La méthode des volumes finis appliquée aux lois de conservation d'ordre un

Julien Vovelle

▶ **To cite this version:**

Julien Vovelle. La méthode des volumes finis appliquée aux lois de conservation d'ordre un. Master. Université Rennes 1, France. 2012. cel-01872732

HAL Id: cel-01872732
https://cel.hal.science/cel-01872732

Submitted on 12 Sep 2018

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

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

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

La méthode des volumes finis appliquée aux lois de conservation d'ordre un

J. Vovelle

September 12, 2018

Abstract

Dans cette troisième partie du cours, on étudie la méthode des volumes finis appliquée aux lois de conservation scalaires d'ordre un.

1 Rappels : Solution faible - solution entropique

1.1 Solution faible

On considère le problème de Cauchy pour une loi de conservation scalaire d'ordre un :

$$\begin{cases} \frac{\partial u}{\partial t} + \mathrm{div} A(u) = 0 \text{ dans } \Omega \times (0,+\infty), \quad \forall i = 1,\dots,N, \\ \\ u|_{t=0} = u_0 \text{ dans } \Omega. \end{cases} \tag{1}$$

Ici Ω est un ouvert de $\mathbb{R}^d$ ou bien le tore $\mathbb{T}^d$, u est l'inconnue. Le flux $A\colon \mathbb{R} \to \mathbb{R}^d$ est supposé localement Lipschitz. Alors $A(u)$ est L^∞ si u l'est et on définit les solutions faibles de la manière suivante.

Définition 1 (Solution faible) *Soit $u_0 \in L^\infty(\Omega)$. On dit que $u \in L^\infty(\Omega \times (0,+\infty))$ est solution faible de (1) si*

$$\int_0^\infty \int_\Omega \left(u(x,t)\varphi_t(x,t) + A(u)(x,t)\cdot\nabla\varphi(x,t)\right) dxdt + \int_\Omega u_0(x)\varphi(x,0)dx = 0,$$

pour tout $\varphi \in C_c^\infty(\Omega \times [0,+\infty))$.

1.2 Solution entropique

On a vu (*cf* Exercice 8 du Cours no 1) qu'il y a parfois non-unicité des solutions faibles. Parmi toutes les solutions faibles, on va distinguer celles qui dissipent l'entropie.

Définition 2 (Entropie) *Soit $\eta\colon \mathbb{R} \to \mathbb{R}$ et $\Phi\colon \mathbb{R} \to \mathbb{R}^d$ des fonctions localement Lipschitziennes. On dit que (η,Φ) est un couple entropie-flux d'entropie (associé à (1)) si η est convexe et $\Phi'(u) = \eta'(u)A'(u)$ pour presque tout $u \in \mathbb{R}$.*

Une classe particulière de couple entropie-flux d'entropie est constituée par les fonctions

$$\eta(u) = (u-a)^{\pm}, \quad \Phi(u) = \mathrm{sgn}_{\pm}(u-a)(A(u)-A(a)), \tag{2}$$

pour a variant dans $\mathbb{R}$. On a noté ici $\mathrm{sgn}_{\pm}(s)$ pour $\pm\mathbf{1}_{]0,+\infty[}(\pm s)$, soit la fonction dérivée de $s \mapsto s^{\pm}$ pour $s \neq 0$, qui vaut 0 en 0. Si η a la forme donnée dans (2), on appelle η une demi-entropie de Kruzhkov. On peut approcher toute fonction convexe par une combinaison de demi-entropies de Kruzhkov (admis), si bien qu'on se restreint à la classe de couples entropie-flux d'entropie (2) dans la définition suivante.

Définition 3 (Solution faible entropique) *Soit $u_0 \in L^\infty(\Omega)$. On dit qu'une fonction $u \in L^\infty(\Omega \times (0,+\infty))$ est solution faible entropique de (1) si*

$$\int_0^\infty \int_\Omega (\eta(u)\varphi_t + \Phi(u)\cdot\nabla\varphi)\,dxdt + \int_\Omega \eta(u_0)\varphi(0)dx \geq 0, \tag{3}$$

pour tout $\varphi \in C_c^\infty(\Omega \times [0,+\infty))$ positif et pour tout couple entropie-flux d'entropie (η,Φ) de type (2).

Remarquer que (si on fait abstraction de la condition initiale), (3) est l'expression faible de l'inégalité

$$\frac{\partial}{\partial t}\eta(u) + \mathrm{div}(\Phi(u)) \leq 0. \tag{4}$$

Attention, le terme entropie est repris de la physique, en choisissant toutefois un signe opposé : sous réserve de la valeur des conditions aux limites

$$t \mapsto \int_\Omega \eta(u(x,t))dx$$

est décroissante et non croissante comme l'entropie physique.

Exercice 1 *Reprendre l'exercice 8 dans le Cours no 1. Montrer que la solution faible proposée n'est pas entropique puis vérifier que la "bonne solution" proposée est solution entropique.*

Supposons que $\Omega = \mathbb{T}^d$. Le problème (1) est bien posé (dans $L^\infty(\mathbb{T}^d)$ muni de la topologie de $L^1(\mathbb{T}^d)$) pour les solutions entropiques : on a en effet le résultat suivant.

Théorème 4 *[Kruzhkov, 1969] Soit $u_0 \in L^\infty(\mathbb{T}^d)$. Il existe un unique $u \in L^\infty(\Omega \times (0,+\infty))$ solution faible entropique de (1). De plus, u est continue en temps à valeurs L^1 : $u \in C([0,+\infty); L^1(\mathbb{T}^d))$, et si $u^\sharp, u^\flat \in L^\infty(\Omega \times (0,+\infty))$ sont les solutions faibles entropique de (1) associées respectivement aux données initiales $u_0^\sharp, u_0^\flat$, alors*

$$\|(u^\sharp(t) - u^\flat(t))^+\|_{L^1(\mathbb{T}^d)} \leq \|(u^\sharp(s) - u^\flat(s))^+\|_{L^1(\mathbb{T}^d)}, \tag{5}$$

pour tout $0 \leq s \leq t$.

Remarque : Lorsqu'on dit que l'unique solution $u \in L^\infty(\Omega \times (0,+\infty))$ vérifie

$$u \in C([0,+\infty); L^1(\mathbb{T}^d)),$$

on signifie que, quitte à modifier u sur un ensemble de mesure nulle, on obtient un représentant de u qui est continu en temps à valeurs dans $L^1(\mathbb{T}^d)$. D'autre part, on peut remplacer $\mathbb{T}^d$ par $\mathbb{R}^d$ dans les énoncés, prendre alors u_0 dans $L^1 \cap L^\infty(\mathbb{R}^d)$. On peut aussi envisager le cas Ω ouvert borné de $\mathbb{R}^d$, mais alors il faut ajouter des conditions au bord pour bien poser le problème, or la question des conditions au bord pose quelques problèmes spécifiques pour les équations d'ordre un, on n'en parlera pas ici.

Exercice 2 *1. En échangeant les rôles de $u^\sharp$ et $u^\flat$ et en additionnant les résultats, montrer que*

$$\|u^\sharp(t) - u^\flat(t)\|_{L^1(\mathbb{T}^d)} \leq \|u^\sharp(s) - u^\flat(s)\|_{L^1(\mathbb{T}^d)}, \tag{6}$$

pour tout $0 \leq s \leq t$ (principe de contraction L^1). Montrer qu'il implique en particulier l'unicité de la solution entropique.

2. Montrer qu'on a le principe de comparaison *:*

$$u_0^\sharp \leq u_0^\flat \text{ p.p.} \implies u^\sharp(t) \leq u^\flat(t) \text{ p.p.} \tag{7}$$

pour tout $t \geq 0$.

Comme les constantes sont solutions de (1) (correspondant aux données initiales constantes), le principe de comparaison (7) implique

$$m \leq u_0 \leq M \text{ p.p.} \implies m \leq u(t) \leq M \text{ p.p.} \tag{8}$$

En particulier, avec

$$M = \|u_0\|_{L^\infty(\mathbb{T}^d)}, \quad m = -M,$$

on a $\|u(t)\|_{L^\infty(\mathbb{T}^d)} \leq \|u_0\|_{L^\infty(\mathbb{T}^d)}$.

Exercice 3 (Approximation parabolique) *Soit $\varepsilon > 0$ et $u^\varepsilon \in C^2(\mathbb{T}^d \times [0,+\infty))$ solution (classique) de l'équation parabolique*

$$u_t^\varepsilon + \operatorname{div}(A(u^\varepsilon)) - \varepsilon \Delta u^\varepsilon = 0,$$

vérifiant $u^\varepsilon|_{t=0} = u_0$ ($u_0 \in C^2(\mathbb{T}^d)$ donnée). Soit (η, Φ) un couple entropie-flux d'entropie associé à (1). Montrer que

$$\frac{\partial}{\partial t}\eta(u^\varepsilon) + \operatorname{div}(\Phi(u^\varepsilon)) = \varepsilon \Delta \eta(u^\varepsilon) - \varepsilon \eta''(u^\varepsilon)|\nabla u^\varepsilon|^2.$$

En déduire

$$\frac{\partial}{\partial t}\eta(u^\varepsilon) + \operatorname{div}(\Phi(u^\varepsilon)) \leq \varepsilon \Delta \eta(u^\varepsilon).$$

Comparer à (4). Que se passe-t-il lorsque $\varepsilon \to 0$?

2 Méthode des volumes finis

2.1 Flux numériques

On va considérer à partir de maintenant le cas $\Omega = \mathbb{T}^d$. Soit $\mathcal{T}$ un maillage de $\mathbb{T}^d$ de taille h (on reporte au paragraphe 7 du cours no 1 pour les définitions et notations concernant les maillages et schémas numériques). Soit k le pas d'espace. On veut que le schéma numérique

$$u_K^{n+1} = u_K^n - \frac{k}{|K|} \sum_{L \in \mathcal{N}(K)} Q_{K\to L}^n, \quad \forall K \in \mathcal{T}, n \in \mathbb{N}, \tag{9}$$

complété avec la donnée initiale discrète

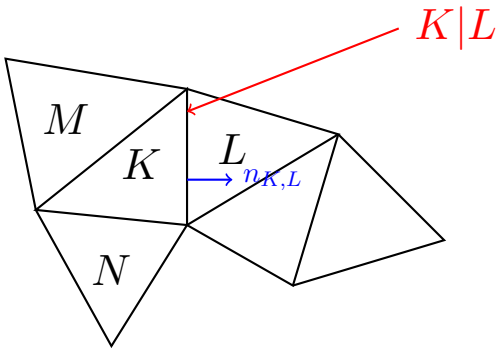
$$u_K^0 = \frac{1}{|K|} \int_K u_0(x)dx, \quad \forall K \in \mathcal{T}, \tag{10}$$

donne une approximation de la solution faible entropique de (1). On rappelle que $Q_{K\to L}^n$ désigne le flux de K vers L sur l'intervalle de temps (t_n, t_{n+1}) et que la propriété de *conservation* impose

$$Q_{K\to L}^n = -Q_{L\to K}^n, \quad \forall K, L \in \mathcal{T}, \quad \forall n \in \mathbb{N}, \tag{11}$$

c'est-à-dire "flux sortant de K entrant en L"$= -$"flux sortant de L entrant en K".

