

These AY Leroux

1979

SERIE : A
N° d'Ordre : 294
N° de Série : 35



THESE

présentée devant

L'UNIVERSITE DE RENNES

pour obtenir

Le Grade de **Docteur es-Sciences Mathématiques**

Option **ANALYSE NUMERIQUE**

par

Alain Yves LE ROUX

Approximation de quelques problèmes hyperboliques non linéaires

2ème Thèse : Réseaux de files d'attente

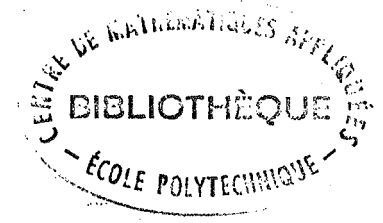
Soutenue le 5 avril 1979 devant la Commission d'Examen

MM. Chr. COATMELEC Président

C. BARDOS
M. CROUZEIX
A. MIGNOT
J.C. NEDELEC
J. PELLAUMAIL
P.A. RAVIART

} Examineurs

SERIE : A
N° d'Ordre : 294
N° de Série : 35



THESE

présentée devant

L'UNIVERSITE DE RENNES

pour obtenir

Le Grade de **Docteur es-Sciences Mathématiques**

Option **ANALYSE NUMERIQUE**

par

Alain Yves LE ROUX

Approximation de quelques problèmes hyperboliques non linéaires

2ème Thèse : Réseaux de files d'attente

Soutenue le 5 avril 1979 devant la Commission d'Examen

MM. Chr. COATMELEC Président

C. BARDOS
M. CROUZEIX
A. MIGNOT
J.C. NEDELEC
J. PELLAUMAIL
P.A. RAVIART

} Examineurs

Le travail qui a abouti à cette thèse a été effectué à l'I.N.S.A. de RENNES, dans le cadre du laboratoire d'Analyse Numérique qui s'est regroupé autour de Monsieur Chr. COATMELEC, dont j'ai apprécié l'aide tant morale que matérielle, et que je remercie très sincèrement pour avoir accepté la responsabilité de présider ce jury.

Je tiens également à exprimer ici toute ma reconnaissance à Monsieur Cl. BARDOS, pour les nombreux conseils et encouragements qu'il a pu me prodiguer, et pour la disponibilité dont il a fait preuve en de nombreuses occasions.

Messieurs M. CROUZEIX et A. MIGNOT ont accepté de participer à ce jury et je leur adresse mes plus sincères remerciements.

Monsieur J.Cl. NEDELEC m'a proposé ce sujet de recherche dès le début de mon troisième cycle, et je veux lui exprimer ici ma plus profonde gratitude. Il n'a jamais cessé de m'encourager, et j'ai apprécié son hospitalité et ses nombreux conseils à l'occasion de ces journées de travail très productives, passées à l'Ecole Polytechnique.

Monsieur J. PELLAUMAIL m'a proposé un second sujet très intéressant ; je l'en remercie, ainsi que pour son étroite et très précieuse collaboration dans l'aboutissement de ce travail.

Monsieur P.A. RAVIART me fait le très grand honneur de participer à ce jury, et je lui exprime ici toute ma reconnaissance, en particulier pour m'avoir suggéré d'étudier les schémas quasi d'ordre deux.

J'assure également de ma considération mes collègues du laboratoire d'Analyse Numérique pour quelques discussions très intéressantes et la bonne entente qui y règne, ainsi qu'à Monsieur P. JOUVE pour son aide très précieuse lors de l'établissement d'un modèle de câble, et Monsieur R. MARIE pour son expérience au niveau du second sujet. J'exprime également ma reconnaissance à Messieurs T. NISHIDA et M. YAMAGUTI pour les conversations très instructives que nous avons eues lors de leurs séjours à Rennes.

Je veux enfin exprimer mes plus sincères remerciements à Madame J. MARTIN pour la qualité de son travail dactylographique.

x x
x

INTRODUCTION

Le problème de Cauchy, constitué de l'équation quasi-linéaire du premier ordre sur $\mathbb{R}^p \times]0, T[$

$$(1) \quad \frac{\partial u}{\partial t} + \text{div}[f(u, x, t)] + g(u, x, t) = 0,$$

et d'une condition initiale donnée sur \mathbb{R}^p

$$(2) \quad u(x, 0) = u_0(x),$$

est étudié dans la première partie. Les hypothèses et quelques généralités sur les solutions discontinues qu'admet ce problème sont regroupées dans le chapitre I. La condition d'entropie y est détaillée, ainsi que la méthode de viscosité et les résultats d'unicité de S.N. KRUKOV. Le chapitre II est consacré à la construction puis à la démonstration de la convergence du schéma de GODUNOV, pour le problème à une seule dimension d'espace ($p=1$). Le chapitre III regroupe quelques résultats sur le problème à plusieurs dimensions et une étude de schémas quasi d'ordre deux. Ces schémas sont construits à partir d'un schéma d'ordre un, le schéma décentré ou le schéma de GODUNOV, dont on diminue le terme de viscosité pour le faire coïncider

avec celui d'un schéma d'ordre deux, le schéma de LAX WENDROFF par exemple, chaque fois que la stabilité n'est pas compromise, ce qui est presque toujours le cas. Cette technique est inspirée des travaux de BOOK, J.P. BORIS, HAIN et VAN LEER, qui utilisent de tels schémas pour résoudre l'équation de transport. On démontre la convergence de la famille de solutions approchées construites à partir de ces schémas, puis après avoir fait quelques restrictions sur les hypothèses et sur le schéma, qui devient légèrement moins souvent d'ordre deux, on établit que la solution obtenue à la limite vérifie la condition d'entropie.

x x
x

Dans la deuxième partie, on introduit un ouvert Ω de \mathbb{R}^D , et on considère le problème constitué de l'équation (1) sur $\Omega \times]0, T[$, de la condition initiale (2) sur Ω , et d'une condition aux limites sur la frontière de Ω . Cette condition est de type "Dirichlet", mais n'intervient que sur certaines parties de la frontière, qui dépendent (de façon peu triviale) de la solution. Au chapitre IV, cette condition aux limites est introduite à partir d'exemples ; on propose ensuite une définition de la solution, qui tient compte à la fois de cette condition aux limites et de la condition d'entropie. En reprenant les résultats de Cl. BARDOS, A.Y. LE ROUX, et J.C. NEDELEC, on prouve l'existence et l'unicité d'une solution satisfaisant aux critères de cette définition.

L'approximation de cette solution est étudiée au chapitre V. On y reprend le schéma de GODUNOV, ainsi que sa version quasi d'ordre deux, pour $p = 1$, et le schéma de LAX pour le problème à plusieurs dimensions. Le chapitre VI traite également de l'approximation de cette solution lorsque la fonction f intervenant dans (1) vérifie une hypothèse de monotonie. On y étudie le schéma décentré en dimension multiple, ainsi que sa version quasi d'ordre deux, et quelques schémas quasi-explicites dont l'utilisation est relativement aisée pour le problème à une dimension.

x x
x

La troisième partie est consacrée à un problème quasi-linéaire d'ordre deux, où intervient l'équation d'onde non linéaire

$$(3) \quad \frac{\partial^2 w}{\partial t^2} = \frac{\partial}{\partial x} \left(f \left(\frac{\partial w}{\partial x} \right) \right), \quad \text{avec } f' \geq 0,$$

qui peut être mise sous la forme d'un système quasi-linéaire de deux équations d'ordre un. Le chapitre VII regroupe quelques généralités sur les systèmes hyperboliques de ce type. On y propose une généralisation de la notion d'invariants de RIEMANN, qui permet de résoudre le problème de RIEMANN sans faire d'autres hypothèses que la monotonie de f , puis de construire quelques schémas numériques, dont le schéma de LAX et des schémas de type GODUNOV.

Au chapitre VIII, un modèle de câble extensible est étudié ; il s'agit d'une équation du type (3) où f est monotone croissante, nulle à l'origine et telle que $|f|$ soit convexe. Ces propriétés assurent la convexité d'un domaine limité par les invariants de RIEMANN, ce qui permet d'établir la stabilité des schémas définis au chapitre précédent. On termine en donnant quelques résultats partiels sur la convergence de ces schémas, et en faisant quelques remarques sur le problème de conditions aux limites associé à (3).

x x
x

La numérotation des formules est interne à chaque chapitre. Pour faire référence à une formule donnée dans un autre chapitre, on précise chaque fois le numéro de cet autre chapitre, devant le numéro de la formule. La bibliographie est classée par ordre alphabétique. Les références bibliographiques dans le texte se font en précisant le nom de l'auteur, suivi d'un numéro lorsque cet auteur est cité plusieurs fois. Un sommaire des titres de chapitre et de paragraphe, situant également les définitions et les théorèmes, est donné à la fin de l'ouvrage.

x x
x

PREMIERE PARTIE
APPROXIMATION POUR LE PROBLEME
DE CAUCHY

Cette première partie traite du problème de condition initiale associé à l'équation quasi-linéaire du premier ordre, sur $\mathbb{R}^p \times]0, T[$

$$\frac{\partial u}{\partial t} + \operatorname{div} [f(u, x, t)] + g(u, x, t) = 0 ,$$

et est divisée en trois chapitres. Quelques généralités sur les solutions discontinues, en particulier la condition d'entropie, ainsi que l'existence et l'unicité sont traitées dans le chapitre I.

Le chapitre II est consacré à la construction puis à la convergence du schéma de GODUNOV pour le cas monodimensionnel.

Dans le chapitre III, on étudie quelques autres schémas, dont un schéma décentré quasi d'ordre deux ; quelques résultats de convergence y sont démontrés, et quelques résultats numériques y sont présentés.

x x
x

CHAPITRE I

QUELQUES GENERALITES POUR LE PROBLEME DE CAUCHY

1 - QUELQUES PROPRIETES DES SOLUTIONS FAIBLES

Soit $T > 0$. On se propose d'étudier l'équation quasi-linéaire du premier ordre sur $\mathbb{R}^p \times]0, T[$

$$(1) \quad \frac{\partial u}{\partial t} + \operatorname{div} [f(u, x, t)] + g(u, x, t) = 0,$$

avec la condition initiale

$$(2) \quad u(x, 0) = u_0(x), \quad x \in \mathbb{R}^p.$$

Cette condition initiale est supposée mesurable et bornée, et à variation localement bornée sur \mathbb{R}^p :

$$(3) \quad u_0 \in L^\infty(\mathbb{R}^p) \cap BV_{loc}(\mathbb{R}^p).$$

La notion de variation bornée est prise ici au sens de TONELLI CESARI, et signifie que pour tout $i \in \{1, \dots, p\}$ et presque tout

$$\hat{x} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_p) \in \mathbb{R}^{p-1}$$

la fonction

$$x_i \rightsquigarrow u_0(\hat{x}, x_i)$$

est à variation bornée sur tout intervalle réel borné. L'extension de cette notion de fonction à variation localement bornée à $\mathbb{R}^p \times]0, T[$ est évidente, et c'est dans cet espace que l'on va chercher u , solution de (1).

Les fonctions f et g satisfont aux hypothèses suivantes, où j est un entier entre 1 et p :

- (i) $f \in \{C^2(\mathbb{R} \times \mathbb{R}^p \times]0, T[) \}^P$, $f = (f_1, \dots, f_p)$
- (ii) $g \in C^1(\mathbb{R} \times \mathbb{R}^p \times]0, T[)$
- (iii) g, f , et chaque $\frac{\partial f_j}{\partial x_j}$, $\frac{\partial f_j}{\partial t}$, sont lipschitziennes en $u \in \mathbb{R}$, uniformément en $(x, t) \in \mathbb{R}^p \times]0, T[$
- (iv) g et chaque $\frac{\partial f_j}{\partial x_j}$ sont lipschitziennes en $x \in \mathbb{R}^p$, uniformément en $(u, t) \in \mathbb{R} \times]0, T[$
- (v) pour $u = 0$, $g(0, \dots)$ et chaque $\frac{\partial f_j}{\partial x_j}(0, \dots)$ sont bornées sur $\mathbb{R}^p \times]0, T[$.

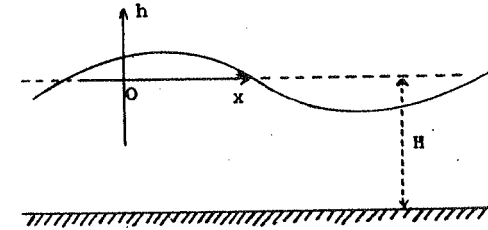
Ces hypothèses sont évidemment beaucoup plus simples dans le cas où g est identiquement nulle et où chaque f_j ne dépend que de u , ce qui est le cas dans de nombreux exemples. Il suffit alors d'écrire :

$$(4') \quad f \in \{C^1(\mathbb{R})\}^P$$

L'équation non linéaire de type (1) la plus connue est l'équation de BURGERS, utilisée comme modèle de traitement des ondes de choc apparaissant lors de l'écoulement unidimensionnel d'un fluide compressible parfait. Cette équation est la suivante

$$(5) \quad \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left(\frac{u^2}{2} \right) = 0$$

On rencontre une équation du même type dans l'étude des ondes apparaissant à la surface d'un fluide incompressible en écoulement infiniment large. En notant H la hauteur de fluide au repos, et $h+H$ la hauteur de fluide en mouvement, l'équation est la suivante



$$\frac{\partial h}{\partial t} + \sqrt{gH} \left(1 + \frac{3h}{2H} \right) \frac{\partial h}{\partial x} = 0$$

où g est l'accélération de la pesanteur.

x x
x

Il n'y a pas d'hypothèse de régularité sur la condition initiale u_0 . L'hypothèse (3) autorise en effet des discontinuités de première espèce. De plus, lorsque f est non linéaire et par exemple pour un problème monodimensionnel, il y a des exemples où une solution u de (1) et (2), au sens classique, peut présenter des singularités à partir d'une certaine valeur de t , et ceci même lorsque u_0 est très régulière. Or, on veut pouvoir construire la solution pour des instants postérieurs à l'apparition d'une éventuelle singularité ; on est ainsi amené à envisager en particulier l'existence de solutions discontinues. Ces discontinuités correspondent d'ailleurs à des chocs, dans le cadre de l'équation de BURGERS.

Si la condition initiale est non bornée, des singularités peuvent apparaître, au-delà desquelles il n'est plus possible de construire la solution. Par exemple, l'équation de BURGERS résolue à partir de la condition initiale

$$u(x, 0) = -x$$

admet la solution

$$u(x, t) = x/(t-1)$$

qui devient infinie en $t = 1$, pour tout $x \neq 0$, et il n'y a plus existence pour $t > 1$. Etant donné un réel positif a , on peut prendre maintenant comme

condition initiale la fonction u_0 définie sur \mathbb{R} par

$$u_0(x) = \begin{cases} \text{Inf}(a, -x) & \text{si } x < 0 \\ \text{Sup}(-a, -x) & \text{si } x > 0 \end{cases}$$

La solution de BURGERS, pour cette condition initiale est

donnée par

$$u(x,t) = \begin{cases} \text{Inf}\left(a, \frac{x}{t-1}\right) & \text{si } x < 0 \\ \text{Sup}\left(-a, \frac{x}{t-1}\right) & \text{si } x > 0 \end{cases}$$

pour $t < 1$. En $t = 1$, cette solution vaut (à la limite)

$$a \operatorname{sg}(x)$$

où sg est la fonction signe définie sur \mathbb{R} par

$$(6) \quad \operatorname{sg}(x) = \begin{cases} 1 & \text{si } x > 0 \\ 0 & \text{si } x = 0 \\ -1 & \text{si } x < 0 \end{cases}$$

et on peut envisager une solution (discontinue) pour les instants $t > 1$. Cette solution est d'ailleurs la solution stationnaire

$$u(x,t) = -a \operatorname{sg}(x)$$

qui vérifie (1) au sens des distributions ; on dit alors qu'il s'agit d'une solution faible, dont la définition est la suivante.

Définition 1 : Une fonction $u \in L^1(\mathbb{R}^p \times]0, T[)$ est dite solution faible du problème (1), (2) lorsque pour toute fonction

$$\phi \in C^2(\mathbb{R}^p \times]0, T[)$$

u vérifie

$$(7) \quad \int_{\mathbb{R}^p} \int_0^T \left\{ u \frac{\partial \phi}{\partial t} + f(u, x, t) \cdot \operatorname{grad} \phi - g(u, x, t) \phi \right\} dx dt = - \int_{\mathbb{R}^p} u_0(x) \phi(x, 0) dx$$

Dans la définition précédente, et aussi dans la suite, le produit scalaire dans \mathbb{R}^p est représenté par un point. De plus l'indice 0 dans C_0^2 signifie "à support compact".

Géométriquement, on remarque que les discontinuités apparaissent à l'intersection des courbes caractéristiques dans l'espace (u, x, t) . En dimension un et en supposant (4') vérifiée et g nulle par exemple, ces courbes sont solutions du système d'équations

$$(8) \quad \frac{du}{dt} = 0 \quad ; \quad \frac{dx}{dt} = f'(u) \quad ;$$

et sont déterminées par un point "initial" $(u_0(x), x, 0)$ où elles passent. La surface formée par ces courbes caractéristiques donne une solution faible du problème (1), (2). On vérifie ainsi qu'une solution au sens classique de (1), (2) satisfait à l'équation (implicite), dans le cas monodimensionnel et avec f vérifiant (4'), et g nulle,

$$(9) \quad u(x,t) = u_0[x - f'(u)t]$$

Cette formule permet d'obtenir des solutions explicites, d'une part, et d'autre part indique que les informations provenant de la condition initiale se propagent avec une vitesse finie, majorée par

$$(10) \quad M = \sup_{|u| \leq |u_0|} |f'(u)| \quad \text{sur } I^\infty(\mathbb{R})$$

On en conclut qu'en un point (x,t) , la valeur prise par la solution est indépendante des données initiales sur

$$] -\infty, x - Mt[U]x + Mt, +\infty [$$

Remarquons également que (9) implique que u reste bornée, ce qui vient d'ailleurs d'être utilisé pour définir M en (10).

Considérons maintenant, dans le cadre monodimensionnel, une solution faible u , admettant une discontinuité le long d'une courbe régulière passant par un point (x,t) . On suppose u de classe C^1 dans un voisinage de ce point (x,t) , et que f vérifie (4'), g nulle. On déduit alors de (7), que la courbe de discontinuité admet une équation de la forme

(* sauf évidemment sur la courbe de discontinuité)

(11) $x = x(t)$

qui satisfait au point (x,t) à la relation de RANKINE-HUGONIOT

(12) $\frac{dx}{dt} = \frac{f(u(x+0,t)) - f(u(x-0,t))}{u(x+0,t) - u(x-0,t)}$

Ce raisonnement peut être mené en sens inverse, et permet d'établir qu'une fonction u bornée, de classe C^1 par morceaux sur un ouvert Omega subset R x]0,T[, dont les lignes de discontinuité peuvent s'exprimer par des équations de la forme (11), qui vérifie (1) au sens classique en tout point de régularité, et (12) en tout point de discontinuité, est solution faible de (1) sur Omega. Cette remarque permet de construire des exemples de solution faible.

On peut faire également un lien entre (8) et (12), et remarquer que (12) est symétrique entre les valeurs à droite et à gauche de la discontinuité. Cette symétrie est la principale cause de non unicité des solutions faibles, et pour établir un résultat d'unicité, il est nécessaire de remplacer (12) par une condition plus restrictive (en particulier non symétrique). Cette condition a été proposée par OLEINIK ; c'est la condition d'entropie.

2 - LA CONDITION D'ENTROPIE

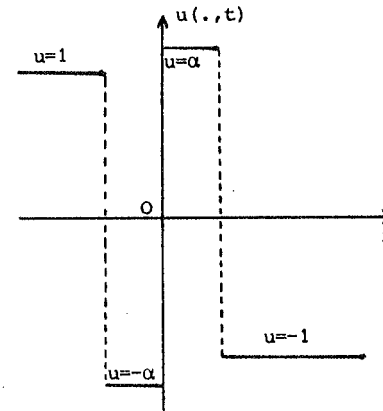
Tout ce paragraphe se situe dans le cadre monodimensionnel, avec une fonction g nulle et une fonction f indépendante de (x,t) et vérifiant (4').

Reprenons l'équation de BURGERS (5), avec une condition initiale discontinue, par exemple

$u(x,0) = -sg(x)$

qui correspond, à une constante multiplicative près, à l'exemple donné dans le paragraphe précédent, et où il était écrit que la solution stationnaire était solution faible. Il se trouve qu'il existe d'autres solutions

faibles vérifiant cette condition initiale ; on peut en effet envisager toutes les solutions faibles suivantes, chacune correspondant à une valeur du paramètre réel alpha > 1,



$u(x,t) = \begin{cases} 1 & \text{si } x < (1-\alpha)t/2 \\ -\alpha & \text{si } (1-\alpha)t/2 < x < 0 \\ -u_\alpha(-x,t) & \text{si } x > 0 \end{cases}$

On vérifie aisément que la condition de RANKINE HUGONIOT (12) est satisfaite sur chacune des discontinuités, et que la solution est constante sur chaque secteur limité par ces lignes de discontinuités, et donc solution de l'équation de BURGERS (5).

La solution stationnaire, c'est-à-dire celle correspondant à alpha = 1, peut être considérée comme étant la plus régulière, dans la mesure qu'elle n'admet qu'une seule discontinuité au lieu de trois, et surtout parce qu'elle conserve la norme uniforme de la condition initiale ; il est donc naturel de vouloir sélectionner cette solution en écrivant qu'elle satisfait à une condition supplémentaire. Cette condition peut être dans le cadre de l'équation de BURGERS

(13) $u(x-0,t) \geq u(x+0,t) \quad , \quad t > 0$

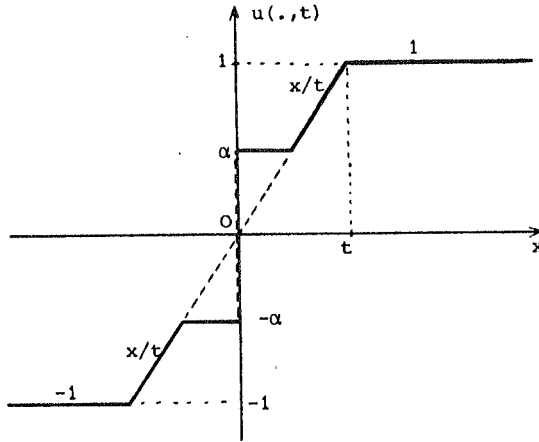
et a pour effet d'interdire les discontinuités croissantes, qui correspondent physiquement à des détonations et sont instationnaires.

Si la condition initiale est discontinue, avec une discontinuité croissante, la condition (13) va fonctionner en régularisant, c'est-à-dire en rendant continue la solution pour t > 0. Ainsi, pour tout alpha in [0,1], la fonction u_alpha définie sur R x]0,T[par

$u_\alpha(x,t) = \begin{cases} 1 & \text{si } x \geq t \\ x/t & \text{si } t \leq x < t \\ \alpha & \text{si } t \leq x < 0 \\ -u_\alpha(-x,t) & \text{si } x > 0 \end{cases}$

est solution faible de l'équation de BURGERS satisfaisant à la condition initiale

$$u(x,0) = \text{sg}(x)$$



Parmi ces solutions faibles, une seule est continue, et c'est aussi la seule qui satisfait à (13). Il s'agit de la solution proposée correspondant à $\alpha = 0$. Toutes les autres admettent une discontinuité croissante en $x = 0$.

Dans le cas général, on ne peut pas toujours trouver une solution continue à partir d'une condition initiale admettant une discontinuité croissante. La condition d'entropie proposée jusqu'ici ne convient en effet que si la fonction f est convexe au niveau de la discontinuité.

Si la fonction f est concave, au niveau de la discontinuité, on prendra une condition interdisant les discontinuités décroissantes. Dans le cas plus général où il n'y a pas d'hypothèse de convexité sur f , on prend la condition suivante, qui est la condition d'entropie proposée par OLEINIK, dans [3] et [4]. On pose

$$u_- = u(x-0,t) \quad , \quad u_+ = u(x+0,t) \quad ,$$

et on introduit l'intervalle fermé limité par u_- et u_+

$$I(u_-, u_+) = [\text{Inf}(u_-, u_+) \quad , \quad \text{Sup}(u_-, u_+)] \quad .$$

La discontinuité satisfait à la condition d'entropie lorsque

$$(14) \quad \frac{f(u_+) - f(u_-)}{u_+ - u_-} = \text{Max}_{k \in I(u_-, u_+)} \left\{ \frac{f(u_+) - f(k)}{u_+ - k} \right\}$$

Cette condition est équivalente à

$$\frac{f(u_+) - f(u_-)}{u_+ - u_-} = \text{Min}_{k \in I(u_-, u_+)} \left\{ \frac{f(u_-) - f(k)}{u_- - k} \right\} \quad ,$$

et en particulier, lorsque f est convexe, on retrouve l'inégalité (13), c'est-à-dire

$$u_- \geq u_+ \quad ,$$

et l'inégalité inverse lorsque f est concave.

On se propose maintenant d'étudier le fonctionnement de cette condition d'entropie sur l'exemple non convexe suivant, auquel ne correspond aucun support physique,

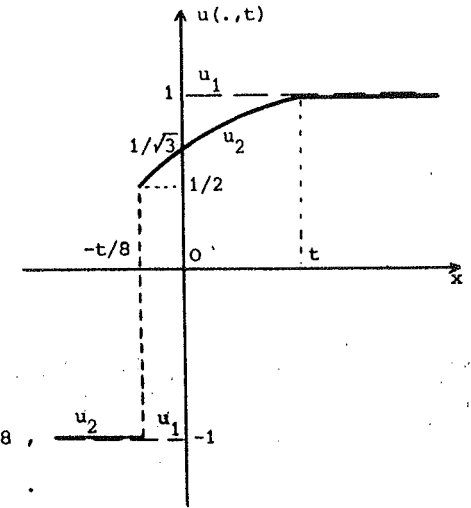
$$\left\{ \begin{array}{l} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left\{ \frac{1}{2} u(u^2 - 1) \right\} \quad , \\ u(x,0) = \text{sg}(x) \quad . \end{array} \right.$$

Ce problème admet les deux solutions faibles suivantes

$$u_1(x,t) = \text{sg}(x) \quad ,$$

solution stationnaire, et

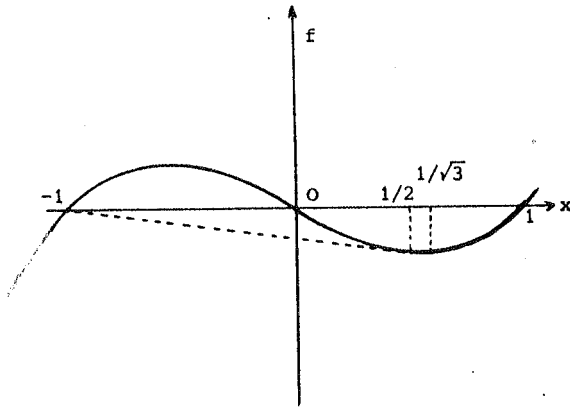
$$u_2(x,t) = \begin{cases} 1 & \text{si } x > t \quad , \\ \sqrt{\frac{2x+t}{3t}} & \text{si } t > x > -t/8 \quad , \\ -1 & \text{si } x < -t/8 \quad . \end{cases}$$



Il est immédiat que u_2 satisfait à la condition d'entropie (14) et que u_1 ne la satisfait pas, ceci bien que toutes deux admettent une discontinuité croissante.

On constate, en ce qui concerne la solution u_2 , que tout se passe comme si on avait remplacé f par son enveloppe convexe sur $[-1,1]$, c'est-à-dire au niveau de la discontinuité initiale.

La discontinuité qui reste correspond exactement à un segment rectiligne de cette enveloppe convexe. Remarquons aussi qu'en remplaçant f par son enveloppe convexe au niveau d'une discontinuité croissante, on s'oppose à l'apparition d'un choc, et de fait cette discontinuité disparaît si aucun segment rectiligne n'apparaît dans le graphe de l'enveloppe convexe. Notons également que dans le cas d'une discontinuité décroissante, la condition d'entropie revient à remplacer f par son enveloppe concave au niveau de la discontinuité, ce qui correspond encore à s'opposer à l'apparition du choc.



En tenant compte de (12) et de (8), il est immédiat que la partie décroissante du graphe de l'enveloppe convexe de f correspond à une propagation négative des données initiales, tandis que la partie croissante de ce graphe correspond à une propagation positive. On obtient ainsi que le minimum de f est réalisé par une valeur stationnaire de la solution. De fait, en $x = 0$, on a bien $u_2(0,t) = 1/\sqrt{3}$, qui réalise le minimum de f . De façon générale, la valeur u qui réalise

$$(15) \quad \text{Min}_{k \in I(u_-, u_+)} \{ \text{sg}(u_+ - u_-) f(k) \}$$

est une valeur stationnaire de la solution si u est situé à l'intérieur de l'intervalle $I(u_-, u_+)$. Cette remarque est fondamentale pour la construction du schéma de GODUNOV (cf. Chapitre II).

3 - LA METHODE DE PSEUDO-VISCOSITE

La méthode de pseudo viscosité permet d'établir l'existence d'une solution faible satisfaisant à la condition d'entropie (14). Elle consiste à introduire un terme d'ordre deux au second membre de (1), de coefficient ϵ positif et destiné à tendre vers zéro. Ce coefficient correspond à un paramètre de viscosité, du moins dans le cadre de l'équation de BURGERS, et le fait qu'il soit strictement positif est essentiel.

L'équation, dont la solution est notée u_ϵ , devient

$$(16) \quad \frac{\partial u_\epsilon}{\partial t} + \text{div} [f(u_\epsilon, x, t)] + g(u_\epsilon, x, t) = \epsilon \Delta u_\epsilon$$

Le terme d'ordre deux a un effet régularisant sur la solution, et pour ϵ petit, u_ϵ reste proche de u , solution faible de (1) satisfaisant à la condition d'entropie. La convergence de u_ϵ vers u est obtenue par un argument de compacité dans l'espace de SOBOLEV $W_{loc}^{1,1}(\mathbb{R}^P \times]0, T[)$, qui s'injecte dans $B V_{loc}(\mathbb{R}^P \times]0, T[)$, et donc u est dans ce dernier espace. Une démonstration détaillée est donnée dans l'article de S.N. KRUKOV [2]; Cl. BARDOS, dans [1] propose une autre démonstration des estimations a priori. On trouve encore une démonstration proche dans l'article de Cl. BARDOS, A.Y. LE ROUX, J.C. NEDELEC, qui est reprise ici dans la seconde partie. De plus, la convergence des schémas numériques va constituer une preuve de l'existence. Il est donc inutile de détailler la démonstration de convergence, et on se contentera d'établir une formulation faible, due à E. HOPF et à S.N. KRUKOV et qui caractérise la solution faible vérifiant la condition d'entropie. On donnera ensuite un résumé de la démonstration du théorème d'unicité de S.N. KRUKOV, à partir de cette formulation.

On introduit l'approximation suivante de la fonction "signe", déjà définie en (6). Pour $\eta > 0$, on pose pour tout $x \in \mathbb{R}$

$$(17) \quad \text{sg}_\eta(x) = \begin{cases} 1 & \text{si } x > \eta \\ x/\eta & \text{si } |x| < \eta \\ -1 & \text{si } x < -\eta \end{cases}$$

On rappelle également le résultat suivant, qui est dû à SAKS, et dont une démonstration est reprise dans Cl. BARDOS [1].

Lemme 1 : Soit v une fonction définie sur \mathbb{R}^q à valeurs dans \mathbb{R} .

On suppose que v est continûment différentiable, et que

$$v \in W^{1,1}(\mathbb{R}^q).$$

Alors

$$(18) \quad \lim_{\eta \rightarrow 0} \int_{\{x \mid |v(x)| < \eta\}} |\text{grad } v| \, dx = 0.$$

Le principe de la démonstration de ce lemme consiste à montrer que sur tout borné, $\{x \mid |x| < R\}$ par exemple, l'ensemble $\{x \mid v(x) = \text{grad } v(x) = 0\}$ a la même mesure de LEBESGUE que l'ensemble $\{x \mid v(x) = 0\}$.

Soit $k \in \mathbb{R}$. On construit une fonction I_η , continûment dérivable sur \mathbb{R} , telle que pour $\eta > 0$, on ait

$$(19) \quad I'_\eta(u_\varepsilon) = \text{sg}_\eta(u_\varepsilon - k) \quad ; \quad I_\eta(k) = 0.$$

Il est immédiat que $I_\eta(u_\varepsilon)$ converge vers $|u_\varepsilon - k|$ lorsque η tend vers zéro. On multiplie (16) par $\text{sg}_\eta(u_\varepsilon - k)$, et il vient

$$(20) \quad \begin{aligned} & \frac{\partial}{\partial t} I_\eta(u_\varepsilon) + \text{div}[\text{sg}_\eta(u_\varepsilon - k) f(u_\varepsilon, x, t)] \\ & + \text{sg}_\eta(u_\varepsilon - k) g(u_\varepsilon, x, t) - \text{grad}(\text{sg}_\eta(u_\varepsilon - k)) \cdot f(u_\varepsilon, x, t) \\ & = \varepsilon \text{div}[\text{sg}_\eta(u_\varepsilon - k) \text{grad } u_\varepsilon] - \varepsilon \text{sg}'_\eta(u_\varepsilon - k) |\text{grad } u_\varepsilon|^2. \end{aligned}$$

On introduit une fonction ϕ de classe C^2 et à support compact sur $\mathbb{R}^p \times]0, T[$, à valeurs réelles et non négatives ; on notera

$$(21) \quad \phi \in C^2_{0+}(\mathbb{R}^p \times]0, T[).$$

En multipliant (20) par ϕ et en intégrant par parties, il vient :

$$\iint_{\mathbb{R}^p \times]0, T[} \left\{ I_\eta(u_\varepsilon) \frac{\partial \phi}{\partial t} + \text{sg}_\eta(u_\varepsilon - k) (f(u_\varepsilon, x, t) - f(k, x, t)) \cdot \text{grad } \phi \right. \quad (a)$$

$$\left. + \text{sg}_\eta(u_\varepsilon - k) (f(k, x, t) \cdot \text{grad } \phi - g(u_\varepsilon, x, t) \phi) \right\} dx \, dt \quad (b)$$

$$+ \iint_{\mathbb{R}^p \times]0, T[} \text{sg}'_\eta(u_\varepsilon - k) \text{grad } u_\varepsilon \cdot (f(u_\varepsilon, x, t) - f(k, x, t)) \phi \, dx \, dt \quad (c)$$

$$(22) \quad + \iint_{\mathbb{R}^p \times]0, T[} \text{grad}(\text{sg}_\eta(u_\varepsilon - k)) \cdot f(k, x, t) \phi \, dx \, dt \quad (d)$$

$$= \varepsilon \iint_{\mathbb{R}^p \times]0, T[} \text{sg}_\eta(u_\varepsilon - k) \text{grad } u_\varepsilon \cdot \text{grad } \phi \, dx \, dt \quad (e)$$

$$+ \varepsilon \iint_{\mathbb{R}^p \times]0, T[} \text{sg}'_\eta(u_\varepsilon - k) |\text{grad } u_\varepsilon|^2 \phi \, dx \, dt \quad (f)$$

On convient de noter (a)...(f) chacun des termes de (22), et on passe à la limite en η . Le terme (c) tend vers zéro d'après le lemme 1 ; en effet, d'après l'hypothèse de LIPSCHITZ uniforme sur f (cf (4) (iii)), on a une majoration de ce terme par

$$\iint_{\{(x,t) \in \text{Supp } \phi, |u_\varepsilon - k| < \eta\}} \frac{c}{\eta} |u_\varepsilon - k| \phi |\text{grad}(u_\varepsilon - k)| \, dx \, dt,$$

et qui tend vers zéro.

Après regroupement, les termes (b) et (d) ont une somme égale à

$$- \iint_{\mathbb{R}^p \times]0, T[} (g(u_\varepsilon, x, t) + \text{div}(f(k, x, t))) \text{sg}_\eta(u_\varepsilon - k) \phi \, dx \, dt.$$

Or la fonction $\text{sg}_\eta(u_\varepsilon - k)$ converge en tout point vers $\text{sg}(u_\varepsilon - k)$, en restant bornée par 1. D'après le théorème de LEBESGUE, il y a convergence dans $L^1_{\text{loc}}(\mathbb{R}^p \times]0, T[)$, et on peut ainsi passer à la limite sur ce terme. Le même argument permet de passer à la limite sur (a) et (e). Le terme (f) est tout simplement minoré par zéro ; l'hypothèse $\varepsilon > 0$ intervient (entre autres) ici.

On obtient

$$(23) \quad \iint_{\mathbb{R}^p \times]0, T[} \left\{ |u_\varepsilon - k| \frac{\partial \phi}{\partial t} + \text{sg}(u_\varepsilon - k) (f(u_\varepsilon, x, t) - f(k, x, t)) \cdot \text{grad } \phi \right. \\ \left. - [g(u_\varepsilon, x, t) + \text{div}(f(k, x, t))] \text{sg}(u_\varepsilon - k) \phi \right\} dx dt \\ \geq \varepsilon \iint_{\mathbb{R}^p \times]0, T[} \text{sg}(u_\varepsilon - k) \text{grad } u_\varepsilon \cdot \text{grad } \phi \quad dx dt$$

On effectue maintenant le passage à la limite en ε . Il est immédiat que le second membre de (23) est nul à la limite, sachant que $|\text{grad } u_\varepsilon|$ est uniformément borné. Le passage à la limite sur $L^1(\text{Supp } \phi)$ le premier terme est techniquement un peu plus compliqué, du fait de la présence de $\text{sg}(u_\varepsilon - k)$, qui est discontinu.

On utilisera le lemme 2 suivant, dont la démonstration est donnée plus loin.

Lemme 2 : Soit Ω un ouvert borné de \mathbb{R}^d , v_m et w_m deux suites de fonctions définies sur Ω et convergentes vers v dans $L^1(\Omega)$ et vers w dans $(L^1(\Omega))^d$ ($d \in \mathbb{N}$) respectivement, et ϕ une fonction localement lipschitzienne sur $\mathbb{R}^{d+1} \times \Omega$. On suppose la suite w_m uniformément bornée sur Ω . Alors, pour tout réel k , on a

$$(24) \quad \lim_{\eta \rightarrow 0} \lim_{m \rightarrow \infty} \int_{\mathbb{R}} \frac{1}{2} \text{sg}'_\eta(\xi - k) \left(\int_{\Omega} \text{sg}(v_m(x) - \xi) \phi(w_m(x), \xi, x) dx \right) d\xi \\ = \int_{\Omega} \text{sg}(v(x) - k) \phi(w(x), k, x) dx$$

On rappelle que u_ε est bornée uniformément d'après le principe du maximum (voir par exemple le livre de A. FRIEDMAN).

Il suffit maintenant de remplacer k par ξ dans (23), de multiplier par $\frac{1}{2} \text{sg}'_\eta(\xi - k)$ et d'intégrer en ξ . Le fait que le second membre tende vers zéro est maintenu, et le passage à la limite en ε puis en η donne, d'après le lemme 2,

$$(25) \quad \iint_{\mathbb{R}^p \times]0, T[} \left\{ |u - k| \frac{\partial \phi}{\partial t} + \text{sg}(u - k) (f(u, x, t) - f(k, x, t)) \cdot \text{grad } \phi \right. \\ \left. - (g(u, x, t) + \text{div}[f(k, x, t)]) \text{sg}(u - k) \phi \right\} dx dt \geq 0.$$

On prend en effet Ω contenant $\text{Supp } \phi$, et $v_m = w_m = u_\varepsilon$. Le lemme 2 est donné ici dans une forme plus générale, qui pourra être appliquée lors de la démonstration de convergence de certains schémas. La formule (25) caractérise la solution faible qui satisfait à la condition d'entropie. En prenant $|k|$ supérieur à $|u|_{L^\infty(\mathbb{R}^p \times]0, T[)}$, il vient, en introduisant k positif, puis k négatif successivement dans (25), que u vérifie (7) et donc est solution faible. Le fait que ϕ soit non négatif ne réduit pas la généralité ; il suffit en effet de prendre ϕ et ψ dans $C^2_0(\mathbb{R}^p \times]0, T[)$, avec ψ égale à 1 sur le support de ϕ et d'écrire (7), avec ϕ^2 et $(\phi - \psi)^2$, pour avoir également (7) avec ϕ .

Supposons maintenant que u soit de classe C^1 , sauf sur un nombre fini de surfaces (de discontinuité) régulières de $\mathbb{R}^p \times]0, T[$. En un point (x, t) de l'une de ces surfaces, on introduit dans (25) une fonction test de la forme $\phi_\eta(x, t)$, de façon à concentrer l'intégrale sur la surface de discontinuité lorsque le paramètre η tend vers zéro. Ceci donne à la limite, en notant

$$\vec{v} = (v_0, \dots, v_p)^t$$

la normale unitaire à la surface de discontinuité, orientée vers les $t < 0$, puis en posant

$$u_- = u((x, t) + 0 \cdot \vec{v}) \quad ; \quad u_+ = u((x, t) - 0 \cdot \vec{v})$$

l'inégalité suivante pour tout $k \in \mathbb{R}$:

$$(26) \quad |u_- - k| v_0 + \sum_{j=1}^p \text{sg}(u_- - k) (f_j(u_-, x, t) - f_j(k, x, t)) v_j \\ \geq |u_+ - k| v_0 + \sum_{j=1}^p \text{sg}(u_+ - k) (f_j(u_+, x, t) - f_j(k, x, t)) v_j$$

c'est-à-dire la condition d'entropie dans le cas général.

En se plaçant dans le cas monodimensionnel, avec f ne dépendant que de u , il vient sachant que la courbe de discontinuité admet une équation de la forme $x = x(t)$, l'inégalité

$$(|u_+ - k| - |u_- - k|)x'(t) \geq \text{sg}(u_+ - k)(f(u_+) - f(k)) - \text{sg}(u_- - k)(f(u_-) - f(k)).$$

On en déduit la relation de RANKINE-HUGONIOT (12) en prenant successivement $k = u_-$ et $k = u_+$, puis la condition d'entropie (14) pour k entre u_- et u_+ , en utilisant (12). Dans le cas général, en faisant successivement $k = u_-$ et $k = u_+$, on obtient la condition de RANKINE HUGONIOT suivante, avec les mêmes notations

$$(27) \quad (u_+ - u_-) v_0 + \sum_{j=1}^p (f_j(u_+, x, t) - f_j(u_-, x, t)) v_j = 0,$$

qui est symétrique en u_+ et u_- . La condition (26) peut également être obtenue en exigeant que l'opérateur de semigroupe associé à l'équation (1) soit une contraction dans L^1 (voir Cl. BARDOS [1], B.L. KEYFITZ [1]).

Démonstration du lemme 2

On a, pour tout $m \in \mathbb{N}$, en notant $\|\cdot\|$ la norme dans \mathbb{R}^d

$$|\phi(w_m(x), \xi, x) - \phi(w(x), k, x)| \leq \alpha \|w_m(x) - w(x)\| + \beta |\xi - k|,$$

où α et β sont les constantes de LIPSCHITZ de ϕ par rapport à ses deux premières variables. On a

$$\frac{1}{2} \int_{\mathbb{R}} \text{sg}'_{\eta}(\xi - k) d\xi = 1 \quad ; \quad \frac{1}{2} \int_{\mathbb{R}} \text{sg}'(\xi - k) |\xi - k| d\xi = \frac{\eta}{2} ;$$

et donc

$$\left| \int_{\mathbb{R}} \frac{1}{2} \text{sg}'_{\eta}(\xi - k) \int_{\Omega} \text{sg}(v_m(x) - \xi) (\phi(v_m(x), \xi, x) - \phi(w(x), k, x)) dx d\xi \right| \leq \alpha \|w_m - w\|_{(L^1(\Omega))^d} + \frac{\beta}{2} \eta \text{mes}(\Omega),$$

qui tend vers zéro lorsque m tend vers l'infini et η vers zéro.

Il reste à considérer

$$\begin{aligned} & \int_{\mathbb{R}} \frac{1}{2} \text{sg}'_{\eta}(\xi - k) \left(\int_{\Omega} \text{sg}(v_m(x) - \xi) \phi(w(x), k, x) dx \right) d\xi, \\ &= \int_{\Omega} \phi(w(x), k, x) \left(\int_{\mathbb{R}} \frac{1}{2} \text{sg}'_{\eta}(\xi - k) \text{sg}(v_m(x) - \xi) d\xi \right) dx, \\ &= \int_{\Omega} \phi(w(x), k, x) \left\{ \int_{-\infty}^{v_m(x)} - \int_{v_m(x)}^{+\infty} \right\} \frac{1}{2} \text{sg}'_{\eta}(\xi - k) d\xi dx, \\ &= \int_{\Omega} \phi(w(x), k, x) \text{sg}_{\eta}(v_m(x) - k) dx. \end{aligned}$$

Pour η fixé, sg_{η} est lipchitzienne et donc $\text{sg}_{\eta}(v_m(x) - k)$ converge vers $\text{sg}_{\eta}(v(x) - k)$ dans $L^1(\Omega)$, lorsque m tend vers l'infini. On obtient donc, jusqu'à présent

$$\begin{aligned} \lim_m \int_{\mathbb{R}} \frac{1}{2} \text{sg}'_{\eta}(\xi - k) \int_{\Omega} \text{sg}(v_m(x) - \xi) \phi(w_m(x), \xi, x) dx d\xi \\ = \int_{\Omega} \text{sg}_{\eta}(v(x) - k) \phi(w(x), k, x) dx. \end{aligned}$$

On fait maintenant tendre η vers zéro. Le terme $\text{sg}_{\eta}(v(x) - k)$ converge vers $\text{sg}(v(x) - k)$ pour tout $x \in \Omega$ (donc presque partout sur Ω), et sg_{η} est uniformément borné par 1. D'après le théorème de la convergence dominée de LEBESGUE, il y a aussi convergence dans $L^1(\Omega)$. On peut donc passer à la limite et obtenir (24), sachant que $\phi(w(\cdot), k, \cdot) \in L^{\infty}(\Omega)$. Le lemme 2 est démontré.

4 - LA DEFINITION DE KRUKOV, ET LE THEOREME D'UNICITE

S.N. KRUKOV propose de définir la solution du problème (1), (2) à partir de (25) et (2), de la façon suivante, dans [2].

Définition 2 : Une fonction $u \in L^{\infty}(\mathbb{R}^p \times]0, T[)$ est solution du problème (1), (2) lorsque

$$\forall k \in \mathbb{R} \quad \forall \phi \in C_{0+}^2(\mathbb{R}^p \times]0, T[)$$

$$(25) \quad \iint_{\mathbb{R}^p \times]0, T[} |u-k| \frac{\partial \phi}{\partial t} + \text{sg}(u-k) (f(u, x, t) - f(k, x, t)) \cdot \text{grad } \phi - \text{sg}(u-k) [\text{div}(f(k, x, t)) + g(u, x, t)] \phi \, dx \, dt \geq 0$$

et lorsqu'il existe un ensemble $\mathcal{E} \subset]0, T[$ de mesure nulle tel que

$$(28) \quad \forall R > 0 \quad \lim_{\substack{t \rightarrow 0 \\ t \notin \mathcal{E}}} \int_{|x| < R} |u(x, t) - u_0(x)| \, dx = 0$$

Remarquons que u étant obtenu par passage à la limite sur les u_ϵ , bornés dans $W_{loc}^{1,1}(\mathbb{R}^p \times]0, T[)$, u appartient en fait à $BV_{loc}(\mathbb{R}^p \times]0, T[)$, et donc admet pour presque tout $x \in \mathbb{R}^p$ une limite en $t = 0$, vers laquelle il converge au sens simple, et que l'on note $u(x, 0)$. On peut donc remplacer (28) par

$$(29) \quad u(x, 0) = u_0(x) \quad \text{pp}(\mathbb{R}^p),$$

c'est-à-dire que (2) est vérifié presque partout.

A partir de la Définition 2, on a le théorème suivant, qui est dû à S.N. KRUKOV.

Théorème 1 : Le problème (1), (2) admet une seule solution, caractérisée par la définition 2. De plus si u et v sont deux solutions, de conditions initiales respectives u_0 et v_0 , on a pour tout $R > 0$

$$(30) \quad \int_{|x| < R} |u(x, t) - v(x, t)| \, dx \leq e^{\alpha_0 t} \int_{|x| < R + Nt} |u_0 - v_0| \, dx,$$

avec N défini par

$$(31) \quad N = \sup_{(x, t)} \left\{ \sum_{j=1}^p \left(\frac{\partial f_j}{\partial u}(u, x, t) \right)^2 \right\}^{1/2},$$

c'est-à-dire la constante de LIPSCHITZ de f , et où α_0 est la constante de LIPSCHITZ de g .

Remarquons que (30) montre que (25) peut être assimilé à un générateur de semi groupe, qui est une contraction dans L^1 du moins lorsque α_0 est nul, c'est-à-dire lorsque la fonction g est constante en u , et ne dépend plus que de x et de t .

Démonstration du théorème 1

On ne donne que les grandes lignes de cette démonstration, qui est détaillée dans l'article de S.N. KRUKOV [2].

Soit u et v deux solutions, de conditions initiales respectives u_0 et v_0 . On se propose de démontrer (30), et l'unicité est alors immédiate. Soit $\phi \in C_{0+}^2(\mathbb{R}^p \times]0, T[)$ et j_0 une fonction de $C_{0+}^2(\mathbb{R})$, telle que

$$j_0(x) = 0 \quad \text{si} \quad |x| \geq 1$$

$$\int_{\mathbb{R}} j_0(x) \, dx = 1.$$

On pose ensuite pour tout $m \in \mathbb{N}^*$

$$j_m(x) = m j_0(mx) \quad ;$$

les fonctions j_m ainsi définies sont de classe C^2 , non négatives, et vérifient

$$j_m(x) = 0 \quad \text{si} \quad x \geq \frac{1}{m},$$

$$j_m(x) \leq C m$$

$$\int_{\mathbb{R}} j_m(x) \, dx = 1$$

$$|j'_m(x)| \leq C' m^2$$

On introduit enfin

$$\phi_m \in C_{0,+}^2 \left\{ (\mathbb{R}^p \times]0, T[) \times (\mathbb{R}^p \times]0, T[) \right\}$$

définie par

$$\phi_m(x, t; y, s) = \phi\left(\frac{x+y}{2}, \frac{t+s}{2}\right) j_m\left(\frac{t-s}{2}\right) \prod_{j=1}^p j_m\left(\frac{x_j - y_j}{2}\right),$$

et on suppose m suffisamment grand pour que le support de ϕ_m soit inclus dans $\{\mathbb{R}^p \times]0, T[\}^2$. On écrit (25) avec $k = v(y, s)$ et en prenant ϕ_m comme fonction test, en tant que fonction de (x, t) , et on intègre en (y, s) . On fait de même avec (25) écrite avec v , en remplaçant les variables (x, t) par les variables (y, s) , en posant $k = u(x, t)$ et en prenant la même fonction test ϕ_m . On ajoute les deux inégalités obtenues, et on obtient, sachant que

$$u = u(x, t) \quad , \quad v = v(y, s)$$

l'inégalité

$$(32) \quad \left\{ \iiint_{(\mathbb{R}^p \times]0, T[)^2} |u - v| \left(\frac{\partial \phi_m}{\partial t} + \frac{\partial \phi_m}{\partial s} \right) + \text{sg}(u-v) (f(u, x, t) - f(v, y, s)) \cdot (\text{grad}_x \phi_m + \text{grad}_y \phi_m) + \text{sg}(u-v) \left[(f(u, y, s) - f(u, x, t)) \cdot \text{grad}_y \phi_m + \text{div}_y (f(u, y, s)) \phi_m \right] + \text{sg}(u-v) \left[(f(v, y, s) - f(v, x, t)) \cdot \text{grad}_x \phi_m - \text{div}_x (f(v, x, t)) \phi_m \right] + \text{sg}(u-v) (g(v, y, s) - g(u, x, t)) \phi_m \right\} dx dt dy ds \geq 0 \quad ,$$

où les indices x ou y indiquent la variable pour laquelle grad ou div est calculé.

On effectue le passage à la limite sur chaque terme de (32), lorsque m tend vers l'infini. On a

$$\frac{\partial \phi_m}{\partial t} + \frac{\partial \phi_m}{\partial s} = \frac{\partial \phi}{\partial t} \left(\frac{x+y}{2}, \frac{t+s}{2} \right) j_m \left(\frac{t-s}{2} \right) \prod_{j=1}^p j_m \left(\frac{x_j - y_j}{2} \right),$$

puisque les dérivées issues de j_m s'éliminent.

Or

$$(33) \quad \left\{ \iiint |v(x, t) - v(y, s)| \frac{\partial \phi}{\partial t} \left(\frac{x+y}{2}, \frac{t+s}{2} \right) j_m \left(\frac{t-s}{2} \right) \prod_{j=1}^p j_m \left(\frac{x_j - y_j}{2} \right) dx dt dy ds + \iint_{\text{Supp } \phi} \left\{ (C_m)^{p+1} \prod_{j=1}^p \int_{|y_j - x_j| < \frac{1}{m}} \int_{|t-s| < \frac{1}{m}} |v(x, t) - v(y, s)| dy ds \right\} dx dt \left| \frac{\partial \phi}{\partial t} \right|_{L^\infty} \right\}$$

et presque tout point $(x, t) \in \mathbb{R}^p \times]0, T[$, est un point de LEBESGUE de v , c'est-à-dire que

$$m^{p+1} \prod_{j=1}^p \int_{|y_j - x_j| < \frac{1}{m}} \int_{|t-s| < \frac{1}{m}} |v(x, t) - v(y, s)| dy ds$$

tend vers zéro lorsque m tend vers l'infini. De plus v est uniformément bornée, et donc l'intégrale en (x, t) sur le support de ϕ tend vers zéro. On obtient ainsi que le premier terme de (32) se comporte comme

$$\iiint |u(x, t) - v(x, t)| \left(\frac{\partial \phi_m}{\partial t} + \frac{\partial \phi_m}{\partial s} \right) dx dt dy ds \quad ,$$

où on fait le changement de variables (à partir de (y, s))

$$\tau = \frac{t-s}{2} \quad , \quad \xi = \frac{x-y}{2}$$

pour obtenir que ce terme est égal à

$$\iiint |u(x, t) - v(x, t)| \int_{\text{Sup } \xi_j < \frac{1}{m}} \int_{|\tau| < \frac{1}{m}} 2^{p+1} \frac{\partial \phi}{\partial t} (x - \xi, t - \tau) j_m(\tau) \prod_{j=1}^p j_m(\xi_j) d\xi d\tau dx dt$$

et il est immédiat que l'intégrale en (ξ, τ) converge vers

$$\frac{\partial \phi}{\partial t} (x, t)$$

en tout point $(x, t) \in \mathbb{R}^p \times]0, T[$. On obtient à la limite en appliquant une nouvelle fois le théorème de LEBESGUE que le premier terme de (32) converge vers

$$2^{p+1} \iiint_{\mathbb{R}^p \times]0, T[} |u-v| \frac{\partial \phi}{\partial t} dx dt \quad .$$

Le second terme de (32) se traite de la même façon, et le même genre d'argument est utilisé pour traiter les autres termes. Ainsi la somme des troisième et quatrième termes tend vers zéro ; la principale difficulté rencontrée est le traitement de termes contenant j'_m , qui sont toujours multipliés par $\frac{t-s}{2}$ ou des $\frac{x_j - y_j}{2}$, si bien qu'il reste la possibilité de les majorer, et ainsi de passer à la limite. Le dernier terme ne présente aucune

difficulté ; il se traite comme (33), ϕ_m jouant le rôle de $\frac{\partial \phi_m}{\partial t}$ et g le rôle de v . On obtient ainsi à la limite sur (32) l'inégalité

$$(34) \quad \iint_{\mathbb{R}^p \times]0, T[} \left\{ |u-v| \frac{\partial \phi}{\partial t} + sg(u-v) (f(u, x, t) - f(v, x, t)) \cdot \text{grad } \phi - sg(u-v) (g(u, x, t) - g(v, x, t)) \phi \right\} dx dt \geq 0,$$

c'est-à-dire exactement (25), v jouant le rôle de k , et sans le terme en $\text{div}(f(k, x, t))$.

Soit $R > 0$; on note K_R le cône $\{(x, t) \mid |x| + Nt < R\}$, ce qui représente le cône caractéristique de base $\{|x| < R\}$ dont on approche la fonction indicatrice de la façon suivante. On pose

$$\psi_m(x) = \int_{-\infty}^x j_m(y) dy,$$

qui est une fonction non décroissante de x , puis on définit

$$\chi_m(x, t) = 1 - \psi_m(|x| + Nt - R + \frac{1}{m}).$$

Soit $\phi_0 \in C_{0+}^2(0, T)$; on prend dans (34) la fonction test égale à $\phi_0(t) \chi_m(x, t)$, puis on passe à la limite, en ayant remarqué que

$$\frac{\partial \chi_m}{\partial t} + \frac{f(u, x, t) - f(v, x, t)}{u - v} \cdot \text{grad } \chi_m$$

$$\leq \frac{\partial \chi_m}{\partial t} + N |\text{grad } \chi_m|$$

$$\leq -N \psi'_m + N |\psi'_m| = 0.$$

On obtient, en notant α_0 la constante de LIPSCHITZ de g , qui est uniforme en (x, t) , l'inégalité

$$(35) \quad \iint_{\mathbb{R}^p \times]0, T[} |u-v| (\phi'_0(t) \chi(x, t) + \alpha_0 \phi_0(t) \chi(x, t)) dx dt \geq 0,$$

où $\chi(x, t)$ est la limite des $\chi_m(x, t)$, c'est-à-dire la caractéristique du cône K_R .

Etant donné $t \in]0, T[$, on note S_t la section du cône K_R au temps t . On se donne ensuite t_0 et t_1 tels que

$$0 < t_0 < t_1 < T,$$

avec S_{t_1} non vide, et on note μ_m une approximation de la fonction indicatrice de l'intervalle $]t_0, t_1[$. On prend alors dans (35) la fonction test ϕ_0 égale à

$$e^{-\alpha_0 t} \mu_m(t),$$

et on fait tendre m vers l'infini. On obtient à la limite, grâce à des arguments semblables à ceux qui ont permis d'établir (34), l'inégalité

$$(36) \quad \int_{S_{t_1}} |u(x, t_1) - v(x, t_1)| dx \leq e^{\alpha_0(t_1 - t_0)} \int_{S_{t_0}} |u(x, t_0) - v(x, t_0)| dx,$$

ce qui permet d'établir (30) en faisant tendre t_0 vers zéro.

Le théorème 2 est démontré, sachant que l'existence a été traitée auparavant. Notons que l'unicité implique la convergence de toute la famille $\{u_\varepsilon\}$.

Une démonstration de (36), avec $\alpha_0 = 0$, a déjà été proposée par B.L. KEYFITZ - QUINN [1], en partant de la condition d'entropie. Inversement, le fait d'avoir (36) entraîne la condition d'entropie (26) sur les surfaces de discontinuités.

x x
x

Dans ce chapitre I, quelques détails de démonstration ont été donnés ; ce sont ceux qui seront utiles dans la suite. C'est le cas par exemple des lemmes 1 et 2, et dans la démonstration du théorème 1, la façon de passer à la limite sur (32). Il existe aussi d'autres démonstrations d'existence ; voir par exemple DOUGLIS [3]. Le lemme 2 est inspiré des techniques de OHARU et TAKAHASHI. On peut remarquer aussi que le semi-groupe généré par (25) est celui donné comme exemple fondamental pour la théorie des semi-groupes non linéaires dans des espaces non réflexifs ; voir à ce sujet CRANDALL ou CRANDALL LIGGETT.

Les hypothèses (4) sur f et g , et l'hypothèse de variation localement bornée faite sur u_0 en (3) ne sont pas exactement celles faites dans S.N. KRUKOV [2], pour la démonstration de l'existence. Il supprime en effet cette hypothèse sur la variation de u_0 , et ne suppose que

$$u_0 \in L^\infty(\mathbb{R}^p).$$

Dans les problèmes pratiques, l'hypothèse de variation bornée est toujours satisfaite, et on l'utilisera au niveau de la démonstration de convergence des schémas numériques. L'estimation de $\text{grad } u_\epsilon$ et de $\frac{\partial u_\epsilon}{\partial t}$ dans L^1 , qui permet d'établir la compacité de u_ϵ va en effet être remplacée par une estimation de la variation en espace et en temps de la solution approchée u_h . Il est immédiat qu'en construisant l'interpolée v_h de type Q_1 de u_h sur le maillage utilisé, il y a coïncidence entre les normes L^1 des dérivées partielles de v_h et de la variation dans les directions correspondants de u_h .

Dans la plupart des exemples classiques, f ne dépend que de la variable u . On se placera donc le plus souvent dans ce cas, après avoir traité le cas général, dans le cadre du schéma de GODUNOV pour le problème à une variable d'espace, pour indiquer la façon de résoudre les difficultés techniques supplémentaires.

x x
x

CHAPITRE II

LE SCHEMA DE GODUNOV

1 - LE PROBLEME DE RIEMANN

Le schéma de GODUNOV est un schéma aux différences finies, d'ordre un, mais bien adapté à la résolution du problème (I.1), (I.2) dans le cas monodimensionnel, particulièrement au niveau des discontinuités. Sa construction utilise localement la résolution du problème de RIEMANN, puis une projection de type L^2 .

Le problème de RIEMANN consiste à résoudre l'équation

$$(1) \quad \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} [f(u,x,t)] + g(u,x,t) = 0 \quad , \quad (x,t) \in \mathbb{R} \times]0,T[,$$

avec une condition initiale discontinue sur \mathbb{R} , de la forme

$$(2) \quad u(x,0) = \begin{cases} \alpha & \text{si } x < 0 , \\ \beta & \text{si } x > 0 , \end{cases}$$

où α et β sont des constantes réelles fixées.

Pour résoudre ce problème, on commence par étudier un cas particulier, celui où g est une fonction nulle, et où f ne dépend que de u ; le cas général sera considéré ensuite. Dans le cas particulier, l'équation est réduite à

$$(3) \quad \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} (f(u)) = 0 \quad , \quad (x,t) \in \mathbb{R} \times]0,T[$$

On a vu, dans le paragraphe 2 du chapitre I, que la condition d'entropie (14) revient à remplacer f par son enveloppe convexe si $\alpha < \beta$, et par son enveloppe concave si $\alpha > \beta$, sur l'intervalle limité par α et β .

Evidemment le cas $\alpha = \beta$ est trivialement résolu, sa solution étant $u = \alpha$. Comme en (I.15), on peut définir une quantité

$$\gamma \in I(\alpha, \beta)$$

par

$$(4) \quad \text{sg}(\beta - \alpha) f(\gamma) = \text{Min}_{k \in I(\alpha, \beta)} \{ \text{sg}(\beta - \alpha) f(k) \} \quad ,$$

où on rappelle que $I(\alpha, \beta)$ désigne l'intervalle fermé dont les bornes sont α et β . La valeur de γ n'est pas unique en général, mais supposons un instant qu'elle le soit. On utilise l'équation caractéristique (I.8), à savoir

$$(5) \quad \frac{dx}{dt} = f'(u) \quad ,$$

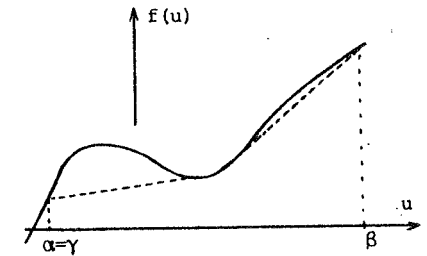
pour suivre l'évolution de cette valeur γ dans la solution, et ainsi avoir une indication sur le lieu des points (x,t) pour lesquels

$$u(x,t) = \gamma \quad .$$

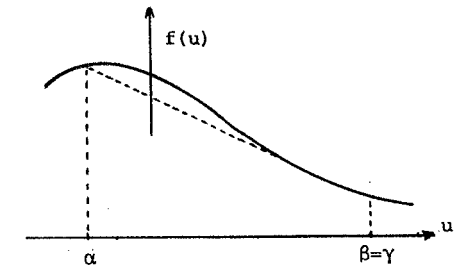
Trois cas peuvent se présenter, suivant la position de γ dans l'intervalle $I(\alpha, \beta)$.

Cas 1. $\gamma = \alpha$. Ceci implique que $f'(\gamma)$ est non négative et d'après (5) la propagation se fait avec une vitesse non négative, c'est-à-dire vers les x positifs. Pour tout $x \leq 0$, on a sûrement $u(x,t) = \alpha$, et ceci pour tout $t > 0$. En particulier, on a

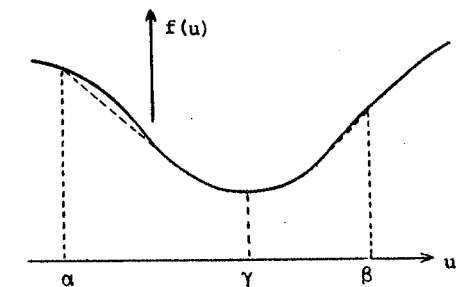
$$(6) \quad \forall t > 0 \quad u(0,t) = \gamma \quad .$$



Cas 2. $\gamma = \beta$. On a cette fois-ci nécessairement $f'(\beta) = f'(\gamma)$ non positive, et la propagation a lieu vers les x négatifs, la vitesse étant négative. On a pour tout $x \geq 0$ et tout $t > 0$, $u(x,t)$ égal à β , et en particulier on a encore (6), avec $\gamma = \beta$.



Cas 3. $\gamma \in \overset{\circ}{I}(\alpha, \beta)$, c'est-à-dire à l'intérieur de l'intervalle. Nécessairement $f'(\gamma)$ est nul, et la valeur γ est une valeur stationnaire de la solution. On a donc encore (6), et on peut noter que les valeurs de u comprises entre α et γ se propagent avec une vitesse négative, et celles comprises entre γ et β avec une vitesse positive.



Le fait de remplacer f par son enveloppe convexe pour $\alpha < \beta$, ou par son enveloppe concave pour $\alpha > \beta$ ne perturbe pas (4), qui définit les mêmes quantités γ . En notant \tilde{f} cette enveloppe, on a

$$sg(f'(\gamma)) = sg(\tilde{f}'(\gamma))$$

et ceci ne change rien aux trois cas considérés plus haut. Pour les quantités $k \in I(\alpha, \beta)$ ne réalisant pas (4), le passage de f à \tilde{f} permet d'introduire la condition d'entropie (14). En effet, les segments rectilignes de \tilde{f} correspondent à des discontinuités, et ailleurs, où nécessairement f coïncide avec \tilde{f} , la solution est régulière. Le signe de \tilde{f}' donne le sens de propagation dans tous les cas.

Lorsque γ n'est pas défini de façon unique par (4), la discontinuité initiale se conserve au moins partiellement et reste stationnaire en $x = 0$. L'enveloppe \tilde{f} est constante sur un intervalle fermé

$$[\gamma_0, \gamma_1] \subset I(\alpha, \beta)$$

avec

$$\tilde{f}(\gamma_0) = f(\gamma_0) = \tilde{f}(\gamma_1) = f(\gamma_1)$$

et

$$\tilde{f}'(\gamma_0) = f'(\gamma_0) = \tilde{f}'(\gamma_1) = f'(\gamma_1)$$

On peut convenir par exemple de prendre $\gamma = \gamma_0$, et aussi décider que (6) est encore vérifié, en sachant que c'est en fait $f(\gamma)$ qui nous servira dans la suite.

Dans le cas général, c'est-à-dire pour l'équation (1), on peut remplacer (4) par

$$(4') \quad \begin{cases} \gamma \in I(\alpha, \beta) \\ sg(\beta - \alpha) f(\gamma, 0, 0) = \text{Min}_{k \in I(\alpha, \beta)} \{ sg(\beta - \alpha) f(k, 0, 0) \} \end{cases}$$

et (6) est devenu

$$(6') \quad \lim_{t \rightarrow 0} u(0, t) = \gamma$$

On peut ainsi résoudre le problème de RIEMANN (ou au moins approcher sa solution sur le segment $x = 0, 0 < t < \delta$ avec δ assez petit) et ceci

permet de construire le schéma de GODUNOV, ainsi que de présenter plusieurs autres schémas, afin de les comparer au schéma de GODUNOV. On vérifie ainsi que le schéma de GODUNOV réalise d'une certaine façon un critère d'optimalité, dans la mesure où il a été tenu compte de (4) ou (4'), et non de façon approchée. Cette comparaison des schémas est faite dans le chapitre III.

Dans le cadre du problème de RIEMANN, et lorsque f ne dépend que de u , c'est-à-dire pour l'équation (3), on peut effectuer le changement de variable $x \rightsquigarrow \lambda x, t \rightsquigarrow \lambda t$ et remarquer que le problème reste invariant. On en déduit que la solution est de la forme

$$u(x, t) = U\left(\frac{x}{t}\right)$$

où U vérifie l'équation

$$[\xi - f'(U)] \frac{dU}{d\xi} = 0$$

et satisfait aux conditions suivantes, provenant de la condition initiale (2)

$$U(-\infty) = \alpha ; \quad U(+\infty) = \beta$$

On peut résoudre

$$f'(U) = \xi$$

tout en vérifiant les conditions en $\pm \infty$ lorsqu'il existe deux valeurs ξ_- et ξ_+ (avec $\xi_- < \xi_+$) pour lesquelles

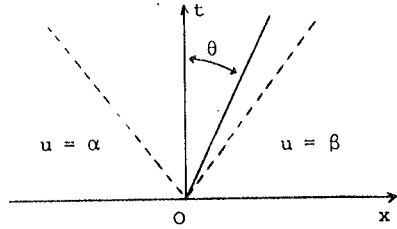
$$f'(\alpha) = \xi_- ; \quad f'(\beta) = \xi_+$$

Ceci est possible lorsque f' est croissante et $\alpha < \beta$, et lorsque f' est décroissante et $\alpha > \beta$. Dans ces deux cas, la solution de (3) est régulière. Dans les autres cas, une discontinuité apparaît, qui vérifie la condition d'entropie. On retrouve également que ces discontinuités restent inchangées lorsqu'on remplace f par son enveloppe convexe (donc de dérivée croissante) si $\alpha < \beta$, ou par son enveloppe concave si $\alpha > \beta$.

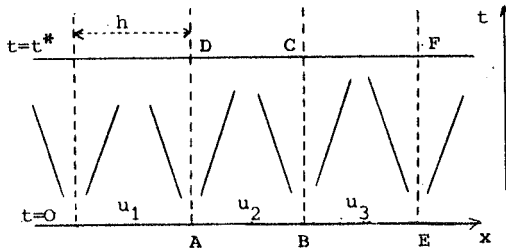
On sait aussi que dans le cas de l'équation (3) les caractéristiques sont des droites, et comme il s'agit d'un problème de RIEMANN, ces droites caractéristiques passent toutes par l'origine.

On note θ l'angle d'une droite caractéristique avec l'axe $\{x = 0\}$. On peut alors majorer la pente $tg\theta$ de la façon suivante

$$|\operatorname{tg} \theta| \leq \sup_{k \in I(\alpha, \beta)} |f'(k)|.$$



Cette remarque permet de résoudre de manière exacte le problème de CAUCHY constitué de l'équation (3) et d'une condition initiale constante par morceaux dont la distance entre deux discontinuités est un multiple d'un réel $h > 0$ fixé, pour $t \in [0, t^*]$, avec t^* vérifiant



$$t^* \cdot \sup_{k \in \mathbb{R}} |f'(k)| \leq h.$$

On peut ainsi utiliser le schéma de GLIMM (voir GLIMM [1]).

On peut également calculer pour chaque discontinuité la valeur γ vérifiant (4), qui coïncide avec la solution sur les segments AD, BC, etc... de la figure ci-dessus. En intégrant ensuite l'équation (3) sur le rectangle ABCD, par une formule de GREEN, on peut calculer la valeur moyenne de la solution sur chaque segment DC, CF, etc... de la figure ci-dessus. Ceci permet de réitérer le procédé, et de connaître une approximation de la solution sur $[t^*, 2t^*]$, etc... On construit ainsi le schéma de GODUNOV. Cette construction est reprise en détail au paragraphe suivant, dans le cas général de l'équation (1).

x x
x

2 - CONSTRUCTION DU SCHEMA DE GODUNOV

Soit h le pas de discrétisation en espace ; h est un réel strictement positif, destiné à tendre vers zéro. On peut donc supposer qu'il existe h_0 fixe tel que

$$0 \leq h \leq h_0.$$

On discrétise ensuite \mathbb{R} , en posant pour tout $i \in \mathbb{Z}$,

$$I_i = \left[\left(i - \frac{1}{2}\right) h, \left(i + \frac{1}{2}\right) h \right].$$

Le pas de discrétisation en temps est pris égal à qh , où q est une constante réelle, strictement positive et qui servira dans la suite à contrôler la stabilité. On recouvre $]0, T[$ par les intervalles

$$J_n = \left[\left(n - \frac{1}{2}\right) qh, \left(n + \frac{1}{2}\right) qh \right],$$

avec $n \in \{0, 1, \dots, N = \lceil T / qh \rceil + 1\}$

La solution approchée est une fonction u_h constante sur chaque pavé $I_i \times J_n$, où sa valeur est notée u_i^n . Ces quantités u_i^n sont construites de la façon suivante, à partir d'une condition initiale

$$u_0 \in L^\infty(\mathbb{R}) \cap BV_{loc}(\mathbb{R}).$$

Pour tout $i \in \mathbb{Z}$, on pose

$$(7) \quad u_i^0 = \frac{1}{h} \int_{I_i} u_0(x) dx,$$

c'est-à-dire qu'on effectue une projection de type L^2 sur chacun des intervalles I_i . Cette approximation est convergente dans $L^1_{loc}(\mathbb{R})$, puisque presque tout $x \in \mathbb{R}$ est point de LEBESGUE pour u_0 .

