ et $||(\partial_{t^a}\partial_{t^b}z)(t=0,x)||<\varepsilon^2C_{a,b}$. Par stabilité L^∞ des systèmes hyperboliques linéaires en 1D ces estimations sont propagées en temps avec de nouvelles constantes.
Le schéma réécrit pour ces variables est

$$\begin{cases} \dfrac{w_j^{n+1}-w_j^n}{\Delta t}+M\dfrac{F_{j+1}^n-F_{j-1}^n}{4\varepsilon\Delta x}-M\dfrac{w_{j+1}^n-w_j^n+w_{j-1}^n}{2\varepsilon\Delta x}=0, \\ \dfrac{F_j^{n+1}-F_j^n}{\Delta t}+M\dfrac{(w+z)_{j+1}^n-(w+z)_{j-1}^n}{2\varepsilon\Delta x}-M\dfrac{F_{j+1}^n-2F_j^n+F_{j-1}^n}{2\varepsilon\Delta x}=-M\dfrac{\sigma_a}{\varepsilon^2}F_j^n, \\ \dfrac{z_j^{n+1}-z_j^n}{\Delta t}+M\dfrac{F_{j+1}^n-F_{j-1}^n}{4\varepsilon\Delta x}-M\dfrac{z_{j+1}^n-2z_j^n+z_{j-1}^n}{2\varepsilon\Delta x}=-M\dfrac{2\sigma_a}{\varepsilon^2}z_j^n. \end{cases} \tag{5.16}$$

De la même façon que pour les autres démonstrations de consistance, on injecte la solution de (5.15) dans le schéma (5.16). On commence par définir l'erreur de consistance Cw_j^n de la première équation

$$Cw_j^n=\partial_t w(x_j,t_n)+O(\Delta t)+\frac{M}{2\varepsilon}\partial_x F(x_j,t_n)+O(M\Delta x^2)-\frac{M\Delta x}{2\varepsilon}\partial_{xx}w(x_j,t_n).$$

On fait intervenir la dérivée seconde sur la quantité z. Cela donne

$$Cw_j^n = \partial_t w(x_j,t_n) + O(\Delta t) + \frac{M}{2\varepsilon}\partial_x F(x_j,t_n) + O(M\Delta x^2) - \frac{M\Delta x}{2\varepsilon}(\partial_{xx} w(x_j,t_n)$$

$$+\partial_{xx} z(x_j,t_n)) + \frac{M\Delta x}{2\varepsilon}\partial_{xx} z(x_j,t_n).$$

On remarque que la solution satisfait $(\partial_{xx} w + \partial_{xx} z)(x_j,t_n) = 2\sigma_a \partial_t w(x_j,t_n) - \varepsilon \partial_{t,x} F(x_j,t_n)$ car $\partial_t w + \frac{1}{2\varepsilon}\partial_x F = 0$. On injecte cette équation dans l'erreur de consistance et on simplifie pour obtenir

$$Cw_j^n = (1-M)\partial_t w(x_j,t_n) + O(\Delta t) + O(M\Delta x^2) - \frac{M\Delta x}{2\varepsilon}\partial_t w(x_j,t_n) + \frac{M\Delta x}{2\varepsilon}\partial_{xx} z(x_j,t_n).$$

Pour conclure, on remarque que $z(x_j,t_n) = O(\varepsilon^2)$, $F(x_j,t_n) = O(\varepsilon)$ et que $1 - M - \frac{M\Delta x}{\varepsilon} = 0$ (sans la modification du coefficient M, cette équation est fausse). Par conséquent on obtient

$$Cw_j^n = O(\Delta t + \varepsilon\Delta x + \Delta x^2).$$

À présent on passe à la seconde équation avec CF_j^n comme erreur de consistance. En injectant la solution exacte on obtient

$$CF_j^n = \partial_t F(x_j,t_n) + O(\Delta t) + \frac{M}{\varepsilon}(\partial_x w(x_j,t_n) + \partial_x z(x_j,t_n)) + O\left(\frac{M\Delta x^2}{2\varepsilon}\right)$$

$$-\frac{M\Delta x}{2\varepsilon}\partial_{xx} F(x_j,t_n) + \frac{M\sigma}{\varepsilon^2} F(x_j,t_n).$$

On obtient

$$CF_j^n = (1-M)\partial_t F(x_j,t_n) + O\left(\frac{M\Delta x^2}{2\varepsilon}\right) - \frac{M\Delta x}{2\varepsilon} F(x_j,t_n).$$

Enfin on utilise $F(x_j,t_n) = O(\varepsilon)$, $\frac{M\Delta x^2}{2\varepsilon} = O(\Delta x)$ et $(1-M)\varepsilon = O(\Delta x)$ ce qui donne

$$CF_j^n = O(\Delta t + \Delta x).$$

On introduit Cz_j^n l'erreur de consistance associée à la troisième équation. On injecte la solution dans l'équation

$$Cz_j^n = (1-M)\partial_t z(x_j,t_n) + O(\Delta t) - \frac{M\Delta x}{2\varepsilon}\partial_{xx} z(x_j,t_n) + O\left(\frac{M\Delta x^2}{2\varepsilon}\right)$$

Il suffit d'utiliser $z(x_j,t_n) = O(\varepsilon^2)$, $F(x_j,t_n) = O(\varepsilon)$ pour conclure. □

Maintenant on propose d'étudier la stabilité pour C_v général, afin de montrer les difficultés induites par la non-linéarité du couplage.

Lemme 5.5. *Le schéma (5.14) satisfait le principe du maximum et stable en norme L^∞ pour les quantités $u = E + F$, $v = E - F$, T sous la CFL*

$$\max\left(\frac{\Delta t}{\varepsilon\Delta x}, \frac{\Delta t}{\varepsilon\Delta x}\frac{\rho C_v}{aT_j^3}\right) \leq 1.$$

Démonstration. On commence par définir les invariants de Riemann $u = E + F$ et $v = E - F$, le système (5.14) se réécrit sous la forme suivante

$$\begin{cases} \dfrac{u_j^{n+1} - u_j^n}{\Delta t} + M\dfrac{u_j^n - u_{j-1}^n}{\varepsilon\Delta x} + M\dfrac{\sigma_a}{2\varepsilon^2}\left(u_j^n - v_j^n\right) = M\dfrac{\sigma_a}{2\varepsilon^2}\left(2aT_j^{4,n} - (u_j^n + v_j^n)\right), \\ \dfrac{v_j^{n+1} - v_j^n}{\Delta t} - M\dfrac{v_{j+1} - v_j^n}{\varepsilon\Delta x} + M\dfrac{\sigma_a}{2\varepsilon^2}\left(-u_j^n + v_j^n\right) = M\dfrac{\sigma_a}{2\varepsilon^2}\left(2aT_j^{4,n} - (u_j^n + v_j^n)\right), \\ \rho C_v\dfrac{T_j^{n+1} - T_j^n}{\Delta t} - M\dfrac{T_{j+1}^{4,n} - 2T_j^{4,n} + T_{j-1}^{4,n}}{2\varepsilon\Delta x} = M\dfrac{\sigma_a}{2\varepsilon^2}\left(u_j^n + v_j^n - 2aT_j^{4,n}\right) \end{cases}$$

par conséquent

$$\begin{cases} u_j^{n+1} = \left(1 - M\dfrac{\Delta t}{\varepsilon\Delta x} - M\dfrac{\sigma_a\Delta t}{\varepsilon^2}\right)u_j^n + M\dfrac{\Delta t}{\varepsilon\Delta x}u_{j-1}^n + M\dfrac{\sigma_a\Delta t}{\varepsilon^2}aT_j^{4,n}, \\ v_j^{n+1} = \left(1 - M\dfrac{\Delta t}{\varepsilon\Delta x} - M\dfrac{\sigma_a\Delta t}{\varepsilon^2}\right)v_j^n + M\dfrac{\Delta t}{\varepsilon\Delta x}v_{j+1}^n + M\dfrac{\sigma_a\Delta t}{\varepsilon^2}aT_j^{4,n}, \\ T_j^{n+1} = \left(1 - \dfrac{M}{\Delta t\varepsilon\rho C_v}aT_j^{3,n} - M\dfrac{\Delta t\sigma_a}{\varepsilon^2\rho C_v}aT_j^{3,n}\right)T_j + M\dfrac{\Delta t\sigma_a}{2\varepsilon^2}u_j^n + M\dfrac{\Delta t\sigma_a}{2\varepsilon^2}v_j^n + \dfrac{\Delta tM}{2\Delta x\varepsilon}aT_{j-1}^{4,n} + \dfrac{\Delta tM}{2\Delta x\varepsilon}aT_{j+1}^{4,n}. \end{cases} \tag{5.17}$$

On remarque que, sous la condition de positivité $\left(1 - M\frac{\Delta t}{\varepsilon\Delta x} - M\frac{\sigma_a\Delta t}{\varepsilon^2}\right) \geq 0$, u_j^{n+1} est une combinaison convexe de u_j^n, u_{j-1}^n et $aT_j^{4,n}$ et v_j^{n+1} est une combinaison convexe de v_j^n, v_{j+1}^n et $aT_j^{4,n}$. Par conséquent u_j^{n+1} est majoré par $\max(v_j^n, v_{j-1}^n, aT_j^{4,n})$ et u_j^{n+1} est majoré par $\max(v_j^n, v_{j+1}^n, aT_j^{4,n})$. Donc on obtient que les quantités u et v soient majorées en tous temps par les données initiales. Pour la température on obtient également une combinaison convexe si $\left(1 - \frac{M}{\Delta t\varepsilon\rho C_v}aT_j^{3,n} - M\frac{\sigma_a\Delta t}{\rho C_v\varepsilon^2}aT_j^{3,n}\right) \geq 0$. □

On remarque que la condition CFL induite par le couplage dépend de T^3 et de ε. Lorsque la température est importante ou que ε est petit la CFL devient rédibitoire. En effet, la présence du terme $\frac{1}{T^3}$ rend cette CFL encore plus contraignante que la CFL associé au modèle sans couplage. On peut s'attendre à ce que cette dépendance existe aussi pour les schémas de Jin-Levermore (5.11) et de Gosse-Toscani (5.12) non modifiés pour le couplage. Par conséquent, il devient absolument nécessaire de définir une discrétisation temporelle semi-implicite ou implicite. On introduira ce type de discrétisation dans le cas de la dimension deux.
Pour résumer, cette étude 1D montre que l'utilisation d'un schéma « asymptotic preserving » pour l'équation de la chaleur hyperbolique avec une discrétisation classique des termes de couplage permet d'obtenir un schéma « asymptotic preserving » sans modification supplémentaire des flux.

5.3 Extension en dimension deux

Les résultats précédents montrent que l'essentiel du travail nécessaire à la construction d'un schéma « asymptotic preserving » porte sur la discrétisation du modèle sans les termes de couplage. Par conséquent, on propose d'utiliser les schémas construits dans les chapitres deux et trois, couplés à une discrétisation en espace classique des termes de relaxation.

5.3.1 Schéma semi-discret et limite asymptotique

On propose de montrer que le schéma JL-(a) semi-discret admet la bonne limite de diffusion d'un point de vue formel. Le principe reste le même pour le schéma JL-(b) ou pour les autres schémas.

Proposition 5.6. *Le schéma JL-(a) pour le système (5.2) est défini par*

$$\begin{cases} |\Omega_j|\,\partial_t E_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r l_{jr}(\mathbf{F}_r, \mathbf{n}_{jr}) = |\Omega_j|\,\dfrac{\sigma_a}{\varepsilon^2}(aT_j^4 - E_j), \\ |\Omega_j|\,\partial_t \mathbf{F}_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r \mathbf{G}_{jr} = -\,|\Omega_j|\,\dfrac{\sigma_a}{\varepsilon^2}\mathbf{F}_j, \\ |\Omega_j|\,\partial_t(\rho C_v T_j) = |\Omega_j|\,\dfrac{\sigma_a}{\varepsilon^2}(E_j - aT_j^4). \end{cases} \tag{5.18}$$

$$\begin{cases} \mathbf{G}_{jr} = l_{jr}E_j\mathbf{n}_{jr} + \widehat{\alpha}_{jr}(\mathbf{F}_j - \mathbf{F}_r) - \dfrac{\sigma_a}{\varepsilon}\widehat{\beta}_{jr}\mathbf{F}_r, \\ \left(\displaystyle\sum_j \widehat{\alpha}_{jr} + \frac{\sigma_a}{\varepsilon}\sum_j \widehat{\beta}_{jr}\right)\mathbf{F}_r = \displaystyle\sum_j l_{jr}E_j\mathbf{n}_{jr} + \sum_j \widehat{\alpha}_{jr}\mathbf{F}_j. \end{cases} \tag{5.19}$$

avec $\widehat{\alpha}_{jr}$, $\widehat{\beta}_{jr}$ définis dans le chapitre un, est « asymptotic preserving » et admet le schéma limite

$$\begin{cases} |\Omega_j|\,\partial_t(\rho C_v T_j + aT_j^4) + \displaystyle\sum_r l_{jr}(\mathbf{F}_r, \mathbf{n}_{jr}) = 0, \\ \sigma_a\left(\displaystyle\sum_j \widehat{\beta}_{jr}\right)\mathbf{F}_r = \displaystyle\sum_j l_{jr}aT_j^4\mathbf{n}_{jr}. \end{cases} \tag{5.20}$$

Démonstration. On écrit le développement de Hilbert :

$$E_j = E_{j,0} + \varepsilon E_{j,1} + \varepsilon^2 E_{j,2} + o(\varepsilon^2),$$
$$E_{jr} = \mathbf{G}_{jr,0} + \varepsilon\mathbf{G}_{jr,1} + \varepsilon^2\mathbf{G}_{jr,2} + o(\varepsilon^2).$$

Le même développement est fait pour toutes les variables. En les incorporant dans (5.33) on obtient le terme proportionnel à $\frac{1}{\varepsilon^2}$:

$$|\Omega_j|\sigma_a(aT_{j,0}^4 - E_{j,0}) = 0. \tag{5.21}$$

Le terme proportionnel à $\frac{1}{\varepsilon}$ est

$$\sum_r l_{jr}(\mathbf{F}_{r,0}, \mathbf{n}_{jr}) = |\Omega_j|\sigma_a(aT_{j,1}^4 - E_{j,1}). \tag{5.22}$$

le terme proportionnel à $O(1)$ est

$$|\Omega_j|\partial_t E_{j,0} + \sum_r l_{jr}(\mathbf{F}_{r,1}, \mathbf{n}_{jr}) = \sigma_a(aT_{j,2}^4 - E_{j,2}). \tag{5.23}$$

En utilisant la même méthode pour la deuxième équation on obtient trois termes

$$\frac{1}{\varepsilon^2}: \qquad \mathbf{F}_{j,0} = \mathbf{0}, \tag{5.24}$$

$$\frac{1}{\varepsilon}: \qquad \sum_r \mathbf{G}_{jr,0} = -|\Omega_j|\sigma_a\mathbf{F}_{j,1}, \tag{5.25}$$

$$\frac{1}{\varepsilon^0}: \qquad |\Omega_j|\partial_t \mathbf{F}_{j,0} + \sum_r \mathbf{G}_{jr,1} = -|\Omega_j|\sigma_a \mathbf{F}_{j,2}. \tag{5.26}$$

Pour l'équation sur la température matière le résultat est immédiat on obtient que

$$-|\Omega_j|\sigma_a(aT_{j,0}^4 - E_{j,0}) = 0, \quad -|\Omega_j|\sigma_a(aT_{j,1}^4 - E_{j,1}) = 0,$$

et que

$$|\Omega_j|\partial_t T_j^0 = -|\Omega_j|\sigma_a(aT_{j,2}^4 - E_{j,2}) = 0. \tag{5.27}$$

On opère aussi le développement classique dans le solveur nodal

$$\frac{1}{\varepsilon}: \qquad \sigma_a \left(\sum_j \widehat{\beta}_{jr} \right) \mathbf{F}_{r,0} = \mathbf{0}, \tag{5.28}$$

et

$$\frac{1}{\varepsilon^0}: \qquad \left(\sum_j \widehat{\alpha}_{jr} \right) \mathbf{F}_{r,0} + \sigma_a \left(\sum_j \widehat{\beta}_{jr} \right) \mathbf{F}_{r,1} \tag{5.29}$$

$$= \sum_j l_{jr} E_{j,0} \mathbf{n}_{jr} + \sum_j l_{jr} \mathbf{n}_{jr} \otimes \mathbf{n}_{jr} \mathbf{F}_{j,0}.$$

En supposant que la matrice dans l'équation (5.28) soit inversible, (5.28) on montre $\mathbf{F}_{r,0} = \mathbf{0}$. Puisque $\mathbf{F}_{j,0} = \mathbf{0}$ on obtient que $\mathbf{F}_{r,1}$ est solution de

$$\sigma_a \left(\sum_j l_{jr} \mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j) \right) \mathbf{F}_{r,1} = \sum_j l_{jr} E_{j,0} \mathbf{n}_{jr}.$$

Enfin il suffit de sommer (5.23) et (5.27) et utiliser le fait que $E_{j,0} = aT_{j,0}^4$ par (5.21). □

De façon formelle on montre que le schéma capture une discrétisation de la limite de diffusion, lorsqu'on ajoute les termes de couplage comme cela avait déja été constaté en dimension un. L'étude du schéma de Gosse-Toscani modifié 1D a montré que la stabilité du schéma était, elle, modifiée contrairement à la consistance. La raideur de la condition CFL nous oblige donc à construire des schémas utilisant des discrétisations temporelles différentes.

5.3.2 Schéma semi-implicite pour le modèle P_1 couplé à la matière

Dans cette section on souhaite étudier la façon de traiter les termes de couplage entre le modèle P_1 et l'équation sur la matière. Comme cela a été montré dans la section sur les schémas 1D pour le couplage avec la matière, la condition CFL induite est très restrictive notamment à cause de sa forte dépendance envers la température de la matière. On propose donc une discrétisation semi-implicite afin de lever cette contrainte. Enfin on étudiera le schéma total obtenu pour le modèle (5.2). Le schéma est basé sur un « splitting » entre le modèle P_1 et les termes de couplage. La partie P_1 est résolue avec les schémas précédemment introduits explicite ou semi-implicite (JL-(a), JL-(b), P1-MPFA etc). Seul le schéma JL-(b) dans sa formulation semi-implicite donne une CFL indépendante de ε. Le couplage est lui résolu en utilisant une méthode de point fixe que l'on détaille maintenant.

Point fixe pour le couplage

Afin d'éviter une condition CFL dépendante des variations de températures, on propose d'utiliser un schéma implicite pour les termes de relaxation en températures du modèle (5.2) :

$$\begin{cases} \partial_t E = \dfrac{\sigma_a}{\varepsilon^2}(aT^4 - E), \\ \rho C_v \partial_t T = \dfrac{\sigma_a}{\varepsilon^2}(E - aT^4). \end{cases} \tag{5.30}$$

La relaxation étant non linéaire on doit coupler un schéma d'Euler implicite avec une procédure de point fixe.
On pose $\Theta = aT^4$, La procédure de point fixe local associé à la relaxation en température s'écrit de la manière suivante

Définition 5.3.1.

$$\begin{cases} \dfrac{E^{q+1} - E^n}{\Delta t} = \dfrac{\sigma_a}{\varepsilon^2}(\Theta^{q+1} - E^{q+1}), \\ \\ \rho C_v \mu^q \dfrac{\Theta^{q+1} - \Theta^n}{\Delta t} = \dfrac{\sigma_a}{\varepsilon^2}(E^{q+1} - \Theta^{q+1}), \end{cases} \tag{5.31}$$

Avec $\mu^q = \frac{T_q - T^n}{\Theta^q - \Theta^n}$.

Avant de continuer, plusieurs remarques. Tout d'abord cette procédure de point fixe est toujours définie, car on peut désinguliariser μ en utilisant le fait que Θ est proportionnelle à T^4. Dans ce cas

$$\mu^q = \frac{1}{T^3 + T^q T^2 + T T^{q,2} + T^{q,3}}.$$

Cette procédure consiste à résoudre le système de taille deux, $A_q x_q = b_q$ avec

$$A = \begin{pmatrix} 1 + \Delta t \dfrac{\sigma_a}{\varepsilon^2} & -\Delta t \dfrac{\sigma_a}{\varepsilon^2} \\ -\Delta t \dfrac{\sigma_a}{\varepsilon^2} & \rho C_v \mu^q + \Delta t \dfrac{\sigma_a}{\varepsilon^2} \end{pmatrix} b = \begin{pmatrix} E \\ \mu^q \rho C_v \Theta \end{pmatrix}.$$

Les coefficients diagonaux de la matrice étant positifs, les coefficients non-diagonaux étant négatifs et la somme des coefficients de chaque colonne étant positive, la matrice A_q est une M-Matrice. Elle est donc inversible et garantit la positivité des variables à chaque itération du point fixe.

5.3.3 Schéma semi-implicite par « splitting » pour le modèle P_1 couplé avec la matière

Comme dans le cas précédent on propose de détailler l'algorithme de résolution du système couplé (5.2) en utilisant le schéma JL-(a) pour la partie hyperbolique. La méthode reste valide pour les autres schémas, par exemple JL-b, P_1-MPFA. On nomme $\overline{E}$ et $\overline{\Theta}$, les valeurs convergées du point fixe.

Schéma de type « splitting » pour le modèle P_1 couplé

Définition 5.3.2. ***Etape1*** *On commence par résoudre le couplage avec l'algorithme*

$$\begin{cases} \dfrac{E_j^{q+1}-E_j^n}{\Delta t} = \dfrac{\sigma_a}{\varepsilon^2}(\Theta_j^{q+1}-E_j^{q+1}), \\ \dfrac{\mathbf{F}_j^{q+1}-\mathbf{F}_j^n}{\Delta t} = \mathbf{0}, \\ \rho C_v \mu^q \dfrac{\Theta_j^{q+1}-\Theta_j^n}{\Delta t} = \dfrac{\sigma_a}{\varepsilon^2}(E_j^{q+1}-\Theta_j^{q+1}), \end{cases} \tag{5.32}$$

avec les quantités E_j^q, Θ_j^q, T_j^q qui convergent vers $\overline{E}_j$, $\overline{\Theta}_j$ et $\overline{T}_j$.

Etape 2 *On résout la partie hyperbolique avec le schéma*

$$\begin{cases} |\Omega_j| \dfrac{E_j^{n+1}-\overline{E}_j}{\Delta t} + \dfrac{1}{\varepsilon}\displaystyle\sum_r l_{jr}(\mathbf{F}_r, \mathbf{n}_{jr}) = 0 \\ |\Omega_j| \dfrac{\mathbf{F}_j^{n+1}-\overline{\mathbf{F}}_j}{\Delta t} + \dfrac{1}{\varepsilon}\displaystyle\sum_r \mathbf{G}_{jr} = -|\Omega_j| \dfrac{\sigma_a}{\varepsilon^2}\overline{\mathbf{F}}_j, \\ |\Omega_j| \rho C_v \dfrac{T_j^{n+1}-\overline{T}_j}{\Delta t} = 0, \end{cases} \tag{5.33}$$

avec

$$\begin{cases} \mathbf{G}_{jr} = l_{jr}\overline{E}_j \mathbf{n}_{jr} + \widehat{\alpha}_{jr}(\overline{\mathbf{F}}_j - \mathbf{F}_r) - \dfrac{\sigma_a}{\varepsilon}\widehat{\beta}_{jr}\mathbf{F}_r, \\ \left(\displaystyle\sum_j \widehat{\alpha}_{jr} + \frac{\sigma_a}{\varepsilon}\sum_j \widehat{\beta}_{jr}\right)\mathbf{F}_r = \displaystyle\sum_j l_{jr}\overline{E}_j\mathbf{n}_{jr} + \sum_j \widehat{\alpha}_{jr}\overline{\mathbf{F}}_j. \end{cases} \tag{5.34}$$

Remarque 5.3.1. *L'algorithme reste valide pour des modèles de plus grande taille comme les modèles P_n. Il suffit d'utiliser un schéma « asymptotic preserving » pour la partie hyperbolique dans l'étape deux.*

5.4 Résultats numériques

Afin d'illustrer cette section nous regardons le comportement du schéma à travers deux cas test. Il s'agit de « Olson-Auer-Hall » [OAH99] et de « Marshak linéarisé » [BCT08]- [B010]- [Tur03]. Il s'agit de deux cas test en milieu semi-transparent (cas où la longueur caractéristique du milieu et σ sont du même ordre. Ces problèmes sont invariants dans la direction y. Ils sont résolus sur un domaine 2D, cependant on ne tracera la solution que pour une seule valeur de y.

Cas test de Olson-Auer-hall

Il s'agit de résoudre le système (5.1). Le domaine est $[0:3] \times [0:0.05]$.

- la densité matière $\rho = 0.38214\ g/cm^3$,
- $T_r(t=0,\mathbf{x}) = T(t=0,\mathbf{x}) = 0.056234$ keV,
- $\sigma = 0$ et $\sigma_a = \frac{3}{c}T^{-3}$,
- $C_v = 0.14361\ Jerk/g/keV$,
- $c = 299.\ cm/sh$, $\varepsilon = \frac{\sqrt{3}}{c}$,
- $a = 0.01372\ g/cm/sh^2/keV^4$.

Le cas consiste à regarder l'évolution des températures de la matière et du rayonnement. Le milieu est chauffé par le rayonnement issu d'un corps noir à gauche. Plus le milieu devient chaud, plus l'opacité d'absorption augmente. Par conséquent, les températures du milieu et du rayonnement s'équilibrent. La condition limite à gauche est une condition de Dirichlet avec $T_r = 1$. Les temps tracés figure (6.1) sont $t = \frac{1}{c}, \frac{3}{c}, \frac{10}{c}, \frac{30}{c}, \frac{100}{c}$. Le schéma semi-implicite et le schéma explicite donnent les mêmes résultats.

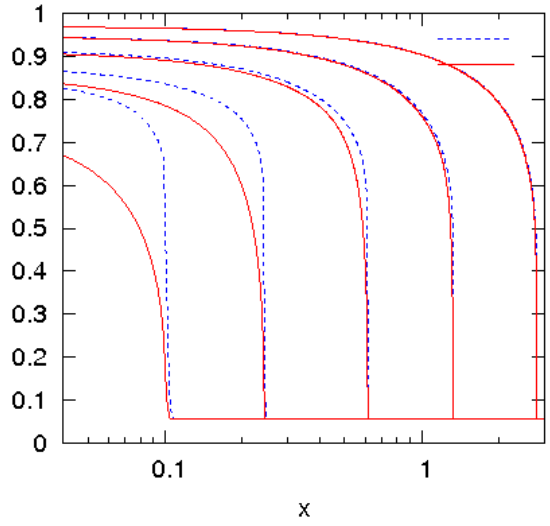


FIGURE 5.1 – La température matière et la température radiative pour les cinq temps choisis précédemment. Les courbes bleu correspondent à la température radiative T_r. Les courbes rouges correspondent à la température matière.

Pour ce cas test le schéma semi-implicite ne présente pas d'avantage particulier car le terme $\frac{aT^3}{\rho C_v}$ est assez faible.

Cas test de Marshak linéaire

Comme précédemment il s'agit d'un milieu chauffé par sa gauche. Le domaine est $[0:10] \times [0:0.05]$

- la densité matière $\rho = 1\ g/cm^3$,
- $T_r(t=0,\mathbf{x}) = T(t=0,\mathbf{x}) = 300$ K,
- $\sigma = 0$ et $\sigma_a = \frac{3}{c}$,
- C_v=0.001 $Jerk/g/K$,
- $c = 2.997910^{10}\ cm/s$, $\varepsilon = \frac{\sqrt{3.}}{c}$,
- $a = 7.55510^{-15}\ g/cm/s^2/K$.

Les temps étudiés sont $t = 1.333 \times 10^{-9}, 1.333 \times 10^{-8}, 1.333 \times 10^{-7}$. On regarde dans ce cas le système (5.2) ainsi que

$$\begin{cases} \partial_t E - \text{div}\left(\dfrac{1}{\sigma_a}\nabla E\right) = \dfrac{\sigma_a}{\varepsilon^2}(aT^4 - E), \\ \partial_t(\rho C_v T) = \dfrac{\sigma_a}{\varepsilon^2}(E - aT^4). \end{cases} \tag{5.35}$$

afin de comparer les deux modèles. On obtient une solution comparable aux solutions obtenues dans

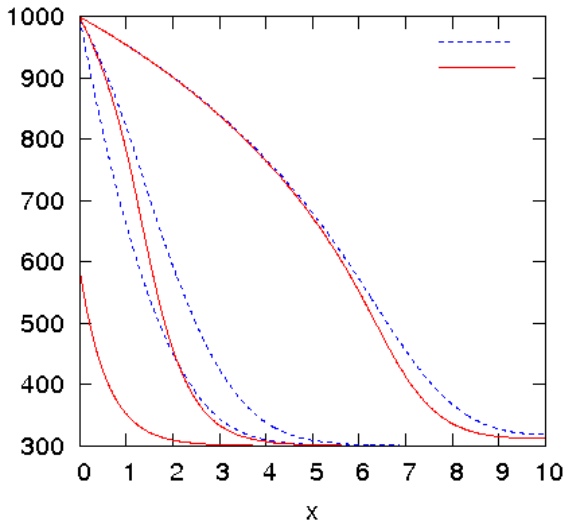


FIGURE 5.2 – Courbes de la température matière et de la température radiative dans le cas du modèle de diffusion. Les courbes bleues correspondent à la température radiative T_r. Les courbes rouges correspondent à la température matière.

les différentes études [Tur03]- [BCT08]- [B010]. Afin d'étudier le caractère « asymptotic preserving » du schéma, on observe le comportement pour un faible nombre de mailles en régimes de diffusion. On obtient un résultat du même type (voir [BCT08]- [B010]). La figure (5.4) montre qu'avec un maillage très grossier, le schéma « asymptotic preserving » capture correctement la dynamique de l'onde de chaleur. En outre les résultats sur maillages aléatoire et Cartésien sont sensiblement équivalents.

Correction géométrique du schéma GLACE

Les schémas aux noeuds basés sur le schéma GLACE [CDDL09] ne dérivent pas sur le schéma acoustique standard (formalisme au arête) lorsqu'on est sur maillage carré ou rectangulaire. En effet, le coefficient de diffusion numérique est différent pour les deux schémas. Pour les cas test précédents cette propriété empêche le schéma de donner des simulations comparables à celles réalisées par le schéma standard. En effet, au premier pas de temps, l'énergie qui rentre dépend uniquement de la diffusion numérique. Pour le schéma classique, le flux rentrant est donné par

$$\frac{h(E_b - E)}{\varepsilon(2 + \frac{\sigma}{\varepsilon}h)}$$

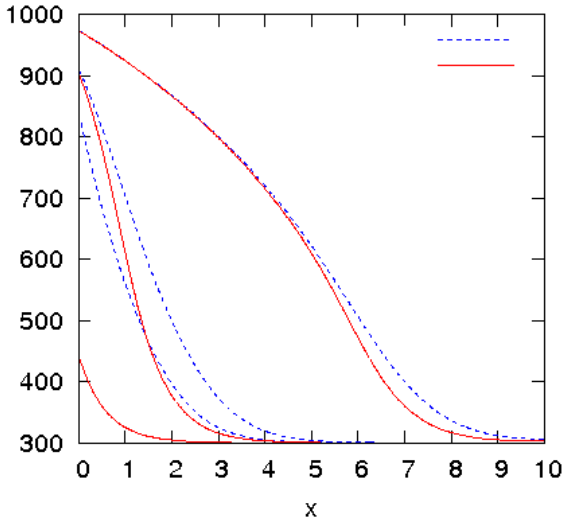


FIGURE 5.3 – Courbes de la température matière et de la température radiative dans le cas du modèle P_1. Les courbes bleues correspondent à la température radiative T_r. Les courbes rouges correspondent à la température matière.

et pour le schéma GLACE le flux entrant est donné par

$$\frac{h(E_b - E)}{\varepsilon(\sqrt{2} + \frac{\sigma}{\varepsilon}h)}.$$

avec E_b la valeur au bord et h le pas de maillage. Par conséquent, il existe un facteur constant entre l'énergie qui entre dans le domaine à la première itération entre le schéma classique et les schémas aux noeuds. Ceci a pour conséquence que les schémas « asymptotic preserving » aux noeuds basés sur le schéma GLACE surestime légèrement, par rapport à la méthode classique, l'énergie qui rentre. Le solveur classique résolvant localement un problème de Riemann exact, on le choisit comme référence pour comparer l'énergie entrante. Afin de corriger ce défaut on propose deux solutions : la première est de calculer exactement le flux à l'interface et de fixer $\mathbf{F}_r$ en conséquence, la seconde est d'utiliser un correcteur géométrique qui permet de dériver sur le schéma acoustique sur grille de carrés ou rectangulaire. Ce correcteur est défini dans [CDDL09] de la façon suivante : il suffit d'utiliser le tenseur

$$\widehat{\alpha}_{jr} = \mu_j l_{jr} \mathbf{n}_{jr} \otimes \mathbf{n}_{jr},$$

avec $\mu_j = \frac{|\sum_r |l_{jr}\mathbf{n}_{jr}||}{\lambda_j}$ et λ_j la plus grande valeur propre de la matrice $\sum_r l_{jr}\mathbf{n}_{jr} \otimes \mathbf{n}_{jr}$. Avec ce correcteur on retrouve les résultats donnés par le schéma classique.