Notons $n_{K,L}$ la normale unitaire sortante à K le long de $K|L$.



Pour (1), le flux sortant de K par $K|L$ est (dans le cas où u est continue par exemple)

$$\int_{K|L} A(u) \cdot n_{K,L} d\sigma(x) \simeq |K|L| A(u) \cdot n_{K,L},$$

où la dernière approximation se justifie en considérant u à peu près constant le long de $K|L$. Pour déterminer $Q_{K\to L}^n$, on va utiliser un schéma à flux monotones : pour $K, L \in \mathcal{T}$ voisins ($K|L \neq \emptyset$) on se donne des fonctions $A_{K\to L} : \mathbb{R}^2 \to \mathbb{R}^d$ vérifiant les propriétés suivantes : ($m < M$ sont fixés)

1. Consistance :
$$A_{K\to L}(v,v) = \int_{K|L} A(v)\cdot n_{K,L}, \tag{12}$$
pour tout $v \in \mathbb{R}$.

2. Régularité : $A_{K\to L}$ est Lipschitz continue en ses deux variables, avec
$$\mathrm{Lip}(A_{K\to L}, [m,M]^2) \leq C|K|L|\mathrm{Lip}(A,[m,M]), \tag{13}$$
pour une certaine constante C (quitte à augmenter $\mathrm{Lip}(A,[m,M])$, on supposera $C=1$ dans la suite), où on a noté
$$\mathrm{Lip}(g,K) = \sup_{x\neq y\in K} \frac{|g(x)-g(y)|}{|x-y|},$$
pour $g\colon \mathbb{R}^p \to \mathbb{R}$ localement lipshitzienne et K compact de $\mathbb{R}^p$.

3. Monotonie : $(v,w) \mapsto A_{K\to L}(v,w)$ est croissante en la première variable v.

On pose alors
$$Q^n_{K\to L} = A_{K\to L}(u^n_K, u^n_L), \quad \forall K,L\in\mathcal{T}, \quad \forall n\in\mathbb{N}. \tag{14}$$
Cela définit un schéma numérique explicite en temps. On peut aussi définir un schéma implicite en temps en posant $Q^n_{K\to L} = A_{K\to L}(u^{n+1}_K, u^{n+1}_L)$. L'analyse des schémas implicites en temps n'est pas plus compliquée que l'analyse des schémas explicites en temps mais on ne traitera que l'explicite ici.

Commentaires : la propriété de consistance est en effet une propriété de consistance du flux numérique avec le flux de l'équation (1). L'hypothèse de régularité est technique (mais nécessaire). L'hypothèse de monotonie est fondamentale. Fixons en effet deux mailles voisines K et L et imaginons que $u^n_K > u^n_L$ au temps $t_n = nk$. Alors
$$Q^n_{K\to L} \geq A_{K\to L}(u^n_L, u^n_L) = \int_{K|L} A(v)\cdot n_{K,L}.$$
Si $A_{K\to L}(u^n_L, u^n_L)$ est vu comme le flux de référence, alors le flux numérique $Q^n_{K\to L}$ est un peu plus important que ce flux de référence. Une manière simple de s'assurer que $A_{K\to L}(u^n_L, u^n_L)$ peut réellement être considéré comme flux de référence, c'est de supposer $A=0$. Dans ce cas, vu le signe "$-$" devant $Q^n_{K\to L}$ dans (9), la valeur de u^{n+1}_K va être abaissée par rapport à celle de u^n_K (attention, pour conclure tout à fait il faut considérer aussi les échanges aves les autres mailles voisines de K).
De même, si $u^n_K < u^n_L$, on va avoir $Q^n_{K\to L} \leq A_{K\to L}(u^n_L, u^n_L)$. En somme, le schéma numérique, en plus de faire passer les flux, tend à modérer les déséquilibres. Or, c'est précisément ce que fait la Loi de Fourier dans les équations d'ordre 2 (paraboliques) ! Comme les équations paraboliques du type de l'équation de la chaleur sont aussi appelées équations de diffusion, on parle ici de *diffusion numérique* : la propriété de monotonie des flux induit de la diffusion numérique. On va revenir sur ce point, mais d'abord étudions le caractère conservatif du schéma.

La propriété de conservation discrète (11) se réecrit

$$A_{K\to L}(u_K^n, u_L^n) = -A_{L\to K}(u_L^n, u_K^n).$$

Comme on envisage toutes les valeurs possibles pour les valeurs u_K^n et u_L^n calculées au temps t_n, on a donc la propriété

4. Conservation :
$$A_{K\to L}(v, w) = -A_{L\to K}(w, v),$$
pour tout $v, w \in \mathbb{R}$.

Avec la propriété de monotonie 3. (appliquée en échangeant K et L), on obtient immédiatement le lemme suivant

Lemme 5 *L'application $(v, w) \mapsto A_{K\to L}(v, w)$ est décroissante en la deuxième variable w.*

2.2 Exemples

Tous les exemples qui suivent vérifient les points 1. à 4. requis pour les flux numériques (le vérifier).

2.2.1 Schéma de Godunov par interface

Les flux $A_{K\to L}$ sont définies par

$$A_{K\to L}(v, w) = \begin{cases} \min\limits_{v\le z\le w} |K|L|A(z)\cdot n_{K,L} & \text{si } v \le w, \\ \max\limits_{w\le z\le v} |K|L|A(z)\cdot n_{K,L} & \text{si } w \le v. \end{cases}$$

2.2.2 Schéma de Engquist - Osher

En notant $a := A'$, les flux $A_{K\to L}$ sont définies par

$$A_{K\to L}(v, w) = \frac{|K|L|}{2}\left(A(v)\cdot n_{K,L} + A(w)\cdot n_{K,L} - \int_v^w |a(z)\cdot n_{K,L}|dz\right).$$

Plus généralement, les schémas appelés "flux splitting scheme" in english sont donnés par

$$A_{K\to L}(v, w) = B_{K\to L}(v) + C_{K\to L}(w)$$

où $|K|L|A(z)\cdot n_{K,L} = B_{K\to L}(z) + C_{K\to L}(z)$, avec B croissante, C décroissante, toutes deux étant $|K|L|\mathrm{Lip}(A)$-lipschtziennes.

2.2.3 Schéma de Lax-Friedrichs (modifié)

Je l'écris dans le cas de la dimension 1 (voir le Cours no 2 pour les notations concernant le maillage en dimension 1, on utilise les mêmes ici). Les flux $A_{i\to i\pm1}$ sont définies par

$$A_{i\to i\pm1}(v,w)=\frac{1}{2}\left\{\pm\left(A(v)+A(w)\right)+D(v-w)\right\}.$$

Ils sont monotones sur $[m,M]^2$ si

$$D\geq \mathrm{Lip}(A,[m,M]).$$

2.3 Équation de transport

Revenons à l'exemple de l'équation de transport, paragraphe 6, Cours no 1. On a vu que le schéma aux différences finies centrées ne peut pas donner une approximation de la solution. Qu'obtient-on par la méthode des volumes finis ?

On a $A(u)=u$ et, en dimension 1, (9) se réécrit

$$u_i^{n+1}=u_i^n-\frac{k}{h_i}(Q^n_{i\to i+1}+Q^n_{i\to i-1}).$$

Avec $Q^n_{i\to j}=A_{i\to j}(u_i^n,u_j^n)$, $j=i\pm1$, on obtient

$$u_i^{n+1}=u_i^n-\frac{k}{h_i}(A_{i\to i+1}(u_i^n,u_{i+1}^n)+A_{i\to i-1}(u_i^n,u_{i-1}^n)),\tag{15}$$

pour tout $i=0,\dots,L-1$. Rappelons que, aux extrémités, pour $i=0$ ou $i=L-1$, les conditions de périodicité définissent

$$u^n_{-1}:=u^n_{L-1},\quad u^n_L:=u^n_0,\quad \forall n\in\mathbb{N}.\tag{16}$$

Etudions les exemples précédents (attention $n_{i,j}=\pm1$ selon que $j=i\pm1$ et la taille des interfaces est toujours 1 car ils sont réduits à des points) : schéma de Godunov :

$$A_{i\to i+1}(v,w)=\left\{\begin{array}{ll}\min\limits_{v\leq z\leq w} z & \text{si } v\leq w\\ \max\limits_{w\leq z\leq v} z & \text{si } w\leq v\end{array}\right\}=v.$$

De même, $A_{i\to i-1}(v,w)=-w$: (15) donne

$$u_i^{n+1}=u_i^n-\frac{k}{h_i}(u_i^n-u_{i-1}^n)=(1-\lambda_i)u_i^n+\lambda_i u_{i-1}^n,\quad \lambda_i:=\frac{k}{h_i}.\tag{17}$$

Schéma d'Engquist-Osher : on calcule (attention à la convention $\int_v^w=-\int_w^v$ si $v>w$)

$$A_{i\to i+1}(v,w)=\frac{1}{2}\left(v+w-\int_v^w dz\right)=\frac{1}{2}(v+w+v-w)=v.$$

De même, $A_{i\to i-1}(v,w)=-w$: on obtient le même schéma numérique que Godunov.

Exercice 4 *Montrer que les schémas de Godunov et Engquist-Osher coïncident dès que A est une fonction monotone.*

Schéma de Lax-Friedrichs : on calcule

$$A_{i\to i\pm1}(v,w)=\frac{1}{2}(\pm(v+w)+D(v-w)),$$

et (15) donne

$$u_i^{n+1}=u_i^n-\frac{k}{h_i}\frac{u_{i+1}^n-u_{i-1}^n}{2}-D\frac{k}{h_i}\frac{2u_i^n-u_{i-1}^n-u_{i+1}^n}{2}. \tag{18}$$

Voici maintenant des essais numériques : sur les deux figures 1-2, on représente la solution exacte en bleu, la solution donnée par le schéma de Engquist-Osher en rouge et la solution donnée par le schéma de Lax-Friedrichs en mauve au temps $t=0.4$. En espace $h_i=h$ pour tout i. Le rapport $\frac{k}{h}=\frac{1}{2}$. Sur la figure 1, on a utilisé 200 mailles (donc $h=0.05$).

