



**HAL**  
open science

# Calcul des lois de conservation scalaires par la méthode de transport-écoulement

Yann Brenier

► **To cite this version:**

Yann Brenier. Calcul des lois de conservation scalaires par la méthode de transport-écoulement. [Rapport de recherche] RR-0053, INRIA. 1981. inria-00076508

**HAL Id: inria-00076508**

**<https://inria.hal.science/inria-00076508>**

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**

Rapports de Recherche

N° 53

**CALCUL DE LOIS  
DE CONSERVATION SCALAIRES  
PAR LA MÉTHODE  
DE TRANSPORT-ÉCROULEMENT**

**Yann BRENIER**

Institut National  
de Recherche  
en Informatique  
et en Automatique

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

Février 1981

CALCUL DE LOIS DE CONSERVATION SCALAIRES

PAR LA METHODE DE TRANSPORT-ECROULEMENT

Yann BRENIER

(Février 1981)

INRIA - Domaine de Voluceau - Rocquencourt - B.P. 105 - 78150 LE CHESNAY - FRANCE

Tel. : 954 90 20

## RESUME

La méthode de transport-écroulement permet le calcul des lois de conservation scalaire :  $\frac{\partial u}{\partial t} + \text{div}(F(u)) = 0$ , en dimension 1 et 2, au moyen de schémas numériques monotones, conservatifs, inconditionnellement stables et complètement affranchis de la condition de Courant-Friedrichs-Lewy, ce qui aboutit à une réduction tant du coût des calculs que de la diffusion numérique. Diverses généralisations de la méthode sont envisagées sur le plan théorique.

## ABSTRACT

The transport and collapse method allows the computation of scalar conservation laws :  $\frac{\partial u}{\partial t} + \text{div}(F(u)) = 0$ , in 1 and 2 dimensional cases. The numerical schemes obtained are monotonic, conservative, unconditionally stable and are not ruled by the Courant-Friedrichs-Lewy condition : therefore the computation cost and the numerical diffusion decrease. Various generalizations are considered from a theoretical point of view.

2 - Description du schéma bidimensionnel	41
<i>a) représentation de <math>u_0</math> par un nombre fini de "particules"</i>	42
<i>b) détermination des <math>k_0(i)</math> par un algorithme de tri</i>	44
<i>c) coût du calcul des <math>k_0(i)</math></i>	47
<i>d) phase de transport : déplacement horizontal des         particules</i>	48
<i>e) phase d'écroutement : déplacement vertical des         particules</i>	48
3 - Coût de l'algorithme bidimensionnel	49
4 - Comparaison avec un schéma aux différences finies	51
<i>a) choix du schéma</i>	51
<i>b) résultats numériques</i>	53
<i>c) conclusion</i>	64
<u>CHAPITRE IV</u> : FORMALISATIONS ET GENERALISATIONS DE LA METHODE	65
1 - Première généralisation : le cas où $F$ dépend de $x$	66
2 - Une première formulation "fonctionnelle"	67
3 - Une seconde formulation "fonctionnelle"	68
4 - Deuxième généralisation : les équations paraboliques semi-linéaires	69
5 - Troisième formulation et troisième généralisation	70
<u>CONCLUSION</u>	72
<u>REFERENCES</u>	74

## PLAN

	<u>Pages</u>
<u>CHAPITRE I</u> : EXPOSE DE LA METHODE DE TRANSPORT-ECROULEMENT	2
1 - Solutions classiques et méthode des caractéristiques	3
2 - Notion de solution entropique	6
3 - Approximation des solutions entropiques par la méthode de transport-écroulement	7
<i>a) construction de l'opérateur <math>T(\delta t)</math></i>	7
<i>b) valeur de la méthode</i>	10
<i>c) autres propriétés de l'opérateur <math>T(\delta t)</math></i>	13
4 - Construction de schémas numériques à partir de la méthode de transport-écroulement : principe général	14
<u>CHAPITRE II</u> : CONSTRUCTION DE SCHEMAS NUMERIQUES A PARTIE DE LA METHODE DE TRANSPORT-ECROULEMENT : LE CAS UNIDIMENSIONNEL	16
1 - Discrétisation des données	17
2 - Description du schéma unidimensionnel	18
<i>a) représentation de <math>u_0</math> par un nombre fini de "particules"</i>	18
<i>b) déplacement horizontal des particules</i>	19
<i>c) déplacement vertical des particules</i>	20
3 - Traitement du problème aux limites	24
4 - Résultats numériques en dimension 1	25
<i>a) le problème de Riemann</i>	26
<i>b) formation d'un choc</i>	33
<i>c) un problème aux limites</i>	35
<u>CHAPITRE III</u> : CONSTRUCTION DE SCHEMAS NUMERIQUES A PARTIR DE LA METHODE DE TRANSPORT-ECROULEMENT : LE CAS BIDIMENSIONNEL	37
1 - Discrétisation des données	38

## INTRODUCTION

Nous exposons la méthode de transport-écroulement dans le cas simple de la loi de conservation scalaire dans  $\mathbb{R}^N$  :

$$(1) \quad \frac{\partial u}{\partial t} + \sum_{i=1}^N \frac{\partial}{\partial x_i} (F_i(u)) = 0 \quad u \in \mathbb{R}, x \in \mathbb{R}^N, t > 0 ;$$

où  $F$  est une application donnée de  $\mathbb{R}$  dans  $\mathbb{R}^N$ .

La méthode la plus classique et la plus ancienne pour résoudre ce type d'équations est la méthode des caractéristiques, qui ne permet que le calcul de solutions régulières. Or le trait marquant des équations hyperboliques non linéaires est l'apparition en temps fini de discontinuités, quelle que soit la régularité des données. La notion de solution classique est inadaptée et doit faire place à celle de solution faible entropique.

Après avoir rappelé ces faits bien connus, nous montrerons comment notre méthode permet d'approcher les solutions entropiques, en généralisant la méthode des caractéristiques par le concept d'écroulement (chapitre 1).

Ensuite, nous verrons qu'à partir de la méthode de transport-écroulement, on peut construire des schémas effectifs de résolution numérique (Chapitres 2 et 3).

Enfin, on exposera plusieurs formulations et généralisations de la méthode, en vue d'autres applications (Chapitre 4).

## CHAPITRE I

EXPOSE DE LA METHODE DE TRANSPORT-ECROULEMENT

## 1 - SOLUTIONS CLASSIQUES ET METHODE DES CARACTERISTIQUES

Il est naturel de chercher d'abord à résoudre l'équation (1) dans la classe des fonctions régulières.

Toute solution classique  $u(t;x)$  vérifie, au sens fort, l'équation :

$$\left(\frac{\partial}{\partial t} + Q(t,x).\text{grad}\right) u(t,x) = 0 \text{ où } Q(t,x) = F'(u(t,x))$$

Par conséquent, on a  $u(t,X(t)) = \text{cte}$  pour toute trajectoire  $t \rightarrow X(t)$  solution de :

$$\dot{X}(t) = Q(t,X(t)) \text{ c'est-à-dire } \dot{X}(t) = F'(u(t,X(t)))$$

De telles trajectoires sont appelées caractéristiques. Leur calcul est immédiat, puisque par définition :

$$F'(u(t,X(t))) = \text{cte} = F'(u_0(X(0))) \text{ où } u_0 \text{ désigne la donnée initiale.}$$

Ce sont donc les droites d'équation :

$$X(t) = X(0) + t F'(u_0(X(0))).$$

On peut ainsi calculer la solution classique au temps  $t$  à l'aide du principe constructif suivant :

Si le graphe de la donnée initiale passe par le point  $(x_0, w_0)$  alors le graphe de la solution au temps  $t$  passe par le point :

$$(x_0 + t F'(w_0), w_0).$$

On peut traduire géométriquement cette construction, en se plaçant dans le "plan du graphe", c'est-à-dire l'espace produit  $\mathbb{R}_x^N \times \mathbb{R}_w$ . On considère plutôt que le graphe, le "sous-graphe" de la donnée initiale  $u_0$  :

$$(2) \quad A = \{(x, w) \in \mathbb{R}^N \times \mathbb{R} ; u_0(x) \geq w\} ;$$

on le découpe en tranches horizontales :

$$A_w = \{x \in \mathbb{R}^N ; u_0(x) \geq w\} , w \text{ décrivant } \mathbb{R} ;$$

A  $w$  fixé, on fait subir à  $A_w$  la translation dans  $\mathbb{R}^N$  de vecteur  $t \cdot F'(w)$ , ce qui nous donne :  $B_w = A_w + t F'(w)$  ; en procédant ainsi pour chaque niveau  $w$ , on obtient finalement l'ensemble  $B$  dans  $\mathbb{R}^N \times \mathbb{R}$  :

$$(3) \quad B = \{(x, w) ; x \in B_w\} = \{(x, w) ; x \in A_w + t F'(w)\} \\ = \{(x, w) ; u_0(x - t F'(w)) \geq w\}$$

(Voir figure 1).

La limitation de cette méthode apparaît clairement :

il existe en général, quelle que soit la régularité des données, un temps critique au delà duquel l'ensemble  $B$  ne définit plus une fonction univoque, ses coupes verticales  $\{w \in \mathbb{R} ; (x, w) \in B\}$  n'étant plus nécessairement connexes (on peut le voir sur la figure 2).

Ce phénomène ne fait que traduire un fait bien connu : le caractère local de la méthode des caractéristiques ou, si l'on préfère, l'inexistence de solutions classiques globales pour l'équation (1).

Pourtant, la "pseudo-solution" que fournit la méthode des caractéristiques n'est pas dénuée d'intérêt, bien au contraire. On le verra plus loin (§ 3).

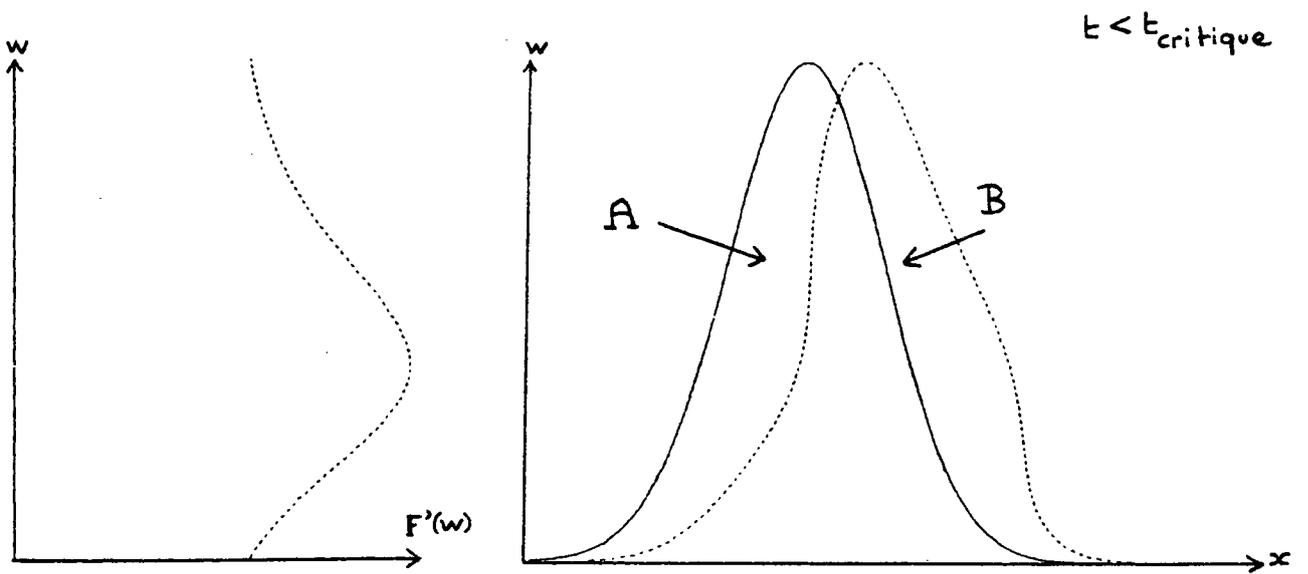


Figure 1

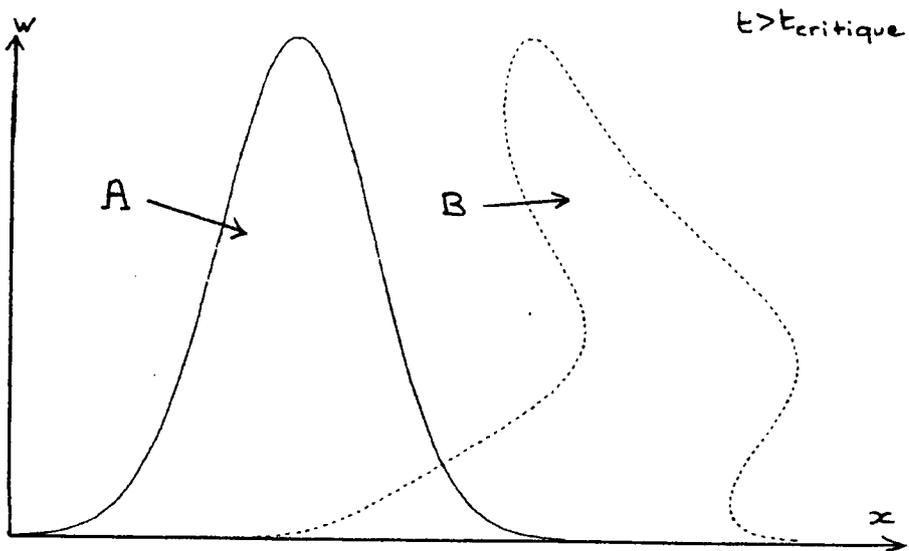


Figure 2

## 2 - NOTION DE SOLUTION ENTROPIQUE

Nous venons de voir que l'équation (1) n'admettait pas, en général, de solution classique globale, quelle que soit la régularité des données  $F$  et  $u_0$ . Il faut donc admettre la possibilité de solutions discontinues, ce qui conduit naturellement à la notion de solution faible, au sens des distributions.

Cette notion assure effectivement l'existence globale, mais pas l'unicité : elle n'est pas assez forte pour sélectionner les solutions physiquement acceptables<sup>(\*)</sup>.

C'est pourquoi a été introduite la notion de solution "entropique" (Oleïnik, 1963), pour aboutir finalement aux formulations de Kruzkov et Volpert (1970), qui assurent l'existence globale, l'unicité, la stabilité par rapport à la donnée initiale et le caractère physique des solutions (voir [7], [10], et [12]).

On peut traduire ces formulations de la manière suivante :

Une fonction  $u(t,x)$  localement intégrable est solution entropique de (1), si elle vérifie, pour chaque fonction  $h : \mathbb{R} \rightarrow \mathbb{R}$  lipschitzienne et convexe, l'inégalité au sens des distributions sur  $\mathbb{R}_+^* \times \mathbb{R}^N$  :

$$(4) \quad \frac{\partial}{\partial t} (h(u)) + \operatorname{div}(F_h(u)) \leq 0,$$

$$\text{où } F_h(w) = \int_0^w F'(s)h'(s)ds.$$

---

(\*) On donne à cette expression le sens suivant :

On désigne par solution physiquement acceptable toute solution faible de (1), telle que la solution du problème "visqueux" associé :

$$\frac{\partial u}{\partial t} + \operatorname{div}(F(u) - \varepsilon \operatorname{grad} u) = 0,$$

converge fortement vers elle - dans un sens à préciser - quand  $\varepsilon$  tend vers 0.

Notons que cette notion est plus forte que celle de solution faible (prendre successivement  $h(w) = w$  et  $h(w) = -w$ ).

### 3 - APPROXIMATION DES SOLUTIONS ENTROPIQUES PAR LA METHODE DE TRANSPORT-ECROULEMENT

Lorsque la solution "entropique" n'est plus régulière et présente des discontinuités, la méthode des caractéristiques n'est plus valable et on rejette habituellement la "pseudo-solution", multivoque, qu'elle fournit<sup>(\*)</sup>.

Notre méthode consiste au contraire à utiliser la pseudo-solution, en la transformant par l'opération dite d'éroulement, pour la rendre univoque.

#### a) Construction de l'opérateur $T(t)$

Reprenons la construction du § 1.

A partir du sous-graphe  $A = \{(x,w) ; u_0(x) \geq w\}$ , on a construit, par la méthode des caractéristiques, l'ensemble :

$$B = \{(x,w) ; u_0(x-t F'(w)) \geq w\}$$

$B$ , on l'a vu, ne définit en général pas une fonction univoque.

La multivocité se traduit par le fait que les coupes verticales :

$$B^X = \{w \in \mathbb{R} ; (x,w) \in B\}$$

---

(\*) les ingénieurs l'utilisent néanmoins pour obtenir dans certains cas la solution exacte à l'aide d'un argument de conservativité (les aires délimitées par la pseudo-solution et par la solution exacte étant égales) - voir [9], p. 73 -.

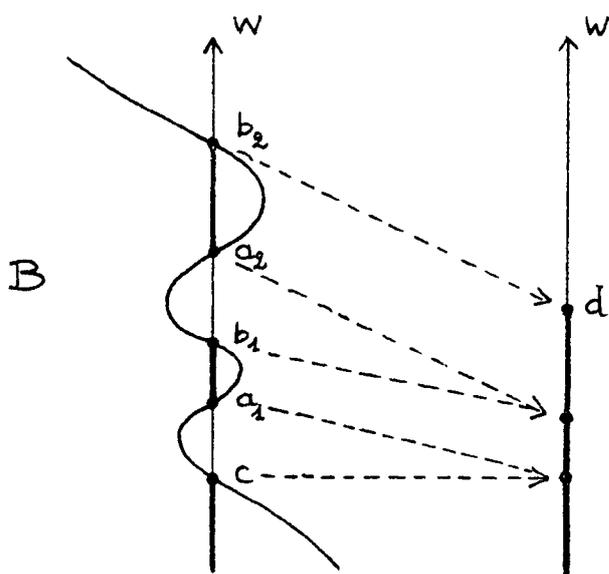
ne sont pas toutes connexes - elles peuvent être de la forme :

$$(5) \quad ]-\infty, c] \cup [a_1, b_1] \cup \dots \cup [a_m, b_m] \text{ avec } -\infty < c < a_1 < b_1 < \dots < a_m < b_m$$

L'opération d'écroutement consiste à remplacer cette coupe par un ensemble connexe "de mesure équivalente". Précisément, on pose :

$$(6) \quad C^x = ]-\infty, d] \text{ où } d = c + (b_1 - a_1) + \dots + (b_m - a_m)$$

De façon imagée, on peut dire qu'on laisse tomber les intervalles  $[a_k, b_k]$  en les empilant les uns sur les autres.



En procédant ainsi pour chaque  $x$ , on obtient un nouvel ensemble :

$$(7) \quad C = \{(x, w) \in \mathbb{R}^N \times \mathbb{R} ; w \in C^x\},$$

qui, par construction, définit une fonction univoque. On notera celle-ci  $T(t)u_0$ .

On peut dire, intuitivement, qu'on a laissé s'effondrer les surplombs de B pour obtenir C : c'est pourquoi nous parlons d'écroulement (voir figure 3).

REMARQUE : l'opération d'écroulement est très proche -à la notion de symétrie près- de la "symétrisation de Steiner", introduite au 19<sup>ème</sup> siècle pour le problème isopérimétrique (cf.[11]) ; les deux opérations ont des propriétés analogues.

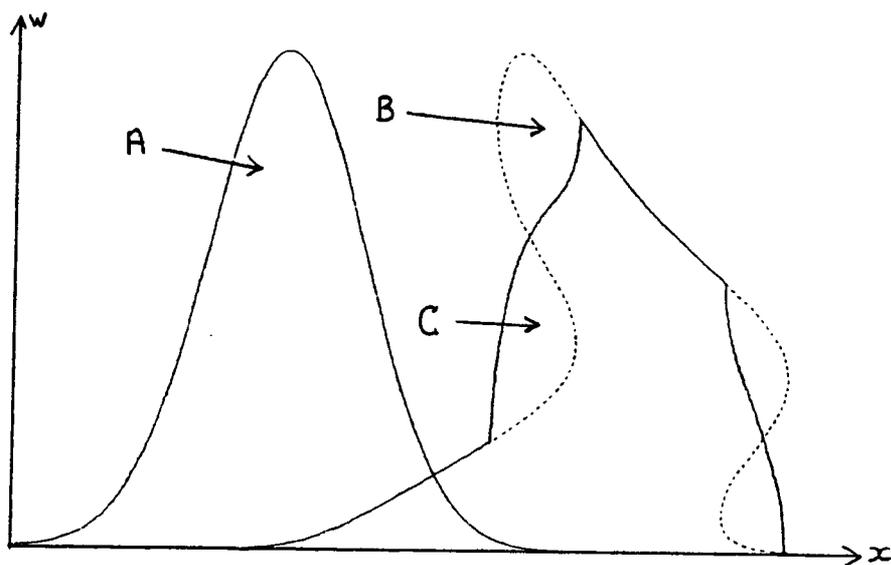


Figure 3

b) Valeur de la méthode

Notre but est de calculer les solutions entropiques. Nous allons donc comparer  $T(t)u_0$  à la solution entropique de (1) au temps  $t$ . Prenons un exemple unidimensionnel simple où l'on connaît la solution entropique<sup>(\*)</sup>.

Sur la figure 4 sont représentées :

- i) la donnée initiale ;
- ii) la pseudo-solution donnée par la méthode des caractéristiques ;
- iii) "l'écroulée" de celle-ci, c'est-à-dire  $T(t)u_0$  ;
- iv) la solution entropique.

On voit immédiatement que  $T(t)u_0$  n'est pas la solution entropique. Néanmoins, si nous découpons  $t$  en 2, on voit, sur la figure 5, que  $T(\frac{t}{2}) T(\frac{t}{2})u_0$  se rapproche de la solution entropique.

