



**HAL**  
open science

# Consistance et stabilite des schemas LRG pour les lois de conservation scalaires

Guy Chavent, B. Cockburn

► **To cite this version:**

Guy Chavent, B. Cockburn. Consistance et stabilite des schemas LRG pour les lois de conservation scalaires. [Rapport de recherche] RR-0710, INRIA. 1987. inria-00075842

**HAL Id: inria-00075842**

**<https://inria.hal.science/inria-00075842>**

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**

UNITÉ DE RECHERCHE  
IRIA-ROCQUENCOURT

Institut National  
de Recherche  
en Informatique  
et en Automatique

Domaine de Voluceau  
Rocquencourt  
BP105  
78153 Le Chesnay Cedex  
France  
Tél.: (1) 39 63 55 11

Rapports de Recherche

N° 710

**CONSISTANCE ET STABILITE DES  
SCHEMAS LRG POUR LES  
LOIS DE CONSERVATION  
SCALAIRES**

**Guy CHAVENT  
Bernardo COCKBURN**

**JUILLET 1987**

**CONSISTANCE ET STABILITE DES SCHEMAS LRG POUR LES LOIS  
DE CONSERVATION SCALAIRES**

---

**CONSISTENCY AND STABILITY OF LRG SCHEMES FOR SCALAR  
CONSERVATION LAWS**

**Guy CHAVENT - Bernardo COCKBURN**

**Résumé :**

On s'intéresse à la résolution numérique de l'équation  $\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0$  par une famille de schémas utilisant la discrétisation en espace par éléments finis discontinus introduit par Lesaint et Raviart, avec le terme de flux de Godunov et une discrétisation en temps explicite par différences finies (schéma LRG).

On montre d'abord, dans le cas linéaire ( $f(u)=u$ ), que les schémas LRG sont consistants, mais on n'obtient la stabilité que sous une condition  $\frac{\Delta t}{h} \leq ch \frac{1}{2}$ .

On améliore la stabilité du schéma LRG en lui adjoignant une procédure de limitation de pente (schéma LRGP), et on étudie, par une méthode numérique, la stabilité, l'aptitude à rendre les chocs et les détente, et la convergence vers la solution entropique des schémas LRG, LRGP, et, à titre de comparaison, Godunov, MUSCL, B1P et G-1/2.

**Summary :**

We are interested in the numerical solution of  $\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} f(u) = 0$  by a family of schemes based on the discontinuous finite element space discretization introduced by Lesaint and Raviart, together with the flux term of Godunov, and a forward finite difference time discretization (LRG schemes).

We first show in the linear case ( $f(u) = u$ ) that LRG schemes are consistent, but that stability holds only under a condition  $\frac{\Delta t}{h} \leq ch \frac{1}{2}$ .

In order to enhance its stability, the LRG scheme is modified by the introduction of a slope limiter (LRGP schemes). Then we study, by numerical methods, the stability, the ability to track shocks and to image fans, and the convergence towards the entropy solution of LRG and LRGP schemes, and, for comparison purposes, Godunov, MUSCL, B1P and G-1/2.

**Mots clefs :**

- Lois de conservations scalaires - équations hyperboliques - Analyse Numérique - .

CONSISTANCE ET STABILITE

DES SCHEMAS LRG

G. Chavent - B. Cockburn

- 1 - LES SCHEMAS POUR L'APPROXIMATION DE  $\partial u / \partial t + \partial u / \partial x = 0$ 
  - 1.1 - Schéma LRG : semi-discrétisation en espace
  - 1.2 - Schéma LRG : discrétisation en temps
  - 1.3 - Deux schémas aux différences finies
  
- 2 - ETUDE DE LA CONSISTANCE DES SCHEMAS
  - 2.1 - Consistance de LRG pour le choix 1  $(a_i, b_i)$  des degrés de liberté
  - 2.2 - Consistance de LRG pour le choix 2  $(a_i, \xi_{i+1/2})$  des degrés de liberté
  - 2.3 - Consistance de LRG pour le choix 3  $(a_i, \tilde{b}_i)$  des degrés de liberté
  - 2.4 - Consistance de LRG pour le choix 4  $(\xi_{i+1/2}, b_i)$  des degrés de liberté
  - 2.5 - Consistance des deux schémas aux différences finies
  
- 3 - ETUDE DE LA STABILITE DES SCHEMAS
  - 3.1 - Introduction
  - 3.2 - Description des méthodes d'étude de la stabilité
    - 3.2.1 - La méthode de Von Neumann
    - 3.2.2 - Une méthode expérimentale numérique

- 3.3 - Etude de la stabilité dans le cas linéaire par la méthode de Von Neumann
    - 3.3.1 - Stabilité linéaire du schéma LRG- $\mu$
    - 3.3.2 - Stabilité linéaire de deux schémas DF
  - 3.4 - Etude numérique expérimentale de la stabilité des schémas LRG- $\mu$ 
    - 3.4.1 - Le cas linéaire
    - 3.4.2 - Le cas non-linéaire
  - 3.5 - Conclusion sur la stabilité du schéma LRG- $\mu$
- 4 - ETUDE NUMERIQUE COMPARATIVE DES PROPRIETES DE CONVERGENCE DU SCHEMA LRG ET DES SCHEMAS LRG-MODIFIES
- 4.1 - Introduction
  - 4.2 - Description de la méthode d'étude de la convergence
    - 4.2.1 - Présentation du schéma numérique
    - 4.2.2 - Description des cas-tests
    - 4.2.3 - Le protocole d'étude numérique
  - 4.3 - Le cas linéaire : une discontinuité de contact
    - 4.3.1 - Le schéma LRG
    - 4.3.2 - Les autres schémas
    - 4.3.3 - Résumé
  - 4.4 - Un cas non-linéaire : l'équation de Buckley-Leverett : un choc
    - 4.4.1 - Le schéma LRG
    - 4.4.2 - Les autres schémas
    - 4.4.3 - Résumé
  - 4.5 - Un cas non-linéaire : la fonction  $f$  est convexe et la solution  $\in C^1$   
 $\forall t > 0$  : une étude de détente
    - 4.5.1 - Le schéma LRG
    - 4.5.2 - Les autres schémas
    - 4.5.3 - Résumé
  - 4.6 - L'influence du paramètre d'antidiffusion  $\alpha$  sur la solution donnée par le schéma G-1/2
    - 4.6.1 - Introduction
    - 4.6.2 - Le cas linéaire : une discontinuité de contact
    - 4.6.3 - L'équation de Buckley-Leverett : un choc

4.6.4 - Le cas d'une onde de détente

4.6.5 - Conclusion

4.7 - Conclusion

ANNEXE 1.

REFERENCES.

#### REMERCIEMENTS

Le deuxième auteur remercie le Laboratorio de Computação Científica, Rio de Janeiro, pour l'accueil et les facilités reçues pour effectuer les expériences numériques du chapitre IV lors de son séjour en août 1984, ainsi que Jim DOUGLAS, Jr. pour le soutien reçu à l'University of Chicago, où le dit chapitre a été rédigé début 1985.

On considère le problème modèle :

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0 \quad \text{sur } [0, L] \times [0, T]$$

(0.1)  $u = u_e \quad \text{en } x = 0$

$u(x, 0) = u_0(x) \quad \text{à } t = 0 .$

On va étudier successivement la consistance et la stabilité du schéma dit "LRC" utilisé dans [ 2 ], ainsi que de quelques schémas aux différences finies. On proposera ensuite des schémas LRG-modifiés ayant de meilleures propriétés de stabilité, et on indiquera comment les étendre au cas bidimensionnel

$$\frac{\partial u}{\partial t} + \text{div}(u \vec{q}) = 0 \quad \text{dans } \Omega \subset \mathbb{R}^2, \quad t \in [0, T]$$

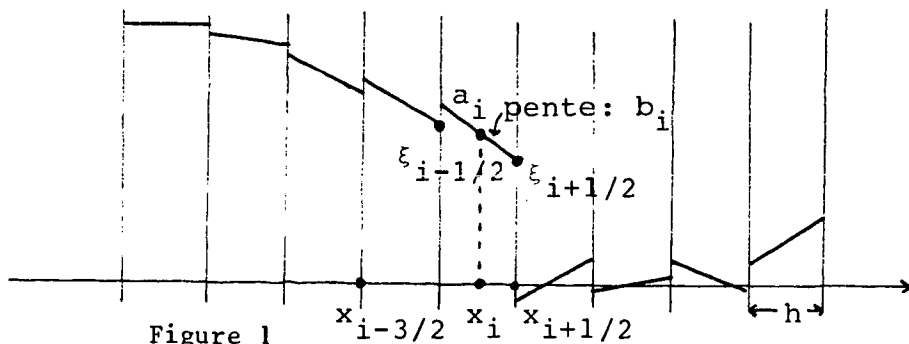
(0.2)  $u = u_e \quad \text{sur } \partial\Omega^-$

$u(x, 0) = u_0(x) \quad \text{à } t = 0 .$

## 1 - LES SCHEMAS POUR L'APPROXIMATION DE $\partial u / \partial t + \partial u / \partial x = 0$

### 1.1 - Schéma LRG : semi-discrétisation en espace

Suivant l'idée introduite par Raviart-Lesaint dans le cas stationnaire, on décide d'approcher  $u$  par (cf. figure 1) :



$$u(\cdot, t) \sim u_h(t) \in V_h = \{v_h \mid |v_h| \}_{x_{i-1/2}, x_{i+1/2}} \in \mathcal{P}^1 \quad i=1, \dots, N\}.$$

En  $x_{1/2}$  et en  $x_{i+1/2}$   $i=1, 2, \dots, N$ , on pose

$$(1.1) \quad \begin{aligned} \xi_{1/2} &= u_e \\ \xi_{i+1/2} &= u_h(x_{i+1/2}^-, t) \end{aligned} \quad (\text{valeurs "amont" de Godunov}).$$

Le schéma LRG s'écrit alors (cf. [ ]) :

$$(1.2) \quad \begin{aligned} &\int_{x_{i-1/2}}^{x_{i+1/2}} \frac{\partial u_h}{\partial t} v + [u_h(x_{i-1/2}^+, t) - (x_{i-1/2}^-, t)] v(x_{i-1/2}) \\ &+ \int_{x_{i-1/2}}^{x_{i+1/2}} \frac{\partial u_h}{\partial x} v = 0 \quad \forall i = 1, \dots, N \quad \forall v \in \mathcal{P}^1 \end{aligned}$$

$$u_h(0) = u_{0h}(x)$$

soit en intégrant par partie en utilisant les  $\xi_{i+1/2}$  :

$$(1.3) \quad \begin{aligned} &\int_{x_{i-1/2}}^{x_{i+1/2}} \frac{\partial u_h}{\partial t} v + \xi_{i+1/2} v(x_{i+1/2}) - \xi_{i-1/2} v(x_{i-1/2}) \\ &- \int_{x_{i-1/2}}^{x_{i+1/2}} u_h \frac{dv}{dx} = 0 \quad \forall i = 1, \dots, N \quad \forall v \in \mathcal{P}^1 \end{aligned}$$

$$u_h(0) = u_{0h}(x).$$

En prenant successivement  $v = \text{cste}$  et  $v = x - x_i$  sur  $]x_{i-1/2}, x_{i+1/2}[$ , il vient, en calculant exactement toutes les intégrales :



$$\frac{da_i}{dt} + \frac{\xi_{i+1/2} - \xi_{i-1/2}}{h} = 0$$

(1.4)

$$\frac{db_i}{dt} + \frac{3}{2} \frac{\xi_{i+1/2}^{-2a_i} + \xi_{i-1/2}}{(h/2)^2} = 0 .$$

En calculant toutes les intégrales par la méthode des trapèzes, il vient :

$$\frac{da_i}{dt} + \frac{\xi_{i+1/2} - \xi_{i-1/2}}{h} = 0$$

(1.5)

$$\frac{db_i}{dt} + \frac{1}{2} \frac{\xi_{i+1/2}^{-2a_i} + \xi_{i-1/2}}{(h/2)^2} = 0 .$$

Pour traiter d'un coup le système (1.4) et le système (1.5), on va considérer dans la suite le système

$$\frac{da_i}{dt} + \frac{\xi_{i+1/2} - \xi_{i-1/2}}{h} = 0$$

(1.6)

$$\frac{db_i}{dt} + \frac{3\mu}{2} \frac{\xi_{i+1/2}^{-2a_i} + \xi_{i-1/2}}{(h/2)^2} = 0$$

où  $\mu$  est un paramètre réel compris entre 0 et 1 :

- pour  $\mu = 1$  le système (1.6) représente le schéma LRG avec intégration exacte
- pour  $\mu = \frac{1}{3}$  le système (1.6) représente le schéma LRG avec intégration numérique par la méthode des trapèzes
- pour  $\mu = 0$  le système (1.6) représente le schéma DF décentré amont.

On obtient une classe de schéma LRG( $\mu$ ) dépendant d'un paramètre lié au type d'intégration numérique utilisé.

On peut écrire le schéma (1.6) en utilisant *divers degrés de liberté* pour représenter la même fonction  $u_h(t) \in V_h$  :

Choix 1 : valeur moyenne  $a_i$  et pente  $b_i$  sur chaque intervalle. Le schéma (1.6) se réécrit

$$\frac{da_i}{dt} + \frac{a_i - a_{i-1}}{h} + \frac{1}{2}(b_i - b_{i-1}) = 0$$

(1.7)

$$\frac{db_i}{dt} + \frac{3\mu}{2} \frac{a_{i-1} - a_i + h/2(b_i + b_{i-1})}{(h/2)^2} = 0 .$$

Choix 2 : valeur moyenne  $a_i$  et valeurs de Godunov  $\xi_{i+1/2}$

$$\frac{da_i}{dt} + \frac{\xi_{i+1/2} - \xi_{i-1/2}}{h} = 0$$

(1.8)

$$\frac{d\xi_{i+1/2}}{dt} + \frac{(3\mu-1)\xi_{i-1/2} - 6\mu a_i + (3\mu+1)\xi_{i+1/2}}{h} = 0 .$$

Le schéma LRG apparaît ici comme un schéma DF sur un pas  $h/2$  avec une discrétisation de  $\frac{\partial u}{\partial x}$  une fois sur deux centrée, une fois sur deux décentrée double-amont avec une pondération dépendant de  $\mu$ .

On remarque que, pour  $\mu = 2/3$  (c'est-à-dire pour un schéma "intermédiaire" entre LRG exact et LRG avec intégration par les trapèzes), la deuxième équation de (1.8) se réécrit :

$$(1.8\text{bis}) \quad \frac{d\xi_{i+1/2}}{dt} + \frac{\xi_{i-1/2} - 4a_i + 3\xi_{i+1/2}}{h} = 0 .$$

Ainsi, pour  $\mu = 2/3$ , la dérivée  $\frac{\partial u}{\partial x}$  en  $x_{i+1/2}$  est estimée à l'aide d'une interpolation parabolique à partir des valeurs en  $x_{i-1/2}$ ,  $x_i$  et  $x_{i+1/2}$ .

Ceci suggère que, pour  $\mu = 2/3$ , la deuxième équation de (1.8) soit

d'un ordre de précision supérieur, ce que l'on vérifiera bien dans la suite.

Choix 3 : valeur moyenne  $a_i$  et pente  $\tilde{b}_i$  en  $x_i$  de la parabole inter-  
polant  $\xi_{i-1/2}$ ,  $a_i$  et  $\xi_{i+1/2}$

On a évidemment, pour un maillage régulier :

$$\tilde{b}_i = \frac{1}{h} (\xi_{i+1/2} - \xi_{i-1/2}) .$$

Le système (1.6) se réécrit alors :

$$\frac{da_i}{dt} + \tilde{b}_i = 0$$

(1.9)

$$\frac{d\tilde{b}_i}{dt} + \frac{h[(3\mu-1)\tilde{b}_{i-1} + (3\mu+1)\tilde{b}_i] - 6\mu a_i - a_{i-1}}{h^2} = 0$$

Choix 4 : valeurs de Godunov  $\xi_{i+1/2}$  et pente  $b_i$  de  $u_h$  en  $x_{i+1/2}$

$$\frac{d\xi_{i+1/2}}{dt} + (1-3\mu) \frac{\xi_{i+1/2} - \xi_{i-1/2}}{h} + 3\mu b_i = 0$$

(1.10)

$$\frac{db_i}{dt} + 2 \frac{hb_i - (\xi_{i+1/2} - \xi_{i-1/2})}{h^2} = 0$$

## 1.2 - Schéma LRG : discrétisation en temps

On découpe l'intervalle  $[0, T]$  en  $M$  intervalles de longueur  $\Delta t$  et on pose :

$$(1.11) \quad t^n = n \Delta t = n \frac{T}{M} \quad n = 0, 1, \dots, M$$

et on utilise un schéma explicite en temps, de sorte que (1.6) est approché par :

$$(1.12) \quad \frac{a_i^{n+1} - a_i^n}{\Delta t} + \frac{\xi_{i+1/2}^n - \xi_{i-1/2}^n}{h} = 0$$

$$\frac{b_i^{n+1} - b_i^n}{\Delta t} + \frac{3\mu}{2} \frac{\xi_{i+1/2}^n - 2a_i^n + \xi_{i-1/2}^n}{(h/2)^2} = 0$$

d'où on déduit immédiatement les approximations de (1.7) (1.8) (1.9) et (1.10) :

Choix 1 : degrés de liberté  $(a_i, b_i)$  :

$$(1.13) \quad \frac{a_i^{n+1} - a_i^n}{\Delta t} + \frac{a_i^n - a_{i-1}^n}{h} + \frac{1}{2}(b_i^n - b_{i-1}^n) = 0$$

$$\frac{b_i^{n+1} - b_i^n}{\Delta t} + \frac{3\mu}{2} \frac{a_{i-1}^n - a_i^n + h/2(b_i^n + b_{i-1}^n)}{(h/2)^2} = 0$$

Choix 2 : degrés de liberté  $(a_i, \xi_{i+1/2})$  :

$$(1.14) \quad \frac{a_i^{n+1} - a_i^n}{\Delta t} + \frac{\xi_{i+1/2}^n - \xi_{i-1/2}^n}{h} = 0$$

$$\frac{\xi_{i+1/2}^{n+1} - \xi_{i+1/2}^n}{\Delta t} + \frac{(3\mu-1)\xi_{i-1/2}^n - 6\mu a_i^n + (3\mu+1)\xi_{i+1/2}^n}{h} = 0$$

Choix 3 : degrés de liberté  $(a_i, \tilde{b}_i)$

$$(1.15) \quad \frac{a_i^{n+1} - a_i^n}{\Delta t} + \tilde{b}_i^n = 0$$

$$\frac{\tilde{b}_i^{n+1} - \tilde{b}_i^n}{\Delta t} + \frac{h[(3\mu-1)\tilde{b}_{i-1}^n + (3\mu+1)\tilde{b}_i^n] - 6\mu(a_i^n - a_{i-1}^n)}{h^2} = 0$$

Choix 4 : degrés de liberté  $(\xi_{i+1/2}, b_i)$

$$(1.16) \quad \frac{\xi_{i+1/2}^{n+1} - \xi_{i+1/2}^n}{\Delta t} + (1-3\mu) \frac{\xi_{i+1/2}^n - \xi_{i-1/2}^n}{h} + 3\mu b_i^n = 0$$

$$\frac{b_i^{n+1} - b_i^n}{\Delta t} + 2 \frac{hb_i^n - (\xi_{i+1/2}^n - \xi_{i-1/2}^n)}{h^2} = 0$$

### 1.3 - Deux schémas aux différences finies

A titre de comparaison, on indique ici deux schémas aux différences finies "d'ordre élevé". Les seuls degrés de liberté sont ici les valeurs moyennes  $a_i$  sur chaque maille.

. Schéma double amont

$$(1.17) \quad \frac{a_i^{n+1} - a_i^n}{\Delta t} + \frac{3a_i^n - 4a_{i-1}^n + a_{i-2}^n}{2h} = 0 .$$

Ce schéma est construit en remarquant que  $\frac{3a_i - 4a_{i-1} + a_{i-2}}{2h}$  est la pente, en  $x_i$  (cf. figure 1), de la parabole interpolant les valeurs  $a_{i-2}$ ,  $a_{i-1}$  et  $a_i$ .

Il peut s'interpréter comme un schéma "élément fini discontinu" dont les degrés de liberté  $(a_i^n, b_i^n)$  sont définis par :

$$(1.18) \quad \frac{a_i^{n+1} - a_i^n}{\Delta t} + \frac{\xi_{i+1/2}^n - \xi_{i-1/2}^n}{h} = 0$$

$$\xi_{i+1/2}^n = u_h^n(x_{i+1/2}) = a_i^n + \frac{h}{2} b_i^n$$

$$b_i^n = \frac{a_i^n - a_{i-1}^n}{h} .$$

. Schéma double-amont avec Lax-Wendroff

En utilisant l'équation (1) pour augmenter la précision de l'approximation de  $\frac{\partial u}{\partial t}$  on obtient le schéma :

$$(1.19) \quad \frac{a_i^{n+1} - a_i^n}{\Delta t} + \frac{3a_i - 4a_{i-1} + a_{i-2}}{2h} - \frac{\Delta t}{2h^2} [a_i - 2a_{i-1} + a_{i-2}] = 0$$

où l'on a rajouté un peu de diffusion. Ce schéma peut aussi s'interpréter en terme d'éléments finis discontinus.

2 - ETUDE DE LA CONSISTANCE DES SCHEMAS

Pour étudier la consistance d'un schéma écrit en termes d'un certain nombre de *degrés de liberté*, il faut commencer par interpréter les *degrés de liberté* en terme de la solution continue, par exemple, dire "le degré de liberté  $a_i$ " de mon schéma approche la valeur  $V_i^n = u(x_i, t^n)$  de la solution  $u$ ".

On porte ensuite ces valeurs de  $u$  ou de sa dérivée dans le schéma numérique, et on évalue, à l'aide de formules de Taylor, le deuxième membre.

Ce deuxième membre, qui n'est en général pas nul, se compose de :

- une équation en  $u$
- des termes tendant vers zéro avec  $\Delta t$  et  $h$ .

Le schéma est consistant si l'équation en  $u$  est l'équation initiale (ou une dérivée de cette équation) que le schéma est censé approcher.

L'ordre ("*en équation*") du schéma est donné par la puissance des premiers termes non nuls en  $\Delta t$  et  $h$ .

On va voir qu'un schéma numérique peut être consistant ou inconsistant suivant le choix fait pour les degrés de liberté et pour leur interprétation : on pourra ainsi mieux cerner quels sont les degrés de liberté qui ont une chance de converger vers la solution exacte lorsque  $\Delta t$  et  $h$  tendent vers zéro.

Soit donc  $u(x,t)$  une fonction régulière. On notera :

$$\begin{aligned}
 V_i^n &= u(x_i, t^n) \\
 V_{i+1/2}^n &= u(x_{i+1/2}, t^n) \\
 B_i^n &= \frac{\partial u}{\partial x}(x_i, t^n) \\
 B_{i+1/2}^n &= \frac{\partial u}{\partial x}(x_{i+1/2}, t^n)
 \end{aligned}
 \tag{2.1}$$

### 2.1 - Consistance de LRG pour le choix 1 $(a_i, b_i)$ des degrés de liberté

L'interprétation naturelle de  $a_i$  et  $b_i$  est :

$$\begin{aligned}
 a_i^n &\text{ est censé approcher } V_i^n \\
 b_i^n &\text{ est censé approcher } B_i^n .
 \end{aligned}
 \tag{2.2}$$

En développant  $u$  au voisinage de  $(x_i, t^n)$  il vient

$$\begin{aligned}
 V_i^{n+1} &= V_i^n + \Delta t \frac{\partial u}{\partial t}(x_i, t^n) + \frac{\Delta t^2}{2} \frac{\partial^2 u}{\partial t^2}(x_i, t^n) + \dots \\
 V_{i-1}^n &= V_i^n - h B_i^n + \frac{h^2}{2} \frac{\partial^2 u}{\partial x^2}(x_i, t^n) - \frac{h^3}{6} \frac{\partial^3 u}{\partial x^3}(x_i, t^n) + \dots \\
 B_{i-1}^n &= B_i^n - h \frac{\partial^2 u}{\partial x^2}(x_i, t^n) + \frac{h^2}{2} \frac{\partial^3 u}{\partial x^3}(x_i, t^n) + \dots \\
 B_i^{n+1} &= B_i^n + \Delta t \frac{\partial^2 u}{\partial x \partial t}(x_i, t^n) + \frac{\Delta t^2}{2} \frac{\partial^3 u}{\partial x \partial t^2}(x_i, t^n) + \dots
 \end{aligned}
 \tag{2.3}$$

En remplaçant dans les premiers membres de (1.13)  $a_i^n$  par  $V_i^n$ ,  $a_{i-1}^n$  par  $V_{i-1}^n$ ,  $b_i^n$  par  $B_i^n$  etc..., il vient, compte tenu de (2.3) :

$$(2.11) \quad \frac{V_i^{n+1} - V_i^n}{\Delta t} + \frac{V_i^n - V_{i-1}^n}{h} + \frac{1}{2}(B_i^n - B_{i-1}^n) =$$

$$\frac{\partial u}{\partial t}(x_i, t^n) + \frac{\partial u}{\partial x}(x_i, t^n) - \frac{h^2}{12} \frac{\partial^3 u}{\partial x^3}(x_i, t^n) + \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2}(x_i, t^n) + \dots$$

$$\frac{B_i^{n+1} - B_i^n}{\Delta t} + \frac{3}{2} \frac{V_{i-1}^n - V_i^n + h/2(B_i^n + B_{i-1}^n)}{(h/2)^2} =$$

$$\frac{\partial^2 u}{\partial x \partial t}(x_i, t^n) + \mu \frac{h}{2} \frac{\partial^3 u}{\partial x^3}(x_i, t^n) + \frac{\Delta t}{2} \frac{\partial^3 u}{\partial t^2 \partial x}(x_i, t^n) + \dots$$

On constate que, si on retrouve bien, au deuxième membre de la première équation (2.4), l'équation initiale  $\partial u / \partial t + \partial u / \partial x$ , il manque, au deuxième membre de la deuxième équation, un terme  $\partial^2 u / \partial x^2(x_i, t^n)$  pour avoir la consistance.

Cette propriété, mise en évidence pour la première fois par G. BARRE, n'implique pas que les schémas LRG( $\mu$ ) soient inconsistants (nous verrons qu'ils le sont pour d'autre degrés de liberté), mais seulement que la pente  $b_i$  de  $u_h(t)$  sur l'intervalle  $]x_{i-1/2}, x_{i+1/2}[$  ne doit pas s'interpréter comme une approximation de la pente de la solution exacte au milieu  $x_i$  de l'intervalle (nous verrons en étudiant la consistance du choix 4 des paramètres que  $b_i$  doit s'interpréter comme une approximation de la pente de la solution exacte au point  $x_{i+1/2}$  !).

## 2.2 - Consistance du schéma LRG pour le choix 2 ( $a_i, \xi_{i+1/2}$ ) des degrés de liberté

L'interprétation choisie de  $a_i$  et  $\xi_{i+1/2}$  est :



(2.5)  $a_i^n$  est censé approcher  $V_i^n$

$\xi_{i+1/2}^n$  est censé approcher  $V_{i+1/2}^n$ .

En développant  $u$  au voisinage de  $(x_i, t^n)$  on vérifie facilement que :

(2.6)

$$V_i^{n+1} = V_i^n + \Delta t \frac{\partial u}{\partial t}(x_i, t^n) + \frac{\Delta t^2}{2} \frac{\partial^2 u}{\partial t^2}(x_i, t^n) + \dots$$

$$V_{i+1/2}^n = V_i^n + \left(\frac{h}{2}\right) B_i^n + \frac{1}{2} \left(\frac{h}{2}\right)^2 \frac{\partial^2 u}{\partial x^2}(x_i, t^n) + \frac{1}{6} \left(\frac{h}{2}\right)^3 \frac{\partial^3 u}{\partial x^3}(x_i, t^n) + \dots$$

$$V_{i-1/2}^n = V_i^n - \left(\frac{h}{2}\right) B_i^n + \frac{1}{2} \left(\frac{h}{2}\right)^2 \frac{\partial^2 u}{\partial x^2}(x_i, t^n) - \frac{1}{6} \left(\frac{h}{2}\right)^3 \frac{\partial^3 u}{\partial x^3}(x_i, t^n) + \dots$$

En remplaçant  $a_i^n$  par  $V_i^n$ ,  $\xi_{i+1/2}^n$  par  $V_{i+1/2}^n$  etc dans la première équation de (1.14), il vient, compte tenu de (2.6) :

(2.7)

$$\frac{V_i^{n+1} - V_i^n}{\Delta t} + \frac{V_{i+1/2}^n - V_{i-1/2}^n}{h} =$$

$$\frac{\partial u}{\partial t}(x_i, t^n) + \frac{\partial u}{\partial x}(x_i, t^n) + \frac{h^2}{24} \frac{\partial^3 u}{\partial x^3}(x_i, t^n) + \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2}(x_i, t^n) + \dots$$

En développant  $u$  au voisinage de  $(x_{i+1/2}, t^n)$  on trouve :

(2.8)

$$V_{i+1/2}^{n+1} = V_{i+1/2}^n + \Delta t \frac{\partial u}{\partial t}(x_{i+1/2}, t^n) + \frac{\Delta t^2}{2} \frac{\partial^2 u}{\partial t^2}(x_{i+1/2}, t^n) + \dots$$

$$V_i^n = V_{i+1/2}^n - \left(\frac{h}{2}\right) \frac{\partial u}{\partial x}(x_{i+1/2}, t^n) + \frac{1}{2} \left(\frac{h}{2}\right)^2 \frac{\partial^2 u}{\partial x^2}(x_{i+1/2}, t^n)$$

$$- \frac{1}{6} \left(\frac{h}{2}\right)^3 \frac{\partial^3 u}{\partial x^3}(x_{i+1/2}, t^n) + \dots$$

$$V_{i-1/2}^n = V_{i+1/2}^n - h \frac{\partial u}{\partial x}(x_{i+1/2}, t^n) + \frac{h^2}{2} \frac{\partial^2 u}{\partial x^2}(x_{i+1/2}, t^n)$$

$$- \frac{h^3}{6} \frac{\partial^3 u}{\partial x^3}(x_{i+1/2}, t^n) + \dots$$

En remplaçant dans la deuxième équation de (1.14)  $\xi_{i+1/2}^n$  par  $V_{i+1/2}^n$ ,  $u_i^n$  par  $V_i^n$  etc..., il vient :

$$\begin{aligned}
 & \frac{v_{i+1/2}^{n+1} - v_{i+1/2}^n}{\Delta t} + \frac{(3\mu-1)v_{i-1/2}^n - 6\mu v_i^n + (3\mu+1)\xi_{i+1/2}^n}{h} = \\
 (2.9) \quad & \frac{\partial u}{\partial t}(x_{i+1/2}, t^n) + \frac{\partial u}{\partial x}(x_{i+1/2}, t^n) + \left(\frac{3\mu}{2} - 1\right) \frac{h}{2} \frac{\partial^2 u}{\partial x^2}(x_{i+1/2}, t^n) + \\
 & + \left(2 - \frac{9\mu}{2}\right) \frac{h^2}{12} \frac{\partial^3 u}{\partial x^3}(x_{i+1/2}, t^n) + \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2}(x_{i+1/2}, t^n) + \dots
 \end{aligned}$$

On retrouve bien l'équation initiale aux seconds membres (2.7) et (2.9).

Quel que soit le paramètre  $\mu$  d'intégration numérique, le schéma LRG( $\mu$ ) est consistant pour l'interprétation (2.5) des degrés de liberté  $a_i^n$  et  $\xi_{i+1/2}^n$ . Il est d'ordre  $O(h^2 + \Delta t)$  aux points  $x_i$  et  $O(h + \Delta t)$  aux points  $x_{i+1/2}$ .

Ceci montre que des valeurs "significatives" de la fonction discontinue  $u_h(t)$  sont : - les valeurs au milieu de chaque intervalle  
 - les valeurs amont en chaque point de discontinuité.  
 En particulier, les valeurs "aval" n'ont aucune raison d'être significatives.

Remarque 1 : Pour  $\mu = \frac{2}{3}$  (ce qui doit correspondre à une autre formule d'intégration numérique), le schéma est en  $O(h^2 + \Delta t)$  en tous les points.

Remarque 2 : La consistance du schéma LRG( $\mu$ ) lorsqu'on interprète  $a_i$  et  $\xi_{i+1/2}$  comme des approximations de  $u(x_i)$  et  $u(x_{i+1/2})$  suggère de remplacer, pour l'interprétation et la représentation graphique des résultats, la fonction  $u_h(t)$  par la fonction  $\hat{u}_h(t)$  continue et linéaire par morceaux interpolant les valeurs  $a_i$  et  $\xi_{i+1/2}$  :

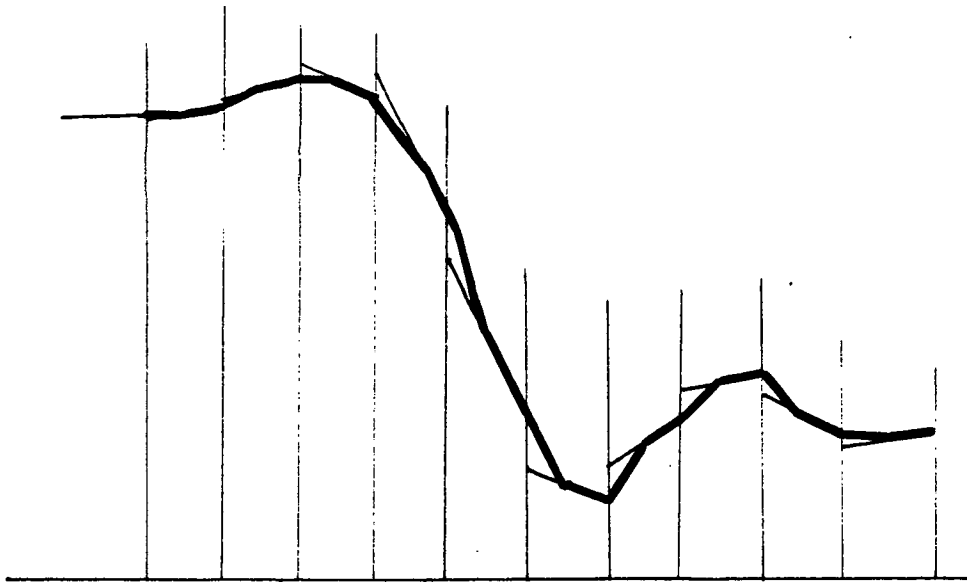


Figure 2

Cette représentation a l'avantage d'être continue mais n'est pas conservative ( $\int_0^t \hat{u}_h(t) dx \neq \int_0^L u_h(t) dx$ ).

