

Méthodes de volumes finis pour les fluides compressibles

Nicolas Seguin
Université Pierre et Marie Curie - Paris 6
Université Pierre et Marie Curie - Paris 6

[Version du 20 juin 2013 ¹]

Résumé. Ce cours est une introduction à l'analyse et à l'approximation des lois de conservation et des systèmes hyperboliques. Une attention particulière est portée sur les méthodes numériques de type volumes finis dans le cas scalaire et dans le cas des systèmes conservatifs.

1. Ceci est une version provisoire. N'hésitez pas à me contacter si vous trouvez des erreurs ou coquilles, ou si vous avez des suggestions d'améliorations : nicolas.seguin@upmc.fr.

Table des matières

1	Introduction	4
1.1	Quelques préliminaires	4
1.1.1	Classement des équations aux dérivées partielles	4
1.1.2	Équations hyperboliques linéaires et non linéaires	5
1.2	Exemples d'équations hyperboliques	5
1.2.1	Équation des ondes	6
1.2.2	Équation de transport	6
1.2.3	EDP quasilineaires	6
1.2.4	Dynamique des gaz	7
1.2.5	Modèle de Saint-Venant	7
1.2.6	Et encore d'autres modèles	8
1.3	Quelques outils d'analyse fonctionnelle	8
1.3.1	Espaces de Lebesgue	8
1.3.2	Distributions	9
1.3.3	Fonctions à variations bornées	10
1.3.4	Transformée de Fourier	11
1.4	Avertissement	11
2	Équations hyperboliques linéaires	12
2.1	Équation de transport	12
2.1.1	Cas unidimensionnel à coefficient constant	12
2.1.2	Cas multidimensionnel à coefficient variable	15
2.2	Systèmes hyperboliques linéaires	15
2.2.1	Cas unidimensionnel	16
2.2.2	Cas multidimensionnel	16
2.3	Conditions aux limites	18
2.3.1	Cas d'une vitesse positive	19
2.3.2	Cas d'une vitesse négative	19
3	Lois de conservation	20
3.1	Solutions régulières, ou pas	20
3.2	Solutions faibles	21
3.3	Problème de Riemann	22
3.4	Entropie et unicité	24
3.5	Problème de Cauchy	26
3.6	Le cas multidimensionnel	30

4	Méthodes des volumes finis pour les lois de conservation	32
4.1	Notations et principes des méthodes de volumes finis	32
4.2	Schémas monotones	34
4.2.1	Estimations <i>a priori</i>	34
4.2.2	Convergence	36
4.2.3	Estimations <i>a posteriori</i>	39
4.3	Cas multidimensionnel	46
4.3.1	Maillage cartésien	46
4.3.2	Maillage non structuré	49
5	Systèmes de lois de conservation	52
5.1	Hyperbolicité et entropie	52
5.2	Solutions faibles	54
5.3	Ondes et caractère non linéaire	55
5.4	Ondes de choc et entropie	57
5.5	Problème de Riemann	57
5.6	Problème de Cauchy	59
5.7	Le problème de Riemann pour Euler barotrope	60
5.7.1	Étude des ondes	60
5.7.2	Résolution du problème de Riemann	63
6	Méthodes de volumes finis pour les systèmes de lois de conservation	64
6.1	Schémas volumes finis et propriétés de base	64
6.1.1	Consistance	65
6.1.2	Préservation de domaine invariant	66
6.1.3	Inégalités d'entropie discrètes	68
6.2	Formalisme de Harten, Lax et Van Leer	68
6.2.1	Cadre général	68
6.2.2	Schéma de Godunov	70
6.2.3	Solveurs simples	71
6.3	Schémas de Godunov approchés	72
6.3.1	Schéma choc-choc	73
6.3.2	Schéma VFRoe	73
6.4	Cas multidimensionnel	73

Chapitre 1

Introduction

Ce polycopié correspond à une compilation de divers résultats concernant l'analyse et l'approximation des lois de conservation et des systèmes hyperboliques.

1.1 Quelques préliminaires

Avant de commencer, posons le contexte. L'idée est d'étudier les solutions de certaines équations aux dérivées partielles (EDP) ; par étudier on entend déterminer l'existence et l'unicité de ces solutions pour des problèmes donnés, comprendre leur structure et leurs propriétés, les approcher par des méthodes numériques.

Soit Ω un ouvert de \mathbb{R}^D , $D > 0$. On considère une fonction

$$\begin{aligned} u : \Omega &\longrightarrow \mathbb{R}^N \\ X &\longmapsto u(X) \end{aligned}$$

où $N > 0$. On appelle *équation aux dérivées partielles* une relation du type

$$F\left(X, u, (\partial_i u_k)_{\substack{1 \leq i \leq D \\ 1 \leq k \leq N}}, (\partial_{ij}^2 u_k)_{\substack{1 \leq i, j \leq D \\ 1 \leq k \leq N}}\right) = f, \quad (1.1)$$

où $F : \mathbb{R}^N \times \mathbb{R}^{ND} \times \mathbb{R}^{ND^2} \rightarrow \mathbb{R}^N$ et $f : \Omega \rightarrow \mathbb{R}^N$ sont deux fonctions données.

Un peu de vocabulaire :

- La fonction f est généralement appelée terme source.
- On dit que l'EDP (1.1) est d'ordre 2.
- Si $D = 1$, on retombe sur le cas des équations différentielles.
- Si $f \equiv 0$, l'EDP est dite homogène.
- L'EDP est dite linéaire si F l'est à X fixé, non linéaire sinon.
- L'EDP est une équation scalaire si $N = 1$ et un système si $N > 1$.
- On dit que l'EDP est une équation d'évolution si une des composantes de X est le temps. Sinon, c'est une équation stationnaire.

Remarque 1. En général, pour étudier (1.1), il est nécessaire d'ajouter des relations complémentaires sur u , bien souvent pour $X \in \partial\Omega$.

1.1.1 Classement des équations aux dérivées partielles

Le classement des EDP est loin d'être exhaustif, valable dans le cas linéaire et pas forcément toujours judicieux. Néanmoins, il permet d'avoir une idée du comportement

des solutions de ces EDP et de leur version “étendue”.

Plaçons-nous dans le cas de l’étude d’une EDP linéaire d’ordre 2, avec $D = 2$ et $N = 1$:

$$A_0 u + \sum_{i=1,2} A_i \partial_i u + \sum_{i,j=1,2} A_{ij} \partial_{ij}^2 u = f, \quad (1.2)$$

où $A_i \in \mathbb{R}$, $i = 0, 1, 2$, et la matrice $A = (A_{ij})_{1 \leq i, j \leq 2}$ est symétrique non nulle. Puisque A est symétrique, elle est diagonalisable dans \mathbb{R} .

La classification des EDP se fait alors suivant le spectre de la matrice A (les termes d’ordre inférieur ou égal à un n’ont pas d’influence) :

- Si $\det(A) > 0$, l’équation est dite *elliptique*. Un exemple est l’équation de Poisson

$$\partial_{xx}^2 u + \partial_{yy}^2 u = f$$

(si $f \equiv 0$, c’est l’équation de Laplace).

- Si $\det(A) = 0$, l’équation est dite *parabolique*. Un exemple classique est l’équation de la chaleur

$$\partial_t u - \partial_{xx}^2 u = f.$$

- Si $\det(A) < 0$, l’équation est dite *hyperbolique*. Un exemple est l’équation des ondes

$$\partial_t^2 u - \partial_{xx}^2 u = f.$$

La terminologie correspond à celle utilisée dans le classement des coniques, mais l’analogie s’arrête là.

Le comportement des solutions est très différent suivant le type d’EDP que l’on étudie. De plus, les données à ajouter pour définir un problème bien posé dépendent aussi du type de l’EDP.

1.1.2 Équations hyperboliques linéaires et non linéaires

On a mentionné le fait que la classification précédente n’était pas exhaustive. On va maintenant préciser ce qu’on appelle une équation hyperbolique. En général, elle sont formulées à l’ordre un, c’est-à-dire dans le cas linéaire

$$\partial_t u + \sum_{i=1}^d A_i \partial_{x_i} u = 0, \quad t > 0, x \in \mathbb{R}^d,$$

où $A_i \in \mathbb{R}^{N \times N}$, et dans le cas non linéaire

$$\partial_t u + \sum_{i=1}^d \partial_{x_i} f_i(u) = 0, \quad t > 0, x \in \mathbb{R}^d,$$

où $f_i : \mathbb{R}^N \rightarrow \mathbb{R}^N$.

1.2 Exemples d’équations hyperboliques

On présente ici quelques exemples d’équations hyperboliques.

1.2.1 Équation des ondes

On l'a vu précédemment, un exemple typique d'EDP hyperbolique est l'équation des ondes, qui s'écrit dans le cas homogène

$$\partial_t^2 u - a^2 \partial_x^2 u = 0.$$

Si on note $v = \partial_t u$ et $w = \partial_x u$, on obtient d'une part

$$\partial_t v - a^2 \partial_x w = 0$$

et l'équation

$$\partial_t w - \partial_x v = 0,$$

correspond au fait que $\partial_{tx}^2 u = \partial_{xt}^2 u$. Maintenant, si on définit $\varphi = v - aw$ et $\psi = v + aw$, on obtient le système équivalent

$$\partial_t \varphi + a \partial_x \varphi = 0,$$

$$\partial_t \psi - a \partial_x \psi = 0,$$

qui est constitué de deux équations découplées.

1.2.2 Équation de transport

On s'intéresse au phénomène de transport à vitesse a d'une densité (de particules, de polluants...) $\rho_0(x)$. Ainsi, si on note $\rho(t, x)$ la densité transportée au temps t et à la position x , on a la formule de translation

$$\rho(t, x) = \rho_0(x - at). \quad (1.3)$$

On peut donc remarquer que

$$\partial_t \rho(t, x) = -a \rho'_0(x - at),$$

$$\partial_x \rho(t, x) = \rho'_0(x - at),$$

dont on peut déduire l'équation

$$\partial_t \rho(t, x) + a \partial_x \rho(t, x) = 0, \quad (1.4)$$

avec pour condition initiale

$$\rho(0, x) = \rho_0(x). \quad (1.5)$$

Ainsi on peut espérer l'équivalence entre la formule (1.3) et le problème (1.4)-(1.5), ce qui permettrait de résoudre aussi l'équation des ondes.

1.2.3 EDP quasilineaires

Les EDP quasilineaires sont de la forme

$$\partial_t u(t, x) + \sum_{i=1}^d \partial_{x_i} f_i(t, x, u(t, x)) = g(t, x, u(t, x))$$

où u est à valeur dans \mathbb{R} (c'est le cas scalaire). Elles correspondent à l'écriture générale des EDP hyperboliques scalaires. Plus particulièrement, on étudie la forme

$$\partial_t u(t, x) + \sum_{i=1}^d \partial_{x_i} f_i(u(t, x)) = 0 \quad (1.6)$$

qui est appelée loi de conservation. En effet, on peut remarquer que si il existe une solution u intégrable et à support compact de cette équation, alors elle vérifie

$$\int_{\mathbb{R}} u(t, x) dx = \int_{\mathbb{R}} u(0, x) dx, \quad \forall t > 0.$$

On a donc conservation de la masse

$$m(t) = \int_{\mathbb{R}} u(t, x) dx.$$

On précisera plus tard la notion de conservation et on verra qu'elle est au cœur de la construction des méthodes de volumes finis. Enfin, on peut remarquer que les fonctions constantes sont des solutions de (1.6).

1.2.4 Dynamique des gaz

On considère un fluide compressible dont la viscosité peut être négligée (air autour d'un avion par exemple). Si on note ρ la masse volumique, u la vitesse, E l'énergie totale et p la pression, les lois de conservation de la masse, de la quantité de mouvement et de l'énergie totale s'écrivent ainsi :

$$\begin{aligned} \partial_t \rho + \partial_x \rho u &= 0, \\ \partial_t \rho u + \partial_x (\rho u^2 + p) &= 0, \\ \partial_t \rho E + \partial_x u (\rho E + p) &= 0, \end{aligned}$$

avec $E = u^2/2 + \varepsilon$, ε étant l'énergie spécifique du fluide. On peut remarquer que ces équations ne sont pas fermées (5 inconnues et 4 équations). Il faut ajouter une relation, c'est l'équation d'état (ou fermeture thermodynamique), elle permet de relier les quantités thermodynamiques en elles :

$$p = \mathcal{P}(\rho, \varepsilon).$$

On nomme ce système les *équations d'Euler* (qui sont en fait les équations de Navier-Stokes sans viscosité dans le cas compressible). Dans le cas d'un gaz parfait, on a la relation $\mathcal{P}(\rho, \varepsilon) = (\gamma - 1)\rho\varepsilon$, où $\gamma > 1$ est une constante (c'est le coefficient adiabatique). On obtient alors un système hyperbolique (on donnera plus tard une définition précise de l'hyperbolicité des systèmes d'EDP d'ordre 1).

1.2.5 Modèle de Saint-Venant

Le modèle de Saint-Venant modélise l'évolution de l'eau (comprise comme un liquide incompressible non visqueux) sur un fond plat. En notant h la hauteur de l'eau, u la vitesse moyenne horizontale et g la constante de la gravité, ce système s'écrit ainsi :

$$\begin{aligned} \partial_t h + \operatorname{div}_x hu &= 0, \\ \partial_t hu + \operatorname{div}_x (hu \otimes u) + \nabla_x (gh^2/2) &= 0, \end{aligned} \quad t > 0, x \in \mathbb{R}^2.$$

Si le fond de l'eau n'est pas plat et que $a(x)$ désigne l'altitude du fond (à une constante additive près), ce système devient :

$$\begin{aligned}\partial_t h + \operatorname{div}_x hu &= 0, \\ \partial_t hu + \operatorname{div}_x(hu \otimes u) + \nabla_x(gh^2/2) &= -gh\nabla_x a.\end{aligned}$$

On peut vérifier dans le cas de solutions régulières que $u \equiv 0$ et $h + a = \text{Cte}$ est une solution (stationnaire) de ce système. [EX]

1.2.6 Et encore d'autres modèles

Il existe de nombreux autres modèles hyperboliques, intervenant dans des domaines très variés : électromagnétisme, écoulements diphasiques, trafic routier, finances... Il peuvent prendre des formes très diverses et l'étude proposée dans ce cours ne prétend pas s'appliquer à tous ces modèles. Néanmoins, elle permettra de comprendre les difficultés associées à ce type de problèmes et on essaiera de donner quelques pistes pour aborder les cas n'entrant pas de le cadre classique de ce cours.

1.3 Quelques outils d'analyse fonctionnelle

Avant de poursuivre pour aborder l'analyse et l'approximation des EDP hyperboliques, nous allons faire quelques rappels d'analyse fonctionnelle et faire un panel des outils qui nous seront nécessaires plus tard. Les définitions qui suivent risquent parfois d'être imprécises pour alléger les notations, bien que sans ambiguïté (il suffira de vous référer aux ouvrages classiques d'analyse fonctionnelle pour obtenir des formulations plus détaillées).

1.3.1 Espaces de Lebesgue

Soit E un espace muni de la mesure de Lebesgue et f une fonction définie de E dans \mathbb{R} mesurable.

On définit la norme pour $1 \leq p < +\infty$ suivante :

$$\|f\|_p = \left(\int_E |f|^p dx \right)^{\frac{1}{p}}$$

et on dit que $f \in \mathbf{L}^p(E)$ si $\|f\|_p < +\infty$ (on travaille en fait modulo l'égalité presque partout).

On définit la norme suivante :

$$\|f\|_\infty = \inf\{C \in \mathbb{R}^+; |f| < C \text{ p.p.}\}$$

et on dit que $f \in \mathbf{L}^\infty(E)$ si $\|f\|_\infty < +\infty$ (on travaille là aussi modulo l'égalité presque partout).

Tous ces espaces sont des espaces de Banach, c'est-à-dire des espaces vectoriels normés complets. Rappelons aussi le résultat classique suivant :

Théorème 1.1 (Convergence dominée). *Soit $1 \leq p < +\infty$ et une suite $(f_n)_{n \in \mathbb{N}} \subset \mathbf{L}^p(E)$ telle que*

$$- f_n \rightarrow f \text{ p.p.,}$$

– il existe $F \in \mathbf{L}^p(E)$ tel que $|f_n| \leq F$ p.p. pour tout $n \in \mathbb{N}$,
alors f_n converge vers f dans $\mathbf{L}^p(E)$ quand $n \rightarrow \infty$.

Ce résultat n'est pas vérifié pour \mathbf{L}^∞ (prendre la suite définie par la fonction caractéristique de l'intervalle $[0, 1/n]$). On pourra néanmoins utiliser ce résultat car on travaillera avec des fonctions de \mathbf{L}^∞ mais *au sens des distributions*, donc localement dans \mathbf{L}^1 .

1.3.2 Distributions

Soit Ω un ouvert de \mathbb{R}^n . On définit tout d'abord l'espace des fonctions régulières à support compact :

$$\mathcal{C}_c^\infty(\Omega) = \{\varphi \in \mathcal{C}^\infty(\Omega) \text{ t.q. } \text{supp } \varphi \text{ compact}\}, \quad (1.7)$$

où

$$\text{supp } \varphi = \overline{\{x \in \Omega \text{ t.q. } \varphi(x) = 0\}}.$$

Une distribution est une forme linéaire continue sur $\mathcal{C}_c^\infty(\Omega)$. On notera l'ensemble des distributions $\mathcal{D}'(\Omega)$ (car c'est le dual topologique de $\mathcal{D}(\Omega) \equiv \mathcal{C}_c^\infty(\Omega)$).

Si $T \in \mathcal{D}'(\Omega)$, alors on note, pour tout $\varphi \in \mathcal{C}_c^\infty(\Omega)$, $T(\varphi) = \langle T, \varphi \rangle$. On peut alors définir une distribution par l'ensemble des valeurs de $\langle T, \varphi \rangle$, pour tout $\varphi \in \mathcal{C}_c^\infty(\Omega)$.

Exemple 1. - Si $f \in \mathbf{L}_{\text{loc}}^1(\Omega)$, c'est-à-dire que f est localement intégrable, alors on peut définir la distribution associée $T_f(\varphi)$ par

$$\langle T_f, \varphi \rangle = \int_{\Omega} f(x) \varphi(x) dx, \quad \forall \varphi \in \mathcal{C}_c^\infty(\Omega).$$

On peut alors identifier directement f et sa distribution.

- La masse (ou distribution, ou mesure) de Dirac δ_0 est l'application qui à $\varphi \in \mathcal{C}_c^\infty(\Omega)$ associe $\varphi(0)$, donc $\langle \delta_0, \varphi \rangle = \varphi(0)$.

- À la fonction d'Heavyside H (qui est la fonction caractéristique de \mathbb{R}^+), on associe

$$\langle H, \varphi \rangle = \int_{\mathbb{R}^+} \varphi(x) dx, \quad \forall \varphi \in \mathcal{C}_c^\infty(\Omega).$$

Si Ω est un ouvert de \mathbb{R} , on définit la dérivée $T' \in \mathcal{D}'(\Omega)$ d'une distribution $T \in \mathcal{C}_c^\infty(\Omega)$ par

$$\langle T', \varphi \rangle = -\langle T, \varphi' \rangle, \quad \forall \varphi \in \mathcal{C}_c^\infty(\Omega). \quad (1.8)$$

Si la distribution est associée à une fonction dérivable et localement intégrable, alors la formule (1.8) est directement donnée par intégration par parties, les termes de bords disparaissant puisque φ est à support compact dans Ω .

Dans le cas $\Omega \subset \mathbb{R}^n$, on en déduit la formule suivante

$$\langle \text{div} T, \varphi \rangle = -\langle T, \nabla \varphi \rangle, \quad \forall \varphi \in \mathcal{C}_c^\infty(\Omega). \quad (1.9)$$

On peut noter qu'une distribution est infiniment dérivable. Rappelons pour mémoire et comparaison la formule de Green (ou son application) :

Proposition 1.2. *Soit un ouvert régulier $\Omega \subset \mathbb{R}^n$. Alors pour tout $u \in \mathcal{C}^1(\bar{\Omega}; \mathbb{R}^n)$ et $v \in \mathcal{C}^1(\bar{\Omega}; \mathbb{R})$:*

$$\int_{\Omega} u \cdot \nabla v dx = \int_{\partial \Omega} v u \cdot n d\gamma - \int_{\Omega} \text{div} u v dx$$

Exemple 2. - La dérivée de la masse de Dirac est donnée par

$$\langle \delta'_0, \varphi \rangle = -\langle \delta, \varphi' \rangle = -\varphi'(0).$$

- La dérivée de la fonction d'Heavyside est alors

$$\langle H', \varphi \rangle = -\langle H, \varphi' \rangle = -\int_{\mathbb{R}^+} \varphi'(x) dx = \varphi(0) = \langle \delta_0, \varphi \rangle.$$

La dérivée au sens des distributions de la fonction d'Heavyside est donc la masse de Dirac.

1.3.3 Fonctions à variations bornées

On va vouloir travailler (dans le cas non linéaire) dans \mathbf{L}^∞ pour pouvoir traiter le cas de fonctions constantes, qui sont des solutions naturelles de (1.6). Néanmoins, on n'a que très peu d'information sur les suites bornées dans \mathbf{L}^∞ . On va donc utiliser en plus les fonctions à variation bornée.

On définit la variation totale d'une fonction $f \in \mathbf{L}^1_{\text{loc}}(\Omega)$, où $\Omega \subset \mathbb{R}^n$, par

$$TV(f) = \sup \left\{ \int_{\mathbb{R}} f(x) \operatorname{div} \varphi(x) dx, \varphi \in \mathcal{C}_c^\infty(\Omega), \|\varphi\|_\infty \leq 1 \right\}. \quad (1.10)$$

Ainsi, on peut définir l'espace des fonctions à variation bornée :

$$\mathbf{BV}(\Omega) = \{f \in \mathbf{L}^1_{\text{loc}}(\Omega), TV(f) < +\infty\} \quad (1.11)$$

et on définit alors la semi-norme associée $|f|_{\mathbf{BV}} = TV(f)$.

Remarque 2. Si $\Omega \subset \mathbb{R}$, on a la définition alternative suivante :

$$TV(f) = \sup \left\{ \sum_{i=1}^N |f(x_i) - f(x_{i-1})|, -\infty < x_1 < x_2 < \dots < x_N < +\infty, N \in \mathbb{N} \right\}.$$

Ainsi, si on considère une suite strictement croissante réelle $(x_{i+1/2})_{i \in \mathbb{Z}}$ et une suite réelle $(f_i)_{i \in \mathbb{Z}}$ et que l'on définit la fonction étagée f par

$$f(x) = \sum_{i \in \mathbb{Z}} f_i \mathbf{1}_{(x_{i-1/2}, x_{i+1/2})}(x),$$

alors on a

$$|f|_{\mathbf{BV}} = \sum_{i \in \mathbb{Z}} |f_{i+1} - f_i|.$$

Enfin, on peut montrer que si $\Omega \subset \mathbb{R}$, alors $\mathbf{BV} \subset \mathbf{L}^\infty$ (mais on utilise souvent la notation $\mathbf{L}^\infty \cap \mathbf{BV}$ pour préciser que la norme utilisée est $\|f\| = \|f\|_\infty + |f|_{\mathbf{BV}}$).

Enfin, citons le théorème de Helly :

Théorème 1.3. *Soit Ω un ouvert de \mathbb{R}^n . Soit une suite $(f_n)_n$ de fonction $\mathbf{L}^1_{\text{loc}}(\Omega)$ telle qu'il existe deux constantes C_1 et C_2 vérifiant pour tout n*

$$\|f_n\|_\infty \leq C_1 \quad \text{et} \quad |f_n|_{\mathbf{BV}(\Omega)} \leq C_2.$$

Alors, il existe une sous-suite de $(f_n)_n$ qui converge vers une fonction $f \in \mathbf{L}^1_{\text{loc}}(\Omega)$ dans $\mathbf{L}^1_{\text{loc}}(\Omega)$. La limite f vérifie en outre

$$\|f\|_\infty \leq C_1 \quad \text{et} \quad |f|_{\mathbf{BV}(\Omega)} \leq C_2.$$

1.3.4 Transformée de Fourier

On définit la transformée de Fourier d'une fonction $u \in \mathbf{L}^1(\mathbb{R}^d) \cap \mathbf{L}^2(\mathbb{R}^d)$ par

$$\mathcal{F}u(\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{-ix \cdot \xi} u(x) dx, \quad \xi \in \mathbb{R}^d.$$

On rappelle la formule de dérivation

$$i\xi_j \mathcal{F}u(\xi) = \mathcal{F}(\partial_{x_j} u)(\xi), \quad \xi \in \mathbb{R}^d,$$

et le théorème de Plancherel :

$$\|u\|_2 = \|\mathcal{F}u\|_2.$$

1.4 Avertissement

La théorie des EDP hyperboliques est loin d'être totalement développée à l'heure actuelle. Dans le cas des équations scalaires, les résultats d'existence, d'unicité et de convergence des méthodes de volumes finis sont acquis depuis les années 70 et 80, même en dimension d'espace supérieure à 1. Concernant les systèmes, les résultats sont très partiels. À part pour des systèmes ayant une structure particulière, les résultats d'existence et d'unicité ne sont valables que pour des données (et des solutions) vivant dans un voisinage d'un état constant, en une dimension d'espace. La convergence des méthodes de volumes finis reste à l'heure actuelle un problème ouvert... ce qui n'empêche pas d'en proposer et de les étudier !

Chapitre 2

Équations hyperboliques linéaires

On étudie dans ce chapitre le cas linéaire. Il correspond au cas le plus simple des équations hyperboliques et permet souvent d'accéder à des formules de représentation des solutions explicites (comme dans le cas de l'équation de transport). La compréhension de ces équations est très avancée, mais elles interviennent souvent de manière couplée à d'autres phénomènes, ce qui rend cette fois leur analyse et leur approximation bien plus difficile.

2.1 Équation de transport

On considère l'équation de transport scalaire suivante :

$$\begin{cases} \partial_t u + a(t, x) \cdot \nabla u = 0, & t > 0, x \in \mathbb{R}^d, \\ u(0, x) = u_0(x), & x \in \mathbb{R}^d, \end{cases} \quad (2.1)$$

où a est une fonction définie de $\mathbb{R}^+ \times \mathbb{R}^d$ dans \mathbb{R}^d et u_0 est la donnée initiale, allant de \mathbb{R}^d dans \mathbb{R} . On cherche donc une fonction u définie de $\mathbb{R}^+ \times \mathbb{R}^d$ dans \mathbb{R} .

Comme dans l'exemple de l'introduction, nous allons définir dans tous les cas les trajectoires suivies par les particules formant la densité u , comme dans (1.3). Ainsi, on va être amené à introduire la notion de *courbes caractéristiques*, c'est-à-dire de courbes dans le plan (x, t) le long desquelles la solution u de (2.1) reste constante.

2.1.1 Cas unidimensionnel à coefficient constant

Supposons que $d = 1$ et que a soit une constante réelle :

$$\begin{cases} \partial_t u + a \partial_x u = 0, & t > 0, x \in \mathbb{R}, \\ u(0, x) = u_0(x), & x \in \mathbb{R}. \end{cases} \quad (2.2)$$

Définissons les courbes caractéristiques associée à ce problème :

Définition 2.1. Les (courbes) caractéristiques de l'équation (2.2) sont les solutions de l'équation différentielle

$$\begin{cases} X'(t; X_0) = a, & t > 0, \\ X(0; X_0) = X_0, \end{cases} \quad (2.3)$$

où $X_0 \in \mathbb{R}$. (La dérivée est en fonction de la première variable.)

On a immédiatement que $X(t; x) = x + at$.

Ces courbes nous permettent de décrire avec précision les solutions de (2.2) :

Théorème 2.2. Supposons que $u_0 \in \mathcal{C}^1(\mathbb{R})$ dans (2.2). Alors, il existe une unique solution $u \in \mathcal{C}^1(\mathbb{R}^+ \times \mathbb{R})$ de (2.2). Elle vérifie

$$\forall t \geq 0, \forall x \in \mathbb{R}, \quad u(t, X(t; x)) = u_0(X(0; x)), \quad (2.4)$$

où X est solution de (2.3). Autrement dit

$$u(t, x) = u_0(x - at).$$

On voit donc que la solution est constante le long d'une caractéristique donnée.

Démonstration. Puisqu'on est dans le cas régulier, on a

$$d_t u(t, X(t; x)) = \partial_t u + X'(t; x) \partial_x u.$$

Si $X'(t; x) = a$, on obtient l'équivalence

$$d_t u(t, X(t; x)) = 0 \iff \partial_t u + X'(t; x) \partial_x u = 0,$$

dont on déduit (2.4). L'unicité de la solution se déduit immédiatement puisque l'on a équivalence. \square

La régularité de la solution est la même que la donnée initiale, puisque la première est la translatée de la dernière. On peut s'attendre à avoir le même comportement même si la donnée initiale est peu régulière. Néanmoins, l'équation (2.2) ne peut plus être comprise au sens classique. On fait alors appel à la théorie des distributions pour définir les dérivées partielles :

Définition 2.3. Soit $u_0 \in \mathbf{L}^\infty(\mathbb{R})$. Une fonction $u \in \mathbf{L}^\infty(\mathbb{R}^+ \times \mathbb{R})$ est une solution faible du problème (2.2) si elle vérifie pour tout $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^+ \times \mathbb{R})$

$$\int_{\mathbb{R}^+} \int_{\mathbb{R}} u(t, x) (\partial_t \varphi(t, x) + a \partial_x \varphi(t, x)) dx dt + \int_{\mathbb{R}} u_0(x) \varphi(0, x) dx = 0. \quad (2.5)$$

On a choisi de se placer dans le cadre \mathbf{L}^∞ , mais on aurait pu se placer dans des cadres plus faibles, comme $\mathbf{L}_{\text{loc}}^1$ par exemple, l'important étant que l'équation (2.5) ait toujours un sens.