En découpant  $t$  en 6 (figure 6), puis en 30 (figure 7) et en 60 (figure 8), on converge rapidement vers la solution entropique.

Ces considérations laissent présumer que la solution entropique s'écrit :

$$(8) \quad u(t,.) = \lim_{n \rightarrow +\infty} T\left(\frac{t}{n}\right)^n \cdot u_0$$

Ce résultat est effectivement vrai.

On tire en effet de [3], le théorème suivant :

---

(\*) Pour le calcul de solutions entropiques unidimensionnelles, à la main, voir par exemple [8].

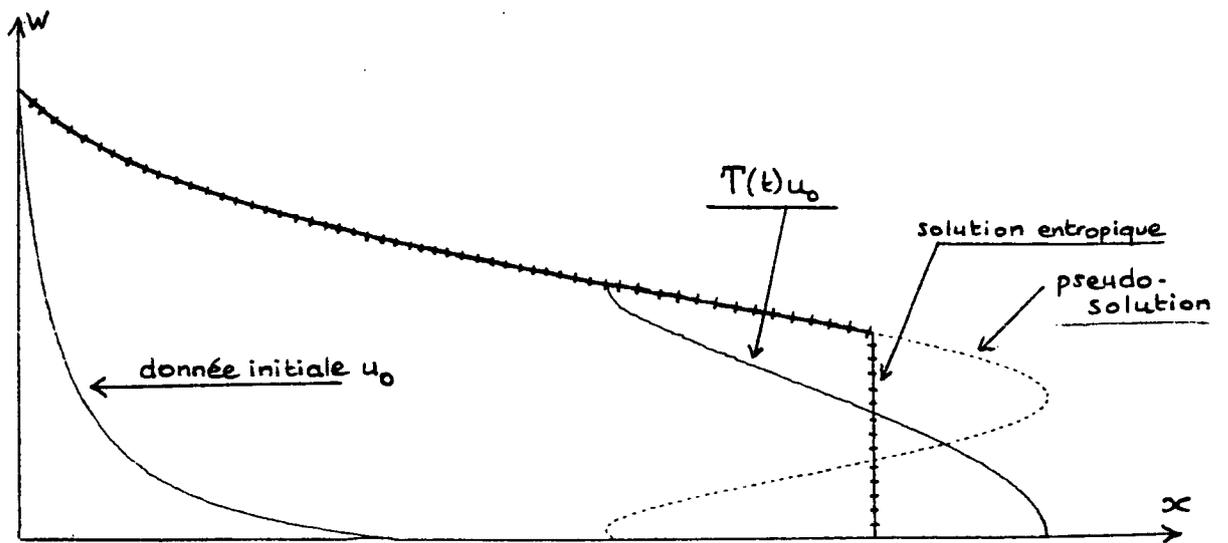


Figure 4

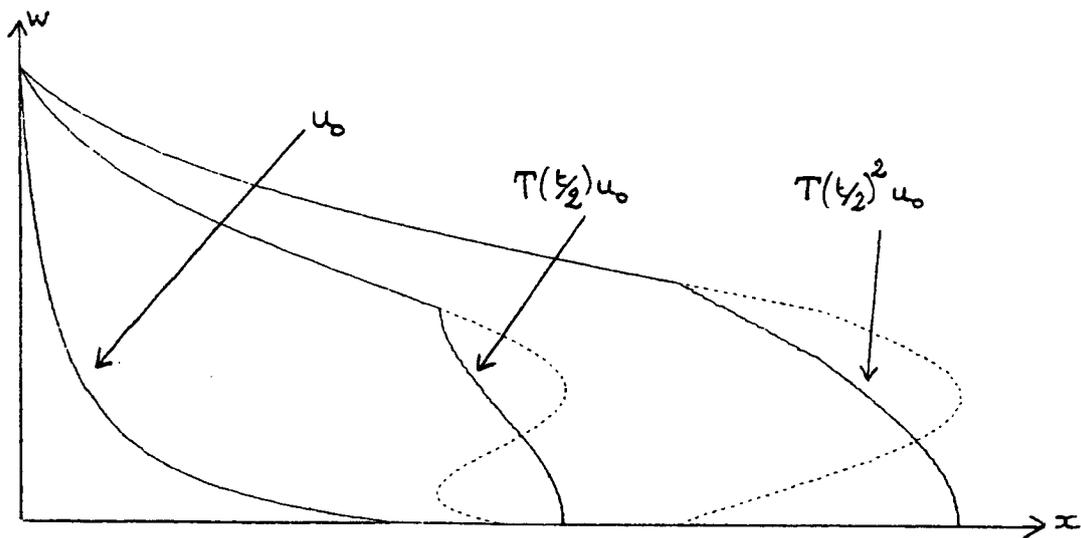


Figure 5

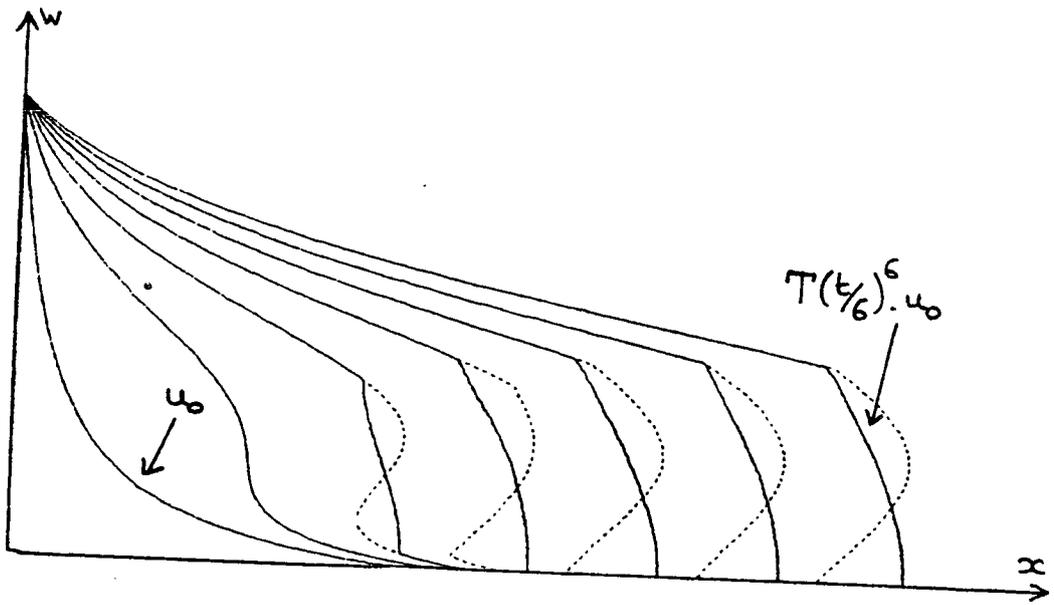


Figure 6

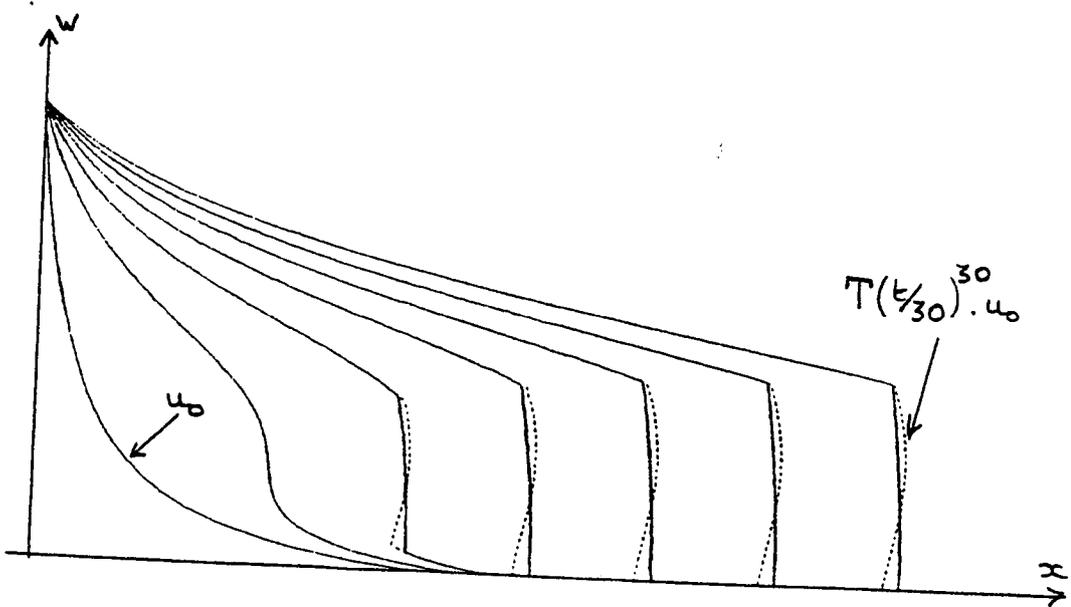


Figure 7

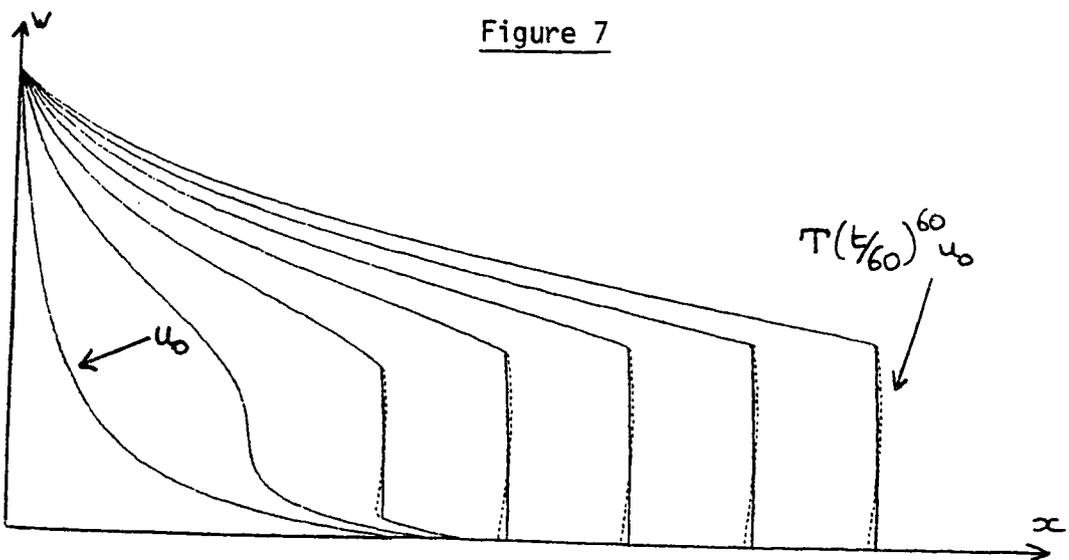


Figure 8

THEOREME 1.1

Si l'application  $F : \mathbb{R} \rightarrow \mathbb{R}^N$  est lipschitzienne, alors, pour une donnée initiale  $u_0 \in L^1(\mathbb{R}^N)$ , l'unique solution entropique de (1), s'écrit sous la forme :

$$u(t, \cdot) = \lim_{n \rightarrow \infty} T\left(\frac{t}{n}\right)^n \cdot u_0 \text{ dans } L^1(\mathbb{R}^N), \text{ quel que soit } t > 0.$$

Il est intéressant d'étudier la continuité du procédé par rapport aux données. On peut énoncer, en utilisant [3] :

THEOREME 1.2

Soit  $(u_{0,m})$  une suite convergeant dans  $L^1(\mathbb{R}^N)$  vers  $u_0$  ; soit  $(F_p)$  une suite d'applications  $\mathbb{R} \rightarrow \mathbb{R}^N$  uniformément lipschitziennes, telles que :  $F'_p(w) \rightarrow F'(w)$  p.p. sur  $\mathbb{R}$  ; si l'on note  $T_p(t)$  l'opérateur associé à  $F_p$ , alors l'unique solution entropique de (1) s'écrit :

$$u(t, \cdot) = \lim_{m, n, p \rightarrow \infty} T_p\left(\frac{t}{n}\right)^n \cdot u_{0,m} \text{ dans } L^1(\mathbb{R}^N), \forall t > 0$$

c) Autres propriétés de l'opérateur  $T(\delta t)$ 

Les solutions entropiques peuvent être approchées, on vient de le voir, à l'aide de l'opérateur  $T(\delta t)$ . Cet opérateur possède d'autres propriétés très intéressantes pour les applications numériques ; énonçons-les :

i) monotonie

$$(9) \quad u \geq v \text{ p.p.} \implies T(\delta t)u \geq T(\delta t)v \text{ p.p.}$$

ii) principe du maximum

$$(10) \quad \inf \operatorname{ess} v \leq \inf \operatorname{ess} T(\delta t)v \leq \sup \operatorname{ess} T(\delta t)v \leq \sup \operatorname{ess} v$$

iii) conservation de l'intégrale

$$(11) \quad \int_{\mathbb{R}^N} T(\delta t)v(x)dx = \int_{\mathbb{R}^N} v(x)dx$$

iv) stabilité inconditionnelle en normes  $L^p$  et en variation totale<sup>(\*)</sup>

$$(12) \quad \| T(\delta t)v \|_{L^p(\mathbb{R}^N)} \leq \| v \|_{L^p(\mathbb{R}^N)}, \quad \forall p \in [1, +\infty]$$

$$(13) \quad TV(T(\delta t)v) \leq TV(v)$$

v) continuité par rapport aux données initiales

$$(14) \quad \| T(\delta t)u - T(\delta t)v \|_{L^1} \leq \| u - v \|_{L^1}$$

#### 4 - CONSTRUCTION DE SCHEMAS NUMERIQUES A PARTIR DE LA METHODE DE TRANSPORT-ÉCROULEMENT : PRINCIPE GENERAL

Jusqu'à maintenant nous n'avons effectué, en introduisant la méthode de transport-écroulement, qu'une semi-discrétisation en temps de l'équation (1).

Pour obtenir un schéma numérique effectif, il faut procéder à une discrétisation complète du problème. Il paraît donc naturel d'introduire un

(\*) La variation totale d'une fonction  $v \in L^1(\mathbb{R}^N)$  peut être définie comme suit :

$$TV(v) = \sup \left\{ \frac{1}{\|h\|} \cdot \int_{\mathbb{R}^N} |v(x+h) - v(x)| dx ; h \in \mathbb{R}^N, h \neq 0 \right\}$$

( $\|\cdot\|$  désignant une norme donnée sur  $\mathbb{R}^N$ ).

un maillage spatial et inévitable de substituer à  $T(\delta t)$  un opérateur discret en espace  $T_h(\delta t)$ .

Or il est possible de ne modifier en rien  $T(\delta t)$  et d'obtenir néanmoins un schéma numérique, à condition de discrétiser convenablement les données  $u_0$  et  $F$  : nous allons le voir dans les chapitres suivants.

Procéder ainsi présente les avantages suivants :

i) le schéma numérique obtenu conserve toutes les propriétés de monotonie, de conservation de l'intégrale, de principe du maximum et de stabilité que nous avons énoncées en § 3c) pour l'opérateur  $T(\delta t)$  ;

ii) la convergence du schéma est assurée, d'après le théorème 1.2, pour peu que les approximations des données  $u_0$  et  $F$  convergent convenablement.

iii) l'erreur introduite par le schéma ne dépend que du pas de temps  $\delta t$  et de l'approximation des données  $u_0$  et  $F$ .

iv) aucune condition de stabilité du type "Courant-Friedrichs-Lewy" liant maillage en espace et pas de temps n'est nécessaire.

En revanche, ce principe conduit, dans le cas bidimensionnel, à une discrétisation drastique de la donnée  $F$  (l'une de ses composantes sera entièrement décrite par des nombres entiers, voir Chapitre III, § 1) :  $F$  ne sera pas toujours très fidèlement représentée.

Enfin, ce principe ne semble pas pouvoir se généraliser au cas où  $F$  dépend des variables spatiales, cas pourtant important en pratique.

## CHAPITRE II

CONSTRUCTION DE SCHEMAS NUMERIQUES A PARTIR DE LA  
METHODE DE TRANSPORT-ECROULEMENT : LE CAS UNIDIMENSIONNEL

Nous renvoyons au chapitre I (§ 4) pour le principe général. On supposera en principe  $F$  lipschitzienne et  $u_0$  intégrable sur  $\mathbb{R}$ , conformément au théorème (1.1) (cf. Chapitre I, § 3b).

En pratique, on considèrera plutôt le cas où  $u_0$  est constante pour les  $|x|$  grands (mais pas nécessairement nulle), et à variation totale finie : l'exemple type est la fonction d'Heavyside.

### 1 - DISCRETISATION DES DONNEES

On considère le "plan du graphe"  $\mathbb{R}_x \times \mathbb{R}_w$  et on introduit un maillage à pas constant de l'axe  $\mathbb{R}_w$  :

$$w_{\text{ref}} + \delta w \mathbb{Z} \text{ (où } w_{\text{ref}} \text{ est une valeur de référence).}$$

Sur chaque maille, on remplace  $F'$  par sa moyenne, ce qui revient à approcher  $F$  par une fonction  $F_{\delta w}$ , affine par morceaux, les points anguleux correspondant aux noeuds du maillage.

On remplace par ailleurs la donnée initiale  $u_0$  par une fonction  $u_{0,\delta w}$  constante par morceaux, à valeurs dans  $w_{\text{ref}} + \delta w \mathbb{Z}$  et qu'on suppose à variation bornée sur  $\mathbb{R}_x$  (ceci afin d'avoir un nombre fini de points de discontinuité).

Ces types d'approximation assurent, pour les deux données  $F$  et  $u_0$ , une convergence convenable, lorsque  $\delta w$  tend vers 0. (\*)

Nous allons maintenant voir comment, grâce à la discrétisation des données, la méthode de transport-écroulement conduit, sans modification, à un schéma numérique effectif. Nous noterons  $F$  et  $u_0$  les données approchées  $F_{\delta w}$  et  $u_{0,\delta w}$  pour alléger l'écriture.

---

(\*) convenable signifie ici conforme aux hypothèses du théorème 1.2, cf. Chapitre 1, § 3b.

2 - DESCRIPTION DU SCHEMA UNIDIMENSIONNELa) Représentation de  $u_0$  par un nombre fini de "particules" (\*)

On suppose donc que  $u_0$  est constante par morceaux et qu'elle a un nombre fini de sauts tous multiples entiers du pas  $\delta w$ .

Chaque arête verticale de son graphe peut donc être décomposée en un nombre entier  $> 0$  de segments élémentaires de hauteur  $\delta w$  (voir figure 2.1). Chacun de ces segments élémentaires sera repéré conventionnellement dans le plan  $(x,w)$  par son barycentre. On obtient ainsi une collection de points dans le plan  $(x,w)$  qu'on appellera (de façon imagée) "particules".

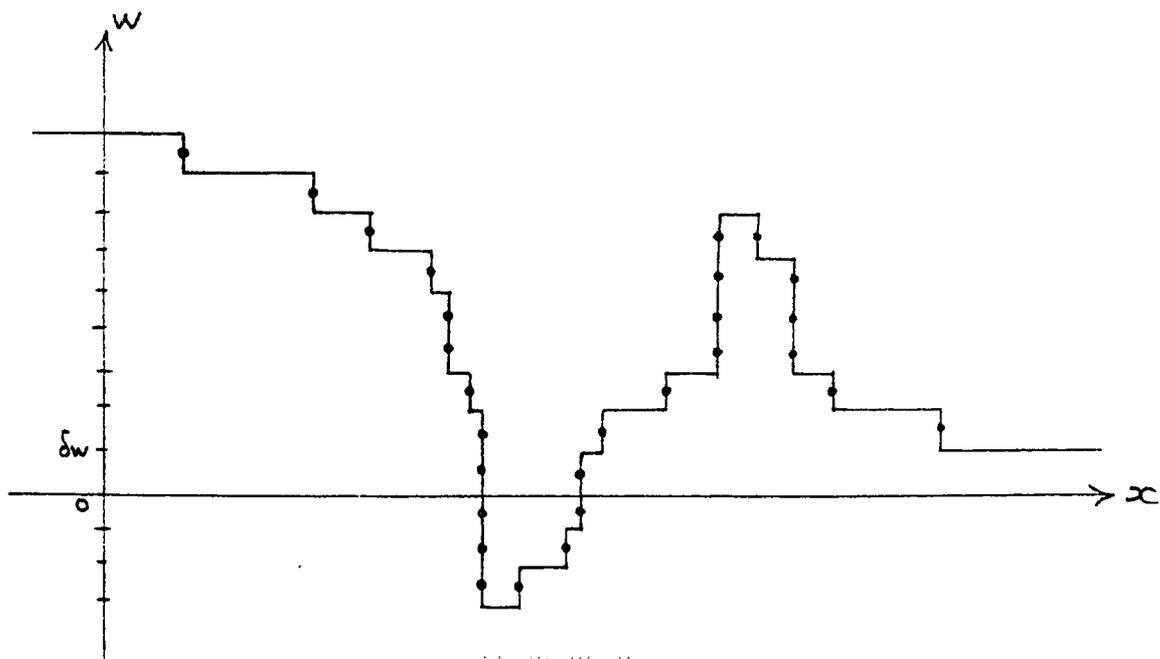


Figure 2.1

(\*) Notre technique est voisine de celle qu'utilise A. CHOPIN pour discrétiser l'équation de la chaleur ([5], § 2).

Ayant numéroté les particules de manière quelconque, on note  $x_0(i)$  et  $w_0(i)$  les coordonnées de la  $i$ ème, on la dote d'un signe  $s(i)$  valant  $-1$  si elle correspond à un saut décroissant de la fonction  $u_0$  et  $+1$  dans le cas contraire.