5.5 Synthèse

Les résultats énoncés dans ce chapitre tendent à montrer que l'ajout du terme de couplage avec la matière, ne pose pas de problème particulier pour la discrétisation spatiale. En particulier, les schémas « asymptotic preserving » construits pour les modèles sans couplage reste « asymptotic preserving » au sens de la définition (1.3.3). Par contre, pour la discrétisation en temps, un travail supplémentaire est nécessaire. En effet, la raideur et la non-linéarité du système de relaxation implique l'utilisation d'une

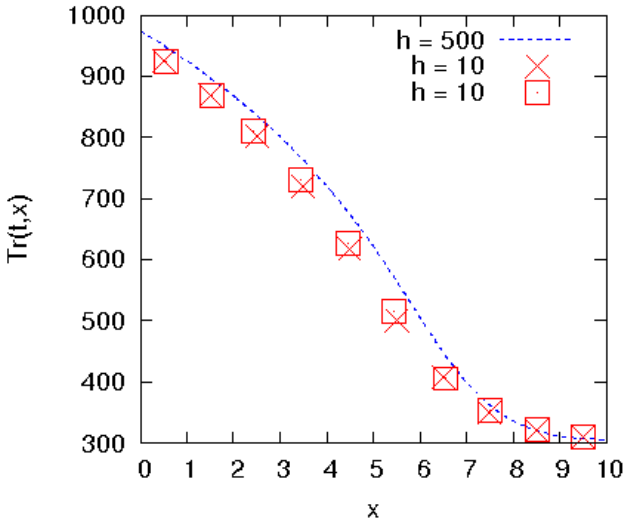


FIGURE 5.4 – Courbes de la température radiative au temps final. La croix rouge représente la solution sur un maillage Cartésien et les carrés représentent la solution sur un maillage aléatoire. h représente le pas du maillage.

discrétisation implicite couplée à une méthode de Newton. La méthode présentée dans ce chapitre, consiste à impliciter uniquement le terme de relaxation, ce qui donne une méthode de Newton locale dans chaque maille. Il serait utile de construire des méthodes complètement implicites et d'étudier l'impact des termes de couplage sur le conditionnement des matrices et sur la convergence des méthodes de Newton.

Chapitre 6

Discrétisation du modèle M_1

Le modèle M_1 est un modèle simplifié, non linéaire pour le rayonnement approchant l'équation de transport (1.6). Contrairement au modèle P_1 qui est obtenu en minimisant la norme L^2 sous contrainte de conservation des moments (la norme L^2 étant une entropie), le modèle M_1 dérive de la minimisation sous les mêmes contraintes de l'entropie physique liée aux particules. Le modèle M_1 n'est pas une bonne approximation angulaire de l'équation de transport, cependant le modèle préserve la positivité de l'énergie, le régime de transport dans le vide et le fait que les photons ne peuvent pas dépasser la vitesse de la lumière. Ces propriétés font du modèle M_1 un candidat intéressant pour approcher le problème de transport des photons. Dans un premier temps on étudie le modèle M_1 sans terme de couplage avec la matière est donné par

$$\begin{cases} \partial_t E + \dfrac{1}{\varepsilon} \operatorname{div} \mathbf{F} = 0, \\ \partial_t \mathbf{F} + \dfrac{1}{\varepsilon} \nabla \widehat{P} = -\dfrac{\sigma}{\epsilon^2} \mathbf{F}, \end{cases} \tag{6.1}$$

avec le tenseur de pression

$$\widehat{P} = \frac{1}{2} \left((1 - \chi(\mathbf{f})) \widehat{I}_d + (3\chi(\mathbf{f}) - 1) \frac{\mathbf{f} \otimes \mathbf{f}}{||\mathbf{f}||} \right) E,$$

le flux normalisé $\mathbf{f} = \mathbf{F}/E$, et le tenseur d'Eddington $\chi(\mathbf{f}) = \dfrac{3 + 4||\mathbf{f}||^2}{5 + 2\sqrt{4 - 3||\mathbf{f}||^2}}$.
Le modèle M_1 est un système hyperbolique, symétrisable. L'idée introduite dans [BD07] pour discrétiser le modèle M_1 consiste à formuler le modèle M_1 comme un système de la dynamique des gaz. Cette formulation nous permet d'utiliser un schéma Lagrange+projection aux noeuds qui est adapté à la construction de schémas « asymptotic preserving » couplé à la méthode de Jin-Levermore et de contrôler le principe du maximum à l'aide de l'inégalité d'entropie. Afin d'obtenir un schéma entropique on utilise la variante tensorielle (définie par [Klu08]) des schémas aux noeuds. La phase Lagrangienne est discrétisée à l'aide du schéma hydrodynamique GLACE [Maz07], [Des09], [CDDL09] couplé à la méthode de Jin-Levermore [JL96]. Pour la phase de projection, on utilise un schéma aux noeuds d'advection. Le schéma obtenu est « asymptotic preserving » avec un schéma limite d'ordre 1, préserve le principe du maximum, le régime de transport pur et l'inégalité d'entropie semi-discrète. Afin d'obtenir un schéma d'ordre deux en limite de diffusion, on complétera le schéma d'advection par une procédure MUSCL active seulement dans le régime de diffusion. Avant d'introduire la formulation sous la forme d'un système de la dynamique des gaz, on propose des rappels sur la construction du modèle M_1 qui seront nécessaires à l'obtention de cette formulation ainsi qu'à la démonstration de ses propriétés.

6.1 Rappels sur le modèle M_1

On part de l'équation de transfert radiatif (1.6), avec $\mathbf{\Omega}$ la direction, ν la fréquence et le nombre de photons $n = \frac{4\pi^5}{15\nu^3} I(t, \mathbf{x}, \mathbf{\Omega}, \nu)$. Cette équation est munie de l'entropie

$$S = -\frac{15}{4\pi^5} \int_{S^2} \int_0^\infty \nu^2 (n \log n - n \log(n+1)) d\mathbf{\Omega} d\nu, \tag{6.2}$$

et du flux d'entropie

$$\mathbf{Q} = -\frac{15}{4\pi^5} \int_{S^2} \int_0^\infty \nu^2 (n \log n - (n+1) \log(n+1)) \mathbf{\Omega} d\mathbf{\Omega} d\nu. \tag{6.3}$$

Le modèle M_1 est obtenu en minimisant l'entropie (6.2) sous la contrainte de conservation des moments. Soit $(E, \mathbf{F}) \in \mathbb{R}^{*,+} \times \mathbb{R}^2$ tel que $||\mathbf{F}|| \leq E$. Les contraintes de conservation s'écrivent

$$\int_{S^2} \int_0^\infty I(t, \mathbf{x}, \mathbf{\Omega}, \nu) d\mathbf{\Omega} d\nu = E, \quad \int_{S^2} \int_0^\infty I(t, \mathbf{x}, \mathbf{\Omega}, \nu) \mathbf{\Omega} d\mathbf{\Omega} d\nu = \mathbf{F}.$$

Les calculs effectués dans [BD04] montre que le minimum est atteint lorsque l'intensité radiative suit une Planckienne généralisée de paramètres T et $\mathbf{b}$ qui s'écrit

$$I(t, \mathbf{x}, \mathbf{\Omega}, \nu) = \frac{15\nu^3}{4\pi^5} \frac{1}{\exp^{\frac{\nu}{T} + \frac{\nu(\mathbf{b}, \mathbf{\Omega})}{T}} - 1},$$

avec $\frac{1}{T}$ et $\frac{\mathbf{b}}{T}$ les multiplicateurs de Lagrange et $L = S - \frac{E}{T} - \frac{\mathbf{b}}{T}\mathbf{F}$ le Lagrangien issu de la minimisation. Le paramètre T peut être associé à une température et le paramètre $\mathbf{b}$ à l'anisotropie du flux radiatif. Ce calcul de minimisation permet d'obtenir le modèle M_1 défini par

$$\begin{cases} \partial_t E + \operatorname{div} \mathbf{F} = 0, \\ \partial_t \mathbf{F} + \nabla \widehat{P} = \mathbf{0}, \end{cases} \tag{6.4}$$

avec

- $E = \frac{3 + ||\mathbf{b}||^2}{3(1 - ||\mathbf{b}||^2)^3} T^4$,
- $\mathbf{F} = -\frac{4\mathbf{b}}{3(1 - ||\mathbf{b}||^2)^3} T^4$,
- $\widehat{P} = \frac{1}{2} \left((1 - \chi(\mathbf{f})) \widehat{I}_d + (3\chi(\mathbf{f}) - 1) \frac{\mathbf{f} \otimes \mathbf{f}}{||\mathbf{f}||} \right) E$,

le flux normalisé $\mathbf{f} = \mathbf{F}/E$, et le tenseur d'Eddington $\chi(\mathbf{f}) = \dfrac{3 + 4||\mathbf{f}||^2}{5 + 2\sqrt{4 - 3||\mathbf{f}||^2}}$.

À partir de là on peut montrer que le modèle M_1 satisfait l'égalité suivante

$$T dS = dE - (\mathbf{b}, d\mathbf{F}). \tag{6.5}$$

Ces résultats et la relation thermodynamique sont explicités par exemple dans [BD04]- [BD06]- [BD07]. On introduit maintenant le modèle avec un terme source issu du « scattering » avec le « rescaling » classique

$$\begin{cases} \partial_t E + \dfrac{1}{\varepsilon} \operatorname{div} \mathbf{F} = 0, \\ \partial_t \mathbf{F} + \dfrac{1}{\varepsilon} \nabla \widehat{P} = -\dfrac{\sigma}{\varepsilon^2} \mathbf{F}. \end{cases} \tag{6.6}$$

Le modèle M_1 admet des propriétés similaires à celles de l'équation de transport que l'ont souhaite préserver au niveau discret.

Proposition 6.1. *Le modèle M_1 satisfait*

- *un principe du maximum : $E > 0, \quad ||\mathbf{f}|| = \frac{||\mathbf{F}||}{E} \leq 1$,*
- *une inégalité d'entropie $\partial_t S + \frac{1}{\varepsilon}\operatorname{div}(\mathbf{Q}) \geq 0$ avec le couple entropie concave, flux d'entropie*

$$S = \frac{4T^3}{3(1-||\mathbf{b}||^2)^3}, \quad \mathbf{Q} = -\mathbf{b}S,$$

- *un régime de diffusion quand ε tend vers zéro : $\partial_t E - \operatorname{div}\left(\frac{1}{3\sigma}\nabla E\right) = 0$,*
- *un régime de transport pur : lorsque $\sigma = 0$, $||\mathbf{f}(t=0)|| = 1$, E et $\mathbf{F}$ sont solutions d'équations d'advection de vitesse 1.*

6.2 Rappels et « formulation comme un système de la dynamique des gaz »

Afin d'utiliser les schémas aux noeuds qui ont une structure compatible avec la construction de schéma « asymptotic preserving » pour l'équation de la chaleur hyperbolique, on a utilisé l'analogie entre l'équation de la chaleur hyperbolique et les équations de l'acoustique (Euler linéarisé, sans équation d'énergie). De la même façon, pour le modèle M_1 il s'agit d'utiliser la proximité avec les équations d'Euler. On considère maintenant E et $\mathbf{F}$ comme variables primitives. On commence par définir la vitesse $\mathbf{u} = -\mathbf{b}$ et une densité artificielle ρ. Le modèle M_1 est équivalent au modèle hydrodynamique suivant :

$$\begin{cases} \partial_t \rho + \operatorname{div}(\rho\mathbf{u}) = 0, \\ \partial_t \rho\mathbf{v} + \operatorname{div}(\rho\mathbf{u}\otimes\mathbf{v}) + \nabla q = \mathbf{0}, \\ \partial_t \rho e + \operatorname{div}(\rho\mathbf{u}e + q\mathbf{u}) = 0, \\ \partial_t \rho s + \operatorname{div}(\rho\mathbf{u}s) \geq 0, \end{cases} \tag{6.7}$$

avec $S = \rho s$ (S l'entropie radiative), $\mathbf{F} = \rho\mathbf{v}$ et $E = \rho e$. Les autres quantités sont définies par

- $q = \frac{1-\chi}{2}E$,
- $\mathbf{u} = \frac{3\chi-1}{2}\frac{\mathbf{f}}{||\mathbf{f}||^2}$ avec $\mathbf{f} = \frac{||\mathbf{F}||}{E} = \frac{||\mathbf{v}||}{e}$,
- $T^4 = \frac{3(1-||\mathbf{u}||^2)^3}{3+||\mathbf{u}||^2}E = \frac{3(1-||\mathbf{u}||^2)^3}{3+||\mathbf{u}||^2}\rho e$,
- $S = \rho s = \frac{4E^{\frac{3}{4}}(1-||\mathbf{u}||^2)^{\frac{1}{4}}}{3^{\frac{1}{4}}(3+||\mathbf{u}||^2)^{\frac{3}{4}}} = \frac{4\rho^{\frac{3}{4}}e^{\frac{3}{4}}(1-||\mathbf{u}||^2)^{\frac{1}{4}}}{3^{\frac{1}{4}}(3+||\mathbf{u}||^2)^{\frac{3}{4}}}$.

Dans cette formulation la densité est artificielle, elle n'est utile que pour définir le schéma. La formulation du modèle M_1 comme un modèle de la dynamique des gaz est basée sur la définition des variables hydrodynamiques précédentes ainsi que deux relations qui relient les variables du modèle M_1 et les nouvelles quantités q et $\mathbf{u}$.

Lemme 6.2. *Les quantités du modèle M_1, E et $\mathbf{F}$ sont reliées aux quantités hydrodynamiques q et $\mathbf{u}$ par les relations suivantes*

$$\mathbf{F} = \mathbf{u}E + q\mathbf{u} \quad et \qquad \widehat{P} = \mathbf{u}\otimes\mathbf{F} + q\widehat{I_d}, \tag{6.8}$$

Démonstration.

$$\mathbf{u}(q+E) = \frac{(3\chi-1)\mathbf{F}\,|E|^2}{2E\,||\mathbf{F}||^2}\frac{(3-\chi)E}{2}.$$

On simplifie afin d'obtenir

$$\mathbf{u}(q+E) = \frac{(3\chi-1)(3-\chi)\mathbf{F}}{4||\mathbf{f}||^2}. \tag{6.9}$$

On commence par réécrire χ sous la forme suivante

$$\chi = \frac{5-2\sqrt{4-3||\mathbf{f}||^2}}{3}.$$

On obtient ainsi que le produit $(3\chi-1)(3-\chi) = 4||\mathbf{f}||^2$, ce qui, injecté dans la relation (6.9) donne le premier résultat.

La seconde identité est aussi rapide à vérifier. En effet

$$u\otimes\mathbf{F} + q\widehat{I_d} = \frac{3\chi-1}{2}\frac{\mathbf{f}\otimes\mathbf{f}}{||\mathbf{f}||^2}E + \frac{1-\chi}{2}E.$$

Il suffit de factoriser par E pour obtenir le résultat. □

On remarque que la seconde identité consiste à décomposer la pression radiative en une partie isotrope et une partie convective. Cela revient à réécrire la pression dans le référentiel de vitesse $\mathbf{u}$ où la pression devient isotrope.

Enfin on vérifie que la formulation, comme un système de la dynamique des gaz, admet une inégalité d'entropie comparable à celle des équations d'Euler ainsi qu'une relation équivalente au second principe de la thermodynamique.

Lemme 6.3. *Le modèle M_1 (6.6) satisfait*

$$ST - q - E + (\mathbf{u},\mathbf{F}) = 0 \Longleftrightarrow sT - q\tau - e + (\mathbf{u},\mathbf{v}) = 0, \tag{6.10}$$

et

$$TdS = dE - (\mathbf{u}, d\mathbf{F}) \Longleftrightarrow Tds = de - (\mathbf{u}, d\mathbf{v}) + qd\tau. \tag{6.11}$$

Démonstration. Par définition de T et de S et puisque $\chi = \frac{1+3||\mathbf{u}||^2}{3+||\mathbf{u}||^2}$ on obtient

$$ST = \frac{4E(1-||\mathbf{u}||^2)}{3+||\mathbf{u}||^2}.$$

On utilise la première équation de (6.8) afin d'obtenir

$$E + q - (\mathbf{F},\mathbf{u}) = (1-||\mathbf{u}^2||)(E+q).$$

Par définition de la pression q on obtient que $E+q-(\mathbf{F},\mathbf{u}) = (1-||\mathbf{u}^2||)(\frac{3-\chi}{2})E$. La proposition précédente implique que

$$\frac{3-\chi}{2} = \frac{4}{3+||\mathbf{u}||^2},$$

on en conclut donc que

$$ST = E + q - (\mathbf{F},\mathbf{u}). \tag{6.12}$$

En utilisant $S=\rho s$, $E=\rho e$, $\mathbf{F}=\rho\mathbf{v}$ et en divisant par ρ l'égalité (6.12), on obtient (6.10). Les relations (6.11) dérivent de la relation thermodynamique (6.5) et de la définition des variables hydrodynamiques. □

Maintenant que la formulation, comme un système de la dynamique des gaz, est rappelée, on propose de discrétiser le système (6.7). Par la suite, on introduira un schéma Lagrange+projection à un pas de temps avec le schéma GLACE défini pour l'hydrodynamique Lagrangienne et un schéma de projection où les flux sont localisés aux noeuds. Ce schéma sera couplé avec la méthode de Jin-Levermore.
En effet, dans le chapitre deux on a montré, que contrairement au schéma classique de type volumes finis, les schémas de volumes finis aux noeuds sont adaptés à la construction de schémas « asymptotic preserving ». Cette formulation permet justement d'utiliser ces schémas aux noeuds ainsi que de contrôler le principe du maximum à l'aide de l'inégalité d'entropie.

6.3 Discrétisation du système Lagrangien associé au modèle M_1

On note la dérivée matérielle $D_t = \partial_t + \mathbf{u}.\nabla$, et τ le volume spécifique. Le modèle (6.7) sous forme Lagrangienne s'écrit

$$\begin{cases} D_t\tau - \dfrac{1}{\varepsilon}\operatorname{div}(\rho\mathbf{u}) = 0, \\ \\ D_t\mathbf{v} + \dfrac{1}{\varepsilon}\nabla q = -\dfrac{\sigma}{\varepsilon^2}\mathbf{v}, \\ \\ D_t e + \dfrac{1}{\varepsilon}\operatorname{div}(e+q\mathbf{u}) = 0, \\ D_t s \geq 0. \end{cases} \tag{6.13}$$

Dans un premier temps, on introduit le schéma GLACE avec terme source. Une discrétisation du terme source localisé aux interfaces pour le modèle (6.13) paraît plus adaptée pour obtenir un schéma entropique. En effet, dans le cas de l'équation de la chaleur hyperbolique, on a réussi à montrer la stabilité uniquement pour le schéma JL-(b) qui utilise une discrétisation localisée aux interfaces du terme source. Pour conclure on propose de formuler le terme source en fonction de $\mathbf{u}$ plutôt que $\mathbf{v}$ afin d'utiliser la vitesse lagrangienne pour discrétiser le terme source. On relie les deux vitesses par $\mathbf{v} = k\mathbf{u}$ ou k est donné par $k = \dfrac{2E||\mathbf{f}||^2}{(3\chi - 1)}$.

6.3.1 Schéma pour le système Lagrangien et limite de diffusion

Le schéma utilisé est la variante tensorielle du schéma GLACE [Klu08] avec un terme source discrétisé aux interfaces :

$$\begin{cases} \rho_j \mid \Omega_j \mid, \partial_t\tau_j - \dfrac{1}{\varepsilon}\displaystyle\sum_r l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) = 0, \\ \\ \rho_j \mid \Omega_j \mid \partial_t\mathbf{v}_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r \mathbf{G}_{jr} = -\dfrac{\sigma}{\epsilon^2}\displaystyle\sum_r \widehat{\beta}_{jr}k_r\mathbf{u}_r, \\ \\ \rho_j \mid \Omega_j \mid \partial_t e_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r (\mathbf{u}_r, \mathbf{G}_{jr}) = 0, \end{cases} \tag{6.14}$$

avec les flux

$$\begin{cases} \mathbf{G}_{jr} = l_{jr}q_j\mathbf{n}_{jr} + r_{jr}\widehat{\alpha}_{jr}(\mathbf{u}_j - \mathbf{u}_r), \\ \left(\displaystyle\sum_j r_{jr}\widehat{\alpha}_{jr}\right)\mathbf{u}_r = \displaystyle\sum_j l_{jr}q_j\mathbf{n}_{jr} + \displaystyle\sum_j r_{jr}\widehat{\alpha}_{jr}\mathbf{u}_j, \end{cases} \tag{6.15}$$

et

$$k_r = \frac{2E_r||\mathbf{f}_r||^2}{(3\chi - 1)}.$$

La quantité r_{jr} correspond à la vitesse d'onde discrète. Elle doit être homogène à

$$\frac{4}{\sqrt{3}}\frac{E}{3 + ||\mathbf{u}||^2}.$$

On choisit de définir r_{jr} par $r_{jr} = r_j = \frac{4}{\sqrt{3}}\frac{E_j}{3+||\mathbf{u}_j||^2}$ ou $r_{jr} = r_r = \frac{1}{N(r)}\sum_j \frac{4}{\sqrt{3}}\frac{E_j}{3+||\mathbf{u}_j||^2}$ avec $N(r)$ le nombre de mailles communes à un noeud.
L'étude du problème de Riemann qui donne les vitesses de propagation des ondes pour ce système de la dynamique des gaz est détaillée dans [BD07]. La conservativité du schéma dérive de l'identité $\sum_j \mathbf{G}_{jr} = \mathbf{0}$.
À présent on propose de modifier le schéma GLACE à l'aide de la méthode de Jin-Levermore qui consiste à introduire l'état d'équilibre $\nabla q = -\frac{\sigma}{\varepsilon}k\mathbf{u}$ du système Lagrangien (6.13) dans les flux. On part du développement de Taylor

$$q(\mathbf{x}_j) \approx q(\mathbf{x}_r) + (\mathbf{x}_j - \mathbf{x}_r, \nabla q(\mathbf{x}_r)).$$

En utilisant l'état stationnaire de la seconde équation de (6.13) $\nabla q = -\frac{\sigma}{\varepsilon}k\mathbf{u}$, on obtient

$$q(\mathbf{x}_j) \approx q(\mathbf{x}_r) - \frac{\sigma}{\varepsilon}(\mathbf{x}_j - \mathbf{x}_r, k_r\mathbf{u}(\mathbf{x}_r)). \tag{6.16}$$

Enfin il suffit de coupler l'équivalent discret de l'équation (6.16) avec les flux du schéma GLACE (6.14)-(6.15).

Définition 6.3.1. *Le schéma GLACE AP pour le système (6.13) est défini par*

$$\begin{cases} \rho_j \mid \Omega_j \mid \partial_t \tau_j - \dfrac{1}{\varepsilon}\displaystyle\sum_r l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) = 0, \\ \\ \rho_j \mid \Omega_j \mid \partial_t \mathbf{v}_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r \mathbf{G}_{jr} = -\dfrac{\sigma}{\epsilon^2}\displaystyle\sum_r \widehat{\beta}_{jr}k_r\mathbf{u}_r, \\ \\ \rho_j \mid \Omega_j \mid \partial_t e_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r (\mathbf{u}_r, \mathbf{G}_{jr}) = 0, \end{cases} \tag{6.17}$$

avec les flux

$$\begin{cases} \mathbf{G}_{jr} = l_{jr}q_j\mathbf{n}_{jr} + r_{jr}\widehat{\alpha}_{jr}(\mathbf{u}_j - \mathbf{u}_r) - \dfrac{\sigma}{\varepsilon}k_r\widehat{\beta}_{jr}\mathbf{u}_r, \\ \left(\displaystyle\sum_j r_{jr}\widehat{\alpha}_{jr} + \dfrac{\sigma}{\varepsilon}k_r\sum_j \widehat{\beta}_{jr}\right)\mathbf{u}_r = \displaystyle\sum_j l_{jr}q_j\mathbf{n}_{jr} + \sum_j r_{jr}\widehat{\alpha}_{jr}\mathbf{u}_j. \end{cases} \tag{6.18}$$

On remarque que le principe est exactement le même que pour les schémas discrétisant l'équation de la chaleur hyperbolique. Afin d'étudier l'asymptotique du schéma (6.17), on rappelle que lorsque $||\mathbf{f}|| \to 0$

$$\frac{3\chi - 1}{2||\mathbf{f}||^2} \to \frac{3}{4}, \quad \frac{1-\chi}{2} \to \frac{1}{3}$$

et

$$\frac{3\chi - 1}{2||\mathbf{f}||^2} \in (\frac{3}{4}, 1)$$

Ces résultats peuvent se dériver à l'aide de $(3\chi-1)(3-\chi) = 4||\mathbf{f}||^2$. À présent on étudie l'asymptotique du schéma (6.17)-(6.18). La technique reste semblable à celle utilisée dans le cas linéaire.

Proposition 6.4. *Le schéma de diffusion limite du schéma Lagrangien (6.17)-(6.18) est :*

$$\begin{cases} \rho_j \mid \Omega_j \mid \partial_t e_j(t) + \sum_r \dfrac{1}{12\sigma} \left((l_{jr} E_j \mathbf{n}_{jr} - \widehat{\beta}_{jr} \mathbf{u}_r), \dfrac{\mathbf{u}_r}{E_r} \right) = 0 \\ \\ \left(\sum_j \hat{\beta}_{jr} \right) \mathbf{u}_r = \sum_j l_{jr} E_j \mathbf{n}_{jr} \end{cases} \tag{6.19}$$

avec E_r une moyenne des quantités localisées aux centres des mailles autour du noeud r.

Démonstration. On commence par réécrire le système précédent de la forme suivante

$$\begin{cases} \rho_j \mid \Omega_j \mid \partial_t \tau_j - \dfrac{1}{\varepsilon} \sum_r l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) = 0, \\ \\ \rho_j \mid \Omega_j \mid \partial_t \mathbf{v}_j + \dfrac{1}{\varepsilon} \sum_r \mathbf{G}^*_{jr} = \mathbf{0}, \\ \\ \rho_j \mid \Omega_j \mid \partial_t e_j + \dfrac{1}{\varepsilon} \sum_r (\mathbf{u}_r, \mathbf{G}_{jr}) = 0, \end{cases} \tag{6.20}$$

avec

$$\mathbf{G}^*_{jr} = l_{jr} q_j \mathbf{n}_{jr} + r_{jr} \widehat{\alpha}_{jr} (\mathbf{u}_j - \mathbf{u}_r).$$

En écrivant le développement de Hilbert pour la première équation on obtient

$$\frac{1}{\varepsilon^1} : \qquad \sum_r l_{jr}(\mathbf{u}^0_r, \mathbf{n}_{jr}) = 0,$$

$$\frac{1}{\varepsilon^0} : \qquad \rho^0_j \mid \Omega_j \mid \partial_t \tau^0_j - \sum_r l_{jr}(\mathbf{u}^1_r, \mathbf{n}_{jr}) = 0. \tag{6.21}$$

De la même façon on développe la seconde équation

$$\frac{1}{\varepsilon^1} : \qquad \sum_r l_{jr}(\mathbf{G}^{*,0}_{jr}, \mathbf{n}_{jr}) = \mathbf{0}, \tag{6.22}$$

$$\frac{1}{\varepsilon^0} : \qquad \rho^0_j \mid \Omega_j \mid \partial_t \mathbf{v}^0_j + \sum_r \mathbf{G}^{1,*}_{jr} = \mathbf{0}. \tag{6.23}$$

Puisque $\sum_r l_{jr} \mathbf{n}_{jr} = \mathbf{0}$ alors (6.22) équivaut a

$$\sum_r r^0_{jr} \widehat{\alpha}_{jr} (\mathbf{u}^0_j - \mathbf{u}^0_r) = \mathbf{0}. \tag{6.24}$$

Reste maintenant la troisième équation, on obtient les égalités suivantes

$$\frac{1}{\varepsilon^2} : \qquad \sum_r k^0_r (\widehat{\beta}_{jr} \mathbf{u}^0_r, \mathbf{u}^0_r) = 0, \tag{6.25}$$

$$\frac{1}{\varepsilon^1}: \qquad \sum_r l_{jr} q_j^0(\mathbf{u}_r^0, \mathbf{n}_{jr}) + \sum_r (r_{jr}\widehat{\alpha}_{jr}^0(\mathbf{u}_j^0 - \mathbf{u}_r^0), \mathbf{u}_r^0) - \sigma \sum_r k_r^0(\widehat{\beta}_{jr}\mathbf{u}_r^0, \mathbf{u}_r^1) \tag{6.26}$$
$$-\sigma \sum_r k_r^0(\widehat{\beta}_{jr}\mathbf{u}_r^1, \mathbf{u}_r^0) - \sigma \sum_r k_r^1(\widehat{\beta}_{jr}\mathbf{u}_r^0, \mathbf{u}_r^0) = 0,$$

$$\frac{1}{\varepsilon^0}: \qquad \rho_j^0 \mid \Omega_j \mid \partial_t e_j^0 + \sum_r l_{jr} q_j^0(\mathbf{u}_r^1, \mathbf{n}_{jr}) + \sum_r r_{jr}(\widehat{\alpha}_{jr}(\mathbf{u}_j^0 - \mathbf{u}_r^0), \mathbf{u}_r^1) + \sum_r l_{jr} q_j^1(\mathbf{u}_r^0, \mathbf{n}_{jr})$$

$$\sum_r r_{jr}^0(\widehat{\alpha}_{jr}(\mathbf{u}_j^1 - \mathbf{u}_r^1), \mathbf{u}_r^0) + \sum_r r_{jr}^1(\widehat{\alpha}_{jr}(\mathbf{u}_j^0 - \mathbf{u}_r^0), \mathbf{u}_r^0) - \sigma \sum_r k_r^0(\widehat{\beta}_{jr}\mathbf{u}_r^1, \mathbf{u}_r^1) = 0.$$

En faisant le même développement dans le solveur nodal que pour les schémas discrétisant l'équation de la chaleur hyperbolique on montre que

$$\sigma k_r^0 \left(\sum_j \hat{\beta}_{jr} \right) \mathbf{u}_r^1 + \sigma k_r^1 \left(\sum_j \hat{\beta}_{jr} \right) \mathbf{u}_r^0 = \sum_j l_{jr} q_j^0 \mathbf{n}_{jr}, \tag{6.27}$$

et que $\mathbf{u}_r^0 = \mathbf{0}$. En utilisant (6.25) on montre que $\mathbf{u}_j^0 = \mathbf{0}$. Par conséquent (6.26) équivaut à

$$\rho_j^0 \mid \Omega_j \mid \partial_t e_j^0 + \sum_r l_{jr} q_j^0(\mathbf{u}_r^1, \mathbf{n}_{jr}) - \sigma \sum_r k_r^0(\widehat{\beta}_{jr}\mathbf{u}_r^1, \mathbf{u}_r^1) = 0.$$

Pour terminer il suffit de remarquer que si $\mathbf{u}_j^0 = \mathbf{0}$ et $\mathbf{u}_r^0 = \mathbf{0}$ on obtient $q_j^0 = \frac{1}{3}E_j^0$ et

$$\begin{gathered} \rho_j^0 \mid \Omega_j \mid \partial_t e_j^0 + \sum_r \left(l_{jr} \frac{1}{3} E_j^0 \mathbf{n}_{jr} - \sigma k_r^0(\widehat{\beta}_{jr}\mathbf{u}_r^1, \mathbf{u}_r^1) \right) = 0, \\ \sigma k_r^0 \left(\sum_j \hat{\beta}_{jr} \right) \mathbf{u}_r^1 = \sum_j l_{jr} \frac{1}{3} E_j^0 \mathbf{n}_{jr}. \end{gathered} \tag{6.28}$$

En faisant un changement de variable dans le système linéaire on obtient

$$\begin{gathered} \rho_j^0 \mid \Omega_j \mid \partial_t e_j^0 + \sum_r \left((l_{jr} \frac{1}{3} E_j^0 \mathbf{n}_{jr} - \frac{1}{3}\widehat{\beta}_{jr}\mathbf{u}_r^1), \frac{1}{3k_r^0\sigma}\mathbf{u}_r^1 \right) = 0, \\ \left(\sum_j \hat{\beta}_{jr} \right) \mathbf{u}_r^1 = \sum_j l_{jr} E_j^0 \mathbf{n}_{jr}. \end{gathered} \tag{6.29}$$

La proposition 7.6 montre que $k_r = \frac{4}{3E_r}$ quand ε tend vers 0, par conséquent

$$\begin{gathered} \rho_j^0 \mid \Omega_j \mid \partial_t e_j^0 + \sum_r \frac{1}{12\sigma} \left((l_{jr} E_j^0 \mathbf{n}_{jr} - \widehat{\beta}_{jr}\mathbf{u}_r^1) \frac{1}{E_r^0}, \mathbf{u}_r^1 \right) = 0, \\ \left(\sum_j \hat{\beta}_{jr} \right) \mathbf{u}_r^1 = \sum_j l_{jr} E_j^0 \mathbf{n}_{jr}. \end{gathered} \tag{6.30}$$

□

L'équation sur E dans la reformulation comme un système de la dynamique des gaz est devenue non linéaire. Par conséquent on obtient un schéma limite de diffusion non linéaire.

6.3.2 Entropie et principe du maximum pour le schéma Lagrangien

À présent nous regardons sous quelles conditions le schéma Lagrangien semi-discret GLACE-AP (6.17)-(6.18) est entropique et vérifie le principe du maximum associé au système Lagrangien (6.13). L'entropie s du système (6.7) est donnée par

$$S = \rho s = \frac{4e^{3/4}(1-||\mathbf{u}||^2)^{\frac{1}{4}}}{3^{\frac{1}{4}}(3+||\mathbf{u}||^2)^{\frac{3}{4}}\tau^{\frac{3}{4}}}.$$

L'inégalité d'entropie associée au modèle M_1 joue un rôle important dans la construction d'un schéma préservant le principe du maximum. En effet, on remarque que la positivité de l'entropie en tout temps permet de préserver le principe du maximum associé au système Lagrangien : $e(t) > 0, 1-||\mathbf{u}(t)||^2 > 0$, $\tau(t) > 0$ et $|\Omega_j(t) > 0|$. Maintenant on montre que le schéma Lagrangien est entropique et préserve le principe du maximum qui lui est associé.