Remarque 3. On a quelques indices de la validité de cette notion de solution :

- Si u_0 est régulière, alors la solution définie par (2.4) est bien une solution faible. Pour s'en convaincre, il suffit de reprendre (2.5) et par intégration par partie, on retombe bien sur (2.2). **[EX]**
- Si on suppose maintenant que $u_0(x) = H(x)$ où H est la fonction d'Heavyside, alors $u(t, x) = H(x - at)$ est une solution faible de (2.2). **[EX]**

On a même mieux que des indices, on a un résultat analogue au théorème 2.2 :

Théorème 2.4. *Supposons que $u_0 \in \mathbf{L}^\infty(\mathbb{R})$. Alors il existe une unique solution faible $u \in \mathbf{L}^\infty(\mathbb{R}^+ \times \mathbb{R})$ de (2.2). Elle vérifie*

$$\text{pour presque tout } t > 0, x \in \mathbb{R}, \quad u(t, X(t; x)) = u_0(X(0; x)), \quad (2.6)$$

Démonstration. Vérifions d'abord l'existence. Pour cela, nous montrons que la fonction définie par (2.6) est bien une solution faible de (2.2) :

$$\begin{aligned} & \int_{\mathbb{R}^+} \int_{\mathbb{R}} u(t, x) (\partial_t \varphi(t, x) + a \partial_x \varphi(t, x)) \, dx \, dt \\ &= \int_{\mathbb{R}^+} \int_{\mathbb{R}} u_0(x - at) (\partial_t \varphi(t, x) + a \partial_x \varphi(t, x)) \, dx \, dt \\ &= \int_{\mathbb{R}^+} \int_{\mathbb{R}} u_0(y) (\partial_t \varphi(t, y + at) + a \partial_x \varphi(t, y + at)) \, dy \, dt \end{aligned}$$

Soit $\psi(t, y) = \varphi(t, y + at)$. Comme $\partial_t \psi(t, y) = \partial_t \varphi(t, y + at) + a \partial_x \varphi(t, y + at)$, on obtient

$$\begin{aligned} & \int_{\mathbb{R}^+} \int_{\mathbb{R}} u(t, x) (\partial_t \varphi(t, x) + a \partial_x \varphi(t, x)) \, dx \, dt \\ &= \int_{\mathbb{R}^+} \int_{\mathbb{R}} u_0(y) \partial_t \psi(t, y) \, dy \, dt = \int_{\mathbb{R}} u_0(y) \int_{\mathbb{R}^+} \partial_t \psi(t, y) \, dt \, dy \\ &= \int_{\mathbb{R}} u_0(y) \psi(0, y) \, dy = \int_{\mathbb{R}} u_0(x) \varphi(0, x) \, dx. \end{aligned}$$

Montrons maintenant l'unicité de la solution. Pour cela, on considère deux solutions faibles, v et w , associées à la même donnée initiale u_0 . Si on note $u = v - w$, alors elle vérifie

$$\int_{\mathbb{R}^+} \int_{\mathbb{R}} u(t, x) (\partial_t \varphi(t, x) + a \partial_x \varphi(t, x)) \, dx \, dt = 0.$$

Montrons que cette équation entraîne que u est nulle presque partout. Pour cela, il suffit de montrer que toute fonction $\psi \in \mathcal{C}_0^\infty(\mathbb{R} + \times \mathbb{R})$, il existe $\varphi \in \mathcal{C}_0^\infty(\mathbb{R} + \times \mathbb{R})$ telle que

$$\psi = \partial_t \varphi + a \partial_x \varphi. \quad (2.7)$$

Si

$$\varphi(t, x) = - \int_t^{+\infty} \psi(s, x + a(s - t)) \, ds, \quad (2.8)$$

on peut montrer que (2.7) est bien vérifiée à l'aide du lemme suivant :

Lemme 2.5. *Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ et $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ deux fonctions régulières. Alors ([EX])*

$$f(t) = \int_T^t g(s, t) \, ds \implies f'(t) = g(t, t) + \int_T^t \partial_t g(s, t) \, ds.$$

Par ailleurs, on voit aisément que $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^+ \times \mathbb{R})$. De plus, l'ensemble des fonctions $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^+ \times \mathbb{R})$ pouvant s'écrire à l'aide de (2.8) est suffisant ([EX]) pour montrer que si

$$\int_{\mathbb{R}^+} \int_{\mathbb{R}} u(t, x) \psi(t, x) \, dx \, dt = 0,$$

alors $u = 0$ presque partout sur $\mathbb{R}^+ \times \mathbb{R}$. On a donc unicité de la solution faible de (2.2). \square

Donc dans le cas non régulier, on a aussi le phénomène de transport qui est vérifié. De plus, l'opérateur $S(t)$ qui à u_0 associe $u(t, \cdot)$ est un semi-groupe de $L^\infty(\mathbb{R})$, au sens où

- $S(0) = Id$,
- $S(t+s) = S(t) \circ S(s)$.

La notion de semi-groupe est très importante, la plupart des équations hyperboliques la vérifiant.

2.1.2 Cas multidimensionnel à coefficient variable

Revenons maintenant au cas (2.1). On peut à nouveau définir les caractéristiques :

Définition 2.6. Les (courbes) caractéristiques de l'équation (2.1) sont les solutions du système différentiel

$$\begin{cases} X'(t; X_0) = a(t, X(t; X_0)), & t > 0, \\ X(0; X_0) = X_0, \end{cases} \quad (2.9)$$

où $X_0 \in \mathbb{R}^d$. (La dérivée est en fonction de la première variable.)

Cette fois, ce système n'admet pas forcément une et une seule solution. Pour remédier à cela, on demande au champ de vitesse a de vérifier les hypothèses du théorème de Cauchy-Lipschitz et on a alors un résultat analogue au théorème 2.2 :

Théorème 2.7. *Supposons que $u_0 \in \mathcal{C}^1(\mathbb{R}^d)$ dans (2.1) et que $a \in \mathcal{C}^0(\mathbb{R}^+ \times \mathbb{R}^d)$ soit lipschitzienne par rapport à sa deuxième variable. Alors, il existe une unique solution $u \in \mathcal{C}^1(\mathbb{R}^+ \times \mathbb{R}^d)$ de (2.1). Elle vérifie*

$$\forall t \geq 0, \forall x \in \mathbb{R}^d, \quad u(t, X(t; x)) = u_0(X(0; x)), \quad (2.10)$$

où X est solution de (2.9).

La démonstration est identique à celle du théorème 2.2.

On voit que la notion de translation a disparu, mais on a cependant quelques propriétés :

Corollaire 2.8. *Sous les hypothèses du théorème 2.7, la solution vérifie les propriétés suivantes : [EX]*

- la solution au point (t, x) ne dépend que de $u_0(y)$ où $|x - y| \leq \|a\|_\infty t$ (propagation de l'information à vitesse finie),
- $\inf u_0 \leq u(t, x) \leq \sup u_0$ pour tout $t > 0$ et $x \in \mathbb{R}^d$ (principe du maximum),
- si $u_0(x) \leq v_0(x)$ pour tout $x \in \mathbb{R}^d$, alors $u(t, x) \leq v(t, x)$ pour tout $t > 0$ et $x \in \mathbb{R}^d$, u et v étant les solutions de (2.1) avec pour données initiales respectives u_0 et v_0 (principe de monotonie).

Concernant le cas des solutions faibles, la tâche s'avère maintenant un peu plus ardue (principalement pour trouver l'analogie de (2.8)), nous ne l'aborderons donc pas ici.

2.2 Systèmes hyperboliques linéaires

On s'intéresse maintenant au cas des systèmes hyperboliques linéaires, à coefficients constants. Un point clé est la notion d'hyperbolicité, permettant d'assurer la stabilité des solutions.

2.2.1 Cas unidimensionnel

On se concentre sur le problème de Cauchy suivant :

$$\begin{cases} \partial_t u + A \partial_x u = 0, & t > 0, x \in \mathbb{R}, \\ u(0, x) = u_0(x), & x \in \mathbb{R}, \end{cases} \quad (2.11)$$

où u est à valeurs dans \mathbb{R}^N et A une matrice carrée de dimension N à coefficients constants.

Pour étudier ce système, on va supposer que la matrice A est diagonalisable :

Définition 2.9. On dit que le système (2.11) est hyperbolique si A est diagonalisable dans \mathbb{R} .

On dit que le système (2.11) est strictement hyperbolique si A est diagonalisable dans \mathbb{R} et si ses valeurs propres sont distinctes.

Si le système est hyperbolique, alors il existe une matrice inversible P telle que $A = P\Lambda P^{-1}$ où Λ est la matrice des valeurs propres de A . Si on multiplie le système (2.11) par P^{-1} et si on définit $v = P^{-1}u$, on obtient

$$\partial_t v + \Lambda \partial_x v = 0,$$

qui est un système de N équations de transport découplées. On peut donc résoudre explicitement le système :

$$u(t, x) = Pv(t, x) = P \begin{pmatrix} v_0^{(1)}(x - \lambda_1 t) \\ \vdots \\ v_0^{(N)}(x - \lambda_N t) \end{pmatrix} \quad (2.12)$$

où $v_0 = P^{-1}u_0$. On a donc le résultat suivant :

Théorème 2.10. *Le problème de Cauchy (2.11) admet une et une seule solution, donnée par la formule explicite (2.12), qui devient sous forme condensée*

$$u(t, x) = \sum_{i=1}^N (l_i \cdot u_0(x - \lambda_i t)) r_i \quad (2.13)$$

où $(l_i)_{1 \leq i \leq N}$ et $(r_i)_{1 \leq i \leq N}$ sont les vecteurs propres (normalisés) à gauche et à droite de la matrice A .

Les calculs permettant d'obtenir (2.12) étant linéaires, le théorème est valable aussi bien pour les solutions régulières que pour les solutions faibles.

2.2.2 Cas multidimensionnel

On regarde maintenant le problème :

$$\begin{cases} \partial_t u + \sum_{i=1}^d A_i \partial_{x_i} u = 0, & t > 0, x \in \mathbb{R}^d, \\ u(0, x) = u_0(x), & x \in \mathbb{R}^d, \end{cases} \quad (2.14)$$

où u est toujours à valeurs dans \mathbb{R}^N et les matrices A_i sont carrées et de dimension N . En général, les matrices A_i ne sont pas diagonalisables (si elle le sont) dans la même base. On est alors contraint de se restreindre à une étude du système le long d'une direction donnée.

Soit $w \in \mathbb{S}^{d-1}$. On considère des solutions le long de la direction w , appelées ondes planes, définies par la forme suivante :

$$u(t, x) = v(\sigma(t, y)), \quad \text{où } y = x \cdot w,$$

v étant une fonction régulière définie de \mathbb{R} dans \mathbb{R}^d non constante (sinon, ça n'a pas vraiment d'intérêt !) et σ de $\mathbb{R}^+ \times \mathbb{R}$ dans \mathbb{R} . On en déduit, en supposant que σ est régulière, que

$$\partial_t \sigma + \sum_{i=1}^d w_i A_i \partial_y \sigma = 0,$$

qui est une équation de transport dans la direction w . On note maintenant $A(w) = \sum_{i=1}^d w_i A_i$ et on en déduit les définitions suivantes :

Définition 2.11. On dit que le système (2.14) est hyperbolique si $\forall w \in \mathbb{S}^{d-1}$, la matrice $A(w)$ est diagonalisable dans \mathbb{R} .

On dit que le système (2.14) est strictement hyperbolique si $\forall w \in \mathbb{S}^{d-1}$, la matrice $A(w)$ est diagonalisable dans \mathbb{R} et si ses valeurs propres sont distinctes.

On dit que le système (2.11) est symétrisable si il existe une matrice symétrique définie positive B telle que $\forall w \in \mathbb{S}^{d-1}$, la matrice $C(w) = BA(w)$ soit symétrique.

On peut remarquer qu'un système symétrisable est hyperbolique. En effet, notons R la matrice symétrique définie positive telle que $R^2 = B^{-1}$, ce qui donne

$$A(w) = B^{-1}C(w) = R^2C(w) = R(RC(w)R)R^{-1}.$$

Or, la matrice $RC(w)R$ est symétrique pour tout w puisque $C(w)$ est symétrique, donc diagonalisable dans \mathbb{R} :

$$A(w) = R(O(w)\Lambda(w)O(w)^T)R^{-1} = P(w)\Lambda(w)P(w)^{-1},$$

où $P(w) = RO(w)$. Donc $A(w)$ est bien diagonalisable pour tout w .

Remarque 4. Si $d = 1$, la réciproque est vraie : un système hyperbolique est symétrisable (il suffit de prendre $B = (P^{-1})^T P^{-1}$ où P est la matrices des vecteurs propres à droite de A). **[EX]**

On se place maintenant dans un cadre nous permettant d'utiliser la transformée de Fourier (disons $\mathbf{L}^1 \cap \mathbf{L}^2$, mais on pourrait aussi regarder le cas des distributions tempérées). On peut alors énoncer le résultat suivant :

Théorème 2.12. *Si le système (2.14) est hyperbolique, alors il existe une constante $C > 0$ telle que*

$$\|u(t, \cdot)\|_2 \leq C_0 \|u_0\|_2, \quad \forall t > 0. \quad (2.15)$$

Si le système (2.14) est symétrisable, alors

$$\|R^{-1}u(t, \cdot)\|_2 = \|R^{-1}u_0\|_2, \quad \forall t > 0. \quad (2.16)$$

Dans les deux cas, le problème de Cauchy (2.14) admet une et une seule solution.

Démonstration. Le système (2.14) devient après la transformation de Fourier :

$$\partial_t \mathcal{F}u(t, \xi) + i \sum_{j=1}^d \xi_j A_j \mathcal{F}u(t, \xi) = 0,$$

c'est-à-dire

$$\partial_t \mathcal{F}u(t, \xi) + iA(\xi) \mathcal{F}u(t, \xi) = 0.$$

Ainsi, si on muni ce système de la donnée initiale de (2.14), on obtient

$$\mathcal{F}u(t, \xi) = e^{-itA(\xi)} \mathcal{F}u_0(\xi).$$

Si (2.14) est hyperbolique, alors

$$\mathcal{F}u(t, \xi) = P(\xi) e^{-it\Lambda(\xi)} P^{-1}(\xi) \mathcal{F}u_0(\xi), \quad (2.17)$$

et

$$\|e^{-it\Lambda(\xi)}\|_2 \leq 1$$

puisque les valeurs propres de $A(\xi)$ sont réelles. On en déduit le caractère bien posé et l'inégalité (2.15), avec $C_0 = \sup_{\xi} \|P^{-1}(\xi)\| \|P(\xi)\|$ (C_0 est bien fini, mais on ne le montrera pas ici). Il reste à vérifier (2.16). Pour cela, on utilise

$$\partial_t R^{-1} \mathcal{F}u(t, \xi) = -iR^{-1}A(\xi) \mathcal{F}u(t, \xi) = -i(RC(\xi)R)R^{-1} \mathcal{F}u(t, \xi).$$

De plus, si $z(t) \in \mathbb{C}^N$, on a $d_t |z|^2 = 2 \Re \langle z | \partial_t z \rangle_{\mathbb{C}^N}$, donc

$$\begin{aligned} \partial_t |R^{-1} \mathcal{F}u(t, \xi)|^2 &= -2 \Re \langle R^{-1} \mathcal{F}u | i(RC(\xi)R)R^{-1} \mathcal{F}u \rangle_{\mathbb{C}^N} \\ &= -2 \Re \langle R^{-1} \mathcal{F}u | iO(\xi)\Lambda(\xi)O(\xi)^T R^{-1} \mathcal{F}u \rangle \\ &= -2 \Re \langle O(\xi)^T R^{-1} \mathcal{F}u | i\Lambda(\xi) O(\xi)^T R^{-1} \mathcal{F}u \rangle \\ &= 0. \end{aligned}$$

On en déduit donc (2.16).

Le résultat d'existence est obtenu à partir de la formule (2.17), qui permet de vérifier que

$$u(t, x) = \mathcal{F}^{-1} \left(P(\xi) e^{-it\Lambda(\xi)} P^{-1}(\xi) \mathcal{F}u_0(\xi) \right),$$

est une solution de (2.14) et le résultat d'unicité vient du fait que si $u_0 \equiv 0$, alors $u \equiv 0$, grâce à (2.15). \square

2.3 Conditions aux limites

On présente ici de manière succincte la notion de condition aux limites de type Dirichlet pour les équations de transport unidimensionnelles à coefficient constant.

Considérons le problème mixte suivant :

$$\begin{cases} \partial_t u + a \partial_x u = 0, & t > 0, x > 0, \\ u(0, x) = u_0(x), & x > 0, \\ u(t, 0) = u_1(t), & t > 0, \end{cases} \quad (2.18)$$

où u_0 et u_1 sont données. En fait, suivant le signe de la vitesse a , le problème est très différent.

2.3.1 Cas d'une vitesse positive

On suppose ici que $a \geq 0$. On sait par l'utilisation des courbes caractéristiques que la solution u vérifie

$$u(t, x) = u(t - \alpha, x - a\alpha).$$

On déduit donc la formule explicite suivante :

$$u(t, x) = \begin{cases} u_0(x - at) & \text{si } x - at > 0, \\ u_1(t - x/a) & \text{si } x - at < 0. \end{cases} \quad (2.19)$$

2.3.2 Cas d'une vitesse négative

On regarde maintenant le cas $a < 0$. Dans ce cas, on a toujours la formule

$$u(t, x) = u(t - \alpha, x - a\alpha)$$

obtenue à l'aide des courbes caractéristiques. Il y a néanmoins une ambiguïté : pour $t > 0$ et $x > 0$, en prenant successivement $\alpha = t$ et $\alpha = x/a$ on obtient simultanément

$$u(t, x) = u_0(x - at) \quad \text{et} \quad u(t, x) = u_1(t - x/a),$$

ce qui est a priori contradictoire ! Étant donné que l'on regarde des problèmes d'évolution, on privilégie la donnée initiale à la donnée au bord. La solution retenue est donc

$$u(t, x) = u_0(x - at). \quad (2.20)$$

On voit donc que la condition au bord n'est pas du tout prise en compte : la solution est totalement indépendante de u_1 .

Chapitre 3

Lois de conservation

Dans la partie précédente, nous avons distingué deux cas : solutions régulières et solutions faibles. On va voir que ce n'est plus pertinent dans le cas non linéaire, il faut se placer dans le cadre des solutions faibles quelle que soit la régularité de la donnée initiale. Malheureusement, ce cadre est trop large et ne garantit pas l'unicité, on doit faire appel à d'autres critères pour aboutir à des problèmes de Cauchy bien posés.

On considère la loi de conservation scalaire unidimensionnelle :

$$\begin{cases} \partial_t u + \partial_x f(u) = 0, & t > 0, x \in \mathbb{R}, \\ u(0, x) = u_0(x), & x \in \mathbb{R}, \end{cases} \quad (3.1)$$

où la fonction $f \in \mathcal{C}^2(\mathbb{R}; \mathbb{R})$ est généralement appelé flux. En effet, si on intègre cette équation sur un intervalle $[x_0, x_1]$, on obtient

$$d_t \int_{x_0}^{x_1} u(t, x) = f(u(t, x_0)) - f(u(t, x_1))$$

ce qui signifie que l'évolution de la masse totale dans l'intervalle $[x_0, x_1]$ est entièrement déterminée par les flux aux bords de cet intervalle. Autrement dit, on a conservation de la masse.

On se restreint volontairement au cas unidimensionnel, bien que l'analyse du problème de Cauchy est aussi valable dans le cas multidimensionnel.

3.1 Solutions régulières, ou pas

On peut définir les courbes caractéristiques associées à (3.1) comme les trajectoires des solutions de l'équation différentielle :

$$\begin{cases} X'(t; X_0) = f'(u(t, X(t; X_0))), & t > 0, \\ X(0; X_0) = X_0. \end{cases} \quad (3.2)$$

Ce problème est bien posé si les hypothèses du théorème de Cauchy-Lipschitz sont vérifiées. La difficulté provient du fait que les caractéristiques dépendent de u , dont on ne connaît pas *a priori* la régularité. On a cependant le résultat suivant :

Proposition 3.1. *Supposons que la donnée initiale est régulière. Alors, pour des temps suffisamment petits, Le problème (3.1) admet une et une seule solution régulière, définie par*

$$\forall x \in \mathbb{R}, \quad u(t, X(t; x)) = u_0(X(0; x)),$$

où X est solution de (3.2). Autrement dit

$$u(t, x + t f'(u_0(x))) = u_0(x).$$

Ce résultat n'est valable qu'en temps petit. En effet, il se peut que des caractéristiques viennent à se croiser, ce qui pose un problème de définition de la solution au point d'intersection :

Proposition 3.2. *Si il existe $x_0 < x_1$ telles que $f'(u_0(x_0)) > f'(u_0(x_1))$, alors la solution régulière définie dans la proposition précédente n'est pas globale. Plus précisément, le temps maximal d'existence de cette solution régulière est*

$$T_{\text{inf}} = \inf_{x_0 < x_1} \frac{x_1 - x_0}{(f'(u_0(x_0)) - f'(u_0(x_1)))^+}$$

où $(\cdot)^+ = \max(0, \cdot)$.

Pour s'en convaincre, un petit dessin faut mieux qu'un long discours...

Exemple 3. Considérons un modèle pour l'évolution d'une densité u de personnes courant vers la droite dans un couloir, en cas d'urgence. Le flux associé à cette équation s'écrit $f(u) = Au(U - u)$ où $U > 0$ est la densité maximale (donc u_0 est à valeurs dans $[0, U]$) et AU la vitesse maximale d'une personne qui court. Si la densité initiale u_0 est croissante par rapport à x , la population à gauche étant moins dense qu'à droite, les personnes la composant rattrapent celles situées devant elles et doivent ralentir : la dérivée par rapport à x de la densité augmente avec le temps t .

Exemple 4. L'équation de Burgers s'écrit

$$\partial_t u + \partial_x (u^2/2) = 0,$$

et peut se comprendre comme la description de l'évolution de la surface d'une vague. Plus la surface est élevée, plus elle va vite.

3.2 Solutions faibles

Pour pouvoir continuer d'étudier les solutions de (3.1) au-delà de T_{inf} , il est nécessaire de considérer des solutions faibles, même si la donnée initiale est régulière :

Définition 3.3. Soit $u_0 \in \mathbf{L}^\infty(\mathbb{R})$. On appelle solution faible de (3.1) une fonction $u \in \mathbf{L}^\infty(\mathbb{R}^+ \times \mathbb{R})$ vérifiant pour tout $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^+ \times \mathbb{R})$

$$\int_{\mathbb{R}^+} \int_{\mathbb{R}} (u(t, x) \partial_t \varphi(t, x) + f(u(t, x)) \partial_x \varphi(t, x)) dx dt + \int_{\mathbb{R}} u_0(x) \varphi(0, x) dx = 0. \quad (3.3)$$

Comme toujours, une solution classique de (3.1) est aussi solution faible (il suffit d'intégrer par parties (3.3)).

Étudions les conditions que doivent vérifier les discontinuités des solutions faibles :

Proposition 3.4. Soit la courbe $\Gamma = \{(t, x) \in \mathbb{R}^+ \times \mathbb{R}, x = \sigma(t)\}$ où $\sigma \in \mathcal{C}^1(\overline{\mathbb{R}^+})$, coupant un ouvert $\Omega \subset \mathbb{R}^+ \times \mathbb{R}$ et soient $\Omega_- = \{(t, x) \in \Omega, x < \Gamma(t)\}$ et $\Omega_+ = \{(t, x) \in \Omega, x > \Gamma(t)\}$. On considère une fonction $u \in \mathcal{C}^1(\overline{\Omega_-}) \cap \mathcal{C}^1(\overline{\Omega_+})$. Alors u est une solution faible de (3.1) sur Ω si et seulement si

$$-\sigma'(U^+ - U^-) + (f(U^+) - f(U^-)) = 0, \quad (3.4)$$

où $U^\pm(t) = u(t, \sigma(t)^\pm)$ sont les limites de u de part et d'autre de la courbe Γ et u est une solution classique de (5.1) sur $\Omega \setminus \Gamma$. On appelle l'équation (3.4) la relation de saut de Rankine-Hugoniot.

Démonstration. Prenons une fonction test φ dont le support est strictement inclus dans Ω . Alors, la formulation faible (3.3) devient

$$\begin{aligned} \int_{\Omega} (u \partial_t + f(u) \partial_x) \varphi dt dx &= \int_{\Omega_-} (u \partial_t + f(u) \partial_x) \varphi dt dx + \int_{\Omega_+} (u \partial_t + f(u) \partial_x) \varphi dt dx, \\ &= \int_{\Omega_-} \begin{pmatrix} u \\ f(u) \end{pmatrix} \cdot \nabla_{t,x} \varphi dt dx + \int_{\Omega_+} \begin{pmatrix} u \\ f(u) \end{pmatrix} \cdot \nabla_{t,x} \varphi dt dx, \\ &= \int_{\Omega \cap \Gamma} \begin{pmatrix} u \\ f(u) \end{pmatrix} \cdot n_- \varphi d\gamma + \int_{\Omega \cap \Gamma} \begin{pmatrix} u \\ f(u) \end{pmatrix} \cdot n_+ \varphi d\gamma \\ &\quad - \int_{\Omega \setminus \Gamma} (\partial_t u + \partial_x f(u)) \varphi dt dx. \end{aligned}$$

Or, les normales vérifient

$$n_-(s) = -n_+(s) = K_0 \begin{pmatrix} -\sigma'(s) \\ 1 \end{pmatrix},$$

où K_0 est un facteur de normalisation. On obtient alors

$$\begin{aligned} \int_{\Omega} (u \partial_t + f(u) \partial_x) \varphi dt dx &= K_0 \int_{\Omega \cap \Gamma} (-\sigma'(s)(U^-(s) - U^+(s)) \\ &\quad + (f(U^-(s)) - f(U^+(s)))) \varphi d\gamma \\ &\quad - \int_{\Omega \setminus \Gamma} (\partial_t u + \partial_x f(u)) \varphi dt dx, \end{aligned}$$

ce qui conclut la démonstration. \square

3.3 Problème de Riemann

Une propriété importante des EDP hyperboliques est le caractère auto-similaire des solutions :

Proposition 3.5. Soit u_0 une fonction telle que

$$\forall \lambda > 0, x \in \mathbb{R}, \quad u_0(\lambda x) = u_0(x).$$

Alors il existe une solution faible du problème de Cauchy (3.1) avec une telle donnée initiale telle que

$$\forall \lambda > 0, x \in \mathbb{R}, t > 0, \quad u(\lambda t, \lambda x) = u(t, x).$$

On peut alors définir une fonction $v \in \mathbf{L}^\infty(\mathbb{R}; \mathbb{R})$ telle que

$$\forall \lambda > 0, x \in \mathbb{R}, t > 0, \quad v(x/t) = u(t, x).$$

Démonstration. Celle-ci est immédiate, en utilisant la proposition précédente et le fait que les courbes de discontinuité sont des droites. \square

Les données initiales vérifiant la propriété d'auto-similarité sont de la forme

$$u_0(x) = \begin{cases} u_L, & \text{si } x < 0, \\ u_R, & \text{si } x > 0, \end{cases} \quad (3.5)$$

où u_L et u_R sont deux états constants et le problème de Cauchy (3.1)-(3.5) s'appelle problème de Riemann. On va en étudier les solutions faibles.

Proposition 3.6. *Le problème de Riemann (3.1)-(3.5) admet pour solution (onde de choc)*

$$u(t,x) = \begin{cases} u_L, & \text{si } x/t < s, \\ u_R, & \text{si } x/t > s, \end{cases} \quad (3.6)$$

où $s = (f(u_R) - f(u_L))/(u_R - u_L)$.

Si le flux f est strictement convexe (respectivement concave) et si $f'(u_L) < f'(u_R)$ (resp. $f'(u_L) > f'(u_R)$) alors, le problème de Riemann (3.1)-(3.5) admet aussi pour solution (onde de détente)

$$u(t,x) = \begin{cases} u_L, & \text{si } x/t < f'(u_L), \\ (f')^{-1}(x/t), & \text{si } f'(u_L) < x/t < f'(u_R), \\ u_R, & \text{si } x/t > f'(u_R). \end{cases} \quad (3.7)$$

Cette solution est continue dès que $t > 0$.

Démonstration. On vérifie aisément que la solution composée d'une onde de choc vérifie la relation de Rankine-Hugoniot (3.4). Pour l'onde de détente, il suffit de chercher les solutions auto-similaires régulières du problème de Riemann. Notons $\xi = x/t$ et $v(x/t) = u(t,x)$. La loi de conservation (3.1) devient alors

$$-\frac{x}{t^2}v'(\xi) + \frac{1}{t}f'(v(\xi))v'(\xi) = 0 \quad \implies \quad (f'(v(\xi)) - \xi)v'(\xi) = 0.$$

On en déduit que la solution composée d'une onde de détente (3.7) est une solution régulière par morceaux de (3.1) et globalement continue, elle est donc solution faible. **[EX]** \square

Remarque 5. Dans le cas d'un flux strictement convexe ou strictement concave, le problème de Riemann peut admettre des solutions composées de plusieurs ondes de choc mais pas de plusieurs ondes de détente. Elle peut de même être composée d'une ou plusieurs ondes de choc et d'une onde de détente. **[EX]**

Les solutions présentées ci-dessus ne sont pas les seules. On ne montrera pas l'unicité pour le problème de Riemann, on montrera seulement que la solution (3.6) n'est pas toujours admissible. L'unicité pour le problème de Riemann sera simplement déduite de l'unicité pour le problème de Cauchy.