2.3 - Consistance du schéma LRG pour le choix 3 ( $a_i, \tilde{b}_i$ ) des degrés de liberté

Nous choisissons pour ( $a_i, \tilde{b}_i$ ) l'interprétation suivante :

$$(2.10) \quad \begin{aligned} a_i^n & \text{ est censé approcher } V_i^n \\ \tilde{b}_i^n & \text{ est censé approcher } B_i^n . \end{aligned}$$

En développant  $u$  au voisinage de  $(x_i, t^n)$ , on a :

$$\begin{aligned}
 V_i^{n+1} &= V_i^n + \Delta t \frac{\partial u}{\partial t}(x_i, t^n) + \frac{\Delta t^2}{2} \frac{\partial^2 u}{\partial t^2}(x_i, t^n) + \dots \\
 B_i^{n+1} &= B_i^n + \Delta t \frac{\partial^2 u}{\partial x \partial t}(x_i, t^n) + \frac{\Delta t^2}{2} \frac{\partial^3 u}{\partial x \partial t^2}(x_i, t^n) + \dots \\
 V_{i-1}^n &= V_i^n - h B_i^n + \frac{h^2}{2} \frac{\partial^2 u}{\partial x^2}(x_i, t^n) - \frac{h^3}{6} \frac{\partial^3 u}{\partial x^3}(x_i, t^n) \\
 &\quad + \frac{h^4}{24} \frac{\partial^4 u}{\partial x^4}(x_i, t^n) + \dots \\
 B_{i-1}^n &= B_i^n - h \frac{\partial^2 u}{\partial x^2}(x_i, t^n) + \frac{h^2}{2} \frac{\partial^3 u}{\partial x^3}(x_i, t^n) - \frac{h^3}{6} \frac{\partial^4 u}{\partial x^4}(x_i, t^n) + \dots
 \end{aligned}
 \tag{2.11}$$

En remplaçant  $a_i^n$  par  $V_i^n$ ,  $\tilde{b}_i^n$  par  $B_i^n$  dans (1.15) etc., il vient, compte tenu de (2.11) :

$$\begin{aligned}
 \frac{V_i^{n+1} - V_i^n}{\Delta t} + B_i^n &= \frac{\partial u}{\partial t}(x_i, t^n) + \frac{\partial u}{\partial x}(x_i, t^n) + \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2}(x_i, t^n) + \dots \\
 \frac{B_i^{n+1} - B_i^n}{\Delta t} + \frac{h[(3\mu-1)B_{i-1}^n + (3\mu+1)B_i^n] - 6\mu(V_i^n - V_{i-1}^n)}{h^2} &= \\
 &= \frac{\partial^2 u}{\partial t \partial x}(x_i, t^n) + \frac{\partial^2 u}{\partial x^2}(x_i, t^n) + (\mu-1) \frac{h}{2} \frac{\partial^3 u}{\partial x^3}(x_i, t^n) - \\
 &\quad \left(\frac{3\mu}{2} - 1\right) \frac{h^2}{6} \frac{\partial^4 u}{\partial x^4}(x_i, t^n) + \frac{\Delta t}{2} \frac{\partial^3 u}{\partial x \partial t^2}(x_i, t^n) + \dots
 \end{aligned}
 \tag{2.12}$$

On constate que :

Quel que soit le paramètre  $\mu$  d'intégration numérique, le schéma LRG( $\mu$ ) est consistant pour l'interprétation (2.10) des degrés de liberté  $a_i^n$  et  $\tilde{b}_i^n$ . Le schéma est :

- en  $O(\Delta t + h^\infty)$  pour l'équation en  $a_i^n$
- en  $O(\Delta t + h)$  pour l'équation en  $\tilde{b}_i^n$  lorsque  $\mu < 1$  (intégration approchée)
- en  $O(\Delta t + h^2)$  pour l'équation en  $\tilde{b}_i^n$  lorsque  $\mu = 1$  (intégration exacte).

Remarque 3 : Ce résultat suggère de remplacer, pour l'interprétation et la

représentation graphique des résultats, la fonction  $u_h(t)$  par la fonction  $\tilde{u}_h(t)$  ayant sur chaque intervalle la même valeur moyenne  $a_i^n$  que  $u_h(t)$  mais pour pente  $\tilde{b}_i^n$  au lieu de  $b_i^n$  (cette représentation conserve les aires) :

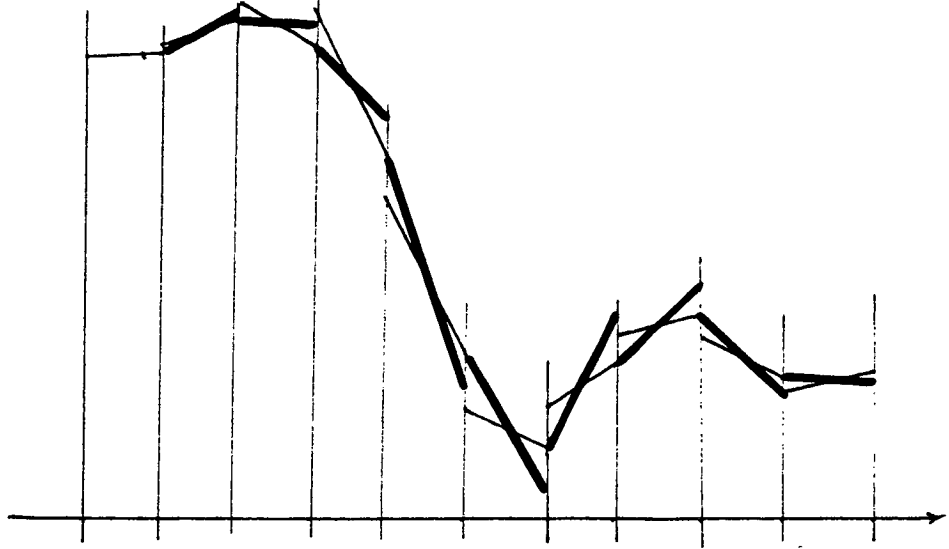


Figure 3

2.4 - Consistance du schéma LRG pour le choix 4  $(\xi_{i+1/2}, b_i)$  des degrés de liberté

Nous choisissons pour  $\xi_{i+1/2}$  et  $b_i$  l'interprétation suivante :

$$(2.13) \quad \xi_{i+1/2}^n \text{ est censé approcher } v_{i+1/2}^n \quad (\text{voir (2.1)})$$

$$b_i^n \text{ est censé approcher } B_{i+1/2}^n .$$

En développant au voisinage de  $x_{i+1/2}, t^n$  on trouve :

$$(2.14) \quad \begin{aligned} v_{i+1/2}^{n+1} &= v_{i+1/2}^n + \Delta t \frac{\partial u}{\partial t}(x_{i+1/2}, t^n) + \frac{\Delta t^2}{2} \frac{\partial^2 u}{\partial t^2}(x_{i+1/2}, t^n) \\ v_{i-1/2}^n &= v_{i+1/2}^n - h B_{i+1/2}^n + \frac{h^2}{2} \frac{\partial^2 u}{\partial x^2}(x_{i+1/2}, t^n) - \frac{h^3}{6} \frac{\partial^3 u}{\partial x^3}(x_{i+1/2}, t^n) \\ &\quad + \dots \\ B_{i+1/2}^{n+1} &= B_{i+1/2}^n + \Delta t \frac{\partial^2 u}{\partial x \partial t}(x_{i+1/2}, t^n) + \frac{\Delta t^2}{2} \frac{\partial^3 u}{\partial x \partial t^2}(x_{i+1/2}, t^n) + \dots \end{aligned}$$

En remplaçant, dans (1.16)  $\xi_{i+1/2}^n$  par  $v_{i+1/2}^n$ ,  $b_i^n$  par  $B_{i+1/2}^n$ , etc..., on trouve, compte tenu de (2.14) :

$$\begin{aligned}
 & \frac{v_{i+1/2}^{n+1} - v_{i+1/2}^n}{\Delta t} + (1-3\mu) \frac{v_{i+1/2}^n - v_{i-1/2}^n}{h} + 3\mu B_{i+1/2}^n = \\
 & = \frac{\partial u}{\partial t}(x_{i+1/2}, t^n) + \frac{\partial u}{\partial x}(x_{i+1/2}, t^n) \\
 & + \frac{3\mu-1}{2} \left[ h \frac{\partial^2 u}{\partial x^2}(x_{i+1/2}, t^n) - \frac{h^2}{3} \frac{\partial^3 u}{\partial x^3}(x_{i+1/2}, t^n) + \dots \right] + \\
 & + \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2}(x_{i+1/2}, t^n) + \dots
 \end{aligned}$$

(2.15)

$$\begin{aligned}
 & \frac{B_{i+1/2}^{n+1} - B_{i+1/2}^n}{\Delta t} + 2 \frac{h B_{i+1/2}^n - (v_{i+1/2}^n - v_{i-1/2}^n)}{h^2} = \\
 & = \frac{\partial^2 u}{\partial x \partial t}(x_{i+1/2}, t^n) + \frac{\partial^2 u}{\partial x^2}(x_{i+1/2}, t^n) - \frac{h}{3} \frac{\partial^3 u}{\partial x^3}(x_{i+1/2}, t^n) \\
 & + \frac{\Delta t}{2} \frac{\partial^3 u}{\partial x \partial t^2}(x_{i+1/2}, t^n) + \dots
 \end{aligned}$$

On constate que le schéma est bien constant pour l'interprétation (2.13) des degrés de liberté  $\xi_{i+1/2}^n$  et  $b_i$ .

Ce résultat suggère d'interpréter la pente  $b_i$  de la solution approchée sur  $]x_{i-1/2}, x_{i+1/2}[$  comme une approximation de la pente de la solution au point  $x_{i+1/2}$ .

degrés de liberté	Interprétation	Schéma consistant à l'équation		Ordre
$a_i^n$	$u(x_i, t^n)$	$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} - \frac{h^2}{12} \frac{\partial^3 u}{\partial x^3} + \dots$	$+ \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2} + \dots = 0$	$O(\Delta t + h^2)$
$b_i^n$	$\frac{\partial u}{\partial x}(x_i, t^n)$	$\frac{\partial^2 u}{\partial t \partial x} + \mu \frac{h}{2} \frac{\partial^3 u}{\partial x^3} + \dots$	$+ \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2} + \dots = 0$	Inconsistant
$a_i^n$	$u(x_i, t^n)$	$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} + \frac{h^2}{24} \frac{\partial^3 u}{\partial x^3} + \dots$	$+ \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2} + \dots = 0$	$O(\Delta t + h^2)$
$\xi_{i+1/2}^n$	$u(x_{i+1/2}, t^n)$	$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} + (\frac{3\mu}{2} - 1) \frac{h}{2} \frac{\partial^2 u}{\partial x^2} + (2 - \frac{9\mu}{2}) \frac{h^2}{12} \frac{\partial^3 u}{\partial x^3} + \dots$	$+ \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2} + \dots = 0$	$O(\Delta t + h)$ $\mu \neq 2/3$
				$O(\Delta t + h^2)$ $\mu = 2/3$
$a_i^n$	$u(x_i, t^n)$	$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x}$	$+ \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2} + \dots = 0$	$O(\Delta t + h^\infty)$
$\tilde{b}_i^n = \frac{\xi_{i+1/2}^n - \xi_{i-1/2}^n}{h}$	$\frac{\partial u}{\partial x}(x_i, t^n)$	$\frac{\partial^2 u}{\partial x \partial t} + \frac{\partial^2 u}{\partial x^2} + (\mu - 1) \frac{h}{2} \frac{\partial^3 u}{\partial x^3} - (\frac{3\mu}{2} - 1) \frac{h^2}{6} \frac{\partial^4 u}{\partial x^4} + \dots$	$+ \frac{\Delta t}{2} \frac{\partial^3 u}{\partial x \partial t^2} + \dots = 0$	$O(\Delta t + h)$ $\mu \neq 1$
				$O(\Delta t + h^2)$ $\mu = 1$
$\xi_{i+1/2}^n$	$u(x_{i+1/2}, t^n)$	$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} + \frac{3\mu - 1}{2} [h \frac{\partial^2 u}{\partial x^2} - \frac{h^2}{3} \frac{\partial^3 u}{\partial x^3} + \dots]$	$+ \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2} + \dots = 0$	$O(\Delta t + h)$ $\mu \neq 1/3$
				$O(\Delta t + h^\infty)$ $\mu = 1/3$
$b_i^n$	$\frac{\partial u}{\partial x}(x_{i+1/2}, t^n)$	$\frac{\partial^2 u}{\partial t \partial x} + \frac{2}{\partial x^2} - \frac{h}{3} \frac{\partial^3 u}{\partial x^3} + \dots$	$+ \frac{\Delta t}{2} \frac{\partial^3 u}{\partial x \partial t^2} + \dots = 0$	$O(\Delta t + h)$

Figure 4 : Consistence des schémas LRG( $\mu$ ) pour divers choix de paramètres d'intégration numériques ( $\mu = 1/3$  correspond à l'intégration par la méthode des trapèzes,  $\mu = 1$  à une intégration numérique exacte)

## 2.5 - Consistance des deux schémas aux différences finies

- Schéma double amont. En faisant le choix, naturel, que  $V_i^n$  approche  $V_i^n = u(x_i, t^n)$ , on vérifie facilement que :

$$(2.16) \quad \frac{V_i^{n+1} - V_i^n}{\Delta t} + \frac{3V_i^n - 4V_{i-1}^n + V_{i-2}^n}{2h} = \frac{\partial u}{\partial t}(x_i, t^n) + \frac{\partial u}{\partial x}(x_i, t^n) - h^2 \frac{\partial^3 u}{\partial x^3}(x_i, t^n) + \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2}(x_i, t^n) + \dots$$

Le schéma double amont est donc d'ordre 2 en espace et 1 en temps.

- Schéma double amont avec Lax Wendroff. On trouve de même

$$(2.17) \quad \frac{V_i^{n+1} - V_i^n}{\Delta t} + \frac{3V_i^n - 4V_{i-1}^n + V_{i-2}^n}{2h} - \frac{\Delta t}{2} \frac{V_i^n - 2V_{i-1}^n + V_{i-2}^n}{h^2} = \frac{\partial u}{\partial t}(x_i, t^n) + \frac{\partial u}{\partial x}(x_i, t^n) - h(h + \frac{\Delta t}{2}) \frac{\partial^3 u}{\partial x^3}(x_i, t^n) + \frac{\Delta t}{2} \frac{\partial^3 u}{\partial t^2} + \dots$$

Le schéma double amont avec Lax Wendroff est donc d'ordre 2 en espace et en temps.

## 3 - ETUDE DE LA STABILITE DU SCHEMA

### 3.1 - Introduction

Soit  $u_h^n$  la solution de l'équation (0.1) obtenue après  $n$  itérations d'un schéma numérique donné. Si  $\Delta t$  est le pas de temps correspondant à une itération, on a :

$$(3.1) \quad u_h^n = u_h(n \Delta t) .$$



L'étude de la stabilité d'un schéma revient à étudier la croissance de la norme de  $u_h^n$  lorsque  $n \rightarrow \infty$ . Il y a naturellement de nombreuses façons de faire tendre  $n$  vers l'infini suivant que l'on fait varier (ou non)  $\Delta t$  et  $h$  simultanément avec  $n$  :

. Lorsqu'on s'intéresse au calcul de la solution  $u(.,T)$  à un instant  $T$  donné, fini ou infini, on est amené à raffiner simultanément le maillage en espace et en temps, de sorte que  $n \rightarrow \infty$  nécessairement puisque

$$n \Delta t = T \quad (T \text{ donné fini ou infini})$$

Si nous posons

$$(3.2) \quad \gamma = \frac{\Delta t}{h}$$

une stratégie de raffinement de la discrétisation se résume à la donnée d'une suite  $\{\gamma_n > 0, n \in \mathbb{N}\}$ .

Ainsi, le calcul de  $u(.,T)$  à l'aide du schéma considéré et de la stratégie  $\{\gamma_n > 0, n \in \mathbb{N}\}$  de raffinement de la discrétisation revient à chercher :

$$\lim_{n \rightarrow \infty} u_h^n \quad (= u(.,T) \text{ si le schéma est convergent !})$$

où

$$(3.3) \quad u_h^n \text{ est calculé avec } \Delta t = \frac{T}{n} \text{ et } h = \frac{\Delta t}{\gamma_n} .$$

Il est alors naturel d'introduire les définitions suivantes :

Définition 1 :

Etant donné une stratégie  $\{\gamma_n > 0\}$  de raffinement de la discrétisation, le schéma sera dit :

.  $L^\infty(0,T;L^2)$ -stable si, pour tout  $T > 0$  donné, il existe  $M(T)$  (en général tendant vers l'infini avec  $T$  !) tel que :

$$(3.4) \quad \forall n, \forall u_{0h}, \quad \|u_h^n\|_{L^2(OL)} \leq M(T) \|u_{0h}\|_{L^2(OL)}$$

où  $u_h^n$  est calculé comme en (3.3).

.  $L^\infty(L^2)$ -stable si il existe  $M$  (indépendante de  $T$ ) tel que :

$$(3.5) \quad \forall T > 0, \forall u_{0h}, \quad \lim_{n \rightarrow \infty} \|u_h^n\|_{L^2(OL)} \leq M \|u_{0h}\|_{L^2(OL)}$$

où  $u_h^n$  est toujours calculé comme en (3.3).

. *Lorsqu'on s'intéresse au calcul de la solution  $u_h(.,\infty)$  pour une discrétisation en espace et en temps donnée, on cherche*

$$\lim_{n \rightarrow \infty} u_h^n \quad (= u_h(.,\infty))$$

où cette fois

$$(3.6) \quad u_h^n \text{ est calculé avec } \Delta t \text{ et } h \text{ fixés.}$$

On introduit alors la

Définition 2 :

Le schéma sera dit asymptotiquement stable si, pour  $\Delta t$  et  $h$  fixés, il existe  $M(\Delta t, h) > 0$  telle que :

$$(3.7) \quad \forall t > 0, \forall u_{0h}, \quad \|u_h^n\|_{L^2(OL)} \leq M(\Delta t, h) \|u_{0h}\|_{L^2(OL)}$$

où  $u_h^n$  est calculé comme en (3.6).

Dans le cas modèle linéaire de l'équation (0.1) considéré jusqu'ici, l'analyse des coefficients d'amplification des coefficients de Fourier d'une solution périodique (Méthode de Von Neumann) permet d'obtenir des conditions nécessaires et suffisantes de stabilité.

Dans les problèmes de déplacements immiscibles, le modèle est en fait non linéaire, et il faut considérer l'équation

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

Les schémas proposés au paragraphe 1 se généralisent sans difficulté au cas de l'équation (3.8), mais la méthode d'analyse évoquée ci-dessus n'est plus utilisable. On abordera alors le problème de la stabilité par une méthode purement expérimentale.

L'étude se décomposera donc en trois étapes :

- . au paragraphe 3.2, nous rappellerons l'adaptation de la méthode de Von Neumann aux schémas Eléments Finis, puis nous établirons un protocole d'étude expérimentale de la stabilité
- . au paragraphe 3.3, nous étudierons les schémas LRG- $\mu$  et différences finies dans le cas linéaire à l'aide de la méthode de Von Neumann
- . au paragraphe 3.4, nous appliquerons la méthode expérimentale d'abord au cas linéaire (vérification des résultats du paragraphe 3.3, puis à quelques cas non-linéaires
- . enfin, le paragraphe 3.5 contiendra les conclusions sur la stabilité des différents schémas.

### 3.2 - Description des méthodes d'étude de la stabilité

#### 3.2.1 - La méthode de Von Neumann

Cette méthode ne s'applique pas directement au système (0.1), mais au système obtenu en remplaçant la condition aux limites de Dirichlet par une condition de périodicité :

$$\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} = 0 \quad \forall x \in \mathbb{R}, \quad \forall t \geq 0$$

$$(3.9) \quad x \rightarrow u(x,t) \text{ p\u00e9riodique de p\u00e9riode } L \quad \forall t > 0$$

$$u(x,0) = u_0(x) \text{ donn\u00e9 p\u00e9riodique de p\u00e9riode } L.$$

Il suffit donc de calculer la solution  $u(x,t)$  pour tout  $x \in [0,L]$  et tout  $t \geq 0$ .

Nous consid\u00e9rerons dans ce paragraphe la classe des sch\u00e9mas aux \u00e9l\u00e9ments finis approchant, \u00e0 chaque instant  $t^n = n \Delta t$ , la solution  $u(.,t^n)$  du syst\u00eame (3.9) par une fonction discontinue lin\u00e9aire par morceaux  $u_h \in V_h$  ( $V_h$  est d\u00e9fini au paragraphe 1.1).

Se donner un tel sch\u00e9ma revient \u00e0 se donner un op\u00e9rateur lin\u00e9aire  $S_h$  de  $V_h$  dans lui-m\u00eame tel que :

$$(3.10) \quad u_h^{n+1} = S_h u_h^n.$$

La fonction  $u_h^n \in V_h$  sera rep\u00e9r\u00e9e par les degr\u00e9s de libert\u00e9 suivants : (cf. figure 1)

$$(3.11) \quad \begin{aligned} a_i^n &= \text{valeur moyenne de } u_h^n \text{ sur } ]x_{i-1/2}, x_{i+1/2}[ \text{ \u00e0 l'instant } t^n \\ c_i^n &= \frac{h b_i^n}{2\sqrt{3}} = \frac{1}{2\sqrt{3}} \{u_h(x_{i+1/2}^-, t^n) - u_h(x_{i-1/2}^+, t^n)\} \end{aligned}$$

qui ont l'avantage que

$$(3.12) \quad \|u_h^n\|_{L^2(0,L)}^2 = h \sum_{i=1}^n \{(a_i^n)^2 + (c_i^n)^2\}.$$

A chaque fonction  $u_h^n$  on associe, de fa\u00e7on biunivoque, les fonctions p\u00e9riodiques et constantes sur chaque intervalle :

$$a_h^n(x) = \sum_{\ell \in \mathbb{Z}} a_\ell^n \chi_\ell(x)$$

$$c_h^n(x) = \sum_{\ell \in \mathbb{Z}} c_\ell^n \chi_\ell(x)$$

où  $\chi_\ell(x)$  est la fonction caractéristique de l'intervalle  $]x_{\ell-1/2}, x_{\ell+1/2}[$ .

La norme de  $u_h^n$  s'écrit alors :

$$(3.13) \quad ||u_h^n||_{L^2(0,L)}^2 = ||a_h^n||_{L^2(0,L)}^2 + ||c_h^n||_{L^2(0,L)}^2 .$$

On décompose ensuite les deux fonctions  $a_h^n$  et  $c_h^n$  en séries de Fourier en posant :

$$(3.11) \quad a_h^n(x) = \sum_{k \in \mathbb{Z}} A_k^n e^{i2k\pi x/L}$$

$$c_h^n(x) = \sum_{k \in \mathbb{Z}} C_k^n e^{i2k\pi x/L} .$$

L'égalité de Bessel-Parseval pour les séries de Fourier donne :

$$\sum_{k \in \mathbb{Z}} |A_k^n|^2 = \frac{1}{L} ||a_h^n||_{L^2(0,L)}^2$$

$$\sum_{k \in \mathbb{Z}} |C_k^n|^2 = \frac{1}{L} ||c_h^n||_{L^2(0,L)}^2$$

d'où on tire immédiatement :

$$(3.14bis) \quad ||u_h^n||_{L^2(0,L)}^2 = L \sum_{k \in \mathbb{Z}} \{ |A_k^n|^2 + |C_k^n|^2 \} .$$

En utilisant le schéma (3.10), les degrés de liberté (3.11) et les transformées de Fourier (3.14), on vérifie facilement que l'évolution des coefficients de Fourier  $A_k^n$  et  $C_k^n$  est donnée par :

$$(3.15) \quad \begin{vmatrix} A_k^{n+1} \\ C_k^{n+1} \end{vmatrix} = G(\theta, \gamma) \begin{vmatrix} A_k^n \\ C_k^n \end{vmatrix} \quad \forall k \in Z \quad \forall n = 0, 1, \dots$$

où  $G(\theta, \gamma)$  est une matrice  $2 \times 2$  à coefficients complexes dépendant des paramètres  $\theta$  et  $\gamma$ , avec :

$$(3.16) \quad \theta = 2k\pi \frac{h}{L} \quad \gamma = \frac{\Delta t}{h} .$$

De (3.15), on tire :

$$(3.17) \quad \begin{vmatrix} A_k^n \\ C_k^n \end{vmatrix} = G(\theta, \gamma)^n \begin{vmatrix} A_k^0 \\ C_k^0 \end{vmatrix} .$$

Pour une matrice  $M$   $2 \times 2$  à coefficients complexes, nous noterons  $\|M\|$  la norme matricielle induite par la norme Euclidienne sur  $C \times C$ . On sait que cette norme est donnée par

$$(3.18) \quad \|M\| = \rho(M^* M)^{1/2} \quad \text{dans le cas général}$$

ou par

$$(3.19) \quad \|M\| = \rho(M) \quad \text{lorsque } M \text{ est normale, c\`ad lorsque } M^* M = M M^*$$

où  $M$  est la matrice adjointe de  $M$  ( $m_{ij}^* = \overline{m_{ji}}$ ) et  $\rho(A)$  désigne le rayon de la matrice  $A$ , c'est-à-dire le Max des modules des valeurs propres de  $A$ .

On déduit donc de (3.17) :

$$(3.20) \quad |A_k^n|^2 + |C_k^n|^2 \leq ||G(\theta, \gamma)^n|| \{ |A_k^0|^2 + |C_k^0|^2 \}$$

$$\forall k \in \mathbb{Z}, \quad \text{avec } \theta = 2k\pi \frac{h}{L}.$$

De (3.14bis), on tire alors :

$$(3.21) \quad ||u_h^n||_{L^2(0,L)} \leq \sup_{-\pi \leq \theta \leq \pi} ||G(\theta, \gamma)^n|| ||u_h^0||$$

On suppose maintenant donnée une suite  $\{\gamma_n > 0\}$  (cf. § 3.1) : par exemple,  $\gamma_n$  pourra tendre vers zero de telle sorte que  $n \Delta t = T$  donné à l'avance lorsqu'on étudiera la stabilité  $L^\infty(0T; L^2)$  ou  $L^\infty(L^2)$ , ou bien  $\gamma_n$  sera choisi constant lorsqu'on étudiera la stabilité asymptotique.

On a alors la proposition suivante :

Proposition 1 :

Pour une suite  $\{\gamma_n > 0\}$  donnée, on a équivalence des deux propriétés :

$$(3.22) \quad ||u_h^n||_{L^2(0,L)} \leq M ||u_{0h}||_{L^2(0,L)} \quad \forall n \in \mathbb{N}, \quad \forall u_{0h}$$

et :

$$(3.23) \quad \sup_{-\pi \leq \theta \leq \pi} ||G(\theta, \gamma_n)^n|| \leq M \quad \forall n \in \mathbb{N}$$

(et les deux constantes  $M$  sont les mêmes).

Démonstration : Il est clair que (3.23) implique (3.22) d'après (3.21). Montrons que (3.22) implique (3.23). Si (3.23) n'avait pas lieu, il existerait  $\tilde{n} \in \mathbb{N}$  et  $\tilde{\theta} \in [-\pi, +\pi]$  tels que

$$(3.24) \quad ||G(\tilde{\theta}, \gamma_n)^{\tilde{n}}|| > M$$

et donc il existerait  $A_k^0$  et  $C_k^0$ , avec  $\tilde{\theta} = 2\tilde{k}\pi \frac{h}{L}$ , tels que :

$$|A_{\tilde{k}}^{\tilde{n}}|^2 + |C_{\tilde{k}}^{\tilde{n}}|^2 = ||G(\tilde{\theta}, \gamma_{\tilde{n}})^{\tilde{n}}||^2 \{ |A_k^0|^2 + |C_k^0|^2 \}.$$

Définissons alors  $u_{0h}$  par ses coefficients de Fourier :

$$(A_k^0, C_k^0) = \begin{cases} 0 & \text{pour } k \neq \tilde{k} \\ (A_{\tilde{k}}^0, C_{\tilde{k}}^0) & \text{pour } k = \tilde{k} \end{cases}.$$

On trouve alors :

$$||u_h^{\tilde{n}}||^2 = L \{ |A_{\tilde{k}}^{\tilde{n}}|^2 + |C_{\tilde{k}}^{\tilde{n}}|^2 \} = L ||\tilde{G}(\tilde{\theta}, \gamma_{\tilde{n}})^{\tilde{n}}||^2 \{ |A_k^0|^2 + |C_k^0|^2 \}$$

$$||u_h^{\tilde{n}}||^2 = ||G(\tilde{\theta}, \gamma_{\tilde{n}})^{\tilde{n}}||^2 ||u_{0h}||^2$$

d'où, grâce à (3.24) :

$$||u_h^{\tilde{n}}|| > M ||u_{0h}||$$

ce qui est en contradiction avec l'hypothèse (3.22). □

On est donc ramené, pour étudier la stabilité du schéma, à rechercher sous quelles conditions la norme de la puissance  $n$ -ième d'une matrice est bornée. Pour une matrice  $M$  régulière, on posera :

$$(3.25) \quad \text{Cond } M = ||M|| ||M^{-1}|| \geq 1 \quad (\text{conditionnement de } M)$$

et on rappelle la

Proposition 2 :

Soit  $G$  une matrice  $2 \times 2$  à coefficients complexes. Il existe une matrice  $P$  régulière et une matrice  $J$  telles que :

$$(3.26) \quad G = P J P^{-1}$$

avec

$$(3.27) \quad J = \begin{vmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{vmatrix} \quad (\text{si } G \text{ diagonalisable})$$



ou bien

$$(3.28) \quad J = \begin{vmatrix} \lambda & 1 \\ 0 & \lambda \end{vmatrix} \quad (\text{si } G \text{ non diagonalisable}).$$

Dans tous les cas, on a :

$$(3.29) \quad (\text{Cond } P)^{-1} ||J^n|| \leq ||G^n|| \quad \text{Cond } P ||J^n||$$

avec :

$$(3.30) \quad ||J^n|| = \begin{cases} \rho(G)^n & \text{si } G \text{ est diagonalisable} \\ \rho(G)^n \left\{ 1 + \frac{n^2}{2\rho(G)^2} (1 + \sqrt{1 + 4\rho(G)^2/n^2}) \right\}^{1/2} & \text{si } G \text{ n'est pas diagonalisable.} \end{cases}$$

Démonstration : Il suffit de démontrer (3.30) dans le cas non diagonalisable; on vérifie d'abord que :

$$(3.31) \quad J^n = \begin{vmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{vmatrix}$$

et que

$$(3.32) \quad (J^n)^* J^n = |\lambda|^{2n} \begin{vmatrix} 1 & \frac{n\bar{\lambda}}{|\lambda|^2} \\ \frac{n\lambda}{|\lambda|^2} & 1 + \frac{n^2}{|\lambda|^2} \end{vmatrix}$$

d'où l'on tire les valeurs propres  $\mu_{\pm}$  de  $(J^n)^* J^n$  :

$$(3.33) \quad \mu_{\pm} = |\lambda|^{2n} \left\{ 1 \pm \frac{n^2}{2|\lambda|^2} (1 + \sqrt{1 + 4|\lambda|^2/n^2}) \right\}$$

d'où (3.30) en prenant la plus grande des deux racines. □

Remarque 4 : on a naturellement

$$(3.34) \quad ||G^n|| \leq ||G||^n$$

mais ceci ne peut donner qu'une condition suffisante de stabilité, car on sait que :

$$(3.35) \quad \|G\| \leq \rho(G)$$

l'égalité n'étant atteinte que pour les matrices normales. Par exemple, si on considère la matrice

$$G = \begin{vmatrix} 0 & 1/2 \\ 2 & 0 \end{vmatrix}$$

on a  $G^2 = I$ ,  $G^3 = G$ ,  $G^4 = I$ , etc... de sorte que l'on a bien  $\|G^n\|$  borné! Par contre, on a :

$$G^*G = \begin{vmatrix} 4 & 0 \\ 0 & 1/4 \end{vmatrix}$$

de sorte que  $\|G\| = \rho(G^*G)^{1/2} = 2 > 1$  de sorte que  $\|G\|^n = 2^n \rightarrow +\infty$ .  $\square$

On aura enfin besoin, pour étudier la limite de  $\rho(G)^n$ , du résultat suivant :

Proposition 3 :

Soit  $\Delta t \rightarrow \varepsilon(\Delta t)$  une fonction de  $\mathbb{R}^{++}$  dans  $\mathbb{R}$ . On a alors, pour tout  $T > 0$  et tout  $\eta > 0$  donnés :

$$(3.36) \quad \left(1 + \varepsilon\left(\frac{T}{n}\right)\right)^n \text{ tend vers } \begin{cases} 0 & \text{si } \varepsilon(\Delta t) \sim -\eta \Delta t^m, \text{ avec } 0 < m < 1 \\ e^{-\eta T} \text{ (par valeurs inférieures)} & \text{si } \varepsilon(\Delta t) \sim -\eta \Delta t \\ 1 & \text{si } \varepsilon(\Delta t) \sim \pm \eta \Delta t^m, \text{ avec } 1 < m \\ e^{\eta T} \text{ (par valeurs inférieures)} & \text{si } \varepsilon(\Delta t) \sim \eta \Delta t \\ +\infty & \text{si } \varepsilon(\Delta t) \sim \eta \Delta t^m, \text{ avec } 0 < m < 1 \end{cases}$$

Nous pouvons maintenant combiner les propositions 1 à 3 pour obtenir des conditions nécessaires et suffisantes de stabilité.

Nous supposons (ce qui est le cas général) que :

$$(3.37) \quad \forall \theta \in [-\pi, +\pi] \quad \forall \gamma > 0 \quad G(\theta, \gamma) \text{ est diagonalisable}$$

et nous noterons :

$$(3.38) \quad P(\theta, \gamma) = \text{matrice de passage de } G(\theta, \gamma) \text{ ayant le plus petit conditionnement}$$

Les matrices de passage de  $G(\theta, \gamma)$  étant toutes engendrées, sous l'hypothèse (3.37), par les deux vecteurs propres distincts de  $G(\theta, \gamma)$ , il est facile de vérifier que celles ayant le meilleur conditionnement sont obtenues lorsque les deux vecteurs propres sont choisis de même norme hermitienne.

Posons alors :

$$(3.39) \quad \rho(\gamma) = \sup_{-\pi \leq \theta \leq \pi} \rho(G(\theta, \gamma)) \quad \forall \gamma > 0$$

et

$$(3.40) \quad v = \sup_{\substack{-\pi \leq \theta \leq \pi \\ 0 < \gamma < 1}} \text{Cond } P(\theta, \gamma) .$$

La proposition 2 donne alors :

$$(3.41) \quad v^{-1} \cdot \rho(\gamma_n)^n \leq \sup_{-\pi \leq \theta \leq \pi} \|G(\theta, \gamma_n)^n\| \leq v \rho(\gamma_n)^n .$$

En combinant alors avec la proposition 1 et 3 on trouve le

Théorème 1 :

Sous l'hypothèse (3.37) et pour une stratégie  $\{\gamma_n > 0, n \in \mathbb{N}\}$  de gestion de la discrétisation, le schéma (3.7) sera :

.  $L^\infty(0T; L^2)$  instable si :

$$(3.42) \quad \rho(\gamma_n) - 1 \sim \eta \Delta t^m \quad \text{avec } \eta > 0 \quad \text{et } 0 < m < 1$$

et alors :

$$(3.43) \quad \forall n, ||u_h^n|| \geq \nu^{-1} \rho(\gamma_n)^n ||u_{0h}|| \rightarrow +\infty$$

.  $L^\infty(0T; L^2)$  stable si :

$$(3.44) \quad \rho(\gamma_n) - 1 \sim \eta \Delta t \quad \text{avec } \eta > 0$$

et alors

$$(3.45) \quad \forall n, ||u_h^n|| \leq \nu e^{\eta T} ||u_{0h}||$$

.  $L^\infty(L^2)$  stable si :

$$(3.46) \quad \rho(\gamma_n) - 1 \sim \eta \Delta t^m \quad \text{avec } \eta > 0 \quad \text{et } m > 1$$

et alors

$$(3.47) \quad \lim_{n \rightarrow \infty} ||u_h^n|| \leq \nu ||u_{0h}||$$

. asymptotiquement stable si :

$$(3.48) \quad \rho(\gamma_n) \leq 1$$

et alors

$$(3.49) \quad \forall n, ||u_h^n|| \leq \nu ||u_{0h}|| .$$

On voit, d'après le théorème 1, que la quantité à étudier pour décider de

la stabilité ou de l'instabilité d'un schéma EF est le rayon spectral  $\rho$  de la matrice d'amplification  $G(\theta, \gamma)$ , et plus précisément son sup en  $\theta$ . Dans de nombreux cas, on trouvera que

$$(3.50) \quad \rho(\gamma) - 1 \sim \delta \gamma^p \quad \delta > 0, p > 0.$$

Le théorème 1 devient alors :

Corollaire 1 :

Sous les hypothèses (3.37) et (3.50) et pour une stratégie  $\{\gamma_n > 0\}$  de gestion de la discrétisation, on a :

- Si  $0 < p < 1$ , le schéma est  $L^\infty(0T; L^2)$  instable pour toutes les stratégies

$$\frac{\Delta t}{h^\alpha} = C \quad \text{c'est-à-dire} \quad \gamma_n = C \Delta t \frac{\alpha-1}{\alpha}$$

correspondant à des valeurs de  $\alpha$  comprises entre 0 et  $+\infty$ . Il est donc inutilisable.