Supposons connaître tous les u_i^n pour $n < N$ fixé ; on veut construire les u_i^{n+1} . En $x = \left(i + \frac{1}{2}\right) h$, la fonction $u_h(\cdot, nqh)$ présente une discontinuité, et on sait résoudre ou au moins approcher le problème de RIEMANN en $x = \left(i + \frac{1}{2}\right) h$ et t assez proche de nqh , pour la condition de CAUCHY en $t = nqh$, définie par

$$u_o(x) = \begin{cases} u_1^n & \text{pour } x < (i + 1/2)h \\ u_{i+1}^n & \text{pour } x > (i + 1/2)h \end{cases}$$

On obtient, conformément à (4) ou (4'),

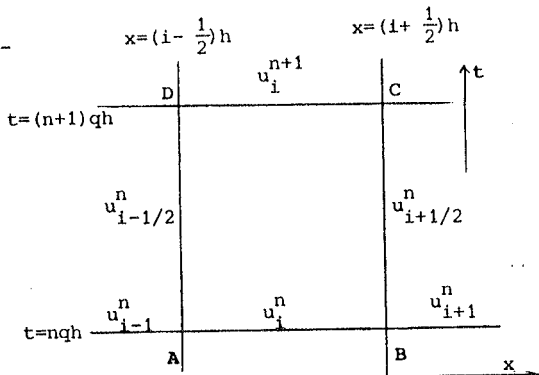
$$(8) \quad \begin{cases} u_{i+1/2}^n \in I(u_i^n, u_{i+1}^n) & \text{réalise} \\ \min_{k \in I(u_{i+1}^n, u_i^n)} \{ \text{sg}(u_{i+1}^n - u_i^n) f(k, (i + \frac{1}{2})h, nqh) \} \end{cases}$$

Connaissant tous les u_i^n , on a ainsi tous les $u_{i+1/2}^n$. En intégrant par parties l'équation

(1) sur le pavé

$$I_i \times [nqh, (n+1)qh]$$

représenté par le rectangle ABCD de la figure ci-contre, et en approchant la solution sur AD par $u_{i-1/2}^n$, et sur BC par $u_{i+1/2}^n$, il vient



$$\begin{aligned} \frac{1}{h} \int_{DC} u(x, (n+1)qh) dx &= \frac{1}{h} \int_{AB} u(x, nqh) dx \\ &+ \frac{1}{qh} \int_{nqh}^{(n+1)qh} q(f(u_{i-1/2}^n, (i-1/2)h, t) - f(u_{i+1/2}^n, (i+1/2)h, t)) dt \\ &- \frac{1}{qh^2} \iint_{I_i \times [nqh, (n+1)qh]} qh g(u(x, t), x, t) dx dt \end{aligned}$$

ce qui peut être approché par

$$(9) \quad \begin{aligned} u_i^{n+1} &= u_i^n - q[f(u_{i+1/2}^n, (i+1/2)h, nqh) - f(u_{i-1/2}^n, (i-1/2)h, nqh)] \\ &- qh g(u_i^n, ih, nqh). \end{aligned}$$

La donnée de (8) et (9) constitue le schéma de GODUNOV. On peut remarquer que u_i^{n+1} est obtenu par une sorte de projection de type L^2 sur chaque I_i ; c'est d'ailleurs exactement le cas pour l'équation (3).

Pour l'équation simple (3), le schéma de GODUNOV s'écrit

$$(10) \quad \begin{cases} u_{i+1/2}^n \in I(u_{i+1}^n, u_i^n) & \text{réalise} \\ \min_{k \in I(u_{i+1}^n, u_i^n)} \{ \text{sg}(u_{i+1}^n - u_i^n) f(k) \} \\ u_i^{n+1} = u_i^n - q(f(u_{i+1/2}^n) - f(u_{i-1/2}^n)) \end{cases}$$

Dans le cas où f est une fonction monotone croissante en u , on a, d'après le paragraphe précédent

$$\forall i \in \mathbb{Z} \quad u_{i+1/2}^n = u_i^n$$

et par exemple (10) est réduit à

$$(11) \quad u_i^{n+1} = u_i^n - q(f(u_i^n) - f(u_{i-1}^n))$$

c'est-à-dire exactement le schéma décentré. Si f est décroissante, on a $u_{i+1/2}^n = u_{i+1}^n$ pour tout $i \in \mathbb{Z}$, c'est-à-dire encore le schéma décentré. Dans OLEINIK [4], le schéma de GODUNOV est donné de façon différente, mais sa présentation n'est pas satisfaisante dans la mesure où elle peut conduire à des solutions qui ne satisfont pas à la condition d'entropie (14). Des contre exemples sont donnés dans LE ROUX [1] et [2]. Remarquons également qu'au cours d'une expérience numérique, (10) se réduit le plus souvent à (11) (ou l'analogue dans le cas f décroissante) et il est immédiat que l'ordre de précision du schéma est alors égal à 1.

Il faut faire une remarque concernant la validité de (8). La donnée du problème de RIEMANN à partir de u_i^n et de u_{i+1}^n n'est effective que sur un intervalle de longueur $2h$, à savoir $I_i \cup I_{i+1}$, et non pas sur \mathbb{R} tout entier. D'après la théorie des caractéristiques, en particulier d'après (I-8) et (I-12), on peut assurer que la solution dépend de u_i^n et de u_{i+1}^n seulement à l'intérieur du triangle

$$\{(x, t) \mid |x - (i+1/2)h| + M(t - nqh) \leq h\}$$

avec, comme en (I-10)

$$M = \text{Sup}_{u, x, t} \left\{ \left| \frac{\partial f}{\partial u}(u, x, t) \right| \right\}$$

ce qui d'après (I-4-iii) est fini. M est d'ailleurs la constante de LIPSCHITZ (uniforme) de f par rapport à u. Ceci va se traduire par une condition de stabilité, liant q et M ; cette condition exprimera que le rapport des pas de discrétisation en temps et en espace ne soit pas trop grand. En fait, si on veut pouvoir atteindre l'instant (n+1)qh en $x = (i + \frac{1}{2})h$, c'est-à-dire si on veut que ce point soit dans le triangle introduit plus haut, il faut que

$$(12) \quad M q \leq 1 .$$

On retrouve ainsi la condition de stabilité de COURANT - FRIEDRICHS - LEWY, qui va servir dans toutes les estimations a priori satisfaites par u_h , pour établir la convergence. Inversement si (12) n'est pas vérifiée, on assiste à des phénomènes d'instabilité, qui se développent souvent de façon catastrophique.

On a le résultat de convergence suivant

Théorème 2 : Sous les hypothèses (I.3) et (I.4), et si la condition de COURANT FRIEDRICHS LEWY (12) est satisfaite, la famille $\{u_h\}_{h>0}$ des solutions approchées construites par le schéma de GODUNOV (8), (9) converge dans $L^1_{loc}(\mathbb{R} \times]0, T[)$, vers

$$u \in L^\infty(\mathbb{R} \times]0, T[) \cap B V_{loc}(\mathbb{R} \times]0, T[) ,$$

solution du problème constitué de (1) et de la donnée initiale u_0 .

x x
x

La démonstration de ce théorème est très longue, et fait l'objet du paragraphe suivant. On ne donnera ici que les grandes lignes de cette démonstration, et ceci pour l'équation simple (3).

Principe de la démonstration

En utilisant (12), on commence par établir les estimations

$$(13) \quad \forall i \in \mathbb{Z} \quad |u_i^{n+1}| \leq \text{Max}(|u_{i-1/2}^n|, |u_i^n|, |u_{i+1/2}^n|) ,$$

$$(14) \quad \sum_{|i| \leq I} |u_{i+1}^{n+1} - u_i^{n+1}| \leq \sum_{|i| \leq I+1} |u_{i+1}^n - u_i^n| ,$$

$$(15) \quad \sum_{|i| \leq I} |u_i^{n+1} - u_i^n| \leq \sum_{|i| \leq I+1} |u_{i+1}^n - u_i^n| ,$$

ce qui permet, par le théorème de HELLY, d'assurer la convergence d'une suite u_h vers $u \in L^\infty(\mathbb{R} \times]0, T[) \cap B V_{loc}(\mathbb{R} \times]0, T[)$, pour la topologie $L^1_{loc}(\mathbb{R} \times]0, T[)$.

On démontre également que pour tout $k \in \mathbb{R}$, on a

$$(16) \quad |u_i^{n+1} - k| \leq |u_i^n - k| - q \left[\text{sg}(u_{i+1/2}^n - k) (f(u_{i+1/2}^n) - f(k)) - \text{sg}(u_{i-1/2}^n - k) (f(u_{i-1/2}^n) - f(k)) \right] ,$$

ce qui permet d'établir (I.25) en introduisant une fonction test $\phi \geq 0$. La démonstration de (I.28), ou mieux de (I.29), peut se faire de la façon suivante. Etant à variation localement bornée, $u(x,t)$ converge vers $u(x,0)$ dans $L^1_{loc}(\mathbb{R})$, lorsque t tend vers zéro, et il reste à vérifier que $u(x,0)$ coïncide avec u_0 presque partout, ce qui peut se faire en montrant que pour toute fonction $\phi \in C^2_0(\mathbb{R})$,

$$(17) \quad \int_{\mathbb{R}} [u(x,0) - u_0(x)] \phi(x) dx = 0 .$$

Dans le cas du schéma décentré, les estimations (13), (14), (15) et (16) sont plus faciles à obtenir. Ce cas est traité dans LE ROUX [1] et [2]. L'estimation (13) exprime la stabilité dans L^∞ du schéma, et (14), (15) signifient que la variation de la condition initiale est conservée au cours de la progression en temps. (16) est l'analogue discret de (I.25), et tient ainsi compte de la condition d'entropie (I.14). Cette condition est introduite dans le schéma par (8), qui fait intervenir une résolution correcte du problème de RIEMANN.

Le schéma de GODUNOV est souvent présenté comme une généralisation du schéma décentré au cas d'une fonction f non monotone. Le schéma suivant

$$(18) \quad u_i^{n+1} = u_i^n \cdot (f(u_{i+1}^n) - f(u_{i-1}^n)) + \frac{q}{2} \operatorname{sg}(u_{i+1}^n - u_i^n) \left| f(u_{i+1}^n) - f(u_i^n) \right| - \frac{q}{2} \operatorname{sg}(u_i^n - u_{i-1}^n) \left| f(u_i^n) - f(u_{i-1}^n) \right|$$

coïncide aussi avec le schéma décentré pour f monotone, mais peut construire des solutions faibles ne vérifiant pas la condition d'entropie (I.14). Il suffit par exemple de prendre l'équation de BURGERS (c'est-à-dire $f(u) = u^2/2$) et la condition initiale discrète définie par

$$(19) \quad u_i^0 = \begin{cases} 1 & \text{si } i \geq 0, \\ -1 & \text{si } i \leq -1, \end{cases}$$

pour constater que (18) se réduit à

$$u_i^{n+1} = u_i^n,$$

ce qui conduit à la solution stationnaire, tandis que la solution correcte est donnée par $u(x,t) = \operatorname{sg}_t(x)$. On peut cependant montrer les estimations (13), (14) et (15) pour le schéma (18) (cf. théorème 3).

x x
x

3 - DEMONSTRATION DU THEOREME 2

On note α_0 et γ_0 les constantes de LIPSCHITZ de g , par rapport aux variables u et x respectivement, puis α_1 et γ_1 les constantes de LIPSCHITZ de $\frac{\partial f}{\partial x}$ par rapport aux mêmes variables, et enfin β_0 et β_1 les bornes uniformes respectives de $g(0, \dots)$ et de $\frac{\partial f}{\partial x}(0, \dots)$, ce qui correspond à l'hypothèse (I.4).

Pour $h > 0$ fixé, on commence par démontrer les estimations analogues à (13), (14), (15). On traitera ensuite du passage à la limite, puis d'une démonstration d'une inégalité en k analogue à (16). On vérifiera enfin chacune des propriétés de la Définition 2.

A - Conservation de la borne uniforme

On va démontrer l'inégalité, pour tout $i \in \mathbb{Z}$, $n \in \mathbb{N}$

$$(20) \quad |u_i^{n+1}| \leq (1 + c_0 h) \sup(|u_{i-1/2}^n|, |u_i^n|, |u_{i+1/2}^n|) + c_0' h$$

avec $c_0 = q(\alpha_0 + \alpha_1)$, $c_0' = q(\beta_0 + \beta_1)$.

Cette estimation se démontre cas par cas.

Cas N° 1 $u_{i-1/2}^n \leq u_i^n \leq u_{i+1/2}^n$.

On introduit, pour tout $j \in \mathbb{Z}$, $\xi_{j\pm 1/4}^n \in I(u_j^n, u_{j\pm 1/2}^n)$ et

vérifiant

$$(21) \quad f(u_{j\pm 1/2}^n, (j\pm 1/2)h, nqh) - f(u_j^n, (j\pm 1/2)h, nqh) = (u_{j\pm 1/2}^n - u_j^n) \frac{\partial f}{\partial u}(\xi_{j\pm 1/4}^n, (j\pm 1/2)h, nqh)$$

On a alors, d'après (8) et (12), les inégalités

$$(22) \quad \begin{aligned} -1 &\leq q \frac{\partial f}{\partial u}(\xi_{j+1/4}^n, (j+1/2)h, nqh) \leq 0, \\ 0 &\leq q \frac{\partial f}{\partial u}(\xi_{j-1/4}^n, (j-1/2)h, nqh) \leq 1, \end{aligned}$$

en définissant $\xi_{j-1/4}^n \in I(u_{j-1/2}^n, u_j^n)$ de façon analogue à (21).

On introduit également $\bar{x}_i^n \in [(i-1/2)h, (i+1/2)h]$ et réalisant

$$(23) \quad f(u_{i+1/2}^n, (i+1/2)h, nqh) - f(u_{i-1/2}^n, (i-1/2)h, nqh) = h \frac{\partial f}{\partial x}(u_i^n, \bar{x}_i^n, nqh)$$

On obtient ainsi, à partir de (9)

$$(24) \quad \begin{aligned} u_i^{n+1} &= u_i^n - q \frac{\partial f}{\partial u}(\xi_{i+1/4}^n, (i+1/2)h, nqh) (u_{i+1/2}^n - u_i^n) \\ &\quad - q \frac{\partial f}{\partial u}(\xi_{i-1/4}^n, (i-1/2)h, nqh) (u_i^n - u_{i-1/2}^n) \\ &\quad - qh \left[\frac{\partial f}{\partial x}(u_i^n, \bar{x}_i^n, nqh) + g(u_i^n, ih, nqh) \right], \end{aligned}$$

Le dernier terme peut être majoré par $qh [(\alpha_0 + \alpha_1) |u_1^n| + \beta_0 + \beta_1]$, et les coefficients de $u_{i+1/2}^n$ et $u_{i-1/2}^n$ sont non négatifs. On majore dans un premier temps $u_{i-1/2}^n$ par u_i^n , et on regroupe le coefficient de u_i^n , qui devient alors non négatif ; on majore alors u_i^n par $u_{i+1/2}^n$. Les minoration s'effectuent en deux temps de façon analogue, et on obtient ainsi les inégalités

$$u_{i-1/2}^n - h(C_0 |u_i^n| + C'_0) \leq u_i^{n+1} \leq u_{i+1/2}^n + h(C_0 |u_i^n| + C'_0),$$

dont on déduit immédiatement (20).

Cas N° 2 $u_{i-1/2}^n \leq u_{i+1/2}^n \leq u_i^n$.

On reprend (24), en utilisant (22), et il vient

$$u_i^{n+1} \leq u_i^n + O.(u_{i+1/2}^n - u_i^n) + O.(u_i^n - u_{i-1/2}^n) - qh \left[\frac{\partial f}{\partial x} (u_i^n, \bar{x}_i^n, nqh) + g(u_i^n, ih, nqh) \right],$$

d'où

$$u_i^{n+1} \leq u_i^n + h(C_0 u_i^n + C'_0).$$

Pour minorer u_i^{n+1} , on introduit $\xi_i^n \in I(u_{i+1/2}^n, u_{i-1/2}^n)$ et le point $\bar{x}_i^n \in [(i-1/2)h, (i+1/2)h]$ tels que

$$(25) \left\{ \begin{aligned} f(u_{i+1/2}^n, (i-1/2)h, nqh) - f(u_{i-1/2}^n, (i-1/2)h, nqh) &= (u_{i+1/2}^n - u_{i-1/2}^n) \frac{\partial f}{\partial u} (\xi_i^n, (i-1/2)h, nqh) \\ f(u_{i+1/2}^n, (i+1/2)h, nqh) - f(u_{i+1/2}^n, (i-1/2)h, nqh) &= h \frac{\partial f}{\partial x} (u_{i+1/2}^n, \bar{x}_i^n, nqh). \end{aligned} \right.$$

On obtient alors, à partir de (9)

$$u_i^{n+1} = (u_i^n - u_{i+1/2}^n) + u_{i+1/2}^n - q \frac{\partial f}{\partial u} (\xi_i^n, (i-1/2)h, nqh) (u_{i+1/2}^n - u_{i-1/2}^n) - qh \left[\frac{\partial f}{\partial x} (u_{i+1/2}^n, \bar{x}_i^n, nqh) + g(u_i^n, ih, nqh) \right],$$

où $u_i^n - u_{i+1/2}^n$ est non négatif (on le minore par zéro), et le coefficient restant de $u_{i+1/2}^n$ est positif. On minore donc $u_{i+1/2}^n$ par $u_{i-1/2}^n$, et il vient, en introduisant les constantes de LIPSCHITZ de g et $\frac{\partial f}{\partial x}$ l'inégalité

$$u_i^{n+1} \geq u_{i-1/2}^n - qh (\alpha_1 |u_{i+1/2}^n| + \alpha_0 |u_i^n| + \beta_0 + \beta_1),$$

ce qui permet d'établir (20) dans ce cas, compte-tenu de la majoration faite précédemment.

Cas N° 3 $u_i^n \leq u_{i-1/2}^n \leq u_{i+1/2}^n$

La minoration de u_i^{n+1} est analogue à la majoration établie dans le cas précédent. Pour majorer u_i^{n+1} , on introduit η_i^n et \bar{x}_i^n analogues aux ξ_i^n et \bar{x}_i^n précédents (cf (25)), ce qui donne

$$u_i^{n+1} = (u_i^n - u_{i-1/2}^n) + u_{i-1/2}^n (1 + q \frac{\partial f}{\partial u} (\eta_i^n, (i+1/2)h, nqh)) - q \frac{\partial f}{\partial u} (\eta_i^n, (i+1/2)h, nqh) u_{i+1/2}^n - qh \left[\frac{\partial f}{\partial x} (u_{i-1/2}^n, \bar{x}_i^n, nqh) + g(u_i^n, ih, nqh) \right].$$

On majore u_i^n par $u_{i-1/2}^n$, puis $u_{i-1/2}^n$ dont le coefficient restant est non négatif par $u_{i+1/2}^n$, et il vient

$$u_i^{n+1} \leq u_{i+1/2}^n + qh [\alpha_1 |u_{i-1/2}^n| + \alpha_0 |u_i^n| + \beta_1 + \beta_0],$$

d'où (20) dans ce cas.

Cas N° 4 $u_i^n \leq u_{i+1/2}^n \leq u_{i-1/2}^n$.

La minoration de u_i^{n+1} est la même que dans le cas N° 3, et la majoration reprend la minoration du cas N° 2, en inversant les inégalités. On établit ainsi (20) dans ce cas.

Cas N° 5 $u_{i+1/2}^n \leq u_{i-1/2}^n \leq u_i^n$

La majoration de u_i^{n+1} est identique à celle du cas N° 2, et la majoration est celle du cas N° 3, d'où (20).

Cas N° 6 $u_{i+1/2}^n \leq u_i^n \leq u_{i-1/2}^n$

On reprend la technique du cas N° 1, en inversant les signes de $(u_{i+1/2}^n - u_i^n)$ et de $(u_{i-1/2}^n - u_i^n)$, et on obtient aussi (20).

Ces six cas recouvrant toutes les possibilités, l'estimation (20) est complètement démontrée.

De (20), on peut déduire

$$\sup_{i \in \mathbb{Z}} |u_i^{n+1}| \leq (1 + C_0 h) \sup_{i \in \mathbb{Z}} |u_i^n| + C'_0 h$$

et de proche en proche, il vient pour tout $n \leq N$

$$\sup_{i \in \mathbb{Z}} |u_i^n| \leq (1 + C_0 h)^n \sup_{i \in \mathbb{Z}} |u_i^0| + C'_0 h (1 + (1 + C_0 h) + \dots + (1 + C_0 h)^{n-1})$$

Or
$$h \sum_{v=0}^{n+1} (1 + C_0 h)^v \leq \frac{T}{q} e^{C_0 T/q}$$

et d'après (7)

$$\sup_{i \in \mathbb{Z}} |u_i^0| \leq |u_0|_{L^\infty(\mathbb{R})}$$

On obtient donc, pour tout n

$$(26) \quad \sup_{i \in \mathbb{Z}} |u_i^n| \leq (|u_0|_{L^\infty(\mathbb{R})} + C'_0 \frac{T}{q}) e^{C_0 T/q}$$

et donc u_n est uniformément borné dans $L^\infty(\mathbb{R} \times]0, T[)$.

B - Conservation de la variation bornée

On va démontrer les deux inégalités, pour tout $n \leq N$ et tout entier i

$$(27) \quad \sum_{|i| \leq I} |u_{i+1}^{n+1} - u_i^{n+1}| \leq (1 + C_1 h) \sum_{|i| \leq I+1} |u_{i+1}^n - u_i^n| + C'_1 I h^2$$

$$(28) \quad \sum_{|i| \leq I} |u_i^{n+1} - u_i^n| \leq \sum_{|i| \leq I+1} |u_{i+1}^n - u_i^n| + C_2 I h (1 + \sup_{i \in \mathbb{Z}} |u_i^n|)$$

avec $C_1 = q(\alpha_1 + \alpha_0) ; C'_1 = 3q(2\gamma_1 + \gamma_0) ;$

$$C_2 = 3q \text{Max}(\beta_0 + \beta_1, \alpha_0 + \alpha_1) = 3 \text{Max}(C_0, C'_0)$$

Pour $n \in \mathbb{N}$ et $i \in \mathbb{Z}$, on a, en utilisant (21) et (23), à partir de (9) écrite en $(i+1)h$ puis en ih , et après soustraction,

$$(29) \quad \begin{aligned} u_{i+1}^{n+1} - u_i^{n+1} &= (u_{i+1}^n - u_{i+1/2}^n) + (u_{i+1/2}^n - u_i^n) \\ &- q \frac{\partial f}{\partial u}(\xi_{i+5/4}^n, (i+3/2)h, nqh) (u_{i+3/2}^n - u_{i+1}^n) \\ &- q \frac{\partial f}{\partial u}(\xi_{i+3/4}^n, (i+1/2)h, nqh) (u_{i+1}^n - u_{i+1/2}^n) \\ &+ q \frac{\partial f}{\partial u}(\xi_{i+1/4}^n, (i+1/2)h, nqh) (u_{i+1/2}^n - u_i^n) \\ &+ q \frac{\partial f}{\partial u}(\xi_{i-1/4}^n, (i-1/2)h, nqh) (u_i^n - u_{i-1/2}^n) \\ &- qh \left[\frac{\partial f}{\partial x}(u_{i+1}^n, x_{i+1}^n, nqh) - \frac{\partial f}{\partial x}(u_i^n, x_i^n, nqh) \right] \\ &- qh [g(u_{i+1}^n, (i+1)h, nqh) - g(u_i^n, ih, nqh)] \end{aligned}$$

Après regroupement et d'après (22), on est assuré que les coefficients respectifs de $(u_{i+3/2}^n - u_{i+1}^n)$, $(u_{i+1}^n - u_{i+1/2}^n)$, $(u_{i+1/2}^n - u_i^n)$, $(u_i^n - u_{i-1/2}^n)$ sont non négatifs. De plus, les deux dernières lignes de (29) peuvent être estimées en utilisant les constantes de LIPSCHITZ de g et $\frac{\partial f}{\partial x}$. On passe en valeurs absolues, puis on remplace les termes dans la même disposition que (29); on aura :

$$\begin{aligned}
 |u_{i+1}^{n+1} - u_{i+1}^n| &\leq |u_{i+1}^n - u_{i+1/2}^n| + |u_{i+1/2}^n - u_i^n| \\
 &- \alpha \frac{\partial f}{\partial u} (\xi_{i+5/4}^n, (i+3/2)h, nqh) |u_{i+3/2}^n - u_{i+1}^n| \\
 &- \alpha \frac{\partial f}{\partial u} (\xi_{i+3/4}^n, (i+1/2)h, nqh) |u_{i+1}^n - u_{i+1/2}^n| \\
 &+ \alpha \frac{\partial f}{\partial u} (\xi_{i+1/4}^n, (i+1/2)h, nqh) |u_{i+1/2}^n - u_i^n| \\
 &+ \alpha \frac{\partial f}{\partial u} (\xi_{i-1/4}^n, (i-1/2)h, nqh) |u_i^n - u_{i-1/2}^n| \\
 &+ qh (\alpha_1 + \alpha_0) |u_{i+1}^n - u_i^n| + (2\gamma_1 + \gamma_0) qh^2.
 \end{aligned}$$

(30)

On effectue maintenant la somme, pour $|i| \leq I$, et on obtient, en tenant compte de (22) pour traiter les coefficients des termes au bord, et puisque

$$|u_{i+1}^n - u_{i+1/2}^n| + |u_{i+1/2}^n - u_i^n| = |u_{i+1}^n - u_i^n|,$$

l'inégalité

$$\sum_{|i| \leq I} |u_{i+1}^{n+1} - u_i^{n+1}| \leq (1 + C_1 h) \sum_{|i| \leq I+1} |u_{i+1}^n - u_i^n| + \frac{C_1}{3} (2I+1)h,$$

dont on déduit (27) de façon immédiate.

Pour établir (28), on reprend (9), où on introduit encore les termes définis en (21) et (23). On obtient, après passage aux valeurs absolues.

$$\begin{aligned}
 |u_i^{n+1} - u_i^n| &\leq \alpha \left| \frac{\partial f}{\partial u} (\xi_{i+1/4}^n, (i+1/2)h, nqh) \right| |u_{i+1/2}^n - u_i^n| \\
 &+ \alpha \left| \frac{\partial f}{\partial u} (\xi_{i-1/4}^n, (i-1/2)h, nqh) \right| |u_i^n - u_{i-1/2}^n| \\
 &+ qh \left| \frac{\partial f}{\partial x} (u_i^n, \bar{x}_i^n, nqh) - \frac{\partial f}{\partial x} (0, \bar{x}_i^n, nqh) \right| \\
 &+ qh |g(u_i^n, ih, nqh) - g(0, ih, nqh)| \\
 &+ qh \left(\left| \frac{\partial f}{\partial x} (0, \bar{x}_i^n, nqh) \right| + |g(0, ih, nqh)| \right).
 \end{aligned}$$

D'après (22), les coefficients des deux premiers termes du second membre sont majorés par 1, et les hypothèses de LIPSCHITZ ou de borne uniforme en $u = 0$ faites sur g et $\frac{\partial f}{\partial x}$ permettent, après sommation en $|i| \leq I$, d'établir

$$\begin{aligned}
 \sum_{|i| \leq I} |u_i^{n+1} - u_i^n| &\leq \sum_{|i| \leq I+1} (|u_{i+1}^n - u_{i+1/2}^n| + |u_{i+1/2}^n - u_i^n|) \\
 &+ q(\beta_0 + \beta_1 + (\alpha_0 + \alpha_1) \text{Sup}_i |u_i^n|) (2I+1)h,
 \end{aligned}$$

dont on déduit (28) immédiatement.

Remarquons qu'au lieu des normes $L^\infty(\mathbb{R} \times]0, T[)$, β_0 et β_1 auraient pu être les normes $L^1(\mathbb{R} \times]0, T[)$ de $g(0, \dots)$ et de $\frac{\partial f}{\partial x}(0, \dots)$, et (28) restait vrai. On peut aussi remarquer qu'en pratique on travaille sur des bornés, et que la continuité implique que ces fonctions sont bornées.

Soit $R > 0$. Pour h fixé, on note J_h le plus petit entier tel que

$$(31) \quad \Omega_R =]-R, R[\times]0, T[\left\{ (x, t) \mid (x, t) \in I_1 \times J_n \Rightarrow |i| + n < J_h \right\}$$

On introduit ensuite la fonction v_h , définie sur $\mathbb{R} \times]0, T[$ comme l'interpolée de degré un de u_h au sommets des rectangles

$$[ih, (i+1)h] \times [nqh, (n+1)qh].$$

Il est immédiat que

$$(32) \quad |v_h|_{L^\infty(\mathbb{R} \times]0, T[)} = |u_h|_{L^\infty(\mathbb{R} \times]0, T[)}$$

et donc est majoré, conformément à (26).

De plus v_h est dérivable sur chaque rectangle, et on obtient, en explicitant sa dérivée

$$\iint_{\Omega_R} \left| \frac{\partial v_h}{\partial t} \right| dx dt \leq \sum_n \sum_{|i| \leq J_h^{n-1}} |u_i^{n+1} - u_i^n| h$$

$$\leq \sum_n \sum_{|i| \leq J_h^n} |u_{i+1}^n - u_i^n| h + C_2 N J_h h^2 (1 + \sup_{i,n} |u_i^n|)$$

d'après (28). On utilise maintenant (27), qui donne de proche en proche, pour tout entier I,

$$(33) \quad \sum_{|i| \leq I} |u_{i+1}^n - u_i^n| \leq e^{C_1 T/q} \left\{ \sum_{|i| \leq I+n} |u_{i+1}^0 - u_i^0| + C_1' T(I+n)h/q \right\},$$

et on obtient

$$\left| \frac{\partial v_h}{\partial t} \right|_{L^1(\Omega_R)} \leq (N+1)h e^{C_1 T/q} \left(\sum_{|i| \leq J_h} |u_{i+1}^0 - u_i^0| + C_1' T(R + T/q) \right).$$

Or d'après (7),

$$(34) \quad \sum_{|i| \leq J_h} |u_{i+1}^0 - u_i^0| \leq \frac{1}{h} \sum_{|i| \leq J_h} \int_{I_i} |u_0(x+h) - u_0(x)| dx$$

$$\leq \frac{1}{h} \int_{|x| < R+T/q} |u_0(x+h) - u_0(x)| dx$$

$$\leq C(R + T/q),$$

car u_0 est à variation bornée localement. La quantité $C(R+T/q)$ peut être considérée comme une fonction réelle et non décroissante de R. On obtient ainsi qu'il existe une fonction Q_1 , réelle et non décroissante telle que

$$(35) \quad \left| \frac{\partial v_h}{\partial t} \right|_{L^1(\Omega_R)} \leq Q_1(R),$$

et de la même façon, il existe une fonction Q_2 non décroissante, telle que

$$(36) \quad \left| \frac{\partial v_h}{\partial x} \right|_{L^1(\Omega_R)} \leq Q_2(R),$$

et ainsi, d'après (32), (35), (36) v_h est uniformément bornée dans $W^{1,1}(\Omega_R)$.

C - Passage à la limite

D'après (32), $\{v_h\}_{h>0}$ est uniformément bornée dans $L^\infty(\mathbb{R} \times]0, T[)$, et on peut donc en extraire une suite $\{v_{h_m}\}$ avec h_m tendant vers zéro lorsque m tend vers l'infini, qui converge $\overset{m}{m}$ au sens de la topologie faible * de $L^\infty(\mathbb{R} \times]0, T[)$, vers une fonction $u \in L^\infty(\mathbb{R} \times]0, T[)$. Cette convergence faible * est insuffisante pour traiter les termes non linéaires, et on a donc besoin d'une convergence forte, qui va être obtenue à partir de (35) et (36).

Pour $R > 0$ fixé, on a, d'après (32), (35) et (36)

$$(37) \quad |v_h|_{W^{1,1}(\Omega_R)} \leq Q_0(R),$$

où Q_0 est une fonction non décroissante de R, donnée par

$$Q_0(R) = 2T \left(|u_0|_{L^\infty(\mathbb{R})} + C_0' \frac{T}{q} \right) C_0^{C_0 T/q} R + Q_1(R) + Q_2(R),$$

et donc indépendante de h. D'après le théorème de RIESZ-TAMARKIN, Ω_R étant borné, $W^{1,1}(\Omega_R)$ s'injecte de façon séquentiellement compacte dans $L^1(\Omega_R)$ vers une certaine fonction de $L^1(\Omega_R)$. On peut donc extraire de la suite $\{v_h\}_m$ introduite précédemment, une sous-suite, notée $\{v_{h_m, R^m}\}_m$, convergente dans $L^1(\Omega_R)$ vers une certaine fonction de $L^1(\Omega_R)$.

On applique ensuite le procédé diagonal, pour rendre cette suite extraite, et donc sa limite, indépendante de R. Partant de $\{v_{h_m, R^m}\}_m$, on peut trouver une estimation analogue à (37) sur Ω_{R+1} , et donc $\{v_{h_m, R^m}\}_m$ en extraire une suite $\{v_{h_m, R^{m+1}}\}_m$ convergente dans $L^1(\Omega_{R+1})$, dont la limite coïncide (presque partout) sur Ω_R avec la limite de $\{v_{h_m, R^m}\}_m$. On réitère ce procédé, ce qui permet pour tout entier k, d'extraire $\{v_{h_m, R^m}\}_m$ une suite $\{v_{h_m, R^{m+k}}\}_m$ convergente vers une fonction $v \in L^1(\Omega_{R+k})$, pour la topologie $L^1(\Omega_{R+k})$. La fonction v peut ainsi être définie en presque tout point de $(\mathbb{R} \times]0, T[)$, et appartient à $L^1_{loc}(\mathbb{R} \times]0, T[)$. On se restreint ensuite à la suite, dite diagonale, $v_{h_m} = v_{h_m, R^{m+1}}$.
et on se donne un ouvert Ω borné de $\mathbb{R} \times]0, T[$. A partir d'une certaine valeur m_0 , Ω est inclus dans Ω_{R+m_0} , et ainsi v_{h_m} converge vers v dans $L^1(\Omega)$.

Or v_{h_m} converge dans $L^\infty(\mathbb{R} \times]0, T[)$ faible * vers une fonction $u \in L^\infty(\mathbb{R} \times]0, T[)$. On en conclue que v coincide avec u presque partout, et est donc bornée, et que toute la suite v_h converge vers u dans $L^1_{loc}(\mathbb{R} \times]0, T[)$.

Il reste à estimer

$$\|u_{h_m} - v_{h_m}\|_{L^1(\Omega)}$$

On a, en décomposant sur chaque pavé

$$\|u_{h_m} - v_{h_m}\|_{L^1(\Omega)} \leq \sum_{n=0}^N \sum_{|i| \leq J_{h_m} - n - 1} (|u_i^{n+1} - u_i^n| + |u_{i+1}^n - u_i^n|) qh^2,$$

où J_{h_m} est un entier défini à partir de Ω_{R+m_0} par (31). En utilisant (27) et (28), qu'on majore comme pour établir (35) et (36), on obtient

$$(38) \quad \|u_{h_m} - v_{h_m}\|_{L^1(\Omega)} \leq \left\{ Q_1(R + m_0) + Q_2(R + m_0) \right\} h,$$

et ainsi u_{h_m} converge vers u dans $L^1(\Omega)$, et dans $L^\infty(\mathbb{R} \times]0, T[)$ faible *.

De plus, pour tout $R > 0$, $W^{1,1}(\Omega_R)$ s'injecte continûment dans $BV(\Omega_R)$, espace de Banach des fonctions à variation bornée au sens de TONELLI CESARI sur Ω_R . On en conclue, par l'extraction d'une sous-suite, que u appartient à $BV(\Omega_R)$, et ceci pour tout Ω_R , donc que u appartient à $BV_{loc}(\mathbb{R} \times]0, T[)$. On pourrait dès maintenant, à partir de (9), vérifier que u est solution faible.

D - Démonstration d'une nouvelle estimation

Soit $k \in \mathbb{R}$ et $h > 0$; on va démontrer l'inégalité suivante, qui est l'analogue discret de (I.25) et correspond à (16) dans le cas général, pour tout $i \in \mathbb{Z}$ et $n < N$

$$(39) \quad \begin{aligned} & |u_i^{n+1} - k| \leq |u_i^n - k| \\ & - q \operatorname{sg}(u_{i+1/2}^n - k) \left(f(u_{i+1/2}^n, (i+1/2)h, nqh) - f(k, (i+1/2)h, nqh) \right) \\ & + q \operatorname{sg}(u_{i-1/2}^n - k) \left(f(u_{i-1/2}^n, (i-1/2)h, nqh) - f(k, (i-1/2)h, nqh) \right) \\ & - qh \operatorname{sg}(u_i^{n+1} - k) \left[g(u_i^n, ih, nqh) \right. \\ & \quad \left. + \frac{1}{h} f(k, (i+1/2)h, nqh) - f(k, (i-1/2)h, nqh) \right] \\ & + Ch |u_i^{n+1} - u_i^n| + C' h^2, \end{aligned}$$

avec

$$C = 2q \alpha_1 ; \quad C' = 2q \gamma_1.$$

Comme pour la stabilité, cette inégalité se démontre cas par cas. Les constantes C et C' n'interviennent que dans le cas N° 2.

$$\text{Cas N° 1} \quad k < \operatorname{Inf} \{ u_i^{n+1}, u_{i-1/2}^n, u_i^n, u_{i+1/2}^n \}$$

$$\text{ou } k > \operatorname{Sup} \{ u_i^{n+1}, u_{i-1/2}^n, u_i^n, u_{i+1/2}^n \}.$$

On introduit k dans (9), qu'on met sous la forme suivante

$$(40) \quad \begin{aligned} u_i^{n+1} - k &= u_i^n - k \\ & - q \left[f(u_{i+1/2}^n, (i+1/2)h, nqh) - f(k, (i+1/2)h, nqh) \right] \\ & + q \left[f(u_{i-1/2}^n, (i-1/2)h, nqh) - f(k, (i-1/2)h, nqh) \right] \\ & - qh \left[g(u_i^n, ih, nqh) + \frac{1}{h} \left\{ f(k, (i+1/2)h, nqh) - f(k, (i-1/2)h, nqh) \right\} \right]. \end{aligned}$$

On multiplie ensuite par $\operatorname{sg}(u_i^{n+1} - k)$, et on obtient (39) avec $C = C' = 0$, donc aussi les constantes données précédemment, en remarquant que

$$\operatorname{sg}(u_i^{n+1} - k) = \operatorname{sg}(u_{i+1/2}^n - k) = \operatorname{sg}(u_i^n - k) = \operatorname{sg}(u_{i-1/2}^n - k).$$

Cas N° 2 $u_i^{n+1} < k < \text{Inf}(u_{i-1/2}^n, u_i^n, u_{i+1/2}^n)$
 ou $u_i^{n+1} > k > \text{Sup}(u_{i-1/2}^n, u_i^n, u_{i+1/2}^n)$.

Ce cas ne peut intervenir que pour l'équation générale (1) ;
 il est impossible dans le cas de l'équation (3). D'après la démonstration
 de (20), on a toujours

$$(41) \quad \begin{aligned} \text{Inf}(u_{i-1/2}^n, u_i^n, u_{i+1/2}^n) &\leq u_i^{n+1} + qh \left[\frac{\partial f}{\partial x}(u_i^n, \bar{x}_i^n, nqh) + g(u_i^n, ih, nqh) \right] \\ &\leq \text{Sup}(u_{i-1/2}^n, u_i^n, u_{i+1/2}^n) . \end{aligned}$$

On avait alors majoré le coefficient de qh par
 $(\alpha_0 + \alpha_1) u_i^n + \beta_0 + \beta_1$. On introduit $y_i^n \in [(i-1/2)h, (i+1/2)h]$ et vérifiant
 pour k fixé

$$f(k, (i+1/2)h, nqh) - f(k, (i-1/2)h, nqh) = h \frac{\partial f}{\partial x}(k, y_i^n, nqh) ,$$

et il vient à partir de (9) et en multipliant par $sg(u_i^{n+1} - k)$,

$$\begin{aligned} &sg(u_i^{n+1} - k) \left[u_i^{n+1} - k + qh \left(\frac{\partial f}{\partial x}(u_i^n, \bar{x}_i^n, nqh) + g(u_i^n, ih, nqh) \right) \right] \\ &\leq sg(u_i^n - k) \left[u_i^{n+1} - k + qh \left(\frac{\partial f}{\partial x}(u_i^n, \bar{x}_i^n, nqh) + g(u_i^n, ih, nqh) \right) \right] \\ &\leq sg(u_i^n - k) \left[u_i^n - k - qh \left(\frac{\partial f}{\partial x}(k, y_i^n, nqh) - \frac{\partial f}{\partial x}(u_i^n, \bar{x}_i^n, nqh) \right) \right. \\ &\quad \left. - q(f(u_{i+1/2}^n, (i+1/2)h, nqh) - f(k, (i+1/2)h, nqh)) \right. \\ &\quad \left. + q(f(u_{i-1/2}^n, (i-1/2)h, nqh) - f(k, (i-1/2)h, nqh)) \right] , \end{aligned}$$

où la première inégalité se déduit de (41). De plus, dans le cas présent,
 on a

$$sg(u_i^n - k) = sg(u_{i+1/2}^n - k) = sg(u_{i-1/2}^n - k) = -sg(u_i^{n+1} - k) ,$$

ce qui permet d'obtenir (39), en utilisant successivement

$$|sg(u_i^{n+1} - k) - sg(u_i^n - k)| = 2 ,$$

$$\left| \frac{\partial f}{\partial x}(k, y_i^n, nqh) - \frac{\partial f}{\partial x}(u_i^n, \bar{x}_i^n, nqh) \right| \leq \alpha_1 |k - u_i^n| + \gamma_1 |y_i^n - \bar{x}_i^n| ,$$

puis

$$|k - u_i^n| \leq |u_i^{n+1} - u_i^n| ; \quad |y_i^n - \bar{x}_i^n| \leq h .$$

Ce cas est le seul pour lequel C et C' sont différents de
 zéro.

Cas N° 3 $u_{i-1/2}^n < k < u_i^n \leq u_{i+1/2}^n$
 ou $u_{i-1/2}^n > k > u_i^n \geq u_{i+1/2}^n$.

On reprend (40), où on introduit les $\xi_{i+1/4}^n$ définis en (21),
 et $k_i^n \in I(k, u_i^n)$, vérifiant

$$f(u_i^n, (i+1/2)h, nqh) - f(k, (i+1/2)h, nqh) = (u_i^n - k) \frac{\partial f}{\partial x}(k_i^n, (i+1/2)h, nqh) .$$

Il vient alors

$$\begin{aligned} &f(u_{i+1/2}^n, (i+1/2)h, nqh) - f(k, (i+1/2)h, nqh) \\ &= (u_{i+1/2}^n - u_i^n) \frac{\partial f}{\partial u}(\xi_{i+1/4}^n, (i+1/2)h, nqh) + (u_i^n - k) \frac{\partial f}{\partial u}(k_i^n, (i+1/2)h, nqh) . \end{aligned}$$

On introduit également $k_{i-1/2}^n \in I(u_{i-1/2}^n, k)$ tel que

$$(42) \quad \begin{aligned} &f(u_{i-1/2}^n, (i-1/2)h, nqh) - f(k, (i-1/2)h, nqh) \\ &= (u_{i-1/2}^n - k) \frac{\partial f}{\partial u}(k_{i-1/2}^n, (i-1/2)h, nqh) . \end{aligned}$$

On obtient alors

$$\begin{aligned} u_i^{n+1} - k &= (u_i^n - k) \left(1 - q \frac{\partial f}{\partial u}(k_i^n, (i+1/2)h, nqh) \right) \\ &\quad + (u_{i+1/2}^n - u_i^n) \left(- q \frac{\partial f}{\partial u}(\xi_{i+1/4}^n, (i+1/2)h, nqh) \right) \\ &\quad + (u_{i-1/2}^n - k) \quad q \frac{\partial f}{\partial u}(k_{i-1/2}^n, (i-1/2)h, nqh) \\ &\quad - q h \left[g(u_i^n, ih, nqh) + \frac{\partial f}{\partial x}(k, y_i^n, nqh) \right] , \end{aligned}$$

où d'après (22), ou par construction de $u_{i+1/2}^n$ et u_i^n , les coefficients de $(u_i^n - k)$, $(u_{i+1/2}^n - u_i^n)$ et $(u_{i-1/2}^n - k)$ sont non négatifs (il faut utiliser pour le dernier de ces coefficients, que $k \in I(u_i^n, u_{i-1/2}^n)$).

On multiplie par $sg(u_i^{n+1} - k)$, pour obtenir

$$\begin{aligned} |u_i^{n+1} - k| &\leq |u_i^n - k| \left(1 - \alpha \frac{\partial f}{\partial u} (k_i^n, (i+1/2)h, nqh) \right) \\ &+ |u_{i+1/2}^n - u_i^n| \left(-\alpha \frac{\partial f}{\partial u} (\xi_{i+1/4}^n, (i+1/2)h, nqh) \right) \\ &+ |u_{i-1/2}^n - k| \alpha \frac{\partial f}{\partial u} (k_{i-1/2}^n, (i-1/2)h, nqh) \\ &- \alpha h \operatorname{sg}(u_i^{n+1} - k) \left[g(u_i^n, ih, nqh) + \frac{\partial f}{\partial x} (k, y_i^n, nqh) \right], \end{aligned}$$

ce qui donne (39) après reconstitution des termes et en utilisant

$$(43) \quad \begin{aligned} &\frac{\partial f}{\partial u} (k_i^n, (i+1/2)h, nqh) |u_i^n - k| + \frac{\partial f}{\partial u} (\xi_{i+1/4}^n, (i+1/2)h, nqh) |u_{i+1/2}^n - u_i^n| \\ &= \operatorname{sg}(u_{i+1/2}^n - k) \left[f(u_{i+1/2}^n, (i+1/2)h, nqh) - f(k, (i+1/2)h, nqh) \right]. \end{aligned}$$

Cas N° 4 $u_{i-1/2}^n \leq u_i^n < k < u_{i+1/2}^n$
ou $u_{i-1/2}^n \geq u_i^n > k > u_{i+1/2}^n$.

La démonstration de (39) pour ce cas correspond à celle faite pour le cas précédent, en inversant les rôles de $u_{i-1/2}^n$ et de $u_{i+1/2}^n$. On introduit les quantités $\xi_{i-1/4}^n$ et k_i^n , ainsi que $k_{i+1/2}^n$ défini de façon analogue à (42), et il vient

$$\begin{aligned} u_i^{n+1} - k &= (u_i^n - k) \left(1 + \alpha \frac{\partial f}{\partial u} (k_i^n, (i-1/2)h, nqh) \right) \\ &+ (u_{i+1/2}^n - k) \left(-\alpha \frac{\partial f}{\partial u} (k_{i+1/2}^n, (i+1/2)h, nqh) \right) \\ &+ (u_{i-1/2}^n - u_i^n) \alpha \frac{\partial f}{\partial u} (\xi_{i-1/4}^n, (i-1/2)h, nqh) \\ &- \alpha h \left[g(u_i^n, ih, nqh) + \frac{\partial f}{\partial x} (k, y_i^n, nqh) \right]. \end{aligned}$$

Les coefficients de $(u_i^n - k)$, $(u_{i+1/2}^n - k)$ et $(u_{i-1/2}^n - u_i^n)$ sont non négatifs ; on multiplie par $sg(u_i^{n+1} - k)$, on majore et on regroupe en utilisant une formule analogue à (43). On obtient ainsi (39).

Cas N° 5 $u_{i-1/2}^n < k < u_{i+1/2}^n \leq u_i^n$
ou $u_{i-1/2}^n > k > u_{i+1/2}^n \geq u_i^n$.

On met (40) sous la forme

$$\begin{aligned} u_i^{n+1} - k &= (u_i^n - u_{i+1/2}^n) + (u_{i+1/2}^n - k) \left(1 - \alpha \frac{\partial f}{\partial u} (k_{i+1/2}^n, (i+1/2)h, nqh) \right) \\ &+ (u_{i-1/2}^n - k) \alpha \frac{\partial f}{\partial u} (k_{i-1/2}^n, (i-1/2)h, nqh) \\ &- \alpha h \left[g(u_i^n, ih, nqh) + \frac{\partial f}{\partial x} (k, y_i^n, nqh) \right], \end{aligned}$$

où les coefficients de $(u_{i+1/2}^n - k)$ et de $(u_{i-1/2}^n - k)$ sont non négatifs, d'après (22) et puisque $k \in I(u_{i-1/2}^n, u_i^n)$. On multiplie par $sg(u_i^{n+1} - k)$, on majore et on regroupe, ce qui donne (39) en remarquant que

$$|u_i^n - u_{i+1/2}^n| + |u_{i+1/2}^n - k| = |u_i^n - k|.$$

Cas N° 6 $u_{i-1/2}^n \leq u_{i+1/2}^n < k < u_i^n$
ou $u_{i-1/2}^n \geq u_{i+1/2}^n > k > u_i^n$.

On reprend (40), où les coefficients de $(u_{i+1/2}^n - k)$ et de $(u_{i-1/2}^n - k)$ sont non négatifs, après avoir fait apparaître $\xi_{i+1/4}^n$ et $\xi_{i-1/4}^n$, car

$$k \in I(u_{i+1/2}^n, u_i^n) \cap I(u_{i-1/2}^n, u_i^n).$$

On multiplie par $sg(u_i^{n+1} - k)$, et on conclue comme dans les cas précédents.

Cas N° 7 $u_i^n < k < u_{i-1/2}^n \leq u_{i+1/2}^n$

ou $u_i^n > k > u_{i-1/2}^n \geq u_{i+1/2}^n$.

Ce cas est analogue au précédent : $k \in I(u_{i+1/2}^n, u_i^n) \cap I(u_{i-1/2}^n, u_i^n)$.

Cas N° 8 $u_i^n \leq u_{i-1/2}^n < k < u_{i+1/2}^n$

ou $u_i^n \geq u_{i-1/2}^n > k > u_{i+1/2}^n$.

Ce cas se traite comme le cas N° 5. Il vient

$$u_i^{n+1} - k = (u_i^n - u_{i-1/2}^n) + (u_{i-1/2}^n - k) \left(1 + q \frac{\partial f}{\partial u} (k, y_i^n, nqh) \right) + (u_{i+1/2}^n - k) \left(-q \frac{\partial f}{\partial u} (k, y_{i+1/2}^n, (i+1/2)h, nqh) \right) - qh \left[g(u_i^n, ih, nqh) + \frac{\partial f}{\partial x} (k, y_i^n, nqh) \right]$$

où les coefficients de $(u_{i-1/2}^n - k)$ et de $(u_{i+1/2}^n - k)$ sont non négatifs. On conclue en multipliant par $sg(u_i^{n+1} - k)$ et en remarquant que

$$|u_i^n - u_{i-1/2}^n| + |u_{i-1/2}^n - k| = |u_i^n - k|$$

On vient de démontrer (39) pour $k \in \{u_{i-1/2}^n, u_i^n, u_{i+1/2}^n, u_i^{n+1}\}$

Il est immédiat, par continuité, que (39) est encore vrai pour $k \neq u_i^{n+1}$, et il ne reste plus qu'à considérer le cas où k et u_i^{n+1} sont égaux.

Tout d'abord, ce cas est trivialement réglé lorsque

$$g(u_i^n, ih, nqh) + \frac{\partial f}{\partial x} (k, y_i^n, nqh) = 0$$

Lorsque ce terme est strictement positif, on approche $k = u_i^{n+1}$ par $k' < k$. Il vient alors, lorsque $k - k'$ est suffisamment petit, en définissant y_i^{n+1} à partir de k' comme l'était y_i^n à partir de k ,

$$- qh sg(u_i^{n+1} - k') \left[g(u_i^n, ih, nqh) + \frac{\partial f}{\partial x} (k', y_i^{n+1}, nqh) \right] < 0,$$

et donc

$$|u_i^{n+1} - k'| \leq |u_i^n - k'| - q sg(u_{i+1/2}^n - k') \left(f(u_{i+1/2}^n, (i+1/2)h, nqh) - f(k', (i+1/2)h, nqh) \right) + q sg(u_{i-1/2}^n - k') \left(f(u_{i-1/2}^n, (i-1/2)h, nqh) - f(k', (i-1/2)h, nqh) \right) + C h |u_i^{n+1} - u_i^n| + C' h^2.$$

On fait tendre k' vers k , et on obtient (39) car $sg(u_i^{n+1} - k)$

est nul par définition de la fonction sg .

Si maintenant

$$g(u_i^n, ih, nqh) + \frac{\partial f}{\partial x} (k, y_i^n, nqh) < 0,$$

on approche k par $k' > k$, et on procède comme ci-dessus, et ainsi (39) est démontré pour tout $k \in \mathbb{R}$.

E - Démonstration de (I.25)

Soit $k \in \mathbb{R}$, $\phi \in C_{0+}^2(\mathbb{R} \times]0, T[)$ et $h \in \{h_m\}_m$. Pour tout $i \in \mathbb{Z}$ et $n < N$, on pose

$$\phi_i^n = \frac{1}{qh^2} \iint_{I_i^n \times J_n} \phi(x, t) dx dt,$$

et on suppose h suffisamment petit pour que tous les ϕ_i^0 soient nuls. On multiplie (39) par $h \phi_i^n$, et on effectue l'équivalent discret d'une intégration par parties sur la variable d'espace, puis sur la variable de temps. Il vient

$$\sum_{n=0}^N \sum_{i \in \mathbb{Z}} \left\{ |u_i^n - k| \frac{\phi_i^{n+1} - \phi_i^n}{qh} + sg(u_{i+1/2}^n - k) \left(f(u_{i+1/2}^n, (i+1/2)h, nqh) - f(k, (i+1/2)h, nqh) \right) \frac{\phi_{i+1}^n - \phi_i^n}{h} \right. \\ \left. - sg(u_i^{n+1} - k) \left(g(u_i^n, ih, nqh) + \frac{\partial f}{\partial x} (k, y_i^n, nqh) \right) \phi_i^n qh^2 \right\} \\ \geq - \sum_{n=0}^N \sum_{i \in \mathbb{Z}} \frac{1}{q} (C |u_i^{n+1} - u_i^n| + C'h) \phi_i^n qh^2.$$

Dans le but d'appliquer le lemme 2 (cf. chap. I), on remplace k par ξ dans (44), on multiplie par

$$\frac{1}{2} \text{sg}'_{\eta} (\xi - k) ,$$

et on intègre en ξ . On passe ensuite à la limite en m, puis en η . On obtient, d'après le lemme 2, avec Ω un ouvert de $]\mathbb{R} \times]0, T[$ contenant le support de ϕ (et de son approximation par les ϕ_i^n),

$$v_m = u_{h_m}, w_m = (u_{h_m}, \bar{u}_{h_m}, x_m, \bar{x}_m, y_m, \phi_m, \phi_{1,m}, \phi_{2,m}, t_m)$$

avec successivement

$$\begin{aligned} \bar{u}_{h_m}(x,t) &= u_{i+1/2}^n \text{ si } (x,t) \in I_i \times J_n, \text{ qui converge vers } u, \\ x_m(x,t) &= ih \text{ si } x \in I_i, \quad " \quad " \quad x, \\ \bar{x}_m(x,t) &= (i+1/2)h \text{ si } x \in I_i, \quad " \quad " \quad x, \\ y_m(x,t) &= y_i^n \text{ si } x \in I_i, \quad " \quad " \quad x, \\ \phi_m(x,t) &= \phi_i^n \text{ si } (x,t) \in I_i \times J_n \quad " \quad " \quad \phi, \\ \phi_{1,m}(x,t) &= \frac{1}{qh} (\phi_i^{n+1} - \phi_i^n) \text{ si } (x,t) \in I_i \times J_n \quad " \quad " \quad \frac{\partial \phi}{\partial t}, \\ \phi_{2,m}(x,t) &= \frac{1}{h} (\phi_{i+1}^n - \phi_i^n) \text{ si } (x,t) \in I_i \times J_n \quad " \quad " \quad \frac{\partial \phi}{\partial x}, \\ t_m(x,t) &= nqh \text{ si } t \in J_n \quad " \quad " \quad t, \end{aligned}$$

en convenant de noter x et t les fonctions "identités" respectives de \mathbb{R} et $]\mathbb{R}, T[$ (D'après le théorème de la convergence dominée de LEBESGUE, toutes ces convergences ont lieu dans $L^1(\Omega \times]0, T[)$), l'inégalité

$$\begin{aligned} (45) \quad & \iint_{\Omega} \left\{ |u - k| \frac{\partial \phi}{\partial t} + \text{sg}(u-k) (f(u, x, t) - f(k, x, t)) \frac{\partial \phi}{\partial x} \right. \\ & \left. - \text{sg}(u-k) \left[g(u, x, t) + \frac{\partial f}{\partial x}(k, x, t) \right] \phi \right\} dx dt \\ & \geq \lim_{\eta \rightarrow 0} \lim_{m \rightarrow \infty} \left(\int_{\mathbb{R}} -\frac{1}{2} \text{sg}'_{\eta}(\xi - k) d\xi \right) \left(\sum_n \sum_i \left(\frac{C}{q} |u_i^{n+1} - u_i^n| + C' \frac{h_m}{q} \right) \phi_i^n q h_m^2 \right) \end{aligned}$$

Le passage à la limite sur le second membre de (45) est immédiat. On a un coefficient égal à 1 provenant de la limite en η , et on sait que

$$\sum_n \sum_i \left(\frac{C}{q} |u_i^{n+1} - u_i^n| + \frac{C' h_m}{q} \right) \phi_i^n q h_m$$

est borné uniformément en h. En effet, si Ω est inclus dans $]-R, R[\times]0, T[$, ce terme est majoré par

$$T \|\phi\|_{L^\infty(\Omega)} \frac{C}{q} Q_1(R) + 2R \frac{C'}{q} .$$

Il ne reste plus qu'un h_m en facteur au second membre de (45), et la limite est donc nulle. (I.25) est démontré.

F - Démonstration de (I. 29)

Soit $R > 0$. On sait que u est à variation bornée sur Ω_R , et ainsi en presque tout $x \in]-R, R[$, u(x,t) admet une limite notée u(x,0) lorsque t tend vers zéro. On veut démontrer que u(x,0) et $u_0(x)$ coïncident pour presque tout $x \in]-R, R[$.

D'après le théorème de LEBESGUE, u(x,t) étant uniformément bornée et convergente presque partout, sa limite est mesurable et bornée, et la convergence a lieu dans $L^1(-R, R)$. Il suffit de montrer que pour toute fonction $\phi \in C^1_0(-R, R)$,

$$(46) \quad \int_{-R}^R (u(x,0) - u_0(x)) \phi(x) dx = 0 .$$

La suite $\{u_{h_m}\}$ converge vers u dans $L^1(\Omega_R)$. On peut donc en extraire une sous-suite $\{u_{h'_m}\}$ convergente vers u presque partout sur Ω_R , lorsque m tend vers l'infini (et donc h'_m vers 0).

D'après le théorème de FUBINI, pour presque tout $t \in]0, T[$, cette sous-suite converge presque partout vers u(.,t), et en restant uniformément bornée. On conclue donc par le théorème de LEBESGUE que pour presque tout $t \in]0, T[$, une sous-suite $u_{h'_m}(.,t)$ vérifie

$$(47) \quad \lim_{m \rightarrow \infty} \|u_{h_m}(\cdot, t) - u(\cdot, t)\|_{L^1(-R, R)} = 0.$$

On sait également que pour $t = 0$, d'après (7), la fonction u_h^0 définie par

$$u_h^0(x) = u_i^0 \quad \text{si } x \in I_i,$$

converge vers u_0 dans $L^1(-R, R)$. Il suffit donc pour avoir (46), de démontrer que lorsque t tend vers zéro,

$$(48) \quad \lim_{m \rightarrow \infty} \int_{-R}^R (u_{h_m}(x, t) - u_{h_m}^0(x)) \phi(x) dx \rightarrow 0.$$

Or pour $I = [R/h]$, le premier membre de (48) vaut, en reprenant la même discrétisation de ϕ que précédemment,

$$\lim_{m \rightarrow \infty} \left\{ \sum_{|i| \leq I} \sum_{v=0}^{n-1} (u_i^{v+1} - u_i^v) \phi_i h \right\},$$

lorsque $t \in J_n$. En utilisant (9), cette expression devient

$$\lim_{m \rightarrow \infty} \sum_{v=0}^{n-1} \left\{ \sum_{|i| \leq I} f(u_{i+1/2}^n, (i+1/2)h, nqh) \frac{\phi_{i+1} - \phi_i}{h} q h^2 - \sum_{|i| \leq I} [g(u_i, ih, nqh) \phi_i] q h^2 \right\},$$

qui peut être majorée par un terme de la forme

$$\lim_{m \rightarrow \infty} \sum_{v=0}^{n-1} (M \|\phi'\|_{L^\infty} + \alpha_0 \|\phi\|_{L^\infty}) R \|u_h\|_{L^\infty} q h,$$

lui-même majoré à la limite par un terme de la forme $C(R)t$ et lorsque t tend vers zéro il est immédiat qu'on a (48).

Or en se restreignant aux t pour lesquels (47) est valable, le premier membre de (48) est égal à

$$\int_{-R}^R (u(x, t) - u_0(x)) \phi(x) dx,$$

et qui converge vers le premier membre de (46). On a ainsi établi (46), et le théorème 2 est démontré, puisque l'unicité implique la convergence de toute la famille $\{u_h\}_{h>0}$.

CHAPITRE III

ETUDE DE QUELQUES SCHEMAS D'ORDRE UN ET D'ORDRE DEUX

1. CONSTRUCTION D'UNE FAMILLE DE SCHEMAS POUR LE PROBLEME MONODIMENSIONNEL

En reprenant les notations et le maillage introduits au chapitre précédent, le schéma de LAX s'écrit, pour le problème monodimensionnel

$$(1) \quad u_i^{n+1} = \frac{1}{2} (u_{i+1}^n + u_{i-1}^n) - \frac{q}{2} (f(u_{i+1}^n, (i+1)h, nqh) - f(u_{i-1}^n, (i-1)h, nq) - q h g(u_i^n, ih, nqh)),$$

ce qui se réduit, dans le cas simple où f ne dépend que de u , et g est nul, à

$$(2) \quad u_i^{n+1} = \frac{1}{2} (u_{i+1}^n + u_{i-1}^n) - \frac{q}{2} (f(u_{i+1}^n) - f(u_{i-1}^n)).$$

Sous la condition de stabilité de COURANT-FRIEDRICHS-LEWY (II.3), ce schéma permet de construire une solution u_h convergente vers la solution du problème (II.3), satisfaisant à une condition initiale donnée u_0 , et vérifiant la condition d'entropie (c'est-à-dire I.25). La démonstration est analogue à celle du théorème 2, en étant plus simple. La convergence vers

une solution faible a été établie par CONWAY et SMOLLER, dans le cas multi-dimensionnel, et la convergence vers la solution au sens de la définition 2 a été démontrée par différents auteurs, dont A. DOUGLIS, OHARU et TAKAHASHI, et LE ROUX [1] ou [2] dans le cadre d'une famille plus générale de schémas.

Comme le schéma de GODUNOV, le schéma de LAX peut être mis sous une forme proche de (II.18), ce qui permet de faire apparaître un terme de viscosité, analogue au second membre de (I.16). Cette forme est la suivante, dans le cas simplifié (II.3)

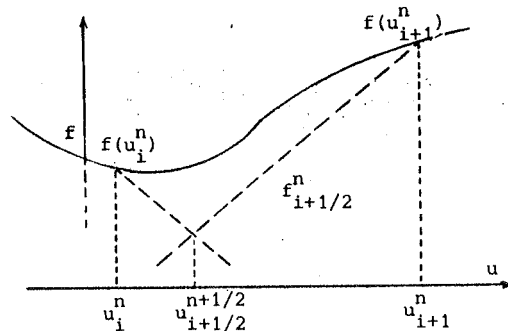
$$(3) \quad u_i^{n+1} = u_i^n - \frac{q}{2} (f(u_{i+1}^n) - f(u_{i-1}^n)) + \frac{1}{2} (u_{i+1}^n - u_i^n) - \frac{1}{2} (u_i^n - u_{i-1}^n).$$

On peut également construire le schéma de LAX d'une façon analogue à celle qui a permis d'écrire le schéma de GODUNOV. On suppose, pour i et n fixés, connaître u_i^n et u_{i+1}^n , et on veut construire une quantité $u_{i+1/2}^n$, solution en $x = (i+1/2)h$ d'un certain problème de RIEMANN, pour $t \leq (n+1)qh$. Dans cette construction, on introduit la condition d'entropie en remplaçant f sur l'intervalle $I(u_i^n, u_{i+1}^n)$ par une approximation de son enveloppe convexe si $u_i^n < u_{i+1}^n$, ou concave si $u_i^n > u_{i+1}^n$, notée $f_{i+1/2}^n$.

Cette enveloppe $f_{i+1/2}^n$ est telle que son graphe soit formé de deux segments rectilignes, dont le premier passe par le point $(u_i^n, f(u_i^n))$, avec la pente $-1/q$, et le second passe en $(u_{i+1}^n, f(u_{i+1}^n))$ avec la pente $1/q$. Ces pentes sont ainsi choisies de façon que sur $I(u_i^n, u_{i+1}^n)$, on ait toujours

$$(4) \quad \text{sg}(u_{i+1}^n - u_i^n) f_{i+1/2}^n(k) \leq \text{sg}(u_{i+1}^n - u_i^n) f(k),$$

lorsque la condition de stabilité de COURANT FRIEDRICHS LEWY est satisfaite. Comme pour le schéma de GODUNOV, la solution du problème de RIEMANN, avec $f_{i+1/2}^n$ au lieu de f dans le cas présent, est la quantité $u_{i+1/2}^{n+1/2}$ vérifiant



$$(5) \quad \min_{k \in I(u_i^n, u_{i+1}^n)} \{ \text{sg}(u_{i+1}^n - u_i^n) f_{i+1/2}^n(k) \},$$

et qui est donnée explicitement par

$$(6) \quad u_{i+1/2}^{n+1/2} = \frac{1}{2} (u_i^n + u_{i+1}^n) - \frac{q}{2} (f(u_{i+1}^n) - f(u_i^n)),$$

c'est-à-dire exactement le schéma de LAX, lorsque les pas de discrétisation en temps et en espace sont divisés par deux.

On pose ensuite, comme en (II.10)

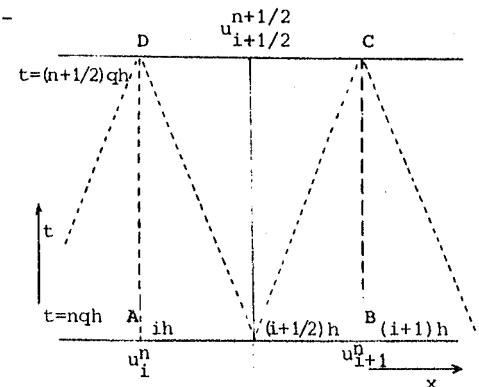
$$(7) \quad u_i^{n+1} = u_i^n - q (f_{i+1/2}^n(u_{i+1/2}^{n+1/2}) - f_{i-1/2}^n(u_{i-1/2}^{n+1/2})),$$

ce qui donne, en remarquant que

$$(8) \quad f_{i+1/2}^n(u_{i+1/2}^{n+1/2}) = \frac{1}{2} (f(u_{i+1}^n) + f(u_i^n)) + \frac{u_i^n - u_{i+1}^n}{2q},$$

la forme classique du schéma de LAX, donnée en (2).

La formule (6) peut être obtenue d'une autre façon, en calculant la moyenne de la solution du problème de RIEMANN associé à u_i^n et u_{i+1}^n , sur l'intervalle $]ih, (i+1)h[$ au temps $t = (n + \frac{1}{2})qh$. Cette moyenne est obtenue en intégrant l'équation (II.3) sur le rectangle ABCD de la figure ci-contre, et sachant que la solution est constante sur les côtés AD et BC, où elle vaut respectivement u_i^n et u_{i+1}^n . Cette interprétation, comparée à la construction du schéma de GODUNOV et rapprochée de la construction du schéma de LAX faite précédemment permet de faire certaines remarques. Tout d'abord, le fait que $u_{i+1/2}^{n+1/2}$ soit une moyenne, donc non calculée exactement, montre



qu'il y a diffusion dans cette première étape pour le schéma de LAX, et que cette diffusion provient essentiellement de l'écart entre les graphes de f et de $f_{i+1/2}^n$.

La deuxième étape, c'est-à-dire (8) pour le schéma de LAX, apporte aussi son lot de diffusion, dans la mesure où on utilise encore $f_{i+1/2}^n$ au lieu de f . En remplaçant (8) par

$$(9) \quad u_i^{n+1} = u_i^n - q(f(u_{i+1/2}^n) - f(u_{i-1/2}^n))$$

c'est-à-dire la seconde étape de (II.10), avec f au lieu de $f_{i+1/2}^n$, on obtient exactement le schéma de LAX WENDROFF à deux pas, constitué de (6) et (9). Or ce schéma ne donne pas toujours la solution faible satisfaisant à la condition d'entropie. Ainsi, sur un exemple déjà proposé au chapitre I avec

$$f(u) = \frac{1}{2} (u^2 - 1)u \quad ; \quad u_0(x) = \text{sg}(x)$$

et en perturbant légèrement la condition initiale (numérique), pour avoir

$$u_i^0 = \begin{cases} 1 & \text{si } i \geq 0 \\ -1 & \text{si } i < 0 \end{cases}$$

on obtient, à partir de (6)

$$u_{i+1/2}^{1/2} \in \{-1, 0, 1\}$$

et donc dans (9), pour tout $i \in \mathbb{Z}$

$$u_i^1 = u_i^0$$

De proche en proche, on construit la solution stationnaire, qui ne satisfait pas à la condition d'entropie (I.14).

Le schéma de LAX WENDROFF offre cependant l'avantage d'être d'ordre deux en précision, puisque (9) est d'ordre deux et (6) d'ordre un. En introduisant $\xi_{i+1/2}^n \in I(u_i^n, u_{i+1}^n)$ et vérifiant

$$(10) \quad f(u_{i+1}^n) - f(u_i^n) = f'(\xi_{i+1/2}^n) (u_{i+1}^n - u_i^n)$$

on peut, en conservant une précision d'ordre deux, mettre le schéma (6),(9) sous la forme suivante

$$(11) \quad u_i^{n+1} = u_i^n - \frac{q}{2} (f(u_{i+1}^n) - f(u_{i-1}^n)) + \frac{q^2}{2} f'(\xi_{i+1/2}^n) (f(u_{i+1}^n) - f(u_i^n)) - \frac{q^2}{2} f'(\xi_{i-1/2}^n) (f(u_i^n) - f(u_{i-1}^n))$$

ce qui constitue le schéma de LAX WENDROFF à un pas. Le schéma (11) est en fait une des formes possibles du schéma de LAX WENDROFF à un pas. On peut construire d'autres formes en imposant un autre choix de $\xi_{i+1/2}^n$ que celui introduit en (10), par exemple la moyenne entre u_i^n et u_{i+1}^n . Tous ces schémas coïncident dans le cas linéaire, et sont d'ordre deux. Cette précision d'ordre deux apparaît de façon plus nette dans (11) ; le dernier terme, qui correspond à une viscosité, provient de la discrétisation de

$$\frac{\partial}{\partial x} \left[f'(u) \frac{\partial}{\partial x} (f(u)) \right]$$

qui est égal à $\frac{\partial^2 u}{\partial t^2}$. On retrouve ainsi tous les termes d'un développement de TAYLOR jusqu'à l'ordre deux. Les schémas (11) et (6),(9) sont stables en norme L^2 , dans le cas linéaire ; voir par exemple le livre de RICHTMYER et MORTON, mais ne sont stables ni dans L^∞ ni dans BV, et il est exclu d'appliquer le résultat de compacité utilisé au théorème 2.

Tous les schémas donnés jusqu'ici peuvent être mis sous la

forme

$$(12) \quad u_i^{n+1} = u_i^n - \frac{q}{2} (f(u_{i+1}^n) - f(u_{i-1}^n)) + \frac{a_{i+1/2}^n}{2} (u_{i+1}^n - u_i^n) - \frac{a_{i-1/2}^n}{2} (u_i^n - u_{i-1}^n)$$

où les coefficients $a_{i+1/2}^n$ représentent une distribution de viscosité, qui peut dépendre de façon non linéaire de la solution. Ainsi, pour le schéma de LAX, cette distribution est uniforme,

$$a_{i+1/2}^n = 1$$

et pour le schéma de GODUNOV, elle prend la forme

$$a_{i+1/2}^n = (1 - \mu) q f'(\xi_{i+3/4}^n) - \mu q f'(\xi_{i+1/4}^n)$$

où $\mu \in [0, 1]$ est défini par

$$u_{i+1/2}^n = (1 - \mu) u_i^n + \mu u_{i+1}^n,$$

et $\xi_{i+1/4}^n, \xi_{i+3/4}^n$ sont les quantités introduites en (II.22). Pour le schéma de LAX WENDROFF, on a

$$a_{i+1/2}^n = q^2 (f'(\xi_{i+1/2}^n))^2$$

On vérifie ainsi que, du point de vue de la viscosité, le schéma de GODUNOV est intermédiaire entre les schémas de LAX et de LAX WENDROFF. On peut également indiquer que dans la mesure où les problèmes de RIEMANN intermédiaires sont résolus de façon exacte, le schéma de GODUNOV est un schéma optimal parmi les schémas convergents admettant la forme (12). Ces remarques sont confirmées par l'expérience numérique.

La famille des schémas de la forme (12) a été étudiée dans LE ROUX [1], [2], ce qui aboutit au théorème suivant, dont la démonstration est analogue à celle du théorème 2.

Théorème 3 : Si la condition de stabilité de COURANT FRIEDRICHS LEWY (II.12) est satisfaite, et si pour tout $h > 0$ le choix des coefficients $a_{i+1/2}^n$ est tel que

$$(13) \quad \forall i \in \mathbb{Z} \quad \forall n \leq N \quad q |f'(\xi_{i+1/2}^n)| \leq a_{i+1/2}^n \leq 1,$$

alors la famille de solutions approchées $\{u_n\}_{n>0}$ contient une suite convergente dans $L_{loc}^1(\mathbb{R} \times]0, T[)$ vers $u \in L^\infty(\mathbb{R} \times]0, T[) \cap B_{loc}(\mathbb{R} \times]0, T[)$, solution faible de (II.3) et vérifiant une condition initiale u_0 donnée dans $L^\infty(\mathbb{R}) \cap B_{loc}(\mathbb{R})$.

Ce théorème est démontré dans LE ROUX [1] ou [2]. Lorsque (13) est satisfait avec $a_{i+1/2}^n = q |f'(\xi_{i+1/2}^n)|$, on retrouve le schéma (II.18), et qui donne quelquefois une solution faible ne satisfaisant pas à la condition d'entropie. On peut également remarquer que le coefficient $a_{i+1/2}^n$ associé au schéma de GODUNOV satisfait à (13).

Pour assurer la convergence vers la solution faible qui vérifie (I.25), on est amené à restreindre le choix des $a_{i+1/2}^n$ dans (13). Dans LE ROUX [1] ou [2], on montre le résultat suivant, en posant pour tout $i \in \mathbb{Z}$ et $n \in \mathbb{N}$

$$(14) \quad \gamma_{i+1/2}^n \begin{cases} q |f'(\xi_{i+1/2}^n)| = q |f'(u_i^n)| & \text{si } u_i^n = u_{i+1}^n, \\ \sup_{k \in I(u_{i+1}^n, u_i^n)} \left\{ q \frac{f(u_{i+1}^n) - f(k)}{u_{i+1}^n - k}, -q \frac{f(u_i^n) - f(k)}{u_i^n - k} \right\} & \text{sinon.} \end{cases}$$

Théorème 4 : Sous les hypothèses du théorème 3, et si

$$(15) \quad \forall i \in \mathbb{Z}, \quad \forall n < N \quad \gamma_{i+1/2}^n \leq a_{i+1/2}^n \leq 1,$$

alors toute la famille $\{u_n\}_{n>0}$ converge vers u solution faible vérifiant (I.25).