3.4 Entropie et unicité

On va maintenant introduire la notion d'entropie, qui permet de définir des solutions admissibles et obtenir l'unicité. On verra dans le cas des systèmes hyperboliques le lien avec l'entropie physique.

Définition 3.7. On dit que (η, F) est un couple entropie-flux d'entropie pour (3.1) si $\eta \in \mathcal{C}^1(\mathbb{R}; \mathbb{R})$ est une fonction convexe et si $F' = \eta' f'$.

Dans le cas de solutions régulières de (3.1), on a alors que

$$\partial_t \eta(u) + \partial_x F(u) = 0.$$

Cette équation n'est plus vérifiée en présence de discontinuités.

Définition 3.8. Soit $u_0 \in \mathbf{L}^\infty(\mathbb{R})$. On appelle solution faible entropique de (3.1) une fonction $u \in \mathbf{L}^\infty(\mathbb{R}^+ \times \mathbb{R})$ vérifiant pour tout couple entropie-flux d'entropie (η, F) et pour tout $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^+ \times \mathbb{R})$ positif

$$\int_{\mathbb{R}^+} \int_{\mathbb{R}} (\eta(u)(t, x) \partial_t \varphi(t, x) + F(u)(t, x) \partial_x \varphi(t, x)) dx dt + \int_{\mathbb{R}} \eta(u_0)(x) \varphi(0, x) dx \geq 0. \quad (3.8)$$

L'inégalité (3.8) est simplement la forme au sens des distributions de

$$\partial_t \eta(u) + \partial_x F(u) \leq 0.$$

Proposition 3.9. Une solution faible entropique est une solution faible.

Démonstration. La démonstration est directe, en prenant dans (3.8) $\eta(u) = \pm u$ et $F(u) = \pm f(u)$. \square

Regardons maintenant ce que cela entraîne dans le cas d'une onde de choc :

Proposition 3.10. Les discontinuités des solutions faibles entropiques vérifient les relations de Rankine-Hugoniot (3.4) et l'inégalité

$$-\sigma'(\eta(U^+) - \eta(U^-)) + (F(U^+) - F(U^-)) \leq 0, \quad (3.9)$$

pour tout couple entropie-flux d'entropie (η, F) , avec les notations de la proposition 3.4.

Dans le cas d'un flux f strictement convexe ou concave, cela signifie que

$$f'(U^-) > \sigma' > f'(U^+). \quad (3.10)$$

C'est le critère de Lax.

Dans le cas d'un flux f quelconque, cela signifie que

$$\sigma' \leq \frac{f(U) - f(U^-)}{U - U^-}, \quad \forall U \in (U^- \perp U^+, U^- \top U^+). \quad (3.11)$$

C'est le critère (étendu) d'Oleinik.

Remarque 6. En fait, on peut encore préciser ces critères :

- Dans le cas d'un flux convexe, une seule entropie strictement convexe suffit pour sélectionner la solution entropique parmi les solutions faibles. **[EX]**
- Le premier critère (3.9) n'est pas exploitable en pratique. Dans le cas de la résolution du problème de Riemann, on utilise plutôt ses versions équivalentes (3.10) et (3.11).
- Bien sûr, (3.11) implique (3.10), mais la réciproque n'est pas vraie. **[EX]**
- On peut en fait montrer que dans le cas de solutions régulières par morceaux, une solution est entropique si et seulement si elle vérifie (3.11) le long des discontinuités et qu'elle est solution classique dans les zones de régularité, à la manière de la proposition 3.4.

Démonstration. La démonstration n'est pas évidente ; le plus simple est de la repousser à plus tard. En effet, la formulation basée sur les entropies de Kruzhkov présentée dans la partie suivante permet une démonstration directe. \square

On peut maintenant sélectionner les solutions entropiques parmi toutes les solutions faibles :

Théorème 3.11. *Le problème de Riemann (3.1)-(3.5) admet une et une seule solution auto-similaire. Dans le cas d'un flux strictement convexe (respectivement strictement concave), cette solution est composée d'une onde de choc si $f'(u_L) > f'(u_R)$ (resp. $f'(u_L) < f'(u_R)$), d'une onde de détente sinon.*

Démonstration. Dans le cas d'un flux strictement convexe ou concave, la démonstration peut se faire directement par construction. **[EX]**

Le cas d'un flux non strictement convexe ou concave se traite à l'aide du

Lemme 3.12. *La fonction $v(\xi)$, $\xi \in \mathbb{R}$, est la solution entropique auto-similaire du problème de Riemann (3.1)-(3.5) si et seulement si pour presque tout $(\xi_L, \xi_R) \in \mathbb{R}^2$, la fonction définie par*

$$\tilde{v}(\xi) = \begin{cases} v(\xi_L), & \text{si } \xi < \xi_L, \\ v(\xi), & \text{si } \xi_L < \xi < \xi_R, \\ v(\xi_R), & \text{si } \xi > \xi_R, \end{cases}$$

est la solution entropique du problème de Riemann (3.1)-(3.5) où la donnée initiale est

$$\tilde{u}_0(x) = \begin{cases} v(\xi_L), & \text{si } x < 0, \\ v(\xi_R), & \text{si } x > 0. \end{cases}$$

Ce lemme est une propriété fondamentale des solutions de problèmes de Riemann et est assez naturel (on ne le démontrera pas).

Le passage au cas d'un flux quelconque utilise le résultat suivant : **[EX]**

Lemme 3.13. *Considérons le problème de Riemann dans le cas d'un flux quelconque. Soit \tilde{f} défini ainsi : si $u_L < u_R$, \tilde{f} est l'enveloppe convexe de f entre u_L et u_R et si $u_L > u_R$, \tilde{f} est l'enveloppe concave du flux entre u_R et u_L . Alors la solution entropique pour (3.1) est identique à la solution entropique de l'équation*

$$\partial_t u + \partial_x \tilde{f}(u) = 0, \tag{3.12}$$

avec la même donnée initiale.

Ce lemme n'est pas évident à démontrer, on ne donne ici que les grandes lignes de sa démonstration. Il faut se placer dans le plan $(u, f(u))$ et décomposer la solution suivant les intervalles pour lesquelles f et \bar{f} coïncident ou pas pour les discontinuités. Ensuite, il faut noter que le critère d'Oleinik revient à considérer l'enveloppe *ad hoc* le long des discontinuités et utiliser le premier lemme.

L'unicité dans le cas d'un flux quelconque peut se faire par l'absurde, en montrant que la courbe paramétrée par ξ décrite par la solution dans le plan $(u, f(u))$ doit être convexe ou concave (selon le signe de $u_R - u_L$), sinon elle deviendrait une solution multivaluée. \square

On peut remarquer que ce dernier lemme donne une méthode constructive pour déterminer la solution entropique du problème de Riemann dans le cas d'un flux quelconque.

On verra par la suite que le problème de Cauchy admet lui aussi une unique solution, ce qui permettra d'enlever l'hypothèse d'auto-similarité de la solution du problème de Riemann (c'est aussi pour cette raison que l'on ne détaille pas plus la démonstration de l'unicité).

3.5 Problème de Cauchy

Avant d'étudier complètement le problème de Cauchy (3.1), on va considérer l'approximation visqueuse suivante :

$$\partial_t u_\varepsilon + \partial_x f(u_\varepsilon) = \varepsilon \partial_{xx}^2 u_\varepsilon, \quad (3.13)$$

où ε est une constante réelle positive. C'est une équation parabolique classique et on supposera l'existence et l'unicité de la solution dans un cadre suffisamment régulier, disons les fonctions uniformément bornées et de classe \mathcal{C}^2 .

Reprenons un couple entropie-flux d'entropie (η, F) avec $\eta \in \mathcal{C}^2$ et multiplions l'équation (3.13) par $\eta'(u_\varepsilon)$:

$$\begin{aligned} & \eta'(u_\varepsilon) \partial_t u_\varepsilon + \eta'(u_\varepsilon) \partial_x f(u_\varepsilon) = \varepsilon \eta'(u_\varepsilon) \partial_{xx}^2 u_\varepsilon, \\ \iff & \partial_t \eta(u_\varepsilon) + \partial_x F(u_\varepsilon) = \varepsilon \partial_{xx}^2 \eta(u_\varepsilon) - \eta''(u_\varepsilon) (\partial_x u_\varepsilon)^2, \\ \implies & \partial_t \eta(u_\varepsilon) + \partial_x F(u_\varepsilon) \leq \varepsilon \partial_{xx}^2 \eta(u_\varepsilon). \end{aligned} \quad (3.14)$$

On en déduit donc le résultat suivant :

Proposition 3.14. *Si on suppose que u_ε converge fortement vers $u \in \mathbf{L}^\infty(\mathbb{R}^+ \times \mathbb{R})$, alors u est une solution faible entropique de (3.1).*

Démonstration. Il suffit d'écrire (3.14) au sens des distributions et d'intégrer par partie. Comme $\eta(u_\varepsilon)$ est uniformément borné, le second membre tend vers 0, ce qui donne (3.8) dans le cas d'une entropie de classe \mathcal{C}^2 . Pour passer aux entropies de classe \mathcal{C}^1 , il suffit de définir une suite de fonctions convexes $\eta_n(u) = \eta(u) \star (n\rho(nu))$ où ρ est une fonction régulière positive à support compact. Le flux d'entropie associé est $F_n = \int_0^s f'(t) E_n'(t) dt$ et comme η_n converge uniformément vers η , il en est de même pour F_n vers F . On peut donc passer à la limite dans (3.8). \square

Les entropies au sens de la définition 3.7 ne sont pas aisément utilisables. L'idée fondamentale de Kruzhkov a été d'utiliser les fonctions suivantes :

Définition 3.15. On dit que (η_κ, F_κ) est un couple entropie-flux d'entropie de Kruzhkov pour (3.1) si $\eta_\kappa(u) = |u - \kappa|$ et si $F_\kappa(u) = \text{sgn}(u - \kappa)(f(u) - f(\kappa))$, où $\kappa \in \mathbb{R}$ et

$$\text{sgn}(s) = \begin{cases} 1, & \text{si } s > 0, \\ 0, & \text{si } s = 0, \\ -1, & \text{si } s < 0. \end{cases}$$

On peut utiliser une formulation uniquement basée sur les entropies de Kruzhkov (on utilisera aussi la notation $F(a, \kappa) = F_\kappa(a)$).

Proposition 3.16. Une fonction $u \in \mathbf{L}^\infty(\mathbb{R}^+ \times \mathbb{R})$ est une solution faible entropique de (3.1) si et seulement si pour tout $\kappa \in \mathbb{R}$ et tout $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^+ \times \mathbb{R})$ positif on a

$$\int_{\mathbb{R}^+} \int_{\mathbb{R}} (\eta_\kappa(u)(t, x) \partial_t \varphi(t, x) + F_\kappa(u)(t, x) \partial_x \varphi(t, x)) dx dt + \int_{\mathbb{R}} \eta_\kappa(u_0)(x) \varphi(0, x) dx \geq 0. \quad (3.15)$$

Démonstration. La formulation initiale est aussi valable pour les entropies convexes et de classe \mathcal{C}^0 , en utilisant le même raisonnement que dans la démonstration précédente. Ensuite, il suffit de remarquer que pour toute fonction η continue convexe, il existe une suite $\eta_\alpha(s) = b_0 + b_1 s + \sum_j a_j |s - \kappa_j|$ où $a_j > 0$ qui converge uniformément vers η (il en est de même pour les flux d'entropie associés). \square

Le fait qu'une solution faible entropique définie par (3.15) soit une solution faible s'obtient directement en prenant $\kappa = a$ puis $\kappa = b$ dans (3.15), où a et b sont les valeurs minimale et maximale prises par u_0 et u . [EX]

L'existence d'une solution faible entropique peut se faire à l'aide de diverses méthodes. En général, elles sont basées sur une approximation de (3.1) pour laquelle on démontre l'existence de solutions ainsi que des estimations *a priori* invariantes par passage à la limite dans l'approximation. Il suffit ensuite de s'assurer de la consistance avec (3.8) ou (3.15) de cette limite. Par exemple, l'approximation visqueuse (3.13) fournit cette possibilité, mais on ne la détaillera pas ici.

Passons à l'unicité.

Théorème 3.17 (Kruzhkov). Soit u_0 et v_0 deux données initiales dans $\mathbf{L}^\infty(\mathbb{R})$ et soit $a, b \in \mathbb{R}$ tels que $a \leq u_0, v_0 \leq b$ p.p. sur \mathbb{R} . On note u et v des solutions faibles entropiques associées respectivement à u_0 et v_0 . Alors, si on définit $L = \sup_{s \in [a, b]} |f'(s)|$, on a pour tout $T > 0$ et $r > 0$

$$\int_{|x| < r} |u(T, x) - v(T, x)| dx \leq \int_{|x| < r + LT} |u_0(x) - v_0(x)| dx. \quad (3.16)$$

Démonstration. L'idée principale est de parvenir à démontrer qu'au sens des distributions, on a

$$\partial_t |u - v| + \partial_x \text{sgn}(u - v)(f(u) - f(v)) \leq 0.$$

Une fois cette inéquation établie, il suffit alors d'intégrer sur le domaine $\{(t, x); t \in [0, T], |x| < r + L(T - t)\}$ pour obtenir (3.16).

En fait, le couple $(|u - v|, \text{sgn}(u - v)(f(u) - f(v)))$ est à la fois un couple entropie-flux d'entropie pour u à v fixé mais aussi pour v à u fixé. Soit $\varphi(x, t, y, s) \in \mathcal{C}_c^\infty((\mathbb{R}^+ \times$

\mathbb{R}^2) positif. On a donc

$$\begin{aligned} \int_{\mathbb{R}^+} \int_{\mathbb{R}} (\eta_{v(s,y)}(u(t,x)) \partial_t \varphi(t,x,s,y) + F_{v(s,y)}(u(t,x)) \partial_x \varphi(t,x,s,y)) dx dt \\ + \int_{\mathbb{R}} \eta_{v(s,y)}(u_0(x)) \varphi(0,x,s,y) dx \geq 0, \end{aligned} \quad (3.17)$$

et d'autre part

$$\begin{aligned} \int_{\mathbb{R}^+} \int_{\mathbb{R}} (\eta_{u(t,x)}(v(s,y)) \partial_s \varphi(t,x,s,y) + F_{u(t,x)}(v(s,y)) \partial_y \varphi(t,x,s,y)) dy ds \\ + \int_{\mathbb{R}} \eta_{u(t,x)}(v_0(y)) \varphi(t,x,0,y) dy \geq 0. \end{aligned} \quad (3.18)$$

En intégrant (3.17) par rapport à (s,y) sur $\mathbb{R}^+ \times \mathbb{R}$ et (3.18) par rapport à (t,x) sur $\mathbb{R}^+ \times \mathbb{R}$ puis en sommant, on obtient

$$\begin{aligned} \int_{\mathbb{R}^+} \int_{\mathbb{R}} \int_{\mathbb{R}^+} \int_{\mathbb{R}} (|u(t,x) - v(s,y)| (\partial_t + \partial_s) \varphi(t,x,s,y) \\ + \operatorname{sgn}(u(t,x) - v(s,y)) (f(u(t,x)) - f(v(s,y))) (\partial_x + \partial_y) \varphi(t,x,s,y)) dx dt dy ds \\ + \int_{\mathbb{R}^+} \int_{\mathbb{R}} \int_{\mathbb{R}} |u(t,x) - v_0(y)| \varphi(t,x,0,y) dy dx dt \\ + \int_{\mathbb{R}^+} \int_{\mathbb{R}} \int_{\mathbb{R}} |u_0(x) - v(s,y)| \varphi(0,x,s,y) dx dy ds \geq 0. \end{aligned} \quad (3.19)$$

On introduit une fonction positive $\rho \in \mathcal{C}_c^\infty(\mathbb{R})$ de masse 1 : $\int_{\mathbb{R}} \rho(z) dz = 1$ et on considère une fonction positive $\psi \in \mathcal{C}_c^\infty(\mathbb{R}^+ \times \mathbb{R})$. On choisit dans (3.19) φ telle que, pour $\varepsilon > 0$ petit,

$$\varphi(t,x,s,y) = \frac{1}{\varepsilon^2} \psi\left(\frac{t+s}{2}, \frac{x+y}{2}\right) \rho\left(\frac{t-s}{2\varepsilon}\right) \rho\left(\frac{x-y}{2\varepsilon}\right).$$

On peut remarquer que

$$\begin{aligned} (\partial_t + \partial_s) \varphi(t,x,s,y) &= \frac{1}{\varepsilon^2} \partial_1 \psi\left(\frac{t+s}{2}, \frac{x+y}{2}\right) \rho\left(\frac{t-s}{2\varepsilon}\right) \rho\left(\frac{x-y}{2\varepsilon}\right), \\ (\partial_x + \partial_y) \varphi(t,x,s,y) &= \frac{1}{\varepsilon^2} \partial_2 \psi\left(\frac{t+s}{2}, \frac{x+y}{2}\right) \rho\left(\frac{t-s}{2\varepsilon}\right) \rho\left(\frac{x-y}{2\varepsilon}\right), \end{aligned}$$

On passe ensuite à la limite $\varepsilon \rightarrow 0$, ce qui donne (on omet quelques détails...)

$$\begin{aligned} \int_{\mathbb{R}^+} \int_{\mathbb{R}} (|u - v|(t,x) \partial_t \psi(t,x) + \operatorname{sgn}(u - v) (f(u) - f(v))(t,x) \partial_x \psi(t,x)) dx dt \\ + \int_{\mathbb{R}} |u_0 - v_0|(x) \psi(0,x) dx \geq 0. \end{aligned} \quad (3.20)$$

On définit maintenant $\psi(t,x) = \chi_\varepsilon(t) \omega_\varepsilon(t,x)$ où

$$\chi_\varepsilon(t) = \begin{cases} 1, & \text{si } 0 \leq t < T, \\ (T-t)/\varepsilon + 1, & \text{si } T \leq t < T + \varepsilon, \\ 0, & \text{si } t \geq T + \varepsilon, \end{cases}$$

et

$$\omega_\varepsilon(t, x) = \begin{cases} 1, & \text{si } |x| \leq r + L(T - t), \\ (r + L(T - t) - |x|)/\varepsilon + 1, & \text{si } r + L(T - t) \leq |x| < r + L(T - t) + \varepsilon, \\ 0, & \text{si } |x| \geq r + L(T - t) + \varepsilon, \end{cases}$$

à une régularisation près pour rester dans \mathcal{C}_c^∞ . L'inégalité (3.20) devient

$$\begin{aligned} & \int_{\mathbb{R}^+} \int_{\mathbb{R}} |u - v|(t, x) (\chi_\varepsilon'(t) \omega_\varepsilon(t, x) + \chi_\varepsilon(t) \partial_t \omega_\varepsilon(t, x)) dx dt \\ & \quad + \int_{\mathbb{R}^+} \int_{\mathbb{R}} \operatorname{sgn}(u - v) (f(u) - f(v))(t, x) \chi_\varepsilon(t) \partial_x \omega_\varepsilon(t, x) dx dt \\ & \quad \quad \quad + \int_{\mathbb{R}} |u_0 - v_0|(x) \psi(0, x) dx \geq 0 \end{aligned}$$

$$\begin{aligned} & - \frac{1}{\varepsilon} \int_T^{T+\varepsilon} \int_{\mathbb{R}} |u - v|(t, x) \omega_\varepsilon(t, x) dx dt \\ & - \frac{L}{\varepsilon} \int_{\mathbb{R}^+} \int_{0 \leq |x| - r - L(T-t) < \varepsilon} |u - v|(t, x) \chi_\varepsilon(t) dx dt \\ & - \frac{1}{\varepsilon} \int_{\mathbb{R}^+} \int_{0 \leq |x| - r - L(T-t) < \varepsilon} \operatorname{sgn}(u - v) (f(u) - f(v))(t, x) \chi_\varepsilon(t) \operatorname{sgn}(x) dx dt \\ & + \int_{\mathbb{R}} |u_0 - v_0|(x) \psi(0, x) dx \geq 0. \end{aligned}$$

Comme $(\operatorname{sgn}(u - v)(f(u) - f(v)) \operatorname{sgn}(x)) \leq L|u - v|$, on obtient

$$\frac{1}{\varepsilon} \int_T^{T+\varepsilon} \int_{|x| < r} |u - v|(t, x) \omega_\varepsilon(t, x) dx dt \leq \int_{|x| < r + LT} |u_0 - v_0| \omega_\varepsilon(0, x) dx$$

et on retrouve bien (3.16) en faisant tendre ε vers 0. \square

Corollaire 3.18. *Soit u_0 et v_0 deux données initiales dans $\mathbf{L}^\infty(\mathbb{R}; [a, b])$ et u et v des solutions faibles entropiques associées. Alors :*

1. *Le problème de Cauchy (3.1) admet une unique solution faible entropique.*
2. *Si $u_0 \in \mathbf{L}^1(\mathbb{R})$, alors $u(t, \cdot) \in \mathbf{L}^1(\mathbb{R})$ et $\|u(t, \cdot)\|_{\mathbf{L}^1(\mathbb{R})} \leq \|u_0\|_{\mathbf{L}^1(\mathbb{R})}$.*
3. *On se place dans \mathbf{L}^1 . Si $u_0(x) \leq v_0(x)$ pour presque tout $x \in \mathbb{R}$, alors $u(t, x) \leq v(t, x)$ pour presque tout $t > 0$.*
4. *Si $u_0 \in \mathbf{BV}(\mathbb{R})$, alors $u(t, \cdot) \in \mathbf{BV}(\mathbb{R})$ et $|u(t, \cdot)|_{\mathbf{BV}(\mathbb{R})} \leq |u_0|_{\mathbf{BV}(\mathbb{R})}$.*
5. *Soit A et B deux constantes réelles telles que $A \leq u_0 \leq B$ p.p. sur \mathbb{R} . Alors, $A \leq u \leq B$ p.p. sur $\mathbb{R}^+ \times \mathbb{R}$.*

Démonstration.

1. On voit bien dans (3.16) que si $u_0 = v_0$, alors $u = v$ (presque partout). Pour avoir l'égalité pour presque tout $x \in \mathbb{R}$, il suffit de remarquer l'invariance par translation du problème de Cauchy.
2. Il suffit de prendre $v_0 \equiv 0$ (donc $v \equiv 0$) dans (3.16).

3. On peut vérifier que $2[z]^+ = |z| + z$, où $[z]^+ = \max(0, z)$. On déduit alors de (3.16) et de la conservation que

$$\int_{\mathbb{R}} [u(t, x) - v(t, x)]^+ dx \leq \int_{\mathbb{R}} [u_0(x) - v_0(x)]^+ dx. \quad (3.21)$$

Cette inégalité assure la monotonie.

4. Soit $h > 0$. On prend $v_0(x) = u_0(x+h)$ et la solution associée $v(t, x) = u(t, x+h)$ dans (3.16). On divise par h et on multiplie par une fonction test comme dans (1.10). Enfin, on « intègre par partie » et on fait ensuite tendre h vers 0.
5. On sait que si u_0 est une constante C , alors la solution faible entropique est égale à C pour tout $t > 0$ et $x \in \mathbb{R}$. Ainsi, en utilisant le point 3, on conclut directement.

Concernant les points 3 et 4, il est important de citer le lemme de Crandall et Tartar qui permet une conclusion (presque) directe :

Lemme 3.19. *Soit T une application de C un convexe de $\mathbf{L}^1(\Omega)$ telle que $\int_{\Omega} T(f) dx = \int_{\Omega} f dx$, $f \in C$. Alors les propriétés suivantes sont équivalentes pour tout $f, g \in C$:*

1. $f \leq g$ p.p. $\implies T(f) \leq T(g)$ p.p.,
2. $\int_{\Omega} |T(f) - T(g)| dx \leq \int_{\Omega} |f - g| dx$,
3. $\int_{\Omega} [T(f) - T(g)]^+ dx \leq \int_{\Omega} [f - g]^+ dx$.

(Ici, T est le semi-groupe $S(t)$ qui à u_0 associe $u(t, \cdot)$.) □

Remarque 7. L'inégalité (3.21) peut être précisée. En effet, tous les calculs de la démonstration de (3.16) peuvent être effectués en considérant les entropies $[u - \kappa]^+$ à la place de $|u - \kappa|$, ce qui donne alors

$$\int_{|x| < r} [u(t, x) - v(t, x)]^+ dx \leq \int_{|x| < r+Lt} [u_0(x) - v_0(x)]^+ dx.$$

3.6 Le cas multidimensionnel

On considère maintenant le cas multidimensionnel, c'est-à-dire le cas de l'équation

$$\partial_t u + \operatorname{div}_x f(u) = 0, \quad t > 0, x \in \mathbb{R}^d, \quad (3.22)$$

où $f: \mathbb{R} \rightarrow \mathbb{R}^d$.

De la même manière que précédemment, définissons la notion de solution faible entropique (on utilise directement les entropies de Kruzhkov) :

Définition 3.20. Une fonction $u \in \mathbf{L}^\infty(\mathbb{R}^+ \times \mathbb{R})$ est une solution faible entropique de (3.22) si pour tout $\kappa \in \mathbb{R}$ et tout $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^+ \times \mathbb{R}^d)$ positif on a

$$\begin{aligned} \int_{\mathbb{R}^+} \int_{\mathbb{R}^d} (|u - \kappa|(t, x) \partial_t \varphi(t, x) + \Phi(u, \kappa)(t, x) \cdot \nabla_x \varphi(t, x)) dx dt \\ + \int_{\mathbb{R}^d} |u_0 - \kappa|(x) \varphi(0, x) dx \geq 0 \end{aligned} \quad (3.23)$$

où $\Phi: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}^d$ et $\Phi_k(a, b) = \operatorname{sgn}(a - b)(f_k(a) - f_k(b))$, pour tout $k = 1, \dots, d$.

L'analyse du problème de Riemann n'est évidemment plus valide car c'est un problème intrinsèquement unidimensionnel. Pour le problème de Cauchy, l'analyse reste identique à la précédente : le théorème 3.17 (et son corollaire 3.18) reste valide et la preuve est identique.

Concernant l'existence de la solution faible entropique, on peut montrer que l'approximation visqueuse

$$\partial_t u_\varepsilon + \operatorname{div}_x f(u_\varepsilon) = \varepsilon \Delta u_\varepsilon,$$

est stable dans $\mathbf{L}^\infty \cap \mathbf{BV}$ et converge quand $\varepsilon \rightarrow 0$ vers la solution faible entropique associée à (3.22).

On aboutit donc au résultat suivant :

Théorème 3.21. *Soit $u_0 \in \mathbf{L}^\infty(\mathbb{R}^d)$. Alors il existe une unique solution faible entropique $u \in \mathbf{L}^\infty(\mathbb{R}^+ \times \mathbb{R})$. De plus, soit u_0 et v_0 deux données initiales dans $\mathbf{L}^\infty(\mathbb{R}^d)$ et soit $a, b \in \mathbb{R}$ tels que $a \leq u_0, v_0 \leq b$ p.p. sur \mathbb{R}^d . On note u et v des solutions faibles entropiques associées respectivement à u_0 et v_0 . Alors, si on définit $L = \sup_k \sup_{s \in [a, b]} |f'_k(s)|$, on a pour tout $T > 0$ et $r > 0$*

$$\int_{|x| < r} |u(T, x) - v(T, x)| \, dx \leq \int_{|x| < r+LT} |u_0(x) - v_0(x)| \, dx. \quad (3.24)$$

Chapitre 4

Méthodes des volumes finis pour les lois de conservation

On s'intéresse maintenant à l'approximation des lois de conservation par les méthodes de volumes finis (VF). On va se placer dans le cadre unidimensionnel (on trouvera en fin de chapitre quelques commentaires sur le cas multidimensionnel).

4.1 Notations et principes des méthodes de volumes finis

Présentons tout d'abord la notion de maillage. Soit une suite réelle strictement croissante $(x_{i+1/2})_{i \in \mathbb{Z}}$ représentant les interfaces entre les mailles M_i . On définit les pas d'espace $\Delta x_i = x_{i+1/2} - x_{i-1/2}$ qui correspondent aux mesures des mailles. On définit en suite le pas de temps Δt et $t^n = n\Delta t$.

L'idée de base des méthode VF est de définir

$$u_i^0 = \frac{1}{\Delta x_i} \int_{M_i} u_0(x) dx \quad (4.1)$$

pour pouvoir calculer la suite $(u_i^n)_{i \in \mathbb{Z}, n \in \mathbb{N}}$ telle que

$$u_i^n \approx \frac{1}{\Delta x_i} \int_{M_i} u(t^n, x) dx \quad (4.2)$$

où u est la solution faible entropique du problème de Cauchy

$$\begin{cases} \partial_t u + \partial_x f(u) = 0, & t > 0, x \in \mathbb{R}, \\ u(0, x) = u_0(x), & x \in \mathbb{R}. \end{cases} \quad (4.3)$$

Pour construire cette suite, on intègre cette loi de conservation sur le carré $(t^n, t^{n+1}) \times M_i$, on a

$$\begin{aligned} \int_{M_i} u(t^{n+1}, x) dx - \int_{M_i} u(t^n, x) dx \\ + \int_{t^n}^{t^{n+1}} f(u(t, x_{i+1/2})) dt - \int_{t^n}^{t^{n+1}} f(u(t, x_{i-1/2})) dt = 0 \end{aligned}$$

qui devient après utilisation de l'approximation (4.2)

$$\Delta x_i (u_i^{n+1} - u_i^n) + \Delta t (f_{i+1/2}^n - f_{i-1/2}^n) \approx 0$$

où $\Delta t f_{i+1/2}^n \approx \int_{t^n}^{t^{n+1}} f(u(t, x_{i+1/2})) dt$. Si on définit $f_{i+1/2}^n$ par une fonction dépendant simplement de $(u_i^n)_{i \in \mathbb{Z}}$, on obtient un schéma explicite à un pas dans la terminologie de l'approximation des équations différentielles.

