



**HAL**  
open science

## Sur un schema pour la resolution numerique des problemes hyperboliques non lineaires

M. Fortin

► **To cite this version:**

M. Fortin. Sur un schema pour la resolution numerique des problemes hyperboliques non lineaires.  
[Rapport de recherche] RR-0212, INRIA. 1983. inria-00076346

**HAL Id: inria-00076346**

**<https://inria.hal.science/inria-00076346>**

Submitted on 24 May 2006

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

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

# IRIA

CENTRE  
SOPHIA ANTIPOLIS

Institut National  
de Recherche  
en Informatique  
et en Automatique

Domaine de Voluceau  
Rocquencourt  
BP 105  
78153 Le Chesnay Cedex  
France  
Tél. 954 90 20

## Rapports de Recherche

N° 212

### **SUR UN SCHEMA POUR LA RESOLUTION NUMERIQUE DES PROBLEMES HYPERBOLIQUES NON-LINEAIRES**

**Michel FORTIN**

**Mai 1983**

SUR UN SCHEMA POUR LA RESOLUTION NUMERIQUE DES PROBLEMES  
HYPERBOLIQUES NON-LINEAIRES

---

Michel FORTIN

Résumé : Au moyen d'une formulation variationnelle faible, on introduit un schéma pour l'approximation de problèmes hyperboliques non-linéaires du type  $u_t + f_x(u) = 0$ . Ce schéma est décrit pour une partition du domaine en quadrilatères ou en triangles. On introduit un procédé de décentrage et on montre, sur un exemple, comment choisir au mieux le paramètre dont ce procédé dépend.

Abstract : By using a weak variational formulation, a finite element scheme is introduced to approximate hyperbolic non-linear equations of type  $u_t + f_x(u) = 0$ . The scheme is described for both quadrilateral and triangular partitions of the domain. An upwinding procedure is also introduced and we show, on an example, how to make an optimal choice of the parameter on which this procedure depends.



## SUR UN SCHEMA POUR LA RESOLUTION DES PROBLEMES HYPERBOLIQUES NON-LINEAIRES

Michel FORTIN

## §0. Introduction

Nous présentons dans ce travail quelques idées sur l'approximation par éléments finis de problèmes hyperboliques non-linéaires, le but ultime étant l'approximation des équations d'Euler des fluides compressibles. Les méthodes que nous proposons font appel à des formulations faibles de ces problèmes; nous ferons d'abord une présentation très élémentaire en rappelant les concepts de base. Nous considérerons donc d'abord des cas très simples: problèmes dont d'équations différentielles ordinaires puis problèmes linéaires bi-dimensionnels. Ces exemples nous serviront de point de départ pour l'attaque de problèmes plus complexes, pour lesquels on doit éventuellement faire intervenir une "condition d'entropie". Les résultats que nous présentons ici sont très partiels et ne constituent qu'une première approche. Nous espérons qu'ils pourront servir de point de départ à d'autres travaux et que l'application numérique en montrera l'utilité.

## §.1 Solutions faibles d'équations différentielles ordinaires; approximation

1.1 Solutions faibles

La notion de solution faible d'une équation aux dérivées partielles et l'approximation numérique d'une telle solution seront au coeur de notre développement. Afin de rendre notre démarche aussi claire que possible, nous allons d'abord considérer le cas le plus simple: soit celui d'une équation différentielle ordinaire,

$$(1.1) \quad \left\{ \begin{array}{l} \frac{du}{dt} = f(t, u), \quad 0 \leq t \leq T, \\ u(0) = u_0. \end{array} \right.$$

Pour une étude détaillée de ce type de problème, par les méthodes décrites ici, nous renvoyons le lecteur au travail de F.Dubeau[2], où l'on trouvera explicitées les relations avec des techniques plus classiques.

Appliquons donc à l'équation (1.1) la méthode des résidus pondérés. Nous la multiplions par une fonction test  $v$  pour obtenir ce que nous appellerons

le formulation forte du problème, soit :

$$(1.2) \quad \int_0^T \frac{du}{dt} v dt = \int_0^T f(t, u(t)) v dt, \quad \forall v.$$

Cette relation doit être complétée par la condition initiale  $u(0)=u_0$ . Pour qu'elle soit bien posée, il faut chercher  $u(t)$  dans un espace de fonctions suffisamment régulières pour que la dérivée  $\frac{du}{dt}$  ait un sens et puisse être mise en dualité avec une fonction test  $v$ . Nous ne chercherons pas ici à préciser le problème davantage, en particulier par le choix d'espaces fonctionnels appropriés. Seule la démarche formelle et sa contrepartie numérique nous intéresseront.

Intégrant par parties le membre de gauche de l'équation (1.2), nous obtenons :

$$(1.3) \quad -\int_0^T u \frac{dv}{dt} dt + u(T) v(T) - u(0) v(0) = \int_0^T f(t, u) v dt$$

Prenant  $v(T)=0$  et utilisant  $u(0)=u_0$ , on obtient

$$(1.4) \quad -\int_0^T u \frac{dv}{dt} dt = \int_0^T f(t, u) v dt + u_0 v(0).$$

Une solution de (1.4) n'a pas à être dérivable au sens usuel. Elle vérifiera cependant l'équation (1.1) au sens des distributions. Cet affaiblissement de la solution est cependant compensé par une plus grande exigence de régularité de la fonction test qui doit maintenant être dérivable.

## 1.2 Approximation

Nous allons maintenant voir comment ces deux formulations peuvent se traduire dans l'approximation numérique. Considérons une partition de  $[0, T]$  en  $N$  intervalles de longueur  $h$  ainsi que les espaces  $L_h$  des fonctions linéaires par morceaux et continues, et  $E_h$  des fonctions étagées (relativement à la partition), (figure 1.1).