Connaissant la valeur de  $u_0$  au voisinage de  $-\infty$ , soit  $u_\infty$ ,<sup>(\*)</sup> on peut déterminer  $u_0$  partout (sauf aux points de discontinuité), grâce à la formule :

$$(2.1) \quad u_0(x) = u_\infty + \delta w \sum_{i \text{ t.q. } x_0(i) < x} s(i)$$

Les "particules" nous serviront donc de degrés de liberté. Remarquons que leur nombre  $N$  est lié à la variation totale de  $u_0$  par la relation :

$$(2.2) \quad TV(u_0) = N \cdot \delta w$$

*b) Déplacement horizontal des "particules" (phase de transport)*

Etant donné un pas de temps  $\delta t$ , on va déplacer chaque particule dans le plan  $(x, w)$ , horizontalement : la  $i$ ème particule de coordonnées  $(x_0(i), w_0(i))$  ira se placer au point :  $(x_1(i), w_0(i))$  où :

$$(2.3) \quad x_1(i) = x_0(i) + \delta t \cdot F'(w_0(i))$$

---

(\*) Rappelons qu'on s'est autorisé à considérer le cas où  $u_0$  est constante à l'infini (et pas forcément nulle), bien que le théorème 1.1 ait été énoncé pour  $u_0$  intégrable sur  $\mathbb{R}$  (chapitre I, § 3b).

On obtient un nouveau profil dans le plan  $(x,w)$ , déformation du graphe de  $u_0$ , qui en général n'est pas lui-même un graphe (voir figure 2.2).

Comme  $F'$  est constante sur chaque maille de l'axe  $w$ , cela revient exactement à effectuer la phase de transport de la méthode de transport-écroulement.

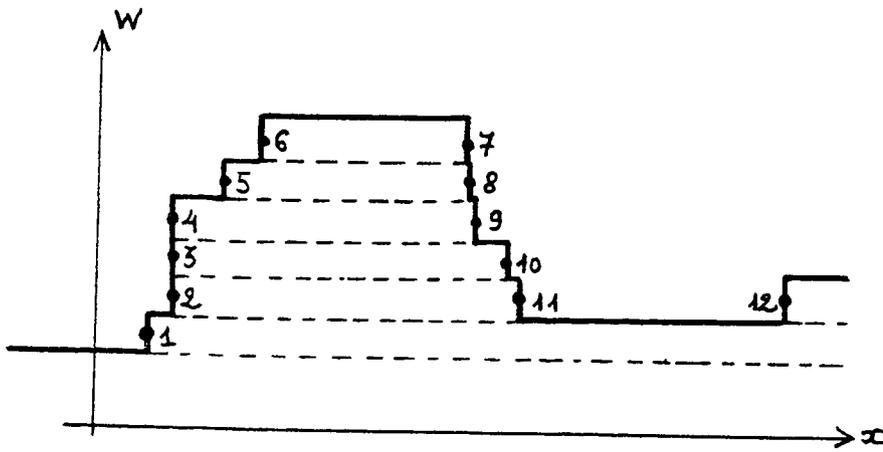
c) Déplacement vertical des "particules" (phase d'écroulement)

Après la phase de transport, on doit effectuer celle d'écroulement pour obtenir  $u_1 = T(\delta t)u_0$ .

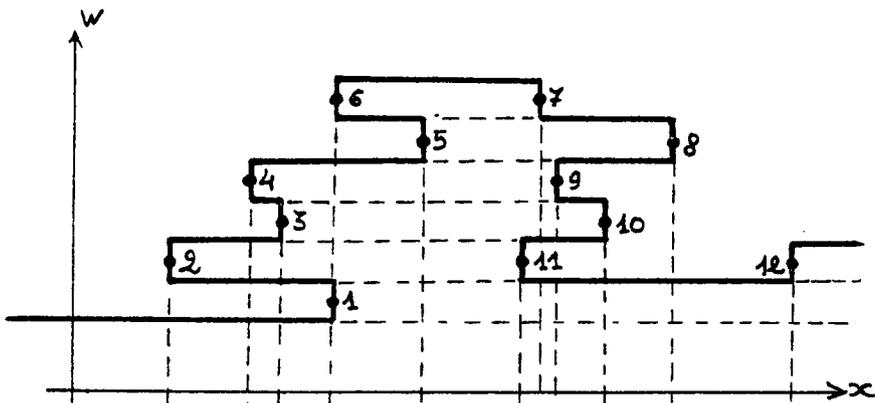
Il est facile de voir que  $u_1(x)$  est donnée, pour presque tout  $x$ , par la formule :

$$(2.4) \quad u_1(x) = u_\infty + \delta w \cdot \sum_{i \text{ t.q. } x_1(i) < x} s(i)$$

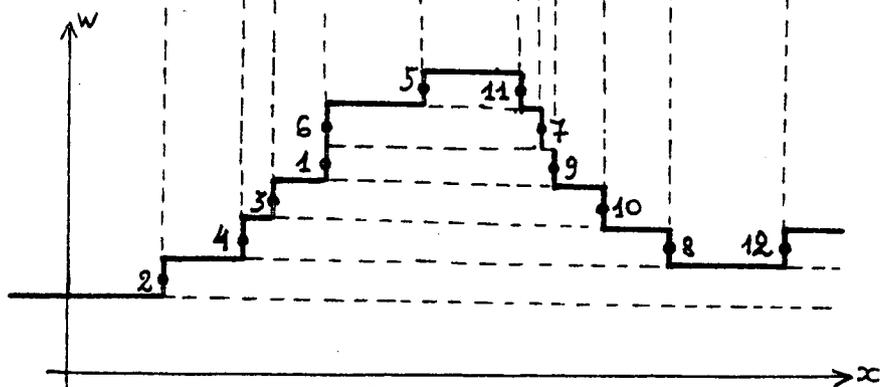
Cette opération peut se traduire en termes de "déplacement de particules" : en effet tout se passe comme si chaque particule subissait un mouvement vertical, ce qu'on voit bien sur la figure 2.2.



avant transport



après transport  
et  
avant écoulement



après écoulement

Figure 2.2

On peut définir précisément ce mouvement comme suit :

1°) On cherche une permutation  $i_1, \dots, i_N$  telle que :

$$(2.5) \quad x_1(i_1) \leq x_1(i_2) \leq \dots \leq x_1(i_N)$$

ce qui revient à faire le tri des  $x_1(i)$ .

2°) Ensuite on pose (en introduisant les quantités auxiliaires  $v_k$ )

$(2.6) \quad \begin{array}{ll} \text{i) } v_1 = u_\infty & \\ \text{ii) } v_{k+1} = v_k + s(i_{k+1})\delta w & k=1, \dots, N-1 \\ \text{iii) } w_1(i_k) = v_k + s(i_k) \frac{\delta w}{2} & k=1, \dots, N \end{array}$
--

et la nouvelle position de la particule  $i$  sera =  $x_1(i)$ ,  $w_1(i)$ .

REMARQUE : La permutation  $i_1, \dots, i_N$  n'est pas forcément unique : il y a ambiguïté dès que l'on a deux particules  $k$  et  $\ell$  telles que :  $k \neq \ell$  et  $x_1(k) = x_1(\ell)$ . Pour le calcul de la solution approchée de  $u_t + F(u)_x = 0$ , cette ambiguïté est sans importance. Par contre, si l'on s'intéresse spécifiquement aux "particules" et à leurs trajectoires, il convient de lever cette ambiguïté. On peut le faire en respectant le logique de l'écroulement : si une particule est située, avant écroulement, à la même abscisse qu'une autre, mais plus haut, cette propriété doit être conservée après écroulement.

(C'est ce qu'on a fait, sur la figure 2.2 pour les particules 1 et 6).

Ainsi le passage de  $u_0$  à  $u_1 = T(\delta t)u_0$  a été entièrement décrit à l'aide de la notion de particules, les phases de transport et d'écroulement se traduisant par le mouvement des particules dans le plan  $(x, w)$ .

Le calcul de  $u_n = T(\delta t)^n \cdot u_0$  à partir de  $u_{n-1}$  s'effectue de la même façon.

REMARQUE : La méthode générale nous fournit l'estimation :

$$TV(u_{n+1}) \leq TV(u_n) \leq \dots \leq TV(u_0)$$

Les  $u_n$  appartiennent toutes à la même classe de fonctions, et le nombre minimal de particules nécessaires pour les représenter, soit  $N_n$ , est lié à leur variation totale par la relation :

$$TV(u_n) = N_n \cdot \delta w$$

Ainsi le nombre (minimal) de degrés de liberté pour représenter les  $u_n$  ne peut pas croître, ce qui explique la "calculabilité" du schéma.

Par ailleurs, dans la description du schéma, le nombre de particules reste constant au cours du temps. Il ne croît pas, et c'est bien en accord avec ce qui vient d'être dit. Mais il ne décroît pas, ce qui semble paradoxal, puisque rien a priori n'empêche la stricte décroissance de la variation totale.

L'explication de ce paradoxe est simple : lorsqu'on passe de  $u_{n-1}$  à  $u_n$ , il se peut que deux particules  $i$  et  $j$  se retrouvent à la même position dans le plan  $(x,w)$ , tout en ayant des signes opposés :

$$s(i) = +1 \quad s(j) = -1 \quad (*)$$

Ces deux particules se neutralisent mutuellement (\*\*), comme le montre bien la formule :

$$u_n(x) = u_\infty + \sum_{i+q \quad x_n(i) < x} s(i)$$

Elles sont donc superflues pour la description de  $u_n$ , et cela explique l'inégalité  $TV(u_n) = N_n \cdot \delta w < N \delta w$  qui traduit la diminution de variation totale.

(\*) à ce propos, remarquons que le signe d'une particule ne change jamais, dans notre construction.

(\*\*) peut-on parler de particules et d'antiparticules ?!

Notons que si toutes les particules décrivant  $u_0$  ont même signe (ce qui signifie que  $u_0$  est monotone), alors la variation totale reste rigoureusement constante.

### 3 - TRAITEMENT DU PROBLEME AUX LIMITES UNIDIMENSIONNEL

Le problème posé est le suivant :

$$(2.7) \quad \begin{aligned} u_t + (F(u))_x &= 0 \quad ; \quad a < x < b \quad ; \quad t > 0 \\ u(x,0) &= u_0(x) \\ \left. \begin{aligned} u(a,t) &= u_a \\ u(b,t) &= u_b \end{aligned} \right\} & \text{au sens de Bardos-Leroux-Nedelec ([1])} \end{aligned}$$

Pour simplifier on suppose que  $a, b, u_a$  et  $u_b$  sont des constantes.

Etant donné un pas de temps  $\delta t$ , on construit par la méthode de transport-écroulement l'opérateur  $T(\delta t)$  associé au problème de Cauchy dans  $\mathbb{R}$  tout entier. On introduit l'opérateur de restriction-prolongement  $p$  défini par :

$$(2.8) \quad pu(x) = \begin{cases} u(x) & \text{si } a < x < b \\ u_a & \text{si } x < a \\ u_b & \text{si } b < x \end{cases}$$

et on pose :

$$(2.9) \quad u_n = (T_{\delta t} \circ p)^n u_0 \quad ; \quad n = 1, 2, 3, \dots \quad (*)$$

En pratique, on reprend exactement l'algorithme qu'on a décrit avec les modifications suivantes :

---

(\*) nous n'avons pas, à ce jour, démontré la convergence (bien qu'elle ne fasse aucun doute) vers la solution du problèmes aux limites.

- i) comme donnée initiale, on prend  $pu_0$  ;
- ii) on effectue une troncature sur le déplacement horizontal des particules, suivant la formule :

$$(2.10) \quad x_{n+1}(i) = \text{Max}\{a, \text{Min}\{b, x_n(i) + \delta t \cdot F'(w_n(i))\}\}$$

#### 4 - RESULTATS NUMERIQUES EN DIMENSION 1

On a choisi, pour illustrer la méthode, les trois cas suivants :

- i) le problème de Riemann<sup>(\*)</sup>

$$(2.11) \quad \begin{cases} u_t + (F(u))_x = 0 & -\infty < x < +\infty ; t > 0 \\ u(t=0, x) = \begin{cases} +1 & \text{si } x < 0 \\ -1 & \text{si } x > 0 \end{cases} \end{cases}$$

pour diverses données  $F$ , de structure plus ou moins compliquée (ni convexe, ni concave) ;

- ii) la formation d'un "choc" à partir d'une donnée initiale régulière ;

- iii) un problème aux limites (cf. § 3)

REMARQUE : Toute solution fournie par l'algorithme unidimensionnel est constante par morceaux et son graphe se présente donc sous la forme d'une succession de marches d'escalier. Graphiquement, nous avons préféré représenter le graphe par la ligne brisée définie par les positions des "particules" (cf. § 2 et fig. 2.3).

---

(\*) On parle de problème de Riemann lorsque la donnée initiale est la fonction d'Heavyside, à une transformation affine près.

On améliore ainsi la qualité des résultats, au moins visuellement.

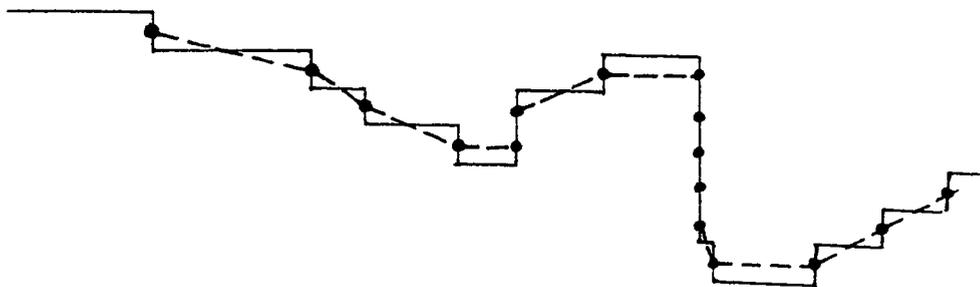


Figure 2.3

a) Le problème de Riemann

Le graphe de la donnée initiale est défini dans le plan  $(x,w)$  par les équations :

$$(2.12) \quad \begin{cases} w = +1 & -\infty < x < 0 \\ x = 0 & -1 < w < +1 \\ w = -1 & 0 < x < +\infty \end{cases}$$

Par la méthode des caractéristiques (voir Chapitre I, § 1), on obtient au temps  $t$  la "pseudo-solution", d'équations :

$$(2.13) \quad \begin{cases} w = +1 & -\infty < x < tF'(+1) \\ x = tF'(w) & -1 < w < +1 \\ w = -1 & tF'(-1) < x < +\infty \end{cases}$$

qui n'est univoque que lorsque  $F$  est concave.

La solution réelle est néanmoins donnée par les formules (2.13), à condition de remplacer  $F$  par son enveloppe concave  $F_c$  sur l'intervalle  $[-1, +1]$  (cf. [8]) :

$$(2.14) \quad \left\{ \begin{array}{ll} w = +1 & -\infty < x < tF'_C(+1) \\ x = tF'_C(w) & -1 < w < +1 \\ w = -1 & tF'_C(-1) < x < +\infty \end{array} \right.$$

On pourra se reporter aux figures (2.4) et (2.5) ou sont représentées respectivement :

i) un exemple de fonction  $F$  (en pointillé) et son enveloppe concave  $F_C$  (en continu) - figure (2.4) ;

ii) la solution exacte au temps  $t = 1$  (en continu) et la pseudo-solution associée (en pointillé) - figure (2.5).

On notera que la figure (2.5) contient toute l'information du problème, étant entendu que, pour  $t = 1$ , la pseudo-solution coïncide avec le graphe de  $F'$  restreinte à  $[-1, +1]$  (formule (2.13)). Aussi représenterons nous les résultats numériques suivant le modèle de la figure (2.5), en figurant en pointillé la donnée  $F'$  (c'est-à-dire la pseudo-solution au temps  $t = 1$ ) et en continue la solution approchée, calculée au temps  $t = 1$  par notre méthode dans les conditions suivantes :

$$(2.15) \quad \delta t = 0.02, \delta w = 0.04$$

Autrement dit, en utilisant 50 pas de temps et 50 mailles sur l'axe des  $w$ .

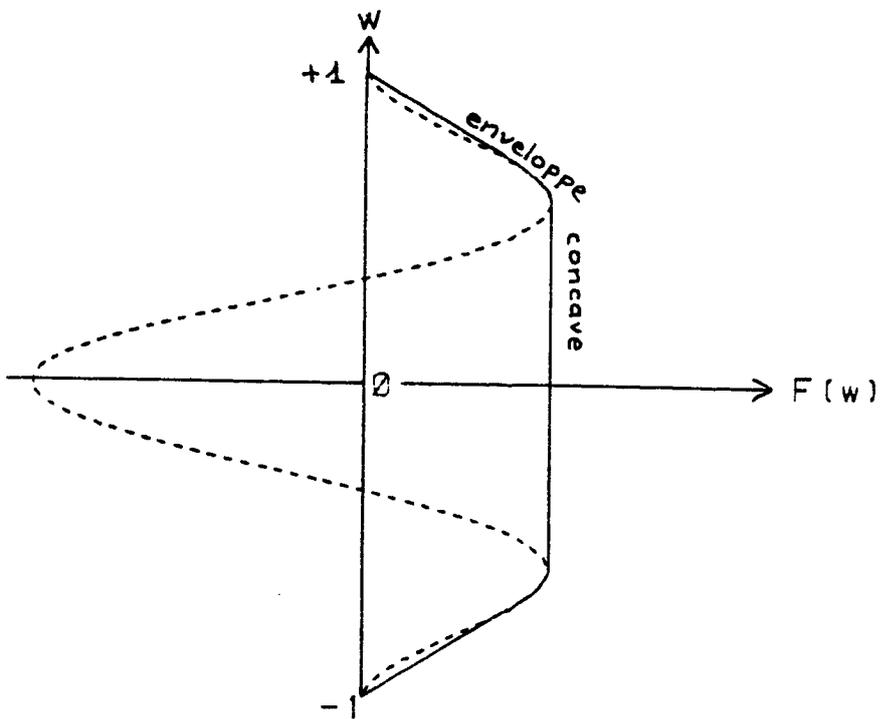


Figure 2.4

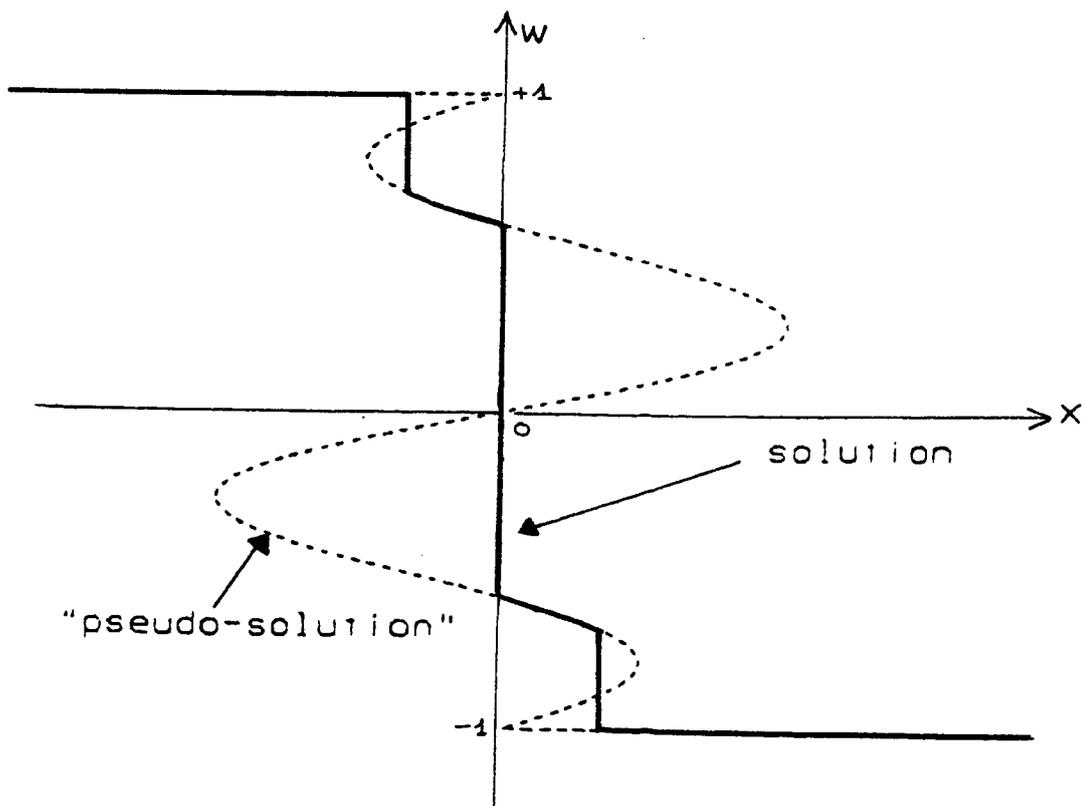


Figure 2.5

On considère quatre données  $F$  différentes sur l'intervalle  $[-1, +1]$  :

$$i) F(w) = \cos(\pi w)$$

$$ii) F(w) = \frac{1}{1 + \frac{1}{2} \left(\frac{1-w}{w+1}\right)^2}$$

$$iii) F(w) = -\cos\left(\frac{\pi}{2} w\right) \cos(\pi w)$$

$$iv) F(w) = -\cos\left(\frac{\pi}{2} w\right) \cos\left(\frac{3\pi}{2} w\right)$$

La complexité des solutions correspondantes va croissant :

- i) un seul choc (figure 2.6)
- ii) un choc et une détente (figure 2.7)
- iii) un choc et deux détente (figure 2.8)
- iv) trois chocs et deux détente (figure 2.9)

Les figures (2.6) à (2.9) montrent que les solutions approchées, calculées sous les conditions (2.15), sont très proches des solutions exactes, la différence étant pratiquement imperceptible visuellement.