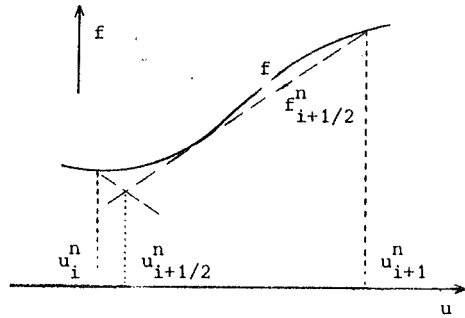
La condition (15) est plus restrictive que (14), et le coefficient associé au schéma de GODUNOV ne vérifie jamais (15) ; en effet, on a toujours dans ce cas

$$a_{i+1/2}^n \leq \gamma_{i+1/2}^n,$$

ce qui montre que le résultat énoncé au théorème 4 n'est pas optimal. Du point de vue numérique, les résultats obtenus à partir du schéma de GODUNOV d'une part, et du schéma correspondant à $a_{i+1/2}^n = \gamma_{i+1/2}^n$ (c'est-à-dire celui pour lequel il y a le moins possible de diffusion, tout en respectant (15)) sont sensiblement les mêmes. Le schéma de LAX correspond à $a_{i+1/2}^n = 1$, et on obtient ainsi la convergence. Le schéma pour lequel $a_{i+1/2}^n = \gamma_{i+1/2}^n$ correspond à une construction analogue à celle proposée pour le schéma de LAX, où $f_{i+1/2}^n$ est constitué de deux segments rectilignes de pentes respectives

$$-\frac{1}{q} a_{i+1/2}^n \text{ et } +\frac{1}{q} a_{i+1/2}^n, \text{ issus des points } (u_i^n, f(u_i^n)) \text{ et } (u_{i+1}^n, f(u_{i+1}^n)).$$

Ce schéma coïncide avec le schéma décentré dans le cas linéaire, et, quand f est monotone, dans certains autres cas, par exemple lorsque f est concave et u_0 croissante. (c'est-à-dire quand il y a choc), ou bien lorsque f est convexe, avec u_0 décroissante.



2. APPROXIMATION POUR LE PROBLEME A PLUSIEURS DIMENSIONS

On se restreint au problème suivant, sur $\mathbb{R}^p \times]0, T[$

$$(16) \quad \begin{cases} \frac{\partial u}{\partial t} + \sum_{j=1}^p \frac{\partial}{\partial x_j} [f_j(u)] = 0 & (x, t) \in \mathbb{R}^p \times]0, T[\\ u(x, 0) = u_0(x) & x \in \mathbb{R}^p \end{cases}$$

sachant que les résultats vont se généraliser au problème associé à l'équation (I.1), à condition de résoudre quelques difficultés techniques, du même type que celles rencontrées au cours de la démonstration du théorème 2.

Le principe général pour discrétiser le problème (16) consiste à mettre l'équation sous la forme

$$(17) \quad \frac{1}{p} \sum_{j=1}^p \left(\frac{\partial u}{\partial t} + p \frac{\partial}{\partial x_j} \{f_j(u)\} \right) = 0$$

puis à discrétiser chaque terme de cette somme en utilisant un schéma monodimensionnel.

Certaines propriétés du schéma monodimensionnel utilisé restent vraies sur le schéma à plusieurs dimensions ainsi obtenu. Ainsi, on

démontre dans LE ROUX [1], que si le schéma monodimensionnel conserve la borne uniforme et la variation bornée en temps, et vérifie l'estimation en k , (II.16), il en est de même pour le schéma à plusieurs dimensions, à condition toutefois de satisfaire à une condition de stabilité. Si le pas de discrétisation en espace, dans la j -ième dimension de \mathbb{R}^p est h_j , et si le pas de discrétisation en temps est t , on suppose constant le rapport

$$(18) \quad q_j = \frac{\Delta t}{h_j}$$

ce qui implique que tous les rapports h_j/h_k sont constants. Si la condition de stabilité (II.12) doit être satisfaite au niveau du schéma monodimensionnel pour établir la convergence, la condition de stabilité que devra vérifier le schéma à plusieurs dimensions est

$$(19) \quad \forall j \in \{1, \dots, p\} \quad q_j \sup_u |f'_j(u)| \leq \frac{1}{p}$$

ce qui est d'autant plus restrictif que p est grand. De plus, ceci n'implique pas que la variation bornée en espace soit conservée. On est amené à faire une hypothèse supplémentaire au niveau du terme de diffusion du schéma, ce qui se traduit en exigeant que les coefficients $a_{i+1/2}^n$ des schémas de la famille de type (12) satisfont à une certaine propriété de continuité en espace. Ceci a pour conséquence essentielle d'uniformiser sensiblement ces coefficients et de se rapprocher ainsi du schéma de LAX. En particulier, on est amené à introduire un terme de viscosité supplémentaire dans le schéma de GODUNOV, et donc à limiter son intérêt.

Le maillage permettant la discrétisation sur \mathbb{R}^p est construit de la façon suivante. \mathbb{R}^p est rapporté à sa base canonique notée $\{1_j\}$ pour $j = \{1, \dots, p\}$, et muni de la norme

$$|x| = \sup_{1 \leq j \leq p} |x_j|$$

Soit $j \in \{1, \dots, p\}$; la j -ième dimension de \mathbb{R}^p est découpée en une infinité d'intervalles de longueur h_j

$$I_{i_j} = \left[\left(i_j - \frac{1}{2}\right) h_j, \left(i_j + \frac{1}{2}\right) h_j \right] \quad \text{pour } i_j \in \mathbb{Z}$$

L'intervalle $]0, T[$ est recouvert par les mêmes intervalles J_n que pour le problème monodimensionnel. On note

$$\forall i \in \mathbb{Z}^p \quad (i = (i_1, \dots, i_p)) \quad I_i = \prod_{j=1}^p I_{i_j}$$

et sur I_i , u_i^0 est approché par la constante

$$(20) \quad u_i^0 = \frac{1}{h_1 \dots h_p} \int_{I_i} u_0(x) dx$$

La solution approchée est encore une fonction u_h , constante sur chaque $I_i \times J_n$, où sa valeur est notée u_i^n . Le vecteur de composantes u_1, \dots, u_p est noté h .

Le schéma de LAX a la forme suivante

$$(21) \quad u_i^{n+1} = \sum_{j=1}^p \left\{ \frac{1}{2p} (u_{i+1_j}^n + u_{i-1_j}^n) - \frac{q_j}{2} (f_j(u_{i+1_j}^n) - f_j(u_{i-1_j}^n)) \right\}$$

qui est obtenue en appliquant le principe précédent, à partir du schéma de LAX monodimensionnel.

Le coefficient de diffusion associée à (21) est uniforme, et ainsi le schéma de LAX conserve la variation bornée en espace, et donc converge dans $L_{loc}^1(\mathbb{R}^p \times]0, T[)$ vers la solution de (16) qui vérifie la définition 2.

Pour les schémas obtenus à partir de (12), on a le résultat suivant

Théorème 5 : Si la condition de stabilité de COURANT FRIEDRICHS LEWY (19) est satisfaite, et si pour tout h, i, j, n les coefficients $a_{i+1_j/2}^n$ vérifient

$$(22) \quad \forall \ell \in \{1, \dots, p\} \quad p q_\ell |f'_\ell(\xi_{i+1_j/2}^n, \ell)| \leq a_{i+1_j/2}^n \leq 1$$

et si il existe $C > 0$ telle que

$$\forall h \quad \forall i, i' \in \mathbb{Z}^p \quad \forall n < N \quad \forall j, j' \in \{1, \dots, p\}$$

$$(23) \quad |i + 1_{j/2} - (i' + 1_{j/2})| \leq \frac{1}{2} \implies |a_{i+1_j/2}^n - a_{i'+1_j/2}^n| \leq C|h|,$$

alors on peut extraire de $\{u_h\}$ une suite convergente dans $L_{loc}^1(\mathbb{R}^p \times]0, T[)$ vers une solution faible de (16).

De plus, si tout en vérifiant (22), $a_{i+1_j/2}^n$ satisfait à

$$(24) \quad a_{i+1_j/2}^n \geq \gamma_{i+1_j/2}^n = \begin{cases} p q_j |f'_j(\xi_{i+1_j/2, j}^n)| & \text{si } u_{i+1_j}^n = u_i^n \\ \sup_{k \in I(u_{i+1_j}^n, u_i^n)} \left\{ \max \left\{ p q_j \frac{f_j(u_{i+1_j}^n) - f(k)}{u_{i+1_j}^n - k}, - p q_j \frac{f_j(u_i^n) - f(k)}{u_i^n - k} \right\} \right\} & \text{sinon} \end{cases}$$

alors il y a convergence vers la solution satisfaisant à la Définition 2.

Remarquons que l'hypothèse (24) n'est pas toujours plus restrictive que l'hypothèse (22). Ces trois hypothèses (22), (23), et (24) tendent à uniformiser sensiblement la diffusion, et en particulier elles sont vérifiées pour le schéma de LAX. La convergence vers une solution faible a été démontrée par CONWAY et SMOLLER, pour le schéma de LAX. Le théorème 5 regroupe quelques résultats démontrés dans LE ROUX [1], au chapitre V.

Dans le cas particulier où toutes les fonctions f_j sont monotones, et ceci sans que la monotonie soit la même pour chaque f_j , on a également convergence pour le schéma construit à partir du schéma décentré monodimensionnel, mais cette propriété ne semble pas se généraliser au schéma de GODUNOV. On peut cependant remarquer que localement le schéma de GODUNOV se réduit presque toujours au schéma décentré, et ainsi la variation bornée se conserve sur la plupart des exemples numériques. De plus, s'il y a convergence, alors la solution obtenue à la limite satisfait à la définition 2, c'est-à-dire à la condition d'entropie. En ce qui concerne le schéma décentré, on a le résultat suivant.

Théorème 6 : On suppose que toutes les fonctions f_j sont monotones, et que la condition de stabilité de COURANT FRIEDRICHS LEWY (19) est satisfaite. Alors la famille $\{u_h^n\}_n$ des solutions approchées construites par le schéma décentré étendu à la dimension p , converge dans $L^1(\mathbb{R}^p \times]0, T[)$ vers la solution de (16) vérifiant la définition 2.

Démonstration

Il suffit de montrer que la variation bornée est conservée.

On note $u_{i+1, j/2}^n$, pour $i \in \mathbb{Z}^p$ et $n < N$ fixés, $j \in \{1, \dots, p\}$, le point de l'intervalle $I(u_{i+1, j}^n, u_i^n)$ où

$$(25) \quad \min_{k \in I(u_{i+1, j}^n, u_i^n)} \{ \text{sg}(u_{i+1, j}^n - u_i^n) f_j(k) \}$$

est réalisé. L'hypothèse de monotonie faite sur f_j assure que ce minimum est atteint en u_i^n si f_j est croissante et en $u_{i+1, j}^n$ si f_j est décroissante. Le schéma s'écrit

$$(26) \quad u_i^{n+1} = u_i^n - \sum_{j=1}^p q_j (f_j(u_{i+1, j/2}^n) - f_j(u_{i-1, j/2}^n)).$$

On a pour $\ell \in \{1, \dots, p\}$

$$(27) \quad u_i^{n+1} = u_i^{n+1} = \frac{1}{p} \left\{ u_{i+1, \ell}^n - u_i^n - p q_\ell (f_\ell(u_{i+3, \ell/2}^n) - f_\ell(u_{i+1, \ell/2}^n)) + p q_\ell (f_\ell(u_{i+1, \ell/2}^n) - f_\ell(u_{i-1, \ell/2}^n)) \right\} + \frac{1}{p} \sum_{j \neq \ell} \left\{ u_{i+1, j}^n - u_i^n - p q_j (f_j(u_{i+1, j/2}^n) - f_j(u_{i-1, j/2}^n)) + p q_j (f_j(u_{i+1, j/2}^n) - f_j(u_{i-1, j/2}^n)) \right\}$$

Le premier terme du second membre peut être traité comme dans le cas monodimensionnel. Supposons par exemple f_j croissante pour traiter un des termes restants; il vient

$$u_{i+1, \ell}^n - u_i^n - p q_j (f_j(u_{i+1, \ell}^n) - f_j(u_i^n)) + p q_j (f_j(u_{i+1, \ell-1, j}^n) - f_j(u_{i-1, j}^n)) = (u_{i+1}^n - u_i^n) (1 - p q_j f_j'(\xi_{i+1, \ell/2}^n)) + p q_j f_j'(\xi_{i-1, j/2}^n) (u_{i+1, \ell-1, j}^n - u_{i-1, j}^n)$$

où les coefficients sont non négatifs. Si maintenant f_j est décroissante on obtient que ce terme vaut

$$(u_{i+1}^n - u_i^n) (1 + p q_j f_j'(\xi_{i+1, \ell/2}^n)) - p q_j f_j'(\xi_{i+1, j/2}^n) (u_{i+1, j}^n - u_{i+1, j}^n).$$

On obtient, en passant aux valeurs absolues, dans chaque cas

$$(28) \quad \left| u_{i+1, \ell}^n - u_i^n - p q_j (f_j(u_{i+1, \ell+1, j/2}^n) - f_j(u_{i+1, j/2}^n)) + p q_j (f_j(u_{i+1, \ell-1, j/2}^n) - f_j(u_{i-1, j/2}^n)) \right| \leq |u_{i+1}^n - u_i^n| - p q_j f_j'(\xi_{i+1, \ell/2+1, j/2}^n) |u_{i+1, \ell+1, j/2}^n - u_{i+1, j/2}^n| + p q_j f_j'(\xi_{i+1, \ell/2-1, j/2}^n) |u_{i+1, \ell-1, j/2}^n - u_{i-1, j/2}^n|$$

Il reste à replacer ces inégalités dans (27), et on obtient immédiatement après sommation en $|i| \leq I$, pour un entier I fixé

$$\sum_{|i| \leq I} |u_{i+1, \ell}^{n+1} - u_i^{n+1}| \leq \sum_{|i| \leq I+1} |u_{i+1, \ell}^n - u_i^n|$$

puisque tous les termes en f_j de l'intérieur disparaissent, et ceux correspondant aux i vérifiant $|i| = I$ peuvent être majorés en utilisant la condition de stabilité.

Le reste de la démonstration est identique à celle donnée dans LE ROUX [1] à l'occasion de la démonstration du théorème 5 ; on sait en effet que le schéma décentré monodimensionnel conserve la norme dans $L^\infty(\mathbb{R} \times]0, T[)$ et la variation en temps, et vérifie une inégalité en k du type (II.16), et que ces propriétés sont maintenues lors de la généralisation du schéma à \mathbb{R}^p . Le théorème 6 est démontré.

Pour le schéma de GODUNOV, l'inégalité (28) n'est pas vraie en général. On peut avoir par exemple

$$u_{i+1, \ell}^n = u_i^n ; \quad u_{i+1, \ell-1/2}^n = u_{i-1, j/2}^n$$

et $f_j(u_{i+1, \ell+1/2}^n) \neq f_j(u_{i+1, j/2}^n)$, avec $f_j'(u_{i+1, \ell/2+1, j/2}^n) > 0$.

Ainsi, en dimension deux, on peut avoir

$$u_{i+1, j}^n = u_{ij}^n = u_{i+1, j-1}^n = u_{i, j-1}^n = 0 ; \quad u_{i, j+1}^n = 1 ;$$

$$u_{i+1, j+1}^n = -\frac{1}{2}$$

alors pour $f(u) = u^2/2$, $u_{i, j+1/2}^n = 0$ et $u_{i+1, j+1/2}^n = -1/2$, ce qui contredit (28).

Cependant, on peut utiliser le résultat précédent pour traiter le problème à p dimensions lorsqu'il n'y a pas d'hypothèse de monotonie sur les f_j . D'après une remarque de L. TARTAR, on peut chercher p réels C_j , les plus petits possible, vérifiant

$$\forall j \in \{1, \dots, p\} \quad |u| \leq |u_0|_{L^\infty(\mathbb{R}^p)} \implies C_j + f_j'(u) \geq 0$$

puis introduire le changement de variable

$$s = t ; \quad y_j = x_j + C_j t ; \quad Y = (y_1, \dots, y_p)$$

L'équation (16) prend alors la forme avec $v(y, s) = u(x, t)$

$$\frac{\partial v}{\partial s} + \sum_{j=1}^p \frac{\partial}{\partial y_j} [C_j v + f_j(v)] = 0 ,$$

où les fonctions de v , $C_j v + f_j(v)$ sont toutes non décroissantes. On peut à partir du théorème 6 établir la convergence vers une solution v vérifiant (II.16) en les variables de (y, s) , puis par un changement de variable pour revenir aux variables (x, t) , et en remarquant que u est solution faible pour traiter les cas où $k \geq |u_0|_{L^\infty(\mathbb{R}^p)}$, obtenir (II.16) en $u(x, t)$.

Cette remarque peut être étendue au schéma d'ordre deux étudié au Théorème 10 ainsi qu'au problème avec conditions aux limites (Théorème 15). Il convient de remarquer que l'introduction des C_j restreint la condition de stabilité de façon importante.

On peut, en utilisant la terminologie de la mécanique des gaz, noter que ce procédé revient à appliquer une méthode utilisant une variable intermédiaire, y , entre la variable d'EULER, x , et la variable de LAGRANGE, qui suit la caractéristique.

3. UN SCHEMA DECENTRE QUASI D'ORDRE DEUX

On se restreint dans ce paragraphe, au problème monodimensionnel

$$(29) \quad \begin{cases} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0 & \text{si } (x, t) \in \mathbb{R} \times]0, T[, \\ u(x, 0) = u_0(x) & \text{si } x \in \mathbb{R} , \end{cases}$$

où la fonction f est supposée non décroissante. Les résultats obtenus seront généralisés dans le paragraphe suivant au cas d'une fonction non nécessairement monotone. On pourra aussi, compte-tenu du théorème 6, étendre les résultats au problème à plusieurs dimensions (16), où les fonctions f_j sont supposées monotones.

On sait qu'un schéma d'ordre deux, le schéma de LAX WENDROFF par exemple, peut construire une solution faible ne satisfaisant pas à la condition d'entropie. L'expérience numérique montre également qu'un tel schéma n'est pas stable dans L^∞ ou dans BV , car on voit apparaître des oscillations lorsque la solution u ou sa dérivée $\partial u / \partial x$ admet une singularité. Ces particularités limitent l'intérêt des schémas d'ordre deux, et les schémas d'ordre un, le schéma décentré par exemple, présentent évidemment l'inconvénient d'être peu précis, mais ont un bon comportement au niveau des singularités. Pour concilier les avantages des deux types de schéma, on va, en empruntant des idées à BORIS, BOOK et HAIN, ainsi qu'à VAN LEER, puis à OVADIA-RAVIART, construire un schéma qui sera d'ordre

deux le plus souvent possible, mais se réduira au schéma décentré dès que la stabilité risque d'être perturbée.

On reprend la discrétisation introduite au chapitre II.

Le schéma décentré, puisque f est supposée non décroissante, admet la forme suivante, déjà introduite en (II.11)

$$(30) \quad u_i^{n+1} = u_i^n - q(f(u_i^n) - f(u_{i-1}^n)) .$$

En reprenant (12), on peut faire apparaître un terme de viscosité, et il vient

$$(31) \quad u_i^{n+1} = u_i^n - \frac{q}{2} (f(u_{i+1}^n) - f(u_{i-1}^n)) + \frac{q}{2} (f(u_{i+1}^n) - f(u_i^n)) - \frac{q}{2} (f(u_i^n) - f(u_{i-1}^n)) .$$

Le schéma ainsi écrit peut être comparé avec le schéma de LAX WENDROFF (11) c'est-à-dire

$$(32) \quad u_i^{n+1} = u_i^n - \frac{q}{2} (f(u_{i+1}^n) - f(u_{i-1}^n)) + \frac{q}{2} [q f'(\xi_{i+1/2}^n)] (f(u_{i+1}^n) - f(u_i^n)) - \frac{q}{2} [q f'(\xi_{i-1/2}^n)] (f(u_i^n) - f(u_{i-1}^n))$$

En supposant satisfaite la condition de stabilité de COURANT FRIEDRICHS LEWY (II;12), il est immédiat que l'on passe de (31) à (32) en diminuant le coefficient de viscosité ; ceci revient en effet à remplacer 1 par $q f'(\xi_{i+1/2}^n)$. On procède ainsi à une phase dite d'antidiffusion, qui consiste à passer de (30) à (32) de la façon suivante. On remplace (30) par

$$(33) \quad u_i^{n+1} = u_i^n - q(f(u_i^n) - f(u_{i-1}^n)) - a_{i+1/2}^n + a_{i-1/2}^n ,$$

où les termes $a_{i+1/2}^n$ sont choisis de telle façon à diminuer la viscosité. Pour tout i , $a_{i+1/2}^n$ est de la forme

$$a_{i+1/2}^n = \alpha_{i+1/2}^n (u_{i+1}^n - u_i^n) ,$$

où le coefficient $\alpha_{i+1/2}^n$ est un réel borné et non négatif. Il y a identité entre (33) et (32) lorsque $a_{i+1/2}^n$ est toujours pris égal à

$$(34) \quad \tilde{a}_{i+1/2}^n = \frac{q}{2} (1 - q f'(\xi_{i+1/2}^n)) (f(u_{i+1}^n) - f(u_i^n)) ,$$

et c'est en faisant ce choix de $\tilde{a}_{i+1/2}^n$ partout que l'on effectue la phase d'antidiffusion. Le schéma (33), (34) est exactement le schéma de LAX WENDROFF, et il faut maintenant le corriger pour éviter tous les inconvénients des schémas d'ordre deux. Le principe de cette correction consiste à diminuer $|a_{i+1/2}^n|$ chaque fois que la stabilité risque d'être perturbée, en effet cette diminution correspond en fait à un accroissement de la viscosité. Ainsi la correction de $\tilde{a}_{i+1/2}^n$ définie en (34) est satisfaisante dans la mesure où la stabilité dans L^∞ et dans $B V_{loc}$ est assurée, ce qui permet d'établir la convergence vers une solution faible. On pose

$$(35) \quad a_{i+1/2}^n = \text{sg}(u_{i+1}^n - u_i^n) \text{Max} \left[0, \min \left\{ \text{sg}(u_{i+1}^n - u_i^n) \frac{u_i^n - u_{i-1}^n}{2}, |a_{i+1/2}^n| \right\} \right]$$

On voit immédiatement que $a_{i+1/2}^n$ devient nul dès que $(u_{i+1}^n - u_i^n)$ et $(u_i^n - u_{i-1}^n)$ sont de signe contraire. La correction apportée par (35) ne semble pas suffisante pour assurer la convergence vers la solution faible satisfaisant à la définition 2, et on sera amené plus tard à restreindre encore un peu $a_{i+1/2}^n$. Remarquons que lorsque la solution est régulière, on a grossièrement

$$(u_{i+1}^n - u_i^n) \approx (u_i^n - u_{i-1}^n)$$

et ainsi interviennent dans le minimum établi dans (35) sensiblement les mêmes termes, avec les coefficients $\frac{1}{2}$ pour le premier, et un coefficient inférieur à $\frac{1}{8}$ pour le second, dans la mesure où la condition de stabilité de COURANT FRIEDRICHS LEWY est satisfaite. En effet, ce coefficient est de la forme

$$\frac{1}{2} q f'(\xi_{i+1/2}^n) (1 - q f'(\xi_{i+1/2}^n)) ,$$

et reste inférieur à $1/8$ puisque $0 \leq q f'(\xi_{i+1/2}^n) \leq 1$. Il est immédiat que le minimum sera ainsi réalisé le plus souvent par le second terme,

c'est-à-dire lorsque le schéma coïncide avec le schéma de LAX WENDROFF. Pour cette raison, le schéma défini par (33), (34) et (35) est dit quasi d'ordre deux. On a le résultat suivant.

Théorème 7 : (f non décroissante)

Sous la condition de stabilité

$$(36) \quad q \sup_{|u| \leq |u_0|} \{ f'(u) \} \leq \frac{1}{2} \quad ,$$

la famille de solutions $\{u_h\}_{h>0}$ construite à partir de (33), (34), (35) contient une suite $\{u_{h_m}\}$ convergente dans $L^1_{loc}(\mathbb{R} \times]0, T[)$ vers une solution^m faible de (29).

On peut remarquer que la condition (36) est plus sévère que la condition de stabilité de COURANT FRIEDRICHS LEWY. Cette condition est aussi celle imposée par BORIS, BOOK et HAIN.

Démonstration

On démontre que le schéma conserve la borne uniforme et la variation bornée. On introduit les quantités $\lambda_{i+1/2}^n$ et $\mu_{i+1/2}^n$ telles que

$$(37) \quad \begin{cases} a_{i+1/2}^n = \lambda_{i+1/2}^n (u_i^n - u_{i-1}^n) / 2 \quad , \\ a_{i+1/2}^n = q \mu_{i+1/2}^n (1 - q f'(\xi_{i+1/2}^n)) (f(u_{i+1}^n) - f(u_i^n)) / 2 \quad . \end{cases}$$

Il est immédiat, d'après la construction de $a_{i+1/2}^n$, que

$$0 \leq \lambda_{i+1/2}^n \leq 1 \quad ; \quad 0 \leq \mu_{i+1/2}^n \leq 1.$$

En introduisant ces quantités dans (33), il vient

$$u_i^{n+1} = u_i^n \left[1 - \frac{\lambda_{i+1/2}^n}{2} - \frac{q}{2} f'(\xi_{i-1/2}^n) (2 - \mu_{i-1/2}^n (1 - q f'(\xi_{i-1/2}^n))) \right] + u_{i-1}^n \left[\frac{\lambda_{i+1/2}^n}{2} + \frac{q}{2} f'(\xi_{i-1/2}^n) (2 - \mu_{i-1/2}^n (1 - q f'(\xi_{i-1/2}^n))) \right],$$

où les coefficients de u_i^n et de u_{i-1}^n sont non négatifs d'après (36). On en déduit que

$$\inf(u_{i-1}^n, u_i^n) \leq u_i^{n+1} \leq \sup(u_{i-1}^n, u_i^n) \quad ,$$

et donc, de proche en proche

$$|u_h|_{L^\infty(\mathbb{R} \times]0, T[)} \leq |u_0|_{L^\infty(\mathbb{R})} \quad .$$

Le schéma conserve également la variation bornée en espace.

On a

$$u_{i+1}^{n+1} - u_i^{n+1} = (u_{i+1}^n - u_i^n) \left[1 - \frac{\lambda_{i+3/2}^n}{2} - \frac{q}{2} f'(\xi_{i+1/2}^n) (2 - \mu_{i+1/2}^n (1 - q f'(\xi_{i+1/2}^n))) \right] + (u_i^n - u_{i-1}^n) \left[\frac{\lambda_{i+1/2}^n}{2} + \frac{q}{2} f'(\xi_{i-1/2}^n) (2 - \mu_{i-1/2}^n (1 - q f'(\xi_{i-1/2}^n))) \right] \quad ,$$

où les coefficients sont encore non négatifs, et inférieurs à 1. On passe aux valeurs absolues, pour obtenir pour tout entier I , l'inégalité

$$(38) \quad \begin{aligned} \sum_{|i| \leq I} |u_{i+1}^{n+1} - u_i^{n+1}| &\leq \sum_{|i| \leq I+1} |u_{i+1}^n - u_i^n| \\ &\leq \sum_{|i| \leq I+n+1} |u_{i+1}^0 - u_i^0| \quad . \end{aligned}$$

Or, pour $R > 0$ fixé, et $I = [R/h]$, cette dernière quantité est bornée, u_0 étant à variation localement bornée. On a également

$$|u_i^{n+1} - u_i^n| \leq |u_i^n - u_{i-1}^n|$$

à partir de (33). Ces estimations permettent d'établir, en utilisant les mêmes arguments que dans la démonstration du théorème 2 (C,F), que $\{u_{h_m}\}$ suite extraite de $\{u_h\}$ converge dans $L^1_{loc}(\mathbb{R} \times]0, T[)$ vers u appartenant à $L^\infty(\mathbb{R} \times]0, T[) \cap BV_{loc}(\mathbb{R} \times]0, T[)$, et vérifiant la condition initiale. Il reste à prouver que u est solution faible de (29). On introduit une fonction-test $\phi \in C^2_0(\mathbb{R} \times]0, T[)$, qu'on approche par des projections de type L^2 sur chaque $I_i \times J_n$ (cf. dém. Th. 2 -E-) et il vient

$$(39) \quad \left| \sum_i \sum_n \left\{ u_i^n \frac{\phi_i^n - \phi_i^{n-1}}{qh} + f(u_i^n) \frac{\phi_{i+1}^n - \phi_i^n}{h} \right\} qh^2 \right| \leq \sum_i \sum_n |a_{i+1/2}^n| \left| \frac{\phi_{i+1}^n - \phi_i^n}{h} \right| h^2$$

après quelques intégrations par parties discrètes. Le second membre de (39) tend vers zéro ; on a en effet

$$|a_{i+1/2}^n| \leq \frac{1}{8} |u_{i+1}^n - u_i^n|$$

ce qui donne, compte-tenu de (38), qu'il existe une constante C dépendante de ϕ telle que

$$\sum_i \sum_n |a_{i+1/2}^n| \left| \frac{\phi_{i+1}^n - \phi_i^n}{h} \right| h^2 \leq \left\{ C(\phi) \frac{T}{q} \sum_{|i| \leq I+N} |u_{i+1}^0 - u_i^0| h \right\}$$

où I est la partie entière de R/h , R étant un réel tel que le support de ϕ soit inclus dans $]-R, R[\times]0, T[$. On sait que la somme en i est bornée, et ainsi le coefficient de h est borné. Il reste à faire tendre h vers zéro pour établir que u est solution faible. En effet, le premier membre de (39) tend vers le premier membre de (I-7), et le second membre de (39) tend vers zéro. La fonction ϕ étant nulle en $t = 0$, le théorème 7 est démontré.

x x
x

Pour assurer la convergence vers la solution vérifiant la définition 2, on est amené à supposer que $a_{i+1/2}^n$ n'est jamais trop grand par rapport à h. On a vu que la précision est d'ordre deux le plus souvent possible, et elle l'est effectivement lorsque la solution est régulière, et dans ce cas $a_{i+1/2}^n$ est de l'ordre de h. Pour ne pas trop perturber (35), on va imposer

$$(40) \quad |a_{i+1/2}^n| \leq C h^\alpha$$

avec $C > 0$ et $\alpha \in]0, 1[$ choisis arbitrairement. Notons que par exemple pour $\alpha = \frac{1}{6}$ et $C = 2$, la contrainte (40) n'est active que pour les valeurs

de h inférieures à 10^{-7} , dans les exemples classiques. Ceci permet, lors de la programmation du schéma, d'éviter de tester (40) sur la plupart des essais numériques.

On va également transformer un peu $a_{i+1/2}^n$ défini dans (35), en faisant intervenir f dans le premier terme du minimum. On prendra

$$(41) \quad a_{i+1/2}^n = \text{sg}(u_{i+1}^n - u_i^n) \cdot \text{Max} \left\{ 0, \min \left\{ \text{sg}(u_{i+1}^n - u_i^n) \right. \right. \\ \left. \left. q(f(u_i^n) - f(u_{i-1}^n)), \tilde{a}_{i+1/2}^n, C h^\alpha \right\} \right\}$$

Le théorème 7 s'applique encore ; on a en effet

$$a_{i+1/2}^n = v_{i+1/2}^n q(f(u_i) - f(u_{i-1}))$$

avec $v_{i+1/2}^n \in]0, 1[$, et on retrouve (37) en posant

$$\lambda_{i+1/2}^n = 2 q f'(\xi_{i-1/2}^n) v_{i+1/2}^n$$

avec encore $\lambda_{i+1/2}^n \in]0, 1[$.

On fait enfin une hypothèse supplémentaire sur u_0 . On convient de dire qu'il y a changement de monotonie chaque fois que la fonction u_0 passe par un minimum local ou un maximum local, et à condition que deux optima locaux consécutifs ne soient pas de même nature. On suppose que u_0 admet sur chaque compact de \mathbb{R} un nombre fini de changements de monotonie : c'est-à-dire que pour tout $R > 0$, il existe une suite finie (la moins nombreuse possible)

$$x_0 = -R < x_1 < x_2 < \dots < x_k = R$$

telle que u_0 soit monotone sur chaque intervalle (x_{i-1}, x_i) . Cette hypothèse n'est pas restrictive dans la pratique. On note $K(R)$ le nombre de changements de monotonie de u_0 sur l'intervalle $]-R, R[$. Il est alors évident, compte-tenu des estimations établies pour obtenir la convergence, que pour tout h et à t fixé, qu'il y a conservation de la monotonie, et

ainsi $u_h(\cdot, t)$ admet au plus $K(R + t/q)$ changements de monotonie sur $] - R, R[$. On a alors le résultat suivant.

Théorème 8 : (f non décroissante)

On suppose vérifiée la condition de stabilité (36), et que u_0 admet un nombre fini de changements de monotonie sur tout compact. Alors la famille $\{u_h\}_{h>0}$ construite à partir du schéma (33), (34), (41) converge dans $L^1_{loc}(\mathbb{R} \times]0, T[)$, lorsque h tend vers zéro, vers une solution faible de (29) vérifiant la définition 2 (c'est-à-dire la condition d'entropie).

Démonstration

On introduit $k \in \mathbb{R}$, et une fonction test ϕ non négative.

Il vient

$$u_i^{n+1} - k = u_i^n - k - q \left[f(u_i^n) - f(k) - f(u_{i-1}^n) + f(k) \right] \left[1 + v_{i+1/2}^n - \frac{\mu_{i-1/2}^n}{2} (1 - qf'(\xi_{i-1/2}^n)) \right]$$

On introduit k_i^n , comme dans la démonstration du théorème 2, tel que

$$f(u_i^n) - f(k) = f'(k_i^n) (u_i^n - k) ,$$

et il vient, pour $|k| \leq |u_0|_{L^\infty(\mathbb{R})}$

$$u_i^{n+1} - k = (u_i^n - k) \left\{ 1 - qf'(k_i^n) \left(1 + v_{i+1/2}^n - \frac{\mu_{i-1/2}^n}{2} (1 - qf'(\xi_{i-1/2}^n)) \right) \right\} + (u_{i-1}^n - k) qf'(k_{i-1}^n) \left\{ 1 + v_{i+1/2}^n - \frac{\mu_{i-1/2}^n}{2} (1 - qf'(\xi_{i-1/2}^n)) \right\} ,$$

où les coefficients respectifs de $(u_i^n - k)$ et de $(u_{i-1}^n - k)$ sont non négatifs. On passe aux valeurs absolues, pour obtenir après reconstitution de certains termes, l'inégalité

$$(42) \quad |u_i^{n+1} - k| \leq |u_i^n - k| - q \left\{ |f(u_i^n) - f(k)| - |f(u_{i-1}^n) - f(k)| \right\} \left\{ 1 + v_{i+1/2}^n - \frac{\mu_{i-1/2}^n}{2} (1 - qf'(\xi_{i-1/2}^n)) \right\} .$$

Pour $|k| > |u_0|_{L^\infty(\mathbb{R})}$, cette inégalité est obtenue directement en multipliant par $sg(u_i^{n+1} - k)$.

On introduit maintenant, pour tout $i \in \mathbb{Z}$ et $n < N$, la quantité $S_{i+1/2}^n \in [0, 1]$, telle que, lorsque $u_i^n \neq u_{i+1}^n$,

$$(43) \quad S_{i+1/2}^n = \begin{cases} 1 & \text{si } k \geq u_{i+1}^n \geq u_i^n \text{ ou } k \leq u_{i+1}^n \leq u_i^n , \\ \frac{f(k) - f(u_i^n)}{f(u_{i+1}^n) - f(u_i^n)} & \text{si } k \in I(u_i^n, u_{i+1}^n) , \\ 0 & \text{si } k \geq u_i^n \geq u_{i+1}^n \text{ ou } k \leq u_i^n \leq u_{i+1}^n . \end{cases}$$

On a alors

$$|f(u_{i+1}^n) - f(k)| - |f(u_i^n) - f(k)| = (1 - 2S_{i+1/2}^n) |f(u_{i+1}^n) - f(u_i^n)| ,$$

et

$$(44) \quad \mu_{i+1/2}^n (S_{i-1/2}^n - S_{i+1/2}^n) \geq 0 .$$

Cette inégalité est vérifiée à l'égalité lorsque $\mu_{i+1/2}^n$ est nul :

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

ou lorsque la différence $S_{i-1/2}^n - S_{i+1/2}^n$ est nulle :

$$k \notin I(u_{i-1}^n, u_i^n) \cup I(u_i^n, u_{i+1}^n) (= I(u_{i-1}^n, u_{i+1}^n)) .$$

Il reste à considérer le cas où $k \in I(u_{i-1}^n, u_i^n)$, puis le cas où $k \in I(u_i^n, u_{i+1}^n)$. Si $k \in I(u_{i-1}^n, u_i^n)$, on a $S_{i+1/2}^n = 0$, et (44) est immédiat, car $S_{i-1/2}^n \geq 0$, $\mu_{i+1/2}^n \geq 0$. Si $k \in I(u_i^n, u_{i+1}^n)$, on a $S_{i-1/2}^n = 1$, donc

$$\mu_{i+1/2}^n (S_{i-1/2}^n - S_{i+1/2}^n) = \mu_{i+1/2}^n (1 - S_{i+1/2}^n) \geq 0$$

puisque $S_{i+1/2}^n$ reste plus petit que 1. On peut remarquer que (44) est vérifiée à l'égalité sauf un nombre fini de fois, qui correspond au plus à deux fois le nombre de changements de monotonie de la condition initiale, sur tout compact, et ceci indépendamment de h.

Lorsque $u_i^n = u_{i+1}^n$, on peut prendre par exemple $S_{i+1/2}^n = \frac{1}{2}$ si $k = u_i^n$, et reprendre (43) dans les autres cas. Dans (44), le premier membre est non nul seulement lorsque k et u_i^n appartiennent à $I(u_{i-1}^n, u_{i+1}^n)$. Compte-tenu du nombre fini de changements de monotonie, ceci n'arrive qu'un nombre fini de fois, sur tout compact.

On fait apparaître dans (42) la viscosité naturelle du schéma décentré, en écrivant

$$\begin{aligned} & |f(u_i^n) - f(k)| - |f(u_{i-1}^n) - f(k)| \\ &= \frac{1}{2} \left[|f(u_{i+1}^n) - f(k)| - |f(u_{i-1}^n) - f(k)| \right] \\ &+ \frac{1}{2} \left[|f(u_i^n) - f(k)| - |f(u_{i-1}^n) - f(k)| \right] \\ &- \frac{1}{2} \left[|f(u_{i+1}^n) - f(k)| - |f(u_i^n) - f(k)| \right], \end{aligned}$$

et il vient

$$\begin{aligned} |u_i^{n+1} - k| &\leq |u_i^n - k| - \frac{q}{2} \left\{ |f(u_{i+1}^n) - f(k)| - |f(u_{i-1}^n) - f(k)| \right\} \\ &+ \frac{q}{2} |f(u_{i+1}^n) - f(u_i^n)| (1 - 2 S_{i+1/2}^n) \\ &- q \nu_{i+1/2}^n |f(u_i^n) - f(u_{i-1}^n)| (1 - 2 S_{i-1/2}^n) \\ &- \frac{q}{2} |f(u_i^n) - f(u_{i-1}^n)| (1 - 2 S_{i-1/2}^n) (1 - \mu_{i-1/2}^n (1 - q f'(\xi_{i-1/2}^n))). \end{aligned}$$

Or d'après la construction de $\nu_{i+1/2}^n$ et de $\mu_{i+1/2}^n$, on a

$$q \nu_{i+1/2}^n |f(u_i^n) - f(u_{i-1}^n)| = \frac{q}{2} |f(u_{i+1}^n) - f(u_i^n)| (1 - q f'(\xi_{i+1/2}^n)) \mu_{i+1/2}^n,$$

et il vient en posant

$$\beta_{i+1/2}^n = \frac{q}{2} f'(\xi_{i+1/2}^n) (1 - \mu_{i+1/2}^n (1 - q f'(\xi_{i+1/2}^n))) (1 - 2 S_{i+1/2}^n),$$

et qui est borné par 1, les égalités

$$\begin{aligned} & \frac{q}{2} |f(u_{i+1}^n) - f(u_i^n)| (1 - 2 S_{i+1/2}^n) - q \nu_{i+1/2}^n |f(u_i^n) - f(u_{i-1}^n)| (1 - 2 S_{i-1/2}^n) \\ &= \frac{q}{2} f'(\xi_{i+1/2}^n) (1 - 2 S_{i+1/2}^n) |u_{i+1}^n - u_i^n| \\ &- \frac{1}{2} \mu_{i+1/2}^n q f'(\xi_{i+1/2}^n) (1 - q f'(\xi_{i+1/2}^n)) (1 - 2 S_{i-1/2}^n) |u_{i+1}^n - u_i^n| \\ &= \beta_{i+1/2}^n |u_{i+1}^n - u_i^n| + \mu_{i+1/2}^n (S_{i-1/2}^n - S_{i+1/2}^n) q f'(\xi_{i+1/2}^n) (1 - q f'(\xi_{i+1/2}^n)). \end{aligned}$$

On obtient donc l'inégalité

$$\begin{aligned} |u_i^{n+1} - k| &\leq |u_i^n - k| - \frac{q}{2} |f(u_{i+1}^n) - f(k)| - |f(u_{i-1}^n) - f(k)| \\ &+ \beta_{i+1/2}^n |u_{i+1}^n - u_i^n| - \beta_{i-1/2}^n |u_i^n - u_{i-1}^n| \\ &+ q |f(u_{i+1}^n) - f(u_i^n)| \mu_{i+1/2}^n (S_{i-1/2}^n - S_{i+1/2}^n) (1 - q f'(\xi_{i+1/2}^n)). \end{aligned} \tag{45}$$

Le passage à la limite sur (45), multipliée par h et ϕ_i^n , projection de type L^2 de ϕ sur $I_i \times J_n$, puis sommée en i et n donne exactement

$$\iint_{\mathbb{R} \times]0, T[} |u-k| \frac{\partial \phi}{\partial t} + \text{sg}(u-k) (f(u) - f(k)) \frac{\partial \phi}{\partial x} \, dx dt$$

$$\geq - \lim_{h \rightarrow 0} \sum_i \sum_n |f(u_{i+1}^n) - f(u_i^n)| (1 - q f'(\xi_{i+1/2}^n)) \mu_{i+1/2}^n (S_{i-1/2}^n - S_{i+1/2}^n) \phi_i^n \leq$$

On ne pourra établir (I.25) que si le second membre de l'inégalité ci-dessus est nul. En effet, d'après (44) ce terme est toujours négatif. On peut cependant remarquer que la somme en i est réduite à un nombre fini de termes, puisque $(S_{i-1/2}^n - S_{i+1/2}^n) \mu_{i+1/2}^n$ est nul sauf lorsque

k et u_i^n appartiennent à $I(u_{i-1}^n, u_{i+1}^n)$, c'est-à-dire au plus deux fois sur chaque portion monotone de u_h , soit au plus

$$2 K(R + T/q)$$

fois, si R est tel que le support de ϕ soit inclus dans $]-R, R[\times]0, T[$. De plus, dans le cas où il est non nul, on a

$$|f(u_{i+1}^n) - f(u_i^n)| (1 - qf'(\xi_{i+1/2}^n)) u_{i+1/2}^n (s_{i-1/2}^n - s_{i+1/2}^n) \leq \frac{2}{q} |a_{i+1/2}^n|$$

$$\leq 2 \frac{C}{q} h^\alpha$$

et ainsi le second membre de (46) peut être minoré par

$$- T 2K(R + T/q) \|\phi\|_{L^\infty} \frac{2C}{q} h^\alpha$$

qui tend vers zéro lorsque h tend vers zéro. Le théorème 8 est démontré.

x x
x

En notant i_n la valeur de i pour laquelle

$k \in I(u_{i_n}^n, u_{i_n+1}^n) - \{u_{i_n+1}^n\}$, on peut obtenir que le second membre de (46) soit minoré par

$$(47) \quad - 2 K(R + T/q) \|\phi\|_{L^\infty} \sum_n |u_{i_n+1}^{n+1} - u_{i_n}^n| h$$

qui fait intervenir la variation de u_h le long d'une "caractéristique approchée", c'est-à-dire d'une courbe de la forme $x = x_h(t)$, avec x_h monotone non décroissante et bornée, le long de laquelle u_h est proche de k . On sait que sur la caractéristique $u = k$, u est de variation nulle, et il serait donc normal que le terme décrit en (47) tende vers zéro, et ceci d'autant plus qu'une variation bornée suffit. On a même, d'après le théorème de FRECHET KOLMOGOROV (Cf. YOSIDA), que ce terme (47) tend vers zéro dès que la fonction étagée construite à partir des $u_{i_n}^n$ converge dans $L^1(0, T)$.

4. GENERALISATION AU CAS f NON NECESSAIREMENT MONOTONE

On ne fait plus l'hypothèse de monotonie sur f ; le schéma décentré doit être remplacé par le schéma de GODUNOV étudié au chapitre II. On procède comme précédemment à une phase d'antidiffusion, puis à une phase de correction. On obtient ainsi, dans le cadre du problème (29), le schéma suivant

$$(48.i) \quad u_{i+1/2}^n \in I(u_i^n, u_{i+1}^n) \quad \text{réalise} \quad \min_{k \in I(u_{i+1}^n, u_i^n)} \{sg(u_{i+1}^n - u_i^n) f(k)\}$$

$$(48.ii) \quad \tilde{a}_{i+1/2}^n = \frac{q}{2} \left[f(u_{i+1}^n) - 2f(u_{i+1/2}^n) + f(u_i^n) - qf'(\xi_{i+1/2}^n) (f(u_{i+1}^n) - f(u_i^n)) \right]$$

qui est de même signe que $(u_{i+1}^n - u_i^n)$ si la condition de stabilité est vérifiée; il s'agit du terme correcteur permettant de passer du schéma de GODUNOV au schéma de LAX WENDROFF. On corrige ce terme de la façon suivante, où C et $\alpha \in]0, 1[$ sont choisis arbitrairement

$$(48.iii) \quad \left. \begin{aligned} a_{i+1/2}^n &= sg(u_{i+1}^n - u_i^n) \max \left\{ 0, \min \left[|a_{i+1/2}^n|, C h^\alpha \right] \right\} \\ &sg(u_{i+2}^n - u_{i+1}^n) sg(u_{i+1}^n - u_i^n) q |f(u_{i+3/2}^n) - f(u_{i+1/2}^n)| \\ &sg(u_i^n - u_{i-1}^n) sg(u_{i+1}^n - u_i^n) q |f(u_{i+1/2}^n) - f(u_{i-1/2}^n)| \end{aligned} \right\}$$

et enfin,

$$(48.iv) \quad u_i^{n+1} = u_i^n - q (f(u_{i+1/2}^n) - f(u_{i-1/2}^n)) - a_{i+1/2}^n + a_{i-1/2}^n$$

On peut remarquer que lorsque f est non décroissante, on a toujours

$$f(u_{i+1/2}^n) = f(u_i^n)$$

et dans le calcul du minimum dans (48.iii), les deux termes suivants interviennent

$$|\hat{a}_{i+1/2}^n| \leq \frac{q}{2} |f(u_{i+1}^n) - f(u_i^n)| (1 - qf'(\xi_{i+1/2}^n))$$

et, si $sg(u_{i+2}^n - u_{i+1}^n)$ et $sg(u_{i+1}^n - u_i^n)$ sont égaux,

$$sg(u_{i+2}^n - u_{i+1}^n) sg(u_{i+1}^n - u_i^n) q |f(u_{i+3/2}^n) - f(u_{i+1/2}^n)| = q |f(u_{i+1}^n) - f(u_i^n)| ;$$

il est immédiat que le premier est toujours inférieur au second, et ainsi on retrouve le schéma (33), (34), (41). On obtient un résultat analogue dans le cas d'une fonction f non croissante, le terme en $sg(u_{i+1}^n - u_i^n) q |f(u_{i+1/2}^n) - f(u_{i-1/2}^n)|$ n'étant pas actif ⁽¹⁾ dans le calcul du minimum, et on retrouve également le schéma (33), (34), (41). On aurait pu également remplacer (48.iii) par une généralisation de (35), de façon à retrouver un résultat proche de celui énoncé au théorème 7, mais le fait de remplacer $qf(u_{i+1/2}^n)$ par $\frac{1}{2} u_{i+1/2}^n$ ne permet plus d'utiliser (48.i), ce qui compromet la stabilité. On peut cependant généraliser le théorème 8 ; c'est l'objet du théorème suivant

Théorème 9 : Sous la condition de stabilité (36), et si la condition initiale admet sur tout compact un nombre fini de changements de monotonie, la famille $\{u_n\}_{n>0}$ construite à partir du schéma (48) converge dans $L_{loc}^1(\mathbb{R} \times]0, T[)$, lorsque h tend vers zéro, vers une solution du problème (29), au sens de la définition 2 (c'est-à-dire vérifiant la condition d'entropie).

Démonstration

On introduit comme en (37) les quantités $\lambda_{i+1/2}^n, \mu_{i+1/2}^n$ et $\nu_{i+1/2}^n$ telles que

$$(49) \begin{cases} a_{i+1/2}^n = \lambda_{i+1/2}^n q (f(u_{i+3/2}^n) - f(u_{i+1/2}^n)), \\ a_{i+1/2}^n = \mu_{i+1/2}^n \frac{q}{2} [f(u_{i+1}^n) - 2f(u_{i+1/2}^n) + f(u_i^n) - qf'(\xi_{i+1/2}^n) (f(u_{i+1}^n) - f(u_i^n))], \\ a_{i+1/2}^n = \nu_{i+1/2}^n q (f(u_{i+1/2}^n) - f(u_{i-1/2}^n)). \end{cases}$$

⁽¹⁾ à condition que $sg(u_{i+1}^n - u_i^n) sg(u_i^n - u_{i-1}^n)$ soit égal à 1.

On a alors

$$0 \leq |\lambda_{i+1/2}^n| \leq 1 ; 0 \leq |\nu_{i+1/2}^n| \leq 1 ; 0 \leq |\mu_{i+1/2}^n| \leq 1,$$

et pour i fixé, $\lambda_{i-1/2}^n$ et $\nu_{i+1/2}^n$ ne peuvent pas être de signe contraire.

De plus, comme les coefficients dans $a_{i+1/2}^n$ s'annulent lors d'un changement de monotonie, on a toujours

$$0 \leq 1 + \nu_{i+1/2}^n - \lambda_{i-1/2}^n \leq 2.$$

D'après (48) et (49), le schéma peut être mis sous la forme

$$u_i^{n+1} = u_i^n - q (f(u_{i+1/2}^n) - f(u_{i-1/2}^n)) (1 + \nu_{i+1/2}^n - \lambda_{i-1/2}^n).$$

La démonstration de la stabilité se réduit à considérer deux cas essentiellement, parmi tous les cas considérés lors de la démonstration du théorème 2.

Si $u_{i-1/2}^n \leq u_i^n \leq u_{i+1/2}^n$, on a de façon analogue à (II.24)

$$(50) \begin{aligned} u_i^{n+1} = & u_i^n - q f'(\xi_{i+1/4}^n) (u_{i+1/2}^n - u_i^n) (1 + \nu_{i+1/2}^n - \lambda_{i-1/2}^n) \\ & - q f'(\xi_{i-1/4}^n) (u_i^n - u_{i-1/2}^n) (1 + \nu_{i+1/2}^n - \lambda_{i-1/2}^n). \end{aligned}$$

On a, d'après (II.22)

$$\begin{aligned} -1 \leq & (1 + \nu_{i+1/2}^n - \lambda_{i-1/2}^n) q f'(\xi_{i+1/4}^n) \leq 0, \\ 0 \leq & (1 + \nu_{i+1/2}^n - \lambda_{i-1/2}^n) q f'(\xi_{i-1/4}^n) \leq 1, \end{aligned}$$

et ainsi les coefficients de $u_{i+1/2}^n$ et de $u_{i-1/2}^n$ dans (50) sont non négatifs. On majore $u_{i-1/2}^n$ par u_i^n ; il vient, après regroupement, comme coefficient de u_i^n un terme non négatif. On majore alors u_i^n par $u_{i+1/2}^n$, et on obtient

$$u_i^{n+1} \leq u_{i+1/2}^n.$$

On procède de même pour effectuer la minoration ; on minore $u_{i+1/2}^n$ par u_i^n , on regroupe le coefficient de u_i^n , qui devient non négatif et enfin on minore u_i^n par $u_{i-1/2}^n$, et il vient

$$u_i^{n+1} \leq u_{i-1/2}^n$$

Le cas où $u_{i-1/2}^n \geq u_i^n \geq u_{i+1/2}^n$ se traite de façon analogue.

Si par exemple, $u_{i-1/2}^n \leq u_{i+1/2}^n \leq u_i^n$, alors $a_{i+1/2}^n$ est nul mais $a_{i-1/2}^n$ peut être non nul, et l'inverse pourra se produire dans un autre cas. On est ainsi amené à procéder comme si les $a_{i+1/2}^n$ étaient non nuls. On reprend (50) pour avoir comme dans le cas N°2 de la démonstration du théorème 2,

$$u_i^{n+1} \leq u_i^n + O(u_{i+1/2}^n - u_i^n) + O(u_i^n - u_{i-1/2}^n)$$

Pour minorer u_i^{n+1} , on commence par minorer u_i^n par $u_{i+1/2}^n$, ce qui donne

$$u_i^{n+1} \geq u_{i+1/2}^n - q(1 + v_{i+1/2}^n - \lambda_{i-1/2}^n)(f(u_{i+1/2}^n) - f(u_{i-1/2}^n)),$$

et d'après (36), la fonction réelle définie par

$$u \rightarrow u - q(1 + v_{i+1/2}^n - \lambda_{i-1/2}^n) f(u)$$

est non décroissante. On peut minorer $u_{i+1/2}^n$ par $u_{i-1/2}^n$ et il vient

$$u_i^{n+1} \geq u_{i-1/2}^n$$

Tous les autres cas où $u_i^n \notin I(u_{i+1/2}^n, u_{i-1/2}^n)$ se démontrent de façon analogue à celui-ci, et on obtient dans chaque cas

$$\inf(u_{i-1/2}^n, u_i^n, u_{i+1/2}^n) \leq u_i^{n+1} \leq \sup(u_{i-1/2}^n, u_i^n, u_{i+1/2}^n)$$

Pour établir la variation bornée, on écrit comme en (II.29),

$$\begin{aligned} u_{i+1}^{n+1} - u_i^{n+1} &= u_{i+1}^n - u_{i+1/2}^n + u_{i+1/2}^n - u_i^n \\ &- q f'(\xi_{i+5/4}^n) (1 + v_{i+3/2}^n - \lambda_{i+1/2}^n) (u_{i+3/2}^n - u_i^n) \\ &- q f'(\xi_{i+3/4}^n) (1 + v_{i+3/2}^n - \lambda_{i+1/2}^n) (u_{i+1}^n - u_{i+1/2}^n) \\ &+ q f'(\xi_{i+1/4}^n) (1 + v_{i+1/2}^n - \lambda_{i-1/2}^n) (u_{i+1/2}^n - u_i^n) \\ &- q f'(\xi_{i-1/4}^n) (1 + v_{i+1/2}^n - \lambda_{i-1/2}^n) (u_i^n - u_{i-1/2}^n) \end{aligned}$$

On regroupe les coefficients respectifs de $(u_{i+1}^n - u_{i+1/2}^n)$ et de $(u_{i+1/2}^n - u_i^n)$ et on obtient que tous les coefficients sont non négatifs, d'après (36) et (48.i). On passe aux valeurs absolues, et on somme en $|i| \leq I$, pour avoir

$$\sum_{|i| \leq I} |u_{i+1}^{n+1} - u_i^{n+1}| \leq \sum_{|i| \leq I+1} (|u_{i+1}^n - u_{i+1/2}^n| + |u_{i+1/2}^n - u_i^n|),$$

les termes en f' disparaissant deux par deux, sauf ceux des extrémités, que l'on majore. On conclut en remarquant que

$$|u_{i+1}^n - u_{i+1/2}^n| + |u_{i+1/2}^n - u_i^n| = |u_{i+1}^n - u_i^n|.$$

L'estimation de la variation bornée en temps est immédiate.

On a en effet

$$|u_i^{n+1} - u_i^n| \leq 2q(|f(u_{i+1/2}^n) - f(u_i^n)| + |f(u_i^n) - f(u_{i-1/2}^n)|),$$

et il reste à sommer en i puis utiliser (36) pour conclure.

On peut maintenant reprendre les arguments développés dans la partie C de la démonstration du théorème 2 et établir que l'on peut extraire une suite u_m^n convergente dans $L_{loc}^1(\mathbb{R} \times]0, T[)$ vers

$$u \in L^\infty(\mathbb{R} \times]0, T[) \cap B V_{loc}(\mathbb{R} \times]0, T[).$$

De plus, comme au théorème 7, on peut établir que u est solution faible

de (29), et en particulier satisfait à la condition initiale. Il reste à prouver une estimation en k analogue à (II.39).

Si $u_{i-1/2}^n < k < u_i^n \leq u_{i+1/2}^n$, on se trouve dans le cas N° 3 décrit dans la partie D de la démonstration du théorème 2, où f doit être remplacée par $(1 + v_{i+1/2}^n - \lambda_{i-1/2}^n) f$. On a, en reprenant les mêmes notations,

$$\begin{aligned} u_i^{n+1} - k &= (u_i^n - k) (1 - q f'(k_i^n) (1 + v_{i+1/2}^n - \lambda_{i-1/2}^n)) \\ &\quad - (u_{i+1/2}^n - u_i^n) (-q f'(\xi_{i+1/4}^n) (1 + v_{i+1/2}^n - \lambda_{i-1/2}^n)) \\ &\quad + (u_{i-1/2}^n - k) (q f'(k_{i-1/2}^n) (1 + v_{i+1/2}^n - \lambda_{i-1/2}^n)), \end{aligned}$$

où les coefficients sont non négatifs ; on passe aux valeurs absolues, et on regroupe, pour obtenir

$$\begin{aligned} |u_i^{n+1} - k| &\leq |u_i^n - k| - q \operatorname{sg}(u_i^n - k) (f(u_i^n) - f(k)) \left[1 + v_{i+1/2}^n - \lambda_{i-1/2}^n \right] \\ &\quad - q \operatorname{sg}(u_{i+1/2}^n - u_i^n) (f(u_{i+1/2}^n) - f(u_i^n)) (1 + v_{i+1/2}^n - \lambda_{i-1/2}^n) \\ &\quad + q \operatorname{sg}(u_{i-1/2}^n - k) (f(u_{i-1/2}^n) - f(k)) (1 + v_{i+1/2}^n - \lambda_{i-1/2}^n), \end{aligned}$$

ce qui donne, en utilisant (II.43)

$$\begin{aligned} |u_i^{n+1} - k| &\leq |u_i^n - k| - q \operatorname{sg}(u_{i+1/2}^n - k) (f(u_{i+1/2}^n) - f(k)) \\ &\quad + q \operatorname{sg}(u_{i-1/2}^n - k) (f(u_{i-1/2}^n) - f(k)) \\ (51) \quad &- q v_{i+1/2}^n \left[\operatorname{sg}(u_{i+1/2}^n - k) (f(u_{i+1/2}^n) - f(k)) - \operatorname{sg}(u_{i-1/2}^n - k) (f(u_{i-1/2}^n) - f(k)) \right] \\ &+ q \lambda_{i-1/2}^n \left[\operatorname{sg}(u_{i+1/2}^n - k) (f(u_{i+1/2}^n) - f(k)) - \operatorname{sg}(u_{i-1/2}^n - k) (f(u_{i-1/2}^n) - f(k)) \right]. \end{aligned}$$

L'inégalité (51) peut être établie de la même façon dans tous les autres cas décrits lors du théorème 2.

On introduit une fonction test ϕ positive, on multiplie par $h \phi_i^n$ comme précédemment, et on somme en i et n . Il vient

$$\begin{aligned} \sum_i \sum_n |u_i^n - k| \frac{\phi_i^n - \phi_i^{n-1}}{qh} qh^2 + \operatorname{sg}(u_{i+1/2}^n - k) (f(u_{i+1/2}^n) - f(k)) \frac{\phi_{i+1}^n - \phi_i^n}{h} qh^2 \\ (52) \quad \geq \sum_i \sum_n qh \phi_i^n (v_{i+1/2}^n - \lambda_{i-1/2}^n) \left[\operatorname{sg}(u_{i+1/2}^n - k) (f(u_{i+1/2}^n) - f(k)) \right. \\ \left. - \operatorname{sg}(u_{i-1/2}^n - k) (f(u_{i-1/2}^n) - f(k)) \right]. \end{aligned}$$

Le passage à la limite sur le premier membre de (52) est immédiat. Pour traiter le second membre, on introduit les quantités $S_i^n \in [0, 1]$, et définies en fonction de k de la façon suivante

$$S_i^n = \begin{cases} 1 & \text{si } k > u_{i+1/2}^n > u_{i-1/2}^n \text{ ou } k < u_{i+1/2}^n < u_{i-1/2}^n \\ \frac{f(k) - f(u_{i-1/2}^n)}{f(u_{i+1/2}^n) - f(u_{i-1/2}^n)} & \text{si } k \in I(u_{i+1/2}^n, u_{i-1/2}^n), \\ 0 & \text{si } k > u_{i-1/2}^n > u_{i+1/2}^n \text{ ou } k < u_{i-1/2}^n < u_{i+1/2}^n. \end{cases}$$

Ces quantités sont analogues aux $S_{i+1/2}^n$ introduites en (43) pour le schéma décentré, au théorème 8. On les note ici S_i^n pour exprimer leur caractère "centré", qui n'est d'ailleurs effectif qu'au niveau de l'écriture. On a $S_i^n = S_{i+1/2}^n$ si f est décroissante, et $S_i^n = S_{i-1/2}^n$ si f est croissante, et ces deux possibilités sont celles qui se réalisent le plus souvent dans le cas général.

Il vient alors

$$\begin{aligned} \operatorname{sg}(u_{i+1/2}^n - k) (f(u_{i+1/2}^n) - f(k)) - \operatorname{sg}(u_{i-1/2}^n - k) (f(u_{i-1/2}^n) - f(k)) \\ = \operatorname{sg}(u_{i+1/2}^n - u_{i-1/2}^n) (f(u_{i+1/2}^n) - f(u_{i-1/2}^n)) (1 - 2S_i^n), \end{aligned}$$

et on a toujours

$$|a_{i+1/2}^n| (S_{i+1}^n - S_i^n) \leq 0.$$

Ainsi le second membre de (52) devient, d'après (49),

$$\sum_i \sum_n qh \phi_i^n \left\{ \begin{aligned} & u_{i+1/2}^n \frac{q}{2} \left[f(u_{i+1}^n) - 2f(u_{i+1/2}^n) + f(u_i^n) - qf'(\xi_{i+1/2}^n) (f(u_{i+1}^n) - f(u_i^n)) \right] \\ & \left[sg(u_{i+3/2}^n - u_{i+1/2}^n) (1-2S_{i+1}^n) + (sg(u_{i+1/2}^n - u_{i-1/2}^n) (1-2S_i^n)) \right. \\ & \left. - sg(u_{i+3/2}^n - u_{i+1/2}^n) (1-2S_{i+1}^n) \right] \\ & - u_{i-1/2}^n \frac{q}{2} \left[f(u_i^n) - 2f(u_{i-1/2}^n) + f(u_{i-1}^n) - qf'(\xi_{i-1/2}^n) (f(u_i^n) - f(u_{i-1}^n)) \right] \\ & \left. sg(u_{i+1/2}^n - u_{i-1/2}^n) (1-2S_i^n) \right\} \end{aligned} \right. \quad (53)$$

où apparaît un terme de diffusion, qui peut être traité par intégration par parties, et il reste un terme où apparaît le coefficient

$$u_{i+1/2}^n \left\{ sg(u_{i+1/2}^n - u_{i-1/2}^n) (1-2S_i^n) - sg(u_{i+3/2}^n - u_{i+1/2}^n) (1-2S_{i+1}^n) \right\}$$

Or

$$u_{i+1/2}^n sg(u_{i+1/2}^n - u_{i-1/2}^n) = u_{i+1/2}^n sg(u_{i+3/2}^n - u_{i+1/2}^n) = u_{i+1/2}^n sg(u_{i+1}^n - u_i^n)$$

par construction, et ainsi, il ne reste plus dans (53), qu'à passer à la limite sur la quantité négative suivante

$$\sum_i \sum_n 2qh \phi_i^n |a_{i+1/2}^n| (S_{i+1}^n - S_i^n)$$

qui se traite comme au théorème 8, en faisant intervenir que la somme en i se réduit à un nombre fini de termes non nuls. En effet, le schéma conserve la monotonie, puisque pour tout i, u_i^{n+1} appartient à l'intervalle $I(u_{i-1/2}^n, u_i^n) \cup I(u_i^n, u_{i+1/2}^n)$, dont l'intersection avec

$$I(u_{i-3/2}^n, u_{i-1}^n) \cup I(u_{i-1}^n, u_{i-1/2}^n) \text{ et } I(u_{i+1/2}^n, u_{i+1}^n) \cup I(u_{i+1}^n, u_{i+3/2}^n)$$

se réduit aux extrémités en cas de monotonie. On utilise ensuite le fait que chaque $|a_{i+1/2}^n|$ est majoré par Ch^α . On établit ainsi que (53) s'annule à la limite, et le théorème 9 est démontré.

x x
x

Le théorème 7 peut être généralisé à la famille de schémas dont la convergence est établie au théorème 4, et donc en particulier pour le schéma de LAX. Ce n'est pas le cas pour le théorème 8, dont la démonstration utilise le fait que le terme de viscosité du schéma considéré fait intervenir explicitement la fonction f.

5. GENERALISATION DU SCHEMA DECENTRE QUASI D'ORDRE DEUX AU PROBLEME A P DIMENSIONS

On sait que le schéma décentré peut être généralisé au problème à p dimensions (16) lorsque chaque f_j est monotone, et le résultat de convergence est énoncé dans le théorème 6. On se propose ici d'étendre le résultat de convergence du théorème 7, c'est-à-dire vers une solution faible, au problème multidimensionnel. L'extension du théorème 8 est plus délicate, dans la mesure où la conservation du nombre de changements de monotonie dans chaque direction n'est pas assurée.

Le schéma à p dimensions obtenu est le suivant

$$\forall i \in \mathbb{Z}^p \quad \forall j \in \{1, \dots, p\} \quad \forall n < N \quad (54.i) \quad u_{i+1_j/2}^n \text{ réalise } \min_{k \in I(u_{i+1_j}^n, u_i^n)} \left[sg(u_{i+1_j}^n - u_i^n) f_j(k) \right],$$

en remarquant que, f_j étant monotone, on a soit $u_{i+1_j/2}^n$ égal à u_i^n (f_j non décroissante), soit $u_{i+1_j/2}^n$ égal à $u_{i+1_j}^n$ (f_j non croissante);

$$(54.ii) \quad a_{i+1_j/2}^n = \frac{q_j}{2} (1 - q_j f_j'(\xi_{i+1_j/2}^n)) (f_j(u_{i+1_j}^n) - f_j(u_i^n))$$

si $f_j' \geq 0$;

$$a_{i+1_j/2}^n = -\frac{q_j}{2} (1 + q_j f_j'(\xi_{i+1_j/2}^n)) (f_j(u_{i+1_j}^n) - f_j(u_i^n))$$

si $f_j' \leq 0$;

ce qui correspond à (48.ii) dans le cas particulier où f_j est monotone, et exprime la différence des viscosités respectives des schémas décentrés et du schéma de LAX WENDROFF.

La phase de correction est la suivante

$$(54.iii) \quad a_{i+1,j/2}^n = \text{sg}(u_{i+1,j}^n - u_i^n) \text{Max} \left\{ 0, \min \left(|a_{i+1,j/2}^n|, \frac{1}{2p} \text{sg}(u_{i+1,j}^n - u_i^n) (u_i^n - u_{i-1,j}^n) \right) \right\} \quad \text{si } f_j' \geq 0 ;$$

$$a_{i+1,j/2}^n = \text{sg}(u_{i+1,j}^n - u_i^n) \text{Max} \left\{ 0, \min \left(|a_{i+1,j/2}^n|, \frac{1}{2p} \text{sg}(u_{i+1,j}^n - u_i^n) (u_{i+2,j}^n - u_{i+1,j}^n) \right) \right\} \quad \text{si } f_j' \leq 0 ;$$

et enfin

$$(54.iv) \quad u_i^{n+1} = u_i^n - \sum_{j=1}^n \left\{ q_j (f_j(u_{i+1,j/2}^n) - f_j(u_{i-1,j/2}^n)) + a_{i+1,j/2}^n - a_{i-1,j/2}^n \right\}.$$

Comme au théorème 5, on est également amené à faire une hypothèse de continuité sur les $a_{i+1,j/2}^n$, afin d'assurer la conservation de la variation bornée en espace. Cette hypothèse peut être de la forme

$$(55) \quad \forall i \in \mathbb{Z}^P \quad \forall \ell, j \in \{1, \dots, p\}, \quad \forall n < N \quad \ell \neq j$$

$$\left| -a_{i+1,\ell+1,j/2}^n + a_{i+1,\ell-1,j/2}^n + a_{i+1,j/2}^n - a_{i-1,j/2}^n \right| \leq C h |u_{i+1,\ell}^n - u_i^n|,$$

où C est une constante positive, indépendante de i, ℓ ou j.

L'hypothèse (55) appelle une remarque ; en effet, les coefficients $a_{i+1,j/2}^n$ ont été définis en (54) de façon monodimensionnelle. De façon à vérifier (55), on aurait pu faire intervenir plus de points, ce qui aurait eu pour conséquence d'alourdir (54) et de rendre le schéma moins souvent d'ordre deux. Or on vérifie qu'en fait les quantités $a_{i+1,j/2}^n$

sont petites, de l'ordre de h, si bien que (55), qui correspond à une différence d'ordre deux entre les $a_{i+1,j/2}^n$, sera presque toujours satisfaite. On peut même remarquer que dans cette inégalité, le premier membre est en fait de l'ordre de h^3 où la solution est régulière. Elle est ainsi largement satisfaite en pratique, et il est inutile de l'imposer au niveau de (54).

Théorème 10 : (Chaque f_j est monotone)

Si la condition (55) est satisfaite, et si la condition de stabilité

$$(56) \quad \forall j \in \{1, \dots, p\} \quad q_j \sup_{|u| \leq |u_0|} |f_j'(u)| \leq \frac{1}{2p}, \quad L^\infty(\mathbb{R}^P)$$

est vérifiée, la famille $\{u_h\}$ des solutions approchées construites à partir du schéma (54) converge dans $L^1_{loc}(\mathbb{R}^P \times]0, T[)$, vers u solution faible de (16).

x x
x

Démonstration

On se ramène au cas monodimensionnel en mettant (54.iv) sous la forme

$$(57) \quad u_i^{n+1} = \frac{1}{p} \sum_{j=1}^p \left\{ u_i^n - p q_j (f_j(u_{i+1,j/2}^n) - f_j(u_{i-1,j/2}^n)) - p a_{i+1,j/2}^n + p a_{i-1,j/2}^n \right\}.$$

Tout se passe comme si, dans chaque terme de la somme, on remplaçait f_j par $p f_j$. En particulier, chaque terme de cette somme est borné, puisque le schéma monodimensionnel conserve la borne uniforme. Comme u_i^{n+1} est la moyenne de ces p termes bornés, $|u_i^{n+1}|$ est majoré par la même borne, et on obtient

$$(58) \quad |u_i^{n+1}| \leq \sup_{i' \in \mathbb{Z}^p} |u_{i'}^{n+1}| \leq \dots \leq |u_0|_{L^\infty(\mathbb{R}^p)}$$

Pour obtenir la variation bornée en espace, on utilise la condition (55). Soit $i \in \mathbb{Z}^p$, $\ell \in \{1, \dots, p\}$; on a

$$(59) \quad u_{i+1}^{n+1} - u_i^{n+1} = \frac{1}{p} \left\{ u_{i+1_\ell}^n - u_i^n - p q_\ell (f_\ell(u_{i+1_\ell+1_{\ell/2}}^n) - f_\ell(u_{i+1_{\ell/2}}^n)) \right. \\ \left. + p q_\ell (f_\ell(u_{i+1_{\ell/2}}^n) - f_\ell(u_{i-1_{\ell/2}}^n)) \right. \\ \left. - p a_{i+1_{\ell+1_{\ell/2}}}^n + 2 p a_{i+1_{\ell/2}}^n - p a_{i-1_{\ell/2}}^n \right. \\ \left. + \sum_{j \neq \ell} \left[\frac{1}{2} (u_{i+1_\ell}^n - u_i^n) - p q_j (f_j(u_{i+1_\ell+1_{j/2}}^n) - f_j(u_{i+1_{j/2}}^n)) \right. \right. \\ \left. \left. + p q_j (f_j(u_{i+1_\ell-1_{j/2}}^n) - f_j(u_{i-1_{j/2}}^n)) \right] \right. \\ \left. + \sum_{j \neq \ell} \left[\frac{1}{2} (u_{i+1_\ell}^n - u_i^n) - p a_{i+1_\ell+1_{j/2}}^n + p a_{i+1_{j/2}}^n \right. \right. \\ \left. \left. + p a_{i+1_\ell-1_{j/2}}^n - p a_{i-1_{j/2}}^n \right] \right\}$$

et on se propose d'utiliser cette expression pour majorer $u_{i+1_\ell}^{n+1} - u_i^{n+1}$. Le second membre de (59) est composé de trois termes, dont le premier est traité comme dans le cas monodimensionnel. Le second terme correspond à un schéma décentré; en effet, si par exemple f_j est non décroissante, on a pour tout $i' \in \mathbb{Z}^p$ et tout $n < N$

$$u_{i'+1_{j/2}}^n = u_{i'}^n,$$

le second terme devient alors majoré par

$$\left| \frac{1}{2} (u_{i+1_\ell}^n - u_i^n) - p q_j (f_j(u_{i+1_\ell}^n) - f_j(u_i^n)) + p q_j (f_j(u_{i+1_\ell-1_j}^n) - f_j(u_{i-1_j}^n)) \right| \\ \leq \frac{1}{2} |u_{i+1_\ell}^n - u_i^n| (1 - 2p q_j f'_j(\xi_{i+1_{\ell/2}}^n)) + p q_j f'_j(\xi_{i-1_{j+1_{\ell/2}}^n}) |u_{i+1_{\ell-1_j}^n} - u_{i-1_j}^n| \leq$$

$$\leq \frac{1}{2} |u_{i+1_\ell}^n - u_i^n| - p q_j |f_j(u_{i+1_\ell}^n) - f_j(u_i^n)| + p q_j |f_j(u_{i-1_{j+1_\ell}}^n) - f_j(u_{i-1_j}^n)|$$

Ces derniers termes disparaissent deux par deux au cours d'une sommation en $i \in \mathbb{Z}^p$. On retrouve le même résultat lorsque f_j est décroissante; on a alors pour tout $i' \in \mathbb{Z}^p$

$$u_{i'+1_{j/2}}^n = u_{i'+1_j}^n.$$

Le troisième terme de (59) se majore en utilisant (55):

$$\left| \frac{1}{2} (u_{i+1_\ell}^n - u_i^n) + p (-a_{i+1_\ell+1_{j/2}}^n + a_{i+1_{j/2}}^n + a_{i+1_\ell-1_{j/2}}^n - a_{i-1_{j/2}}^n) \right| \\ \leq \frac{1}{2} |u_{i+1_\ell}^n - u_i^n| (1 + 2 p C h).$$

On peut remarquer qu'on dispose dans ce troisième terme, de la quantité

$$\frac{p-1}{2} (u_{i+1_\ell}^n - u_i^n),$$

qui peut être utilisée dans les premier et second termes pour réduire la condition de stabilité. Le troisième terme peut également être traité en faisant sur les $a_{i'+1_{j/2}}^n$ les hypothèses suffisantes pour avoir

$$\left| \frac{1}{2} (u_{i+1_\ell}^n - u_i^n) - p (a_{i+1_\ell+1_{j/2}}^n - a_{i+1_{j/2}}^n) + p (a_{i+1_\ell-1_{j/2}}^n - a_{i-1_{j/2}}^n) \right| \\ \leq \frac{1}{2} |u_{i+1_\ell}^n - u_i^n| - p |a_{i+1_\ell+1_{j/2}}^n - a_{i+1_{j/2}}^n| + p |a_{i+1_\ell-1_{j/2}}^n - a_{i-1_{j/2}}^n|,$$

où une sommation en i élimine les différences en $a_{i+1_{j/2}}^n$. Dans ce cas, la différence $(u_{i+1_\ell}^n - u_i^n)$ est utilisée, et la condition de stabilité (56) ne peut être améliorée.