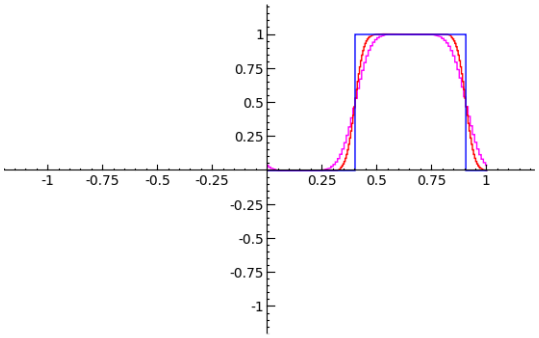


Figure 1: Solutions à $t=0.4$, $h=0.05$

Sur la figure 2, on a utilisé 700 mailles (donc $h=0.014$).

Bien sûr, les résultats sont incomparablement meilleurs que ceux donnés par le schéma aux différences finies centrées (*cf* Cours no 1). Noter les phénomènes de diffusion numérique (qui ont pour effet de lisser le créneau). Ils sont plus important dans le schéma de Lax-Friedrichs. Toujours au sujet de la diffusion numérique, observer que le schéma de Lax-Friedrichs s'écrit exactement comme le schéma du Cours no 1 (le schéma aux différences finies centrées qui produit de si mauvais résultats) *plus* une correction

$$-D\frac{k}{h}\frac{2u_i^n-u_{i-1}^n-u_{i+1}^n}{2}.$$

Or cette correction est l'approximation par différences finies de $\frac{Dkh}{2}u_{xx}(x_i,t_n)$, d'après la formule de Taylor à l'ordre 2, qui donne

$$u(x_{i-1},t_n)+u(x_{i+1},t_n)-2u(x_i,t_n)=h^2u_{xx}(x_i,t_n)+o(h^2),\quad x_i:=ih.$$

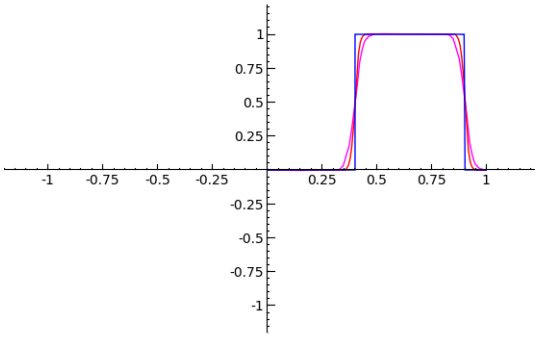


Figure 2: Solutions à $t = 0.4$, $h = 0.014$

Par conséquent, le schéma de Lax-Friedrichs est en réalité une approximation de l'équation parabolique

$$u_t + u_x - \varepsilon u_{xx} = 0,$$

où on a posé $\varepsilon = \frac{Dh}{2}$. On poursuit l'étude de la diffusion numérique avec celle de la stabilité dans le paragraphe suivant.

2.4 Diffusion numérique

2.4.1 Transport par Godunov

Notre but est d'évaluer la diffusion numérique dans le schéma (17), dans le cas $h_i = h$, $\forall i$. Pour cela on observe que, pour une équation de type

$$u_t + au_x - \eta u_{xx} = 0,$$

on peut

1. Faire la translation (*i.e.* transport par $-a$) $x := x - at$ pour obtenir l'équation

$$u_t - \eta u_{xx} = 0,$$

   sans terme d'ordre un.

2. Identifier le coefficient de diffusion : η.

On veut procéder de même pour (17). L'équivalent de l'étape 1. ci-dessus va alors être discret : on va utiliser le schéma de Godunov pour l'équation $v_t - v_x = 0$, à savoir (après calcul des flux numériques)

$$v_i^{n+1} = v_i^n - \frac{k}{h_i}(-v_{i+1}^n + v_i^n) = (1-\lambda)v_i^n + \lambda v_{i+1}^n, \quad \lambda := \frac{k}{h}. \tag{19}$$

Corrigeons donc chaque étape de (17) par (19) : on remplace la valeur u_i^{n+1} calculée grâce à (17) par la valeur

$$\begin{aligned}(1-\lambda)u_i^{n+1}+\lambda u_{i+1}^{n+1} &= (1-\lambda)((1-\lambda)u_i^n+\lambda u_{i-1}^n)+\lambda((1-\lambda)u_{i+1}^n+\lambda u_i^n)\\ &= [(1-\lambda)^2+\lambda^2]u_i^n+\lambda(1-\lambda)(u_{i-1}^n+u_{i+1}^n)\\ &= (1-2\lambda+2\lambda^2)u_i^n+\lambda(1-\lambda)(u_{i-1}^n+u_{i+1}^n)\\ &= u_i^n-\lambda(1-\lambda)(2u_i^n-u_{i-1}^n-u_{i+1}^n).\end{aligned}$$

On obtient donc

$$\frac{u_i^{n+1}-u_i^n}{k}-h(1-\lambda)\frac{u_{i-1}^n+u_{i+1}^n-2u_i^n}{h^2}=0,$$

qui est une approximation de l'équation $u_t-\eta u_{xx}=0$ avec $\eta:=h(1-\lambda)$. Attention cependant, si le schéma (17) induit de la diffusion numérique, le schéma (19) aussi ! (Avec la même amplitude). En faisant la translation discrete en espace, on a donc multiplié par 2 la valeur du coefficient de diffusion, soit une valeur finale

$$\eta:=\frac{h(1-\lambda)}{2}.$$

On peut comparer η au coefficient de diffusion $\varepsilon=\frac{Dh}{2}$ obtenu pour le schéma de Lax-Friedrichs :

$$\frac{\eta}{\varepsilon}=\frac{1-\lambda}{D}.$$

Sous la condition de CFL $\lambda\leq 1$ et l'hypothèse $D\geq 1$ (nécessaire pour assurer la monotonie du flux Lax-Friedrichs), on a donc $\eta\leq\varepsilon$: la diffusion induite par le schéma de Lax-Friedrichs est toujours plus importante.

Exercice 5 (Diffusion numérique non-linéaire/linéaire) *Voici, sur la figure 3, le calcul de la solution du problème de Cauchy pour l'équation de Burgers non visqueuse*

$$u_t+(u^2/2)_x=0, \tag{20}$$

avec donnée

$$u_0(x)=\begin{cases}2 & si\ x<0\\ -1 & si\ x>0\end{cases}. \tag{21}$$

En rouge est représentée la solution calculée au temps $T=1$ avec $h=0.01$, $\frac{k}{h}=\frac{1}{2}$, en bleu la donnée initiale (attention, ne pas tenir compte de l'échelle de l'axe des abscisses, qui est fausse).

1. *Interpréter en terme de particules/champ de vitesse le problème (20)-(21). Montrer que la solution entropique faible est "un choc se déplaçant à vitesse $1/2$", soit $u(x,t)=u_0(x-t/2)$.*
2. *A l'oeil nu, on n'observe pas de diffusion numérique sur la figure 3, au contraire de ce qu'il se produit dans l'équation de transport. Pouvez-vous expliquer cela ?*

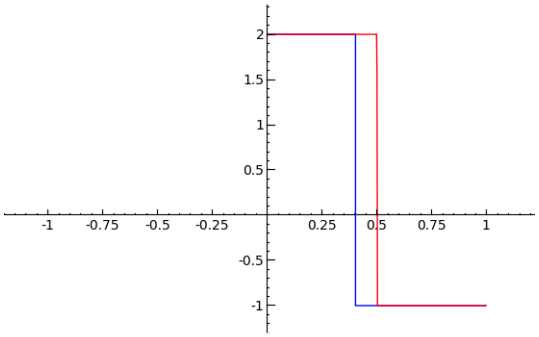


Figure 3: Solution à $T = 1$, $h = 0.01$

2.4.2 Solutions constantes

Les constantes sont solutions du problème continu (1) (correspondant à des conditions initiales constantes), elles sont aussi solutions du schéma volume fini (9)-(10)-(14). En effet, par la propriété de consistance des flux numériques, on a, pour toute constante $v \in \mathbb{R}$, pour tout $K \in \mathcal{T}$, d'après la formule de Green-Gauss,

$$\begin{aligned}\sum_{L\in\mathcal{N}(K)} A_{K\to L}(v,v) &= \sum_{L\in\mathcal{N}(K)} \int_{K|L} A(v)\cdot n_{K,L} d\sigma \\ &= \int_{\partial K} A(v)\cdot n d\sigma = \int_K \operatorname{div}(A(v))dx = 0. \end{aligned} \tag{22}$$

On a noté σ la mesure de surface sur ∂K et n la normale unitaire sortante à ∂K qui est définie σ-presque-partout et coïncide avec $n_{K,L}$ sur $K|L$.

2.4.3 Comparaison discrète

Notation On note $u_{h,k}$ la fonction constante par morceaux définie (p.p.) par

$$u_{h,k}(x,t) = u_K^n, \quad (x,t) \in K \times [t_n, t_{n+1}[. \tag{23}$$

Une conséquence de la monotonie des flux est une propriété de stabilité dans L^∞, elle-même conséquence de la proposition suivante. On se donne $m < M$ et on note $u \in L^\infty(\mathbb{T}^d, [m,M])$ si $u\colon \mathbb{T}^d \to \mathbb{R}$ est mesurable avec $m \le u \le M$ p.p.

Proposition 6 (Principe de comparaison discret) *Soit $u_{h,k}^\sharp$ et $u_{h,k}^\flat$ deux solutions du schéma volume fini (9)-(10)-(14) correspondant aux données initiales $u_0^\sharp$, $u_0^\flat$ dans l'espace $L^\infty(\mathbb{T}^d, [m,M])$ respectivement. Sous la condition dite de CFL (pour Courant - Friedrichs - Lewy)*[1]

$$k\frac{|\partial K|}{|K|}\operatorname{Lip}(A,[m,M]) \le 1, \quad \forall K \in \mathcal{T}, \tag{24}$$

[1] voir http://en.wikipedia.org/wiki/Courant-Friedrichs-Lewy_condition

on a le principe de comparaison discret

$$u_0^\sharp \leq u_0^\flat \ p.p. \Longrightarrow u_{h,k}^\sharp \leq u_{h,k}^\flat \ p.p.$$

Corollaire 7 (Stabilité L^∞) *Soit $u_0 \in L^\infty(\mathbb{T}^d, [m, M])$. On suppose la condition de CFL (24) vérifiée. On a alors l'estimation L^∞*

$$\|u_{h,k}\|_{L^\infty(\mathbb{T}^d \times (0,+\infty))} \leq \|u_0\|_{L^\infty(\mathbb{T}^d)},$$

pour la solution $u_{h,k}$ du schéma volume fini (9)-(10)-(14).