- Si  $p = 1$  le schéma est  $L^\infty(0T; L^2)$  stable sous la condition de stabilité :

$$(3.51) \quad \frac{1}{h} \leq C \quad \text{c'est-à-dire} \quad \gamma_n \leq C \Delta t$$

et la constante  $M(T)$  est donnée par :

$$(3.52) \quad M(T) = v \exp(\delta C T).$$

- Si  $p > 1$  le schéma est :
  - i)  $L^\infty(0T; L^2)$  stable sous la condition de stabilité

$$(3.53) \quad \frac{\Delta t}{h^{p/p-1}} \leq C^{p/p-1} \quad \text{c'est-à-dire} \quad \gamma_n \leq C \Delta t^{1/p}$$

et la constante  $M(T)$  est donnée par

$$(3.54) \quad M(T) = \nu \exp(\delta C^p T)$$

ii)  $L^\infty(L^2)$  stable sous toute condition de la forme

$$(3.55) \quad \frac{\Delta t}{h^{p/p-m}} \leq C^{p/p-m} \quad \text{c'est-à-dire} \quad \gamma_n \leq C \Delta t^{m/p}$$

où  $m$  est un nombre à choisir dans  $]1, p[$ . La constante  $M$  est alors donnée par :

$$(3.56) \quad M = \nu .$$

Remarque 5 : La condition de stabilité (3.51) des schémas avec  $p = 1$  ne permet pas de faire tendre simultanément vers zéro  $\Delta t$  et  $h$  ! De tels schémas sont donc peu recommandables. □

Dans un nombre de cas plus restreint, l'étude du rayon spectral de la matrice  $G(\theta, \gamma)$  montrera que l'on a :

$$(3.57) \quad \rho(\gamma) \leq 1 \quad \text{pour} \quad \gamma \leq \bar{\gamma} .$$

Le théorème 1 donne alors le

Corollaire 2 :

Sous les hypothèses (3.37) et (3.57), le schéma sera asymptotiquement stable pour toute stratégie  $\{\gamma_n > 0\}$  vérifiant la condition de stabilité

$$(3.58) \quad \frac{\Delta t}{h} \leq \bar{\gamma} \quad \text{c'est-à-dire} \quad \gamma \leq \bar{\gamma}$$

et la constante  $M(\Delta t, h)$  est donnée par

$$(3.59) \quad M(\Delta t, h) = \nu .$$

Les conditions de stabilité dans les divers cas sont résumés sur la figure 5.

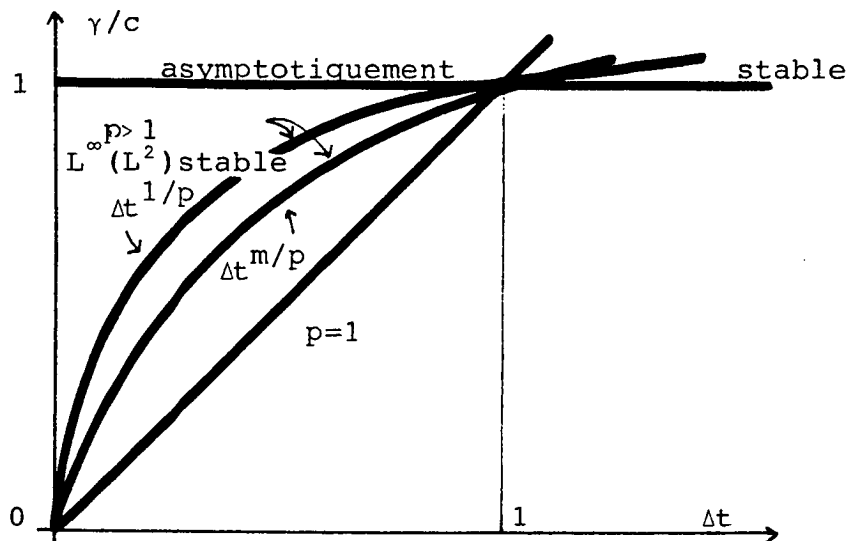


Figure 5 : Les conditions de stabilité sur  $\gamma$  pour diverses valeurs de  $p$  dans (3.50)

Remarque 6 : Une étude plus précise des vecteurs propres de la matrice d'amplification  $G(\theta, \gamma)$  permettrait de déterminer des conditions assurant que, lorsque  $n \rightarrow \infty$   $\|G^n\|/\rho(G) \rightarrow 1$  et donc de remplacer  $\nu$  par 1 dans les résultats précédents. Compte tenu des faibles valeurs de  $\nu$  que l'on trouvera plus loin (cf. figure 10), ce raffinement n'a pas été incorporé dans ce travail. □

Nous venons de voir que, dans de nombreux cas, la condition de stabilité d'un schéma est de la forme

$$(3.60) \quad \frac{\Delta t}{h^\alpha} \leq \text{constante} .$$

Nous poserons donc :

Définition 3 :

Le nombre réel  $\alpha \geq 1$  dans la condition de stabilité (3.60) sera appelé *l'exposant de stabilité du schéma considéré.*

Comme on l'a vu dans les corollaires 1 et 2, on n'a  $\alpha = 1$  (c'est-à-dire la condition la moins contraignante) que lorsque  $\rho(\gamma) \leq 1$ , les autres conditions menant toujours à des exposants  $\alpha > 1$ .

3.2.2 - Une méthode expérimentale numérique

On pourra ici étudier la stabilité pour des normes  $L^p$  avec  $p \neq 2$ . Les définitions 1 et 2 se généralisent immédiatement.

Choisissons :

$$(3.61) \quad \Delta t > 0, \quad \gamma = \frac{\Delta t}{h}, \quad u_{0h}.$$

Alors, lorsque  $n \rightarrow \infty$ , la norme de  $u_h^n$  dans  $L^p(OL)$  soit reste bornée (si le schéma est asymptotiquement stable), soit tend vers l'infini (dans tous les autres cas), c'est-à-dire que le schéma "explose". Dans ce dernier cas, la "vitesse d'explosion" va nous permettre de déterminer l'exposant de stabilité  $\alpha$  du schéma ou l'exposant  $p$  avec lequel  $\rho(\gamma)-1$  tend vers 0 avec  $\gamma$  (cf. (3.50)).

Nous introduisons pour cela un nombre  $\varepsilon > 1$  et nous définissons le nombre de pas de temps  $n^*$  nécessaire pour faire exploser le schéma par :

$$(3.62) \quad n^*(\gamma) = \min\{n \mid \|u_h^n\|_{L^p(OL)} > \varepsilon \|u(n \Delta t)\|_{L^p(OL)}\}$$

où  $u(t)$  est la solution exacte (calculée analytiquement lorsque c'est possible).



L'application  $\gamma \rightarrow n^*(\gamma)$  dépend de

- i) la condition initiale et les conditions aux limites
- ii) le paramètre  $\varepsilon$
- iii) le choix de la norme  $\| \cdot \|_p$
- iv) le pas de temps  $\Delta t$ .

Mais on constatera que certains de ces paramètres ont peu d'influence.

C'est la connaissance de cette application  $\gamma \rightarrow n^*(\gamma)$  qui, sous l'hypothèse que le rayon spectral de  $G$  est de la forme

$$(3.50) \quad \rho(\gamma) - 1 \sim \delta \gamma^{p''} \quad \delta > 0, p'' > 1$$

et donc (cf. corollaire 1) que les conditions de stabilité sont de la forme :

$$(3.60) \quad \frac{\Delta t}{h^\alpha} \leq \text{constante} \quad \text{avec} \quad \alpha = \frac{p''}{p''-m}, \quad 1 \leq m < p''$$

permettra de trouver des estimations de  $\alpha$  et  $p''$ .

Supposons en effet que l'on ait constaté que

$$(3.63) \quad n^*(\gamma) \sim \text{cte} / \gamma^{p'} \quad \text{lorsque} \quad \gamma \ll 1 \quad (p' \geq 1).$$

Alors, l'instant d'explosion  $T^*$  est donné par :

$$(3.64) \quad T^* = n^* \Delta t = \text{cte} \left( \frac{\Delta t}{h^{\alpha'}} \right)^{1-p'} \quad \text{avec} \quad \alpha' = \frac{p'}{p'-1}.$$

Supposons que  $\alpha' > \alpha$ . Alors (3.64) se réécrit

$$T^* = \text{cte} \times \left( \frac{\Delta t}{h^\alpha} \right)^{1-p'} (h^{\alpha'-\alpha})^{p'-1}$$

qui tend vers zéro lorsque  $\Delta t$  et  $h$  tendent vers zéro tout en vérifiant la condition de stabilité (3.60), ce qui est contradictoire. Donc nécessairement

$$\alpha' \leq \alpha.$$

Pour déterminer expérimentalement  $\alpha'$  on suppose que  $\rho(\gamma)$  est de la forme

$$(3.64\text{bis}) \quad \rho(\gamma) - 1 \sim \delta' \gamma^{p'} \quad \delta' > 0, p' > 1.$$

Or, d'après (3.41),  $\|u_h^n\|_{L^2(OL)} \sim \rho(\gamma)^n$  lorsque  $n \rightarrow \infty$ . On obtient donc, d'après (3.62)

$$(1 + \delta' \gamma^{p'})^{n^*} \approx \varepsilon \|u_{0h}\|_{L^2(OL)}$$

d'où, pour  $\gamma \sim 0$  :

$$n^* = \text{cte} / \gamma^{p'}$$

qui est bien (3.63). Si  $n_1^*$  et  $n_2^*$  sont deux instants d'explosion correspondant à deux valeurs  $\gamma_1$  et  $\gamma_2$  de  $\gamma$ , on en déduit

$$(3.65) \quad \begin{cases} p' = -\log(n_1^*/n_2^*) / \log(\gamma_1/\gamma_2) \\ \alpha' = \frac{p'}{p'-1} = \frac{\log(n_1^*/n_2^*)}{\log(n_1^*/n_2^*) + \log(\gamma_1/\gamma_2)} \end{cases}$$

Pour l'étude de la stabilité dans  $L^p(OL)$   $p \neq 2$ , on utilisera, de façon heuristique, les formules (3.65).

Pour le cas linéaire, cette façon d'estimer l'exposant de stabilité  $\alpha$  s'avérera satisfaisante, on obtiendra en effet des valeurs de  $\alpha'$  très proches, et en même temps plus petites que la valeur de  $\alpha$  obtenue par le calcul analytique.

Il sera donc raisonnable d'appliquer cette méthode au cas non linéaire.

### 3.3 - Etude de la stabilité dans le cas linéaire par la méthode de Von Neumann

#### 3.3.1 - Stabilité linéaire du schéma LRG- $\mu$

On vérifie facilement que la matrice d'amplification  $G$  introduite en (3.15) est ici :

$$(3.66) \quad G(\gamma, \theta) = \begin{vmatrix} 1 - \gamma(1 - e^{-i\theta}) & -\sqrt{3} \gamma(1 - e^{-i\theta}) \\ \sqrt{3} \mu \gamma(1 - e^{-i\theta}) & 1 - 3\mu \gamma(1 + e^{-i\theta}) \end{vmatrix}$$

qui peut aussi s'écrire de la façon suivante :

$$(3.67) \quad G(\gamma, \theta) = \begin{vmatrix} 1 - \beta & -\sqrt{3} \beta \\ \sqrt{3} \mu \beta & 1 + 3\mu i \cotg \psi \beta \end{vmatrix}$$

où :  $\psi = \theta/2$  et  $\beta = \gamma(1 - e^{-i\theta}) = 2i \sin \psi e^{-i\psi} \gamma$ .

#### Calcul des valeurs propres

Si l'on peut écrire

$$G(\gamma, \theta) = \text{Id} + \gamma \begin{vmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{vmatrix}$$

$$= \text{Id} + \gamma g$$

alors, les valeurs propres de  $G(\gamma, \theta)$  sont :

$$(3.68) \quad \lambda_{\pm} = 1 + \left\{ \left( \frac{1}{2} \text{tr } g \right) \pm \left\{ \left( \frac{1}{2} \text{tr } g \right)^2 - \det g \right\}^{1/2} \right\} \gamma$$

$$\lambda_{\pm} = 1 + \gamma A_{\pm}(\theta, \mu)$$

où l'on a posé :

$$(3.69) \quad A_{\pm}(\theta, \mu) = \left\{ -\frac{1}{2}((1+3\mu)+(3\mu-1)e^{-i\theta}) \right. \\ \left. \pm \left[ \frac{1}{4}((1+3\mu)+(3\mu-1)e^{-i\theta})^2 - 6\mu(1-e^{-i\theta}) \right]^{1/2} \right\} \gamma$$

. Vérification de la condition (3.37) :  $G(\theta, \gamma)$  diagonalisable pour tout  $\theta \in [-\Pi, \Pi]$  :

On vérifie facilement que le discriminant  $\Delta$  de l'équation caractéristique de  $G(\theta, \gamma)$  vaut

$$\Delta = \frac{1}{4} \beta^2 [(1+3i\mu \cotg\psi)^2 - 12\mu]$$

où  $\beta$  et  $\psi$  sont définis en (3.67). On a donc

$$\begin{aligned} \gamma &= 0 \\ \text{ou bien} \\ \Delta = 0 &\Leftrightarrow \mu = 0 \text{ et } \theta = 0 \\ \text{ou bien} \\ \mu &= 1/12 \text{ et } \theta = \Pi. \end{aligned}$$

Dans les deux premiers cas, la matrice  $G(\theta, \gamma)$  est diagonale, tandis que dans le dernier cas on a :

$$G(\theta, \gamma) = \begin{vmatrix} 1-2\gamma & -2\sqrt{3} \gamma \\ \gamma/2\sqrt{3} & 1 \end{vmatrix}$$

qui n'est pas diagonalisable. La condition (3.37) est donc vérifiée dès que  $(\theta, \mu) \neq (\Pi, 1/12)$ .

. Etude graphique des courbes  $\theta \rightarrow A_{\pm}(\theta, \mu)$  pour différentes valeurs de  $\mu$ .

On a tracé sur les figures 6 et 7 les courbes  $\theta \rightarrow A_{+}(\theta, \mu)$  et  $\theta \rightarrow A_{-}(\theta, \mu)$  dans le plan complexe, ainsi que leur réunion  $A$ , pour différentes valeurs du paramètre  $\mu$ .

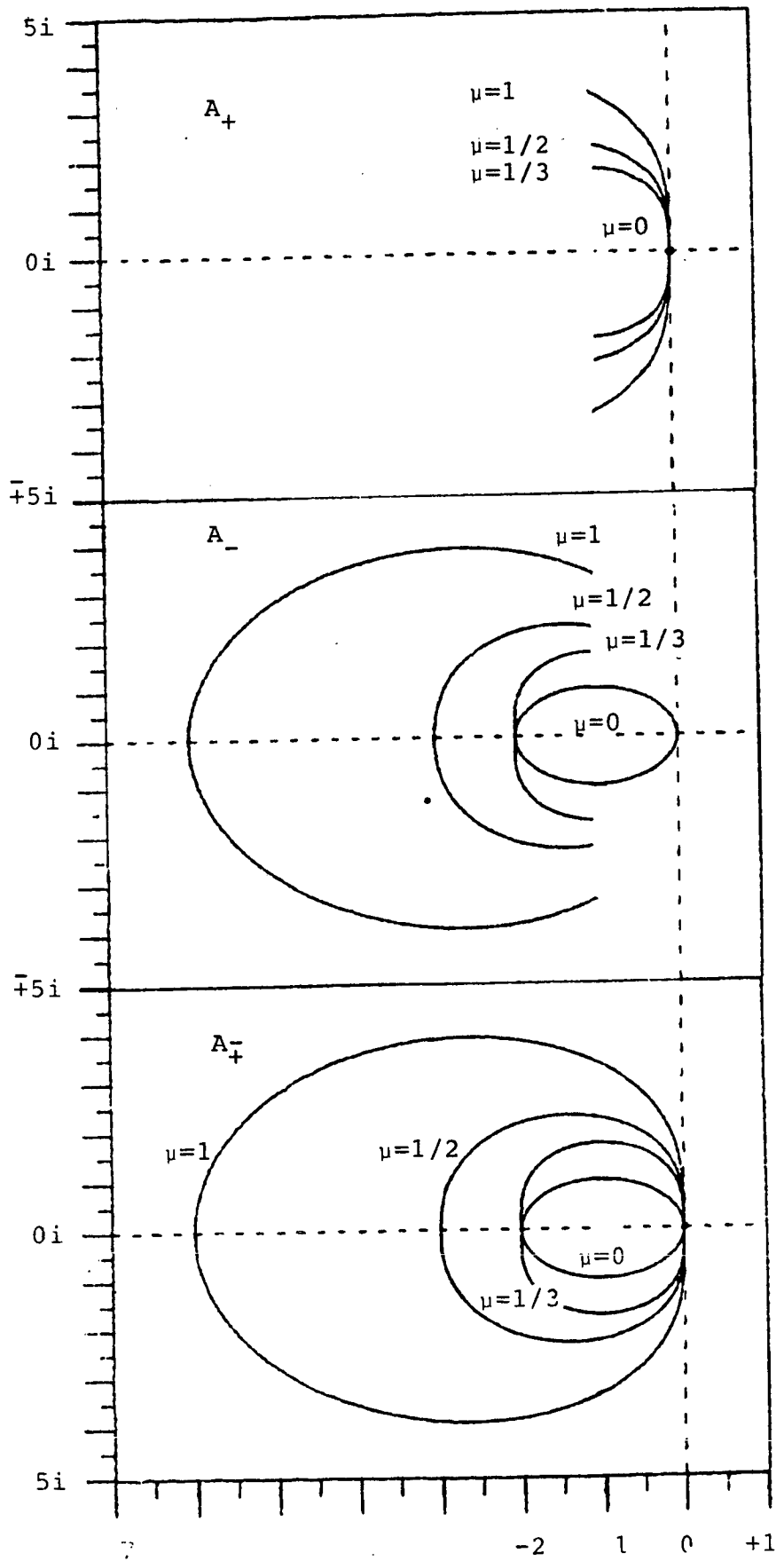


Figure 6 : Les courbes  $A_{\pm}(\mu, \cdot)$  pour  $\mu = 1, 1/2, 1/3$  et 0

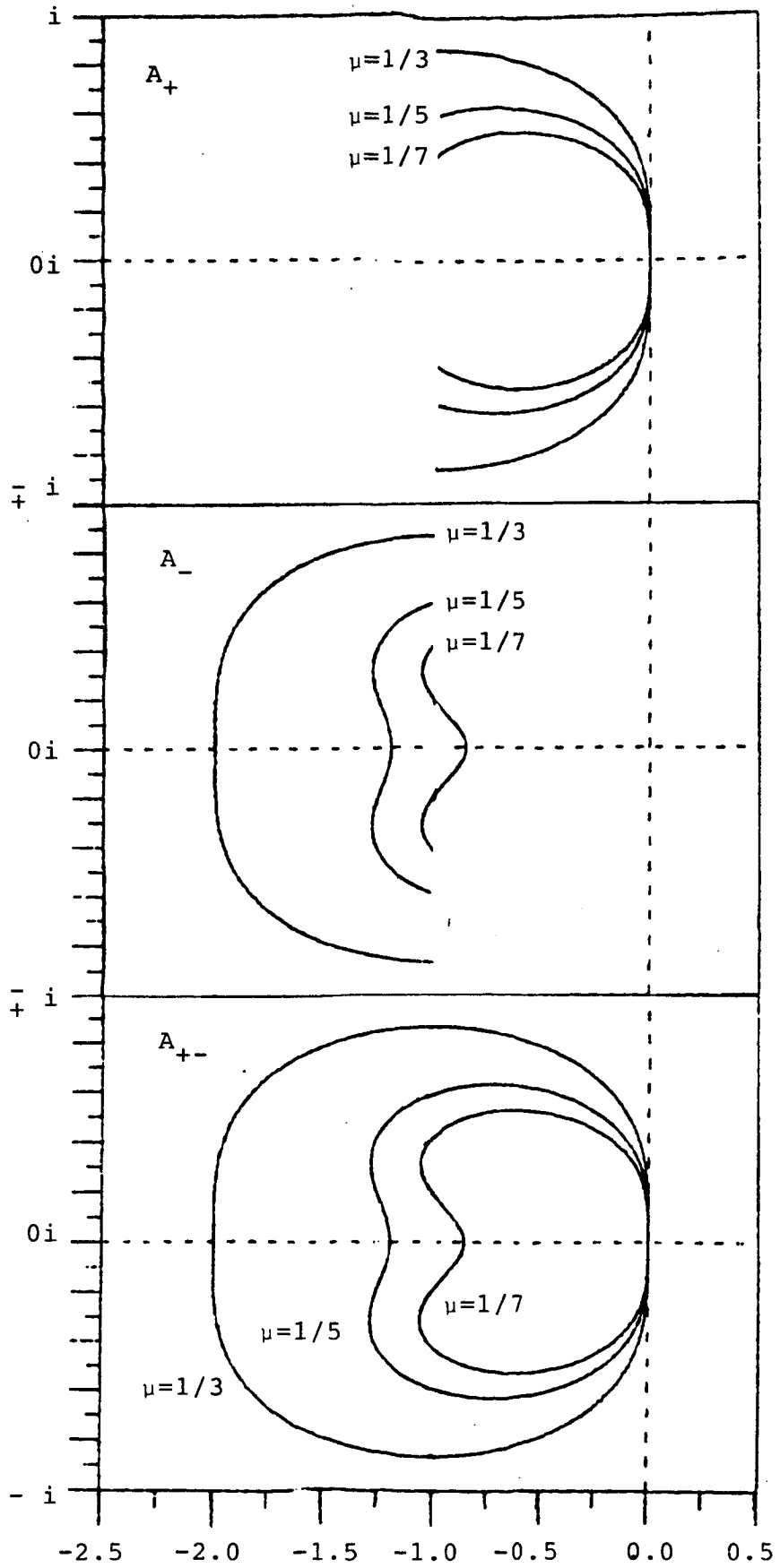


Figure 7 : Les courbes  $A_{\pm}(\mu, \cdot)$  pour  $\mu = 1/3, 1/5$  et  $1/7$ .

On constate que toutes ces courbes ont leur partie réelle négative, et que leur contact en 0 avec l'axe imaginaire semble être supérieur à deux (on verra plus loin qu'il est d'ordre 4).

On conçoit donc que, lorsque  $\gamma \rightarrow 0$ , la courbe

$$\theta \rightarrow 1 + \gamma A_{\pm}(\theta, \mu)$$

finisse par être toute entière à l'intérieur du disque unité, à l'exception d'une petite partie qui, au voisinage de l'origine, sera nécessairement à l'extérieur du cercle unité. On étudiera plus loin analytiquement ce dépassement et son influence sur le rayon spectral de la matrice  $G(\theta, \gamma)$ .

A titre de curiosité, on montre sur les figures 8 a) b) et c) ce qui se passe au voisinage de la valeur  $\mu = 1/12$  pour laquelle la matrice n'est plus diagonalisable : pour  $\mu > 1/12$ , la courbe  $\theta \rightarrow A_{\pm}(\theta, \mu)$  possède une seule composante connexe; pour  $0 < \mu < 1/12$ , elle possède deux composantes connexes, dont une tend vers zéro et l'autre vers le cercle unité centré au point d'affixe  $-1$  lorsque  $\mu \rightarrow 0$  (cf. figure 9). On obtient bien à la limite le schéma de stabilité du schéma DF décentré amont, dont les valeurs propres sont données par

$$\theta \rightarrow 1 + \gamma A_0(\theta)$$

où  $A_0(\theta) = -1 + e^{-i\theta}$  a justement pour image le cercle limite de l'une des composantes connexes de  $A_{\pm}$ .

. Etude du rayon spectral de  $G(\theta, \gamma)$

Les valeurs propres  $\lambda_+$  et  $\lambda_-$  de  $G(\theta, \gamma)$  étant données par

$$\lambda_{\pm} = 1 + \gamma A_{\pm}(\theta, \mu)$$

il résulte clairement de l'étude graphique précédente qu'il existe  $\gamma^* = \gamma^*(\mu)$  tel que :

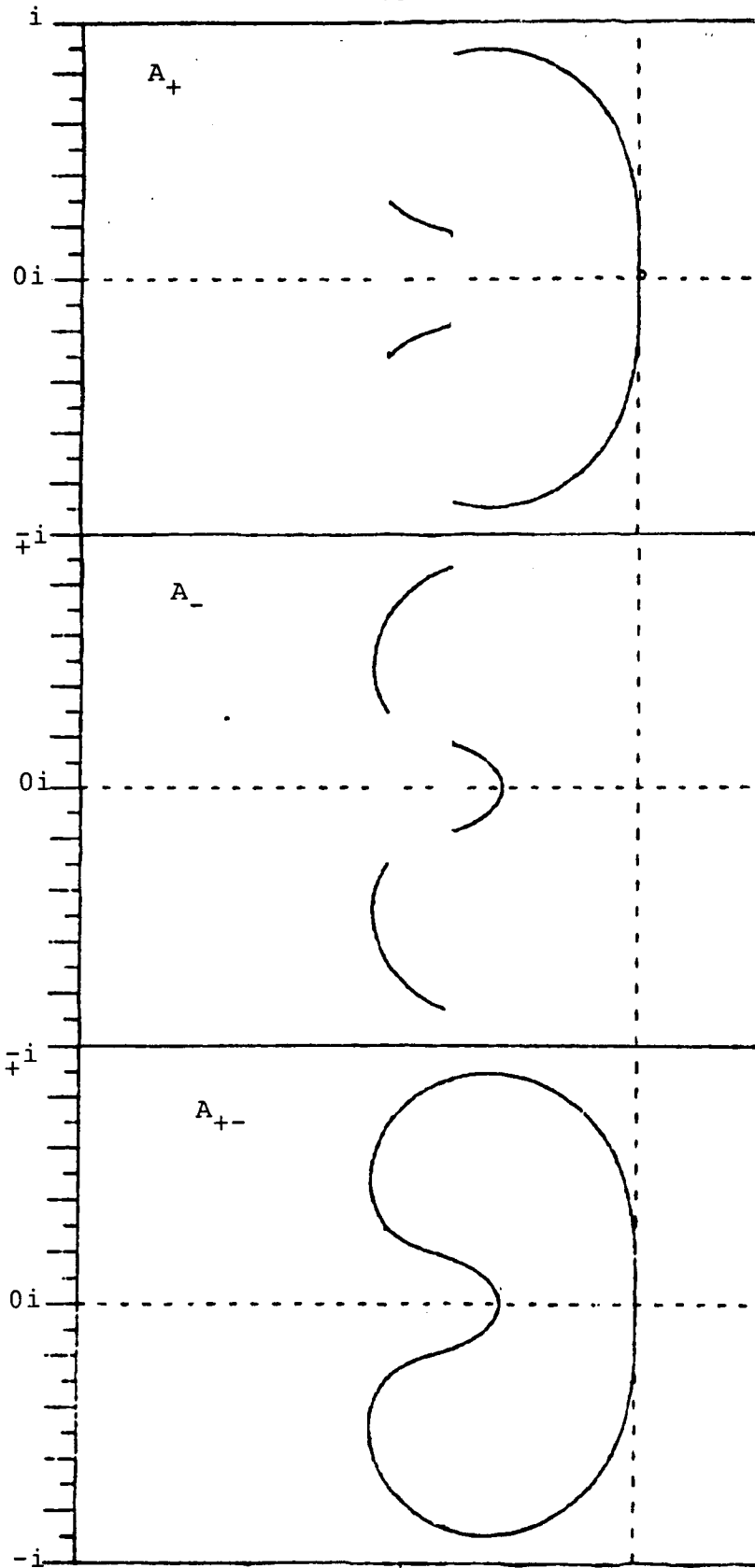


Figure 8 : perte de la connexité des courbes  $A_{\pm}(\mu,.)$

a) le cas  $\mu = 1/11 < 1/12$  : une seule composante connexe



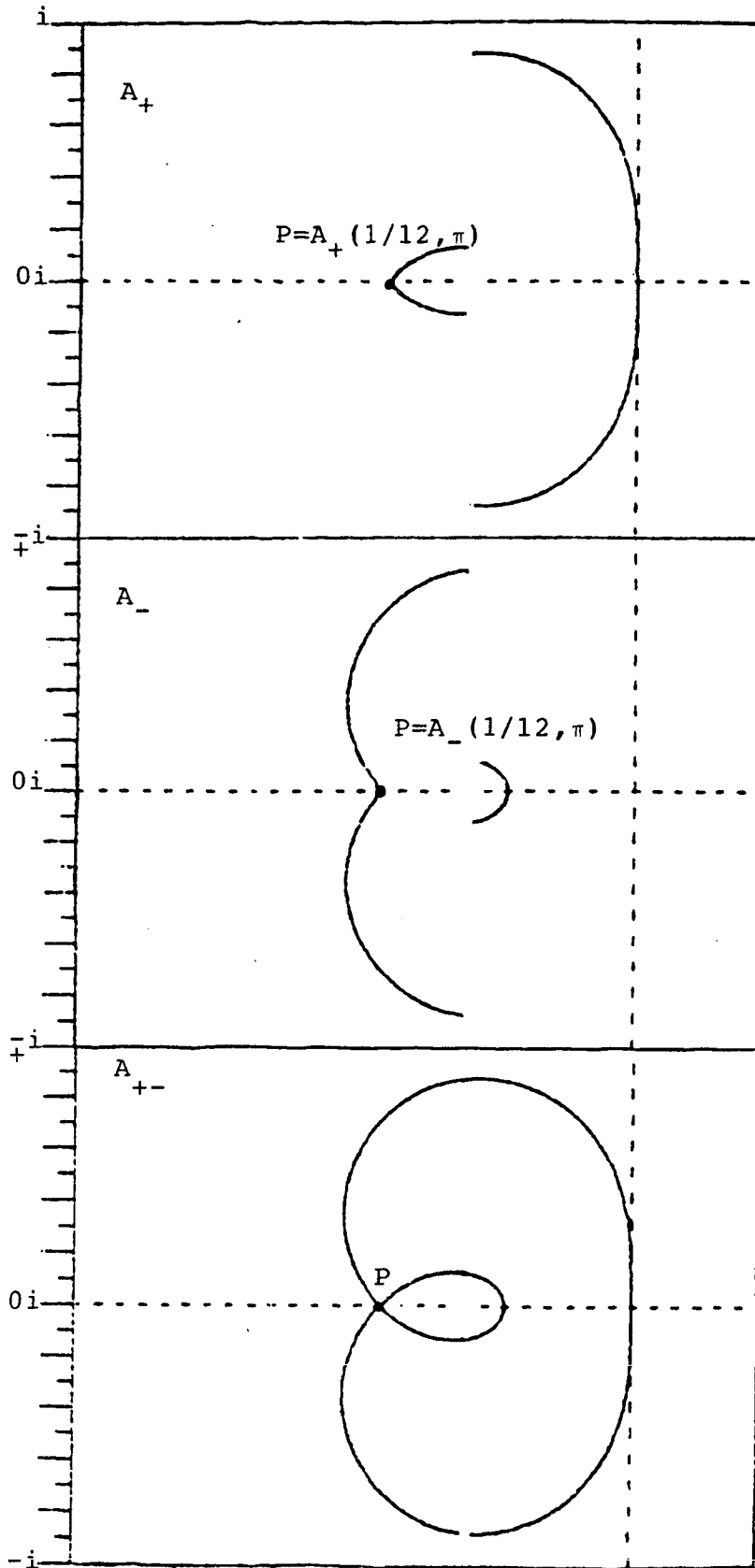


Figure 8 : perte de la connexité des courbes  $A_{\pm}(\mu, \cdot)$

b) le cas critique  $\mu = 1/12$  : pour  $\mu = \pi$  la matrice  $G(\theta)$  n'est pas diagonalisable