Proposition 6.5. *Soit $e_j(t) > 0$, $\rho_j(t) > 0$, $1 - ||\mathbf{u}_j(t)||^2 > 0$ et $\mid \Omega_j(t) \mid > 0$, le schéma JL-(b) pour la phase Lagrangienne est entropique au sens où*

$$\partial_t s_j = \frac{1}{\varepsilon \mid \Omega_j \mid T_j}\left(\sum_r r_{jr}(\mathbf{n}_{jr}, (\mathbf{u}_j - \mathbf{u}_r))^2 + \frac{\sigma}{\epsilon}\sum_r (k_r \hat{\beta}_{jr}\mathbf{u}_r, \mathbf{u}_r)\right) \geq 0,$$

si $\widehat{\beta}_{jr}$ est défini positif.

Démonstration. En utilisant le second principe de la thermodynamique introduit dans le lemme 6.4, on obtient la relation suivante

$$T_j ds_j = de_j - (\mathbf{u}_j, d\mathbf{v}_j) + q_j d\tau_j. \tag{6.31}$$

On va étudier l'entropie du schéma JL-(b) réécrit sous la forme

$$\begin{cases} \rho_j \mid \Omega_j \mid \partial_t \tau_j - \dfrac{1}{\varepsilon}\displaystyle\sum_r l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) = 0, \\ \rho_j \mid \Omega_j \mid \partial_t \mathbf{v}_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r \mathbf{G}_{jr} = -\dfrac{\sigma}{\epsilon^2}\displaystyle\sum_r k_r \widehat{\beta}_{jr}, \mathbf{u}_r \\ \rho_j \mid \Omega_j \mid \partial_t e_j + \dfrac{1}{\varepsilon}\displaystyle\sum_r (\mathbf{u}_r, \mathbf{G}_{jr}) = 0. \end{cases} \tag{6.32}$$

On écrit la dérivée de l'entropie terme à terme

$$\rho_j \mid \Omega_j \mid \partial_t e_j = -\sum_r l_{jr} q_j(\mathbf{n}_{jr}, \mathbf{u}_r) - \sum_r r_{jr}(\widehat{\alpha}_{jr}(\mathbf{u}_j - \mathbf{u}_r), \mathbf{u}_r) + \frac{\sigma}{\varepsilon}\sum_r (k_r \widehat{\beta}_{jr}\mathbf{u}_r, \mathbf{u}_r), \tag{6.33}$$

$$\rho_j \mid \Omega_j \mid \mathbf{u}_j \partial_t \mathbf{v}_j = -\sum_r l_{jr} q_j(\mathbf{n}_{jr}, \mathbf{u}_j) - \sum_j r_{jr}(\widehat{\alpha}_{jr}(\mathbf{u}_j - \mathbf{u}_r), \mathbf{u}_j), \tag{6.34}$$

et

$$\rho_j \mid \Omega_j \mid q_j \partial_t \tau_j = \sum_r l_{jr} q_j(\mathbf{n}_{jr}, \mathbf{u}_r). \tag{6.35}$$

Le premier terme de (6.34) est nul. On somme (6.33)-(6.34)-(6.35), on obtient

$$\rho_j \mid \Omega_j \mid T_j \partial_t s_j = \sum_r r_{jr}(\widehat{\alpha}_{jr}(\mathbf{u}_j - \mathbf{u}_r), \mathbf{u}_j - \mathbf{u}_r) + \frac{\sigma}{\varepsilon} \sum_r (k_r \widehat{\beta}_{jr} \mathbf{u}_r, \mathbf{u}_r), \tag{6.36}$$

ce qui équivaut à

$$\rho_j \mid \Omega_j \mid T_j \partial_t s_j = \sum_r r_{jr}(\mathbf{n}_{jr}(\mathbf{u}_j - \mathbf{u}_r))^2 + \frac{\sigma}{\epsilon} \sum_r (k_r \widehat{\beta}_{jr} \mathbf{u}_r, \mathbf{u}_r). \tag{6.37}$$

Si $T_j > 0$, $\rho_j > 0$ et $(\widehat{\beta}_{jr}\mathbf{u}_r, \mathbf{u}_r) \geq 0$, l'entropie est croissante. □

Proposition 6.6. *Soit $\tau_j(0) > 0$, $e_j(0) > 0$, $1 - ||\mathbf{u}_j(0)||^2 > 0$ et $\mid \Omega_j(0) \mid > 0$ alors la solution de schéma semi-discret (6.17)-(6.18) est entropique si le tenseur $\widehat{\beta}_{jr}$ est positif. En outre, le schéma semi-discret préverse le principe du maximum au sens où*

$$\tau_j(t) > 0,\ e(t) > 0,\ 1 - ||\mathbf{u}_j(t)||^2 > 0,\ \mid \Omega_j(t) \mid > 0, \quad \forall \varepsilon > 0$$

pour tout temps.

Démonstration. On réécrit le schéma sous la forme de système dynamique. Soit le vecteur

$$\mathbf{X}(t) = (\tau_1(t), ...\tau_n(t), \mathbf{v}_1(t), ..., \mathbf{v}_n(t), e_1(t)...e_n(t), \mathbf{x}_1(t), ...\mathbf{x}_{nv}(t))$$

et le domaine $D = \left\{\tau_j(t) > 0, e_j(t) > 0, 1 - ||\mathbf{u}_j(t)||^2 > 0, |\Omega_j(t)| > 0\right\}$. Les hypothèses de la proposition implique que $\mathbf{X}(0)$ appartient à D. On introduit le flux

$$\mathbf{H}(\mathbf{X}(t)) = \begin{cases} \frac{1}{\varepsilon} \sum_r l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}), \\ .. \\ -\frac{1}{\varepsilon} \sum_r \mathbf{G}_{jr} - \frac{\sigma}{\varepsilon^2} \sum_r k_r \widehat{\beta}_{jr} \mathbf{u}_r, \\ .. \\ -\frac{1}{\varepsilon} \sum_r (\mathbf{u}_r, \mathbf{G}_{jr}), \\ .. \\ \mathbf{u}_r. \end{cases}$$

Le schéma semi-discret s'écrit sous la forme suivante

$$\mathbf{X}'(t) = \mathbf{H}(\mathbf{X}(t)).$$

Une rapide étude du schéma montre que $\mathbf{H}(\mathbf{X}(t))$ est une fonction C^1 car les dérivées temporelles ne dépendent des variables que linéairement ou quadratiquement. On utilise le théorème de Cauchy-Paeno qui montre qu'il existe une la solution $\mathbf{X}(t)$ existe dans le domaine invariant D sur l'intervalle de temps $[0, T[$. À présent on souhaite utiliser le théorème du bout (voir annexe du chapitre). On définit le bout droit $(\lim t, \lim \mathbf{X}(t))$ lorsque $t \to T$. D'après le théorème des bouts, soit ce bout est vide soit il appartient à la frontière de $\mathbb{R}^+ \times D$. Si, sur l'intervalle $[0, T[$, le vecteur $\mathbf{X}(t)$ appartient à un compact, alors il existe une valeur d'adhérence ce qui empêche le bout droit d'être vide. Dans ce cas, soit $T = +\infty$ soit $\lim \mathbf{X}(t)$ lorsque $t \to T$ appartient au bord de D. Puisque, sur l'intervalle de temps $[0, T[$, $\mathbf{X}(t)$ appartient à un espace compact inclus dans D, la frontière ne peut pas être atteinte, par conséquent $T = +\infty$. Il suffit donc de montrer que sur $[0, T[$ la trajectoire $\mathbf{X}(t)$ reste dans un espace compact. Puisque $\mathbf{X}(t) \in D$ sur $[0, T[$ les quantités $e(t)$ et $\rho(t)$ et donc $T(t)$ sont positives, par

conséquent en utilisant la proposition 6.5 on montre que l'entropie est croissante. On commence par utiliser la conservativité du schéma pour montrer que $e(t)$ et $\tau(t)$ sont majorés. En effet

$$\sum_j \mid \Omega_j \mid^0 \rho_j^0 e_j(t) = \sum_j \mid \Omega_j \mid^0 \rho_j^0 e_j(0) - \int_0^t \sum_j \sum_r (\mathbf{G}_{jr}, \mathbf{u}_r). \tag{6.38}$$

Puisque $\sum_j \mathbf{G}_{jr} = \mathbf{0}$ on constate que le dernier terme de la relation (6.38) est nul et donc

$$e_j(t) \leq \frac{||e(0)||_{L^1}}{\mid \Omega_j(0) \mid \rho_j(0)}.$$

On montre aussi que $\tau_j(t)$ est majoré avec $\tau_j(t) \leq \frac{||\tau(0)||_{L^1}}{|\Omega_j(0)|\rho_j(0)}$ en utilisant le même argument. Sur l'intervalle de temps $]0,T]$ $\mathbf{X}(t) \in D$. En utilisant la proposition 6.5, on obtient que $\partial_t s_j(t) \geq 0$, ce qui implique que

$$s_j(t) = \frac{4 e_j(t)^{\frac{3}{4}} (1 - ||\mathbf{u}_j(t)||^2)^{\frac{1}{4}} \tau_j(t)^{\frac{1}{4}}}{3^{\frac{1}{4}} (3 + ||\mathbf{u}_j(t)||^2)^{\frac{3}{4}}} \geq s_j(0).$$

Puisque $e_j(t)$ et $1 - ||\mathbf{u}_j(t)||^2$ sont majorés on constate que $\tau_j(t)$ ne peut être que minoré. Par conséquent $\rho_j(t)$ est minoré et majoré. De plus

$$e_j(t)^{\frac{3}{4}} \geq \frac{3^{\frac{1}{4}} (3 + ||\mathbf{u}_j(t)||^2)^{\frac{3}{4}} \rho_j(t)^{\frac{1}{4}}}{4 (1 - ||\mathbf{u}_j(t)||^2)^{\frac{1}{4}}} s_j(0).$$

En remarquant que les fonctions de $||\mathbf{u}_j(t)||$ ainsi que la densité sont minorées on peut prouver que la quantité $e_j(t)$ est minorée. De la même façon la quantité $1 - ||\mathbf{u}_j(t)||$ est aussi minorée. Enfin on montre que le volume de maille est borné. Le volume de maille est majoré par la mesure du domaine. Pour la minoration on utilise la conservation de la masse $\mid \Omega_j(t) \mid = \mid \Omega_j(0) \mid \rho_j(0) \tau_j(t)$ et la minoration de $\tau_j(t)$. Pour conclure, les quantités $e_j(t)$, $\tau_j(t)$, $1 - ||\mathbf{u}_j(t)||$ et $\mid \Omega_j(t) \mid$ sont bornées dans l'ouvert D. La solution $\mathbf{X}(t)$ du système dynamique $\mathbf{X}^{'}(t) = \mathbf{H}(\mathbf{X}(t))$ est donc incluse dans un espace compact de D sur le temps d'existence $[0, T[$. En utilisant le théorème des bouts on montre ainsi que $T = +\infty$. Le schéma semi-discret préserve donc le principe du maximum pour tout temps. □

6.4 Étape de projection

On commencera par introduire un schéma d'advection aux noeuds pour la phase de projection, puis on étend à l'ordre deux le schéma d'advection à l'aide d'une technique MUSCL. Afin d'obtenir une discrétisation complète de (6.7) on doit ajouter une phase de projection à la phase Lagrange.

6.4.1 Rappels sur le schéma aux noeuds d'advection

On étudie l'équation d'advection suivante

$$\begin{cases} \partial_t f + \nabla.(\mathbf{a} f) = 0, \\ f(0, \mathbf{x}) = f_0(\mathbf{x}). \end{cases} \tag{6.39}$$

Afin de pouvoir définir un schéma ALE (Arbitrary Lagrange Euler) un schéma d'advection aux noeuds a été défini dans [Pa08]. On rappelle ce schéma ainsi que ses propriétés de base.

Proposition 6.7. *Le schéma upwind aux noeuds pour l'équation d'advection est*

$$\mid \Omega_j \mid \partial_t f_j + \sum_{R^+} l_{jr}(\mathbf{a}_r, \mathbf{n}_{jr}) f_j + \sum_{R^-} l_{jr}(\mathbf{a}_r, \mathbf{n}_{jr}) f_{k(r)} = 0 \tag{6.40}$$

avec $R^+ = \{r / l_{jr}(\mathbf{a}_r, \mathbf{n}_{jr}) > 0\}$, $R^- = \{r / l_{jr}(\mathbf{a}_r, \mathbf{n}_{jr}) < 0\}$

$$f_{k(r)} = \frac{\sum_{j / l_{jr}(\mathbf{a}_r, \mathbf{n}_{jr}) > 0} l_{jr}(\mathbf{a}_r, \mathbf{n}_{jr}) f_j}{\sum_{j / l_{jr}(\mathbf{a}_r, \mathbf{n}_{jr}) > 0} l_{jr}(\mathbf{a}_r, \mathbf{n}_{jr})}.$$

On rappelle maintenant les résultats théoriques associés.

Proposition 6.8. *Le schéma upwind (6.39) pour une vitesse* $\mathbf{a}$ *constante avec une discrétisation explicite en temps est stable pour les normes* $L^p(\Omega)$ *et préserve le principe du maximum sous la CFL*

$$\frac{\Delta t}{\mid \Omega_j \mid} \sum_{R^+} (\mathbf{a}, \mathbf{n}_{jr}) \leq 1.$$

En outre le schéma est convergent.

6.4.2 Projection et limite de diffusion

La phase de projection consiste à advecter à la vitesse $\mathbf{u}_r$ les différentes quantités Euleriennes du modèle (6.7). On introduit $R^+ = \{r / l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) > 0\}$ et $R^- = \{r / l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) > 0\}$. La phase de projection s'écrie

$$\left\{ \begin{array}{l} \mid \Omega_j \mid \partial_t \rho_j + \dfrac{1}{\varepsilon} \left(\displaystyle\sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) \rho_j + \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) \rho_{k(r)} \right) = 0, \\ \mid \Omega_j \mid \partial_t (\rho_j \mathbf{v}_j) + \dfrac{1}{\varepsilon} \left(\displaystyle\sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) \rho_j \mathbf{v}_j + \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) (\rho \mathbf{v})_{k(r)} \right) = \mathbf{0}, \\ \mid \Omega_j \mid \partial_t (\rho_j e_j) + \dfrac{1}{\varepsilon} \left(\displaystyle\sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) \rho_j e_j + \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) (\rho e)_{k(r)} \right) = 0, \end{array} \right. \tag{6.41}$$

avec la vitesse $\mathbf{u}_r$ donnée par les flux Lagrangien (6.18). L'étude asymptotique du solveur nodal montre que $\mathbf{u}_r^0 = \mathbf{0}$ et

$$\sigma k_r \left(\sum_j \hat{\beta}_{jr} \right) \mathbf{u}_r^1 = \sum_j l_{jr} q_j^0 \mathbf{n}_{jr}.$$

La vitesse d'advection $\mathbf{u}_r = \mathbf{u}_r^0 + \varepsilon \mathbf{u}_r^1$, est équivalente en régime de diffusion à $\mathbf{u}_r = \varepsilon \mathbf{u}_r^1$. Par conséquent l'équation sur $E = \rho e$ s'écrit

$$\mid \Omega_j \mid \partial_t E_j + \frac{1}{4\sigma} \left(\sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) \frac{E_j}{E_r} + \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) \frac{E_{k(r)}}{E_r} \right) = 0, \tag{6.42}$$

avec

$$\left(\sum_j \hat{\beta}_{jr} \right) \mathbf{u}_r = \sum_j l_{jr} E_j \mathbf{n}_{jr},$$

avec donc $\mathbf{u}_r$ homogène à $-\nabla E(\mathbf{x}_r, t)$.

Remarque 6.4.1.
- *Le coefficient de diffusion donné par la phase Lagrange est* $\frac{1}{12\sigma}$ *et le coefficient de diffusion associé à la phase de projection est* $\frac{1}{4\sigma}$. *Le schéma limite total préserve le coefficient de diffusion* $\frac{1}{3\sigma}$.
- *Le schéma de diffusion (6.42) est une réécriture de l'équation de diffusion sous forme d'une équation d'advection de vitesse* $\mathbf{a} = \nabla E/E$. *Comme tout schéma de transport le schéma upwind aux noeuds est d'ordre* 1. *Par conséquent le schéma de diffusion issu de la phase de projection, et donc le schéma de diffusion final, sont d'ordre* 1.

6.4.3 Principe du maximum pour la phase de projection

Maintenant on souhaite montré que la phase de projection préserve elle aussi le principe du maximum aussi au modèle M_1.

Proposition 6.9. *On suppose* $(\rho_j e_j)(0) = E_j(0) > 0$ *et* $||(\rho_j \mathbf{v}_j)(0) = \mathbf{F}_j(0)|| < (\rho_j e_j)(0) = E_j(0)$ *alors*

$$E_j(t) > 0 \text{ et } ||\mathbf{F}_j(t)|| < E_j(t), \forall t.$$

Le schéma d'advection semi-discret préserve le principe du maximum associé au modèle M_1.

Démonstration. Le schéma semi-discret s'écrit

$$\partial_t E_j(t) = -\left(\sum_{R^+} a^*_{jr}\right) E_j(t) - \left(\sum_{R^-} a^*_{jr}\right) E_{k(r)}(t),$$

avec $a^*_{jr} = l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr})$. On utilise

$$\left(\sum_{R^-} a^*_{jr}\right) = -\left(\sum_{R^+} a^*_{jr}\right),$$

puis on pose $\alpha_j = \sum_{R^+} a^*_{jr}$, $\alpha_i = \sum_{R^+} a^*_{jr}(\frac{a^*_{ir}}{\sum_{R^+} a^*_{ir}})$. On peut donc réécrire le schéma sous la forme

$$\partial_t E_j(t) = -\alpha_j E_j(t) + \sum_i \alpha_i E_i(t).$$

On a supposé que $E_j(0) > 0$. Par continuité il existe τ et un i_0 tels que

$$E_j(t) > 0, \text{pour tout } t \in [0, \tau[\text{ et } E_{i_0}(\tau) = 0.$$

Soit $i_0 \neq j$ par conséquent $\sum_i \alpha_i E_i > 0$ ce qui implique

$$\frac{dE_j(\tau)}{dt} > -\alpha E_j(\tau) \longrightarrow E_j(\tau) = e^{-\alpha\tau} E_j(0) > 0.$$

Dans le cas où $j = i_0$ on obtient $\frac{dE_j(t)}{dt} > 0, \quad \forall j \in [0, \tau]$ et $E_j(t) > 0, \quad \forall j \in [0, \tau[$, ce qui implique $E_j(\tau) > 0$. On obtient une contradiction. Par conséquent $E_j(\tau) > 0$. On prolonge sur des intervalles de temps $[n\tau, (n+1)\tau]$ pour conclure. En utilisant le même principe on montre que si $E_j(0)+(\mathbf{n}, \mathbf{F}_j)(0) > 0$ avec $\mathbf{n} \in \mathbb{R}^2$ alors

$$E_j(t) + (\mathbf{n}, \mathbf{F}_j)(t) > 0,$$

car cette quantité satisfait le même schéma d'advection. Reste à montrer l'équivalence entre le principe du maximum pour M_1 et la relation $E_j(t) + (\mathbf{n}, \mathbf{F}_j)(t) > 0$. Pour cela il suffit de remarquer que

$$E_j(t) + (\mathbf{n}, \mathbf{F}_j)(t) > E_j(t) - ||(\mathbf{n}, \mathbf{F}_j)(t)|| > E_j(t) - ||\mathbf{F}_j(t)||.$$

Pour le sens inverse le bon choix de $\mathbf{n}$ conclut naturellement la démonstration. □

6.4.4 Projection d'ordre 2, méthode MUSCL

Le schéma de diffusion issu de la phase de projection est un schéma d'ordre 1. En effet le schéma de diffusion peut être assimilé à un schéma d'advection de vitesse $\frac{\nabla E}{E}$. Les schémas de diffusion, ainsi que le schéma limite de la phase Lagrange étant d'ordre deux, on souhaite obtenir l'ordre deux également pour le schéma limite issu de la phase de projection. Pour cela on propose d'utiliser une technique de type MUSCL sur le schéma d'advection. La méthode MUSCL introduite par Van Leer consiste à interpoler une valeur de la solution à l'interface (ici un noeud) et à utiliser le schéma avec ces valeurs interpolées au lieu des valeurs aux centres des mailles. Cette méthode est souvent couplée avec des limiteurs de pentes qui permettent de pouvoir conserver la positivité, d'éviter de grandes disparités entre des pentes associées à une même interface et de stabiliser le schéma. La méthode MUSCL consiste donc à remplacer la quantité f_j par un état interpolé f_{jr} construit à l'aide de l'approximation d'ordre un, $f_{jr} = f_j - ((\mathbf{x}_r - \mathbf{x}_j), \nabla \mathbf{f}_r)$ où $\nabla \mathbf{f}_r$ est une approximation du gradient aux noeuds r.

Définition 6.4.1. *Le schéma d'advection MUSCL est défini par*

$$\mid \Omega_j \mid \partial_t f_j(t) + \sum_{R^+} l_{jr}(\mathbf{a}_r, \mathbf{n}_{jr}) f_{jr} + \sum_{R^-} l_{jr}(\mathbf{a}_r, \mathbf{n}_{jr}) f_{k(r)} = 0, \tag{6.43}$$

avec

$$f_{k(r)} = \frac{\sum_{j/l_{jr}(\mathbf{a}_r, \mathbf{n}_{jr})>0} l_{jr}(\mathbf{a}_r, \mathbf{n}_{jr}) f_{jr}}{\sum_{j/l_{jr}(\mathbf{a}_r, \mathbf{n}_{jr})>0} l_{jr}(\mathbf{a}_r, \mathbf{n}_{jr})}.$$

Les valeurs interpolées issues de la méthode MUSCL sont

$$\begin{cases} f_{jr} = f_j - \phi(m)((\mathbf{x}_r - \mathbf{x}_j), \nabla \mathbf{f}_r), \\ \left(\sum_j \hat{\beta}_{jr} \right) \nabla \mathbf{f}_r = \sum_j l_{jr} f_j \mathbf{n}_{jr}, \end{cases} \tag{6.44}$$

et avec la pente

$$m = \min(m_1, m_2), \quad m_1 = \frac{\min(f_i) - f_j}{\min(f_r) - f_j}, \quad m_2 = \frac{\max(f_i) - f_j}{\max(f_r) - f_j}$$

ϕ un limiteur de pente type minmod et $f_r = f_j - ((\mathbf{x}_r - \mathbf{x}_j), \nabla \mathbf{f}_r)$.

Les schémas MUSCL d'ordre deux préservent la positivité mais ne préservent pas à priori le principe du maximum. On constate numériquement que $||\mathbf{f}|| < 1$ ne peut pas être préservé par la méthode. Cependant l'analyse asymptotique montre que $||\mathbf{f}||$ est homogène à ε. Par conséquent, on propose que la reconstruction d'ordre deux intervienne uniquement en régime de diffusion. Grâce à cela on pourra conserver le principe du maximum en transport et obtenir un schéma d'ordre deux en régime de diffusion. Par exemple, le limiteur suivant $\phi(m) = \min(\sigma, 1-\varepsilon)\max(0, \min(m, 1))$, permet d'activer la procédure MUSCL seulement dans le régime de diffusion.

6.4.5 Autre schéma d'advection

Chaque schéma d'advection a des propriétés caractéristiques différentes. Par exemple la diffusion numérique du schéma aux noeuds est nulle dans la direction perpendiculaire à l'écoulement pour un écoulement oblique. Pour le schéma d'advection classique aux arêtes la diffusion numérique est nulle dans la direction perpendiculaire à l'écoulement pour un écoulement horizontal. Selon les problèmes

considérés, il peut être intéressant d'utiliser l'un ou l'autre schéma d'advection. Une alternative à la phase d'advection présentée ci-dessus est d'utiliser le schéma classique aux arêtes

$$|\Omega|\,\partial_t f_j + \sum_{k|l_{jk}(\mathbf{u}_{jk},\mathbf{n}_{jk})>0} l_{jk}(\mathbf{u}_{jk},\mathbf{n}_{jk}) f_j + \sum_{k|l_{jk}(\mathbf{u}_{jk},\mathbf{n}_{jk})>0} l_{jk}(\mathbf{u}_{jk},\mathbf{n}_{jk}) f_k = 0 \tag{6.45}$$

où $\mathbf{u}_{jk}$ la vitesse associée à l'arête $\partial\Omega_{jk}$ peut être reconstruit à partir des vitesses aux noeuds composant l'arête. Comme dans le cas du schéma aux noeuds, on peut construire un schéma d'ordre deux avec une technique MUSCL. L'utilisation d'un schéma aux arêtes en phase d'advection permet de limiter les modes parasites inhérents aux schémas aux noeuds.

6.5 Schéma Lagrange+Projection à un pas

6.5.1 Schéma Lagrange+Projection et limite de diffusion

En couplant la phase Lagrange et la phase projection on obtient le schéma Eulerien pour le modèle (6.7). Ici on propose de ne pas « splitter » les deux phases. Le schéma obtenu peut être vu comme la somme des flux Lagrangien et des flux de projection.

Définition 6.5.1. *Le schéma Lagrange+projection est*

$$\left\{\begin{array}{l} |\,\Omega_j\,|\,\partial_t\rho_j + \dfrac{1}{\varepsilon}\left(\displaystyle\sum_{R^+} l_{jr}(\mathbf{u}_r,\mathbf{n}_{jr})\rho_j + \sum_{R^-} l_{jr}(\mathbf{u}_r,\mathbf{n}_{jr})\rho_{k(r)}\right) = 0 \\ |\,\Omega_j\,|\,\partial_t(\rho_j\mathbf{v}_j) + \dfrac{1}{\varepsilon}\left(\displaystyle\sum_{R^+} l_{jr}(\mathbf{u}_r,\mathbf{n}_{jr})\rho_j\mathbf{v}_j + \sum_{R^-} l_{jr}(\mathbf{u}_r,\mathbf{n}_{jr})(\rho\mathbf{v})_{k(r)}\right) + \dfrac{1}{\varepsilon}\displaystyle\sum_r \mathbf{G}_{jr} \\ = -\displaystyle\sum_r \frac{\sigma}{\varepsilon^2}\widehat{\beta}_{jr}k_r\mathbf{u}_r \\ |\,\Omega_j\,|\,\partial_t(\rho_j e_j) + \dfrac{1}{\varepsilon}\left(\displaystyle\sum_{R^+} l_{jr}(\mathbf{u}_r,\mathbf{n}_{jr})\rho_j e_j + \sum_{R^-} l_{jr}(\mathbf{u}_r,\mathbf{n}_{jr})(\rho e)_{k(r)}\right) + \dfrac{1}{\varepsilon}\displaystyle\sum_r (\mathbf{u}_r,\mathbf{G}_{jr}) = 0 \end{array}\right. \tag{6.46}$$

avec comme flux

$$\left\{\begin{array}{l} \mathbf{G}_{jr} = l_{jr}q_j\mathbf{n}_{jr} + r_{jr}\widehat{\alpha}_{jr}(\mathbf{u}_j - \mathbf{u}_r) - \dfrac{\sigma}{\varepsilon}k_r\widehat{\beta}_{jr}\mathbf{u}_r, \\ \left(\displaystyle\sum_j r_{jr}\widehat{\alpha}_{jr} + \frac{\sigma}{\varepsilon}k_r\sum_j\widehat{\beta}_{jr}\right)\mathbf{u}_r = \displaystyle\sum_j l_{jr}q_j\mathbf{n}_{jr} + \sum_j r_{jr}\hat{\alpha}_{jr}\mathbf{u}_j, \end{array}\right. \tag{6.47}$$

$k_r = \dfrac{2||\mathbf{f}_r||^2(\rho e)_r}{(3\chi-1)}$, *où* $(\rho e)_r$ *et* $\mathbf{f}_r$ *sont des moyennes,* $R^\pm = \{r/l_{jr}(\mathbf{u}_r,\mathbf{n}_{jr}) \lessgtr 0\}$ *et*

$$\rho_{k(r)} = \frac{\sum_{j/l_{jr}(\mathbf{u}_r,\mathbf{n}_{jr})>0} l_{jr}(\mathbf{u}_r,\mathbf{n}_{jr})\rho_j}{\sum_{j/l_{jr}(\mathbf{u}_r,\mathbf{n}_{jr})>0} l_{jr}(\mathbf{u}_r,\mathbf{n}_{jr})}.$$

On définit $(\rho\mathbf{v})_{k(r)}$ *et* $(\rho e)_{k(r)}$ *de la même façon.*

Ce schéma est simplement une unification de la phase Lagrangienne et de la phase de projection.