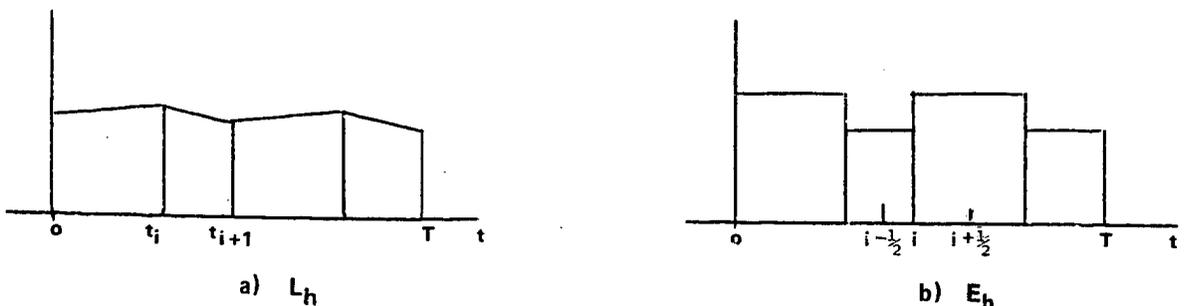


Figure 1.1 Les espaces  $L_h$  et  $E_h$

Pour l'approximation de la solution forte, nous cherchons  $u_h \in L_h$ , vérifiant la condition initiale  $u_h(0) = u_0$ . Sur un intervalle  $(t_i, t_{i+1})$ ,  $u_h(t_i)$  étant supposé connu, nous pouvons prendre une fonction test  $v \in E_h$ , constante sur  $(t_i, t_{i+1})$  et nulle ailleurs. On obtient ainsi

$$(1.5) \quad u_h(t_{i+1}) - u_h(t_i) = \int_{t_i}^{t_{i+1}} f(t, u_h(t)) dt.$$

En évaluant le membre de droite par des formules de quadrature (comme la règle du point milieu ou celle du trapèze), on obtient des schémas classiques. Avec la règle du trapèze on obtient par exemple, notant  $u_i$  pour  $u_h(t_i)$

$$(1.6) \quad u_{i+1} - u_i = \frac{h}{2} [f(t_{i+1}, u_{i+1}) + f(t_i, u_i)].$$

On peut étendre ce procédé à des polynômes de degré plus élevé (cf. Dubeau[2]).

Pour l'approximation de la solution faible, nous considérons symétriquement une fonction test dans  $L_h$  alors que nous recherchons la solution dans  $E_h$ . Prenant en particulier pour  $v_h$  la fonction valant 1 en  $t_i$  et zéro aux autres noeuds, nous trouvons :

$$(1.7) \quad u_{i+\frac{1}{2}} - u_{i-\frac{1}{2}} = \int_0^T f(t, u_h) v_h dt,$$

et en évaluant l'intégrale par la règle du point-milieu dans chacun des intervalles  $(t_{i-1}, t_i)$  et  $(t_i, t_{i+1})$

$$(1.8) \quad u_{i+\frac{1}{2}} - u_{i-\frac{1}{2}} = \frac{h}{2} [f(t_{i+\frac{1}{2}}, u_{i+\frac{1}{2}}) + f(t_{i-\frac{1}{2}}, u_{i-\frac{1}{2}})].$$

Si la formule est la même, son interprétation n'est pas la même. On peut d'ailleurs pousser un peu plus loin en utilisant directement (1.3) au niveau d'un intervalle  $(t_i, t_{i+1})$ . On est alors conduit à chercher une solution  $u_h$  ayant deux types de composantes, soit d'une part des valeurs constantes par morceaux sur chacun des intervalles et d'autre part des valeurs ponctuelles définies aux noeuds de la partition (figure 1.2).

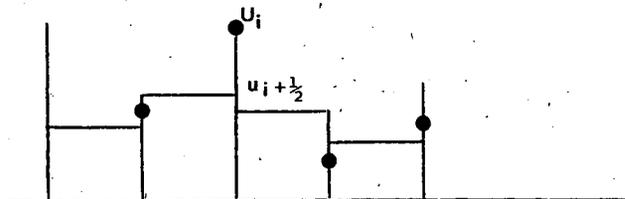


Figure 1.2 La solution approchée

Nous noterons  $u_{i+\frac{1}{2}}$  la valeur de la solution  $u_h$  dans l'intervalle  $(t_i, t_{i+1})$

et  $U_i$  la valeur ponctuelle au point  $t_i$ . Il n'y a a priori aucun lien entre les valeurs modales  $U_i$  et les valeurs  $u_{i+\frac{1}{2}}$  des intervalles. Supposons  $U_i$  connu et appliquons (1.3) dans  $(t_i, t_{i+1})$ . Posant  $v=1$  nous obtenons

$$(1.9) \quad U_{i+1} - U_i = \int_{\tau_i}^{\tau_{i+1}} f(t_i, u_h) dt,$$

d'où avec la règle du point milieu,

$$(1.10) \quad U_{i+1} - U_i = h f(t_i, u_{i+\frac{1}{2}}).$$

Prenant par ailleurs  $v=1$  en  $t_{i+1}$  et  $v=-1$  en  $t_i$ ,  $v$  étant linéaire, on obtient

$$(1.11) \quad u_{i+\frac{1}{2}} = \frac{U_{i+1} + U_i}{2} + \int_{\tau_i}^{\tau_{i+1}} f(t_i, u_h) dt.$$

La règle du point milieu donnerait simplement

$$(1.12) \quad u_{i+\frac{1}{2}} = \frac{U_{i+1} + U_i}{2},$$

ce qui conduit évidemment à un schéma classique. A noter cependant : cette formulation est essentiellement la même que la précédente, on peut recoller deux intervalles, utiliser une fonction test continue et retomber ainsi sur le schéma précédent.

Dans le cas des équations différentielles ordinaires, F.Dubeau[2] a montré que les valeurs ponctuelles  $U_i$  sont super-convergentes par rapport aux approximations constantes (ou plus généralement polynomiales) de chacun des intervalles.

Le fait le plus étonnant est sans doute ici l'apparition d'une solution numérique comportant des valeurs ponctuelles, indépendantes en quelque sorte du reste de l'approximation.