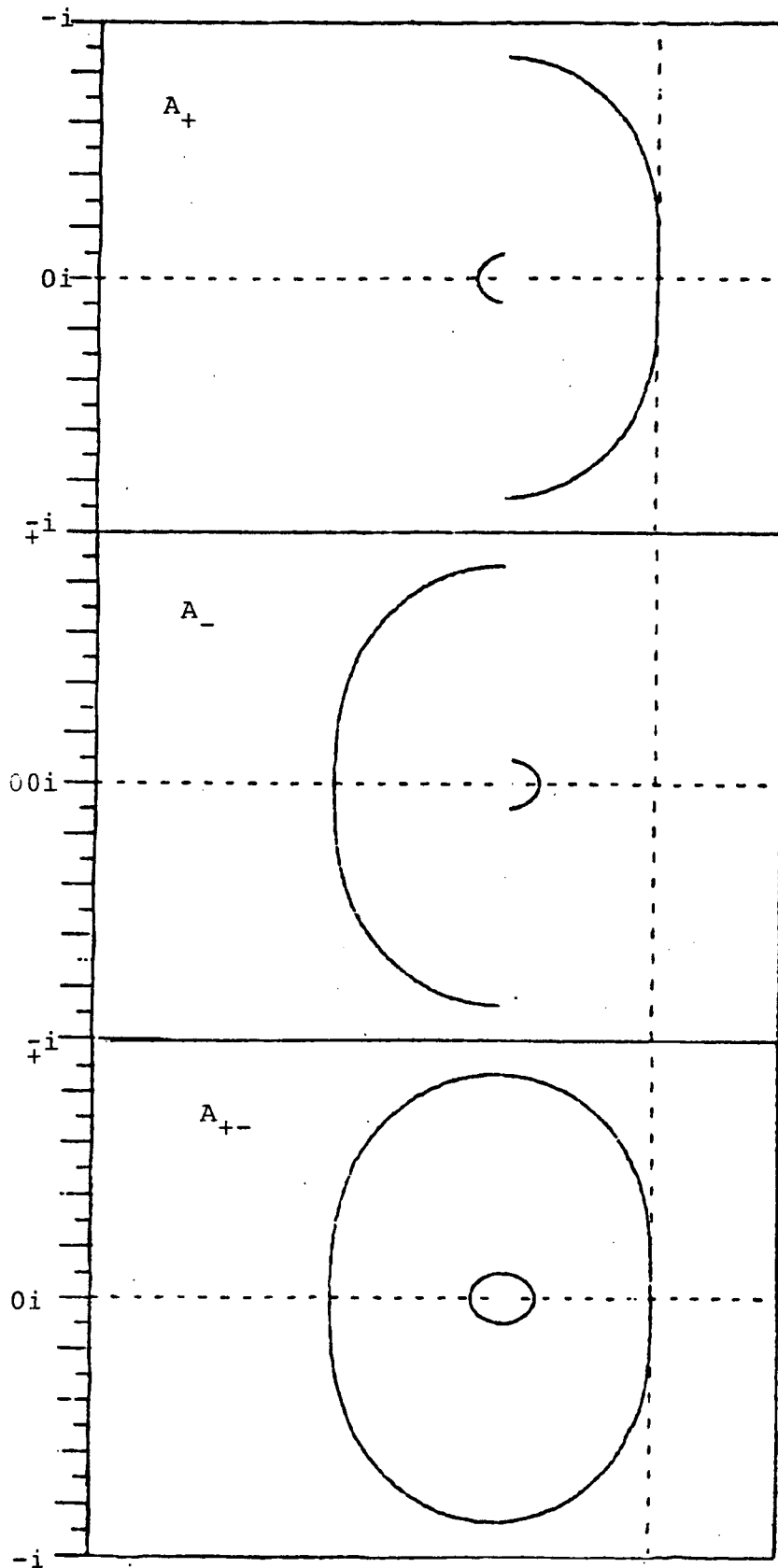


Figure 8 : perte de la connexité des courbes  $A_{\pm}(\mu, \cdot)$

c) la cas  $\mu = 1/13 > 1/12$  : deux composantes connexes

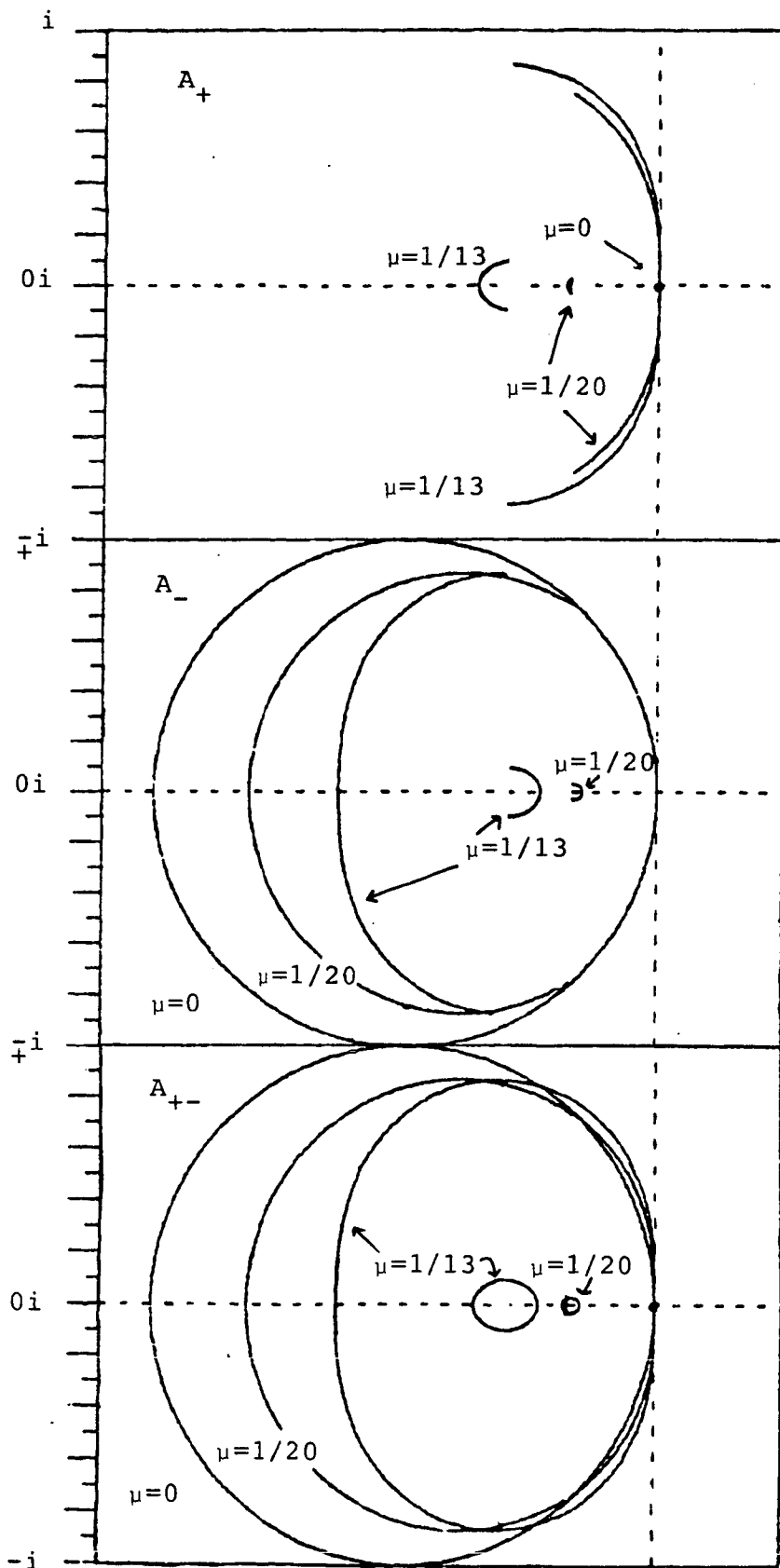


Figure 9 : évolution des courbes  $A_{\pm}(\mu,.)$  lorsque  $\mu \neq 0$ .  
 $A_{+}(\mu,.)$  tend vers l'origine et  $A_{-}(\mu,.)$  tend vers le cercle unité. L'ordre de contact de  $A_{+}(\mu \neq 0,.)$  avec l'axe imaginaire est toujours 4.

$$(3.70) \quad \forall \gamma \in [0, \gamma^*] \quad , \quad \forall \theta \in [-\pi, \pi] \quad , \quad |\lambda_-(\theta)| \leq 1 \quad .$$

Par contre, ce résultat ne sera pas vrai pour la valeur propre  $\lambda_+$ . Suivant les indications données par l'étude graphique, étudions le comportement de  $\lambda_+$  pour des  $\theta$  voisins de 0 :

Proposition 4 : (sur la valeur propre  $\lambda_+(\gamma, \theta, \mu)$ )

Si  $\mu = 0$ , alors

$$\lambda_+(\theta, \mu) = 1 \quad \forall \theta \in [-\pi, \pi]$$

Si  $\mu \neq 0$ , alors  $\lambda_+(\theta, \mu)$  admet le développement limité suivant lorsque  $\theta \rightarrow 0$  :

$$\operatorname{Re}(\lambda_+(\gamma, \theta, \mu)) = 1 - \frac{\gamma^2}{72\mu^2} \theta^4 + o(\theta^6)$$

(3.71)

$$\operatorname{Im}(\lambda_+(\gamma, \theta, \mu)) = -\gamma\theta + o(\theta^3) \quad .$$

Démonstration : cf. Annexe I. □

Il résulte de la proposition 4 que les ordres de contact du cercle unité (centré en -1) et de la courbe  $\theta \rightarrow \lambda_+(\theta, \mu)$  avec l'axe imaginaire sont respectivement 2 et 4, indépendamment de  $\mu$ . Par conséquent, pour n'importe quel  $\gamma > 0$  fixé, il existe toujours un intervalle de la forme  $]0, \theta_0(\gamma)[$  (où  $\theta_0(\gamma) \rightarrow 0$  avec  $\gamma!$ ) sur lequel le module de  $\lambda_+(\gamma, \theta, \mu)$  est strictement plus grand que 1.

Estimons maintenant le module de  $\lambda_+(\theta)$  dans un voisinage de  $\theta = 0$ . D'après la proposition précédente, on peut écrire :

$$e^{-i\theta} = (1 - \theta^2/2) + (-i\theta) + o(\theta^3)$$

$$\lambda_+(\theta) = (1 - \frac{1}{72\mu^2} \gamma\theta^4) + (-i\gamma\theta) + o(\theta^3)$$

autrement dit, près de l'origine le cercle unité et  $\lambda_+$  coïncident, respectivement avec les courbes suivantes

$$y \rightarrow (1 - y^2/2, y)$$

$$y \rightarrow (1 - (72\mu^2\gamma^3)^{-1}y^4, y) .$$

Un calcul simple montre que pour  $y \approx 0$  la distance entre ces deux courbes,

$$d = y^2/2 - (72\mu^2\gamma^3)^{-1}y^4$$

est maximale quand  $y^2 = 18\mu^2\gamma^3$  et donc

$$d \leq d_{\max} = \frac{9}{2} \mu^2 \gamma^3$$

par conséquent, lorsque  $\gamma$  est petit, on a

$$(3.72) \quad \sup_{-\pi \leq \theta \leq \pi} \lambda_+(\theta) = 1 + \frac{9}{2} \mu^2 \gamma^3 + \dots$$

De (3.70) et (3.72), on déduit le

Théorème 2 :

Le rayon spectral de la matrice d'amplification  $G(\theta, \gamma)$  du schéma LRG- $\mu$  vérifie, pour tout  $\mu \in ]0, 1]$  :

$$(3.73) \quad \rho(\gamma) - 1 \approx \frac{9}{2} \mu^2 \gamma^3$$

où

$$(3.39) \quad \rho(\gamma) = \sup_{-\pi \leq \theta \leq \pi} \rho(G(\theta, \gamma)) .$$

. Conditions de stabilité des schémas LRG- $\mu$

D'après la forme (3.66) de  $G(\theta, \gamma)$ , la matrice de passage  $P$  diagonalisant  $G$  est visiblement indépendante de  $\gamma$ . La définition (3.40) se réécrit donc :

$$(3.74) \quad \nu = \sup_{-\pi \leq \theta \leq \pi} \text{Cond } P(\theta) .$$

On a représenté sur la figure 10 la variation de  $\nu$  en fonction du paramètre  $\mu$  d'intégration numérique du schéma. On constate que pour des valeurs de  $\mu$  comprises entre  $1/3$  et  $1$ ,  $\nu$  reste inférieur à  $3$ . On retrouve bien aussi le fait que pour  $\mu = 1/12$ , la matrice n'est plus diagonalisable.

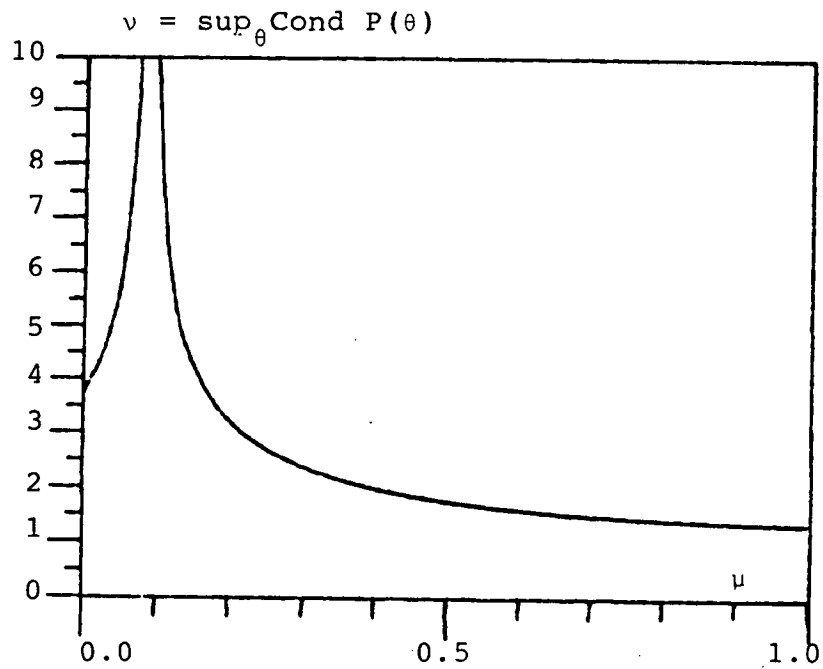


Figure 10 : variation de  $\nu$  avec le paramètre  $\mu$

On peut alors appliquer le corollaire 1 avec  $\delta = \frac{9}{2} \mu^2$  et  $p = 3$  ce qui donne le

Théorème 2 :

Le schéma LRG- $\mu$  est

- . asymptotiquement stable pour  $\mu = 0$  (schéma DF décentré amont)
- . asymptotiquement instable pour  $\mu > 0$
- .  $L^\infty(0T; L^2)$ -stable pour  $\mu > 0$  sous la condition

$$(3.75) \quad \frac{\Delta t}{h^{3/2}} \leq C$$

et alors :

$$(3.76) \quad \|u_h(T)\|_{L^2(OL)} \leq v(\mu) \exp\left(\frac{9}{2} \mu^2 C^2 T\right) \|u_{0h}\|_{L^2(OL)}$$

- .  $L^\infty(L^2)$ -stable pour  $\mu > 0$  sous toute condition

$$(3.77) \quad \frac{\Delta t}{h^\alpha} \leq C, \quad \alpha > 3/2$$

et alors :

$$(3.78) \quad \lim_{n \rightarrow \infty} \|u_h^n\|_{L^2(OL)} \leq v(\mu) \|u_{0h}\|_{L^2(OL)}$$

où  $v(\mu)$  est donné sur la figure 10.

La condition (3.75) montre que, pour calculer la solution  $u(.,T)$ , lorsqu'on divise le pas d'espace  $h$  par  $m$ , il faut diviser le pas de temps par  $m^{1,5}$  et donc faire  $m^{1,5}$  fois plus de pas de temps pour arriver au temps  $T$ .

Remarque 7 : On peut vérifier aisément que l'on peut encore garantir la stabilité sous la seule condition

$$(3.79) \quad \frac{\Delta t}{h} \leq C$$

à condition de diminuer  $\mu$  (c'est-à-dire changer de schéma en se rapprochant du schéma DF décentré amont) en même temps que  $\Delta t$  en respectant la condition

$$(3.80) \quad \frac{\mu^2}{\Delta t} \leq C'$$

On obtient alors :

$$(3.81) \quad \| |u_h^n| \|_{L^2(OL)} \leq \nu \exp\left(\frac{9}{2} C' C^3 T\right) \| |u_{Oh}| \|_{L^2(OL)} . \quad \square$$

Remarque 8 : On a vu au § 2 que le schéma LRG- $\mu$  pouvait s'écrire en termes de degrés de liberté  $(a_i, \tilde{b}_i)$  (cf. (1.9)) au lieu de  $(a_i, b_i)$ . Si l'on peut toujours calculer  $(a_i, \tilde{b}_i)$  en fonction de  $(a_i, b_i)$ , il est impossible par contre d'exprimer  $(a_i, b_i)$  en fonction de  $(a_i, \tilde{b}_i)$  dans le cas d'un problème périodique, car on a seulement

$$a_i = a_i$$

$$b_i - b_{i-1} = 2\tilde{b}_i - 2 \frac{a_i - a_{i-1}}{h}$$

ce qui ne détermine  $b_i$  que si une des pentes est fixée.

Il semble donc que les schémas LRG- $\mu$  exprimés en termes des degrés de liberté  $(a_i, b_i)$  ou  $(a_i, \tilde{b}_i)$  ne soient pas complètement équivalents. On a donc calculé la matrice d'amplification  $\tilde{G}(\theta, \gamma)$  correspondant au schéma (1.15) (exprimé en terme de  $a_i^n$  et  $\tilde{c}_i^n = h/2\sqrt{3} \tilde{b}_i^n$ ) :



$$(3.82) \quad \tilde{G}(\theta, \gamma) = \begin{vmatrix} 1 & -2\sqrt{3} \gamma \\ \sqrt{3} \gamma \mu (1 - e^{-i\theta}) & 1 - \gamma (3\mu (1 + 2^{-i\theta}) + (1 - e^{-i\theta})) \end{vmatrix}$$

On vérifie alors que les valeurs propres de  $\tilde{G}(\theta, \gamma)$  sont exactement les mêmes que celle de  $G(\theta, \gamma)$  donnée en (3.66). Il n'y a donc aucun gain en stabilité à utiliser  $(a_i, \tilde{b}_i)$  au lieu de  $(a_i, b_i)$ .  $\square$

### 3.3.2 - Stabilité linéaire de deux schémas DF

- Commençons par le schéma double-amont (2.16), qui est rappelons-le d'ordre 2 en espace et 1 en temps.

Il est facile de voir que le coefficient d'amplification dans la méthode de Von Neumann est

$$(3.83) \quad G(\theta, \gamma) = 1 - \frac{\gamma}{2}(3 - 4e^{-i\theta} + e^{-2i\theta})$$

d'où on déduit

$$|G(\theta, \gamma)|^2 = 1 + 4\gamma \sin^2 \frac{\theta}{2} [\gamma - (2 - 3\gamma) \sin^2 \frac{\theta}{2}] .$$

Pour  $0 < \gamma \leq \frac{2}{3}$  on voit sur la figure 11 que  $|G(\theta, \gamma)|^2$  est maximum pour  $\sin^2 \frac{\theta}{2} = \gamma / (2 - 3\gamma)$  de sorte que

$$(3.84) \quad \rho(\gamma) = \max_{-\pi \leq \theta \leq \pi} G(\theta, \gamma) = 1 + \frac{\gamma^3}{2 - 3\gamma} .$$

Cette formule montre que le schéma double-amont n'est pas asymptotiquement stable, mais qu'il a des propriétés de stabilité tout à fait comparables à celles des schémas LRG- $\mu$  (comparer à cet effet (3.84) et (3.73)). En particulier pour  $\mu = 1/5$  (intégration numérique par la méthode des trapèzes),

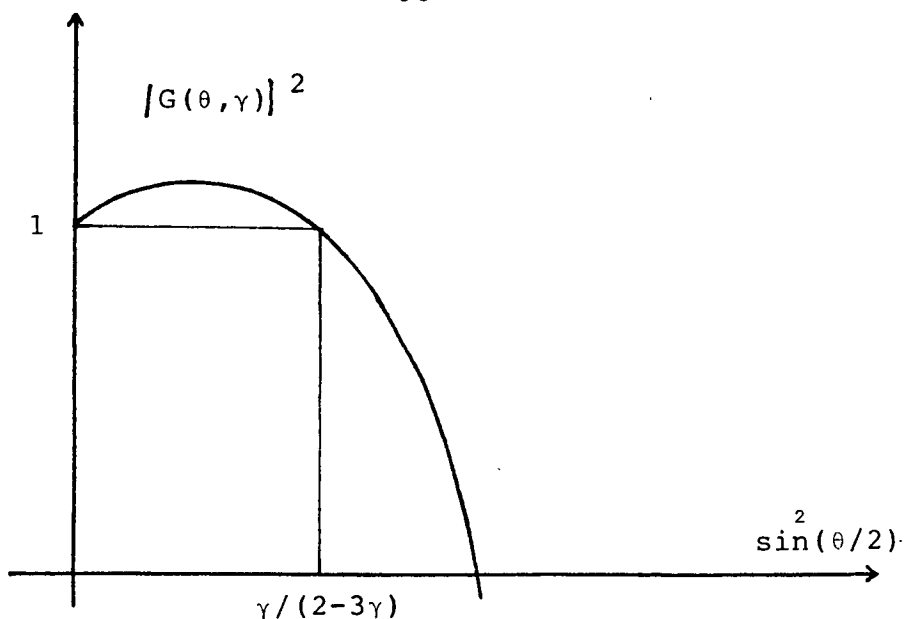


Figure 11 : coefficient d'amplification du schéma double amont

on a  $\rho(\gamma) - 1 \approx 1/2 \gamma^3$  tant pour  $LRG - \frac{1}{3}$  que pour le schéma double-amont.

Si l'on veut mettre en évidence numériquement l'instabilité asymptotique de ce schéma, il suffit de considérer, pour  $\gamma = \frac{1}{2}$ , la solution correspondant à la donnée initiale de la figure 12, dont la norme  $L^2$  doit augmenter avec le coefficient d'amplification  $1 + \frac{1}{4}$ .

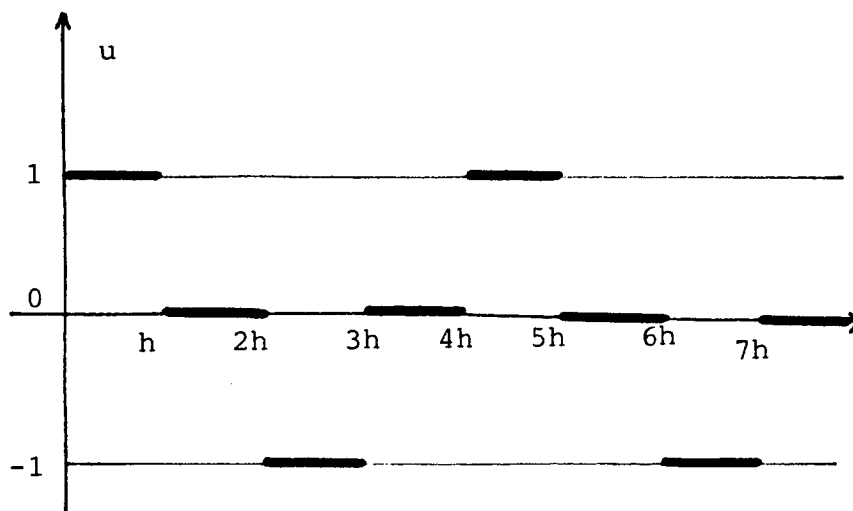


Figure 12 : condition initiale à utiliser avec  $\gamma = 1/2$  pour tester l'instabilité asymptotique du schéma double amont.

- Etudions maintenant la stabilité du schéma double-amont avec Lax Wendroff (2.17)

Le coefficient d'amplification  $G$  dépend maintenant de  $\theta$ ,  $\gamma$  et  $h$  :

$$(3.85) \quad G(\theta, \gamma, h) = 1 - \frac{\gamma}{2} \left[ 3 - \frac{1}{h} - 2 \left( 2 - \frac{1}{h} \right) e^{-i\theta} + \left( 1 - \frac{1}{h} \right) e^{-2i\theta} \right]$$

d'où on déduit (après quelques calculs fastidieux) :

$$(3.86) \quad |G|^2 = 1 + 4\gamma \sin^2 \frac{\theta}{2} \left\{ \gamma - \frac{1}{h} + \left( \frac{1}{h} - 1 \right) \left[ 2 + \left( \frac{1}{h} - 3 \right) \gamma \right] \sin^2 \frac{\theta}{2} \right\} .$$

Si on suppose

$$(3.87) \quad h \leq 1 \quad \text{et} \quad 0 < \gamma \leq 1$$

on vérifie que  $|G|^2$  dépend de  $\sin^2 \frac{\theta}{2}$  comme indiqué sur la figure 13.

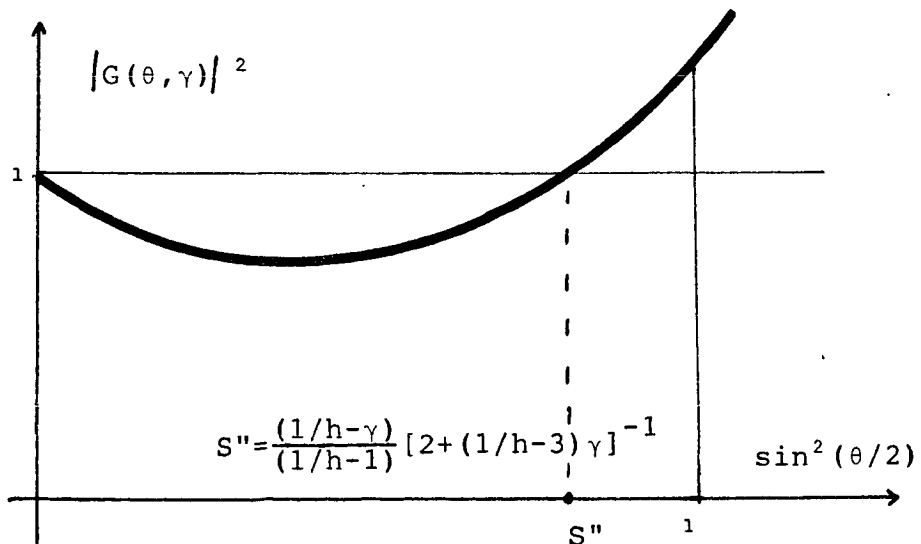


Figure 13 : le coefficient d'amplification du schéma double amont avec Lax Wendroff

Ainsi,  $|G|^2$  est maximum pour  $\sin^2 \frac{\theta}{2} = 1$  et vaut :

$$(3.88) \quad \rho(\gamma) = \max_{-\pi \leq \theta \leq \pi} |G|^2 = 1 + 4\gamma \left\{ \frac{1}{h} - 2 + \gamma \left[ 1 + \left( \frac{1}{h} - 1 \right) \left( \frac{1}{h} - 3 \right) \right] \right\} .$$

On constate, d'après le corollaire 1, que le schéma double amont avec Lax-Wendroff n'est pas asymptotiquement stable (on s'y attendait !) mais surtout qu'il n'est  $L^\infty(0T; L^2)$ -stable que sous la condition de stabilité très forte

$$\frac{1}{h} \leq C$$

qui interdit de raffiner simultanément le pas de temps et le pas d'espace. Ici l'amélioration du schéma en temps ( $O(\Delta t^2)$  au lieu de  $O(\Delta t)$ ) par la technique de Lax-Wendroff s'est faite au détriment de la stabilité ( $\rho(\gamma) - 1$  en  $O(\frac{\gamma}{h})$  au lieu de  $O(\gamma^3)$ ).

Par exemple, avec la solution initiale de la figure 14, correspondant à  $\theta = \pi$ , la norme  $L^2$  de la solution doit croître à chaque pas de temps avec le facteur  $\rho(\gamma)$  donné par (3.88), qui vaut par exemple 41 lorsqu'on prend  $\gamma = 1/4$  et  $h = 1/10$ .

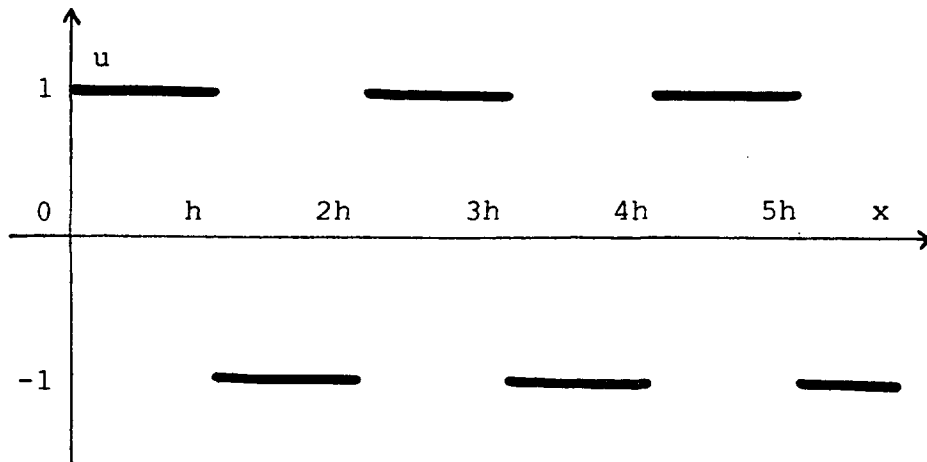


Figure 14 : condition initiale à utiliser pour tester l'instabilité asymptotique du schéma double amont avec Lax Wendroff.

### 3.4 - Etude numérique expérimentale de la stabilité des schémas LRG- $\mu$

On utilisera ici la méthodologie et les notations développées au § 3.2.2.

#### 3.4.1 - Le cas linéaire

Pour faire l'étude numérique de la stabilité des schémas LRG- $\mu$ , on a pris comme condition initiale :

$$u_0(x) = \begin{cases} 1 & \text{si } x = 0^- \\ 0 & \text{si } x \in ]0,1] \end{cases}$$

et comme conditions aux bords

$$\begin{aligned} u_0(t) &= 1 & \forall t \in [0,1] \\ u_1(t) &= 0 & \forall t \in [0,1] . \end{aligned}$$

Avec ce choix on est sûr d'exciter *toutes* les fréquences et par conséquent, au moins pour le cas linéaire, on doit retrouver l'exposant de stabilité  $\alpha = 3/2$ .

On avait aussi signalé, au paragraphe 3.2.2, que l'application

$$\gamma \rightarrow n^*(\gamma)$$

dépendait du paramètre  $\varepsilon$ , de la norme  $L^p$  et du pas de temps  $\Delta t$ , car

$$n^* = \min\{m : \|u_h^m\|_{L^p(0,1)} > \|u(m\Delta t)\|_{L^p} \varepsilon\} .$$

En fait, les expériences numériques montrent que la variation de  $n^*$  par rapport à  $\Delta t$  n'est pas significative.

Par contre  $n^*$  varie beaucoup avec  $\varepsilon$  et la norme  $L^p$ , mais, et c'est ce qui est important, l'estimation de l'exposant de stabilité est à peu près le même.

En effet, pour  $\mu = 1$ , on a pris successivement

$$\varepsilon = \sqrt{2}, 2, 3$$

$$L^p = L^\infty, L^2.$$

Pour chacun de ces cas, on a calculé  $p'$  et  $\alpha'$  à l'aide des formules (3.65) pour différentes valeurs de  $\gamma$ . Les résultats figurent dans les tableaux 1 et 2 ainsi que sur la figure 15. On constate que *l'exposant  $p'$  varie, pour les plus petites valeurs de  $\gamma$ , entre 3,17 et 3,28, alors que la valeur  $p$  donnée par la méthode de Von Neumann est de 3.*

D'autre part, la courbe  $n^*(\gamma)$  définie en (3.63) doit avoir, en données logarithmiques, une pente égale à  $-p'$  et doit donc être une droite si l'hypothèse (3.63) ou (3.64bis) est vérifiée. C'est bien ce que l'on constate sur la figure 15 pour les valeurs de  $\gamma$  suffisamment faibles.

Finalement, on voit que :

- l'hypothèse (3.63), c'est-à-dire  $n^* = C/\gamma^{p'}$  est assez raisonnable pour  $\gamma$  petit
- le choix du paramètre  $\varepsilon$  et de la norme  $L^p$  ont une influence importante sur la constante  $C$ , mais négligeable sur l'exposant  $p'$ .

$\gamma$	$n^*$	$p'$	$\alpha'$	
1	1	-	-	
1/2	1	0	0	
1/3	2	1.71	2.41	
1/4	6	3.82	1.35	( $\mu = 1$ )
1/5	13	3.46	1.41	
1/6	28	4.21	1.31	$\epsilon = \sqrt{2}$
1/7	48	3.50	1.40	
1/8	79	3.73	1.37	$L^p = L^\infty$
1/9	116	3.26	1.44	
1/10	169	3.57	1.39	

$\gamma$	$n^*$	$p'$	$\alpha'$	
1	1	-	-	
1/2	2	1.00	00	
1/3	5	2.26	1.79	
1/4	14	3.58	1.39	( $\mu = 1$ )
1/5	33	3.84	1.35	
1/6	63	3.55	1.39	$\epsilon = 2$
1/7	109	3.56	1.39	
1/8	172	3.42	1.41	$L^p = L^\infty$
1/9	242	2.90	1.52	
1/10	338	3.17	1.46	

$\gamma$	$n^*$	$p'$	$\alpha'$	
1	1	-	-	
1/2	3	1.58	2.71	
1/3	8	2.42	1.70	
1/4	21	3.35	1.42	( $\mu = 1$ )
1/5	46	3.51	1.40	
1/6	85	3.37	1.42	$\epsilon = 3$
1/7	161	4.14	1.32	
1/8	220	2.34	1.75	$L^p = L^\infty$
1/9	327	3.36	1.42	
1/10	462	3.28	1.44	

Tableau 1 : l'application  $\gamma \rightarrow n^*$  pour 3 valeurs de  $\epsilon$ ,  
pour  $L^p = L^\infty$ .

$\gamma$	$n^*$	$p'$	$\alpha'$	
1	1	-	-	
1/2	2	1.00	00	
1/3	5	2.26	1.79	$(\mu = 1)$
1/4	17	4.25	1.31	
1/5	43	2.61	1.04	$\epsilon = \sqrt{2}$
1/6	89	3.99	1.33	
1/7	156	3.64	1.38	$L = L^2$
1/8	256	3.71	1.37	
1/9	395	3.68	1.37	
1/10	576	3.58	1.38	

$\gamma$	$n^*$	$p'$	$\alpha'$	
1	1	-	-	
1/2	3	1.58	2.71	
1/3	10	2.97	1.51	$(\mu = 1)$
1/4	25	3.19	1.46	
1/5	58	3.77	1.36	$\epsilon = 2$
1/6	119	3.71	1.37	
1/7	200	3.65	1.38	$L = L^2$
1/8	324	3.61	1.38	
1/9	488	3.98	1.40	
1/10	706	3.51	1.40	

$\gamma$	$n^*$	$p'$	$\alpha'$	
1	1	-	-	
1/2	4	2.00	2.00	
1/3	11	2.99	1.67	$(\mu = 1)$
1/4	32	3.71	1.37	
1/5	71	3.57	1.39	$\epsilon = 3$
1/6	136	3.56	1.39	
1/7	253	4.03	1.33	$L = L^2$
1/8	376	2.97	1.51	
1/9	565	3.46	1.41	
1/10	816	3.99	1.40	

Tableau 2 : l'application  $\gamma \rightarrow n^*$  pour 3 valeurs de  $\epsilon$ ,  
pour  $L^p = L^2$



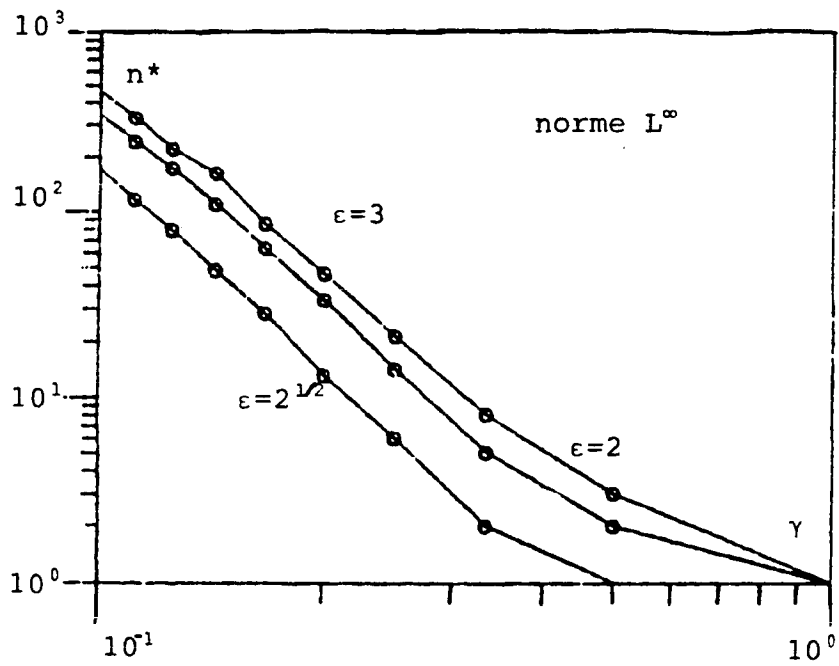
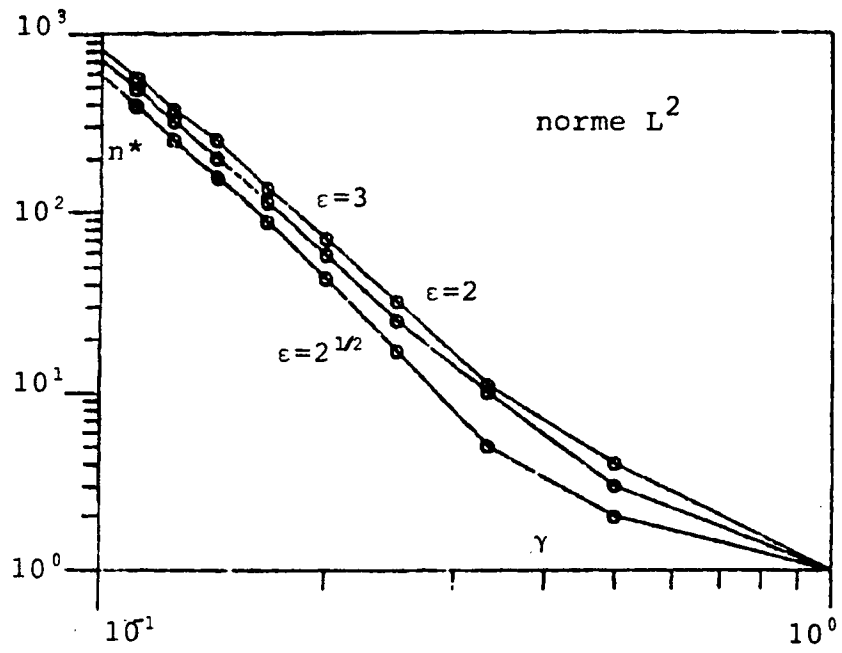


Figure 15 : nombre d'itérations  $n^*$  en fonction du rapport  $\gamma = \Delta t / \Delta x$

Maintenant, et pour finir l'étude du cas linéaire, on s'intéresse à la dépendance de l'application  $\gamma \rightarrow n^*$  par rapport au paramètre  $\mu$ .