On ne va s'intéresser ici qu'aux schémas à trois points, c'est-à-dire que le flux numérique est défini par $f_{i+1/2}^n = g(u_i^n, u_{i+1}^n)$, ce qui donne

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x_i} (g(u_i^n, u_{i+1}^n) - g(u_{i-1}^n, u_i^n)). \quad (4.4)$$

On a comme conséquence immédiate de cette écriture :

Proposition 4.1. *Le schéma VF (4.4) est conservatif, c'est-à-dire que si $u_0 \in \mathbf{L}^1(\mathbb{R})$, alors pour tout $n > 0$,*

$$\sum_{i \in \mathbb{Z}} u_i^n \Delta x_i = \int_{\mathbb{R}} u_0(x) dx.$$

Démonstration. La démonstration est directe en sommant (4.4) pour $i \in \mathbb{Z}$ et en utilisant (4.1). \square

Donnons quelques exemples de schémas VF :

– Schéma de Rusanov

$$g(u, v) = \frac{f(u) + f(v)}{2} - \frac{a(u, v)}{2} (v - u)$$

où $a(u, v) = \max(|f'(u)|, |f'(v)|)$.

– Schéma de Godunov

$$g(u, v) = \begin{cases} \min_{a \in [u, v]} f(a) & \text{si } u < v, \\ \max_{a \in [v, u]} f(a) & \text{sinon.} \end{cases}$$

Ce schéma peut se réinterpréter comme un algorithme de transport-projection pourvu que $\Delta t \leq \Delta x / (2L)$.

– Schéma d'Engquist-Osher

$$g(u, v) = \frac{f(u) + f(v)}{2} - \frac{1}{2} \int_u^v |f'(s)| ds.$$

– Schéma de Murman-Roe

$$g(u, v) = \frac{f(u) + f(v)}{2} - \frac{a(u, v)}{2} (v - u)$$

où

$$a(u, v) = \begin{cases} (f(v) - f(u)) / (v - u) & \text{si } u \neq v, \\ f'(u) & \text{sinon.} \end{cases}$$

– Schéma de Lax-Wendroff

$$g(u, v) = \frac{f(u) + f(v)}{2} - \frac{\Delta t}{\Delta x} f'((u+v)/2) (f(v) - f(u)).$$

4.2 Schémas monotones

Il est clair que tout schéma de la forme (4.4) ne converge pas vers la solution entropique de (4.3). On va introduire quelques hypothèses sur le flux numérique g :

Définition 4.2. Soit g une fonction lipschitzienne de classe $\mathcal{C}^1(\mathbb{R}^2; \mathbb{R})$ (avec la même constante de Lipschitz que f). On dit que :

- g est consistant si pour tout $u \in \mathbb{R}$, $g(u, u) = f(u)$,
- g est monotone si c'est une fonction croissante par rapport à sa première variable et décroissante par rapport à sa deuxième variable.

On suppose à partir de maintenant que le flux numérique g vérifie ces hypothèses.

On va vouloir démontrer que la suite $(u_i^n)_{i \in \mathbb{Z}, n \in \mathbb{N}}$ converge vers la solution entropique. Pour cela, on va obtenir des bornes sur cette suite (généralement appelées estimations *a priori*), ce qui va nous permettre de montrer qu'à une sous-suite près, elle converge. Ensuite, la consistance du schéma VF avec l'équation (4.3) nous assurera que la limite sera bien la solution entropique. Enfin, comme la solution entropique est unique, toute la suite converge.

4.2.1 Estimations *a priori*

Pour commencer, on définit la fonction u_Δ par

$$u_\Delta(t, x) = \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{Z}} u_i^n \mathbf{1}_{[n\Delta, (n+1)\Delta] \times M_i}(t, x).$$

On doit donc montrer que cette fonction converge vers la solution entropique quand Δt et Δx tendent vers 0.

Le cadre fonctionnel dans lequel il est agréable de travailler quand on traite de lois de conservation *unidimensionnelles* est le cadre $\mathbf{L}^\infty \cap \mathbf{BV}$, pour lequel le théorème de Helly permet de passer à la limite.

Proposition 4.3 (Estimation \mathbf{L}^∞ et monotonie). *Sous la condition CFL (Courant-Friedrichs-Lewy)*

$$\Delta t \leq \frac{\min_{i \in \mathbb{Z}} \Delta x_i}{2L} \quad (4.5)$$

le schéma (4.4) est monotone, c'est-à-dire que la fonction H définie par

$$u_i^{n+1} = H(u_{i-1}^n, u_i^n, u_{i+1}^n)$$

est croissante par rapport à ses trois variables. De plus, si on note $A, B \in \mathbb{R}$ tels que $A \leq u_0 \leq B$ p.p., alors

$$A \leq u_\Delta(t, x) \leq B \quad (4.6)$$

pour presque tout $t > 0$ et $x \in \mathbb{R}$.

Démonstration. On a

$$\begin{aligned} \partial_1 H(u, v, w) &= \frac{\Delta t}{\Delta x_i} \partial_1 g(u, v), \\ \partial_3 H(u, v, w) &= -\frac{\Delta t}{\Delta x_i} \partial_2 g(v, w), \\ \partial_2 H(u, v, w) &= 1 - \frac{\Delta t}{\Delta x_i} (\partial_1 g(v, w) - \partial_2 g(u, v)). \end{aligned}$$

On voit immédiatement que sous la condition CFL, $\partial_i H$ est bien positif. Par ailleurs, on a que

$$H(A, A, A) = A \quad \text{et} \quad H(B, B, B) = B.$$

Le schéma étant monotone sous la condition CFL (4.5), on en déduit la stabilité \mathbf{L}^∞ (4.6). \square

Proposition 4.4. *Sous la condition CFL (4.5),*

$$\sum_{i \in \mathbb{Z}} |u_{i+1}^{n+1} - u_i^{n+1}| \leq \sum_{i \in \mathbb{Z}} |u_{i+1}^n - u_i^n|, \quad \forall n \in \mathbb{N}. \quad (4.7)$$

On dit que le schéma est à variation totale décroissante (schéma TVD).

Démonstration. On définit

$$b_{i+1/2}^n = \begin{cases} \frac{\Delta t}{\Delta x_i} \frac{g(u_i^n, u_{i+1}^n) - f(u_i^n)}{u_i^n - u_{i+1}^n} & \text{si } u_i^n \neq u_{i+1}^n, \\ 0 & \text{sinon,} \end{cases}$$

et

$$a_{i-1/2}^n = \begin{cases} \frac{\Delta t}{\Delta x_i} \frac{g(u_{i-1}^n, u_i^n) - f(u_i^n)}{u_{i-1}^n - u_i^n} & \text{si } u_i^n \neq u_{i-1}^n, \\ 0 & \text{sinon,} \end{cases}$$

alors le schéma (4.4) peut s'écrire

$$u_i^{n+1} = (1 - b_{i+1/2}^n - a_{i-1/2}^n)u_i^n + b_{i+1/2}^n u_{i+1}^n + a_{i-1/2}^n u_{i-1}^n. \quad (4.8)$$

Tout d'abord, comme g est décroissant (respectivement croissant) par rapport à sa deuxième variable (resp. première variable), alors $b_{i+1/2}^n \geq 0$ (resp. $a_{i-1/2}^n \geq 0$). De plus,

$$a_{i-1/2}^n + b_{i+1/2}^n \leq 2L \frac{\Delta t}{\Delta x_i} \leq 1$$

sous la condition CFL (4.5). La forme (4.8) revient donc à écrire u_i^{n+1} comme une combinaison convexe de u_{i-1}^n , u_i^n et u_{i+1}^n (les coefficients associés sont compris entre 0 et 1 et leur somme vaut 1). Ainsi, si on écrit (4.8) en $i+1$ et que l'on y soustrait (4.8), on obtient

$$\begin{aligned} u_{i+1}^{n+1} - u_i^{n+1} &= (u_{i+1}^n - u_i^n)(1 - b_{i+1/2}^n - a_{i+1/2}^n) \\ &\quad + (u_{i+2}^n - u_{i+1}^n)b_{i+3/2}^n + (u_i^n - u_{i-1}^n)a_{i-1/2}^n \end{aligned}$$

dont on déduit sous la condition (4.5) que

$$\begin{aligned} |u_{i+1}^{n+1} - u_i^{n+1}| &\leq |u_{i+1}^n - u_i^n|(1 - b_{i+1/2}^n - a_{i+1/2}^n) \\ &\quad + |u_{i+2}^n - u_{i+1}^n|b_{i+3/2}^n + |u_i^n - u_{i-1}^n|a_{i-1/2}^n. \end{aligned}$$

Il suffit maintenant de sommer sur $i \in \mathbb{Z}$ pour obtenir (4.7). \square

Remarque 8. La forme (4.8) est due à LeRoux et Harten. Celle-ci est très utile puisqu'elle peut aussi permettre d'obtenir simplement la stabilité \mathbf{L}^∞ étant donné que sous la condition CFL u_i^{n+1} est une combinaison convexe de u_{i-1}^n , u_i^n et u_{i+1}^n .

Nous avons donc une borne BV en espace. Si on se place sur l'ouvert $(-T, T) \times \mathbb{R}$, $T > 0$, on va voir que celle-ci implique une borne $BV(((-T, T) \times \mathbb{R}))$ sur u_Δ . En fait, on se place sur $(-T, T)$ mais il suffit de se placer sur $(-\varepsilon, T)$ où $\varepsilon > 0$, pour pouvoir inclure la condition initiale dans le passage à la limite.

Corollaire 4.5 (Estimation BV). *Supposons que la donnée initiale est à variation bornée, $u_0 \in BV(\mathbb{R})$. Sous la condition CFL (4.5), il existe une constante C ne dépendant que de T , u_0 et g telle que*

$$|u_\Delta|_{BV(((-T, T) \times \mathbb{R}))} \leq C \quad (4.9)$$

où $u_\Delta(t, \cdot) = u_\Delta(0^+, \cdot)$ pour $t < 0$.

Démonstration. Tout d'abord, notons que

$$|u_\Delta|_{BV(((-T, T) \times \mathbb{R}))} = \sum_n \sum_i \Delta t |u_i^n - u_{i-1}^n| + \sum_n \sum_i \Delta x_i |u_i^{n+1} - u_i^n|$$

De plus, $\sum_{i \in \mathbb{Z}} |u_{i+1}^0 - u_i^0| \leq |u_0|_{BV(\mathbb{R})}$.

Supposons que $T \leq \Delta t$, alors $|u_\Delta|_{BV(((-T, T) \times \mathbb{R}))} \leq 2T \sum_{i \in \mathbb{Z}} |u_{i+1}^0 - u_i^0|$. Donc (4.9) est vérifié avec $C = 2T|u_0|_{BV(\mathbb{R})}$.

Supposons maintenant que $T > \Delta t$ et définissons $N > 0$ par $N\Delta t < T \leq (N+1)\Delta t$. On sait que

$$\begin{aligned} |u_\Delta|_{BV(((-T, T) \times \mathbb{R}))} &\leq T \sum_{i \in \mathbb{Z}} |u_{i+1}^0 - u_i^0| + \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} \Delta t |u_{i+1}^n - u_i^n| \\ &\quad + (T - N\Delta t) \sum_{i \in \mathbb{Z}} |u_{i+1}^N - u_i^N| + \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} \Delta x_i |u_i^{n+1} - u_i^n|. \end{aligned} \quad (4.10)$$

Les trois premiers termes peuvent être bornés par $2T|u_0|_{BV(\mathbb{R})}$ en utilisant la proposition 4.4. De plus, on peut écrire en utilisant directement la forme (4.4) que

$$|u_i^{n+1} - u_i^n| \leq 2L \frac{\Delta t}{\Delta x_i} (|u_i^n - u_{i-1}^n| + |u_{i+1}^n - u_i^n|)$$

qui devient après sommation sur $i \in \mathbb{Z}$

$$\sum_{i \in \mathbb{Z}} \Delta x_i |u_i^{n+1} - u_i^n| \leq 2L\Delta t \sum_{i \in \mathbb{Z}} |u_i^n - u_{i-1}^n|.$$

Ainsi, puisque $N\Delta t < T$ le dernier terme de (4.10) est borné par $2LT|u_0|_{BV(\mathbb{R})}$. On a finalement l'inégalité (4.9) avec $C = 2T(1+L)|u_0|_{BV(\mathbb{R})}$. \square

4.2.2 Convergence

Comme la suite $(u_\Delta)_\Delta$ est bornée dans $L^\infty \cap BV$ sous la condition CFL (4.5), le théorème de Helly assure à une sous-suite près la convergence L^1_{loc} de $(u_\Delta)_\Delta$ vers une fonction $u \in L^\infty \cap BV$.

Il reste maintenant à vérifier que la limite ainsi obtenue est la solution entropique. Pour cela, il suffit d'utiliser le théorème de Lax-Wendroff (on se place dans le cadre d'un maillage régulier pour simplifier) :

Théorème 4.6 (Lax-Wendroff). *Supposons que la suite $(u_\Delta)_\Delta$ est uniformément bornée dans $L^\infty(\mathbb{R}^+ \times \mathbb{R})$ et qu'elle converge vers u dans $L^1_{loc}(\mathbb{R}^+ \times \mathbb{R})$ quand $\sup_i \Delta x_i$ et Δt tendent vers 0, alors u est solution faible du problème de Cauchy (4.3).*

Démonstration. L'idée de base de la démonstration n'est pas très compliquée, mais les calculs sont fastidieux. On ne va donner que les étapes principales.

Soit $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^+ \times \mathbb{R})$ une fonction test positive. On multiplie le schéma (4.4) par $\int_{M_i} \varphi(n\Delta t, x) dx$ et on somme sur $i \in \mathbb{Z}$ et $n \in \mathbb{N}$, ce qui donne

$$A_\Delta + B_\Delta = 0$$

où

$$\begin{aligned} A_\Delta &= \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{Z}} (u_i^{n+1} - u_i^n) \int_{M_i} \varphi(n\Delta t, x) dx, \\ B_\Delta &= \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{Z}} \frac{\Delta t}{\Delta x} (g(u_i^n, u_{i+1}^n) - g(u_{i-1}^n, u_i^n)) \int_{M_i} \varphi(n\Delta t, x) dx. \end{aligned}$$

Il faut donc démontrer que si $\Delta x \rightarrow 0$,

$$A_\Delta \longrightarrow - \int_{\mathbb{R}} \int_{\mathbb{R}^+} u \partial_t \varphi dt dx - \int u_0(x) \varphi(0, x) dx, \quad (4.11)$$

$$B_\Delta \longrightarrow - \int_{\mathbb{R}} \int_{\mathbb{R}^+} f(u) \partial_x \varphi dt dx. \quad (4.12)$$

Cela se fait simplement à l'aide de la règle d'Abel (intégration par parties discrète) et du théorème de convergence dominé ! Regardons le premier terme :

$$\begin{aligned} A_\Delta &= \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{Z}} (u_i^{n+1} - u_i^n) \int_{M_i} \varphi(n\Delta t, x) dx, \\ &= - \sum_{n \in \mathbb{N}^*} \sum_{i \in \mathbb{Z}} u_i^n \int_{M_i} (\varphi(n\Delta t, x) - \varphi((n-1)\Delta t, x)) dx - \sum_{i \in \mathbb{Z}} u_i^0 \int_{M_i} \varphi(0, x) dx, \\ &= - \sum_{n \in \mathbb{N}^*} \sum_{i \in \mathbb{Z}} \int_{M_i} \int_{n\Delta t}^{(n+1)\Delta t} u_\Delta(t, x) \frac{\varphi(n\Delta t, x) - \varphi((n-1)\Delta t, x)}{\Delta t} dt dx, \\ &\quad - \sum_{i \in \mathbb{Z}} \int_{M_i} u_\Delta(0, x) \varphi(0, x) dx \\ &= - \sum_{n \in \mathbb{N}^*} \sum_{i \in \mathbb{Z}} \int_{M_i} \int_{n\Delta t}^{(n+1)\Delta t} u_\Delta(t, x) \left(\frac{\varphi(n\Delta t, x) - \varphi((n-1)\Delta t, x)}{\Delta t} - \partial_t \varphi(t, x) \right) dt dx \\ &\quad - \sum_{n \in \mathbb{N}^*} \sum_{i \in \mathbb{Z}} \int_{M_i} \int_{n\Delta t}^{(n+1)\Delta t} u_\Delta(t, x) \partial_t \varphi(t, x) dt dx \\ &\quad - \sum_{i \in \mathbb{Z}} \int_{M_i} u_\Delta(0, x) \varphi(0, x) dx \end{aligned}$$

La suite (u_Δ) étant bornée, le premier terme tend vers 0 et les deux autres termes assurent (4.11) grâce aux hypothèses sur la suite (u_Δ) . Passons maintenant au terme B_Δ . Avant tout calcul, introduisons B'_Δ

$$B'_\Delta = - \sum_{n \in \mathbb{N}} \int_{n\Delta t}^{(n+1)\Delta t} \int_{\mathbb{R}} f(u_\Delta(t, x)) \partial_x \varphi(n\Delta t, x) dt dx$$

qui converge bien vers la limite (4.12), notamment car f est Lipschitz-continu. Ce

terme peut aussi s'écrire

$$\begin{aligned} B'_\Delta &= \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{Z}} \Delta t (f(u_i^n) - f(u_{i-1}^n)) \varphi(n\Delta t, x_{i-1/2}) \\ &= \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{Z}} \Delta t (f(u_i^n) - g(u_{i-1}^n, u_i^n)) \varphi(n\Delta t, x_{i-1/2}) \\ &\quad + \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{Z}} \Delta t (g(u_{i-1}^n, u_i^n) - f(u_{i-1}^n)) \varphi(n\Delta t, x_{i-1/2}). \end{aligned}$$

De même, on peut écrire

$$\begin{aligned} B_\Delta &= \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{Z}} \Delta t (f(u_i^n) - g(u_{i-1}^n, u_i^n)) \frac{1}{\Delta x} \int_{M_i} \varphi(n\Delta t, x) dx \\ &\quad + \sum_{n \in \mathbb{N}} \sum_{i \in \mathbb{Z}} \Delta t (g(u_i^n, u_{i+1}^n) - f(u_i^n)) \frac{1}{\Delta x} \int_{M_i} \varphi(n\Delta t, x) dx. \end{aligned}$$

Grâce au caractère lipschitzien du flux et à la consistance, on peut majorer les termes en $g(u, v) - f(u)$, ce qui donne (I et N vérifient $\text{supp } \varphi \subset [0, N\Delta t) \times (x_{-I+1/2}, x_{I-1/2})$)

$$|B_\Delta - B'_\Delta| \leq C \Delta x \Delta t \sum_{n=0}^N \sum_{i=-I}^I |u_\Delta(n\Delta t, x_{i-1/2} + \Delta x_i) - u_\Delta(n\Delta t, x_{i-1/2})|$$

où C ne dépend que de L et de $\|\varphi\|_\infty$ et N et I sont définis par le support de φ . Par continuité de la translation, le second membre tend vers 0 quand $\sup_i \Delta x_i$ tend vers 0. \square

Il est important de noter que cette démonstration n'utilise que la régularité et la consistance du flux numérique g . Si on utilise la monotonie, on peut montrer le théorème suivant :

Théorème 4.7. *On considère le schéma VF (4.4) et sa représentation $(u_\Delta)_\Delta$. On suppose que la condition CFL (4.5) est vérifiée (on peut prendre $\Delta t = \alpha \sup_i \Delta x_i / (2L)$ par exemple, avec $0 < \alpha < 1$ fixé). Alors le schéma VF converge vers u dans $\mathbf{L}_{\text{loc}}^1(\mathbb{R} \times \mathbb{R}^+)$, où u est la solution entropique du problème de Cauchy (4.3).*

Démonstration. Le point important est d'obtenir ce qu'on appelle des inégalités d'entropie discrètes, c'est-à-dire des équations discrètes vérifiées par le schéma numérique analogue à la formulation entropique. On peut en fait vérifier :

Lemme 4.8. *Sous la condition CFL (4.5), le schéma (4.4) vérifie les inégalités suivantes :*

$$\frac{1}{\Delta t} (|u_i^{n+1} - \kappa| - |u_i^n - \kappa|) + \frac{1}{\Delta x_i} (G_{i+1/2}^n - G_{i-1/2}^n) \leq 0, \quad (4.13)$$

$\forall n \in \mathbb{N}, \forall i \in \mathbb{Z}, \forall \kappa \in \mathbb{R}$, où

$$G_{i+1/2}^n = g(u_i^n \top \kappa, u_{i+1}^n \top \kappa) - g(u_i^n \perp \kappa, u_{i+1}^n \perp \kappa).$$

Démonstration. La démonstration repose sur la décomposition $|u_i^{n+1} - \kappa| = u_i^{n+1} \top \kappa - u_i^{n+1} \perp \kappa$ et sur la monotonie de H . En effet,

$$\begin{aligned} u_i^{n+1} \top \kappa &= H(u_{i-1}^n, u_i^n, u_{i+1}^n) \top \kappa = H(u_{i-1}^n, u_i^n, u_{i+1}^n) \top H(\kappa, \kappa, \kappa) \\ &\leq H(u_{i-1}^n \top \kappa, u_i^n \top \kappa, u_{i+1}^n \top \kappa) \end{aligned}$$

et

$$\begin{aligned} u_i^{n+1} \perp \kappa &= H(u_{i-1}^n, u_i^n, u_{i+1}^n) \perp \kappa = H(u_{i-1}^n, u_i^n, u_{i+1}^n) \perp H(\kappa, \kappa, \kappa) \\ &\geq H(u_{i-1}^n \perp \kappa, u_i^n \perp \kappa, u_{i+1}^n \perp \kappa) \end{aligned}$$

donc

$$|u_i^{n+1} - \kappa| \leq H(u_{i-1}^n \top \kappa, u_i^n \top \kappa, u_{i+1}^n \top \kappa) - H(u_{i-1}^n \perp \kappa, u_i^n \perp \kappa, u_{i+1}^n \perp \kappa)$$

qui est exactement (4.13). \square

On sait grâce aux estimations (4.6) et (4.9) et par le théorème de Helly que la suite définie à partir du schéma numérique converge dans $\mathbf{L}_{\text{loc}}^1()$, à une extraction de sous-suite près. Ensuite, on applique aux inégalités (4.13) le même raisonnement que celui utilisé dans le théorème de Lax-Wendroff et on passe à la limite (voir le Lemme 4.10). On peut donc extraire une sous-suite qui converge et la limite est solution faible entropique. De plus, comme la solution faible entropique est unique, on obtient la convergence de toute la suite vers cette limite. \square

Corollaire 4.9. *Le problème de Cauchy (4.3) admet une et une seule solution faible entropique dans $\mathbf{L}^\infty(\mathbb{R}^+ \times \mathbb{R})$.*

Démonstration. On savait déjà que le problème de Cauchy (4.3) admettait au plus une solution. L'approximation numérique par le schéma VF (4.4) nous fournit une démonstration de l'existence d'une telle solution, pour autant que la donnée initiale soit dans $\mathbf{L}^\infty(\mathbb{R}) \cap \text{BV}(\mathbb{R})$. Pour passer de ce cadre au cadre $u_0 \in \mathbf{L}^\infty(\mathbb{R})$, on définit la suite $u_0^\alpha = \rho_\alpha \star (\mathbf{1}_{(-1/\alpha, 1/\alpha)} u_0)$ où $\rho_\alpha \in \mathcal{C}_c^\infty(\mathbb{R})$ positif tel que $\text{supp } \rho_\alpha \subset [-1/\alpha, 1/\alpha]$. Cette suite est bien dans $\mathbf{L}^\infty(\mathbb{R}) \cap \text{BV}(\mathbb{R})$ et converge vers u_0 dans $\mathbf{L}_{\text{loc}}^1(\mathbb{R})$. Le théorème de Kruzhkov assure que

$$\int_{|x|<r} |u^{\alpha_1}(T, x) - u^{\alpha_2}(T, x)| dx \leq \int_{|x|<r+LT} |u_0^{\alpha_1}(x) - u_0^{\alpha_2}(x)| dx.$$

Cette inégalité entraîne que la suite $(u^\alpha)_\alpha$ est une suite de Cauchy dans $\mathbf{L}^\infty(\mathbb{R}^+ \times \mathbb{R})$, donc elle converge vers une limite u , qui n'est autre que la solution faible entropique associée à la donnée initiale $u_0 \in \mathbf{L}^\infty(\mathbb{R})$. \square

4.2.3 Estimations *a posteriori*

On s'intéresse maintenant à étudier l'erreur commise par le schéma numérique. L'idée est d'obtenir un résultat de comparaison entre la solution approchée u_Δ et la solution entropique u d'un même problème de Cauchy. On verra lors des calculs qu'il est nécessaire de se placer dans BV pour obtenir les estimations qui vont suivre. On supposera aussi que la condition CFL (4.5) est vérifiée et que le rapport $\Delta t/\Delta x$ est constant.

Tout d'abord, énonçons le résultat de consistance suivant :

Lemme 4.10. *Soit $u_0 \in \text{BV}(\mathbb{R})$. Alors pour toute fonction $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^+ \times \mathbb{R})$ positive, le schéma monotone définissant u_Δ vérifie pour tout $t^N = N\Delta t$ et $\kappa \in \mathbb{R}$*

$$\begin{aligned} - \int_0^{t^N} \int_{\mathbb{R}} (|u_\Delta(t, x) - \kappa| \partial_t + F(u_\Delta, \kappa) \partial_x) \varphi(t, x) dt dx - \int_{\mathbb{R}} |u_\Delta(0, x) - \kappa| \varphi(0, x) dx \\ + \int_{\mathbb{R}} |u_\Delta(t^N, x) - \kappa| \varphi(t^N, x) dx \leq C \Delta x \text{TV}(u_0) t^N, \end{aligned} \quad (4.14)$$

où C ne dépend que des constantes de Lipschitz du flux numérique g et de φ .

On peut noter que ce résultat permet de démontrer que si le schéma numérique converge presque partout, alors il converge vers la solution faible entropique.

Démonstration. On considère l'inégalité d'entropie discrète (4.13) que l'on multiplie par $\Delta t \int_{M_i} \varphi(t^{n+1}, x) dx$ et on somme sur $i \in \mathbb{Z}$ et sur $n \in [0, N-1]$. On note

$$\begin{aligned} A_\Delta &= \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} (|u_i^{n+1} - \kappa| - |u_i^n - \kappa|) \int_{M_i} \varphi(t^{n+1}, x) dx, \\ B_\Delta &= \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} \frac{\Delta t}{\Delta x} (G_{i+1/2}^n - G_{i-1/2}^n) \int_{M_i} \varphi(t^{n+1}, x) dx, \end{aligned}$$

ce qui donne $A_\Delta + B_\Delta \leq 0$. En appliquant la règle d'Abel, on obtient

$$\begin{aligned} A_\Delta &= \sum_{i \in \mathbb{Z}} \left(- \sum_{n=0}^{N-1} |u_i^n - \kappa| \int_{M_i} (\varphi(t^{n+1}, x) - \varphi(t^n, x)) dx \right. \\ &\quad \left. - |u_i^0 - \kappa| \int_{M_i} \varphi(0, x) dx + |u_i^N - \kappa| \int_{M_i} \varphi(t^N, x) dx \right) \\ &= \sum_{i \in \mathbb{Z}} \left(- \sum_{n=0}^{N-1} |u_i^n - \kappa| \int_{M_i} \int_{t^n}^{t^{n+1}} \partial_t \varphi(t, x) dt dx \right. \\ &\quad \left. - \int_{M_i} |u_\Delta(0, x) - \kappa| \varphi(0, x) dx + \int_{M_i} |u_\Delta(t^N, x) - \kappa| \varphi(t^N, x) dx \right) \\ &= - \int_0^{t^N} \int_{\mathbb{R}} |u_\Delta(t, x) - \kappa| \partial_t \varphi(t, x) dx dt - \int_{\mathbb{R}} |u_\Delta(0, x) - \kappa| \varphi(0, x) dx \\ &\quad + \int_{\mathbb{R}} |u_\Delta(t^N, x) - \kappa| \varphi(t^N, x) dx \end{aligned}$$

Concernant le terme B_Δ , on a

$$\begin{aligned} B_\Delta &= \Delta t \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} (G_{i+1/2}^n - F(u_i^n, \kappa)) \frac{1}{\Delta x} \int_{M_i} \varphi(t^{n+1}, x) dx \\ &\quad + \Delta t \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} (F(u_i^n, \kappa) - G_{i-1/2}^n) \frac{1}{\Delta x} \int_{M_i} \varphi(t^{n+1}, x) dx \end{aligned}$$

On introduit maintenant le terme $B'_\Delta = \int_0^{t^N} \int_{\mathbb{R}} F(u_\Delta, \kappa) \partial_x \varphi \, dt \, dx$, qui s'écrit aussi :

$$\begin{aligned}
B'_\Delta &= - \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} \int_{M_i} \int_{t^n}^{t^{n+1}} F(u_\Delta, \kappa) \partial_x \varphi(t, x) \, dt \, dx \\
&= - \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} \int_{t^n}^{t^{n+1}} F(u_i^n, \kappa) (\varphi(t, x_{i+1/2}) - \varphi(t, x_{i-1/2})) \, dt \\
&= - \Delta t \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} F(u_i^n, \kappa) (\varphi(\tau^n, x_{i+1/2}) - \varphi(\tau^n, x_{i-1/2})) \\
&= \Delta t \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} (F(u_{i+1}^n, \kappa) - F(u_i^n, \kappa)) \varphi(\tau^n, x_{i+1/2}) \\
&= \Delta t \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} (F(u_i^n, \kappa) - G_{i-1/2}^n) \varphi(\tau^n, x_{i-1/2}) \\
&\quad + \Delta t \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} (G_{i+1/2}^n - F(u_i^n, \kappa)) \varphi(\tau^n, x_{i+1/2})
\end{aligned}$$

où $\tau_n \in [t^n, t^{n+1}[$ est défini par $\Delta t \varphi(\tau^n, x) = \int_{t^n}^{t^{n+1}} \varphi(t, x) \, dt$. On obtient donc

$$\begin{aligned}
B_\Delta &= - \int_0^{t^N} \int_{\mathbb{R}} F(u_\Delta, \kappa) \partial_x \varphi(t, x) \, dt \, dx \\
&\quad - \Delta t \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} (G_{i-1/2}^n - F(u_i^n, \kappa)) \left(\frac{1}{\Delta x} \int_{M_i} \varphi(t^{n+1}, x) \, dx - \varphi(\tau^n, x_{i-1/2}) \right) \\
&\quad - \Delta t \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} (F(u_i^n, \kappa) - G_{i+1/2}^n) \left(\frac{1}{\Delta x} \int_{M_i} \varphi(t^{n+1}, x) \, dx - \varphi(\tau^n, x_{i+1/2}) \right)
\end{aligned}$$

En utilisant (4.13), c'est-à-dire que $A_\Delta + B_\Delta \leq 0$, on a

$$\begin{aligned}
&- \int_0^{t^N} \int_{\mathbb{R}} |u_\Delta(t, x) - \kappa| \partial_t \varphi(t, x) \, dx \, dt - \int_0^{t^N} \int_{\mathbb{R}} F(u_\Delta, \kappa) \partial_x \varphi(t, x) \, dt \, dx \\
&- \int_{\mathbb{R}} |u_\Delta(0, x) - \kappa| \varphi(0, x) \, dx + \int_{\mathbb{R}} |u_\Delta(t^N, x) - \kappa| \varphi(t^N, x) \, dx \\
&= \Delta t \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} (G_{i-1/2}^n - F(u_i^n, \kappa)) \left(\frac{1}{\Delta x} \int_{M_i} \varphi(t^{n+1}, x) \, dx - \varphi(\tau^n, x_{i-1/2}) \right) \\
&\quad + \Delta t \sum_{n=0}^{N-1} \sum_{i \in \mathbb{Z}} (F(u_i^n, \kappa) - G_{i+1/2}^n) \left(\frac{1}{\Delta x} \int_{M_i} \varphi(t^{n+1}, x) \, dx - \varphi(\tau^n, x_{i+1/2}) \right) \\
&\leq 2 \Delta t N L TV(u_0) C(\Delta t + \Delta x). \tag{4.15}
\end{aligned}$$

En supposant que le rapport $\Delta t / \Delta x$ est constant, on en déduit l'inégalité (4.14). \square

On désire maintenant avoir une véritable estimation de l'erreur commise par le schéma numérique, c'est-à-dire comparer u_Δ et u . La première étape est de comparer u_Δ à une fonction $v \in \mathbf{L}^\infty(\mathbb{R}^+ \times \mathbb{R})$.