On peut montrer que ce schéma est indépendant de la densité qui permet d'obtenir un schéma utilisant uniquement les variables radiatives E et $\mathbf{F}$. On obtient

Définition 6.5.2. *Schéma pour le modèle M_1*

$$\begin{cases} \mid \Omega_j \mid \partial_t E_j + \dfrac{1}{\epsilon}\left(\displaystyle\sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr})E_j + \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr})E_{k(r)} + \sum_r (\mathbf{u}_r, \mathbf{G}_{jr})\right) = 0 \\ \mid \Omega_j \mid \partial_t \mathbf{F}_j + \dfrac{1}{\epsilon}\left(\displaystyle\sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr})\mathbf{F}_j + \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr})\mathbf{F}_{k(r)} + \sum_r \mathbf{G}_{jr}\right) = -\dfrac{\sigma}{\varepsilon^2}\displaystyle\sum_r \widehat{\beta}_{jr} k_r \mathbf{u}_r \end{cases} \tag{6.48}$$

avec comme flux

$$\begin{cases} \mathbf{G}_{jr} = l_{jr} q_j \mathbf{n}_{jr} + r_{jr}\widehat{\alpha}_{jr}(\mathbf{u}_j - \mathbf{u}_r) - \dfrac{\sigma}{\varepsilon} k_r \widehat{\beta}_{jr}\mathbf{u}_r, \\ \left(\displaystyle\sum_j r_{jr}\widehat{\alpha}_{jr} + \frac{\sigma}{\varepsilon}k_r \sum_j \widehat{\beta}_{jr}\right)\mathbf{u}_r = \displaystyle\sum_j l_{jr} q_j \mathbf{n}_{jr} + \sum_j r_{jr}\widehat{\alpha}_{jr}\mathbf{u}_j, \end{cases} \tag{6.49}$$

$k_r = \dfrac{2||\mathbf{f}_r||^2 E_r}{(3\chi - 1)}$, *où* E_r *et* $\mathbf{f}_r$ *sont des moyennes autour d'un noeud et*

$$q_j = \frac{1-\chi}{2}E_j,\ \mathbf{u}_j = \frac{3\chi - 1}{2}\frac{\mathbf{f}_j}{||\mathbf{f}_j||^2}.$$

Proposition 6.10. *La limite de diffusion du schéma (6.48)-(6.49) est*

$$|\Omega_j|\,\partial_t E_j + \frac{1}{12\sigma}\sum_r \left((l_{jr}E_j\mathbf{n}_{jr} - \sigma k_r \widehat{\beta}_{jr}\mathbf{u}_r), \frac{\mathbf{u}_r}{E_r}\right) + \tag{6.50}$$

$$\frac{1}{4\sigma}\left(\sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr})\frac{E_j}{E_r} + \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr})\frac{E_{k(r)}}{E_r}\right) = 0.$$

avec

$$\left(\sum_j \hat{\beta}_{jr}\right)\mathbf{u}_r = \sum_j l_{jr}E_j\mathbf{n}_{jr}$$

Démonstration. On utilise un développement de Hilbert. Le développement du solveur nodal est le même que pour les démonstrations précédentes, on ne détaille pas l'analyse asymptotique, on sait que

$$\sigma k_r^0 \left(\sum_j \hat{\beta}_{jr}\right)\mathbf{u}_r^1 = \sum_j l_{jr}q_j\mathbf{n}_{jr}$$

et que $\mathbf{u}_r^0 = \mathbf{0}$. Puisque $\varepsilon > 0$ et $\mathbf{u}$ est nul à l'ordre 0, on en déduit que le signe de $l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr})$ est déterminé par le signe de la quantité $l_{jr}(\mathbf{u}_r^1, \mathbf{n}_{jr})$. On pose donc $R^+ = \left\{r | l_{jr}(\mathbf{u}_r^1, \mathbf{n}_{jr}) > 0\right\}$ et

$R^- = \big\{r | l_{jr}(\mathbf{u}_r^1, \mathbf{n}_{jr}) < 0\big\}$. À partir de cette remarque le reste de la démonstration devient classique. On étudie les termes du développement de Hilbert de la seconde équation du schéma (6.48)-(6.49)

$$\frac{1}{\varepsilon^1} : \quad \left(\sum_{R^+} l_{jr}(\mathbf{u}_r^0, \mathbf{n}_{jr})\mathbf{F}_j^0 + \sum_{R^-} l_{jr}(\mathbf{u}_r^0, \mathbf{n}_{jr})\mathbf{F}_{k(r)}^0\right) + \sum_r \mathbf{G}_{jr}^{0,*} = \mathbf{0}, \tag{6.51}$$

$$\frac{1}{\varepsilon^0} : \quad |\Omega_j|\partial_t \mathbf{F}_{j,0} + \left(\sum_{R^+} l_{jr}(\mathbf{u}_r^1, \mathbf{n}_{jr})\mathbf{F}_j^0 + \sum_{R^-} l_{jr}(\mathbf{u}_r^1, \mathbf{n}_{jr})\mathbf{F}_{k(r)}^0\right) + \sum_r \mathbf{G}_{jr}^{1,*} = \mathbf{0}. \tag{6.52}$$

Puisque $\mathbf{u}_r^0 = \mathbf{0}$ et $\sum_j l_{jr}\mathbf{n}_{jr} = \mathbf{0}$, le terme (6.51) est égal à $\sum_r r_{jr}^0 \widehat{\alpha}_{jr}\mathbf{u}_j^0 = \mathbf{0}$ et comme pour les cas précédents on peut déduire que $\mathbf{u}_j^0 = \mathbf{0}$. On étudie maintenant les termes issus du développement de Hilbert de la première équation de (6.48)-(6.49) en utilisant $\mathbf{u}_r^0$

$$\frac{1}{\varepsilon^0} : \quad |\Omega_j|\,\partial_t E_j(t) + \sum_r l_{jr} q_j^0(\mathbf{u}_r^1, \mathbf{n}_{jr}) - \sigma \sum_r k_r^0(\widehat{\beta}_{jr}\mathbf{u}_r^1, \mathbf{u}_r^1) \tag{6.53}$$

$$+ \sum_{R^+} l_{jr}(\mathbf{u}_r^1, \mathbf{n}_{jr})E_j^0 + \sum_{R^-} l_{jr}(\mathbf{u}_r^1, \mathbf{n}_{jr})E_{k(r)}^0 = 0,$$

avec

$$\sigma k_r^0 \left(\sum_j \hat{\beta}_{jr}\right)\mathbf{u}_r^1 = \sum_j l_{jr} q_j^0 \mathbf{n}_{jr}.$$

Pour terminer, en utilisant que $\mathbf{u}_j^0 = \mathbf{0}$ et $q_j^0 = \frac{1}{3}E_j^0$, on montre que (6.53) est équivalent à

$$\begin{gathered} |\,\Omega_j\,|\,\partial_t E_j^0 + \sum_r \left(l_{jr}\frac{1}{3}E_j^0\mathbf{n}_{jr} - \sigma k_r^0(\widehat{\beta}_{jr}\mathbf{u}_r^1, \mathbf{u}_r^1)\right) \\ + \left(\sum_{R^+} \frac{1}{3\sigma k_r} l_{jr}(\mathbf{u}_r^1, \mathbf{n}_{jr})E_j^0 + \sum_{R^-} \frac{1}{3\sigma k_r} l_{jr}(\mathbf{u}_r^1, \mathbf{n}_{jr})E_{k(r)}^0\right) \\ \left(\sum_j \hat{\beta}_{jr}\right)\mathbf{u}_r^1 = \sum_j l_{jr}E_j^0\mathbf{n}_{jr}. \end{gathered} \tag{6.54}$$

Puisque $k_r = \frac{4}{3E_r}$ quand ε tend vers 0 le schéma (6.54) se réécrit sous la forme

$$\begin{gathered} |\,\Omega_j\,|\,\partial_t E_j^0 + \sum_r \frac{1}{12\sigma}\left((l_{jr}E_j^0\mathbf{n}_{jr} - \widehat{\beta}_{jr}\mathbf{u}_r^1), \frac{\mathbf{u}_r^1}{E_r^0}\right) \\ + \frac{1}{4\sigma}\left(\sum_{R^+} l_{jr}(\mathbf{u}_r^1, \mathbf{n}_{jr})\frac{E_j^0}{E_r^0} + \sum_{R^-} l_{jr}(\mathbf{u}_r^1, \mathbf{n}_{jr})\frac{E_{k(r)}^0}{E_r^0}\right) = 0, \\ \left(\sum_j \hat{\beta}_{jr}\right)\mathbf{u}_r^1 = \sum_j l_{jr}E_j^0\mathbf{n}_{jr}. \end{gathered} \tag{6.55}$$

□

6.5.2 Schéma limite de diffusion : interprétation et remarques

Le schéma limite est un schéma d'ordre un (d'ordre deux, si on utilise une méthode MUSCL en phase de projection) non linéaire. Cependant, son principe de construction repose sur une analogie avec le transport. En effet, on peut écrire l'équation de diffusion comme une équation de transport

$$\partial_t E + \mathrm{div}(\mathbf{a}E) = 0, \text{ avec } \mathbf{a} = -\frac{\nabla E}{E}.$$

La partie du schéma de la phase de projection est exactement la discrétisation de l'EDP précédente. La vitesse Lagrangienne est donnée par un gradient approximé aux noeuds r et une reconstruction de E au point r et la partie divergence est discrétisée avec une schéma upwind. Le schéma upwind préservant la positivité sous la condition CFL on obtient un schéma positif d'ordre 1. En effet, le Laplacien est discrétisé par $-\sum_{R^+} E_j(-\frac{\nabla E_r}{E_r}, \mathbf{n}_{jr}) + \sum_{R^+} E_{k(r)}(-\frac{\nabla E_r}{E_r}, \mathbf{n}_{jr})$, on voit bien que l'on perd ainsi un ordre par rapport au schéma classique de diffusion en raison du ratio $\frac{E_j}{E_r} = O(h)$. Par conséquent, en utilisant une méthode MUSCL qui consiste à remplacer E_j par une valeur interpolée linéairement E_{jr}, on peut récupérer l'ordre deux attendu pour un schéma de diffusion. Le schéma de diffusion limite de la phase de projection est donc un schéma de transport d'ordre deux à la vitesse $-\frac{\nabla E}{E}$. Le schéma limite issu de la phase Lagrange peut aussi être interprété comme un schéma de transport. On part du schéma limite de la phase de projection

$$|\Omega_j|\partial_t E_j + \sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr})\frac{E_j}{E_r} + \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr})\frac{E_{k(r)}}{E_r}.$$

On peut la réécrire sous la forme

$$|\Omega_j|\partial_t E_j + \sum_{R^+}((l_{jr}E_j\mathbf{n}_{jr} - \sigma k_r\widehat{\beta}_{jr}\mathbf{u}_r), \frac{\mathbf{u}_r}{E_r}) + \sum_{R^-}((l_{jr}E_{k(r)}\mathbf{n}_{jr} - \sigma k_r\widehat{\beta}_{jr}\mathbf{u}_r), \frac{\mathbf{u}_r}{E_r}).$$

avec $\widehat{\beta}_{jr} = \phi(m)l_{jr}\mathbf{n}_{jr}\otimes(\mathbf{x}_r - \mathbf{x}_j)$ avec $\phi(m)$ le limiteur de pente. Cette reformulation proche du schéma obtenu par la phase Lagrange montre que le schéma issu de la phase Lagrange peut s'interpréter aussi comme un schéma de transport d'ordre deux où la vitesse est donnée par $-\frac{\nabla E}{E}$. Le schéma total peut être interprété comme un schéma de transport MUSCL avec la vitesse donnée $-\frac{\nabla E}{E}$.

Remarque 6.5.1.
- *Le schéma limite (6.50) est positif si le schéma (6.48)-(6.49) préserve le principe du maximum pour tout ε.*
- *On peut remarquer que la discrétisation du gradient ne change pas par rapport à celle du schéma linéaire de diffusion. Par conséquent elle est consistante (théorème 2.11).*
- *On remarque aussi que le schéma n'est pas défini si $E_r = 0$ et $\mathbf{u}_r$ (homogène au gradient) est non nul et $E_r = 0$ le schéma n'est pas défini. Ce schéma est valide pour la discrétisation de l'équation de diffusion si la quantité diffusée reste positive (transfert radiatif, écoulement en milieu poreux).*

6.5.3 Condition d'entropie pour le schéma total

Il s'agit de montrer que le schéma total est entropique et que cette propriété permet d'obtenir le principe du maximum. Préserver le principe du maximum est intéressant pour obtenir des simulations réalistes mais est aussi nécessaire afin que le schéma soit correctement défini. En effet, si $E = \rho e < 0$ la vitesse d'onde r_{jr} est négative et les matrices nodales ne sont plus inversibles. En outre, si $||\mathbf{u}|| \leq 1$ la pression et la température ne restent pas positives.

Proposition 6.11. *Si* $E_j(t) > 0$, $1 - ||\mathbf{u}_j(t)|| > 0$, $(\mathbf{x}, \widehat{\beta}_{jr}\mathbf{x}) \geq 0, \quad \forall \mathbf{x} \in \mathbb{R}^2$ *alors le schéma (6.48)-(6.49) est entropique au sens ou*

$$| \Omega_j | \partial_t S_j + \frac{1}{\varepsilon} \left(\sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) S_j + \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) S_{k(r)} \right) \geq 0. \tag{6.56}$$

Démonstration. On utilise l'équivalent du second principe de la thermodynamique pour le modèle M_1

$$TdS = dE - (\mathbf{b}, d\mathbf{F}).$$

Ce principe thermodynamique nous permet d'obtenir

$$T\partial_t S = \partial_t E - (\mathbf{u}, \partial_t \mathbf{F}). \tag{6.57}$$

On étudie maintenant le schéma semi-discret, l'équation (6.56) couplée avec (6.57) donne une quantité nommée T_s et définie par

$$T_s = | \Omega_j | \frac{1}{T_j}(\partial_t E_j - (\mathbf{u}_j, \mathbf{F}_j)) + \frac{1}{\varepsilon} \left(\sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) S_j + \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) S_{k(r)} \right). \tag{6.58}$$

On remplace les dérivées temporelles par les discrétisations spatiales à l'aide de (6.58). Les relations obtenues peuvent être scindées en deux groupes de termes, il y a ceux qui viennent de la phase Lagrangienne et ceux de la phase de projection. Par conséquent, $T_s = (L + P)$ où L les termes issus de la phase Lagrangienne et P ceux de la phase projection.

$$L = \frac{1}{T_j \varepsilon} \left(\sum_r \mathbf{G}_{jr} - \frac{\sigma}{\varepsilon} \sum_r \widehat{\beta}_{jr} \mathbf{u}_r - \sum_r (\mathbf{G}_{jr}, \mathbf{u}_r) + \sum_r l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) q_j - \sum_r l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) q_j \right).$$

Les mêmes calculs que pour la phase Lagrangienne donnent que

$$L = \frac{1}{T_j \varepsilon} \left(\sum_r r_{jr}(l_{jr}\mathbf{n}_{jr}, (\mathbf{u}_j - \mathbf{u}_r))^2 + \frac{\sigma}{\varepsilon} \sum_r (\widehat{\beta}_{jr}\mathbf{u}_r, \mathbf{u}_r) \right) - \frac{q_j}{T_j \varepsilon} \sum_r l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}).$$

À présent on étudie la phase de projection avec $-\frac{q_j}{T_j \varepsilon} \sum_r l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr})$. Ce terme est proche d'un schéma upwind mais avec q_j qui remplace $q_{k(r)}$. À partir de là on peut écrire une factorisation qui donne

$$P - \frac{q_j}{T_j \varepsilon} \sum_r l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) = \frac{1}{T_j \varepsilon} \sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) \left(T_j S_j - E_j + (\mathbf{u}_j, \mathbf{F}_j) - q_j \right)$$

$$+ \frac{1}{T_j \varepsilon} \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) \left(T_j S_{k(r)} - E_{k(r)} + (\mathbf{u}_j, \mathbf{F}_{k(r)}) - q_j \right).$$

On a deux cas : $l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr})$ est positif ou négatif. Dans le cas positif on constate que le premier terme est nul par la formule (6.12). Reste le cas du second terme

$$\sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) \left(T_j S_{k(r)} - E_{k(r)} + (\mathbf{u}_j, \mathbf{F}_{k(r)}) - q_j \right). \tag{6.59}$$

Maintenant on utilise la relation

$$(T_j S_j - E_j + (\mathbf{u}_j, \mathbf{F}_j) - q_j) = 0.$$

En retranchant cette équation qui vaut zéro à (6.59) on obtient

$$(6.59) = \frac{1}{T_j \varepsilon} \left(S_{k(r)} - S_j - E_{k(r)} - E_j + (\mathbf{u}_j, \mathbf{F}_{k(r)} - \mathbf{F}_j) \right). \tag{6.60}$$

On pose $g(\mathbf{x}_j) = S_j$ et $\mathbf{x}_j = (\mathbf{F}_j, E_j)$. On remarque $\nabla g(\mathbf{x}_j) = \frac{1}{T_j}(-\mathbf{u}_j, 1)$. Puisque on a une entropie concave alors

$$g(\mathbf{x}_k) - g(\mathbf{x}_j) - (\mathbf{x}_k - \mathbf{x}_j, \nabla g(\mathbf{x}_j)) \leq 0.$$

On voit que c'est exactement cette relation qui est visible dans l'inégalité (6.60). Par conséquent ce terme est négatif, or $(\mathbf{u}_r, \mathbf{n}_{jr}) < 0$, ce qui implique que le terme (6.59) est positif. Si $E_j(t) > 0$ et $1 - ||\mathbf{u}_j(t)||^2 > 0$ (donc $T_j(t) > 0$) et $\widehat{\beta}_{jr}$ est positif alors $T_s = (L + P)$ est positif. □

Proposition 6.12. *On suppose que* $r_{jr} = \frac{1}{N(r)} \sum_j \frac{4}{\sqrt{3}} \frac{E_j}{3+||\mathbf{u}_j||^2}$, $(\mathbf{x}, \widehat{\beta}_{jr}\mathbf{x}) > 0$ *et* $(\mathbf{x}A_r, \mathbf{x}) > V_r||\mathbf{x}||^2$ *pour* $\mathbf{x} \in \mathbb{R}^2$. *Soit* $E(0) > 0$ *et* $||\mathbf{F}(0)|| < E(0)$ *alors le schéma (6.48)-(6.49) est entropique et le principe du maximum est préservé au sens où*

$$E_j(t) > 0,\ ||\mathbf{F}_j(t)|| < E_j(t), \quad \forall \varepsilon > 0.$$

Démonstration. On réécrit le schéma sous la forme de système dynamique. Soit le vecteur

$$\mathbf{X}(t) = (E_1(t), ...E_n(t), \mathbf{F}_1(t), ..., \mathbf{F}_n(t)),$$

le domaine $D = \left\{ E(t) > 0, 1 - ||\mathbf{u}_j(t)||^2 > 0 \right\}$ et

$$\mathbf{H}(\mathbf{X}(t)) = \begin{pmatrix} -\frac{1}{\varepsilon} \left(\sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) E_j(t) + \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) E_{k(r)}(t) \right) - \frac{1}{\varepsilon} \sum_r (\mathbf{u}_r, \mathbf{G}_{jr}) \\ .. \\ -\frac{1}{\varepsilon} \left(\sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) \mathbf{F}_j(t) + \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) \mathbf{F}_{k(r)}(t) \right) - \frac{1}{\varepsilon} \sum_r \mathbf{G}_{jr} - \frac{\sigma}{\varepsilon^2} \sum_r (\widehat{\beta}_{jr}\mathbf{u}_r, \mathbf{u}_r) \end{pmatrix}$$

On a le système

$$\mathbf{X}'(t) = \mathbf{H}(\mathbf{X}(t)).$$

Une rapide étude des schémas montre que $\mathbf{H}(\mathbf{X}(t))$ est une fonction C^1 par morceaux car les dérivées temporelles ne dépendent des variables que linéairement ou quadratiquement. On utilise le théorème de Cauchy-Paeno : par conséquent on voit que $\mathbf{X}(t)$ existe et reste dans D sur $[0, T[$. À présent on souhaite utiliser le théorème du bout. On définit le bout droit $(\lim t, \lim \mathbf{X}(t))$ lorsque $t \to T$. D'après le théorème des bouts, soit ce bout est vide soit il appartient à la frontière de $\mathbb{R}^+ \times D$. Si, sur l'intervalle $[0, T[$, $\mathbf{X}(t)$ appartient à un compact de D, alors il existe une valeur d'adhérence, ce qui empêche le bout d'être vide. Dans ce cas, soit $T = +\infty$ soit $\lim \mathbf{X}(t)$ lorsque $t \to T$ appartient au bord de D.
Si, sur l'intervalle $[0, T[$ $\mathbf{X}(t)$, appartient à un compact inclus dans D, la frontière ne peut pas être atteinte, par conséquent $T = +\infty$. Il suffit donc de montrer que sur $[0, T[$ la trajectoire reste dans un compact. On commence par remarquer que $||\mathbf{F}_j(t)|| < E_j(t)$ est équivalent à $1 - ||\mathbf{u}_j(t)||^2 > 0$. Puisque $\mathbf{X}(t) \in D$ sur $[0, T[$ on a $E_j(t)$ et $1 - ||\mathbf{u}_j(t)||^2$ sont donc positifs par conséquent, la proposition 6.11 est satisfaite.

Majoration des quantités centrées

On commence par majorer $E_j(t)$ et $1-||\mathbf{u}_j(t)||^2$. Il est évident que $1-||\mathbf{u}_j(t)||^2<1$. Pour majorer la quantité $E_j(t)$ on utilise la conservativité du schéma.

$$\sum_j \mid \Omega_j \mid E_j(t) = \sum_j \mid \Omega_j \mid E_j(0) - \int_0^t \sum_j \sum_r (G_{jr}, \mathbf{u}_r) - \int_0^t \sum_j \left(\sum_{R^+} l_{jr}(\mathbf{n}_{jr}, \mathbf{u}_r) E_j + \sum_{R^-} l_{jr}(\mathbf{n}_{jr}, \mathbf{u}_r) E_{k(r)} \right).$$

Par conservativité on voit que les termes de flux sont nuls. Par conséquent on voit que

$$E_j(t) \leq \frac{||E(0)||_{L^1}}{\mid \Omega_j \mid}.$$

Afin d'obtenir une minoration de $E_j(t)$ et $1-||\mathbf{u}_j(t)||^2$.

Minoration de l'entropie

Pour cela on souhaite minorer l'entropie $S_j(t)$ du modèle M_1. La proposition 6.11 montre que $S_j(t)$ est solution de

$$\mid \Omega_j \mid \partial_t S_j(t) + \frac{1}{\epsilon} \left(\sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) S_j(t) + \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) S_{k(r)}(t) \right) \geq 0, \tag{6.61}$$

on pose $a_j(t) = \frac{1}{\varepsilon} \sum_{R^+} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr})$ et

$$a_i = -\frac{1}{\varepsilon} \sum_{R^-} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) \frac{l_{ir}(\mathbf{u}_r, \mathbf{n}_{ir})}{\sum_{R^+} l_{ir}(\mathbf{u}_r, \mathbf{n}_{ir})}.$$

Les coefficients a_i et a_j sont positifs. Par conséquent la condition d'entropie (6.61) s'écrit

$$\partial_t S_j(t) \geq -a_j S_j(t) + \sum_i a_i S_i(t).$$

Puisque $E_j(t)>0$ et $1-||\mathbf{u}_j(t)||^2>0$ sur $[0,T[$ l'entropie est positive dans chaque maille on obtient donc que $S_j(t) \geq e^{-a_j t} S_j(0)$. Par conséquent si $||\mathbf{u}_r|| \leq \varepsilon C$, le coefficient a_j est majoré et l'entropie est minorée sur $[0,T[$.

Majoration de $\mathbf{u}_r$ sur $[0,T[$

On part du solveur nodal avec $r_{jr} = r_r = \frac{1}{N(r)} \sum_j r_j$ et $r_j = \frac{4}{\sqrt{3}} \frac{E_j}{3+||\mathbf{u}_j||^2}$. On multiplie par $\mathbf{u}_r$ le solveur nodal. Le second membre est majoré. On obtient

$$\sum_j (q_j l_{jr} \mathbf{n}_{jr}, \mathbf{u}_r) \leq C_1 \sum_j |q_j| ||\mathbf{u}_r||,$$

avec $C_1 = \max_j(||l_{jr}\mathbf{n}_{jr}||)$ et

$$\sum_j r_r(\mathbf{u}_j \widehat{\alpha}_{jr}, \mathbf{u}_r) \leq C_2 r_r ||\mathbf{u}_r||,$$

avec $C_2 = \max_j(||\widehat{\alpha}_{jr}||)$ car $||\mathbf{u}_j|| < 1$. Maintenant on minore les matrices. On obtient

$$\left(\mathbf{u}_r, (\frac{\sigma k_r}{\varepsilon} \sum_j \widehat{\beta}_{jr}) \mathbf{u}_r \right) \geq \frac{\sigma k_r}{\varepsilon} V_r ||\mathbf{u}_r||^2,$$

et

$$\left(\mathbf{u}_r, (r_r \sum_j \widehat{\alpha}_{jr}) \mathbf{u}_r \right) \geq r_r \lambda_{jr} ||\mathbf{u}_r||^2,$$

avec λ_{jr} la valeur propre minimale de la matrice $\sum_j \widehat{\alpha}_{jr}$. En associant les minorations et les majorations précédentes, on majore $||\mathbf{u}_r||$ par

$$||\mathbf{u}_r|| \leq C \frac{1}{r_r \lambda_{jr} + \frac{\sigma k_r}{\varepsilon} V_r} \left(C_1 \sum_j |q_j| + C_2 r_r \right). \tag{6.62}$$

Puisque $r_r = \sum_j \frac{4}{\sqrt{3}} \frac{E_j(t)}{3+||\mathbf{u}_j(t)||^2}$ et $||\mathbf{u}_j(t)|| < 1$ on obtient les bornes $\frac{E_r}{4} < r_r < \frac{E_r}{3}$. Par ailleurs sur l'intervalle de temps $[0, T[$, on sait que $||\mathbf{u}_j(t)|| < 1$ qui implique que $\frac{1}{3} < \chi < 1$. Par conséquent

$$0 < |q_j| < \frac{1}{3} E_j, \quad 1 < k_r < \frac{4}{3} E_r.$$

Ces bornes permettent d'obtenir

$$||\mathbf{u}_r|| \leq C \frac{4}{(\lambda_{jr} + \frac{\sigma}{\varepsilon} V_r) E_r} \left(C_1 \sum_j |q_j| + C_2 r_r \right). \tag{6.63}$$

Pour conclure on utilise $\sum_j |q_j| < \frac{1}{3} \sum_j E_j$, ce qui donne

$$||\mathbf{u}_r|| < C \frac{4}{3\lambda_{jr} + \frac{\sigma}{\varepsilon} V_r}.$$

Cette majoration peut se réécrire sous la forme

$$||\mathbf{u}_r|| < C \frac{4\varepsilon}{3\lambda_{jr}\varepsilon + \sigma V_r}.$$

Une remarque : les majorations sont valides pour différentes définitions de la moyenne E_r. Ici on considère une moyenne arithmétique.

Minoration des quantités centrées

Puisque la vitesse nodale $\mathbf{u}_r$ est majorée par une constante homogène à ε on sait que $S_j(t) \geq e^{-a_j t} S_j(0) > 0$. On commence par remarquer que $\frac{(1-||\mathbf{u}_j(t)||^2)^{\frac{1}{4}}}{(3+||\mathbf{u}_j||^2)^{\frac{3}{4}}}$ est borné par $\frac{1}{3^{\frac{3}{4}}}$ alors par définition de l'entropie

$$E_j \geq 3^{\frac{3}{4}} (e^{-a_j(t)t} S_j(0))^{\frac{4}{3}}.$$

En reprenant la définition de l'entropie et la majoration de E on obtient que

$$(1 - ||\mathbf{u}_j||^2)^{\frac{1}{4}} \geq \frac{3^{\frac{1}{4}} S_j(0) (3 + ||\mathbf{u}_j||^2)^{\frac{3}{4}}}{4 E_j^{3/4}},$$

$$(1 - ||\mathbf{u}_j||^2)^{\frac{1}{4}} \geq \frac{3^{\frac{1}{4}} S_j(0) \mid \Omega_j \mid^{3/4}}{4 ||E_j(0)||_{L^1}^{3/4}} (3 + ||\mathbf{u}_j||^2)^{\frac{3}{4}} \geq \frac{9 S_j(0) \mid \Omega_j \mid^{3/4}}{||E_j(0)||_{L^1}^{3/4}}.$$

On a ainsi montré que les quantités $1 - ||\mathbf{u}_j(t)||^2$ et $E_j(t)$ sont incluses dans un compact. Par conséquent en appliquant le théorème des bouts on obtient que $\mathbf{X}(t)$ existe en tout temps et reste dans le domaine D pour tout temps. Il reste juste à montrer que $(1 - ||\mathbf{u}_j||^2) > 0$ est équivalent au fait que $||\mathbf{F}_j|| < E_j$. La définition de $\mathbf{u}_j$ montre que

$$(1 - ||\mathbf{u}_j||^2) > 0,$$

équivaut à

$$E_j^2 - ||\mathbf{F}_j||^2 \geq \left(\frac{3\chi - 1}{2||\mathbf{f}_j(t)||^2} - 1 \right) ||\mathbf{F}_j||^2.$$

Puisque $\frac{3\chi-1}{2||\mathbf{f}_j(t)||^2} \leq 1$ alors

$$E_j^2 - ||\mathbf{F}_j||^2 > 0.$$

□

Remarque 6.5.2. *Pour le cas où $r_{jr} = r_j = \frac{4}{\sqrt{3}} \frac{E_j}{3+||\mathbf{u}_j||^2}$ les arguments précédents sont valides sauf pour la majoration de $\mathbf{u}_r$. En effet la matrice*

$$\sum_j r_j \widehat{\alpha}_{jr}$$

est plus difficile à minorer sans supposer que les vitesses d'onde r_j sont elles mêmes-minorées ce que l'on ne peut supposer. Dans le cas où $\sigma > 0$ on peut arriver au résultat final. En effet dans ce cas, on peut minorer la matrice totale

$$\sum_j r_j \widehat{\alpha}_{jr} + \frac{\sigma k_r}{\varepsilon} \sum_j \widehat{\beta}_{jr}$$

par la minoration de $\frac{\sigma k_r}{\varepsilon} \sum_j \widehat{\beta}_{jr}$. On peut dérouler la démonstration précédente sans difficulté. Par conséquent, pour $r_{jr} = r_j$, le schéma (6.48)-(6.49) préserve la principe du maximum si $(\mathbf{x}, \widehat{\beta}_{jr}\mathbf{x}) > 0$ et $(\mathbf{x}A_r, \mathbf{x}) > V_r||\mathbf{x}||^2$ pour $\mathbf{x} \in \mathbb{R}^2$ et $\sigma > 0$.

Ces résultats montrent que le choix du tenseur $\widehat{\beta}_{jr}$ est essentiel pour préserver le principe du maximum. Par conséquent on souhaite proposer un tenseur $\widehat{\beta}_{jr}$ positif sur le plus de maillages possibles. Sauf dans le cas d'un maillage Cartésien, le tenseur $\widehat{\beta}^I_{jr} = l_{jr}\mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)$ n'est pas défini positif. Quant au tenseur $\widehat{\beta}^{II}_{jr} = V_{jr}\widehat{I}_d$ il génère une matrice nodale qui ne permet d'obtenir un flux $\mathbf{F}_r$ consistant avec le gradient. Par conséquent on souhaite trouver une décomposition de la matrice A_r où le tenseur est défini positif. Nous proposons deux nouvelles variantes. La première est basée sur la remarque suivante

$$A_r = \sum_j \left(|\widetilde{V}_{jr}|\widehat{I}_d + \frac{1}{2}(\mathbf{w}^{\perp}_{j-1/2} \otimes \mathbf{w}_{j-1/2} - \mathbf{v}^{\perp}_{j+1/2} \otimes \mathbf{v}_{j+1/2}) \right) = \sum_j \left(|\widetilde{V}_{jr}|\widehat{I}_d + P_j \right) \tag{6.64}$$

et sur le fait que chaque terme de cette somme est défini positif sous des conditions géométriques. En conséquent on souhaite utiliser le tenseur suivant

$$\widehat{\beta}^{III}_{jr} = \left(|\widetilde{V}_{jr}|\widehat{I}_d + \frac{1}{2}(\mathbf{w}^{\perp}_{j-1/2} \otimes \mathbf{w}_{j-1/2} - \mathbf{v}^{\perp}_{j+1/2} \otimes \mathbf{v}_{j+1/2}) \right).$$

la somme de ce tenseur donne la matrice $A_r = \sum_j l_{jr}\mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)$, on assure donc la consistance du gradient. La deuxième variante consiste simplement à répartir la matrice A_r de telle façon que chaque bout soit défini positif, on défini donc

$$\widehat{\beta}^{IV}_{jr} = \frac{\widetilde{V}_{jr}}{V_r} A_r.$$

Ce tenseur permet de garder la consistance du gradient et de définir un tenseur positif si la matrice A_r l'est.

Définition 6.5.3. *Les tenseurs possibles de diffusion sont*

- $\widehat{\beta}^{I}_{jr} = l_{jr}\mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)$,
- $\widehat{\beta}^{II}_{jr} = \widetilde{V}_{jr}\widehat{I}_d$,
- $\widehat{\beta}^{III}_{jr} = \Big(|\widetilde{V}_{jr}|\widehat{I}_d + \frac{1}{2}(\mathbf{w}^{\perp}_{j-1/2} \otimes \mathbf{w}_{j-1/2} - \mathbf{v}^{\perp}_{j+1/2} \otimes \mathbf{v}_{j+1/2})\Big)$,
- $\widehat{\beta}^{IV}_{jr} = \frac{\widetilde{V}_{jr}}{V_r}A_r$.

6.6 Discrétisation semi-implicite en temps

Comme pour l'équation de la chaleur hyperbolique, on souhaite définir un schéma semi-implicite avec une CFL optimale (somme de la CFL hyperbolique et de la CFL parabolique). On part pour cela du schéma (6.48)-(6.49) pour le modèle M_1, en introduisant la matrice

$$M_r = \left(\sum_j r_{jr} \widehat{\alpha}_{jr} + \frac{k_r \sigma}{\varepsilon} \sum_j \widehat{\beta}_{jr} \right)^{-1} \left(\sum_j r_{jr} \widehat{\alpha}_{jr} \right),$$

on obtient

$$\mathbf{u}_r = M_r \mathbf{u}_r^*,$$

avec $\mathbf{u}_r^*$ la solution du solveur classique du schéma GLACE défini par $\left(\sum_j r_{jr}\widehat{\alpha}_{jr}\right) \mathbf{u}_r = \sum_j l_{jr} E_j \mathbf{n}_{jr} + \sum_j r_{jr}\widehat{\alpha}_{jr}\mathbf{F}_j$. À partir des manipulations algébriques détaillées dans le chapitre 2, on peut formuler le schéma (6.48)-(6.49) de façon à obtenir un terme source local et un schéma semi-implicite.

Définition 6.6.1. *Le schéma semi-implicite en temps pour le modèle M_1 est donné par*

$$\left\{ \begin{array}{l} \mid \Omega_j \mid \dfrac{E_j^{n+1} - E_j^n}{\Delta t} + \dfrac{1}{\varepsilon} \left(\displaystyle\sum_{R^+} l_{jr}(M_r \mathbf{u}_r^*, \mathbf{n}_{jr}) E_j^n + \sum_{R^-} l_{jr}(M_r \mathbf{u}_r^*, \mathbf{n}_{jr}) E_{k(r)}^n \right) + \dfrac{1}{\varepsilon} \displaystyle\sum_r (M_r \mathbf{u}_r^*, \mathbf{G}_{jr}) = 0, \\ \mid \Omega_j \mid \dfrac{\mathbf{F}_j^{n+1} - \mathbf{F}_j^n}{\Delta t} + \dfrac{1}{\varepsilon} \left(\displaystyle\sum_{R^+} l_{jr}(M_r \mathbf{u}_r^*, \mathbf{n}_{jr}) \mathbf{F}_j^n + \sum_{R^-} l_{jr}(M_r \mathbf{u}_r^*, \mathbf{n}_{jr}) \mathbf{F}_{k(r)}^n \right), \\ + \dfrac{1}{\varepsilon} \displaystyle\sum_r \mathbf{G}_{jr}^* = -\dfrac{1}{\varepsilon} \left(\displaystyle\sum_r r_{jr}^n \widehat{\alpha}_{jr} (\widehat{I}_d - M_r) \right) \mathbf{u}_j^{n+1}, \end{array} \right. \tag{6.65}$$

avec comme flux

$$\left\{ \begin{array}{l} \mathbf{G}_{jr} = l_{jr} q_j^n \mathbf{n}_{jr} + r_{jr}^n \hat{\alpha}_{jr} (\mathbf{u}_j^n - M_r \mathbf{u}_r^*) - \dfrac{\sigma}{\epsilon} k_r^n \hat{\beta}_{jr} M_r \mathbf{u}_r^*, \\ \mathbf{G}_{jr}^* = l_{jr} q_j^n \mathbf{n}_{jr} + r_{jr}^n \hat{\alpha}_{jr} M_r (\mathbf{u}_j^n - \mathbf{u}_r^*), \\ \left(\displaystyle\sum_j r_{jr}^n \hat{\alpha_{jr}} \right) \mathbf{u}_r^* = \displaystyle\sum_j l_{jr} q_j^n \mathbf{n}_{jr} + \sum_j r_{jr}^n \hat{\alpha}_{jr} \mathbf{u}_j, \end{array} \right. \tag{6.66}$$

Les différentes quantités sont : $k_r^n = \dfrac{2||\mathbf{f}_r^n||^2 E_r^n}{(3\chi - 1)}$, ou E_r^n et $\mathbf{f}_r^n$ sont des moyennes des quantités E et $\mathbf{F}$ autour du noeud r, $R^{\pm} = \{r / l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) \lessgtr 0\}$ et

$$E_{k(r)}^n = \frac{\sum_{j/l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) > 0} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) E_j^n}{\sum_{j/l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr}) > 0} l_{jr}(\mathbf{u}_r, \mathbf{n}_{jr})}$$

On définit $\mathbf{F}_{k(r)}^n$ de la même façon. La matrice M_r est définie par

$$M_r = \left(\sum_j r_{jr}^n \widehat{\alpha}_{jr} + \frac{\sigma}{\varepsilon} k_r^n \sum_j \widehat{\beta}_{jr} \right)^{-1} \left(\sum_j r_{jr}^n \widehat{\alpha}_{jr} \right)$$

En remarquant que $M_r = O(\varepsilon)$, $\sum_r r_{jr}^n \widehat{\alpha}_{jr}(\widehat{I}_d - M_r) = O(\frac{\sigma \sum_r \widehat{\beta}_{jr} M_r k_r^n}{\varepsilon})$ et $\sum_r \widehat{\beta}_{jr} = |\Omega_j \mid \widehat{I}_d$ on constate que $\sum_r r_{jr}^n \widehat{\alpha}_{jr}(\widehat{I}_d - M_r)$ est donc homogène à une quantité $Mk|\Omega|$ associée à la maille j. Le terme source est donc homogène à $M_j|\Omega_j \mid \mathbf{F}_j^n$ car $\mathbf{u}_j^n = \frac{\mathbf{F}_j^n}{k_j^n}$.
Les tests numériques montre que la CFL du schéma semi-implicite (6.65)-(6.66) est indépendante de ε.