## §.2 Equation de transport linéaire

### 2.1 Position du problème

Nous allons maintenant chercher à appliquer le même type de méthode à un problème de transport linéaire. Considérons dans  $\mathbb{R}^2$  un champ de vecteur  $a(x)$  à divergence nulle (nous verrons plus loin comment on peut relaxer cette hypothèse). Soit  $u(x)$  la concentration d'une substance ou d'une quantité quelconque transportée par le champ  $u$  selon l'équation,

$$(2.1) \quad \operatorname{div}(a u) + a_0 u = 0,$$

le terme  $a_0 u$  représentant une absorption.

L'équation la plus simple est bien sûr

$$(2.2) \quad \frac{\partial u}{\partial x_1} + c \frac{\partial u}{\partial x_2} = 0,$$

dont la solution est constante le long des caractéristiques  $x_2 - c_1 x_1 = 0$ .

Il est bien connu que pour spécifier la solution d'une équation telle que (2.1) dans un domaine  $\Omega$ , il faut imposer des conditions aux limites sur la partie entrante  $\Gamma_-$  (i.e. où  $a \cdot n < 0$ ) de la frontière  $\Gamma$  de  $\Omega$ .

Ici encore, nous pouvons appliquer la méthode des résidus pondérés en multipliant (2.1) par une fonction test  $v$  et en indiquant sur  $\Omega$ ,

$$(2.3) \quad \int_{\Omega} (\text{div}(a u) + a_0 u) v \, dx = 0, \quad \forall v.$$

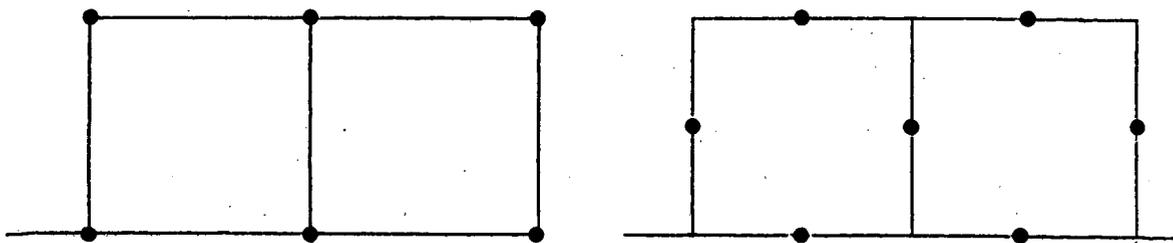
On intègre (2.3) par parties pour obtenir une solution faible

$$(2.4) \quad -\int_{\Omega} u \, \text{grad} v \, dx + \int_{\Omega} a_0 u v \, dx + \int_{\Gamma} a \cdot n u v \, d\gamma = 0, \quad \forall v.$$

Dans la formulation (2.3) la solution  $u$  doit être assez régulière pour que  $\text{div}(a u)$  ait un sens alors que dans la formulation (2.4) la régularité doit porter sur la fonction test. Les méthodes numériques qu'on en déduira seront par conséquent différentes.

## 2.2 L'approximation des solutions "fortes"

L'approximation des solutions "fortes" a été considérée par Lesaint et Raviart[5][4] qui ont généralisé, en particulier, quelques résultats de Keller[3]; on trouvera dans leur article une analyse détaillée de l'erreur. Pour notre part nous nous contenterons ici de présenter le principe de leur méthode. Considérons pour simplifier, un maillage rectangulaire. Nous utiliserons pour l'approximation de  $u$  des éléments  $Q_1$  conformes ou non-conformes selon que le raccord d'un élément à l'autre sera complet ou limité au milieu des côtés, (figure 2.1).



a) éléments conformes

b) éléments non conformes

Figure 2.1

Il est bien connu que les éléments non-conformes de degré 1 ne sont pas déterminés par les valeurs aux milieux des côtés, la fonction  $\xi\eta$  s'annulant en ces points sur l'élément de référence  $[-1,1] \times [-1,1]$ . Cette difficulté peut être levée en utilisant, dans le cas qui nous intéresse des fonctions linéaires ( $P_1$ ) par morceaux et non bilinéaires ( $Q_1$ ). Pour le cas général on pourra consulter Lesaint et Raviart[5].

Considérons d'abord une approximation conforme. Pour fixer les idées, nous allons expliciter les calculs dans le cas de l'équation (2.2) avec  $c=1$ . Considérons donc un élément dans lequel nous supposons  $u$  donné sur la partie entrante de la frontière. Nous supposons pour simplifier l'élément orienté selon les axes.

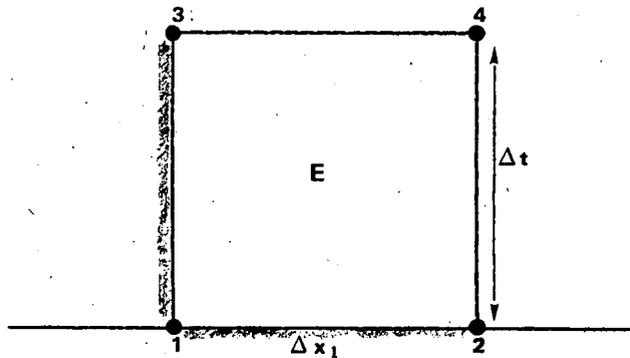


Figure 2.2

Sur la figure 2.2, nous avons hachuré la partie entrante de  $\partial K$ . Il nous faut donc supposer connues les valeurs de  $u$  aux noeuds 1,2,3. Pour déterminer la quatrième valeur, nous utilisons dans (2.3) une fonction test constante d'où pour le cas particulier considéré

$$(2.5) \quad \int_E \left( \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} \right) dx = 0,$$

ce qui s'écrit, posant  $K = \frac{\Delta t}{\Delta x}$

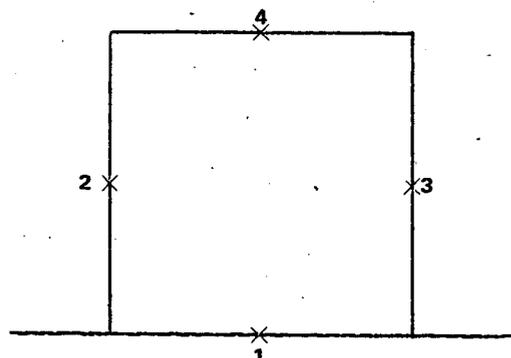
$$(2.6) \quad u_4 = u_1 + (u_2 - u_3) \left( \frac{1-K}{1+K} \right).$$

On remarque que ce schéma peut être utilisé à condition de balayer le maillage dans un bon ordre.