Lemme 4.11. Soit $u_0 \in \text{BV}(\mathbb{R})$. Alors, pour tout $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^+ \times \mathbb{R})$ positive symétrique et pour tout $v \in \mathbf{L}^\infty(\mathbb{R}^+ \times \mathbb{R})$, on a pour tout $t^N = N\Delta t$,

$$\begin{aligned} & - \int_0^{t^N} \int_{\mathbb{R}} \int_0^{t^N} \int_{\mathbb{R}} (|u_\Delta(t,x) - v(s,y)| \partial_t + F(u_\Delta(t,x), v(s,y)) \partial_x) \psi(t,x,s,y) dx dt dy ds \\ & - \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} (|u_\Delta(0,x) - v(s,y)| \psi(0,x,s,y) - |u_\Delta(t^N,x) - v(s,y)| \psi(t^N,x,s,y)) dx dy ds \\ & \leq C \Delta x \text{TV}(u_0) t^N |\varphi|_{1,1} \quad (4.16) \end{aligned}$$

où $\psi(t,x,s,y) = \varphi(t-s, x-y)$.

Démonstration. Il suffit, comme dans le dédoublement de variable, de remplacer κ par $v(s,y)$ et $\varphi(t,x)$ par $\psi(t,x,s,y)$ dans l'inégalité (4.14) et d'intégrer pour $s \in [0, t^N]$ et $y \in \mathbb{R}$. Les calculs sont identiques à ceux de la démonstration du lemme 4.10, excepté pour la majoration du terme d'erreur (4.15) pour lequel on utilise :

$$\begin{aligned} & \int_0^{t^N} \int_{\mathbb{R}} \left(\frac{1}{\Delta x} \int_{M_i} \psi(t^{n+1}, x, s, y) dx - \psi(\tau^n, x_{i-1/2}, s, y) \right) dy ds \\ & = \int_0^{t^N} \int_{\mathbb{R}} \left(\frac{1}{\Delta x} \int_{M_i} \varphi(t^{n+1} - s, x - y) dx - \varphi(\tau^n - s, x_{i-1/2} - y) \right) dy ds \\ & \leq (\Delta t + \Delta x) |\varphi|_{1,1}, \end{aligned}$$

ce qui donne bien (4.16). \square

Passons maintenant au résultat principal, dû à Kuznetsov :

Théorème 4.12. Soit $u_0 \in \text{BV}(\mathbb{R})$. Alors, pour presque tout $t > \sqrt{\Delta t}$, le schéma numérique u_Δ et la solution faible entropique vérifient l'inégalité

$$\|u(t, \cdot) - u_\Delta(t, \cdot)\|_{\mathbf{L}^1(\mathbb{R})} \leq \|u_0 - u_\Delta(0, \cdot)\|_{\mathbf{L}^1(\mathbb{R})} + C t \text{TV}(u_0) \sqrt{\Delta t}. \quad (4.17)$$

On appelle ce genre de résultat estimation *a posteriori* car il permet d'estimer l'erreur commise par le schéma numérique, sans connaître la solution (le membre de droite est indépendant de la solution u).

Démonstration. On ne démontrera ce résultat que pour $t = t^N := N\Delta t$.

Tout d'abord, remarquons que la solution faible entropique u vérifie

$$\begin{aligned} & - \int_0^{t^N} \int_{\mathbb{R}} \int_0^{t^N} \int_{\mathbb{R}} (|u_\Delta(t,x) - u(s,y)| \partial_s + F(u_\Delta(t,x), u(s,y)) \partial_y) \psi(t,x,s,y) dx dt dy ds \\ & - \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} (|u_\Delta(t,x) - u(0,y)| \psi(t,x,0,y) - |u_\Delta(t,x) - u(t^N,y)| \psi(t,x,t^N,y)) dx dy ds \\ & \leq 0. \quad (4.18) \end{aligned}$$

Comme φ est symétrique, on a $\partial_t \psi + \partial_s \psi = \partial_x \psi + \partial_y \psi = 0$. Ainsi, si on remplace v par u dans (4.16) puis que l'on additionne (4.16) et (4.18), on obtient

$$\begin{aligned} & \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} (|u_\Delta(t,x) - u(t^N,y)| \psi_{|s=t^N} - |u_\Delta(t,x) - u(0,y)| \psi_{|s=0}) dx dy ds \\ & + \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} (|u_\Delta(t^N,x) - u(s,y)| \psi_{|t=t^N} - |u_\Delta(0,x) - u(s,y)| \psi_{|t=0}) dx dy ds \\ & \leq C \Delta x \text{TV}(u_0) t^N |\varphi|_{1,1} \end{aligned}$$

qui peut aussi se réécrire

$$\begin{aligned}
& \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} (|u_{\Delta}(t,x) - u(t^N,y)| \psi_{|s=t^N} + |u_{\Delta}(t^N,x) - u(t,y)| \psi_{|s=t^N}) dx dy dt \\
& \leq \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} (|u_{\Delta}(t,x) - u(0,y)| \psi_{|s=0} + |u_{\Delta}(0,x) - u(t,y)| \psi_{|s=0}) dx dy dt \\
& \quad + C \Delta x TV(u_0) t^N |\varphi|_{1,1}.
\end{aligned} \tag{4.19}$$

Pour s'approcher de l'estimation finale (4.17), on veut que t (resp. y) et t^N (resp. x) soient « proches à ε près ». Pour cela, on utilise une partition de l'unité ζ , c'est-à-dire que $\zeta \in \mathcal{C}_c^\infty(\mathbb{R})$, $\zeta(z) = \zeta(-z)$, $\text{supp} \zeta \subset [-1, 1]$ et $\int_{\mathbb{R}} \zeta(z) dz = 1$. On définit donc $\zeta_\varepsilon(z) = \zeta(z/\varepsilon)/\varepsilon$ et $\varphi(t,x) = \zeta_\varepsilon(t)\zeta_\varepsilon(x)$, ce qui entraîne $|\varphi|_{1,1} \leq C/\varepsilon$ (et rappelons que $\psi(t,s,x,y) = \varphi(t-s,x-y)$). On a de plus

$$\int_{\mathbb{R}} \zeta_\varepsilon(x-y) dy = 2 \int_0^{t^N} \zeta_\varepsilon(t-t^N) dt = 1. \tag{4.20}$$

Revenons maintenant à l'estimation complète :

$$\begin{aligned}
& \|u_{\Delta}(t^N, \cdot) - u(t^N, \cdot)\|_{\mathbf{L}^1(\mathbb{R})} \\
& \leq 2 \int_0^{t^N} \zeta_\varepsilon(t-t^N) dt \int_{\mathbb{R}} |u_{\Delta}(t^N, x) - u(t^N, x)| dx \\
& \leq 2 \int_0^{t^N} \zeta_\varepsilon(t-t^N) dt \int_{\mathbb{R}} \int_{\mathbb{R}} \zeta_\varepsilon(x-y) dy |u_{\Delta}(t^N, x) - u(t^N, x)| dx
\end{aligned}$$

grâce à (4.20), ce qui donne par définition de ψ

$$\begin{aligned}
& \leq 2 \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} |u_{\Delta}(t^N, x) - u(t^N, x)| \psi_{|s=t^N} dx dy dt \\
& \leq \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} \left(|u_{\Delta}(t^N, x) - u_{\Delta}(t, x)| + |u_{\Delta}(t, x) - u(t^N, y)| + |u(t^N, y) - u(t^N, x)| \right. \\
& \quad \left. + |u_{\Delta}(t^N, x) - u_{\Delta}(t, y)| + |u(t, y) - u(t, x)| + |u(t, x) - u(t^N, x)| \right) \\
& \quad \times \psi_{|s=t^N} dx dy dt \\
& \leq \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} (|u_{\Delta}(t, x) - u(t^N, y)| + |u_{\Delta}(t^N, x) - u_{\Delta}(t, y)|) \psi_{|s=t^N} dx dy dt \\
& + \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} \left(|u_{\Delta}(t^N, x) - u_{\Delta}(t, x)| + |u(t^N, y) - u(t^N, x)| \right. \\
& \quad \left. + |u(t, y) - u(t, x)| + |u(t, x) - u(t^N, x)| \right) \psi_{|s=t^N} dx dy dt
\end{aligned}$$

ce qui grâce à (4.19) devient

$$\begin{aligned}
&\leq \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} (|u_{\Delta}(t,x) - u(0,y)| + |u_{\Delta}(0,x) - u(t,y)|) \psi_{|s=0} dx dy dt \\
&+ C \Delta x TV(u_0) t^N / \varepsilon \\
&+ \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} (|u_{\Delta}(t^N,x) - u_{\Delta}(t,x)| + |u(t^N,y) - u(t,x)| \\
&\quad + |u(t,y) - u(t,x)| + |u(t,x) - u(t^N,x)|) \psi_{|s=t^N} dx dy dt \\
&\leq \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} (|u_{\Delta}(0,x) - u(0,x)| + |u_{\Delta}(0,x) - u(0,x)|) \psi_{|s=0} dx dy dt \\
&+ C \Delta x TV(u_0) t^N / \varepsilon \\
&+ \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} (|u_{\Delta}(t,x) - u_{\Delta}(0,x)| + |u(0,x) - u(0,y)| \\
&\quad + |u(0,x) - u(0,y)| + |u(0,y) - u(t,y)|) \psi_{|s=t^N} dx dy dt \\
&+ \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} (|u_{\Delta}(t^N,x) - u_{\Delta}(t,x)| + |u(t^N,y) - u(t^N,x)| \\
&\quad + |u(t,y) - u(t,x)| + |u(t,x) - u(t^N,x)|) \psi_{|s=t^N} dx dy dt.
\end{aligned}$$

Pour les deux dernières intégrales, on utilise le lemme suivant :

Lemme 4.13. *En reprenant les notations précédentes, on a $\forall (t,x) \in \mathbb{R}^+ \times \mathbb{R}$*

$$\forall \tau > 0, \quad \int_{\mathbb{R}} |u(t+\tau,x) - u(t,x)| dx \leq C\tau TV(u_0), \quad (4.21)$$

$$\forall \theta > 0, \quad \int_{\mathbb{R}} |u(t,x+\theta) - u(t,x)| dx \leq \theta TV(u_0), \quad (4.22)$$

$$\forall \tau > 0, \quad \int_{\mathbb{R}} |u_{\Delta}(t+\tau,x) - u_{\Delta}(t,x)| dx \leq C(\tau + \Delta t) TV(u_0). \quad (4.23)$$

En effet, on peut déduire du lemme 4.13 que

$$\begin{aligned}
&\int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} |u_{\Delta}(t^N,x) - u_{\Delta}(t,x)| \psi_{|s=t^N} dx dy dt \\
&= \int_0^{t^N} \int_{\mathbb{R}} |u_{\Delta}(t^N,x) - u_{\Delta}(t,x)| \int_{\mathbb{R}} \zeta_{\varepsilon}(x-y) dy \zeta_{\varepsilon}(t-t^N) dx dt \\
&= \int_0^{t^N} \int_{\mathbb{R}} |u_{\Delta}(t^N,x) - u_{\Delta}(t^N+\tau,x)| \zeta_{\varepsilon}(\tau) dx d\tau \\
&= \int_0^{\varepsilon} \int_{\mathbb{R}} |u_{\Delta}(t^N,x) - u_{\Delta}(t^N+\tau,x)| \zeta_{\varepsilon}(\tau) dx d\tau \\
&\leq \int_0^{\varepsilon} \zeta_{\varepsilon}(\tau) d\tau \sup_{\tau \in [0,\varepsilon]} \int_{\mathbb{R}} |u_{\Delta}(t^N,x) - u_{\Delta}(t^N+\tau,x)| dx \\
&\leq C(\varepsilon + \Delta t) TV(u_0).
\end{aligned}$$

On aboutit par le même type de calcul aux estimations suivantes :

$$\begin{aligned}
& \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} |u(t,x) - u(t^N,x)| \psi_{|s=t^N} dx dy dt \leq C\varepsilon TV(u_0), \\
& \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} |u(t,y) - u(t,x)| \psi_{|s=t^N} dx dy dt \leq \varepsilon TV(u_0), \\
& \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} |u_{\Delta}(t,x) - u_{\Delta}(0,x)| \psi_{|s=0} dx dy dt \leq C(\varepsilon + \Delta t) TV(u_0), \\
& \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} |u(0,x) - u(0,y)| \psi_{|s=0} dx dy dt \leq \varepsilon TV(u_0), \\
& \int_0^{t^N} \int_{\mathbb{R}} \int_{\mathbb{R}} |u(0,y) - u(t,y)| \psi_{|s=0} dx dy dt \leq C\varepsilon \Delta t TV(u_0).
\end{aligned}$$

On obtient donc l'estimation

$$\|u(t, \cdot) - u_{\Delta}(t, \cdot)\|_{L^1(\mathbb{R})} \leq \|u_0 - u_{\Delta}(0, \cdot)\|_{L^1(\mathbb{R})} + C t^N TV(u_0) [\varepsilon + \Delta t + \Delta t/\varepsilon]$$

qui est optimale pour $\varepsilon = \sqrt{\Delta t}$. \square

Concernant le Lemme 4.13, on ne le démontrera pas ici. Il peut être déduit des estimations uniformes BV qui avait été obtenue sur le schéma numérique et donc qui sont préservés pour la solution, à la limite.

En pratique, l'estimation (4.17) est optimale dès lors que le flux f est linéaire (ou s'il existe des intervalles sur lesquels il est linéaire). Cependant, si f'' est uniformément non nul, l'ordre généralement mesuré est $\mathcal{O}(\Delta x)$ (on ne sait pas le démontrer en tout généralité).

Remarque 9. Résumons les différentes étapes ayant mené à l'estimation *a posteriori* (4.17).

1. Tout d'abord, nous avons écrit les inégalités d'entropie de Kruzhkov $(\eta_{\kappa}, F_{\kappa})$ vérifiées par le schéma : c'est l'inégalité (4.14), qui comporte donc un terme d'erreur en second membre (qui tend vers 0 à convergence, ce qui permet notamment de démontrer la convergence du schéma vers la solution entropique).
2. Ensuite, on effectue le dédoublement de variable, ce qui permet de comparer l'approximation numérique à la solution entropique avec des variables différentes. En prenant comme fonction test $\psi(t, x, s, y)$ une fonction test symétrique, seuls les termes de bord en temps (c'est-à-dire en 0 et t^N) persistent (voir (4.19)).
3. On choisit ensuite la fonction test $\psi(t, x, s, y)$ comme une approximation symétrique de $\delta_0(x-y)\delta_0(t-s)$, à ε près. En utilisant la régularité BV de l'approximation numérique et de la solution numérique, on peut majorer les écarts en $|x-y|$ et $|t-s|$ en fonction de ε .
4. Enfin, on ajuste ε en fonction de Δt pour aboutir à une erreur minimale.

Ce procédé n'est pas restreint aux estimations de l'erreur produite par le schéma numérique. On peut aussi l'étendre à d'autres approximations : approximation visqueuse, approximation du flux f ... En fait, cette approche peut même être systématisée.

4.3 Cas multidimensionnel

On a évoqué précédemment le fait que le problème de Cauchy

$$\begin{cases} \partial_t u + \operatorname{div}_x f(u) = 0, & t > 0, x \in \mathbb{R}^d, \\ u(0, x) = u_0(x), & x \in \mathbb{R}^d, \end{cases} \quad (4.24)$$

où $f : \mathbb{R} \rightarrow \mathbb{R}^d$, admettait aussi une et une seule solution dans $\mathbf{L}^\infty(\mathbb{R}^+ \times \mathbb{R}^d)$. On peut démontrer l'existence d'une solution en utilisant l'approximation visqueuse et l'unicité par le dédoublement de variable de Kruzhkov. Voyons comment définir des schémas VF pour le cas multidimensionnel.

On se donne un maillage \mathcal{M} de \mathbb{R}^d , c'est-à-dire une partition d'ouverts polygonaux disjoints (mailles). Si K et L sont deux mailles voisines (on pourra noter $L \in \mathcal{V}(K)$ et inversement $L \in \mathcal{V}(K)$), on définit l'arête en 2D ou la face en 3D $e_{KL} = \bar{K} \cap \bar{L}$ ainsi que la normale unitaire n_{KL} orientée de K vers L .

On définit alors

$$u_K^0 = \frac{1}{|K|} \int_K u_0(x) dx.$$

On intègre maintenant (4.24) sur $(t^n, t^{n+1}) \times K$, ce qui donne

$$\int_K u(t^{n+1}, x) dx - \int_K u(t^n, x) dx + \int_{t^n}^{t^{n+1}} \sum_{L \in \mathcal{V}(K)} \int_{e_{KL}} f(u(t, x)) \cdot n_{KL} d\gamma dt = 0.$$

On en déduit alors le schéma VF suivant

$$u_K^{n+1} = u_K^n - \frac{\Delta t}{|K|} \sum_{L \in \mathcal{V}(K)} |e_{KL}| g(u_K^n, u_L^n; n_{KL}) \quad (4.25)$$

où le flux numérique g est une fonction définie de $\mathbb{R}^2 \times \mathbf{S}^{d-1}$ dans \mathbb{R} . Ce flux numérique doit alors vérifier les propriétés de base suivantes :

- conservation : $g(u, v; n) = -g(v, u; -n)$ pour tout $(u, v; n) \in \mathbb{R}^2 \times \mathbf{S}^{d-1}$,
- consistance : $g(u, u; n) = f(u) \cdot n$ pour tout $(u, n) \in \mathbb{R} \times \mathbf{S}^{d-1}$,
- monotonie : g est croissante par rapport à sa première variable et décroissante par rapport à la deuxième.

Deux cas très différents en terme d'analyse se présente.

4.3.1 Maillage cartésien

On appelle maillage cartésien un maillage dont les mailles sont des rectangles en 2D ou des parallélépipèdes en 3D. On peut alors réutiliser l'analyse faite en 1D. Pour simplifier la présentation, regardons le cas 2D. On considère le problème de Cauchy

$$\begin{cases} \partial_t u(t, x, y) + \partial_x f_1(u(t, x, y)) + \partial_y f_2(u(t, x, y)) = 0, & t > 0, (x, y) \in \mathbb{R}^2, \\ u(0, x, y) = u_0(x, y), & (x, y) \in \mathbb{R}^2, \end{cases} \quad (4.26)$$

où $u_0 \in \mathbf{L}^\infty(\mathbb{R}^2) \cap \mathbf{L}^1(\mathbb{R}^2) \cap \mathbf{BV}(\mathbb{R}^2)$ et $f_1, f_2 \in \mathcal{C}^2(\mathbb{R}; \mathbb{R})$ sont lipschitziennes. Soit Δt le pas de temps et Δx et Δy les pas d'espace. On introduit les points $x_{i+1/2} = (i + 1/2)\Delta x$, $y_{j+1/2} = (j + 1/2)\Delta y$ et les mailles $M_{i,j} = [x_{i-1/2}, x_{i+1/2}] \times [y_{j-1/2}, y_{j+1/2}]$, pour tout $(i, j) \in \mathbb{Z}^2$. On définit

$$u_{i,j}^0 = \frac{1}{\Delta x \Delta y} \int_{M_{i,j}} u_0(t, x, y) dx dy \quad \forall (i, j) \in \mathbb{Z}^2.$$

On construit alors le schéma volumes finis suivant, $\forall n \in \mathbb{N}, (i, j) \in \mathbb{Z}^2$:

$$u_{i,j}^{n+1} = u_{i,j}^n - \frac{\Delta t}{\Delta x} \left(g_1(u_{i,j}^n, u_{i+1,j}^n) - g_1(u_{i-1,j}^n, u_{i,j}^n) \right) - \frac{\Delta t}{\Delta y} \left(g_2(u_{i,j}^n, u_{i,j+1}^n) - g_2(u_{i,j-1}^n, u_{i,j}^n) \right) \quad (4.27)$$

où $g_k \in \mathcal{C}^1(\mathbb{R}^2; \mathbb{R})$ est un flux numérique lipschitzien (avec la même constante de Lipschitz que f_k), consistant avec f_k et monotone, $k = 1, 2$. On utilisera aussi la forme condensée

$$u_{i,j}^{n+1} = G(u_{i,j}^n, u_{i-1,j}^n, u_{i+1,j}^n, u_{i,j-1}^n, u_{i,j+1}^n). \quad (4.28)$$

On a alors le résultat suivant :

Lemme 4.14. *Sous la condition CFL*

$$\Delta t \leq \frac{\min(\Delta x, \Delta y)}{2(L_1 + L_2)} \quad (4.29)$$

où L_1 et L_2 sont les constantes de Lipschitz de g_1 et g_2 , le schéma (4.28) est monotone, c'est-à-dire que G est une fonction croissante par rapport à toutes ses variables.

Démonstration. La démonstration est la même qu'en 2D. On a tout d'abord grâce à la monotonie de g_1 et g_2 :

$$\begin{aligned} \partial_{\alpha_1} G(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) &= 1 - \frac{\Delta t}{\Delta x} \partial_1 g_1(\alpha_1, \alpha_3) + \frac{\Delta t}{\Delta x} \partial_2 g_1(\alpha_2, \alpha_1) \\ &\quad - \frac{\Delta t}{\Delta x} \partial_1 g_2(\alpha_1, \alpha_5) + \frac{\Delta t}{\Delta x} \partial_2 g_2(\alpha_4, \alpha_1) \\ &= 1 - \frac{\Delta t}{\Delta x} |\partial_1 g_1(\alpha_1, \alpha_3)| - \frac{\Delta t}{\Delta x} |\partial_2 g_1(\alpha_2, \alpha_1)| \\ &\quad - \frac{\Delta t}{\Delta x} |\partial_1 g_2(\alpha_1, \alpha_5)| + \frac{\Delta t}{\Delta x} |\partial_2 g_2(\alpha_4, \alpha_1)| \\ &\geq 1 - 2 \frac{\Delta t}{\Delta x} L_1 - \frac{\Delta t}{\Delta y} L_2, \end{aligned}$$

Par ailleurs, toujours en utilisant la monotonie de g_1 et g_2 , on a :

$$\begin{aligned} \partial_{\alpha_2} G(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) &= \frac{\Delta t}{\Delta x} \partial_1 g_1(\alpha_2, \alpha_1) = \frac{\Delta t}{\Delta x} |\partial_1 g_1(\alpha_2, \alpha_1)| \\ \partial_{\alpha_3} G(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) &= -\frac{\Delta t}{\Delta x} \partial_2 g_1(\alpha_1, \alpha_3) = \frac{\Delta t}{\Delta x} |\partial_2 g_1(\alpha_1, \alpha_3)| \\ \partial_{\alpha_4} G(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) &= \frac{\Delta t}{\Delta y} \partial_1 g_2(\alpha_4, \alpha_1) = \frac{\Delta t}{\Delta y} |\partial_1 g_2(\alpha_4, \alpha_1)| \\ \partial_{\alpha_5} G(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5) &= -\frac{\Delta t}{\Delta y} \partial_2 g_2(\alpha_1, \alpha_5) = \frac{\Delta t}{\Delta y} |\partial_2 g_2(\alpha_1, \alpha_5)|. \end{aligned}$$

Donc G est inconditionnellement croissante par rapport à ses quatre dernières variables et croissante par rapport à sa première variable sous la condition (4.29). \square

Proposition 4.15. *Soit A et B deux constantes réelles telles que $A \leq u_0 \leq B$ presque partout et on suppose que la condition CFL (4.29) est vérifiée. Alors*

$$\forall n \in \mathbb{N}, \forall (i, j) \in \mathbb{Z}^2, \quad A \leq u_{i,j}^n \leq B$$

et

$$\forall n \in \mathbb{N}, \quad |U^{n+1}|_{\text{BV}(\mathbb{R}^2)} \leq |u_0|_{\text{BV}(\mathbb{R}^2)}$$

où $U^n(x, y) = \sum_{i,j} u_{i,j}^n \mathbf{1}_{M_{i,j}}(x, y)$.