*On a fait deux expériences numériques. Dans la première, en prenant toujours  $\varepsilon = \sqrt{2}$  et  $L^p = L^\infty$ , on a obtenu l'application  $\gamma \rightarrow n^*$  pour  $\mu = 1/3$  et on l'a comparée avec l'application correspondante pour  $\mu = 1$ .*

Les résultats se trouvent sur le tableau 3 et la figure 16. On peut constater que :

- l'estimation de l'exposant de stabilité est à peu près le même,
- avec  $\mu = 1/3$ , on peut faire environ 3 fois plus de pas de temps qu'avec  $\mu = 1$ .

*Dans la deuxième expérience, on a fixé  $\gamma = 1/10$  et on a évalué  $n$  pour plusieurs valeurs de  $\mu$ . On a ainsi obtenu l'application*

$$\mu \rightarrow n^*(\gamma = 1/10, \mu)$$

que l'on montre sur le tableau 4 et la figure 17. Si, guidé par le résultat théorique (3.73), on suppose que, pour  $\gamma$  fixé, on a

$$\rho(\gamma) - 1 \sim C \mu^{q'} \quad q' > 0$$

on trouve facilement, suivant une analyse semblable à celle faite au § 3.2.2 pour évaluer  $p'$  et  $\mu'$ , que  $q'$  peut être estimé par la formule :

$$q' = -\log(n_1^*/n_2^*)/\log(\mu_1/\mu_2)$$

c'est donc cette valeur qui apparaît dans la dernière colonne du tableau 4. On constate que la valeur  $q' = 2,19$  trouvée pour les faibles valeurs de  $\mu$  est proche de l'exposant  $q = 2$  de  $\mu$  dans la formule théorique (3.73).

$\gamma$	$n^*(\mu=1/3)$	$n^*(\mu=1/3)/n^*(\mu=1)$	$\alpha$
1	1	1	-
1/2	2	2	
1/3	8	4	1.41
1/4	21	3.5	1.42
1/5	44	3.4	1.43
1/6	81	2.89	1.43
1/7	140	2.92	1.39
1/8	221	2.80	1.41
1/9	325	2.80	1.44
1/10	458	2.71	1.44

Tableau 3 : l'application  $\gamma \rightarrow n^*$  pour  $\mu = 1/3$   
Comparaison avec  $\gamma \rightarrow n^*$  pour  $\mu = 1$ .

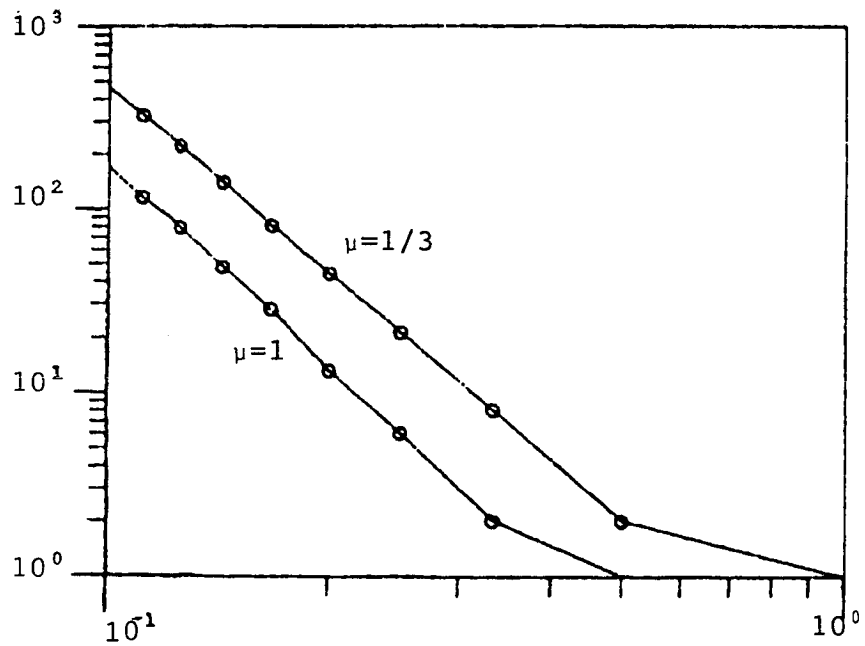


Figure 16 : comparaison des courbes  $(\gamma, n^*(\gamma))$  pour  $\mu = 1/3$  et  $\mu = 1$ .

$\mu$	$n^*(\mu=1/10, \mu)$	$q'$
1	169	
1/2	222	+2.54
1/3	458	+1.79
1/4	977	+2.63
1/5	1625	+2.28
1/6	2424	+2.19

Tableau 4 : l'application  $\mu \rightarrow n^*(\gamma=1/10, \mu)$

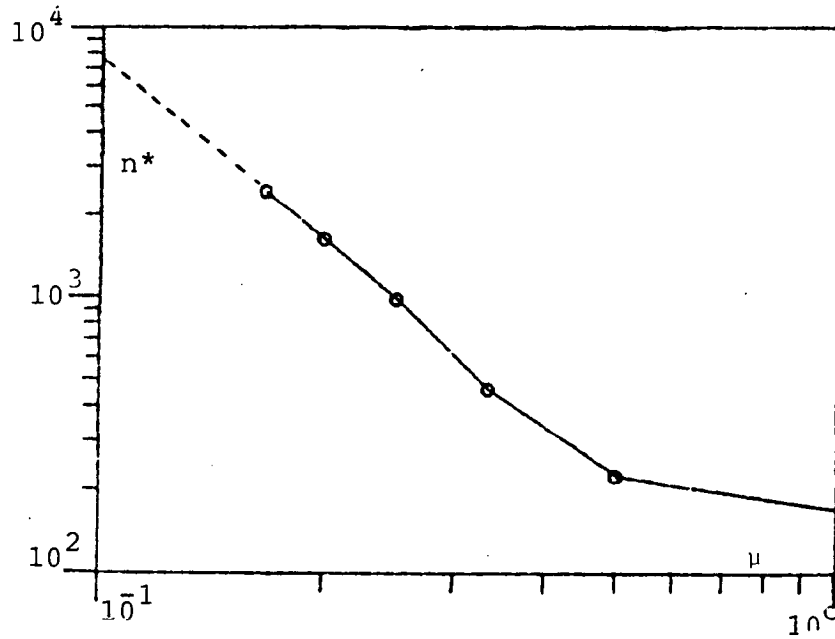


Figure 17 : la courbe  $\mu \rightarrow n^*(\gamma=1/10, \mu)$

En conclusion de cette étude expérimentale de la stabilité linéaire, on peut dire que les résultats théoriques obtenus au § 3.3 sont confirmés expérimentalement :

	Théorique	Expérimental
(3.89) $\rho(\gamma)-1$	$3\mu^2\gamma^3$	$C \mu^2,19, \gamma 3,17 \text{ à } 3,57$

### 3.4.2 - Le cas non linéaire

On a considéré ici l'approximation de l'équation (3.8) où l'on a pris pour  $f(u)$  une fonction coïncidant avec  $u$  dans l'intervalle  $[0,1]$ ; précisément :

$$(3.90) \quad f(u) = \begin{cases} 0 & \forall u < 0 \\ u & \forall u \in [0,1] \\ 1+30(u-1) & \forall u > 1 \end{cases}$$

avec les mêmes conditions aux bords et la même condition initiale que pour le cas linéaire.

Pour l'obtention de  $n^*$  on a encore pris  $\varepsilon = \sqrt{2}$  et  $L^P = \bar{L}^\infty$ .

Les expériences numériques que l'on a faites sont différentes de celles du cas précédent. En effet, pour  $\mu = 1$  et  $\mu = 1/3$ , valeurs de correspondant à l'intégration par Simpson et par la méthode des trapèzes, respectivement, du terme

$$\int_{x_i}^{x_{i+1}} \frac{\partial u_h}{\partial t} v \, dx \quad , \quad v \text{ étant une fonction test}$$

on a fait des essais en changeant la façon avec laquelle on intègre le terme non-linéaire,

$$\int_{x_i}^{x_{i+1}} f(u_h) \frac{\partial v}{\partial x} dx$$

pour lequel on a testé quatre méthodes d'intégration

- la méthode de Gauss
- la méthode de Simpson
- la méthode des trapèzes
- la méthode des carrés.

Les résultats se trouvent sur les tableaux que l'on présente à la suite (rappelons que  $n^*$  et  $\alpha'$  sont définis en (3.62), (3.63) et (3.64)) :

$\gamma$	$n^*(\mu=1)$	$\alpha'(\mu=1)$	$n^*(\mu=1/3)$	$\alpha'(\mu=1/3)$
1	1	-	1	-
1/2	1	0	2	-
1/3	2	2.41	6	1.59
1/4	2	0	18	1.36
1/5	4	1.47	>1500	-
1/6	10	1.25	-	-
1/7	79	1.08	-	-
1/8	76	0.23	-	-
1/9	88	5.09	-	-
1/10	97	-12.20	-	-
1/11	> 2200	-	-	-

Tableau 5 : l'application  $\gamma \rightarrow n^*$ .

Pour la méthode d'intégration de Simpson

$\gamma$	$n^*(\mu=1)$	$\alpha'(\mu=1)$	$n^*(\mu=1/3)$	$\alpha'(\mu=1/3)$
1/5	3	-	7	-
1/6	5	-	63	1.090
1/7	4	-	64	0.093
1/8	5	-	286	1.098
1/9	-	-	122	0.261
1/10	73	-	116	0.324
1/11	-	-	366	1.098
1/12	-	-	239	0.830
1/13	-	-	522	1.114
1/14	100	-	764	1.260
1/15	>1200	-	224	-
1/16	>1200	-	-	-

Tableau 6 : l'application  $\gamma \rightarrow n^*$

Pour la méthode d'intégration de Gauss.

$\gamma$	$n^*(\mu=1)$	$\alpha'(\mu=1)$	$n^*(\mu=1/3)$	$\alpha'(\mu=1/3)$
1/7	10	-	87	-
1/8	11	0.42	89	-0.21
1/9	17	1.37	103	5.16
1/10	104	1.06	108	1.94
1/11	119	1.02	>550	-
1/12	128	1.04	-	-
1/13	143	3.60	-	-
1/14	>700	-	-	-

Tableau 7 : l'application  $\gamma \rightarrow n^*$   
Pour la méthode des trapèzes

$\gamma$	$n^*(\mu=1)$	$n^*(\mu=1/3)$
1/20	13	25
1/24	-	199
1/25	17	>1300
1/30	>1500	

Tableau 8 : l'application  $\gamma \rightarrow n^*$   
Pour la méthode d'intégration des carrés

Deux remarques importantes sont à faire. D'abord le comportement de  $n^*$  par rapport à  $\gamma$ , et par conséquent celui de  $\alpha$  par rapport à  $\gamma$ , est *très irrégulier*, contrairement à ce qui se passe pour le cas linéaire. En plus, il paraît exister un  $\gamma^* = \gamma^*(\mu)$  tel que pour  $\gamma \in [0, \gamma^*]$  le schéma est stable!

D'autre part, comme dans le cas linéaire, on peut pour  $\gamma$  fixé, faire plus de pas de temps pour  $\mu = 1/3$  que pour  $\mu = 1$ .

Dans le tableau que l'on montre à la suite, on trouve les  $\gamma^*$  en fonction de  $\mu$  et de la méthode d'intégration.

Méthode d'intégration	$\mu = 1$	$\mu = 1/3$
Simpson	1/11	1/5
Gauss	1/15	>1/15
trapèzes	1/14	1/11
carrés	1/30	1/25

Tableau 9 : le paramètre  $\gamma^*$  en fonction de  $\mu$  et des méthodes d'intégration numérique du terme non linéaire.

On y constate que, sauf pour le cas de Gauss,  $\gamma^*(\mu=1)$  est plus petit que  $\gamma^*(\mu=1/3)$ . D'autre part, à  $\mu$  fixé, c'est la méthode de Simpson pour laquelle on obtient le plus petit  $\gamma^*$ .

On peut donc conclure que :

- les propriétés de stabilité du schéma LRG- $\mu$  du cas linéaire sont très différentes de celles du cas non linéaire.
- si pour le cas linéaire l'exposant de stabilité était  $\alpha = 3/2$  pour le cas non linéaire considéré, il semble être seulement  $\alpha = 1$ .

La façon adoptée de prolonger  $u \rightarrow f(u) = u$  hors l'intervalle  $[0,1]$  s'avère donc très efficace pour l'amélioration des propriétés de stabilité du schéma (dans ce cas précis)

- pour  $\mu$  fixé, c'est la méthode de Simpson qui donne les résultats de stabilité les meilleurs
- pour la méthode d'intégration fixée, c'est avec  $\mu = 1/3$  que l'on obtient les résultats les meilleurs.

### 3.5 - Conclusion sur la stabilité du schéma LRG- $\mu$

L'étude mathématique des schémas LRG- $\mu$  dans le cas linéaire, c'est-à-dire



$$u \rightarrow f(u) = u$$

montre que ces schémas sont, au sens des définitions 1 et 2 du § 3.1, *asymptotiquement instables* (quel que soit  $\mu$ ), mais sont par contre  $L^\infty(L^2)$ -*stables* sous toute condition de stabilité de la forme

$$\frac{\Delta t}{h^\alpha} \leq C \quad , \quad \alpha > 3/2$$

et  $L^\infty(0T;L^2)$ -*stables* sous la condition de stabilité

$$\frac{\Delta t}{h^{3/2}} \leq C \quad .$$

Dans ce dernier cas (cf. (3.75) (3.76)) le facteur d'amplification de la norme  $L^2$  de la solution approchée entre les instants 0 et T est proportionnel à  $\exp(9/2 \mu^2 C^2 T)$  qui tend vers 1 lorsque  $\mu$  décroît.

Par contre, l'étude numérique du cas

$$(3.90) \quad u \rightarrow f(u) = \begin{cases} 0 & \forall u < 0 \\ u & \forall u \in [0,1] \\ 1+30(u-1) & \forall u > 1 \end{cases}$$

qui peut être considéré comme une perturbation du cas linéaire, donne des *résultats très différents*.

En effet, la méthode numérique utilisée, qui dans le cas linéaire confirme tout à fait les résultats précédents, montre que, dans le cas non linéaire considéré, on obtient la *stabilité asymptotique* sous une condition de stabilité du type :

$$(3.91) \quad \frac{\Delta t}{h} \leq \gamma^*$$

où  $\gamma^*$  est une constante expérimentale qui est maximum lorsque le terme non linéaire est intégré par la méthode de Simpson, et qui, dans ce cas, croît de  $1/11$  à  $1/5$  lorsque  $\mu$  décroît de  $1$  (intégration exacte du terme d'évolution) à  $1/3$  (intégration par la méthode des trapèzes du terme d'évolution). Ce  $\gamma^*$  est très sensible au choix de la formule de quadrature utilisée dans le terme non linéaire : il ne vaut que  $1/30$  lorsqu'on utilise la méthode des carrés.

En conclusion, on constate que l'introduction de la non-linéarité (3.90) permet de stabiliser les schémas LRG- $\mu$ , mais au prix d'une condition de stabilité (3.91) très contraignante en pratique puisque  $\gamma^*$  est assez petit devant  $1$ .

On va donc proposer au § 4 des modifications des schémas LRG- $\mu$  qui en augmentent la stabilité tout en conservant la possibilité de passage à la dimension 2.

#### 4 - ETUDE NUMERIQUE COMPARATIVE DES PROPRIETES DE CONVERGENCE DU SCHEMA LRG ET DES SCHEMAS LRG-MODIFIES

##### 4.1 - Introduction

Dans ce chapitre, on fait une étude purement numérique des *propriétés de convergence* du schéma LRG et d'autres schémas, qui possèdent de meilleures *propriétés de stabilité*, que l'on a réuni sous le nom de schémas "LRG-modifiés".

Dans le cas linéaire, c'est-à-dire quand  $f(u) = c u$ ,  $c \in \mathbb{R}$ , le théorème d'équivalence de Lax, voir par exemple [ 5 ], établit l'équivalence entre la stabilité et la convergence d'un schéma consistant. On connaît

donc, d'après les résultats des chapitres précédents, les conditions sous lesquelles notre schéma LRG est convergent, et par conséquent l'unique intérêt d'une étude numérique de sa convergence serait celui de déterminer l'ordre de convergence lorsque les données initiales ne sont pas "régulières" (il n'y a presque pas de résultats théoriques dans cette direction).

Mais il faut remarquer que :

- i) le théorème d'équivalence de Lax n'est plus valable dans le cas non-linéaire, et par conséquent, faute de résultats théoriques dans ce cas, l'unique moyen de connaître les propriétés de convergence des schémas considérés est via une étude numérique.
- ii) Si le schéma est "vraiment non-linéaire" (vnl), c'est-à-dire, si la relation entre  $u_h^{n+1}$  et  $u_h^n$  n'est pas linéaire lorsque  $f(u) = c u$ ,  $c \in \mathbb{R}$ , alors le théorème de Lax n'est plus valable et, encore une fois, il ne nous reste que la méthode expérimentale pour étudier la convergence de ces schémas. En fait, tous les schémas LRG-modifiés que l'on considère ici sont vnl.

Il est donc vraiment nécessaire de réaliser cette étude. Le plan du chapitre est le suivant : au paragraphe 4.2, on présente les schémas groupés sous le nom "LRG-modifié", et on établit le protocole de l'étude numérique de la convergence. Au paragraphe 4.3, on étudie le *cas linéaire*  $f(u) = u$  avec un échelon comme condition initiale. La solution exacte est le *transport à vitesse constante de cet échelon*. C'est un des cas les plus difficiles à traiter numériquement. Au paragraphe 4.4, on considère un *cas non linéaire* avec la même condition initiale. La fonction  $f$  possède un point d'inflexion et dans ce cas la solution présente *un choc et une région dans laquelle elle est de classe  $\mathcal{C}^1$* . On est dans le cas d'une équation de Buckley-Leverett. Au paragraphe 4.5, on considère un autre *cas non linéaire*. La condition initiale est aussi un échelon, déplacé, mais dans ce cas, *la solution est  $\mathcal{C}^0$  pour  $t > 0$* . Le paragraphe 4.6 est dédié à l'étude de l'effet sur la convergence d'un paramètre "d'antidiffusion" du schéma dit G-1/2. Finalement, au paragraphe 4.7, on trouve les conclusions de cette étude.

## 4.2 - Description de la méthode d'étude de la convergence

### 4.2.1 - Présentation des schémas numériques

Tous les schémas que l'on va considérer sont des schémas explicites à deux niveaux de temps. Deux d'entre eux, le schéma de Godunov, G, et notre schéma LRG sont des schémas linéaires (c'est-à-dire, appliqués au cas  $f(u) = c u$  ils donnent lieu à une relation linéaire entre  $u_h^{n+1}$  et  $u_h^n$ ). Le reste sont des schémas vraiment non-linéaires.

On a voulu considérer le schéma de Godunov pour deux raisons principalement

- i) c'est un schéma très robuste et convergent (sous une certaine condition de stabilité). Malgré le fait qu'il diffuse beaucoup au niveau des chocs, il est monotone et positif.
- ii) Notre schéma LRG peut être considéré comme la version d'ordre 2 du schéma de Godunov (d'ordre 1 seulement).

On va donc inclure le schéma G dans notre étude, comme point de base pour les comparaisons à faire.

Avant de présenter les schémas "LRG-modifiés", il est convenable d'introduire quelques notations et de décrire les schémas G et LRG.

On va utiliser une discrétisation uniforme en temps et en espace. Comme d'habitude  $\Delta t$  représente le pas de temps et  $h$  le pas d'espace. On prendra  $x_i = i h$  et  $t^n = n \Delta t$ .  $u_h^n$  représente la solution du problème discret, donnée par un schéma numérique. Sur un élément  $]x_i, x_{i+1}[$  on peut écrire  $u_h^n$  de la façon suivante :

$$u_h^n(x) = \frac{u_{i+1/2}^n}{2} + \frac{2}{h}(x-x_{i+1/2}) \frac{u_{i+1/2}^n}{2}$$

où  $x_{i+1/2} = \frac{1}{2}(x_i + x_{i+1})$ . Il est donc clair que :

i)  $\underline{u}_{i+1/2}^n$  est la valeur moyenne de  $u_h^n$  dans l'élément  $]x_i, x_{i+1}[$ .

ii)  $u_h^n(x_{i+1}^-) = \underline{u}_{i+1/2}^n + \tilde{u}_{i+1/2}^n \equiv u_{i+1}^{-n}$

$u_h^n(x_i^+) = \underline{u}_{i+1/2}^n - \tilde{u}_{i+1/2}^n \equiv u_i^{+n}$ .

Avec cette notation, on a (pour simplifier, on ne va écrire que les équations!),

Le schéma de Godunov (G)

$$\star. \underline{u}_{i+1/2}^{n+1} = \underline{u}_{i+1/2}^n - \frac{\Delta t}{h} (f(\xi_{i+1}^n) - f(\xi_i^n))$$

où  $\xi_i^n$  est défini par

$$(G) \quad \text{signe}(\underline{u}_{i+1/2}^n - \underline{u}_{i-1/2}^n) (f(\xi_i^n) - f(k)) \leq 0$$

$$k = \theta \underline{u}_{i+1/2}^n + (1-\theta) \underline{u}_{i-1/2}^n \quad \forall \theta \in [0,1]$$

$$\star. \tilde{u}_{i+1/2}^{n+1} = 0$$

Il faut remarquer que ce schéma est convergent si la condition CFL  $\leq 1$  est vérifiée. Par contre, le schéma LRG est convergent (dans le cas linéaire) sous la condition de stabilité, beaucoup plus sévère, CFL  $\leq c h^{1/2}$ .

Le schéma LRG

$$\star. \underline{u}_{i+1/2}^{n+1} = \underline{u}_{i+1/2}^n - \frac{\Delta t}{h} (f(\xi_{i+1}^n) - f(\xi_i^n))$$

(LRG) où  $\xi_i^n$  est définie par

$$\text{signe}(u_{i+1}^{-n} - u_i^{+n}) (f(\xi_i^n) - f(k)) \leq 0$$

$$k = \theta u_{i+1}^{-n} + (1-\theta) u_i^{+n} \quad \forall \theta \in [0,1]$$

$$\begin{aligned} * \cdot \underline{u}_{i+1/2}^{n+1} &= \underline{u}_{i+1/2}^n - 3 \frac{\Delta t}{h} (f(\xi_{i+1}^n) + f(\xi_i^n)) \\ &+ 6 \Delta t \int_{x_i}^{x_{i+1}} f(u_h^n(x)) dx . \end{aligned}$$

Maintenant, on va introduire les schémas "LRG-modifiés". Comme on l'avait dit, ces schémas sont vraiment non-linéaires et possèdent deux étapes : dans la première étape, on calcule les moyennes  $\underline{u}_h$ , et dans la deuxième le reste des degrés de liberté  $\underline{u}_h$ . Les schémas que l'on va décrire sont MUSCL, G-1/2, LGRP et BIP.

Le schéma MUSCL (cf. [4])

$$* \cdot \underline{u}_{i+1/2}^{n+1} = \underline{u}_{i+1/2}^n - \frac{\Delta t}{h} (f(\xi_{i+1}^n) - f(\xi_i^n))$$

où  $\xi_i^n$  est défini comme dans LRG.

$$\begin{aligned} \text{(MUSCL)} \quad * \cdot \underline{u}_{i+1/2}^{n+1} &= \text{signe}(\underline{u}_{i+3/2}^{n+1} - \underline{u}_{i+1/2}^{n+1}) \max \{0, \psi_{i+1/2}^{n+1}/2\} \\ \psi_{i+1/2}^{n+1} &= \min \{ |\underline{u}_{i+3/2}^{n+1} - \underline{u}_{i+1/2}^{n+1}|, |\underline{u}_{i+1/2}^{n+1} - \underline{u}_{i-1/2}^{n+1}| \cdot \\ &\cdot \text{signe}((\underline{u}_{i+3/2}^{n+1} - \underline{u}_{i+1/2}^{n+1})(\underline{u}_{i+1/2}^{n+1} - \underline{u}_{i-1/2}^{n+1})) \}. \end{aligned}$$

La façon avec laquelle on calcule  $\underline{u}_h$  rend ce schéma vraiment non-linéaire. Il faut remarquer que :

$$\text{i) si } (\underline{u}_{i+3/2}^{n+1} - \underline{u}_{i+1/2}^{n+1})(\underline{u}_{i+1/2}^{n+1} - \underline{u}_{i-1/2}^{n+1}) \leq 0$$

$$\text{alors } \underline{u}_{i+1/2}^{n+1} = 0, \text{ c.à.d. si } \underline{u}_h^{n+1}(x) \in [a, b] \quad \forall x$$

$$\text{alors } \underline{u}_h^{n+1}(x) \in [a, b] \quad \forall x .$$

$$\text{ii) Supposons } \underline{u}_{i+3/2}^{n+1} \geq \underline{u}_{i+1/2}^{n+1}, \text{ alors}$$

$$\underline{u}_{i+1}^{n+1} = \underline{u}_{i+3/2}^{n+1} - \underline{u}_{i+3/2}^{n+1} \geq \underline{u}_{i+1/2}^{n+1} + \underline{u}_{i+1/2}^{n+1} = \underline{u}_{i+1}^{n+1} .$$

Il n'est pas difficile de voir que ceci implique que la variation de  $u_h^{n+1}$  est exactement égale à la variation de  $u_h^n$ .

La convergence de ce schéma a été établie, [ 4 ], sous une condition additionnelle que l'on ne va pas considérer ici.

Le calcul de  $u_h^{n+1}$  se fait à partir de  $u_h^{n+1}$  d'une façon très simple : la pente de  $u_h^{n+1}$  sur l'élément  $]x_i, x_{i+1}[$  est donc choisie comme la plus petite des pentes déterminées par  $u_{i+3/2}^{n+1}$ ,  $u_{i+1/2}^{n+1}$  et  $u_{i-1/2}^{n+1}$ . Un autre choix, qui semble favorable, est celui de prendre cette pente égale à la pente déterminée cette fois-ci par des *valeurs de Godunov* associées à  $u_h^{n+1}$ . On obtient ainsi le schéma BIP.

Le schéma BIP

$$* \quad u_{i+1/2}^{n+1} = u_{i+1/2}^n - \frac{\Delta t}{h} (f(\xi_{i+1}^n) - f(\xi_i^n))$$

où  $\xi_i^n$  est défini comme dans LRG.

$$* \quad u_{i+1/2}^{n+1} = \text{signe}(u_{i+3/2}^{n+1} - u_{i+1/2}^{n+1}) \max\{0, \psi_{i+1/2}^{n+1}\}$$

(BIP)

$$\psi_{i+1/2}^{n+1} = \min\{ |\xi_{i+1}^{n+1} - \xi_i^{n+1}|/2, |u_{i+3/2}^{n+1} - u_{i+1/2}^{n+1}|, |u_{i+1/2}^{n+1} - u_{i-1/2}^{n+1}| \cdot \text{signe}((u_{i+3/2}^{n+1} - u_{i+1/2}^{n+1})(u_{i+1/2}^{n+1} - u_{i-1/2}^{n+1})) \}$$

où  $\xi_{i+1}^{n+1}$  est défini comme dans G.

Encore une fois, on a un schéma vraiment non-linéaire. Il faut remarquer que :

i) si  $(u_{i+3/2}^{n+1} - u_{i+1/2}^{n+1})(u_{i+1/2}^{n+1} - u_{i-1/2}^{n+1}) \leq 0$

alors  $u_{i+1/2}^{n+1} = 0$ . Comme pour le schéma MUSCL, si  $u_h^{n+1}(x) \in [a, b] \quad \forall x$ ,

alors  $u_h^{n+1}(x) \in [a, b] \quad \forall x$ .

ii) Supposons  $\underline{u}_{i+3/2}^{n+1} \geq \underline{u}_{i+1/2}^{n+1}$ , alors, on a

$$u_{i+1}^+, u_{i+1}^- \in [\underline{u}_{i+1/2}^{n+1}, \underline{u}_{i+3/2}^{n+1}]$$

iii) Ceci veut dire qu'en général, la variation de  $u_h^{n+1}$  est plus grande que celle de  $\underline{u}_h^{n+1}$ .

On ne connaît aucun résultat sur la convergence de ce schéma.

Avant de passer à la description des schémas LRGP et G-1/2, il faut remarquer un point important : le calcul de  $\underline{u}_{i+1/2}^{n+1}$  se fait via une projection. En effet, on peut réécrire ce calcul de la façon suivante :

. calculer  $\underline{v}_{i+1/2}^{n+1} = \frac{1}{2} (\xi_{i+1}^{n+1} - \xi_i^{n+1})$

. calculer l'ensemble des valeurs admissibles  $\mathcal{A}_{i+1/2}^{n+1}$  de la façon suivante :

$$\mathcal{A}_{i+1/2}^{n+1} = I(0, \underline{u}_{i+3/2}^{n+1} - \underline{u}_{i+1/2}^{n+1}) \cap I(0, \underline{u}_{i+1/2}^{n+1} - \underline{u}_{i-1/2}^{n+1})$$

où  $I(a,b) = [a,b]$  si  $a \leq b$ ,  $= [b,a]$  si non.

. prendre  $\underline{u}_{i+1/2}^{n+1}$  égal à la projection de  $\underline{v}_{i+1/2}^{n+1}$  sur  $\mathcal{A}_{i+1/2}^{n+1} = [\alpha, \beta]$ , c'est-à-dire

$$\underline{u}_{i+1/2}^{n+1} = \begin{cases} \alpha, & \text{si } \underline{v}_{i+1/2}^{n+1} \\ \underline{v}_{i+1/2}^{n+1}, & \text{si } \underline{v}_{i+1/2}^{n+1} \in [\alpha, \beta] \\ \beta & \text{autrement.} \end{cases}$$

On peut utiliser cette idée pour définir un autre schéma qui calcule  $\underline{u}_{i+1/2}^{n+1}$  avec plus de précision. Et une idée naturelle est celle



de prendre  $u_{i+1/2}^{n+1}$  égale à la valeur obtenue par le schéma LRG lui-même. Si on fait ceci, on obtient ce que l'on va appeler le schéma LRGP (P = projection).

Le schéma LRGP

$$\star. \underline{u}_{i+1/2}^{n+1} = \underline{u}_{i+1/2}^n - \frac{\Delta t}{h} (f(\xi_{i+1}^n) - f(\xi_i^n))$$

où  $\xi_i^n$  est défini comme dans LRG

$$\star. \underline{u}_{i+1/2}^{n+1} = \text{signe}(u_{i+3/2}^{n+1} - u_{i+1/2}^{n+1}) \max\{0, \psi_{i+1/2}^{n+1}\}$$

(LRG)

$$\psi_{i+1/2}^{n+1} = \min\{|\underline{u}_{i+1/2}^{n+1}|, |u_{i+3/2}^{n+1} - u_{i+1/2}^{n+1}|, |u_{i+1/2}^{n+1} - u_{i-1/2}^{n+1}| \cdot \text{signe}((u_{i+3/2}^{n+1} - u_{i+1/2}^{n+1})(u_{i+1/2}^{n+1} - u_{i-1/2}^{n+1}))\}$$

où  $\underline{u}_{i+1/2}^{n+1} = (\underline{u}_{i+1/2}^{n+1})$  calculée par LRG).

Il est facile de voir que ce schéma possède les propriétés i), ii) et iii) du schéma BLP. Il est difficile d'autre part de déterminer lequel des schémas est le meilleur. En principe, tenant compte du fait que  $\underline{u}_{i+1/2}^{n+1}$  est calculé en utilisant un schéma d'ordre 2 en espace, il est raisonnable de penser que LRGP est "meilleur" que BLP, mais comme on verra sur les expériences numériques, ceci n'est pas toujours vrai.

Encore une fois, on ne possède pas d'information sur les propriétés de convergence de ce schéma. Par contre, on sait démontrer que sous la condition  $CFL \leq 1/2$ , le schéma G-1/2, que l'on présente ci dessous, converge vers la solution entropique.

Le schéma G-1/2

$$\star. \underline{u}_{i+1/2}^{n+1} = \underline{u}_{i+1/2}^n - \frac{\Delta t}{h} (f(\xi_{i+1}^n) - f(\xi_i^n)) - \frac{\Delta t}{h} (a_{i+1}^n - a_i^n)$$

où  $\xi_i^n$  est définie comme dans G,

$$a_i^n = \text{signe}(\underline{u}_{i+1/2}^n - \underline{u}_{i-1/2}^n) \cdot \max\{0, \psi_{i+1/2}, \psi_{i-1/2}\}$$

$$(G-1/2) \quad \psi_{i+1/2} = \min\{ |f(\xi_i^n) - f(\xi_i^n)|, |c_0 h^\alpha |f(\xi_{i+1}^n) - f(\xi_i^n)|, \\ c_0 h^\alpha |f(\xi_i^n) - f(\xi_{i-1}^n)| \cdot \text{signe}((\underline{u}_{i+3/2}^n - \underline{u}_{i+1/2}^n)(\underline{u}_{i+1/2}^n - \underline{u}_{i-1/2}^n)) \}$$

et  $c_0 h^\alpha < 1$ ,  $\alpha$  une constante positive *arbitraire*.