Preuve du Corollaire 7 : On applique la proposition 6 avec $u_{h,k}^\sharp = u_{h,k}$ et $u_{h,k}^\flat = M_1 = \|u_0\|_{L^\infty(\mathbb{T}^d)}$, qui est une solution particulière de (9)-(10)-(14) correspondant à la donnée initiale $u_0^\flat = M_1$. On obtient $u_{h,k} \leq \|u_0\|_{L^\infty(\mathbb{T}^d)}$ p.p. De même, $u_{h,k} \geq -\|u_0\|_{L^\infty(\mathbb{T}^d)}$ p.p., d'où le résultat. ■

Preuve de la Proposition 6 : montrons d'abord que, sous la condition de CFL (24), l'équation (9)-(14) s'écrit

$$u_K^{n+1} = H\left(u_K^n, \{u_L^n\}_{L \in \mathcal{N}(K)}\right), \tag{25}$$

où H est croissante en ses arguments. On a

$$u_K^{n+1} = u_K^n - \frac{k}{|K|} \sum_{L \in \mathcal{N}(K)} A_{K \to L}(u_K^n, u_L^n). \tag{26}$$

Comme $A_{K \to L}$ est décroissante en sa deuxième variable d'après le lemme 5, le membre de gauche est croissant en les $\{u_L^n\}_{L \in \mathcal{N}(K)}$. Pour la croissance en u_K^n, on remarque que $\mathrm{Id} - \alpha f$ est croissante sur $[m, M]$ dès que

$$\alpha \mathrm{Lip}(f, [m, M]) \leq 1,$$

ce qui donne (25) sous (24) en utilisant l'hypothèse de régularité des flux numériques et l'identité

$$\sum_{L \in \mathcal{N}(K)} |K|L| = |\partial K|.$$

On montre alors par récurrence sur n l'inégalité

$$m \leq u_K^{\sharp,n} \leq u_K^{\flat,n} \leq M,$$

pour tout $K \in \mathcal{T}$. C'est vérifié pour $n = 0$ et, en admettant le rang n, on a, d'après (25),

$$u_K^{\sharp,n+1} = H\left(u_K^{\sharp,n}, \{u_L^{\sharp,n}\}_{L \in \mathcal{N}(K)}\right) \leq H\left(u_K^{\flat,n}, \{u_L^{\flat,n}\}_{L \in \mathcal{N}(K)}\right) = u_K^{\flat,n+1},$$

ainsi que

$$m = H\left(m, (m)_{L \in \mathcal{N}(K)}\right) \leq H\left(u_K^{\sharp,n}, \{u_L^{\sharp,n}\}_{L \in \mathcal{N}(K)}\right) = u_K^{\sharp,n+1},$$

et, de même, $u_K^{\flat,n+1} \leq M$, ce qui donne le rang $n+1$. ■

2.4.4 Contraction L^1 discrète

Un autre corollaire de la Proposition 6 est le principe de contraction L^1 :

Proposition 8 (Principe de contraction L^1) *Soit $u_{h,k}^\sharp$ et $u_{h,k}^\flat$ deux solutions des équations du schéma volume fini (9)-(10)-(14) correspondant aux données initiales $u_0^\sharp$, $u_0^\flat$ dans l'espace $L^\infty(\mathbb{T}^d,[m,M])$ respectivement. Sous la condition de CFL (24), on a le principe de contraction L^1 :*

$$\|u_{h,k}^\sharp(t_n)-u_{h,k}^\flat(t_n)\|_{L^1(\mathbb{T}^d)} \leq \|u_0^\sharp-u_0^\flat\|_{L^1(\mathbb{T}^d)},$$

pour tout $n\in\mathbb{N}$.

Preuve : On montre par récurrence sur n

$$\|u_{h,k}^\sharp(t_{n+1})-u_{h,k}^\flat(t_{n+1})\|_{L^1(\mathbb{T}^d)} \leq \|u_{h,k}^\sharp(t_n)-u_{h,k}^\flat(t_n)\|_{L^1(\mathbb{T}^d)}. \tag{27}$$

On déduit en effet de (27) que

$$\begin{aligned}
\|u_{h,k}^\sharp(t_n)-u_{h,k}^\flat(t_n)\|_{L^1(\mathbb{T}^d)} &\leq \|u_{h,k}^\sharp(0)-u_{h,k}^\flat(0)\|_{L^1(\mathbb{T}^d)}\\
&= \sum_{K\in\mathcal{T}} |K||u_K^{\sharp,0}-u_K^{\flat,0}|\\
&= \sum_{K\in\mathcal{T}} |K|\left|\frac{1}{|K|}\int_K (u_0^\sharp(x)-u_0^\flat(x))dx\right|\\
&\leq \sum_{K\in\mathcal{T}} \int_K \left|u_0^\sharp(x)-u_0^\flat(x)\right| dx\\
&= \int_{\mathbb{T}^d} \left|u_0^\sharp(x)-u_0^\flat(x)\right| dx = \|u_0^\sharp-u_0^\flat\|_{L^1(\mathbb{T}^d)}.
\end{aligned}$$

Pour montrer (27), on introduit l'ensemble $\mathcal{M}\subset L^1(\mathbb{T}^d)$ des fonctions constantes par mailles $K\in\mathcal{T}$ et l'opérateur

$$T\colon \mathcal{M}\to\mathcal{M},$$

qui à $(u_K^n)_{K\in\mathcal{T}}$ associe $(u_K^{n+1})_{K\in\mathcal{T}}$ et on applique le lemme suivant à T.

Lemme 9 (Lemme de Crandall-Tartar) *Soit $(X,\mathcal{A},\mu)$ un espace mesuré, soit $\mathcal{M}$ un sous-ensemble de $L^1(X,\mu)$ stable par l'opération*

$$(f,g)\mapsto f\top g := \max(f,g),$$

et soit un opérateur $T\colon\mathcal{M}\to\mathcal{M}$ monotone, au sens

$$f\leq g \text{ p.p.} \implies T(f)\leq T(g) \text{ p.p.}, \quad \forall f,g\in\mathcal{M}, \tag{28}$$

et préservant l'intégrale, au sens où

$$\int_X T(f)d\mu = \int_X f d\mu, \quad \forall f\in L^1(X,\mu).$$

On a alors

$$\|(T(f)-T(g))^+\|_{L^1(X,\mu)} \leq \|(f-g)^+\|_{L^1(X,\mu)}, \quad \forall f,g \in \mathcal{M}.$$

En admettant provisoirement ce lemme, et en l'appliquant à $(u_K^{\sharp,n}, u_K^{\flat,n})$ et $(u_K^{\flat,n}, u_K^{\sharp,n})$ respectivement, on obtient

$$\|(u_{h,k}^{\sharp}(t_{n+1}) - u_{h,k}^{\flat}(t_{n+1}))^+\|_{L^1(\mathbb{T}^d)} \leq \|(u_{h,k}^{\sharp}(t_n) - u_{h,k}^{\flat}(t_n))^+\|_{L^1(\mathbb{T}^d)}$$

et

$$\|(u_{h,k}^{\flat}(t_{n+1}) - u_{h,k}^{\sharp}(t_{n+1}))^+\|_{L^1(\mathbb{T}^d)} \leq \|(u_{h,k}^{\flat}(t_n) - u_{h,k}^{\sharp}(t_n))^+\|_{L^1(\mathbb{T}^d)},$$

d'où (27) en additionnant les deux résultats. ■

Preuve du Lemme 9 : On note que $(f-g)^+ = f\top g - g$. On a $f \leq f\top g$ *p.p.* donc

$$T(f) \leq T(f\top g) \text{ p.p.},$$

par la propriété de monotonie (28). De même, $T(g) \leq T(f\top g)$ p.p. donc

$$T(f)\top T(g) \leq T(f\top g) \text{ p.p.}$$

En intégrant cette relation, on obtient

$$\int_X T(f)\top T(g) d\mu \leq \int_X T(f\top g) d\mu = \int_X f\top g d\mu,$$

où on a utilisé le fait que T préserve l'intégrale. En soustrayant la relation $\int_X T(g)d\mu = \int_X T(g)d\mu$, on obtient

$$\int_X (T(f)\top T(g) - T(g)) d\mu \leq \int_X (f\top g - g) d\mu,$$

soit $\|(T(f)-T(g))^+\|_{L^1(X,\mu)} \leq \|(f-g)^+\|_{L^1(X,\mu)}$. ■

2.4.5 Estimation d'énergie

Pour l'équation parabolique

$$u_t + \operatorname{div}(A(u)) - \varepsilon\Delta u = 0 \text{ dans } \mathbb{T}^d \times (0,+\infty),$$

on a, en multipliant par u (supposée régulière) et en intégrant sur $\mathbb{T}^d \times (0,t)$,

$$\int_0^t \frac{1}{2}\frac{d}{dt}\int_{\mathbb{T}^d} u^2 dxds + \int_0^t \int_{\mathbb{T}^d} u\operatorname{div}(A(u)) dxds - \varepsilon \int_0^t \int_{\mathbb{T}^d} u\Delta u dxds = 0. \tag{29}$$

En développant

$$u\operatorname{div}(A(u)) = uA'(u)\cdot\nabla u = B'(u)\cdot\nabla u = \operatorname{div}(B(u)), \quad B'(u) := uA'(u),$$

on obtient

$$\int_{\mathbb{T}^d} u\mathrm{div}(A(u))dx = \int_{\mathbb{T}^d} \mathrm{div}(B(u))dx = 0,$$

par périodicité[2]. D'autre part, une intégration par parties[3] donne

$$-\varepsilon \int_{\mathbb{T}^d} u\Delta u = \varepsilon \int_{\mathbb{T}^d} \nabla u \cdot \nabla u dx = \varepsilon \int_{\mathbb{T}^d} |\nabla u|^2 dx.$$

De (29), on déduit donc *l'estimation d'énergie* suivante :

$$\frac{1}{2}\int_{\mathbb{T}^d} |u(x,t)|^2 dx + \varepsilon \int_0^t \int_{\mathbb{T}^d} |\nabla u|^2 dx \leq \frac{1}{2}\int_{\mathbb{T}^d} |u_0|^2 dx.$$

En se restreignant aux termes issus de la diffusion, on a en particulier l'estimation

$$\varepsilon \int_0^t \int_{\mathbb{T}^d} |\nabla u|^2 dx \leq \frac{1}{2}\int_{\mathbb{T}^d} |u_0|^2 dx.$$

L'objet de la proposition suivante est une estimation d'énergie similaire dans le cadre discret (on va utiliser cette estimation d'énergie dans la preuve de convergence de la méthode). La diffusion numérique étant beaucoup plus compliquée que la diffusion représentée par le terme $-\varepsilon\Delta u$, on obtient un énoncé plus complexe.