On obtient donc quasiment la solution exacte, dans les conditions (2.15).

Etudions maintenant le comportement de l'algorithme dans des conditions plus défavorables, en considérant, par exemple, le cas de la figure (2.8).

En reprenant le même maillage en  $w$  (50 mailles,  $\delta w = 0.04$ ), on a représenté sur la figure (2.10)<sup>(\*)</sup> les résultats obtenus pour les valeurs suivantes de  $\delta t$  : 0,2 ; 0,1 ; 0,05 ; 0,02<sup>(\*\*)</sup>.

On a fait de même sur la figure (2.11), mais avec un maillage en  $w$  moins fin (25 mailles,  $\delta w = 0,08$ ).

---

(\*) l'échelle étant distordue par rapport à la figure (2.8)

(\*\*) c'est-à-dire, respectivement, 5, 10, 20 et 50 pas de temps.

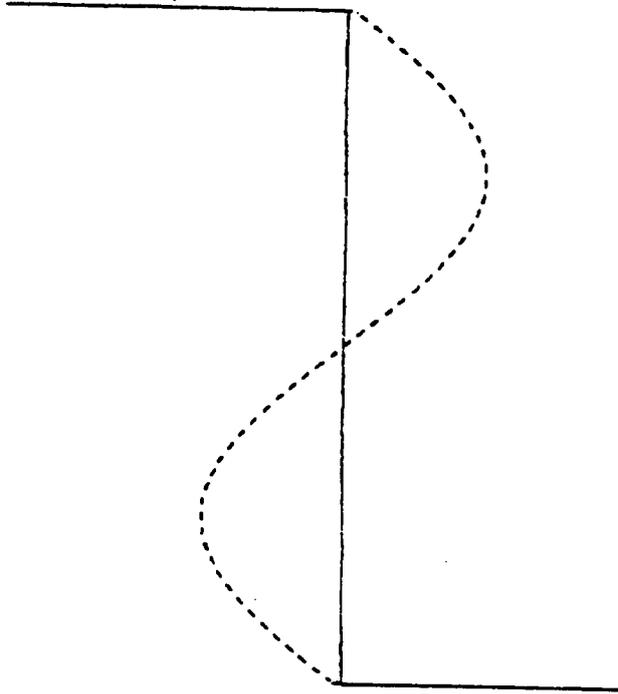


Figure 2.6

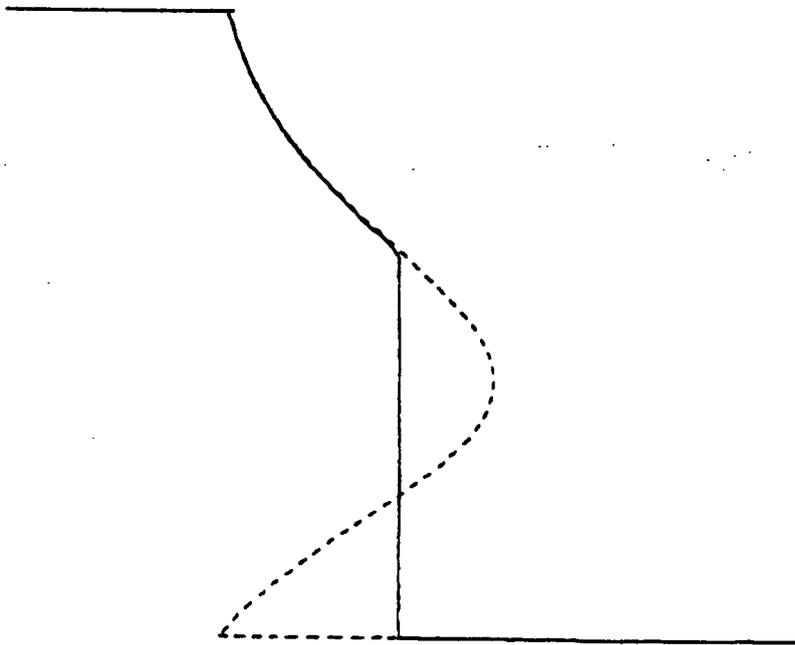


Figure 2.7

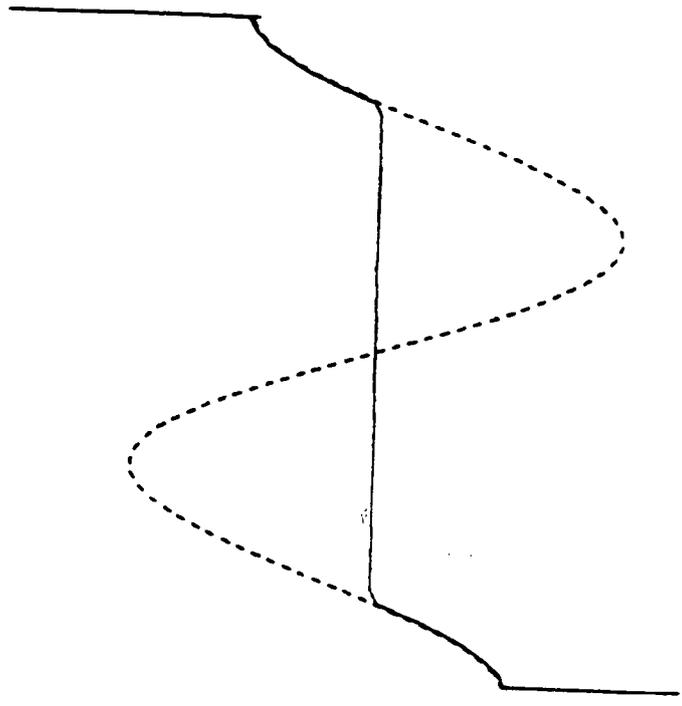


Figure 2.8

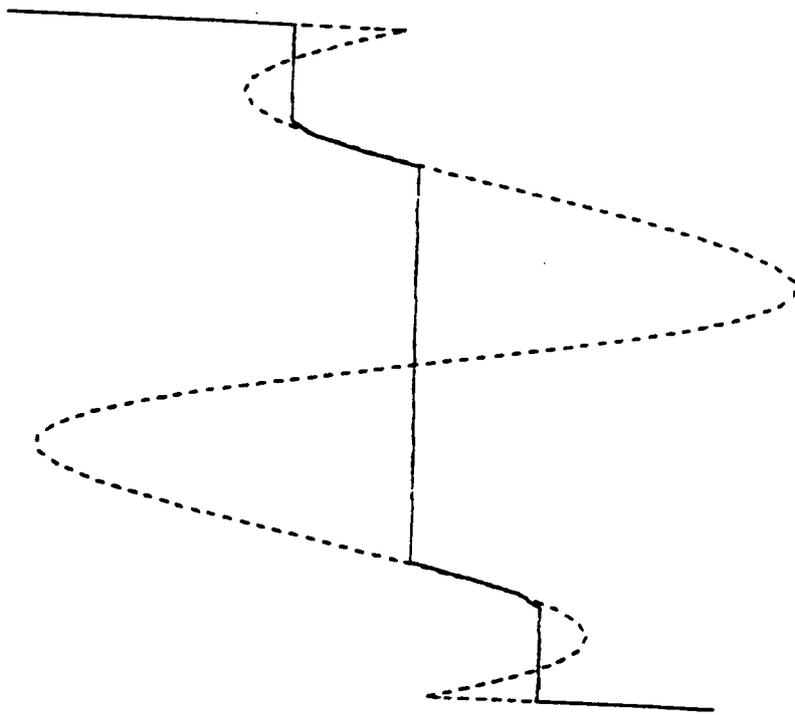


Figure 2.9

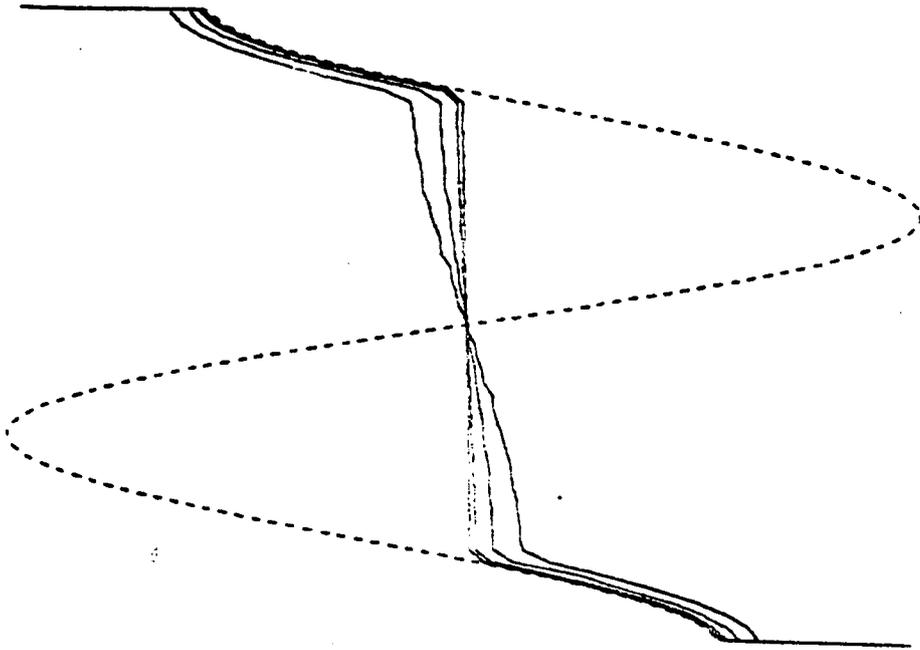


Figure 2.10

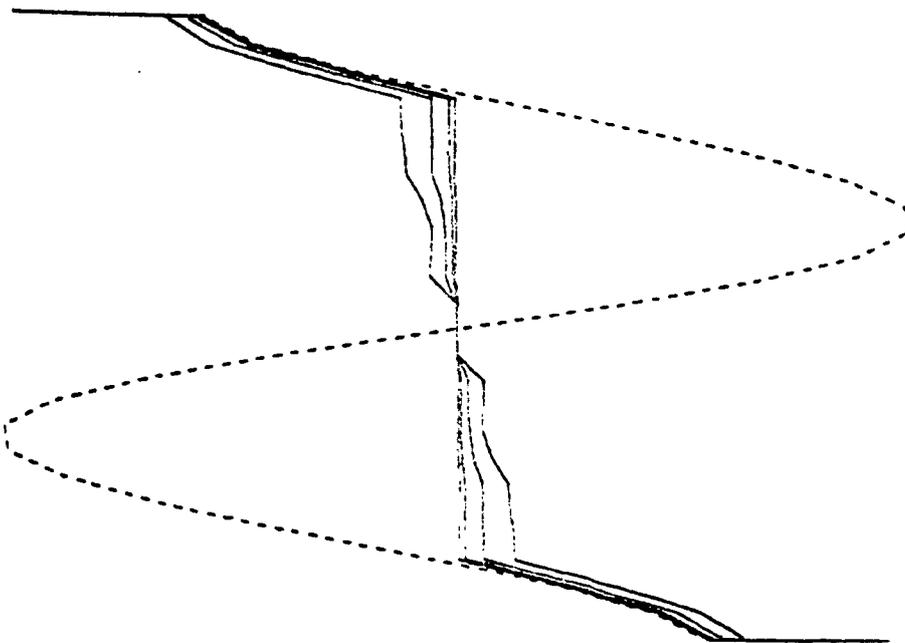


Figure 2.11

b) Formation d'un choc

L'exemple est tiré de [6].

Ici  $w$  varie entre 0 et 1 et on a :

$$F(w) = \frac{w^2}{w^2 + \frac{1}{2}(1-w)^2}$$

$$u_0(x) = \begin{cases} 1 & \text{si } x < 0 \\ \frac{\varepsilon}{x+\varepsilon} & \text{si } 0 < x < 1 \\ \frac{\varepsilon}{1+\varepsilon} & \text{si } 1 < x \end{cases} \quad \text{avec } \varepsilon = 0.1$$

Cet exemple va nous permettre de comparer (figure 2.12) notre méthode à deux autres méthodes, pour lesquelles nous disposons des résultats numériques de F. FORGES<sup>(\*)</sup> :

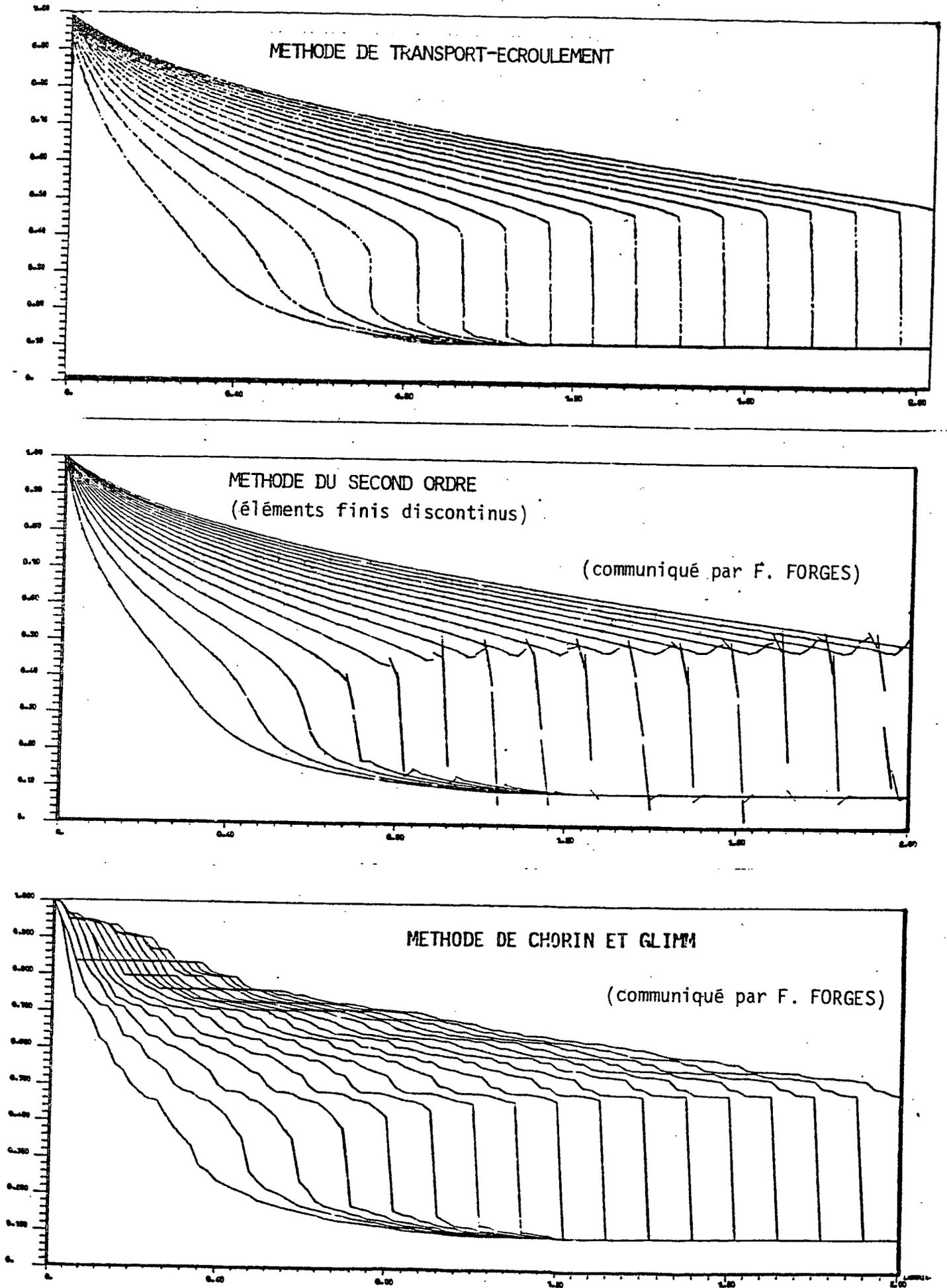
- i) le schéma de Glimm, utilisé par CONCUS [6],
- ii) un schéma du second ordre, de type "élément fini discontinu" (cf. [4]).

Dans tous les cas, on utilise le même pas de temps :  $\delta t = 0.009$ . Pour notre méthode, on prend  $\delta w = 0.02$  (c'est-à-dire 50 mailles en  $w$ ). Les autres méthodes utilisent un pas d'espace  $\delta x = 0.02$  choisi de façon à respecter au mieux la condition de Courant-Friedrichs-Lewy<sup>(\*\*)</sup>.

On représente l'évolution de la solution tous les 7 pas de temps :

(\*) Thèse de 3ème cycle, Université Paris IX

(\*\*) laquelle, rappelons le, n'a aucun sens pour notre algorithme (cf. Chapitre I § 4).



c) Un problème aux limites<sup>(\*)</sup>

On a pour données (cf. § 3).

$$a = 0 ; u_a = 1 ; b = 1 ; u_b = 0$$

$$u_0(x) = x \quad 0 < x < 1$$

$$d'où : pu_0(x) = \begin{cases} 1 & \text{si } x < 0 \\ x & \text{si } 0 < x < 1 \\ 0 & \text{si } 1 < x \end{cases}$$

et successivement :

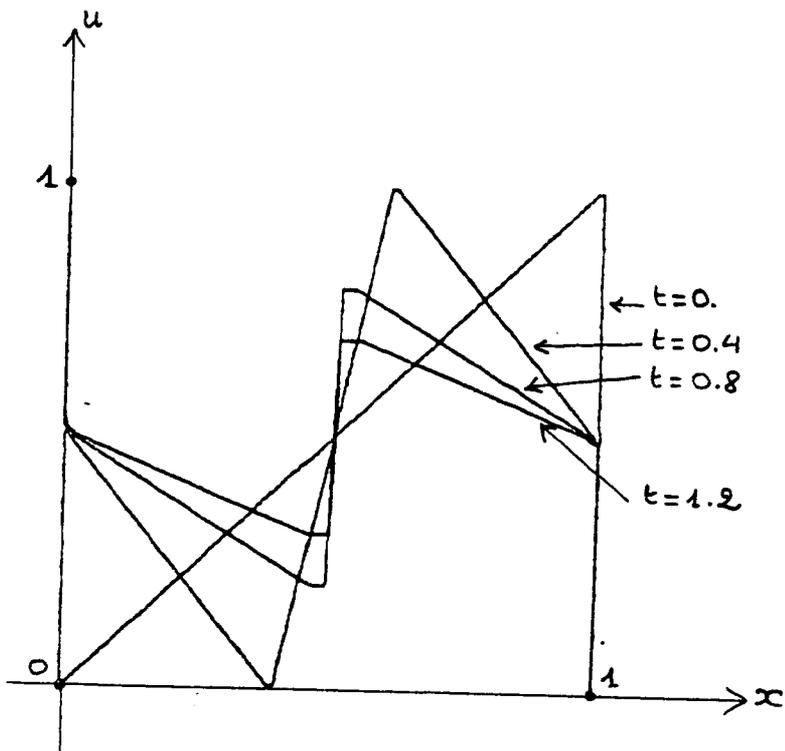
$$F(w) = w(1-w) \text{ et } F(w) = -w(1-w)$$

La solution de ces deux problèmes aux limites peut être calculée "à la main" (voir [4]). Sur les figures (2.13) et (2.14), nous avons représenté les solutions approchées, pour diverses valeurs du temps, calculées sous les conditions suivantes :

$$\delta t = 0.04 \text{ et } \delta w = 0.02 \quad (\text{i.e. } 50 \text{ mailles en } w) ;$$

elles diffèrent très peu des solutions exactes (les angles sont légèrement arrondis).

(\*) Une interprétation physique de ce problème aux limites est la suivante : On considère un barreau poreux vertical (représenté par  $[a,b]$ ) et deux phases liquides (eau et huile), avec une certaine répartition initiale de saturation en eau  $u_0(x)$ . Les extrémités a et b du barreau sont en contact respectivement avec l'eau ( $u_a = 1$ ) et l'huile ( $u_b = 0$ ). On calcule l'évolution au cours du temps de la saturation en eau  $u(t,x)$ , à l'aide du modèle de Muskat, sous le seul effet de la gravité, qu'on suppose dirigée de a vers b dans le premier cas (huile en bas, eau en haut) et de b vers a dans le second. Voir l'étude en détail dans [4].

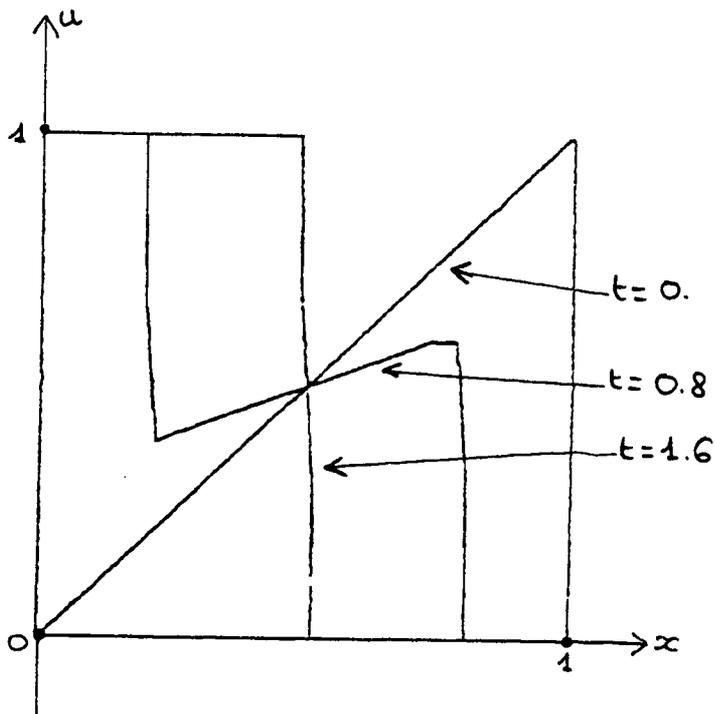


eau

$\xrightarrow{\text{gravité}}$

huile

Figure 2.13



eau

$\xleftarrow{\text{gravité}}$

huile

Figure 2.14

### CHAPITRE III

CONSTRUCTION DE SCHEMAS NUMERIQUES A PARTIR DE LA  
METHODE DE TRANSPORT-ECROULEMENT : LE CAS BIDIMENSIONNEL

Nous renvoyons le lecteur au Chapitre 1, § 4 pour le principe général. Dans ce qui suit, on notera  $x$  et  $y$  les variables d'espaces, et :

$$w \in \mathbb{R} \rightarrow F(w) = (f(w), g(w)) \in \mathbb{R}^2.$$

On supposera que  $F$  est lipschitzienne sur  $\mathbb{R}$  et  $u_0$  intégrable sur  $\mathbb{R}^2$ . En pratique  $u_0$  sera à support compact et à variation bornée.