On obtient, à partir de (59), après un passage aux valeurs absolues, et une sommation en $|i| \leq I$, pour tout entier I fixé,

$$\sum_{|i| \leq I} |u_{i+1}^{n+1} - u_i^{n+1}| \leq \frac{1}{p} \sum_{|i| \leq I+1} |u_{i+1}^n - u_i^n| \left[1 + (p-1) \left(\frac{1}{2} + \frac{1}{2} + p C h \right) \right]$$

$$\leq (1 + (p-1) C h) \sum_{|i| \leq I+1} |u_{i+1}^n - u_i^n|$$

On déduit de (60) la conservation de la variation bornée en espace, de la même façon qu'au théorème 2.

L'estimation de la variation bornée en temps est plus immédiate. On a, comme dans le cas monodimensionnel

$$|u_i^{n+1} - u_i^n| \leq \sum_{j=1}^p q_j |f_j(u_{i+1, j/2}^n) - f_j(u_{i-1, j/2}^n)| + a_{i+1, j/2}^n - a_{i-1, j/2}^n$$

$$\leq \frac{5}{4p} \sum_{j=1}^p |u_{i+1, j/2}^n - u_{i-1, j/2}^n|$$

En effet, si par exemple f_j est non décroissante, on a

$$|a_{i+1, j/2}^n| \leq \left(\frac{1}{2p} |u_i^n - u_{i-1, j}^n| \right) = \frac{1}{2p} |u_{i+1, j/2}^n - u_{i-1, j/2}^n|$$

$$|a_{i-1, j/2}^n| \leq \left(\frac{1}{4p} (1 - q f_j'(u_{i-1, j/2}^n)) |u_i^n - u_{i-1, j}^n| \right) \leq \frac{1}{4p} |u_{i+1, j/2}^n - u_{i-1, j/2}^n|$$

et on utilise la condition de stabilité pour estimer la différence en f_j . On a les mêmes inégalités lorsque f_j est décroissante.

Il vient, après sommation en $|i| < I$, sachant que les indices $i+1, j/2$ correspondent en fait à des indices entiers, i ou $i+1, j$, suivant la monotonie de f_j , l'estimation

$$\sum_{|i| < I} |u_i^{n+1} - u_i^n| \leq \frac{5}{4p} \sum_{|i| \leq I} \sum_j |u_{i+1, j}^n - u_i^n|$$

qui est borné, d'après (60).

A partir de ces trois estimations, on peut établir que la solution approchée u_h converge dans $L_{loc}^1(\mathbb{R}^p \times]0, T[)$, lorsque $|h|$ tend vers zéro, vers une solution faible de (16). On procède de la même façon qu'au théorème 6, en remarquant que la contribution des $a_{i+1, j/2}^n$ converge faiblement vers zéro. Le théorème 10 est démontré.

x x
x

Pour établir la convergence vers la solution satisfaisant à la définition 2, il faut établir une estimation en k , analogue à (45), dans chaque direction, et introduire une différence en $f_j(u_{i+1, j/2}^n)$ dans la définition de $a_{i+1, j/2}^n$, comme en (41).

On a effectivement, en posant

$$u_{i, j}^{n+1} = u_i^n - p \left\{ q_j (f_j(u_{i+1, j/2}^n) - f_j(u_{i-1, j/2}^n)) + a_{i+1, j/2}^n - a_{i-1, j/2}^n \right\}$$

l'inégalité

$$|u_{i, j}^{n+1} - k| \leq |u_i^n - k| - \frac{1}{2} p q_j |f_j(u_{i+1, j/2}^n) - f_j(k)|$$

$$+ \frac{1}{2} p q_j |f_j(u_{i-1, j/2}^n) - f_j(k)| + p \beta_{i+1, j/2}^n |u_{i+1, j}^n - u_i^n| - p \beta_{i-1, j/2}^n |u_i^n - u_{i-1, j}^n|$$

$$+ 2 p |a_{i+1, j/2}^n| (S_{i-1, j/2}^n - S_{i+1, j/2}^n)$$

en procédant comme en (45), et en généralisant de façon immédiate les notations introduites au théorème 8.

En effectuant une sommation en j , il vient

$$|u_i^{n+1} - k| \leq \frac{1}{p} \sum_{j=1}^p |u_{i, j}^n - k|$$

$$\leq |u_i^n - k| - \sum_{j=1}^p \frac{q_j}{2} \left\{ |f_j(u_{i+1, j/2}^n) - f_j(k)| - |f_j(u_{i-1, j/2}^n) - f_j(k)| \right\}$$

$$+ \sum_{j=1}^p \beta_{i+1, j/2}^n |u_{i+1, j}^n - u_i^n| - \beta_{i-1, j/2}^n |u_i^n - u_{i-1, j}^n| + 2 \sum_{j=1}^p |a_{i+1, j/2}^n| (S_{i-1, j/2}^n - S_{i+1, j/2}^n)$$

Il reste à introduire une fonction test $\phi \geq 0$, à multiplier par

$$h_1 \dots h_p \phi_1^n$$

comme au théorème 5, puis à sommer en i et n . On obtient à la limite en h :

$$(61) \quad \int_{\mathbb{R}^p} \int_0^T \left\{ |u-k| \frac{\partial \phi}{\partial t} + \sum_{j=1}^p |f_j(u) - f_j(k)| \frac{\partial \phi}{\partial x_j} \right\} dx dt \geq$$

$$- \lim_{h \rightarrow 0} \left\{ 2 \sum_{i,n} \sum_j |a_{i+1,j/2}^n| (S_{i-1,j/2}^n - S_{i+1,j/2}^n) h_1 \dots h_p \phi_1^n \right\}$$

et on est amené à montrer que ce second membre est non négatif. Or d'après (44), ce terme ne peut pas être positif ; il faut donc établir que cette limite est nulle. Si dans chaque direction 1_j , il y a un nombre fini de changements de monotonie, et si on suppose

$$|a_{i+1,j/2}^n| \leq C h^\alpha,$$

on peut procéder comme dans le cas monodimensionnel, à condition que ces hypothèses soient uniformes en i et en h . On n'est pas assuré de cette uniformité en ce qui concerne le nombre fini (localement) de changements de monotonie dans chaque direction. En effet, il ne semble pas que dans le cas général, il y ait conservation du nombre de changements de monotonie dans chaque direction de la condition initiale.

On peut cependant préciser que dans les exemples utilisés en pratique, ce nombre de changements de monotonie est fini, et alors le second membre de (61) tend vers zéro. On est ainsi assuré que la solution obtenue à la limite satisfait à la définition 2.

6. QUELQUES REMARQUES ET RESULTATS NUMERIQUES

Ce paragraphe regroupe quelques remarques provenant des expériences numériques réalisées à partir de certains schémas étudiés dans ce chapitre. Tout d'abord, la condition de COURANT FRIEDRICHS LEWY n'est pas seulement une condition de stabilité ; en ne la respectant pas, on peut

réaliser la convergence vers une solution faible ne vérifiant pas la condition d'entropie. Ainsi pour le schéma décentré appliqué à l'équation de BURGERS ($f(u) = \frac{u^2}{2}$), avec la condition initiale (non négative)

$$u_1^0 = \begin{cases} 1 & \text{si } i \leq 0 \\ 0 & \text{si } i \geq 1 \end{cases},$$

donne, pour tout $n < N$, et $q = 2$

$$u_1^{n+1} = u_{i-1}^n,$$

c'est-à-dire la solution exacte, qui vérifie la condition d'entropie, mais en l'appliquant avec la condition initiale non négative

$$u_1^0 = \begin{cases} 0 & \text{si } i \leq 0 \\ 1 & \text{si } i \geq 1 \end{cases},$$

on obtient encore

$$u_1^{n+1} = u_{i-1}^n,$$

qui ne vérifie plus la condition d'entropie.

On rencontre des situations analogues, de convergence vers une solution faible ne vérifiant pas la condition d'entropie, avec les schémas d'ordre deux classiques, même lorsque la condition de stabilité est vérifiée. Si on applique le schéma décentré avec une fonction f non monotone, on retrouve aussi ce phénomène. Enfin en appliquant la méthode Antidiffusion-Correction du paragraphe 3 à un schéma initialement non convergent vers la solution faible vérifiant la condition d'entropie, on arrive en général à trouver des contre-exemples prouvant que cette situation ne s'est pas améliorée : la solution faible peut être mauvaise. On a même la situation inverse, la phase correction ne rétablit pas toujours la convergence vers la bonne solution. Ainsi le coefficient $\frac{1}{2}$ dans (35) est essentiel d'une part pour la stabilité, mais aussi pour la convergence vers la solution faible vérifiant la condition d'entropie, ce qui se voit davantage sur les schémas proposés par BOOK, BORIS et HAIN.

Dans le cas d'une fonction f quelconque, ce schéma peut s'écrire ainsi (par exemple)

$$(62.i) \quad \hat{u}_i^n = u_i^n - q(f(u_{i+1/2}^n) - f(u_{i-1/2}^n)), \quad (\text{GODUNOV})$$

$$(62.ii) \quad a_{i+1/2}^n = \frac{q}{2} \int_{u_i^n}^{\hat{u}_{i+1}^n} f'(\xi) (1 - qf'(\xi)) d\xi,$$

$$(62.iii) \quad a_{i+1/2}^n = \text{sg}(u_{i+1}^n - u_i^n) \text{Max} \left\{ 0, \min \left\{ \text{sg}(u_{i+1}^n - u_i^n) (u_{i+2}^n - u_{i+1}^n), \right. \right. \\ \left. \left. |a_{i+1/2}^n|, \text{sg}(u_{i+1}^n - u_i^n) (u_i^n - u_{i-1}^n) \right\} \right\},$$

$$(62.iv) \quad u_i^{n+1} = u_i^n - a_{i+1/2}^n + a_{i-1/2}^n.$$

Ce schéma effectue la correction et l'antidiffusion à partir de quantités déjà évaluées par le schéma de GODUNOV, ce qui peut faciliter la programmation de ce schéma, en séparant les différentes phases sans retenir trop de mémoires.

En prenant la condition initiale

$$u_i^0 = \begin{cases} 1 & \text{si } i \geq 1, \\ -1 & \text{si } i \leq -1, \end{cases}$$

et l'équation de BURGERS, avec un rapport q assez petit ($q = \frac{1}{4}$ par exemple), on réalise la convergence vers la solution stationnaire. Ceci n'a plus lieu dès que, dans (62.iii), $(u_{i+2}^n - u_{i+1}^n)$ et $(u_i^n - u_{i-1}^n)$ sont affectés d'un coefficient $\frac{1}{2}$.

Bien qu'il en existe, on n'a pas étudié ici de version quasi d'ordre deux pour le schéma de LAX. La différence entre les viscosités respectives du schéma de LAX et du schéma de LAX WENDROFF est telle qu'en pratique le terme qui rend le schéma d'ordre deux, réalisera moins souvent le minimum au niveau de la phase de correction. Ainsi l'analogie de (41) serait, par exemple lorsque f est non décroissante

$$a_{i+1/2}^n = \text{sg}(u_{i+1}^n - u_i^n) \text{Max} \left\{ 0, \min \left[\text{sg}(u_{i+1}^n - u_i^n) q(f(u_i^n) - f(u_{i-1}^n)), \right. \right. \\ \left. \left. \left(\frac{1}{2} - \frac{q^2 f'^2(\xi_{i+1/2}^n)}{2} \right) |u_{i+1}^n - u_i^n|, \text{sg}(u_{i+1}^n - u_i^n) q(f(u_{i+2}^n) - f(u_{i+1}^n)), \text{Ch}^\alpha \right] \right\}.$$

Pour des différences en u_i^n voisines, on est amené à comparer $qf'(\xi)$ avec $\frac{1}{2} (1 - q^2 f'^2(\xi))$, et ce dernier est le plus petit seulement pour

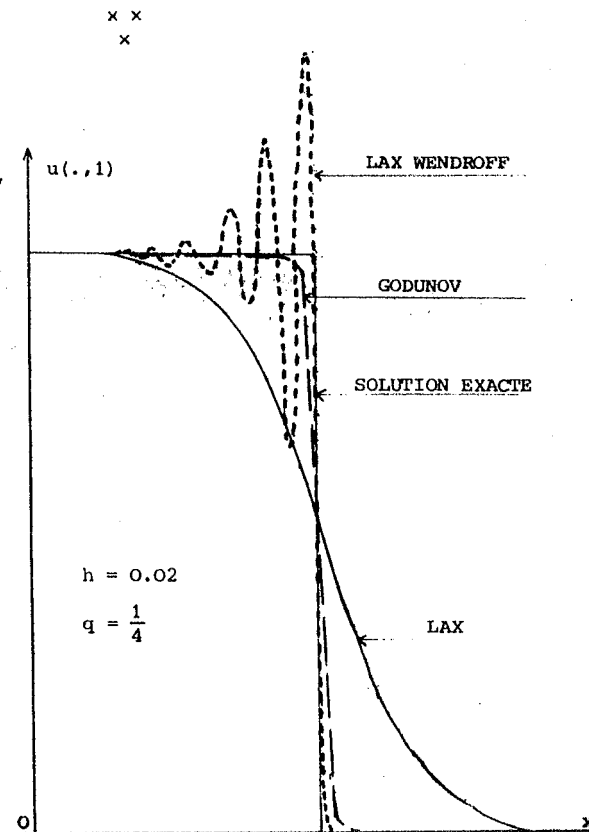
$$-1 + \sqrt{2} \leq q f'(\xi) \leq \frac{1}{2} \quad (0.414 < qf'(\xi) < 0.5).$$

Dans le cas général (f non monotone), on définit $a_{i+1/2}^n$ comme en (48.iii) avec un produit de signes.

Dans un premier exemple, les schémas de LAX, de LAX WENDROFF et de GODUNOV ont été testés dans les mêmes conditions, sur l'équation de BURGERS, avec la condition initiale

$$u(x,0) = \begin{cases} 1 & x < 0, \\ 0 & x > 0. \end{cases}$$

Le schéma de GODUNOV coïncide alors avec le schéma décentré. En comparant les termes de viscosité de chacun de ces schémas, on retient que cette viscosité est proportionnelle à un terme de la forme $\frac{1}{2q} |qf'(\xi_{i+1/2}^n)|^\alpha$



entre u_i^n et u_{i+1}^n , avec $\alpha = 0$ pour le schéma de LAX, $\alpha = 1$ pour le schéma de GODUNOV, et $\alpha = 2$ pour le schéma de LAX WENDROFF. Ainsi pour une valeur de q assez petite ($q = 1/4$ pour l'exemple numérique testé ici), la viscosité pour le schéma de LAX est importante, ce qui se traduit par un étalement important du choc. La viscosité pour le schéma de LAX WENDROFF est faible, et la solution approchée oscille au niveau des singularités. La viscosité pour le schéma de GODUNOV est indépendante de q , et le comportement au niveau d'un choc est très satisfaisant. La solution exacte pour cet exemple est donnée par $u(x,t) = u(x - \frac{t}{2}, 0)$, et est représentée sur la figure ci-dessus au temps $t = 1$.

Dans un second exemple, on compare entre eux quelques schémas quasi d'ordre deux, et les schémas de LAX et de GODUNOV sur l'équation de BURGERS, et la condition initiale, pour $a > 0$

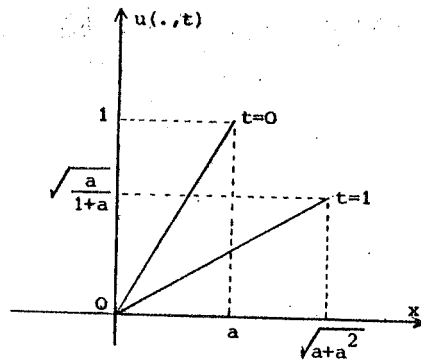
$$u(x,0) = \begin{cases} x/a & \text{si } 0 < x < a, \\ 0 & \text{sinon,} \end{cases}$$

dont la solution est donnée par

$$u(x,t) = \begin{cases} x/(t+a) & \text{si } 0 < x < \sqrt{at+a^2}, \\ 0 & \text{sinon.} \end{cases}$$

Les schémas de LAX et de GODUNOV ont été testés avec $q = 1$, de façon à avoir le meilleur résultat possible pour le schéma de LAX, et avec $q = 1/2$ pour les schémas décentrés quasi d'ordre deux, et pour le schéma proposé par BOOK, BORIS et HAIN, c'est-à-dire (63). La valeur de u au maximum est donnée par

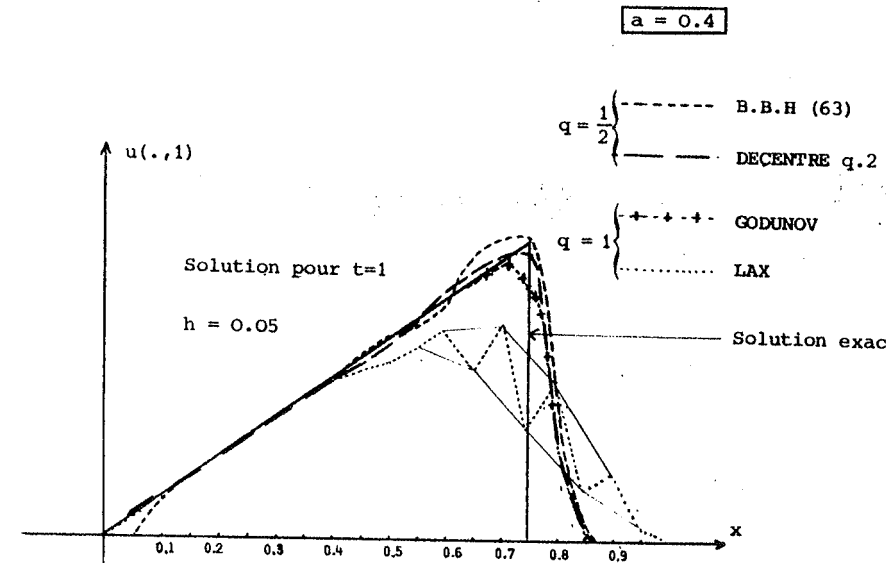
$$u = \sqrt{\frac{a}{t+a}}, \quad x = \sqrt{at + a^2}$$



et on obtient numériquement, en $t = \frac{1}{2}$ et en $t = 1$ les différents résultats

a = 0.4	t = 0.5		t = 1	
	maxi	x =	maxi	x =
solution exacte	0.6666	0.6	0.5345	0.7483
LAX	0.503	0.5	0.375*	0.7*
GODUNOV	0.600	0.55	0.491	0.7
Décentré Q ordre 2	0.612	0.6	0.514	0.75
BORIS BOOK HAIN (63)	0.652	0.6	0.546	0.75

(* voir remarque page suivante)



La différence entre les schémas de GODUNOV et les deux schémas quasi d'ordre deux n'est pas très importante, surtout si on tient compte du fait qu'il a fallu deux fois plus d'étapes de calcul en temps, compte-tenu de la condition de stabilité. L'oscillation qui apparaît pour la solution calculée par le schéma de LAX provient de la discrétisation de la condition initiale ; on a pris au niveau du choc, en $t = 0$, $u_i^0 = 1$ en $ih = 0.4$ et $u_{i+1}^0 = 0$, et non la projection de type L^2 . Comme pour calculer u_i^{n+1} on n'utilise pas u_i^n , tout se passe comme s'il y avait deux maillages et deux solutions indépendantes. Ce phénomène disparaît si la discrétisation de la condition initiale respecte (II.7). Quelques autres résultats numériques, où ces schémas interviennent seront encore détaillés au cours du chapitre IV, où les conditions aux limites interviennent.

DEUXIEME PARTIE

LE PROBLEME AUX LIMITES

Cette seconde partie traite du problème mixte, c'est-à-dire avec conditions initiales et conditions aux limites, associé à la même équation quasi-linéaire du premier ordre, que dans la première partie, sur $\Omega \times]0, T[$, où Ω est un ouvert borné de \mathbb{R}^p , dont la frontière Γ est supposée très régulière.

Cette partie est divisée en trois chapitres, numérotés IV, V et VI.

x x
x

Dans le chapitre IV, on propose une définition convenable de la solution, et l'existence et l'unicité de cette solution sont démontrées ; ce chapitre reprend les résultats de BARDOS - LE ROUX - NEDELEC.

Les deux chapitres suivants sont consacrés à l'approximation.

Le chapitre V est consacré à la convergence de quelques schémas, dont le schéma de GODUNOV et le schéma de LAX dans le cas général.

Le chapitre VI traite du problème avec hypothèse de monotonie sur la fonction f, qui permet d'utiliser quelques schémas quasi explicites.

x x
x

CHAPITRE IV

LE PROBLEME AVEC CONDITIONS AUX LIMITES

1. INTRODUCTION DES CONDITIONS AUX LIMITES

On se restreint dans un premier temps au problème monodimensionnel sur $\Omega =]0,1[$, déterminé par l'équation simple

$$(1) \quad \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0 \quad \text{si } (x,t) \in]0,1[\times]0,T[,$$

la condition initiale

$$(2) \quad u(x,0) = u_0(x) \quad \text{si } x \in]0,1[,$$

et enfin, si elles sont nécessaires, les conditions aux limites

$$(3) \quad u(0,t) = a(t) \quad \text{si } t \in]0,T[,$$

$$(4) \quad u(1,t) = b(t) \quad \text{si } t \in]0,T[.$$

On suppose encore

$$f \in C^1(\mathbb{R})$$

et, pour simplifier, car ces hypothèses sont excessives,

$$u_0 \in C^2(\bar{\Omega}) \quad , \quad a, b \in C^2([0, T])$$

Les conditions aux limites (3) et (4) ne sont pas toujours nécessaires, et la définition que l'on proposera, devra les exclure dès qu'elles seront inutiles. Ainsi, on ne peut pas envisager de généraliser simplement la notion de solution faible, en écrivant que pour tout fonction test

$$\phi \in C^1([0, 1] \times [0, T])$$

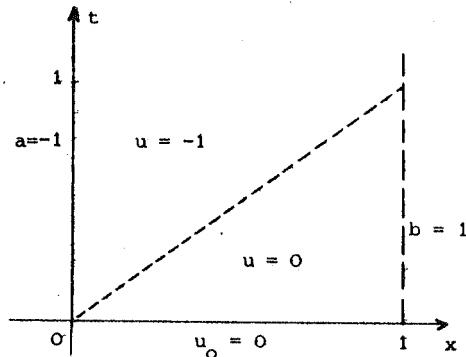
on a

$$(5) \quad \iint_{\Omega \times]0, T[} (u \frac{\partial \phi}{\partial t} + f(u) \frac{\partial \phi}{\partial x}) dx dt + \int_{\Omega} u_0(x) \phi(x, 0) dx + \int_0^T (a(t) \phi(0, t) - b(t) \phi(1, t)) dt = 0$$

car il n'y aura pas toujours existence, ni unicité. Des contre exemples d'unicités ont déjà été présentés au chapitre I, et ils se généralisent encore ici. En prenant $f(u) = u$, et les conditions initiales et aux limites

$$\begin{aligned} u_0(x) &= 0 \quad , \\ a(t) &= -1 \quad , \\ b(t) &= +1 \quad , \end{aligned}$$

on trouve en utilisant la condition de RANKINE HUGONIOT et la théorie des caractéristiques, et sachant que la condition d'entropie est toujours vérifiée dans le cas linéaire, que la solution convenable est donnée par



$$u(x, t) = \begin{cases} -1 & \text{si } x < t \quad , \\ 0 & \text{si } x > t \text{ et } t < 1 \quad , \end{cases}$$

bien que la condition limite donnée en $x = 1$ ne soit pas satisfaite. Or u ne satisfait pas à (5) si $\phi(1, t)$ est non nulle ; il ne peut donc y avoir existence.

Le fait que la condition en $x = 1$ ne soit pas vérifiée s'explique par la théorie des caractéristiques, en particulier par l'équation (I.8.) qui se réduit ici à

$$\frac{dx}{dt} = 1 \quad ,$$

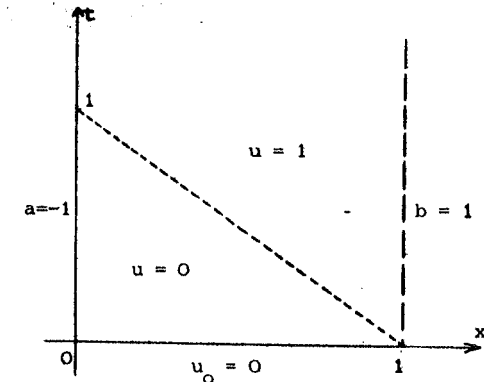
et signifie que les données (initiales et aux limites) se propagent vers les x positifs ; ces données ont ainsi tendance à sortir du domaine en $x = 1$.

Inversement, il est impossible à une donnée fixée en $x = 1$ de pénétrer à l'intérieur du domaine.

En prenant $f(u) = -u$, on change le sens de la propagation, et on obtient la solution

$$u(x, t) = \begin{cases} 0 & \text{si } 0 < x < 1-t \quad , \\ 1 & \text{si } 1-t < x < 1 \quad , \end{cases}$$

qui ne se raccorde pas avec la condition donnée en $x = 0$, et donc ne vérifie pas (5).

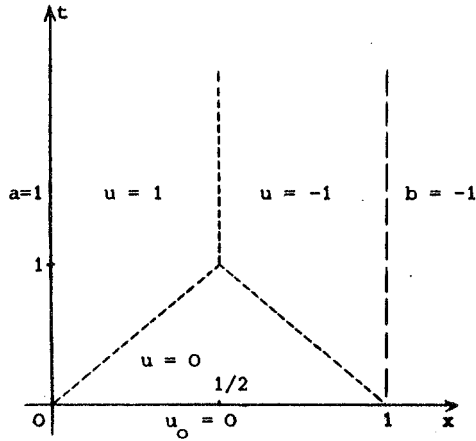


Il arrive aussi que toutes les conditions aux limites soient prises en compte. C'est par exemple le cas pour l'équation de BURGERS

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0$$

avec les conditions

$$\begin{aligned} a(t) &= 1, \\ b(t) &= -1, \\ u_0(x) &= 0. \end{aligned}$$



On obtient la solution

$$u(x,t) = \begin{cases} 1 & \text{si } x < \text{Inf} \left(\frac{t}{2}, \frac{1}{2} \right) \\ 0 & \text{si } t < 1 \text{ et } \frac{t}{2} < x < 1 - \frac{t}{2} \\ -1 & \text{si } x > \text{Sup} \left(1 - \frac{t}{2}, \frac{1}{2} \right) \end{cases}$$

qui vérifie (5). Deux discontinuités, issues de 0 et de 1 à l'instant initial, se rencontrent en $x = 1/2$ à l'instant $t = 1$, pour ne former qu'une seule discontinuité stationnaire. La condition initiale n'est pratiquement plus influente au-delà de $t = 1$, elle n'aura servi qu'à fixer la discontinuité en $x = 1/2$.

En prenant dans l'exemple précédent

$$a(t) = -1 ; \quad b(t) = 1 ;$$

aucune des conditions aux limites n'intervient, et la solution obtenue est identiquement nulle.

Ces exemples montrent que le raccordement de la solution avec les conditions aux limites est fonction du signe de f' , et des valeurs de la solution et de la condition-limite. Ainsi, il y a raccordement en $x=0$ lorsque $f'(a)$ est strictement positif, et il y a raccordement en $x = 1$ lorsque $f'(b)$ est strictement négatif, en général, mais il ne s'agit pas là d'une condition suffisante. Ceci est cependant toujours vrai lorsque

f est monotone ; si f est décroissante, la condition en $x = 0$ est inutile, et si f est croissante la condition en $x = 1$ est inutile.

Lorsque le raccordement n'est pas assuré, on peut estimer que la solution présente une discontinuité le long de la frontière du domaine $\Omega \times]0, T[$, et étant en présence d'une discontinuité, on peut soulever le problème de l'unicité, et ceci va permettre à partir d'un exemple de définir en quel sens la solution doit satisfaire à la condition-limite. Il se trouve en effet qu'il n'y a pas unicité si on ne fait qu'éliminer purement et simplement la condition-limite, et que la condition d'entropie devra être adaptée à ce genre de discontinuité. Donnons d'abord un exemple, encore avec l'équation de BURGERS, c'est-à-dire

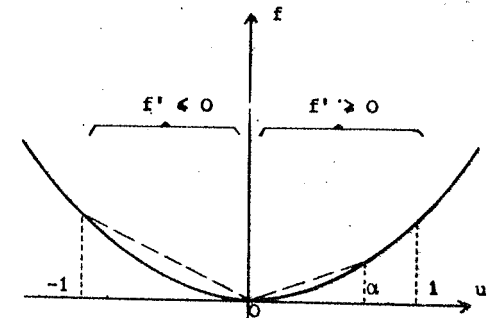
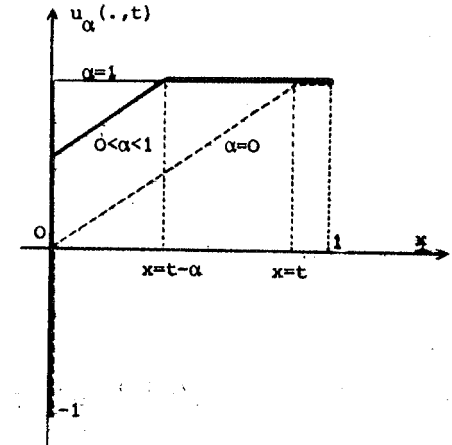
$$f(u) = u^2/2.$$

Les conditions initiales et aux limites sont les suivantes

$$\begin{aligned} a(t) &= -1, \\ b(t) &= 1, \\ u_0(x) &= 1. \end{aligned}$$

On ne peut pas trouver de solution vérifiant ces trois conditions ; en effet la condition en $x = 0$ n'est pas réalisée. On peut cependant, pour tout $\alpha \in [0, 1]$ concevoir la solution définie par

$$u_\alpha(x,t) = \begin{cases} 1 & \text{si } x > t - \alpha \\ \frac{x + \alpha}{t} & \text{si } x < t - \alpha \end{cases}$$



La solution stationnaire, qui correspond à $\alpha = 1$, doit être écartée parce que pour $u = 1$, $f'(u) = 1$ et la propagation, dont le sens est déterminé par l'équation des caractéristiques (I.8) et la condition de RANKINE HUGONIOT (I.12), se fait vers les $x > 0$. En effet, en retenant

cette solution ($\alpha = 1$), on introduit une discontinuité au niveau de $x = 0$ entre la solution et la condition-limite, qui ne satisfait pas à la condition d'entropie (I.14). Il en est de même pour toutes les solutions u_α proposées ci-avant, avec $\alpha \in]0,1[$, et il ne reste plus que la solution correspondant à $\alpha = 0$. Pour celle-ci, on a

$$(6) \quad \sup_{k \in I(u(0,t), a(t))} \left\{ \frac{f(u(0,t)) - f(k)}{u(0,t) - k} \right\} \leq 0 \quad (\text{en fait} = 0) ,$$

où $u(0,t)$ est la limite de $u(x,t)$ lorsque x tend vers zéro en décroissant.

La discontinuité qui subsiste au niveau de la limite $x = 0$ ne satisfait toujours pas à la condition d'entropie, mais d'après la condition de RANKINE HUGONIOT et la théorie des caractéristiques, la propagation de cette discontinuité a lieu vers les $x < 0$, et a ainsi tendance à sortir du domaine $\Omega \times]0,T[$. On peut remarquer que la condition-limite apparaît dans (6).

Pour n'avoir qu'une seule condition, exprimant à la fois les cas où il y a raccordement et les cas où il n'y a pas raccordement, on multiplie dans (6) par $|u(0,t) - k|$ avant de prendre la borne supérieure, ce qui ne change pas le signe. Pour le raccordement en $x = 1$, on a une condition analogue, qui exprime qu'une éventuelle discontinuité restante a une vitesse de propagation non négative, de façon à ne pas pouvoir pénétrer à l'intérieur de $\Omega \times]0,T[$.

On exprimera donc les conditions aux limites de la façon suivante, en remarquant que pour les valeurs k considérées, $(u(0,t) - k)$ et $(u(0,t) - a(t))$ sont de même signe,

$$(7) \quad \sup_{k \in I(u(0,t), a(t))} [sg(u(0,t) - a(t)) \{f(u(0,t)) - f(k)\}] = 0 \quad \text{pp }]0,T[,$$

$$(8) \quad \inf_{k \in I(u(1,t), b(t))} [sg(u(1,t) - b(t)) \{f(u(1,t)) - f(k)\}] = 0 \quad \text{pp }]0,T[,$$

c'est-à-dire presque partout sur $]0,T[$.

Dans le cas linéaire, pour $f(u) = \lambda u$, (7) et (8) deviennent

$$\sup_{k \in I(u(0,t), a(t))} \left\{ \lambda |u(0,t) - k| \right\} = 0 ,$$

$$\inf_{k \in I(u(1,t), b(t))} \left\{ \lambda |u(1,t) - k| \right\} = 0 .$$

Si λ est strictement positif, pour que le "Sup" soit nul, il faut que l'intervalle $I(u(0,t), a(t))$ se réduise à un seul point, c'est-à-dire que

$$u(0,t) = a(t) ,$$

ce qui exprime le raccordement. La borne "Inf" est réalisée pour $k = u(1,t)$, sans que le raccordement soit nécessaire. Si λ est strictement négatif, c'est l'inverse qui se produit.

Pour avoir existence et unicité d'une solution, on peut reprendre la définition 2 avec ϕ à support compact dans $\Omega \times]0,T[$, et inter-préter les conditions aux limites par (7) et (8). C'est ce qui a été fait dans LE ROUX [5]. On peut également construire une définition qui inclut dans une même formulation faible l'inégalité en k (I.25) et les conditions aux limites (7) et (8).

Dans le cas multidimensionnel, on note ν la normale unitaire extérieure à Γ , et les conditions aux limites s'écrivent, presque partout sur $\Gamma \times]0,T[$ et dans le cadre du problème (III.16)

$$(9) \quad \inf_{k \in I(\gamma u(x,t), a(x,t))} \left\{ sg(\gamma u(x,t) - a(x,t)) (f(\gamma u(x,t)) - f(k)) \cdot \nu \right\} = 0$$

où $\gamma u(x,t)$ est la trace de la solution u sur la frontière, à l'instant t et $a(x,t)$ est la condition limite. Comme dans le cas du problème de CAUCHI la solution est à variation bornée sur $\Omega \times]0,T[$, et ceci permet de définir sa trace sur $\Gamma \times]0,T[$. C'est l'objet du lemme suivant.

Lemme 3 : Soit $u \in BV(\Omega \times]0, T[$; on peut définir sa trace γu dans $L^\infty(\Omega)$ pour $t = 0$, et dans $L^\infty(\Gamma \times]0, T[$.

Cette trace γu est atteinte par une convergence en norme L^1 .

De plus, pour toute fonction $h \in C^1(\mathbb{R})$

$$\gamma(h(u)) = h(\gamma u)$$

presque partout sur $\Gamma \times]0, T[$.

x x
x

Démonstration

Comme $u \in BV(\Omega \times]0, T[$, pour presque tout $x \in \Omega$, la fonction $u(x, \cdot)$ appartient à $BV(]0, T[$ et donc admet en tout point une limite à droite et une limite à gauche. On peut en particulier passer à la limite lorsque t tend vers zéro en décroissant, et noter $\gamma u(x, 0)$ cette limite. Or $u(x, t)$ est uniformément bornée sur $\Omega \times]0, T[$, car Ω étant borné,

$$L^\infty(\Omega \times]0, T[\supset BV(\Omega \times]0, T[$$

et ainsi $u(\cdot, t)$ converge presque partout sur Ω vers $\gamma u(\cdot, 0)$, en restant borné. D'après le théorème de LEBESGUE, $\gamma u(\cdot, 0)$ est mesurable et $u(\cdot, t)$ converge vers $\gamma u(\cdot, 0)$ dans $L^1(\Omega)$. Il est immédiat qu'on a aussi $\gamma u(\cdot, 0) \in L^\infty(\Omega)$.

On procède de même sur $\Gamma \times]0, T[$. Soit $\{e_1, \dots, e_p\}$ la base canonique de \mathbb{R}^p . En presque tout point $(x, t) \in \Gamma \times]0, T[$ tel que le vecteur de base e_i ne soit pas tangent à Γ en ce point, on peut comme précédemment définir une trace $\gamma_i u(x, t)$, atteinte par une convergence dans L^1 sur un voisinage du point considéré, tel que e_i n'y soit jamais tangent à Γ . De plus $\gamma_i u(x, t)$ est mesurable et bornée sur ce voisinage de (x, t) .

Il reste à rendre cette trace indépendante de la base. On introduit $du(x, \cdot)$ la mesure de RADON associée à la distribution

$$\text{grad } u(x, \cdot) = \left(\frac{\partial u}{\partial x_1}(x, \cdot), \dots, \frac{\partial u}{\partial x_p}(x, \cdot) \right).$$

On peut en effet identifier l'espace (de BANACH) des mesures de RADON à l'espace $BV(\Omega \times]0, T[/ \mathbb{R}$ (voir par exemple YOSIDA). Il vient alors, pour toute fonction

$$\phi \in (C^1(\bar{\Omega}))^p$$

et tout vecteur de base e_i , choisis tels que e_i ne soit jamais tangent à Γ sur le support de ϕ , la formule de GREEN suivante

$$(10) \quad \int_{\Gamma} \gamma_i u \nu \cdot \phi \, ds = \int_{\Omega} u \text{div } \phi \, dx + \int_{\Omega} \phi \cdot du(x, \cdot),$$

pour presque tout $t \in]0, T[$ (cf. théorie de l'intégrale de STIELTJES).

Si un autre vecteur e_j est non tangent à Γ sur le support de ϕ , on déduit immédiatement de (10) que $\gamma_i u$ et $\gamma_j u$ coïncident presque partout. On peut donc définir une trace unique, atteinte par exemple le long de la normale par une convergence de type L^1 . Cette trace est notée γu , et appartient à $L^\infty(\Gamma \times]0, T[$. Le lemme 3 est démontré.

x x
x

On peut également envisager le problème aux limites associé à l'équation complète (I.1). La condition (9) se généralise sans problème, et l'existence d'une trace sur Ω en $t = 0$ et sur $\Gamma \times]0, T[$ subsiste. On peut aussi remarquer que dans ce cas, le problème peut être simplifié en prenant la condition nulle au bord.

Etant donné $a \in C^2(\Gamma \times]0, T[$, on peut en effet prolonger a de façon régulière à $\Omega \times]0, T[$. On note \bar{a} ce prolongement. Si v est la solution de l'équation (I.16) satisfaisant à une certaine condition initiale v_0 et à la condition aux limites a , au sens de (10), on pose

$$u = v - \bar{a},$$

et on vérifie que u satisfait à une équation du même type que (I.16),

à savoir

$$\frac{\partial u}{\partial t} + \operatorname{div}(f(u+\bar{a}(x,t), x, t)) + g(u+\bar{a}(x,t), x, t) + \frac{\partial \bar{a}}{\partial t}(x, t) = 0$$

en laissant inchangée la constante de LIPSCHITZ de g en u , ce qui confère le même type de continuité dans $L^1(\Omega)$ des opérateurs de semigroupe associés respectivement aux équations en v et en u .

La condition (8) devient

$$\operatorname{Min}_{k \in I(\gamma v, a)} \left\{ \operatorname{sg}(\gamma u) (f(\gamma u + a, x, t) - f(k, x, t)) \cdot v \right\} = 0 ,$$

et il suffit de poser $k' = k - a$, qui décrit $I(\gamma v, a) - \{a\} = I(\gamma u, 0)$ pour obtenir (9) avec la nouvelle fonction f , et une condition nulle au bord.

Cette remarque permet de ne considérer lors de la démonstration de l'existence et de l'unicité que le problème avec condition nulle au bord, ce qui est fait dans le prochain paragraphe.

2. EXISTENCE ET UNICITE POUR LE PROBLEME AUX LIMITES

On considère le problème suivant, sur $\Omega \times]0, T[$

$$(11) \quad \frac{\partial u}{\partial t} + \operatorname{div}(f(u, x, t)) + g(u, x, t) = 0 \quad \text{si } (x, t) \in \Omega \times]0, T[,$$

$$(12) \quad u(x, 0) = u_0(x) \quad \text{si } x \in \Omega ,$$

$$(13) \quad \operatorname{Min}_{k \in I(\gamma u, 0)} \left\{ \operatorname{sg}(\gamma u) (f(\gamma u, x, t) - f(k, x, t)) \cdot v \right\} = 0 ,$$

presque partout sur $\Gamma \times]0, T[$.

Les hypothèses sur f et g sont les mêmes que (I.4). On suppose u_0 de classe C^2 sur $\bar{\Omega}$. On reprend la fonction sg_η définie en (I.17), et on pose, pour $\xi \in \mathbb{R}^p$

$$||\xi|| = (\xi_1^2 + \dots + \xi_p^2)^{1/2} ,$$

et

$$I_\eta(\xi) = \int_0^{||\xi||} \operatorname{sg}_\eta(y) dy .$$

On a alors, pour tout $j \in \{1, \dots, p\}$

$$\frac{\partial I_\eta}{\partial \xi_j}(\xi) = \operatorname{sg}_\eta(||\xi||) \frac{\xi_j}{||\xi||} ,$$

et I_η est une fonction convexe, puisque l'application de \mathbb{R}^+ dans \mathbb{R}^+ définie par

$$z \longrightarrow \int_0^z \operatorname{sg}_\eta(y) dy$$

est elle-même convexe. Ceci assure que pour tout ξ et tout vecteur $w \in \mathbb{R}^p$,

on a

$$(15) \quad \sum_{k, j=1}^p \frac{\partial^2 I_\eta}{\partial \xi_k \partial \xi_j}(\xi) w_k w_j \geq 0 .$$

La fonction I_η introduite ici généralise celle introduite en (I.19), et permet mieux de traiter le laplacien, pour la méthode de pseudo-viscosité, ayant le même comportement dans chaque direction.

Pour établir l'existence, on utilise la méthode de pseudoviscosité. On considère $u_\varepsilon \in C^2(\bar{\Omega} \times]0, T[)$, solution unique de l'équation parabolique quasilineaire sur $\Omega \times]0, T[$

$$(16) \quad \frac{\partial u_\varepsilon}{\partial t} + \operatorname{div}(f(u_\varepsilon, x, t)) + g(u_\varepsilon, x, t) = \varepsilon \Delta u_\varepsilon$$

avec les conditions initiales et aux limites

$$(17) \quad \begin{cases} u_\varepsilon(x, 0) = u_0(x) & \text{si } x \in \Omega , \\ u_\varepsilon(x, t) = 0 & \text{si } x \in \Gamma, t \in]0, T[. \end{cases}$$

Le second membre de (16) correspond encore à un terme de viscosité, de coefficient ε strictement positif, et destiné à tendre vers zéro. L'existence et l'unicité pour u_ε solution de (16), (17) sont prouvées

dans la littérature ; on peut voir par exemple l'article de M.I. VISIK, ou l'article de O.A. LADYZENSKAYA et N.N. URAL'CEVA. On étudie ici la convergence de u_ϵ lorsque ϵ tend vers zéro en restant positif, ce qui, par un argument de compacité dans $W^{1,1}(\Omega \times]0,T[)$, permet d'établir l'existence d'une solution au problème (11), (12), (13) vérifiant une certaine condition d'entropie qui assure l'unicité. Cette solution est caractérisée par la définition suivante.

Définition 3 : Une fonction $u \in B V(\Omega \times]0,T[)$ est solution du problème (11), (12), (13) si, pour tout réel k et toute fonction test non négative

$$\phi \in C_{0+}^2(\bar{\Omega} \times]0,T[),$$

elle vérifie l'inégalité suivante

$$(18) \quad \int_{\Omega} \int_0^T \left\{ |u-k| \frac{\partial \phi}{\partial t} + \text{sg}(u-k) (f(u,x,t) - f(k,x,t)) \cdot \text{grad } \phi - \left[g(u,x,t) + \sum_i \frac{\partial f_i}{\partial x_i}(k,x,t) \right] \text{sg}(u-k) \phi \right\} dx dt + \int_{\Gamma} \int_0^T \text{sg}(k) (f(\gamma u,s,t) - f(k,s,t)) \cdot \nu \phi(s,t) ds dt \geq 0,$$

et la condition initiale (12) presque partout sur Ω , c'est-à-dire

$$(19) \quad \gamma u(x,0) = u_0(x) \quad \text{pp } (\Omega).$$

x x
x

Le lemme 3 assure l'existence de la trace γu , sur $\Gamma \times]0,T[$ et sur Ω pour $t = 0$, qui est uniformément bornée et mesurable, et ainsi l'inéquation (18) a un sens, de même que (19). On a le résultat suivant, dû à Cl. BARDOS, A.Y. LE ROUX et J.C. NEDELEC.

Théorème 11 : Le problème (11), (12), (13) admet une solution unique caractérisée par la définition 3. Cette solution est la limite dans $L^1(\Omega \times]0,T[)$ de $\{u_\epsilon\}_{\epsilon > 0}$, famille des solutions de (16), (17).

x x
x

Démonstration

On procède en trois étapes, en commençant par établir les estimations a priori qui vont assurer la convergence d'une suite $\{u_{\epsilon_m}\}$ vers une fonction $u \in B V(\Omega \times]0,T[)$. On vérifiera ensuite que u satisfait à la Définition 3. On établira enfin l'unicité, à partir d'un résultat de continuité dans $L^1(\Omega)$ du semi-groupe associé à (11) et (13). Certaines parties coïncident avec des passages de la démonstration du théorème 1.

A - Les estimations a priori et la convergence

On va estimer u_ϵ dans $L^\infty(\Omega \times]0,T[) \cap W^{1,1}(\Omega \times]0,T[)$, de façon à établir la convergence. D'après le principe du maximum (voir par exemple O.A. LADYZENSKAJA et N.N. URAL'CEVA), il existe une constante M_0 ne dépendant que des données, c'est-à-dire de f, g, u_0 et du domaine $\Omega \times]0,T[$, telle que pour tout $\epsilon > 0$,

$$(22) \quad \|u_\epsilon\|_{L^\infty(\Omega \times]0,T[)} \leq M_0.$$

Cette inégalité implique que u_ϵ est uniformément borné dans $L^1(\Omega \times]0,T[)$. L'estimation de la norme L^1 peut aussi être obtenue directement.

Pour estimer la norme L^1 de $\frac{\partial u_\epsilon}{\partial t}$, on dérive l'équation (16) par rapport à t et on la multiplie par $\text{sg}_\eta(\frac{\partial u_\epsilon}{\partial t})$, pour $\eta > 0$, destiné à tendre vers zéro. On intègre ensuite sur Ω pour obtenir

$$\begin{aligned}
 (23) \quad & \frac{d}{dt} \int_{\Omega} \left(\int_0^{\frac{\partial u_{\epsilon}}{\partial t}} sg_{\eta}(v) dv \right) dx \\
 & + \int_{\Omega} \operatorname{div} \left[\frac{\partial f}{\partial u}(u_{\epsilon}, x, t) \frac{\partial u_{\epsilon}}{\partial t} \right] sg_{\eta} \left(\frac{\partial u_{\epsilon}}{\partial t} \right) dx \\
 & + \int_{\Omega} \operatorname{div} \left[\frac{\partial f}{\partial t}(u_{\epsilon}, x, t) \right] + \frac{\partial}{\partial t} \left[g(u_{\epsilon}, x, t) \right] sg_{\eta} \left(\frac{\partial u_{\epsilon}}{\partial t} \right) dx \\
 & = \epsilon \int_{\Omega} \Delta \left(\frac{\partial u_{\epsilon}}{\partial t} \right) sg_{\eta} \left(\frac{\partial u_{\epsilon}}{\partial t} \right) dx .
 \end{aligned}$$

On intègre par parties le second et le dernier terme . Il

vient

$$\begin{aligned}
 & \int_{\Omega} \operatorname{div} \left[\frac{\partial f}{\partial u}(u_{\epsilon}, x, t) \frac{\partial u_{\epsilon}}{\partial t} \right] sg_{\eta} \left(\frac{\partial u_{\epsilon}}{\partial t} \right) dx \\
 & = - \int_{\Omega} \frac{\partial f}{\partial u}(u_{\epsilon}, x, t) \left(\frac{\partial u_{\epsilon}}{\partial t} sg'_{\eta} \left(\frac{\partial u_{\epsilon}}{\partial t} \right) \right) \cdot \operatorname{grad} \left(\frac{\partial u_{\epsilon}}{\partial t} \right) dx ,
 \end{aligned}$$

la condition au bord étant nulle. Lorsque η tend vers zéro, d'après le lemme 1 (SAKS), ce terme tend vers zéro, en restant uniformément borné sur $]0, T[$. D'après le théorème de LEBESGUE, ce terme tend encore vers zéro avec η , une fois qu'il est intégré sur $]0, t[$. On a en effet besoin d'intégrer en t avant le passage à la limite en η , pour pouvoir traiter le premier terme de (23). Dans le second terme de (23), on peut majorer $sg_{\eta} \left(\frac{\partial u_{\epsilon}}{\partial t} \right)$ par 1. Le second membre de (23) n'est pas positif. En effet, en intégrant par parties, il vient :

$$\epsilon \int_{\Omega} \Delta \left(\frac{\partial u_{\epsilon}}{\partial t} \right) sg_{\eta} \left(\frac{\partial u_{\epsilon}}{\partial t} \right) dx = - \epsilon \int_{\Omega} sg'_{\eta} \left(\frac{\partial u_{\epsilon}}{\partial t} \right) \left\| \operatorname{grad} \left(\frac{\partial u_{\epsilon}}{\partial t} \right) \right\|^2 dx ,$$

et qui reste négatif après intégration en t . On obtient ainsi après intégration sur $]0, t[$ de (23), et passage à la limite en η , l'inégalité :

$$\begin{aligned}
 & \left| \frac{\partial u_{\epsilon}}{\partial t}(\cdot, t) \right|_{L^1(\Omega)} - \left| \frac{\partial u_{\epsilon}}{\partial t}(\cdot, 0) \right|_{L^1(\Omega)} \\
 & \leq \int_0^t \int_{\Omega} \left\{ \left| \operatorname{div} \left(\frac{\partial f}{\partial t}(u_{\epsilon}, x, s) \right) \right| + \left| \frac{\partial}{\partial t} g(u_{\epsilon}, x, s) \right| \right\} dx ds
 \end{aligned}$$

$$\begin{aligned}
 & \leq \int_0^t \int_{\Omega} \left[\left| \sum_k \frac{\partial^2 f_k}{\partial x_k \partial t}(u_{\epsilon}, x, s) \right| + \left| \frac{\partial g}{\partial t}(u_{\epsilon}, x, s) \right| \right. \\
 & \quad \left. + \left| \sum_k \frac{\partial^2 f_k}{\partial u \partial t}(u_{\epsilon}, x, s) \frac{\partial u_{\epsilon}}{\partial x_k} \right| + \left| \frac{\partial g}{\partial u}(u_{\epsilon}, x, s) \frac{\partial u_{\epsilon}}{\partial t} \right| \right] dx ds .
 \end{aligned}$$

Cette dernière intégrale peut être majorée en utilisant les hypothèses (I.4) faites sur f et g . On note α_0 et α_2 les constantes de LIPSCHITZ respectives de g et de la fonction (à valeurs vectorielles) $\frac{\partial f}{\partial t}$, puis γ_2 le maximum de la fonction continue

$$\left| \frac{\partial g}{\partial t} \right| + \left| \sum_k \frac{\partial^2 f_k}{\partial x_k \partial t} \right|$$

sur le compact $[-M_0, M_0] \times \bar{\Omega} \times [0, T]$. Il vient

$$\begin{aligned}
 & \left| \frac{\partial u_{\epsilon}}{\partial t}(\cdot, t) \right|_{L^1(\Omega)} \leq \left| \frac{\partial u_{\epsilon}}{\partial t}(\cdot, 0) \right|_{L^1(\Omega)} \\
 & + \int_0^t \left(\gamma_2 \operatorname{mes}(\Omega) + \alpha_0 \left| \frac{\partial u_{\epsilon}}{\partial t}(\cdot, s) \right|_{L^1(\Omega)} + \alpha_2 \left\| \operatorname{grad} u_{\epsilon} \right\|_{L^1(\Omega)} \right) ds ,
 \end{aligned}$$

en notant

$$\left\| \operatorname{grad} u_{\epsilon} \right\|_{L^1(\Omega)} = \left(\sum_k \left| \frac{\partial u_{\epsilon}}{\partial x_k}(\cdot, s) \right|_{L^1(\Omega)}^2 \right)^{1/2} .$$

On a également, en reprenant l'équation (16) en $t = 0$,

$$\left| \frac{\partial u_{\epsilon}}{\partial t}(\cdot, 0) \right|_{L^1(\Omega)} \leq \left| \operatorname{div}[f(u_0, x, 0)] + g(u_0, x, 0) - \epsilon \Delta u_0 \right|_{L^1(\Omega)}$$

qu'on peut majorer par un terme de la forme

$$C_1' + C_2' \left| u_0 \right|_{W^{2,1}(\Omega)} ,$$

où C_1' et C_2' sont des constantes indépendantes de ϵ (dans la mesure où ϵ est borné par 1 par exemple), et qui ne dépendent que de $f(0, \cdot, 0)$, de

$g(0, \dots, 0)$ et de Ω pour C_1^1 , et de f, g et de leurs dérivées pour C_2^1 .

On obtient finalement

$$(24) \quad \left| \frac{\partial u_\epsilon}{\partial t} (\dots, t) \right|_{L^1(\Omega)} \leq C_1^1 + C_2^1 |u_0|_{W^{2,1}(\Omega)} + \gamma_2 t \text{mes}(\Omega) + \int_0^t \left(\alpha_0 \left| \frac{\partial u_\epsilon}{\partial t} (\dots, s) \right|_{L^1(\Omega)} + \alpha_2 \| |\text{grad } u_\epsilon(\dots, s)| \|_{L^1(\Omega)} \right) ds$$

et on conclura en utilisant le lemme de GRONWALL-BELLMANN après avoir établi une majoration analogue pour $\text{grad } u_\epsilon$.

Pour obtenir cette majoration, on prend le gradient de chaque membre de l'équation (16), puis on effectue le produit scalaire dans \mathbb{R}^p par le vecteur dont les composantes sont données par

$$\frac{\partial I_\eta}{\partial \xi_j} (\text{grad } u_\epsilon) \quad 1 \leq j \leq p.$$

On intègre ensuite sur Ω l'équation obtenue. En remarquant que par intégration par parties, il vient

$$\begin{aligned} & \sum_j \epsilon \int_\Omega \frac{\partial I_\eta}{\partial \xi_j} (\text{grad } u_\epsilon) \Delta \left(\frac{\partial u_\epsilon}{\partial x_j} \right) dx \\ &= -\epsilon \sum_j \sum_i \int_\Omega \frac{\partial}{\partial x_i} \left(\frac{\partial I_\eta}{\partial \xi_j} (\text{grad } u_\epsilon) \right) \frac{\partial^2 u_\epsilon}{\partial x_i \partial x_j} dx \\ & \quad + \epsilon \int_\Gamma \sum_j \frac{\partial I_\eta}{\partial \xi_j} (\text{grad } u_\epsilon) \frac{\partial}{\partial \nu} \left(\frac{\partial u_\epsilon}{\partial x_j} \right) ds. \end{aligned}$$

Or, sur la frontière $\text{grad } u_\epsilon = \frac{\partial u_\epsilon}{\partial \nu} \nu$, et en effectuant la dérivation en x_i dans l'intégrale sur Ω , puis en regroupant les termes dans l'intégrale sur la frontière pour retrouver une dérivée normale, on obtient que ce second membre est égal à

$$\begin{aligned} & -\epsilon \sum_i \int_\Omega \sum_{j,k} \frac{\partial^2 I_\eta}{\partial \xi_j \partial \xi_k} (\text{grad } u_\epsilon) \frac{\partial^2 u_\epsilon}{\partial x_i \partial x_k} \frac{\partial^2 u_\epsilon}{\partial x_i \partial x_j} dx \\ & \quad + \epsilon \int_\Gamma \frac{\partial}{\partial \nu} \left(I_\eta \left(\frac{\partial u_\epsilon}{\partial \nu} \nu \right) \right) ds. \end{aligned}$$

On effectue également la dérivation de $\text{div } f(u_\epsilon, x, t)$ en x_j , puis une intégration par parties analogue à la précédente sur le terme en $\frac{\partial f}{\partial u} \cdot \text{grad } I_\eta (\text{grad } u_\epsilon)$ qui apparaît. En tenant compte de l'expression de $\frac{\partial I_\eta}{\partial \xi_j}$, il vient

$$(25) \quad \begin{aligned} & \frac{d}{dt} \int_\Omega I_\eta (\text{grad } u_\epsilon(x, t)) dx \\ &= -\sum_i \epsilon \int_\Omega \sum_{j,k} \frac{\partial^2 I_\eta}{\partial \xi_j \partial \xi_k} (\text{grad } u_\epsilon) \frac{\partial^2 u_\epsilon}{\partial x_i \partial x_j} \frac{\partial^2 u_\epsilon}{\partial x_i \partial x_k} dx \\ & \quad - \int_\Omega \text{div} \left[\frac{\partial f}{\partial u} (u_\epsilon, x, t) \right] \left[\| |\text{grad } u_\epsilon| \| \text{sg}_\eta (\| |\text{grad } u_\epsilon| \|) - I_\eta (\text{grad } u_\epsilon) \right] dx \\ & \quad - \sum_j \int_\Omega \left\{ \text{div} \left[\frac{\partial f}{\partial x_j} (u_\epsilon, x, t) \right] + \frac{\partial}{\partial x_j} \left[g(u_\epsilon, x, t) \right] \right\} \frac{\partial u_\epsilon}{\partial x_j} \frac{\text{sg}_\eta (\| |\text{grad } u_\epsilon| \|)}{\| |\text{grad } u_\epsilon| \|} dx \\ & \quad + \int_\Gamma \left\{ \epsilon \frac{\partial}{\partial \nu} \left[I_\eta \left(\frac{\partial u_\epsilon}{\partial \nu} \nu \right) \right] - \frac{\partial f}{\partial u} (0, \sigma, t) \cdot \nu I_\eta \left(\frac{\partial u_\epsilon}{\partial \nu} \nu \right) \right\} d\sigma. \end{aligned}$$

Comme précédemment, on intègre sur $]0, t[$ avec $0 < t < T$, de façon à pouvoir faire tendre η vers zéro dans le premier membre, puis on considère séparément chaque terme du second membre.

En appliquant (15), qui exprime la convexité de I_η , au vecteur de \mathbb{R}^p défini par

$$w = \text{grad} \left(\frac{\partial u_\epsilon}{\partial x_i} \right)$$

pour chaque $i \in \{1, \dots, p\}$, il vient

$$-\sum_i \int_0^t \int_\Omega \sum_{j,k} \frac{\partial^2 I_\eta}{\partial \xi_j \partial \xi_k} (\text{grad } u_\epsilon) \frac{\partial^2 u_\epsilon}{\partial x_i \partial x_j} \frac{\partial^2 u_\epsilon}{\partial x_i \partial x_k} dx ds \leq 0,$$

et ce terme disparaît de (25).

Dans le second terme,

$$\| |\text{grad } u_\epsilon| \| \text{sg}_\eta (\| |\text{grad } u_\epsilon| \|) - I_\eta (\text{grad } u_\epsilon)$$

converge vers zéro lorsque η tend vers zéro, partout sur $\Omega \times]0, T[$.

En appliquant le théorème de LEBESGUE, on obtient que ce second terme tend vers zéro avec η .

Le troisième terme fait apparaître la norme de $\text{grad } u_\epsilon$ dans $L^1(\Omega)$. On a

$$\begin{aligned} & \left| \int_0^t \int_\Omega \left\{ \text{div} \left[\frac{\partial f}{\partial x_j} (u_\epsilon, x, s) \right] + \frac{\partial}{\partial x_j} [g(u_\epsilon, x, s)] \right\} \frac{\partial u_\epsilon}{\partial x_j} \frac{\text{sg}_\eta (||\text{grad } u_\epsilon||)}{||\text{grad } u_\epsilon||} dx ds \right. \\ & \leq \int_0^t \int_\Omega \left| \frac{\partial g}{\partial u} (u_\epsilon, x, s) \right| ||\text{grad } u_\epsilon|| dx ds \\ & + \int_0^t \int_\Omega \sum_{j,k} \left| \frac{\partial^2 f_k}{\partial u \partial x_j} (u_\epsilon, x, s) \right| ||\text{grad } u_\epsilon|| dx ds \\ & + \int_0^t \int_\Omega \left(||\text{grad}(g)(u_\epsilon, x, s)|| + ||\text{grad}(\text{div } f)(u_\epsilon, x, s)|| \right) dx ds. \end{aligned}$$

Ce dernier terme est borné ; en effet $\text{grad}(g)$ et $\text{grad}(\text{div } f)$ sont des fonctions continues sur le compact $[-M_0, M_0] \times \bar{\Omega} \times [0, T]$. Il peut être majoré par un terme de la forme $\gamma_3 \text{mes}(\Omega)$. Les deux autres termes peuvent être estimés en utilisant les constantes de LIPSCHITZ (uniforme) de g et f , et on réalise ainsi une majoration globale du troisième terme par

$$\int_0^t \left(\gamma_3 \text{mes}(\Omega) + (\alpha_0 + \alpha_2) ||\text{grad } u_\epsilon(\cdot, s)||_{L^1(\Omega)} \right) ds.$$

On considère maintenant l'intégrale sur la frontière, où n'apparaissent que des dérivations par rapport à la direction normale, parce que la condition nulle au bord fait disparaître toutes les dérivées tangentielles. On a, en remarquant que sur la frontière, l'équation (16) s'écrit

$$\frac{\partial f}{\partial u} (0, x, t) \cdot \nu \frac{\partial u_\epsilon}{\partial \nu} = \epsilon \Delta u_\epsilon - g(0, x, t) - (\text{div } f)(0, x, t),$$

la relation

$$\begin{aligned} & \int_0^t \int_\Gamma \left\{ \epsilon \frac{\partial}{\partial \nu} (I_\eta \left(\frac{\partial u_\epsilon}{\partial \nu} \right)) - \frac{\partial f}{\partial u} (0, \sigma, s) \cdot \nu I_\eta \left(\frac{\partial u_\epsilon}{\partial \nu} \right) \right\} d\sigma ds \\ & = \epsilon \int_0^t \int_\Gamma \left\{ \frac{\partial^2 u_\epsilon}{\partial \nu^2} \text{sg}_\eta \left(\frac{\partial u_\epsilon}{\partial \nu} \right) - \Delta u_\epsilon I_\eta \left(\frac{\partial u_\epsilon}{\partial \nu} \right) \left(\frac{\partial u_\epsilon}{\partial \nu} \right)^{-1} \right\} d\sigma ds \\ & + \int_0^t \int_\Gamma \left\{ g(0, \sigma, s) + \text{div } f(0, \sigma, s) I \left(\frac{\partial u_\epsilon}{\partial \nu} \right) \left(\frac{\partial u_\epsilon}{\partial \nu} \right)^{-1} \right\} d\sigma ds. \end{aligned}$$

Ce dernier terme peut être borné uniformément par rapport à ϵ et η ; en effet, on a toujours

$$\left| I_\eta \left(\frac{\partial u_\epsilon}{\partial \nu} \right) \left(\frac{\partial u_\epsilon}{\partial \nu} \right)^{-1} \right| \leq 1.$$

En faisant tendre η vers zéro sur l'autre terme, en remarquant que pour tout y ,

$$\lim_{\eta \rightarrow 0} \frac{1}{y} \int_0^{|y|} \text{sg}_\eta(y') dy' = \text{sg}(y),$$

on obtient à la limite

$$\epsilon \int_0^t \int_\Gamma \left| \frac{\partial^2 u_\epsilon}{\partial \nu^2} - \Delta u_\epsilon \right| d\sigma ds.$$

En exprimant Δu_ϵ en coordonnées tangentielles et normales, on vérifie qu'il existe une constante $C(\Gamma)$ qui ne dépend que de Γ (en fait de son rayon de courbure), telle que, en tout point de la frontière Γ et pour tout $s \in]0, t[$,

$$\left| \frac{\partial^2 u_\epsilon}{\partial \nu^2} - \Delta u_\epsilon \right| \leq C(\Gamma) \left| \frac{\partial u_\epsilon}{\partial \nu} \right|.$$

On utilise maintenant un résultat dû à Cl. BARDOS, D. BREZIS et H. BREZIS, qui permet d'exprimer, en norme L^1 , le gradient sur la frontière par le Laplacien à l'intérieur de Ω , lorsque la fonction considérée vérifie une condition nulle au bord. On a alors pour tout s ,

$$\varepsilon \left| \frac{\partial u_\varepsilon}{\partial \nu}(\cdot, s) \right|_{L^1(\Gamma)} \leq \left| \varepsilon \Delta u_\varepsilon(\cdot, s) \right|_{L^1(\Omega)}$$

On reprend maintenant l'équation (16) pour exprimer $\varepsilon \Delta u_\varepsilon$ en fonction de u_ε et de ses dérivées premières, et il vient une majoration de la forme

$$\varepsilon \left| \frac{\partial u_\varepsilon}{\partial \nu}(\cdot, s) \right|_{L^1(\Gamma)} \leq \left| \frac{\partial u_\varepsilon}{\partial t}(\cdot, s) \right|_{L^1(\Omega)} + \alpha_2 \|\text{grad } u_\varepsilon(\cdot, s)\|_{L^1(\Omega)} + \gamma_2$$

avec

$$\gamma_2 = \left| \text{div } f + g \right|_{L^\infty(\cdot]_{-M_0, M_0}[\times \Omega \times]0, T[}$$

On obtient finalement, à partir de (25), la majoration

$$(26) \quad \begin{aligned} & \|\text{grad } u_\varepsilon(\cdot, t)\|_{L^1(\Omega)} \leq \|\text{grad } u_0\|_{L^1(\Omega)} \\ & + \int_0^t \left\{ (\gamma_2 + \gamma_3) \text{mes}(\Omega) + [\alpha_0 + \alpha_2(1+C(\Gamma))] \|\text{grad } u_\varepsilon(\cdot, s)\|_{L^1(\Omega)} \right. \\ & \left. + C(\Gamma) \left| \frac{\partial u_\varepsilon}{\partial t}(\cdot, s) \right|_{L^1(\Omega)} \right\} ds \end{aligned}$$

On additionne (24) et (26), pour obtenir l'inégalité

$$(27) \quad \begin{aligned} & \left| \frac{\partial u_\varepsilon}{\partial t}(\cdot, t) \right|_{L^1(\Omega)} + \|\text{grad } u_\varepsilon(\cdot, t)\|_{L^1(\Omega)} \leq C_1 + C_2 |u_0|_{W^{2,1}(\Omega)} \\ & + C_3 \int_0^t \left\{ \left| \frac{\partial u_\varepsilon}{\partial t}(\cdot, s) \right|_{L^1(\Omega)} + \|\text{grad } u_\varepsilon(\cdot, s)\|_{L^1(\Omega)} \right\} ds \end{aligned}$$

valable pour tout $t \in]0, T[$, et où C_1, C_2 et C_3 sont des constantes indépendantes de ε ; elles ne dépendent que du domaine $\Omega \times]0, T[$, et des fonctions f, g et de leurs dérivées sur $\cdot]_{M_0, M_0}[\times \Omega \times]0, T[$. On peut remarquer que, dépendant de M_0 , elles dépendent d'une certaine façon de u_0 , mais la régularité de u_0 n'intervient pas.

On applique le lemme de GRONWALL BELLMANN (cf. H. BREZIS par exemple), et on obtient la majoration suivante, pour tout $t \in]0, T[$

$$\left| \frac{\partial u_\varepsilon}{\partial t}(\cdot, t) \right|_{L^1(\Omega)} + \|\text{grad } u_\varepsilon(\cdot, t)\|_{L^1(\Omega)} \leq \left(C_1 + C_2 |u_0|_{W^{2,1}(\Omega)} \right) e^{C_3 t}$$

Cette estimation et (22) permettent de conclure qu'il existe une constante K_0 , indépendante de ε , telle que

$$(28) \quad |u_\varepsilon|_{L^\infty(\Omega \times]0, T[)} + \left| \frac{\partial u_\varepsilon}{\partial t} \right|_{L^1(\Omega \times]0, T[)} + \|\text{grad } u_\varepsilon\|_{L^1(\Omega \times]0, T[)} \leq K_0$$

Ainsi, la famille $\{u_\varepsilon\}_{\varepsilon > 0}$ est bornée dans $W^{1,1}(\Omega \times]0, T[)$, et d'après le théorème de RIESZ-TAMARKIN, on peut en extraire une suite $\{u_{\varepsilon_m}\}_m$ convergente (fortement) dans $L^1(\Omega \times]0, T[)$. On note u sa limite. D'après (22), u_{ε_m} converge aussi vers u dans L^∞ faible *, donc u est borné, et de plus $W^{1,1}(\Omega \times]0, T[)$ s'injecte continûment dans $BV(\Omega \times]0, T[)$, on en déduit que u appartient à $BV(\Omega \times]0, T[)$. On peut ajouter également qu'il y a convergence faible dans l'espace des mesures de RADON.

B - Existence

On démontre maintenant que u satisfait aux critères de la Définition 3. D'après le lemme 3, u admet une trace notée $\gamma u(\cdot, 0)$ en $t = 0$ qui appartient à $L^\infty(\Omega)$ et est la limite dans $L^1(\Omega)$ de $u(\cdot, t)$ lorsque t tend vers zéro. D'après le théorème de FUBINI, pour presque tout t , $u_{\varepsilon_m}(\cdot, t)$ converge vers $u(\cdot, t)$ presque partout sur Ω (il s'agit en fait d'une sous-suite extraite de u_{ε_m}), et par le théorème de LEBESGUE, en utilisant (22), on établit que pour presque tout t , $u_{\varepsilon_m}(\cdot, t)$ converge vers $u(\cdot, t)$ dans $L^1(\Omega)$. De plus, u_ε vérifiant la condition initiale, on a

$$|u_\varepsilon(\cdot, t) - u_0|_{L^1(\Omega)} \leq K_0 t$$

d'après (28). On peut donc, pour presque tout $t \in]0, T[$, passer à la limite en m sur chaque terme du second membre de l'inégalité

$$\begin{aligned} |\gamma u(\cdot, 0) - u_0|_{L^1(\Omega)} &\leq |\gamma u(\cdot, 0) - u(\cdot, t)|_{L^1(\Omega)} \\ &+ |u(\cdot, t) - u_m(\cdot, t)|_{L^1(\Omega)} \\ &+ |u_{em}(\cdot, t) - u_0|_{L^1(\Omega)} \quad (\leq K_0 t) . \end{aligned}$$

On obtient

$$|\gamma u(\cdot, 0) - u_0|_{L^1(\Omega)} \leq K_0 t ,$$

et ceci pour presque tout $t \in]0, T[$. On fait tendre t vers zéro et on obtient (19).

Il reste à prouver (18). On introduit $k \in \mathbb{R}$, et une fonction test

$$\phi \in C_{0,+}^2(\bar{\Omega} \times]0, T[) .$$

On multiplie, pour $\eta > 0$ et $\varepsilon \in \{\varepsilon_m\}_m$, l'équation (16) par $\phi \operatorname{sg}_\eta(u_\varepsilon - k)$, et on intègre par parties sur $\bar{\Omega} \times]0, T[$. Il vient l'expression suivante (donc chaque terme est noté T_i , $i = 1, \dots, 8$)

$$\begin{aligned} &\int_{\Omega} \int_0^T \left\{ \int_k^{u_\varepsilon} \operatorname{sg}_\eta(v-k) \, dv \right\} \frac{\partial \phi}{\partial t} \, dx \, dt \quad (T_1) \\ &+ \int_{\Omega} \int_0^T \operatorname{sg}_\eta(u_\varepsilon - k) (f(u_\varepsilon, x, t) - f(k, x, t)) \cdot \operatorname{grad} \phi \, dx \, dt \quad (T_2) \\ &- \int_{\Omega} \int_0^T \operatorname{sg}_\eta(u_\varepsilon - k) (\operatorname{div} f(k, x, t) + \operatorname{grad}(u_\varepsilon, x, t)) \phi \, dx \, dt \quad (T_3) \\ &+ \int_{\Omega} \int_0^T (f(u_\varepsilon, x, t) - f(k, x, t)) \cdot \operatorname{grad} u_\varepsilon \operatorname{sg}'_\eta(u_\varepsilon - k) \phi \, dx \, dt \quad (T_4) \\ (29) \quad &= \varepsilon \int_{\Omega} \int_0^T \operatorname{sg}'_\eta(u_\varepsilon - k) |\operatorname{grad} u_\varepsilon|^2 \phi \, dx \, dt \quad (T_5) \\ &+ \varepsilon \int_{\Omega} \int_0^T \operatorname{sg}_\eta(u_\varepsilon - k) \operatorname{grad} u_\varepsilon \cdot \operatorname{grad} \phi \, dx \, dt \quad (T_6) \\ &+ \varepsilon \int_{\Gamma} \int_0^T \operatorname{sg}_\eta(k) \frac{\partial u_\varepsilon}{\partial \nu} \phi \, d\sigma \, dt \quad (T_7) \\ &- \int_{\Gamma} \int_0^T \operatorname{sg}_\eta(k) (f(0, \sigma, t) - f(k, \sigma, t)) \cdot \nu \phi \, d\sigma \, dt \quad (T_8) \end{aligned}$$

On fait tendre η vers zéro. Les termes T_1 , T_2 et T_3 donnent à la limite

$$(30) \quad \int_{\Omega} \int_0^T \left[|u_\varepsilon - k| \frac{\partial \phi}{\partial t} + \operatorname{sg}(u_\varepsilon - k) (f(u_\varepsilon, x, t) - f(k, x, t)) \cdot \operatorname{grad} \phi - \operatorname{sg}(u_\varepsilon - k) (\operatorname{div} f(k, x, t) + \operatorname{grad}(u_\varepsilon, x, t)) \phi \right] dx \, dt ,$$

c'est-à-dire le premier terme de (18) après passage à la limite en ε , ce qui sera fait plus tard.