$\star$ . On calcule  $\underline{u}_{i+1/2}^{n+1}$  exactement comme dans LRGP.

Une description détaillée de ce schéma peut être trouvée dans [1]. Ici, on va faire quelques remarques.

i) Si  $(\underline{u}_{i+3/2}^n - \underline{u}_{i+1/2}^n)(\underline{u}_{i+1/2}^n - \underline{u}_{i-1/2}^n) \leq 0$ , alors  $\psi_{i+1/2} = 0$  et par conséquent  $a_i = a_{i+1} = 0$ . Ceci veut dire que, pour les moyennes, notre schéma devient dans ce cas le schéma de Godunov, G (cette situation est produite quand la positivité du schéma est en danger).

ii) Quand ceci n'est pas le cas, on a

$$a_j^n = f(\xi_j^n) - f(\xi_j^n) \quad j=i, i+1$$

alors, le schéma pour les moyennes coïncide avec le schéma LRG (ceci est souvent le cas dans les régions où la solution est "régulière").

iii) Il est bien connu que le schéma G produit beaucoup de diffusion au niveau des discontinuités et que LRG diffuse, par contre, très peu

à ces niveaux. Le terme  $-\frac{1}{h}(a_{i+1}-a_i)$  peut être donc interprété comme un terme d'antidiffusion.  $G-1/2$  produit de la diffusion au niveau des chocs et l'enlève là où la solution est régulière.

- iv) Le paramètre  $\alpha$  contrôle la taille des  $a_i$  et c'est pour cela qu'on l'a appelé *le paramètre d'antidiffusion*. (Aux paragraphes 4.3, 4.4 et 4.5, on prendra  $\alpha = 0$ , au paragraphe 4.6, on étudie l'effet de  $\alpha$  sur les propriétés de convergence du schéma).

On a donc présenté les 6 schémas dont les propriétés de convergence seront étudiées numériquement. Avant de passer à la description des cas-tests que l'on va considérer, on montre un tableau dans lequel on résume les principales caractéristiques, connues, de ces schémas.

	G	MUSCL	B1P	LRGP	G-1/2	LRG
linéaire	oui	non	non	non	non	oui
conservatif	oui	oui	oui	oui	oui	oui
positif	oui	?	?	?	oui	non
monotone	oui	?	?	?	oui	non
diffuse au niveau des chocs	beaucoup	?	?	?	peu	très peu
oscillations au niveau des chocs	non	?	?	?	non	oui
condition de stabilité qui garantit la convergence	$CFL \leq 1$	?	?	?	$CFL \leq 1/2$	$CFL \leq ch^{1/2}$ (cas linéaire seulement)

Schémas "LRG-modifiés"

Tableau 1 : propriétés connues des schémas

#### 4.2.2 - Description des cas-tests

On va tester ces schémas sur trois problèmes de Riemann, c'est-à-dire sur trois problèmes de la forme :

$$(P) \quad \left\{ \begin{array}{l} \partial_t u + \partial_x f(u) = 0 \quad \text{dans } \mathbb{R}^+ \times \mathbb{R} \\ u(t=0, x) = \begin{cases} u_g & \forall x < x_0 \\ u_d & \forall x > x_0 \end{cases} \quad x_0 \in \mathbb{R} \end{array} \right.$$

ce choix est convenable pour deux raisons principalement :

i) tout d'abord, parce qu'on connaît explicitement la solution exacte, ce qui nous permet de calculer l'erreur.

En effet, si l'on pose

$$\hat{f}(u) = \begin{cases} \text{l'enveloppe convexe de } f \text{ sur } [u_g, u_d] \\ \quad \quad \quad (\text{si } u_g \leq u_d) \\ \text{l'enveloppe concave de } f \text{ sur } [u_d, u_g] \\ \quad \quad \quad (\text{si } u_d \leq u_g) \end{cases}$$

et on définit  $\xi_g, \xi_d$  de la façon suivante :

$$\xi_g = \hat{f}'(u_g)$$

$$\xi_d = \hat{f}'(u_d)$$

alors la solution entropique de (P) , u, est donnée par :

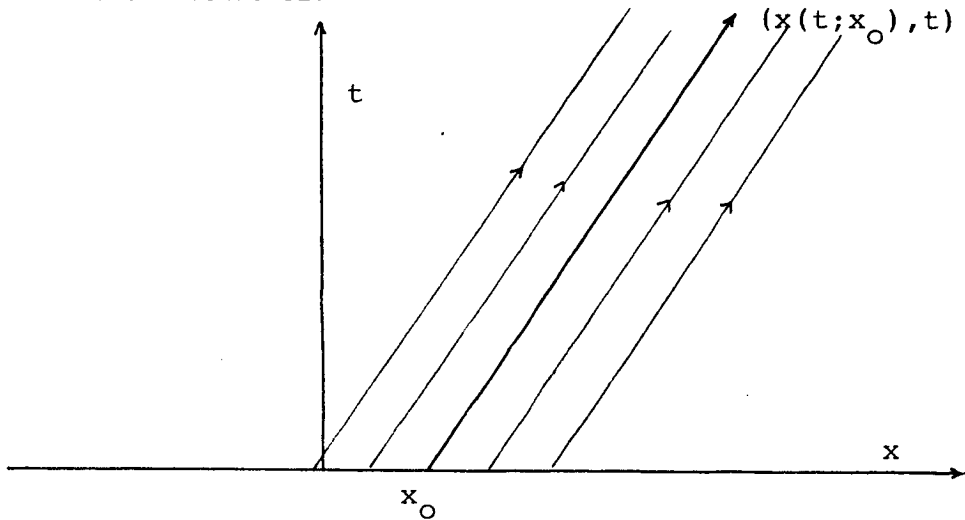
$$u(x, t) = \begin{cases} u_g & \forall \frac{x-x_0}{t} < \xi_g \\ u_d & \forall \frac{x-x_0}{t} < \xi_d \\ (\hat{f}')^{-1} \left( \frac{x-x_0}{t} \right) & \text{autrement} \end{cases}$$

Il faut prendre ici  $(\hat{f}')^{-1}$  dans le sens généralisé. On a toujours soit  $\hat{f}'' \geq 0$  soit  $\hat{f}'' \leq 0$ . Si on a l'inégalité stricte, alors  $(\hat{f}')^{-1}$  est univoquement définie, mais l'égalité peut avoir lieu et alors  $(\hat{f}')^{-1}$  est multivoque. C'est justement dans cette situation que les discontinuités de  $u$  apparaissent.

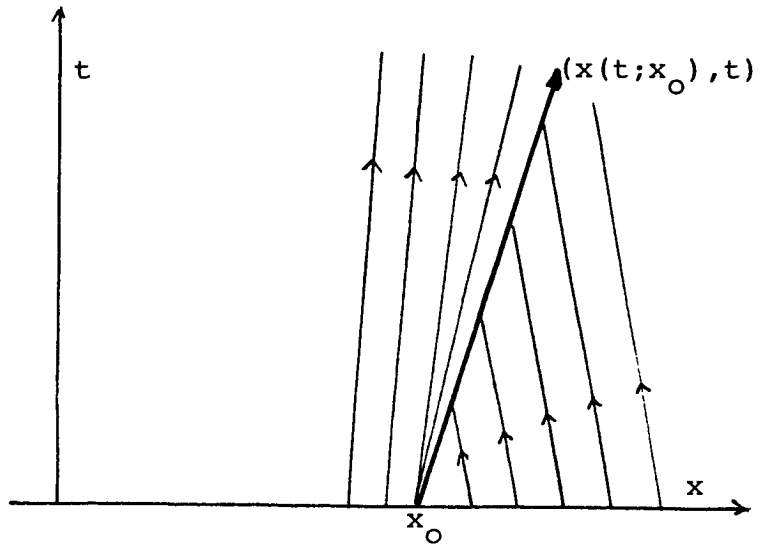
- ii) Il est raisonnable de penser que qualitativement le comportement du schéma, c'est-à-dire ses propriétés d'approximation et de convergence, sera le même si on change la condition initiale de (P) par une autre "un peu plus générale" ( $\mathcal{C}^1$  par morceaux par exemple).

Maintenant il faut choisir les problèmes (P). Le critère qui va nous guider en ce choix est le suivant : la solution entropique de (P) est  $\mathcal{C}^1$  par morceaux, et ses discontinuités sont de deux classes :

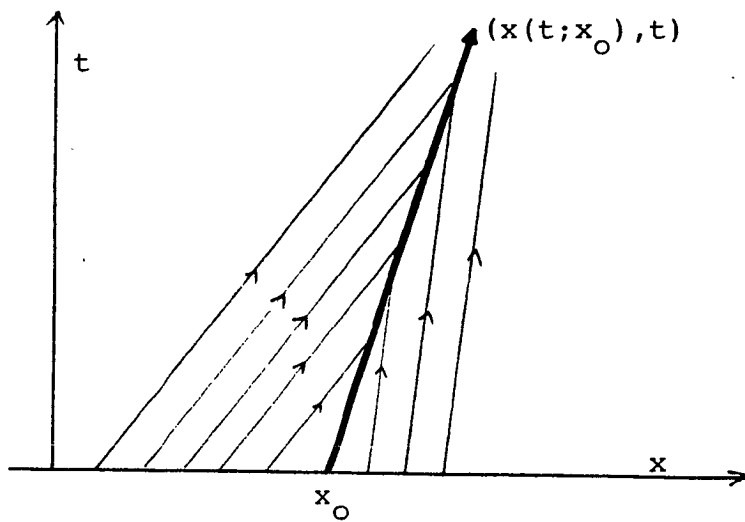
- i) une discontinuité de contact : si le long de la courbe  $(x(t), t)$  sur laquelle est placée la discontinuité les caractéristiques demeurent parallèles à celle-ci.



- ii) un choc : s'il existe au moins une caractéristique "rentrant" dans la courbe de discontinuité  $(x(t), t)$



(a)



(b)

D'autre part, il faut remarquer qu'il serait très intéressant d'inclure *le cas linéaire* dans ces expériences.

On va ainsi choisir les trois cas-tests suivants :

Le cas (a) : le cas linéaire, une discontinuité de contact

$$\left\{ \begin{array}{l} f(u) = u \\ u_g = 1 \\ u_d = 0 \\ x_0 = 0 \end{array} \right.$$

Le cas (b) : le cas Buckley-Leverett : un choc

$$\left\{ \begin{array}{l} f(u) = \frac{1}{2} \frac{u^2}{u^2 + (1-u)^2} \\ u_g = 1 \\ u_d = 0 \\ x_0 = 0 \end{array} \right.$$

Le cas (c) : la solution est de classe  $\mathcal{C}^0$ ,  $\forall t > 0$

$$\left\{ \begin{array}{l} f(u) = u(1-u) \\ u_g = 1 \\ u_d = 0 \\ x_0 = 1/2 \end{array} \right.$$

Le cas (a), où la solution est le transport de la condition initiale à vitesse constante égale à 1, nous servira pour tester nos schémas sur le cas linéaire et sur un des cas les plus difficiles : le transport d'une discontinuité de contact.

Le cas (b), où les solutions possèdent un choc du type ii)(a), nous servira pour tester les schémas sur un cas non-linéaire où la fonction  $f$  possède un point d'inflexion. Le choix se situe à peu près vers  $u = 0.70$  et se déplace avec une vitesse constante égale à 0.604.

Finalement, le cas (c) où  $f$  est concave, nous servira pour tester le cas où la solution est de classe  $\mathcal{C}^0$  pour tout  $t$  strictement positif.

Ainsi, on a envisagé d'isoler dans chaque cas les différents types de singularités et de régularité de la solution de notre problème.

Avant de passer à la description du protocole de l'étude numérique, on remarque que le domaine spatial sur lequel la solution  $u_h$  sera calculée sera  $]0,1[$ . Des conditions aux bords adéquates seront imposées sur  $x = 0$



et  $x = 1$ . On a pris le soin de vérifier que ces conditions produisent pour les temps considérés, sur  $]0,1[$  les mêmes solutions  $u_h$  obtenues en considérant un intervalle plus grand  $] -1,2[$ .

#### 4.2.3 - Le protocole d'étude numérique

Pour chacun des trois cas, on va procéder de la même façon. En fait, étant donné que la condition de stabilité de LRG est

$$CFL \leq c h^{1/2}$$

et que, par contre, pour  $G$  et  $G-1/2$ , on a

$$CFL \leq 1,1/2 \text{ respectivement}$$

(et une condition de ce type doit garantir la stabilité de MUSCL, B1P et LRGP!) on va séparer l'étude de LRG de l'étude des autres schémas.

Ainsi,

(I) i) Pour  $G$ , MUSCL, B1P, LRGP et  $G-1/2$ , on va faire des expériences pour les valeurs de CFL suivantes :

$$CFL = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}.$$

ii) Pour LRG, on prendra simplement

$$h^{-1/2} CFL = \frac{\sqrt{10}}{4}$$

(on a pris cette valeur car on va prendre une partition initiale dont les éléments ont un diamètre égal à  $h = 1/10$  et parce qu'avec  $CFL = 1/4$  le schéma LRG n'explose pas, comme pour le cas  $CFL = 1/2$ ).

maintenant, à une valeur fixée de la condition de stabilité,

- (II) i) Pour  $G$  on prend  $h_0 = 1/20$  et on raffine *trois* fois en divisant  $h$  par 2 à chaque fois.
- ii) Pour MUSCL, B1P, LRGP et  $G-1/2$ , on prend  $h_0 = 1/10$  (ils ont le double de degrés de liberté que  $G$ !) et on fait la même chose qu'en i).
- iii) Pour LRG, on prend  $CFL = 1/4$  et  $h_0 = 1/10$  comme on avait dit, et on raffine *deux* fois en divisant  $h$  par 4 à chaque fois.

Il faut remarquer que lorsqu'on raffine en utilisant  $G$ , MUSCL, B1P, LRGP ou  $G-1/2$ , on divise  $\Delta t$  par deux à chaque fois que  $h$  est divisé par 2. Mais lorsqu'il s'agit de LRG, il faut diviser  $\Delta t$  par *huit* à chaque fois que  $h$  est divisé par 4 !

Finalement,

- (III) i) On fixe un temps  $t_0$  et on calcule pour chacun de ces cas la solution  $u_h(t_0)$ ; on calcule alors l'erreur  $e(h)$  :

$$e(h) = \|u_h(t_0) - u(t_0)\|_{L^1(0,1)}$$

(la norme  $L^1$  est la "bonne" norme pour étudier la convergence, [ 3 ]).

- ii) On dessine alors *la courbe (discrète) des erreurs*

$$\{(h, e(h)) , h=h_0, h_0/2, h_0/4, h_0/8 \}$$

pour  $G$ , MUSCL, B1P, LRGP et  $G-1/2$

$$\{(h, e(h)) , h=h_0, h_0/4, h_0/16 \}$$

pour LRG

- iii) On dessine finalement *la courbe des ordres de convergence*

$$\{(h, \alpha(h) , h=h_0/\sqrt{2}, h_0/2\sqrt{2}, h_0/4\sqrt{2} \}$$

pour  $G$ , MUSCL, B1P, LRGP et  $G-1/2$

où,

$$\alpha(h/\sqrt{p}) = \log\left(\frac{e(h)}{e\left(\frac{h}{p}\right)}\right) / \log(p)$$

$p = 4$  pour LRG,  $= 2$  autrement.

Il faut rendre évident le fait que la fonction  $\alpha$  est une mesure de l'ordre de convergence. Supposons que la méthode est d'ordre  $\alpha$ ,

$$e(h) \approx c h^\alpha$$

alors,

$$\frac{e(h)}{e\left(\frac{h}{p}\right)} \approx p^\alpha$$

$$\alpha \approx \log\left(\frac{e(h)}{e\left(\frac{h}{p}\right)}\right) / \log p$$

Dans (I), (II) et (III) notre protocole d'étude numérique de la convergence est décrit. Il ne nous reste qu'à faire quelques remarques. Pour chaque expérience numérique, on va dessiner la solution  $u_h$  en au moins 4 instants équidistants (la condition initiale mise à part). Les courbes d'erreur des schémas G, MUSCL, BIP, LRGP et G-1/2 seront dessinées ensemble pour faciliter la comparaison. On fera de même pour les courbes d'ordre de convergence.

#### 4.3 - Le cas linéaire : une discontinuité de contact

##### 4.3.1 - Le schéma LRG

Les résultats des expériences numériques se trouvent sur les figures 1 et 2. Sur la figure 1, comme on l'avait dit au paragraphe 4.2.3, on a pris :

$$\text{CFL} = 1/4 \quad , \quad h = 1/10$$

$$\text{CFL} = 1/8 \quad , \quad h = 1/40$$

$$\text{CFL} = 1/16 \quad , \quad h = 1/160$$

de façon telle à maintenir constante la condition de stabilité, pour le cas linéaire,

$$h^{-1/2} \text{ CFL} = \text{constante} \left( = \frac{\sqrt{10}}{4} \right) .$$

Sur cette figure, on voit bien que LRG n'est ni positif ni monotone. De très fortes oscillations se présentent près de la discontinuité, ceci étant dû à l'absence de diffusion du schéma. On voit tout de même que la solution converge, comme prédit par la théorie.

Ceci peut être mieux évalué sur les figures 2, où l'on voit, non seulement que l'erreur est monotone décroissante lorsque  $h \downarrow 0$ , mais que l'ordre de convergence est  $\approx 1$ . (Si la condition initiale est très régulière, l'ordre de convergence est 2, la discontinuité produit donc la perte de 1 dans l'ordre de convergence!).

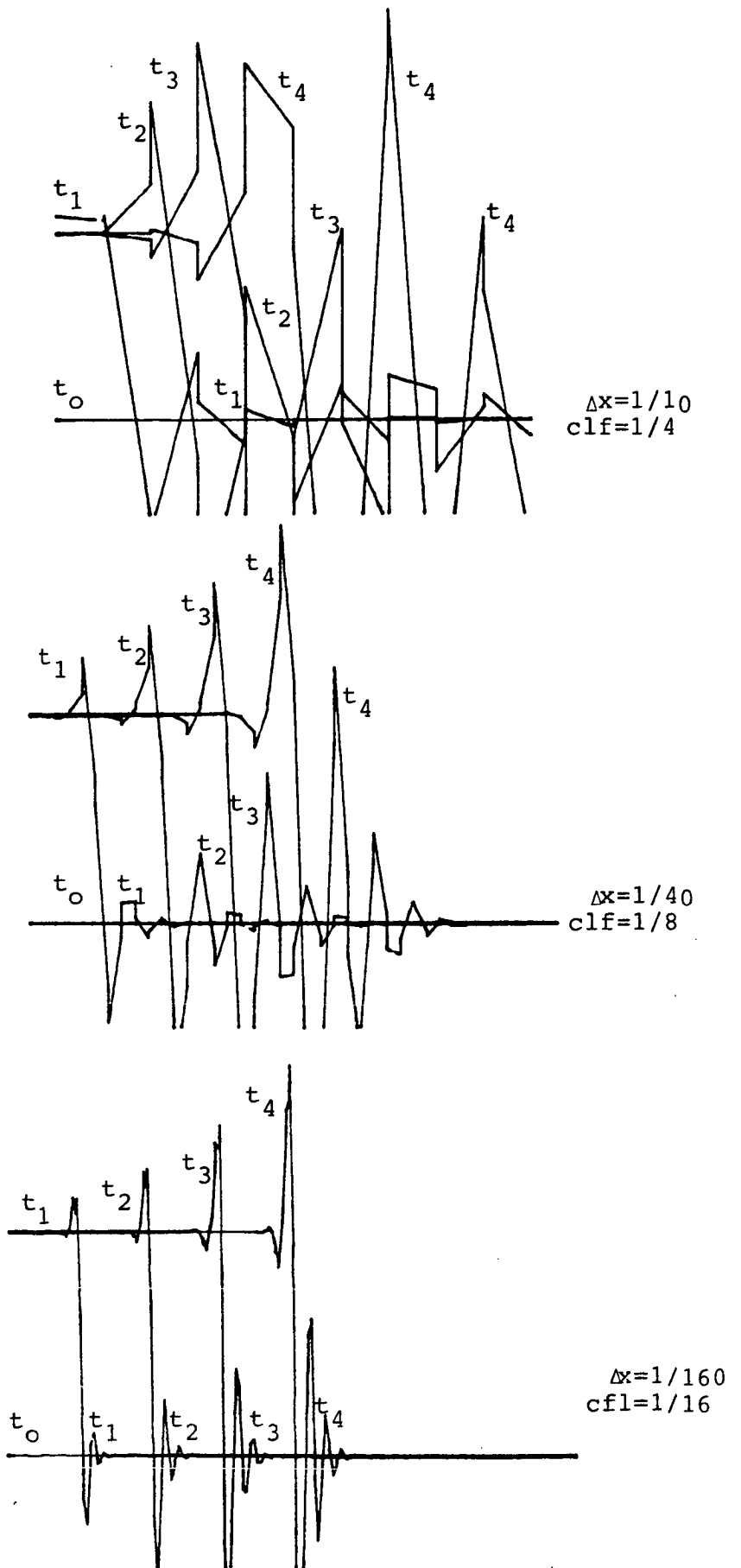


Figure 1 : Cas d'une discontinuité de contact : schéma LRG. Raffinement du maillage sous la condition de stabilité :  $\Delta t = \text{constante } (\Delta x)^{3/2}$ .

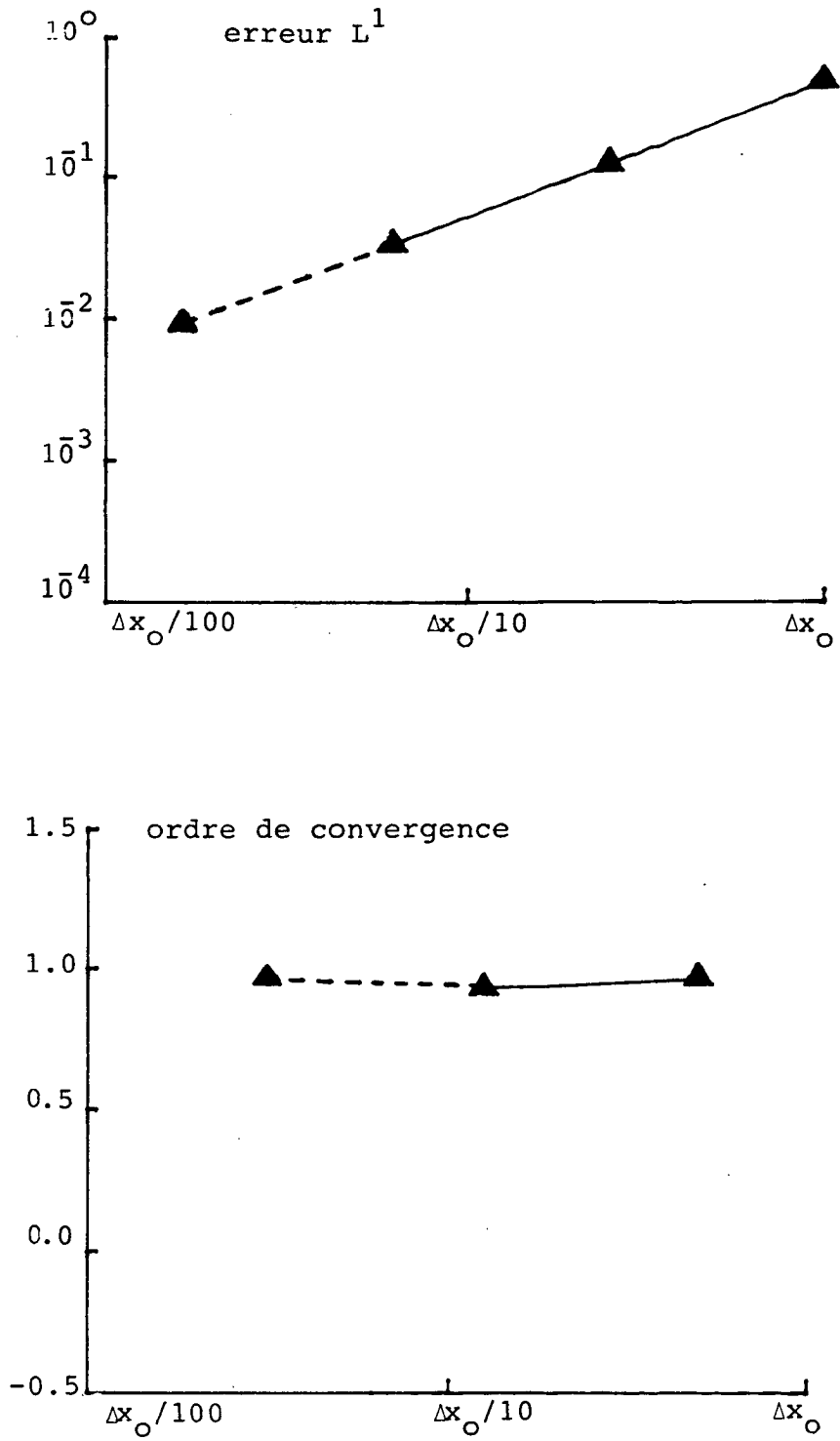


Figure 2 : Cas d'une discontinuité de contact : schéma LRG. Résultats numériques de convergence sous la condition de stabilité  $\Delta t = \text{const.} (\Delta x)^{3/2}$ . On a pris  $\Delta x_0 = 1/10$  (voir la figure précédente). En haut, l'évolution de l'erreur L<sup>1</sup> en fonction du  $\Delta x$ ; en bas l'ordre de convergence du schéma est estimé à partir des pentes de la courbe des erreurs (voir ( . )). Le trait discontinu représente une extrapolation probable. (La convergence du schéma, dans ce cas, est d'ordre  $\approx 1$ ).

#### 4.3.2 - Les autres schémas

Comme on l'avait dit au paragraphe 4.2.3, on va considérer trois cas :

$$CFL = 1/2 , 1/4 , 1/8$$

et pour chacun de ces cas, on va prendre

$$h = 1/10 , 1/20 , 1/40 , 1/80$$

(pour G on prend  $h = 1/20 , 1/40 , 1/80 , 1/160$ , car les autres schémas doublent le nombre de degrés de liberté).

##### i) Le cas CFL = 1/2

Sur les figures 3 à 7, on montre les solutions  $u_h$  à différents instants de temps, obtenues par les schémas G, MUSCL, BIP, LRGP et G-1/2 dans cet ordre.

On y observe que :

- i) tous les schémas sont positifs dans ce cas
- ii) ils sont aussi monotones, sauf G-1/2 qui est monotone seulement pour les moyennes,
- iii) on voit bien l'effet de la diffusion "contenue" dans le schéma G :  
le front raide de la discontinuité est établi, et en plus, *cet étalement augmente avec le temps,*
- iv) les schémas MUSCL et BIP ont aussi cette propriété, bien que chez eux la diffusion ne soit pas si grande que celle de G. BIP présente légèrement moins de diffusion que MUSCL,
- v) dans le cas LRGP donné, *une fois sur deux,* la solution est exacte,
- vi) le comportement de G-1/2 dans ce cas, est très intéressant. Ce schéma diffuse aussi, mais à la différence des schémas G, MUSCL et BIP, *cette diffusion diminue avec le temps.*

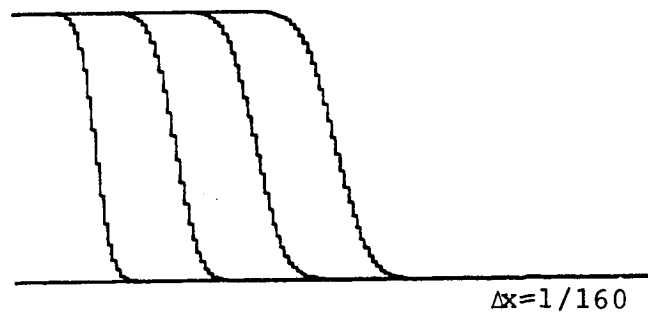
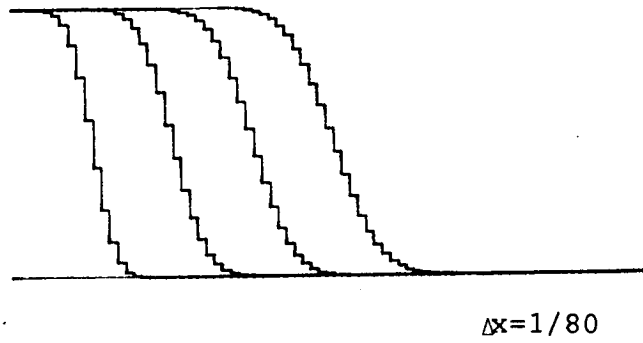
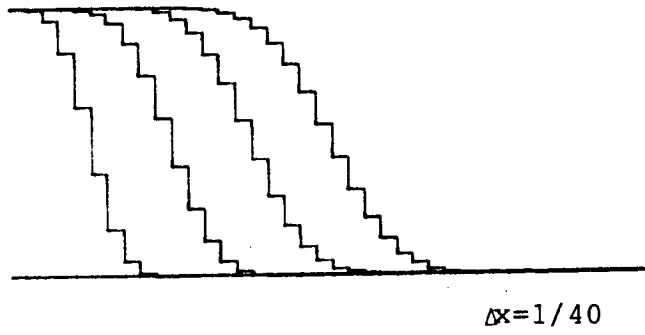
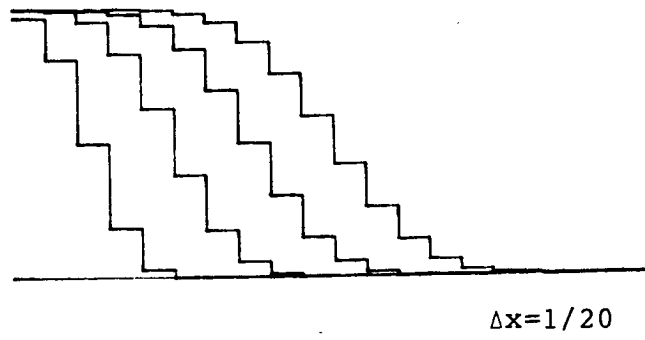


Figure 3 : Cas d'une discontinuité de contact : schéma de Godunov. Raffinement du maillage sous la condition  $CFL = 1/2$



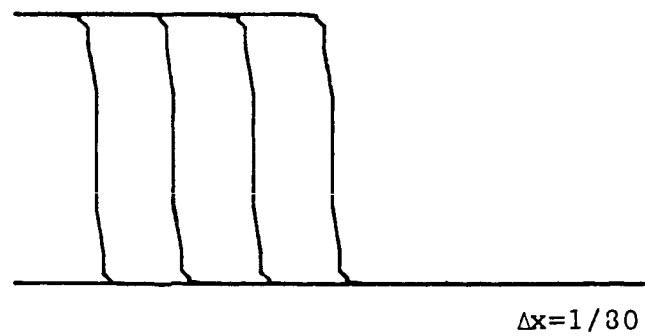
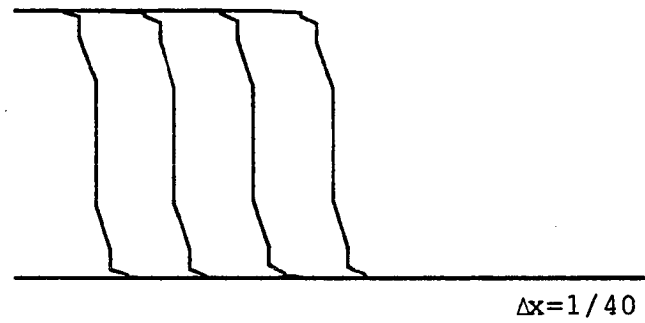
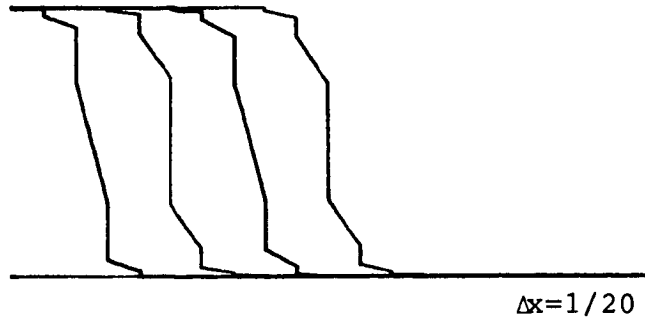
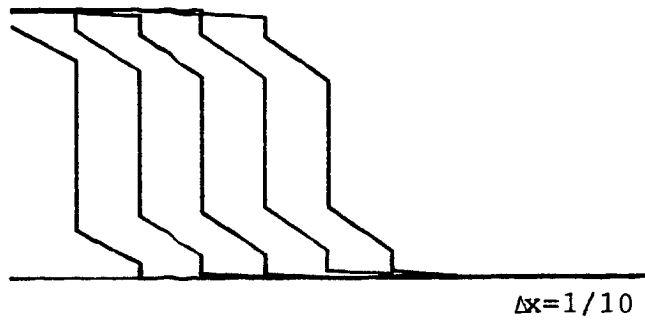


Figure 4 : Cas d'une discontinuité de contact : schéma MUSCL. Raffinement du maillage sous la condition  $CFL = 1/2$ .

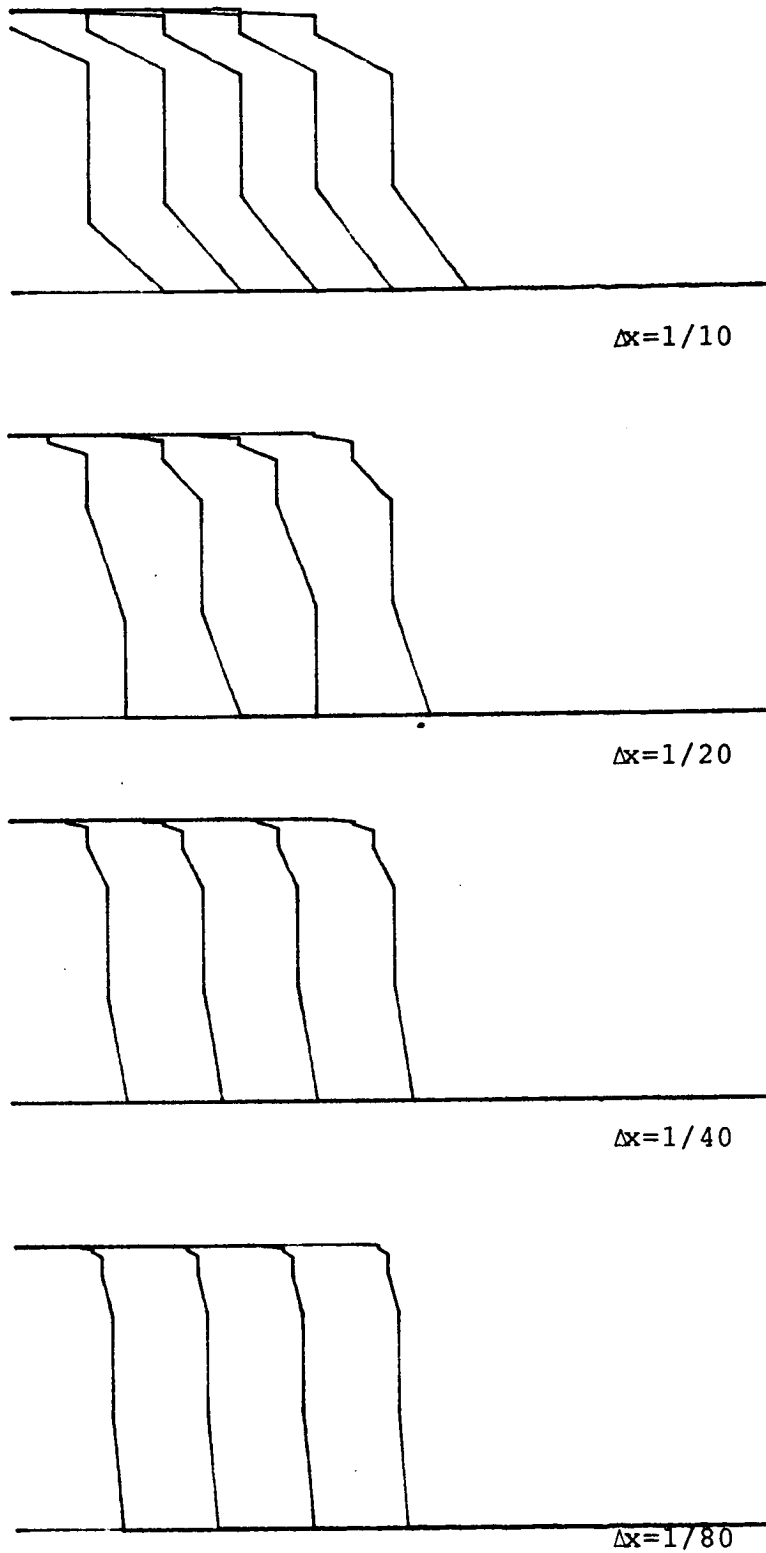


Figure 5 : Cas d'une discontinuité de contact : schéma BIP. Raffinement du maillage sous la condition  $CFL = 1/2$ .