Proposition 10 (Estimation d'énergie) *Soit $u_0 \in L^\infty(\mathbb{T}^d, [m, M])$. Soit $u_{h,k}$ la solution du schéma volume fini (9)-(10)-(14), soit*

$$\mathcal{E}(t_N) = \frac{1}{2}\sum_{n=0}^{N-1} k \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \int_{u_L^n}^{u_K^n} \{A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(z,z)\}dz.$$

Sous la condition de CFL renforcée : il existe $\xi \in]0,1]$ tel que

$$k\frac{|\partial K|}{|K|}\mathrm{Lip}(A,[m,M]) \leq 1-\xi, \quad \forall K \in \mathcal{T}, \tag{30}$$

on a l'estimation d'énergie

$$\frac{1}{2}\|u_{h,k}(t_N)\|^2_{L^2(\mathbb{T}^d)} + \xi\mathcal{E}(t_N) \leq \frac{1}{2}\|u_0\|^2_{L^2(\mathbb{T}^d)}, \tag{31}$$

pour tout $N \geq 1$.

Remarque : le terme $\mathcal{E}(t_N)$ est positif ; c'est une somme de termes positifs. En effet, par monotonie de $A_{K\to L}$, on a

$$\begin{aligned} A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(z,z) &\geq A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(u_K^n, z) \\ &\geq A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(u_K^n, u_L^n) = 0, \end{aligned} \tag{32}$$

[2] on utilise la formule $\int_{\mathbb{T}^d} \lambda\mathrm{div}(\psi)dx = -\int_{\mathbb{T}^d} \nabla\lambda\cdot\psi$ avec $\lambda = 1$ et $\psi = B(u)$

[3] de même avec $\lambda = u$ et $\psi = -\nabla u$

pour $u_L^n \le z \le u_K^n$. De même,

$$A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(z,z) \le 0,$$

si $u_K^n \le z \le u_L^n$. En particulier, on a toujours

$$\int_{u_L^n}^{u_K^n} A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(z,z)dz \ge 0.$$

Preuve de la Proposition 10 : En prenant $v = u_K^n$ dans (22), on a

$$\sum_{L\in\mathcal{N}(K)} A_{K\to L}(u_K^n, u_K^n) = 0,$$

si bien que (9)-(14) se réécrit

$$\frac{u_K^{n+1} - u_K^n}{k} + \frac{1}{|K|}\sum_{L\in\mathcal{N}(K)} (A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(u_K^n, u_K^n)) = 0. \tag{33}$$

Soit $N \in \mathbb{N}^*$. On multiplie (33) par u_K^n. En sommant $\sum_n k \sum_K |K|[\cdots]$, on obtient

$$E_1 + E_2 = 0, \tag{34}$$

où

$$E_1 = \sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}} |K| u_K^n (u_K^{n+1} - u_K^n),$$

et

$$E_2 = \sum_{n=0}^{N-1} k \sum_{K\in\mathcal{T}}\sum_{L\in\mathcal{N}(K)} u_K^n \{A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(u_K^n, u_K^n)\}.$$

En utilisant la formule

$$a(b-a) = \frac{b^2-a^2}{2} - \frac{(a-b)^2}{2},$$

avec $a = u_K^n$, $b = u_K^{n+1}$ dans E_1 de manière à faire une intégration discrète, on obtient

$$\begin{aligned} E_1 &= \frac{1}{2}\sum_{K\in\mathcal{T}} |K||u_K^N|^2 - \frac{1}{2}\sum_{K\in\mathcal{T}} |K||u_K^0|^2 - \frac{1}{2}\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}} |K||u_K^{n+1} - u_K^n|^2 \\ &= \frac{1}{2}\|u_{h,k}(t_N)\|^2_{L^2(\mathbb{T}^d)} - \frac{1}{2}\|u_{h,k}(0)\|^2_{L^2(\mathbb{T}^d)} - \frac{1}{2}\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}} |K||u_K^{n+1} - u_K^n|^2. \end{aligned} \tag{35}$$

On compensera le dernier terme (qui a un mauvais signe) à l'aide de la condition de CFL et du terme E_2. Pour réécrire E_2, on utilise le

Lemme 11 *Soit une fonction* $a\colon \mathcal{T}^2 \to \mathbb{R}$, *alors*

$$\sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} a_{K,L} = \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} a_{K,L} + a_{L,K} \tag{36}$$

Preuve du lemme : pour K^*, L^* voisins l'un de l'autre, figurent a_{K^*,L^*} et a_{L^*,K^*} dans la somme de gauche : d'une part lorsque $K = K^*$, $L = L^*$, d'autre part lorsque $K = L^*$ et $L = K^*$. ■

D'après (36), on a

$$E_2 = \frac{1}{2} \sum_{n=0}^{N-1} k \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} u_K^n \{A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(u_K^n, u_K^n)\} \\ + u_L^n \{A_{L\to K}(u_L^n, u_K^n) - A_{L\to K}(u_L^n, u_L^n)\}.$$

En utilisant la propriété de *conservation*

$$A_{K\to L}(v, w) = -A_{L\to K}(w, v),$$

on obtient

$$E_2 = \frac{1}{2} \sum_{n=0}^{N-1} k \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} u_K^n \{A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(u_K^n, u_K^n)\} \\ - u_L^n \{A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(u_L^n, u_L^n)\}. \tag{37}$$

Soit maintenant $\psi_{K\to L}$ une primitive de $z \mapsto z\frac{d}{dz} A_{K\to L}(z, z)$. Par intégration par partie et introduction du terme $A_{K\to L}(v, w)$, on a

$$\psi_{K\to L}(v) - \psi_{K\to L}(w) = v\{A_{K\to L}(v, v) - A_{K\to L}(v, w)\} \\ - w\{A_{K\to L}(w, w) - A_{K\to L}(v, w)\} + \int_w^v \{A_{K\to L}(v, w) - A_{K\to L}(z, z)\} dz.$$

En prenant $w = u_L^n$, $v = u_K^n$, on en déduit

$$E_2 = \frac{1}{2} \sum_{n=0}^{N-1} k \left\{ \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \int_{u_L^n}^{u_K^n} \{A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(z, z)\} dz + R_2^n \right\} \tag{38}$$

où

$$R_2^n = -\frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \psi_{K\to L}(u_K^n) - \psi_{K\to L}(u_L^n).$$

On vérifie que $\psi_{K\to L}(u_K^n)$ est anti-symétrique en (K, L), si bien que d'après (36) (lue de droite à gauche),

$$R_2^n = -\sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \psi_{K\to L}(u_K^n).$$

Ce terme est nul car

$$\sum_{L\in\mathcal{N}(K)} \psi_{K\to L}(z) = 0,$$

pour tout z et pour tout K d'après (22). On a donc obtenu

$$E_2 = \mathcal{E}(t_N). \tag{39}$$

Revenons maintenant à l'estimation du terme de droite dans le membre de droite de (35). D'après l'équation du schéma (33) et l'inégalité de Cauchy-Schwarz, on a

$$\begin{aligned}
&\left|u_K^{n+1} - u_K^n\right|^2 \\
&\qquad \le \frac{k^2}{|K|^2} \sum_{L\in\mathcal{N}(K)} |K|L| \sum_{L\in\mathcal{N}(K)} \frac{1}{|K|L|} |A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(u_K^n, u_K^n)|^2 \\
&\qquad\qquad = \frac{k^2|\partial K|}{|K|^2} \sum_{L\in\mathcal{N}(K)} \frac{1}{|K|L|} |A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(u_K^n, u_K^n)|^2.
\end{aligned}$$

Sous la condition de CFL (30), et d'après (36), on a donc

$$\frac{1}{2} \sum_{n=0}^{N-1} \sum_{K\in\mathcal{T}} |K| \left|u_K^{n+1} - u_K^n\right|^2 \le (1-\xi) E_2^*, \tag{40}$$

avec

$$E_2^* := \frac{1}{2\mathrm{Lip}(A)} \sum_{n=0}^{N} k \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \frac{1}{|K|L|} |A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(u_K^n, u_K^n)|^2.$$

Reste à montrer $E_2^* \le \mathcal{E}(t_N)$. Supposons par exemple $u_K^n \ge u_L^n$: on applique le lemme suivant (démontré plus loin) à $B(z) := A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(u_K^n, z + u_L^n)$.

Lemme 12 *Soit $B\colon \mathbb{R}_+ \to \mathbb{R}$ une fonction lipschitzienne croissante telle que $B(0) = 0$. On a*

$$\int_0^r B(z)dz \ge \frac{1}{2\mathrm{Lip}(B)} B(r)^2,$$

pour tout $r \ge 0$.

En prenant $r = u_K^n - u_L^n$, en utilisant la propriété de régularité 2. des flux numériques, et (32), on obtient

$$\begin{aligned}
&\int_{u_L^n}^{u_K^n} \{A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(z,z)\}dz \\
&\qquad\qquad \ge \frac{1}{2|K|L|\mathrm{Lip}(A)} |A_{K\to L}(u_K^n, u_L^n) - A_{K\to L}(u_K^n, u_K^n)|^2.
\end{aligned} \tag{41}$$

Si $u_L^n \geq u_K^n$, appliquer le Lemme 12 à $B(z) := A_{K\to L}(z + u_K^n, u_L^n) - A_{K\to L}(u_K^n, u_L^n)$. On en déduit $E_2^* \leq \mathcal{E}(t_N)$. De (34), (35) et (40), on conclutt donc que

$$\frac{1}{2}\|u_{h,k}(t_N)\|^2_{L^2(\mathbb{T}^d)} + \mathcal{E}(t_N) \leq (1-\xi)\mathcal{E}(t_N) + \frac{1}{2}\|u_{h,k}(0)\|^2_{L^2(\mathbb{T}^d)}.$$

Comme $\|u_{h,k}(0)\|_{L^2(\mathbb{T}^d)} \leq \|u_0\|_{L^2(\mathbb{T}^d)}$ (utiliser l'inégalité de Jensen), on a finalement (31). ■

Preuve du Lemme 12 : On a $0 \leq B$ car B est croissante et nulle en 0 ainsi que $0 \leq B'(r) \leq \mathrm{Lip(B)}$ donc

$$\frac{B(r)^2}{2} = \int_0^r B(z)B'(z)dz \leq \mathrm{Lip}(B)\int_0^r B(z)dz,$$

pour tout $r \geq 0$. ■

Remarque 13 *Au cours de la preuve on a aussi démontré une estimation H^1 discrète sur les dérivées en temps de $u_{h,k}$. En effet, d'après (40), on a*

$$\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}} |K|\left|u_K^{n+1} - u_K^n\right|^2 \leq \frac{1-\xi}{\xi}\|u_0\|^2_{L^2(\mathbb{T}^d)}, \tag{42}$$

pour tout $N \geq 1$.