D'après le lemme 1, le terme T_3 tend vers zéro.

Le terme T_5 est non négatif ; on le minore par zéro.

D'après (28), le terme T_6 peut être majoré par

$$(31) \quad \varepsilon \|\operatorname{grad} \phi\|_{L^\infty(\Omega)} K_0 ,$$

et va tendre vers zéro avec ε .

Il reste les termes sur la frontière, où le passage à la limite en η est immédiat.

On passe maintenant à la limite en ε ; pour passer à la limite sur (30), on procède par régularisation, en utilisant le lemme 2. On obtient bien le premier membre de (18) à la limite. Ceci ne perturbe pas le terme (31), qui donne encore zéro à la limite. On considère maintenant les termes sur la frontière, c'est-à-dire, après régularisation l'expression suivante (voir lemme 2), où η est le nouveau paramètre de régularisation, destiné à tendre vers zéro

$$(32) \quad \int_{\mathbb{R}} \frac{1}{2} \operatorname{sg}'_\eta(\xi - k) \left[\int_{\Gamma'} \int_0^T \varepsilon \operatorname{sg}(\xi) \frac{\partial u_\varepsilon}{\partial \nu} \phi \, d\sigma \, dt + \int_{\Gamma} \int_0^T \operatorname{sg}'(\xi) (f(\xi, \sigma, t) - f(0, \sigma, t)) \cdot \nu \phi \, d\sigma \, dt \right] d\xi .$$

Il est immédiat que lorsque η tend vers zéro,

$$\int_{\mathbb{R}} \frac{1}{2} \operatorname{sg}'_{\eta}(\xi-k) \int_{\Gamma} \int_0^T \operatorname{sg}(\xi) (f(\xi, \sigma, t) - f(0, \sigma, t)) \cdot \nu \phi \, d\sigma \, dt \, d\xi$$

converge vers

$$\int_{\Gamma} \int_0^T \operatorname{sg}(k) (f(k, \sigma, t) - f(0, \sigma, t)) \cdot \nu \phi \, d\sigma \, dt \quad ;$$

il suffit en effet d'introduire $(-f(k, \sigma, t) + f(k, \sigma, t))$ dans l'intégrale sur $\Gamma \times]0, T[$, et d'utiliser la constante de LIPSCHITZ de f en u (ici en ξ ou k), comme dans la démonstration du lemme 2.

Dans l'autre intégrale, on commence par effectuer le passage à la limite en ε . En supposant qu'on ait obtenu une limite, le passage à la limite en η devient immédiat, la limite obtenue étant indépendante de ξ . On considère donc

$$\varepsilon \int_{\Gamma} \int_0^T \phi \frac{\partial u_{\varepsilon}}{\partial \nu} \, d\sigma \, dt \quad .$$

On introduit une fonction $\rho_{\delta} \in C^2(\bar{\Omega})$, non négative, dépendant d'un paramètre δ destiné à tendre vers zéro et telle que

$$(33) \quad \begin{aligned} \rho_{\delta}(\sigma) &= 1 & \text{si } \sigma \in \Gamma, \\ \rho_{\delta}(x) &= 0 & \text{si } x \in \Omega \text{ et } \operatorname{dist}(x, \Gamma) \geq \delta, \\ 0 \leq \rho_{\delta} &\leq 1 & \text{sur } \bar{\Omega}, \\ \|\operatorname{grad} \rho_{\delta}\|_{L^{\infty}(\Omega)} &\leq C / \delta, \end{aligned}$$

où C est une constante indépendante de δ . Une telle fonction peut se construire par exemple en régularisant la fonction indicatrice du complémentaire d'un ouvert Ω_{δ} , défini par

$$\Omega_{\delta} = \{ x \in \Omega \mid \operatorname{dist}(x, \Gamma) \geq \delta/2 \},$$

puis en utilisant la suite j_m introduite au début de la démonstration du théorème 1. On a alors, en écrivant que $\phi(\sigma, t)$ est la trace sur $\Gamma \times]0, T[$ de $\rho_{\delta}(x) \phi(x, t)$, l'égalité

$$\begin{aligned} &\int_{\Gamma} \int_0^T \varepsilon \frac{\partial u_{\varepsilon}}{\partial \nu} \phi \, d\sigma \, dt \\ &= \int_{\Omega} \int_0^T \varepsilon \left\{ \Delta u_{\varepsilon} \phi \rho_{\delta} + \operatorname{grad} u_{\varepsilon} \cdot \operatorname{grad}(\phi \rho_{\delta}) \right\} dx \, dt, \end{aligned}$$

c'est-à-dire, en utilisant l'équation (16), et en intégrant par parties

$$\begin{aligned} &= - \int_{\Omega} \int_0^T \left\{ u_{\varepsilon} \frac{\partial \phi}{\partial t} + f(u_{\varepsilon}, x, t) \cdot \operatorname{grad} \phi - g(u_{\varepsilon}, x, t) \phi \right\} \rho_{\delta} dx \, dt \\ &\quad - \int_{\Omega} \int_0^T \phi f(u_{\varepsilon}, x, t) \cdot \operatorname{grad} \rho_{\delta} \, dx \, dt \\ &\quad + \int_{\Gamma} \int_0^T \phi f(0, \sigma, t) \cdot \nu \, d\sigma \, dt \\ &\quad + \int_{\Omega} \int_0^T \varepsilon \operatorname{grad} u_{\varepsilon} \cdot \operatorname{grad}(\phi \rho_{\delta}) \, dx \, dt. \end{aligned}$$

On fait tendre ε vers zéro, le passage à la limite est immédiat sur chaque terme de l'expression précédente, car il suffit d'y changer u_{ε} en u , sauf peut-être pour le dernier où il faut remarquer que

$$\varepsilon \|\operatorname{grad} u_{\varepsilon}\|_{L^1(\Omega)}$$

tend vers zéro avec ε , d'après l'estimation (28). On fait tendre ensuite δ vers zéro; le premier terme tend vers zéro, car il s'agit de l'intégrale d'une fonction de $L^1(\Omega \times]0, T[)$, sur une partie de $\Omega \times]0, T[$ dont la mesure de LEBESGUE tend vers zéro. Le dernier terme est indépendant de δ . En notant $dv(x, t)$ la mesure de RADON associée à $\operatorname{div} [f(u, x, t)]$, on a la formule de GREEN suivante, analogue à (10), et en utilisant le lemme 3, car f est continue en u ,

$$\begin{aligned} \int_{\Omega} \int_0^T \operatorname{grad} \rho_{\delta} \cdot f(u, x, t) \phi \, dx \, dt &= \int_{\Gamma} \int_0^T f(\gamma u, \sigma, t) \cdot \nu \phi \, d\sigma \, dt \\ &\quad - \int_{\Omega} \int_0^T \rho_{\delta} \phi \, dv(x, t) - \int_{\Omega} \int_0^T \rho_{\delta} f(u, x, t) \cdot \operatorname{grad} \phi \, dx \, dt. \end{aligned}$$

Cette dernière intégrale tend vers zéro, comme précédemment, et il en est de même de la précédente. En effet, ρ_{δ} tend vers zéro partout sur $\Omega \times]0, T[$, donc presque partout pour la mesure $dv(x, t)$, et

on conclut par le théorème de LEBESGUE. On obtient donc finalement le résultat de convergence faible suivant

$$(34) \quad \lim_{\varepsilon \rightarrow 0} \int_{\Gamma} \int_0^T \varepsilon \frac{\partial u_\varepsilon}{\partial v} \phi \, d\sigma \, dt = - \int_{\Gamma} \int_0^T f(\gamma u, \sigma, t) \cdot v \phi \, d\sigma \, dt + \int_{\Gamma} \int_0^T \phi f(0, \sigma, t) \cdot v \, d\sigma \, dt .$$

On introduit ce résultat dans (32), où le passage à la limite ne pose aucun problème en utilisant le théorème de LEBESGUE une fois de plus. On effectue ensuite le passage à la limite en η , et on obtient, à partir de (29)

$$\int_{\Omega} \int_0^T \left\{ |u-k| \frac{\partial \phi}{\partial t} + \text{sg}(u-k) (f(u, x, t) - f(k, x, t)) \cdot \text{grad } \phi - \text{sg}(u-k) (\text{div } f(k, x, t) + g(u, x, t)) \phi \right\} dx \cdot dt \\ \geq \int_{\Gamma} \int_0^T \phi \cdot v \cdot [-f(\gamma u, \sigma, t) + f(0, \sigma, t) - f(0, \sigma, t) + f(k, \sigma, t)] \text{sg}(k) \, d\sigma \, dt ,$$

c'est-à-dire exactement (18), puisque $f(0, \sigma, t)$ disparaît du second membre. L'existence d'une solution vérifiant les critères de la Définition 3 est établie.

C - Unicité

Soient u et v deux fonctions appartenant à $BV(\Omega \times]0, T[)$, et vérifiant la relation (18). Ces deux fonctions sont ainsi supposées satisfaire aux mêmes conditions aux limites.

En prenant dans (18) une fonction ϕ à support compact dans $\Omega \times]0, T[$, on obtient exactement (I.25), dont on peut déduire (I.34) de la même façon que dans l'article de S.N. KRUIZKOV. On a donc

$$(35) \quad \int_{\Omega} \int_0^T \left\{ |u-v| \frac{\partial \phi}{\partial t} + \text{sg}(u-v) (f(u, x, t) - f(v, x, t)) \cdot \text{grad } \phi - \text{sg}(u-v) (g(u, x, t) - g(v, x, t)) \phi \right\} dx \, dt \geq 0 ,$$

pour toute fonction $\phi \in C_{0+}^2(\Omega \times]0, T[)$.

Soit $\psi \in C_{0+}^2(]0, T[)$ et $\delta > 0$; on reprend une fonction de la forme ρ_δ introduite en (33), et on utilise dans (35) la fonction test définie sur $\Omega \times]0, T[$ par

$$\phi(x, t) = \psi(t) (1 - \rho_\delta(x)) ,$$

de façon à faire apparaître la trace sur la frontière lorsque δ tend vers zéro. Ce passage à la limite fait apparaître une intégrale sur la frontière où interviennent les traces respectives de u et v , qui existent d'après le lemme 3, et on obtient, comme pour établir (34), l'inégalité suivante, valable pour toute fonction $\psi \in C_{0+}^2(]0, T[)$,

$$(36) \quad \int_{\Omega} \int_0^T \left\{ |u-v| \psi'(t) - \text{sg}(u-v) (g(u, x, t) - g(v, x, t)) \psi(t) \right\} dx \, dt \\ \geq \int_{\Gamma} \int_0^T \text{sg}(\gamma u - \gamma v) (f(\gamma u, \sigma, t) - f(\gamma v, \sigma, t)) \cdot v \psi(t) \, d\sigma \, dt .$$

On veut maintenant prouver que le second membre de (36) est non négatif. En un point (σ, t) de $\Gamma \times]0, T[$, on définit le nombre réel $k(\sigma, t)$ par

$$(37) \quad k(\sigma, t) = \begin{cases} \gamma u(\sigma, t) & \text{si } \gamma u(\sigma, t) \in I(0, \gamma v(\sigma, t)) , \\ 0 & \text{si } 0 \in I(\gamma u(\sigma, t), \gamma v(\sigma, t)) , \\ \gamma v(\sigma, t) & \text{si } \gamma v(\sigma, t) \in I(0, \gamma u(\sigma, t)) . \end{cases}$$

Ceci signifie que des trois quantités $\gamma u(\sigma, t)$, $\gamma v(\sigma, t)$ et 0, on note $k(\sigma, t)$ celle qui appartient à l'intervalle fermé limité par les deux autres. On a alors la relation suivante, valable presque partout sur $\Gamma \times]0, T[$, pour $k = k(\sigma, t)$

$$(38) \quad \text{sg}(\gamma u - \gamma v) (f(\gamma u, \sigma, t) - f(\gamma v, \sigma, t)) = \\ \text{sg}(\gamma u) (f(\gamma u, \sigma, t) - f(k, \sigma, t)) + \text{sg}(\gamma v) (f(\gamma v, \sigma, t) - f(k, \sigma, t))$$

On établit cette relation en considérant l'une après l'autre chaque valeur possible de $k(\sigma, t)$ donnée dans (37).

Prenons maintenant dans (18), la fonction test ϕ définie par

$$\phi(x,t) = \psi(t) \rho_\delta(x) ,$$

où ψ et ρ_δ sont les mêmes fonctions que celles introduites précédemment. On obtient, en passant à la limite en δ , pour tout $k \in \mathbb{R}$ et presque partout sur $\Gamma \times]0, T[$

$$(f(\gamma u, \sigma, t) - f(k, \sigma, t)) \cdot \nu (sg(\gamma u - k) + sg(k)) \geq 0.$$

En prenant

$$k \in I(\gamma u, 0) ,$$

il vient

$$sg(\gamma u - k) = sg(k) = sg(\gamma u)$$

si k est différent de 0 ou de γu . Dans les autres cas, on a toujours

$$sg(\gamma u - k) + sg(k) = sg(\gamma u) ,$$

et ainsi, pour tout $k \in I(\gamma u, 0)$, on retrouve (13), c'est-à-dire

$$sg(\gamma u) (f(\gamma u, \sigma, t) - f(k, \sigma, t)) \cdot \nu \geq 0.$$

On a la même inégalité pour γv et $k \in I(\gamma v, 0)$. Or, dans (37),

on a

$$k = k(\sigma, t) \in I(\gamma u, 0) \cap I(\gamma v, 0) .$$

En multipliant par n scalairement, on vérifie ainsi, à partir de (38), que pour presque tout $(\sigma, t) \in \Gamma \times]0, T[$,

$$sg(\gamma u - \gamma v) (f(\gamma u, \sigma, t) - f(\gamma v, \sigma, t)) \cdot \nu \geq 0.$$

On déduit donc de (36) que

$$\int_{\Omega} \int_0^T \left\{ |u-v| \psi'(t) - sg(u-v) (g(u, x, t) - g(v, x, t)) \psi(t) \right\} dx dt \geq 0.$$

Comme au chapitre I, on prend la fonction ψ de la forme

$$\psi(t) = e^{-\alpha_0 t} \mu_m(t) , \quad (\alpha_0 \text{ constante de LIPSCHITZ de } g)$$

où μ_m est une approximation régulière de la fonction indicatrice d'un intervalle $[t_0, t_1] \subset]0, T[$. On obtient une inégalité analogue à (I.36) en passant à la limite en m , qui s'écrit ici

$$(39) \quad |u(\cdot, t_1) - v(\cdot, t_1)|_{L^1(\Omega)} \leq e^{\alpha_0(t_1 - t_0)} |u(\cdot, t_0) - v(\cdot, t_0)|_{L^1(\Omega)} .$$

On en déduit l'unicité en faisant tendre t_0 vers zéro, lorsque u et v satisfont à la même condition initiale. Le théorème 11 est démontré.

x x
x

La relation (39) appelle une remarque ; elle exprime que le semigroupe associé à l'équation (11) et à la condition aux limites nulle par l'inéquation (18) est une application continue dans $L^1(\Omega)$. Lorsque g ne dépend pas de la variable u , cette application est une contraction. On retrouve ainsi les mêmes propriétés que pour le problème de CAUCHY.

On peut traiter également le problème (11), (12), (13) sur un ouvert Ω non borné. La formulation de la définition reste la même, en précisant que

$$u \in L^\infty(\Omega \times]0, T[) \cap BV_{loc}(\Omega \times]0, T[) ,$$

et le théorème d'existence et d'unicité reste vrai. On procède comme au théorème 11 près de la frontière, et comme au théorème 1 en faisant intervenir le cône caractéristique, sur les parties de Ω ouvertes sur l'infini. Ce problème est traité dans LE ROUX [6] dans le cadre monodimensionnel

La formulation de la solution donnée dans la Définition 3 se généralise de la façon suivante, dans le cas d'une condition aux limites $a(\sigma, t)$ non nécessairement nulle, c'est-à-dire pour le problème (11), (12), (9). On conserve (19), et on remplace (18) par

$$(40) \quad \int_{\Omega} \int_0^T \left\{ |u-k| \frac{\partial \phi}{\partial t} + \text{sg}(u-k) (f(u, x, t) - f(k, x, t)) \cdot \text{grad } \phi - \left[g(u, x, t) + \text{div } f(k, x, t) \right] \text{sg}(u-k) \phi \right\} dx dt + \int_{\Gamma} \int_0^T \text{sg}(k-a) (f(\gamma u, \sigma, t) - f(k, \sigma, t)) \cdot \nu \phi \, d\sigma dt \geq 0 ;$$

le théorème 11 reste vrai.

On peut remarquer également que (9) peut être déduit de (40), de la même façon que (13) a été déduit de (18).

CHAPITRE V

APPROXIMATION POUR LE PROBLEME AUX LIMITES

1. LE SCHEMA DE GODUNOV

On considère le problème suivant pour $\Omega =]0, 1[$

$$(1) \quad \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0 \quad \text{si } (x, t) \in \Omega \times]0, T[,$$

avec pour condition initiale

$$(2) \quad u(x, 0) = u_0(x) \quad \text{si } x \in \Omega ,$$

et les conditions aux limites, presque partout sur $]0, T[$

$$(3) \quad \sup_{k \in I(\gamma u(0, t), a(t))} \left\{ \text{sg}(\gamma u(0, t) - a(t)) (f(\gamma u(0, t)) - f(k)) \right\} = 0 ,$$

$$(4) \quad \inf_{k \in I(\gamma u(1, t), b(t))} \left\{ \text{sg}(\gamma u(1, t) - b(t)) (f(\gamma u(1, t)) - f(k)) \right\} = 0 .$$

On suppose, ce qui est une hypothèse plus faible que dans le chapitre précédent

$$u_0 \in BV(\Omega) \quad ; \quad a, b \in BV(0, T) \quad ;$$

et on se propose de démontrer la convergence du schéma de GODUNOV sans faire de relèvement, qui permettrait d'annuler a et b.

Soit I un entier, destiné à tendre vers l'infini ; on pose

$$h = 1/I \quad ,$$

puis on reprend le maillage introduit au chapitre II, et les mêmes notations. Pour tout $n < N$, on note

$$(5) \quad u_0^n = \frac{1}{qh} \int_{J_n} a(t) dt \quad ; \quad u_I^n = \frac{1}{qh} \int_{J_n} b(t) dt \quad ;$$

et pour $i \leq I-1$

$$(6) \quad u_i^0 = \frac{1}{h} \int_{I_i} u_0(x) dx \quad ,$$

au besoin en prolongeant a, b et u_0 par parité, au voisinage de 0 et 1.

Pour $i \in \{1, \dots, I-1\}$ et $n < N$, chaque u_i^{n+1} est calculé par le schéma de GODUNOV, construit au chapitre II

$$(7) \quad u_{i+1/2}^n \in I(u_{i+1}^n, u_i^n) \quad \text{réalise} \\ \text{Min}_{k \in I(u_{i+1}^n, u_i^n)} \left\{ \text{sg}(u_{i+1}^n - u_i^n) f(k) \right\} \quad ,$$

et

$$(8) \quad u_i^{n+1} = u_i^n - q(f(u_{i+1/2}^n) - f(u_{i-1/2}^n))$$

On note u_n la fonction égale à la constante u_i^n sur chaque pavé $I_i \times J_n$; on a le résultat de convergence suivant.

Théorème 12 : Sous la condition de stabilité de Courant Friedrichs Lewy

$$(9) \quad q \sup_{|u| \leq M} |f'(u)| \leq 1 \quad , \quad M = \text{Max}(|u_0|_{L^\infty}, |a|_{L^\infty}, |b|_{L^\infty}) \quad ,$$

la famille $\{u_n\}_{n>0}$ des solutions approchées construites par le schéma de GODUNOV (6), (7) converge vers $u \in BV(\Omega \times]0, T[)$, solution du problème (1), (2), (3), (4) et caractérisée par la définition 3, dans $L^1(\Omega \times]0, T[)$, lorsque h tend vers zéro.

x x
x

Démonstration

On utilise de nombreux résultats de la démonstration du théorème II. Tout d'abord, d'après (II.20), on a pour tout couple (i, n)

$$|u_i^{n+1}| \leq \text{Sup}(|u_{i-1/2}^n|, |u_i^n|, |u_{i+1/2}^n|)$$

et la stabilité dans $L^\infty(\Omega)$ est assurée ; il vient de proche en proche

$$(10) \quad |u_n|_{L^\infty(\Omega \times]0, T[)} \leq \text{Max} \left\{ |u_0|_{L^\infty(\Omega)}, |a|_{L^\infty(0, T)}, |b|_{L^\infty(0, T)} \right\} = M.$$

On a également l'analogue de (II.28) qui exprime la conservation de la variation bornée en temps

$$\sum_{i=1}^{I-1} |u_i^{n+1} - u_i^n| \leq \sum_{i=0}^I |u_{i+1}^n - u_i^n|$$

dont on déduit

$$(11) \quad \sum_{i=0}^I |u_i^{n+1} - u_i^n| \leq \sum_{i=0}^I |u_{i+1}^n - u_i^n| + |u_0^{n+1} - u_0^n| + |u_I^{n+1} - u_I^n| \quad ,$$

qui fait apparaître les variations respectives de a et de b.

La conservation de la variation bornée en espace se démontre à partir de (II-30), c'est-à-dire, pour $i \in \{1, \dots, I-2\}$

$$(12) \quad \begin{aligned} |u_{i+1}^{n+1} - u_i^{n+1}| &\leq |u_{i+1}^n - u_i^n| - q f'(\xi_{i+5/4}^n) |u_{i+3/2}^n - u_{i+1}^n| \\ &+ q f'(\xi_{i-1/2}^n) |u_i^n - u_{i-1/2}^n| - q f'(\xi_{i+3/4}^n) |u_{i+1}^n - u_{i+1/2}^n| \\ &+ q f'(\xi_{i+1/4}^n) |u_{i+1/2}^n - u_i^n|. \end{aligned}$$

D'autre part,

$$(13) \quad \begin{aligned} |u_1^{n+1} - u_0^{n+1}| &\leq |u_1^n - u_0^n - q(f(u_{3/2}^n) - f(u_{1/2}^n))| + |u_0^{n+1} - u_0^n| \\ &\leq |u_1^n - u_0^n| - q f'(\xi_{3/4}^n) |u_1^n - u_{1/2}^n| + q f'(\xi_{1/4}^n) |u_{1/2}^n - u_0^n| \\ &\quad + |u_0^{n+1} - u_0^n|, \end{aligned}$$

en procédant comme pour établir (II.28).

De la même façon

$$(14) \quad \begin{aligned} |u_I^{n+1} - u_{I-1}^{n+1}| &\leq |u_I^n - u_{I-1}^n| - q f'(\xi_{I-3/4}^n) |u_{I-1/2}^n - u_{I-1}^n| \\ &+ q f'(\xi_{I-5/4}^n) |u_{I-1}^n - u_{I-3/2}^n| + |u_I^{n+1} - u_I^n|. \end{aligned}$$

En sommant (12) en $i \in \{1, \dots, I-2\}$, puis en y ajoutant (13) et (14), il vient

$$(15) \quad \sum_{i=0}^{I-1} |u_{i+1}^{n+1} - u_i^{n+1}| \leq \sum_{i=0}^{I-1} |u_{i+1}^n - u_i^n| + |u_I^{n+1} - u_I^n| + |u_0^{n+1} - u_0^n|.$$

Les estimations (10), (11), et (15), ainsi que les hypothèses de variation bornée faites sur a et b , assurent que l'on peut extraire de la famille $\{u_h\}$ une suite $\{u_n\}$ convergente vers $u \in BV(\Omega \times]0, T[)$, dans $L^1(\Omega \times]0, T[)$. La démonstration est identique à celle de la démonstration du théorème II, paragraphe D. On sait, d'après le lemme 3 que u admet

une trace $\gamma u(x)$ en $t = 0$, et en reprenant le paragraphe F de la démonstration du théorème II, on établit que u satisfait à (IV.19).

Il reste à vérifier (IV.18). On reprend (II.39), qui s'écrit ici, pour tout $k \in \mathbb{R}$ et pour $i \in \{1, \dots, I-1\}$

$$(16) \quad \begin{aligned} |u_i^{n+1} - k| &\leq |u_i^n - k| - q \operatorname{sg}(u_{i+1/2}^n - k) (f(u_{i+1/2}^n) - f(k)) \\ &\quad + q \operatorname{sg}(u_{i-1/2}^n - k) (f(u_{i-1/2}^n) - f(k)). \end{aligned}$$

On introduit une fonction test

$$\phi \in C_{0,+}^2(\Omega \times]0, T[),$$

qu'on approche par des projections de type L^2 , comme précédemment, et on obtient en multipliant (16) par $h_m \phi_1^n$, en sommant en i et n puis en passant à la limite en h_m , l'inéquation suivante

$$\int_{\Omega} \int_0^T \left\{ |u-k| \frac{\partial \phi}{\partial t} + \operatorname{sg}(u-k) (f(u) - f(k)) \frac{\partial \phi}{\partial x} \right\} dx dt \geq 0,$$

c'est-à-dire (18) sans le terme sur la frontière.

Pour établir (18), on va maintenant vérifier (3) et (4).

Soit $h \in \{h_m\}$ et une fonction test

$$\psi \in C_{0,+}^1(]0, T[).$$

$$\text{On note } \psi^n = \frac{1}{qh} \int_J \psi(t) dt,$$

et pour $k \in \mathbb{R}$, $i \in \{1, \dots, I-1\}$ fixés, on multiplie (16) par $h \psi^n$, puis on somme en n . Il vient, si h est assez petit,

$$\begin{aligned} &\sum_{n=0}^N \operatorname{sg}(u_{i+1/2}^n - h) (f(u_{i+1/2}^n) - f(k)) \psi^n qh \\ &\leq \sum_{n=0}^N \left\{ \operatorname{sg}(u_{i-1/2}^n - k) (f(u_{i-1/2}^n) - f(k)) \psi^n qh + |u_i^n - k| \frac{\psi^n - \psi^{n-1}}{qh} qh^2 \right\}. \end{aligned}$$

On recommence l'opération pour $i-1, i-2, \dots, \text{etc...}$ et il vient, de proche en proche, pour tout $i \in \{1, \dots, I-1\}$,

$$(17) \quad \sum_{n=0}^N \text{sg}(u_{i+1/2}^n - k) (f(u_{i+1/2}^n) - f(k)) \psi^n qh \\ \leq \sum_{n=0}^N \text{sg}(u_{1/2}^n - k) (f(u_{1/2}^n) - f(k)) \psi^n qh \\ + \sum_{n=0}^N \sum_{j=1}^i |u_j^n - k| \frac{\psi^n - \psi^{n-1}}{qh} qh^2 .$$

On définit \tilde{u}_h par $\tilde{u}_h(x, t) = u_{i+1/2}^n$ si $(x, t) \in I_i \times J_n$. Comme $\{u_h\}_m$ converge vers u , on sait que $\{\tilde{u}_h\}_m$ converge aussi vers u dans $L^1(\Omega \times]0, T[)$; on peut en extraire une sous-suite convergente presque partout sur $\Omega \times]0, T[$ vers u , et ainsi pour presque tout $x \in \Omega$, cette sous-suite encore notée $\tilde{u}_h(x, \cdot)$ converge vers $u(x, \cdot)$ presque partout sur $]0, T[$, d'après le théorème de FUBINI. De plus, $\tilde{u}_h(x, \cdot)$ reste bornée, et donc, d'après le théorème de LEBESGUE, pour presque tout $x \in \Omega$

$$(18) \quad \lim_{m \rightarrow \infty} \int_0^T |\tilde{u}_h(x, \cdot) - u(x, \cdot)|_{L^1(0, T)} = 0 .$$

Soit $x \in \Omega$, tel que (18) soit vérifié. On note i l'entier tel que $x \in I_i$, et on vérifie qu'alors le premier membre de (17) s'écrit

$$\int_0^T \text{sg}(\tilde{u}_h(x, t) - k) (f(\tilde{u}_h(x, t)) - f(k)) \psi(t) dt ,$$

qui converge vers

$$\int_0^T \text{sg}(u(x, t) - k) (f(u(x, t)) - f(k)) \psi(t) dt .$$

Le dernier terme de (17) peut être majoré ainsi

$$\left| \sum_n \sum_{j < i} |u_j^n - k| \frac{\psi^n - \psi^{n-1}}{qh} qh^2 \right| \leq (M + |k|) |\psi'|_{L^\infty T(ih_m)}$$

où ih_m converge vers x .

Pour estimer l'autre terme de (17), on remarque que pour tout $n < N$, on a

$$(19) \quad \text{sg}(u_{1/2}^n - k) (f(u_{1/2}^n) - f(k)) \leq \text{sg}(u_0^n - k) (f(u_{1/2}^n) - f(k)) .$$

En effet, si $\text{sg}(u_{1/2}^n - k) = \text{sg}(u_0^n - k)$, la démonstration est évidente. Elle l'est encore si $k = u_{1/2}^n$. Si $k \in I(u_{1/2}^n, u_0^n)$, alors, par construction de $u_{1/2}^n$, d'après (7),

$$\text{sg}(u_{1/2}^n - k) (f(u_{1/2}^n) - f(k)) \leq 0 ,$$

et (19) est vrai si $k = u_0^n$. Si $k \neq u_0^n$, $k \neq u_{1/2}^n$, alors

$$\text{sg}(u_{1/2}^n - k) = - \text{sg}(u_0^n - k)$$

si $k \in I(u_{1/2}^n, u_0^n)$; le premier membre de (19) est non positif, et le second est non négatif.

On définit maintenant sur $]0, T[$, pour tout m , la fonction β_m par

$$\beta_m(t) = u_{1/2}^n \quad \text{si} \quad t \in J_n ,$$

puis la fonction a_m par

$$a_m(t) = u_0^n \quad \text{si} \quad t \in J_n .$$

Il est évident que a_m converge vers a dans $L^1(0, T)$. La suite β_m est uniformément bornée, donc aussi la suite $f(\beta_m)$. On peut en extraire une sous-suite, encore notée $\{\beta_m\}_m$, qui converge vers une fonction $\lambda \in L^\infty(0, T)$ dans $L^\infty(0, T)$ faible *. Il vient alors

$$\sum_n \text{sg}(u_{1/2}^n - k) (f(u_{1/2}^n) - f(k)) \psi^n qh \\ \leq \int_0^T \text{sg}(a_m - k) (f(\beta_m) - f(k)) \psi^n(t) dt .$$

Pour passer à la limite sur $sg(a_m - k)$, on régularise et on utilise le lemme 2. Les passages à la limite sur les premiers et derniers termes de (17) restent inchangés, et il reste à passer à la limite sur

$$\begin{aligned} & \int_{\mathbb{R}} \frac{1}{2} sg'_\eta(\xi-k) \int_0^T sg(a_m - \xi) (f(\beta_m) - f(\xi)) \psi(t) dt \\ &= \int_{\mathbb{R}} \frac{1}{2} sg'_\eta(\xi-k) \int_0^T sg(a_m - \xi) f(\beta_m) \psi(t) dt \\ &- \int_{\mathbb{R}} \frac{1}{2} sg'_\eta(\xi-k) \int_0^T sg(a_m - \xi) f(\xi) \psi(t) dt . \end{aligned}$$

Le passage à la limite sur ce dernier terme est immédiat, d'après le lemme 2. L'autre terme vaut

$$\int_0^T f(\beta_m) \psi(t) sg_\eta(a_m(t)-k) dt ,$$

dont la limite en m est

$$\int_0^T \lambda(t) \psi(t) sg_\eta(a(t)-k) dt .$$

On passe maintenant à la limite en η en utilisant le théorème de LEBESGUE, et il vient, après passage à la limite, à partir de (17), en m puis en η , l'inégalité :

$$\begin{aligned} & \int_0^T sg u(x,t)-k (f(u(x,t))-f(k)) \psi(t) dt \\ & \leq \int_0^T sg(a(t)-k) (\lambda(t)-f(k)) \psi(t) dt + (M + |k|) |\psi'|_{L^\infty} T x \end{aligned}$$

On fait tendre x vers 0, en évitant l'ensemble de mesure nulle de $]0,1[$ pour lequel (18) n'est pas valable. Il vient, pour la trace de u en $x = 0$,

$$\int_0^T sg(\gamma u - k) (f(\gamma u) - f(k)) \psi(t) dt \leq \int_0^T sg(a(t)-k) (\lambda(t)-f(k)) \psi(t) dt .$$

En prenant successivement $k < -M$ et $k > M$, on obtient

$$(20) \quad \lambda(t) = f(\gamma u(0,t))$$

puis pour $k \in I(\gamma u(0,t), a(t))$

$$sg(\gamma u - a) (f(\gamma u) - f(k)) \leq 0 \quad \text{pp } (0,T)$$

dont on déduit (3) immédiatement.

La démonstration de (4) se fait de la même façon. Il reste à prouver que (IV.18) est équivalent à (3), (4) et (IV.18) avec ϕ nulle sur la frontière. On sait, depuis le théorème 11 (unicité) que (IV.18) implique (3) et (4).

Soit $\phi \in C^2_{0,+}([0,1] \times]0,T])$; on écrit (IV.18) avec une fonction test de la forme $\phi(1-\rho_\delta)$, donc nulle sur la frontière, puis on passe à la limite en δ , pour obtenir

$$\begin{aligned} & \int_{\Omega} \int_0^T [|u-k| \frac{\partial \phi}{\partial t} + sg(u-k) (f(u)-f(k)) \frac{\partial \phi}{\partial x}] dx dt \geq \\ (21) \quad & - \int_0^T sg(\gamma u(0,t)-k) (f(\gamma u(0,t)) - f(k)) \phi(0,t) dt \\ & + \int_0^T sg(\gamma u(1,t)-k) (f(\gamma u(1,t)) - f(k)) \phi(1,t) dt . \end{aligned}$$

On ajoute

$$\begin{aligned} & - \int_0^T sg(k-a(t)) (f(\gamma u(0,t)) - f(k)) \phi(0,t) dt \\ & + \int_0^T sg(k-b(t)) (f(\gamma u(1,t)) - f(k)) \phi(1,t) dt \end{aligned}$$

à chaque membre de l'inégalité (21), de façon à faire apparaître le pre-

mier membre de (IV.40). Il reste à prouver que le second membre, c'est-à-dire

$$- \int_0^T [\text{sg}(\gamma u(0,t)-k) + \text{sg}(k-a(t))] [f(\gamma u(0,t))-f(k)] \phi(0,t) dt + \int_0^T [\text{sg}(\gamma u(1,t)-k) + \text{sg}(k-b(t))] [f(\gamma u(1,t))-f(k)] \phi(1,t) dt.$$

est non négatif. Il s'agit de la somme de deux termes, dont le premier est l'intégrale sur]0,T[de la fonction

$$- [\text{sg}(\gamma u(0,t)-k) + \text{sg}(k-a(t))] [f(\gamma u(0,t))-f(k)] \phi(0,t).$$

Si $k \notin I(\gamma u(0,t), a(t))$, cette expression est nulle, et elle est non négative si $k \in I(\gamma u(0,t), a(t))$, d'après (3). Il en est de même de l'autre terme, et on obtient ainsi (IV.40). Le théorème 12 est démontré.

2. LE SCHEMA (III.48) (GODUNOV QUASI D'ORDRE DEUX)

On se propose d'étendre la conclusion du théorème 12 au schéma quasi d'ordre deux construit à partir du schéma de GODUNOV. Ce schéma est décrit en (III.48), et sa convergence est démontrée au théorème 9, dans le cadre du problème de CAUCHY ; il s'agit du schéma suivant

$$(22.i) \quad u_{i+1/2}^n \in I(u_i^n, u_{i+1}^n) \quad \text{réalise} \quad \min_{k \in I(u_i^n, u_{i+1}^n)} \left\{ \text{sg}(u_{i+1}^n - u_i^n) f(k) \right\}$$

$$(22.ii) \quad a_{i+1/2}^n = \frac{q}{2} \left[f(u_{i+1}^n) - 2f(u_{i+1/2}^n) + f(u_i^n) - qf'(\xi_{i+1/2}^n) (f(u_{i+1}^n) - f(u_i^n)) \right]$$

$$(22.iii) \quad a_{i+1/2}^n = \text{sg}(u_{i+1}^n - u_i^n) \text{Max} \left\{ 0, \text{Min} \left[a_{i+1/2}^n, C h^\alpha, \text{sg}(u_{i+2}^n - u_{i+1}^n) \text{sg}(u_{i+1}^n - u_i^n) q |f(u_{i+3/2}^n) - f(u_{i+1/2}^n)|, \text{sg}(u_i^n - u_{i-1}^n) \text{sg}(u_{i+1}^n - u_i^n) q |f(u_{i+1/2}^n) - f(u_{i-1/2}^n)| \right] \right\}$$

$$(22.iv) \quad u_i^{n+1} = u_i^n - q(f(u_{i+1/2}^n) - f(u_{i-1/2}^n)) - a_{i+1/2}^n + a_{i-1/2}^n,$$

avec $\alpha \in]0,1[$ et $C > 0$ arbitraires dans (22.iii). On a le résultat de convergence suivant.

Théorème 13 : Sous la condition de stabilité

$$(23) \quad q \sup_{|u| \leq M} |f'(u)| \leq \frac{1}{2}$$

$$\text{avec } M = \text{Max} \left\{ \|u_0\|_{L^\infty(0,1)}, \|a\|_{L^\infty(0,T)}, \|b\|_{L^\infty(0,T)} \right\},$$

et si la condition initiale u_0 et les conditions aux limites a et b admettent un nombre fini de changements de monotonie, la famille $\{u_h\}_{h>0}$ des solutions approchées construites par le schéma (22), converge lorsque h tend vers zéro, vers u , solution du problème (1), (2), (3), (4) et caractérisée par la définition 3, dans $L^1(\Omega \times]0,T[)$.

x x
x

Démonstration

En reprenant les démonstrations du théorème 12 et du théorème 9 on établit que la famille $\{u_h\}_{h>0}$ contient une suite $\{u_{h_m}\}_{m \in \mathbb{N}}$ convergente dans $L^1(\Omega \times]0,T[)$, vers une fonction $u \in BV(\Omega \times]0,T[)$, et qui satisfait à (IV.19) et à (IV.18) pour une fonction test ϕ nulle en $x = 0$ et en $x = 1$ pour tout $t \in]0,T[$. L'hypothèse du nombre fini de changements de monotonie faite sur a et sur b intervient au même titre que celle faite sur u_0 , pour

assurer que pour tout $t \in]0, T[$, la solution approchée $u_h(\cdot, t)$ admet un nombre de changements de monotonie, qui est fini et indépendant de h .

Il reste à établir (3) et (4). Soit $k \in \mathbb{R}$ et $i \in \{1, \dots, I-1\}$, on a d'après (III.51), l'inégalité suivante, avec $\psi \in C^1_{0,+}([0, T])$

$$\begin{aligned} & \sum_n (\text{sg}(u_{i+1/2}^n - k) (f(u_{i+1/2}^n) - f(k)) + \beta_{i+1/2}^n) \psi^n qh \\ & \leq \sum_n (\text{sg}(u_{i-1/2}^n - k) (f(u_{i-1/2}^n) - f(k)) + \beta_{i-1/2}^n) \psi^n qh \\ & \quad + \sum_n \frac{|u_i^n - k| - |u_i^{n+1} - k|}{qh} qh^2 \psi^n \\ & \quad + 2 qh \sum_n |a_{i+1/2}^n| (S_{i+1}^n - S_i^n) \psi^n, \end{aligned}$$

où $\beta_{i+1/2}^n$ est défini par

$$\beta_{i+1/2}^n = u_{i+1/2}^n a_{i+1/2}^n \text{sg}(u_{i+3/2}^n - u_{i+1/2}^n) (1 - 2 S_{i+1}^n),$$

et ψ^n est la projection L^2 de ψ sur J_n .

On obtient de proche en proche une inégalité analogue à (17),

$$\begin{aligned} & \sum_n (\text{sg}(u_{i+1/2}^n - k) (f(u_{i+1/2}^n) - f(k)) + \beta_{i+1/2}^n) \psi^n qh \\ & \leq \sum_n (\text{sg}(u_{1/2}^n - k) (f(u_{1/2}^n) - f(k)) + \beta_{1/2}^n) \psi^n qh \\ & \quad + \sum_n \sum_{j \leq i} 2 qh |a_{i+1/2}^n| (S_{i+1}^n - S_i^n) \psi^n + |u_i^n - k| \frac{\psi_i^n - \psi_i^{n-1}}{qh} qh^2. \end{aligned}$$

Ce dernier terme peut être majoré par (K = nombre de changements de monotonie)

$$2 |\psi|_{L^1(0, T)} K C h^\alpha + (M + |k|) |\psi'|_{L^\infty T(ih)}.$$

On a également pour tout j

$$\left| \sum_n \beta_{j+1/2}^n \psi^n qh \right| \leq C h^\alpha |\psi|_{L^1(0, T)}.$$

On passe à la limite en $h \in \{h_m\}$, après s'être donné $x \in]0, 1[$, où la suite \tilde{u}_h^n construite à partir de $u_{i+1/2}^n$ converge vers $u(x, \cdot)$ dans $L^1(0, T)$, avec $x \in I_1$ (cf. (18)), et après avoir utilisé (19) et introduit la limite dans L^∞ faible * d'une sous-suite construite à partir des $f(u_{1/2}^n)$. On obtient, à la limite, pour presque tout $x \in]0, 1[$, et après régularisation,

$$\begin{aligned} & \int_0^T \text{sg}(u(x, t) - k) (f(u(x, t)) - f(k)) \psi(t) dt \\ & \leq \int_0^T \text{sg}(a(t) - k) (\lambda(t) - f(k)) \psi(t) dt + T (M + |k|) |\psi'|_{L^\infty x}. \end{aligned}$$

On fait tendre ensuite x vers zéro, et on obtient (20); le reste de la démonstration est identique à celle du théorème 12. Le théorème 13 est démontré.

3. LE SCHEMA DE LAX, A PLUSIEURS DIMENSIONS

On reprend l'ouvert Ω de \mathbb{R}^p , où on considère le problème

$$(24) \quad \frac{\partial u}{\partial t} + \text{div} [f(u)] = 0 \quad \text{si } (x, t) \in \Omega \times]0, T[,$$

$$(25) \quad u(x, 0) = u_0(x) \quad \text{si } x \in \Omega ,$$

$$(26) \quad \text{Min}_{k \in I(\gamma u, 0)} \left\{ \text{sg}(\gamma u) (f(\gamma u) - f(k)) \cdot \nu \right\} = 0 \quad \text{pp } (\Gamma \times]0, T[),$$

c'est-à-dire que la condition aux limites est prise identiquement nulle.

On reprend le maillage utilisé au chapitre III, en notant

$$(27) \quad \begin{aligned} \mathcal{J}_h &= \left\{ i \in \mathbb{Z}^p, \quad I_i \subset \Omega \right\} \\ \Gamma_h &= \left\{ i \in \mathbb{Z}^p, \quad I_i \cap \Gamma \neq \emptyset \text{ et } \exists j \in \mathcal{J}_h \right. \\ &\quad \left. I_i \text{ et } I_j \text{ ont une face commune} \right\}. \end{aligned}$$

On suppose, ce qui n'est pas restrictif, que $f(0)$ est nul.

On considère le schéma de LAX, c'est-à-dire

$$(28) \quad \begin{aligned} u_i^{n+1} &= 0 \quad \text{si } i \in \Gamma_h \\ u_i^{n+1} &= \sum_{j=1}^p (u_{i+1j}^n + u_{i-1j}^n) - \frac{q_j}{2} [f_j(u_{i+1j}^n) - f_j(u_{i-1j}^n)] \\ &\quad \text{si } i \in \mathcal{J}_h. \end{aligned}$$

Pour simplifier la démonstration des estimations a priori,

on suppose que l'intersection de Ω par toute droite parallèle à l'une des vecteurs de base est connexe. Cette hypothèse est réalisée par exemple si Ω est convexe. Le résultat de convergence est le suivant.

théorème 14 : Si la condition de stabilité de COURANT FRIEDRICHS LEWY

$$(29) \quad \forall j \in \{1 \dots p\} \quad q_j \sup_{|u| \leq |u_0|_{L^\infty}} |f'_j(u)| \leq \frac{1}{p},$$

est satisfaite, la famille $\{u_h\}$ des solutions approchées construites par le schéma de LAX (28) converge dans $L^1(\Omega \times]0, T[)$ vers une solution du problème (24), (25), (26) et caractérisée par la définition 3.

x x
x

Démonstration

Soit $i \in \mathcal{J}_h$, $n < N$; on a, comme pour le problème de CAUCHY, la stabilité de la norme uniforme (cf. CONWAY-SMOLLER, ou LE ROUX [1])

$$(30) \quad |u_i^{n+1}| \leq \sup_{|i-j| \leq 1} |u_j^n|$$

On convient, pour alléger les notations, de prolonger u_h par zéro à l'extérieur de Ω , c'est-à-dire de poser

$$u_i^n = 0 \quad \text{si } i \notin \mathcal{J}_h.$$

On a également, comme pour le problème de CAUCHY, l'inégalité

$$|u_i^{n+1} - u_i^n| \leq \frac{1}{p} \sum_{j=1}^p |u_{i+1j}^n - u_i^n|,$$

qui exprime la conservation de la variation en temps.

On considère maintenant la variation en espace. Soit $\ell \in \{1 \dots p\}$, $i \in \mathcal{J}_h$. On a, si $i+1 \in \mathcal{J}_h$, la majoration

$$(31) \quad \begin{aligned} |u_{i+1\ell}^{n+1} - u_i^{n+1}| &\leq \frac{1}{2p} \sum_{j=1}^p |u_{i+1j+1\ell}^n - u_{i+1j\ell}^n| (1-p q_j f'_j(\xi_{i+1j+1\ell/2}^n, j)) \\ &\quad + |u_{i-1j+1\ell}^n - u_{i-1j\ell}^n| (1+p q_j f'_j(\xi_{i-1j+1\ell/2}^n, j)), \end{aligned}$$

où $\xi_{i+1\ell/2}^n$ représente un point de l'intervalle $I(u_{i+1\ell}^n, u_{i+1\ell}^n)$ tel que

$$f_j(u_{i+1\ell}^n) - f_j(u_i^n) = f'_j(\xi_{i+1\ell/2}^n, j) (u_{i+1\ell}^n - u_i^n).$$

On veut une majoration analogue pour les i voisins de Γ_h . Si par exemple $i \in \mathcal{J}_h$, $i+1 \in \Gamma_h$, on a $u_{i+1\ell}^{n+1} = 0$, donc

$$|u_{i+1\ell}^{n+1} - u_i^{n+1}| = |u_i^{n+1}| = u_i^n - \sum_{j=1}^p \frac{q_j}{2} (f_j(u_{i+1j}^n) - f_j(u_{i-1j}^n)) +$$

$$+ \frac{1}{2p} (u_{i+1_j}^n - u_i^n) - \frac{1}{2p} (u_i^n - u_{i-1_j}^n) \Big|$$

qu'on majore par

$$\frac{1}{p} \sum_{j=1}^p \left| u_i^n - \frac{pq_j}{2} (f_j(u_{i+1_j}^n) - f_j(u_{i-1_j}^n)) + \frac{1}{2} (u_{i+1_j}^n - u_i^n) - (u_i^n - u_{i-1_j}^n) \right|$$

Dans cette somme, le terme en $j = \ell$ peut être majoré comme pour un problème à une dimension, sachant que $u_{i+1_\ell}^n$ est nul. On obtient comme majorant :

$$|u_{i+1_\ell}^n - u_i^n| \left(\frac{1}{2} - \frac{pq_\ell}{2} f'_\ell(\xi_{i+1_{\ell/2}}^n) \right) + |u_i^n - u_{i-1_\ell}^n| \left(\frac{1}{2} + \frac{pq_\ell}{2} f'_\ell(\xi_{i-1_{\ell/2}}^n) \right),$$

ce qui correspond au second membre de (31).

Pour un terme en $j \neq \ell$, on est assuré, d'après l'hypothèse faite sur Ω , qu'il existe $\mu_1 \in \mathbb{N}$, $\mu_2 \in \mathbb{Z}$ tels que

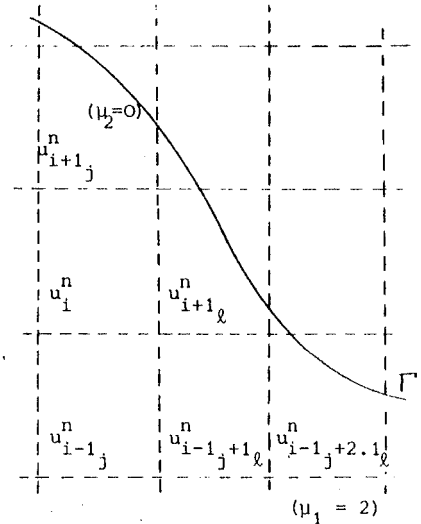
$$i + 1_j + \mu_1 1_\ell \in \Gamma_h \quad ; \quad i - 1_j + \mu_2 1_\ell \in \Gamma_h.$$

On écrit alors ce terme en $j \neq \ell$ de la façon suivante

$$\begin{aligned} & u_{i+1_j}^n + u_{i-1_j}^n - pq_j (f_j(u_{i+1_j}^n) - f_j(u_{i-1_j}^n)) \\ = & u_{i+1_j}^n - u_{i+1_{j+1_\ell}}^n + u_{i+1_{j+1_\ell}}^n - \dots - u_{i+1_{j+\mu_1 1_\ell}}^n \\ & + u_{i-1_j}^n - u_{i-1_{j+1_\ell}}^n + u_{i-1_{j+1_\ell}}^n - \dots - u_{i-1_{j+\mu_2 1_\ell}}^n \\ (32) \quad & - pq_j (f_j(u_{i+1_j}^n) - f_j(u_{i+1_{j+1_\ell}}^n)) + \dots - f_j(u_{i+1_{j+\mu_1 1_\ell}}^n) \\ & + pq_j (f_j(u_{i-1_j}^n) - f_j(u_{i-1_{j+1_\ell}}^n)) + \dots - f_j(u_{i-1_{j+\mu_2 1_\ell}}^n) \\ = & \sum_{k=0}^{\mu_1} \left[u_{i+1_{j+k \cdot 1_\ell}}^n - u_{i+1_{j+(k+1) 1_\ell}}^n \right] \left[1 - pq_j (f'_j(\xi_{i+1_{j+k \cdot 1_\ell}^n})) \right] \\ & + \sum_{k=0}^{\mu_2} \left[u_{i-1_{j+k \cdot 1_\ell}}^n - u_{i-1_{j+(k+1) 1_\ell}}^n \right] \left[1 + pq_j (f'_j(\xi_{i-1_{j+k \cdot 1_\ell}^n})) \right], \end{aligned}$$

où les coefficients sont non négatifs. On passe aux valeurs absolues et on obtient une majoration analogue à (31), où interviennent des sommes de μ_1 termes en $i+1_j$ et de μ_2 termes en $i-1_j$; ces sommes n'interviennent que lorsqu'elles ne peuvent pas intervenir dans le cadre de (31), en provenant d'une majoration de différences $|u_{i+1_\ell}^{n+1} - u_i^{n+1}|$ telles que $i \in \Gamma_h$ et $i+1_\ell \in \mathcal{J}_h$, car elles correspondent à des pavés adjacents à la frontière approchée, c'est-à-dire la réunion des pavés I_i , où i appartient à Γ_h . On peut remarquer également que μ_1 et μ_2 sont de signe contraire, et qu'on ne l'a pas utilisé ici.

Si maintenant $i \in \Gamma_h$ et $i+1_\ell \in \mathcal{J}_h$, on a une estimation analogue à (31), qui provient d'un regroupement de termes semblable à (32), après avoir fait apparaître les différences en $(u_{i+1_{j+k \cdot 1_\ell}}^n - u_{i+1_{j+(k+1) 1_\ell}}^n)$ jusqu'à la frontière.



On ajoute toutes ces estimations, et on obtient la majoration suivante, qui exprime la variation bornée en espace au temps $(n+1)\Delta t$ en fonction de la même au temps $n \Delta t$, dans la direction 1_ℓ

$$\sum_{i \in \mathcal{J}_h \cup \Gamma_h} |u_{i+1_\ell}^{n+1} - u_i^{n+1}| \leq \sum_{i \in \mathcal{J}_h \cup \Gamma_h} |u_{i+1_\ell}^n - u_i^n| ;$$

tous les termes en f'_j disparaissent, puisqu'ils apparaissent chacun deux fois, avec des signes opposés.

On en déduit, comme pour le problème de CAUCHY, que la famille $\{u_h\}_h$ contient une suite $\{u_h\}$ convergente dans $L^1(\Omega \times]0, T[)$ vers une fonction $u \in BV(\Omega \times]0, T[)$, m lorsque m tend vers l'infini, et u vérifie (IV.19). Comme au théorème 5 (cf. LE ROUX [1]), on vérifie que u satisfait

V.18) pour une fonction ϕ nulle sur $\Gamma \times]0, T[$. Ceci provient de égalité

$$|u_i^{n+1} - k| \leq \frac{1}{2p} \sum_{j=1}^p |u_{i+1_j}^n - k| + |u_{i-1_j}^n - k|$$

$$+ \text{sg}(u_{i+1_j}^n - k) (f_j(u_{i+1_j}^n) - f(k)) + p q_j \text{sg}(u_{i-1_j}^n - k) (f_j(u_{i-1_j}^n) - f(k))$$

se démontre de façon immédiate.

On se propose maintenant de démontrer (IV.13). On va considérer un premier temps le problème à une seule dimension, pour $\Omega =]0, 1[$, généraliser la technique utilisée au problème à plusieurs dimensions.

A une dimension, on procède comme au théorème 12, ou du moins façon analogue, car il est plus difficile d'établir un résultat de convergence de $u_h(x, \cdot)$ vers $u(x, \cdot)$ dans $L^1(0, T)$ pour presque tout x , où il s'agit des différences de la forme $|u_{i+1}^n - k| - |u_i^n - k|$. On introduit une fonction test $\phi \in C_{0+}^2([0, 1[\times]0, T[)$, de la forme

$$\phi(x, t) = \begin{cases} 0 & , \text{ si } x > 1/2 \\ \rho_\delta(x) \psi(t) & , \text{ si } x < 1/2 \end{cases}$$

ρ_δ est la fonction introduite en (IV.33). On effectue une projection de type L^2 de ϕ , sur chaque $I_i \times J_n$, et on multiplie l'inégalité

$$|u_i^{n+1} - k| \leq |u_i^n - k| + \frac{1}{2} (|u_{i+1}^n - k| - 2|u_i^n - k| + |u_{i-1}^n - k|)$$

$$\text{sg}(u_{i+1}^n - k) (f(u_{i+1}^n) - f(k)) + \frac{q}{2} \text{sg}(u_{i-1}^n - k) (f(u_{i-1}^n) - f(k))$$

en ϕ_i^n , puis on somme en i et n ; il vient après quelques regroupements

$$0 \leq \sum_{n,i} \frac{1}{2h} (\phi_{i+1}^n - \phi_{i-1}^n) \text{sg}(u_i^n - k) (f(u_i^n) - f(k)) qh^2$$

$$+ \sum_{n,i} \frac{1}{qh} |u_i^n - k| (\phi_i^n - \phi_i^{n-1}) qh^2 +$$

$$(34) \quad \begin{aligned} & + \sum_{n,i} \frac{1}{2q} |u_i^n - k| \frac{1}{h^2} (\phi_{i+1}^n - 2\phi_i^n + \phi_{i-1}^n) qh^3 \\ & + \sum_n \frac{1}{2} \phi_1^n \left[\text{sg}(u_0^n - k) (f(u_0^n) - f(k)) + \text{sg}(u_1^n - k) (f(u_1^n) - f(k)) \right] qh \\ & + \sum_n \frac{1}{2q} \phi_1^n (|u_0^n - k| - |u_1^n - k|) qh + C |\phi'|_{L^\infty T(M+|k|)} h \end{aligned}$$

On regroupe ces deux derniers termes; on a pour tout $k \in \mathbb{R}$

$$(35) \quad \begin{aligned} & \text{sg}(u_0^n - k) (qf(u_0^n) - qf(k)) + \text{sg}(u_1^n - k) (qf(u_1^n) - f(k)) \\ & + |u_0^n - k| - |u_1^n - k| \leq \text{sg}(u_0^n - k) (qf(u_0^n) - 2qf(k) + qf(u_1^n) + u_0^n - u_1^n) \end{aligned}$$

En effet, cette inégalité est satisfaite à l'égalité si $k \in I(u_1^n, u_0^n)$ puisque dans ce cas $\text{sg}(u_0^n - k)$ et $\text{sg}(u_1^n - k)$ sont égaux. Elle est encore vraie pour $k = u_1^n$, de façon évidente. Si $k = u_0^n$ ou si k appartient à l'intérieur de l'intervalle $I(u_0^n, u_1^n)$, on remarque, compte-tenu de la condition de stabilité, que

$$\text{sg}(u_1^n - k) (qf(u_1^n) - qf(k)) - |u_1^n - k| \leq 0$$

Or $\text{sg}(u_0^n - k)$ vaut soit zéro, soit $-\text{sg}(u_1^n - k)$. On a donc

$$\text{sg}(u_0^n - k) (qf(u_1^n) - qf(k) - u_1^n + k) \geq 0, \text{ d'où (35).}$$

On peut maintenant passer à la limite sur chaque terme de (34), en utilisant (35) pour majorer la somme des deux derniers termes, et en notant comme au théorème 12, $\lambda(t)$ la limite dans L^∞ faible * sur $]0, T[$ de la suite β_m définie sur $]0, T[$ par

$$\beta_m(t) = f(u_1^n) - \frac{u_1^n}{q} \quad \text{si } t \in J_n$$

On obtient à la limite, en rappelant que u_0^n est nul pour

tout n

$$\int_{\Omega} \int_0^T \frac{\partial \phi}{\partial x} \operatorname{sg}(u-k) (f(u)-f(k)) \, dx \, dt \leq \int_{\Omega} \int_0^T \frac{\partial \phi}{\partial t} |u-k| \, dx \, dt + \int_0^T \frac{1}{2} \left\{ \operatorname{sg}(k) f(k) + \operatorname{sg}(k) (f(k) - \lambda(t)) \right\} \psi(t) \, dt,$$

car le terme

$$\sum_{n,i} \frac{1}{2q} |u_1^n - k| \frac{1}{h} (\phi_{i+1}^n - 2\phi_i^n + \phi_{i-1}^n) \, qh^3$$

peut être majoré par $\frac{T}{2q} \left| \frac{\partial^2 \phi}{\partial x^2} \right|_{L^\infty} (M+|k|) h$, et tend vers zéro avec h .

Or $\phi(x,t) = \psi(t) \rho_\delta(x)$; on fait tendre δ vers zéro, et il vient, pour presque tout $t \in]0, T[$

$$(36) \quad \operatorname{sg}(\gamma u(0,t) - k) (f(\gamma u(0,t)) - f(k)) \leq \operatorname{sg}(k) \left\{ f(k) - \frac{\lambda(t)}{2} \right\}.$$

En prenant $k > \operatorname{Max}(\gamma u(0,t), 0)$ puis $k < \operatorname{Min}(\gamma u(0,t), 0)$,

on obtient

$$\lambda(t) = 2 f(\gamma u(0,t)) \quad \text{pp }]0, t[$$

c'est-à-dire que la suite construite à partir des

$$\frac{1}{2} \left(f(u_0^n) + \frac{u_0^n}{q} + f(u_1^n) - \frac{u_1^n}{q} \right)$$

converge vers $f(\gamma u(0,t))$. Ce dernier résultat est conservé même lorsque la condition aux limites $a(t)$, c'est-à-dire les u_0^n , est non nécessairement nulle. On obtient donc, à partir de (36), que pour $k \in I(\gamma u(0,t), 0)$,

$$\operatorname{sg}(u(0,t) - 0) (f(\gamma u(0,t)) - f(k)) \quad \text{pp } (]0, T[),$$

c'est-à-dire (3). On obtient (4) de la même façon. Ayant également (IV.18) lorsque la fonction test ϕ est nulle sur $\Gamma \times]0, T[$, on en déduit comme au théorème 12 que (IV.18) est vraie.

On revient maintenant au problème multidimensionnel. On introduit la fonction test $\phi \in C_{0,+}^2(\bar{\Omega} \times]0, T[)$, telle que

$$\phi(x,t) = \psi(t, x) \rho_\delta(x),$$

qu'on approche par projection L^2 sur chaque $I_i \times J_n$, puis on multiplie (33) par $h_1 \dots h_p \phi_i^n$ et on somme en $i \in J_h^n$, $n < N$. Il vient

$$(37) \quad \begin{aligned} & \sum_{i \in J_h^n} \sum_n \left[|u_i^{n+1} - k| - |u_i^n - k| \right] \phi_i^n h_1 \dots h_p \\ & \leq \frac{1}{2p} \sum_{j=1}^p \sum_{i \in J_h^n} (|u_{i+1_j}^n - k| - 2|u_i^n - k| + |u_{i-1_j}^n - k|) \phi_i^n h_1 \dots h_p \\ & + \frac{1}{2p} \sum_{j=1}^p \sum_{i \in J_h^n} \left(-pq_j (f_j(u_{i+1_j}^n) - f_j(k)) \operatorname{sg}(u_{i+1_j}^n - k) \right. \\ & \quad \left. + pq_j (f_j(u_{i-1_j}^n) - f_j(k)) \operatorname{sg}(u_{i-1_j}^n - k) \right) \phi_i^n h_1 \dots h_p. \end{aligned}$$

Pour j fixé, on note $i_{0,j}$ et $i_{1,j}$ les indices de Γ_h tels que $(i_{1,j} - i_{0,j})$ soit un nombre entier, non négatif, de fois le vecteur i_j . On a alors

$$\begin{aligned} & \sum_{j=1}^p \sum_{i \in J_h^n} (|u_{i+1_j}^n - k| - 2|u_i^n - k| + |u_{i-1_j}^n - k|) \phi_i^n h_1 \dots h_p \\ & = \sum_{j=1}^p \sum_{i_1 \dots i_{j-1}} \sum_{i_{j+1} \dots i_p} \sum_{\substack{k=i_{0,j} \\ (i_j = k)}}^{i_{1,j}-1} (|u_{i+1_j}^n - k| - 2|u_i^n - k| + |u_{i-1_j}^n - k|) \phi_i^n h_1 \dots h_p \\ & = \sum_j \sum_{(i_j \text{ fixé})} \sum_{\substack{k=i_{0,j} \\ (i_j = k)}}^{i_{1,j}-1} |u_i^n - k| (\phi_{i+1_j}^n - 2\phi_i^n + \phi_{i-1_j}^n) \frac{1}{h_j} \Delta t h_1 \dots h_p \cdot \frac{h_j}{q_j} \\ & + \sum_j \sum_{(i_j \text{ fixé})} (|u_{i_{1,j}}^n - k| - |u_{i_{1,j}-1}^n - k|) \phi_{i_{1,j}-1}^n h_1 \dots h_p + \end{aligned}$$

$$+ \sum_j \sum_{(i_j \text{ fixé})} (|u_{i_o, j}^n - k| - |u_{i_o, j+1}^n - k|) \phi_{i_o, j+1}^n h_1 \dots h_p$$

quel s'ajoute un terme majoré par

$$C |\text{grad } \phi| T |h| (M + |k|),$$

qui tend vers zéro avec h.

De la même façon,

$$\sum_{i \in J_h} \left\{ -pq_j \text{sg}(u_{i+1_j}^n - k) (f_j(u_{i+1_j}^n) - f_j(k)) + pq_j \text{sg}(u_{i-1_j}^n - k) (f_j(u_{i-1_j}^n) - f_j(k)) \right\}$$

$$\phi_i^n h_1 \dots h_p$$

$$\sum_{i=1} \sum_{(i_j \text{ fixé})} \sum_k p \text{sg}(u_i^n - k) (f_j(u_i^n) - f_j(k)) \frac{\phi_{i+1_j}^n - \phi_{i-1_j}^n}{2 h_j} 2 \Delta t h_1 \dots h_p$$

$$\sum_{i=1} \sum_{(i_j \text{ fixé})} -p q_j \left\{ \text{sg}(u_{i-1_j}^n - k) (f_j(u_{i-1_j}^n) - f_j(k)) \right.$$

$$\left. + \text{sg}(u_{i+1_j}^n - k) (f_j(u_{i+1_j}^n) - f_j(k)) \right\} \phi_{i+1_j}^n h_1 \dots h_p$$

$$q_j \left\{ \text{sg}(u_{i_o, j}^n - k) (f_j(u_{i_o, j}^n) - f_j(k)) + \text{sg}(u_{i_o, j+1}^n - k) (f_j(u_{i_o, j+1}^n) - f_j(k)) \right\}$$

$$\phi_{i_o, j}^n h_1 \dots h_p$$

quel on doit ajouter un terme majoré par (38).

De façon analogue à (35), on démontre que pour tout $i_{o, j} \in \Gamma_h$, $c j \in \{1, \dots, p\}$,

$$\text{sg}(u_{i_o, j}^n - k) (q_j f_j(u_{i_o, j}^n) - q_j f_j(k)) + |u_{i_o, j}^n - k|$$

$$+ \text{sg}(u_{i_o, j+1}^n - k) (q_j f_j(u_{i_o, j+1}^n) - q_j f_j(k)) - |u_{i_o, j+1}^n - k|$$

$$\text{sg}(u_{i_o, j}^n - k) (q_j f_j(u_{i_o, j}^n) - 2 q_j f_j(k) + q_j f_j(u_{i_o, j+1}^n) + u_{i_o, j}^n - u_{i_o, j+1}^n)$$

et pour tout $i_{1, j} \in \Gamma_h$, $j \in \{1 \dots p\}$

$$\text{sg}(u_{i_{1, j}}^n - k) (q_j f_j(u_{i_{1, j}}^n) - q_j f_j(k)) - |u_{i_{1, j}}^n - k|$$

$$(40) \quad + \text{sg}(u_{i_{1, j-1}}^n - k) (q_j f_j(u_{i_{1, j-1}}^n) - q_j f_j(k)) + |u_{i_{1, j-1}}^n - k|$$

$$\geq \text{sg}(u_{i_{1, j}}^n - k) (q_j f_j(u_{i_{1, j}}^n) - 2 q_j f_j(k) + q_j f_j(u_{i_{1, j-1}}^n) + u_{i_{1, j}}^n - u_{i_{1, j-1}}^n).$$

On est maintenant amené à sommer (39) et (40) en n et i (i_j fixé) c'est-à-dire à sommer dans un premier temps,

$$(41) \quad \sum_j \sum_n \sum_{i \in \Gamma_h} p \text{sg}(u_{i_o, j}^n - k) \left(q_j f_j(u_{i_o, j}^n) - 2 q_j f_j(k) \right. \\ \left. + q_j f_j(u_{i_o, j+1}^n) + u_{i_o, j}^n - u_{i_o, j+1}^n \right) \phi_{i_o, j}^n h_1 \dots h_p$$

et ensuite un terme correspondant en $i_{1, j}$, affecté du signe moins.

On note $\lambda_j(t, \sigma)$ la limite dans $L^\infty[0, T] \times \Gamma$ faible * de la suite β_m définie par

$$\beta_m(t, \sigma) = f_j(u_{i_o, j+1}^n) - \frac{1}{q_j} \cdot u_{i_o, j+1}^n \quad \text{si } \sigma \in I_{i_o, j},$$

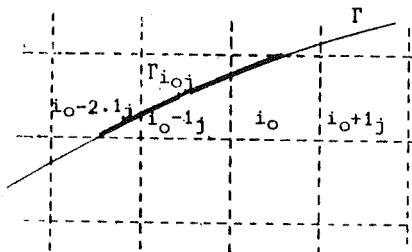
$$\beta_m(t, \sigma) = f_j(u_{i_{1, j}}^n - 1) - \frac{1}{q_j} \cdot u_{i_{1, j-1}}^n \quad \text{si } \sigma \in I_{i_{1, j}}.$$

Il reste à faire apparaître l'élément de frontière dans (41), avant de passer à la limite en h; pour cela on introduit la normale extérieure "approchée", qu'on définit ainsi.

Soit $j \in \{1 \dots p\}$. Etant donné $i_{o, j} \in \Gamma_h$, on note $\gamma_{i_{o, j}}$ la mesure de LEBESGUE superficielle de l'ensemble

$$\Gamma \cap \left\{ \bigcup_{i_j \leq i_{o, j}} I_i \right\} = \Gamma_{i_{o, j}}.$$

Cet ensemble représente l'intersection de la frontière Γ par une réunion de pavés I_i "parallèle" à la direction i_j . On note ensuite



$$v_{i_o,j} = - \prod_{k \neq j} h_k (\gamma_{i_o,j})^{-1}$$

La normale extérieure "approchée" est le vecteur dont les composantes sur $\Gamma \cap I_o$ sont les $v_{i_o,j}$, mais on n'utilisera ici que les composantes séparément, la sommation en j dans (41) étant effectuée la dernière.

Pour $i_{1,j} \in \Gamma_h$, on construit $v_{i_{1,j}}$ de la même façon, en l'affectant du signe +. Pour un indice $i \in \Gamma_h$, tel que i ne soit ni de la forme i_o , ni de la forme i_1 , on pose $v_{i,j} = 0$, ce qui arrive lorsque les tangentes à la frontière sont parallèles à la direction i_j .

Si v_j est la composante le long de la direction i_j de la normale extérieure, il y a convergence forte, dans $L^1(\Gamma)$ de la famille $\{v_{j,k}\}$ définie par

$$v_{j,k}(\sigma) = v_{i_o,j} \quad \text{si} \quad \sigma \in \Gamma_{i_o,j}$$

vers v_j . En effet, les $v_{j,h}$ sont uniformément bornés par un, et la frontière est supposée régulière : $v_{j,h}$ converge en tout point de Γ vers v_j . on conclut qu'il y a convergence dans $L^1(\Gamma)$ par le théorème de LEBESGUE.

On obtient, dans (41), sachant que $u_{i_o,j}^n$ est nul, et dans l'inégalité analogue en $i_{1,j}$, sachant que $u_{i_{1,j}}^n$ est nul,

$$\begin{aligned} & \sum_j \sum_n \left\{ \sum_{\substack{i \in \Gamma_h \\ i=i_{o,j}}} p \operatorname{sg}(k) (2q_j f_j(k) - q_j f_j(u_{i_{o,j+1}}^n + u_{i_{o,j+1}}^n)) \phi_{i_{o,j}}^n h_1 \dots h_p \right. \\ & \quad \left. - \sum_{\substack{i \in \Gamma_h \\ i=i_{o,j}}} p \operatorname{sg}(k) (2q_j f_j(k) - q_j f_j(u_{i_{1,j-1}}^n + u_{i_{1,j-1}}^n)) \phi_{i_{1,j}}^n h_1 \dots h_p \right\} \\ & = \sum_j \sum_n \left\{ \sum_{\substack{i \in \Gamma_h \\ (i=i_{o,j} \text{ ou } i_{1,j})}} p \operatorname{sg}(k) \left[-2 f_j(k) + \beta_m(t, \sigma) \right] \phi_i^n(q_j h_j) v_{i,j} \gamma_{i,j} \right\} \end{aligned}$$

qui converge vers

$$\sum_{j=1}^p \int_0^T \int_{\Gamma} p \operatorname{sg}(k) (-2 f_j(k) + \lambda_j(t, \sigma)) \psi(t) \rho_\delta(\sigma) v_j \, d\sigma \, dt.$$

On obtient, en passant à la limite sur les autres termes provenant de (37), qui se traitent comme dans le cas monodimensionnel

$$\begin{aligned} & \int_{\Omega} \int_0^T \operatorname{sg}(u-k) (f(u)-f(k)) \cdot \operatorname{grad}(\rho_\delta \psi(t, x)) \, dx \, dt \\ & + \int_0^T \int_{\Gamma} \operatorname{sg}(k) \left(\frac{\lambda(t, \sigma)}{2} - f(k) \right) \cdot v \, \psi(t, \sigma) \, d\sigma \, dt \\ & + C T \left[M + |k| \right] \left| \frac{\partial \psi}{\partial t} \right|_{L^\infty([0, T] \times \Omega)} |\rho_\delta|_{L^1(\Omega)} \geq 0 \end{aligned}$$

On passe maintenant à la limite en δ ; le dernier terme tend vers zéro, et le second reste inchangé. Le premier terme fait apparaître la trace de u (qui appartient à $BV(\Omega \times]0, T[)$, et dont la trace existe d'après le lemme 3). On obtient

$$\int_{\Gamma} \int_0^T \psi(t, \sigma) \left[\operatorname{sg}(\gamma u - k) (f(\gamma u) - f(k)) + \operatorname{sg}(k) \left(\frac{\lambda(t, \sigma)}{2} - f(k) \right) \right] \cdot v \, d\sigma \, dt \geq 0$$

Comme dans le cas monodimensionnel, on prend successivement k supérieur à $\text{Max}(\gamma u, 0)$ et inférieur à $\text{Min}(\gamma u, 0)$ dans l'inégalité suivante, qui est vérifiée presque partout sur $\Gamma \times]0, T[$

$$\left[\text{sg}(\gamma u - k) (f(\gamma u) - f(k)) + \text{sg}(k) \left(\frac{\lambda}{2} - f(k) \right) \right] \cdot v \geq 0,$$

on obtient que

$$(42) \quad f(\gamma u(\sigma, t)) = \frac{1}{2} \lambda(\sigma, t) \quad \text{pp } (\Gamma \times]0, T[),$$

et donc pour tout $k \in \mathbb{R}$

$$[\text{sg}(\gamma u - k) + \text{sg}(k)] [f(\gamma u(\sigma, t)) - f(k)] \cdot v \geq 0.$$

Cette dernière inégalité implique (IV.13), si on la restreint à $k \in I(\gamma u, 0)$. Comme précédemment, (IV.18) est satisfaite dès que (IV.13) est vraie et (IV.18) pour les fonctions test ϕ nulles sur la frontière. La démonstration du théorème 14 est terminée.

4. QUELQUES REMARQUES ET RESULTATS NUMERIQUES

L'hypothèse d'intersection connexe par les droites parallèles aux vecteurs de base a permis de simplifier la démonstration en deux endroits différents. Dans la démonstration de la conservation de la variation bornée en espace, on était assuré lorsque $i \in \mathcal{J}_h, i+1_\ell \in \Gamma_h, j < p$ qu'il existait un $\mu_j \in \mathbb{Z}$ tel que pour tout $k \in \{0, \dots, \mu_j\}$, les indices $i-1_j + k \cdot 1_\ell$ appartenait à \mathcal{J}_h et les $\Gamma_{i-1_j+k \cdot 1_\ell}$ étaient adjacents à des pavés indicés par des éléments de Γ_h . Si ce n'est plus le cas, on peut concevoir que l'existence d'un tel μ_j n'est plus assurée qu'un nombre fini (et indépendant de h) de fois, ce qui ne perturbe pas trop la variation en espace. En effet, si on a

$$\sum_{i \in \mathcal{J}_h \cup \Gamma_h} |u_{i+1_\ell}^{n+1} - u_i^{n+1}| \leq \sum_{i \in \mathcal{J}_h \cup \Gamma_h} |u_{i+1_\ell}^n - u_i^n| + C,$$

on doit encore le multiplier par $\prod_{k \neq \ell} h_k$ pour retrouver la variation de la condition initiale.