6.7 Résultats numériques

6.7.1 Résultats numériques pour les schémas limites

Pour commencer, on propose de comparer le schéma limite de la phase Lagrange avec les différents tenseurs $\widehat{\beta}_{jr}$ afin de sélectionner le tenseur le plus pertinent. Ensuite on donnera les résultats numériques du schéma de diffusion limite total pour le tenseur choisi.

Cas test : Solution fondamentale

On prend le même cas test que pour le schéma de diffusion linéaire. Soit $SF(t)$ la solution fondamentale de la chaleur avec le coefficient de diffusion $\frac{1}{3}$. Les conditions initiales sont définies par $E(t=0) = SF(0.01)$, $\mathbf{F}(t=0) = \mathbf{0}$. On compare la solution numérique avec la solution théorique de diffusion $E(t) = SF(0.01 + T_f)$, $\mathbf{F}(t) = \mathbf{0}$. Deux aspects nous intéressent, l'ordre 2 et la préservation de la positivité. **Résultats pour le schéma issu de la phase Lagrange**
On rappelle le schéma limite de la phase Lagrangienne.

Définition 6.7.1. *Le schéma limite issu de la phase Lagrange est*

$$\begin{cases} \partial_t E_j^0(t) + \displaystyle\sum_r \frac{1}{12\sigma}\left((l_{jr}E_j\mathbf{n}_{jr} - \sigma\widehat{\beta}_{jr}\mathbf{u}_r), \frac{\mathbf{u}_r}{E_r}\right) = 0 \\ \sigma\left(\displaystyle\sum_j \hat{\beta}_{jr}\right)\mathbf{u}_r = \displaystyle\sum_j l_{jr}E_j\mathbf{n}_{jr} \end{cases} \tag{6.67}$$

Remarque 6.7.1. *Le nombre de coefficients négatifs est calculé sur un maillage de 1600 mailles et sur 310 itérations en temps environ.*

Le tableau (6.1) donne le nombre de coefficients négatifs pour chaque variante tensorielle et chaque maillage. Les tableaux (6.2)-(6.3) donnent les ordres de convergence pour chaque tenseur et chaque maillage.

Ces résultats numériques montrent que seul le tenseur IV préserve la positivité et la consistance sur tous les maillages proposés. Les résultats pour les tenseurs I et II étaient attendus. En effet le tenseur I n'est pas défini positif, ce qui est nécessaire pour préserver la positivité, mais aussi la convergence. Quant au tenseur II, il ne permet pas de construire un gradient consistant comme cela a été déjà montré dans le chapitre deux. Le tenseur III, vient de la condition géométrique de positivité de la matrice $\sum_j l_{jr}\mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)$ introduite dans le chapitre deux. Puisque le tenseur III n'est pas positif sur tous les maillages alors que le tenseur IV l'est, cela montre que cette condition géométrique était bien sous optimale. Le tenseur IV paraît donc le plus le pertinent car il permet d'obtenir la matrice $\sum_j l_{jr}\mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)$ qui génère un gradient consistant et reste positif tant que la matrice est positive.
Résultat pour le schéma limite du schéma Lagrange+projection avec MUSCL
Pour terminer on observe la combinaison du schéma limite de la phase Lagrange avec un coefficient

maillage variante du tenseur $\widehat{\beta}_{jr}$	I	II	III	IV
Cartesien	220	0	0	0
maillage bruité quad	66845	0	0	0
maillage smooth quad	58834	0	0	0
maillage triang 1	41643	0	0	0
maillage triang 2	744178	0	0	0
maillage triang bruité	64135	0	0	0
Kershaw K=1	26019	0	1976656	0

TABLE 6.1 – Nombre de coefficients négatifs pour le schéma de diffusion limite Lagrangien pour chaque tenseur $\widehat{\beta}_{jr}$.

maillage/ variante du tenseur $\widehat{\beta}_{jr}$	I	II	III	IV
Cartésien 40-80	1.87	1.83	1.83	1.83
Cartésien 80- 160	1.92	1.93	1.93	1.93
Cartésien 160- 320	1.97	1.98	1.98	1.98
Aléatoire quad. 40-80	2.1	1.64	1.84	1.94
Aléatoire quad. 180- 160	Nc	1.28	1.98	1.99
Aléatoire quad. 160- 320	Nc	0.70	1.96	1.96
« Smooth » 40-80	2	1.75	1.71	1.73
« Smooth » 80- 160	nc	1.86	1.84	1.86
« Smooth » 160- 320	nc	1.95	1.94	1.95
« Kershaw » 40-80	nc	1.92	nc	1.91
« Kershaw » 80- 160	nc	1.88	nc	1.88
« Kershaw » 160- 320	nc	1.94	nc	1.94

TABLE 6.2 – Ordre de convergence sur maillages quadrangulaires. Lorsque l'erreur explose on note *nc* pour non-convergence.

maillage / variante du tenseur $\widehat{\beta}_{jr}$	I	II	III	IV
Triangulaire 1 40-80	1.11	1.88	1.2	1.88
Triangulaire 1 80- 160	1.04	1.95	1.08	1.95
Triangulaire 1 160- 320	nc	1.98	1.02	1.98
Triangulaire 2 40-80	nc	1.85	1.84	1.85
Triangulaire 2 180- 160	nc	1.95	1.92	1.95
Triangulaire 2 160- 320	nc	1.98	1.96	1.98
Triangulaire alea. 40-80	nc	1.82	1.76	1.82
Triangulaire alea. 80- 160	nc	1.86	1.69	1.86
Triangulaire alea. 160- 320	nc	1.66	1.53	1.66

TABLE 6.3 – Ordre de convergence sur maillages triangulaires. Lorsque l'erreur explose on note *nc* pour non-convergence

$\frac{1}{4}$ et du schéma de projection MUSCL avec un coefficient $\frac{3}{4}$. Il s'agit du schéma limite du schéma (6.48)-(6.49). On utilise la variante IV du tenseur

$$\widehat{\beta}_{jr} = \sum_j V_{jr} \frac{A_r}{V_r},$$

avec $A_r = \sum_j l_{jr} \mathbf{n}_{jr} \otimes (\mathbf{x}_r - \mathbf{x}_j)$

maillage	ordre	nb coef $E_j < 0$
Cartésien 40-80	1.92	0
Cartésien 1 80- 160	1.95	0
Cartésien 1 160- 320	1.98	0
Aléatoire quad. 40-80	2.20	0
Aléatoire quad. 180- 160	2.05	0
Aléatoire quad. 160- 320	1.97	0
« Smooth » 40-80	1.78	0
« Smooth » 80- 160	1.87	0
« Smooth » 160- 320	1.94	0
« Kershaw » 40-80	2.00	0
« Kershaw » 80- 160	1.86	0
« Kershaw » 160- 320	1.93	0

TABLE 6.4 – Ordre de convergence sur maillages quadrangulaires

maillage	ordre	nb coef $E_j < 0$
Triangulaire 1 40-80	2.1	0
Triangulaire 1 80- 160	2.24	0
Triangulaire 1 160- 320	2.13	0
Triangulaire 2 40-80	2.01	0
Triangulaire 2 180- 160	2.18	0
Triangulaire 2 160- 320	2.1	0
Triangulaire alea. 40-80	2.13	0
Triangulaire alea. 80- 160	2.21	0
Triangulaire alea. 160- 320	2.00	0

TABLE 6.5 – Ordre de convergence sur maillages triangulaires

Ces résultats montrent que le schéma limite du schéma (6.48)-(6.49) est bien positif et d'ordre deux. L'ajout d'une méthode MUSCL permet bien d'obtenir un schéma limite d'ordre deux pour la phase de projection. Des tests sur des maillages de Kershaw plus déformés que celui présenté montre que le tenseur $\widehat{\beta}_{jr}^{IV}$ reste défini positif et que le schéma reste positif et convergent.

6.7.2 Résultats numériques pour le modèle M_1

Cas test de transport

Dans un premier temps on regarde le schéma pour M_1 sans terme source. On étudie un cas test sur ce que l'on appelle le régime libre, c'est à dire quand $||\mathbf{f}|| = 1$ et $\sigma = 0$. Dans ce cas la pression q est nulle et finalement le modèle M_1 va se comporter comme l'équation de transport classique.

test 1

On commence par un cas test d'advection d'un créneau, d'abord en 1D puis en 2D. Les conditions initiales sont données par : E et F_x valent 1 pour $x \in [0.4, 0.6]$ et 0.0001 ailleurs. F_y vaut zéro. La solution au temps 0.2 est un créneau non nul dans l'intervalle $x \in [0.6, 0.8]$. La solution est donnée par la condition initiale exactement transportée à la vitesse 1.

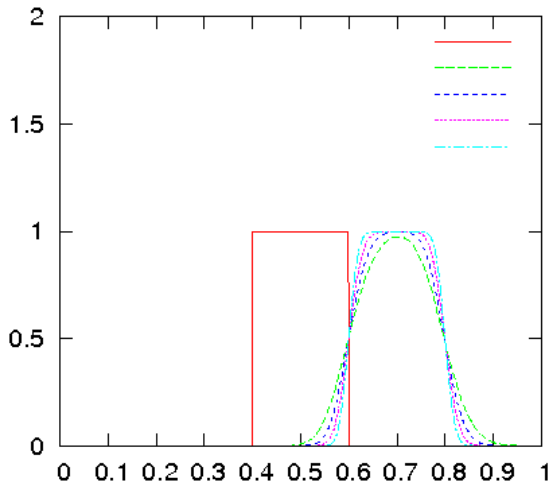


FIGURE 6.1 – La condition initiale et la solution au temps $t = 0.2$ pour différents pas d'espace. La courbe rouge représente $E(t, \mathbf{x})$ au temps initial. Les courbes verte, bleu, magenta et cyan représentent la solution numérique pour la quantité E et pour 100, 200, 400 et 800 mailles.

L'ordre de convergence constaté est approximativement de 0.5, ce qui est cohérent avec les résultats classiques de volumes finis pour le transport pour donnée initiale discontinue. Maintenant on propose le même cas test en 2D. Le tableau (6.6) donne les ordres de convergence pour différents maillages ainsi

Maillage	40-80	80-160	160-320	Nb coef $E < 0$	Nb coef $\| \mathbf{f} \| > 1$
Maillage cartésien	0.475	0.49	0.496	0	0
Maillage aléa. quad	0.475	0.494	0.496	0	0
Maillage de Kershaw	0.42	0.465	0.493	0	0
Maillage triang. reg.	0.474	0.491	0.496	0	0
Maillage aléa. triang.	0.477	0.492	0.496	0	0

TABLE 6.6 – Ordre de convergence pour une solution discontinue (trois premières colonnes) et nombre de coefficient ne respectant pas le principe du maximum (deux dernières colonnes).

que le nombre de quantités ne respectant pas le principe du maximum. En affinant on se rapproche de

l'ordre 0.5. Si on, prend maintenant comme condition initiale une Gaussienne à la place du créneau, les résultats numériques montre que le schéma converge à l'ordre 1.

Cas test en régime de diffusion

On prend le même cas test que pour le schéma de diffusion linéaire. Soit $SF(t)$ la solution fondamentale de la chaleur avec le coefficient de diffusion $\frac{1}{3}$. Les conditions initiales sont définies par $E(t=0) = SF(0.01)$, $\mathbf{F}(t=0) = \mathbf{0}$. On compare la solution numérique avec la solution théorique de diffusion $E(t) = SF(0.01 + T_f)$, $\mathbf{F}(t) = \mathbf{0}$.

Maillage/nb mailles	80	160	320	Nb $E_j < 0$
Cartésien	1.79	1.91	1.90	0
Aléatoire quad.	1.85	1.92	1.92	0
Cartésien tri.	1.90	1.95	1.92	0
Aléatoire tri.	1.92	1.95	1.92	0
Kershaw	1.68	1.67	1.87	0

Maillage/nb mailles	80	160	320	Nb $E_j < 0$
Cartesien	1.79	1.94	1.98	0
Aléatoire quad.	1.85	1.94	2.	0
Cartesien tri.	1.90	1.96	2.05	0
Aléatoire tri.	1.92	1.96	1.94	0
Kershaw	1.69	1.67	1.89	0

TABLE 6.7 – Ordre de convergence du schéma M_1 pour $\varepsilon = 0.0001$

Les ordres de convergence constatés dans le tableau (6.7) sont très proches des ordres de convergence du schéma limite de diffusion. Le schéma pour le modèle M_1 converge avec un ordre proche de deux pour une petite valeur de ε. Ces résultats valident numériquement le caractère « asymptotic preserving » du schéma.

Résultats numériques pour le modèle M_1 couplé avec la matière

On souhaite étudier le couplage du modèle M_1 avec la matière, c'est à dire résoudre le système suivant

$$\begin{cases} \partial_t E + \dfrac{1}{\epsilon} \operatorname{div} \mathbf{F} = \dfrac{\sigma_a}{\varepsilon^2}(aT^4 - E), \\ \partial_t \mathbf{F} + \dfrac{1}{\epsilon} \nabla(\hat{P}) = -\dfrac{\sigma_a}{\epsilon^2}\mathbf{F}, \\ \rho C_v \partial_t T = \dfrac{\sigma_a}{\varepsilon^2}(E - aT^4). \end{cases} \tag{6.68}$$

La stratégie est de procéder par « splitting ». On scinde donc le système en deux systèmes résolus successivement. Dans un premier temps on résout le modèle M_1 (voir (6.6)), puis le système issu de couplage :

$$\begin{cases} \partial_t E = \dfrac{\sigma_a}{\varepsilon^2}(aT^4 - E), \\ \partial_t \mathbf{F} = \mathbf{0}, \\ \rho C_v \partial_t T = \dfrac{\sigma_a}{\varepsilon^2}(E - aT^4). \end{cases} \tag{6.69}$$

Ce système dit de relaxation en température peut être résolu avec des schémas temporels classiques. Cependant, les calculs de stabilité du chapitre précédent montrent que l'on aura une condition sur le pas de temps de type CFL dépendant de $\frac{\sigma_a}{\varepsilon^2}$ mais aussi de T^3. Par conséquent il est difficile de traiter cette phase de façon explicite. Nous utilisons le schéma semi-implicite défini pour P_1 dans le chapitre précédent.

Cas test de Marshak linéaire Ici il s'agit de reprendre le cas de Marshak linéaire étudié dans le chapitre précédent. La température n'étant pas trop élevée, on est dans un cadre quasi isotrope. En conséquence on devrais s'attendre à ce que les résultats soient proches de ceux obtenus pour P_1. Pour ce test c'est un schéma d'advection d'ordre 1 qui est utilisé. La solution correspond à celle introduite

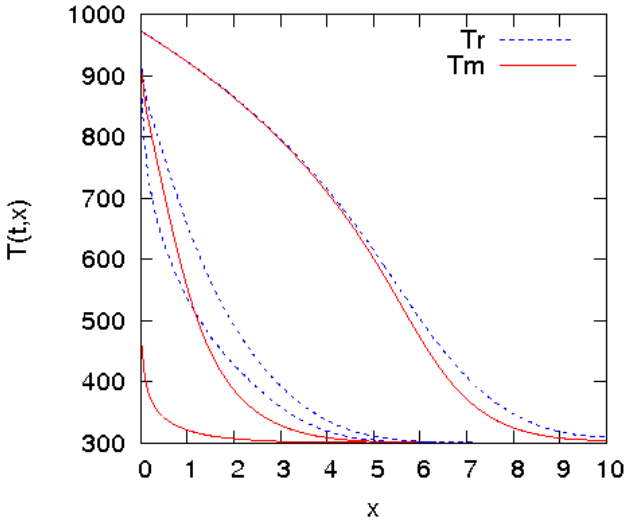


FIGURE 6.2 – La température matière et la température radiative dans le cas du modèle P_1.

dans [BCT08]- [B010]. Afin d'observer le caractère « asymptotic preserving » du schéma on regarde si malgré un faible nombre de mailles on obtient un résultat du même type (voir [BCT08]- [B010]). Pour terminer on trace la solution 2D sur maillage Cartésien et aléatoire (voir figure 6.4).

Remarque 6.7.2. *Ce n'est pas visible pour ce cas test, mais le schéma (6.48)-(6.49) couplé avec la matière, préserve la positivité de E mais ne préserve pas au niveau discret l'inégalité $||\mathbf{F}|| < E$.*

6.8 Autres tenseurs d'Eddington

La méthode proposée dans ce chapitre possède le désavantage de ne pas être naturellement extensible à tout type de tenseur d'Eddington. Considérons le tenseur dit de Kershaw

$$\chi = \frac{1 + 2||\mathbf{f}||^2}{3}.$$

Pour obtenir la décomposition de la pression $\widehat{P} = \mathbf{u} \otimes \mathbf{F} + q\widehat{I}_d$, on introduit

$$\mathbf{u} = \mathbf{f}$$

et

$$q = \frac{1 - ||\mathbf{f}||^2}{3} E.$$

Par contre la décomposition du flux $\mathbf{F}$ est différente. En effet, on obtient $\mathbf{F} = \mathbf{u}E$. En posant $E = \rho$ et $\mathbf{F} = \rho\mathbf{u}$, on obtient donc le système de la dynamique des gaz suivant

$$\begin{cases} \partial_t \rho + \dfrac{1}{\varepsilon} \operatorname{div}(\rho \mathbf{u}) = 0 \\ \partial_t \rho \mathbf{u} + \dfrac{1}{\varepsilon} \operatorname{div}(\rho \mathbf{u} \otimes \mathbf{u}) + \dfrac{1}{\varepsilon} \nabla q = -\dfrac{\sigma}{\varepsilon^2} \rho \mathbf{u}. \end{cases} \tag{6.70}$$

Il s'agit des équations d'Euler barotrope avec friction. On peut utiliser, comme précédemment, un schéma Lagrange+projection « asymptotic preserving » . Dans ce cas l'équation sur E est linéaire mais

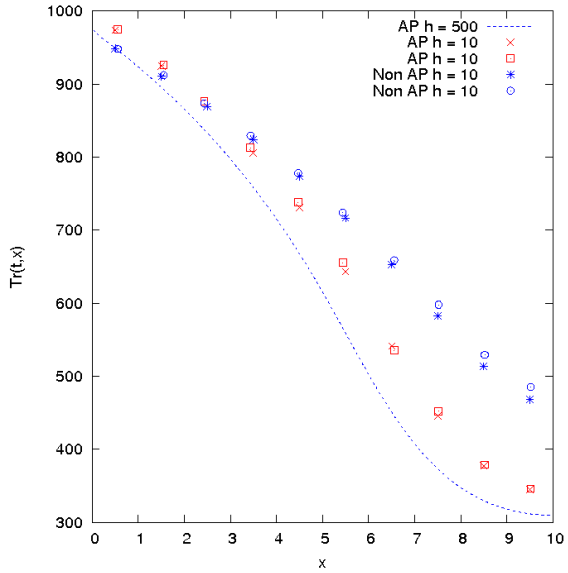


FIGURE 6.3 – La température radiative au temps final. Les croix rouges représentent la solution sur un maillage Cartésien et les carrés représentent la solution sur un maillage aléatoire

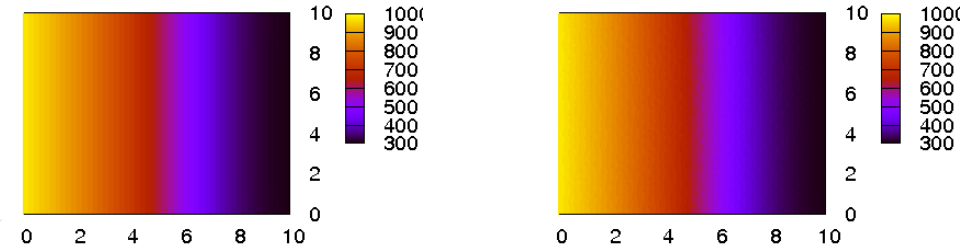


FIGURE 6.4 – Trace de la température radiative sur maillage Cartésien (à gauche) et sur maillage aléatoire (à droite)

le schéma limite obtenu reste non linéaire. Il correspond au schéma limite de la phase de projection dans le cas du précédent modèle. Cet exemple montre que le système de la dynamique des gaz associé au modèle M_1 diffère selon le facteur d'Eddington choisi. Par conséquent, il n'est pas possible de proposer de façon générale cette méthode pour tout facteur d'Eddington, il faut procéder au cas par cas.
Il serait intéressant d'étudier cette méthode dans la cadre multi-groupes. Une première approximation des modèles multi-groupes consiste à utiliser le facteur d'Eddington classique pour chaque groupe. Dans ce cas on peut reformuler chaque modèle M_1 comme un système d'Euler. Avec une bonne formulation du couplage entre les modèles on pourrait obtenir une méthode valide. Par contre, pour les facteurs d'Eddington réels associés à la méthode multi-groupes (voir [Tur03]), il n'est pas évident qu'on obtienne des systèmes de la dynamique des gaz correctement définis.

6.9 Synthèse

On propose un résumé des propriétés du schéma (6.48)-(6.49) construit pour le modèle M_1. Le schéma GLACE-AP pour le modèle M_1

- est entropique.
- préserve le principe du maximum pour le modèle M_1 sans couplage.
- préserve la positivité de E pour le modèle M_1 avec ou sans couplage.
- admet une variante semi-implicite, stable sous une condition CFL indépendante de ε.
- converge numériquement sur maillages non structurés.
- admet un schéma limite de diffusion préservant la positivité.
- admet un schéma limite convergent avec le même ordre que celui du schéma d'advection.

Cependant le schéma admet aussi quelques défauts,

- il exhibe des modes parasites.
- il n'est pas valable dans le vide ($E = 0$).

Afin de faire disparaître les modes parasites on peut utiliser d'autres définitions des normales comme cela a d'ailleurs été introduit dans le chapitre sur la discrétisation de l'équation de la chaleur hyperbolique. On peut également utiliser un schéma d'advection aux arêtes. Cependant, tous les modes parasites ne disparaissent pas avec cette méthode.

Annexe : rappels sur les systèmes dynamiques.

Dans cette annexe on propose de rappeler des résultats d'existence pour des systèmes dynamiques. Ces définitions et théorèmes sont issus de [Dem06]. On commence par introduire le problème de Cauchy

$$\begin{cases} \mathbf{x}'(t) = f(t, \mathbf{x}(t)) \\ \mathbf{x}'(t_0) = \mathbf{x}_0 \end{cases} \tag{6.71}$$

et le cylindre $C_0 = [t_0 - T_0, t_0 + T_0] \times \overline{B}(y_0, r_0)$. On propose de donner des résultats d'existence pour le problème de Cauchy (6.71).

Définition
On dit que $C = [t_0 - T, t_0 + T] \times \overline{B}(y_0, r_0) \subset C_0$ *avec* $T \leq \min(T_0, \frac{r_0}{M})$ *est un cylindre de sécurité pour le problème de Cauchy (6.71) si toute solution* $\mathbf{x}(t) : I \to \mathbb{R}^n$ *avec* $I \subset [t_0 - T, t_0 + T]$ *reste continu dans* $\overline{B}(\mathbf{x}_0, r_0)$.

La constante M est définie par $M = \sup_{(t,\mathbf{x}_0)\in C_0} ||f(t,\mathbf{x}(t))|| < \infty$.

Définition
Soit l'EDO (6.71) avec $f : D \subset \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$ Une solution maximale $\mathbf{x}(t) : [t_0 - T, t_0 + T] \to D$ est une solution qui n'est pas la restriction d'une solution définie sur un intervalle plus grand.

Théorème de Cauchy Lipschitz
Soit $f(t,\mathbf{x}(t)) : [t_0 - T, t_0 + T] \times D$. Soit l'EDO (6.71) Si $f(t,\mathbf{x}(t))$ est continue et localement Lipschitz par rapport à $\mathbf{x}$ alors il existe une et une seule solution maximale $\mathbf{x}(t) : [t_0 - T, t_0 + T] \longrightarrow D$ avec $T \leq \min(T_0, \frac{r_0}{M})$.

Théorème de Cauchy-Peano
Soit l'EDO (6.71). Si $f(t,\mathbf{x}(t))$ est continue et bornée alors il existe une solution maximale $\mathbf{x}(t) : [t_0 - T, t_0 + T] \to \overline{B}(\mathbf{x}_0, r_0)$ avec $T \leq \min(T_0, \frac{r_0}{M})$.

Définition *Soit $\mathbf{x}(t) :]a, b[\to \Omega$ solution maximale du système (6.71) Le bout droit est le couple défini par la valeur d'adherence de $(t,\mathbf{x}(t))$ lorsque $t \to b$ et que le bout gauche est le couple défini par la valeur d'adhérence de $(t,\mathbf{x}(t))$ lorsque $t \to a$.*

Corollaire
Soit $\mathbf{x}(t)$ solution maximale de

$$\begin{cases} \mathbf{x}'(t) = f(t,\mathbf{x}(t)), \\ \mathbf{x}(t=0) = \mathbf{x}_0 \in D. \end{cases}$$

Si la fonction $f(t,\mathbf{x}(t))$ est continue et bornée par rapport à $\mathbf{x}$ alors le bout droit est soit vide soit contenu sur la frontière de $]a, b[\times D$.

Quatrième partie

Perspectives

Chapitre 7

Schémas positifs pour les modèles S_n d'ordonnées discrètes

Dans ce chapitre prospectif, on propose une réflexion sur une discrétisation positive et « asymptotic preserving » de la méthode des ordonnées discrètes (modèles S_n). Cette discrétisation est basée sur un formalisme aux arêtes de la méthode des volumes finis couplée avec une formulation « pair - impair » des modèles.

Afin d'obtenir des simulations réalistes en transport linéaire dans tous les régimes liés à la matière, il est intéressant de construire des schémas qui respectent les propriétés physiques et mathématiques des modèles. Les modèles S_n (méthode des ordonnées discrètes) qui correspondent à une discrétisation en vitesse de l'équation de transport préservent à la fois la limite de diffusion de l'équation de transport et la positivité de la fonction de distribution discrète et du premier moment discret. La méthode de décomposition « diffusive - non diffusive » introduite lors du chapitre quatre, préserve le régime asymptotique, mais ne préserve pas la positivité des variables. On souhaite donc proposer des nouveaux schémas sur un formalisme aux arêtes de la méthode des volumes finis corrigeant ce défaut. L'idée de départ est assez simple : on propose de partir d'un schéma positif et « asymptotic preserving » en dimension un sur grille non uniforme, puis d'étendre le schéma sur maillage de Delaunay en écrivant les flux 1D dans la direction normale. Le schéma obtenu serait « asymptotic preserving » et admettrait le schéma TPFA (1.69) comme schéma limite. Enfin on propose de modifier la viscosité numérique en introduisant le schéma non linéaire de diffusion de Droniou-Le Potier [DlP09] qui est convergent et positif sur maillages non structurés. Les schémas « asymptotic preserving » 1D et positif existants ne paraissant pas satisfaisants pour un extension en dimension deux, ce chapitre portera principalement sur la construction d'un nouveau schéma pour les modèles S_n en dimension un. L'extension en dimension deux sera rapidement abordée à la fin du chapitre.

7.1 Schémas positifs, « asymptotic preserving » pour les modèles S_n en dimension un

Pour construire un schéma « asymptotic preserving » en dimension deux on doit commencer par choisir un schéma en dimension un sur grille non uniforme en 1D. Il existe, à notre connaissance, deux schémas « asymptotic preserving » -« cell centered » pour les modèles S_n qui satisfont la propriété de positivité : le schéma de Gosse [Gos10] et le schéma de Berthon-Turpault [BT10]. Cependant, le schéma de Berthon-Turpault n'est pas conservatif et le schéma de Gosse est trop coûteux sur grille non uniforme ; le problème stationnaire est en effet résolu exactement par le schéma. Ce problème étant

dépendant de la géométrie locale de la grille et de la valeur de l'opacité, on doit résoudre un problème stationnaire différent à chaque interface et à chaque temps ce qui génère un coût de calcul important. Par conséquent, on propose une nouvelle méthode basée sur une formulation dite « pair - impair » de la méthode des ordonnées discrètes.

7.1.1 Modèles S_n en dimension un

On considère l'équation de transport linéaire, mono-groupe, avec un « scattering » isotrope.

$$\partial_t f + \frac{1}{\varepsilon}\mu\partial_x f = \frac{\sigma}{\varepsilon^2}\left(\frac{1}{4\pi}\int_{\mu\in[-1,1]} f d\mu - f\right), \text{ avec } \mu = (\mathbf{e}_x, \mathbf{\Omega}) \text{ et } \mathbf{\Omega} \in \mathbb{R}^3, \tag{7.1}$$

avec $\mathbf{e}_x$ l'axe des abscisses. La méthode des ordonnées discrètes pour l'équation de transport (7.1) est définie par un ensemble de vitesses discrètes $\mathbf{\Omega}_k$ et de poids de quadrature $w_k \geq 0, k \in \{-N, N\}^3, k \neq 0$ qui sont choisis tels que

- $\sum_k w_k = 1$
- $\mathbf{\Omega}_{-k} = -\mathbf{\Omega}_k, \quad w_k = w_{-k} \Longrightarrow \sum_k w_k \mathbf{\Omega}_k = \mathbf{0}$.

Les points de quadrature sont choisis de façon symétrique par rapport aux plans de symétrie de la sphère.
La méthode des ordonnées discrètes pour l'équation (7.1) est donnée par le système

$$\partial_t f_k + \frac{1}{\varepsilon}\mu_k\partial_x f = \frac{\sigma}{\varepsilon^2}\left(\sum_l w_l f_l - f_k\right), k \in \{-N, N\}, k \neq 0, \tag{7.2}$$

avec $\mathbf{f} = \{f_k = f(\mu_k)\}$. Le système (7.2) s'écrit sous la forme matricielle suivante

$$\partial_t \mathbf{f} + \frac{1}{\varepsilon}\mu.\nabla\mathbf{f} = \frac{\sigma}{\varepsilon^2} R\mathbf{f}, \tag{7.3}$$

avec $\mu = \{\mu_k\}$, $\mathbf{w} = \{w_k\}$ et un terme source défini par $R = \mathbf{1} \otimes \mathbf{w} - \widehat{I}_d$. Dans le régime limite, lorsque ε tend vers 0, l'énergie E satisfait l'équation de diffusion suivante

$$\partial_t E - \text{div}\left(\frac{1}{D\sigma}\nabla E\right) = 0, \tag{7.4}$$

avec la matrice de diffusion donnée par $D = \sum w_k \,\text{diag}\{\mu_k : \mu_k\}$. Les vitesses discrètes μ_k et les poids associés w_k sont choisis tel que $D = \frac{1}{3}\widehat{I}_d$ si les vitesses appartiennent à la sphère S^2 et $D = \frac{1}{2}\widehat{I}_d$ si les vitesses appartiennent à la sphère S^1. Dans la suite du chapitre, on se place dans le cas $D = \frac{1}{3}$.