3 Inégalité entropique

Soit $u_0 \in L^\infty(\mathbb{T}^d)$ une donnée initiale fixée. Le problème (1) admet une unique solution faible entropique (paragraphe 1). Dans quelle mesure est-elle bien approchée par la méthode volume fini ? Exception faite de la dernière étape de la preuve, on va démontrer une estimation d'erreur (en norme $L^\infty_t L^1_x$) dite d'ordre $h^{1/4}$ (Théorème 16).

3.1 Inégalité entropique discrète

On considère ici un couple entropie-flux d'entropie (η, ϕ) de la forme (2). Pour fixer les idées, choisissons la partie positive :

$$\eta(u) = (u-a)^+, \quad \phi(u) = \mathrm{sgn}_+(u-a)(A(u) - A(a)). \tag{43}$$

On a aussi l'expression plus algbrique

$$\eta(u) = u\top a - a, \quad \phi(u) = A(u\top a) - A(a), \quad u\top a := \max(u, a).$$

Posons

$$\Phi_{K\to L}(v, w) = A_{K\to L}(v\top a, w\top a) - A_{K\to L}(a, a). \tag{44}$$

On vérifie facilement que $\Phi_{K\to L}(v, w)$ est un flux numérique associé à ϕ, c'est-à-dire qu'on a consistance avec $|K|L|\phi(u)\cdot n_{K,L}$, régularité, monotonie et conservativité ; $\Phi_{K\to L}(v, w)$ est le flux entropique numérique associé à η.

Proposition 14 (Inégalité entropique discrète) *Soit $u_0 \in L^\infty(\mathbb{T}^d)$ telle que $m \leq u_0 \leq M$ p.p. Soit $u_{h,k}$ solution du schéma volume fini (9)-(10)-(14). Alors, sous la condition de CFL (24), $\eta(u_{h,k})$ vérifie l'inégalité entropique*

$$\frac{\eta(u_K^{n+1}) - \eta(u_K^n)}{k} + \frac{1}{|K|} \sum_{L \in \mathcal{N}(K)} \Phi_{K\to L}(u_K^n, u_L^n) \leq 0, \tag{45}$$

pour tout $K \in \mathcal{T}$, $n \in \mathbb{N}$.

Preuve : On part de la propriété de monotonie (25) (qui découle de la monotonie des flux numériques et de la condition de CFL), soit

$$u_K^{n+1} = H\left(u_K^n, (u_L^n)_{L\in\mathcal{N}(K)}\right),$$

avec H monotone croissante en ses arguments. En particulier,

$$u_K^{n+1} \leq H\left(u_K^n \top a, (u_L^n \top a)_{L\in\mathcal{N}(K)}\right).$$

Les constantes sont solutions du schéma d'après (22), on a donc aussi

$$a = H\left(a, (a)_{L\in\mathcal{N}(K)}\right) \leq H\left(u_K^n \top a, (u_L^n \top a)_{L\in\mathcal{N}(K)}\right).$$

Par conséquent, $u_K^{n+1} \top a \leq H\left(u_K^n \top a, (u_L^n \top a)_{L\in\mathcal{N}(K)}\right)$. En soustrayant a,

$$\eta(u_K^{n+1}) \leq H\left(u_K^n \top a, (u_L^n \top a)_{L\in\mathcal{N}(K)}\right) - H\left(a, (a)_{L\in\mathcal{N}(K)}\right).$$

En développant

$$H\left(u_K^n \top a, (u_L^n \top a)_{L\in\mathcal{N}(K)}\right) = u_K^n \top a - \sum_{L\in\mathcal{N}(K)} A_{K\to L}(u_K^n \top a, u_L^n \top a),$$

on obtient (45). ■

3.2 Inégalité entropique approchée

On fixe de nouveau le couple (η, ϕ) comme en (43). On va faire l'hypothèse suivante sur le maillage (qui signifie qu'il n'y a pas de mailles trop petites ni de mailles trop allongées). On rappelle que $h = \sup_{K\in\mathcal{T}}(\mathrm{diam}(K))$. On suppose qu'il existe $\alpha > 0$ tel que

$$\alpha h^d \leq |K|, \tag{46a}$$

$$|\partial K| \leq \frac{1}{\alpha} h^{d-1}, \tag{46b}$$

pour tout $K \in \mathcal{T}$.

Théorème 15 (Inégalité entropique approchée) *Soit $u_0 \in L^\infty(\mathbb{T}^d)$ telle que $m \leq u_0 \leq M$ p.p. Soit $u_{h,k}$ solution du schéma volume fini (9)-(10)-(14). Sous la condition de CFL renforcée (30) et sous l'hypothèse (46), on a l'inégalité entropique approchée*

$$\int_0^\infty \int_{T^d} (\eta(u_{h,k})\varphi_t + \Phi(u_{h,k}) \cdot \nabla\varphi)\, dxdt + \int_{T^d} \eta(u_{0,h})\varphi(0)dx \\ \geq -\int_0^\infty \int_{T^d} \{|\varphi_t| + |\nabla\varphi|\} d\mu_{h,k} \quad (47)$$

pour tout $\varphi \in C_c^\infty(\Omega \times [0, +\infty))$ positif, où $\mu_{h,k}$ est une mesure positive sur $\mathbb{T}^d \times [0, +\infty)$ satisfaisant l'estimation

$$\mu_{h,k}(\mathbb{T}^d \times [0, t_N]) \leq C t_N^{1/2}(k^{1/2} + h^{1/2}), \quad (48)$$

pour tout $N \geq 1$, où C est une constante ne dépendant que de $\|u_0\|_{L^2(\mathbb{T}^d)}$, ξ, α, d et $\mathrm{Lip}(A, [m, M])$.

On a noté $u_{0,h}$ la donnée initiale discrète définie par

$$u_{0,h}(x) = u_K^0 = \frac{1}{K} \int_K u_0(x)dx,$$

pour tout $K \in \mathcal{T}$.

Preuve : Soit $\varphi \in C_c^\infty(\Omega \times [0, +\infty))$ positif. En développant le membre de gauche de (47), on a

$$\begin{aligned} T_1 &:= \int_0^\infty \int_{T^d} \eta(u_{h,k})\varphi_t dxdt + \int_{T^d} \eta(u_{0,h})\varphi(0)dx \\ &= \sum_{n\geq 0} \sum_{K\in\mathcal{T}} \int_{t_n}^{t_{n+1}} \int_K \eta(u_K^n)\varphi_t dxdt + \sum_{K\in\mathcal{T}} \int_K \eta(u_K^0)\varphi(0,x)dx, \\ &= \sum_{n\geq 0} \sum_{K\in\mathcal{T}} |K|\eta(u_K^n)\{\varphi_K(t_{n+1}) - \varphi_K(t_n)\} + \sum_{K\in\mathcal{T}} |K|\eta(u_K^0)\varphi_K(t_0), \end{aligned}$$

pour le terme en temps (*i.e.* correspondant à la partie dérivée en temps de l'équation), en posant $\varphi_K(t) := \frac{1}{|K|}\int_K \varphi(x,t)dx$. Pour le terme en espace,

$$\begin{aligned}
T_2 &:= \int_0^\infty \int_{T^d} \phi(u_{h,k}) \cdot \varphi dx dt \\
&= \sum_{n\geq 0} \sum_{K\in\mathcal{T}} \int_{t_n}^{t_{n+1}} \int_K \phi(u_K^n) \cdot \varphi dx dt \\
&= \sum_{n\geq 0} \sum_{K\in\mathcal{T}} \int_{t_n}^{t_{n+1}} \phi(u_K^n) \cdot \int_{\partial K} \varphi\, n d\sigma(x) dt \\
&= \sum_{n\geq 0} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \int_{t_n}^{t_{n+1}} \phi(u_K^n) \cdot \int_{K|L} \varphi\, n_{K,L} d\sigma(x) dt \\
&= \sum_{n\geq 0} k \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} |K|L|\phi(u_K^n) \cdot n_{K,L}\varphi_{K|L}^n,
\end{aligned}$$

en posant

$$\varphi_{K|L}^n := \frac{1}{k|K|L|} \int_{t_n}^{t_{n+1}} \int_{K|L} \varphi d\sigma(x) dt.$$

Etudions T_1 tout d'abord : par intégration par partie discrète,

$$\begin{aligned}
T_1 &= -\sum_{n\geq 0} \sum_{K\in\mathcal{T}} |K|\{\eta(u_K^n) - \eta(u_K^{n+1})\}\varphi_K(t_{n+1}) \\
&= -\sum_{n\geq 0} \sum_{K\in\mathcal{T}} |K|\{\eta(u_K^n) - \eta(u_K^{n+1})\}\varphi_K^n + \varepsilon_1(\varphi), \qquad (49)
\end{aligned}$$

avec $\varphi_K^n := \frac{1}{k|K|}\int_{t_n}^{t_{n+1}} \int_K \varphi dx dt$ et

$$\varepsilon_1(\varphi) := \sum_{n\geq 0} \sum_{K\in\mathcal{T}} |K|\{\eta(u_K^n) - \eta(u_K^{n+1})\}\{\varphi_K(t_{n+1}) - \varphi_K^n)\}.$$

D'autre part, $\Phi_{K\to L}(u_K^n, u_L^n)$ étant anti-symétrique en (K,L), tandis que $\varphi_{K|L}^n$ est symétrique, on a d'après (36)

$$\sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \Phi_{K\to L}(u_K^n, u_L^n)\varphi_{K|L}^n = 0.$$

Par consistance du flux entropique numérique $\Phi_{K\to L}$, on en déduit

$$T_2 = -\sum_{n\geq 0} k \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \{\Phi_{K\to L}(u_K^n, u_L^n) - \Phi_{K\to L}(u_K^n, u_K^n)\}\varphi_{K|L}^n,$$

et

$$T_2 = -\sum_{n\geq 0} k \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \{\Phi_{K\to L}(u_K^n, u_L^n) - \Phi_{K\to L}(u_K^n, u_K^n)\}\varphi_K^n + \varepsilon_2(\varphi), \qquad (50)$$

où

$$\varepsilon_2(\varphi) = \sum_{n\geq 0} k \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \{\Phi_{K\to L}(u_K^n, u_L^n) - \Phi_{K\to L}(u_K^n, u_K^n)\}\{\varphi_{K|L}^n - \varphi_K^n\}.$$

Par définition, *cf* (44), le flux entropique numérique $\Phi_{K\to L}$ vérifie la même propriété d'annulation (22) que le flux numérique :

$$\sum_{L\in\mathcal{N}(K)} \Phi_{K\to L}(v, v) = 0, \quad \forall v \in \mathbb{R}.$$