Démonstration. Puisque $A \leq u_0 \leq B$, on a, $\forall (i, j) \in \mathbb{Z}^2, A \leq u_{i,j}^0 \leq B$. Supposons maintenant que $\forall (i, j) \in \mathbb{Z}^2, A \leq u_{i,j}^n \leq B$, montrons que l'encadrement est vrai au rang $n+1$. Comme G est croissante par rapport à toute ses variables (on a supposé que (4.29) est vérifiée), alors

$$\begin{aligned} G(A, A, A, A, A) &\leq G(u_{i,j}^n, u_{i-1,j}^n, u_{i+1,j}^n, u_{i,j-1}^n, u_{i,j+1}^n) \\ \text{et } G(B, B, B, B, B) &\geq G(u_{i,j}^n, u_{i-1,j}^n, u_{i+1,j}^n, u_{i,j-1}^n, u_{i,j+1}^n). \end{aligned}$$

Or, $G(A, A, A, A, A) = A$ et $G(B, B, B, B, B) = B$, donc

$$A \leq G(u_{i,j}^n, u_{i-1,j}^n, u_{i+1,j}^n, u_{i,j-1}^n, u_{i,j+1}^n) \leq B \quad \Leftrightarrow \quad A \leq u_{i,j}^{n+1} \leq B.$$

On rappelle maintenant que la semi-norme BV peut s'écrire ainsi

$$|U^n|_{\text{BV}(\mathbb{R}^2)} = \Delta y \sum_{(i,j) \in \mathbb{Z}^2} |u_{i+1,j}^n - u_{i,j}^n| + \Delta x \sum_{(i,j) \in \mathbb{Z}^2} |u_{i,j+1}^n - u_{i,j}^n|.$$

Considérons tout d'abord le premier terme :

$$\begin{aligned} \sum_{(i,j) \in \mathbb{Z}^2} |u_{i+1,j}^{n+1} - u_{i,j}^{n+1}| &= \sum_{(i,j) \in \mathbb{Z}^2} |G(u_{i+1,j}^n, u_{i,j}^n, u_{i+2,j}^n, u_{i+1,j-1}^n, u_{i+1,j+1}^n) \\ &\quad - G(u_{i,j}^n, u_{i-1,j}^n, u_{i+1,j}^n, u_{i,j-1}^n, u_{i,j+1}^n)|. \end{aligned}$$

On va utiliser les identités $|a - b| = [a - b]^+ + [b - a]^+$ et $[a - b]^+ = \max(a, b) - b$ ainsi que la conservation du schéma numérique, qui assure pour toute suite réelle $(\alpha_{i,j})_{(i,j) \in \mathbb{Z}^2}$ que

$$\sum_{i,j \in \mathbb{Z}^2} G(\alpha_{i,j}, \alpha_{i-1,j}, \alpha_{i+1,j}, \alpha_{i,j-1}, \alpha_{i,j+1}) = \sum_{i,j \in \mathbb{Z}^2} \alpha_{i,j}. \quad (4.30)$$

Ainsi, si on considère deux suites $(\alpha_{i,j})_{(i,j) \in \mathbb{Z}^2}$ et $(\beta_{i,j})_{(i,j) \in \mathbb{Z}^2}$, on a

$$\begin{aligned} &\sum_{(i,j) \in \mathbb{Z}^2} [G(\alpha_{i,j} \dots) - G(\beta_{i,j} \dots)]^+ \\ &= \sum_{(i,j) \in \mathbb{Z}^2} \max(G(\alpha_{i,j} \dots), G(\beta_{i,j} \dots)) - \sum_{(i,j) \in \mathbb{Z}^2} G(\beta_{i,j} \dots) \quad (\text{grâce à (4.30)}) \\ &\leq \sum_{(i,j) \in \mathbb{Z}^2} G(\max(\alpha_{i,j}, \beta_{i,j}) \dots) - \sum_{(i,j) \in \mathbb{Z}^2} \beta_{i,j} \quad (G \text{ est croissante}) \\ &\leq \sum_{(i,j) \in \mathbb{Z}^2} \max(\alpha_{i,j}, \beta_{i,j}) - \sum_{(i,j) \in \mathbb{Z}^2} \beta_{i,j} \quad (\text{grâce à (4.30)}) \\ &\leq \sum_{(i,j) \in \mathbb{Z}^2} [\alpha_{i,j} - \beta_{i,j}]^+. \end{aligned}$$

On en déduit donc

$$\begin{aligned}
\sum_{(i,j) \in \mathbb{Z}^2} |G(\alpha_{i,j,\dots}) - G(\beta_{i,j,\dots})| &= \sum_{(i,j) \in \mathbb{Z}^2} [G(\alpha_{i,j,\dots}) - G(\beta_{i,j,\dots})]^+ \\
&\quad + \sum_{(i,j) \in \mathbb{Z}^2} [G(\beta_{i,j,\dots}) - G(\alpha_{i,j,\dots})]^+ \\
&\leq \sum_{(i,j) \in \mathbb{Z}^2} [\alpha_{i,j} - \beta_{i,j}]^+ + \sum_{(i,j) \in \mathbb{Z}^2} [\beta_{i,j} - \alpha_{i,j}]^+ \\
&\leq \sum_{(i,j) \in \mathbb{Z}^2} |\alpha_{i,j} - \beta_{i,j}|.
\end{aligned}$$

Par conséquent, on obtient en prenant $\alpha_{i,j} = u_{i+1,j}^n$ et $\beta_{i,j} = u_{i,j}^n$

$$\begin{aligned}
\sum_{(i,j) \in \mathbb{Z}^2} |u_{i+1,j}^{n+1} - u_{i,j}^{n+1}| &= \sum_{(i,j) \in \mathbb{Z}^2} |G(u_{i+1,j}^n, u_{i,j}^n, u_{i+2,j}^n, u_{i+1,j-1}^n, u_{i+1,j+1}^n) \\
&\quad - G(u_{i,j}^n, u_{i-1,j}^n, u_{i+1,j}^n, u_{i,j-1}^n, u_{i,j+1}^n)| \\
&\leq \sum_{(i,j) \in \mathbb{Z}^2} |u_{i+1,j}^n - u_{i,j}^n|,
\end{aligned}$$

ce qui implique donc l'estimation BV. \square

On peut remarquer que cette démonstration revient finalement à appliquer le lemme de Crandall-Tartar 3.19. On aurait pu cependant faire comme en 1D et utiliser la décomposition de $u_{i,j}^n$ en combinaison convexe.

Pour aboutir à la convergence vers la solution faible entropique associée à (4.26), il reste tout d'abord à obtenir une borne $BV(\mathbb{R}^+ \times \mathbb{R}^2)$, ce qui se fait comme dans le cas unidimensionnel. Ensuite, on peut appliquer le théorème de Helly et conclure à la convergence du schéma, à une sous-suite près.

Pour terminer, il faut obtenir des inégalités d'entropie discrètes, ce qui se démontre aisément grâce à la monotonie du schéma (lemme 4.14), puis passer à la limite en imitant le théorème de Lax-Wendroff (là aussi, c'est possible en suivant directement le raisonnement unidimensionnel).

4.3.2 Maillage non structuré

Dans le cas de maillages généraux, l'analyse s'avère beaucoup plus difficile. Tout d'abord, il faut définir les maillages que l'on peut considérer. La forme des mailles peut être quelconque, mais il est absolument nécessaire que lors du passage à la limite, les mailles ne « dégénère » pas, c'est-à-dire que si $h = \sup_{K \in \mathcal{M}} (\text{diam}(K))$, il existe alors $\alpha > 0$ tel que

$$\forall K \in \mathcal{M}, \quad \alpha |\partial K| h \leq h^d \leq \frac{1}{\alpha} |K|.$$

Concernant maintenant la convergence, la difficulté est « simplement » le manque de compacité : on ne peut pas obtenir d'estimation BV. Prenons par exemple le cas d'un maillage triangulaire dont les sommets sont les points $x_{i,j} = (i\Delta x, j\Delta y)$ et les mailles sont les triangles $M_{i,j}^+ = (x_{i,j}, x_{i+1,j}, x_{i,j+1})$ et $M_{i,j}^- = (x_{i,j}, x_{i-1,j}, x_{i,j-1})$. On considère alors l'équation

$$\partial_t u(t, x, y) + \partial_x u(t, x, y) = 0$$

et la donnée initiale $u_0(x, y) = \mathbf{1}_{\mathbb{R}^+ \times \mathbb{R}}(x, y)$. Ainsi, $u_K^0 = 1$ si $K = M_{i,j}^+$ avec $i \geq 0$ ou $K = M_{i,j}^-$ avec $i \geq 1$ et $u_K^0 = 0$ pour les autres mailles K . Après la première itération, seules les valeurs aux mailles $M_{0,j}^\pm$ sont modifiées et on a $u_{M_{0,j}^\pm}^1 \in]0, 1[$, pour peu que l'on considère un schéma monotone. La variation totale suivant la direction x , pour tout y , ne change donc pas. Par contre, pour $x \in]0, \Delta x[$, l'approximation $U^1(x, y)$ prend pour valeur alternativement $u_{M_{0,j}^+}^1$ et $u_{M_{1,j}^-}^1 = 0$. La variation totale suivant la direction y est donc infinie.

Les conséquences de ce manque de compacité sont lourdes : le passage à la limite sera (très...) faible. Plus précisément, on ne peut plus appliquer le théorème de Helly (ou un résultat du même genre), on doit utiliser la notion de convergence non linéaire faible \star :

Définition 4.16. Soit Ω un ouvert de \mathbb{R}^N , $N \geq 1$, et on considère $(u_m)_{m \in \mathbb{N}} \subset \mathbf{L}^\infty(\Omega)$ and $\mu \in \mathbf{L}^\infty(\Omega \times (0, 1))$. La suite $(u_m)_{m \in \mathbb{N}}$ converge vers μ au sens *non linéaire faible \star* si

$$\int_{\Omega} \theta(u_m(y)) \varphi(y) dy \xrightarrow{m \rightarrow \infty} \int_{\Omega} \int_0^1 \theta(\mu(y, \alpha)) \varphi(y) dy d\alpha$$

pour toute fonction $\varphi \in \mathbf{L}^1(\Omega)$ et tout $\theta \in \mathcal{C}_c^\infty(\Omega)$.

Cette notion est en fait une interprétation de la convergence au sens des mesures de Young. Elle permet de pouvoir passer à la limite en utilisant simplement une estimation \mathbf{L}^∞ :

Théorème 4.17. Soit Ω un ouvert de \mathbb{R}^N , $N \geq 1$. On considère une suite $(u_m)_{m \in \mathbb{N}}$ bornée dans $\mathbf{L}^\infty(\Omega)$. Alors, on peut extraire une sous-suite de $(u_m)_{m \in \mathbb{N}}$ qui converge au sens *non linéaire faible \star* , vers une fonction $\mu \in \mathbf{L}^\infty(\Omega \times (0, 1))$.

De plus, la convergence est forte (au sens $\mathbf{L}_{loc}^1(\Omega)$) si et seulement si la limite *non linéaire faible \star* μ de $(u_m)_{m \in \mathbb{N}}$ est indépendante de α .

Revenons maintenant aux estimations que l'on peut obtenir dans le cas de maillages non structurés. On réécrit de nouveau le schéma (4.25) sous la forme

$$u_K^{n+1} = G(u_K^n, (u_L^n)_{L \in \mathcal{V}(K)}). \quad (4.31)$$

On a à nouveau le résultat suivant :

Lemme 4.18. *Sous la condition CFL*

$$\Delta t \leq \frac{\alpha^2 h}{L_g} \quad (4.32)$$

où L_g est la constante de Lipschitz de $g(\cdot, v; n)$ (uniforme en $v \in \mathbb{R}$ et $n \in \mathbf{S}^{d-1}$), le schéma (4.31) est monotone, c'est-à-dire que G est une fonction croissante par rapport à toutes ses variables.

Démonstration. Pour cette démonstration, plusieurs options sont possibles. On peut utiliser une extension de la forme incrémentale (4.8). On va ici procéder comme dans la démonstration de la proposition 4.3, en supposant que g est dérivable. Comme g est décroissante par rapport à sa deuxième variable, on a pour $m > 1$

$$\partial_m G(u_K, (u_L)_{L \in \mathcal{V}(K)}) = - \frac{\Delta t |e_{KL_m}|}{|K|} \partial_2 g(u_K, u_{L_m}; n_{KL_m}) \geq 0,$$

L_m étant la m -ième maille voisine de la maille K . On a de plus

$$\begin{aligned}\partial_1 G(u_0, (u_L)_{L \in \mathcal{V}(K)}) &= 1 - \frac{\Delta t}{|K|} \sum_{L \in \mathcal{V}(K)} |e_{KL}| \partial_1 g(u_K, u_L; n_{KL}) \\ &\geq 1 - \frac{\Delta t}{|K|} \sum_{L \in \mathcal{V}(K)} |e_{KL}| L_g \\ &\geq 1 - \frac{\Delta t}{|K|} |\partial K| L_g\end{aligned}$$

qui est bien positif ou nul sous la condition (4.32). \square

On peut donc en déduire :

Proposition 4.19. *Soit A et B deux constantes réelles telles que $A \leq u_0 \leq B$ presque partout et on suppose que la condition CFL (4.32) est vérifiée. Alors*

$$\forall n \in \mathbb{N}, \forall (i, j) \in \mathbb{Z}^2, \quad A \leq u_{i,j}^n \leq B. \quad (4.33)$$

De plus, toujours sous la condition CFL (4.32), le schéma (4.31) vérifie les inégalités d'entropie discrètes suivantes :

$$\forall \kappa \in \mathbb{R}, \quad |u_K^{n+1} - \kappa| - |u_K^n - \kappa| + \frac{\Delta t}{|K|} \sum_{L \in \mathcal{V}(K)} |e_{KL}| G_\kappa(u_K^n, u_L^n; n_{KL}) \leq 0 \quad (4.34)$$

où $G_\kappa(u_K^n, u_L^n; n_{KL}) = g(u_K^n \top \kappa, u_L^n \top \kappa; n_{KL}) - g(u_K^n \perp \kappa, u_L^n \perp \kappa; n_{KL})$.

Démonstration. La borne L^∞ se déduit directement de la monotonie et de la préservation par le schéma des solutions constantes.

Concernant la démonstration des inégalités d'entropie (4.34), il suffit de suivre exactement le raisonnement de la démonstration en 1D (voir lemme 4.8). \square

Passons maintenant à la convergence. Il s'agit d'une part de montrer que la fonction constante par maille définie par le schéma numérique vérifie des inégalités d'entropie du type (3.23) avec un terme de reste (comme dans le lemme 4.10).

Chapitre 5

Systèmes de lois de conservation

On vient de voir que l'analyse des équations scalaires non linéaires est très différente de celle des équations linéaires. Évidemment, il en est de même dans le cas des systèmes.

On va s'intéresser dans ce chapitre à l'étude du problème de Cauchy suivant :

$$\begin{cases} \partial_t U + \partial_x f(U) = 0, & t > 0, x \in \mathbb{R}, \\ U(0, x) = U_0(x), & x \in \mathbb{R}, \end{cases} \quad (5.1)$$

où $U : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}^N$ et où le flux vérifie $f \in \mathcal{C}^2(\mathbb{R}^N; \mathbb{R}^N)$, $N > 1$. En général, la solution ne balaye pas tout \mathbb{R}^N , mais seulement un sous-ensemble de \mathbb{R}^N . On notera $\Omega \subset \mathbb{R}^N$ l'espace des états, c'est-à-dire l'ensemble (convexe) des valeurs de U physiquement admissibles.

On utilisera aussi la notation $U = (u_1, \dots, u_N)^T$ et $f = (f_1, \dots, f_N)^T$, qui permet d'avoir la forme développée

$$\begin{cases} \partial_t u_1 + \partial_x f_1(u_1, \dots, u_N) = 0, \\ \dots \\ \partial_t u_N + \partial_x f_N(u_1, \dots, u_N) = 0. \end{cases}$$

5.1 Hyperbolicité et entropie

Par analogie avec le cas linéaire, on propose les définitions suivantes :

Définition 5.1. On dit que le système (5.1) est hyperbolique si la matrice $\nabla f(U)$ est diagonalisable dans \mathbb{R} , pour tout $U \in \Omega$.

On dit que le système (5.1) est strictement hyperbolique si $\nabla f(U)$ est diagonalisable dans \mathbb{R} et si ses valeurs propres sont distinctes, pour tout $U \in \Omega$.

On peut remarquer que ces propriétés sont invariantes par changement de variable, même non linéaire. Soit $V = \varphi(U)$ et $B(V) = \nabla \varphi(U) \nabla f(U) \nabla \varphi(U)^{-1}$, on a alors

$$\partial_t V + B(V) \partial_x V = 0.$$

Si le changement de variable φ permet de diagonaliser le système (5.1), on a

$$\begin{cases} \partial_t v_1 + \lambda_1(V) \partial_x v_1 = 0, \\ \dots \\ \partial_t v_N + \lambda_N(V) \partial_x v_N = 0. \end{cases} \quad (5.2)$$

Contrairement au cas linéaire, les équations restent couplées et on ne peut pas résoudre directement le système.

De plus, si il existe un changement de variable tel que le système (5.1) puisse se réécrire

$$A_0(V)\partial_t V + A_1(V)\partial_x V = 0 \quad (5.3)$$

où A_0 est une matrice symétrique définie positive et A_1 une matrice symétrique, alors on dit que le système est *symétrisable*. On va voir dans la suite comment relier la symétrisabilité d'un système avec la notion d'entropie.

Tout d'abord, introduisons l'entropie pour les systèmes de lois de conservation.

Définition 5.2. On dit que (η, F) est un couple entropie-flux d'entropie pour (5.1) si $\eta \in \mathcal{C}^2(\mathbb{R}^N; \mathbb{R})$ est une fonction strictement convexe (au sens où la matrice $\nabla^2 \eta(U)$ est symétrique définie positive) et si $\nabla F(U)^T = \nabla \eta(U)^T \nabla f(U)$ pour tout $U \in \Omega$.

Ainsi, les solutions régulières de (5.1) vérifient

$$\partial_t \eta(U) + \partial_x F(U) = 0.$$

Contrairement au cas scalaire, l'existence d'un couple entropie-flux d'entropie n'est pas toujours assuré, bien que ce soit souvent le cas pour les systèmes issus de la physique.

Revenons maintenant au problème de la symétrisation du système (5.1). Une première étape est de trouver une variable V telle qu'il existe une matrice $A_0(V)$ symétrique définie positive. Pour cela, on cherche à obtenir une matrice A_0 qui est la matrice hessienne d'une fonction strictement convexe de V . Pour cela, on prend $V = \eta'(U)$, qui est la *variable entropique*. Comme η est strictement convexe, cela définit un difféomorphisme. Notons Φ l'application inverse de η' , donc $\Phi(V) = U$. On introduit maintenant la transformée de Legendre de l'entropie :

$$\begin{aligned} \eta^*(V) &= \sup_{U \in \Omega} (U^T V - \eta(U)) \\ &= \Phi(V)^T V - \eta(\Phi(V)) \end{aligned} \quad (5.4)$$

puisque l'entropie est strictement convexe. On peut tout d'abord vérifier que $\nabla \eta^*(V) = \Phi(V)$ et de plus que η^* est une fonction strictement convexe de V . On a donc

$$\partial_t U = \partial_t \Phi(V) = \nabla \Phi(V)^T \partial_t V = \nabla^2 \eta^*(V) \partial_t V.$$

De même, on définit

$$F^*(V) = f(\Phi(V))^T V - F(\Phi(V))$$

(attention, malgré l'utilisation de la même notation, F^* n'est pas la transformée de Legendre de F puisque qu'il n'y a pas ici de propriété de convexité). On peut là aussi vérifier que $\nabla F^*(V) = f(\Phi(V))$, donc

$$\partial_x f(U) = \partial_x f(\Phi(V)) = \partial_x \nabla F^*(V) = \nabla^2 F^*(V) \partial_x V.$$

On obtient au final

$$\nabla^2 \eta^*(V) \partial_t V + \nabla^2 F^*(V) \partial_x V = 0 \quad (5.5)$$

où $V = \eta'(U)$. Comme $\nabla^2 \eta^*(V)$ est une matrice définie positive et $\nabla^2 F^*(V)$ une matrice symétrique, la forme (5.5) assure donc que le système (5.1) est symétrisable.

D'autre part, si dans le système (5.3) on suppose que $A_0(V)$ et $A_1(V)$ sont symétriques, alors il existe deux fonctions η^* et F^* telles que $\nabla^2 \eta^*(V) = A_0(V)$ et $\nabla^2 F^*(V) = A_1(V)$. On définit alors

$$\begin{aligned}\eta(U) &= U^T \Psi(U) - \eta^*(\Psi(U)) \\ F(U) &= f(U)^T \Psi(U) - F^*(\Psi(U))\end{aligned}$$

où Ψ est définie par $V = \Psi(U)$.

Théorème 5.3. *Si un système de loi de conservation admet un couple entropie-flux d'entropie, alors il est symétrisable.*

Il existe en fait une autre manière de démontrer ce théorème, sans passer par la transformée de Legendre de l'entropie. En dérivant la caractérisation $\nabla F^T = \nabla \eta^T \nabla f$ et sa transposée $\nabla F = \nabla f^T \nabla \eta$, on a

$$\begin{aligned}\nabla^2 F &= \nabla^2 \eta \nabla f + \nabla \eta^T \nabla^2 f \\ \nabla^2 F &= \nabla^2 f \nabla \eta + \nabla f^T \nabla^2 \eta,\end{aligned}$$

ce qui, après soustraction, donne

$$\nabla^2 \eta \nabla f = \nabla f^T \nabla^2 \eta. \quad (5.6)$$

Si maintenant on applique la matrice $\nabla^2 \eta(U)$ à (5.1), on obtient

$$\nabla^2 \eta(U) \partial_t U + \nabla^2 \eta(U) \nabla f(U) \partial_x U = 0.$$

On obtient donc une forme symétrique puisque $\nabla^2 \eta(U)$ est symétrique définie positive et $\nabla^2 \eta(U) \nabla f(U)$ est symétrique grâce à (5.6).

Remarque 10. On peut montrer que si la relation (5.6) est vraie, alors il existe un flux d'entropie F vérifiant $\nabla F^T = \nabla \eta^T \nabla f$ par application du lemme de Poincaré.

5.2 Solutions faibles

On a vu que le cas non linéaire impose de considérer des solutions admettant des discontinuités, quelle que soit la régularité de la donnée initiale. Il en est de même dans le cas des système, mais le temps d'apparition d'une discontinuité est plus difficile à calculer puisque les caractéristiques sont toutes couplées (voir l'équation (5.2) pour s'en convaincre).

En calquant ce qui est fait dans le cas scalaire, on introduit :

Définition 5.4. Soit $u_0 \in \mathbf{L}^\infty(\mathbb{R})^N$. On appelle solution faible de (5.1) une fonction $U \in \mathbf{L}^\infty(\mathbb{R}^+ \times \mathbb{R})^N$ vérifiant pour tout $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^+ \times \mathbb{R})^N$

$$\int_{\mathbb{R}^+} \int_{\mathbb{R}} (U(t,x) \partial_t \varphi(t,x) + f(U(t,x)) \partial_x \varphi(t,x)) dx dt + \int_{\mathbb{R}} U_0(x) \varphi(0,x) dx = 0. \quad (5.7)$$

De nouveau, une solution classique de (5.1) est aussi solution faible et on peut en déduire les relations de saut admissibles le long des discontinuités :

Proposition 5.5. Soit la courbe $\Gamma = \{(t, x) \in \mathbb{R}^+ \times \mathbb{R}, x = \sigma(t)\}$ où $\sigma \in \mathcal{C}^1(\overline{\mathbb{R}^+})$, coupant un ouvert $\Omega \subset \mathbb{R}^+ \times \mathbb{R}$ et soient $\Omega_- = \{(t, x) \in \Omega, x < \Gamma(t)\}$ et $\Omega_+ = \{(t, x) \in \Omega, x > \Gamma(t)\}$. On considère une fonction $U \in \mathcal{C}^1(\overline{\Omega_-})^N \cap \mathcal{C}^1(\overline{\Omega_+})^N$. Alors U est une solution faible de (5.1) si et seulement si

$$-\sigma'(U^+ - U^-) + (f(U^+) - f(U^-)) = 0, \quad (5.8)$$

où $U^\pm(t) = U(t, \sigma(t)^\pm)$ sont les limites de U de part et d'autre de la courbe Γ et U est une solution classique de (5.1) sur $\Omega \setminus \Gamma$. On appelle le système d'équations (5.8) les relations de saut de Rankine-Hugoniot.

Démonstration. La démonstration est strictement identique au cas scalaire. \square

Dans le cas scalaire, étant donné un état constant à gauche de la discontinuité, on obtient une famille à un paramètre d'états à droite joignables par les relations de Rankine-Hugoniot. Dans le cas d'un système, si on suppose que l'état de gauche est connu, le système (5.8) est composé de N équations et contient $N + 1$ inconnues (la vitesse de la discontinuité et l'état de droite). On aboutit donc là aussi à une famille à un paramètre.

5.3 Ondes et caractère non linéaire

On a vu dans le cas des systèmes linéaires qu'à un changement de variable près, on aboutit à N équations de transport découplées les unes des autres, dont les vitesses sont les valeurs propres de la matrices, ce qui donnait la forme de solution (2.13).

Pour mieux comprendre les phénomènes de transport sous-jacent à (5.1), on considère une solution de la forme

$$U(t, x) = V(\sigma(t, x))$$

où $\sigma : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}$ et $V : \mathbb{R} \rightarrow \Omega$ sont deux fonctions régulières. En injectant cette solution dans (5.1), on obtient

$$[\partial_t \sigma \mathbf{I} + \partial_x \sigma \nabla f(V(\sigma))] V'(\sigma) = 0.$$

On suppose que les solutions que l'on considère ne sont pas des constantes, donc $V'(\sigma)$ et $\partial_x \sigma$ sont non nuls. On obtient alors

$$\nabla f(V(\sigma)) V'(\sigma) = -\frac{\partial_t \sigma}{\partial_x \sigma} V'(\sigma),$$

c'est-à-dire que, en notant λ_i et r_i les valeurs propres et vecteurs propres de ∇f , on a

$$V'(\sigma) = r_i(V(\sigma)), \quad (5.9)$$

$$\partial_t \sigma + \lambda_i(V(\sigma)) \partial_x \sigma = 0. \quad (5.10)$$

L'équation (5.9) est un système d'équations différentielles permettant de calculer V en fonction de σ et l'équation (5.10) est une équation de transport qui est *a priori* non linéaire. Pour en savoir plus sur cette équation, on est amené à introduire les définitions suivantes :

Définition 5.6. L'onde associé à la valeur propre λ_i est linéairement dégénérée si

$$\nabla \lambda_i(U) \cdot r_i(U) = 0, \quad \forall U \in \Omega.$$

L'onde associé à la valeur propre λ_i est vraiment non linéaire si

$$\nabla \lambda_i(U) \cdot r_i(U) \neq 0, \quad \forall U \in \Omega.$$

(Comme précédemment, ces définitions sont invariantes par changement de variable.)

La quantité $\nabla \lambda_i \cdot r_i$ intervenant dans ces définitions est tout simplement la dérivée de la vitesse de transport $\lambda_i(V(\sigma))$ de l'équation (5.10) :

$$\begin{aligned} d_\sigma \lambda_i(V(\sigma)) &= \nabla \lambda_i(V(\sigma)) \cdot V'(\sigma) \\ &= \nabla \lambda_i(V(\sigma)) \cdot r_i(V(\sigma)) \end{aligned}$$

par (5.9). On en déduit donc que l'équation (5.10) est une équation de transport linéaire si l'onde associée à λ_i est linéairement dégénérée. De même, si l'onde associée à λ_i est vraiment non linéaire, alors l'équation (5.10) est une équation scalaire non linéaire dont le flux est strictement convexe ou strictement concave (puisque la dérivée seconde de ce flux est de signe fixé).

Remarque 11. On ne va pas considérer les cas où $\nabla \lambda_i \cdot r_i$ n'est ni non nul ni de signe donné, ça devient très compliqué (même si c'est faisable).

Proposition 5.7. Soit $\Phi \in \mathcal{C}^1(\Omega; \mathbb{R})$ un i -invariant de Riemann, c'est-à-dire que Φ vérifie

$$\nabla \Phi(U) \cdot r_i(U) = 0, \quad \forall U \in \Omega. \quad (5.11)$$

Alors, $\Phi(V(\sigma))$ est constant si V est solution de (5.9).

Démonstration. On a directement $d_\sigma \Phi(V(\sigma)) = \nabla \Phi(V(\sigma)) \cdot r_i(V(\sigma)) = 0$. \square

En fait, on peut montrer qu'il existe exactement $(N - 1)$ i -invariants de Riemann dont les gradients sont linéairement indépendants (cela assure qu'un invariant de Riemann n'est pas une fonction des autres). On retombe alors sur un découplage qui peut rappeler le cas linéaire diagonalisé puisque qu'on a une variable convectée à la vitesse λ_i et $N - 1$ autres (les i -invariants de Riemann) qui restent constantes. Cela entraîne la définition suivante :

Définition 5.8. Une solution régulière U de (5.1) définie sur un domaine D de $\mathbb{R}^+ \times \mathbb{R}$ est une i -onde simple si $\Phi(U(t, x))$ est constant dans D pour tout i -invariant de Riemann Φ .

Comment construire une onde simple ? Soit un état $U_0 \in \Omega$ et un profil initial $\sigma_0(x)$ et on note $\sigma_0(0) = \xi_0$. Alors, V est solution du problème de Cauchy

$$\begin{cases} V'(\xi) = r_i(V(\xi)), \\ V(\xi_0) = U_0. \end{cases} \quad (5.12)$$

Sous des hypothèses de régularité, il existe une unique solution à ce problème, au moins pour ξ proche de ξ_0 . De plus, l'équation (5.10) donne directement $\sigma(t, x) = \sigma_0(x - \lambda_i(V(\sigma(t, x)))t)$ si on suppose que $\sigma(0, x) = \sigma_0(x)$. On obtient donc que

$$U(t, x) = V(\sigma_0(x - \lambda_i(U(t, x))t)),$$

dont on déduit de nouveau la notion de courbe caractéristique, c'est-à-dire de courbe sur laquelle la solution est constante :

Définition 5.9. Les courbes caractéristiques dans une i -onde simple sont les solutions de l'équation différentielle

$$\begin{cases} X'(t; X_0) = \lambda_i(U(t, x)), & t > 0, \\ X(0; X_0) = X_0, \end{cases} \quad (5.13)$$

où $(0, X_0) \in D$.

On obtient donc

Proposition 5.10. Soit U une i -onde simple. Alors, les caractéristiques sont des lignes droites de pente $\lambda_i(U(t, x))$ et U est constant le long de celles-ci.

Il est important de noter que jusqu'à présent, on a toujours supposé que la solution était régulière. Dans le cas linéairement dégénéré, il suffit que σ_0 soit régulier. En revanche, dans le cas vraiment non linéaire, on peut avoir apparition de chocs.

5.4 Ondes de choc et entropie

On considère maintenant des solutions discontinues. On sait qu'elles doivent vérifier les relations de saut de Rankine-Hugoniot (5.8). Néanmoins, ces relations ne suffisait pas dans le cas scalaire, au sens où l'unicité des solutions n'était pas garantie. On faisait alors appel à la notion d'entropie pour sélectionner les solutions « physiquement » admissibles. Pour cela, on peut ajouter un terme de diffusion

$$\partial_t U_\varepsilon + \partial_x f(U_\varepsilon) = \varepsilon \partial_{xx}^2 U_\varepsilon,$$

qui par le même calcul que dans le cas scalaire conduit à considérer les solutions faibles vérifiant

$$\partial_t \eta(U) + \partial_x F(U) \leq 0.$$

Cependant, l'utilisation de l'entropie ne peut plus être systématique : il faut d'une part être sûr qu'il en existe une et surtout être capable de toutes les connaître. En outre, le terme de diffusion $\varepsilon \partial_{xx}^2 U_\varepsilon$ n'est en général pas justifié physiquement. On devrait plutôt ajouter les termes correspondant aux effets visqueux mais la régularité de la solution et son caractère entropique est beaucoup moins clair. Pour éviter ces difficultés, on fait appel aux critères d'admissibilité locaux, comme (3.10) et (3.11).

On dira dans la suite qu'une i -onde est discontinue si dans les relations de Rankine-Hugoniot (5.8) la vitesse σ' tend vers $\lambda_i(U^-)$ quand U^+ tend U^- .

Comme dans le cas des ondes simples (donc régulières), on va considérer deux cas :

Définition 5.11. Une i -onde discontinue est appelée discontinuité de contact si elle est linéairement dégénérée et onde de choc si elle est vraiment non linéaire.