### 1 - DISCRETISATION DES DONNEES $u_0$ , $f$ ET $g$

Comme dans le cas unidimensionnel, on introduit un maillage à pas constant de l'axe des  $w$  :

$$(3.1) \quad w_{\text{ref}} + \delta w \cdot \mathbb{Z}$$

où  $w_{\text{ref}}$  est une valeur de référence, qu'on supposera nulle.

On approche  $f'$  et  $g'$  par des fonctions  $f'_{\delta w}$ ,  $g'_{\delta w}$  constantes sur chaque maille, et  $u_0$  par une fonction  $u_{0,\delta w}$  à support compact et variation totale finie, constante par morceaux et à valeurs dans  $\delta w \cdot \mathbb{Z}$ .

On supposera en outre que les lignes de niveau de  $u_{0,\delta w}$  dans  $\mathbb{R}^2$  sont formées de tronçons parallèles à l'un des deux axes de coordonnées  $x$  et  $y$ .

Jusqu'ici, les hypothèses ne font que généraliser assez naturellement celles faites dans le cas unidimensionnel (cf. Chapitre II, § 1).

Pour être en mesure d'appliquer les principes du Chapitre I, § 4, nous sommes conduits à introduire de nouvelles hypothèses, indiscutablement plus contraignantes qu'en dimension 1.

On introduit pour l'un des axes de coordonnées spatiales,  $y$  pour fixer les idées, un maillage à pas constant :

$$(3.2) \quad y_{\text{ref}} + \delta y \cdot \mathbb{Z}$$

On fait alors les hypothèses suivantes :

- (3.3) Les tronçons parallèles à l'axe  $x$  des lignes de discontinuité de  $u_{0,\delta w}$  -qu'on notera désormais  $u_{0,\delta w,\delta y}$ - sont entièrement contenus dans l'ensemble :

$$\mathbb{R}_x \times (y_{\text{ref}} + \delta y \mathbb{Z}) \quad (\text{voir figure 3.1})$$

- (3.4) La fonction  $g'_{\delta w}$  prend ses valeurs dans l'ensemble :

$$\frac{\delta y}{\delta t} \cdot \mathbb{Z}, \text{ on la notera donc } g'_{\delta w,\delta y,\delta t} \quad (\text{voir figure 3.2})$$

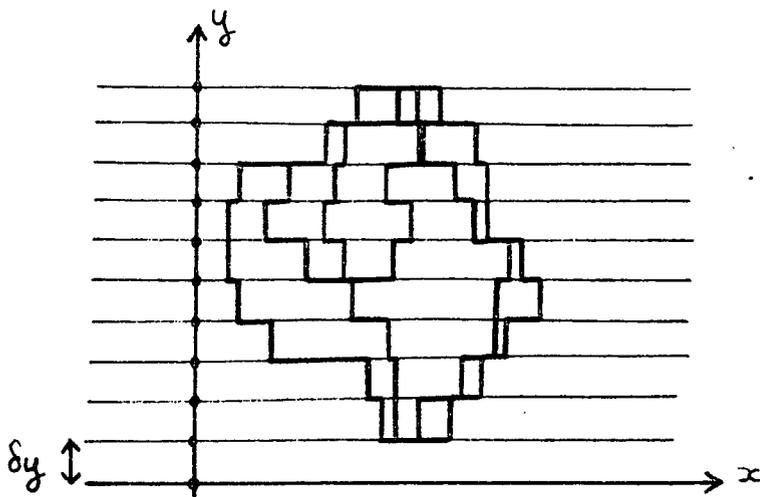


Figure 3.1

Lignes de discontinuité de  $u_{0,\delta w,\delta y}$

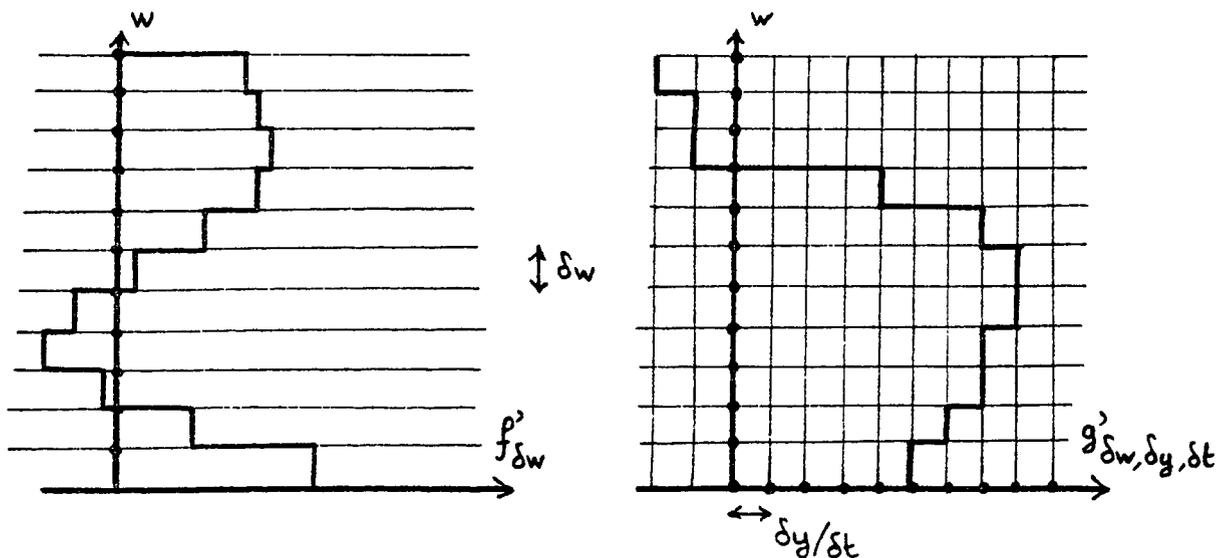


Figure 3.2

Par la première hypothèse, on suppose en fait que les coupes parallèles à l'axe des  $y$  du graphe de  $u_0$  épousent les contours du maillage  $(y_{\text{ref}}, w_{\text{ref}}) + \delta y \mathbb{Z} \times \delta w \mathbb{Z}$  du plan  $\mathbb{R}_y \times \mathbb{R}_w$  (voir figure 3.3).

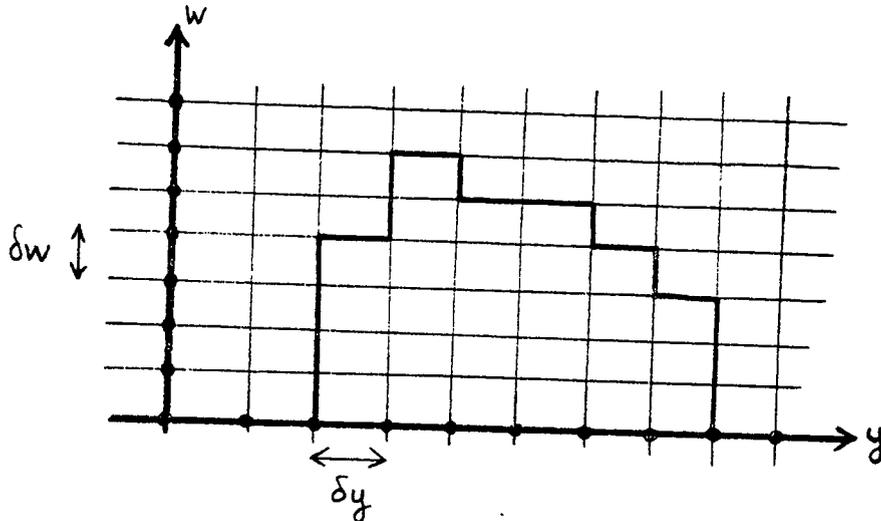


Figure 3.3

Coupe parallèle à  $y$  de  $u_0$ .

C'est cette hypothèse qui assure la possibilité de représenter  $u_0$  par un nombre fini de degrés de liberté (voir §2,a).

La seconde hypothèse est une hypothèse de compatibilité : elle assure qu'après la phase de transport les coupes parallèles à l'axe des  $y$  du graphe déformé épouseront encore les contours du maillage (cf. § 2,d et voir figure 3.4) ; ainsi, après écroulement, on obtiendra une fonction  $u_1$  du même type que  $u_0$ , représentable par un nombre fini de degrés de liberté.

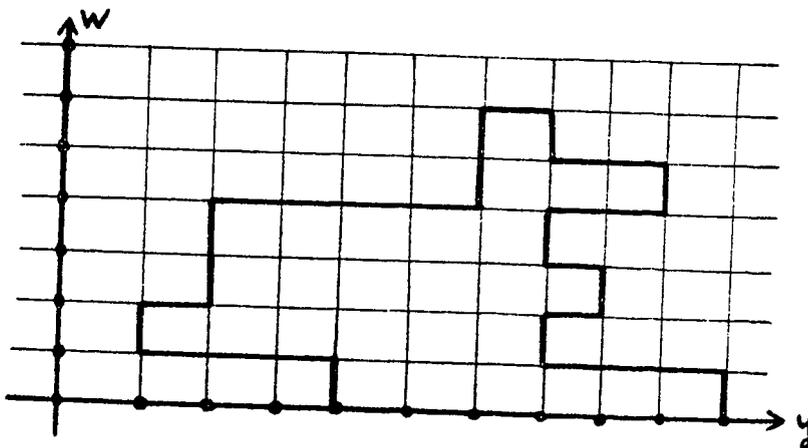


Figure 3.4

après transport

Dans ces conditions, en appliquant aux données approchées  $u_{0,\delta w,\delta y}$ ,  $f_{\delta w}$  et  $g_{\delta w,\delta y,\delta t}$  l'opérateur  $T(\delta t)$  de la méthode de transport-écroulement, on obtient effectivement un schéma numérique, comme nous le verrons bientôt.

Auparavant, il faut nous assurer que les données approchées convergent conformément aux hypothèses du théorème 1.2 (cf. Chapitre 1, § 3b), pour être certain de la convergence du schéma.

Les données  $u_{0,\delta w,\delta y}$  et  $f_{\delta w}$  ne posent pas de problème : en les choisissant convenablement on peut les faire converger vers les données  $u_0$  et  $f$  (supposées respectivement intégrable sur  $\mathbb{R}^2$  et lipschitzienne sur  $\mathbb{R}$ ), en faisant tendre vers 0 les pas  $\delta w$  et  $\delta y$ .

En revanche, l'hypothèse (3.4) -inhabituelle en analyse numérique- fait qu'en général le rapport  $\frac{\delta y}{\delta t}$  doit tendre vers 0 pour que  $g_{\delta w,\delta y,\delta t}$  converge comme il faut vers  $g$ .

Une contrainte apparaît donc, liant pas d'espace et pas de temps, ce qui n'était pas le cas en dimension un (en raison de l'absence de tout maillage spatial). Cette contrainte est complètement différente de la condition de Courant-Friedrichs-Lewy (CFL), qui apparaît dans la quasi-totalité des schémas numériques classiques ; là où la condition CFL nous imposerait :  $\delta t \leq C\delta y$ , notre contrainte s'écrit :

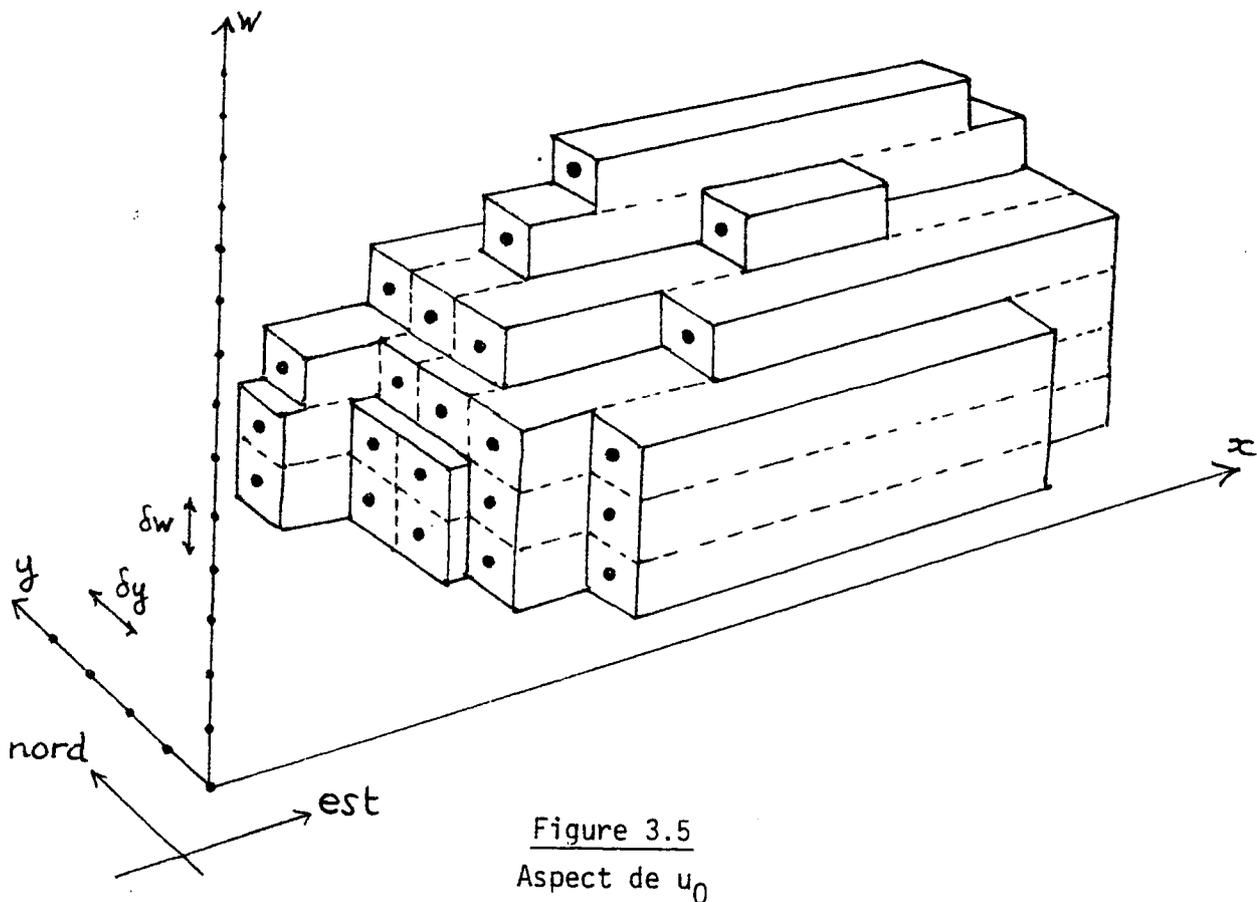
$$(3.5) \quad \delta y \ll \delta t$$

## 2 - DESCRIPTION DU SCHEMA BIDIMENSIONNEL

Pour alléger les notations, nous noterons  $u_0$ ,  $f$  et  $g$  les données approchées  $u_{0,\delta w,\delta y}$ ,  $f_{\delta w}$  et  $g_{\delta w,\delta y,\delta t}$ .

a) Représentation de  $u_0$  par un nombre fini de "particules"

- Plaçons-nous dans l'espace  $\mathbb{R}_x \times \mathbb{R}_y \times \mathbb{R}_w$  ; nous appellerons conventionnellement Est la direction des  $x$  croissants et Nord celle des  $y$  croissants. Compte tenu des hypothèses faites au paragraphe précédent, on peut décomposer chaque facette du graphe de  $u_0$  tournée vers l'Est ou l'Ouest en un nombre fini de facettes élémentaires de hauteur  $\delta w$  (suivant l'axe des  $w$ ) et de largeur  $\delta y$  (suivant l'axe  $y$ ) (voire figure 3.5).



Chacune de ces facettes sera repérée par son barycentre, nommé conventionnellement "particule".

- A l'instar de ce que nous avons fait en dimension 1, nous numérotons ces particules de façon arbitraire ; nous notons  $x_0(i)$ ,  $y_0(i)$  et  $w_0(i)$  les coordonnées de la  $i$ ème et nous lui affectons le signe  $s(i)$  qui vaut  $-1$  et  $+1$  selon que la facette correspondante regarde vers l'Est ou vers l'Ouest.

Comme les coordonnées  $y_0(i)$  et  $w_0(i)$  ont toujours des valeurs discrètes, en vertu des hypothèses (3.3) et (3.4), respectivement dans :

$$y_{\text{ref}} + \frac{\delta y}{2} + \delta y \mathbb{Z} \text{ et } w_{\text{ref}} + \frac{\delta w}{2} + \delta w \mathbb{Z} ,$$

on considèrera de préférence les entiers relatifs  $j_0(i)$  et  $k_0(i)$  :

$$(3.6) \quad j_0(i) = \frac{y_0(i) - y_{\text{ref}}}{\delta y} - \frac{1}{2} \quad k_0(i) = \frac{w_0(i) - w_{\text{ref}}}{\delta w} - \frac{1}{2}$$

- La donnée des  $x_0(i)$ ,  $j_0(i)$  et  $s(i)$  suffit à déterminer presque partout  $u_0$ , suivant la formule :

$$(3.7) \quad u_0(x,y) = \delta w \sum_i s(i)$$

$i$  décrivant l'ensemble des numéros de particules tels que :

$$x_0(i) < x \text{ et } j_0(i) = E\left(\frac{y - y_{\text{ref}}}{\delta y}\right) \quad (E \text{ désignant la partie entière})$$

Cette formule n'est rien d'autre que l'analogie discret de la formule continue :

$$u(x,y) = \int_{-\infty}^x \partial_1 u(\xi,y) d\xi,$$

valable pour toute fonction régulière à support compact.

- Le nombre de particules nécessaires pour représenter  $u_0$  est donné par la formule :

$$(3.8) \quad N = \frac{TV_x(u_0)}{\delta y \delta w} ,$$

où  $TV_x(u_0)$  désigne la variation totale en  $x$  de  $u_0$  ; c'est-à-dire :

$$\sup_{h \neq 0} \frac{1}{|h|} \iint |u_0(x+h, y) - u_0(x, y)| dx dy .$$

b) Détermination des  $k_0(i)$  par un algorithme de tri

o Lorsqu'on se donne  $u_0$ , on connaît évidemment à la fois les  $s(i)$ ,  $x_0(i)$ ,  $j_0(i)$  et  $k_0(i)$ .

Or on a vu que la seule donnée des  $x_0(i)$ ,  $j_0(i)$  et  $s(i)$  suffit à déterminer  $u_0$  presque partout.

Les quantités  $k_0(i)$  ne sont donc pas nécessaires à la description de  $u_0$ , et on peut même les reconstituer, connaissant les  $x_0(i)$ ,  $j_0(i)$  et  $s(i)$ , par un algorithme de tri. Ce calcul qui paraît superflu pour l'instant est très important pour la suite, et nous allons maintenant le décrire en détail :

o On classe les particules d'après leur position et pour cela on effectue le tri des couples  $(x_0(i), j_0(i))$  dans  $\mathbb{R} \times \mathbb{Z}$  pour la relation d'ordre :

$$(x, j) \leq (x', j') \text{ si et seulement si } \begin{cases} j < j' \\ \text{ou} \\ j = j' \text{ et } x \leq x' \end{cases}$$

Il s'agit donc d'un tri lexicographique, qui se décompose en deux étapes :

i) on regroupe les particules par classes d'équivalence :

$$(3.9) \quad C(j) = \{i, j_0(i) = j\} \quad , \quad j \text{ décrivant } \mathbb{Z}$$

On note  $m(j)$  le cardinal de  $C(j)$  ;

Le coût de cette étape est de l'ordre de  $N$  opérations élémentaires ;

ii) dans chaque classe  $C(j)$ , on trie les particules par leurs coordonnées  $x_0(i)$  ; le coût de cette opération est, pour un bon algorithme de tri, de l'ordre de  $m(j) \log_2 m(j)$  opérations élémentaires où  $\log_2$  désigne le logarithme binaire.

On peut alors stocker le résultat du tri dans un tableau

$$(3.1.) \quad ((in(j,m) ; m = 1, \dots, m(j)) ; j \in \mathbb{Z})$$

de longueur  $N$ , indiquant le numéro de la particule classée  $m^{\text{ième}}$  dans la classe  $C(j)$ .

o Une fois effectué le tri, on peut procéder au calcul des  $k_0(i)$  comme suit :

pour chaque  $j \in \mathbb{Z}$  tel que  $m(j) \neq 0$ , on pose :

$$(3.11) \quad \begin{aligned} S_m &= s(in(j,m)) \text{ si } m = 1, \dots, m(j) \\ S_0 &= -1 \text{ et } K_0 = +1 \end{aligned}$$

puis on calcule par récurrence :

$$(3.12) \quad K_m = K_{m-1} + \frac{1}{2} (S_m + S_{m-1}) \quad (*)$$

et on pose finalement :

$$(3.13) \quad k_0(in(j,m)) = K_m$$

---

(\*) Cette formule suppose  $w_{ref} = 0$ , ce qui est naturel, puisque  $u_0$  est supposée nulle à l'infini.