En prenant $v = u_K^n$, et d'après l'inégalité entropique discrète (45),

$$\frac{\eta(u_K^{n+1}) - \eta(u_K^n)}{k} + \frac{1}{|K|} \sum_{L\in\mathcal{N}(K)} \{\Phi_{K\to L}(u_K^n, u_L^n) - \Phi_{K\to L}(u_K^n, u_K^n)\} \leq 0. \tag{51}$$

En multipliant (51) par φ_K^n en sommant sur $n \in \mathbb{N}$, $K \in \mathcal{T}$, (49) et (50) donnent

$$T_1 + T_2 \geq \varepsilon_1(\varphi) + \varepsilon_2(\varphi).$$

Il reste maintenant à estimer les termes d'erreur $\varepsilon(\varphi)$. Par Fubini, on a

$$\int_{t_n}^{t_{n+1}} \{\varphi(t_{n+1}) - \varphi(t)\}dt = \int_{t_n}^{t_{n+1}} \int_t^{t_{n+1}} \varphi_t(s)dsdt = \int_{t_n}^{t_{n+1}} (t_{n+1} - s)\varphi_t(s)ds,$$

d'où

$$\begin{aligned} |\varphi_K(t_{n+1}) - \varphi_K^n| &= \frac{1}{k|K|} \left| \int_{t_n}^{t_{n+1}} \int_K (\varphi(x, t_{n+1}) - \varphi(x, t))dxdt \right| \\ &\leq \frac{1}{|K|} \int_{t_n}^{t_{n+1}} \int_K |\varphi_t(x, t)|\, dxdt \end{aligned}$$

et, en utilisant le fait que η est 1-lipschitzienne,

$$|\varepsilon_1(\varphi)| \leq \int_0^\infty \int_{\mathbb{T}^d} |\varphi_t| d\mu_{h,k}^1,$$

avec

$$\langle \psi, \mu_{h,k}^1 \rangle := \sum_{n\geq 0} \sum_{K\in\mathcal{T}} |u_K^n - u_K^{n+1}| \int_{t_n}^{t_{n+1}} \int_K \psi(x, t)dxdt,$$

pour tout $\psi \in C_c(\mathbb{T}^d \times [0, +\infty))$. D'après l'estimation d'énergie en temps (42) et l'inégalité de Cauchy-Schwarz,

$$\begin{aligned} \mu_{h,k}^1(\mathbb{T}^d \times [0, t_N]) &= \sum_{n=0}^{N-1} \sum_{K\in\mathcal{T}} k|K||u_K^n - u_K^{n+1}| \\ &\leq \left| \sum_{n=0}^{N-1} \sum_{K\in\mathcal{T}} |K||u_K^n - u_K^{n+1}|^2 \right|^{1/2} \left| \sum_{n=0}^{N-1} \sum_{K\in\mathcal{T}} k^2|K| \right|^{1/2} \\ &\leq \left| \frac{1-\xi}{\xi} \right|^{1/2} \|u_0\|_{L^2(\mathbb{T}^d)} \left| kt_N|\mathbb{T}^d| \right|^{1/2} \leq \frac{\|u_0\|_{L^2(\mathbb{T}^d)}}{\sqrt{\xi}} t_N^{1/2} k^{1/2}, \end{aligned}$$

car $|\mathbb{T}^d| = 1$. De même, pour les termes en espace, on a

$$\begin{aligned}|\varphi_{K|L} - \varphi_K| &\leq \frac{1}{|K|L||K|} \int_{K|L} \int_K |\varphi(x) - \varphi(y)| dx d\sigma(y) \\ &\leq \frac{1}{|K|L||K|} \int_{K|L} \int_K \int_0^1 |\nabla\varphi(ry + (1-r)x)||x-y| dr dx d\sigma(y),\end{aligned}$$

donc

$$|\varepsilon_2(\varphi)| \leq \int_0^\infty \int_{\mathbb{T}^d} |\nabla\varphi| d\mu_{h,k}^2,$$

où

$$\begin{aligned}\langle \psi, \mu_{h,k}^2 \rangle := &\sum_{n\geq 0} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} |\Phi_{K\to L}(u_K^n, u_K^n) - \Phi_{K\to L}(u_K^n, u_L^n)| \\ &\times \frac{1}{|K|L||K|} \int_{t_n}^{t_{n+1}} \int_{K|L} \int_K \int_0^1 \psi(ry + (1-r)x, t)|x-y| dr dx d\sigma(y) dt,\end{aligned}$$

pour tout $\psi \in C_c(\mathbb{T}^d \times [0, +\infty))$. Si $x \in K$, $y \in K|L$, alors $|x-y| \leq h$, si bien que

$$\begin{aligned}\mu_{h,k}^2(\mathbb{T}^d \times [0, t_N]) &\leq h \sum_{n=0}^{N-1} k \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} |\Phi_{K\to L}(u_K^n, u_K^n) - \Phi_{K\to L}(u_K^n, u_L^n)| \\ &\leq h \left| \sum_{n=0}^{N-1} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} k|K|L| \right|^{1/2} \\ &\quad \times \left| \sum_{n=0}^{N-1} k \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \frac{1}{|K|L|} |\Phi_{K\to L}(u_K^n, u_K^n) - \Phi_{K\to L}(u_K^n, u_L^n)|^2 \right|^{1/2}.\end{aligned} \tag{52}$$

Le premier facteur dans (52) s'écrit

$$h \left| \sum_{n=0}^{N-1} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} k|K|L| \right|^{1/2} = h \left| t_N \sum_{K\in\mathcal{T}} |\partial K| \right|^{1/2}.$$

Par (46), on a donc

$$h \left| \sum_{n=0}^{N-1} \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} k|K|L| \right|^{1/2} \leq h t_N^{1/2} \left| \sum_{K\in\mathcal{T}} \frac{h^{d-1}}{\alpha} \frac{1}{\alpha h^d} |K| \right|^{1/2} \leq \frac{h^{1/2} t_N^{1/2}}{\alpha}.$$

Pour estimer le second facteur dans (52), on utilise l'estimation suivante (démontrée ci-dessous):

$$\sum_{n=0}^{N-1} k \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \frac{1}{|K|L|} |\Phi_{K\to L}(u_K^n, u_K^n) - \Phi_{K\to L}(u_K^n, u_L^n)|^2$$
$$\leq 2\mathrm{Lip}(A) \sum_{n=0}^{N-1} k \sum_{K\in\mathcal{T}} \sum_{L\in\mathcal{N}(K)} \int_{u_L^n}^{u_K^n} \{A_{K\to L}(u_K^n, u_K^n) - A_{K\to L}(z,z)\} dz. \quad (53)$$

Par l'estimation d'énergie (31), on en déduit finalement

$$\mu_{h,k}^2(\mathbb{T}^d \times [0, t_N]) \leq \left(\frac{\mathrm{Lip}(A) t_N}{\xi}\right)^{1/2} \|u_0\|_{L^2(\mathbb{T}^d)} h^{1/2}.$$

On conclut en posant $\mu_{h,k} = \mu_{h,k}^1 + \mu_{h,k}^2$. Il reste à montrer (53). Supposons par exemple $u_L^n \geq a \geq u_K^n$ (les autres cas sont plus faciles à traiter). On a alors

$$|\Phi_{K\to L}(u_K^n, u_L^n) - \Phi_{K\to L}(u_K^n, u_K^n)| = |A_{K\to L}(a, u_L^n) - A_{K\to L}(a,a)|.$$

On a (*cf.* (41), en remplaçant u_K^n par a)

$$|A_{K\to L}(a, u_L^n) - A_{K\to L}(a,a)|^2$$
$$\leq \frac{1}{2\mathrm{Lip}(A)|K|L|} \int_a^{u_L^n} \{A_{K\to L}(z,z) - A_{K\to L}(a, u_L^n)\} dz. \quad (54)$$

Par monotonie de $A_{K\to L}$,

$$A_{K\to L}(z,z) - A_{K\to L}(a, u_L^n) \leq A_{K\to L}(z,z) - A_{K\to L}(u_K^n, u_L^n),$$

et comme $A_{K\to L}(z,z) - A_{K\to L}(u_K^n, u_L^n) \geq 0$ pour $z \in [u_K^n, u_L^n]$, on a finalement

$$\int_a^{u_L^n} \{A_{K\to L}(z,z) - A_{K\to L}(a, u_L^n)\} dz \leq \int_a^{u_L^n} \{A_{K\to L}(z,z) - A_{K\to L}(u_K^n, u_L^n)\} dz$$
$$\leq \int_{u_K^n}^{u_L^n} \{A_{K\to L}(z,z) - A_{K\to L}(u_K^n, u_L^n)\} dz.$$

On conclut alors avec (54). ■

3.3 Estimation d'erreur

Pour conclure, on admet d'abord le résultat suivant : soit $W^{1,1}(\mathbb{T}^d)$ l'espace de Sobolev des fonctions L^1 à dérivées L^1 sur $\mathbb{T}^d$. Si $u_0 \in W^{1,1}(\mathbb{T}^d)$, alors

$$\|u_{0,h} - u_0\|_{L^1(\mathbb{T}^d)} \leq C \|u_0\|_{W^{1,1}(\mathbb{T}^d)} h, \quad (55)$$

où la constante C dépend de la dimension d et de la constante de régularité α du maillage (*cf* (46)) seulement. Ce résultat (non trivial mais pas trop compliqué à démontrer non plus) s'étend au cas $u_0 \in BV(\Omega)$ (non détaillé ici). Comme η est 1-lipschitzienne, on déduit de (55) et du Théorème 15, que $u_{h,k}$ vérifie les *inégalités entropiques approchées*

$$\int_0^\infty \int_{T^d} (\eta(u_{h,k})\varphi_t + \Phi(u_{h,k}) \cdot \nabla\varphi)\, dxdt + \int_{T^d} \eta(u_0)\varphi(0)dx \\ \geq -\int_0^\infty \int_{T^d} \{|\varphi_t| + |\nabla\varphi|\} d\mu_{h,k} - \int_{\mathbb{T}^d} \varphi d\mu_0, \quad (56)$$

où $\mu_{h,k}$ vérifie (57) et μ_0 est une mesure positive sur $\mathbb{T}^d$ satisfaisant l'estimation

$$\mu_0(\mathbb{T}^d) \leq C\|u_0\|_{W^{1,1}(\mathbb{T}^d)} h, \quad (57)$$

où la constante C dépend de la dimension d et de la constante de régularité α du maillage. De (56), on déduit alors (voir [CH99, EGH95a, EGGH98, EGH95b]) l'estimation d'erreur suivante.