Proposition 5.12. La caractérisation d'une discontinuité de contact par les relations de Rankine-Hugoniot (5.8) est équivalente à la caractérisation par ses $N - 1$ invariants de Riemann (dont les gradients sont linéairement indépendants).

Cela est simplement dû au fait que les discontinuités de contact se comporte comme des ondes linéaires. Ainsi, le cas régulier et le cas discontinu sont identiques pour ces ondes.

Définition 5.13. Une i -onde de choc est admissible au sens de Lax si elle vérifie

$$\lambda_i(U^-) > \sigma' > \lambda_i(U^+). \quad (5.14)$$

On peut étendre de même le critère d'Oleinik au cas des systèmes (on l'appelle alors critère de Liu), mais celui-ci n'est utile que dans les cas où $\nabla \lambda_i \cdot r_i$ peut changer de signe.

De plus on peut démontrer que ce critère est équivalent au critère entropique.

5.5 Problème de Riemann

Passons maintenant à la résolution du problème de Riemann

$$\begin{cases} \partial_t U + \partial_x f(U) = 0, & t > 0, x \in \mathbb{R}, \\ U(0, x) = \begin{cases} U_L & \text{si } x < 0, \\ U_R & \text{si } x > 0. \end{cases} \end{cases} \quad (5.15)$$

Comme dans le cas scalaire, la propriété d'auto-similarité de la solution est vérifiée :

Proposition 5.14. Soit U_0 une fonction telle que

$$\forall \lambda > 0, x \in \mathbb{R}, \quad U_0(\lambda x) = U_0(x).$$

Alors il existe une solution faible U du problème de Cauchy (3.1) avec une telle donnée initiale telle que

$$\forall \lambda > 0, x \in \mathbb{R}, t > 0, \quad U(\lambda t, \lambda x) = U(t, x).$$

On peut alors définir une fonction $V \in \mathbf{L}^\infty(\mathbb{R}; \mathbb{R})$ telle que

$$\forall \lambda > 0, x \in \mathbb{R}, t > 0, \quad V(x/t) = U(t, x).$$

On va supposer dans la suite que le système est strictement hyperbolique et que ses valeurs propres sont ordonnées :

$$\lambda_1(U) < \lambda_2(U) < \dots < \lambda_N(U), \quad \forall U \in \Omega.$$

On s'attend donc à avoir une solution composée de $N + 1$ états constants $U_1 = U_L, U_2, \dots, U_{N+1} = U_R$ séparés par des ondes : les états U_i et U_{i+1} sont séparés par une i -onde.

Avant de continuer, penchons-nous sur le cas du problème de Riemann pour un système linéaire.

Théorème 5.15. Supposons que $f(U) = AU$ où $A \in \mathbb{R}^{N \times N}$ est diagonalisable dans \mathbb{R} et ses valeurs propres sont de multiplicité 1. Alors, l'unique solution du problème de Riemann (5.15) est donnée par

$$U(t, x) = U_L + \sum_{i \text{ t.q. } x/t < \lambda_i} (l_i \cdot (U_R - U_L)) r_i \quad (5.16)$$

où $(l_i)_{1 \leq i \leq N}$ et $(r_i)_{1 \leq i \leq N}$ sont les vecteurs propres à gauche et à droite de A .

Démonstration. [EX] □

On voit donc que, en se plaçant dans Ω , le vecteur $U_{i+1} - U_i$ est colinéaire à r_i . On peut remarquer une nouvelle fois l'importance de l'hypothèse de l'hyperbolicité, qui assure que $(r_i)_{1 \leq i \leq N}$ est une base de \mathbb{R}^N , donc que le vecteur $U_R - U_L$ se décompose de manière unique dans la base $(r_i)_{1 \leq i \leq N}$. De plus, les états intermédiaires sont donc donnés par $U_i = U_L + \sum_{j=1}^{i-1} (l_j \cdot (U_R - U_L)) r_j$.

Dans le cas des systèmes non linéaires, la situation n'est pas bien différente. En effet, les discontinuités de contact étant assimilables à des ondes simples, (5.12) assure que pour ξ proche de ξ_0 , on suit la direction $r_i(V_0)$. Concernant les ondes vraiment non linéaires, on utilise le résultat suivant :

Proposition 5.16. *Supposons que la i -onde soit vraiment non linéaire.*

Soit il existe $s \in \mathbb{R}$ tel que les états U_i et U_{i+1} vérifient les relations de Rankine-Hugoniot

$$-s(U_{i+1} - U_i) + (f(U_{i+1}) - f(U_i)) = 0$$

et le critère de Lax $\lambda_i(U_i) > s > \lambda_i(U_{i+1})$. Les états U_i et U_{i+1} sont alors séparés par une onde de choc. De plus, quand $|U_{i+1} - U_i|$ tend vers 0, le vecteur $U_{i+1} - U_i$ tend vers $r_i(U_i)$ (à l'ordre 1).

Soit les états U_i et U_{i+1} vérifient pour tout i -invariant de Riemann

$$\Phi(U_i) = \Phi(U_{i+1})$$

et $\lambda_i(U_i) \leq \lambda_i(U_{i+1})$. Les états U_i et U_{i+1} sont alors séparés par une onde de détente, définie localement (en x/t) par

$$U(t, x) = \begin{cases} U_i & \text{si } x/t < \lambda_i(U_i), \\ V(x/t) & \text{si } \lambda_i(U_i) < x/t < \lambda_i(U_{i+1}), \\ U_{i+1} & \text{si } x/t > \lambda_i(U_{i+1}), \end{cases} \quad (5.17)$$

où V est solution de (5.12) avec $U_0 = U_i$ et $\xi_0 = \lambda_i(U_i)$ (c'est en fait une onde simple).

Il n'y a pas d'autre cas.

Démonstration. La démonstration de la propriété des ondes de choc est assez fastidieuse et n'est pas présentée ici. Concernant l'onde de détente, le fait que ce soit une onde simple suffit pour conclure que c'est bien une solution de (5.15). Le fait qu'il n'y ait pas d'autre possibilité n'est pas non plus immédiat à démontrer. \square

On peut déduire de ce résultat que l'ensemble des états U_{i+1} que l'on peut connecter à U_i à travers une i -onde vraiment non linéaire est une courbe de dimension 1 dans Ω passant par U_i et tangente à $r_i(U_i)$ au point U_i . D'un côté elle correspond à l'onde de détente et de l'autre à la courbe de choc. Dans le cas d'une i -onde linéairement dégénérée, on obtient aussi une courbe de dimension 1 passant par U_i et tangente à $r_i(U_i)$ au point U_i . On peut donc en déduire le résultat suivant :

Théorème 5.17. *Considérons le problème de Riemann (5.15) avec U_L proche de U_R . Alors, si le système n'admet que des ondes vraiment non linéaires ou linéairement dégénérées, il existe une et une seule solution faible admissible au sens de Lax composée de N ondes (éventuellement d'amplitude nulle) séparant $N + 1$ états constants.*

Démonstration. Notons $\mathcal{C}_i(U_i, \cdot)$ l'ensemble des états de Ω que l'on peut connecter à U_i à travers une i -onde. Cette ensemble est une famille à un paramètre, noté ici ε_i ,

tel que $U_i = \mathcal{C}_i(U_i, 0)$. Ainsi, $U_2 = \mathcal{C}_1(U_L, \varepsilon_1)$, $U_3 = \mathcal{C}_2(\mathcal{C}_1(U_L, \varepsilon_1), \varepsilon_2)$, etc. Soit $\varepsilon = (\varepsilon_1, \dots, \varepsilon_N) \in \mathbb{R}^N$. Alors, on peut définir la fonction \mathcal{C} telle que

$$U_R = \mathcal{C}(U_L, \varepsilon) = \mathcal{C}_N(\mathcal{C}_{N-1}(\dots \mathcal{C}_1(U_L, \varepsilon_1) \dots, \varepsilon_{N-1}), \varepsilon_N).$$

Si U_L et U_R sont suffisamment proches, les états intermédiaires U_i appartiennent à un voisinage de U_L et donc $\mathcal{C}_i(U_i, \varepsilon_i) \approx U_i + \varepsilon r_i(U_i)$. Comme le système est hyperbolique, $(r_i(U_L))_{1 \leq i \leq N}$ est une base de \mathbb{R}^N et par ailleurs $r_i(U_i) \approx r_i(U_L)$, donc $(r_i(U_L))_{1 \leq i \leq N}$ est aussi une base de \mathbb{R}^N . On obtient donc l'existence et l'unicité de $\mathcal{C}(U_L, \varepsilon)$, qui est voisin de (5.16). \square

Cette démonstration est totalement imprécise bien sûr, il faudrait mieux estimer l'impact des non linéarités sur les courbes d'ondes \mathcal{C}_i .

Ce théorème peut paraître un peu « léger » car il n'est valable que pour des données initiales proches (on a donc un résultat local), mais au cas par cas, on peut cependant obtenir des résultats globaux, c'est-à-dire pour tout $U_L, U_R \in \Omega$.

5.6 Problème de Cauchy

Concernant l'existence d'une solution, il est difficile de déterminer des estimations *a priori* vérifiées par les solutions faibles admissibles au sens de Lax, même dans le cas de données initiales presque constantes.

Concernant l'unicité, le critère entropique

$$\partial_t \eta(U) + \partial_x F(U) \leq 0 \quad (5.18)$$

(au sens faible) n'est pas forcément suffisant, car pour appliquer la technique de Kruzhkov il est nécessaire d'avoir un très grand nombre d'entropie, ce qui n'est pas le cas ici, excepté pour des systèmes particuliers.

5.7 Le problème de Riemann pour Euler barotrope

Passons maintenant à un exemple de résolution de problème de Riemann. On regarde le système

$$\partial_t \rho + \partial_x \rho u = 0 \quad (5.19)$$

$$\partial_t \rho u + \partial_x (\rho u^2 + p) = 0 \quad (5.20)$$

où $p = \mathcal{P}(\rho)$ telle que $\mathcal{P}(\rho) > 0$ et $\mathcal{P}'(\rho) > 0$ pour tout $\rho > 0$. On supposera que $\mathcal{P}(\rho) = a\rho^\gamma$, $\gamma > 1$ et $a > 0$. On se place dans le domaine $\Omega = \mathbb{R}^{*+} \times \mathbb{R}$. Ce système s'écrit aussi pour les solutions régulières

$$\partial_t \begin{pmatrix} \rho \\ u \end{pmatrix} + \begin{pmatrix} u & \rho \\ \mathcal{P}'(\rho)/\rho & u \end{pmatrix} \partial_x \begin{pmatrix} \rho \\ u \end{pmatrix} = 0.$$

Si on note $c(\rho) = \sqrt{\mathcal{P}'(\rho)} = \sqrt{a\gamma\rho^{(\gamma-1)/2}}$ (la vitesse du son), alors les valeurs propres du système sont $u \pm c$ et

$$r_- = \begin{pmatrix} 1 \\ -c/\rho \end{pmatrix} \quad \text{et} \quad r_+ = \begin{pmatrix} 1 \\ c/\rho \end{pmatrix}.$$

Le système est donc strictement hyperbolique dans Ω et on peut vérifier aisément que les deux ondes sont vraiment non linéaires (ce sont soit des ondes de détente, soit des ondes de choc, selon les données initiales).

Vu que l'on a un système 2×2 , on s'attend à avoir deux ondes, avec un état constant au milieu. La première étape est d'étudier ces ondes, c'est-à-dire, en se donnant un état $U_0 \in \Omega$, déterminer tous les états que l'on peut y connecter à travers une 1- ou une 2-onde.

Cela permet de déterminer l'état intermédiaire par intersection de la courbe de 1-onde issue de U_L et de la courbe de 2-onde issue de U_R . Enfin, connaissant cet état et la nature des ondes le séparant de U_L et U_R , on peut déterminer la solution du problème de Riemann.

5.7.1 Étude des ondes

Plaçons-nous dans le cas de l'onde de vitesse $u - c$ et regardons les ondes de détente. Si on reprend (5.12), on a

$$\begin{cases} \rho'(\xi) = 1 \\ u'(\xi) = -c/\rho = -\sqrt{a\gamma} \rho^{(\gamma-3)/2} \\ (\rho, u)(0) = (\rho_0, u_0) \end{cases}.$$

La première équation donne $\rho(\xi) = \xi + \rho_0$, ce qui permet d'écrire $\xi(\rho) = \rho - \rho_0$. Ainsi, la deuxième équation devient $u'(\rho) = -\sqrt{a\gamma} \rho^{(\gamma-3)/2}$, donc

$$u(\rho) = -\frac{2}{\gamma-1} \sqrt{a\gamma} \rho^{(\gamma-1)/2} + K_0 = -2c/(\gamma-1) + K_0.$$

Avec la donnée initiale, on obtient donc

$$\begin{aligned} \rho(\xi) &= \xi + \rho_0 \\ u(\rho(\xi)) &= u_0 - \frac{2}{\gamma-1} (c(\rho) - c(\rho_0)). \end{aligned}$$

On peut aussi passer directement par les invariants de Riemann. En effet, comme le système est de dimension 2, il suffit d'en trouver un seul, notons-le $\Phi_-(\rho, u)$. L'équation (5.11) donne

$$\partial_\rho \Phi_- - \frac{c}{\rho} \partial_u \Phi_- = 0.$$

Si on suppose qu'il est de la forme $\Phi_-(\rho, u) = u + \varphi_-(\rho)$, alors on a $\varphi'_-(\rho) = c(\rho)/\rho$, ce qui donne

$$\Phi_-(\rho, u) = u + \frac{2}{\gamma-1} c(\rho)$$

et on aboutit à la même conclusion.

Lemme 5.18. *La courbe de 1-onde de détente dans le plan (ρ, u) , notée $\mathcal{R}_-(\rho_0, u_0)$ est définie par*

$$u = u_0 - \frac{2}{\gamma-1} (c(\rho) - c(\rho_0)), \quad 0 < \rho \leq \rho_0. \quad (5.21)$$

Elle est strictement décroissante et sur cette courbe, la fonction $u - c$ est monotone décroissante (par rapport à ρ). Cette courbe correspond à l'ensemble des états U que l'on peut atteindre de (ρ_0, u_0) à travers une onde de détente de vitesse $u - c$.

Passons maintenant aux courbes de choc. Pour cela, il faut tout d'abord écrire les relations de Rankine-Hugoniot entre U_- et U_+ (avec $\Delta\alpha = \alpha_+ - \alpha_-$ et s la vitesse du choc) :

$$\begin{aligned} -s\Delta\rho + \Delta(\rho u) &= 0, \\ -s\Delta(\rho u) + \Delta(\rho u^2 + p) &= 0. \end{aligned}$$

On définit maintenant $v = u - s$. La première équation devient directement

$$\Delta(\rho v) = 0$$

et après quelques calculs :

$$\begin{aligned} -s\Delta(\rho u) + \Delta(\rho u^2 + p) &= -s(-s\Delta\rho + \Delta(\rho u)) - s\Delta(\rho u) + \Delta(\rho u^2 + p) \\ &= \Delta(\rho s^2) - \Delta(2\rho u) + \Delta(\rho u^2) + \Delta p \\ &= \Delta(\rho(u-s)^2) + \Delta p, \end{aligned}$$

la seconde équation devient

$$\Delta(\rho v^2) + \Delta p = 0.$$

Notons maintenant $M = \rho_- v_- = \rho_+ v_+$. Par définition de v , on obtient $s = u_- - M\tau_- = u_+ - M\tau_+$, où $\tau = 1/\rho$. On déduit alors

$$M = \frac{\Delta u}{\Delta\tau}.$$

De même, on a $Mv_- + p_- = Mv_+ + p_+$, c'est-à-dire $-M\Delta u = \Delta p$, donc

$$M^2 = -\frac{\Delta p}{\Delta\tau}.$$

(Cette relation est bien définie car p est décroissant par rapport à τ , donc le second membre est bien négatif.) Ainsi, on obtient la relation

$$(\Delta u)^2 = -\Delta\tau \Delta p.$$

Pour obtenir une équation du même type que (5.21) (u en fonction seulement de u_0 , ρ et ρ_0), on désire prendre la racine carrée, mais le signe de Δu n'est pas connu. En fait, suivant le signe, cela correspond à un 1- ou un 2-choc.

Prenons le cas d'un 1-choc stationnaire de petite amplitude. Alors $s = 0$ et $u - c \approx 0$, donc $u \approx c > 0$. Comme $M = \rho u$, on en déduit que $M > 0$. On obtient donc que pour un 1-choc, $\Delta u = \Delta\tau \sqrt{-\Delta p/\Delta\tau}$.

Il reste à inclure la condition d'entropie pour sélectionner les chocs admissibles. On va utiliser la condition de Lax (5.14), qui pour la 1-onde, oblige à avoir $\Delta(u - c) < 0$. Supposons que $\Delta p > 0$. Alors $\Delta\tau = -\Delta\rho/\rho_- \rho_+ < 0$ et donc $\Delta u = -\sqrt{-\Delta p \Delta\tau} < 0$. Comme $\Delta c > 0$, on en déduit que $\Delta(u - c) < 0$. Regardons maintenant le cas $\Delta p < 0$. Dans ce cas, $\Delta\tau > 0$ et $\Delta u = \sqrt{-\Delta p \Delta\tau} > 0$, ce qui donne $\Delta(u - c) > 0$. On en déduit donc :

Lemme 5.19. *La courbe de 1-onde de choc entropique dans le plan (ρ, u) , notée $\mathcal{S}_-(\rho_0, u_0)$, est strictement décroissante et définie par*

$$u = u_0 - \sqrt{-(p(\rho) - p(\rho_0))(1/\rho - 1/\rho_0)}, \quad \rho > \rho_0. \quad (5.22)$$

Cette courbe correspond à l'ensemble des états U que l'on peut atteindre de (ρ_0, u_0) à travers une onde de choc de vitesse $u - c$ (U_0 est à gauche du choc et U à droite).

Pour calculer les courbes associées à 2-onde, on procède de même. Il faut cependant noter que cette fois, on suppose que l'état que l'on se donne U_0 est à droite de l'onde et on cherche l'ensemble des états U admissibles à gauche de cette 2-onde. On aboutit alors à :

Lemme 5.20. *La courbe de 2-onde de détente dans le plan (ρ, u) , notée $\mathcal{R}_+(\rho_0, u_0)$, est définie par*

$$u = u_0 + \frac{2}{\gamma-1}(c(\rho) - c(\rho_0)), \quad 0 < \rho \leq \rho_0. \quad (5.23)$$

Elle est strictement croissante et sur cette courbe, la fonction $u + c$ est monotone décroissante (par rapport à ρ). Cette courbe correspond à l'ensemble des états U que l'on peut atteindre de (ρ_0, u_0) à travers une onde de détente de vitesse $u + c$.

La courbe de 2-onde de choc entropique dans le plan (ρ, u) , notée $\mathcal{S}_+(\rho_0, u_0)$, est définie par

$$u = u_0 + \sqrt{-(p(\rho) - p(\rho_0))(1/\rho - 1/\rho_0)}, \quad \rho > \rho_0. \quad (5.24)$$

Cette courbe correspond à l'ensemble des états U que l'on peut atteindre de (ρ_0, u_0) à travers une onde de choc de vitesse $u + c$ (U_0 est à droite du choc et U à gauche).

5.7.2 Résolution du problème de Riemann

On définit maintenant les courbes d'onde $\mathcal{C}_-(\rho_L, u_L) = \mathcal{R}_-(\rho_L, u_L) \cup \mathcal{S}_-(\rho_L, u_L)$ et $\mathcal{C}_+(\rho_R, u_R) = \mathcal{R}_+(\rho_R, u_R) \cup \mathcal{S}_+(\rho_R, u_R)$. L'état intermédiaire est donc l'intersection de ces deux courbes, pour peu qu'il appartienne à Ω . Ainsi, en étudiant $\mathcal{C}_-(\rho_L, u_L)$ et $\mathcal{C}_+(\rho_R, u_R)$ pour $\rho = 0$ et $\rho \rightarrow +\infty$, on obtient :

Théorème 5.21. *Si $U_L, U_R \in \Omega$ vérifient*

$$u_R - u_L < \frac{2}{\gamma-1}(c_L + c_R) \quad (5.25)$$

alors $U^ = \mathcal{C}_-(\rho_L, u_L) \cap \mathcal{C}_+(\rho_R, u_R)$ existe et appartient à Ω . On en déduit que le problème de Riemann admet une et une seule solution sous cette condition.*

En général, si cette condition n'est pas vérifiée, on prolonge la solution par le vide, c'est-à-dire que l'état intermédiaire vérifie $(\rho, \rho u) = (0, 0)$, ce qui permet d'étendre le résultat précédent à tout $U_L, U_R \in \Omega$ (de la même manière, on peut aussi étendre ce résultat à tout $U_L, U_R \in \bar{\Omega}$).

Remarque 12. Pour terminer, on va montrer comment obtenir directement les résultats (5.12) et la proposition 5.10 en supposant l'auto-similarité de la solution. Soit $\xi = x/t$ et $V(\xi) = U(t, x)$. Alors, les solutions régulières du systèmes vérifient :

$$-\frac{x}{t^2}V'(\xi) + F'(V(\xi))\frac{1}{t}V'(\xi) = 0$$

ce qui donne

$$F'(V(\xi))V'(\xi) = \xi V'(\xi).$$

Si on suppose qu'on regarde les solutions non constantes, alors on obtient directement que $\xi = u \pm c$ et que V' est égal au vecteur propre associé à $u \pm c$.

Chapitre 6

Méthodes de volumes finis pour les systèmes de lois de conservation

Bien que l'on ne sache pas en général étudier théoriquement le problème de Cauchy

$$\begin{cases} \partial_t U + \partial_x f(U) = 0, & t > 0, x \in \mathbb{R}, \\ U(0, x) = U_0(x), & x \in \mathbb{R}, \end{cases} \quad (6.1)$$

on peut être amené à étudier son approximation numérique par des schémas VF. Contrairement au cas scalaire, on ne sait pas, pour l'instant, trouver des bornes sur la solution approchée pour passer à la limite (on aurait dans ce cas une démonstration de l'existence pour (6.1)). On va toutefois tenter d'imiter le cas scalaire dans la construction des schémas VF, à défaut d'une meilleure compréhension du cas des systèmes.

6.1 Schémas volumes finis et propriétés de base

Rappelons tout d'abord la notion de maillage. Soit une suite réelle strictement croissante $(x_{i+1/2})_{i \in \mathbb{Z}}$ représentant les interfaces entre les mailles M_i . On définit les pas d'espace $\Delta x_i = x_{i+1/2} - x_{i-1/2}$ qui correspondent aux mesures des mailles. On définit en suite le pas de temps Δt et $t^n = n\Delta t$.

Comme dans le cas scalaire, on définit

$$U_i^0 = \frac{1}{\Delta x_i} \int_{M_i} U_0(x) dx \quad (6.2)$$

et on cherche à calculer la suite $(U_i^n)_{i \in \mathbb{Z}, n \in \mathbb{N}}$. Toujours par intégration du système (6.1) sur $(t^n, t^{n+1}) \times M_i$, on a

$$\begin{aligned} \int_{M_i} U(t^{n+1}, x) dx - \int_{M_i} U(t^n, x) dx \\ + \int_{t^n}^{t^{n+1}} f(U(t, x_{i+1/2})) dt - \int_{t^n}^{t^{n+1}} f(U(t, x_{i-1/2})) dt = 0 \end{aligned}$$

qui devient

$$\Delta x_i (U_i^{n+1} - U_i^n) + \Delta t (f_{i+1/2}^n - f_{i-1/2}^n) \approx 0$$

où $\Delta t f_{i+1/2}^n \approx \int_{t^n}^{t^{n+1}} f(U(t, x_{i+1/2})) dt$. Si on définit $f_{i+1/2}^n$ par une fonction dépendant simplement de $(U_i^n)_{i \in \mathbb{Z}}$, on obtient un schéma explicite à un pas dans la terminologie de l'approximation des équations différentielles.

On ne va s'intéresser encore une fois qu'aux schémas à trois points, c'est-à-dire que le flux numérique est défini par $f_{i+1/2}^n = g(U_i^n, U_{i+1}^n)$, ce qui donne

$$U_i^{n+1} = U_i^n - \frac{\Delta t}{\Delta x_i} (g(U_i^n, U_{i+1}^n) - g(U_{i-1}^n, U_i^n)). \quad (6.3)$$

On suppose dans la suite que le flux numérique g est Lipschitz continu. De nouveau, on a la propriété de base des schémas VF :

Proposition 6.1. *Le schéma VF (6.3) est conservatif, c'est-à-dire que si $U_0 \in \mathbf{L}^1(\mathbb{R})$, alors pour tout $n > 0$,*

$$\sum_{i \in \mathbb{Z}} U_i^n \Delta x_i = \int_{\mathbb{R}} U_0(x) dx.$$

Démonstration. La démonstration est directe en sommant (6.3) pour $i \in \mathbb{Z}$ et en utilisant (6.2). \square

6.1.1 Consistance

Comme dans le cas scalaire, plusieurs propriétés sont nécessaires sur le flux numérique pour avoir la convergence du schéma vers la solution faible entropique. Ici, on ne pourra pas aboutir à un tel résultat, donc on détaille les propriétés requises et leurs conséquences.

Définition 6.2. Le schéma VF (6.3) est consistant si pour tout $U \in \Omega$

$$g(U, U) = f(U).$$

C'est la même notion de consistance que dans le cas scalaire, qui permet notamment d'assurer que si pour tout x $U_0(x) = \bar{U}$ où $\bar{U} \in \Omega$ est un état constant, alors pour tout i et n , $U_i^n = \bar{U}$. On a le résultat suivant de consistance avec la solution classique :

Proposition 6.3. *Si on suppose que pour tout $i \in \mathbb{Z}$*

$$U_i^n = \frac{1}{\Delta x_i} \int_{M_i} U(t^n, x) dx$$

où U est une solution régulière de (6.1), alors U_i^{n+1} défini par le schéma VF (6.3) vérifie

$$U_i^{n+1} - \frac{1}{\Delta x_i} \int_{M_i} U(t^{n+1}, x) dx \rightarrow 0$$

quand $\Delta t, \sup_i \Delta x_i \rightarrow 0$ dès que le schéma VF est consistant.

Démonstration. On intègre le système sur $(t^n, t^{n+1}) \times M_i$ et on soustrait le schéma numérique (6.3), ce qui donne

$$\frac{1}{\Delta x_i} \int_{M_i} U(t^{n+1}, x) dx - U_i^{n+1} + \frac{\Delta t}{\Delta x_i} (F_{i+1/2}^n - F_{i-1/2}^n) = 0$$

où $F_{i+1/2}^n = (1/\Delta t) \int_{t^n}^{t^{n+1}} f(U(t, x_{i+1/2})) dt - g(U_i^n, U_{i+1}^n)$. Par continuité de g et U ,

$$\begin{aligned} g(U_i^n, U_{i+1}^n) &= g(U(t^n, x_{i+1/2}), U(t^n, x_{i+1/2})) + \mathcal{O}(\Delta x_i) + \mathcal{O}(\Delta x_{i+1}) \\ &= f(U(t^n, x_{i+1/2})) + \mathcal{O}(\Delta x_i) + \mathcal{O}(\Delta x_{i+1}) \end{aligned}$$

par consistance du schéma. On en déduit donc (la solution étant régulière) que $F_{i+1/2}^n = \mathcal{O}(\Delta t) + \mathcal{O}(\Delta x_i) + \mathcal{O}(\Delta x_{i+1})$, ce qui permet de conclure. \square

Remarque 13. En fait, le théorème de Lax-Wendroff peut être appliqué au cas des systèmes. Néanmoins, l'hypothèse de la borne \mathbf{L}^∞ n'est en général pas réaliste.

L'autre hypothèse cruciale dans le cas scalaire est la monotonie du flux numérique. Il est clair que cette propriété n'a pas d'analogue direct dans le cas des systèmes. Si on reprend l'analyse des schémas VF dans le cas scalaire, on peut remarquer que la monotonie permet d'obtenir d'une part des bornes dans $\mathbf{L}^\infty \cap \mathbf{BV}$ (inégalités (4.6) et (4.9)) et d'autre part des inégalités d'entropie discrètes (4.13). On va voir comment imiter ce type de propriétés dans le cas des systèmes. Pour cela, on va travailler avec la notion de demi-maille. On définit les fonctions $U^\pm : \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^+ \rightarrow \mathbb{R}^N$ par

$$\begin{aligned} U^-(U_l, U_r, \sigma) &= U_l - \frac{1}{\sigma_l(U_l, U_r)} (g(U_l, U_r) - f(U_l)), \\ U^+(U_l, U_r, \sigma) &= U_r - \frac{1}{\sigma_r(U_l, U_r)} (f(U_r) - g(U_l, U_r)). \end{aligned} \quad (6.4)$$

Celles-ci vont nous permettre de séparer chaque interface, sous condition CFL.

Lemme 6.4. *Le schéma numérique (6.3) peut s'écrire*

$$U_i^{n+1} = \frac{U_{i-}^{n+1} + U_{i+}^{n+1}}{2}$$

où

$$\begin{aligned} U_{i-}^{n+1} &= U^+(U_{i-1}^n, U_i^n, \Delta x_i / (2\Delta t)), \\ U_{i+}^{n+1} &= U^-(U_i^n, U_{i+1}^n, \Delta x_i / (2\Delta t)). \end{aligned}$$

Démonstration. En effet, on a par définition

$$\begin{aligned} U_{i-}^{n+1} &= U_i^n - \frac{2\Delta t}{\Delta x_i} (f(U_i^n) - g(U_{i-1}^n, U_i^n)), \\ U_{i+}^{n+1} &= U_i^n - \frac{2\Delta t}{\Delta x_i} (g(U_i^n, U_{i+1}^n) - f(U_i^n)), \end{aligned}$$

ce qui donne bien (6.3) quand on prend la moyenne arithmétique. \square

Attention, les $-$ et $+$ ne signifient pas la même chose (pour U^\pm , c'est de part et d'autre de l'interface, alors que pour $U_{i\pm}^n$, c'est de part et d'autre du centre de la maille).