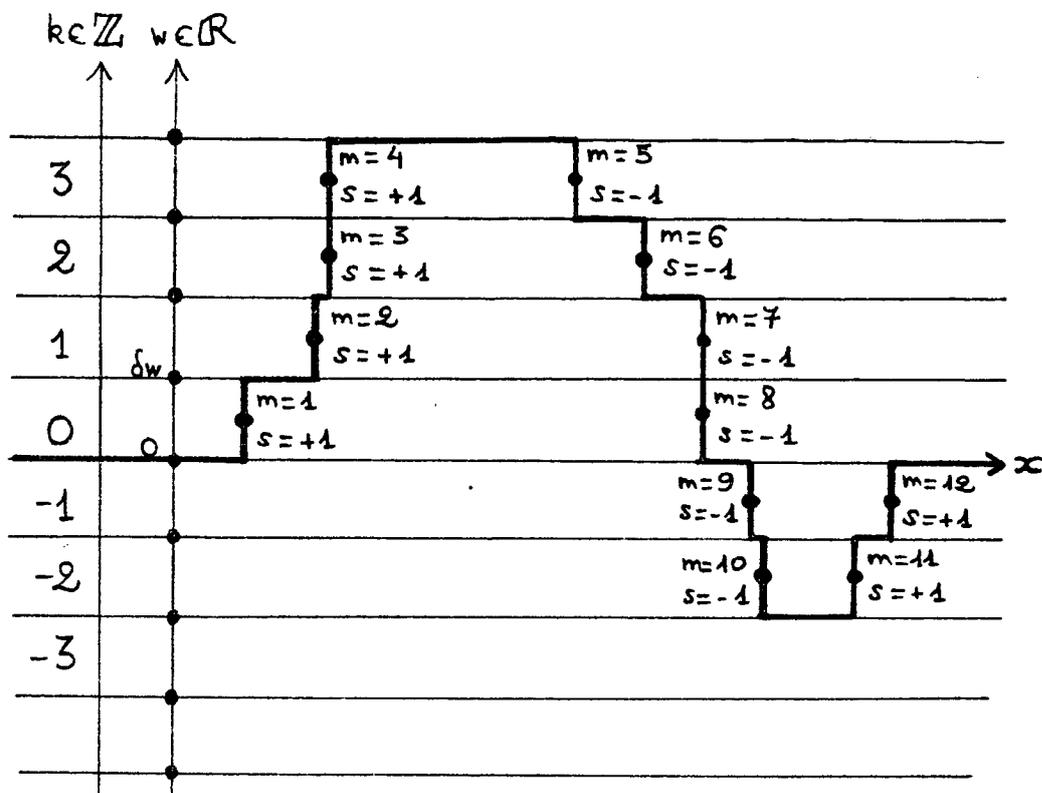


Figure 3.6

On peut s'assurer de la validité de ces formules à l'aide de la figure 3.6, où est représentée une coupe du graphe de  $u_0$  suivant l'axe  $x$ .

Le coût de ce calcul est, pour chaque  $j$ , de l'ordre de  $m(j)$  opérations élémentaires, ce qui fait en tout  $N$ .

c) Coût du calcul des  $k_0(i)$

- En effectuant le calcul des  $k_0(i)$  à partir de la seule donnée des  $x_0(i)$ ,  $j_0(i)$  et  $s(i)$  nous avons eu besoin :

- i) d'un tableau de longueur  $N$  pour stocker le résultat du tri ;
- ii) d'un nombre d'opérations élémentaires (transferts, tris et additions d'entiers) de l'ordre de  $N$ , pour effectuer le tri par rapport à la variable  $j$  et pour calculer les  $k_0(i)$  après le tri ;
- iii) d'un nombre d'opérations de l'ordre de :

$$\sum_{j \in \mathbb{Z}} m(j) \text{Log}_2 m(j)$$

pour effectuer le tri par rapport à la variable  $x$ .

On peut évaluer plus finement ce dernier nombre à l'aide des taux :

$$\theta(j) = m(j)/N \quad , \quad j \in \mathbb{Z},$$

qui sont compris entre 0 et 1, et dont la somme vaut 1.

On trouve :

$$N \sum_j \theta(j) \cdot \{\text{Log}_2 \theta(j) + \text{Log}_2 N\}$$

soit :

$$(3.14) \quad N(\text{Log}_2 N - h(\theta))$$

où  $h(\theta)$  désigne l'entropie (au sens de Kolmogorov, et non d'Oleïnik !) :

$$(3.15) \quad - \sum_j \theta(j) \text{Log}_2 \theta(j)$$

de la décomposition  $(C(j) ; j \in \mathbb{Z})$  (voir [0], p. 33-34).

d) Phase de transport : déplacement horizontal des particules

En raison de l'hypothèse (3.4), on peut introduire les quantités entières :

$$(3.16) \quad \delta j(k) = \frac{\delta t}{\delta y} \cdot g'(w_{\text{ref}} + k\delta w) \in \mathbb{Z}$$

Comme en dimension 1, la phase de transport de la méthode de transport-écroulement se traduit par le mouvement horizontal des particules.

Après déplacement, les nouvelles positions sont :

$$(3.17) \quad \begin{aligned} x_1(i) &= x_0(i) + \delta t f'(w_{\text{ref}} + k_0(i) \delta w) \in \mathbb{R}, \text{ suivant l'axe des } x, \\ j_1(i) &= j_0(i) + \delta j(k_0(i)) \in \mathbb{Z}, \text{ suivant l'axe des } y. \end{aligned}$$

Du fait de l'hypothèse (3.4), les déplacements sont donc discrets dans la direction des  $y$ .

e) Phase d'écroulement : déplacement vertical des particules

L'écroulement se traduit, exactement comme en dimension 1, par le mouvement vertical des particules, suivant l'axe  $w$ .

La transformée  $u_1 = T(\delta t)u_0$  est déterminée par la seule donnée des  $x_1(i)$ ,  $j_1(i)$  et  $s(i)$  exactement comme l'était  $u_0$  par celle des  $x_0(i)$ ,  $j_0(i)$  et  $s(i)$ , suivant la formule, valable presque partout :

$$(3.18) \quad u_1(x,y) = \delta w \sum_i s(i)$$

où les  $i$  sont tels que :  $x_1(i) < x$  et  $j_1(i) = E\left(\frac{y-y_{\text{ref}}}{\delta y}\right)$

La nouvelle position  $k_1(i)$  des particules, suivant l'axe  $w$ , en fonction des  $x_1(i)$ ,  $j_1(i)$  et  $s(i)$  peut donc être calculée à l'aide de l'algorithme introduit en b).

Connaissant les  $x_1(i)$ ,  $j_1(i)$  et  $s(i)$ , on peut recommencer les processus et calculer  $u_2 = T(\delta t)u_1$  etc...

### 3 - COUT DE L'ALGORITHME BIDIMENSIONNEL

Si l'on se réfère à la description du paragraphe 2, on peut évaluer le coût du schéma bidimensionnel en fonction des paramètres  $\delta t$ ,  $\delta w$  et  $\delta y$  destinés à tendre vers 0 (et supposés "adimensionnels").

(3.19)

La quantité d'éléments à stocker est de l'ordre de :

$$(\delta y \delta w)^{-1}$$

Il faut en effet stocker au minimum les quatre tableaux  $s(i)$ ,  $x(i)$ ,  $j(i)$ ,  $in(j,m)$  (1 de type "logique", 1 réel et 2 entiers), qui sont de longueur :

$$N = TV_x(u_0) \cdot (\delta y \delta w)^{-1}$$

Le stockage des  $k_0(i)$  n'est pas indispensable et les autres tableaux requis sont de longueur négligeable devant  $N$ .

(3.20)

Le nombre d'opérations élémentaires à effectuer est de l'ordre de :

$$(\delta t \delta y \delta w)^{-1} \cdot (-\log_2 \delta w)$$

Le facteur  $(\delta t)^{-1}$  provient de la discrétisation en temps.

A chaque pas de temps  $n$ , l'opération la plus coûteuse est le calcul des  $k_n(i)$  à partir des  $s_n(i)$ ,  $x_n(j)$  et  $j_n(i)$ .

On en connaît le coût (cf. § 2c) :

$$\sum_j m_n(j) \text{Log}_2 m_n(j) ,$$

où  $m_n(j)$  est le nombre de particules  $i$  telles que  $j_n(i) = j$ .

Comme  $\delta w \cdot m_n(j)$  est exactement la variation totale en  $x$  de la restriction de  $u_n(x,y)$  à l'axe :  $y = y_{\text{ref}} + j \cdot \delta y$ , on peut supposer, grossièrement, que  $m_n(j)$  est soit nul, soit de l'ordre de  $(\delta w)^{-1}$ .

Le coût est donc de l'ordre de :

$$-\text{Log}_2 \delta w \cdot \sum_j m_n(j) \text{ soit } -(\delta y \delta w)^{-1} \text{Log}_2 \delta w$$

puisque :  $\sum_j m_n(j) = N \sim (\delta y \delta w)^{-1}$

Il faut noter que les opérations élémentaires sont de l'un des types suivants : transferts de mémoires, tests entre réels ou additions.

Il n'y a donc ni multiplications, ni divisions, ni appels de fonctions, sauf à l'initialisation, pour calculer une fois pour toutes les  $\delta t \cdot f'(k\delta w)$  et  $\delta j(k)$ ,  $k \in \mathbb{Z}$  (cf. § 2d)

Rappelons enfin, qu'on doit vérifier les conditions de convergence :

$$(3.21) \quad \boxed{\begin{array}{l} \delta y \ll \delta t \ll 1 \\ \delta w \ll 1 \end{array}} \quad (\text{cf. § 1})$$

4 - COMPARAISON AVEC UN SCHEMA AUX DIFFERENCES FINIESa) Choix du schéma

Notre algorithme s'applique à la résolution du problème de Cauchy dans  $\mathbb{R}^2$  :

$$(3.22) \quad \begin{cases} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) + \frac{\partial}{\partial y} g(u) = 0 & , \quad u \in \mathbb{R}, (x,y) \in \mathbb{R}^2, t > 0 \\ u(t=0, x,y) = u_0(x,y) \end{cases}$$

Ce problème, relativement simple (absence de conditions aux limites, indépendance de  $f$  et  $g$  par rapport aux variables d'espace), peut être traité aisément, moyennant la technique des directions alternées ("splitting"), à l'aide d'un schéma aux différences finies tel que le schéma décentré ("upwind") unidimensionnel. Ce schéma consiste à discrétiser l'équation à une dimension d'espace :  $s_t + (h(s))_x = 0$  lorsque la fonction  $h$  est monotone croissante, de la façon suivante :

$$s_i^{n+1} = s_i^n - \frac{\delta t}{\delta x} \cdot \{h(s_i^n) - h(s_{i-1}^n)\} ,$$

où  $i$  et  $n$  désignent les indices de discrétisation en espace et en temps,  $\delta x$  et  $\delta t$  notant les pas d'espace et de temps.

Dans le cas de l'équation (3.22)  $f$  et  $g$  ne sont pas forcément monotones croissantes, mais on peut se ramener à ce cas de la manière suivante :

On note :

$$(3.23) \quad I = [\inf_{x,y} u_0(x,y) ; \sup_{x,y} u_0(x,y)] \subset \mathbb{R}$$

$$a = \text{ess inf}\{f'(w) ; w \in I\} ; b = \text{ess inf}\{g'(w) ; w \in I\}$$

( $f$  et  $g$  étant supposées lipschitziennes de  $I$  dans  $\mathbb{R}$ ), il est facile de vérifier que la fonction :

$$(3.24) \quad \hat{u}(t,x,y) = u(t,x+at,y+bt) \quad (*)$$

est solution de l'équation :

$$(3.25) \quad \frac{\partial \hat{u}}{\partial t} + \frac{\partial}{\partial x} \hat{f}(\hat{u}) + \frac{\partial}{\partial y} \hat{g}(\hat{u}) = 0$$

où  $\hat{f}$  et  $\hat{g}$  sont monotones croissantes sur  $I$ , définies par :

$$(3.26) \quad \hat{f}(w) = f(w) - aw \quad ; \quad \hat{g}(w) = g(w) - bw$$

C'est donc  $\hat{u}$  (et non  $u$ ) que nous calculerons à l'aide du schéma décentré unidimensionnel, par la technique du "splitting" en espace.

Le problème (3.22) est ainsi particulièrement favorable aux différences finies, puisqu'on peut le résoudre très élémentairement à l'aide du seul schéma décentré unidimensionnel : cela se traduit concrètement par une grande économie de moyens informatique.

Il est facile d'évaluer les qualités et le coût du schéma aux différences finies ainsi obtenu :

i) il est conservatif (malgré le splitting, en raison de la non-dépendance en  $x, y$  de  $f$  et  $g$ ) et respecte le principe du maximum ;

ii) la "place mémoire" est de l'ordre de :

$$(3.27) \quad \boxed{(\delta x)^{-2} \text{ éléments à stocker}}$$

---

(\*) il ne s'agit que d'un changement de repère qui ne modifie en rien le caractère entropique des solutions.

iii) le volume des calculs est de l'ordre de :

$$(3.28) \quad (\delta x)^{-3} \text{ opérations élémentaires}$$

ceci compte tenu de la condition de Courant-Friedrichs-Lewy :

$$(3.29) \quad Q\delta t \leq \delta x$$

$$\text{où : } Q = \text{Max} \left\{ \underset{I}{\text{ess sup } f'} - \underset{I}{\text{ess inf } f'} ; \underset{I}{\text{ess sup } g'} - \underset{I}{\text{ess inf } g'} \right\}$$

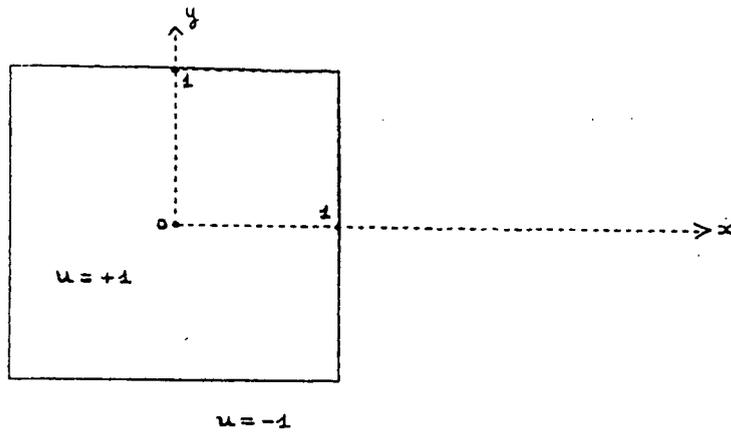
b) Résultats numériques

- Pour comparer notre méthode au schéma de différences finies, nous considérons le problème (3.22) avec la donnée initiale :

$$(3.30) \quad u_0(x,y) = \begin{cases} +1 & \text{si } |x| \leq 1 \text{ et } |y| \leq 1 \\ -1 & \text{sinon} \end{cases} \quad (\text{voir figure (3.7)})$$

et nous cherchons à en calculer la solution au temps  $t = 2$ , pour divers jeux de données  $f, g$  :

$$\left. \begin{array}{l} \text{cas 1 : } f(u) = \frac{u^2}{2} \\ \text{cas 2 : } f(u) = u \left( \frac{u^2 - 1}{2} \right) \\ \text{cas 3 : } f(u) = \frac{(u_+)^2}{2} \\ \text{cas 4 : } f(u) = \frac{u^4}{4} \end{array} \right\} g(u) = \frac{u^2}{2}$$



donnée initiale

figure 3.7

REMARQUE : La solution du cas numéro 1 peut être calculée à la main.

Il s'agit en effet d'un cas pseudo-bidimensionnel : on peut se ramener par rotation de  $45^\circ$  à l'équation de Burgers unidimensionnelle, qu'il est facile d'intégrer.

Moyennant quelques calculs, on obtient la solution explicite du problème de Cauchy, par exemple pour  $t \geq 2$  :

$$u(x,y,t) = \begin{cases} -1 + 2 \cdot \frac{\eta - a(\xi,t)}{b(\xi,t) - a(\xi,t)} \cdot v(\xi,t) & \text{pour } |\xi| \leq 1 \text{ et } a(\xi,t) \leq \eta \leq b(\xi,t) \\ -1 & \text{ailleurs} \end{cases}$$

$$\text{où } \xi = \frac{x - y}{2}$$

$$\eta = \frac{x + y}{2}$$

$$a(\xi,t) = -1 + |\xi| - t$$

$$b(\xi,t) = -1 + |\xi| - t + 2\sqrt{2t(1 - |\xi|)}$$

$$v(\xi,t) = \sqrt{\frac{2(1 - |\xi|)}{t}}$$

- Le calcul par différences finies est effectué sur une grille carrée de 100 mailles de côté, avec pour pas d'espace  $\delta x = 0.04$ . La condition (3.29) détermine le pas de temps maximal (et optimal), le même dans les quatre cas :  $\delta t = 0.02$ .

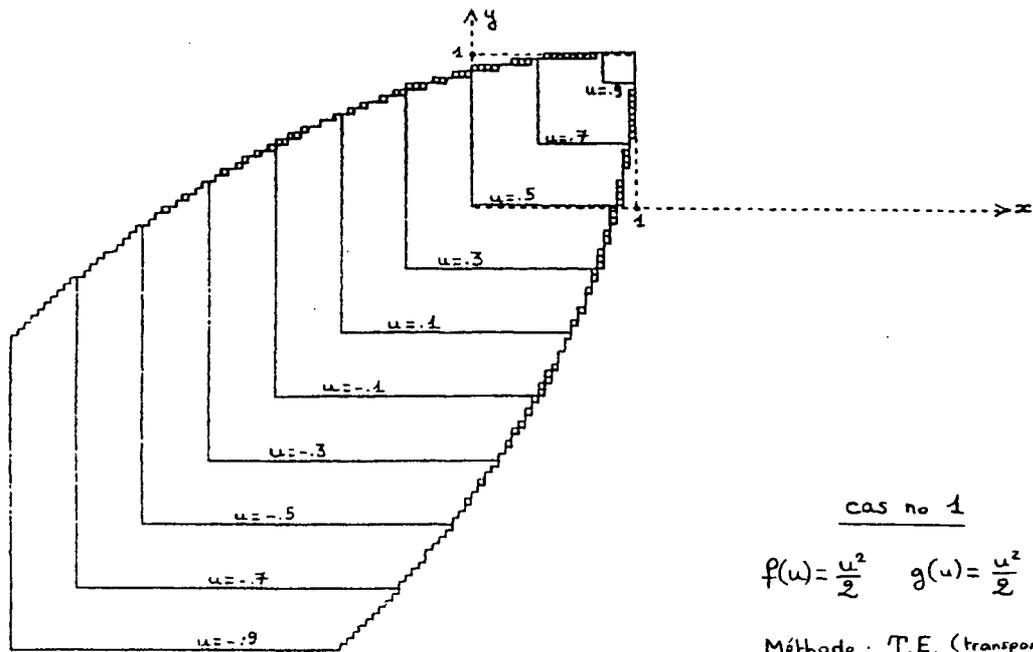
Il faut donc 100 pas de temps pour atteindre  $t = 2$  et 10 000 degrés de liberté pour représenter la solution.

En comparaison, on utilise pour la méthode de transport-écroulement les pas de discrétisation suivants :

$\delta w = 0.2$  (c'est-à-dire 10 mailles sur l'axe des  $u$ ),  $\delta y = 0.02$  et  $\delta t = 0.2$ . Il suffit donc de 10 pas de temps pour atteindre  $t = 2$  et il faut 2000 degrés de liberté pour représenter la solution.

On pourra comparer les résultats obtenus dans les quatre cas considérés sur les figures (3.8) à (3.11), où sont dessinées les lignes de niveau  $u = -0.9, -0.7, -0.5, -0.3, -0.1, +0.1, +0.3, +0.5, +0.7, +0.9$ , des différentes solutions approchées (transport-écroulement en haut, différences finies en bas).

Si la précision des calculs est comparable, le coût est très supérieur dans le cas des différences finies (le rapport est de 1 à 15 !), essentiellement à cause de la condition de Courant-Friedrichs-Lewy qui impose 100 pas de temps là où la méthode de transport-écroulement se contente de 10.



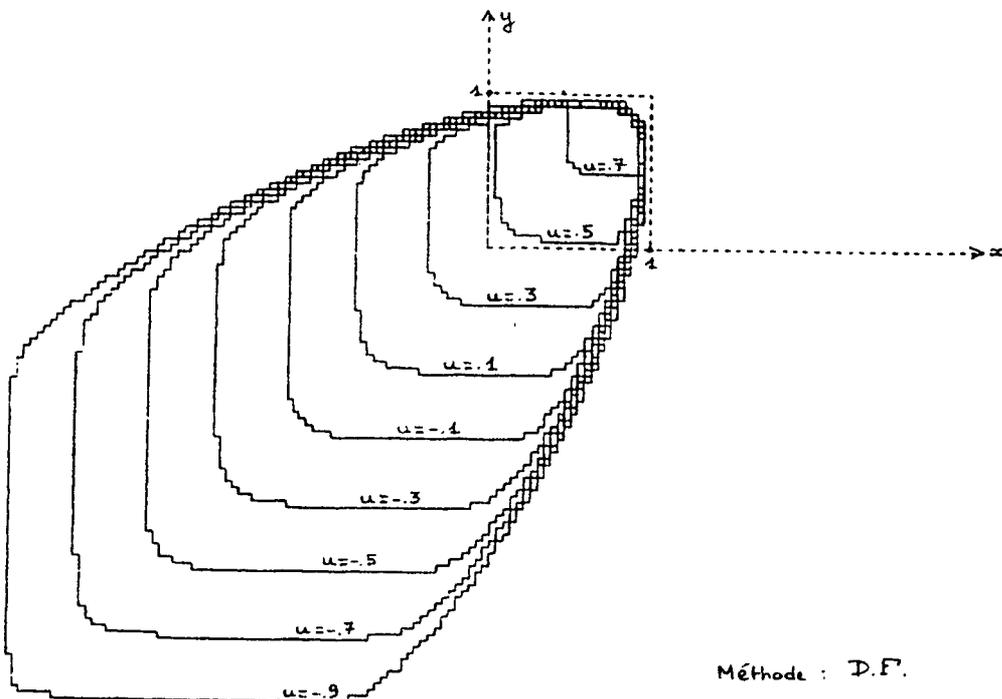
cas no 1

$$f(u) = \frac{u^2}{2} \quad g(u) = \frac{u^2}{2}$$

Méthode : T.E. (transport-écroulement)

10 pas de temps  
2000 degrés de liberté

coût : 5 s.