Théorème 16 (Estimation d'erreur) *Soit $u_0 \in L^\infty(\mathbb{T}^d, [m, M])$ avec $u_0 \in W^{1,1}(\mathbb{T}^d)$. Soit $u_{h,k}$ solution du schéma volume fini (9)-(10)-(14) et soit $u \in L^\infty(\mathbb{T}^d \times (0, +\infty))$ la solution faible entropique de (1). Sous la condition de CFL renforcée (30) et sous l'hypothèse (46), on a l'estimation d'erreur*

$$\|u_{h,k}(t) - u(t)\|_{L^1(\mathbb{T}^d)} \leq Ch^{1/4}, \quad \forall 0 \leq t \leq T, \quad (58)$$

où la constante C dépend de $\|u_0\|_{L^\infty(\mathbb{T}^d)}$, $\|u_0\|_{W^{1,1}(\mathbb{T}^d)}$, ξ, α, d, $\mathrm{Lip}(A, [m, M])$ et T.

Exercice 6 *Montrer qu'on peut améliorer (58) pour avoir une estimation en $Ch^{1/2}$.*

4 Solution des exercices

4.1 Exercice 1

Pour $\rho(x,t) = \mathbf{1}_{x<0}$, (η, ϕ) couple entropie-flux d'entropie et $\varphi \in C_c^\infty(\mathbb{R} \times (0, +\infty))$ positive (attention $\varphi|_{t=0}$ est nulle), on calcule

$$\int_0^\infty \int_{\mathbb{R}} \eta(\rho)\varphi_t + \phi(\rho)\varphi_x dxdt = \int_0^\infty (\phi(1) - \phi(0))\varphi(0,t)dt.$$

Le flux $A(\rho) = \rho(1-\rho)$ est une fonction convexe. En prenant $\eta = -A$, on a $\phi' = -|A'|^2$ donc ϕ est strictement décroisante et $\phi(1) < \phi(0)$: ρ n'est donc pas solution entropique. La "bonne" solution

$$\rho(x,t) = \begin{cases} 1 & \text{si } x < -v_{\max}t \\ \dfrac{1}{2} - \dfrac{x}{2v_{\max}t} & \text{si } -v_{\max}t < x < v_{\max}t \\ 0 & \text{si } x > v_{\max}t \end{cases}$$

est C^1 par morceaux dans $\mathbb{R}\times(0,+\infty)$, ce qui est une régularité suffisante pour utiliser le lemme de dérivation des fonctions composées et justifier le calcul suivant :

$$\begin{aligned}0&=\eta'(\rho)(\rho_t+A(\rho)_x)\\&=\eta'(\rho)\rho_t+\eta'(\rho)A'(\rho)\rho_x\\&=\eta'(\rho)\rho_t+\phi'(\rho)\rho_x\\&=\eta(\rho)_t+\phi(\rho)_x.\end{aligned}$$

En utilisant la continuité en temps de ρ, à valeur $L^1(\mathbb{R})_{\text{loc}}$, on prouve que l'inégalité entropique faible est vérifié. C'est en fait une égalité ici, et dans cette exemple ce sont les solutions qui ne dissipent pas l'entropie qui sont rejetés.

4.2 Exercice 2

Le point 1. est évident. On obtient l'unicité en prenant $s=0$. Pour le point 2. on utilise l'équivalence

$$u^\sharp\le u^\flat \text{ p.p. } \Leftrightarrow \|(u^\sharp-u^\flat)^+\|_{L^1(\mathbb{T}^d)}=0.$$

4.3 Exercice 3

On montre

$$\frac{\partial}{\partial t}\eta(u^\varepsilon)+\operatorname{div}(\Phi(u^\varepsilon))=\varepsilon\Delta\eta(u^\varepsilon)-\varepsilon\eta''(u^\varepsilon)|\nabla u^\varepsilon|^2$$

en multipliant l'équation en u^ε par $\eta'(u^\varepsilon)$, en utilisant la règle de dérivation des fonctions composées et l'identité

$$\operatorname{div}(\lambda\Psi)=\lambda\operatorname{div}(\Psi)+\nabla\lambda\cdot\Psi,$$

avec $\lambda=\eta'(u^\varepsilon)$, $\Psi=\nabla u^\varepsilon$. On en déduit

$$\frac{\partial}{\partial t}\eta(u^\varepsilon)+\operatorname{div}(\Phi(u^\varepsilon))\le\varepsilon\Delta\eta(u^\varepsilon),$$

puisque $\eta''\ge 0$ (η convexe). Si $u^\varepsilon\to u$ dans $L^1_{\text{loc}}(\mathbb{T}^d\times(0,+\infty))$ avec (u^ε) bornée L^∞, alors $\eta(u^\varepsilon)$ et $\phi(u^\varepsilon)$ convergent vers $\eta(u)$ et $\phi(u)$ dans $L^1_{\text{loc}}(\mathbb{T}^d\times(0,+\infty))$ respectivement (par convergence dominée). En particulier, $\eta(u^\varepsilon)$ converge (vers $\eta(u)$) au sens des distributions, de même pour $\Delta\eta(u^\varepsilon)$, donc $\varepsilon\Delta\eta(u^\varepsilon)$ converge vers 0. A la limite $\varepsilon\to 0$, on obtient donc (4). Moyennant le problème de convergence des données initiales, la limite u est solution entropique.

4.4 Exercice 4

Il suffit de le vérifier en apliquant les définitions.

4.5 Exercice 5

1. A $t = 0$ des particules de vitesse $+2$ sont disposées sur l'axe des $x < 0$, des particules de vitesse -1 sont situées sur l'axe des $x > 0$. On vérifie que la solution entropique faible est "un choc se déplaçant à vitesse $1/2$". Soit $u(x,t) = u_0(x - t/2)$, soit les sous-ensembles de $\mathbb{R} \times (0, +\infty)$

$$\omega^+ = \{x > t/2\}, \quad \omega^- = \{x < t/2\}, \quad \Gamma = \{x = t/2\},$$

et soit ν la normale unitaire sortante à ω^+. En notant

$$y = (x,t), \quad Z(u) := \begin{pmatrix} \phi(u) \\ \eta(u) \end{pmatrix},$$

on a, pour $\varphi \in C_c^\infty(\mathbb{R} \times (0, +\infty))$,

$$\begin{aligned} &\int_0^\infty \int_{\mathbb{R}} \eta(u)\varphi_t + \phi(u)\varphi_x dx dt \\ &= \int_0^\infty \int_{\mathbb{R}} Z(u) \cdot \nabla_y \varphi dy \\ &= \iint_{\omega^+} Z(u^+) \cdot \nabla_y \varphi dy + \iint_{\omega^-} Z(u^-) \cdot \nabla_y \varphi dy, \end{aligned}$$

en notant $u^+ = -1$, $u^- = 2$. Comme la normale unitaire sortante à ω^- est $-\nu$, on en déduit

$$\int_0^\infty \int_{\mathbb{R}} \eta(u)\varphi_t + \phi(u)\varphi_x dx dt = \int_\Gamma (Z(u^+) - Z(u^-)) \cdot \nu \varphi d\sigma(y)$$

A un facteur multiplicatif près, $\nu = (-2,\, 1)^t$, donc

$$(Z(u^+) - Z(u^-)) \cdot \nu = 2(\phi(2) - \phi(-1)) - (\eta(2) - \eta(-1)).$$

Prenons maintenant (η, ϕ) demi-entropie de Kruzhkov :

$$\eta(u) = (u - a)^+, \quad \phi(u) = \operatorname{sgn}_+(u - a)(A(u) - A(a)), \quad A(u) = u^2/2.$$

On calcule

$$2(\phi(2) - \phi(-1)) - (\eta(2) - \eta(-1)) = 0,$$

si $a \leq -1$ ou $2 \leq a$. Si $-1 < a < 2$, on a

$$\begin{aligned} 2(\phi(2) - \phi(-1)) - (\eta(2) - \eta(-1)) &= 2\phi(2) - \eta(2) \\ &= (2^2 - a^2) - (2 - a) \\ &= (2 - a)(2 + a - 1) \\ &= (2 - a)(1 + a) \geq 0. \end{aligned}$$

De même dans le cas sgn_. Comme de plus $u(t)$ tends vers u_0 dans $L^1_{\text{loc}}(\mathbb{R})$, u est la solution entropique du problème (20)-(21).

2. Contrairement à la discontinuité considérée dans l'équation de transport, qui est passive (elle est simplement translatée au cours du temps), la discontinuité observée ici est active (on parle de choc ou même de choc compressif). En effet, les particules rentrent dans le choc et contribuent à le stabiliser. Si le choc se déplace, c'est parce que le groupe de particules de gauche a plus de poids (vitesse plus grande) que le groupe des particules de droite. Le choc étant actif, la diffusion numérique le régularise dans une moindre mesure, et bien qu'il y ait effectivement un effet de régularisation, il n'apparaît pas sur la figure 3, car le pas d'espace a été choisi suffisament petit.

4.6 Exercice 6

Exercice extrêmement difficile ! En fait l'estimation d'erreur en $h^{1/2}$ est une conjecture. On l'observe en pratique numériquement. On sait la démontrer :

1. en dimension $d = 1$,
2. en dimension $d > 1$ dans le cas où le maillage est structuré (ce qui signifie à peu de chose près que toutes les mailles sont les copies d'un même pavé, et dans ce cas on se ramène à la situation en dimension 1),
3. en dimension $d > 1$ sur maillage quelconque dans le cas *linéaire* $A(u) = au$ (avec a autorisé à dépendre de (x, t)), [MV07, DL10].

Montrer l'estimation d'erreur en $h^{1/2}$ dans le cas général ($d > 1$, A quelconque, maillage quelconque) est un problème ouvert. A vous d'essayer !

References

[CH99] C. Chainais-Hillairet, *Finite volume schemes for a nonlinear hyperbolic equation. Convergence towards the entropy solution and error estimate*, M2AN Math. Model. Numer. Anal. **33** (1999), no. 1, 129–156.

[DL10] F. Delarue and F. Lagoutière, *Probabilistic analysis of the upwind scheme for transport.*

[EGGH98] R. Eymard, T. Gallouët, M. Ghilani, and R. Herbin, *Error estimates for the approximate solutions of a nonlinear hyperbolic equation given by finite volume schemes*, IMA J. Numer. Anal. **18** (1998), no. 4, 563–594.

[EGH95a] R. Eymard, T. Gallouët, and R. Herbin, *Convergence of a finite volume scheme for a nonlinear hyperbolic equation*, Proceedings of the Third International Colloquium on Numerical Analysis (Plovdiv, 1994) (Utrecht), VSP, 1995, pp. 61–70.

[EGH95b] ———, *Existence and uniqueness of the entropy solution to a nonlinear hyperbolic equation*, Chinese Ann. Math. Ser. B **16** (1995), no. 1, 1–14, A Chinese summary appears in Chinese Ann. Math. Ser. A **1**6 (1995), no. 1, 119.

[MV07] B. Merlet and J. Vovelle, *Error estimate for finite volume scheme*, Numer. Math. **106** (2007), no. 1, 129–155.