Dans le cas non conforme :

on utilise une fonction linéaire dont la valeur est donnée aux deux points numérotés 1 et 2 sur la figure 2.3.

Figure 2.3



Les valeurs aux noeuds 2 et 3 sont donc liées par la relation :

$$(2.7) \quad u_4 = u_3 + u_2 - u_1 .$$

Utilisant toujours une fonction test constante, on trouve le schéma,

$$(2.8) \quad (u_4 - u_1) + K(u_3 - u_2) = 0 ,$$

d'où l'on tire, moyennant (2.7)

$$(2.9) \quad \begin{cases} u_3 = \frac{2u_1 - (1-K)u_2}{1+K} \\ u_4 = \frac{2K u_2 + (1-K)u_1}{1+K} \end{cases}$$

Le schéma ainsi construit, qui est en partie orienté selon les caractéristiques donne en général de meilleurs résultats que le précédent. L'extension au cas de l'équation (1.1) est un exercice immédiat. On peut par ailleurs considérer comme dans Lesaint et Raviart[5] des approximations d'ordre plus élevé. Notre but est autre : nous allons maintenant voir comment on peut transposer les schémas précédents pour approcher la formulation faible du problème.

### 2.3 Approximation des solutions faibles

Nous allons encore ici nous contenter de considérer l'équation de transport la plus simple. La formulation faible du problème s'écrit, posant  $a=\{1,1\}$ ,

$$(2.10) \quad - \int_{\Omega} \left( \frac{\partial v}{\partial t} + \frac{\partial v}{\partial x} \right) dx + \int_{\Gamma} a \cdot n u v d_{\gamma}$$

Conformément à la technique qui nous avons introduite pour les équations différentielles, nous allons maintenant rechercher une solution constante par élément, en utilisant pour fonctions tests, des polynômes  $Q_1$ , conformes ou non-conformes.

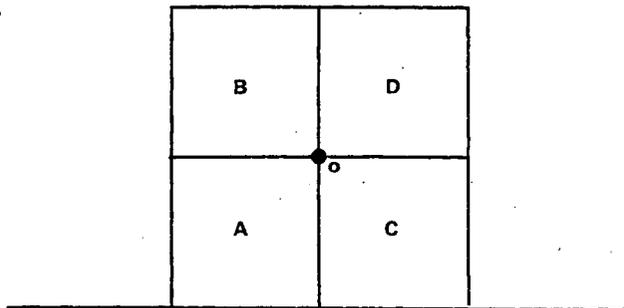


Figure 2.4

Nous considérons toujours le même maillage orienté selon les axes. La figure 2.4 présente un assemblage de quatre éléments. Nous pouvons supposer connues les valeurs de  $u$  dans les éléments A,B,C, seule la valeur D étant à déterminer. Prenant pour fonction test  $v$  la fonction dont la valeur est 1 en o,

nulle hors de  $A \cup B \cup C \cup D$ , bilinéaire et continue, on trouve

$$(2.11) \quad u_D = u_A + (u_C - u_B) \left( \frac{1-K}{1+K} \right),$$

c'est-à-dire, le même schéma qu'en (2.4), à une translation près.

On peut transposer la formulation non conforme en prenant pour fonction test  $v$  celle dont les courbes de niveau sont présentées à la figure 2.5,

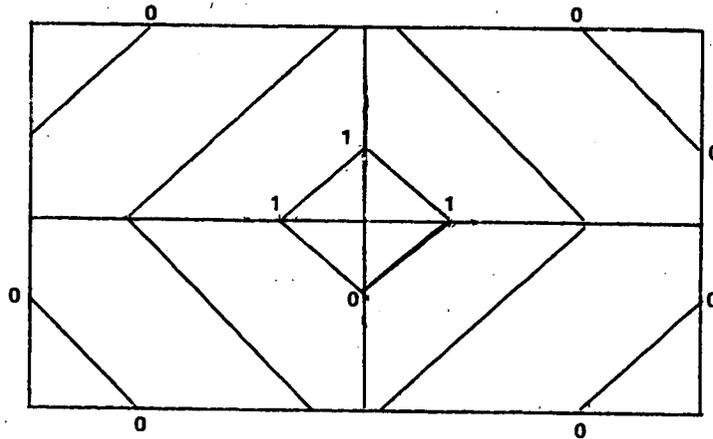


Figure 2.5

et qui prend la valeur 1 aux milieux des côtés  $AB, AC, BD, DC$ . On trouve de nouveau comme schéma l'expression (2.11). A noter que les schémas seraient différents pour un second membre non nul ou des coefficients variables. Pour obtenir un résultat réellement intéressant, nous allons maintenant restreindre (2.10) à un seul élément en introduisant une solution formée d'une *valeur constante* sur l'élément et de *traces au bord*. Nous présentons uniquement le cas correspondant à l'approximation non-conforme, celle correspondant au cas conforme étant plus difficile à généraliser.

Nous cherchons donc une valeur  $u$  sur  $K$  et quatre valeurs  $U_N, U_S, U_E, U_O$  (numérotées "géographiquement" nord, sud, est, ouest), (figure 2.6).