On a, de proche en proche

$$\sum_i |u_{i+1_\ell}^n - u_i^n| \prod_{k \neq \ell} h_k \leq \left(C n + \sum_i |u_{i+1_\ell}^0 - u_i^0| \right) \prod_{k \neq \ell} h_k,$$

$$\text{où } \sum_i |u_{i+1_\ell}^0 - u_i^0| \prod_{k \neq \ell} h_k \leq \frac{1}{h_\ell} \int_\Omega |u_0(x + h_\ell) - u_0(x)| dx \leq C^{te}$$

et

$$(43) \quad C n \prod_{k \neq \ell} h_k \leq C_0 |h|^{p-1},$$

et donc tend vers zéro avec $|h|$ lorsque $p \geq 2$.

Cette propriété intervient également pour construire $i_{0,j}$ et $i_{1,j}$ uniques, mais s'ils existent en nombre fini, la démonstration reste la même.

Le théorème 14 peut être étendu à la famille de schémas dont la convergence est établie au théorème 5. La démonstration est la même, sauf pour établir la conservation de la variation bornée en espace. Le principe reste le même, mais si la quantité μ_j qu'on fait apparaître au cours de la démonstration ne peut pas être uniformément bornée par une constante indépendante de h , on est amené soit à exiger que les coefficients $a_{i+1_j}^n$ sont constants au voisinage de la frontière, soit à faire une $i+1_j/2$ étude plus précise de la variation au voisinage de la frontière. Dans le premier cas, la démonstration reste inchangée. Dans l'autre cas, on fait apparaître que la perturbation de la variation est bornée par un terme de la forme

$$C_1 / |h|^{p-1}, \quad p \geq 2$$

ce qui a pour effet de conserver une variation bornée en espace (cf. (43)).

On peut aussi éviter l'hypothèse d'intersection connexe par les droites parallèles aux vecteurs de base, en reprenant la remarque analogue faite pour le schéma de LAX.

Le résultat (42) exprime un passage à la limite sur la solution au voisinage de la frontière. La fonction $\lambda(t)$ provient d'une convergence faible * au niveau de la couche limite, lorsqu'il n'y a pas raccordement de la condition aux bords avec la solution. Cette convergence faible doit être rapprochée de (IV.34), et fait intervenir la viscosité. Ceci est immédiat en comparant les éléments des suites β_m respectives construites pour le schéma de GODUNOV ou pour le schéma de LAX.

Numériquement, on remarque que plus le terme de viscosité est important, par exemple pour le schéma de LAX, et plus la couche limite est large. Ceci est illustré par l'exemple suivant, sur $]0,1[x]0,T[, T > 1$.

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0,$$

$$u(x,0) = \begin{cases} 1 & \text{si } x < 1/2, \\ 0 & \text{si } x > 1/2, \end{cases}$$

$$a(t) = 1; \quad b(t) = 0.$$

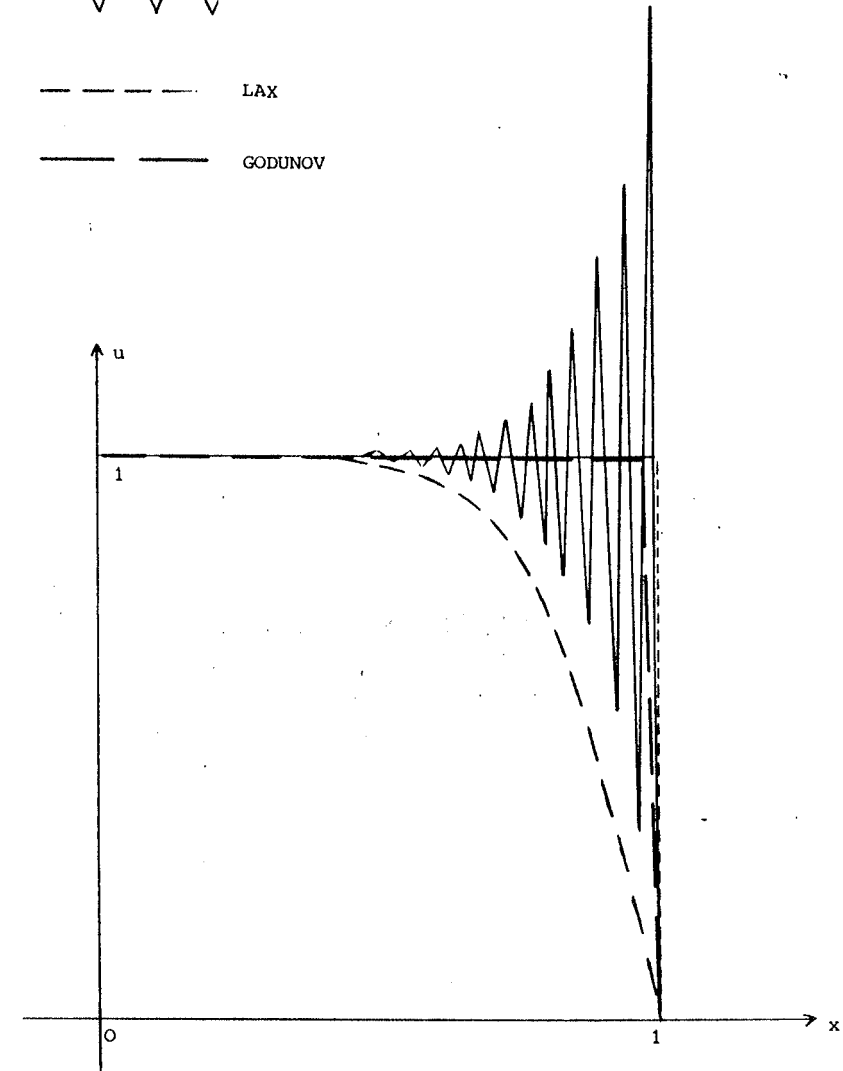
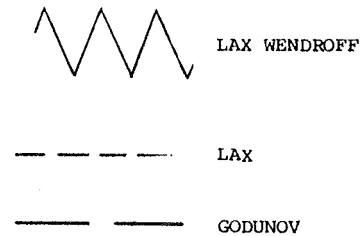
La solution est vérifiée

$$u(x,t) = 1 \quad \text{lorsque } t > 1, \quad x < 1.$$

Le rapport des pas en temps et en espace est pris petit, ce qui a pour effet d'augmenter considérablement le coefficient de viscosité pour le schéma de LAX.

Ceci se traduit par une couche limite importante, elle s'étend sur $[\frac{1}{2}, 1]$, à savoir la moitié de Ω . Le schéma de GODUNOV, dont la viscosité est indépendante de q , donne la solution exacte.

$$\begin{aligned} h = \Delta x &= 0.02 \\ qh = \Delta t &= 0.002 \\ q &= 0.1 \end{aligned}$$



Solution pour $t > 1$

Le schéma de LAX WENDROFF a également été testé ; la faible valeur de q a réduit d'autant sa viscosité, ce qui provoque de nombreuses oscillations sur un voisinage important de la frontière.

Le schéma de GODUNOV coïncide ici avec le schéma décentré ; il tient compte de la propagation et il devient normal que la couche limite ait disparu.

Les schémas quasi d'ordre deux présentent également un bon comportement au niveau de la couche limite, lorsque le schéma d'ordre un à partir duquel ils sont construits se comporte bien. Quelques résultats numériques sont présentés à la fin du chapitre suivant.

Dans l'exemple précédent, la condition $a(t) = 1$ a été prise en compte telle quelle dans le schéma, et non sous la forme d'un relèvement. On peut procéder ainsi pour tous les schémas, même à plusieurs dimensions, et démontrer la convergence. Dans la démonstration du théorème 14, on a écrit que cette condition était nulle au dernier moment. Pour une condition non nulle, on doit procéder à une régularisation, comme aux théorèmes 12 et 13, et on obtient (IV.9), dont on déduit (IV.40), ayant déjà (IV.40) pour les fonctions test ϕ nulles sur la frontière.

x x
x

CHAPITRE VI

APPROXIMATION DU PROBLEME

AUX LIMITES

DANS LE CAS MONOTONE A PLUSIEURS DIMENSIONS

On suppose dans ce chapitre que f dépend de façon monotone de u , dans le cadre du problème monodimensionnel. Pour le problème à plusieurs dimensions, chaque f_j est supposée monotone.

Afin de simplifier les démonstrations, on supposera f indépendante de x et de t . Tous les théorèmes se généralisent, en reprenant les détails de la démonstration du théorème 2. On suppose également que $f(0)$ est nul, ce qui ne réduit pas la généralité.

1. LE SCHEMA DECENTRE, A UNE OU PLUSIEURS VARIABLES

On considère le problème sur $\Omega \subset \mathbb{R}^p$

$$(1) \quad \frac{\partial u}{\partial t} + \operatorname{div} [f(u)] = 0 \quad \text{sur} \quad \Omega \times]0, T[,$$

$$(2) \quad u(x, 0) = u_0(x) \quad \text{sur} \quad \Omega ,$$

$$(3) \quad \text{Min}_{k \in I(\gamma u, 0)} \left\{ \text{sg}(\gamma u) (f(\gamma u) - f(k)) \cdot v \right\} = 0 \quad \text{pp } (\Gamma \times]0, T[),$$

c'est-à-dire que la condition aux bords est prise égale à zéro. Ceci peut être généralisé, au prix de difficultés techniques (celles rencontrées au théorème 12 (schéma de GODUNOV)).

On introduit le maillage utilisé au chapitre III pour approcher le problème à plusieurs dimensions (paragraphe 2), et on se propose de généraliser le théorème 6. On considère le schéma (III.26), c'est-à-dire

$$(4.i) \quad u_{i+1, j/2}^n \quad \text{réalise} \quad \text{Min}_{k \in I(u_{i+1, j}^n, u_i^n)} \left\{ \text{sg}(u_{i+1, j}^n - u_i^n) f_j(k) \right\}$$

$$(4.ii) \quad u_i^{n+1} = u_i^n - \sum_{j=1}^p q_j (f_j(u_{i+1, j/2}^n) - f_j(u_{i-1, j/2}^n)),$$

pour $i \in \mathcal{J}_h$, et $u_i^{n+1} = 0$ si $i \in \Gamma_h$, \mathcal{J}_h et Γ_h étant les mêmes ensembles d'indice que précédemment. Chaque f_j étant monotone, on a toujours

$$(5) \quad \begin{aligned} \text{soit} \quad u_{i+1, j/2}^n &= u_i^n && (f_j \text{ non décroissante}) \\ \text{soit} \quad u_{i+1, j/2}^n &= u_{i+1, j}^n && (f_j \text{ non croissante}) \end{aligned}$$

Théorème 15 : On suppose que toutes les fonctions f_j sont monotones, et que la condition de stabilité de COURANT FRIEDRICHS LEWY

$$(6) \quad \forall j \in \{1 \dots p\} \quad q_j \sup_{|k| \leq |u_0|_{L^\infty(\Omega)}} |f_j'(k)| \leq \frac{1}{p}$$

est satisfaite. Alors la famille $\{u_h\}_h$ des solutions construites par le schéma décentré (4.i), (4.ii) converge dans $L^1(\Omega \times]0, T[)$, vers une solution du problème (1), (2), (3) et caractérisée par la définition 3.

Démonstration

On suppose encore que $f_j(0)$ est nul pour tout j , et que l'intersection de Ω par toute droite parallèle à un des vecteurs de base est connexe. Cette hypothèse permet de simplifier la démonstration, et l'extension du résultat pour les ouverts Ω ne vérifiant pas cette hypothèse fait l'objet d'une remarque au chapitre précédent, paragraphe 4.

La technique utilisée est la même que celle du théorème 14, et lorsque $p = 1$, ce théorème est exactement le théorème 12.

Soit $i \in \mathcal{J}_h$; on a pour tout n

$$(7) \quad |u_i^{n+1}| \leq \max_{j \in \mathcal{I}^p |i-j| \leq 1} |u_i^n|,$$

et

$$(8) \quad \sum_{i \in \mathcal{J}_h} |u_i^{n+1} - u_i^n| \leq \sum_{j=1}^p \sum_{i \in \mathcal{J}_h} q_j |f_j(u_{i+1, j/2}^n) - f_j(u_{i-1, j/2}^n)|,$$

qui se majore immédiatement en utilisant (5) et (6), et il vient

$$\sum_{i \in \mathcal{J}_h} |u_i^{n+1} - u_i^n| \leq \sum_{j=1}^p \sum_{i \in \mathcal{J}_h \cup \Gamma_h} |u_{i+1, j}^n - u_i^n|,$$

en convenant de prendre toujours $u_i^n = 0$ si $i \notin \mathcal{J}_h$, même lorsque $i \in \Gamma_h$, pour faciliter l'écriture.

Soit $i \in \mathcal{J}_h$; si $i+l \in \mathcal{J}_h$, avec $l \in \{1 \dots p\}$ donné, on a encore (III.28); si $i+l \in \Gamma_h$, on a $u_{i+l}^{n+1} = u_{i+l}^n = 0$, donc

$$\begin{aligned} |u_{i+l}^{n+1} - u_i^{n+1}| &= |u_i^{n+1}| = |u_i^n - \sum_j q_j (f_j(u_{i+1, j/2}^n) - f_j(u_{i-1, j/2}^n))| \\ &\leq \frac{1}{p} \left(|u_i^n| + p q_l |f_l(u_{i+1, l/2}^n) - f_l(u_{i-1, l/2}^n)| + \sum_{j \neq l} |u_i^n - p q_j (f_j(u_{i+1, j/2}^n) - f_j(u_{i-1, j/2}^n))| \right) \end{aligned}$$

Pour $j \neq \ell$ fixé, on suppose par exemple f_j non décroissante,

et on a

$$(9) \quad \begin{aligned} & |u_i^n - pq_j (f_j(u_i^n) - f_j(u_{i-1_j}^n))| \\ & \leq |u_i^n| - pq_j |f_j(u_i^n)| + pq_j |f_j(u_{i-1_j}^n)| \end{aligned}$$

Or, on a toujours

$$|u_i^n| = |u_{i+1_\ell}^n - u_i^n|$$

et

$$|f_j(u_i^n)| = |f_j(u_{i+1_\ell}^n) - f_j(u_i^n)|,$$

puisque'on a supposé que $f_j(0)$ est nul.

De plus, d'après l'hypothèse faite sur Ω , il existe un entier relatif μ_1 tel que

$$u_{i-1_j + \mu_1 1_\ell}^n = 0 \quad \text{avec} \quad i-1_j + \mu_1 1_\ell \in \Gamma_h.$$

On a alors

$$|f_j(u_{i-1_j}^n)| \leq \sum_{k=0}^{\mu_1} |f_j(u_{i+k 1_\ell - 1_j}^n) - f_j(u_{i-1_j + (k+1) 1_\ell}^n)|.$$

Il vient alors une expression semblable à (III.28) prolongée jusqu'à la frontière, et tous les termes ajoutés ici vont disparaître avec des termes provenant de l'évaluation de

$$|u_{i+k 1_\ell - 1_j}^{n+1} - u_{i+(k+1) 1_\ell - 1_j}^{n+1}|,$$

par (III.28). Si f_j est croissante, (9) doit être remplacée par

$$|u_i^n - pq_j (f_j(u_{i+1_{j/2}}^n) - f_j(u_{i-1_{j/2}}^n))|$$

$$\leq |u_i^n| + pq_j |f_j(u_{i+1_{j/2}}^n)| - pq_j |f_j(u_i^n)|,$$

où $|u_i^n|$ et $|f_j(u_i^n)|$ sont traités comme précédemment, et $|f_j(u_{i+1_{j/2}}^n)|$ est estimé en atteignant la frontière parallèlement à la direction 1_ℓ , comme $|f_j(u_{i-1_j}^n)|$ précédemment. On traite de même le cas $i \in \Gamma_h, (i+1_\ell) \in \Gamma_h$.

On obtient, en sommant toutes les inégalités (III.28) obtenues

$$\sum_{i \in \mathcal{J}_h \cup \Gamma_h} |u_{i+1_\ell}^{n+1} - u_i^{n+1}| \leq \sum_{i \in \mathcal{J}_h \cup \Gamma_h} |u_{i+1_\ell}^n - u_i^n|.$$

Cette dernière estimation ainsi que (7) et (8) assurent la convergence dans $L^1(\Omega \times]0, T[)$ d'une suite $\{u_h\}$ vers une fonction $u \in BV(\Omega \times]0, T[)$ et qui satisfait à la condition initiale (IV.19) d'une part, et à (IV.18) pour les fonctions test nulles sur Γ .

Il reste à prouver (3), c'est-à-dire (IV.13) ; on procède comme au théorème 14, à partir de l'inégalité suivante, qui se déduit de la même inégalité vérifiée par le schéma décentré monodimensionnel.

Pour tout $k \in \mathbb{R}, i \in \mathcal{J}_h, n \leq N$, on a

$$(10) \quad \begin{aligned} |u_i^{n+1} - k| \leq \frac{1}{p} \sum_{j=1}^p \left\{ |u_i^{n-k}| - pq_j \operatorname{sg}(u_{i+1_{j/2}}^n - k) (f_j(u_{i+1_{j/2}}^n) - f_j(k)) \right. \\ \left. + pq_j \operatorname{sg}(u_{i-1_{j/2}}^n - k) (f_j(u_{i-1_{j/2}}^n) - f_j(k)) \right\}, \end{aligned}$$

ce qui est analogue à (V.33). On introduit une fonction test

$$\phi \in C_{0,+}^2(\bar{\Omega} \times]0, T[),$$

de la forme $\phi(x, t) = \psi(t, x) \rho_\delta(x)$, et on multiplie (10) par

$$\phi_i^n h_1 \dots h_p.$$

On obtient, avec les mêmes notations $i_{0,j}$ et $i_{1,j}$ qu'au théorème 14 précédent, l'inégalité suivante

$$\begin{aligned} & \sum_{j=1}^p \sum_n \sum_{i \in \Gamma_h} \sum_{k=i_{0,j+1}}^{i_{1,j}-1} (f_j(u_{i+1,j/2}^n) - f_j(k)) \operatorname{sg}(u_{i+1,j/2}^n - k) \frac{\phi_{i+1,j}^n \phi_i^n}{h_j} (q_{j,h_j}) h_1 \dots h_p \\ & \quad (i_j \text{ fixé}) \\ & + \sum_{j=1}^p \sum_n \sum_{i \in \Gamma_h} \operatorname{sg}(u_{i_0+1,j/2}^n - k) (f_j(u_{i_0+1,j/2}^n) - f_j(k)) q_j h_1 \dots h_p \\ & \quad i_j = i_0 \\ & - \sum_{j=1}^p \sum_n \sum_{i \in \Gamma_h} \operatorname{sg}(u_{i_1-1,j/2}^n - k) (f_j(u_{i_1-1,j/2}^n) - f_j(k)) q_j h_1 \dots h_p \\ & + C T(M + |k|) \left| \frac{\partial \psi}{\partial t} \right|_{L^\infty(\Gamma \times \Omega)} \left| \rho_\delta \right|_{L^1(\Omega)} > 0 \end{aligned}$$

On utilise sur la frontière une estimation du type (V.19) en i_0 , et du type correspondant en i_1 , puis on note $\lambda_j(t, \sigma)$ la limite dans $L^\infty \Gamma \times \Gamma$ faible * de la suite construite à partir des $f_j(u_{i_0+1,j/2}^n)$ ou des $f_j(u_{i_1-1,j/2}^n)$. On passe à la limite en h , et il vient

$$\begin{aligned} & \int_{\Omega} \int_0^T \operatorname{sg}(u-k) (f(u) - f(k)) \cdot \operatorname{grad}(\rho_\delta \psi(t,x)) \, dx \, dt \\ & + \int_{\Gamma} \int_0^T \operatorname{sg}(k) (\lambda(t,\sigma) - f(k)) \cdot v \, \psi(t,\sigma) \, d\sigma \, dt \geq 0, \end{aligned}$$

en traitant comme au chapitre V l'intégrale sur la frontière.

On passe maintenant à la limite en δ sur l'inégalité obtenue, et il vient, presque partout sur la frontière et pour presque tout $t \in]0, T[$, et tout $k \in \mathbb{R}$, l'inégalité

$$(11) \quad \begin{aligned} & \operatorname{sg}(\gamma u(\sigma, t) - k) (f(\gamma u(\sigma, t)) - f(k)) \cdot v \\ & + \operatorname{sg}(k) (\lambda(t, \sigma) - f(k)) \cdot v \geq 0. \end{aligned}$$

En prenant k très grand puis très petit ($-k$ grand), il vient

$$f(\gamma u(\sigma, t)) = \lambda(t, \sigma) \quad \text{pp } (\Gamma \times]0, T[),$$

et pour $k \in I(\gamma u, 0)$, l'inégalité (11) donne (3). Le théorème 15 est démontré.

2. LE SCHEMA QUASI D'ORDRE DEUX, A UNE OU PLUSIEURS VARIABLES

Le résultat précédent peut être étendu au schéma quasi d'ordre deux construit à partir du schéma (4i), (4.ii). On a le résultat suivant, en convenant de prolonger u_n par zéro à l'extérieur de Ω , de façon à donner un sens à (III.54.iii) près de la frontière, en annulant $a_{i+1,j/2}^n$ le plus souvent. On suppose également, pour le problème à plusieurs dimensions que la condition initiale satisfait, dans chaque direction parallèle aux axes, à l'hypothèse du nombre fini de changements de monotonie, et que cette hypothèse est conservée par le schéma. On a vu au chapitre III, du fait de la conservation de la monotonie, qu'il y a conservation de ce nombre en dimension 1.

Théorème 16 : On suppose que tous les f_j sont monotones, et que la condition de stabilité

$$(12) \quad \forall j \in \{1, \dots, p\} \quad q_j \sup_{|k| \leq |u_0|} |f_j'(k)| \leq \frac{1}{2p}, \quad L^\infty(\Omega)$$

et l'hypothèse du nombre fini de changements de monotonie sur u_0 sont satisfaites, ainsi que sa conservation, pour le problème à plusieurs dimensions ; alors la famille $\{u_n\}$ construite par le schéma (III.54.55) converge dans $L^1(\Omega \times]0, T[)$ vers u solution du problème (1), (2), (3), et caractérisée par la définition 3.

Démonstration

La démonstration réunit des arguments déjà utilisés dans les théorèmes 10, 13, 14 et 15, et on n'en donne ici que les étapes essentielles.

En procédant comme aux théorèmes 10 et 14, on établit la conservation de la borne uniforme, et de la variation bornée en temps. On a aussi l'estimation suivante de la variation bornée en espace, pour tout $\ell \in \{1, \dots, p\}$,

$$i \in \mathcal{J}_h \cup \Gamma_h \quad |u_{i+1_\ell}^{n+1} - u_i^{n+1}| \leq i \in \mathcal{J}_h \cup \Gamma_h \quad |u_{i+1_\ell}^n - u_i^n|,$$

en procédant comme au théorème 10, si i et $(i+1)_\ell$ appartiennent à \mathcal{J}_h , et comme aux théorèmes 14 ou 15 si i ou $(i+1)_\ell \in \Gamma_h$. On a par exemple, si $i \in \mathcal{J}_h$, $(i+1)_\ell \in \Gamma_h$, f_j croissante, $j \neq \ell$

$$\begin{aligned} & \left| u_i^n - p q_j (f_j(u_i^n) - f_j(u_{i-1_j}^n)) - p a_{i+1_{j/2}}^n + p a_{i-1_{j/2}}^n \right| \\ & \leq \frac{1}{2} |u_i^n| - p q_j |f_j(u_i^n)| + p q_j |f_j(u_{i+1_j}^n)| \\ & + \left| \frac{1}{2} u_i^n - p a_{i+1_{j/2}}^n + p a_{i-1_{j/2}}^n \right|, \end{aligned}$$

et ce dernier terme peut être majoré, d'après (III.55), ou bien par une estimation de la forme (cf. remarque théorème 10)

$$\frac{1}{2} |u_i^n| - p |a_{i+1_{j/2}}^n| + p |a_{i-1_{j/2}}^n|.$$

Tous ces termes disparaissent dans une éventuelle sommation, sauf $|u_i^n|$.

On obtient ainsi la convergence d'une suite $\{u_h^n\}$ dans $L^1(\Omega \times]0, T[)$ vers $u \in BV(\Omega \times]0, T[)$, qui vérifie (IV.18) pour les

fonctions ϕ nulles sur $\Gamma \times]0, T[$, et (IV.19). Il reste à prouver (3), et la technique étant la même que précédemment, il suffit de le montrer pour le problème monodimensionnel, ce qui revient à reprendre certains passages du théorème 13. On suppose, par exemple, que f est non décroissante.

On a en reprenant (III.45), pour tout $i \in \{1 \dots I-1\}$,

$$(13) \quad \left\{ \begin{aligned} |u_i^{n+1} - k| & \leq |u_i^n - k| - \frac{q}{2} \left\{ |f(u_{i+1}^n) - f(k)| - |f(u_{i-1}^n) - f(k)| \right\} \\ & + \beta_{i+1/2}^n |u_{i+1}^n - u_i^n| - \beta_{i-1/2}^n |u_i^n - u_{i-1}^n| + 2 |a_{i+1/2}^n| |S_{i-1/2}^n - S_{i+1/2}^n|, \end{aligned} \right.$$

où $\beta_{i+1/2}^n$ est donné par

$$(14) \quad \beta_{i+1/2}^n = \frac{q}{2} f'(\xi_{i+1/2}^n) (1 - 2S_{i+1/2}^n) (1 - u_{i+1/2}^n (1 - q f'(\xi_{i+1/2}^n))),$$

et $S_{i+1/2}^n$ vérifie $+1 \geq S_{i+1/2}^n \geq 0$ (cf (III.43))

donc $|1 - 2 S_{i+1/2}^n| \leq 1$.

On introduit une fonction test ϕ de la forme

$$\phi(x, t) = \begin{cases} \psi_1(t) \rho_\delta(x) & \text{si } x > 1/2, \\ \psi_0(t) \rho_\delta(x) & \text{si } x < 1/2, \end{cases}$$

ainsi que les ϕ_i^n en projetant ϕ sur chaque pavé. On multiplie (13) par $h \phi_i^n$ et on somme en i et n pour obtenir

$$\begin{aligned} 0 \leq \sum_n \sum_{i=2}^{I-2} q |f(u_i^n) - f(k)| \frac{\phi_{i+1}^n - \phi_{i-1}^n}{2h} h^2 \\ + \sum_n \sum_{i=1}^I \frac{1}{q} \beta_{i+1/2}^n |u_{i+1}^n - u_i^n| \frac{\phi_{i+1}^n - \phi_i^n}{h} q h^2 \end{aligned}$$

$$\begin{aligned}
 & - \sum_n (|f(u_1^n) - f(k)| + |f(u_{I-1}^n) - f(k)|) \phi_I^n \frac{q}{2} h \\
 & + \sum_n (|f(u_0^n) - f(k)| + |f(u_1^n) - f(k)|) \phi_0^n \frac{q}{2} h \\
 & + \sum_n \beta_{I-1/2}^n |u_I^n - u_{I-1}^n| \phi_I^n h \\
 & - \sum_n \beta_{1/2}^n |u_1^n - u_0^n| \phi_0^n h \\
 & + T (M+|k|) |\psi'|_{L^\infty} |\rho_\delta|_{L^1(\Omega)} + C |\phi|_L h M T \\
 & + 2 C h^\alpha K |\psi|_{L^1(0,T)}.
 \end{aligned}$$

On doit maintenant estimer les sommes sur la frontière.

D'après (III.54), on a, en utilisant les propriétés du schéma décentré,

$$|u_1^{n+1} - k| \leq |u_1^n - k| - q |f(u_1^n) - f(k)| + q |f(u_0^n) - f(k)| + 2 C h^\alpha,$$

d'où en sommant en n,

$$\begin{aligned}
 \sum_n q |f(u_1^n) - f(k)| \phi_0^n h & \leq \sum_n q |f(u_0^n) - f(k)| \phi_0^n h \\
 (16) \quad & + 2 C T h^\alpha + |\psi'|_{L^\infty} (M+|k|) T h.
 \end{aligned}$$

On n'a pas, le schéma d'origine étant décentré, d'inégalité analogue en $x = 1$ ($i = I$). On majore donc le terme correspondant, qui est non positif, dans (15) par zéro.

On a également, pour tout n

$$|\beta_{1/2}^n| |u_1^n - u_0^n| \leq C h^\alpha; \quad |\beta_{I-1/2}^n| |u_I^n - u_{I-1}^n| \leq C h^\alpha.$$

On peut maintenant passer à la limite en h sur (15) ainsi majoré, et obtenir, sachant que u_0^n est nul pour tout n,

$$(17) \quad 0 \leq \int_{\Omega} \int_0^T |f(u) - f(k)| \frac{\partial \phi}{\partial x} dx dt + \int_0^T |f(k)| \psi_0(t) dt;$$

on a en effet

$$\sum_n \sum_{i=1}^I \frac{1}{q} \beta_{i+1/2}^n |u_{i+1}^n - u_i^n| \frac{\phi_{i+1}^n - \phi_i^n}{h} q h^2 \leq \frac{T}{q} \text{Var}(u_0; [0,1]) \left| \frac{\partial \phi}{\partial x} \right|_{L^\infty} h$$

en utilisant la variation bornée en espace. On peut également majorer cette quantité en utilisant le fait que $a_{i+1/2}^n$ soit borné par $C h^\alpha$, et obtenir un majorant de la forme

$$C \frac{T}{q} \left| \frac{\partial \phi}{\partial x} \right|_{L^\infty} h^\alpha.$$

On fait tendre δ vers zéro dans (17), sachant que ρ_δ est en facteur dans ϕ , et on obtient presque partout sur $]0, T[$,

$$|f(\gamma u(0,t)) - f(k)| - |f(k)| \leq 0,$$

$$|f(\gamma u(1,t)) - f(k)| \geq 0.$$

On en déduit immédiatement (3), avec $f(\gamma u(0,t)) = f(0) = 0$.

Si f est décroissante on a une estimation analogue à (16) en $x = 1$, dont on déduit en passant à la limite en h et δ que $f(\gamma u(1,t))$ est nul presque partout. On élimine le terme en zéro en prenant par exemple $\psi_0(t) = 0$. On peut également, et c'est ce qui se généralise à plusieurs dimensions, noter $\lambda(t)$ la limite dans $L^\infty(0,T)$ faible * de la suite (en fait la sous-suite extraite) construite à partir des $f(u_1^n)$, ce qui donnera à la limite.

$$- 2 |f(\gamma u(0,t)) - f(k)| + |f(k)| + |\lambda(t) - f(k)| \geq 0 \quad \text{pp }]0, T[.$$

En prenant $k > \text{Max}(\lambda(t), \gamma u(0,t), 0)$ puis $k < \text{Min}(\lambda(t), \gamma u(0,t), 0)$, il vient

$$2 f(\gamma(0,t)) = \lambda(t) \quad \text{pp }]0, T[,$$

et l'inégalité se réduit ensuite à $|f(k)| \geq 0$.

Cette technique se généralise à plusieurs dimensions, et on obtient ainsi (3). Le théorème 16 est démontré.

x x
x

On peut remarquer, dans le cas monodimensionnel, qu'il est inutile de vérifier (V.4) lorsque f est non décroissante ; cette relation est toujours satisfaite. Il en est de même de (V.3) lorsque f est non croissante.

3. UN SCHEMA QUASI EXPLICITE D'ORDRE DEUX EN TEMPS

On suppose f non décroissante, et on considère le problème (1), (2), (3), qui peut être écrit de la façon suivante, pour $\Omega =]0, 1[$,

$$(18) \quad \left\{ \begin{array}{l} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0 \quad \text{sur } \Omega \times]0, T[, \\ u(x, 0) = u_0(x) \quad x \in \Omega , \\ f(u(0, t)) = f(a(t)) \quad t \in]0, T[. \end{array} \right.$$

On fait sur u_0 et a les mêmes hypothèses qu'au chapitre V.

On veut construire un schéma quasi explicite, c'est-à-dire un schéma dans lequel le calcul de u_i^{n+1} est obtenu en résolvant une équation non linéaire de la forme

$$(19) \quad u + \alpha f(u) = C^{te} ,$$

où α est une constante positive fixe. Il est immédiat, f étant non décroissante, que (19) admet une solution unique.

Pour obtenir un tel schéma, on considère un schéma d'ordre un (ici le schéma décentré), et on discrétise l'équation selon ce schéma, de façon à obtenir un reste provenant des différents développements de TAYLOR. On discrétise ensuite ce reste, pour obtenir l'ordre deux en temps ; le schéma ainsi construit possède des propriétés de stabilité dans L^∞ et dans $BV(\Omega)$ analogues à celles du schéma décentré.

En prenant le schéma décentré, et en notant \tilde{u}_i^n la valeur de la solution de (18) au point $(x, t) = (ih, nqh)$, qu'on suppose suffisamment régulière dans un voisinage de ce point, on obtient en utilisant des développements de TAYLOR, au point (a, t) ,

$$(20) \quad \begin{aligned} & \frac{1}{qh} (\tilde{u}_i^{n+1} - \tilde{u}_i^n) + \frac{1}{h} (f(\tilde{u}_i^n) - f(\tilde{u}_{i-1}^n)) \\ & \approx - \frac{qh}{2} \frac{\partial}{\partial x} \left\{ \frac{\partial}{\partial t} f(u) \right\}_i^n - \frac{h}{2} \left\{ \frac{\partial^2}{\partial x^2} f(u) \right\}_i^n . \end{aligned}$$

Le schéma de LAX WENDROFF est obtenu en écrivant, dans le second membre de (20) que

$$\frac{\partial}{\partial t} f(u) = f'(u) \frac{\partial u}{\partial t} = - f'(u)^2 \frac{\partial u}{\partial x} ,$$

et en exprimant l'erreur par une discrétisation en espace.

Ici on va conserver la dérivée en temps, et écrire

$$\left\{ \frac{\partial^2}{\partial x \partial t} f(u) \right\}_i^n \approx \frac{1}{qh^2} \left\{ f(\tilde{u}_i^{n+1}) - f(\tilde{u}_{i-1}^{n+1}) - f(\tilde{u}_i^n) + f(\tilde{u}_{i-1}^n) \right\} .$$

Le schéma obtenu est le suivant, avec les mêmes notations que précédemment,

$$(21) \quad u_i^{n+1} + \frac{q}{2} f(u_i^{n+1}) = u_i^n - \frac{q}{2} f(u_i^n) + \frac{q}{2} f(u_{i-1}^{n+1}) + \frac{q}{2} f(u_{i-1}^n),$$

et il est d'ordre deux en temps, et l'ordre un en espace. De plus, le coefficient q est toujours multiplié par $\frac{1}{2}$, et ceci permet d'avoir une condition de stabilité plus large ; on aura deux fois moins d'étapes de calcul en temps que pour le schéma décentré.

Théorème 17 : Sous la condition de stabilité

$$(22) \quad 0 \leq qf'(u) \leq 2$$

$$\text{si } |u| \leq \text{Max} \left\{ |u_0|_{L^\infty(\Omega)}, |a|_{L^\infty(0,T)} \right\},$$

la famille $\{u_h\}_{h>0}$ des solutions approchées construites par (21) converge dans $L^1(\Omega \times]0, T[)$, vers u solution du problème (18), et caractérisée par la définition 3.

x x
x

Démonstration

On note $[\alpha, \beta]$ le plus petit intervalle fermé contenant u_{i-1}^n , u_i^n et u_{i-1}^{n+1} , supposés connus pour $n < N$ et $i \in \{1 \dots I\}$ fixés. On a alors, la fonction

$$u \rightsquigarrow u - \frac{q}{2} f(u)$$

étant non décroissante, les deux inégalités

$$\alpha - \frac{q}{2} f(\alpha) \leq u_i^n - \frac{q}{2} f(u_i^n) \leq \beta - \frac{q}{2} f(\beta)$$

Il vient alors, dans (21), f étant monotone

$$\alpha + \frac{q}{2} f(\alpha) \leq u_i^{n+1} + \frac{q}{2} f(u_i^{n+1}) \leq \beta + \frac{q}{2} f(\beta),$$

et donc

$$u_i^{n+1} \in [\alpha, \beta],$$

et on en déduit la stabilité dans $L^\infty(\Omega \times]0, T[)$:

$$(23) \quad \text{Sup}_i |u_i^{n+1}| \leq \text{Max} \left\{ |u_0^{n+1}|, \text{Sup}_i |u_i^n| \right\}.$$

En écrivant (21) en $i+1$ et en i , il vient

$$\begin{aligned} u_{i+1}^{n+1} - u_i^{n+1} + \frac{q}{2} (f(u_{i+1}^{n+1}) - f(u_i^{n+1})) &= u_{i+1}^n - u_i^n - \frac{q}{2} (f(u_{i+1}^n) - f(u_i^n)) \\ &+ \frac{q}{2} (f(u_i^{n+1}) - f(u_{i-1}^{n+1})) + \frac{q}{2} (f(u_i^n) - f(u_{i-1}^n)), \end{aligned}$$

et on passe aux valeurs absolues, en remarquant que

$$\begin{aligned} |u_{i+1}^n - u_i^n - \frac{q}{2} (f(u_{i+1}^n) - f(u_i^n))| &= |u_{i+1}^n - u_i^n| \left(1 - \frac{q}{2} f'(\xi_{i+1/2}^n)\right) \\ &= |u_{i+1}^n - u_i^n| - \frac{q}{2} |f(u_{i+1}^n) - f(u_i^n)| \end{aligned}$$

et de la même façon,

$$|u_{i+1}^{n+1} - u_i^{n+1} + \frac{q}{2} (f(u_{i+1}^{n+1}) - f(u_i^{n+1}))| = |u_{i+1}^{n+1} - u_i^{n+1}| + \frac{q}{2} |f(u_{i+1}^{n+1}) - f(u_i^{n+1})|.$$

On obtient ainsi

$$\begin{aligned} |u_{i+1}^{n+1} - u_i^{n+1}| + \frac{q}{2} |f(u_{i+1}^{n+1}) - f(u_i^{n+1})| &\leq |u_{i+1}^n - u_i^n| \\ &- \frac{q}{2} |f(u_{i+1}^n) - f(u_i^n)| + \frac{q}{2} |f(u_i^n) - f(u_{i-1}^n)| + \frac{q}{2} |f(u_i^{n+1}) - f(u_{i-1}^{n+1})| \end{aligned}$$

On somme en $i \in \{1 \dots I-1\}$, et il vient, après élimination des termes en $|f(u_{i+1}^{n+1}) - f(u_i^{n+1})|$ d'une part et des termes en $|f(u_{i+1}^n) - f(u_i^n)|$ d'autre part,

$$\sum_{i=1}^{I-1} |u_{i+1}^{n+1} - u_i^{n+1}| \leq \sum_{i=1}^{I-1} |u_{i+1}^n - u_i^n| + \frac{q}{2} \left\{ |f(u_1^n) - f(u_0^n)| + |f(u_1^{n+1}) - f(u_0^{n+1})| \right\}.$$

On ajoute $|u_1^{n+1} - u_0^{n+1}|$, majoré par,

$$|u_1^{n+1} - u_0^{n+1}| \leq |u_1^n - u_0^n| - \frac{q}{2} |f(u_1^n) - f(u_0^n)| - \frac{q}{2} |f(u_1^{n+1}) - f(u_0^{n+1})| + |u_0^{n+1} - u_0^n|,$$

en utilisant (21) pour $i = 1$ et la condition de stabilité, et on obtient pour tout $n < N$

$$(24) \quad \sum_{i=0}^{I-1} |u_{i+1}^{n+1} - u_i^{n+1}| \leq \sum_{i=0}^{I-1} |u_{i+1}^n - u_i^n| + |u_0^{n+1} - u_0^n|.$$

On a également conservation de la variation bornée en temps.

On a, en reprenant (21), pour tout i et tout n ,

$$|u_i^{n+1} - u_i^n| \leq \frac{q}{2} \left\{ |f(u_i^n) - f(u_{i-1}^n)| + |f(u_i^{n+1}) - f(u_{i-1}^{n+1})| \right\},$$

qu'on majore en utilisant (22), et (24). On obtient

$$(25) \quad \sum_{i=0}^I |u_i^{n+1} - u_i^n| = 2|u_0^{n+1} - u_0^n| + \sum_{i=0}^{I-1} |u_{i+1}^n - u_i^n|.$$

De proche en proche, à partir de (24) et de (25), on a

$$\sum_{i=0}^{I-1} |u_{i+1}^{n+1} - u_i^{n+1}| \leq \sum_{v=0}^n |u_0^{v+1} - u_0^v| + \sum_{i=0}^{I-1} |u_{i+1}^0 - u_i^0|$$

et

$$\sum_{i=0}^I |u_i^{n+1} - u_i^n| \leq 2 \sum_{v=0}^n |u_0^{v+1} - u_0^v| + \sum_{i=0}^{I-1} |u_{i+1}^0 - u_i^0|.$$

La condition initiale u_0 et la condition en $x = 0, a$, étant à variation bornée, ces quantités sont bornées ; on avait déjà (23) et donc $\{u_h\}$ contient une suite $\{u_{h_n}\}$ convergente dans $L^1(\Omega \times]0, T[)$ vers une fonction $u \in BV(\Omega \times]0, T[)$.

En procédant comme au chapitre II, partie F, on vérifie que la trace de u en $t = 0$ est égale à u_0 presque partout sur $]0, 1[$.

Soit $k \in \mathbb{R}$; on introduit k dans (21), pour obtenir

$$u_i^{n+1} - k + \frac{q}{2} (f(u_i^{n+1}) - f(k)) = u_i^n - k - \frac{q}{2} (f(u_i^n) - f(k)) + \frac{q}{2} (f(u_{i-1}^{n+1}) - f(k)) + \frac{q}{2} (f(u_{i-1}^n) - f(k)).$$

On passe aux valeurs absolues, et il vient, compte-tenu de la condition (22),

$$(26) \quad |u_i^{n+1} - k| + \frac{q}{2} |f(u_i^{n+1}) - f(k)| \leq |u_i^n - k| - \frac{q}{2} |f(u_i^n) - f(k)| + \frac{q}{2} |f(u_{i-1}^n) - f(k)| + \frac{q}{2} |f(u_{i-1}^{n+1}) - f(k)|.$$

On introduit $\phi \in C_{0,+}^2(]0, 1[\times]0, T[)$, nulle en $x = 1$ puisqu'on sait d'avance que (IV.8) est vérifié. On effectue des projections de type L^2 sur chaque pavé $I_i \times J_n$, comme précédemment, et on multiplie (26) par $h \phi_i^n$. On somme en i et n , pour obtenir, pour

$$\phi_{I-1}^n = \phi_I^n = \phi_i^0 = \phi_i^1 = \phi_i^{N-1} = \phi_i^N = 0 \quad \text{pour tout } i \text{ et } n,$$

$$\sum_{i=1}^{I-1} \sum_n |u_i^n - k| \frac{\phi_i^n - \phi_i^{n-1}}{qh} qh^2$$

$$+ \frac{1}{2} \sum_{i=1}^{I-1} \sum_n \left\{ |f(u_i^{n+1}) - f(k)| + |f(u_i^n) - f(k)| \right\} \frac{\phi_{i+1}^n - \phi_i^n}{h} qh^2$$

$$+ \frac{1}{2} \sum_n \left[|f(u_0^{n+1}) - f(k)| + |f(u_0^n) - f(k)| \right] \phi_1^n qh \geq 0.$$

On fait tendre h vers zéro, et on obtient

$$(27) \quad \int_{\Omega} \int_0^T |u-k| \frac{\partial \phi}{\partial t} + |f(u) - f(k)| \frac{\partial \phi}{\partial x} \, dx \, dt + \int_0^T |f(a(t)) - f(k)| \phi(0,t) \, dt \geq 0.$$

En prenant ϕ de la forme

$$\phi(x,t) = \begin{cases} \psi(t) \rho_{\delta}(x), & \text{si } x < 1/2, \\ 0, & \text{si } x > 1/2, \end{cases}$$

on obtient en passant à la limite en δ , l'inégalité

$$\forall k \in \mathbb{R} \quad |f(a(t)) - f(k)| \geq |f(\gamma u(0,t)) - f(k)|,$$

pour presque tout $t \in]0, T[$. En prenant $k = a(t)$, on a immédiatement

$$f(\gamma u(0,t)) = f(a(t)) \quad \text{pp }]0, T[,$$

c'est-à-dire la condition aux limites donnée dans (18). La condition (IV.7) est maintenant évidente, et le théorème 17 est démontré.

x x
x

4. UN SCHEMA QUASI EXPLICITE POUR LES EQUATIONS AVEC MIGRATION

Parmi les équations du type

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u),$$

il en existe où la fonction f est strictement croissante, et même où il existe une constante $\alpha_0 > 0$ telle que pour tout u ,

$$(28) \quad f'(u) \geq \alpha_0.$$

Ce n'est pas le cas pour l'équation de BURGERS, mais ceci est vrai pour l'équation d'onde de surface d'un fluide incompressible, donnée en deuxième exemple au chapitre I, la solution cherchée (notée h) étant toujours positive. C'est encore le cas pour certaines équations où apparaît un terme de migration, c'est-à-dire un terme de la forme

$$\frac{\partial}{\partial x} (\alpha u),$$

avec α strictement plus grand qu'un α_0 fixe, et qui incite la solution à se propager avec une vitesse supérieure à α_0 . On peut citer l'exemple du vent, et l'exemple de certaines équations de la convection.

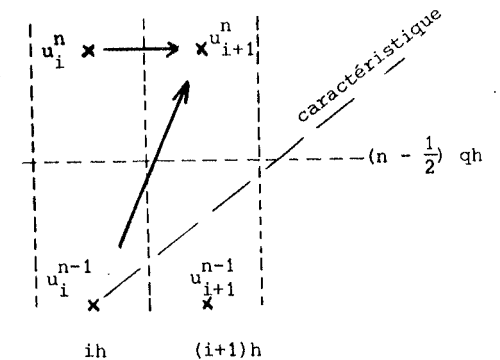
Il arrive également que α_0 soit relativement grand, ce qui se traduit par une grande vitesse de propagation. Dans ces cas la condition de stabilité de COURANT FRIEDRICHS LEWY rend très petit le pas de discrétisation en temps, par rapport à h.

On propose ici un schéma quasi explicite, qui inverse la condition de COURANT FRIEDRICHS LEWY. Ce schéma est construit en considérant à la fois x et t comme des variables d'évolution, et ne sera stable, ayant (28), que lorsque $q \alpha_0$ est plus grand que 1.

Ce schéma est le suivant, pour $i \geq 1, n < N$

$$(29) \quad q f(u_{i+1}^n) = q f(u_i^n) - u_i^n + u_i^{n-1}.$$

Il exprime u_{i+1}^n en fonction de u_i^n et de u_i^{n-1} . On n'utilise pas u_{i+1}^{n-1} , et il faut donc que $I_{i+1} \times J_n$ ne soit jamais traversé par une caractéristique issue de u_i^{n-1} , ce qu'exprime la condition $q \alpha_0 \geq 1$.



On a le résultat suivant

Théorème 18 : Sous la condition de stabilité

$$(30) \quad |u| \leq \text{Max} \left(|u_0|_{\infty}, |a|_{\infty} \right) \quad \text{Inf}_{L} \quad qf'(u) > 1$$

la famille $\{u_h\}$ des solutions approchées construites à partir du schéma (29), converge dans $L^1(\Omega \times]0, T[)$ vers u solution du problème (18), caractérisée par la définition 3.

x x
x

Démonstration

On procède comme au théorème 17.

Si u_i^{n-1} est plus petit que u_i^n , on a d'une part

$$q f(u_{i+1}^n) \leq q f(u_i^n) - u_i^n + u_i^n = q f(u_i^n)$$

et d'autre part, f étant monotone croissante, ainsi que $qf(u) - u$,

$$q f(u_{i+1}^n) \geq q f(u_i^{n-1}) - u_i^{n-1} + u_i^{n-1} = q f(u_i^{n-1})$$

On en déduit immédiatement

$$q f(u_i^{n-1}) \leq q f(u_{i+1}^n) \leq q f(u_i^n)$$

et on a un résultat analogue si $u_i^n \leq u_i^{n-1}$. On obtient ainsi

$$(31) \quad u_{i+1}^n \in I(u_i^n, u_i^{n-1})$$

On a également la conservation de la variation bornée ; on écrit (29) en i et i-1, et il vient

$$q(f(u_{i+1}^n) - f(u_i^n)) = q(f(u_i^n) - f(u_{i-1}^n)) - (u_i^n - u_{i-1}^n) + (u_i^{n-1} - u_{i-1}^{n-1})$$

Or

$$\begin{aligned} |q(f(u_i^n) - f(u_{i-1}^n)) - (u_i^n - u_{i-1}^n)| &= (q f'(\xi_{i-1/2}^n) - 1) |u_i^n - u_{i-1}^n| \\ &= q |f(u_i^n) - f(u_{i-1}^n)| - |u_i^n - u_{i-1}^n| \end{aligned}$$

donc en passant aux valeurs absolues,

$$q |f(u_{i+1}^n) - f(u_i^n)| \leq q |f(u_i^n) - f(u_{i-1}^n)| - |u_i^n - u_{i-1}^n| + |u_i^{n-1} - u_{i-1}^{n-1}|$$

On effectue la somme en $i \in \{1, \dots, I-1\}$, et il vient, les termes en $|f(u_{i+1}^n) - f(u_i^n)|$ disparaissant,

$$\sum_{i=1}^{I-1} |u_i^n - u_{i-1}^n| \leq \sum_{i=1}^{I-1} |u_i^{n-1} - u_{i-1}^{n-1}| + q |f(u_1^n) - f(u_0^n)|$$

Or, en reprenant le schéma,

$$q |f(u_1^n) - f(u_0^n)| = |u_0^n - u_0^{n-1}|$$

On obtient donc (24). On a également une estimation analogue à (25) ; en effet

$$|u_i^n - u_i^{n-1}| = q |f(u_{i+1}^n) - f(u_i^n)|$$

(32)

$$\leq \text{Max}_{|u| \leq \text{Max}(|u_0|_{\infty}, |a|_{\infty})} \{q f'(u)\} |u_{i+1}^n - u_i^n|$$

et il reste à faire la somme en i pour avoir la variation bornée en temps.

Ayant (31), (32) et (24), la famille $\{u_h\}$ contient une suite $\{u_h\}$ convergente dans $L^1(\Omega \times]0, T[)$ vers $u \in BV(\Omega \times]0, T[)$, qui vérifie la condition initiale. Il reste à prouver que u satisfait à (27).

Soit $k \in \mathbb{R}$; on a, à partir de (29),

$$q(f(u_{i+1}^n) - f(k)) = q(f(u_i^n) - f(k)) - (u_i^n - k) + (u_i^{n-1} - k) .$$

Or, d'après la condition (30),

$$\begin{aligned} |q(f(u_i^n) - f(k)) - (u_i^n - k)| &= (q f'(k_i^n) - 1) |u_i^n - k| \\ &= q |f(u_i^n) - f(k)| - |u_i^n - k| , \end{aligned}$$

et donc

$$(33) \quad q |f(u_{i+1}^n) - f(k)| \leq q |f(u_i^n) - f(k)| - |u_i^n - k| + |u_i^{n-1} - k| .$$

On introduit $\phi \in C_{0,+}^2([0,1] \times]0,T[)$, comme au théorème précédent, on multiplie par $h\phi_i^n$ et on somme en i et n pour obtenir

$$\begin{aligned} \sum_n \sum_{i=1}^I \left\{ q |f(u_i^n) - f(k)| \frac{\phi_i^n - \phi_{i-1}^n}{h} h^2 + |u_i^n - k| \frac{\phi_i^{n+1} - \phi_i^n}{qh} q h^2 \right\} \\ + \sum_n q |f(u_0^n) - f(k)| \phi_0^n \geq 0 . \end{aligned}$$

En passant à la limite en h , il vient exactement (27).

Le reste de la démonstration est identique à la fin de la démonstration du théorème 17. Le théorème 18 est démontré, la condition limite (IV.8) en $x = 1$ étant trivialement vérifiée.

5. QUELQUES REMARQUES ET RESULTATS NUMERIQUES

Il est possible, en utilisant le principe d'antidiffusion - correction, de construire des versions quasi d'ordre deux pour les schémas (21) et (29). En ce qui concerne le schéma (21), on est amené à discrétiser

$$- \frac{h^2}{2} \frac{\partial^2}{\partial x^2} \{q f(u)\}$$

au point (ih, nqh) , ce qui, compte-tenu de la correction qu'on doit ensuite opérer, revient à ajouter au second membre de (15), $- a_{i+1/2}^n + a_{i-1/2}^n$,

$$\text{avec } a_{i+1/2}^n = \text{sg}(u_{i+1}^n - u_i^n) \text{Max} \left[0, \min \left[\frac{q}{2} (f(u_i^n) - f(u_{i+1}^n)), \frac{q}{2} (f(u_{i+1}^n) - f(u_i^n)) \right] \right] .$$

On remarque que les deux termes dans ce minimum sont de même ordre, et multipliés par la même constante, et donc le schéma ne sera pas "très souvent" d'ordre deux. De plus, la condition de stabilité redevient la condition de COURANT FRIEDRICHS LEWY (second membre égal à 1), et le schéma, étant quasi explicite, devient moins intéressant.

On peut également opérer de façon à retrouver le schéma de CRANK NICHOLSON, en écrivant

$$- q h^2 \frac{\partial^2}{\partial x^2} (f(u)) = q h^2 \frac{\partial^2 u}{\partial x \partial t} \approx u_i^{n+1} - u_i^n - u_{i-1}^{n+1} + u_{i-1}^n .$$

On retrouve alors un schéma qui n'est plus décentré, mais qui reste quasi explicite si au niveau de la résolution numérique on tient compte du sens de propagation.

Pour le schéma (29), on est amené à discrétiser en (ih, nqh)

$$q \frac{h^2}{2} \left(\frac{\partial^2 f(u)}{\partial x^2} - q \frac{\partial^2 u}{\partial t^2} \right) ,$$

ce qui rend la résolution numérique de plus en plus implicite, $f(u_{i+1}^n)$ intervenant dans le terme de correction, et d'autre part introduit $qf(u_{i-1}^n)$ avec un coefficient négatif si (30) est satisfaite, et le schéma est instable en général.

On peut, sur certains exemples, utiliser les schémas (21) et (29) pour des fonctions f non monotones croissantes. Il est immédiat qu'on a de tels schémas lorsque f est décroissante, en tenant compte de la condition limite en $x = 1$. On peut donc utiliser ces schémas à plusieurs dimensions, en les généralisant de la même façon que le schéma décentré,

lorsque chaque f_j est monotone. La condition de stabilité rend alors ces schémas assez "économiques" en pratique.

On peut également envisager d'utiliser le schéma (21) pour l'équation de BURGERS, quand la condition initiale et les conditions aux limites vérifient certaines propriétés. Par exemple, pour le problème de CAUCHY, avec une condition initiale u_0 vérifiant

$$(34) \quad \exists x_0 \in \mathbb{R} \begin{cases} u_0(x) \geq 0 & \text{si } x \geq x_0, \\ u_0(x) \leq 0 & \text{si } x \leq x_0, \end{cases}$$

on peut résoudre séparément les problèmes sur $]x_0, +\infty[$ avec f croissante, et une condition nulle en x_0 , et sur $] -\infty, x_0[$ avec f décroissante, et une condition nulle en x_0 .

Si maintenant, sur un intervalle ouvert Ω ,

$$(35) \quad \exists x_0 \in \Omega \begin{cases} u_0(x) \leq 0 & \text{si } x \geq x_0, \\ u_0(x) \geq 0 & \text{si } x \leq x_0, \end{cases}$$

avec des conditions aux limites nulles, on peut résoudre séparément, pour tout n ($t = nqh$) les problèmes sur $\{x < x_n\}$ et sur $\{x > x_n\}$, puis par exemple prendre la moyenne en x_n , x_n étant, au temps nqh , un point de Ω tel que $u_n(\cdot, nqh)$ est non positif pour $x \geq x_n$ et non négatif pour $x \leq x_n$. Au bout d'un temps fini, il apparaît d'ailleurs un choc le long du lieu des x_n .

On peut encore généraliser : s'il existe plusieurs x_0 vérifiant (34) et (35), on peut découper \mathbb{R} ou Ω en plusieurs intervalles, et utiliser les remarques précédentes sur chacun de ces intervalles.

Quelques expériences numériques sur des schémas, dont la convergence est établie dans ce chapitre, ont été menées.

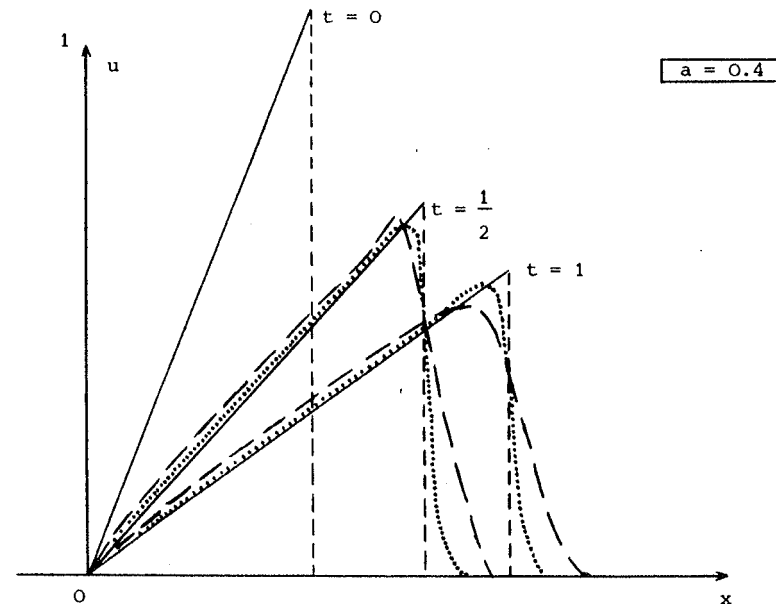
On reprend tout d'abord pour le schéma quasi explicite (21) l'exemple

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0 \quad ; \quad u(x,0) = \begin{cases} x/a & \text{si } 0 < x < a, \\ 0 & \text{si } a < x < 1, \end{cases}$$

les deux conditions aux limites étant nulles.

Le pas de discrétisation en espace a été pris successivement égal à 0.05 et à 0.025, et le rapport q était toujours maximal : $q = 2$.

<u>Schéma quasi explicite (21)</u>	—————	Solution exacte
	- - - - -	$h = 0.05, q = 2$
	$h = 0.025, q = 2$



On obtient pour $t = \frac{1}{2}$ et $t = 1$ les résultats représentés ci-avant. On peut remarquer que l'erreur diminue également au niveau de la discontinuité. Les pas de discrétisation en temps sont respectivement 0.1 et 0.05.

Ce schéma étant décentré, les couches limites éventuelles se réduisent à un pas du maillage, comme pour le schéma décentré classique.

Ces différents schémas ont également été testés sur l'exemple suivant, qui représente une détente légèrement perturbée en h . La condition initiale prise numériquement était

$$u_i^0 = \begin{cases} 0 & i = 0 \text{ ou } 1 \\ 0,75 & i = 2 \\ 1 & i \geq 3 \end{cases}$$

ce qui correspond, pour $h = 0.05$, à

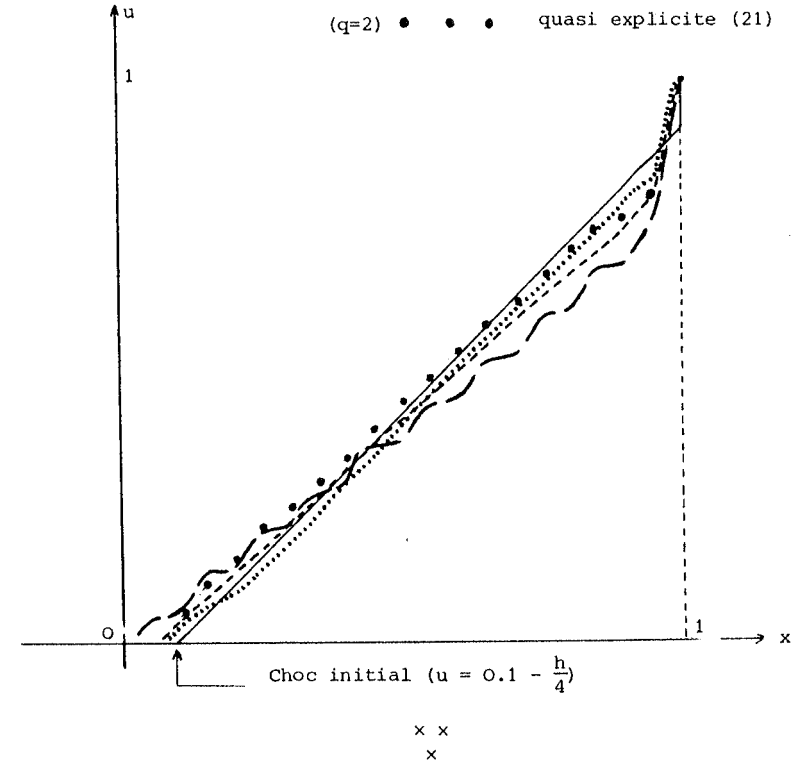
$$(36) \quad u_0(x) = \begin{cases} 0 & x < 0,1 - \frac{h}{4} \\ 1 & x > 0,1 - \frac{h}{4} \end{cases}$$

et les conditions aux limites $a = 0$, $b = 1$.

On a choisi le même pas de discrétisation en espace, et le rapport q le plus grand possible pour chacun de ces schémas. On retrouve un comportement "en escalier" pour le schéma de LAX, dû à la perturbation initiale, à laquelle les autres schémas sont pratiquement insensibles.

Au temps $t=1$, représenté sur la figure ci-contre, les résultats obtenus par le schéma quasi explicite (21), après 10 pas de temps sont de qualité comparable à ceux obtenus par le schéma décentré quasi d'ordre deux, après 40 pas de temps, sauf peut-être au niveau de la légère couche limite qui apparaît en $x = 1$. La condition en $x = 1$ n'est en effet plus satisfaite pour $t > 1$, ce qui apparaît déjà à cause de la perturbation initiale.

----- solution exacte
 Solution au temps $t=1$
 $h = 0.05$
 (q=1) --- LAX
 (q=1) - - - - - décentré
 (q=1/2) décentré quasi d'ordre deux
 (q=2) • • • quasi explicite (21)



Le schéma décentré généralisé à deux dimensions, étudié au théorème 15 a été testé sur le problème suivant

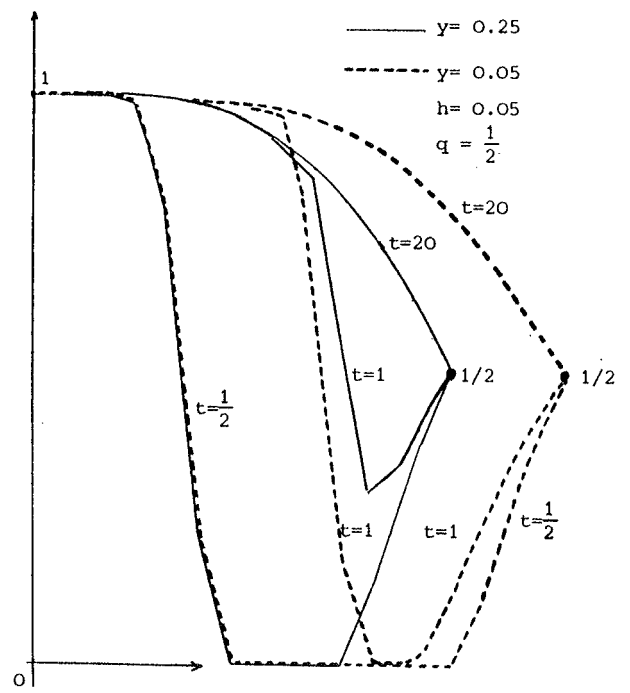
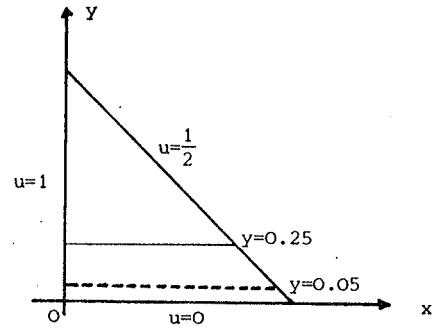
$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left(\frac{u^2}{2} \right) - \frac{\partial}{\partial y} \left(\frac{u^2}{2} \right) = 0, \quad (x,t) \in \Omega \times]0,T[$$

où Ω est l'intérieur du triangle rectangle isocèle déterminé par les axes et la droite d'équation $x+y = 1$. Les conditions aux limites sont non nulles, et données par

$u(0,y) = 1,$
 $u(x,x) = 0.5,$
 $u(x,0) = 0$ (non active).

La condition initiale est identiquement nulle. La condition en $x=0$ pénètre dans Ω sous la forme d'un choc, et la condition en $x=1-y$ sous la forme d'une détente.

Ce choc et cette détente se rencontrent pour ne former qu'une seule détente, dont la pente est de signe différent de la précédente. On a représenté ci-contre la solution pour $y = \frac{1}{4}$ et pour $y = h$. On constate qu'il n'y a pas de couche limite en $y = 0$, et que le choc obtenu a l'allure d'un choc calculé pour une équation mono-dimensionnelle, par un schéma décentré.



TROISIEME PARTIE

UNE EQUATION DES ONDES NON LINEAIRE

On applique, dans cette dernière partie, quelques idées développées précédemment, pour approcher une solution (faible) de l'équation des ondes non linéaire

$$\frac{\partial^2 w}{\partial t^2} = \frac{\partial}{\partial x} \left\{ f \left(\frac{\partial w}{\partial x} \right) \right\}.$$

Dans un premier chapitre, on construit une généralisation des invariants de RIEMANN, tenant compte à la fois des chocs et des détente, et on l'utilise pour résoudre le problème de RIEMANN, puis pour construire quelques schémas numériques.

Au chapitre suivant, on démontre la stabilité de ces schémas, sur une équation des ondes intervenant (comme modèle simplifié) en dynamique des câbles.

On fait ensuite quelques remarques sur les conditions aux limites et sur certains résultats numériques.

x x
x

CHAPITRE VII

APPROXIMATION DE L'EQUATION DES ONDES NON LINEAIRE

Pour les systèmes hyperboliques de la forme

$$\frac{\partial u}{\partial t} = \frac{\partial v}{\partial x} ; \quad \frac{\partial v}{\partial t} = \frac{\partial}{\partial x} f(u) ; \quad f' \geq 0$$

on suppose habituellement que f'' est de signe constant (genuinely non linear systems = systèmes vraiment non linéaires). Cette hypothèse permet de résoudre de façon simple le problème de RIEMANN associé. On étudie au chapitre suivant un modèle de câble extensible, où f'' change de signe pour $u = 0$, qui correspond à la position du câble au repos et qu'il est hors de question d'exclure...

Pour résoudre correctement le problème de RIEMANN dans le cas général, on commence par proposer une généralisation des invariants de RIEMANN, qui décrivent les détenteurs lorsqu'elles sont possibles, et les chocs lorsqu'ils satisfont à la condition d'entropie (cf. T.P. LIU [3], [4]. Cette construction utilise la condition d'entropie liée à l'équation scalaire et notamment son interprétation géométrique détaillée aux chapitres I et II.

Une fois le problème de RIEMANN résolu dans le cas où aucune hypothèse de convexité n'est faite sur f , on propose deux schémas, construits de façon analogue au schéma de GODUNOV, au chapitre II, et une construction du schéma de LAX. La stabilité de ces différents schémas est établie au chapitre VIII, dans le cadre du modèle d'équation de câble extensible.

1. UNE GENERALISATION DES INVARIANTS DE RIEMANN

On considère le problème de CAUCHY

$$(1) \quad \frac{\partial^2 w}{\partial t^2} = \frac{\partial}{\partial x} \left[f \left(\frac{\partial w}{\partial x} \right) \right], \quad x \in \mathbb{R}, \quad t \in]0, T[,$$

avec la condition initiale

$$(2) \quad w(x,0) = w_0(x), \quad x \in \mathbb{R},$$

$$(3) \quad \frac{\partial w}{\partial t}(x,0) = v_0(x), \quad x \in \mathbb{R},$$

On suppose que f est de classe $C^1(\mathbb{R})$, et non décroissante. Il s'agit là d'une hypothèse d'hyperbolicité, qui est fondamentale ici. En prenant par exemple $f(u) = -u$, l'équation (1) devient identique à l'équation de LAPLACE - $\Delta w = 0$, en les variables (x,t) , et une des deux conditions (2) et (3) est de trop. On utilisera dans la suite le fait que f' est non négative.

On suppose que v_0 et $u_0 = \frac{d}{dx} w_0$ sont à variation localement bornée sur \mathbb{R} , et uniformément bornées sur \mathbb{R} .

On transforme (1) en un système hyperbolique non linéaire, en posant $v = \frac{\partial w}{\partial t}$ et $u = \frac{\partial w}{\partial x}$. On obtient, en écrivant que le vecteur (u,v) est un gradient, et en reprenant l'équation (1), le système

$$(4) \quad \frac{\partial u}{\partial t} = \frac{\partial v}{\partial x},$$

$$(5) \quad \frac{\partial v}{\partial t} = \frac{\partial}{\partial x} f(u),$$

avec les conditions initiales

$$(6) \quad u(x,0) = u_0(x) \quad (= w'_0(x)),$$

$$(7) \quad v(x,0) = v_0(x).$$

On suppose aussi que $f(0)$ est nul, ce qui ne change en rien la généralité de (5).

Il est immédiat que les constantes sont solutions de (4), (5), et si par exemple, u ne dépend que de l'une des variables, v ne dépend que de l'autre, et on a nécessairement pour $a \neq 0$ quelconque

$$\text{soit } u = at + b, \quad v = ax + b',$$

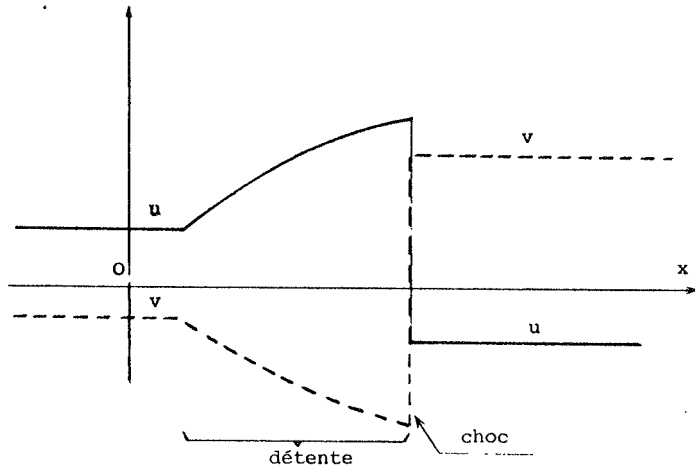
$$\text{soit } f(u) = ax + b, \quad v = at + b',$$

ce qui est exclu par les hypothèses faites sur les conditions initiales. Cette remarque étant faite, on ne rencontrera localement que des états constants, des chocs ou des détente.

On appellera "détente" une partie du graphe de (u,v) , pour t fixé, qui est connexe (pas de choc) et strictement monotone.

Sur l'intervalle considéré, u est alors une fonction en x strictement monotone, croissante ou décroissante, et régulière, et il en est de même de v . On peut donc exprimer v comme une fonction de u , en éliminant x , et obtenir une relation de la forme

$$(8) \quad v = g(u).$$



On peut connaître g en introduisant (8) dans (4) et (5), et il vient

$$\frac{\partial u}{\partial t} = g'(u) \frac{\partial u}{\partial x} ; \quad g'(u) \frac{\partial u}{\partial t} = f'(u) \frac{\partial u}{\partial x} ;$$

d'où

$$g'(u) = \pm \sqrt{f'(u)} ,$$

et on retrouve la condition d'hyperbolicité $f'(u) \geq 0$. On notera

$$(9) \quad g(u) = \int_0^u \sqrt{f'(y)} dy ,$$

et on remplacera (8) par $v = \pm g(u)$.

On peut remarquer qu'on se ramène ainsi à une équation scalaire en u . Par exemple, pour $f(u) = u^3/3$, on obtient $g(u) = u|u|/2$, et on retrouve l'équation de BURGERS. Cette remarque permet de construire des exemples explicites.

Lorsqu'il y a choc, on peut établir une condition de RANKINE HUGONIOT pour chacune des équations (4) et (5), et les lier par une condition de compatibilité en écrivant que la vitesse de propagation est la même pour les deux conditions obtenues. Il vient les équations suivantes, en notant (u_0, v_0) et (u_1, v_1) les valeurs de chaque côté de la ligne de discontinuité d'équation $x = x(t)$,

$$-\frac{dx}{dt} = \frac{v_1 - v_0}{u_1 - u_0} = \frac{f(u_1) - f(u_0)}{v_1 - v_0} ,$$

dont on déduit la condition de compatibilité

$$(10) \quad v_1 - v_0 = \pm \sqrt{(f(u_1) - f(u_0))(u_1 - u_0)}$$

qu'il convient de rapprocher de (8) et de (9).

Dans le plan de phase, c'est-à-dire dans le plan (u, v) , on appelle invariant de RIEMANN une courbe de la forme

$$(11) \quad v \pm g(u) = c^{te}$$

avec $g(u)$ donné en (9).

L'importance de ces invariants de RIEMANN est liée à la diagonalisation du système (4), (5) mis sous forme non conservative (en effectuant la dérivation $\frac{\partial}{\partial x} f(u) = f'(u) \frac{\partial u}{\partial x}$) ; on obtient que les couples (u, v) vérifiant (11) sont des solutions particulières. Bien que le système soit alors non conservatif dans le cas non linéaire, le fait que les invariants de RIEMANN satisfassent au système diagonalisé, incite à les utiliser au niveau de la stabilité des schémas ; on y reviendra plus loin, en (33). On trouve dans A. JEFFREY une étude assez complète de ces invariants de RIEMANN.

On appellera *courbe de choc* un arc de courbe issu de (u_1, v_1) (ou de (u_0, v_0)), d'équation (10) en (u_0, v_0) (ou en (u_1, v_1)). En particulier si (u_0, v_0) est relié par une courbe de choc à (u_1, v_1) , la réciproque est autorisée par (10), (u_1, v_1) peut être relié par le même choc à (u_0, v_0) ;

on retrouve la symétrie de la condition de RANKINE HUGONIOT, et il faudra faire intervenir une condition d'entropie pour la détruire. Cette condition d'entropie va être introduite géométriquement. On va réunir les courbes définies par (10) et (9), (11) en une seule équation, celle des invariants de RIEMANN généralisés.