6.1.2 Préservation de domaine invariant

Les estimations *a priori* n'ont pas lieu d'être dans le cas des systèmes puisqu'on ne sait pas si elles sont vérifiées par la (ou les) solution(s). Néanmoins, on peut supposer que l'ensemble convexe $\Omega \subset \mathbb{R}^N$ est un *domaine invariant* pour (6.1), c'est-à-dire que

si $U_0(x) \in \Omega$ pour tout $x \in \mathbb{R}$, alors $U(t, x) \in \Omega$ pour tout $t > 0$ et $x \in \mathbb{R}$.

On va donc définir l'analogue discret de cette propriété (qui est à rapprocher de la stabilité L^∞) :

Définition 6.5. Soit Ω un domaine invariant pour (6.1). On dit que le schéma VF (6.3) préserve invariant le domaine Ω

$$\text{si } U_i^n \in \Omega \text{ pour tout } i \in \mathbb{Z}, \text{ alors } U_i^{n+1} \in \Omega \text{ pour tout } i \in \mathbb{Z},$$

pour tout $n \in \mathbb{N}$.

Cette propriété est difficile à interpréter directement. L'idée est alors de travailler par demi-maillages pour obtenir plutôt des conditions sur le flux numérique, donc localisées aux interfaces :

Définition 6.6. Soit Ω un domaine invariant pour (6.1). On dit que le flux numérique g préserve invariant le domaine Ω si il existe $-\sigma_l(U_l, U_r) < 0 < \sigma_r(U_l, U_r)$ tels que

$$U_l, U_r \in \Omega \implies \begin{cases} U_-(U_l, U_r, \sigma_l(U_l, U_r)) \in \Omega, \\ U_+(U_l, U_r, \sigma_r(U_l, U_r)) \in \Omega. \end{cases} \quad (6.5)$$

En fait, la propriété est encore vraie pour n'importe quels $-\tilde{\sigma}_l \leq -\sigma_l$ et $\sigma_r \leq \tilde{\sigma}_r$.
On peut alors lier les propriétés de préservation de domaine invariant ainsi :

Proposition 6.7. Soit Ω un domaine invariant pour (6.1).

Si le schéma VF préserve invariant le domaine Ω , alors le flux numérique associé préserve invariant le domaine Ω en prenant $\sigma_l = \Delta x_i / \Delta t$ et $\sigma_r = \Delta x_{i+1} / \Delta t$.

Si le flux numérique g préserve invariant le domaine Ω , alors le schéma VF préserve invariant le domaine Ω sous les conditions CFL

$$\sigma_l(U_i^n, U_{i+1}^n) \Delta t \leq \Delta x_i / 2 \quad \text{et} \quad \sigma_r(U_i^n, U_{i+1}^n) \Delta t \leq \Delta x_{i+1} / 2, \quad (6.6)$$

pour tout $i \in \mathbb{Z}$.

Démonstration. Le premier cas se démontre en deux temps. Si on prend $U_{i-1}^n = U_i^n = U_l$ et $U_{i+1}^n = U_r$ dans le schéma VF, la consistance du flux implique que U_i^{n+1} est égal à la première expression de (6.5). On fait de même en prenant cette fois $U_{i-1}^n = U_l$ et $U_i^n = U_{i+1}^n = U_r$ et on obtient bien la propriété pour le flux numérique.

Le deuxième cas se traite en considérant les demi-maillages. Soit

$$\begin{aligned} U_{i-}^{n+1} &= U_i^n - \frac{2\Delta t}{\Delta x_i} (f(U_i^n) - g(U_{i-1}^n, U_i^n)), \\ U_{i+}^{n+1} &= U_i^n - \frac{2\Delta t}{\Delta x_i} (g(U_i^n, U_{i+1}^n) - f(U_i^n)). \end{aligned}$$

On a $U_i^{n+1} = (U_{i-}^{n+1} + U_{i+}^{n+1})/2$. Étant données les conditions CFL sur σ_l et σ_r , on déduit de (6.5) que U_{i-}^{n+1} et U_{i+}^{n+1} appartiennent à Ω . Comme Ω est convexe, $U_i^{n+1} \in \Omega$. \square

Il n'est pas rare que les schémas numériques ne vérifient pas cette condition (même les schémas VF utilisés dans le milieu industriel). En effet, les conditions de calcul sont souvent telles que la solution approchée est à valeurs dans un compact loin du bord de Ω . Ce n'est que lorsque la solution approchée admet des valeurs au voisinage de $\partial\Omega$ que la notion de préservation de domaine invariant devient importante.

6.1.3 Inégalités d'entropie discrètes

Définition 6.8. Soit (η, F) un couple entropie-flux d'entropie pour le système (6.1). On dit que le schéma VF (6.3) est entropique si il existe un flux d'entropie numérique $G(U, V)$ consistant avec le flux d'entropie, c'est-à-dire que $G(U, U) = F(U)$, tel que sous une certaine condition CFL, le schéma VF vérifie

$$\eta(U_i^{n+1}) - \eta(U_i^n) + \frac{\Delta t}{\Delta x_i} (G(U_i^n, U_{i+1}^n) - G(U_{i-1}^n, U_i^n)) \leq 0. \quad (6.7)$$

Comme dans le cas précédent, on va travailler par interfaces :

Définition 6.9. Soit (η, F) un couple entropie-flux d'entropie pour le système (6.1). On dit que le flux numérique g est entropique si il existe un flux d'entropie numérique $G(U, V)$ consistant avec le flux d'entropie, c'est-à-dire que $G(U, U) = F(U)$, tel qu'il existe $-\sigma_l(U_l, U_r) < 0 < \sigma_r(U_l, U_r)$ vérifiant

$$\begin{aligned} \eta(U_-(U_l, U_r, \sigma_l)) - \eta(U_l) - \frac{1}{\sigma_l} (G(U_l, U_r) - F(U_l)) &\leq 0, \\ \eta(U_+(U_l, U_r, \sigma_l)) - \eta(U_r) - \frac{1}{\sigma_r} (F(U_r) - G(U_l, U_r)) &\leq 0. \end{aligned} \quad (6.8)$$

Là aussi, la propriété reste vraie pour tout $-\tilde{\sigma}_l \leq -\sigma_l$ et $\sigma_r \leq \tilde{\sigma}_r$.

On a alors le résultat suivant permettant de relier les deux notions :

Proposition 6.10. Soit (η, F) un couple entropie-flux d'entropie pour le système (6.1). Si le schéma VF (6.3) est entropique, alors le flux numérique associé est entropique en prenant $\sigma_l = \Delta x_i / \Delta t$ et $\sigma_r = \Delta x_{i+1} / \Delta t$.

Si le flux numérique g est entropique, alors le schéma VF est entropique sous les conditions CFL (6.6).

Démonstration. La démonstration est la même que celle de la proposition 6.7, en utilisant en plus la convexité de l'entropie. \square

De nouveau, il n'est pas rare de rencontrer des schémas numériques non entropiques. En effet, cette propriété n'est pas toujours cruciale car la diffusion numérique permet dans la plupart des cas de bien approcher les solutions vérifiant (5.18). De plus, il existe des techniques de correction permettant, à partir d'un schéma non entropique, d'obtenir un schéma entropique.

6.2 Formalisme de Harten, Lax et Van Leer

Le formalisme de Harten, Lax et Van Leer permet d'inclure de nombreux schémas existants. On va d'abord se placer dans un cadre général puis voir plusieurs exemples de tels schémas.

6.2.1 Cadre général

Il se base sur la notion de solveur de Riemann approché, qui correspond à l'utilisation d'une approximation de la solution auto-similaire du problème de Riemann (5.15).

Définition 6.11. Un solveur de Riemann approché pour (5.15) est une fonction $R : \mathbb{R} \times \Omega^2 \rightarrow \Omega$ vérifiant la propriété de consistance pour tout $U \in \Omega$ et $x/t \in \mathbb{R}$

$$R(x/t; U, U) = U.$$

On y associe le flux numérique suivant

$$\begin{aligned} g(U_l, U_r) &= f(U_l) - \int_{-\infty}^0 (R(\xi; U_l, U_r) - U_l) d\xi. \\ &\left(= f(U_r) + \int_0^{\infty} (R(\xi; U_l, U_r) - U_r) d\xi. \right) \end{aligned} \quad (6.9)$$

On appelle le schéma VF associé à un tel flux numérique un schéma de type Godunov.

Proposition 6.12. *Un schéma de type Godunov est conservatif et consistant.*

Démonstration. La propriété de conservation est immédiatement vérifiée et la consistance du schéma est donnée par la consistance du solveur de Riemann approché. \square

En fait, ce type schéma se rapproche de la vision transport-projection du schéma de Godunov, à ceci près que la solution de chaque problème de Riemann est maintenant donnée par R .

Proposition 6.13. *On considère un schéma de type Godunov. Alors*

$$U_i^{n+1} = \frac{1}{\Delta x_i} \left[\int_0^{\Delta x_i/2} R(x/\Delta t; U_{i-1}^n, U_i^n) dx + \int_{-\Delta x_i/2}^0 R(x/\Delta t; U_i^n, U_{i+1}^n) dx \right] \quad (6.10)$$

sous les conditions CFL $\sigma_l \Delta t \leq \Delta x_i/2$ et $\sigma_r \Delta t \leq \Delta x_{i+1}/2$, où σ_l et σ_r vérifient

$$\begin{aligned} R(\xi; U_l, U_r) &= U_l \quad \forall \xi < -\sigma_l, \\ R(\xi; U_l, U_r) &= U_r \quad \forall \xi > \sigma_r. \end{aligned}$$

Démonstration. On a

$$\begin{aligned} &\frac{1}{\Delta x_i} \left[\int_0^{\Delta x_i/2} R(x/\Delta t; U_{i-1}^n, U_i^n) dx + \int_{-\Delta x_i/2}^0 R(x/\Delta t; U_i^n, U_{i+1}^n) dx \right] \\ &= U_i^n + \frac{1}{\Delta x_i} \int_0^{\Delta x_i/2} (R(x/\Delta t; U_{i-1}^n, U_i^n) - U_i^n) dx \\ &\quad + \frac{1}{\Delta x_i} \int_{-\Delta x_i/2}^0 (R(x/\Delta t; U_i^n, U_{i+1}^n) - U_i^n) dx \\ &= U_i^n + \frac{\Delta t}{\Delta x_i} \int_0^{\infty} (R(\xi; U_{i-1}^n, U_i^n) - U_i^n) d\xi + \frac{\Delta t}{\Delta x_i} \int_{-\infty}^0 (R(\xi; U_i^n, U_{i+1}^n) - U_i^n) d\xi \\ &= U_i^n + \frac{\Delta t}{\Delta x_i} (g(U_{i-1}^n, U_i^n) - f(U_i^n)) + \frac{\Delta t}{\Delta x_i} (f(U_i^n) - g(U_i^n, U_{i+1}^n)) \\ &= U_i^n - \frac{\Delta t}{\Delta x_i} (g(U_i^n, U_{i+1}^n) - g(U_{i-1}^n, U_i^n)) = U_i^{n+1}. \end{aligned}$$

On a donc le résultat escompté (la condition CFL a été utilisée dans la deuxième égalité). \square

Proposition 6.14. *Si le solveur de Riemann R préserve invariant le domaine Ω , c'est-à-dire que*

si $U_l, U_r \in \Omega$, alors $R(\xi; U_l, U_r) \in \Omega$ pour tout $\xi \in \mathbb{R}$,

alors le schéma de type Godunov associé à R préserve invariant le domaine Ω , sous la condition CFL de la proposition 6.13.

Démonstration. Ce résultat est immédiat grâce à la convexité de la formule (6.10) (on pourrait aussi utiliser la forme du flux numérique et la proposition 6.7). \square

Proposition 6.15. Soit les fonctions

$$\begin{aligned} G_l(U_l, U_r) &= F(U_l) - \int_{-\infty}^0 (\eta(R(\xi; U_l, U_r)) - \eta(U_l)) d\xi, \\ G_r(U_l, U_r) &= F(U_r) + \int_0^{\infty} (\eta(R(\xi; U_l, U_r)) - \eta(U_r)) d\xi. \end{aligned}$$

Si

$$G_r(U_l, U_r) \leq G_l(U_l, U_r), \quad (6.11)$$

alors toute fonction $G(U_l, U_r) = \alpha G_r(U_l, U_r) + (1 - \alpha) G_l(U_l, U_r)$ pour $\alpha \in [0, 1]$ est un flux d'entropie numérique et le schéma de type Godunov associé à R est entropique sous la condition CFL de la proposition 6.13.

Démonstration. Ici aussi, on pourrait utiliser la proposition 6.10, mais on va plutôt se baser sur (6.10). Par convexité de l'entropie η et par l'inégalité de Jensen, on a

$$\begin{aligned} \eta(U_i^{n+1}) &\leq \frac{1}{\Delta x_i} \int_0^{\Delta x_i/2} \eta(R(x/\Delta t; U_{i-1}^n, U_i^n)) dx \\ &\quad + \frac{1}{\Delta x_i} \int_{-\Delta x_i/2}^0 \eta(R(x/\Delta t; U_i^n, U_{i+1}^n)) dx \\ &\leq \eta(U_i^n) - \frac{\Delta t}{\Delta x_i} (G_l(U_i^n, U_{i+1}^n) - G_r(U_{i-1}^n, U_i^n)) \\ &\leq \eta(U_i^n) - \frac{\Delta t}{\Delta x_i} (G(U_i^n, U_{i+1}^n) - G(U_{i-1}^n, U_i^n)) \end{aligned}$$

grâce à (6.11). Le schéma vérifie donc bien (6.7). \square

6.2.2 Schéma de Godunov

Ce schéma entre dans le formalisme de Harten, Lax et Van Leer. Il se base sur la solution exacte du problème de Riemann (5.15).

Proposition 6.16. Le schéma de Godunov préserve invariant le domaine Ω et il est entropique.

Démonstration. La propriété de préservation du domaine Ω est immédiate grâce à la proposition 6.14 puisque $R \in \Omega$.

Le flux d'entropie numérique associé au schéma de Godunov est $\eta(R(0; U_l, U_r))$, ce qui permet de conclure directement. \square

En pratique, le schéma de Godunov est peu utilisé car il nécessite la résolution exacte du problème de Riemann (5.15). Cette résolution est vite compliquée car les courbes d'ondes dans Ω sont non linéaires et il faut en calculer l'intersection.

6.2.3 Solveurs simples

Les solveurs simples sont des solveurs de Riemann approchés composés de M ondes discontinues de vitesse (ordonnées) σ_k , séparant $M + 1$ états constants U_k : $R(\xi; U_L, U_R) = U_k$ si $\sigma_k < x/t < \sigma_{k+1}$ (avec pour conventions $\sigma_0 = -\infty$, $\sigma_{M+1} = \infty$, $U_0 = U_L$ et $U_M = U_R$). Il est important de noter qu'il n'est pas nécessaire d'avoir $M = N$.

En intégrant le schéma associé sur le maillage décalé, on obtient la condition de conservation

$$\sum_{k=1}^M \sigma_k (U_k - U_{k-1}) = f(U_R) - f(U_L).$$

On peut donc définir naturellement les flux intermédiaires f_k (avec $f_0 = f(U_L)$ et $f_M = f(U_R)$), ce qui donne les relations

$$-\sigma_k (U_k - U_{k-1}) + (f_k - f_{k-1}) = 0$$

que l'on peut assimiler à des relations de saut de Rankine-Hugoniot (mais on ne connaît pas le système associé et en général $f(U_k) \neq f_k$). Le flux numérique est alors donné par

$$g(U_L, U_R) = f_k,$$

où k est tel que $\sigma_k \leq 0 \leq \sigma_{k+1}$. Autrement dit, on a

$$\begin{aligned} g(U_L, U_R) &= f(U_L) + \sum_{k \text{ t.q. } \sigma_k < 0} \sigma_k (U_k - U_{k-1}) \\ &= f(U_R) - \sum_{k \text{ t.q. } \sigma_k > 0} \sigma_k (U_k - U_{k-1}), \end{aligned}$$

ce qui donne

$$g(U_L, U_R) = \frac{1}{2} (f(U_L) + f(U_R)) - \frac{1}{2} \sum_{k=1}^M |\sigma_k| (U_k - U_{k-1}).$$

Passons maintenant en revue quelques solveurs simples.

Schéma décentré

Ce schéma serait le plus simple, il consisterait à n'introduire qu'une seule onde ($M = 1$). Nécessairement, la vitesse de cette onde serait donnée par la relation de Rankine-Hugoniot

$$-\sigma (U_R - U_L) + (f(U_R) - f(U_L)) = 0,$$

ce qui en général n'a pas de solution (N équations pour une seule inconnue !).

Schéma de Rusanov

Ce schéma, déjà présenté dans le cas scalaire, peut être réinterprété comme un schéma de type Godunov basé sur un solveur simple à deux ondes, de vitesse $-a$ et a , où a est une constante positive. Le flux s'écrit alors **[EX]**

$$g(U_l, U_r) = \frac{f(U_l) + f(U_r)}{2} - \frac{1}{2} A(U_l, U_r) (U_r - U_l).$$

Pour peu que $A(U_l, U_r) = \max_{U=U_l, U_r} \max_i |\lambda_i(U)|$, ce schéma préserve invariant Ω et est entropique (sous la condition CFL $\Delta t \leq \inf_i (\Delta x_i) / (2a)$).

Schéma HLL

Ce schéma, présenté en même temps que le formalisme de Harten, Lax et Van Leer, se base lui aussi sur deux ondes, mais de vitesse $a_1 \leq a_2$ *a priori* différentes. Toujours en utilisant les relations précédentes, on déduit le flux numérique suivant : [EX]

$$g(U_l, U_r) = \begin{cases} f(U_l) & \text{si } 0 < a_1, \\ \frac{a_2 f(U_l) - a_1 f(U_r)}{a_2 - a_1} + \frac{a_1 a_2}{a_2 - a_1} (U_r - U_l) & \text{si } a_1 < 0 < a_2, \\ f(U_r) & \text{si } a_2 < 0. \end{cases}$$

Ce schéma aussi est consistant, préserve invariant Ω et est entropique, avec pour condition sur a_1 et a_2

$$a_1 = \min_{U=U_L, U_R} \min_i \lambda_i(U), \quad a_2 = \max_{U=U_L, U_R} \max_i \lambda_i(U).$$

Schéma de Roe

Le schéma de Roe est plus compliqué que les deux précédents, mais il est aussi plus précis en pratique. Il se base sur la construction d'une matrice $A(U_l, U_r) \in \mathbb{R}^{N \times N}$, dite matrice de Roe, qui vérifie pour tout $U_l, U_r \in \Omega$:

1. la matrice $A(U_l, U_r)$ est diagonalisable dans \mathbb{R} ,
2. consistence : $A(U_l, U_l) = f'(U_l)$,
3. $f(U_r) - f(U_l) = A(U_l, U_r)(U_r - U_l)$.

Le solveur simple associé est la solution du problème de Riemann linéaire

$$\begin{cases} \partial_t U + A(U_L, U_R) \partial_x U = 0, & t > 0, x \in \mathbb{R}, \\ U(0, x) = \begin{cases} U_L & \text{si } x < 0, \\ U_R & \text{si } x > 0. \end{cases} \end{cases}$$

Après quelques calculs, on peut en déduire le flux numérique [EX]

$$g(U_l, U_r) = \frac{f(U_l) + f(U_r)}{2} - \frac{1}{2} |A(U_l, U_r)| (U_r - U_l)$$

où $|A| = P \text{diag}(|\lambda_1|, \dots, |\lambda_N|) P^{-1}$, P étant la matrice des vecteurs propres à droite de A et λ_i les valeurs propres de A .

Cette matrice peut en pratique être difficile à trouver. On peut cependant montrer que si le système hyperbolique (6.1) admet une entropie strictement convexe, alors il existe une matrice de Roe.

Ce schéma, bien que plus précis que les précédents, ne préserve pas invariant Ω et n'est pas entropique.

6.3 Schémas de Godunov approchés

Ces schémas sont des versions simplifiées du schéma de Godunov mais ils n'entrent pas dans le formalisme de Harten, Lax et Van Leer. Tout comme le schéma de Godunov, le flux numérique de tels schémas s'écrit

$$g(U_l, U_r) = f(R(0; U_l, U_r)) \quad (6.12)$$

à ceci près que R n'est pas la solution exacte du problème de Riemann.

Il est important de noter que même si $R(0; U_l, U_r) \in \Omega$, le schéma ne préserve pas forcément invariant Ω .

6.3.1 Schéma choc-choc

Cette méthode se base sur une approximation des courbes d'ondes dans la résolution du problème de Riemann : toutes les ondes sont supposées être des discontinuités, les détetes sont donc remplacées par des ondes de choc qui ne vérifient donc pas le critère de Lax. Cette approximation n'est pas aberrante car en général, les courbes d'onde de détente et les courbes d'onde de choc sont très proches.

Avec cette approximation, le schéma ne préserve pas invariant Ω et n'est pas entropique. Cependant, il permet de résoudre des solutions avec des ondes de choc très précisément. Il est important de noter que ce schéma peut aussi approcher des solutions comportant des ondes de détente, grâce à la diffusion numérique.

6.3.2 Schéma VFRoe

Ce schéma se base sur une linéarisation de la résolution du problème de Riemann. Soit un changement de variable $V = \varphi(U)$ et $B(V) = \nabla \varphi(U) \nabla f(U) \nabla \varphi(U)^{-1}$. Le solveur de Riemann définissant le flux numérique (6.12) est la solution du problème de Riemann linéarisé suivant

$$\begin{cases} \partial_t V + B((\varphi(U_L) + \varphi(U_R))/2) \partial_x V = 0, & t > 0, x \in \mathbb{R}, \\ U(0, x) = \begin{cases} U_L & \text{si } x < 0, \\ U_R & \text{si } x > 0. \end{cases} \end{cases}$$

Suivant le changement de variable utilisé, les propriétés du schéma changent. Mais en général, il ne préserve pas invariant Ω et n'est pas entropique (quoique...).

Remarque 14. On peut se demander quel est l'intérêt de schémas qui ne préservent pas invariant Ω et ne sont pas entropique. Ils sont en général plus précis et souvent facile à mettre en œuvre, en particulier le schéma VFRoe. De plus, beaucoup de configurations de simulations ne nécessitent pas ces propriétés, tous les schémas paraissant converger vers la même solution.

Il existe des schémas préservant invariant Ω et entropique dont la précision est comparable au schéma de Godunov (qui est la référence) et la complexité est comparable aux schémas basés sur un solveur linéarisé. Mais en général ils sont définis au cas par cas, suivant le système considéré.

6.4 Cas multidimensionnel

On reprend les notations du cas scalaire et le schéma se dérive de la même manière, pour obtenir :

$$U_K^{n+1} = U_K^n - \frac{\Delta t}{|K|} \sum_{L \in \mathcal{V}(K)} |e_{KL}| g(U_K^n, U_L^n; n_{KL}) \quad (6.13)$$

où le flux numérique g vérifie les propriétés de base suivantes :

- conservation : $g(U, V; n) = -g(V, U; -n)$ pour tout $(U, V; n) \in (\mathbb{R}^N)^2 \times \mathbf{S}^{d-1}$,
- consistence : $g(U, U; n) = f(U) \cdot n$ pour tout $(U, n) \in \mathbb{R}^N \times \mathbf{S}^{d-1}$.

Comme en une dimension, on va s'attacher à étudier les propriétés de préservation de domaine invariant et d'entropie. En fait, on peut étendre ces propriétés quasiment directement, à l'aide du calcul suivant.

De manière analogue aux U^\pm , on définit

$$\bar{U}(U, V, n, \sigma) = U - \frac{1}{\sigma_l(U, V)} (g(U, V; n) - f(U) \cdot n). \quad (6.14)$$

De plus, on a comme propriété classique que $\sum_{L \in \mathcal{V}(K)} |e_{KL}| n_{KL} = 0$, donc

$$\frac{\Delta t}{|K|} \sum_{L \in \mathcal{V}(K)} |e_{KL}| f(U) \cdot n_{KL} = 0.$$

Soit $K \in \mathcal{M}$ et $\alpha_{KL} \in (0, 1)$ tels que $\sum_{L \in \mathcal{V}(K)} \alpha_{KL} = 1$. On peut donc écrire

$$\begin{aligned} U_K^{n+1} &= U_K^n - \frac{\Delta t}{|K|} \sum_{L \in \mathcal{V}(K)} |e_{KL}| g(U_K^n, U_L^n; n_{KL}) + \frac{\Delta t}{|K|} \sum_{L \in \mathcal{V}(K)} |e_{KL}| f(U) \cdot n_{KL} \\ &= \sum_{L \in \mathcal{V}(K)} \alpha_{KL} \left(U_K^n - \frac{\Delta t |e_{KL}|}{\alpha_{KL} |K|} (g(U_K^n, U_L^n; n_{KL}) - f(U_K^n) \cdot n_{KL}) \right) \\ &= \sum_{L \in \mathcal{V}(K)} \alpha_{KL} \bar{U}(U_K^n, U_L^n, n_{KL}, \alpha_{KL} |K| / (\Delta t |e_{KL}|)). \end{aligned}$$

Pour simplifier l'écriture, on peut prendre $\alpha_{KL} = |e_{KL}| / |\partial K|$, ce qui donne

$$U_K^{n+1} = \sum_{L \in \mathcal{V}(K)} \alpha_{KL} \bar{U}(U_K^n, U_L^n, n_{KL}, |K| / (\Delta t |\partial K|)).$$

On obtient donc une combinaison convexe et les propriétés du schéma numérique sont à nouveau sous condition sur \bar{U} (et l'existence de σ) et sous la condition CFL

$$\Delta t \leq \frac{\alpha^2 h}{\max \sigma(U_K, U_L)}.$$

Ainsi, si on considère un schéma préservant les domaines invariants et entropique en 1D pour un système invariant galiléen (donc ces propriétés sont vraies pour toute normale n), alors il s'étend naturellement en multidimension.

Un peu de bibliographie

Tout d'abord, les livres d'Edwige Godlewski et Pierre-Arnaud Raviart : le premier [GR91] porte sur l'analyse et l'approximation des lois de conservation scalaires (difficile à trouver car il n'est plus édité) et le second [GR96] traite des systèmes de lois de conservation (analyse et approximation). Ces deux livres sont très complets et bien détaillés.

Ensuite, on peut citer les livres de Denis Serre, [Ser96a] et [Ser96b], peut-être plus difficiles que les deux précédents car ils sont destinés à un public plus spécialisé (ils existent aussi en version anglaise).

D'un point de vue numérique, on peut citer le livre de Randall J. LeVeque [LeV02], celui d'Eleuterio F. Toro [Tor99] et celui de François Bouchut [Bou04]. Tous ceux-ci concernent principalement les systèmes, pour le cas scalaire on peut se référer au livre de Robert Eymard, Thierry Gallouët et Raphaële Herbin [EGH00].

Par ailleurs, le livre de Joel Smoller [Smo83] est particulièrement pédagogique et contient des développements très intéressants.

Enfin, le livre de Constantine M. Dafermos [Daf05] constitue l'ouvrage de référence à l'heure actuelle, il dresse un état de l'art sur le sujet véritablement impressionnant.

Pour terminer, voici quelques liens de pages et polycopiés sur le sujet :

http://www.ann.jussieu.fr/~despres/BD_fichiers/seism.htm

http://www.ann.jussieu.fr/~perthame/cours_hyp.pdf

<http://www.cmi.univ-mrs.fr/~herbin/PUBLI/bookevol.pdf>

http://www.cmap.polytechnique.fr/~allaire/cours_master.html

<http://www-gm3.univ-mrs.fr/polys/gm3-08/index.html>

Bibliographie

- [Bou04] F. Bouchut. *Nonlinear stability of finite volume methods for hyperbolic conservation laws and well-balanced schemes for sources*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2004. 6.4
- [Daf05] C. M. Dafermos. *Hyperbolic conservation laws in continuum physics*, volume 325 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2005. 6.4
- [EGH00] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In *Handbook of numerical analysis, Vol. VII*, Handb. Numer. Anal., VII, pages 713–1020. North-Holland, Amsterdam, 2000. 6.4
- [GR91] E. Godlewski and P.-A. Raviart. *Hyperbolic systems of conservation laws*, volume 3/4 of *Mathématiques & Applications (Paris) [Mathematics and Applications]*. Ellipses, Paris, 1991. 6.4
- [GR96] E. Godlewski and P.-A. Raviart. *Numerical approximation of hyperbolic systems of conservation laws*, volume 118 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1996. 6.4
- [LeV02] R. J. LeVeque. *Finite volume methods for hyperbolic problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2002. 6.4
- [Ser96a] D. Serre. *Systèmes de lois de conservation. I. Fondations*. [Foundations]. Diderot Editeur, Paris, 1996. Hyperbolicité, entropies, ondes de choc. [Hyperbolicity, entropies, shock waves]. 6.4
- [Ser96b] D. Serre. *Systèmes de lois de conservation. II. Fondations*. [Foundations]. Diderot Editeur, Paris, 1996. Structures géométriques, oscillation et problèmes mixtes. [Geometric structures, oscillation and mixed problems]. 6.4
- [Smo83] J. Smoller. *Shock waves and reaction-diffusion equations*, volume 258 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science]*. Springer-Verlag, New York, 1983. 6.4
- [Tor99] E. F. Toro. *Riemann solvers and numerical methods for fluid dynamics*. Springer-Verlag, Berlin, second edition, 1999. A practical introduction. 6.4