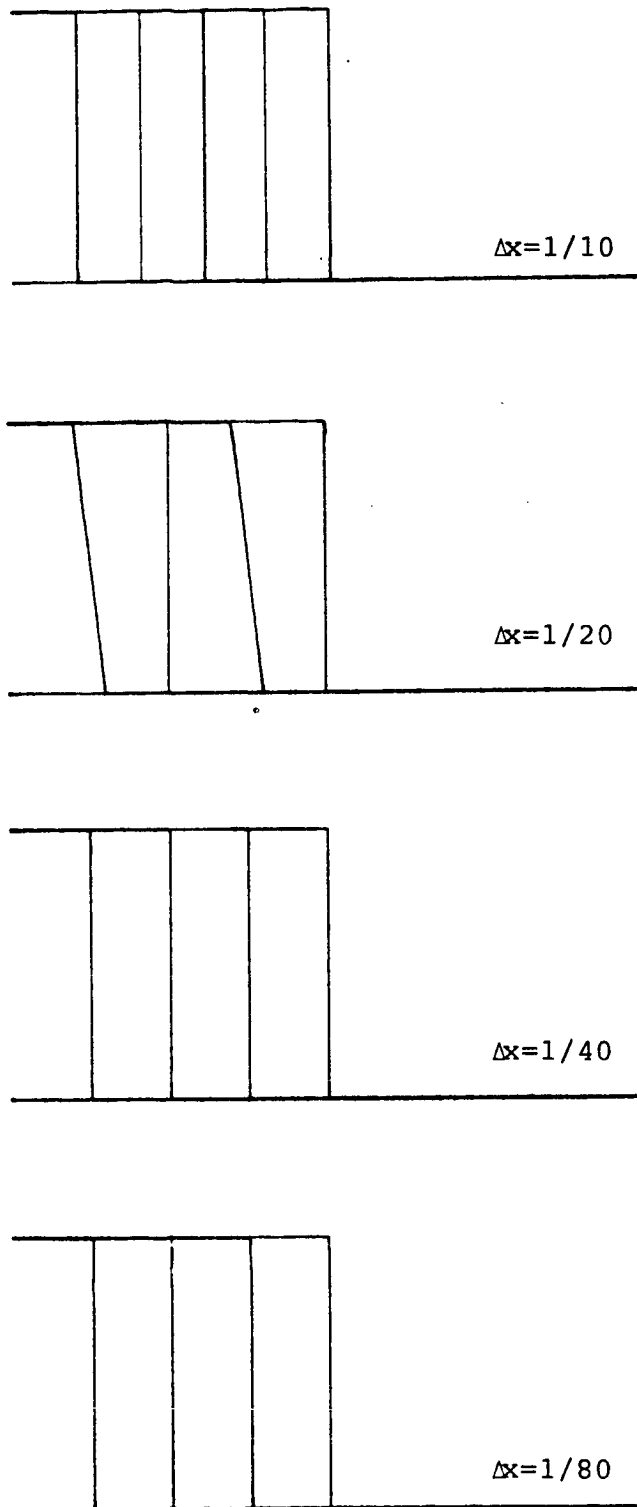


Figure 6 : Cas d'une discontinuité de contact : schéma LRGP. Raffinement du maillage sous la condition  $CFL = 1/2$ .

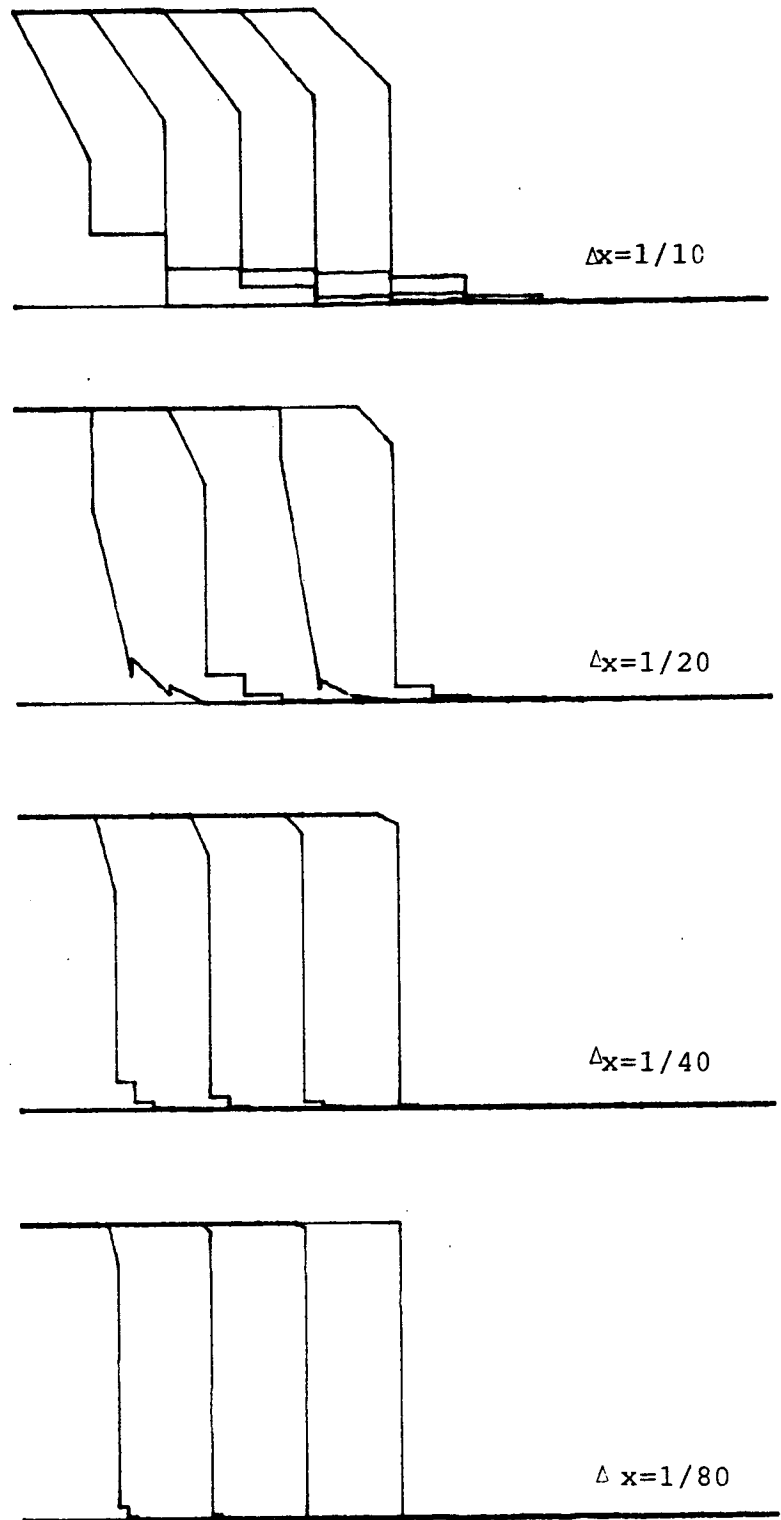


Figure 7 : Cas d'une discontinuité de contact : schéma G-1/2. Raffinement du maillage sous la condition CFL = 1/2.

Maintenant, considérons les propriétés de convergence de ces schémas. Sur la figure 8, on a mesuré les courbes d'erreur  $L^1$  et d'ordre de convergence de ces schémas. (Evidemment, les courbes de LRGP n'apparaissent pas dans cette figure, aussi la courbe d'ordre de convergence de G-1/2 n'a pas pu être dessinée car on ne voulait pas changer d'échelle!)

On y constate que (LRGP mis à part),

- i) G-1/2 produit des erreurs plus petites que le reste de schémas. Ensuite, BIP produit des résultats légèrement meilleurs que MUSCL. Et celui-ci produit des erreurs bien plus petites que G;
- ii) l'ordre de convergence pour BIP et MUSCL semble être très près de 1, il est très légèrement supérieur pour BIP;
- iii) pour G l'ordre de convergence semble être (exactement égal à) 1/2;
- iv) pour G-1/2, il se passe quelque chose de "vraiment non linéaire". En effet, l'ordre de convergence augmente plus vite que linéairement!, lorsque  $h \rightarrow 0$ , si  $h_0 = 1/10$ , on a

$$(h_0/\sqrt{2}) = 2.056$$

$$(h_0/2\sqrt{2}) = 2.963$$

$$(h_0/4\sqrt{2}) = 4.855 .$$

En conclusion, dans ce cas, on peut ordonner les schémas "du meilleur au moins bon" de la façon suivante

$$\text{LRGP} \geq \text{G-1/2} \geq \text{BIP} \geq \text{MUSCL} \geq \text{G}$$

où "meilleur" signifie qu'il produit des erreurs plus petites et possède un ordre de convergence plus grand.

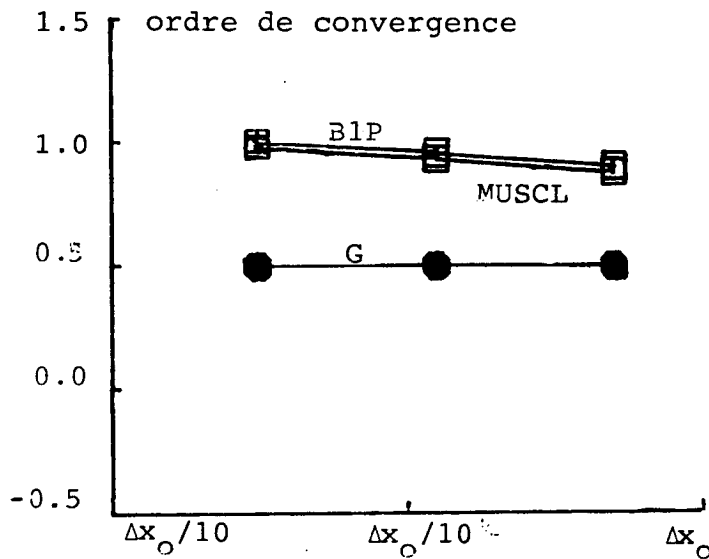
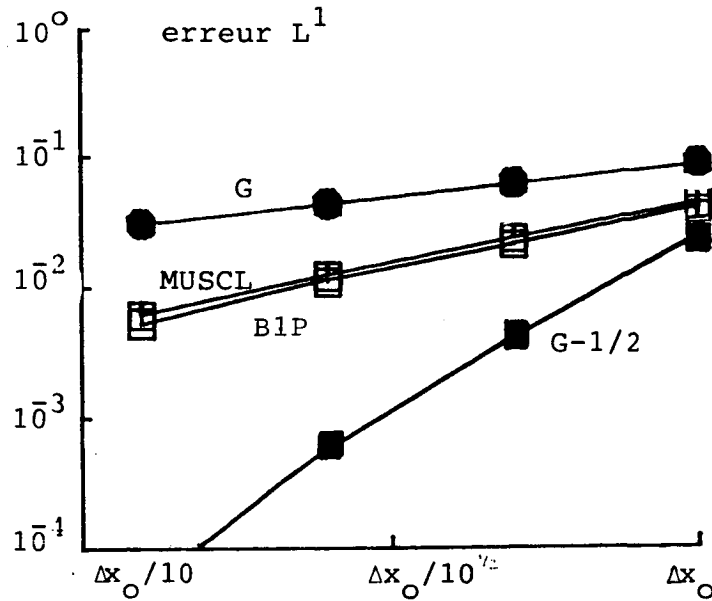


Figure 8 : Cas d'une discontinuité de contact : comparaison numérique de convergence  $L^1$  des différents schémas sous la condition CFL = 1/2.

Les schémas LRGP et BIP donnent la *solution exacte*. Le schéma G-1/2 possède dans ce cas une propriété de superconvergence très intéressante: les valeurs de  $\alpha$ , pour ce schéma, qui n'apparaissent pas sur la figure, sont, respectivement (du plus grand  $\Delta x$  au plus petit) : 2.056, 2.963, 4.855!  
 (On a pris  $\Delta x_0 = 1/10$ , sauf pour le schéma de Godunov, pour lequel on a :  $\Delta x_0 = 1/20$ ).

ii) Le cas CFL = 1/4

Sur les figures 9 à 13, on montre les solutions  $u_h$  obtenues par les schémas G, MUSCL, BIP, LRGP et G-1/2, dans cet ordre. Les mêmes propriétés de monotonie, positivité et de "diffusion" trouvées dans le cas CFL = 1/2 apparaissent aussi dans ce cas (il faut signaler que LRGP ne donne plus la solution exacte...).

Sur la figure 14, on montre les propriétés de convergence des schémas considérés. Par rapport au cas précédent, on constate les différences suivantes :

- i) l'ordre de convergence de G-1/2 n'a plus le comportement non-linéaire observé dans le cas précédent. Dans ce cas, l'ordre semble être près de 1;
- ii) LRGP produit des erreurs très légèrement inférieures à celles produites par G-1/2. D'autre part, l'ordre de convergence de ces deux schémas semble être pratiquement le même;
- iii) au fur et à mesure que l'on raffine le maillage, les schémas LRGP et G-1/2 semblent devenir "les mêmes".

Ceci est un point très intéressant à discuter, car cela veut dire que (voir la définition de G-1/2) :

$$\frac{1}{N_x} \text{Card}\{i = a_i \neq (f(\xi_i) - f(\xi_i))\}$$

où  $N_x$  est le nombre d'éléments, et est un nombre qui va vers zéro lorsque  $h \rightarrow 0$ . Ceci est très raisonnable car, en fait, les  $a_i$  sont différents de  $(f(\xi_i) - f(\xi_i))$  seulement dans un petit voisinage de la discontinuité.

Et, en plus, d'après cette expérience numérique, il est clair que ce voisinage dépend de  $h$ , plus précisément on conjecture que sa mesure doit être  $O(h)$ . (Sur la figure 13, où on montre que la solution donnée par le schéma G-1/2, on constate que pour tout temps et pour toutes les valeurs de  $h$ , le profil

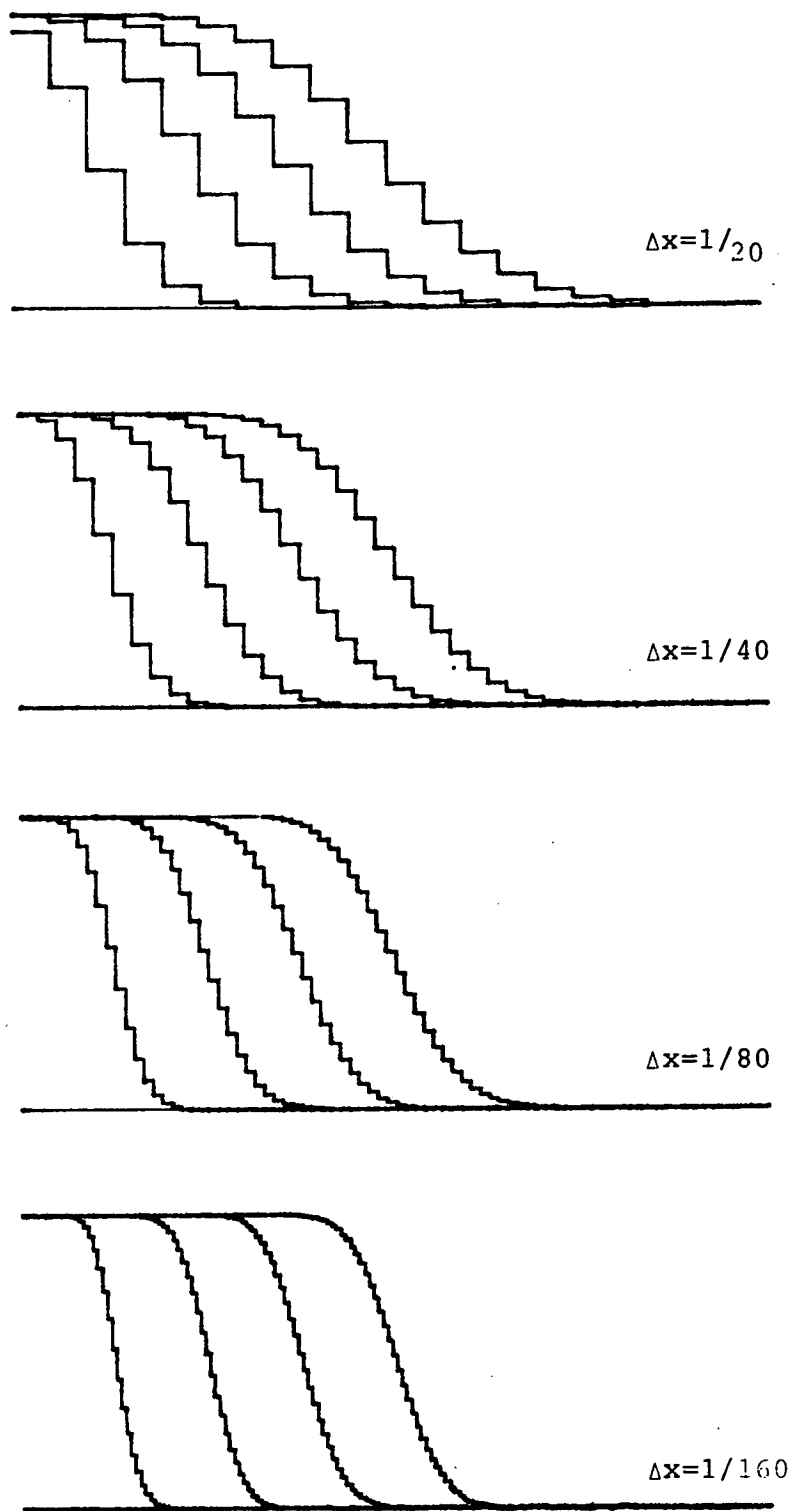


Figure 9 : Cas d'une discontinuité de contact : schéma de Godunov. Raffinement du maillage sous la condition CFL = 1/4.



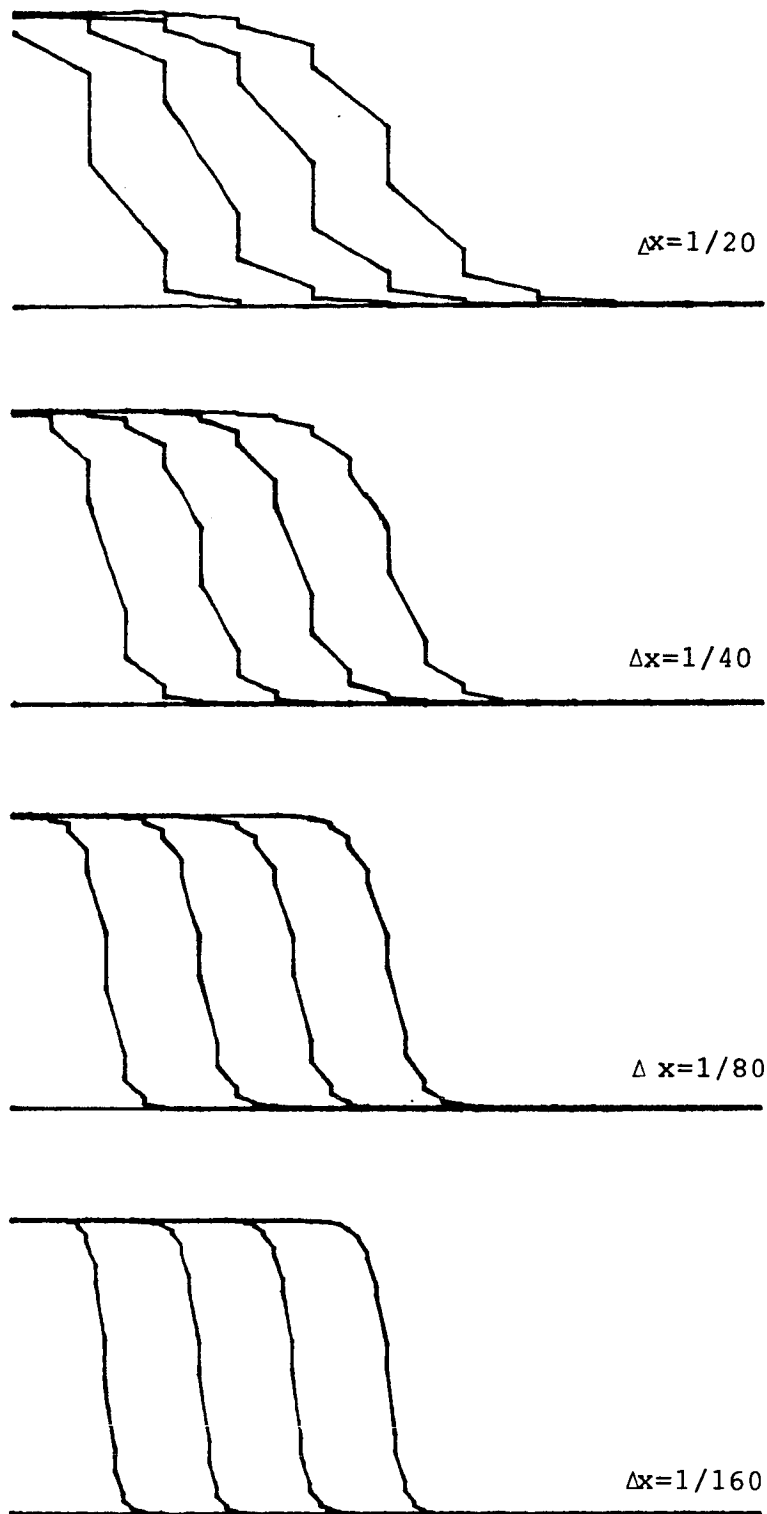


Figure 10 : Cas d'une discontinuité de contact : schéma MUSCL. Raffinement du maillage sous la condition CFL = 1/4.

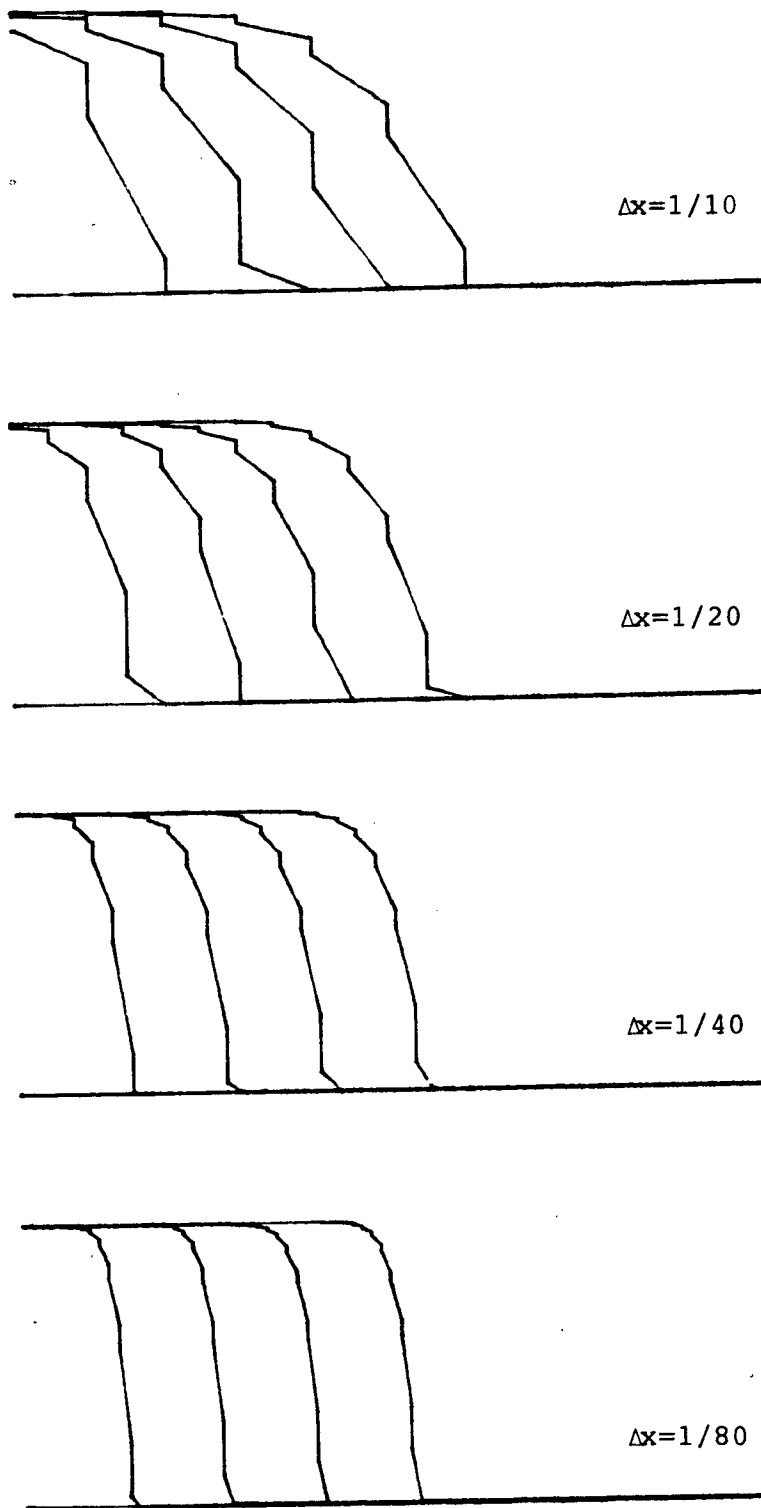


Figure 11 : Cas d'une discontinuité de contact : schéma BIP. Raffinement du maillage sous la condition CFL = 1/4.

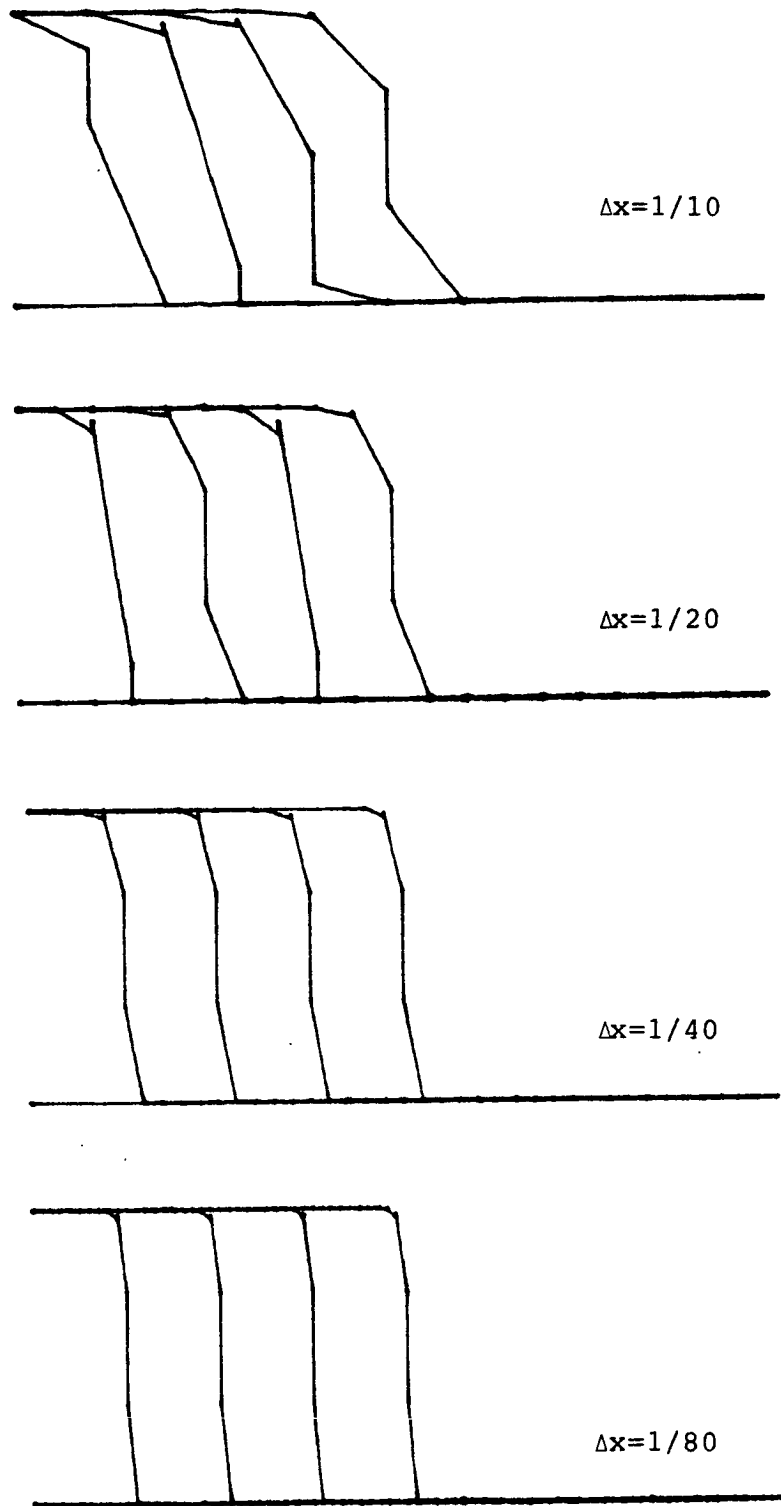


Figure 12 : Cas d'une discontinuité de contact : schéma LRGP. Raffinement du maillage sous la condition  $CFL = 1/4$ .

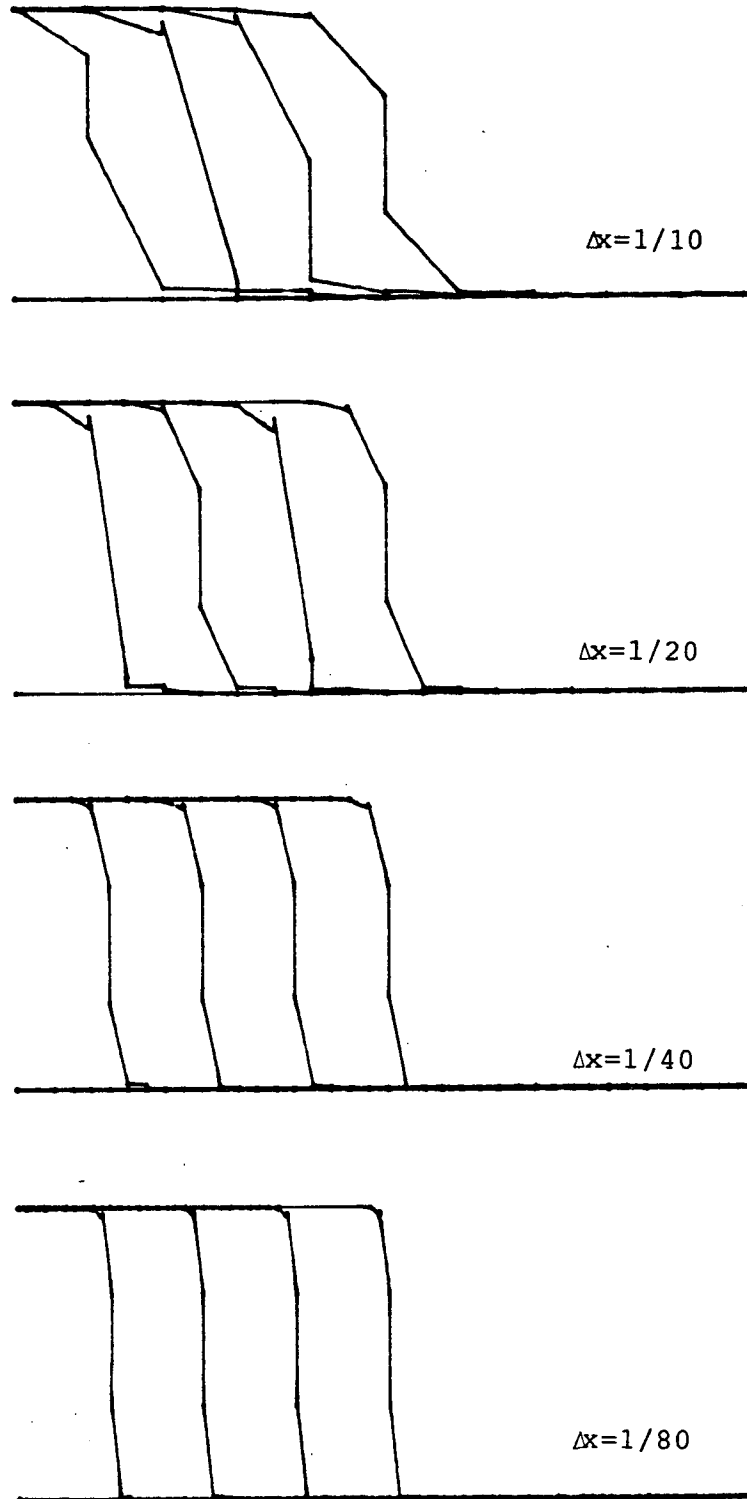


Figure 13 : Cas d'une discontinuité de contact : schéma G-1/2. Raffinement du maillage sous la condition CFL = 1/4.