On considère d'abord le cas où v est une fonction croissante de u , valant v_0 pour $u = u_0$. On a, pour une détente

$$(12) \quad v = v_0 + \int_{u_0}^u \sqrt{f'(y)} dy = v_0 + g(u) - g(u_0) \quad ,$$

et u vérifie l'équation scalaire

$$\frac{\partial u}{\partial t} - g'(u) \frac{\partial u}{\partial x} = 0.$$

La propagation, dont le sens est donné par le signe de $-g'$, est négative. De façon à s'opposer le plus possible à l'apparition d'éventuelles discontinuités, on pourrait, comme pour l'équation scalaire, remplacer $-g$ par son enveloppe convexe sur $[u_0, u]$ si $u > u_0$, ou par son enveloppe concave sur $[u, u_0]$ si $u < u_0$. Cependant, ceci ne permet pas, lorsqu'un choc subsiste, de retrouver la relation (10) sur ce choc. On est en effet amené à faire intervenir f plutôt que g dans la construction de ces différentes enveloppes, ce qui peut être effectué de la façon suivante, lorsque le sens de propagation est négatif.

On remplace f par son enveloppe convexe sur $[u_0, u]$ si $u_0 < u$, et par son enveloppe concave sur $[u, u_0]$ si $u_0 > u$, et on note $\tilde{f}_+(y; u_0, u)$ la valeur de l'enveloppe ainsi construite, au point $y \in I(u_0, u)$. Tout comme f , $\tilde{f}_+(\cdot; u_0, u)$ est croissante et de classe C^1 sur $I(u_0, u)$. On lui associe une fonction \tilde{g} définie sur $I(u_0, u)$ par

$$(13) \quad \tilde{g}(z; u_0, u) = \int_{u_0}^z \sqrt{\frac{d}{dy} \tilde{f}_+(y; u_0, u)} dy \quad ,$$

puis on pose

$$(14) \quad g(u; u_0) = \tilde{g}(u; u_0, u) \quad .$$

Rappelons que le choix du sens de la convexité de l'enveloppe \tilde{f}_+ en fonction du signe de $(u_0 - u)$ est lié au sens de propagation qu'on s'est fixé a priori.

Si par exemple f est convexe sur $[u_0, u]$, il y a coïncidence entre $\tilde{f}_+(\cdot; u_0, u)$ et f , et donc

$$g(u; u_0) = \int_{u_0}^u \sqrt{f'(u)} du \quad .$$

On retrouve la fonction g définie par (9). Si maintenant f est concave, on obtient

$$\frac{d}{dy} \tilde{f}_+(y; u_0, u) = \frac{f(u_0) - f(u)}{u_0 - u}$$

puis

$$g(u; u_0) = \sqrt{(f(u_0) - f(u)) (u_0 - u)} \quad ,$$

c'est-à-dire le second membre de (10).

Par exemple, pour $f(u) = u^3/3$, on obtient

$$(15) \quad g(u; u_0) = \begin{cases} (u - u_0) \sqrt{\frac{u^2 + u u_0 + u_0^2}{3}} & \text{si } u \in I(-u_0/2, u_0) \text{ ,} \\ \text{sg}(u_0) (u^2 - u_0^2) / 2 & \text{si } u < u_0 < 0 \text{ ou } u > u_0 > 0 \text{ ,} \\ -\text{sg}(u_0) \left(\frac{u_0^2}{2} + \frac{5u^2}{8} \right) & \text{si } u < -\frac{u_0}{2} < 0 \text{ ou } u > -\frac{u_0}{2} > 0 \text{ .} \end{cases}$$

Remarquons que g croît en u et décroît en u_0 .

L'invariant de RIEMANN croissant généralisé, issu de (u_0, v_0) , est l'arc d'équation suivante dans le plan de phase :

$$(16) \quad v = v_0 + g(u; u_0) .$$

On considère maintenant le cas où v s'exprime comme fonction décroissante de u . La fonction $(-g)$ est maintenant croissante, et la propagation est de vitesse positive ; les rôles de u et u_0 sont cependant échangés car u est la valeur à droite, et u_0 , le point d'appui, la valeur à gauche. On retrouve donc la même fonction g que précédemment.

L'invariant de RIEMANN décroissant généralisé issu du point (u_0, v_0) est, dans le plan de phase, l'arc d'équation

$$(17) \quad v = v_0 - g(u; u_0) .$$

Si f est convexe sur $[u_0, u]$, on retrouve l'invariant de RIEMANN classique, et si f est concave sur $[u_0, u]$, on obtient la condition de choc (10).

On peut remarquer que pour u et v fixés, (16) donne v_0 comme une fonction croissante de u_0 , et (17) donne v_0 comme une fonction décroissante de u_0 .

2. LE PROBLEME DE RIEMANN

On se propose de résoudre le système (4), (5), pour les données initiales

$$(18) \quad u(x,0) = \begin{cases} u_+ & \text{si } x > 0 , \\ u_- & \text{si } x < 0 , \end{cases}$$

$$v(x,0) = \begin{cases} v_+ & \text{si } x > 0 , \\ v_- & \text{si } x < 0 , \end{cases}$$

où (u_+, v_+) et (u_-, v_-) représentent les points notés A_+ et A_- dans le plan

de phase. Ces points sont supposés fixes. On cherche particulièrement à connaître la valeur stationnaire (u_0, v_0) que prendra la solution le long de la demi-droite $\{x=0, t>0\}$, afin de construire un schéma qui généralisera le schéma de GODUNOV. Cette valeur stationnaire servira également de point d'appui pour construire la solution sur les $x > 0$ et les $x < 0$, à partir des invariants de RIEMANN généralisés qui en sont issus.

Ainsi le point A_+ doit être situé sur un invariant de RIEMANN généralisé décroissant, issu de (u_0, v_0) , et on a donc

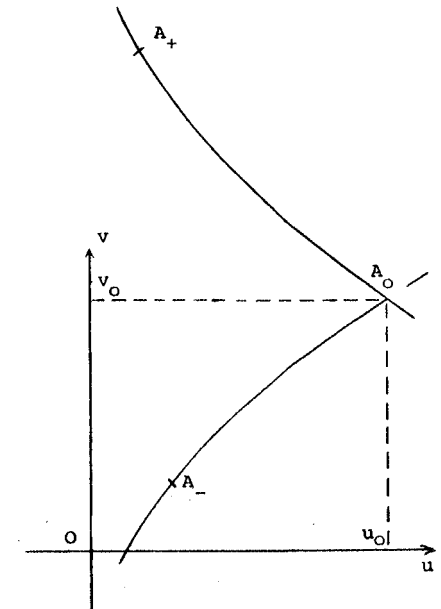
$$(19) \quad v_+ = v_0 - g(u_+; u_0) ,$$

pour assurer une propagation de sens positif.

De la même façon, A_- est situé sur un invariant de RIEMANN généralisé croissant, et on a

$$(20) \quad v_- = v_0 + g(u_-; u_0) ,$$

de façon à avoir une propagation vers les x négatifs.



Ces deux équations (19) et (20) vont permettre d'obtenir (u_0, v_0) à condition que les graphes aient une intersection, ce qui est assuré sous l'hypothèse de stricte hyperbolicité suivante.

On dira que le système (4), (5) est *strictement hyperbolique* lorsque f est strictement croissante, et

$$(21) \quad \lim_{R \rightarrow \pm \infty} \int_0^R \sqrt{f'(y)} dy = \infty .$$

Le terme "strictement" peut être justifié en précisant que (21) et la stricte croissance de f , expriment que le système (4), (5) ne se rapproche jamais d'un système parabolique ou elliptique. En effet, f' n'est pas trop petit ni sur un ensemble de mesure non nulle, ni à l'infini, et $f' \leq 0$ signifie que le système est elliptique, $f' = 0$ que le système peut être interprété comme étant parabolique.

Ainsi, pour les systèmes étudiés par NISHIDA [1] et NISHIDA et SMOLLER [1], [2] où $f(u) = -1/u^{1+\varepsilon}$, avec $u > 0$ et $\varepsilon \geq 0$, (21) n'est vérifié que pour $\varepsilon = 0$. Pour $\varepsilon > 0$, ces auteurs sont amenés à faire une hypothèse sur la variation de la condition initiale, à savoir que $[v_+ - v_-]$ est petit, ce qui est une façon d'assurer l'existence de (u_0, v_0) .

Si le système (4), (5) est strictement hyperbolique, le système d'équations (19), (20) admet une solution unique (u_0, v_0) . La solution du problème de RIEMANN est obtenue, sur $\{x \geq 0\}$ en résolvant lorsque l'invariant de RIEMANN coïncide avec l'invariant généralisé, le problème

$$\frac{\partial u}{\partial t} + \sqrt{f'(u)} \frac{\partial u}{\partial x} = 0, \quad u(0, t) = u_0, \quad u(x, 0) = u_+,$$

pour avoir u , et v est obtenu à partir de (19). Si l'invariant de RIEMANN coïncide avec (10), un choc apparaît, et sa vitesse est déterminée par la condition de RANKINE HUGONIOT. Son intensité dépend de la distance entre les points de contact de la fonction avec son enveloppe.

Dans le cas général, on commence par séparer les chocs des détenteurs en utilisant (19) et (20), puis on résout le problème séparément pour chaque cas. On procède de la même façon sur $\{x < 0\}$. On peut remarquer que dans la pratique, la fonction f n'est pas trop compliquée, et il n'y a souvent soit qu'un choc, soit qu'une détente sur $\{x \geq 0\}$, et de même sur $\{x \leq 0\}$. On reviendra sur cette partie au chapitre VIII pour une fonction f plus particulière.

Donnons deux exemples, pour $f(u) = u^3/3$. On prend tout d'abord $(u_-, v_-) = (0, 0)$, puis $(u_+, v_+) = (0, 1)$. Alors u_0 et v_0 doivent vérifier

$$1 = v_0 - g(0; u_0),$$

$$0 = v_0 + g(0; u_0).$$

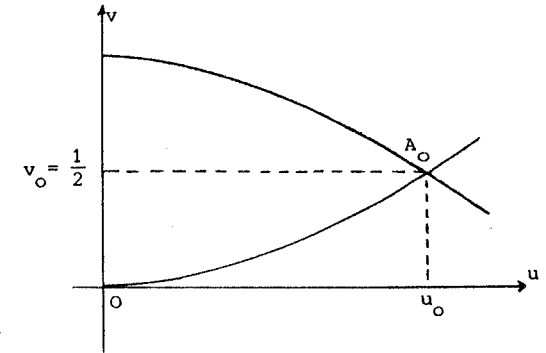
On a immédiatement

$$v_0 = \frac{1}{2}, \quad u_0 > 0$$

et u_0 vérifie l'équation

$$u_0^2 = \frac{\sqrt{3}}{2},$$

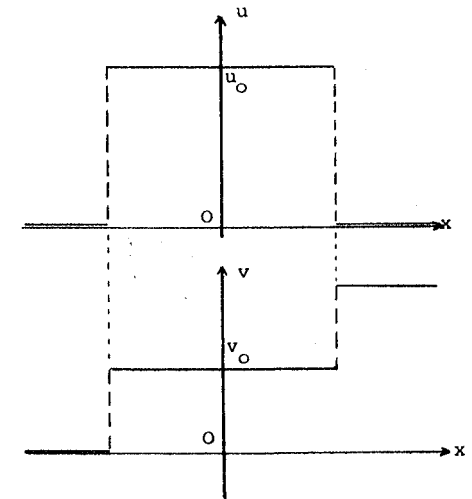
$$\text{d'où } u_0 = \frac{\sqrt{\sqrt{3}}}{2} \sim 0.9306.$$



On voit apparaître deux chocs, de vitesses respectives $-v_0/u_0$ et v_0/u_0 . La solution est donnée par

$$u(x, t) = \begin{cases} 0 & \text{si } x < -\frac{v_0}{u_0} t, \\ u_0 & \text{si } -\frac{v_0}{u_0} t < x < \frac{v_0}{u_0} t, \\ 0 & \text{si } x > \frac{v_0}{u_0} t, \end{cases}$$

$$v(x, t) = \begin{cases} 0 & \text{si } x < -\frac{v_0}{u_0} t, \\ v_0 & \text{si } -\frac{v_0}{u_0} t < x < \frac{v_0}{u_0} t, \\ 1 & \text{si } x > \frac{v_0}{u_0} t. \end{cases}$$



Pour $(u_-, v_-) = (0, 0)$ et $(u_+, v_+) = (1, 0)$, on obtient un choc de vitesse négative, et une détente dont la vitesse de propagation est positive. Il vient

$$u_0 = \sqrt{\frac{\sqrt{3}}{2 + \sqrt{3}}} \sim 0.681,$$

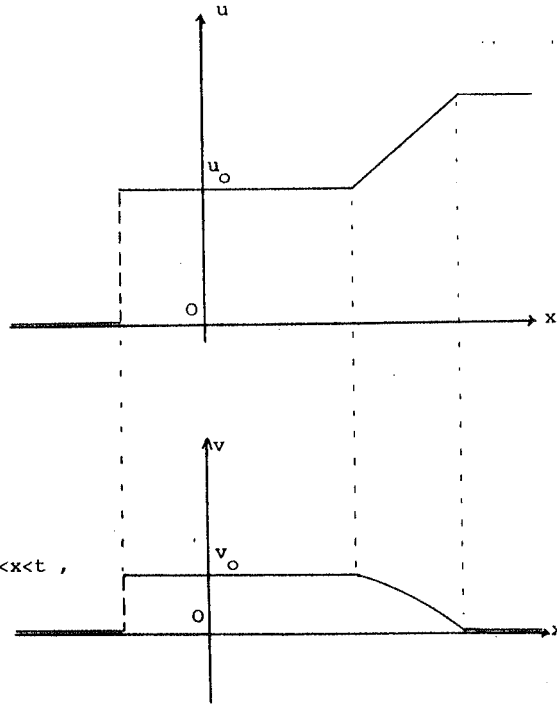
$$v_0 = \frac{1}{2 + \sqrt{3}} \sim 0.268,$$

et la solution est donnée par

$$u(x,t) = \begin{cases} 0 & \text{si } x < \alpha t, \\ u_0 & \text{si } \alpha t < x < u_0 t, \\ x/t & \text{si } u_0 t < x < t, \\ 1 & \text{si } x > t, \end{cases}$$

$$v(x,t) = \begin{cases} 0 & \text{si } x < \alpha t, \\ v_0 & \text{si } \alpha t < x < u_0 t, \\ v_0 - \frac{x^2}{2t^2} + \frac{u_0^2}{2} & \text{si } u_0 t < x < t, \\ 0 & \text{si } x > t, \end{cases}$$

avec $\alpha = -\frac{v_0}{u_0} \sim -0.393$.



On vérifie, et ceci est général pour le problème de RIEMANN, que la solution s'exprime en fonction de la variable $\xi = x/t$, comme dans le cas scalaire (cf. Chapitre II, paragraphe 1). En reprenant les articles [1] et [4] de C.M. DAFERMOS et [3] de T.P. LIU (entre autres), cette remarque permet d'obtenir une condition d'entropie, qui correspond à celle incluse implicitement dans la construction géométrique des invariants de RIEMANN généralisés qu'on vient de faire.

En effet, les chocs obtenus satisfont à la condition d'entropie proposée par T.P. LIU [3] ; en notant (u_-, v_-) et (u_+, v_+) les valeurs en $x = x(t) - 0$ et $x = x(t) + 0$ de la solution de chaque côté d'une ligne de discontinuité d'équation $x = x(t)$, on vérifie qu'on a, pour tout (k, ℓ) entre (u_-, v_-) et (u_+, v_+) tel que

$$(22) \quad \frac{f(k) - f(u_-)}{\ell - v_-} = \frac{\ell - v_+}{k - u_-} = \sigma(k, \ell; u_-, v_-),$$

l'inégalité

$$(23) \quad \sigma(k, \ell; u_-, v_-) \leq \sigma(u_+, v_+; u_-, v_-) \quad (= -\frac{dx}{dt}).$$

Supposons par exemple que $\sigma(k, \ell; u_-, v_-)$ soit non négatif.

On a alors

$$[\sigma(k, \ell; u_-, v_-)]^2 = \frac{f(k) - f(u_-)}{k - u_-},$$

et la condition (23) exprime que

$$\text{Max}_{k \in I(u_+, u_-)} \left\{ \frac{f(k) - f(u_-)}{k - u_-} \right\} = \frac{f(u_+) - f(u_-)}{u_+ - u_-}.$$

(Il apparaît un minimum pour un choc de vitesse positive)

On retrouve une condition d'entropie analogue au cas scalaire. En se plaçant par exemple dans le cas où f est convexe, on a alors $u_+ > u_-$, et c'est ce qu'on obtient sur les exemples ci-dessus. On retrouve encore (23) dans les autres cas de figure possibles.

On peut démontrer un théorème d'existence, et de stabilité pour le problème de RIEMANN dans le cas général. L'existence est obtenue par la construction précédente. La stabilité exprime que la solution est à valeurs dans le domaine limité par les invariants de RIEMANN contenant la condition initiale. Ce résultat est établi au théorème 19 dans le cadre de l'équation des câbles, pour limiter le nombre de cas à étudier, mais la même démonstration s'adapte au cas général. Notons également que la solution obtenue satisfait à la condition d'entropie de T.P. LIU, ce qui constitue un critère d'unicité.

L'utilisation essentielle que l'on fera, de cette résolution du problème de RIEMANN, est la construction de (u_0, v_0) à partir de (19)

et (20). Numériquement, ces deux équations sont difficiles à résoudre ; en remarquant que l'invariant de RIEMANN classique est tangent en (u_+, v_+) et en (u_-, v_-) aux invariants généralisés (19) et (20) respectivement, et ont même type de convexité, on peut ne prendre dans (19) et (20) que les invariants de RIEMANN classiques, sans commettre une trop grande erreur lorsque (u_+, v_+) et (u_-, v_-) ne sont pas trop éloignés. Par exemple, ces invariants de RIEMANN classiques donnent pour le premier exemple

$$1 = \bar{v}_0 + \frac{\bar{u}_0^2}{2} ; \quad 0 = \bar{v}_0 - \frac{\bar{u}_0^2}{2} ;$$

donc $\bar{v}_0 = \frac{1}{2}$ ($= v_0$) et $\bar{u}_0 = 1$ (on avait $u_0 = 0.9306$) .

Pour le second exemple, il vient

$$\bar{v}_0 = \frac{1}{4} \quad (v_0 = 0.268) \quad \bar{u}_0 = \frac{1}{\sqrt{2}} \sim 0.707 \quad (\text{on avait } u_0 = 0.681).$$

La différence est encore plus réduite lorsque (u_-, v_-) et (u_+, v_+) représentent la valeur de la solution approchée en deux pas consécutifs du maillage.

Cette remarque permet de construire un schéma numérique d'utilisation plus facile.

3. CONSTRUCTION D'UN SCHEMA NUMERIQUE

Soit $h > 0$ le pas de discrétisation en espace ; le pas de discrétisation en temps est encore pris égal à qh . On introduit également les mêmes intervalles que précédemment

$$I_i = \left[\left(i - \frac{1}{2}\right) h, \left(i + \frac{1}{2}\right) h \right] \quad i \in \mathbb{Z},$$

$$J_n = \left[\left(n - \frac{1}{2}\right) qh, \left(n + \frac{1}{2}\right) qh \right] \quad n \in \left\{ 0, \dots, N = \left[\frac{T}{qh} \right] \right\}.$$

On approche les conditions initiales u_0 et v_0 par

$$(24) \quad u_i^0 = \frac{1}{h} \int_{I_i} u_0(x) dx \quad i \in \mathbb{Z},$$

et

$$(25) \quad v_i^0 = \frac{1}{h} \int_{I_i} v_0(x) dx \quad i \in \mathbb{Z}.$$

On appellera solution approchée, un couple (u_h, v_h) , constant sur chaque pavé $I_i \times J_n$, où sa valeur est notée (u_i^n, v_i^n) . Pour $n = 0$, ces valeurs sont données par (24) et (25). On suppose connaître tous les u_i^n et tous les v_i^n pour n fixé, et on construit les u_i^{n+1} et v_i^{n+1} de la façon suivante.

On résout le problème de RIEMANN (4), (5) avec les conditions initiales $(u_-, v_-) = (u_i^n, v_i^n)$ et $(u_+, v_+) = (u_{i+1}^n, v_{i+1}^n)$, pour chaque $i \in \mathbb{Z}$, sur le segment

$$\left\{ x = \left(i + \frac{1}{2}\right) h, t \geq nqh \right\},$$

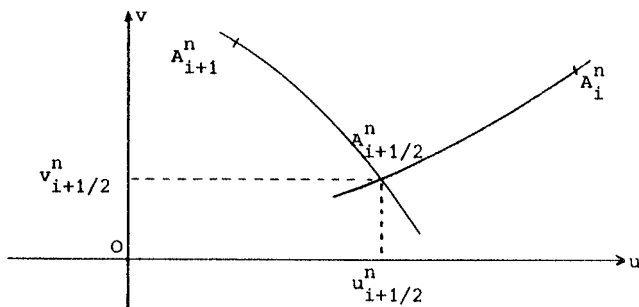
comme on l'a fait au chapitre II, pour le schéma de GODUNOV. Cette valeur est constante pour $(t - nqh)$ assez petit, de façon à éviter l'influence des u_j^n pour j différent de i et $i+1$, et ceci va constituer une condition de stabilité. La valeur cherchée est notée $(u_{i+1/2}^n, v_{i+1/2}^n)$; elle est solution de (19) et (20), c'est-à-dire

$$(26) \quad v_{i+1}^n = v_{i+1/2}^n - g(u_{i+1}^n ; u_{i+1/2}^n),$$

$$v_i^n = v_{i+1/2}^n + g(u_i^n ; u_{i+1/2}^n),$$

ce qui est assez difficile à calculer en pratique, sauf peut être dans certains cas où la fonction f s'y prête bien. Comme au cours de l'expérience numérique, (u_{i+1}^n, v_{i+1}^n) et (u_i^n, v_i^n) ne sont pas très éloignés dans le plan de phase, on ne commet pas une très grande erreur en la calculant

non plus avec les invariants de RIEMANN généralisés, mais avec les invariants de RIEMANN classiques, conformément à la remarque faite à la fin du paragraphe précédent, où on a vu sur quelques exemples que cette erreur est assez faible.



On prendra donc

$$(27) \quad \begin{cases} v_{i+1}^n = v_{i+1/2}^n - g(u_{i+1}^n) + g(u_{i+1/2}^n) \\ v_i^n = v_{i+1/2}^n + g(u_i^n) - g(u_{i+1/2}^n) \end{cases}$$

au lieu de (26), avec g définie par

$$(28) \quad g(u) = \int_0^u \sqrt{f'(\xi)} d\xi$$

Le calcul de \$(u_{i+1/2}^n, v_{i+1/2}^n)\$ par (27) est nettement plus facile à programmer ; il suffit en effet de résoudre

$$(29) \quad g(u_{i+1/2}^n) = \frac{1}{2} \left\{ v_{i+1}^n - v_i^n + g(u_{i+1}^n) + g(u_i^n) \right\}$$

soit directement si g est assez simple, soit par la méthode de NEWTON qui est bien adaptée ici, la dérivée de g étant connue et égale à \$\sqrt{f'}\$.

Le calcul de \$v_{i+1/2}^n\$ est encore plus simple :

$$(30) \quad v_{i+1/2}^n = \frac{1}{2} \left\{ v_{i+1}^n + v_i^n + g(u_{i+1}^n) - g(u_i^n) \right\}$$

Ayant \$u_{i+1/2}^n\$ et \$v_{i+1/2}^n\$ pour tout \$i \in \mathbb{Z}\$, on intègre les équations (4) et (5) sur le pavé

$$\left] (i - \frac{1}{2})h, (i + \frac{1}{2})h \right[\times \left] nqh, (n+1)qh \right[,$$

en utilisant une formule de GREEN. On note ensuite \$u_i^{n+1}\$ la moyenne de cette solution sur \$I_i\$, au temps \$t = (n+1)qh\$. Il vient, comme pour le schéma de GODUNOV dans le cadre des équations scalaires,

$$(31) \quad u_i^{n+1} = u_i^n + q(v_{i+1/2}^n - v_{i-1/2}^n)$$

$$(32) \quad v_i^{n+1} = v_i^n + q(f(u_{i+1/2}^n) - f(u_{i-1/2}^n))$$

On notera \$S_0\$ le schéma (27,31,32) et \$S_1\$ le schéma (26,31,32).

La condition de stabilité peut être obtenue en considérant les vitesses de propagation respectives de u et v, qui sont obtenues en diagonalisant le système (4), (5) mis sous la forme non conservative

$$\frac{\partial u}{\partial t} = \frac{\partial v}{\partial x}$$

$$\frac{\partial v}{\partial t} = f'(u) \frac{\partial u}{\partial x}$$

c'est-à-dire

$$\frac{\partial}{\partial t} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ f'(u) & 0 \end{pmatrix} \frac{\partial}{\partial x} \begin{pmatrix} u \\ v \end{pmatrix}$$

Le système diagonalisé est non conservatif ; il est donné par

$$(33) \quad \begin{aligned} \frac{\partial}{\partial t} \{v + g(u)\} &= \sqrt{f'(u)} \frac{\partial}{\partial x} \{v + g(u)\} \\ \frac{\partial}{\partial t} \{v - g(u)\} &= -\sqrt{f'(u)} \frac{\partial}{\partial x} \{v - g(u)\} \end{aligned}$$

qu'on peut obtenir directement en introduisant les invariants de RIEMANN dans (4) et (5). Ce système est à l'origine de l'utilisation des invariants de RIEMANN ; il exprime que les quantités $v + g(u)$ et $v - g(u)$ restent bornées dans L^∞ , et se propagent avec les vitesses respectives $\sqrt{f'(u)}$ et $-\sqrt{f'(u)}$. La condition de stabilité peut donc s'écrire

$$(34) \quad q \sup_{u \in \mathbb{R}} \sqrt{f'(u)} \leq 1 .$$

La diagonalisation (33) n'est autorisée que lorsque la solution est régulière ; s'il y a choc, on reprend la condition (10), pour établir que sa vitesse de propagation vérifie

$$(35) \quad \left| \frac{dx}{dt} \right| = \sqrt{\frac{f(u_1) - f(u_0)}{u_1 - u_0}} ,$$

et conduit à la même condition de stabilité que (34).

Pour le schéma S_0 ou pour le schéma S_1 , il n'y a pas stabilité de type L^∞ sur u et v séparément ; en effet, sur un des exemples donnés au paragraphe précédent, on avait à la fois u_0 constant et égal à zéro et $u(\cdot, t)$ non nul. On ne peut concevoir de stabilité que sur les invariants de RIEMANN (généralisés ou non). Etant donnés u_0 et v_0 dans $L^\infty(\mathbb{R})$, on note

$$(36) \quad K_0 = \left\{ (u, v) \mid |v| + |g(u)| \leq M_0 = \max_{x \in \mathbb{R}} \left[|v_0(x)| + |g[u_0(x)]| \right] \right\} .$$

Cet ensemble est un compact du plan de phase. On dira qu'il y a stabilité lorsque la solution (u, v) est à valeurs dans K_0 .

Cette solution est également envisagée dans un sens faible comme au chapitre I, on appellera solution faible une solution de (4), (5) au sens des distributions, qui est mesurable et bornée. On aura pour toute fonction test

$$(\psi, \phi) \left\{ C_0^2(\mathbb{R} \times [0, T]) \right\}^2 ,$$

les relations

$$(37) \quad \int_{\mathbb{R}} \int_0^T (u \frac{\partial \psi}{\partial t} - v \frac{\partial \psi}{\partial x}) dx dt + \int_{\mathbb{R}} u_0(x) \psi(x, 0) dx = 0 ,$$

$$\int_{\mathbb{R}} \int_0^T (v \frac{\partial \phi}{\partial t} - f(u) \frac{\partial \phi}{\partial x}) dx dt + \int_{\mathbb{R}} v_0(x) \phi(x, 0) dx = 0 .$$

Il n'y a pas unicité des solutions faibles, en général ; l'exemple suivant le montre. On prend $f(u) = u^3/3$, puis on résout le problème de RIEMANN avec $(u_-, v_-) = (0, 0)$ et $(u_+, v_+) = (1, 0)$, déjà traité au paragraphe précédent. On trouve la solution déjà donnée, constituée d'un choc et d'une détente. On peut également concevoir la suivante, constituée de deux chocs. On choisit u_1 égal à l'unique solution dans $]0, 1[$ de l'équation

$$u^2 = (1-u) \sqrt{1 + u + u^2}$$

donc $u_1^3 + u_1 - 1 = 0$, $u_1 \approx 0.6823$,

puis $v_1 = u_1^2/\sqrt{3}$, $v_1 \approx 0.2688$,

et on a pour solution faible, la solution suivante, qui vérifie la condition de RANKINE HUGONIOT (10), et dont la valeur en $x = 0, t > 0$ diffère de la valeur correspondante u_0 pour l'autre solution faible proposée.

$$u(x, t) = \begin{cases} 0 & \text{si } x < \alpha t \\ u_1 & \text{si } \alpha t < x < \beta t \\ 1 & \text{si } x > \beta t \end{cases} ;$$

$$v(x, t) = \begin{cases} 0 & \text{si } x < \alpha t \\ v_1 & \text{si } \alpha t < x < \beta t \\ 0 & \text{si } x > \beta t \end{cases} ;$$

avec $\alpha = -\frac{v_1}{u_1}$, $\beta = \frac{v_1}{1-u_1}$.

($\alpha \approx 0.394$) ($\beta \approx 0.846$)

4. QUELQUES REMARQUES ET RESULTATS NUMERIQUES

Ce paragraphe fait l'objet de quelques remarques concernant le schéma S_0 . On y donnera également quelques résultats numériques sur ce schéma, et sur le schéma de LAX. La convergence pour ces schémas est considérée au chapitre VIII, dans un cadre assez particulier qui permet d'avoir K_0 convexe.

On fait souvent une hypothèse sur la convexité de f . Ainsi LAX a introduit la notion de "genuinely non linear systems", qu'on peut traduire par "systèmes vraiment non linéaires" bien qu'il s'agisse d'une définition plus restrictive, lorsque dans (5) la fonction f admet une dérivée monotone, croissante ou décroissante. Ce n'est pas le cas par exemple de $f(u) = u^3/3$. Pour les systèmes vraiment non linéaires, le problème de RIEMANN est plus facile à résoudre, car $g(u; u_0)$ est alors soit donné par l'invariant de RIEMANN classique, soit par la condition de compatibilité (10) sur les chocs. On ne fait pas cette hypothèse ici ; on étudie en effet au chapitre suivant un modèle de câble qui généralise le cas où $f(u) = u^3/3$ et où $u = 0$ correspond à un changement de monotonie de f' et également à une position d'équilibre, si bien qu'il est exclu de supposer u de signe constant.

x x
x

Dans le cas linéaire, l'équation (1) correspond à l'équation des ondes

$$\frac{\partial^2 w}{\partial t^2} = k^2 \frac{\partial^2 w}{\partial x^2}$$

où k est une constante positive fixée ; on a alors $f(u) = k^2 u$, et le système (4), (5) devient

$$(38) \quad \frac{\partial u}{\partial t} = \frac{\partial v}{\partial x} ; \quad \frac{\partial v}{\partial t} = k^2 \frac{\partial u}{\partial x}$$

Les invariants de RIEMANN classiques et généralisés coïncident, et le schéma S_0 ($= S_1$) s'écrit :

$$v_{i+1}^n = v_{i+1/2}^n - k u_{i+1}^n + k u_{i+1/2}^n ,$$

$$v_i^n = v_{i+1/2}^n + k u_i^n - k u_{i+1/2}^n ,$$

d'où

$$v_{i+1/2}^n = \frac{1}{2} (v_{i+1}^n + v_i^n) + \frac{k}{2} (u_{i+1}^n - u_i^n) ,$$

$$u_{i+1/2}^n = \frac{1}{2} (u_{i+1}^n + u_i^n) + \frac{1}{2k} (v_{i+1}^n - v_i^n) .$$

En introduisant ces valeurs dans (31) et (32), il vient

$$u_i^{n+1} = u_i^n + \frac{g}{2} (v_{i+1}^n - v_{i-1}^n) + \frac{gk}{2} (u_{i+1}^n - 2u_i^n + u_{i-1}^n) ,$$

$$v_i^{n+1} = v_i^n + \frac{gk^2}{2} (u_{i+1}^n - u_{i-1}^n) + \frac{gk}{2} (v_{i+1}^n - 2v_i^n + v_{i-1}^n) .$$

Ces deux relations font apparaître des termes de viscosité. On peut également calculer $(v + ku)_i^{n+1}$ et $(v - ku)_i^{n+1}$; on obtient

$$\begin{aligned} (v + ku)_i^{n+1} &= v_i^n + k u_i^n + \frac{gk}{2} \left[v_{i+1}^n - 2v_i^n + v_{i-1}^n + v_{i+1}^n - v_{i-1}^n \right] \\ &\quad + \frac{gk^2}{2} \left[u_{i+1}^n - u_{i-1}^n + u_{i+1}^n - 2u_i^n + u_{i-1}^n \right] \\ &= (v + ku)_i^n + gk \left[(v+ku)_{i+1}^n - (v + ku)_i^n \right] . \end{aligned}$$

De la même façon,

$$(v - ku)_i^{n+1} = (v - ku)_i^n - gk \left[(v - ku)_i^n - (v - ku)_{i-1}^n \right] .$$

On retrouve dans chaque cas le schéma décentré utilisé pour l'équation scalaire. On peut conclure, par le théorème 2 que $z_h = v_h + ku_h$ et $y_h = v_h - ku_h$ convergent vers z et y respectivement, dans

$L^1_{loc}(\mathbb{R} \times]0, T[)$, puis en posant

$$v_h = \frac{1}{2} (z_h + y_h) \quad ; \quad u_h = \frac{1}{2k} (z_h - y_h) \quad ;$$

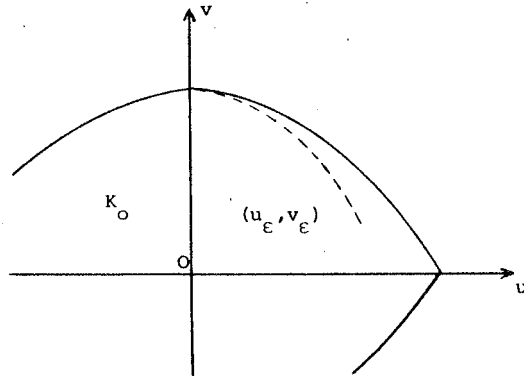
établir la convergence de (u_h, v_h) , vers $(u, v) = \frac{1}{2} \left(\frac{z-y}{k}, z+y \right)$ qui est solution faible de (38).

x x
x

On peut faire une autre remarque concernant le choix de l'invariant de RIEMANN classique dans (27), plutôt par exemple que la courbe de choc (10). Le principe de la condition d'entropie consiste essentiellement à s'opposer à l'apparition de chocs. Dans le cadre de l'équation scalaire, on a utilisé ce principe pour introduire les enveloppes convexes et concaves de f , suivant le sens de la discontinuité initiale. Ici on va s'opposer à l'apparition d'un choc en traitant a priori chaque problème de RIEMANN comme si sa solution n'était constituée que de détente. Ce procédé n'exclut pas d'obtenir des chocs, mais se traduit par un apport de viscosité. Ainsi les chocs qui resteront après cette opération vont satisfaire à la condition d'entropie de T.P. LIU (23).

L'application de ce principe doit être rapproché d'un article de CHUEH, CONLAY et SMOLLER, dans lequel le système parabolique du type

$$(39) \quad \begin{aligned} \frac{\partial u_\epsilon}{\partial t} &= \frac{\partial v_\epsilon}{\partial x} + \epsilon \frac{\partial^2 u_\epsilon}{\partial x^2} \\ \frac{\partial v_\epsilon}{\partial t} &= \frac{\partial}{\partial x} (f(u_\epsilon)) + \epsilon \frac{\partial^2 v_\epsilon}{\partial x^2} \end{aligned}$$



est étudié. Il apparaît que le fait que ϵ soit positif suffit pour que les valeurs de (u_ϵ, v_ϵ) soient toutes situées à l'intérieur du compact K_0 , limité par les

invariants de RIEMANN classiques. Lorsque ϵ tend vers zéro, ces valeurs (u_ϵ, v_ϵ) restent évidemment dans K_0 , et ceci implique en particulier que les courbes de chocs physiquement acceptables ne peuvent pas être situées à l'extérieur de K_0 . (27) peut ainsi être interprété comme une projection sur K_0 , au cas où le problème de RIEMANN donnerait un point (u_0, v_0) n'appartenant pas à K_0 .

x x
x

Le schéma (27), (31), (32) ($=S_0$) se différencie du schéma de GLIMM qui est habituellement employé. Tout d'abord on n'utilise pas de générateur de nombres aléatoires, qu'on remplace par une projection de type L^2 , au niveau de (31) et de (32). Ensuite, dans (27), on ne calcule que deux valeurs $u_{i+1/2}^n$ et $v_{i+1/2}^n$, au lieu de déterminer complètement la solution du problème de RIEMANN à chaque fois. Ce schéma de GLIMM permet cependant d'établir la convergence dans le cadre d'un système "vraiment non linéaire", et donne des résultats numériques satisfaisants. On peut voir à ce sujet les travaux de NISHIDA [1], [2], de NISHIDA et SMOLLER [1], et de CHORIN [1], [2].

x x
x

L'existence et l'unicité d'une solution faible vérifiant une certaine condition d'entropie ont été établies notamment par R. DIPERNA [1,2,4,5] dans le cadre d'un système vraiment non linéaire. La démarche employée est constructive, mais il ne s'agit pas d'un schéma numérique. On peut citer également les travaux de T.P. LIU ou de J.A. SMOLLER sur ce sujet. Tous ces travaux portent sur le système plus général

$$(40) \quad \begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial}{\partial x} (f(u, v)) \quad , \\ \frac{\partial v}{\partial t} &= \frac{\partial}{\partial x} (g(u, v)) \quad , \end{aligned}$$

pour lequel le schéma S_0 est inadapté dans la mesure où la solution en $x=0$ du problème de RIEMANN n'est pas en général obtenue aussi facilement qu'en (27), les invariants de RIEMANN étant connus souvent soit sous forme implicite, soit sous une forme paramétrée.

x x
x

Le schéma de LAX peut aussi être utilisé pour le système (4), (5). En généralisant de façon immédiate le cas scalaire, on obtient

$$(41) \quad u_i^{n+1} = \frac{1}{2} (u_{i+1}^n + u_{i-1}^n) + \frac{q}{2} [v_{i+1}^n - v_{i-1}^n] ,$$

$$(42) \quad v_i^{n+1} = \frac{1}{2} (v_{i+1}^n + v_{i-1}^n) + \frac{q}{2} [f(u_{i+1}^n) - f(u_{i-1}^n)] ,$$

avec pour condition de stabilité

$$q \sup_{u \in \mathbb{R}} \sqrt{f'(u)} \leq 1 .$$

On peut également construire ce schéma de façon analogue au schéma S_0 , en écrivant, chaque fois, que f est approchée par une fonction affine de pente $\frac{1}{2}$, de façon à vérifier la condition de stabilité. Ainsi, (27) doit être remplacé par

$$(43.i) \quad v_{i+1}^n = v_{i+1/2}^n - \frac{1}{q} (u_{i+1}^n - u_{i+1/2}^n)$$

$$(43.ii) \quad v_i^n = v_{i+1/2}^n + \frac{1}{q} (u_i^n - u_{i+1/2}^n)$$

d'où, en faisant la somme et la différence,

$$(44) \quad \begin{aligned} u_{i+1/2}^n &= \frac{1}{2} (u_{i+1}^n + u_i^n) + \frac{q}{2} (v_{i+1}^n - v_i^n) , \\ v_{i+1/2}^n &= \frac{1}{2} (v_{i+1}^n + v_i^n) + \frac{1}{2q} (u_{i+1}^n - u_i^n) . \end{aligned}$$

On introduit ensuite les $v_{i+1/2}^n$ dans (31), et il vient

donc

$$\begin{aligned} u_i^{n+1} &= u_i^n + q(v_{i+1/2}^n - v_{i-1/2}^n) \\ u_i^{n+1} &= u_i^n + \frac{q}{2} (v_{i+1}^n - v_{i-1}^n) + \frac{1}{2} (u_{i+1}^n - 2u_i^n + u_{i-1}^n) , \end{aligned}$$

et on retrouve exactement (41), en regroupant u_i^n et le terme de viscosité qui apparaît. On procède de même avec (32), en approchant $f(u_{i+1/2}^n)$ par

$$(45) \quad f(u_{i+1/2}^n) = \frac{1}{2} [f(u_{i+1}^n) + f(u_i^n)] + \frac{1}{2} \left[u_{i+1/2}^n - \frac{u_{i+1}^n + u_i^n}{2} \right] ,$$

et on obtient exactement (42).

x x
x

On peut également introduire des conditions aux limites adaptées à l'équation (1), sur un domaine borné. Par exemple sur $]0,1[\times]0,T[$, une condition du type DIRICHLET

$$w(0,t) = a(t) \quad ; \quad w(1,t) = b(t) \quad ; \quad t \in]0,T[$$

se traduit en

$$v(0,t) = a'(t) \quad ; \quad v(1,t) = b'(t) \quad ; \quad t \in]0,T[$$

sans condition sur u . Un problème de ce type est traité dans NISHIDA et SMOLLER [2]. On obtient des conditions sur u dans le cadre d'un problème de type NEUMANN. On peut également concevoir des conditions de type mixte, portant à la fois sur u et v . Compte-tenu du problème de RIEMANN, ces conditions, qui peuvent être écrites

$$A(u,v) = c^{te} \quad \text{en } x = 0 \quad , \quad B(u,v) = c^{te} \quad \text{en } x = 1 \quad ,$$

devront être telles que A exprime implicitement v comme fonction non décroissante de u et que B exprime v comme fonction non croissante de u, pour assurer l'existence d'une solution et conserver la stabilité dans K_0 . On observe des phénomènes de réflexion d'onde sur la frontière pour le problème de DIRICHLET (v donné) ou de NEUMANN (u donné). Ce n'est pas le cas en général pour le problème mixte, car en prenant par exemple pour l'équation linéaire, une condition portant sur $(v - ku)$ en $x = 0$ et sur $(v + ku)$ en $x = 1$, on retrouve, en découplant, le problème aux limites pour l'équation scalaire, pour lequel il n'y a pas réflexion des ondes.

x x
x

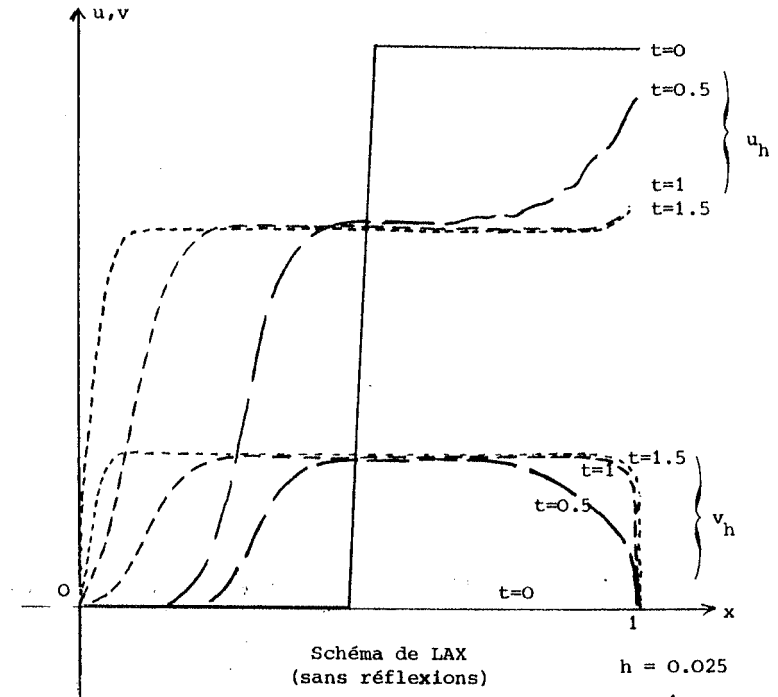
Quelques expériences numériques ont été menées sur le schéma de LAX et le schéma S_0 , avec, pour conditions aux limites $v = 0$ en $x = 0$ et $x = 1$, qui est de type DIRICHLET, pour ce dernier schéma. L'exemple représenté ici est le second décrit pages 213-214, avec la discontinuité initiale en $x = 0.5$. On obtient pour $t = 0,5$ la solution proposée comme exemple de solution du problème de RIEMANN, avec une légère réflexion en $x = 1$ due au fait que la vitesse de propagation de la détente est égale à $1/2$, avec un choc initial "situé" sur $\left] \frac{1}{2} - h, \frac{1}{2} + h \right]$. On peut comparer ce résultat avec la solution obtenue à partir du schéma de LAX dans les mêmes conditions (figure page 230). Les chocs sont plus raides, et les détenteurs plus régulières que pour le schéma de LAX.

Pour $t = 1$ et $t = 1.5$, les réflexions d'onde interviennent, pour le schéma S_0 ; le programme a en effet tenu compte de la condition $v = 0$, en calculant $v_{1/2}^n$ comme intersection des courbes

$$v = 0 \quad \text{et} \quad v_{1/2}^n = v - g(u_1^n) + g(u).$$

Pour $v_{I-1/2}^n$, à l'autre extrémité, on effectue un calcul analogue.

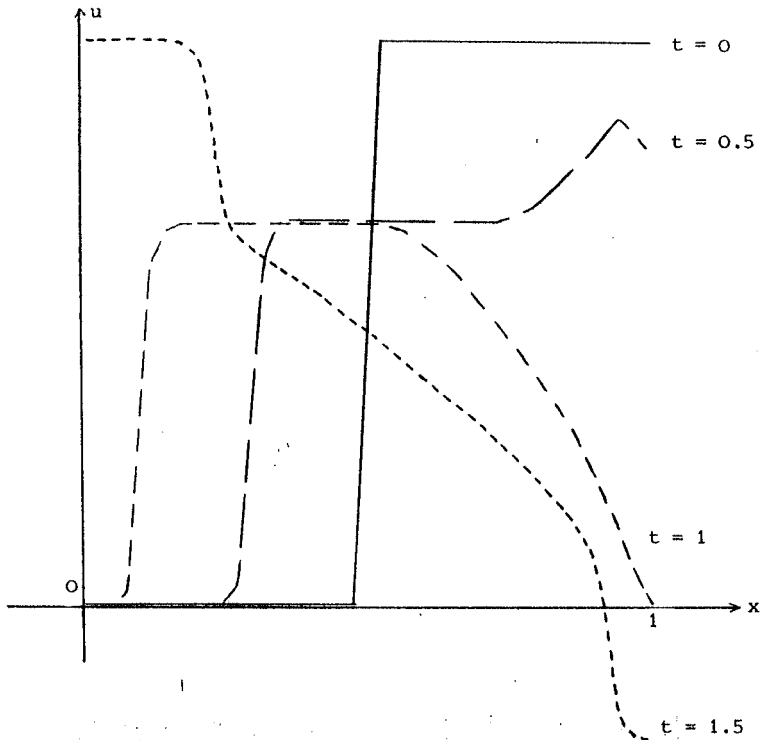
Le schéma de LAX est moins adapté à l'introduction des conditions aux limites, car $u_{i-1}^n, u_{i+1}^n, v_{i-1}^n, v_{i+1}^n$ apparaissent explicitement dans le calcul de u_i^{n+1} et v_i^{n+1} . On peut remarquer que les conditions aux limites effectivement utilisées sont $u = 1$ et $v = 0$ en $x = 1$, et $u = v = 0$ en $x = 0$, ce qui se traduit par des couches limites relativement étroites. On n'observe aucun phénomène de réflexion d'ondes.



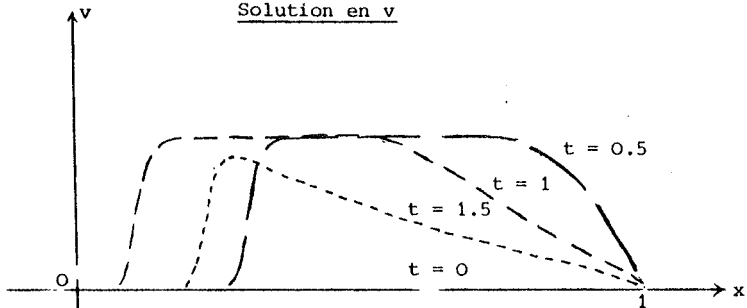
Pour introduire des conditions aux limites de type DIRICHLET ou de type NEUMANN, il faut changer de schéma au voisinage de la frontière (sur la première et sur la dernière maille), par exemple en reprenant l'interprétation faite précédemment. Cette technique est détaillée à la fin du chapitre VIII, pour le problème du câble extensible fixé à chaque extrémité, mais peut faire l'objet d'une généralisation.

x x
x

Solution en u



Solution en v



$$h = 0.025$$

$$q = 1$$

$$f(u) = \frac{u^3}{3}$$

CHAPITRE VIII

APPLICATION

A UN MODELE DE CABLE EXTENSIBLE

On montre dans un premier paragraphe, qu'un modèle, assez simplifié, en dynamique des câbles extensibles correspond à l'équation

$$(1) \quad \frac{\partial^2 w}{\partial t^2} = \frac{\partial}{\partial x} \left[f \left(\frac{\partial w}{\partial x} \right) \right]$$

où f est une fonction croissante, nulle à l'origine, convexe sur $]0, +\infty[$ et concave sur $] -\infty, 0[$. On étudie ensuite la stabilité des schémas introduits au chapitre VII pour la classe de fonctions f ainsi définie.

1. UN MODELE DE DYNAMIQUE DES CABLES EXTENSIBLES

On considère un câble, de longueur ℓ à l'état neutre ; on note A l'aire de la section du câble, et E son module d'élasticité. Ce câble est supposé fixé en $x = 0$, et on applique à l'autre extrémité une certaine tension initiale constante par une force d'intensité $N > 0$.

Le déplacement vertical est noté w , et le déplacement horizontal est noté z . On suppose que le mouvement reste plan, ce qui constitue une hypothèse assez restrictive, qui interdit les trop grandes déformations. Pour écrire les équations du mouvement, et obtenir une formule du type de (1), on utilise une technique analogue à celle de P. JOUVE dans le cadre des structures à barres, en statique.

L'énergie cinétique est

$$T = \frac{1}{2} \int_0^l (\dot{z}^2 + \dot{w}^2) \rho \, dx,$$

où ρ est la densité du câble.

On calcule également l'énergie potentielle du câble, somme des forces internes V_1 et des forces extérieures V_2 . On note ds l'élément de longueur dans l'état déformé, qui correspond à dx dans l'état non déformé et non tendu (état neutre). On a alors

$$(z' = \frac{\partial z}{\partial x}, \quad w' = \frac{\partial w}{\partial x})$$

$$V_1 = \int_0^l \frac{EA}{2} \left[\frac{ds}{dx} - 1 \right]^2 dx = \int_0^l \left\{ \frac{EA}{2} \sqrt{(1+z')^2 + w'^2} - 1 \right\}^2 dx,$$

$$V_2 = - \int_0^l w g \rho \, dx \quad (g = \text{gravité}).$$

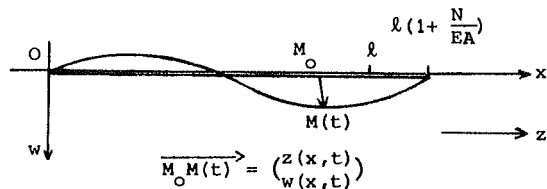
Les conditions aux limites sont les suivantes

$$z(0, t) = 0 \quad ; \quad z(l, t) = \frac{Nl}{EA} \quad (= \text{constante}) \quad ;$$

$$w(0, t) = 0 \quad ; \quad w(l, t) = 0 \quad ;$$

le câble étant supposé fixé à chaque extrémité.

Supposons un instant que pour $t_1 < t_2$, on connaisse les valeurs respectives de z et w à l'instant t_1 et à l'instant t_2 :



$$z(x, t_1) = z_1(x) \quad ; \quad z(x, t_2) = z_2(x) \quad ;$$

$$w(x, t_1) = w_1(x) \quad ; \quad w(x, t_2) = w_2(x) \quad .$$

On applique le principe d'HAMILTON : les fonctions z et w doivent rendre stationnaire l'expression

$$\int_{t_1}^{t_2} (T - V_1 - V_2) \, dt \quad .$$

Pour simplifier le problème, on va supposer que z_1 et z_2 sont des fonctions linéaires de x , et ceci correspond à écrire que w reste la seule inconnue du problème. On a en effet, pour vérifier les conditions aux limites

$$z_1(x) = z_2(x) = \frac{Nx}{EA} \quad ,$$

et on peut admettre, qu'il en est de même pendant la durée $t_2 - t_1$:

$$z(x, t) = \frac{Nx}{EA} \quad .$$

On est ainsi amené à chercher w , rendant stationnaire l'expression

$$\begin{aligned} \Pi(w) = \int_{t_1}^{t_2} & \left[\frac{1}{2} \int_0^l \dot{w}^2 \rho \, dx - \int_0^l \frac{EA}{2} \left[\sqrt{(1+z')^2 + w'^2} - 1 \right]^2 dx \right. \\ & \left. + \int_0^l w \rho g \, dx \right] dt \quad , \end{aligned}$$

où $z' = \frac{N}{EA}$ est une constante.

On introduit une variation Δw , telle que

$$\Delta w(0, t) = 0 \quad ; \quad \Delta w(l, t) = 0 \quad ;$$

$$\Delta w(x, t_1) = 0 \quad ; \quad w(x, t_2) = 0 \quad ;$$

puis on développe $\Pi(w + \Delta w) - \Pi(w)$, pour obtenir au premier ordre

$$\int_{t_1}^{t_2} \int_0^l \left\{ \dot{w} \Delta \dot{w} \rho - \frac{EA}{2} (F'(w') \Delta w') + \Delta w \rho g \right\} dx dt = 0 \quad ,$$

avec

$$F(u) = \left(\lambda \sqrt{1 + \left(\frac{u}{\lambda}\right)^2} - 1 \right)^2$$

et en posant $\lambda = 1 + z' = 1 + \frac{N}{EA} \geq 1$.

On intègre par parties, et il vient

$$\int_{t_1}^{t_2} \int_0^l \left\{ -\ddot{w} \rho + \frac{EA}{2} \frac{\partial}{\partial x} \left\{ F'(w') \right\} + \rho g \right\} \Delta w dx dt = 0 \quad ,$$

et l'accroissement Δw étant quelconque, on a immédiatement

$$(2.1) \quad \frac{\partial^2 w}{\partial t^2} = \frac{EA}{2\rho} \frac{\partial}{\partial x} \left\{ F' \left(\frac{\partial w}{\partial x} \right) \right\} + g \quad ,$$

dont on peut faire disparaître g en posant $w_1 = w - g t^2/2$.

On pose pour $\lambda \geq 1$ $(\lambda = 1 + \frac{N}{EA})$

$$(2.11) \quad f(u) = F'(u) = 2u \left\{ 1 - \frac{1}{\lambda \sqrt{1 + \left(\frac{u}{\lambda}\right)^2}} \right\}$$

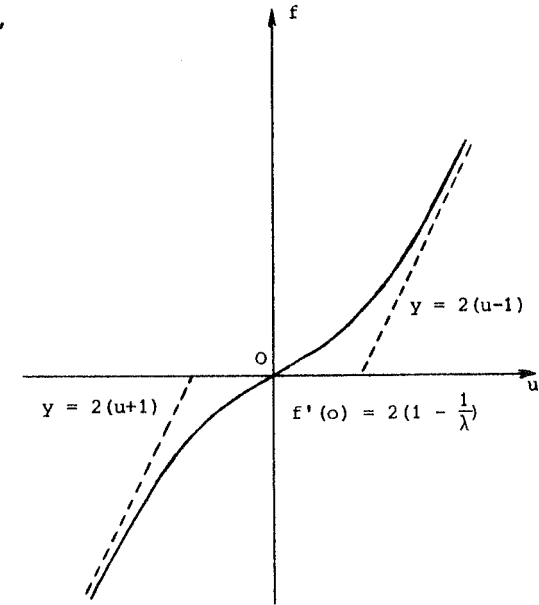
La fonction f ainsi définie est strictement croissante et impaire. Sa dérivée est donnée par

$$f'(u) = 2 \left[1 - \frac{1}{\lambda \left(1 + \left(\frac{u}{\lambda}\right)^2 \right)^{3/2}} \right],$$

qui est croissante, pour $u \geq 0$ et donc f est convexe sur $]0, +\infty[$.

Dans l'hypothèse de déplacements assez petits, on peut développer $f(u)$ au voisinage de l'origine, et il vient

$$f(u) \sim 2 \left(1 - \frac{1}{\lambda} \right) u + \left(\frac{u}{\lambda} \right)^3$$



On retrouve l'équation linéaire en négligeant le terme en u^3 , et à condition que λ soit différent de 1, c'est-à-dire lorsque la tension initiale du câble est non nulle. Pour une tension initiale nulle, on trouve $f(u) = u^3$, ce qui correspond aux exemples donnés précédemment. Un développement à l'ordre supérieur donne un équivalent de f qui n'est pas monotone ; il apparaît un terme de la forme $-\frac{3}{4} \left(\frac{u}{\lambda}\right)^5$.

Le calcul de

$$g(u) = \int_0^u \sqrt{f'(y)} dy$$

n'est pas immédiat, et on est amené à utiliser une technique d'intégration numérique, pour résoudre VII-27 par une méthode de NEWTON. Lorsque f est de la forme

$$f(u) = \frac{u^3}{3} + k^2 u \quad ,$$

on a

$$g(u) = \frac{1}{2} \left[u \sqrt{u^2 + k^2} + k^2 \operatorname{Arg sh} \left(\frac{u}{k} \right) \right].$$

Les conditions aux limites sont nulles en w , ce qui donne

pour $v = \frac{\partial w}{\partial t}$:

$$v(0,t) = v(0,l) = 0$$

et on doit ajouter une condition de compatibilité sur $u = \frac{\partial w}{\partial x}$, qui est donnée par

$$\int_0^l u(x,0) dx = \int_0^l \frac{\partial w}{\partial x}(x,0) dx = 0.$$

Cette condition est conservée pour tout t ; en effet

$$\begin{aligned} \frac{d}{dt} \int_0^l u(x,t) dx &= \int_0^l \frac{\partial u}{\partial t}(x,t) dx \\ &= \int_0^l \frac{\partial v}{\partial x}(x,t) dx \\ &= 0. \end{aligned}$$

Du point de vue numérique, le schéma devra également vérifier cette condition, de façon à assurer que u et v sont les composantes du gradient d'une fonction nulle aux limites.

Les équations générales pour le calcul des câbles sont données dans CRISTESCU, qui considère le problème à trois dimensions. Des schémas numériques sont proposés dans P. ROUSSEL, également pour le problème à trois dimensions. L'introduction de nouvelles forces extérieures se fait en ajoutant un second membre dépendant de (u,v,x,t) à (2.i).

2. STABILITE POUR LE PROBLEME DE RIEMANN

On se propose, dans le cadre de l'équation (1), avec une fonction $f \in C^1(\mathbb{R})$, croissante, concave sur $]-\infty, 0[$, et convexe sur $]0, +\infty[$ de

montrer en détails l'existence d'une solution au problème de RIEMANN, et que cette solution est à valeurs dans K_0 , défini par

$$(3) \quad K_0 = \left\{ (u,v) ; |v| + |g(u)| \leq M_0 \right\}$$

avec $M_0 = \operatorname{Max} \left\{ |v_-| + |g(u_-)|, |v_+| + |g(u_+)| \right\},$

et

$$(4) \quad g(u) = \int_0^u \sqrt{f'(y)} dy.$$

Notons que les hypothèses faites sur f assurent que K_0 est convexe.

Le problème de RIEMANN considéré est le suivant,

$$(5) \quad \frac{\partial v}{\partial t} = \frac{\partial}{\partial x} (f(u)) ; \quad \frac{\partial u}{\partial t} = \frac{\partial v}{\partial x} ; \quad \text{avec } v = \frac{\partial w}{\partial t}, \quad u = \frac{\partial w}{\partial x},$$

$$u(x,0) = \begin{cases} u_+ & x > 0 ; \\ u_- & x < 0 ; \end{cases}$$

(6)

$$v(x,0) = \begin{cases} v_+ & x > 0 ; \\ v_- & x < 0 ; \end{cases}$$

et était déjà introduit au chapitre VII, paragraphe 2. On a le résultat suivant, analogue aux résultats de T.P. LIU, J. SMOLLER ou B.L. KEYFITZ et H.C. KRANZER.

Théorème 19 : Le problème de RIEMANN (5), (6) admet une solution faible u , à valeurs dans K_0 . Cette solution satisfait à la condition d'entropie de T.P. LIU (VII.22.23) qui assure l'unicité.

Démonstration

On distingue quatre cas différents, suivant les positions respectives dans le plan de phase, de $A_- = (u_-, v_-)$ et $A_+ = (u_+, v_+)$.

L'hypothèse de convexité sur $]0, +\infty[$ faite sur f , assure que le système (5) est strictement hyperbolique au sens de (VII.21), et on est assuré de l'existence du point $A_0 = (u_0, v_0)$ tel que (VII.19) et (VII.20), c'est-à-dire

$$(7) \quad v_+ = v_0 - g(u_+; u_0)$$

$$(8) \quad v_- = v_0 + g(u_-; u_0)$$

soient vérifiées, avec $g(\cdot; \cdot)$ définie en (VII.13), (VII.14). On veut prouver ici que les arcs de courbe

$$v = v_0 - g(u; u_0) \quad v \in I(v_+, v_-), \quad u \in I(u_+, u_0)$$

et

$$v = v_0 + g(u; u_0) \quad v \in I(v_-, v_+), \quad u \in I(u_-, u_0)$$

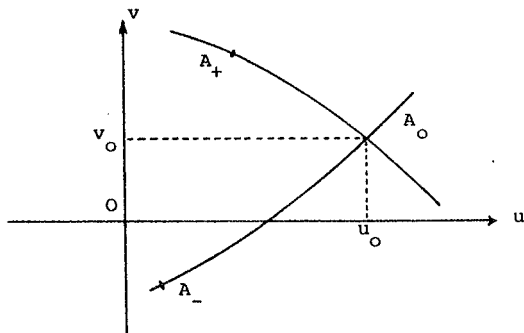
sont situés entièrement dans K_0 , et décrivent effectivement les valeurs de la solution.

Cas N° 1 : $v_0 \in I(v_+, v_-)$, $u \geq \text{Max}(u_+, u_-)$.

On a nécessairement

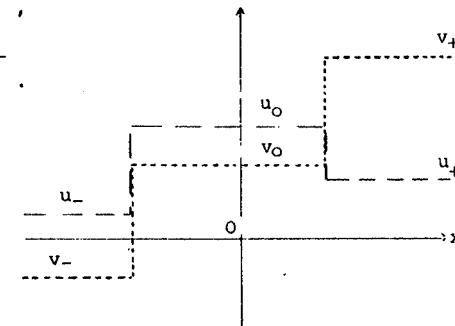
$$v_- \leq v_0 \leq v_+$$

et par construction de $g(u; u_0)$, à partir de l'enveloppe concave de f , on obtient en supposant $\text{Min}(u_+, u_-) \geq 0$, dans un premier temps,



$$(9) \quad \begin{aligned} g(u_-, u_0) &= -\sqrt{(f(u_-) - f(u_0))(u_- - u_0)}, \\ g(u_+, u_0) &= -\sqrt{(f(u_+) - f(u_0))(u_+ - u_0)}. \end{aligned}$$

La solution est constituée de deux chocs, et prend les seules valeurs possibles (u_-, v_-) , (u_0, v_0) et (u_+, v_+) . On peut même préciser explicitement ses valeurs



$$(u(x,t), v(x,t)) = \begin{cases} u_-, v_- & \text{si } x < -\alpha t, \\ u_0, v_0 & \text{si } x \in]-\alpha t, -\beta t[, \\ u_+, v_+ & \text{si } x > -\beta t, \end{cases}$$

$$\text{avec } \alpha = \frac{v_- - v_0}{u_- - u_0}, \quad \beta = \frac{v_+ - v_0}{u_+ - u_0}.$$

On a $(u_0, v_0) \in K_0$. En effet, en appliquant l'inégalité de CAUCHY-SCHWARTZ, on a

$$(10) \quad \begin{aligned} g(u_0) - g(u_+) &= \int_{u_+}^{u_0} \sqrt{f'(y)} dy \\ &\leq \left[\int_{u_+}^{u_0} 1 \cdot dy \right]^{1/2} \left[\int_{u_+}^{u_0} f'(y) dy \right]^{1/2} \\ &\leq \sqrt{u_0 - u_+} \sqrt{f(u_0) - f(u_+)} = -g(u_+; u_0). \end{aligned}$$

Donc, d'après (7)

$$v_+ \geq v_0 + g(u_0) - g(u_+),$$

c'est-à-dire

$$v_0 + g(u_0) \leq v_+ + g(u_+) \leq M_0.$$

D'autre part, g est croissante, donc

$$g(u_+) \leq g(u_0) \quad \text{et} \quad g(u_-) \leq g(u_0) \quad +$$

On a également $v_+ \geq v_0 \geq v_-$, donc

$$v_0 + g(u_0) \geq v_- + g(u_-) \geq -M_0,$$

et

$$v_0 - g(u_0) \leq v_+ - g(u_+) \leq M_0.$$

En reprenant l'inégalité de CAUCHY SCHWARTZ, comme en (8), on montre également

$$g(u_0) - g(u_-) \leq -g(u_-; u_0),$$

d'où en utilisant (9),

$$v_0 - g(u_0) \geq v_- - g(u_-) \geq -M_0.$$

On obtient donc

$$|v_0| + |g(u_0)| \leq M_0$$

c'est-à-dire que $(u_0, v_0) \in K_0$.

Il reste à montrer que la condition d'entropie est satisfaite sur chacun des chocs. Sur le choc de vitesse négative, la condition de T.P. LIU s'écrit

$$(11) \quad \max_{k \in I(u_0, u_-)} \frac{f(k) - f(u_-)}{k - u_-} = \frac{f(u_0) - f(u_-)}{u_0 - u_-},$$

ce qui est vérifié, car f est convexe et $u_0 > u_-$. Pour le choc de vitesse positive, cette condition devient

$$(12) \quad \min_{k \in I(u_0, u_+)} \frac{f(k) - f(u_0)}{k - u_0} = \frac{f(u_+) - f(u_0)}{u_+ - u_0},$$

ce qui est vrai, f étant convexe et $u_0 > u_+$.

Considérons maintenant le cas général où u_- et u_+ sont de signe non nécessairement positif. Il existe alors $a \in [u_+, u_0]$ tel que l'enveloppe concave de f coïncide avec f sur $[u_+, a]$ et avec une droite de pente

$$\frac{f(u_0) - f(a)}{u_0 - a}$$

sur $[a, u_0]$. Lorsque u_+ est non négatif, on a $u_+ = a$, et lorsque u_0 est négatif, on a $u_0 = a$.

On a alors

$$\begin{aligned} g(u_+; u_0) &= - \int_{u_+}^a \sqrt{f'(\xi)} d\xi + \int_{u_0}^a \sqrt{\frac{f(u_0) - f(a)}{u_0 - a}} d\xi \\ &= g(u_+) - g(a) + g(a; u_0). \end{aligned}$$

Comme précédemment, d'après l'inégalité de CAUCHY SCHWARTZ

$$g(a; u_0) \leq g(a) - g(u_0),$$

donc

$$g(u_+; u_0) \leq g(u_+) - g(u_0).$$

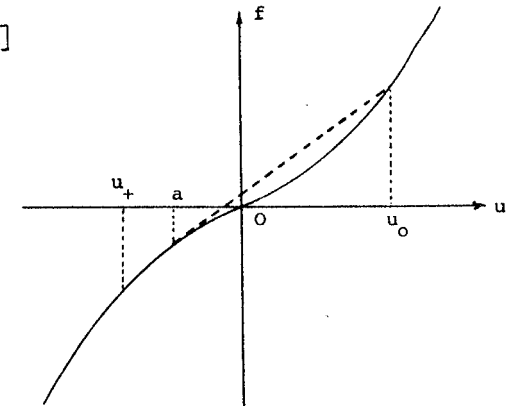
On prouve de la même façon qu'on a encore, en introduisant $b \in [u_-, u_0]$, de façon analogue à a précédemment,

$$g(u_-; u_0) = g(u_-) - g(b) + g(b; u_0),$$

et donc

$$g(u_-; u_0) \leq g(u_-) - g(u_0),$$

et on obtient la stabilité : $(u_0, v_0) \in K_0$, puisque g est toujours croissante,



et $v_0 \in [v_-, v_+]$. On peut remarquer que nécessairement, a et b sont négatifs. Sur $[a, u_0]$ et $[b, u_0]$, on a encore (11) et (12) respectivement bien que f ne soit pas convexe sur ces intervalles ; le graphe de f ne traverse cependant pas la corde. Sur $[u_+, a]$ et $[u_-, b]$, la solution se comporte comme une détente ; u coïncide alors avec la solution du problème de RIEMANN scalaire suivant,

$$\frac{\partial u}{\partial t} + \sqrt{F'(u)} \frac{\partial u}{\partial x} \left(= \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} (g(u)) \right) = 0 ,$$

(13)

$$u(x,0) = \begin{cases} u_+ & \text{si } x > 0 , \\ a & \text{si } x \leq 0 , \end{cases}$$

sur $\left\{ (x,t) ; T > t, x > -\frac{C - v_0}{a - u_0} t \right\}$, où

(14) $C = v_0 - g(a ; u_0) \quad , \quad (C > v_0)$,

qui est la valeur de v lorsque u est égal à a . Le problème (13) admet une solution régulière sur $\mathbb{R} \times]0, T[$; en effet g , de dérivée décroissante $\sqrt{F'}$, est concave et $0 \geq a \geq u_+$. On procède de même sur $[u_-, b]$, pour aboutir à une détente vérifiant l'équation

$$\frac{\partial u}{\partial t} - \sqrt{F'(u)} \frac{\partial u}{\partial x} = 0 .$$

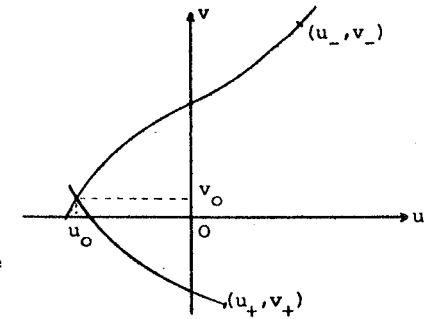
La valeur de v est chaque fois obtenue par (14) (où l'équivalent pour (u_-, u_0) , au niveau du choc, et en écrivant au niveau de la détente l'équation de l'invariant de RIEMANN classique décroissant qui passe par (u_+, v_+) , dans le plan de phase (ou par (u_-, v_-) avec un invariant de RIEMANN croissant).

Cas N° 2 : $v_0 \in I(v_+, v_-)$, $u \leq \text{Min}(u_+, u_-)$.

On a alors nécessairement

$$v_+ \leq v_0 \leq v_- .$$

Chaque fois, $g(u; u_0)$ est obtenu en prenant l'enveloppe convexe de f sur $[u_0, u_+]$ et sur $[u_0, u_-]$ respectivement. Lorsque u_+ et u_- sont tous les deux négatifs, on voit apparaître deux chocs, l'un de vitesse positive, l'autre de vitesse négative.



Dans le cas général, il existe $a \in [u_0, u_+]$ et $b \in [u_0, u_-]$ tels que la solution coïncide avec un choc sur $[u_0, a]$ et sur $[u_0, b]$, et avec une détente sur $[a, u_+]$ et sur $[b, u_-]$. On obtient encore en utilisant l'inégalité de CAUCHY SCHWARTZ, la croissance de g et le fait que v_0 appartienne à $[v_+, v_-]$, les inégalités,

$$v_+ + g(u_+) \leq v_0 + g(u_0) \leq v_- + g(u_-) ,$$

$$v_+ - g(u_+) \leq v_0 - g(u_0) \leq v_- - g(u_-) .$$

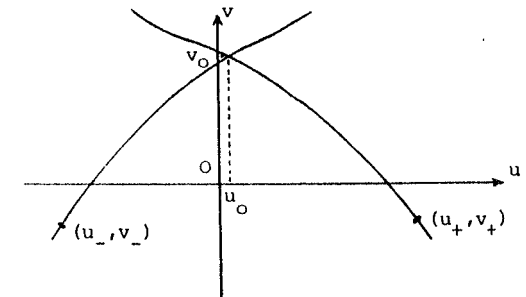
On en déduit immédiatement la stabilité dans K_0 . Le choc, au niveau de $[u_0, a]$, admet une vitesse positive, et la condition d'entropie de T.P. LIU revient à écrire (12), ce qui est vérifié, car f est concave sur $[u_0, a]$ (en effet, si $a \neq u_0$, nécessairement, u_0 et a sont négatifs), et $u_0 < a$. On vérifie également que (11) est vérifié au niveau de $[u_0, b]$.

Cas N° 3 : $u_0 \in I(u_-, u_+)$ et $v_0 \geq \text{Max}(v_-, v_+)$.

On a nécessairement

$$u_- \leq u_0 \leq u_+ .$$

Si u_- est positif, on voit apparaître un choc de vitesse négative, et une détente de vitesse positive. Il se produit l'inverse si u_+ est



négatif. Dans chacun des cas,

on a soit $u_0 \geq 0$ et $v_0 + g(u_0) = v_+ + g(u_+)$, car il y a détente, soit $u_0 < 0$ et $v_0 - g(u_0) = v_- - g(u_-)$, car il y a également détente. Dans le premier cas, si v_0 est non négatif, l'appartenance à K_0 est immédiate, car $v_0 + g(u_0) = |v_0| + |g(u_0)|$. Si v_0 est négatif, on a, en minorant v_0 par v_+ et $-g(u_0)$ par $-g(u_+)$, l'inégalité

$$0 \geq v_0 - g(u_0) \quad (= -|v_0| - |g(u_0)|) \geq v_+ - g(u_+).$$

Dans l'autre cas, u_0 est négatif ou nul, et si v_0 est positif l'appartenance à K_0 est immédiate. Si v_0 est négatif, on a

$$0 \leq -v_0 - g(u_0) \leq -v_- - g(u_-).$$

Dans le cas général, on a l'alternative suivante.

- ou bien u_0 est positif, et il existe $b \in [u_-, u_0]$ tel que

$$g(u_-; u_0) = g(u_-) - g(b) + g(b; u_0).$$

Il y a alors détente au niveau de (u_0, u_+) , et on procède comme dans le premier cas précédent.

- ou bien u_0 est négatif, et il existe $a \in [u_0, u_+]$ tel que

$$g(u_+; u_0) = g(u_+) - g(a) + g(a; u_0).$$

Il y a alors détente au niveau de (u_-, u_0) , et on procède comme dans le second cas précédent.

On obtient ainsi la stabilité dans K_0 dans chaque cas.

La condition d'entropie sur les chocs se vérifie comme dans les cas 1 et 2.