7.1.2 Schéma sur grille non uniforme pour l'équation de la chaleur hyperbolique

Avant d'introduire la nouvelle discrétisation « asymptotic preserving » et positive pour les modèles S_n, on propose un rappel sur le schéma de Gosse-Toscani [GT01] pour l'équation de la chaleur hyperbolique pour des grilles non uniformes. Le maillage est défini par un ensemble de points $...x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}, x_{j+\frac{3}{2}}...$ et de mailles $\Omega_j = \left[x_{j+\frac{1}{2}}, x_{j-\frac{1}{2}}\right]$ de volume $\Delta x_j = |x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}|$. Le centre de la maille Ω_j est noté x_j. On introduit pour finir la maille duale $\Omega_{j+\frac{1}{2}} = [x_j, x_{j+1}]$ de volume $\Delta x_{j+\frac{1}{2}} = |x_{j+1} - x_j|$.
On part de l'équation de la chaleur hyperbolique sous sa forme diagonalisée

$$\begin{cases} \partial_t u + \dfrac{\mu}{\varepsilon}\partial_x u = \dfrac{\sigma}{2\varepsilon^2}(v - u), \\ \partial_t v - \dfrac{\mu}{\varepsilon}\partial_x v = \dfrac{\sigma}{2\varepsilon^2}(u - v). \end{cases} \tag{7.5}$$

Le principe de construction est le même que sur grille uniforme. On localise les termes sources aux interfaces (voir (1.38)), on résout le problème stationnaire (voir (1.40)) pour obtenir les flux. La solution du problème de Riemann à l'interface $x_{j+\frac{1}{2}}$ est donnée par

$$\left\{\begin{array}{l} u^G_{j+\frac{1}{2}} = u_j, \quad v^G_{j+\frac{1}{2}} = M_{j+\frac{1}{2}} v_{j+1} + (1 - M_{j+\frac{1}{2}}) u_j = v_{j+1} + (1 - M_{j+\frac{1}{2}})(u_j - v_{j+1}), \\ u^D_{j+\frac{1}{2}} = M_{j+\frac{1}{2}} u_j + (1 - M_{j+\frac{1}{2}}) v_{j+1} = u_j + (1 - M_{j+\frac{1}{2}})(v_{j+1} - u_j), \quad v^D_{j+\frac{1}{2}} = v_{j+1}, \end{array}\right. \tag{7.6}$$

avec

$$M_{j+\frac{1}{2}} = \frac{2\mu\varepsilon}{\sigma_{j+\frac{1}{2}} \Delta x_{j+\frac{1}{2}} + 2\mu\varepsilon}. \tag{7.7}$$

À partir des flux (7.6), on dérive le schéma de Gosse-Toscani. Après de rapides calculs on obtient une forme du schéma faisant intervenir les termes sources :

$$\left\{\begin{array}{l} \dfrac{du_j}{dt} + \dfrac{\mu}{\varepsilon} \dfrac{M_{j+\frac{1}{2}} u_{j+\frac{1}{2}} - M_{j-\frac{1}{2}} u_{j-\frac{1}{2}}}{\Delta x_j} = \dfrac{1}{4} \left(M_{j+\frac{1}{2}} \dfrac{\Delta x_{j+\frac{1}{2}} \sigma_{j+\frac{1}{2}}}{\Delta x_j \varepsilon^2} + M_{j-\frac{1}{2}} \dfrac{\Delta x_{j-\frac{1}{2}} \sigma_{j-\frac{1}{2}}}{\Delta x_j \varepsilon^2} \right) (v_j - u_j) \\ + \mu \dfrac{M_{j+\frac{1}{2}} - M_{j-\frac{1}{2}}}{\varepsilon \Delta x_j} \dfrac{(u_j + v_j)}{2}, \\ \\ \dfrac{dv_j}{dt} + \dfrac{\mu}{\varepsilon} \dfrac{M_{j+\frac{1}{2}} v_{j+\frac{1}{2}} - M_{j-\frac{1}{2}} v_{j-\frac{1}{2}}}{\Delta x_j} = \dfrac{1}{4} \left(M_{j+\frac{1}{2}} \dfrac{\Delta x_{j+\frac{1}{2}} \sigma_{j+\frac{1}{2}}}{\Delta x_j \varepsilon^2} + M_{j-\frac{1}{2}} \dfrac{\Delta x_{j-\frac{1}{2}} \sigma_{j-\frac{1}{2}}}{\Delta x_j \varepsilon^2} \right) (u_j - v_j) \\ - \mu \dfrac{M_{j+\frac{1}{2}} - M_{j-\frac{1}{2}}}{\varepsilon \Delta x_j} \dfrac{(u_j + v_j)}{2}, \end{array}\right. \tag{7.8}$$

avec les flux

$$\left\{\begin{array}{l} u_{j+\frac{1}{2}} = \dfrac{u_j + u_{j+1}}{2} + \dfrac{u_j - u_{j+1}}{2}, \\ \\ v_{j+\frac{1}{2}} = \dfrac{v_j + v_{j+1}}{2} - \dfrac{v_j - v_{j+1}}{2}. \end{array}\right. \tag{7.9}$$

On sait que le changement de variable $E = u + v$, $F = u - v$ dans le modèle (7.5) permet d'obtenir l'équation de la chaleur hyperbolique (1.20). Le schéma de Gosse-Toscani sur grille non uniforme pour le système (1.20) s'écrit

$$\left\{\begin{array}{l} \dfrac{dE_j}{dt} + \dfrac{\mu}{\varepsilon} \dfrac{M_{j+\frac{1}{2}} F_{j+\frac{1}{2}} - M_{j-\frac{1}{2}} F_{j-\frac{1}{2}}}{\Delta x_j} = 0, \\ \\ \dfrac{dF_j}{dt} + \dfrac{\mu}{\varepsilon} \dfrac{M_{j+\frac{1}{2}} E_{j+\frac{1}{2}} - M_{j-\frac{1}{2}} E_{j-\frac{1}{2}}}{\Delta x_j} = -\dfrac{1}{2} \left(M_{j+\frac{1}{2}} \dfrac{\Delta x_{j+\frac{1}{2}}}{\Delta x_j} \dfrac{\sigma_{j+\frac{1}{2}}}{\varepsilon^2} + M_{j-\frac{1}{2}} \dfrac{\Delta x_{j-\frac{1}{2}}}{\Delta x_j} \dfrac{\sigma_{j-\frac{1}{2}}}{\varepsilon^2} \right) F_j \\ + \dfrac{M_{j+\frac{1}{2}} - M_{j-\frac{1}{2}}}{\varepsilon \Delta x_j} \mu E_j, \end{array}\right. \tag{7.10}$$

avec les flux donnés par

$$\left\{\begin{array}{l} E_{j+\frac{1}{2}} = \dfrac{E_j + E_{j+1}}{2} + \dfrac{F_j - F_{j+1}}{2}, \\ \\ F_{j+\frac{1}{2}} = \dfrac{F_j + F_{j+1}}{2} + \dfrac{E_j - E_{j+1}}{2}. \end{array}\right. \tag{7.11}$$

Ce schéma est donc positif, « asymptotic preserving », conservatif et « well-balanced » . La discrétisation du terme source $\frac{1}{2}\left(M_{j+\frac{1}{2}}\frac{\Delta x_{j+\frac{1}{2}}}{\Delta x_j}\frac{\sigma_{j+\frac{1}{2}}}{\varepsilon^2}+M_{j-\frac{1}{2}}\frac{\Delta x_{j-\frac{1}{2}}}{\Delta x_j}\frac{\sigma_{j-\frac{1}{2}}}{\varepsilon^2}\right)F_j$ faisant intervenir un ratio entre le volume de la maille et ceux des mailles duales associées, est nécessaire pour conserver exactement les états stationnaires du système (un schéma « well-balanced »).

7.1.3 Schéma « flux pair-impair » basé sur le schéma de Gosse-Toscani

On considère le cas 1D-plan. Les vitesses discrètes sont données par $\mu_k = (\mathbf{e}_x, \mathbf{\Omega}_k)$ où $\mathbf{e}_x$ l'axe des abscisses. En régime de diffusion, le coefficient de diffusion est défini par $D = \sum w_k \mu_k^2$. La construction du schéma « asymptotic preserving » et positif est basé sur une formulation « pair - impair » . Il s'agit d'écrire les modèles S_n comme une collection d'équations de chaleur hyperbolique de vitesse μ_k couplée avec un terme source sur les équations d'énergie (E^k). Soit $E^k = f_k + f_{-k}$ et $F^k = f_k - f_{-k}$, les équations (7.3) peuvent se réécrire sous la forme

$$\left\{\begin{array}{l} k \in \{1, N\} \\ \partial_t E^k + \dfrac{1}{\varepsilon}\mu_k \partial_x F^k = \dfrac{\sigma}{\varepsilon^2}\left(\displaystyle\sum_{l=1}^{N} 2w_l (E^l - E^k)\right), \\ \partial_t F^k + \dfrac{1}{\varepsilon}\mu_k \partial_x E^k = -\dfrac{\sigma}{\varepsilon^2} F^k. \end{array}\right. \tag{7.12}$$

La première étape de construction du schéma est d'étendre le schéma de Gosse-Toscani à chaque équation de la chaleur hyperbolique qui compose la formulation « flux pair - impair » , pour cela on introduit un facteur M pour chaque vitesse discrète

$$M^k_{j+\frac{1}{2}} = \frac{2\mu_k \varepsilon}{\sigma_{j+\frac{1}{2}} \Delta x_{j+\frac{1}{2}} + 2\mu_k \varepsilon}. \tag{7.13}$$

La seconde étape consiste à coupler les différentes équations à l'aide d'une discrétisation qui préserve le principe du maximum et la conservativité. Si on veut préserver le principe du maximum, il est essentiel de faire intervenir le facteur M dans le terme source qui couple les équations de la chaleur hyperbolique. Pour cela on propose d'introduire le coefficient

$$N^{k,l}_j = \min\left(\frac{\Delta x_{j+\frac{1}{2}}}{\Delta x_j}\frac{\sigma_{j+\frac{1}{2}}}{\varepsilon^2}M^k_{j+\frac{1}{2}}, \frac{\Delta x_{j-\frac{1}{2}}}{\Delta x_j}\frac{\sigma_{j-\frac{1}{2}}}{\varepsilon^2}M^k_{j-\frac{1}{2}}, \frac{\Delta x_{j+\frac{1}{2}}}{\Delta x_j}\frac{\sigma_{j+\frac{1}{2}}}{\varepsilon^2}M^l_{j+\frac{1}{2}}, \frac{\Delta x_{j-\frac{1}{2}}}{\Delta x_j}\frac{\sigma_{j-\frac{1}{2}}}{\varepsilon^2}M^l_{j-\frac{1}{2}}\right). \tag{7.14}$$

A partir de ces quantités on obtient un nouveau schéma pour les modèles S_n définis par

$$\left\{\begin{array}{l} \forall k \in \{1, N\} \\ \dfrac{dE^k_j}{dt} + \dfrac{\mu_k}{\varepsilon}\dfrac{M^k_{j+\frac{1}{2}}F^k_{j+\frac{1}{2}} - M^k_{j-\frac{1}{2}}F^k_{j-\frac{1}{2}}}{\Delta x_j} = \displaystyle\sum_{l=1}^{N} 2w_l N^{k,l}_j (E^l_j - E^k_j), \\ \\ \dfrac{dF^k_j}{dt} + \dfrac{\mu_k}{\varepsilon}\dfrac{M^k_{j+\frac{1}{2}}E^k_{j+\frac{1}{2}} - M^k_{j-\frac{1}{2}}E^k_{j-\frac{1}{2}}}{\Delta x_j} = -\dfrac{1}{2}\left(M^k_{j+\frac{1}{2}}\dfrac{\Delta x_{j+\frac{1}{2}}}{\Delta x_j}\dfrac{\sigma_{j+\frac{1}{2}}}{\varepsilon^2} + \dfrac{\Delta x_{j-\frac{1}{2}}}{\Delta x_j}M^k_{j-\frac{1}{2}}\dfrac{\sigma_{j-\frac{1}{2}}}{\varepsilon^2}\right)F^k_j \\ + \dfrac{M^k_{j+\frac{1}{2}} - M^k_{j-\frac{1}{2}}}{\varepsilon \Delta x_j}\mu_k E^k_j. \end{array}\right. \tag{7.15}$$

avec les flux

$$\begin{cases} E_{j+\frac{1}{2}} = \frac{E_j^k + E_{j+1}^k}{2} + \frac{F_j^k - F_{j+1}^k}{2}, \\ F_{j+\frac{1}{2}} = \frac{F_j^k + F_{j+1}^k}{2} + \frac{E_j^k - E_{j+1}^k}{2}. \end{cases} \tag{7.16}$$

En diagonalisant le schéma (7.15)-(7.16), on obtient un schéma sur les quantités $u_k = f_k$ et $v_k = f_{-k}$

$$\begin{cases} \frac{du_j^k}{dt} + \frac{\mu_k}{\varepsilon} \frac{M_{j+\frac{1}{2}}^k u_{j+\frac{1}{2}}^k - M_{j-\frac{1}{2}}^k u_{j-\frac{1}{2}}^k}{\Delta x_j} = \sum_{l=1}^{N} w_l N_j^{k,l}((u_j^l + v_j^l) - (u_j^k + v_j^k)) + \\ \frac{1}{4}\left(\frac{\Delta x_{j+\frac{1}{2}}}{\Delta x_j} M_{j+\frac{1}{2}}^k \frac{\sigma_{j+\frac{1}{2}}}{\varepsilon^2} + \frac{\Delta x_{j-\frac{1}{2}}}{\Delta x_j} M_{j-\frac{1}{2}}^k \frac{\sigma_{j-\frac{1}{2}}}{\varepsilon^2} \right)(v_j^k - u_j^k) + \frac{M_{j+\frac{1}{2}}^k - M_{j-\frac{1}{2}}^k}{\varepsilon \Delta x} \frac{\mu_k}{2}(u_i^k + v_j^k), \\ \frac{dv_j^k}{dt} + \frac{\mu_k}{\varepsilon} \frac{M_{j+\frac{1}{2}}^k v_{j+\frac{1}{2}}^k - M_{j-\frac{1}{2}}^k v_{j-\frac{1}{2}}^k}{\Delta x_j} = \sum_{l=1}^{N} w_l N_j^{k,l}((u_j^l + v_j^l) - (u_j^k + v_j^k)) + \\ \frac{1}{4}\left(\frac{\Delta x_{j+\frac{1}{2}}}{\Delta x_j} M_{j+\frac{1}{2}}^k \frac{\sigma_{j+\frac{1}{2}}}{\varepsilon^2} + \frac{\Delta x_{j-\frac{1}{2}}}{\Delta x_j} M_{j-\frac{1}{2}}^k \frac{\sigma_{j-\frac{1}{2}}}{\varepsilon^2} \right)(u_j^k - v_j^k) - \frac{M_{j+\frac{1}{2}}^k - M_{j-\frac{1}{2}}^k}{\varepsilon \Delta x} \frac{\mu_k}{2}(u_j^k + v_j^k). \end{cases} \tag{7.17}$$

avec les flux

$$\begin{cases} u_{j+\frac{1}{2}}^k = \frac{u_j^k + u_{j+1}^k}{2} + \frac{u_j^k - u_{j+1}^k}{2}, \\ v_{j+\frac{1}{2}}^k = \frac{v_j^k + v_{j+1}^k}{2} - \frac{v_j^k - v_{j+1}^k}{2}. \end{cases} \tag{7.18}$$

Propriétés

Maintenant que le schéma « pair - impair » (7.15)-(7.16) est obtenu on propose de valider le caractère « asymptotic preserving » à l'aide d'une étude asymptotique formelle et le caractère positif à l'aide du principe du maximum. L'étude de la limite asymptotique au sens de Hilbert permet d'obtenir la limite du schéma (7.15)-(7.16) sur maillage grossier ($\frac{\Delta x}{\varepsilon} >> 1$). La limite sur maillage fin ne peut être obtenue qu'avec une preuve de convergence ou de consistance.

Proposition 7.1. *$\forall k$, $\mu_k \neq 0$, le schéma (7.15)-(7.16) admet le schéma limite suivant*

$$\frac{dE_j}{dt} - \frac{1}{\Delta x_j}\left(D_{j+\frac{1}{2}}\left(\frac{E_{j+1} - E_j}{\Delta x_{j+\frac{1}{2}}} \right) - D_{j-\frac{1}{2}}\left(\frac{E_j - E_{j-1}}{\Delta x_{j-\frac{1}{2}}} \right) \right) = 0, \tag{7.19}$$

avec $E_j = \sum_k w_k E_j^k$ et le coefficient de diffusion $D_{j+\frac{1}{2}}$ défini par

$$D_{j+\frac{1}{2}} = \frac{\sum_k w_k \mu_k^2}{\sigma_{j+\frac{1}{2}}}. \tag{7.20}$$

Démonstration. On utilise un développement de Hilbert. Auparavant, on remarque que $M^{k,*}_{j+\frac{1}{2}} = \frac{1}{\varepsilon}M^k_{j+\frac{1}{2}} \simeq 0$. Par conséquent les termes homogènes à $\frac{1}{\varepsilon}$ sont

$$\frac{1}{2}\left(\frac{\Delta x_{j+\frac{1}{2}}}{\Delta x_j}M^k_{j+\frac{1}{2}}\sigma_{j+\frac{1}{2}} + \frac{\Delta x_{j-\frac{1}{2}}}{\Delta x_j}M^k_{j-\frac{1}{2}}\sigma_{j-\frac{1}{2}}\right)F^k_j = 0, \tag{7.21}$$

ce qui montre que $F^k_j = 0$ et

$$\sum_{l=1}^{N} 2w_l N^{*,k,l}_j (E^l_j - E^k_j) = 0, \tag{7.22}$$

avec

$$N^{*,k,l}_j = \min\left(\frac{\Delta x_{j+\frac{1}{2}}}{\Delta x_j}\sigma_{j+\frac{1}{2}}M^{k,*}_{j+\frac{1}{2}}, \frac{\Delta x_{j-\frac{1}{2}}}{\Delta x_j}\sigma_{j-\frac{1}{2}}M^{k,*}_{j-\frac{1}{2}}, \frac{\Delta x_{j+\frac{1}{2}}}{\Delta x_j}\sigma_{j+\frac{1}{2}}M^{l,*}_{j+\frac{1}{2}}, \frac{\Delta x_{j-\frac{1}{2}}}{\Delta x_j}\sigma_{j-\frac{1}{2}}M^{l,*}_{j-\frac{1}{2}}\right).$$

Lorsque ε tend vers zéro on constate que la relation (7.22) équivaut à

$$\forall k \qquad \sum_{l=1}^{N} 2w_l \min(\mu_k, \mu_l)(E^l_j - E^k_j) = 0. \tag{7.23}$$

À présent on introduit la matrice

$$\begin{cases} S_{kl} = 2w_l \min(\mu_k, \mu_l) \text{ si } k \neq l, \\ S_{kk} = -\sum_{l=1}^{N} 2w_l \min(\mu_k, \mu_l) + 2w_l \min(\mu_k, \mu_k). \end{cases} \tag{7.24}$$

Le noyau de la matrice S, comme celui de R, est $\operatorname{Ker} S = \{\lambda \overrightarrow{1}\}$, donc l'équation (7.22) implique que $E_k = E_l$, $\forall l, k \in \{1..N\}$. Dans un second temps on étudie les termes homogènes à $O(1)$ dans le développement de Hilbert :

$$\begin{cases} \partial_t E^k_j + \dfrac{1}{\Delta x_j}\left(\dfrac{2\mu_k^2}{\sigma_{j+\frac{1}{2}}\Delta x_{j+\frac{1}{2}} + 2\varepsilon}F^k_{j+\frac{1}{2}} - \dfrac{2\mu_k^2}{\sigma_{j-\frac{1}{2}}\Delta x_{j-\frac{1}{2}} + 2\varepsilon}F^k_{j-\frac{1}{2}}\right) = 0, \\ \partial_t F^k_j + \dfrac{1}{\Delta x_j}\left(\dfrac{2\mu_k^2}{\sigma_{j+\frac{1}{2}}\Delta x_{j+\frac{1}{2}} + 2\varepsilon}E^k_{j+\frac{1}{2}} - \dfrac{2\mu_k^2}{\sigma_{j-\frac{1}{2}}\Delta x_{j-\frac{1}{2}} + 2\varepsilon}E^k_{j-\frac{1}{2}}\right) \\ + \dfrac{1}{\Delta x_j}\left(\dfrac{2\mu_k^2}{\sigma_{j+\frac{1}{2}}\Delta x_{j+\frac{1}{2}} + 2\varepsilon} - \dfrac{2\mu_k^2}{\sigma_{j-\frac{1}{2}}\Delta x_{j-\frac{1}{2}} + 2\varepsilon}\right)E^k_j = 0. \end{cases} \tag{7.25}$$

Lorsque ε tend vers zéro, le système (7.25) est équivalent à

$$\begin{cases} \partial_t E^k_j + \dfrac{1}{\Delta x_j}\left(\dfrac{2\mu_k^2}{\sigma_{j+\frac{1}{2}}\Delta x_{j+\frac{1}{2}}}F^k_{j+\frac{1}{2}} - \dfrac{2\mu_k^2}{\sigma_{j-\frac{1}{2}}\Delta x_{j-\frac{1}{2}}}F^k_{j-\frac{1}{2}}\right) = 0, \\ \partial_t F^k_j + \dfrac{1}{\Delta x_j}\left(\dfrac{2\mu_k^2}{\sigma_{j+\frac{1}{2}}\Delta x_{j+\frac{1}{2}}}E^k_{j+\frac{1}{2}} - \dfrac{2\mu_k^2}{\sigma_{j-\frac{1}{2}}\Delta x_{j-\frac{1}{2}}}E^k_{j-\frac{1}{2}}\right) \\ + \dfrac{1}{\Delta x_j}\left(\dfrac{2\mu_k^2}{\sigma_{j+\frac{1}{2}}\Delta x_{j+\frac{1}{2}}} - \dfrac{2\mu_k^2}{\sigma_{j-\frac{1}{2}}\Delta x_{j-\frac{1}{2}}}\right)E^k_j = 0. \end{cases} \tag{7.26}$$

Puisque $E_j^k = E_j^l \quad \forall l,k \in \{1..N\}$ alors $E_j^k = E_j^* \quad \forall k \in \{1..N\}$. On part de l'équation

$$\partial_t E_j^k + \frac{1}{\Delta x_j}\left(\frac{2\mu_k^2}{\sigma_{j+\frac{1}{2}}\Delta x_{j+\frac{1}{2}}}F_{j+\frac{1}{2}}^k - \frac{2\mu_k^2}{\sigma_{j-\frac{1}{2}}\Delta x_{j-\frac{1}{2}}}F_{j-\frac{1}{2}}^k\right) = 0 \quad \forall k \in \{1..N\}\,. \tag{7.27}$$

Puisque $E_j^k = E_j^*$, $\forall k \in \{1..N\}$, l'énergie E_j est définie par $E_j = \sum_{k=-N}^N w_k E_j^k = \sum_{k=1}^N 2w_k E_j^k = E_j^*(\sum_{k=1}^N 2w_k) = E_j^*$. Enfin on somme $\forall k \in \{1..N\}$ les équations (7.27) multipliées par w_k, on utilise la définition de E_j^* et $F_j^k = 0$, ce qui donne

$$\frac{dE_j}{dt} - \frac{1}{\Delta x_j}\left(D_{j+\frac{1}{2}}\left(\frac{E_{j+1}-E_j}{\Delta x_{j+\frac{1}{2}}}\right) - D_{j-\frac{1}{2}}\left(\frac{E_j - E_{j-1}}{\Delta x_{j-\frac{1}{2}}}\right)\right) = 0, \tag{7.28}$$

avec le coefficient de diffusion

$$D_{j+\frac{1}{2}} = \frac{\sum_k w_k \mu_k^2}{\sigma_{j+\frac{1}{2}}}. \tag{7.29}$$

□

Ce résultat montre que le schéma « pair - impair » sur maillage grossier admet un schéma de diffusion consistant avec l'équation de diffusion limite. Le résultat n'est cependant pas suffisant pour montrer que le schéma est uniformément « asymptotic preserving ». Pour cela il faut obtenir des estimations de convergence et consistance indépendantes de ε. Pour l'instant ce point est bloquant. Pour démontrer la posivité, on montre que le schéma préserve le principe du maximum sous une certaine condition CFL.

Proposition 7.2. *Le schéma (7.15)-(7.16) préserve le principe du maximum pour les quantités $u^k = E^k + F^k$ et $v^k = E^k - F^k$ sous la condition CFL*

$$\Delta t\left(\frac{\mu_k M_{j-\frac{1}{2}}^k}{\varepsilon \Delta x_j} + \frac{\sigma_{j-\frac{1}{2}} M_{j-\frac{1}{2}}^k}{2\varepsilon^2} + \sum_{l=1}^N w_l N_j^{k,l}\right) \leq 1$$

Démonstration. Pour vérifier cette propriété, il est plus simple de partir de la forme (7.17)-(7.18) du schéma. On note $K = \sum_{l=1}^N w_l N_j^{k,l}(v_j^{l,n} + u_j^{l,n})$. Le schéma (7.17)-(7.18) peut se réécrire après simplification sous la forme suivante

$$\begin{cases} \dfrac{u_j^{k,n+1} - u_j^{k,}}{\Delta t} + \dfrac{\mu^k M_{j-\frac{1}{2}}^k}{\varepsilon}\dfrac{u_j^{k,n} - u_{j-1}^{k,n}}{\Delta x_j} = M_{j-\frac{1}{2}}^k \dfrac{\Delta x_{j-\frac{1}{2}}\sigma_{j-\frac{1}{2}}}{2\varepsilon^2 \Delta x_j}(v_j^{k,n} - u_j^{k,n}) + K - (\sum_{l=1}^N w_l N_j^{k,l})(v_j^{k,n} + u_j^{k,n}), \\ \dfrac{v_j^{k,n+1} - v_j^{k,n}}{\Delta t} - \dfrac{\mu^k M_{j+\frac{1}{2}}^k}{\varepsilon}\dfrac{v_{j+1}^{k,n} - v_j^{k,n}}{\Delta x_j} = M_{j+\frac{1}{2}}^k \dfrac{\Delta x_{j+\frac{1}{2}}\sigma_{j+\frac{1}{2}}}{2\varepsilon^2 \Delta x_j}(u_j^{k,n} - v_j^{k,n}) + K - (\sum_{l=1}^N w_l N_j^{k,l})(v_j^{k,n} + u_j^{k,n}). \end{cases} \tag{7.30}$$

Cette forme est obtenue grâce à la propriété $(1 - M_{j+\frac{1}{2}})\mu^k = \frac{M_{j+\frac{1}{2}}^k \sigma_{j+\frac{1}{2}} \Delta x_{j+\frac{1}{2}}}{2\varepsilon}$. Pour l'équation sur

$u_j^{k,n+1}$, les coefficients associés aux inconnues sont

$$\left\{\begin{array}{c} u_j^{k,n}\left(1-\frac{\Delta t\mu_k M^k_{j-\frac{1}{2}}}{\varepsilon\Delta x_j}-\frac{\Delta t\sigma_{j-\frac{1}{2}}M^k_{j-\frac{1}{2}}}{2\varepsilon^2}-\Delta t\sum_{l=1}^N w_l N_j^{k,l}\right), \\ u_{j-1}^{k,n}\left(\frac{\Delta t\mu_k M^k_{j-\frac{1}{2}}}{\varepsilon\Delta x_j}\right), \\ u_j^{l,n}\left(\Delta t\sum_{l=1}^N w_l N_j^{k,l}\right), \\ v_j^{l,n}\left(\Delta t\sum_{l=1}^N w_l N_j^{k,l}\right), \\ v_j^{k,n}\left(\Delta t\frac{\Delta x_{j-\frac{1}{2}}\sigma_{j-\frac{1}{2}}M^k_{j-\frac{1}{2}}}{2\varepsilon^2\Delta x_j}-\Delta t\sum_{l=1}^N w_l N_j^{k,l}\right). \end{array}\right. \tag{7.31}$$

La somme de ces coefficients de (7.31) est égale à 1. Tous les coefficients non diagonaux sont positifs. C'est trivial pour tous les coefficients excepté celui lié v_j^k. On rappelle que $\sum_{k=1}^N 2w_k = 1$ ce qui implique

$$M^k_{j-\frac{1}{2}}\frac{\Delta x_{j-\frac{1}{2}}\sigma_{j-\frac{1}{2}}}{2\varepsilon^2\Delta x_j}-(\sum_{l=1}^N w_l N_j^{k,l})=\sum_{l=1}^N 2w_l\left(\frac{\Delta x_{j-\frac{1}{2}}}{\Delta x_j}\frac{\sigma_{j-\frac{1}{2}}}{\varepsilon^2}M^k_{j-\frac{1}{2}}\right.$$

$$\left.-\min\left(\frac{\Delta x_{j+\frac{1}{2}}}{\Delta x_j}\frac{\sigma_{j+\frac{1}{2}}}{2\varepsilon^2}M^k_{j+\frac{1}{2}},\frac{\Delta x_{j-\frac{1}{2}}}{\Delta x_j}\frac{\sigma_{j-\frac{1}{2}}}{2\varepsilon^2}M^k_{j-\frac{1}{2}},\frac{\Delta x_{j+\frac{1}{2}}}{\Delta x_j}\frac{\sigma_{j+\frac{1}{2}}}{2\varepsilon^2}M^l_{j+\frac{1}{2}},\frac{\Delta x_{j-\frac{1}{2}}}{\Delta x_j}\frac{\sigma_{j-\frac{1}{2}}}{2\varepsilon^2}M^l_{j-\frac{1}{2}}\right)\right)$$

Avec la définition de $N_j^{k,l}$ (7.14), et la définition des poids de quadrature w_k, on remarque que ce terme est positif. Pour le coefficient diagonal, il est positif sous la condition CFL

$$\Delta t\left(\frac{\mu_k M^k_{j-\frac{1}{2}}}{\varepsilon\Delta x_j}+\frac{\Delta x_{j-\frac{1}{2}}}{\Delta x_j}\frac{\sigma_{j-\frac{1}{2}}M^k_{j-\frac{1}{2}}}{2\varepsilon^2}+\sum_{l=1}^N w_l N_j^{k,l}\right)\leq 1.$$

Puisque tous les coefficients sont positifs et forment une combinaison convexe, le schéma préserve le principe du maximum. □

Remarque 7.1.1.
- *On remarque que le schéma « pair-impair » dérive sur le schéma upwind pour la formulation « pair - impair » lorsque $\sigma = 0$.*
- *Le choix du couplage à travers le facteur $N_j^{k,l}$ entre différentes équations de la chaleur hyperbolique par le minimum des coefficients M permet de conserver le principe du maximum.*

7.1.4 Résultats numériques

Dans cette section, on propose de valider numériquement, le schéma « pair - impair » comme discrétisation des modèles S_n. On commence par donner un exemple de quadrature utilisable pour construire les modèles S_n. Dans la suite on pose n le nombre de vitesses positives associé au modèles S_n. Par conséquent un modèle de type S_n admet $2n$ inconnues. Les formules de quadrature sont issues de [JL91]. Ces formules de quadrature sont construites tel que $\sum_{k=-N}^N w_k\mu_k^2 = 2\sum_{k=1}^N w_k\mu_k^2 = \frac{1}{3}$. On donne uniquement les vitesses μ_k pour $k > 0$ et les poids associés.

n=2	μ	w
	0.253619	0.2756145
	0.814716	0.2243855

n=4	μ	w
	0.112164	0.130395
	0.436752	0.173514
	0.751377	0.133596
	0.950828	0.0624995

n=6	μ	w
	0.072101	0.084985
	0.290789	0.1213175
	0.528456	0.113072
	0.734471	0.0913535
	0.888694	0.0618245
	0.978528	0.027447

Cas test de transport pur

Dans un premier temps, on souhaite vérifier que le schéma dérive sur le schéma upwind lorsque $\sigma = 0$. On considère un modèle S_n quelconque. Les paramètres sont donnés par $\sigma = 0$ et $\varepsilon = 1$. Les conditions initiales sont définies par $u^0(t = 0, x) = \chi_{[0.4,0.6]}(x)$, avec χ la fonction indicatrice, $v^0(t = 0, x) = 0$, $u^i(t = 0, x)$ et $v^i(t = 0, x) = 0$. La solution est donnée par $u^0(t = 0, x) = \chi_{[0.4,0.6]}(x - \mu_0 t)$, $v^0(t = 0, x) = 0$, $u^i(t = 0, x)$ et $v^i(t = 0, x) = 0$. Le schéma converge donc avec un ordre très proche

Maillage/ modèle	S_2	S_6
Uniforme 50-100	0.48	0.46
Uniforme 100- 200	0.49	0.47
Uniforme 200- 400	0.49	0.48
Aléatoire 50-100	0.49	0.56
Aléatoire 100- 200	0.48	0.45
Aléatoire 200- 400	0.49	0.46

TABLE 7.1 – Ordre de convergence pour le cas test de transport

de 0.5 ce qui est classique pour les solutions discontinues en transport.

Cas test de transport

On propose deux cas test dans le régime de transport ($\sigma = 1$, $\varepsilon = 1$), un cas test stationnaire et un autre instationnaire. On résout le modèle S_n sous la forme

$$\begin{cases} k \in \{1, N\} \\ \partial_t u^k + \dfrac{1}{\varepsilon}\mu_k \partial_x u^k = \dfrac{\sigma}{2\varepsilon^2}(v_k - u_k) + \dfrac{\sigma}{\varepsilon^2}\left(\displaystyle\sum_{l=1} w^l((u^l + v^l) - (u^k + v^k)\right), \\ \partial_t v^k - \dfrac{1}{\varepsilon}\mu_k \partial_x v^k = \dfrac{\sigma}{2\varepsilon^2}(u_k - v_k) + \dfrac{\sigma}{\varepsilon^2}\left(\displaystyle\sum_{l=1} w^l((u^l + v^l) - (u^k + v^k)\right). \end{cases} \tag{7.32}$$

Cas test stationnaire

La solution stationnaire est donnée par $u^k(t, x) = 1 + \frac{1}{2\mu_k}(1 - x)$, $v^k(t, x) = \frac{1}{2\mu_k}(1 - x)$ pour $k \in \{1..N\}$. La condition initiale est donnée par la solution stationnaire. Le schéma de Gosse-Toscani pour le modèle S_1 étant « well-balanced » il était attendu que le schéma préserve exactement la solution stationnaire. Le schéma « pair - impair » pour des modèles S_n avec $n > 1$ préserve aussi cette état stationnaire affine, cependant pour $n > 1$ il n'est pas à priori « well-balanced »

Maillage/ modèle	S_1	S_2	S_4
Uniforme 100	0	0	0
Uniforme 200	0	0	0
Aléatoire 50-100	0	0	0
Aléatoire 100- 200	0	0	0

TABLE 7.2 – Erreur en norme L^1 entre les solutions numérique et exacte pour le cas test stationnaire.

Cas test non stationnaire
La solution est donnée par $u^k(t,x) = \left(1 + \frac{\sigma x}{\mu_k}\right) e^{-\sigma t}$, $v^k(t,x) = \left(1 - \frac{\sigma x}{\mu_k}\right) e^{-\sigma t}$ pour $k \in \{1..N\}$. La condition initiale est apportée par la solution au temps $t = 0$. Les résultats donnés dans le tableau

Maillage/ modèle	S_1	S_2	S_4
Uniforme 50-100	1.0	0.99	0.96
Uniforme 100-200	0.99	0.99	0.98
Uniforme 200-400	1	1	0.99
Aléatoire 50-100	0.5	0.4	0.32
Aléatoire 100-200	0.61	0.47	0.37
Aléatoire 200-400	0.78	0.63	0.52
Aléatoire 400-800	0.95	0.82	0.69
Aléatoire 400-1600	0.98	0.98	0.96

TABLE 7.3 – Ordres de convergence associés au cas test non stationnaire.