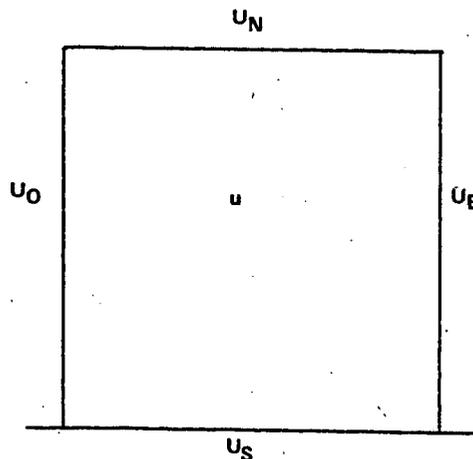


Figure 2.6

Ces quatre dernières valeurs sont supposées constantes sur les côtés qui les portent. Les fonctions tests admissibles sont les fonctions linéaires sur  $K$ . Les valeurs  $U_S$  et  $U_O$ , situées sur des côtés entrants sont supposées connues. En prenant pour fonction test une constante, on trouve

$$(2.12) \quad U_N - U_S + K(U_E - U_O) = 0.$$

Par ailleurs prenant  $v = x$  et  $v = y$  (dans  $[-1, +1]^2$ ), on obtient

$$(2.13) \quad \begin{cases} u = \frac{U_E + U_O}{2}, \\ u = \frac{U_N + U_S}{2}. \end{cases}$$

Si on compare ce résultat avec (2.7)-(2.8), on s'aperçoit qu'on a trouvé le même schéma (à son interprétation près). On a donc

$$(2.14) \quad \begin{cases} U_E = \frac{2U_S - (1-K)U_O}{(1+K)}, \\ U_N = \frac{2KU_O + (1-K)U_S}{(1+K)}, \\ u = \frac{U_S + KU_O}{(1+K)}. \end{cases}$$

On peut évidemment généraliser en utilisant des polynômes d'ordre plus élevé.

#### 2.4 Schémas décentrés

On peut introduire dans les schémas précédents, en particulier dans (2.12)-(2.13) un paramètre de décentrage. Prenons en effet pour fonction test au lieu de  $v = x$  la fonction :

$$(2.15) \quad v = x + 3\left(\alpha - \frac{1}{2}\right)\left(x^2 - \frac{1}{3}\right)$$

On peut même introduire un décentrage nord-sud dépendant d'un paramètre  $\beta$ . On trouve alors au lieu de (2.13)

$$(2.16) \quad \begin{cases} u = \alpha U_O + (1-\alpha) U_E, \\ u = \beta U_N + (1-\beta) U_S. \end{cases}$$

On en tire,

$$(2.17) \quad \begin{cases} U_N = U_S + \frac{K}{(\alpha + K\beta)}(U_O - U_S), \\ U_E = U_O + \frac{1}{(\alpha + K\beta)}(U_S - U_O), \\ u = \frac{U_S + K\beta/\alpha U_O}{1 + K\beta/\alpha}. \end{cases}$$

Un cas particulier intéressant a priori, est celui où  $\beta = \alpha$  pour lequel la valeur de  $u$  est indépendante du décentrage. On a alors

$$(2.18) \quad \left\{ \begin{array}{l} U_N = U_S + \frac{K}{\alpha(1+K)}(U_0 - U_S), \\ U_E = U_0 + \frac{1}{\alpha(1+K)}(U_S - U_0), \\ u = \frac{U_S + K U_0}{1+K}. \end{array} \right.$$

Un des problèmes fondamentaux dans l'usage des schémas décentrés est de choisir la valeur des paramètres. Un décentrage trop violent introduit en effet une dissipation importante dans le schéma. Nous allons maintenant considérer un exemple montrant qu'un choix judicieux d'une suite de paramètres de décentrage permet d'obtenir un schéma très efficace.

Nous considérons la propagation par l'équation  $\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x}$  d'un front valant au temps  $t=0$ ,

$$(2.19) \quad \begin{array}{l} u(x,0) = 1, \quad x < 0, \\ u(x,0) = 0, \quad x > 0. \end{array}$$

On sait que pour  $K=1$  et  $\alpha = \frac{1}{2}$ , on obtient la solution exacte, c'est-à-dire que le front avance exactement d'un élément à chaque pas de temps. Nous allons montrer qu'on peut obtenir un schéma exact pour toute valeur de  $K < 1$ , à condition de choisir convenablement à chaque pas le paramètre de décentrage. Pour ce faire, considérons l'élément situé au temps  $t$  à la droite du front. On y a  $U_0=1$  et  $U_S=0$ , d'où en appliquant les formules (2.18)

$$(2.20) \quad \left\{ \begin{array}{l} U_N = \frac{K}{\alpha(1+K)}, \\ U_E = 1 - \frac{K}{\alpha(1+K)}. \end{array} \right.$$

La valeur de  $U_E$  devient négative, ce qui est ici inadmissible, dès que  $\alpha < \frac{1}{1+K}$ . En choisissant donc la plus petite valeur correcte de  $\alpha$ , soit  $\frac{1}{1+K}$ , on trouve au premier pas,  $U_E=0$  et  $U_N=K$ , ce qui est intuitivement correct pour  $K < 1$ .

Considérons le pas suivant; on a maintenant pour le même élément  $U_S=K$  et  $U_0=1$ . On doit maintenant distinguer deux cas selon que  $K < \frac{1}{2}$  ou  $K \geq \frac{1}{2}$ . Pour  $K < \frac{1}{2}$  on vérifie facilement que  $\alpha = \frac{1-K}{1+K}$  fournit  $U_N=2K$  et  $U_0=0$  ce qui est encore correct. Pour  $K \geq \frac{1}{2}$  on a pour  $\alpha = \frac{K}{1+K}$ ,  $U_N=1$  et  $U_E=2 - \frac{1}{K}$ . On doit donc maintenant calculer la valeur dans l'élément de droite dans lequel nous avons  $U_S=0$  et  $U_0=2 - \frac{1}{K}$ . Pour  $\alpha = \frac{1}{1+K}$ , on y trouve  $U_N=2K-1$  et  $U_E=0$  ce qui

est de nouveau correct.