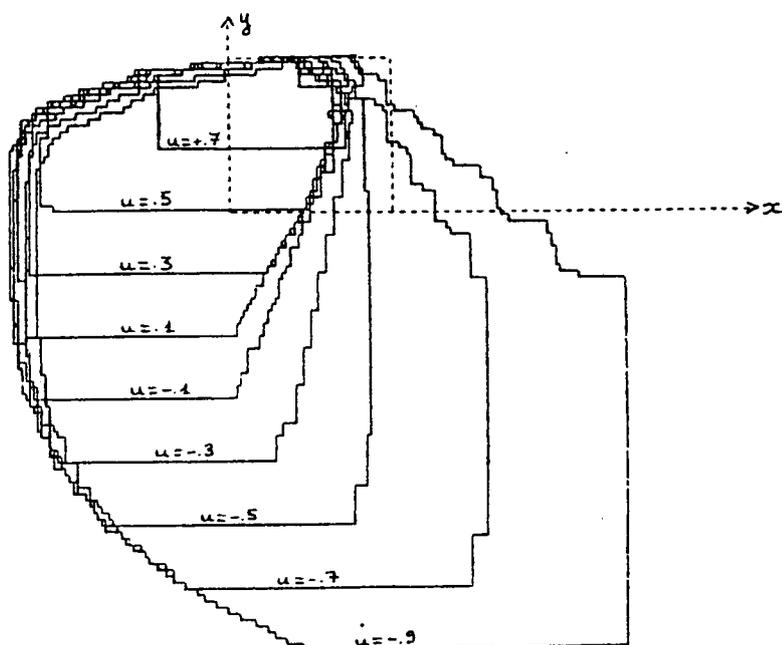


Méthode : D.F.

100 pas de temps  
10000 degrés de liberté

coût : 72 s.

Figure 3.8



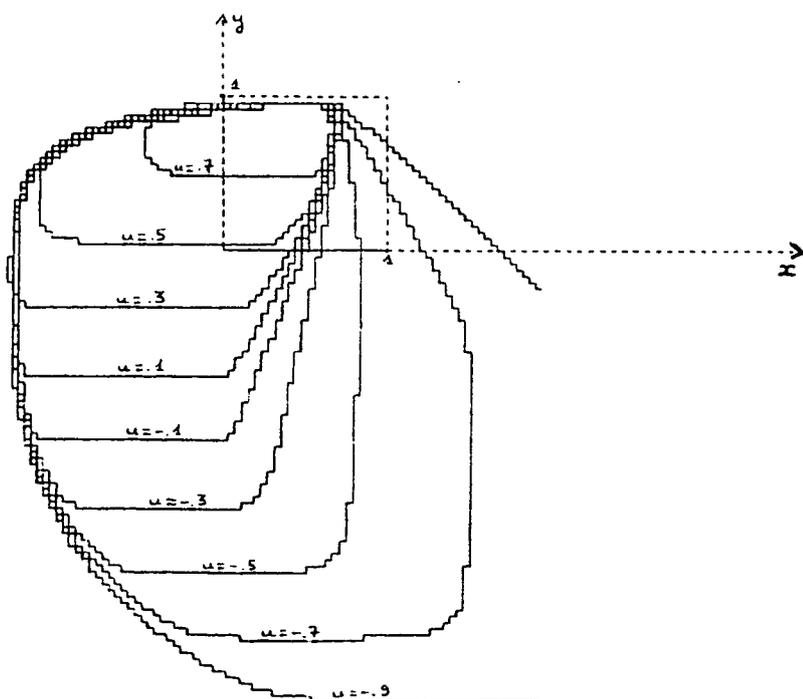
cas no 2

$$f(u) = u \left( \frac{u^2}{2} - \frac{1}{2} \right) \quad g(u) = \frac{u^2}{2}$$

Méthode : T.E.

10 pas de temps  
2000 degrés de liberté

coût : 5 s.

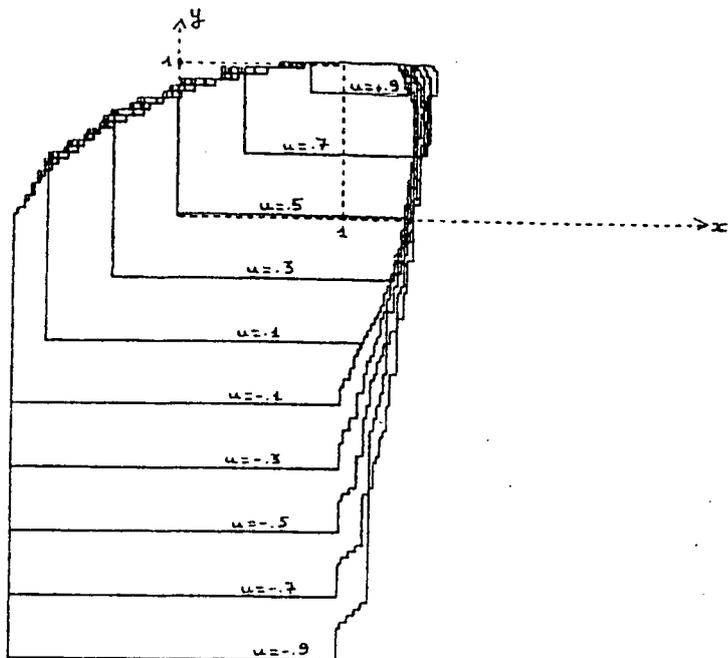


Méthode : D.F.

100 pas de temps  
10000 degrés de liberté

coût : 75 s.

Figure 3.9



cas no 3

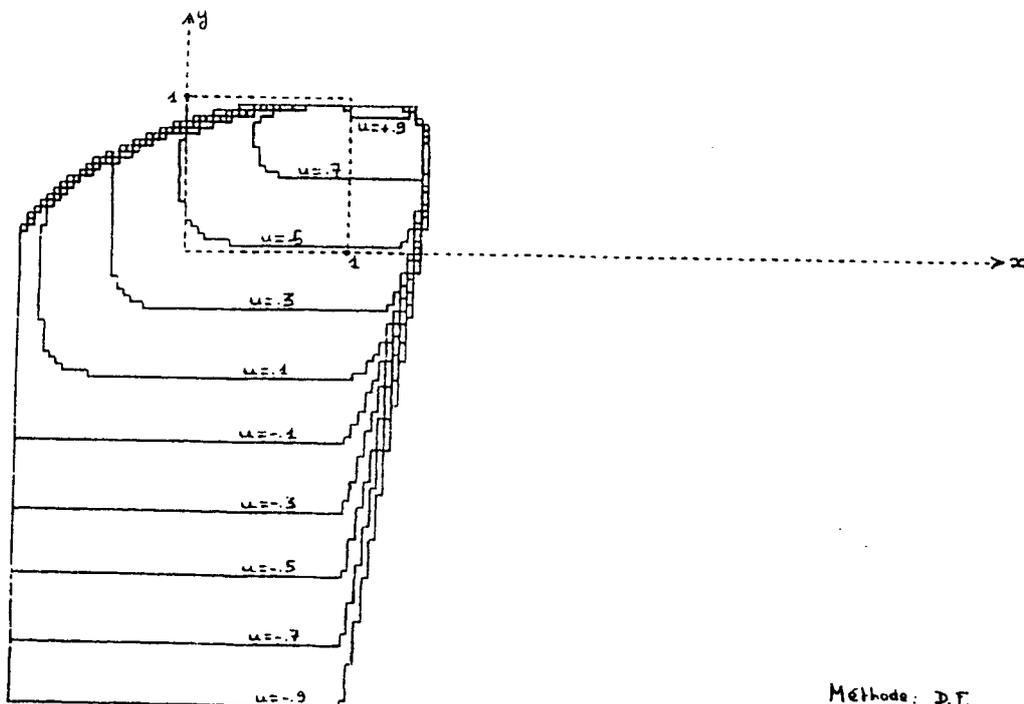
$$f(u) = \frac{(u_x)^2}{2} \quad g(u) = \frac{u^2}{2}$$

Méthode : T.E.

10 pas de temps

2000 degrés de liberté

coût : 5 s.



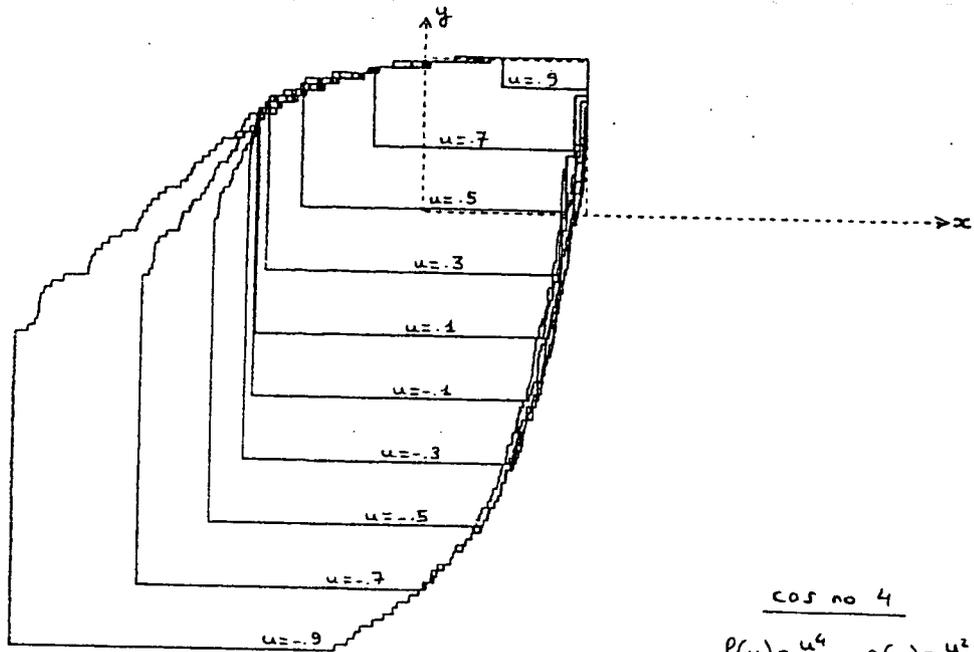
Méthode : D.F.

100 pas de temps

10000 degrés de liberté

coût : 74 s.

Figure 3.10



cas no 4

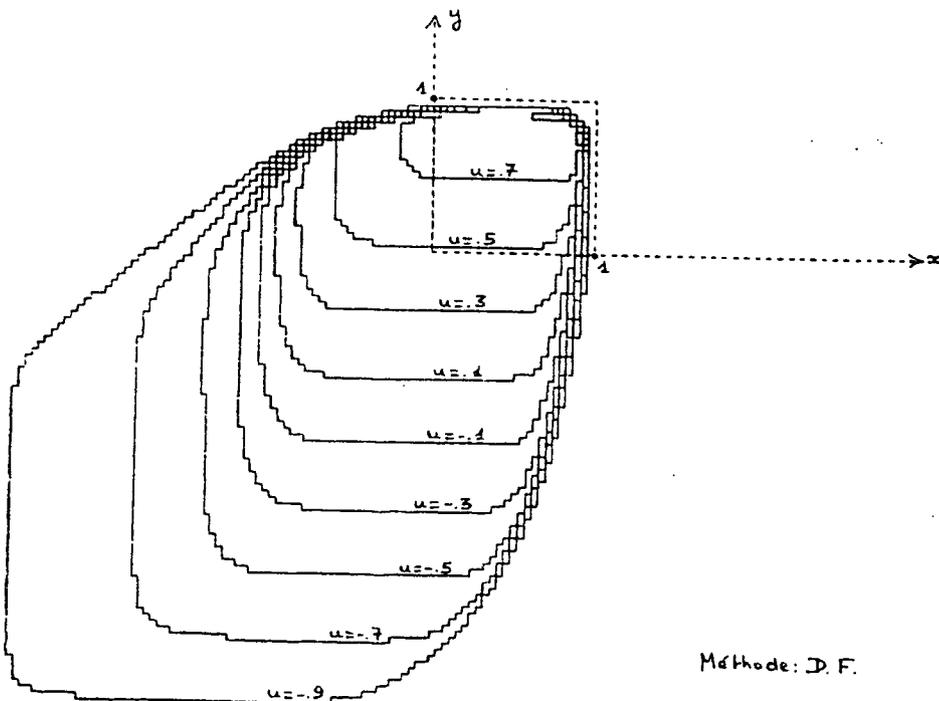
$$f(u) = \frac{u^4}{4} \quad g(u) = \frac{u^2}{2}$$

Méthode: TE

10 pas de temps

2000 degrés de liberté

coût: 5 s.



Méthode: D. F.

100 pas de temps

10000 degrés de liberté

coût: 77 s.

Figure 3.11

Pour évaluer plus en détail la qualité et le coût des deux méthodes, nous avons reconsidéré le cas numéro 4 en jouant sur les divers pas de discrétisation, de façon à obtenir pour les deux méthodes des temps de calcul comparables, variant de 5 à 80 s. (\*)

Le résultat de la comparaison est résumé dans les tableaux suivants :

DIFFERENCES FINIES	DFa	DFb	DFc
Pas d'espace $\delta x$	0.08	0.0615	0.04
Nombre de pas de temps	50	65	100
Nombre de degrés de liberté	2500	4225	10 000
Figure correspondante	(3.12)	(3.14)	(3.16)
Temps calcul (*)	10	22	77

TRANSPORT-ECROULEMENT	TEa	TEb	TEc
Pas d'espace $\delta y$	0.02	0.01	0.005
Nombre de pas de temps	10	20	40
Nombre de mailles en u	10	10	10
Nombre de degrés de liberté	2000	4000	8000
Figure correspondante	(3.13)	(3.15)	(3.17)
Temps calcul (*)	5	16	61

(\*) l'unité est la seconde de temps CPU sur CII-HB DPS68.