(7.3) montrent que le schéma converge à l'ordre un pour une solution régulière. Cela confirme que la modification induite par le caractère « asymptotic preserving » du schéma ne modifie pas la consistance pour le régime de transport. Cependant on remarque que sur maillage aléatoire l'ordre un est atteint pour un nombre important de mailles. Pour ce cas test les conditions limites sont définies à l'aide de mailles fantômes dans lequel on injecte la solution exacte à chaque temps.

Cas test de diffusion

On propose un cas test pour valider le régime limite de diffusion pour le schéma « asymptotic preserving » pair-impair pour les modèles S_n. L'opacité est constante et égale à 1. Les conditions initiales sont définies par $u^i(t = 0, x) = v^i(t = 0, x) = G(0.01, x)$ avec $G(t, x)$ la solution fondamentale associée à l'équation de la chaleur de coefficient $\frac{1}{3}$. Par conséquent $E(t = 0, x) = \sum_{i=1}^{N} w_i(u^i + v^i)(t = 0, x) = G(0.01, x)$. La solution est donnée par $E(t, x) = \sum_{i=1}^{N} w_i(u^i + v^i)(t = 0, x) = G(0.01 + t, x)$. Les ordres de convergence détaillés dans les tableaux (7.4)-(7.5) montrent que le schéma converge en régime de diffusion avec un ordre proche de deux. La décroissance de l'ordre pour les deux premières valeurs de ε s'explique par le fait qu'on compare la solution exacte de diffusion à la solution numérique du modèle S_n. Afin d'étudier le comportement du schéma sur grille fine ou intermédiaire, on propose de faire un test de convergence avec $\frac{\Delta x}{\varepsilon}$ fixé et ε qui tend vers zéro. Les résultats sont répertoriés dans le tableau (7.6). On note que le schéma converge pour la limite $\frac{\Delta x}{\varepsilon}$. La différence importante entre les trois premières erreurs et les deux dernières provient du fait que l'on compare la solution numérique à la solution exacte de diffusion et non à la solution exacte du modèle S_2.

Maillage	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$E_j < 0$
Uniforme 50-100	1.63	1.94	2.00	0
Uniforme 100- 200	1.46	1.90	2.00	0
Uniforme 200- 400	1.28	1.82	1.99	0
Aléatoire 50-100	1.68	2.00	2.05	0
Aléatoire 100- 200	1.45	1.88	1.97	0
Aléatoire 200- 400	1.28	1.80	1.98	0

TABLE 7.4 – Ordre de convergence pour le modèle S_2

maillage	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-6}$	$E_j < 0$
uniforme 50-100	1.57	1.93	2.00	0
uniforme 100- 200	1.40	1.88	2.00	0
uniforme 200- 400	1.23	1.78	1.99	0
aléatoire 50-100	1.69	2.00	2.05	0
Aléatoire 100- 200	1.45	1.88	1.98	0
Aléatoire 200- 400	1.28	1.80	1.98	0

TABLE 7.5 – Ordre de convergence pour le modèle S_6

valeurs des paramètres/maillage	Cartésien	Aléatoire
$\Delta x = 10$, $\varepsilon = 0.1$	0.8580	0.7806
$\Delta x = 50$, $\varepsilon = 0.02$	0.0694	0.0712
$\Delta x = 100$, $\varepsilon = 0.01$	0.0178	0.0181
$\Delta x = 500$, $\varepsilon = 0.002$	0.00071	0.000739
$\Delta x = 1000$, $\varepsilon = 0.001$	0.000177	0.000184

TABLE 7.6 – Erreur entre la solution numérique du modèle S_2 et la solution exacte de diffusion.

7.2 Extension sur maillages non structurés

Pour terminer on propose une rapide réflexion sur l'extension en dimension deux du schéma « pair - impair » . Afin d'obtenir un schéma « asymptotic preserving » sur maillages non structurés, on doit commencer par étendre le schéma de Gosse-Toscani ainsi que la méthode « pair - impair » sur maillages de Delaunay.

7.2.1 Extension de la méthode « pair - impair » sur maillages de Delaunay

Soit $\mathbf{\Omega}_k$, $k \in \{0...N\}$ les vitesses d'ondes discrètes et w_k les poids de quadrature. Les modèles S_n pour l'équation de transport 2D s'écrivent

$$\begin{cases} \partial_t f_k + \dfrac{1}{\varepsilon}\Omega_k^x \partial_x f_k + \dfrac{1}{\varepsilon}\Omega_k^y \partial_y f_k = \dfrac{\sigma}{\varepsilon^2}\left(\displaystyle\sum_{l=1}^{N} w_l (f_l + f_{-l}) - f_k\right), \\ \partial_t f_{-k} - \dfrac{1}{\varepsilon}\Omega_k^x \partial_x f_{-k} - \dfrac{1}{\varepsilon}\Omega_k^y \partial_y f_{-k} = \dfrac{\sigma}{\varepsilon^2}\left(\displaystyle\sum_{l=1}^{N} w_l (f_l + f_{-l}) - f_{-k}\right). \end{cases}$$

Ce système admet également une formulation « pair - impair » qui est donnée par

$$\left\{\begin{array}{l} \partial_t E_k + \dfrac{1}{\varepsilon}\Omega_k^x \partial_x F_k + \dfrac{1}{\varepsilon}\Omega_k^y \partial_y F_k = \dfrac{\sigma}{\varepsilon^2}(\displaystyle\sum_{l=1}^{N} 2w_l E_l - E_k), \\ \partial_t F_k - \dfrac{1}{\varepsilon}\Omega_k^x \partial_x E_k - \dfrac{1}{\varepsilon}\Omega_k^y \partial_y E_k = -\dfrac{\sigma}{\varepsilon^2} F_k. \end{array}\right.$$

Afin d'obtenir l'extension en dimension deux de la méthode « pair - impair » on propose une extension du schéma de Gosse-Toscani en dimension deux. La construction du schéma pour le modèle S_2 est basée sur l'écriture des flux du schéma de Gosse-Toscani dans les directions normales associées aux faces du maillages. On part du système

$$\partial_t \mathbf{U} + \frac{1}{\varepsilon} A \partial_x \mathbf{U} + \frac{1}{\varepsilon} B \partial_y \mathbf{U} = -\frac{\sigma}{\varepsilon^2} D\mathbf{U}, \tag{7.33}$$

avec $\mathbf{U} = (u, v)$ et

$$D = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad A = \begin{pmatrix} \Omega^x & 0 \\ 0 & -\Omega^x \end{pmatrix}, \quad B = \begin{pmatrix} \Omega^y & 0 \\ 0 & -\Omega^y \end{pmatrix}. \tag{7.34}$$

Afin de pouvoir construire le schéma de Gosse-Toscani en dimension deux, on étend les flux 1D dans la direction normale. En conséquence on se place sur le segment $[\mathbf{x}_j, \mathbf{x}_k]$ avec $\mathbf{x}_j$, $\mathbf{x}_k$ les centres des mailles Ω_j et Ω_k. On écrit l'équation de la chaleur hyperbolique diagonalisée dans le repère tangente, normale, puis on projette dans la direction normale. On obtient

$$\left\{\begin{array}{l} \partial_t u + \dfrac{(\mathbf{\Omega}, \mathbf{n}_{jk})}{\varepsilon} \partial_{x'} u = \dfrac{\sigma}{2\varepsilon^2}(v - u), \\ \partial_t v - \dfrac{(\mathbf{\Omega}, \mathbf{n}_{jk})}{\varepsilon} \partial_{x'} v = \dfrac{\sigma}{2\varepsilon^2}(u - v). \end{array}\right. \tag{7.35}$$

À présent on localise le terme source à l'interface. On obtient

$$\left\{\begin{array}{l} \partial_t u + \dfrac{(\mathbf{\Omega}, \mathbf{n}_{jk})}{\varepsilon} \partial_{x'} u = \dfrac{\sigma}{2\varepsilon^2} h_{jk} \delta(x_{jk})(v - u), \\ \partial_t v - \dfrac{(\mathbf{\Omega}, \mathbf{n}_{jk})}{\varepsilon} \partial_{x'} v = \dfrac{\sigma}{2\varepsilon^2} h_{jk} \delta(x_{jk})(u - v). \end{array}\right. \tag{7.36}$$

avec $h_{jk} = d(\mathbf{x}_j, \mathbf{x}_k)$. Une fois ce modèle introduit, il suffit de résoudre le problème stationnaire à l'interface $\partial\Omega_{jk}$ pour construire le flux associé aux $\partial\Omega_{jk}$. Le problème stationnaire est défini par

$$\left\{\begin{array}{l} \dfrac{(\mathbf{\Omega}, \mathbf{n}_{jk})}{\varepsilon} \partial_{x'} u^* = \dfrac{\sigma}{2\varepsilon^2} h_{jk}(v^* - u^*), \\ \\ -\dfrac{(\mathbf{\Omega}, \mathbf{n}_{jk})}{\varepsilon} \partial_{x'} v^* = \dfrac{\sigma}{2\varepsilon^2} h_{jk}(u^* - v^*), \\ u^*(0) = u_L, \quad v^*(h_{jk}) = v_R. \end{array}\right. \tag{7.37}$$

La solution de ce système stationnaire est donnée par

$$\left\{\begin{array}{l} u^* = u_L + (1 - M_{jk})(v_R - u_L)\chi_{(\mathbf{\Omega}, \mathbf{n}_{jk})>0} + u_R - (1 - M_{jk})(u_R - v_L)\chi_{(\mathbf{\Omega}, \mathbf{n}_{jk})<0}, \\ v^* = v_R - (1 - M_{jk})(v_R - u_L)\chi_{(\mathbf{\Omega}, \mathbf{n}_{jk})>0} + v_L - (1 - M_{jk})(u_R - v_L)\chi_{(\mathbf{\Omega}, \mathbf{n}_{jk})<0}, \end{array}\right. \tag{7.38}$$

avec $M_{jk} = \frac{2\varepsilon|(\mathbf{\Omega}, \mathbf{n}_{jk})|}{2\varepsilon|(\mathbf{\Omega}, \mathbf{n}_{jk})| + \sigma_{jk} h_{jk}}$. L'étude du problème de Riemann montre que le flux est défini par

$$F_{jk} = \begin{pmatrix} (\mathbf{\Omega}, \mathbf{n}_{jk}) u_j \\ -(\mathbf{\Omega}, \mathbf{n}_{jk}) v^* \end{pmatrix} \chi_{(\mathbf{\Omega}, \mathbf{n}_{jk})>0} + \begin{pmatrix} (\mathbf{\Omega}, \mathbf{n}_{jk}) u^* \\ -(\mathbf{\Omega}, \mathbf{n}_{jk}) v_j \end{pmatrix} \chi_{(\mathbf{\Omega}, \mathbf{n}_{jk})<0},$$

$$F_{jk} = \left(\begin{array}{l} (\mathbf{\Omega}, \mathbf{n}_{jk}) u_j \\ -(\mathbf{\Omega}, \mathbf{n}_{jk})(v_k - (1 - M_{jk})(v_k - u_j)) \end{array} \right) \chi_{(\mathbf{\Omega}, \mathbf{n}_{jk})>0}$$
$$+ \left(\begin{array}{l} (\mathbf{\Omega}, \mathbf{n}_{jk})(u_k - (1 - M_{jk})(u_k - v_j)) \\ -(\mathbf{\Omega}, \mathbf{n}_{jk}) v_j \end{array} \right) \chi_{(\mathbf{\Omega}, \mathbf{n}_{jk})<0}.$$

Une fois les flux construits à chaque interface, on dérive le schéma

$$\left\{ \begin{array}{l} |\Omega_j| \partial_t \left(\begin{array}{c} u_j \\ v_j \end{array} \right) + \dfrac{1}{\varepsilon} \displaystyle\sum_k l_{jk} \mathbf{F}_{jk} = \mathbf{0} \\ \mathbf{F}_{jk} = G\mathbf{n}_{jk}^+ \left(\begin{array}{c} u_j \\ v_j \end{array} \right) + G\mathbf{n}_{jk}^- \left(\begin{array}{c} u_k - (1 - M_{jk})(u_k - v_j) \\ v_k - (1 - M_{jk})(v_k - u_j) \end{array} \right) \end{array} \right. \tag{7.39}$$

avec $G\mathbf{n}_{jk} = An_{jk}^x + Bn_{jk}^y$ et avec $M_{jk} = \frac{2\varepsilon |(\mathbf{\Omega}, \mathbf{n}_{jk})|}{2\varepsilon |(\mathbf{\Omega}, \mathbf{n}_{jk})| + \sigma_{jk} h_{jk}}$. Enfin on réécrit ce dernier afin de faire apparaître le terme source.

Proposition 7.3. *Le schéma (7.39) est équivalent à*

$$\left\{ \begin{array}{l} |\Omega_j| \partial_t \left(\begin{array}{c} u_j \\ v_j \end{array} \right) + \dfrac{1}{\varepsilon} \displaystyle\sum_k l_{jk} M_{jk} \mathbf{F}_{jk} = \sum_k \frac{M_{jk} \sigma_{jk}}{2\varepsilon^2} l_{jk} h_{jk} \left(\left(\begin{array}{c} \frac{v_j - u_j}{2} \\ \frac{u_j - v_j}{2} \end{array} \right) \right), \\ \mathbf{F}_{jk} = \frac{G\mathbf{n}_{jk}}{2} \left(\begin{array}{c} u_j + u_k \\ v_j + v_k \end{array} \right) + \frac{|G\mathbf{n}_{jk}|}{2} \left(\begin{array}{c} u_j - u_k \\ v_j - v_k \end{array} \right) - \frac{G\mathbf{n}_{jk}}{2} \left(\begin{array}{c} u_j + v_j \\ u_j + v_j \end{array} \right), \end{array} \right. \tag{7.40}$$

avec $G\mathbf{n}_{jk} = An_{jk}^x + Bn_{jk}^y$ *et* $M_{jk} = \frac{2\varepsilon |(\mathbf{\Omega}, \mathbf{n}_{jk})|}{2\varepsilon |(\mathbf{\Omega}, \mathbf{n}_{jk})| + \sigma_{jk} h_{jk}}$.

Démonstration. Le but est d'écrire le schéma sous une forme proche de la forme (7.8)-(7.9) introduite en 1D, avec un terme source centré dans les mailles.
On commence par remarquer que la relation

$$\mathbf{F}_{jk} = G\mathbf{n}_{jk}^+ \left(\begin{array}{c} u_j \\ v_j \end{array} \right) + G\mathbf{n}_{jk}^- \left(\begin{array}{c} M_{jk} u_k \\ M_{jk} v_k \end{array} \right) + G\mathbf{n}_{jk}^- \left(\begin{array}{c} (1 - M_{jk}) v_j \\ (1 - M_{jk}) u_j \end{array} \right),$$

est équivalente à

$$\mathbf{F}_{jk} = M_{jk} \left(G\mathbf{n}_{jk}^+ \left(\begin{array}{c} u_j \\ v_j \end{array} \right) + G\mathbf{n}_{jk}^- \left(\begin{array}{c} u_k \\ v_k \end{array} \right) \right) + G\mathbf{n}_{jk}^+ \left(\begin{array}{c} (1 - M_{jk}) u_j \\ (1 - M_{jk}) v_j \end{array} \right) + G\mathbf{n}_{jk}^- \left(\begin{array}{c} (1 - M_{jk}) v_j \\ (1 - M_{jk}) u_j \end{array} \right).$$

Maintenant on étudie le terme $T_{jk} = G\mathbf{n}_{jk}^+ \left(\begin{array}{c} (1 - M_{jk}) u_j \\ (1 - M_{jk}) v_j \end{array} \right) + G\mathbf{n}_{jk}^- \left(\begin{array}{c} (1 - M_{jk}) v_j \\ (1 - M_{jk}) u_j \end{array} \right)$. De rapides calculs montrent que le terme T_{jk} s'écrit

$$T_{jk} = (1 - M_{jk}) G\mathbf{n}_{jk}^+ \left(\begin{array}{c} \frac{u_j - v_j}{2} \\ \frac{v_j - u_j}{2} \end{array} \right) + (1 - M_{jk}) G\mathbf{n}_{jk}^- \left(\begin{array}{c} \frac{v_j - u_j}{2} \\ \frac{u_j - v_j}{2} \end{array} \right) + (1 - M_{jk}) G\mathbf{n}_{jk} \left(\begin{array}{c} \frac{u_j + v_j}{2} \\ \frac{u_j + v_j}{2} \end{array} \right).$$

Afin d'obtenir l'équation précédente on a utilisé $G\mathbf{n}_{jk} = G\mathbf{n}_{jk}^+ + G\mathbf{n}_{jk}^-$. On injecte ces nouvelles formules de flux dans le schéma et on utilise le fait que $\sum_k l_{jk} G\mathbf{n}_{jk} = \mathbf{0}$ ce qui donne

$$\partial_t \left(\begin{array}{c} u_j \\ v_j \end{array} \right) + \frac{1}{\varepsilon} \sum_k l_{jk} M_{jk} \left(G\mathbf{n}_{jk}^+ \left(\begin{array}{c} u_j \\ v_j \end{array} \right) + G\mathbf{n}_{jk}^- \left(\begin{array}{c} u_k \\ v_k \end{array} \right) \right) = \frac{1}{\varepsilon} \sum_k l_{jk} M_{jk} G\mathbf{n}_{jk} \left(\begin{array}{c} \frac{u_j + v_j}{2} \\ \frac{u_j + v_j}{2} \end{array} \right) \tag{7.41}$$

$$-\frac{1}{\varepsilon}\sum_k l_{jk}(1-M_{jk})\left(G\mathbf{n}_{jk}^+\begin{pmatrix}\frac{u_j-v_j}{2}\\ \frac{v_j-u_j}{2}\end{pmatrix}+G\mathbf{n}_{jk}^-\begin{pmatrix}\frac{v_j-u_j}{2}\\ \frac{u_j-v_j}{2}\end{pmatrix}\right) \tag{7.42}$$

On note $|G\mathbf{n}_{jk}| = G\mathbf{n}_{jk}^+ - G\mathbf{n}_{jk}^-$, ce qui permet de simplifier le terme source pour obtenir

$$\partial_t\begin{pmatrix}u_j\\ v_j\end{pmatrix}+\frac{1}{\varepsilon}\sum_k l_{jk}M_{jk}\left(G\mathbf{n}_{jk}^+\begin{pmatrix}u_j\\ v_j\end{pmatrix}+G\mathbf{n}_{jk}^-\begin{pmatrix}u_k\\ v_k\end{pmatrix}\right)=\frac{1}{\varepsilon}\sum_k l_{jk}M_{jk}G\mathbf{n}_{jk}\begin{pmatrix}\frac{u_j+v_j}{2}\\ \frac{u_j+v_j}{2}\end{pmatrix} \tag{7.43}$$

$$-\frac{1}{\varepsilon}\sum_k l_{jk}(1-M_{jk})\left(|G\mathbf{n}_{jk}|\begin{pmatrix}\frac{v_j-u_j}{2}\\ \frac{u_j-v_j}{2}\end{pmatrix}\right). \tag{7.44}$$

Enfin on utilise $l_{jk}\frac{(1-M_{jk})}{\varepsilon}|(\mathbf{\Omega},\mathbf{n}_{jk})| = M_{jk}\frac{\sigma_{jk}}{2\varepsilon^2}l_{jk}h_{jk}$ pour obtenir le résultat final. □

7.2.2 Schéma « pair-impair » sur maillages de Delaunay et propriétés

Maintenant que le schéma de Gosse-Toscani est étendu en dimension deux, on propose d'étendre la formulation « pair - impair » en dimension deux, puis d' étudier le régime de diffusion.

Définition 7.2.1. *Le schéma « pair - impair » pour les modèles S_n sur maillages de Delaunay est défini par*

$$\begin{cases}|\Omega_j|\partial_t\begin{pmatrix}u_j^l\\ v_j^l\end{pmatrix}+\frac{1}{\varepsilon}\sum_k l_{jk}M_{jk}^l\mathbf{F}_{jk}=\sum_k\frac{M_{jk}^l\sigma_{jk}}{2\varepsilon^2}l_{jk}h_{jk}\begin{pmatrix}\frac{v_j^l-u_j^l}{2}\\ \frac{u_j^l-v_j^l}{2}\end{pmatrix}\\ \\ +\begin{pmatrix}\sum_{m=1}^N 2w_mN_j^{l,m}(u_j^m+v_j^m-(u_j^l+v_j^l))\\ \sum_{m=1}^N 2w_mN_j^{l,m}(u_j^m+v_j^m-(u_j^l+v_j^l))\end{pmatrix},\\ \\ \mathbf{F}_{jk}=\frac{G\mathbf{n}_{jk}}{2}\begin{pmatrix}u_j^l+u_k^l\\ v_j^l+v_k^l\end{pmatrix}+\frac{|G\mathbf{n}_{jk}|}{2}\begin{pmatrix}u_j^l-u_k^l\\ v_j^l-v_k^l\end{pmatrix}-\frac{G\mathbf{n}_{jk}}{2}\begin{pmatrix}u_j^l+v_j^l\\ u_j^l+v_j^l\end{pmatrix}.\end{cases} \tag{7.45}$$

avec $G\mathbf{n}_{jk} = An_{jk}^x + Bn_{jk}^y$ et $M_{jk}^l = \frac{2\varepsilon|(\mathbf{\Omega}_l,\mathbf{n}_{jk})|}{2\varepsilon|(\mathbf{\Omega}_l,\mathbf{n}_{jk})|+\sigma_{jk}h_{jk}}$. Le coefficient $N_j^{l,m}$ est défini par

$$N_j^{l,m} = \min\left(\min_k M_{jk}^l\frac{\sigma_{jk}h_{jk}}{\varepsilon^2}, \min_k M_{jk}^m\frac{\sigma_{jk}h_{jk}}{\varepsilon^2}\right).$$

Enfin montre que le schéma (7.45) admet au sens de Hilbert le schéma limite sur maillages de Delaunay.

Proposition 7.4. *Le schéma « pair-impair » (7.45) admet le schéma limite*

$$|\Omega_j|\partial_t E_j - \sum_k l_{jk}\frac{E_k-E_j}{3\sigma_{jk}} = 0, \tag{7.46}$$

avec $E_j = \sum_{l=1}^N w_lE_j^l = \sum_{l=1}^N 2w_lE_j^l = \sum_{l=1}^N 2w_l(u_j^l+v_j^l)$.

Démonstration. On commence par réécrire le schéma (7.40) avec les variables $E^l = u^l + v^l$ et $E^l = u^l - v^l$. On obtient

$$\begin{cases}|\Omega_j|\partial_t\begin{pmatrix}E_j^l\\ F_j^l\end{pmatrix}+\frac{1}{\varepsilon}\sum_k l_{jk}M_{jk}^l\mathbf{F}_{jk}=\begin{pmatrix}\sum_{m=1}^N 2w_mN_j^{l,m}(E_j^m-E_j^l)\\ -\sum_k\frac{M_{jk}^l\sigma_{jk}}{2\varepsilon^2}l_{jk}h_{jk}F_j^l\end{pmatrix},\\ \\ \mathbf{F}_{jk}=\frac{(\mathbf{\Omega}_l,\mathbf{n}_{jk})}{2}\begin{pmatrix}F_j^l+F_k^l\\ E_j^l+E_k^l\end{pmatrix}+\frac{|(\mathbf{\Omega}_l,\mathbf{n}_{jk})|}{2}\begin{pmatrix}E_j^l-E_k^l\\ F_j^l-F_k^l\end{pmatrix}-(\mathbf{\Omega}_l,\mathbf{n}_{jk})\begin{pmatrix}0\\ E_j^l\end{pmatrix}.\end{cases} \tag{7.47}$$

avec $M_{jk}^l = \frac{2\varepsilon|(\mathbf{\Omega}_l,\mathbf{n}_{jk})|}{2\varepsilon|(\mathbf{\Omega}_l,\mathbf{n}_{jk})|+\sigma_{jk}h_{jk}}$. On commence par poser $M_{jk}^{l,*} = \frac{M_{jk}^l}{\varepsilon} = \frac{2|(\mathbf{\Omega}_l,\mathbf{n}_{jk})|}{2\varepsilon|(\mathbf{\Omega}_l,\mathbf{n}_{jk})|+\sigma_{jk}h_{jk}}$, puis on applique un développement de Hilbert sur le schéma (7.47).
Les termes homogènes à $\frac{1}{\varepsilon}$ sont :

- $\sum_{m=1}^N 2w_m N_j^{l,m}(u_j^m+v_j^m-(u_j^l+v_j^l)) = 0$, avec $N_j^{l,m} = \min\left(\min_k M_{jk}^{l,*}\frac{\sigma_{jk}h_{jk}}{\varepsilon^2}, \min_k M_{jk}^{m,*}\frac{\sigma_{jk}h_{jk}}{\varepsilon^2}\right)$,
- $-\sum_k \frac{M_{jk}^l\sigma_{jk}}{2\varepsilon^2} l_{jk}h_{jk}F_j^l = \mathbf{0}$, ce qui implique que $F_j^l = \mathbf{0}$.

On remarque que lorsque ε tend vers 0, le terme $\sum_{m=1}^N 2w_m N_j^{l,m}(E_j^m - E_j^l) = 0$ est équivalent à

$$\sum_{m=1}^N 2w_m \min\left(\min_k(\mathbf{\Omega}_l,\mathbf{n}_{jk}), \min_k(\mathbf{\Omega}_m,\mathbf{n}_{jk})\right)(E_j^m - E_j^l) = 0.$$

Avec le même type d'argument que pour la preuve en 1D, on obtient $E_j^m - E_j^l = 0, \forall l,m$. À présent on étudie le terme d'ordre $\frac{1}{\varepsilon^0}$ associé à la première équation en utilisant $F_j^l = 0$. On obtient

$$|\Omega_j|\partial_t E_j^l + \sum_k l_{jk}M_{jk}^{l,*}|(\mathbf{\Omega}_l,\mathbf{n}_{jk})|\frac{E_j^l - E_k^l}{2} = 0. \tag{7.48}$$

On pose $E_j^* = E_j^l$ pour tout l. En utilisant $E_j^m - E_j^l = 0$, on obtient $E_j = \sum_{m=1}^N 2w_m E_j^m = E_j^*(\sum_{m=1}^N 2w_m) = E_j^*$ car $\sum_{m=1}^N 2w_m = 1$. Afin d'obtenir la limite de diffusion, on multiplie par w_l et on somme sur l les équations (7.48),

$$|\Omega_j|\partial_t E_j + \sum_l w_l\left(\sum_k l_{jk}M_{jk}^{l,*}|(\mathbf{\Omega}_l,\mathbf{n}_{jk})|\frac{E_j - E_k}{2}\right) = 0. \tag{7.49}$$

Lorsque ε tend vers 0 l'équation précédente est équivalente à

$$|\Omega_j|\partial_t E_j - \sum_k l_{jk}\left(\sum_l w_l|(\mathbf{\Omega}_l,\mathbf{n}_{jk})|^2\right)\frac{E_k - E_j}{\sigma_{jk}h_{jk}} = 0. \tag{7.50}$$

Enfin il suffit de montrer que le coefficient de diffusion est valide. A cet effet on n'utilise que

$$\sum_l w_l(\mathbf{\Omega}_l,\mathbf{n}_{jk})^2 = \left(\mathbf{n}_{jk}(\sum_l w_l\mathbf{\Omega}_l\otimes\mathbf{\Omega}_l),\mathbf{n}_{jk}\right) = \frac{1}{3}(\mathbf{n}_{jk},\mathbf{n}_{jk}) = \frac{1}{3}$$

et on obtient donc la limite de diffusion suivante

$$|\Omega_j|\partial_t E_j - \sum_k l_{jk}\frac{E_k - E_j}{3\sigma_{jk}h_{jk}} = 0. \tag{7.51}$$

□

Remarques

La robustesse de la méthode « pair - impair », utilisée pour discrétiser les modèles S_n dépend de la relation entre les vitesses discrètes et les normales associées au maillage. Par exemple, pour le modèle S_2 (quatre vitesses), si les vitesses discrètes sont alignées avec les normales associées à une maille, le schéma n'est pas convergent. En effet, lorsque $\sigma = 0$, le coefficient M_{jk} n'est pas défini. Lorsque σ est différent de zéro, le coefficient $N_j^{l,m}$ est nul. Par conséquent, on obtient des équations de diffusion sur E_j^l découplées entre elles, et l'équilibre entre les fonctions E_j^l n'est pas atteint. Dans ce cas, le schéma ne capture pas la limite de diffusion du modèle continu. Pour compléter cette étude il faut :

- montrer que le principe du maximum est préservé par le schéma (7.45) (même principe de preuve qu'en 1D).
- étudier la validité de la discrétisation du terme source, notamment le terme de couplage entre les différents modèles de tailles deux.

7.2.3 Correction du schéma sur maillage non structurés

Idée : *Le schéma TPFA (qui est la limite du schéma de Gosse-Toscani (7.45)) n'est par convergent. Par conséquent on propose donc de modifier la viscosité numérique du schéma (7.45) en utilisant un schéma convergent sur maillages non structurés, positif, appelé le schéma LMP [DlP09] qui peut s'interpréter comme une correction non linéaire du schéma TPFA.*

Le flux du schéma TPFA pour une variable u s'écrit

$$\sum_k F_{jk}(u)^{TPFA} = -\sum_k \frac{l_{jk}}{d(\mathbf{x}_j, \mathbf{x}_k)}(u_k - u_j).$$

De son coté, le flux du schéma LMP pour une variable u s'écrit

$$\sum_k F_{jk}(u)^{lmp} = \sum_k -a(u_k - u_j) + G(u),$$

avec $G(u)$ un terme non linéaire. Sur maillage Cartésien $G(u) = 0$ et $a = \frac{l_{jk}}{d(\mathbf{x}_j,\mathbf{x}_k)}$, par conséquent schéma LMP dérive sur le schéma TPFA pour un maillage Cartésien. A partir de cette remarque, on propose un couplage entre le schéma (7.45) pour les modèles S_n et le schéma de diffusion LMP [DlP09].

$$\left\{ \begin{array}{l} |\Omega_j|\partial_t \begin{pmatrix} u_j^l \\ v_j^l \end{pmatrix} + \frac{1}{\varepsilon}\sum_k l_{jk} M_{jk}^l \mathbf{F}_{jk} = \sum_k \frac{M_{jk}^l \sigma_{jk}}{2\varepsilon^2} l_{jk} h_{jk} \left(\begin{pmatrix} \frac{v_j^l - u_j^l}{2} \\ \frac{u_j^l - v_j^l}{2} \end{pmatrix} \right), \\ \\ + \begin{pmatrix} \sum_{m=1}^N 2w_m N_j^{l,m}(u_j^m + v_j^m - (u_j^l + v_j^l)) \\ \sum_{m=1}^N 2w_m N_j^{l,m}(u_j^m + v_j^m - (u_j^l + v_j^l)) \end{pmatrix} \\ \\ \mathbf{F}_{jk} = \frac{G\mathbf{n}_{jk}}{2} \begin{pmatrix} u_j^l + u_k^l \\ v_j^l + v_k^l \end{pmatrix} + \frac{|G\mathbf{n}_{jk}| h_{jk}}{2 l_{jk}} \begin{pmatrix} F_{jk}(u^l)^{lmp} \\ F_{jk}(v^l)^{lmp} \end{pmatrix} - \frac{G\mathbf{n}_{jk}}{2} \begin{pmatrix} u_j^l + v_j^l \\ u_j^l + v_j^l \end{pmatrix}, \end{array} \right. \tag{7.52}$$

avec $G\mathbf{n}_{jk} = An_{jk}^x + Bn_{jk}^y$ et $M_{jk} = \frac{2\varepsilon|(\mathbf{\Omega},\mathbf{n}_{jk})|}{2\varepsilon|(\mathbf{\Omega},\mathbf{n}_{jk})| + \sigma_{jk} h_{jk}}$. Ce couplage permet de dériver sur le schéma LMP en régime de diffusion sur maillage grossier. On peut donc s'attendre à ce que le schéma modifié converge sur des maillages non structurés grossiers. Cependant, cette modification ne donne pas d'indication sur le comportement de la méthode sur maillages fins.

7.3 Synthèse et perspectives

Dans ce chapitre, on a proposé un schéma « asymptotic preserving » et positif pour les modèles S_n sur grilles 1D non uniformes. Les résultats numériques montrent que le schéma est « asymptotic preserving » positif et capable de préserver certains états stationnaires. Cette méthode nous paraît être un bon candidat pour une extension en dimension deux. Pour obtenir une discrétisation convergente sur maillages non structurés, on propose d'étendre sur maillages de Delaunay la méthode proposée en

écrivant les flux 1D dans la direction normale, puis de modifier la viscosité numérique à l'aide d'un schéma de diffusion positif qui peut être vu comme une correction non linéaire du schéma TPFA. Avant d'obtenir une méthode valide, il reste cependant du travail à accomplir. D'abord il faut vérifier numériquement le caractère « asymptotic preserving » du schéma. Il existe dans ce sens des résultats encourageants au moins sur maillages grossiers. Sur maillage fin, cela semble plus délicat et une étude précise de la discrétisation géométrique des termes sources semble nécessaire. Pour conclure il est important de montrer que le schéma final pour les modèles S_n est positif et stable en norme L^2.