Pour fixer les idées, nous donnons ci-dessous les suites optimales de valeurs de  $\alpha$  pour quelques valeurs de  $K$ .

$$\begin{aligned} K = \frac{1}{2}, & \quad \alpha = \frac{2}{3}, \frac{1}{3} \\ K = \frac{1}{3}, & \quad \alpha = \frac{3}{4}, \frac{1}{2}, \frac{1}{4} \\ K = \frac{2}{3}, & \quad \alpha = \frac{3}{5}, \frac{2}{5} \\ K = \frac{2}{5}, & \quad \alpha = \frac{5}{7}, \frac{3}{7}, \frac{2}{7} \text{ dans le premier élément,} \\ & \quad \alpha = \frac{5}{7}, \frac{4}{7}, \frac{2}{7} \text{ dans l'élément suivant.} \end{aligned}$$

On remarque donc une alternance de valeurs de  $\alpha$  supérieures à  $\frac{1}{2}$ , pour lesquelles le schéma est dissipatif, avec des valeurs inférieures à  $\frac{1}{2}$  pour lesquelles le schéma est anti-dissipatif. On peut évidemment objecter que nous avons utilisé ici notre connaissance de la solution exacte pour construire ces valeurs. On peut cependant remarquer qu'on peut les obtenir en utilisant un critère plus objectif. Dans tous les cas, *la valeur optimale du décentrage est la valeur minimum de  $\alpha$  conservant la monotonie du schéma.* Pour juger de la monotonie nous demandons que la valeur  $U$  à l'interface de deux éléments A et B soit comprise dans l'intervalle  $[u_B, u_A]$  où  $u_B$  est la valeur constante de la solution dans l'élément A. En appliquant ce critère on obtient pour  $K = \frac{m}{n}$  une propagation exacte de  $m$  éléments de toute solution  $u_0(x)$  après  $n$  pas de temps.

L'alternance de la diffusion et de l'anti-diffusion rejoint, dans un cadre différent les idées de Boris, Book et Hain[1], reprises entre autres dans la thèse de Leroux[4]. La mise en oeuvre effective demandera une étude spéciale. En effet le schéma devient non-linéaire et il faudra envisager une itération appropriée pour trouver les valeurs optimales des  $\alpha$ .

## 2.5 Variantes

Ces résultats s'étendent sans difficulté au cas d'une équation de la forme:

$$(2.21) \quad \frac{\partial u}{\partial t} + a_1 \frac{\partial u}{\partial x} + a_2 \frac{\partial u}{\partial y} .$$

Considérons un élément rectangulaire orienté selon les axes

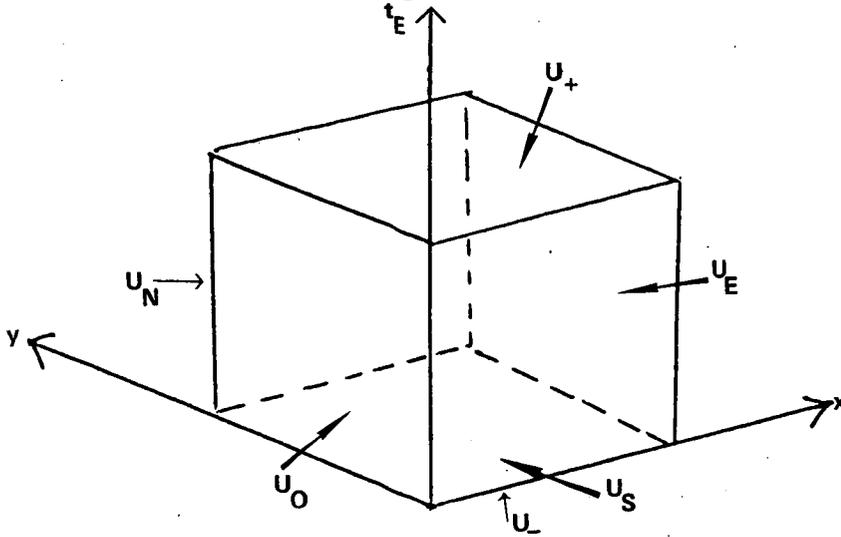


Figure 2.7

Notant  $u$  la valeur de la solution dans l'élément et  $U_+, U_-, U_E, U_O, U_N, U_S$  les valeurs sur les faces comme à la figure 2.7. On trouve facilement les relations

$$(2.2) \quad \begin{cases} U_+ - U_- + K_1(U_E - U_O) + K_2(U_N - U_S), \\ u = \alpha U_+ + (1 - \alpha)U_-, \\ u = \beta U_N + (1 - \beta)U_S, \\ u = \gamma U_E + (1 - \gamma)U_O, \end{cases}$$

où  $K_1 = a_1 \frac{\Delta t}{\Delta x}$  et  $K_2 = a_2 \frac{\Delta t}{\Delta y}$ . On tire de (2.22) la relation

$$(2.23) \quad u = \frac{U_- + K_1 \frac{\alpha}{\gamma} U_O + K_2 \frac{\alpha}{\beta} U_S}{1 + K_1 \frac{\alpha}{\gamma} + K_2 \frac{\alpha}{\beta}}$$

En prenant  $\alpha = \beta = \gamma$ , la valeur de  $u$  est indépendante du décentrage. Le choix des  $\alpha$  optimaux est évidemment plus complexe.

On peut aussi considérer des éléments plus généraux. Nous allons expliciter un schéma applicable à l'équation (2.21) lorsque les éléments sont des prismes triangulaires. Pour simplifier l'exposé, nous allons considérer le cas d'un problème bi-dimensionnel

$$(2.24) \quad a_1 \frac{\partial u}{\partial x_1} + a_2 \frac{\partial u}{\partial x_2} = 0.$$

La formulation faible s'écrit :

$$(2.25) \quad - \int_{\Omega} u (a_1 \frac{\partial v}{\partial x_1} + a_2 \frac{\partial v}{\partial x_2}) dx + \int_{\partial \Omega} a \cdot n u v d\gamma.$$

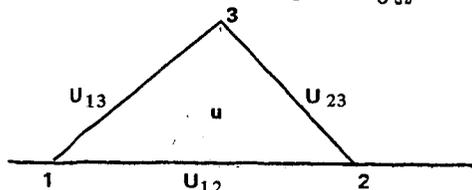


Figure 2.8

Dans un élément triangulaire, les inconnues seront une valeur constante  $u$  dans  $K$  et les valeurs au bord  $U_{12}, U_{23}, U_{13}$  comme à la figure 2.8. Pour poser les équations on détermine d'abord les côtés entrants ( $a \cdot n < 0$ ) et sortants ( $a \cdot n > 0$ ) de l'élément. Les côtés pour lesquels  $a \cdot n = 0$  disparaissent de (2.25) et la valeur de  $U$  y restera indéterminée. Les valeurs sont supposées connues sur les côtés entrants; elles sont à calculer sur les côtés sortants. Si le champ  $a$  est à divergence nulle, au moins un des côtés est entrant et au moins sortant.