Cas N° 4 : $u_0 \in I(u_-, u_+)$, $v_0 \leq \text{Min}(v_-, v_+)$.

On a nécessairement $u_+ \leq u_0 \leq u_-$, et ce cas se traite exactement comme le précédent, en inversant les rôles de u_- et de u_+ .

Il resterait à considérer le cas où u_0 et v_0 appartiennent respectivement à l'intérieur des intervalles $I(u_-, u_+)$, $I(v_-, v_+)$, mais le fait que g soit croissante exclut cette possibilité, et le théorème 19 est vérifié.

Comme prévu au chapitre précédent, ce théorème 19 peut être généralisé au cas où f est monotone et le problème strictement hyperbolique. La démonstration devient cependant plus technique car il faut tenir compte de chaque changement de convexité de f , et écrire qu'au niveau des détente la solution est décrite par un invariant de RIEMANN classique, donc parallèle à la frontière, et qu'au niveau des chocs, cette solution est décrite par une courbe de choc et donc pénètre davantage dans K_0 , dont la convexité n'est pas utilisée.

On peut en effet remarquer qu'on ne s'est pas servi de la convexité de K_0 , au cours de cette démonstration. Cette propriété intervient au niveau de la projection de type L^2 effectuée par les schémas.

3. STABILITE POUR LES SCHEMAS DE LAX, S_0 et S_1 .

On considère le problème général constitué du système (5) et de la condition initiale sur \mathbb{R} , supposée uniformément bornée et mesurable

$$(15) \quad u(x, 0) = u_0(x) \quad ; \quad v(x, 0) = v_0(x) .$$

On prend désormais pour ensemble K_0 l'ensemble suivant, dans le plan de phase

$$K_0 = \left\{ (u, v) , \quad |v| + |g(u)| \leq |v_0|_{L^\infty(\mathbb{R})} + |g(u_0)|_{L^\infty(\mathbb{R})} = M_0 \right\} .$$

Cet ensemble est borné, et l'hypothèse sur la convexité ou la concavité de f implique qu'il est convexe. On a le résultat suivant, (en notant S_1 le schéma (VII.26, 31, 32) et S_0 le schéma (VII.27, 31, 32)).

Théorème 20 : Pour tout $h > 0$, les solutions approchées construites par le schéma de LAX, ou par l'un des deux schémas S_0 et S_1 , sont stables dans K_0 , lorsque la condition de stabilité

$$(16) \quad q \sup_{|g(u)| \leq M_0} \sqrt{f'(u)} \leq \alpha$$

est satisfaite, avec $\alpha = \frac{1}{2}$ pour les schémas S_0 et S_1 , et $\alpha = 1$ pour le schéma de LAX.

x x
x

On peut déduire de ce théorème la convergence dans $\{L^\infty(\mathbb{R} \times]0, T[) \}^2$ faible * d'une suite (u_n, v_n) vers $(u, v) \in \{L^\infty(\mathbb{R} \times]0, T[) \}^2$, qui satisfait au sens des distributions à $\frac{1}{m} \frac{1}{h}$ la première équation, c'est-à-dire $\partial u / \partial t = \partial v / \partial x$, et est à valeurs dans K_0 . Ce théorème peut également s'adapter pour le système (5) avec un second membre dépendant de (u, v, x, t) . On reprend les techniques du théorème 2, ce second membre intervenant dans (VII.32) ; le compact K_0 est remplacé par un compact $K(t)$, de diamètre $2M_0 e^{ct}$.

x x
x

Démonstration

A. Le schéma S_1

Si la condition de stabilité (16) est satisfaite, tous les u_i^n étant connus pour $n < N$ fixé, on est assuré que le problème de type

RIEMANN

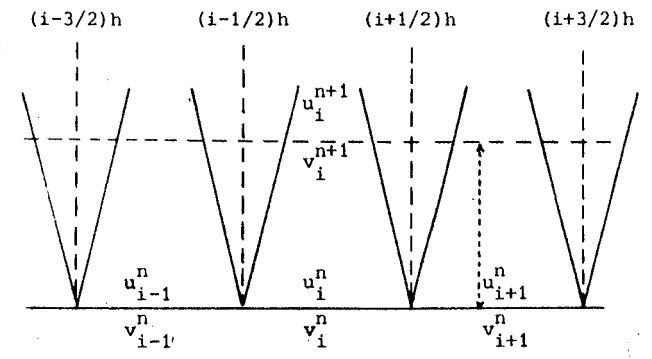
$$\frac{\partial u}{\partial t} = \frac{\partial v}{\partial x} ; \quad \frac{\partial v}{\partial t} = \frac{\partial}{\partial x} [f(u)] ;$$

$$(17) \quad \begin{aligned} u(x,0) &= \begin{cases} u_{i+1}^n & \text{si } x > (i+1/2)h \\ u_i^n & \text{si } x < (i+1/2)h \end{cases} \\ v(x,0) &= \begin{cases} v_{i+1}^n & \text{si } x > (i+1/2)h \\ v_i^n & \text{si } x < (i+1/2)h \end{cases} \end{aligned}$$

admet une solution unique, donnée par le théorème 19, sur le triangle

$$(18) \quad \left\{ (x,t) ; |x - (i+1/2)h| + \frac{t}{q} \leq h \right\}.$$

En effet, la condition de stabilité adapte le maillage à la vitesse maximale de propagation des données, et les valeurs données initialement pour $|x - (i+1/2)h| > h$ ne peuvent pas intervenir pour modifier la solution du problème (17).



En reprenant la construction du schéma S_1 et le théorème 19, on a nécessairement l'appartenance à K_0 des valeurs prises par les solutions de problèmes élémentaires de type (17), et ceci est vrai même lorsque le second membre de (16) vaut 1.

On a une précision supplémentaire lorsque ce second membre est égal à 1/2 ; la solution de (17) est alors étendue au pavé $[ih, (i+1)h] \times [0, qh]$, et est toujours à valeurs dans K_0 . Il en est de même sur le pavé $[(i-1)h, ih] \times [0, qh]$.

D'après (VII.31) et (VII.32), u_i^{n+1} et v_i^{n+1} sont obtenus en effectuant des moyennes de fonctions à valeurs dans K_0 , qui est convexe. D'après l'inégalité de JENSEN :

(19) Si ϕ est une fonction convexe définie sur \mathbb{R} , et $a < b$,

$$\phi \left(\frac{1}{b-a} \int_a^b u(x) dx \right) \leq \frac{1}{b-a} \int_a^b \phi(u(x)) dx ;$$

en prenant pour fonction ϕ successivement la frontière inférieure ($v + |g(u)| = M_0$ avec $v \leq 0$) puis la frontière supérieure ($v \geq 0$) de K_0 , on vérifie immédiatement que

(20) $\forall i \in \mathbb{Z} \quad (u_i^{n+1}, v_i^{n+1}) \in K_0 .$

B. Le schéma de LAX

Supposons par exemple i pair ; le calcul des u_i^{n+1} pour i pair ne fait intervenir que les u_j^n pour j impair, dans le schéma de LAX. On peut donc procéder comme si chaque u_j^n correspondait à la moyenne sur un intervalle de longueur $2h$, en évitant les interférences entre deux problèmes de RIEMANN élémentaires, lorsque la condition de stabilité (16) est vérifiée, avec $\alpha = 1$. On peut faire un raisonnement analogue lorsque i est impair.

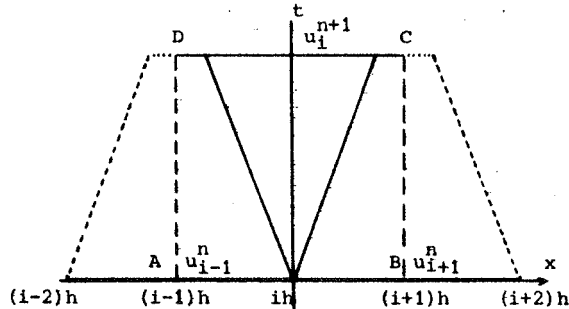
Le problème de RIEMANN

$$\frac{\partial u}{\partial t} = \frac{\partial v}{\partial x} ,$$

$$\frac{\partial v}{\partial t} = \frac{\partial}{\partial x} [f(u)] ,$$

(21)

$$(u(x,0), v(x,0)) = \begin{cases} (u_{i-1}^n, v_{i-1}^n) & \text{si } x < ih, \\ (u_{i+1}^n, v_{i+1}^n) & \text{si } x > ih, \end{cases}$$



admet une solution faible unique dont la restriction au pavé défini par $](i-1)h, (i+1)h[\times]0, qh[$ reste inchangée si la condition initiale change sur $\{x \mid |x-ih| \geq 2h\}$.

En intégrant par une formule de GREEN le système (21) sur le rectangle ABCD de la figure ci-dessus, qui correspond au pavé $](i-1)h, (i+1)h[\times]0, qh[$, et sachant que sur AD et BC la solution est constante, on obtient exactement le schéma de LAX lorsque les moyennes en u et v de la solution du problème de RIEMANN (21) sur $](i-1)h, (i+1)h[$ au temps $t = qh$ sont notées u_i^{n+1} et v_i^{n+1} respectivement. En appliquant l'inégalité de JENSEN (19), on établit que le couple (u_i^{n+1}, v_i^{n+1}) appartient à K_0 , puisque (u_{i-1}^n, v_{i-1}^n) et (u_{i+1}^n, v_{i+1}^n) y sont.

C. Le schéma S_0

Ce schéma ne correspond pas en général à la résolution de problèmes de RIEMANN particuliers. On est ainsi amené à prouver directement la stabilité. Pour limiter le nombre de cas à envisager, on met (31), (32) sous la forme

(22)
$$u_{i,1}^{n+1} = u_i^n + 2q(v_{i+1/2}^n - v_i^n) ,$$

$$v_{i,1}^{n+1} = v_i^n + 2q(f(u_{i+1/2}^n) - f(u_i^n)) ,$$

(23)
$$u_{i,2}^{n+1} = u_i^n + 2q(v_i^n - v_{i-1/2}^n) ,$$

$$v_{i,2}^{n+1} = v_i^n + 2q(f(u_i^n) - f(u_{i-1/2}^n)) ,$$

puis

$$u_i^{n+1} = \frac{1}{2} (u_{i,1}^{n+1} + u_{i,2}^{n+1})$$

$$v_i^{n+1} = \frac{1}{2} (v_{i,1}^{n+1} + v_{i,2}^{n+1}) .$$

K_0 étant convexe, il est immédiat que le couple (u_i^{n+1}, v_i^{n+1}) appartient à K_0 dès que $(u_{i,1}^{n+1}, v_{i,1}^{n+1})$ et $(u_{i,2}^{n+1}, v_{i,2}^{n+1})$ y sont.

Dans chacune des expressions (22), (23), le rapport q se trouve remplacé par $2q$, ce qui explique que la condition de stabilité soit plus restrictive pour le schéma S_3 .

On suppose que (u_{i-1}^n, v_{i-1}^n) , (u_i^n, v_i^n) et (u_{i+1}^n, v_{i+1}^n) appartiennent à K_0 . Il est alors immédiat, d'après (VII.27), c'est-à-dire

$$(24) \quad \begin{aligned} v_{i+1/2}^n + g(u_{i+1/2}^n) &= v_{i+1}^n + g(u_{i+1}^n) \\ v_{i+1/2}^n - g(u_{i+1/2}^n) &= v_i^n - g(u_i^n) \end{aligned}$$

que $(u_{i+1/2}^n, v_{i+1/2}^n)$ et de la même façon $(u_{i-1/2}^n, v_{i-1/2}^n)$ appartiennent également à K_0 .

On considère maintenant (22), cas par cas en ce qui concerne les positions respectives de u_i^n , $u_{i+1/2}^n$ et 0 , le changement de convexité de f .

1°/ Si $0 \leq u_i^n \leq u_{i+1/2}^n$, on a, d'après (24)

$$v_{i+1/2}^n - v_i^n = g(u_{i+1/2}^n) - g(u_i^n),$$

donc

$$u_{i,1}^{n+1} = u_i^n + 2q(g(u_{i+1/2}^n) - g(u_i^n))$$

et par la condition de stabilité, il vient

$$(25) \quad \begin{aligned} u_{i,1}^{n+1} &= u_i^n (1 - 2q \sqrt{f'(\eta_{i+1/4}^n)}) + u_{i+1/2}^n (2q \sqrt{f'(\eta_{i+1/4}^n)}) \\ \eta_{i+1/4}^n &\in I(u_i^n, u_{i+1/2}^n) = [u_i^n, u_{i+1/2}^n], \end{aligned}$$

$$\text{et donc, } u_{i,1}^{n+1} \in [u_i^n, u_{i+1/2}^n].$$

On a également, d'après la formule de la moyenne

$$\begin{aligned} f(u_{i+1/2}^n) - f(u_i^n) &= \int_{u_i^n}^{u_{i+1/2}^n} f'(\xi) d\xi \\ &= \sqrt{f'(\xi_{i+1/4}^n)} \int_{u_i^n}^{u_{i+1/2}^n} \sqrt{f'(\xi)} d\xi \\ &= \sqrt{f'(\xi_{i+1/4}^n)} (g(u_{i+1/2}^n) - g(u_i^n)) \\ &= \sqrt{f'(\xi_{i+1/4}^n)} (v_{i+1/2}^n - v_i^n), \end{aligned}$$

où $\xi_{i+1/4}^n \in I(u_i^n, u_{i+1/2}^n)$.

On a donc

$$v_{i,1}^{n+1} = v_i^n + 2q \sqrt{f'(\xi_{i+1/4}^n)} (v_{i+1/2}^n - v_i^n),$$

dont on déduit

$$v_{i,1}^{n+1} \in I(v_i^n, v_{i+1/2}^n)$$

avec nécessairement $v_i^n \leq v_{i+1/2}^n$ car la fonction g est croissante et $u_i^n \leq u_{i+1/2}^n$.

Il vient immédiatement, en tenant compte de la croissance de g

$$\begin{aligned} v_{i,1}^{n+1} + g(u_{i,1}^{n+1}) &\leq v_{i+1/2}^n + g(u_{i+1/2}^n), \\ v_{i,1}^{n+1} + g(u_{i,1}^{n+1}) &\geq v_i^n + g(u_i^n). \end{aligned}$$

On a également, $u_{i,1}^{n+1}$ étant positif

$$\begin{aligned} v_{i,1}^{n+1} - g(u_{i,1}^{n+1}) &\leq v_{i,1}^{n+1} + g(u_{i,1}^{n+1}) \\ &\leq v_{i+1/2}^n + g(u_{i+1/2}^n). \end{aligned}$$

Il reste à minorer $v_{i,1}^{n+1} - g(u_{i,1}^{n+1})$. On a

$$g(u_{i,1}^{n+1}) = g(u_i^n) + g'(\xi_i^n) (u_{i,1}^{n+1} - u_i^n) ,$$

avec $\xi_i^n \in I(u_i^n, u_{i,1}^{n+1})$, donc

$$(26) \quad v_{i,1}^{n+1} - g(u_{i,1}^{n+1}) = v_i^n - g(u_i^n) + 2q (\sqrt{f(\xi_{i+1/4}^n)} - g'(\xi_i^n)) (v_{i+1/2}^n - v_i^n) .$$

Or
$$\sqrt{f'(\xi_{i+1/4}^n)} = \frac{f(u_{i+1/2}^n) - f(u_i^n)}{g(u_{i+1/2}^n) - g(u_i^n)} ,$$

et d'après l'inégalité de CAUCHY SCHWARTZ :

$$g(u_{i+1/2}^n) - g(u_i^n) = \int_{u_i^n}^{u_{i+1/2}^n} \sqrt{f'(y)} dy$$

$$\leq \sqrt{u_{i+1/2}^n - u_i^n} \sqrt{f(u_{i+1/2}^n) - f(u_i^n)}$$

On a donc, en élevant au carré

$$\frac{f(u_{i+1/2}^n) - f(u_i^n)}{g(u_{i+1/2}^n) - g(u_i^n)} \geq \frac{g(u_{i+1/2}^n) - g(u_i^n)}{u_{i+1/2}^n - u_i^n}$$

De plus, g est convexe, et $u_i^n < u_{i+1/2}^n$, donc pour tout $u \in [u_i^n, u_{i+1/2}^n]$

$$\frac{g(u) - g(u_i^n)}{u - u_i^n} \leq \frac{g(u_{i+1/2}^n) - g(u_i^n)}{u_{i+1/2}^n - u_i^n}$$

On en déduit, en prenant $u = u_{i,1}^{n+1}$, que

$$\sqrt{f'(\xi_{i+1/4}^n)} > g'(\xi_i^n) ,$$

et comme $v_{i+1/2}^n - v_i^n$ est positif, (26) donne

$$(27) \quad v_i^{n+1} - g(u_{i,1}^{n+1}) \geq v_i^n - g(u_i^n) .$$

2°/ Si $u_i^n \leq u_{i+1/2}^n \leq 0$

La fonction f est concave, sur l'intervalle considéré, et donc

$$g(u_i^n ; u_{i+1/2}^n) = g(u_i^n) - g(u_{i+1/2}^n) .$$

En prenant par exemple $u_{i+1}^n = u_{i+1/2}^n$, on est dans le cadre du schéma S_1 , avec $u_{i-1}^n = u_i^n$ et q remplacé par $2q$. On peut appliquer le résultat obtenu pour le schéma S_1 , la valeur u_i^n étant prise sur un intervalle de longueur $2h$. On a donc

$$(u_{i,1}^{n+1}, v_{i,1}^{n+1}) \in K_0 ,$$

et en particulier, en appliquant l'inégalité de JENSEN

$$(28) \quad v_{i,1}^{n+1} - g(u_{i,1}^{n+1}) \leq v_i^n - g(u_i^n) \quad (= v_{i+1/2}^n - g(u_{i+1/2}^n)) .$$

On peut également faire une démonstration directe pour ce cas, mais on sera amené pour établir (28) à utiliser le fait que $(u_{i,1}^{n+1}, v_{i,1}^{n+1})$ est obtenu à partir d'un problème de RIEMANN particulier.

3°/ Si $u_i^n \leq 0 \leq u_{i+1/2}^n$

On aura comme précédemment $u_{i,1}^{n+1} \in I(u_i^n, u_{i+1/2}^n)$. Supposons tout d'abord $u_{i,1}^{n+1}$ négatif.

On a, comme au cas N° 1, $v_{i,1}^{n+1} \in [v_i^n, v_{i+1/2}^n]$, donc

$$(29) \quad v_{i,1}^{n+1} + g(u_{i,1}^{n+1}) \leq v_{i+1/2}^n + g(u_{i+1/2}^n)$$

$$v_{i,1}^{n+1} + g(u_{i,1}^{n+1}) \geq v_i^n + g(u_i^n)$$

et puisque $u_{i,1}^{n+1}$ est négatif,

$$v_{i,1}^{n+1} - g(u_{i,1}^{n+1}) \geq v_{i,1}^{n+1} + g(u_{i,1}^{n+1}) \geq v_i^n + g(u_i^n)$$

Il reste à majorer $v_{i,1}^{n+1} - g(u_{i,1}^{n+1})$; on va démontrer l'inégalité suivante

$$(30) \quad v_{i,1}^{n+1} - g(u_{i,1}^{n+1}) \leq v_{i+1/2}^n$$

En posant $x = u_{i+1/2}^n$,

on a

$$v_i^{n+1} - g(u_{i,1}^{n+1}) = v_i^n - g(u_i^n) + 2q(f(x) - f(u_i^n))$$

$$- 2q \left\{ g \left[u_i^n + 2q[g(x) - g(u_i^n)] \right] - g(u_i^n) \right\}$$

c'est-à-dire

$$v_i^{n+1} - g(u_{i,1}^{n+1}) = v_{i+1/2}^n - g(x)$$

$$+ 2q \left\{ f(x) - f(u_i^n) - \frac{1}{2q} \int_{u_i^n}^{u_i^n + 2q[g(x) - g(u_i^n)]} \sqrt{f'(y)} dy \right\}$$

On veut établir, pour $a (= u_i^n) \leq 0$ et $x \geq 0$

$$(31) \quad 2q \left\{ f(x) - f(a) - \frac{1}{2q} \int_a^{a+2q(g(x)-g(a))} \sqrt{f'(y)} dy \right\} - g(x) \leq 0$$

Pour $x = 0$, cette inégalité est vraie, car on est alors dans le cadre du cas N° 2, et on a d'après (28)

$$(32) \quad 2q \left\{ -f(a) - \frac{1}{2q} \int_a^{a-2q g(a)} \sqrt{f'(y)} dy \right\} \leq 0$$

On dérive en x le premier membre de (31) ; il vient

$$\left[2q \left\{ \sqrt{f'(x)} - \frac{1}{2q} \sqrt{f'(a + 2q(g(x) - g(a)))} \cdot 2q \right\} - 1 \right] \sqrt{f'(x)}$$

puisque

$$\sqrt{f'(x)} = g'(x)$$

D'après la condition de stabilité, ce terme est négatif ou nul ; on intègre donc l'inégalité ainsi réalisée sur $[0, x]$, avec $x \geq 0$, et il vient

$$2q \left\{ f(x) - \frac{1}{2q} \int_{a-2q g(a)}^{a+2q(g(x)-g(a))} \sqrt{f'(y)} dy \right\} - g(x) \leq 0$$

dont on déduit (31) en lui ajoutant (32).

On a donc (30), qui entraîne car $u_{i+1/2}^n \geq 0$,

$$v_{i,1}^{n+1} - g(u_{i,1}^{n+1}) \leq v_{i+1/2}^n \leq v_{i+1/2}^n + g(u_{i+1/2}^n)$$

Tous les autres cas se ramènent à l'un des trois cas considérés ci-dessus.

Si $0 \leq u_{i+1/2}^n \leq u_i^n$, f est convexe sur $[u_{i+1/2}^n, u_i^n]$ et on applique le résultat obtenu pour le schéma S_1 .

Si $u_{i+1/2}^n \leq 0 \leq u_i^n$, on procède comme au cas N° 3 précédent, en remarquant qu'une inégalité du type (31) est vraie pour $u_{i+1/2}^n = 0$; on la dérive en $u_{i+1/2}^n$ et on la réintègre sur $[u_{i+1/2}^n, 0]$.

Si $u_{i+1/2}^n \leq u_i^n \leq 0$, on procède comme au cas N°1, en disposant d'une inégalité entre $\sqrt{f'(\xi_{i+1/4}^n)}$ et $g'(\xi_i^n)$.

Une technique analogue permet d'établir l'appartenance à K_0 de $(u_{i,2}^{n+1}, v_{i,2}^{n+1})$ déterminés par (23) ; les inégalités sur $v_{i,2}^{n+1} - g(u_{i,2}^{n+1})$ seront immédiates, comme l'ont été précédemment les inégalités sur $v_{i,1}^{n+1} + g(u_{i,1}^{n+1})$. On a cette fois-ci :

$$u_{i,2}^{n+1} = u_i^n + 2q (v_i^n - v_{i-1/2}^n) ,$$

$$u_{i,2}^{n+1} = u_i^n + 2q (g(u_{i-1/2}^n) - g(u_i^n)) ,$$

$$v_{i,2}^{n+1} = v_i^n + 2q (f(u_i^n) - f(u_{i-1/2}^n)) ,$$

$$v_{i,2}^{n+1} = v_i^n + 2q \sqrt{f'(\xi_{i-1/4}^n)} (v_{i-1/2}^n - v_i^n) ,$$

puisque

$$g(u_i^n) - g(u_{i-1/2}^n) = - (v_i^n - v_{i-1/2}^n) .$$

Si $0 \leq u_i^n \leq u_{i-1/2}^n$, on a $v_i^n \geq v_{i-1/2}^n$, et

$$v_{i,2}^{n+1} + g(u_{i,2}^{n+1}) = v_i^n + g(u_i^n) + 2q (\sqrt{f'(\xi_{i-1/4}^n)} - g'(\xi_i^n)) (v_{i-1/2}^n - v_i^n) .$$

où $\sqrt{f'(\xi_{i-1/4}^n)} \geq g'(\xi_i^n)$,

donc $v_{i,2}^{n+1} + g(u_{i,2}^{n+1}) \leq v_i^n + g(u_i^n)$.

Les autres estimations sont immédiates.

Si $u_i^n \leq u_{i-1/2}^n \leq 0$, on est dans le cadre du schéma S_1 , et le résultat est donc déjà acquis.

Si $u_i^n \leq 0 \leq u_{i-1/2}^n$, on reprend exactement (31) avec $x = u_{i-1/2}^n$.

Les autres cas se démontrent de façon identique. Le théorème 20 est démontré.

x x
x

On peut faire plusieurs remarques sur le théorème 20.

La condition $\alpha = 1/2$ n'est sûrement pas optimale pour les schémas S_0 et S_1 . Elle n'a servi qu'à éviter les interactions d'ondes entre deux problèmes de RIEMANN élémentaires, et ainsi faciliter la démonstration. L'expérience numérique montre que la stabilité est effective pour $\alpha = 1$, et qu'elle est perdue de façon souvent catastrophique dès que $q\sqrt{f'(u)}$ devient strictement supérieur à 1.

Remarquons également que la démonstration proposée ici pour le schéma S_1 aurait pu être faite au théorème 2 pour le schéma de GODUNOV, mais la condition de stabilité s'en serait trouvée affaiblie : $\alpha = \frac{1}{2}$.

L'hypothèse de convexité sur $|f|$ sert essentiellement à deux niveaux. D'une part au théorème 19, elle limite les cas pour la stabilité dans K_0 de la solution du problème de RIEMANN élémentaire, mais ceci est vrai même lorsque K_0 n'est pas convexe. La convexité de $|f|$, et donc celle de K_0 , est par contre fondamentale au niveau de l'utilisation de l'inégalité de JENSEN.

4. QUELQUES REMARQUES ET RESULTATS NUMERIQUES

Pour établir un résultat de convergence pour l'un des schémas considérés ci-dessus, ce qui constituerait également un résultat d'existence globale pour le problème (5), (6), on a besoin de la convergence forte de u_n , pour traiter le terme non linéaire. Il semble que, comme dans le cas de l'équation scalaire, cette convergence puisse avoir lieu dans $L^1_{loc}(\mathbb{R} \times]0, T[)$, et en conséquence, on suppose désormais que

$$(33) \quad (u_0, v_0) \in \left\{ BV_{loc}(\mathbb{R}) \cap L^\infty(\mathbb{R}) \right\}^2 .$$

On sait, par exemple à partir du problème de RIEMANN que la variation totale de u_0 et de v_0 n'est pas conservée par la solution ; ce n'est également pas le cas pour la solution approchée calculée à partir

de l'un des schémas proposés ci-dessus. La variation sur les invariants de RIEMANN n'est pas toujours conservée, par exemple lorsqu'un choc apparaît, mais on n'enregistre alors qu'une perturbation faible de la variation des invariants de RIEMANN. Pour la solution approchée, la variation sur ces invariants de RIEMANN est également perturbée par les projections de type L^2 qu'on effectue à chaque pas, et ceci même lorsque la solution exacte est décrite par un invariant de RIEMANN classique (solution régulière). Ces différentes remarques montrent la difficulté, qui n'a pas été surmontée, que l'on éprouve en voulant établir la conservation de la variation bornée par le schéma numérique.

En supposant qu'il y ait une estimation de la variation sur la solution, et donc séparément sur u_h et sur v_h , on est assuré de la convergence dans $L^1_{loc}(\mathbb{R} \times]0, T[)$ d'une suite extraite (u_h, v_h) . Il reste à vérifier que dans ce cas, la limite (u, v) est solution m -faible de (5). D'après la stabilité dans K_0 , il est immédiat que u et v sont uniformément bornés, et il en est de même du couple de fonctions (\bar{u}_h, \bar{v}_h) , défini sur $\mathbb{R} \times]0, T[$ par les quantités intermédiaires construites par les schémas S_0 ou S_1 :

$$(34) \quad \begin{aligned} \bar{u}_h(x, t) &= u_{i+1/2}^n & \text{si } (x, t) \in I_i \times J_n \\ \bar{v}_h(x, t) &= v_{i+1/2}^n & \text{si } (x, t) \in I_i \times J_n \end{aligned}$$

Il est immédiat également que, la variation bornée étant admise, ce couple (\bar{u}_h, \bar{v}_h) converge vers la même limite (u, v) que (u_h, v_h) .

En introduisant deux fonctions test ϕ et ψ de $C^1_0(\mathbb{R} \times]0, T[)$, qu'on approche sur chaque pavé $I_i \times J_n$ par une projection de type L^2 , il vient

$$(35) \quad \sum_{i,n} \frac{1}{qh} (u_i^{n+1} - u_i^n) \phi_i^n qh^2 = \sum_{i,n} \frac{1}{h} (v_{i+1/2}^n - v_{i-1/2}^n) \psi_i^n qh^2,$$

ce qui donne à la limite, après un regroupement des termes reconstituant une intégration par parties

$$(36) \quad \iint_{\mathbb{R} \times]0, T[} (u \frac{\partial \phi}{\partial t} - v \frac{\partial \phi}{\partial x}) dx dt = 0.$$

De la même façon, à partir de

$$\sum_{i,n} \frac{v_i^{n+1} - v_i^n}{qh} \psi_i^n qh^2 = \sum_{i,n} \frac{f(u_{i+1/2}^n) - f(u_{i-1/2}^n)}{h} \psi_i^n qh^2,$$

on établit

$$(37) \quad \iint_{\mathbb{R} \times]0, T[} (v \frac{\partial \psi}{\partial t} - f(u) \frac{\partial \psi}{\partial x}) dx dt = 0.$$

On est assuré également que u et v satisfont aux conditions initiales respectives u_0 et v_0 ; la démonstration est identique au cas scalaire. On procède de la même façon pour le schéma de LAX.

x x
x

Il resterait encore à montrer que la limite satisfait à une certaine condition d'entropie, qui serait exprimée comme pour l'équation scalaire dans une formulation faible. On ne dispose pas pour encore d'une telle formulation, mais on peut remarquer, comme BARDOS dans [1], CRANDALL ou B. KEYFITZ [1], qu'une telle formulation, pour l'équation scalaire, peut être établie en exprimant que le semi-groupe associé est une contraction pour la norme L^1 . Les schémas considérés ici ne semblent pas satisfaire de façon immédiate à une telle propriété, et c'est d'ailleurs une des raisons pour laquelle la conservation de la variation bornée en espace n'est pas établie en général, la difficulté étant du même type.

Il existe également une inégalité d'énergie, qui correspond à une inégalité supplémentaire d'entropie proposée par LAX [4,5]. On l'obtient en multipliant les équations du système respectivement par $f(u)$ et

par v , en intégrant, après avoir supposé par exemple v nulle à l'infini, et en introduisant la fonction convexe

$$F(u) = \int_0^u f(y) dy .$$

On a alors

$$\frac{d}{dt} \int_{\mathbb{R}} \left(\frac{v^2}{2} + F(u) \right) dx = 0 .$$

L'inégalité d'entropie correspondante est

$$(38) \quad \frac{d}{dt} \int_{\mathbb{R}} \left(\frac{v^2}{2} + F(u) \right) dx \leq 0 ,$$

qui exprime que l'énergie va en décroissant. On ne peut pas non plus établir cette inégalité à partir des schémas numériques considérés ici ; ceci reviendrait à vouloir établir une stabilité dans L^2 , qu'on n'obtient pas en général dans le cas non linéaire. De plus, il n'est pas certain que (38) assure l'unicité, l'hypothèse de vraie non linéarité préconisée dans LAX [4] n'étant pas satisfaite.

x x
x

Dans le cadre du câble attaché à chaque extrémité, les conditions aux limites s'écrivent

$$(39) \quad v(0,t) = v(l,t) = 0 , \text{ pour tout } t \geq 0$$

La condition initiale doit également satisfaire à une condition de compatibilité, à savoir

$$(40) \quad \int_0^l u_0(x) dx = 0 .$$

Pour les schémas S_0 et S_1 , les conditions (39) sont introduites en posant pour $i = 0$ et pour $i = L = \left[\frac{l}{h} \right] - 1$, à la place de (VII.26) ou de (VII.27), respectivement

$$(41.i) \quad v_{1/2}^n = 0 \quad v_1^n + g(u_1^n ; u_{1/2}^n) = 0 ;$$

$$(41.ii) \quad v_{1/2}^n = 0 \quad v_1^n + g(u_1^n) - g(u_{1/2}^n) = 0 ;$$

$$(42.i) \quad v_{L+1/2}^n = 0 \quad v_L^n - g(u_L^n ; u_{L+1/2}^n) = 0 ;$$

$$(42.ii) \quad v_{L+1/2}^n = 0 \quad v_L^n - g(u_L^n) + g(u_{L+1/2}^n) = 0 .$$

On peut remarquer que ces diverses conditions ne perturbent pas la stabilité dans K_0 , et ainsi le théorème 20 reste vrai.

De plus, si la condition initiale discrétisée satisfait à

$$(43) \quad \sum_{i=1}^L u_i^0 = 0 ,$$

de façon à vérifier la condition de compatibilité (40), on a de proche en proche, pour tout n

$$\sum_{i=1}^L u_i^{n+1} = \sum_{i=1}^L u_i^n + q(v_{i+1/2}^n - v_{i-1/2}^n)$$

$$= \sum_{i=1}^L u_i^n ,$$

car

$$v_{1/2}^n = v_{L+1/2}^n = 0 .$$

On obtient ainsi que la condition (43) est conservée.

Pour approcher la fonction w , solution de l'équation des ondes (1), on utilisera des sommations de u_i^n . Pour plus de commodité, en précisant qu'il ne s'agit pas là de nouvelles conditions aux limites, on pose

pour tout n

$$u_0^n = u_{L+1}^n = 0$$

puis pour tout $i \in \{1, \dots, L+1\}$

$$(44) \quad w_i^n = \sum_{j=1}^i \frac{h}{2} (u_j^n + u_{j-1}^n)$$

$$w_0^n = 0$$

On retrouve, la condition (43) étant conservée, que la condition limite en $x = 1$ est satisfaite pour tout n :

$$w_{L+1}^n = 0$$

On peut procéder de façon analogue à (41), (42) pour introduire les conditions aux limites dans le schéma de LAX. On fait alors appel à la construction de ce schéma proposée au chapitre précédent en (VII.43.44.45). Il faut remarquer que cette construction est différente de celle utilisée au niveau de la démonstration de la stabilité dans K_0 .

En $x = 0$, on impose, comme en (41)

$$v_{1/2}^n = 0$$

et on calcule $u_{1/2}^n$ en utilisant (VII.43.i), c'est-à-dire

$$v_1^n = v_{1/2}^n - \frac{1}{q} (u_1^n - u_{1/2}^n)$$

ce qui donne, sachant que $v_{1/2}^n$ est nul

$$u_{1/2}^n = u_1^n + q v_1^n$$

On introduit ces valeurs $(u_{1/2}^n, v_{1/2}^n)$ dans (VII.31) et dans (VII.32), pour obtenir les formules suivantes en approchant f comme en (VII.45), ou de façon analogue pour évaluer $f(u_{1/2}^n)$:

$$(45) \quad \begin{aligned} u_1^{n+1} &= \frac{u_1^n + u_2^n}{2} + \frac{q}{2} (v_2^n + v_1^n) \\ v_1^{n+1} &= \frac{v_2^n - v_1^n}{2} + \frac{q}{2} (f(u_2^n) - f(u_1^n)) \end{aligned}$$

En effectuant une opération analogue en $x = 1$, il vient successivement

$$v_{L+1/2}^n = 0$$

puis en utilisant (VII.43.ii)

$$u_{L+1/2}^n = u_L^n - q v_L^n$$

On obtient

$$(46) \quad \begin{aligned} u_L^{n+1} &= \frac{u_L^n + u_{L-1}^n}{2} - \frac{q}{2} (v_L^n + v_{L-1}^n) \\ v_L^{n+1} &= -\frac{1}{2} (v_L^n - v_{L-1}^n) + \frac{q}{2} (f(u_L^n) - f(u_{L-1}^n)) \end{aligned}$$

On obtient, en calculant la somme des u_i^{n+1} donnés par (45), (46) et (VII.41)

$$\begin{aligned} \sum_{i=1}^L u_i^{n+1} &= \sum_{i=2}^{L-1} \left[\frac{1}{2} (u_{i+1}^n + u_{i-1}^n) \right] + \frac{u_1^n + u_2^n}{2} + \frac{u_i^n + u_{i-1}^n}{2} \\ &+ \frac{q}{2} \left\{ \sum_{i=2}^{L-1} (v_{i+1}^n - v_{i-1}^n) + \frac{v_1^n + v_2^n}{2} - \frac{v_L^n + v_{L-1}^n}{2} \right\} \\ &= \sum_{i=1}^L u_i^n \end{aligned}$$

en remarquant que chaque u_i^n apparaît deux fois, et que tous les v_i^n disparaissent deux à deux. La condition de compatibilité (43) est ainsi conservée. Le calcul des w_i^n peut être effectué en utilisant (44) comme indiqué précédemment.

Le schéma S_0 et le schéma de LAX ont été testés dans les mêmes conditions pour le problème constitué de l'équation (1) avec

$$f(w) = \frac{w^3}{3}$$

et les conditions initiales

$$(47) \quad w(x,0) = \begin{cases} x/2 & \text{si } 0 < x < 1/2 \\ (1-x)/2 & \text{si } 1/2 < x < 1 \end{cases}$$

$$\frac{\partial w}{\partial t}(x,0) = 0$$

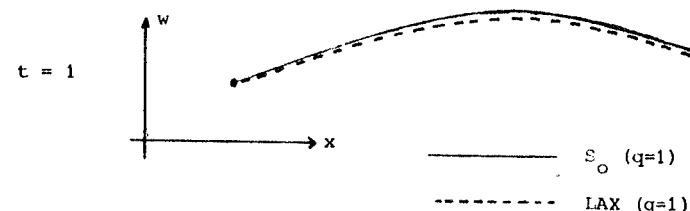
Les conditions aux limites correspondent au câble fixé à chaque extrémité.

$$w(0,t) = w(1,t) = 0$$

qui est compatible avec la donnée initiale (47). Ces conditions aux limites ont été interprétées numériquement comme indiqué ci-dessus.

L'expérience numérique a été menée conjointement pour le schéma de LAX et le schéma S_0 , avec deux valeurs différentes du rapport q : $q = 1$ et $q = \frac{1}{2}$. Le pas de discrétisation en espace h , est égal à 0.025, et les calculs ont été menés jusqu'à $T = 12.5$, soit 500 ou 1000 pas en temps, suivant les valeurs de q . Pour $t = 12.5$, les résultats donnés par S_0 pour $q=1$ et 0.5 sont très proches ; ceci permet de retrouver une propriété du schéma de GODUNOV pour l'équation scalaire. Comme pour l'équation scalaire également, le schéma de LAX donne des résultats très différents pour ces deux valeurs de q . Cette différence se traduit pour $q = \frac{1}{2}$ par un plus grand amortissement des oscillations, que pour $q = 1$.

Au temps $t = 1$, soit après 40 (ou 80) pas en temps, on obtient le résultat suivant en w :



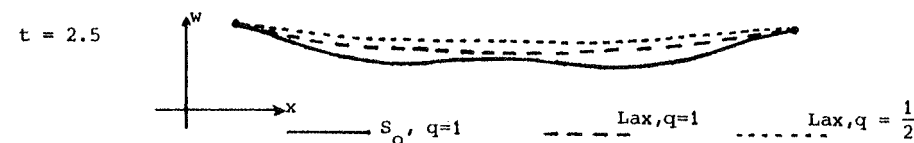
On peut remarquer, la solution étant constituée de détentes, que les schémas S_0 et S_1 coïncident, d'une part, et d'autre part en revenant aux invariants de RIEMANN établir que w est solution d'une équation scalaire de type BURGERS. Ceci permet de connaître explicitement la solution pour $t < 1$, c'est-à-dire avant que la détente n'atteigne l'extrémité fixe et ne s'y réfléchisse. On obtient en particulier

$$w(x,1) = \frac{1}{2} x(1-x)$$

ce qui donne en $x = \frac{1}{2}$, $w(\frac{1}{2}, 1) = 0.125$. Le schéma S_0 donne 0.126 et le schéma de LAX donne, pour $q = 1$, 0.118.

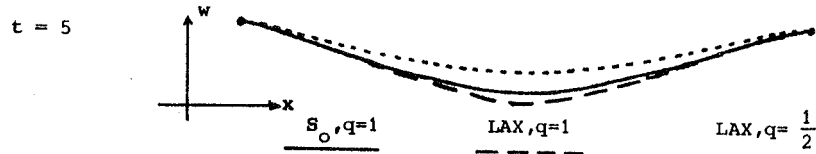
On n'a pas représenté ici la solution obtenue pour $q = \frac{1}{2}$ par le schéma de LAX. La différence n'est sensible que pour de plus grandes valeurs de t . La vitesse est de signe négatif ; ainsi la vitesse en $x = \frac{1}{2}$, $t = 1$, vaut - 0.089 pour le schéma de LAX ($q = 1$), et - 0.118 pour le schéma S_0 ($q = 1$).

Au temps $t = 2.5$, le câble a traversé sa position au repos, et w est devenu négatif. La vitesse est toujours négative et un certain retard est déjà accusé pour la solution construite par le schéma de LAX pour $q = 1/2$.



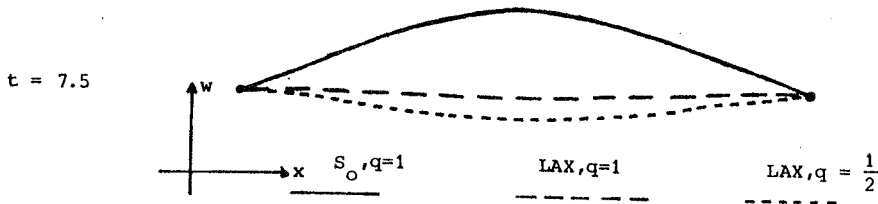
Ce retard est destiné à se creuser ; en effet la vitesse au milieu vaut - 0.122 pour le schéma S_0 , - 0.091 pour le schéma de LAX avec $q = 1$, et - 0.0604 pour le schéma de LAX avec $q = 1/2$. De plus, c'est au milieu du câble qu'on trouve la vitesse maximale.

Au temps $t = 5$, le câble a commencé à remonter pour le schéma S_0 , et est proche de son minimum pour le schéma de LAX.



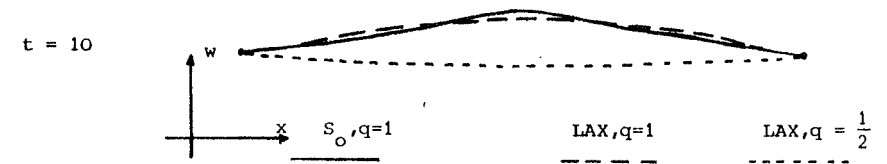
Les vitesses au milieu sont 0.094 pour S_0 , 0.0062 pour LAX ($q = 1$) et - 0.0203 pour LAX ($q = 1/2$). On peut remarquer que pour le schéma de LAX, $q = 1$, la vitesse est minimale au milieu et maximale en $x = 0.675$ ou 0.325 , où elle vaut 0.0216.

Au temps $t = 7.5$, la différence est très nette, entre le schéma S_0 et le schéma de LAX : la position calculée par S_0 est proche de la position au maximum, tandis que la position calculée pour le schéma de LAX n'a pas encore traversé la position du câble au repos.



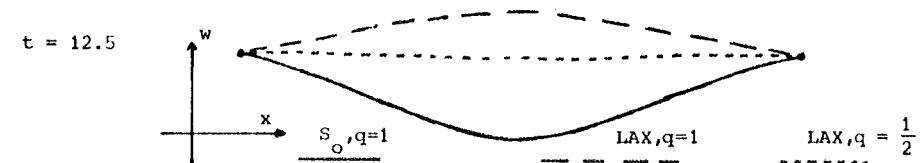
Les vitesses au milieu sont respectivement 0.105 pour S_0 , 0.037 pour LAX ($q = 1$) et 0.0025 pour LAX ($q = 1/2$). Pour ce dernier, la vitesse est maximale pour $x = 0.275$ ou 0.725 , et vaut 0.0035, c'est-à-dire que le câble ne bouge presque plus. Pour les autres schémas, la vitesse est maximale au milieu, et de signe constant.

Au temps $t = 10$, le câble descend pour le schéma S_0 , et monte pour le schéma de LAX, $q = 1$. La position est presque stationnaire pour le schéma de LAX, $q = 1/2$.



Les vitesses sont maximales au milieu, et valent respectivement - 0.086 pour S_0 , + 0.026 pour LAX, $q = 1$, et 0.0031 pour LAX, $q = 1/2$. On peut remarquer que les positions calculées par les schémas de LAX, $q = 1$, et par S_0 sont proches, avec des vitesses différentes.

Au temps $t = 12.5$, la position calculée par S_0 est proche du minimum, et celle calculée par le schéma de LAX, $q = 1$ est proche du maximum, cette position minimale n'étant pas encore atteinte pour S_0 . On peut remarquer la différence d'amplitude entre les deux positions. Les amplitudes au milieu sont respectivement -0.155 pour S_0 , 0.069 pour LAX, $q=1$ et -0.004 pour LAX, $q = 1/2$.



En fait, cette différence est légèrement plus importante si on tient compte des vitesses au milieu, qui sont maximales. Elles valent - 0.09 pour S_0 , + 0.0037 pour le schéma de LAX ($q = 1$), et 0.0017 pour le schéma de LAX ($q = \frac{1}{2}$). Ceci traduit un amortissement assez sensible pour le schéma de LAX ($q = 1$) : on peut en effet estimer qu'il atteint ici une position très proche de la position maximale, la vitesse n'étant pas de signe constant. Elle vaut par exemple - 0.0028 en $x = 0.225$.

Pour le schéma S_0 , avec $q = \frac{1}{2}$, on obtient pour $t = 12.5$, une amplitude au milieu égale à - 0.143, avec la vitesse - 0.088. La différence avec le schéma S_0 , $q = 1$ peut être estimée comme très faible, au bout de 1000 pas de temps.

Cette expérience numérique a fait également apparaître le maintien de la discontinuité initiale en u , avec cependant une intensité allant en décroissant, pour le schéma S_0 . Cette discontinuité a disparu dès que l'onde réfléchie a atteint pour la première fois le milieu du câble, c'est-à-dire vers $t = 2.5$. On peut rapprocher ce phénomène de la remarque suivante : au niveau de la stabilité pour le problème de RIEMANN, seuls les chocs permettent de pénétrer davantage dans K_0 et au niveau de l'approximation par S_0 , c'est la projection de type L^2 qui joue ce rôle. Pour l'équation scalaire, choc et projection de type L^2 , c'est-à-dire viscosité, avaient des effets opposés.

x x
x

CONCLUSION

La construction du schéma de GODUNOV, proposée au chapitre II, et dont le principe se résume en une résolution de problèmes de RIEMANN puis à des projections de type L^2 , s'adapte au système d'ordre deux étudié en troisième partie. La généralisation de cette technique au problème à plusieurs dimensions est liée à la résolution, et avant tout à une définition convenable du problème de RIEMANN à plusieurs dimensions.

Au cours de l'étude des schémas quasi d'ordre deux, le fait que le schéma de LAX WENDROFF soit d'ordre deux en précision, n'est pas fondamental. Cette remarque doit permettre de généraliser le résultat à des schémas plus précis. Cette technique peut également s'adapter à d'autres schémas d'ordre deux, tels que ceux décrits par R. PEYRET. Il faut également préciser qu'on ne connaît pas d'exemple de solutions faibles ne vérifiant pas la condition d'entropie, et construites par le schéma étudié au théorème 7.

Le problème avec condition aux limites étudié en seconde partie peut être rapproché de certains problèmes de perturbations singulières ou de couches limites. La définition 3 caractérise la solution obtenue par la méthode de viscosité, c'est-à-dire le problème perturbé. Au niveau de l'approximation, cette perturbation est rétablie par de la viscosité numérique, dépendante de h , et qui disparaît lorsque h tend vers zéro. L'habitude, en matière de perturbations singulières, consiste souvent à supposer "ε petit par rapport à h", et cette hypothèse se réduit à ε nul, au niveau de la convergence, c'est-à-dire lorsque h tend vers zéro. On pourrait peut-être comparer certains résultats obtenus ici, avec les techniques classiques d'approximation de couche limites, pour de très petites valeurs du paramètre ε.

En ce qui concerne la troisième partie, le résultat de stabilité pour le problème de RIEMANN reste vrai lorsque l'ensemble K_0 n'est pas convexe. Ce n'est plus le cas au niveau de la stabilité des schémas, les projections de type L^2 risquant de faire sortir de K_0 la solution approchée, lorsque K_0 est non convexe, et ceci à chaque pas, compromettant ainsi la stabilité. Cette projection fonctionne alors en sens inverse de la condition d'entropie. Pour éviter d'effectuer cette projection, et à condition de savoir résoudre le problème de RIEMANN, ce qui peut se faire en utilisant la généralisation de l'invariant de RIEMANN proposée au chapitre VII, on peut adopter le schéma de GLIMM. De nombreux auteurs l'ont utilisé pour établir des résultats d'existence (GLIMM, NISHIDA, SMOLLER, etc...) ou au niveau de l'approximation proprement dite, en ayant des résultats satisfaisants (CHORIN).

Même lorsque la solution est régulière, c'est-à-dire régie par un invariant de RIEMANN (classique), le fait de projeter à chaque pas a pour conséquence de faire s'éloigner la solution approchée de cet invariant de RIEMANN. C'est une des raisons pour laquelle il nous manque une estimation sur les dérivées des composantes de la solution approchée, de façon à établir la convergence.

Le modèle de câble peut être généralisé, en introduisant un déplacement variable dans chaque direction de l'espace. Le problème obtenu est constitué de trois équations d'ondes, couplés par leurs seconds membres non linéaires. Les techniques d'approximation étudiées ici peuvent également être étendues à des problèmes hyperboliques d'ordre deux plus généraux. Bien que la recherche des invariants de RIEMANN soit souvent plus difficile, il existe des exemples où le schéma S_0 est facilement applicable et pour lesquels l'ensemble K_0 est convexe.

x x
x

BIBLIOGRAPHIE

G.A. BAKER A finite element method for first order hyperbolic equations. Math. Comp. 29 (132) (1975), p. 995-1006.

N.S. BAKVAROV On the existence in the large for quasilinear systems, Z. Vycisl. Fiz. 10 (1970), p. 969-980 (en russe).

Cl. BARDOS (1) Problèmes d'évolution non linéaires. Séminaire de Nice, (1974) (1975) (1976).

Cl. BARDOS, D. BREZIS, H. BREZIS Perturbations singulières et prolongements maximaux d'opérateurs positifs. Arch. Rat. Mech. An., 53, 1 (1973), p. 69-100.

Cl. BARDOS, A.Y. LE ROUX, J.Cl. NEDELEC First order quasi linear equations with boundary conditions. A paraître + Rapport interne N° 38, Centre de Mathématiques Appliquées, Ecole Polytechnique (1978).

D.L. BOOK, J.P. BORIS, K. HAIN Flux corrected transport II : Generalization of the method. J. Comp. Physics. 18 (1975), p. 248-283.

D.L. BOOK Voir J.P. BORIS, D.L. BOOK.

J.P. BORIS, D.L. BOOK (1) Flux corrected transport I : SHASTA a fluid transport algorithm that works. J. Comp. Physics, 11 (1973), p. 38-69.

J.P. BORIS, D.L. BOOK (2) Flux corrected transport III : Minimal error FCT algorithms. J. Comp. Physics, 20 (1976), p. 397-431.

- D. BREZIS Voir Cl. BARDOS, D. BREZIS, H. BREZIS.
- H. BREZIS Opérateurs maximaux monotones. North Holland (1973).
- H. BREZIS Voir Cl. BARDOS, D. BREZIS, H. BREZIS.
- P.H. CHANG (1) On the break down phenomena of solutions of quasilinear wave equations. Michigan Math. Journal, 23 (1976) p.277-287.
- P.H. CHANG (2) On the existence of shock curves of quasilinear wave equations. Indiana Univ. Math. J. 26 (3) (1977), p. 605-622.
- A.J. CHORIN Random choice solution of hyperbolic systems. J. Comp. Physics 22 (4). (1976), p. 517-533.
- K.N. CHUEH, C.C. CONLEY, J.A. SMOLLER Positively invariant regions for systems of nonlinear diffusion equations. Indiana Univ. Math. J. Vol. 26 (2) (1977), p. 373-392.
- C.C. CONLEY Voir K.N. CHUEH, C.C. CONLEY, J.A. SMOLLER.
- C.C. CONLEY, J.A. SMOLLER (1) Viscosity Matrices for twodimensional nonlinear hyperbolic systems. Comm. Pure Applied. Math 23 (1970), pp. 867-884.
- C.C. CONLEY, J.A. SMOLLER (2) Shock waves as limits of progressive wave solutions of higher order equations. Comm. Pure Applied. Math. 24 (1971), p. 459-472.
- E. CONWAY, J.A. SMOLLER Global solution of the Cauchy problem for quasilinear first order equations in several space variables. Comm. Pure Applied. Math. 19 (1966), p. 95-105.
- R. COURANT, K.O. FRIEDRICHS, H. LEWY Über die partielle Differenzengleichungen der Mathematischen Physik. Math. Ann. 100 (1928), p. 32.

- M.G. CRANDALL The semi group approach to first order quasilinear equations in several space variables. Israël J. of Math. 12 (1972) p. 108-132.
- M.G. CRANDALL, T.M. LIGGETT Generation of semi group of non linear transformation on general Banach spaces. Amer. J. of Math. 93 (1971), p. 265-298.
- N. CRISTESCU Dynamic Plasticity. North Holland (1967).
- C.M. DAFERMOS (1) Solution of the Riemann problem for a class of hyperbolic systems by the viscosity method. Arch. Rat. Mech. Anal. 52 (1974), p. 1-9.
- C.M. DAFERMOS (2) Structure of solutions of the Riemann problem for hyperbolic systems of conservation laws. Arch. Rat. Mech. Anal. 53 (1973), p. 203-217.
- C.M. DAFERMOS (3) Polygonal approximations of solutions to the initial value problem for a conservation law. J. Math. Anal. Appl. 38 (1972), p. 33-41.
- C.M. DAFERMOS (4) The entropy rate admissibility criterion for solutions of hyperbolic conservation laws. J. Diff. Eq. 14 (1973), p. 202-212.
- C.M. DAFERMOS, R.J. DI PERNA The Riemann problem for certain classes of hyperbolic systems of conservation laws. J. Diff. Eq. 20 (1976), p. 90-114.
- R.J. DI PERNA (1) Existence in the large for quasilinear hyperbolic conservation laws. Arch. Rat. Mech. Anal. 52 (1973), p.244-257.
- R.J. DI PERNA (2) Global solutions to a class of nonlinear hyperbolic systems of equations. Comm. Pure. Applied. Math. 26 (1973), p. 1-28.

- R.J. DI PERNA (3) Singularities of solutions of nonlinear hyperbolic systems of conservation laws. Arch. Rat. Mech. Anal. 60 (1976), p. 75-100.
- R.J. DI PERNA (4) Global existence of solutions to nonlinear hyperbolic systems of conservation laws. Preprint.
- R.J. DI PERNA (5) Uniqueness of solutions to hyperbolic conservation laws. Preprint.
- R.J. DI PERNA Voir C.M. DAFERMOS, R.J. DI PERNA.
- A. DOUGLIS (1) On calculating weak solutions of quasilinear first order partial differential equations. Contrib. to Diff. Eq. 1.1. (1963), p. 59-94.
- A. DOUGLIS (2) Layering methods to nonlinear partial differential equations of first order. Annales de l'Institut Fourier. 22 (3), (1972), p. 141-227.
- A. DOUGLIS (3) Lectures on discontinuous solutions of first order nonlinear partial differential equations. North British Symp. on P.D.E. (1972).
- A. DOUGLIS (4) The continuous dependence of generalized solutions of nonlinear partial differential equations upon initial data. Comm. Pure Applied. Math. 14 (1961), p. 267-287.
- A. FRIEDMAN Partial differential equations of parabolic type. Prentice Hall, New York (1964).
- K.O. FRIEDRICHS Voir R. COURANT, K.O. FRIEDRICHS, H. LEWY.
- J. GLIMM Solutions in the large for nonlinear hyperbolic systems of equations. Comm. Pure Applied. Math. 18 (1965), p.697-715.
- J. GLIMM, P.D. LAX Decay of solutions of systems of nonlinear hyperbolic conservation laws. Memo. Amer. Math. Soc. 101 (1970).

- S.K. GODUNOV (1) Equations de la physique mathématique. ED. MIR. Moscou (1973).
- S.K. GODUNOV (2) On the uniqueness of the solution of the equations of hydrodynamics. Mat. Sb. 40 (1956), p. 467-478 (en russe).
- S.K. GODUNOV (3) A difference method for the calculation of shock waves.
- S.K. GODUNOV (4) Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics Math. Sb. 47 (3), (1959), p. 271.
- J.M. GREENBERG (1) On the existence, uniqueness and stability of solutions of the equation $\rho_0 \chi_{tt} = E(\chi_x) \chi_{xx} + \lambda \chi_{xxt}$. J. Math. Anal. Applic. 25 (1969), p. 575-591.
- J.M. GREENBERG (2) Estimates for fully developed shock solution of the equation $\frac{\partial u}{\partial t} - \frac{\partial v}{\partial x} = 0$ and $\frac{\partial v}{\partial t} - \frac{\partial \sigma(u)}{\partial x} = 0$. Indiana Univ. Math. J. 22 (10) (1973), p. 989-1003.
- K. HAIN Voir D.L. BOOK, J.P. BORIS, K. HAIN.
- P.R. HALMOS Measure theory, Van Nostrand. New York (1950).
- A. HARTEN, J.M. HYMAN, P.D. LAX On difference approximations and entropy conditions for shocks. Comm. Pure Applied. Math. 29 (1976) p. 297-322.
- E. HOPF (1) The partial differential equation $u_t + uu_x = \mu u_{xx}$. Comm. Pure Applied Math. 3 (1950), p. 201-230.
- E. HOPF (2) Generalized solutions of nonlinear equation of first order. J. Math. Mech. 14 (1965), p. 951-972.
- E. HOPF (3) On the right weak solution of the Cauchy problem for a quasilinear equation of first order. J. Math. Mech. 19 (1969), p. 483-487.

- A.E. HURD A uniqueness theorem for second order quasilinear hyperbolic equations. Pacific J. of Math. 32 (2) (1970), p.415-427.
- J.M. HYMAN Voir A. HARTEN, J.M. HYMAN, P.D. LAX.
- A. JEFFREY Quasilinear hyperbolic systems and waves. Pitman Publishing, London (1976).
- A. JEFFREY, T. TANIUTI Nonlinear wave propagation. Academic Press NY 1964.
- F. JOHN Formation of singularities in one dimensional nonlinear wave propagation. Comm. Pure Applied Math 27 (1974), p. 377-405.
- J.L. JOHNSON, J.A. SMOLLER (1) Global solutions of certain systems of quasilinear hyperbolic equations. J. Math. Mech. 17 (1967), p. 561-576.
- J.L. JOHNSON, J.A. SMOLLER (2) Global solutions for an extended class of hyperbolic systems of conservation laws. Arch. Rat. Mech. Anal. 32 (1969), p. 169-189, + Erratum. Arch. Rat. Mech. 37 (1970), p. 399-400.
- P. JOUVE Contribution à l'étude du comportement non linéaire des structures à barres. Thèse d'Etat. Rennes, 1976.
- B.L. KEYFITZ (1) (= B.K. QUINN) Solution with shocks : An example of an L_1 -contractive Semigroup. Comm. Pure Applied Math. 24 (1971), p. 125-132.
- B.L. KEYFITZ (2) Voir Appendice de l'article de A. HARTEN, J. HYMAN, P.D. LAX.
- B.L. KEYFITZ, H.C. KRANZER Existence and uniqueness of entropy solutions to the Riemann Problem for hyperbolic systems of two non linear conservation laws. J. of Diff. Eq. 27 (1978), p. 444-476.

- H.C. KRANZER Voir B.L. KEYFITZ, H.C. KRANZER.
- S.N. KRUKOV (1) Generalized solutions for the Cauchy problem in the large for nonlinear equation of first order. Soviet Math. Dokl. 10 (1969), p. 785-788.
- S.N. KRUKOV (2) First order quasilinear equations in several independent variables. Math. USSR Sb. 10 (1970), p. 217-243.
- O.A. LADYZENSKAJA, N.N. URAL'CEVA Boundary problems for linear and quasilinear parabolic equations. AMS Transl. ser. 2, Vol. 47, P. 217-299 (1965).
- P.D. LAX (1) Weak solutions of nonlinear hyperbolic equations and their numerical computation. Comm. Pure Appl. Math. 7 (1954), p. 159-193.
- P.D. LAX (2) Hyperbolic systems of conservation laws II. Comm. Pure Applied Math. 10 (1957), p. 537-566.
- P.D. LAX (3) The formation and decay of shock waves. Amer. Math. Monthly 79 (1972), p. 227-241.
- P.D. LAX (4) A concept of Entropy. Contribution to nonlinear functional Analysis. E.H. Zarantonello ed. Academic Press NY (1971).
- P.D. LAX (5) Shock waves and entropy = autre titre quelquefois donné au précédent.
- P.D. LAX (6) Development of singularities of solution of nonlinear hyperbolic partial differential equations. J. Math. Phys. 5 (1964), p. 611-613.
- P.D. LAX Voir J. GLIMM et P.D. LAX ; A. HARTEN, J. HYMAN et P.D. LAX
- P.D. LAX, B. WENDROFF (1) Systems of conservation laws. Comm. Pure Applied Math. 13 (1960), p. 217-237.

- P.D. LAX, B. WENDROFF (2) Difference schemes for hyperbolic equations with high order of accuracy. *Comm. Pure Applied. Math* 17 (1964), p. 381-398.
- L. LEIBOVICH Solutions of the Riemann problem for hyperbolic systems of quasilinear equations without convexity conditions. *J. Math. Anal. Appl.*
- A.Y. LE ROUX (1) Résolution numérique du problème de Cauchy pour une équation quasilineaire du premier ordre à une ou plusieurs variables d'espace. Thèse 3e Cycle, Rennes (1974).
- A.Y. LE ROUX (2) A numerical conception of entropy for quasilinear equations. *Math. Comp.* 31 (140) (1977), p. 848-872.
- A.Y. LE ROUX (3) Problèmes hyperboliques d'évolution. Note LA 44. CELAR. RENNES (1977).
- A.Y. LE ROUX (4) Convergence of the Godunov scheme for first order quasilinear equations. *Proc. Japan Acad.* 52 (9) (1976). p. 488-491.
- A.Y. LE ROUX (5) Etude du problème mixte pour une équation quasi linéaire du premier ordre. *CRAS, Ser. A.* 285 (1977), p.351-354.
- A.Y. LE ROUX (6) Problème mixte pour une équation quasi linéaire du premier ordre. *Publications des laboratoires d'Analyse Numérique et Mécanique.* RENNES, 1977.
- A.Y. LE ROUX (7) Problème aux limites pour une équation quasi linéaire du premier ordre. *Journées Eléments Finis.* RENNES (1978).
- A.Y. LE ROUX (8) Vanishing viscosity method for a quasilinear first order equation with boundary conditions. *Conference of the Numerical Analysis of Singular Perturbations problems.* Nijmegen (1978), à paraître à Academic Press (J.J.H. Miller ed.).

- A.Y. LE ROUX (9) Convergence d'un schéma quasi d'ordre deux pour une équation quasi linéaire du premier ordre. A paraître *CRAS* (1979).
- A.Y. LE ROUX Voir Cl. BARDOS, A.Y. LE ROUX, J.C. NEDELEC.
- P. LESAINTE Numerical solution of the equation of continuity. *Topics in numerical analysis III* (J.J.H. Miller ed.). Academic Press (1977).
- H. LEWY Voir R. COURANT, K.O. FRIEDRICHS, H. LEWY.
- T.M. LIGGETT Voir M.C. GRANDALL, T.M. LIGGETT.
- T.P. LIU (1) The Riemann problem for general 2×2 conservation laws. *Trans. of the Amer. Math. Soc.* 199 (1974), p. 89-112.
- T.P. LIU (2) Existence and uniqueness theorems for Riemann problems. *Trans. of the Amer. Math. Soc.* 212 (1975), p. 375-382.
- T.P. LIU (3) Uniqueness of weak solutions of the Cauchy problem for general 2×2 conservation laws. *J. of Diff. Eq.* 20 (1976), p. 369-388.
- T.P. LIU (4) The entropy condition and the admissibility of shocks. *J. of Math. Anal. and Appl.* 53 (1976), p. 78-88.
- T.P. LIU (5) Decay to N-waves of solutions of general systems of non linear hyperbolic conservation laws. Preprint (1977).
- K.W. MORTON Voir R.D. RICHTMYER et K.W. MORTON.
- J.P. NATANSON *Theory of functions of a real variable.* Vol I. II, 3e Ed. (1964), Fred Ungar Publ. New York.
- J.C. NEDELEC Voir Cl. BARDOS, A.Y. LE ROUX, J.C. NEDELEC.

- . NISHIDA (1) Global solutions for an initial boundary value problem of a quasilinear hyperbolic system. Proc. Japan Acad. 44 (1968), p. 642-648.
- . NISHIDA (2) Nonlinear hyperbolic equations and related topics in fluid dynamics. Cours de D.E.A. Orsay, 1977.
- . NISHIDA, J.A. SMOLLER (1) Solutions in the large for some nonlinear hyperbolic conservation laws. Comm. Pure Applied Math. 26 (1973), p. 183-200.
- . NISHIDA, J.A. SMOLLER (2) Mixed problems for nonlinear conservation laws. J. of Diff. Eq. 23 (2), (1977), p. 244-269.
- . NISHIDA, M. YAMAGUTI On some global solution for the quasilinear hyperbolic equations. Funkcial. Ekvac. 2 (1968), p. 51-57.
- . OHARU, T. TAKAHASHI A convergence theorem for nonlinear semigroup and its applications to first order quasilinear equation.
- . O.A. OLEINIK (1) On discontinuous solutions of nonlinear differential equations. Dokl. Akad. Nauk SSSR 109 (1956), p. 1098-1101.
- . O.A. OLEINIK (2) Construction of a generalized solution of the Cauchy problem for a quasilinear equation of first order, by the introduction of vanishing viscosity. Amer. Math. Soc. Transl. Ser. 2, 33 (1963), p. 277-283.
- . O.A. OLEINIK (3) Uniqueness and stability of the generalized solution of the Cauchy problem for a quasilinear equation. Amer. Math. Soc. Transl. Ser. 2, 33 (1963), p. 285-290.
- . O.A. OLEINIK (4) Discontinuous solution of nonlinear differential equations. Amer. Math. Soc. Transl. Ser. 2, 26 (1963), p. 95-172.
- . OVADIA, P.A. RAVIART Nouvelles méthodes numériques pour le traitement de la convection. CEA Limeil (1978).

- R. PEYRET Résolution numérique des systèmes hyperboliques ; application à la dynamique des gaz. ONERA, Paris, 5 (1977).
- B.K. QUINN Voir B.L. KEYFITZ.
- P.A. RAVIART Pseudo viscosity methods and nonlinear hyperbolic equations. Proc. Roy. Soc. Lond. 1, 323 (1971), p. 277-283.
- P.A. RAVIART Voir J. OVADIA, P.A. RAVIART.
- R.D. RICHTMYER, K.W. MORTON Difference methods for initial value problems. Interscience Publ. & J. Wiley (2e ed.) New York (1967).
- B. RIEMANN Über die Fortpflanzung ebener Luftwellen von endlicher Schwingungsweite. Gesammelte Mathematische Werke (H. Weber ed.), p. 145-164. Leipzig, B.G. Teubner (1876).
- P. ROUSSEL Numerical solution of static and dynamic equations of cables. Comp. Methods in Appl. Mech. Ing. 9 (1976), p. 65-74.
- S. SAKS Theory of the Integral. Warsaw (1937).
- D. SHAEFFER A regularity theorem for conservation laws. Report NSF, Contact GP 22927, (1975).
- J.A. SMOLLER (1) On the solution of the Riemann problem with general step data for an extended class of hyperbolic systems. Michigan Math. J. 16 (1969), p. 201-210.
- J.A. SMOLLER (2) A uniqueness theorem for Riemann problems. Arch. Rat. Mech. Anal. 33 (1969), p. 110-115.
- J.A. SMOLLER (3) Contact discontinuities in quasilinear hyperbolic systems. Comm. Pure Applied. Math. 23 (1970), p. 791-801.
- J.A. SMOLLER Voir K.N. CHUEH, C.C. CONLEY et J.A. SMOLLER ; C.C. CONLEY et J.A. SMOLLER ; E. CONWAY et J.A. SMOLLER ; J.L. JOHNSON et J.A. SMOLLER ; T. NISHIDA et J.A. SMOLLER.

T. TAKAHASHI Voir S. OHARU, T. TAKAHASHI.

T. TANIUTI Voir A. JEFFREY, T. TANIUTI.

L. TARTAR Communication orale. 1979.

N.N. URAL'CEVA Voir O.A. LADYZENSKAJA, N.N. URAL'CEVA.

B. VAN LEER (1) Towards the ultimate conservative difference scheme III. Upstream centered finite difference schemes for Ideal compressible flow. J. of. Comp. Phys. 23 (1977), p. 263-275.

B. VAN LEER (2) Towards the ultimate conservative difference scheme IV. A New approach of numerical convection 23 (1977), p. 276-299.

B. VAN LEER (3) Towards the ultimate conservative difference scheme V. A second order sequel to Godunov's method (in preparation).

M.I. VIZIK Solvability of boundary problems for quasilinear parabolic equations of higher order. Mat. Sb. 59 (101) (1962), p. 289-325.
- Amer. Math. Soc. Transl. Ser. 2 (65), (1967), p. 1-40.

A.I. VOL'PERT The space B.V. and quasilinear equations. Math. Sb. 73 (115) (1967), p. 255-302 = Math. USSR Sb. 2 (1967), p. 225-267.

B. WENDROFF The Riemann problem for Materials with nonconvex equation of state I. Isentropic Flow. J. Math. Anal. Appl. 38 (1977) p. 454-466. + II. General Flow. J. Math. Anal. Appl. 38 (1977), p. 640-658.

B. WENDROFF Voir P.D. LAX et B. WENDROFF.

M. YAMAGUTI Voir T. NISHIDA et M. YAMAGUTI.

K. YOSHIDA Functional Analysis. Springer Verlag Berlin (1965).

SOMMAIRE

INTRODUCTION 1

PREMIERE PARTIE : APPROXIMATION POUR LE PROBLEME DE CAUCHY 5

CHAPITRE I : QUELQUES GENERALITES SUR LE PROBLEME DE CAUCHY 7

1. Quelques propriétés des solutions faibles 7

 Définition 1 (solution faible) 10

2. La condition d'entropie 12

3. La méthode de pseudo-viscosité 17

 Lemme 1 (SAKS) 18

 Lemme 2 (Régularisation par sg_η) 20

4. La définition de KRUKOV, et le théorème d'unicité 23

 Définition 2 (KRUKOV) 23

 Théorème 1 (Existence et unicité) 24

x

CHAPITRE II : LE SCHEMA DE GODUNOV 31

1. Le problème de RIEMANN 31

2. Construction du schéma de GODUNOV 37

 Théorème 2 (Convergence) 40

3. Démonstration du théorème 2 42

CHAPITRE III : ETUDE DE QUELQUES SCHEMAS D'ORDRE UN
 ET D'ORDRE DEUX 63

1. Construction d'une famille de schémas pour le problème
 monodimensionnel 63
 Théorème 3 (Convergence vers une solution faible) 68
 Théorème 4 (Convergence + condition d'entropie) 69

2. Approximation pour le problème à plusieurs dimensions 70
 Théorème 5 (Convergence) 72
 Théorème 6 (Schéma décentré) 74

3. Un schéma décentré quasi d'ordre deux 77
 Théorème 7 (Convergence vers une solution faible) 80
 Théorème 8 (Convergence + condition d'entropie) 84

4. Généralisation au cas f non nécessairement monotone 89
 Théorème 9 (GODUNOV quasi d'ordre deux) 90

5. Généralisation du schéma décentré au problème à p dimensions 97
 Théorème 10 (Schéma quasi d'ordre deux) 99

6. Quelques remarques et résultats numériques 104

x x
 x

DU XIEME PARTIE : LE PROBLEME AUX LIMITES 111

CHAPITRE IV : LE PROBLEME AVEC CONDITIONS AUX LIMITES 113

1. Introduction des conditions aux limites 113
 Lemme 3 (Théorème de trace sur $BV(\Omega)$) 120

2. Existence et unicité pour le problème aux limites 122
 Définition 3 (Entropie + conditions aux limites) 124
 Théorème 11 (Existence et unicité) 125

x

CHAPITRE V : APPROXIMATION POUR LE PROBLEME AUX LIMITES 143

1. Le schéma de GODUNOV 143
 Théorème 12 (Convergence) 145

2. Le schéma de GODUNOV quasi d'ordre deux 152
 Théorème 13 (Convergence) 153

3. Le schéma de LAX à plusieurs dimensions 155
 Théorème 14 (Convergence) 156

4. Quelques remarques et résultats numériques 168

x

CHAPITRE VI : APPROXIMATION DU PROBLEME AUX LIMITES DANS
 LE CAS MONOTONE 173

1. Le schéma décentré, à une ou plusieurs variables 173
 Théorème 15 (Convergence) 174

2. Le schéma quasi d'ordre deux, à une ou plusieurs variables 179
 Théorème 16 (Convergence) 179

3. Un schéma quasi explicite d'ordre deux en temps 184
 Théorème 17 (Convergence) 186

4. Un schéma quasi explicite pour les équations avec
 migration 190
 Théorème 18 (Convergence) 192

5. Quelques remarques et résultats numériques 194

x x
 x

TROISIEME PARTIE : UNE EQUATION DES ONDES NON LINEAIRE 201

CHAPITRE VII : APPROXIMATION DE L'EQUATION DES ONDES
 NON LINEAIRE 203

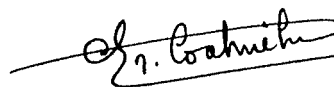
1. Une généralisation des invariants de RIEMANN 204

2. Le problème de RIEMANN 210

3. Construction d'un schéma numérique	216
4. Quelques remarques et résultats numériques	222
x	
<u>CHAPITRE VIII</u> : APPLICATION A UN MODELE DE CABLE EXTENSIBLE	231
1. Un modèle de dynamique des câbles extensibles	231
2. Stabilité pour le problème de RIEMANN	236
Théorème 19 (Existence, unicité, stabilité)	237
3. Stabilité pour les schémas de LAX, S_0 et S_1	245
Théorème 20 (Stabilité)	246
4. Quelques remarques et résultats numériques	257
x x	
x	
CONCLUSION	269
BIBLIOGRAPHIE	271

x x
x

VU :
Le Président de la Thèse,



Chr. COATMELEC

VU :
Le Directeur de Thèse,



J.C. NEDELEC

VU pour autorisation de soutenance N° 79/S/4 n°7
RENNES, le 5 MARS 1979
Le Président de l'Université de RENNES,



R. DABARD