Conclusion et Perspectives

Conclusion

Schémas « asymptotic preserving » pour l'équation de la chaleur hyperbolique

Dans les chapitres deux et trois, nous avons proposé des schémas « asymptotic preserving » pour l'équation de la chaleur hyperbolique sur maillages non structurés. Après avoir montré dans le chapitre un que l'extension des schémas « asymptotic preserving » en dimension deux, en utilisant un formalisme aux arêtes, n'était pas convergente dans le régime de diffusion, on a proposé deux différentes stratégies pour obtenir des discrétisations valides en régime de diffusion.
La première stratégie est basée sur l'utilisation d'un formalisme aux noeuds des méthodes de volumes finis. Ce formalisme aux noeuds [CDDL09]- [Des09], couplé avec la méthode de Jin-Levermore [JL96] nous a permis de construire deux schémas « asymptotic preserving » pour lesquels on a pu donner un certain nombre de résultats théoriques. En outre, un de ces schémas admet une variante semi-implicite stable indépendamment du paramètre de relaxation ε.
La seconde stratégie utilise un point de vue inverse. On commence par choisir un schéma cible de diffusion et par construire un schéma hyperbolique compatible géométriquement avec le schéma de diffusion. Enfin on modifie le schéma hyperbolique pour obtenir une discrétisation « asymptotic preserving » . À l'aide de cette stratégie on a obtenu deux nouvelles méthodes basées sur les schémas [AE06]- [BM06].

Discrétisation de systèmes linéaires généraux

Nous avons proposé une méthode permettant d'obtenir des discrétisations « asymptotic preserving » pour des modèles linéaires généraux issus notamment de l'approximation angulaire des équations de transport linéaires. Cette décomposition consiste à isoler la partie dite « diffusive » qui donne la limite de diffusion dans le système hyperbolique. À partir de cette décomposition on constate que seule la partie intervenant en limite de diffusion nécessite une discrétisation « asymptotic preserving » . La partie « diffusive » étant homogène à l'équation de la chaleur hyperbolique, cette décomposition permet de faire porter la difficulté uniquement sur la discrétisation de l'équation de la chaleur hyperbolique. On a donc proposé une méthode simple permettant de construire des discrétisations « asymptotic preserving » pour des modèles linéaires généraux.

Discrétisation de modèles hyperboliques non linéaires

Les deux derniers chapitres ont étudié la discrétisation de modèles non linéaires issus de la photonique. Dans un premier temps nous avons montré que les discrétisations spatiales « asymptotic preserving » proposés précédemment restent valides pour le problème de couplage non-linéaire entre le modèle P_1 et une équation sur la matière. Dans un second temps on a étudié la discrétisation d'un modèle hyperbolique non-linéaire. Ce modèle admet un limite de diffusion et un principe du maximum. En utilisant une formulation différente du système, un schéma Lagrange+projection aux noeuds et la

méthode Jin-Levermore, on a obtenu un schéma « asymptotic preserving » et préservant le principe du maximum. Cette étude montre que les idées introduites, notamment dans le chapitre deux, peuvent être généralisées pour des modèles non linéaires. Une étude extension de ces résultats aux équations d'Euler et Saint Venant avec friction et gravité afin d'obtenir des schémas « asymptotic preserving » et « well-balanced » paraît possible.

Perspectives

Méthodes numériques introduites dans cette étude

Perspectives autour des maillages et de l'extension en 3D

Nous avons proposé plusieurs méthodes valides sur des maillages polygonaux non structurés. Cependant il reste une certain nombre de questions ouvertes. Par exemple il serait intéressant d'obtenir des conditions d'inversibilité de la matrice nodale nécessaires et suffisantes. En effet, nous n'avons pu donner qu'une condition suffisante et une étude numérique sur l'inversibilité de ces matrices locales. L'obtention de ces conditions nécessaires et suffisantes permettrait de déterminer les limites du schéma. Une autre question que nous n'avons que partiellement traitée est celle de la stabilisation des modes parasites. En effet nous avons proposé une méthode basée sur la définition de plusieurs jeux de normales que l'on applique aux schémas avant de moyenniser. Cette méthode est certes efficace mais introduit un surcout dans la construction de la matrice du schéma qui devient importante lorsqu'elle varie à chaque pas de temps. Il serait intéressant de construire un terme de stabilisation simple qui ne modifie pas les propriétés des schémas.
Dans cette étude nous n'avons raisonné que sur des maillages polygonaux 2D. Il serait donc utile d'étudier comment les méthodes proposées peuvent être importées sur maillages coniques (les arêtes sont définies par des coniques) ou sur maillages 3D. Pour les maillages coniques, plusieurs travaux ont récemment été réalisés dans le cadre du schéma GLACE ([CDLL11])-([BEHN10]). Actuellement P. Hoch and G. Samba cherchent à étendre ces résultats aux schémas de diffusion et « asymptotic preserving » aux noeuds présentés dans cette thèse. Pour la dimension trois, les cas sont différents. Le schéma GLACE-hydrodynamique est un schéma décrit en dimension quelconque grâce aux $\mathbf{C}_{jr} = \nabla_{\mathbf{x}_r}\Omega_j$ qui généralisent le produit $l_{jr}\mathbf{n}_{jr}$ en dimension supérieur. A priori les schémas « asymptotic preserving » basés sur le formalisme GLACE devraient se généraliser de la même façon en dimension supérieure que le schéma hydrodynamique. Bien sur, cela reste à vérifier. Pour les schémas utilisant une localisation des flux aux arêtes l'extension 3D paraît moins naturelle.

Perspectives autour de la discrétisation temporelle

Il y a deux points concernant la discrétisation temporelle que nous n'avons pas réellement traités dans cette étude. Le premier est l'obtention de condition CFL rigoureuse pour les schémas « asymptotic preserving » explicite et semi-implicite sur maillage non structurés. Le second point que nous n'avons pas abordé est la discrétisation implicite des schémas non linéaires (qui interviennent dans le cadre du couplage avec la matière ou pour la discrétisation du modèle M_1). Puisque les opacités peuvent varier de façon importante la CFL parabolique peut devenir très contraignante et l'utilisation de schéma implicite devient nécessaire. Les schémas précédemment décrits s'écrivent de façon implicite de façon naturelle, seules les questions de la linéarisation et du point fixe à adopter restent en suspens.

Perspectives autour de la discrétisation d'ordre élevé

Une autre question intéressante et rarement étudiée dans le cadre de l'étude des schémas « asymptotic preserving » est celle des schémas d'ordre élevé en espace (on peut citer cependant les travaux de R. Natalini et M. Ribot). En effet il est montré dans [McHL10] qu'une technique classique de type MUSCL ne donne pas de bon résultat en régime de diffusion. Les auteurs montrent que le choix du limiteur est très important et peut modifier fortement le résultat. Il serait donc utile d'étudier comment se comportent les différents limiteurs de pentes en 2D sur maillages non structurés.

Perspectives autour de l'étude théorique des méthodes proposées

D'un point de vue théorique, il conviendrait de démontrer la convergence des schémas « asymptotic preserving » introduits dans les chapitres deux et trois avec des estimations d'erreur indépendantes de ε. Il s'agit d'un problème difficile, il serait donc utile de commencer par démontrer la convergence du schéma de Gosse-Toscani ou de Jin-Levermore en 1D sur grille non uniforme afin de comprendre les difficultés inhérentes à ce type de preuve. En ce qui concerne les schémas et modèles non linéaires introduits dans les chapitres 6 et 7, on pourrait se baser sur le résultat de consistance faible obtenu dans [Des09] pour le schéma GLACE Lagrangien.

Applications des méthodes « asymptotic preserving » à la physique

Modèles multi-groupes en vitesse et fréquence

Si l'on veut obtenir des approximations précises des problèmes de neutronique où de transfert radiatif, il faut discrétiser des modèles multi-groupes en fréquence pour la photonique et en énergie pour la neutronique. En effet il peut y avoir de grandes variations selon les valeurs des sections efficaces et des opacités. Pour l'extension multi-groupes des modèles P_n et S_n la principale difficulté porte sur les spatiale te temporelle des termes sources. Pour un système tel que le modèle M_1 la situation est par contre plus compliquée. En effet, l'extension multi-groupes en fréquence ou en vitesse modifie le facteur d'Eddington. Or la reformulation du modèle M_1 comme un système de la dynamique des gaz n'est à priori valable que pour le tenseur issu de l'équation de transport mono-groupe. Il est donc important de construire des schémas valables pour n'importe quel tenseur d'Eddington. On pourrait s'inspirer de la démarche utilisée pour obtenir des schémas positifs pour les modèles S_n.

Systèmes dérivés des modèles P_n

Dans la première partie de ce travail nous avons introduit des modèles dérivés des systèmes P_n tels que les modèles PP_n, FP_n. Là aussi il serait intéressant de regarder si ces différents modèles rentrent dans le cadre des méthodes proposées ici. En effet, la positivité, notamment de la fonction de répartition et du premier moment sont importantes.

Bibliographie

[ABBM96] I. Aavatsmark, T. Barkve, O. Boe, T. Mannseth, *Discretization on non-orthogonal, quadrilateral grids for inhomogeneous, anisotropic media*, J. Comput. Phys. vol. 127, 1996, p. 2-14.

[AE06] I. Aavatsmark, G. Eigestad, *Numerical convergence of the MPFA O-method and U-method for general quadrilateral grids*, Int J. Numer. Meth. Fluids 2006, 51 939-961.

[AG11] G. Allaire, F. Golse *Transport et diffusion*, cours de l'école Polytechnique.

[Al05] G. Allaire *analyse numérique et optimisation*, Éditions de l'École Polytechnique 2005, ISBN : 2-7302-1255-8 .

[AM08] L. Angelas, R. Masson, *Convergence of the finite volume MPFA O scheme for heterogeneous anisotropic diffusion problems on general meshes*, C. R. math. 346, 1007-1012 (2008).

[AEH] L. Angelas, R. Eymard, R. Herbin ,*A nine-point finite volume scheme for the simulation of diffusion in heterogeneous media*, Comptes Rendus Mathématiques. Volume 347, Issues 11-12, June 2009, Pages 673-676.

[BM06] J. Breil, P-H. Maire, *A cell-centered diffusion scheme on two-dimensional unstructured meshes*, JCP 224 (2007) 785-823

[AMTV04] S. Atzeni, J. Meyer-Ter-Vehn, *The physics of inertial fusion*, Oxford Science publications, Oxford University Press, 2004.

[BD07] C. Berthon, C. Buet, J.-F. Coulombel, B. Desprès, J. Dubois, T. Goudon, J.E. Morel et R. Turpault, *Mathematical Models and Numerical Methods for Radiative Transfer*, Panoramas et Synthèses 28.

[BCT08] C. Berthon, P. Charrier, R. Turpault, *An HLLC scheme to solve the M_1 model of radiative transfer in two dimensions*, J. Scie. Comput., J. Sci. Comput., 31, no. 3, pp. 347-389, 2007.

[B010] C. Berthon, J. Dubois, B. Dubroca, T.-H. Nguyen-Bui, R. Turpault, *A free streaming contact preserving scheme for the M_1 model*, Advances in Applied Mathematics and Mechanics Adv. Appl. Math. Mech., 3 (2010), pp. 259-285.

[BLeFT11] C. Berthon, P. LeFloch, R. Turpault, *Late-time relaxation limits of nonlinear hyperbolic systems. A general framework*, Math. of Comp.

[BMT11] C. Berthon, F. Marche, R. Turpault, *An efficient scheme on wet/dry transitions for shallow water equations with friction*, PComputer and Fluids 48 (2011) pp 192-201.

[BT10] C. Berthon, R. Turpault, *Asymptotic preserving HLL schemes*, Numerical Methods for Partial Differential Equations, 27 (6) (2011) pp 1396-1422.

[BOP07] F. Bouchut, H. Ounaissa, B. Perthame, *Upwinding of source term at interfaces for Euler equations with high friction*, Computers and Mathematics with Applications, 53(3-4) :361–375, 2007.

[Bou04] F. Bouchut *Nonlinear stability of finite volumes methods for hyperbolic conservation laws and well-balanced schemes for sources* Frontiers in mathematics, 2004

[BEHN10] B. Boutin, E. Deriaz, P. Hoch, P. Navarro, *Extention of ALE methodology to unstructured conical meshes*, ESAIM : Proceedings, Octobre 2011, Vol 32, p 31-55.

[BLS05] F. Brezzi, K. Lipnikov, V. Simoncini *A family of mimetic finite difference methods on polygonal and polyhedral meshes*, Math. Models Methods Appl. Sci. 15 (2005) 1533-1553.

[Bru05] T. Brunner, *Riemann solvers for time-dependant transport based on the maximum entropy and spherical harmonics closures*, Ph.D thesis Los alamos.

[Bru02] T. Brunner, *Form of approximate radiation transport*, Sandia report, 2002.

[Bru05] T. Brunner, J. P. Holloway, *Two-dimensional time dependent Riemann solvers for neutron transport*, JCP 210 386-399, 2005.

[BC07] C. Buet, S. Cordier, *An asymptotic preserving scheme for hydrodynamics radiative transfer models : numerics for radiative transfer*, Numer. Math., 108, 199-221 (2007).

[BCLM02] C. Buet, S. Cordier,B. Lucquin-Desreux, S. Mancini, *Diffusion limit of the lorentz model : asymptotic preserving schemes*, ESAIM : M2AN vol. 36, 4, 2002, pp 631-655.

[BD06] C. Buet, B. Després, *Asymptotic preserving and positive schemes for radiation hydrodynamics*, JCP 2006, vol. 215, no2, pp. 717-740.

[BDF11] C. Buet, B. Després, E. Franck, *Design of asymptotic preserving schemes for the hyperbolic heat equation on unstructured meshes* Numerish Mathematik, en ligne.

[BDFCras] C. Buet, B. Després, E. Franck, *Asymptotic preserving scheme with maximum principle for non linear radiative transfer model on unstructured meshes*, Note Cras, accepté.

[BDFproc] C. Buet, B. Després, E. Franck, *Asymptotic preserving finite volumes discretization for non-linear moment model on unstructured meshes*, Proceedings FVCA6 2011.

[BD04] C. Buet, B. Desprès, *Asymptotic analysis of fluid models for the coupling of radiation and hydrodynamics*, Journal of Quantitative Spectroscopy and Radiative Transfer 85 (2004) 385-418.

[CS10] P. Cargo, G. Samba, *Resolution of the time dependant P_n equations by a godunov scheme having diffusion limit*, ESAIM : M2AN 44 (2010) 1193-1224.

[CDDL09] G. Carré, S. Del Pino, B. Desprès, E. Labourasse, *A Cell-centered lagrangian hydrodynamics scheme on general unstructured meshes in arbitrary dimension*, JCP vol. 228 (2009) no14, pp. 5160-518.

[CCGRS010] C. Chalons F. Coquel, E. Godlewski, P.-A. Raviart, N. Seguin, *Godunov-type schemes for hyperbolic systems with parameter dependent source. the case of euler system with friction*, Mathematical Models and Methods in Applied Sciences (M3AS), volume 20(11), pages 2109-2166, 2010.

[Chan60] S. Chandrasekhar *Radiative transfer*, Dover Publications Inc. 1960.

[CLS04] A. Chinnayya, A-Y. Leroux, N. Seguin, *A well-balanced numerical scheme for the approximation of the shallow-water equations with topography : the resonance phenomenon* , Int. J. Finite Volumes, pages 1-33, 2004. 2.1.2, 3.3.1.

[CDLL11] A. Claisse, B. Després, E. Labourasse, F. Ledoux, *A new exceptional points method with application to cell-centered Lagrangian schemes and curved meshes*, JCP, volume 231, Issue 11, 1 June 2012, Pages 4324-4354.

[CVV00] Y. Coudière, J. P. Vila, and P. Villedieu, *Convergence d'un schéma volumes finis explicite en temps pour les systèmes hyperboliques linéaires symétriques en domaines bornés (convergence of a finite volume time-explicit scheme for symmetric linear hyperbolic systems on bounded domains)*, C.R. Acad. Sci., Paris, Sr. I, Math., 331 :95-100, 2000.

[GC010] G. Coppa, *Deduction of a symmetric tensor formulation of the P_n method for the linear transport equation*, Progress in Nuclear Energy 52 (2010) 474-752.

[CGRS08] F. Coquel, E. Godlewski, P.-A. Raviart, N. Seguin, *An asymptotic preserving scheme for Euler equations with gravity and friction*, Finite volumes for complex applications V, Aussois, France, 8-13 juin 2008, pages 305-312, ISTE, Londres.

[CL11] N. Crouseilles, M. Lemou, *An asymptotic preserving scheme based on a micro-macro decomposition for collisional Vlasov equations : diffusion and high-field scaling limits*, Kinetic and related models 4, 2 (2011) 441-477.

[DW94] R. Dautray, J. P. Watteau, *Fusion thermonucléaire inertielle par laser* Eyrolles, 1994.

[DDSV09] P. Degond, F. Deluzet, A. Sangam, M-H. Vignal, *An asymptotic preserving scheme for the Euler equations in a strong magnetic field*, J. Comput. Phys., 228 (2009), pp. 3540-3558.

[DGM07] P. Degond, S. Gallego, F. Méhats, *An asymptotic-preserving scheme for the Schrodinger equation in the semi-classical limit*, C. R. Acad. Sci. Paris, Ser I. 345 (2007), pp. 531-536.

[DLV08] P. Degond, J-G. Liu, M-H. Vignal, *Analysis of an asymptotic preserving scheme for the Euler-Poison system in the quasineutral limit*, SIAM J. Numer. Anal. 46 (2008), pp. 1298-1322.

[Dem06] J. P. Demailly *Analyse numérique et équations différentielles*, collection Grenobles sciences, EDP sciences.

[Des09] B. Després *Weak consistancy of the cell centered lagrangian GLACE scheme on general mesh in any dimension*, Comput. Methods Appl. Mech. Engrg. 199 (2010) 2669-2679.

[Des010] B. Després *Lois de Conservations Euleriennes, Lagrangiennes et Méthodes numériques*, Springer, 2010.

[DM03] B. Després, C. Mazeran *Symmetrization of Lagrangian gas dynamics and Lagrangian solvers*, Arch. Ration. Mech. Anal. 178 (2005) 327-372. Sci.

[DM05] B. Després, C. Mazeran *Lagrangian gas dynamics in 2D and Lagrangian systems*, C. R. Acad. Sci (Paris) 331 (2003) 475-480.

[DM04] D. Donatelli and P. Marcati *Convergence of singular limits for multi-D semilinear hyperbolic systems to parabolic systems.* Trans. Amer. Math. Soc. 356 (2004), no. 5, 2093–2121

[DEGH08] J. Droniou, R. Eymard, T. Gallouet, R. Herbin *A unified approach to mimetic finite difference, hydrid finite volume and mixed finite volume methods*, Math. Models Methods Appl. Sci. (M3AS), 20 (2010), no. 2, 265-295.

[DlP09] J. Droniou, C. Le-Potier *Construction and convergence study of local-maximum-principe preserving schemes for elliptic equations*, SIAM journal on numerical analysis, 2011, vol. 49, no1-2, pp. 459-490.

[DF99] B. Dubroca, J L.Feugeas *Hiérarchie des modèles aux moments pour le transfert radiatif*, C. R. Acad. Sci. Paris t.329, serie I, p.915-920, 1999.

[En07] C. Enaux *Analyse mathématique et numérique d'un modèle multifluide multivitesse pour l'interpénétration de fluides miscibles*, Thèse, Centrale Paris, 2007.

[EG06I] A. Ern and J.-L. Guermond, *Discontinuous Galerkin methods for Friedrichs' systems. I. General theory*, SIAM J. Numer. Anal., 44(2), 753–778 (2006).

[EG06II] A. Ern and J.-L. Guermond, *Discontinuous Galerkin methods for Friedrichs' systems. II. Second-order elliptic PDE's*, SIAM J. Numer. Anal., 44(6), 2363–2388 (2006).

[EGH00] R. Eymard, T. Gallouet, R. Herbin, *The finite volume method*, Handbook for Numerical Analysis, Ph. Ciarlet J.L. Lions eds, North Holland, 2000, 715-1022.

[EGH09] R. Eymard, T. Gallouet, R. Herbin, *Discretisation of heteregeneous and anisotropic diffusion problem on general noncorforming meshes*, IMA J. Num. Anal, 2009.

[FJ10] F. Filbet, S. Jin, *A class of asymptotic preserving schemes for kinetic equations and related problems with stiff sources*, JCP 229 (2010) 7625-7648.

[FHSN11] E. Franck, P. Hoch, G. Samba, P. Navarro, *An asymptotic preserving scheme for P_1 model using classical diffusion schemes on unstructured polygonal meshes*, ESAIM : Proceedings, Octobre 2011, Vol 32, p 56-75.

[Fr07] M. Frank, *Approximate models for radiative transfer*, Bull. Inst. Math. Acad. Sinica (New Series) 2 (2007) 409-432.

[FKLY07] M. Frank, A. Klar, E W. Larsen, S. Yasuda, *Time-dependent simplified P_n approximation to the equations of radiative transfer*, Journal of Computational Physics (206) 2289-2305, 2007.

[Gal02] G. Gallice, *Solveurs simples positifs et entropiques pour les systèmes hyperboliques avec term source*, C. R. Acad. Sci. Paris, Ser. I 334 (2002) 713-716.

[GHS03] T. Gallouet, J. M. Herard, N. Seguin *Some approximate Godunov schemes to compute shallow water equations with topography*, Comput. and Fluids, 32(4) :479-513, 2003. 3.3.1

[GR91] E. Godlewski, P. A Raviart, *Hyperbolic systems of conservation laws*, volume 3/4 de matéhmatiques et applications (Paris).

[Gos10] L. Gosse, *Asymptotic preserving and well-balanced scheme for the 1D Cattaneo model of chemotaxis movement in both hyperbolic and diffusive regimes*, Journal of Mathematical Analysis and Applications, Volume 388, Issue 2, 15 April 2012, Pages 964-983.

[Gos10] L. Gosse, *Transient radiative transfer in the grey case : well-balanced and asymptotic preserving schemes built on Case's elementary solutions*, Journal of Quantitative Spectroscopy and Radiative Transfer 112 (2011) 1995-2012.

[Gos10] L. Gosse, *Well-Balanced schemes using elementary solutions for linear models of the Boltzmann equation in one space dimension*,

[GT01] L. Gosse, G. Toscani, *An asymptotic-preserving well-balanced scheme for the hyperbolic heat equations*, C. R. Acad. Sci Paris,Ser. I 334 (2002) 337-342.

[GT04] L. Gosse, G. Toscani, *An asymptotic-preserving and well-balanced schemes for radiative transfer and rosseland approximation*, Numer. Math. (2004) 98 : 223-250

[GL96] 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.

[HLMc10] C. D. Hauck, R. B. Lowrie, R. G McClarren, *Methods for diffusive relaxation in P_n equations*, Numerical methods for relaxation systems and balance equations.

[HMc11] C. D. Hauck, R. G McClarren, *Positive P_n closures*, Siam, Journal on scientific computing, 2010

[Her00] F. Hermeline, *A finite volume method for the approximation of diffusions operators on distorted meshes*, Journal of computational Physics 160(2), 481-499 (2000)

[Her07] F. Hermeline, *Approximation of 2-D and 3-D diffusion operators with variable full tensor coefficiens on arbitrary meshes*, Computer Methods in Applied Mechanics and Engineering 196(21-24), 2497-2526 (2007).

[JL91] S. Jin, D. Levermore, *The discrete-ordinate method in diffusive regime*, Transport Theory and Statistical Physics, vol. 20, issue 5, pp. 413-439, 1991.

[JL96] S. Jin, D. Levermore, *Numerical schemes for hyperbolic conservation laws with stiff relaxation terms*, JCP 126,449-467, 1996.

[JX95] S. Jin, Z. Xin, *The relaxation schemes for systems of conservation laws in arbitrary space dimensions*, Comm. Pure Appl. Math 48 (1995) 235-277.

[Jin99] S. Jin, *Efficient Asymptotic-Preserving (AP) Schemes for Some Multiscale Kinetic Equations*, SIAM J. Sci. Comp. 21, 441-454, 1999

[JY11] S. Jin, B. Yan, *A class of asymmptotic-preserving schemes for the Fokker-Planck-Landau equation*, J. Comp. Phys. 230, 6420-6437, 2011.

[Jin11] S. Jin, *Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations : a review*, Lecture Notes for Summer School on « Methods and Models of Kinetic Theory » (M and MKT), Porto Ercole (Grosseto, Italy), Juin 2010.

[Ker81] D. Kershaw, *Differencing of the diffusion equation in lagrangian hydrodynamic codes*, JCP 39 375-395 (1981).

[Klu08] G. Kluth, *Analyse mathématique et numérique de systèmes hyperélastiques et introduction áÂ la plasticité.* Ph. D thesis, University Pierre and Marie Curie, 2008.

[KD10] G. Kluth, B. Després, *Discretization of hyperelasticity on unstructured meshes with a cell-centered Lagrangian scheme* J. Comp. Phys., 229, (2010), 9092-9118.

[KLS04] Y. Kuznetsov, K. Lipnikov, M. Shashkov, *Mimetic finite difference method on polygonal meshes for diffusion-type problems*, Comp. Geosciences V.8, No.4, pp. 301–324.

[LM07] M. Lemou, L.Mieussens, *A new symptotic preserving scheme based on micro-macro formulation for linear kinetic equations inthe diffusion limit*, SIAM J. Sci. COMPUT. Vol. 31, 1. 334-368

[Lemou10] M. Lemou, *Relaxed micro-macro schemes for kinetic equations*, C. R. Acad. Sci. Paris, Ser. I 348 (2010) 455-460.

[LePo09] C. Le Potier *Un schéma linéaire vérifiant le principe du maximum pour des opérateurs de diffusion très anisotropes sur des maillages déformés*, Comptes Rendus Mathématique, Volume 347, Issues 1-2, 2009, Pages 105-110.

[LePo09] C. Le Potier, A. Mahamane *Correction non linéaire et principe du maximum avec des schémas hybrides pour la discrétisation d'opérateurs de diffusion*, Comptes Rendus Mathématique, Volume 350, Issues 1-2, 2012, Pages 101-106.

[LMS04] K. Lipnikov, J. Morel, M. Shashkov, *Mimetic finite difference methods for diffusion equations on non-orthogonal non-conformal meshes* , Journal of Computational Physics 199 (2004) 589-597.

[LSSV07] K. Lipnikov, M. Shashkov, D. Svyatskiy, Y. Vassilevski, *Monotone finite volume schemes for diffusion euqations on unstructured triangular and shape-regular polygonal meshes*, JCP 227 (2007). 492-512.

[Maz07] C. Mazeran, *Sur la structure mathématique et l'approximation numérique de l'hydrodynamique Lagrangienne bidimensionelle*, PhD thesis, University of Bordeaux, 2007.

[MABO07] P-H. Maire, R. Abgrall, J. Breil, J. Ovadia, *A cell-centered lagragian scheme for two-dimensional compressible flow problems*, SIAM J. Sci. Comput. Vol 29,No. 4, pp. 1781-1824. 2007.

[Mai11] P-H. Maire, *Contribution to the numerical modeling of inertial confinement fusion*, Hdr, CEA and university of Bordeaux.

[McH10] R. McClarren, C. D. Hauck, *Simulating radiative transfer with flitered spherical harmonics*, Physics Letters A 374, 2290-2296, 2010.

[McHL10] R. McClarren, C. D. Hauck, R. B Lowrie, *Filtered spherical harmonics methods for transports problems* , Proceedings of the 2009 international conference on mathematics and computational methods and reactor physics, 2008.

[McL08] R. McClarren, R. B Lowrie, *Manufactured solutions for the P_1 radiation-hydrodynamics equation*, Journal of Quantitative Spectroscopy and Radiative Transfer 109 (2008) 2590-2602.

[McHL10] R. McClarren, R. B Lowrie, *The effects of slope limiting on asymptotic-preserving numerical methods for hyperbolic conservation laws*, Journal of Computational Physics 227 (2008) 9711-9726.

[MF11] P. Monreal, M. Franck, *Higher order minimum entropy approximations in radiative transfer*, Prépublication.

[NR12] R. Natalini, M. Ribot, *Asymptotic high order mass-preserving schemes for a hyperbolic models of chemotaxis.* SIAM J. Num. Anal (2012).

[OAH99] G.L. Olson, L.H Auer, M.L. Hall, *Diffusion, P_1, and other approximate forms of radiation transport*, JSQRT 64,1999.

[Pa08] S. Paul, *Etude d'un schéma numérique volumes finis avec bilan des flux aux coins pour l'ALE*, Rapport stage M2, UPMC 2008.

[Pl95] J. Planchard, *Méthode mathématique en neutronique*, Eyrolles, 1995.

[Reu03] P. Reuss, *Précis de neutronique* , EDP sciences 2003.

[Sah11] S. Sahmin, *Un schéma aux volumes finis avec matrice signe pour les systèmes non homogènes.* Thèse, Université Paris 13, 2005.

[SFL11] M. Schafer, M. Frank, C. D. Levermore, *Diffusive corrections to P_n approximations*, Multiscale Model. Simul. 9 1-28, 2011.

[SYY09] Z. Sheng, J. Yue, G. Yuan, *Monotone finite volume schemes of nonequilibrium radiation diffuison equations of distortes meshes* SIAM J. Sci. Comput. Vol. 31,No. 4,pp 2915-2934.

[Tor28] E. F. Toro *Riemann Solvers and Numerical Methods for Fluid Dynamics : A Practical Introduction*, Springer 2008.

[Tur03] R. Turpault, *Modélisation, approximation numérique et applications du transfert radiatif en déséquilibre spectral couplé avec l'hydrodynamique.* Thèse, Université Bordeaux, CEA CESTA, 2003.

[MV06] M. Vohralik, *Equivalence between lowest-order mixed finite element and multi-point finite volume methods on simplicial meshes.* ESAIM : M2AN Vol. 40, No 2, 2006, pp. 367-391.

Résumé

L'équation de transport, dans le régime fortement collisionnel admet une limite asymptotique de diffusion. Les discrétisations angulaires comme la méthode des ordonnées discrètes S_n où le développement tronqué en harmonique sphérique préservent aussi cette limite de diffusion. Par conséquent, il est intéressant de construire pour de tels systèmes des méthodes de volumes finis sur maillages non structurés qui préservent cette limite de diffusion pour des grilles grossières. En effet, ces modèles peuvent être couplés avec des codes hydrodynamiques Lagrangiens qui génèrent des maillages très tordus.
Pour commencer, on considère la discrétisation angulaire la plus simple de l'équation de transport appelée le modèle P_1. Après une rapide introduction sur les méthodes 1D, on commence par modifier le schéma acoustique en dimension deux avec la méthode de Jin-Levermore. Le schéma ainsi obtenu n'est pas convergent dans le régime de diffusion car le schéma de diffusion valide n'est pas consistant sur maillages non structurés. Pour résoudre ce problème, on a proposé de nouvelles méthodes valides sur maillages non structurés. Ces méthodes sont basées sur un autre formalisme des méthodes de volumes finis ou les flux sont localisés aux interfaces, couplé avec la méthode de Jin-Levermore. On obtient deux schémas convergents qui dérivent sur les schémas « asymptotic preserving » 1D. Le schéma limite de diffusion obtenu est un nouveau schéma pour lequel on a donné une preuve de convergence. Dans un second temps, on a proposé une extension du travail réalisé pour le modèle P_1 dans le cadre des discrétisations angulaires d'ordres élevés. Pour obtenir une discrétisation « asymptotic preserving » pour ces modèles on a utilisé une décomposition entre la discrétisation angulaire de premièr ordre et les discrétisations angulaires d'ordres supérieurs. Enfin on a étudié la discrétisation du problème d'absorption/émission présent en transfert radiatif ainsi que la discrétisation du modèle non linéaire M_1. L'approximation du modèle M_1 est basé sur un couplage entre un schéma Lagrange+projection pour une reformulation du modèle M_1 et la méthode de Jin-Levermore. La méthode numérique obtenue préserve la limite asymptotique, l'inégalité d'entropie et le principe du maximum associé au système sur maillages non structurés.

Mots-clés: méthode de volumes finis, schémas « asymptotic preserving » , méthode de Jin-Levermore, discrétisation aux noeuds, schéma GLACE, maillages non structurés, systèmes de Friedrichs, schéma de Gosse-Toscani.

Abstract

The transport equation in highly scattering regimes has a limit in which the dominant behavior is given by the solution of a diffusion equation. The angular discretizations like the discrete ordinate method S_n or the truncated spherical harmonic expansion P_n have the same property. For such systems it would be interesting to construct finite volume schemes on unstructured meshes which have the same dominant behavior even if the mesh is coarse (these schemes are called asymptotic preserving schemes). Indeed these models can be coupled with Lagrangian hydrodynamics codes which generate very distorted meshes. To begin we consider the lowest order angular discretization of the transport equation that is the P_1 model also called the hyperbolic heat equation. After an introduction of 1D methods, we start by modify the classical edge scheme with the Jin-Levermore procedure, this scheme is not valid in the diffusion regime because the limit diffusion scheme (Two Points Flux Approximation)

is not consistent on unstructured meshes. To solve this problem we propose news schemes valid on unstructured meshes. These methods are based on the nodal scheme (GLACE scheme) designed for the acoustic and dynamic gas problems, coupled with the Jin-Levermore procedure. We obtain two schemes valid on unstructured meshes which give in 1D on the Jin-Levermore scheme and the Gosse-Toscani scheme. The limit diffusion scheme obtained is a new nodal scheme. Convergence and stability proofs have been exhibited for these schemes. In a second time, these methods have been extended to higher order angular discretisation like the P_n and S_n models using a splitting strategy between the lowest order angular discretization and the higher order angular discretization. To finish we will propose to study the discretization of the absorption/emision problem in radiative transfer and a non-linear moment model called M_1 model. To treat the M_1 model we propose to use a formulation like a dynamic gas system coupled with a Lagrange+remap nodal scheme and the Jin-Levermore method. The numerical method obtained preserve the asymptotic limit, the maximum principle, and the entropy inequality on unstructured meshes.

Keywords: Finite volume method, asymptotic preserving schemes, Jin-Levermore method, nodal discretization, unstructured meshes, radiative transfer, Friedrichs systems, Gosse-Toscani scheme.