En général un triangle quelconque aura, soit un, soit deux côtés entrants.

Dans le premier cas nous avons 3 inconnues à déterminer et nous pouvons utiliser comme fonction test une fonction linéaire générale.

Dans le second cas nous n'avons plus que 2 inconnues à déterminer. Nous suggérons d'utiliser comme fonction test une constante et la coordonnée barycentrique  $\lambda_i$  du sommet opposé à un côté sortant.

On peut décentrer cette dernière fonction test en lui ajoutant un terme de la forme  $\alpha(\lambda_i^2 - \lambda_i + \frac{1}{6})$  dont l'intégrale est nulle sur les côtés adjacents au sommet  $i$ . Cette fonction s'annule en effet aux points de Gauss-Legendre des côtés. Il serait aussi possible d'utiliser les points de Gauss-Radan.

Quant à l'introduction du temps, elle se fait sans problèmes comme dans le cas de l'exemple précédent. L'analyse détaillée et l'expérimentation d'un tel schéma reste à faire.

### §.3 Problèmes non linéaires

#### 3.1 Le schéma de base

Nous allons maintenant chercher à étendre nos résultats au cas d'une équation non-linéaire de la forme classique

$$(3.1) \quad \frac{\partial u}{\partial t} + \operatorname{div}(f(u)) = 0$$

où  $f$  est une fonction de  $\mathbb{R}$  dans  $\mathbb{R}^n$ . Nous considérons uniquement l'approximation des solutions faibles, voie qui nous paraît la plus prometteuse.

Nous supposons ici que les composantes de  $f$  sont monotones (du moins pour  $u > 0$ ), de façon à simplifier l'application de la condition d'entropie.

Nous introduisons, pour définir une solution faible un domaine  $\Omega \times ]0, T[ = \Sigma$ .

Nous multiplions (3.1) par une fonction test "appropriée" et nous intégrons par parties. On trouve, notant  $B = \partial \Omega \times ]0, T[$

$$(3.2) \quad -\int_{\Sigma} u \frac{\partial v}{\partial t} dx dt - \int_{\Sigma} f(u) \cdot \text{grad } v dx dt + \int_{\Omega} u(T) v(T) dx - \int_{\Omega} u(0) v(0) dx + \int_B f(u) \cdot n v dT.$$

Pour obtenir une approximation de ce problème, nous considérons une partition de  $\Omega$  en éléments rectangulaires. Sur chaque élément, on recherchera comme dans le cas linéaire une valeur constante  $u$ , dans l'élément, une valeur finale  $U_+$ , la valeur initiale  $U_-$  étant supposée connue et les valeurs des flux normaux  $f \cdot n$  sur  $B$ , (figure 3.1).

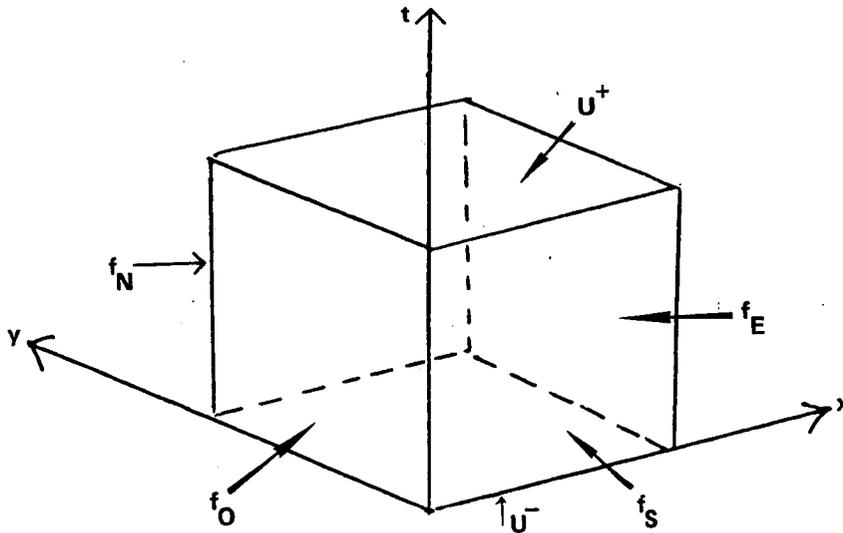


Figure 3.1

Nous noterons  $f_E, f_O, f_N, f_S$  les valeurs des flux normaux sur les faces. On respecte donc ainsi, en un sens, la condition de Rankine-Hugoniot qui demande la continuité des flux aux interfaces. Pour respecter la condition d'entropie nous raisonnerons sur chaque élément en supposant connues les valeurs sur les faces éclairées  $(f(u) \cdot n) > 0$  du prisme  $K$ . Pour éviter des calculs trop lourds, nous nous restreindrons par la suite au cas d'une seule dimension d'espace en considérant l'équation :

$$(3.3) \quad \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0.$$

Nous n'avons plus qu'à déterminer la valeur de  $u$ , les flux gauche et droit, la valeur initiale et la valeur finale, (figure 3.2).