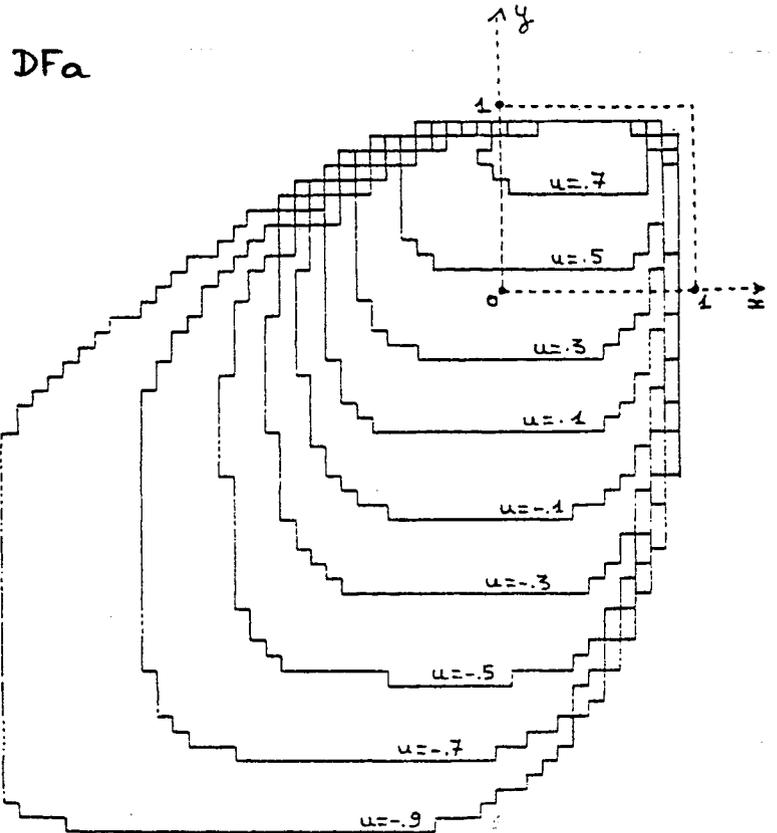


Figure 3.12

Différences Finies

50 pas de temps

2500 degrés de liberté

coût: 10s. HB-DP568

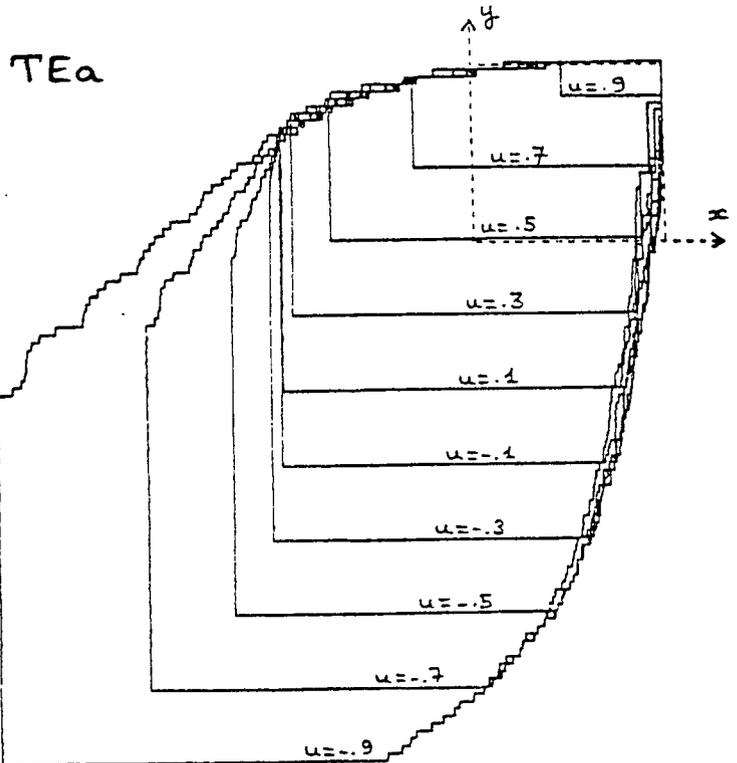


Figure 3.13

Transport - Ecroutement

10 pas de temps

2000 degrés de liberté

coût: 5s. HB-DP568

DFB

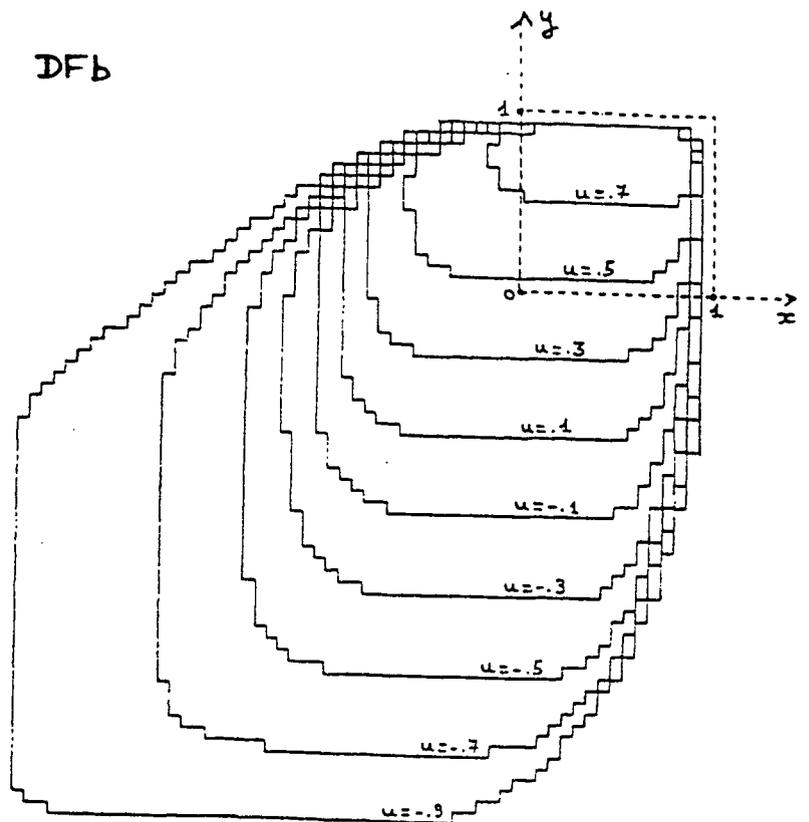


Figure 3.14

Différences Finies

65 pas de temps

4225 degrés de liberté

coût: 22s. MB-DPSC8

TEB

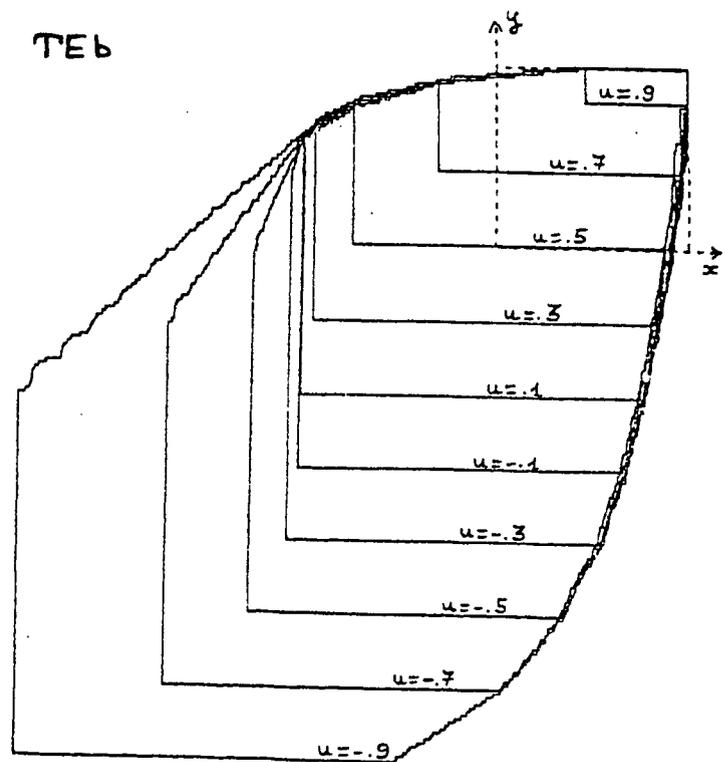


Figure 3.15

Transport-Ecroulement

20 pas de temps

4000 degrés de liberté

coût: 16s MB-DPSC8

DFc

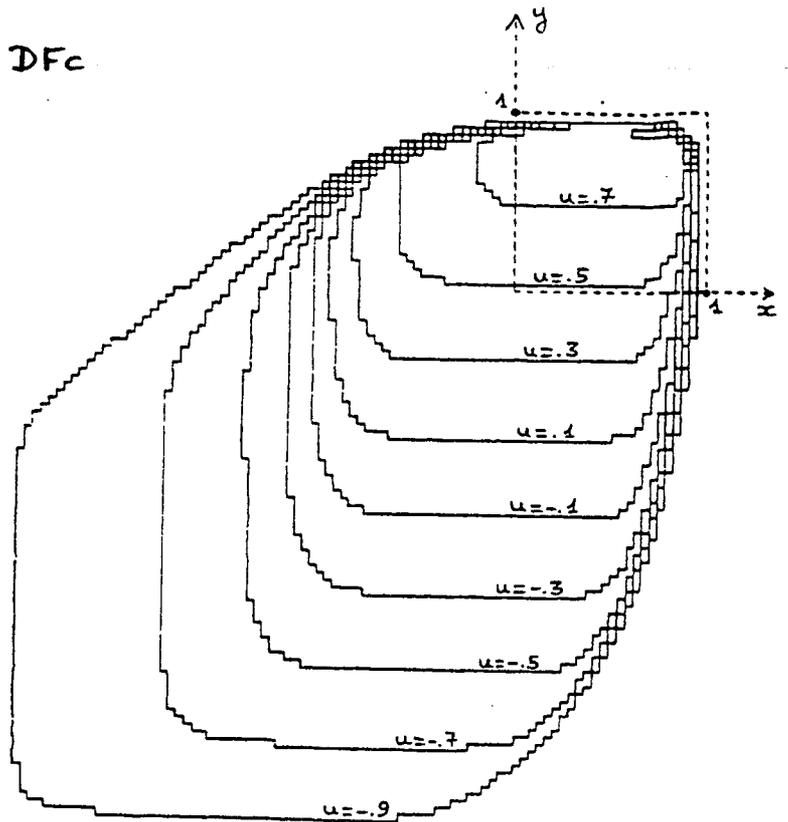


Figure 3.16

Différences Finies

100 pas de temps

10000 degrés de liberté

coût: 77s. HB-PP68

TEc

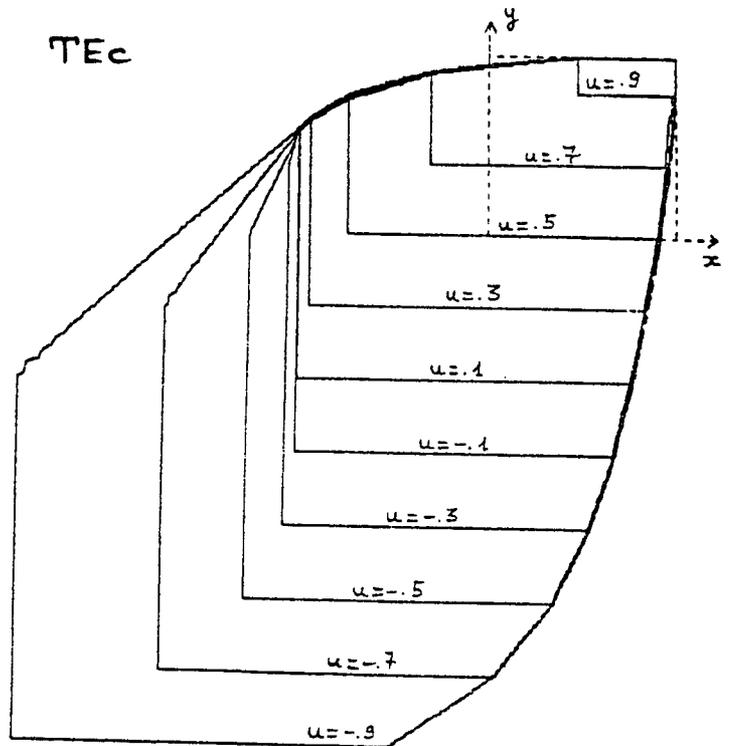


Figure 3.17

Transport-Écoulement

40 pas de temps

8000 degrés de liberté

coût: 61s. HB-PP68

c) Conclusion

La comparaison que nous avons effectuée montre que la méthode de transport-écroulement est très efficace pour la résolution du problème (3.22) : elle donne de bien meilleurs résultats que le schéma "décentré" pour un coût sensiblement moindre.

Il faut cependant rappeler que le problème (3.22) est très largement un cas d'école. L'adaptation de la méthode à des cas plus complexes et plus proches des applications (dépendance de  $f$  et  $g$  par rapport aux variables d'espace, conditions aux limites...) reste entièrement à faire et promet beaucoup de difficultés. La souplesse des différences finies (ou éléments finis) reste, à cet égard, bien plus grande.

## CHAPITRE IV

FORMALISATIONS ET GENERALISATIONS DE LA

METHODE DE TRANSPORT-ECROULEMENT

1 - PREMIERE GENERALISATION : LE CAS OU F DEPEND DE x

• Reportons-nous à la construction de l'opérateur  $T(\delta t)$ , au chapitre I (§ 1 et 3).

A partir du sous-graphe :

$$(4.1) \quad A = \{(x, w) \in \mathbb{R}^N \times \mathbb{R} ; u(x) \geq w\}$$

nous avons construit, par translation des tranches  $A_w$ , le nouvel ensemble

$$B = \{(x, w) \in \mathbb{R}^N \times \mathbb{R} ; x - \delta t F'(w) \in A_w\}$$

En appliquant à B l'opération d'écroutement, on obtenait ensuite l'ensemble C dont le profil définit la transformée  $T(\delta t)u$ .

• Nous allons maintenant donner une première généralisation de cette construction, pour l'équation :

$$(4.2) \quad \frac{\partial u}{\partial t} + \sum_{i=1}^N \frac{\partial}{\partial x_i} (F_i(u, x)) = 0 ; \quad u \in \mathbb{R}, x \in \mathbb{R}^N, t > 0$$

On introduit, pour chaque  $w \in \mathbb{R}$ , le champ de vecteurs sur  $\mathbb{R}^N$

$$(4.3) \quad Q(w; x) = \frac{\partial F}{\partial u}(w, x), \text{ qu'on supposera à divergence nulle,}$$

et le flot associé  $g(w; x, t)$  défini par =

$$(4.4) \quad \frac{\partial}{\partial t} g(w; x, t) = Q(w; g(w; x, t)) \text{ et } g(w; x, 0) = x$$

Nous définissons alors B à partir de A, comme suit :

$$(4.5) \quad B = \{(x, w) ; g(w; x, -\delta t) \in A_w\}$$

et nous ne changeons rien d'autre à la construction de  $T(\delta t)$ .

Dans le cas où  $F$  est indépendant de  $x$ , on a :

$$Q(w;x) = F'(w) \text{ et } g(w;x,t) = x + t F'(w) ;$$

On retrouve donc la construction précédente.

• On peut justifier cette généralisation en prouvant, sous des hypothèses convenables, la convergence de la méthode. C'est d'ailleurs sous cette forme que la méthode de transport-écroulement est exposée dans [2].

## 2 - UNE PREMIERE FORMULATION "FONCTIONNELLE"

Jusqu'à maintenant, nous avons décrit la méthode de manière très géométrique, en opérant des transformations sur les parties de  $\mathbb{R}^N \times \mathbb{R}$ . Il est possible d'en donner une formulation n'utilisant que le concept de fonctions.

Ainsi, nous remplaçons  $A$  par sa fonction caractéristique  $\chi$  :

$$(4.6) \quad \chi(x,w) = \begin{cases} 1 & \text{si } u(x) \geq w \\ 0 & \text{sinon} \end{cases}$$

Ensuite, plutôt que de considérer le flot  $g(w;x,t)$  défini par (4.4), nous utilisons l'opérateur d'évolution linéaire :

$$(4.7) \quad \frac{\partial}{\partial t} + Q(w,x) \cdot \text{grad} ,$$

et le semi-groupe associé (qui est en l'occurrence un groupe !) :  $S_w(t)$ .

Il est alors facile de voir que :

i) la fonction de deux variables :

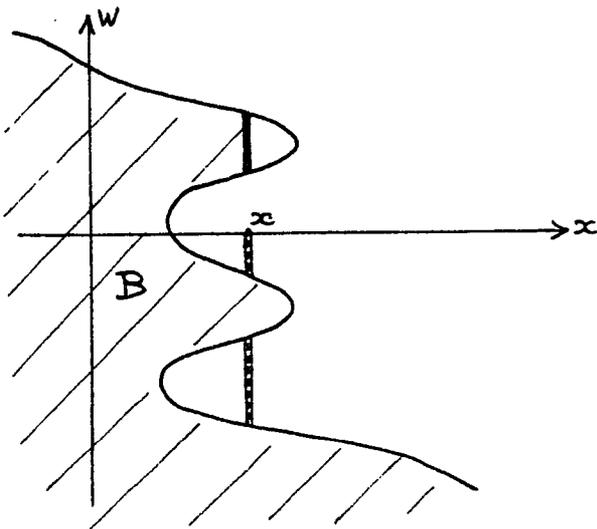
$$(4.8) \quad \psi(x,w) = (S_w(\delta t)\chi(\cdot, w))(x),$$

n'est rien d'autre que la fonction caractéristique de B ;

ii) l'opération d'écroutement se traduit par la formule :

$$(4.9) \quad (T(\delta t)u)(x) = \int_{-\infty}^{+\infty} s(x,w)dw \text{ où : } s(x,w) = \begin{cases} \psi(x,w) & \text{si } w \geq 0 \\ -(1-\psi(x,w)) & \text{sinon} \end{cases}$$

(voir figure 4.1).



La valeur de  $T(\delta t)$  au point  $x$  est donnée par  $S_1 - S_2$  où  $S_1$  (resp.  $S_2$ ) est la somme des longueurs des traits tracés en gras (resp. en hachures) à la verticale de  $x$ .

Figure 4.1

Figure 4.1

Cette présentation rejoint en gros celle donnée dans [2].

### 3 - UNE SECONDE FORMULATION "FONCTIONNELLE"

Donnons maintenant la formulation la plus pratique pour les démonstrations, celle utilisée dans [3].

Reprenons les formules du paragraphe précédent. Nous avons :

$$(T(\delta t)u)(x) = \int_{-\infty}^{+\infty} s(w,x)dx$$

où :

$$s(w,x) = \begin{cases} \psi(x,w) & \text{si } w \geq 0 \\ -(1-\psi(x,w)) & \text{sinon} \end{cases}$$

et : 
$$\psi(x,w) = (S_w(\delta t) \chi(\cdot, w))(x)$$

Comme  $S_w(\delta t)$  est linéaire, on peut donc écrire :

$$s(w,x) = (S_w(\delta t)c(\cdot, w))(x) \text{ où : } c(x,w) = \begin{cases} \chi(x,w) & \text{si } w \geq 0 \\ -(1-\chi(x,w)) & \text{sinon} \end{cases}$$

Ceci nous amène à introduire :

$$(4.10) \quad ju(x,w) = \begin{cases} 1 & \text{si } 0 < w < u(x) \\ -1 & \text{si } 0 > w > u(x) \\ 0 & \text{sinon} \end{cases}$$

et à poser :

$$(4.11) \quad T(\delta t)u(x) = \int_{-\infty}^{+\infty} (S_w(\delta t)ju(\cdot, w))(x)dw ,$$

(On a en effet  $c(x,w) = ju(x,w)$  p.p. dans  $\mathbb{R}^N \times \mathbb{R}$ ).

Cette formulation présente l'avantage de décrire toute la méthode avec les deux formules (4.11) et (4.10) !

#### 4 - DEUXIEME GENERALISATION : LES EQUATIONS PARABOLIQUES SEMI-LINEAIRES

En utilisant les notations du paragraphe précédent (4.11 et 4.10), et en considérant le semi-groupe  $S_w(t)$  associé à l'opérateur d'évolution linéaire :

$$(4.12) \quad \frac{\partial}{\partial t} + B'(w)\text{grad} - A'(w)\Delta, \text{ ou } A'(w) \text{ est supposé positif ou nul,}$$

on dispose, avec la formule (4.11), d'une semi-discrétisation de l'équation parabolique semi-linéaire, éventuellement dégénérée :

$$(4.13) \quad \frac{\partial u}{\partial t} + \text{div}(B(u) - \text{grad} A(u)) = 0 ,$$

où A et B sont supposés assez régulières, avec la condition :

$$(4.14) \quad A'(w) \geq 0 \quad , \quad \forall w \in \mathbb{R}$$

Les lois de conservation scalaire n'en sont qu'un cas particulier  
(A = 0 et B = F).

La démonstration de la convergence du procédé vers les solutions au sens de Volpert (cf. [13]) fait l'objet de [3].

### 5 - TROISIEME FORMULATION ET TROISIEME GENERALISATION

Ce paragraphe est tout à fait indépendant des précédents.

Reprenons de nouveau la construction de l'opération  $T(\delta t)$  au chapitre 1, où l'on construisait successivement les ensembles A, B et C. La première phase (dite de transport) peut se traduire comme suit :

i) on paramètre le graphe de  $u$  par des applications régulières ( $C^1$ ) :

$$(4.15) \quad s \in \mathbb{R}^N \rightarrow (X(s), U(s))$$

(ce qui semble possible même quand  $u$  présente des discontinuités : c'est le cas de la fonction d'Heavyside, pour  $N = 1$ ).

ii) on pose :

$$(4.16) \quad X_{\delta t}(s) = X(s) + \delta t Q(U(s)) \quad (\text{où } Q = F').$$

$s \in \mathbb{R}^N \rightarrow (X_{\delta t}(s), U(s))$  n'est alors rien d'autre qu'un paramétrage du profil délimité par B (cf. Chapitre 1, § 1).

L'opération d'éroulement se traduit ensuite par la formule :

$$(4.17) \quad (T_{\delta t}u)(x) = \sum_s U(s) \operatorname{sgn}(\det DX_{\delta t}(s))$$

où la somme s'effectue sur l'ensemble (génériquement fini) des  $s$  tels que  $X_{\delta t}(s) = x$ , et où  $DX_{\delta t}(s)$  désigne la matrice jacobienne de  $X_{\delta t}$  au point  $s$ .

L'intérêt de cette formulation, outre son allure "topologique" (\*), vient de ce qu'elle n'est pas liée au caractère scalaire de l'équation. Elle permettrait -mais c'est encore pure conjecture- l'approximation d'équations vectorielles du type :

$$(4.18) \quad \left( \frac{\partial}{\partial t} + \sum_{j=1}^N Q_j(u) \frac{\partial}{\partial x_j} \right) u_i = 0 \quad , \quad i = 1, \dots, M$$

où  $Q$  est donnée, de  $\mathbb{R}^M$  dans  $\mathbb{R}^N$ , et  $u$  recherchée, de  $\mathbb{R}^N$  dans  $\mathbb{R}^M$ .

Ce type d'équations (dont la formulation n'est pas très claire, quand les solutions ne sont plus régulières) semble lié d'une part avec les équations de la Mécanique des Fluides (Euler, Navier Stokes) et d'autre part avec l'équation d'Hamilton-Jacobi.

---

(\*) Rappelons que le degré topologique de  $X_{\delta t}$  par rapport à  $x$  est exactement :  $\sum_s \text{sgn}(\det DX_{\delta t}(s))$ , la somme s'effectuant sur les  $s$  tels que :  $X_{\delta t}(s) = x$ .

## CONCLUSION

De nombreux schémas couramment utilisés pour résoudre les équations hyperboliques non-linéaires -en particulier ceux de Godunov, de Glimm et toute la famille des schémas décentrés- reviennent en définitive, mais plus ou moins consciemment, à résoudre autant de problèmes de Riemann locaux qu'il y a de mailles en espace ; puis à en déduire, par des techniques variables (interpolation, projection  $L^2$  ou échantillonnage aléatoire), l'évolution de la solution autant qu'il est possible, c'est-à-dire tant que les problèmes locaux n'interagissent pas. La condition de Courant-Friedrichs-Lewy s'ensuit naturellement qui interdit toute interaction en limitant le pas de temps en fonction du maillage spatial.

Les schémas qu'on peut construire à partir de la méthode de transport-écroulement sont en revanche complètement affranchis de toute contrainte de cette nature. La méthode peut être donc qualifiée, sinon de "globale" -ce qu'elle n'est assurément pas, vu la nécessité d'introduire un pas de temps destiné à tendre vers 0-, du moins de "semi-globale", en ce sens que le pas de temps n'est pas limité par le choix de la discrétisation en espace.

L'avantage qui en résulte est double : on peut espérer ainsi réduire à la fois le coût des calculs et la diffusion numérique, comme le montre déjà les résultats numériques obtenus dans le cas de l'équation (1).

Une autre différence doit être notée entre les schémas usuels et ceux que nous avons présentés ici, concernant les techniques numériques utilisées de part et d'autre : contrairement aux premiers, les seconds n'utilisent ni interpolation, ni splitting (ni même de maillage en espace dans le cas unidimensionnel), et se ramènent, en définitive, à des problèmes de tris de nombre réels. Sur ce plan, il convient de les rapprocher de certains algorithmes proposés par A. Chorin dans [5] pour d'autres équations (problèmes de couche limite en

hydrodynamique, équation de la chaleur, équations de réaction-diffusion)<sup>(\*)</sup>, et dont l'étude a été déterminante pour notre propre réflexion.

Cette différence est néanmoins de moindre portée que la précédente, car elle semble liée à la structure somme toute assez simple de l'équation (1). Il apparaît indispensable pour aborder des équations plus complexes et plus proches des applications, d'adapter la méthode en utilisant les techniques plus classiques d'interpolation et d'éléments finis.

---

(\*) comparer notamment notre algorithme unidimensionnel à l'algorithme proposé par Chorin pour l'équation de la chaleur dans [5], § 2.

BIBLIOGRAPHIE

- [0] V.I. ARNOLD et A. AVEZ  
 "Problèmes ergodiques de la mécanique classique" - Gauthier-Villars - Paris 1967 - p. 33.
- [1] Cl. BARDOS, A.Y. LEROUX, J.Cl. NEDELEC  
 "First order quasilinear equations with boundary conditions" - Rapport interne n° 38 - Centre de Mathématiques Appliquées - Ecole Polytechnique - 1978.
- [2] Y. BRENIER  
 "Une application de la symétrisation de Steiner aux équations hyperboliques : la méthode de transport-écroulement" - A paraître aux C.R.A.S.
- [3] Y. BRENIER  
 "Approximation des solutions entropiques d'équations aux dérivées partielles quasilinéaires" - A paraître.
- [4] G. CHAVENT, G. SALZANO  
 "A finite element method for the one-dimensional water flooding problem with gravity" - Soumis au J. Comp. Phys.
- [5] A.J. CHORIN  
 "Numerical methods for use in combustion modeling" - Proceedings of the fourth international Symposium on computing methods in applied sciences and engineering, Versailles, dec. 1979 - (Glowinski, Lions Ed.), North-Holland, 1980.
- [6] P. CONCUS, W. PROSKUROWSKI  
 "Numerical solution of a non-linear hyperbolic equation by the random choice method" - J. Comp. Phys., 30 - 1979 - p. 153-166.
- [7] S.N. KRUKOV  
 "First order quasilinear equations with several space variables" - Math. U.S.S.R. SB.10 - 1970 - p. 217-243.
- [8] A.Y. LEROUX  
 Thèse d'Etat - Rennes - 1979.
- [9] C. MARLE  
 Cours de production - Tome IV - Technip - 1972.

- [10] O.A. OLEÏNIK  
"Discontinuous solution of non-linear differential equations" - Amer. Math. Soc. Transl. Ser. 2, 26 (1963) - p. 95-172.
- [11] G. POLYA, A. SZEGÖ  
"Isoperimetric inequalities in mathematical physics" - Princeton University Press - 1951.
- [12] A.I. VOLPERT  
"The space BV and quasilinear equations" - Math. U.S.S.R. Sb.2 - 1967 - p. 225-267.
- [13] A.I. VOLPERT  
"Cauchy's problem for degenerate second order quasilinear parabolic equations" - Math. U.S.S.R. Sb.7 - 1969 - p. 365-387.