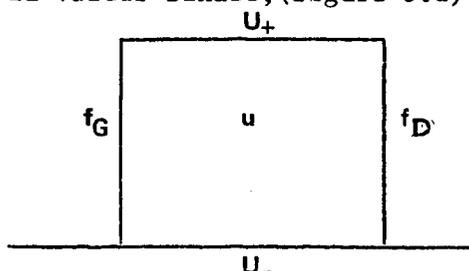


Figure 3.2

Nous supposons comme la valeur initiale  $U_-$  et nous supposons pour fixer les idées que le côté gauche est entrant de sorte qu'on suppose connue la valeur  $f_G$ . (Cette valeur est cependant une inconnue dans l'élément adjacent). En utilisant le même procédé que dans le cas linéaire, on trouve facilement

$$(3.4) \quad U_+ - U_- + K(f_D - f_G) = 0,$$

$$(3.5) \quad u = \alpha U_+ + (1-\alpha)U_-,$$

$$(3.6) \quad f(u) = \beta f_D + (1-\beta)f_G.$$

En reportant (3.5) et (3.6) dans (3.4), il vient

$$(3.7) \quad u + K \frac{\beta}{\alpha} f(u) - (U_- + K \frac{\beta}{\alpha} f_G) = 0.$$

Dans le cas linéaire on avait trouvé une équation analogue avec  $f(u) = u$ ,  $f_G = U_G$ . L'équation définissant  $u$  est ici, comme on pouvait s'y attendre, non-linéaire. Le choix  $\frac{\beta}{\alpha}$  est encore spécialement agréable car  $u$  devient indépendant du décentrage.

La valeur  $u$  étant connue, on pourra déterminer  $U_+$  et  $f_D$  par les équations (3.5) et (3.6). Pour un schéma général, il faudra prévoir le cas où le nombre de flux entrants est différente du nombre de flux sortants et prévoir un choix approprié de fonctions tests.

Nous allons maintenant considérer, à titre d'exemple l'équation de Burgers,

$$(3.8) \quad \frac{\partial u}{\partial t} + \frac{1}{2} \frac{\partial}{\partial x}(u^2) = 0.$$

On a donc  $f(u) = \frac{u^2}{2}$ . Considérant l'équation (3.7) avec  $\beta = \alpha$ , il vient

$$(3.9) \quad u = \frac{-1 \pm \sqrt{1 + 2K(U_- + K \frac{U_G^2}{2})}}{K}.$$

Nous avons noté  $\frac{U_G^2}{2}$  la valeur du flux  $f_G$ . A noter cependant qu'on ne cherche pas à déterminer les valeurs  $U_G$  et  $U_D$  elles-mêmes, car elles peuvent devenir imaginaires si le schéma n'est pas monotone.

Considérons pour cette équation le problème déjà traité dans le cas linéaire, à savoir la propagation d'un front. Posant  $U_- = 0$  et  $\frac{U_G}{2} = \frac{1}{2}$ , il vient

$$(3.10) \quad u = \frac{-1 + \sqrt{1 + K^2}}{K}$$

d'où  $f(u) = \frac{u^2}{2} = -\frac{u}{K} + \frac{1}{2}$ .

On vérifie que pour  $\alpha = \frac{2u}{K}$ , on trouve  $f_D = 0$ ,  $U_+ = \frac{K}{2}$ , ce qui est la meilleure solution possible dans le contexte de cette approximation. Cette valeur de  $\alpha$  est la plus petite conservant la monotonie du schéma : pour  $\alpha$  plus petit on trouve des valeurs négatives de  $-\frac{U_D^2}{2}$ . L'apparition d'une valeur imaginaire de  $U_D$  n'entrave pas la marche du schéma car la seule valeur de  $f_D = \frac{U_D^2}{2}$  est utilisée. Cependant elle rend dans ce cas l'application de la condition d'entropie impossible car celle-ci dépend du signe de U, la fonction  $f$  n'étant pas monotone.

Encore une fois nous trouvons une stratégie possible : utiliser le  $\alpha$  minimum tel que le schéma reste monotone. Nous voyons que le  $\alpha$  dépend ici de la solution elle-même, et le problème de le déterminer de façon automatique reste entièrement ouvert.

### 3.2 Variantes et extensions

On étendra sans difficultés les résultats précédents à un problème en deux (et même en trois) dimensions d'espace. Le cas des éléments triangulaires ne pose pas de difficulté nouvelle par rapport au cas linéaire.

Dans le cas des équations d'Euler on aura par le procédé que nous avons introduit un schéma dont les inconnues seront dans chaque élément

- la densité  $\rho$ , la densité de quantité de mouvement  $\rho u$ , la densité d'énergie, dans l'élément,
- les valeurs initiales et finales de ces quantités,
- les flux  $\rho u \cdot n$ ,  $\rho(u \cdot n)^2 + p$ , ....., aus interfaces

Le schéma sera donc conservatif et nous croyons que ses qualités le rendront compétitif en comparaison avec d'autres, apparemment plus simples mais dans lesquels la qualité inférieure de l'approximation oblige à augmenter le nombre d'éléments.

**Bibliographie**

- [1] Book, D.L., Boris, J.P., Hain, K., *Flux connected transport II. Generalisation of the method*, J. Conf. Phys., 18, (1975), pp 248-283
- [2] Dubeau, F., Thèse, Montréal 1981
- [3] Keller, H.B., *A new difference scheme for parabolic problems*, Numerical Solution of Partial Differential Equations II, Academic Press, New York, (1971), pp 327-350
- [4] Le Roux, A.Y., *Approximation de quelques problèmes hyperboliques non-linéaires*, Thèse d'Etat, Univ. de Rennes, (1979)
- [5] Lesaint, P., Raviart, P.A., *Finite Element Colloration Methods for First Order Systems*, Maths of Conf., 33, (1979), pp 891-918
- [6] Lesaint, P., Raviart, P.A., *On a Finite Element Method for Solving the Newton Transport Equation*, Mathematical aspects of Finite Elements in Partial Differential Equations, Acad. Press., New York, (1974), pp 89-123.

Michel Fortin  
Département de Mathématiques  
I.M.S.P. - Parc Valrose  
06034 NICE CEDEX (France)

MF/da  
03.83

Imprimé en France

par

l'Institut National de Recherche en Informatique et en Automatique