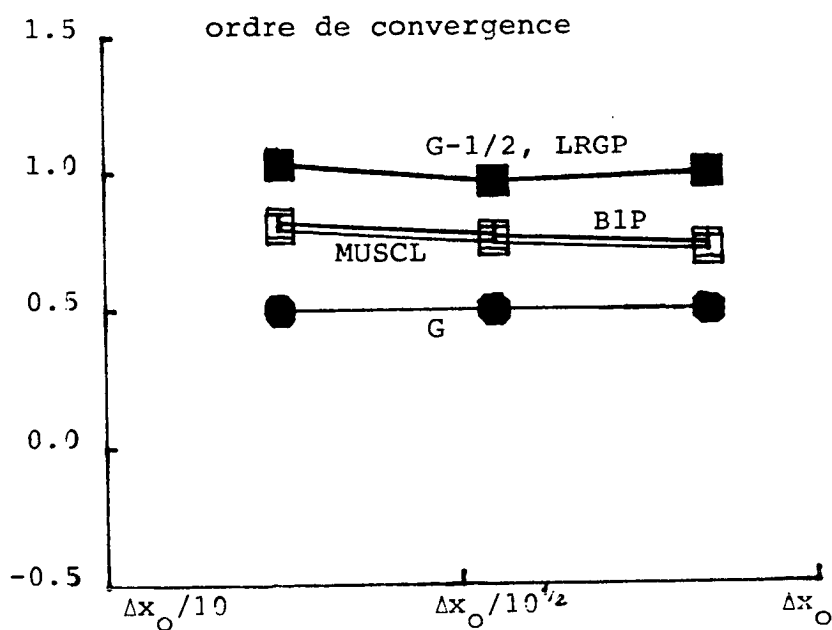
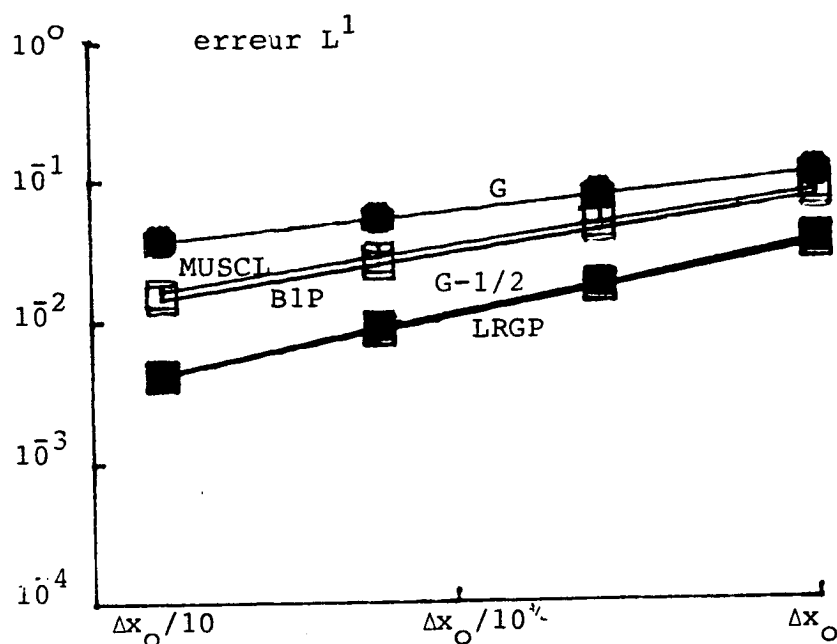


Figure 14 : Cas d'une discontinuité de contact : comparaison numérique des différents schémas sous la condition  $CFL = 1/4$ . Les schémas LRGP et B1P coïncident pratiquement dans ce cas. Le schéma G-1/2, pour  $\Delta x$  "grand" donne des erreurs légèrement plus grandes que celles produites par les schémas LRGP et B1P (On a pris  $\Delta x_0 = 1/10$ , sauf pour le schéma de Godunov, pour lequel on a pris  $\Delta x_0 = 1/20$ ).

de la discontinuité est contenu en au plus 4 éléments! ).

Ceci explique iii), et, en partie, ii). En fait, LRGP produit des meilleurs résultats que G-1/2, car il diffuse moins. D'autre part, la diminution du CFL introduit de la diffusion dans les schémas. Ceci explique le point i).

iii) Le cas CFL = 1/8

Les résultats sont montrés sur les figures 15 à 20. Les seules différences par rapport au cas précédent sont :

- i) la diffusion a augmenté (ceci est l'effet de la diminution de CFL);
- ii) les erreurs sont aussi augmentées (ceci est aussi le fruit de la diminution du CFL);
- iii) l'ordre de convergence pour G-1/2, LRGP a très légèrement diminué, il a diminué beaucoup plus pour MUSCL et BIP, et est demeuré invariable pour G.

iv) Quelques commentaires

Si on ne considère pas le comportement spécial de LRGP et celui de G-1/2, dans le cas CFL = 1/2, on constate que, indépendamment de CFL :

- i) tous les schémas semblent être positifs et monotones (pour LRGP et G-1/2 la monotonie est au niveau des moyennes seulement);
- ii) la diffusion *augmente avec le temps* pour G, MUSCL et BIP, elle *n'augmente pas* (et souvent elle *diminue*) pour LRGP et G-1/2;
- iii) pour cet exemple, on constate que si un schéma produit moins de diffusion, il produit des erreurs plus petites et possède un ordre de convergence plus grand. Ainsi, on peut ordonner nos schémas de la façon suivante

$$\text{LRGP} \tilde{\approx} \text{G-1/2} \gg \text{BIP} \tilde{\approx} \text{MUSCL} \gg \text{G}$$

où  $\tilde{\approx}$  (resp.  $\gg$ ) signifie que les schémas ont un comportement très très semblable (resp. assez différencié).

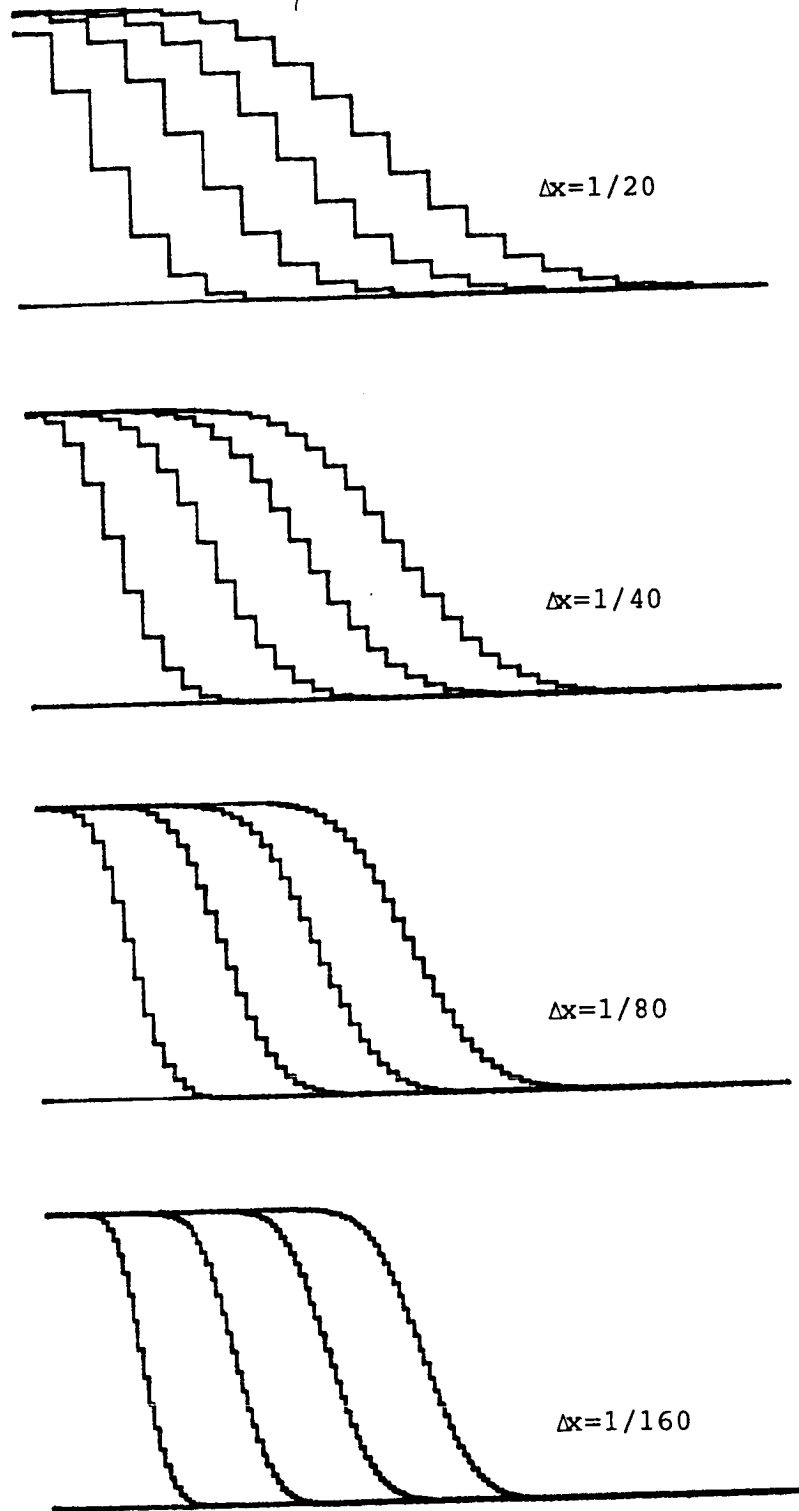


Figure 15 : Cas d'une discontinuité de contact : schéma de Godunov.  
Raffinement du maillage sous la condition CFL = 1/8.

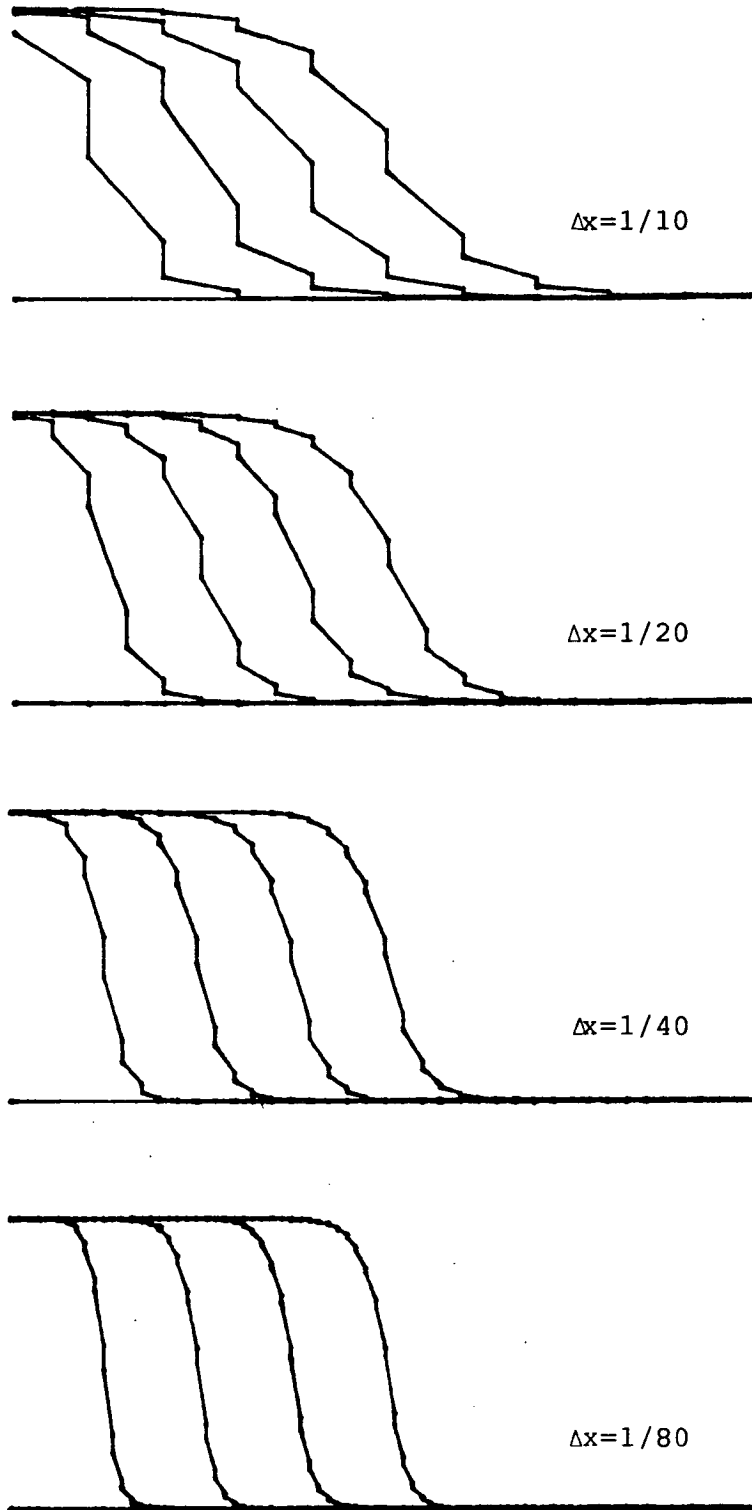


Figure 16 : Cas d'une discontinuité de contact : schéma MUSCL. Raffinement du maillage sous la condition CFL = 1/8.



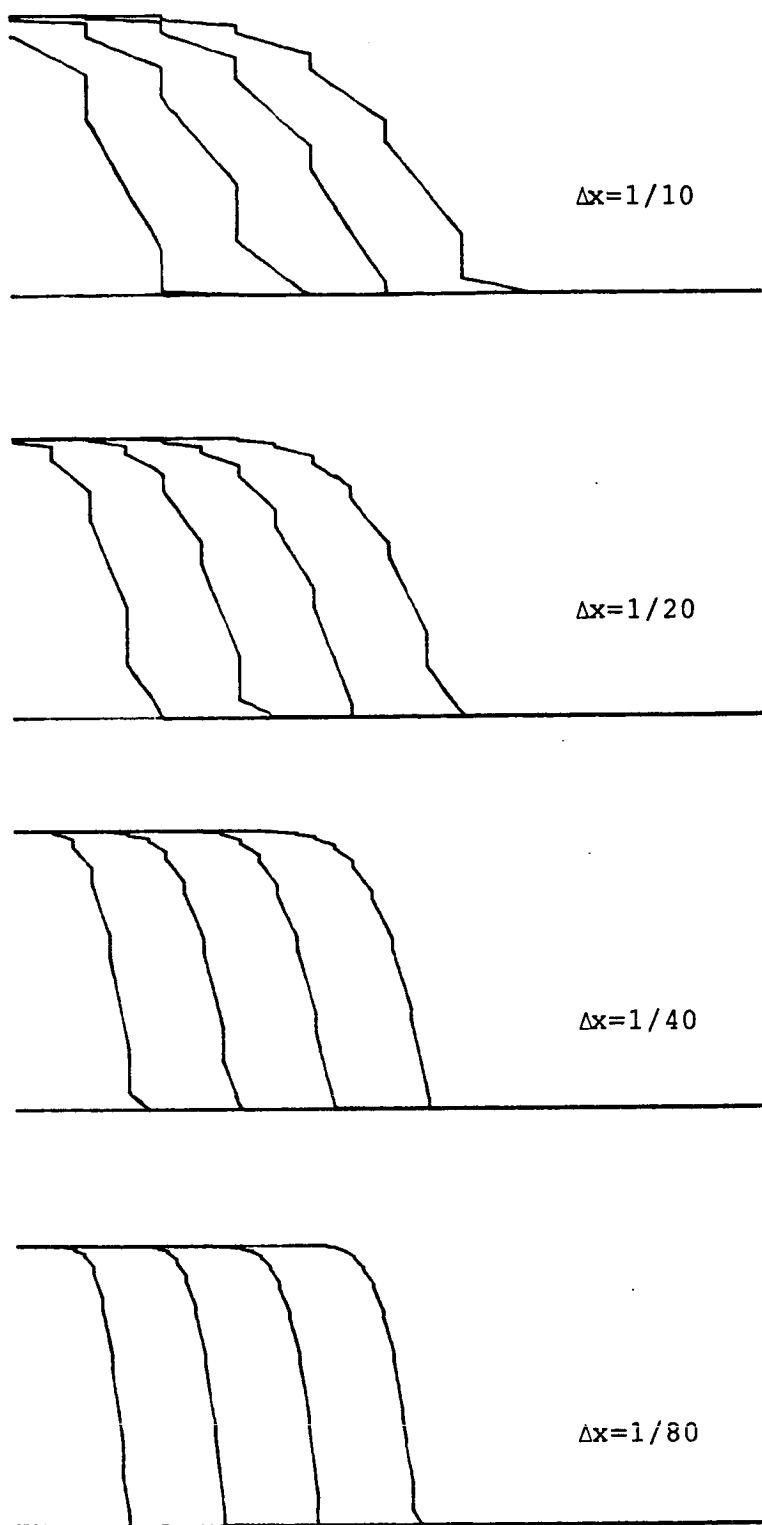


Figure 17 : Cas d'une discontinuité de contact : schéma BIP. Raffinement du maillage sous la condition CFL = 1/8.

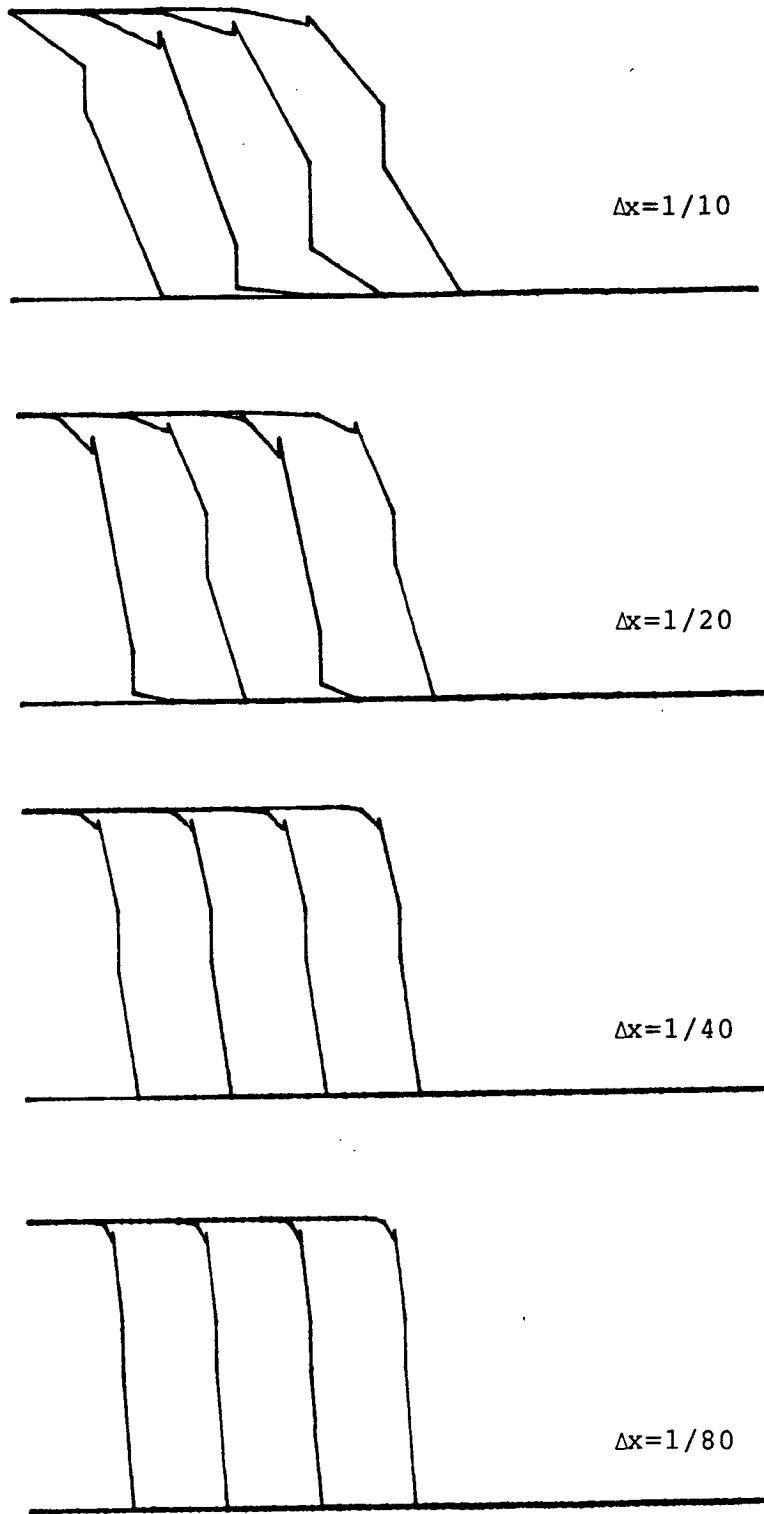


Figure 18 : Cas d'une discontinuité de contact : schéma LRGP. Raffinement du maillage sous la condition CFL = 1/8:

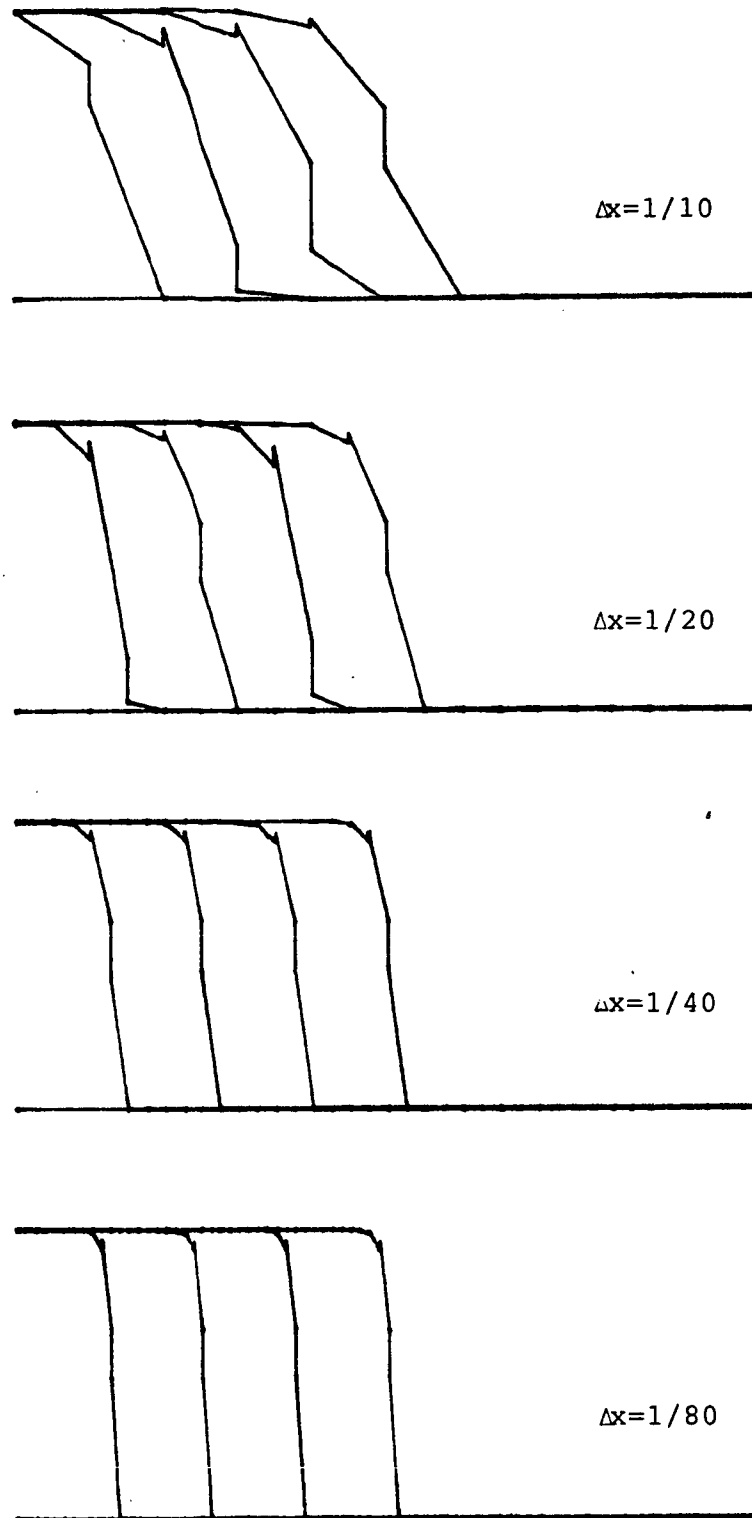
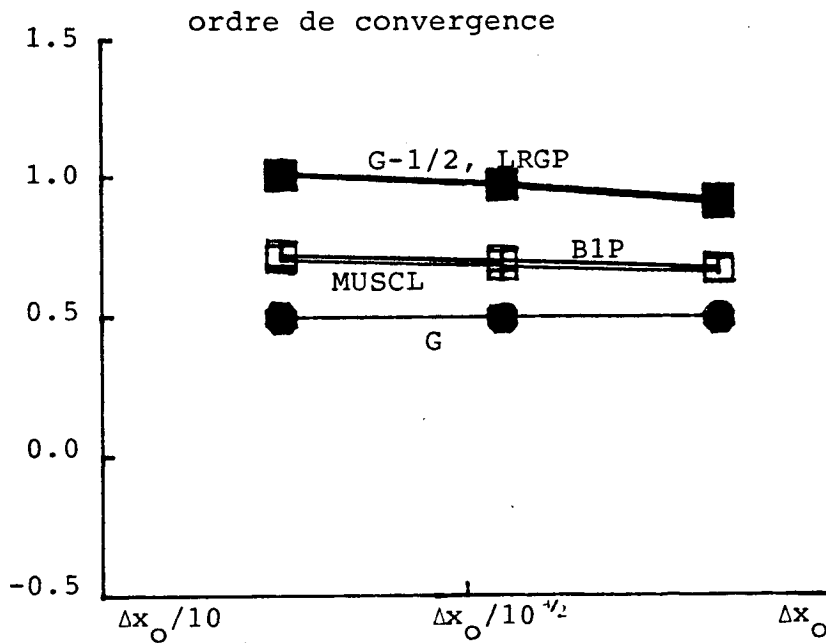
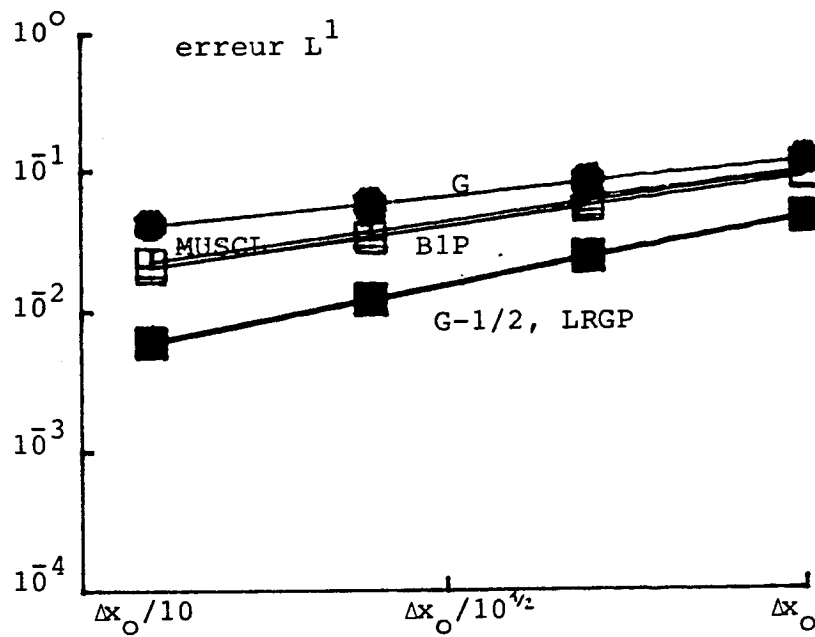


Figure 19 : Cas d'une discontinuité de contact : schéma G-1/2. Raffinement du maillage sous la condition CFL = 1/8.



**Figure 20** : Cas d'une discontinuité de contact : comparaison numérique des propriétés de convergence des différents schémas sous la condition  $CFL = 1/8$ . Les schémas G-1/2, LRGP, B1P coïncident pratiquement dans ce cas. (On a pris  $\Delta x_0 = 1/10$ , sauf pour le schéma de Godunov, pour lequel on a pris  $\Delta x_0 = 1/20$ ).

Ceci veut dire que pour ce cas, le meilleur choix est le schéma LRGP.

Avant de passer au paragraphe suivant, il faut remarquer que la variation de CFL a une influence très nette sur le comportement de nos schémas. En effet, une diminution de CFL produit :

- i) l'augmentation de la diffusion, c'est-à-dire, l'étalement du profil du front approchant la discontinuité de contact;
- ii) l'augmentation des erreurs  $L^1$  (comparer les figures 8, 14 et 20);
- iii) la diminution de l'ordre de convergence. Diminution *imperceptible* pour G, *très nette* pour MUSCL et B1P (on passe de 1, pour CFL = 1/2, à moins de 3/4 pour CFL = 1/8), *très nette* pour LRGP et G-1/2 quand on passe de CFL = 1/2 à CFL = 1/4 et *très faible* quand on passe de CFL = 1/4 à CFL = 1/8 (elle demeure à peu près égale à 1).

#### 4.3.3 - Résumé

La conclusion principale de ces expériences est que *le meilleur choix pour approcher une discontinuité de contact est le schéma LRGP*. En effet, il diffuse moins, il produit des erreurs plus petites et possède un ordre de convergence plus grand que les schémas G-1/2, B1P, MUSCL, G. D'autre part, comme le schéma LRG, il diffuse très peu et converge avec un ordre de convergence "égal" à 1, mais sous une condition de stabilité remarquablement moins restrictive (en plus, il ne présente pas des oscillations près de la discontinuité).

Dans le tableau 2, on résume les différentes propriétés des schémas observées sur les expériences faites.

Tableau 2 (propriétés théoriques et "expérimentales" des schémas, pour le cas (a)) Comparer au tableau 1

	G	MUSCL	BIP	LRGP	G-1/2	LRG
positif	oui	oui	oui	oui	oui	non
monotone	oui	oui	oui	au niveau des moyennes	au niveau des moyennes	non
diffusion	beaucoup	beaucoup moins que G	légèrement moins que MUSCL	légèrement moins que G-1/2	beaucoup moins que BIP	moins que LRG
oscillations au niveau de la discontinuité	non	non	non	non	non	oui
condition de stabilité	$CFL \leq 1$	$CFL \leq 1/2$	$CFL \leq 1/2$	$CFL \leq 1/2$	$CFL \leq 1/2$	$CFL \leq c_0 h^{1/2}$
ordre de convergence	1/2	$\approx 1.0$	pour $CFL = 1/2$	sol. exacte 1 fois sur 2 pour $CFL=1/2$	1'ordre augmente rapidement qd $h \rightarrow 0$ pour $CFL=1/2$	$\approx 1$
		$\approx 0.8$	pour $CFL = 1/4$	$\approx 1$	$\approx 1$	$\approx 1$
		$\approx 0.7$	pour $CFL = 1/2$	$(CFL < 1/2)$	$(CFL < 1/2)$	

Schémas "LRG-modifiés"

#### 4.4 - Un cas non-linéaire : l'équation de Buckley-Leverett : un choc

##### 4.4.1 - Le schéma LRG

Etant donné que ce schéma n'est pas positif, il est nécessaire de définir la fonction  $f$  pour des points hors de l'intervalle  $[0,1]$ . Dans les expériences numériques réalisées, on a pris la forme suivante de  $f$  :

$$P_2 u \quad \forall u < 0$$
$$f(u) = \frac{1}{2} u^2 (u^2 + (1-u)^2)^{-1} \quad \forall u \in [0,1]$$
$$\frac{1}{2} + P_2(u-1) \quad \forall u > 1$$

où  $P_1, P_2$  sont des constantes réelles.

On rappelle que lors de l'étude numérique *de la condition de stabilité* dans le cas non-linéaire, faite au paragraphe 3.4.2. on a utilisé un prolongement de ce type. Voir (3.90). Dans ce cas, on a pris  $P_1 = 0$  et  $P_2 = 30$ . Il est donc raisonnable d'utiliser ces valeurs pour ces expériences aussi.

Mais avec ce choix, et en prenant, comme dans 4.3.1 :

$$\begin{aligned} \text{CFL} &= 1/4 \quad , \quad h = 1/10 \\ \text{CFL} &= 1/8 \quad , \quad h = 1/40 \\ \text{CFL} &= 1/16 \quad , \quad h = 1/160 \end{aligned}$$

on trouve les erreurs suivantes :

$$e(h=1/10) = + \infty$$

$$e(h=1/40) = 0.18$$

$$e(h=1/160) = 0.22$$

(les résultats se trouvent sur la figure 21). Ceci montre qu'avec ce choix de  $P_1, P_2$  ce schéma ne converge pas. Ici, il faut faire quelques remarques importantes. On pourrait penser que ce résultat peut être dû

- i) à un mauvais choix de la méthode d'intégration numérique, méthode utilisée pour le calcul de l'erreur;
- ii) à un mauvais choix de la méthode de calcul de la solution exacte (on n'a pas d'expression explicite dans ce cas et il nous faut utiliser une fonction non-linéaire! )

mais on a utilisé trois méthodes d'intégration : la méthode des trapèzes, la méthode de Simpson, la méthode de Gauss (deux points), en variant la taille de l'intervalle d'intégration de  $h$  à  $h/20!$ , et on a obtenu le même comportement de l'erreur. Et d'autre part, on est sûr d'avoir calculé la solution exacte avec la précision limite de la machine.

En plus, et ceci est le point le plus important, avec la même méthode de calcul de l'erreur, et avec la même méthode de calcul de la solution exacte on a montré, comme on le verra, que *le reste des schémas converge!* On peut donc être sûr que c'est le mauvais choix des paramètres  $P_1, P_2$  la cause de cette propriété inconvenable.

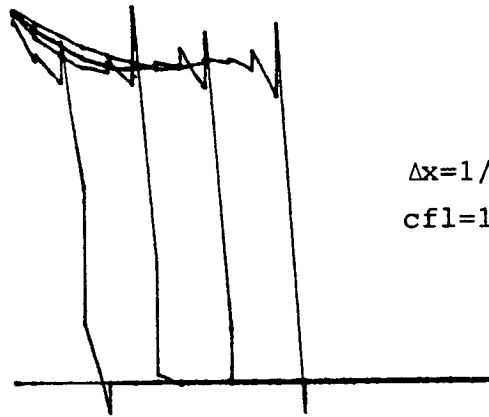
La question immédiate est donc la suivante : existe-t-il une paire de paramètres  $P_1, P_2$  pour lequel LRG converge ?



(explode)

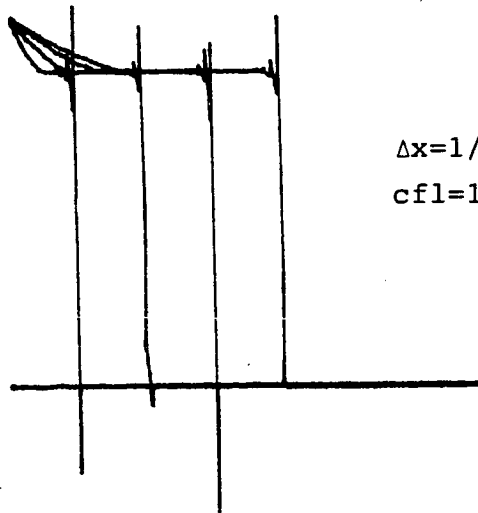
$$\Delta x = 1/10$$

$$cfl = 1/4$$



$$\Delta x = 1/40$$

$$cfl = 1/8$$



$$\Delta x = 1/160$$

$$cfl = 1/16$$

Figure 21 : Cas d'un choc : schéma LRG. Raffinement du maillage sous la condition de stabilité (du cas linéaire)  $\Delta t = \text{const } (\Delta x)^{3/2}$ . Premier prolongement de la fonction  $f$  (le schéma ne converge pas!)

On a fait varier  $P_2$  de 0 à 300, tout en prenant  $P_1=0$ , et les résultats ont été qualitativement égaux.

Sur la figure 22, on montre les résultats du cas  $P_1 = -10$ ,  $P_2 = 30$ , les erreurs produites dans ce cas sont :

$$e(h=1/10) = 0.17$$

$$e(h=1/40) = 0.16$$

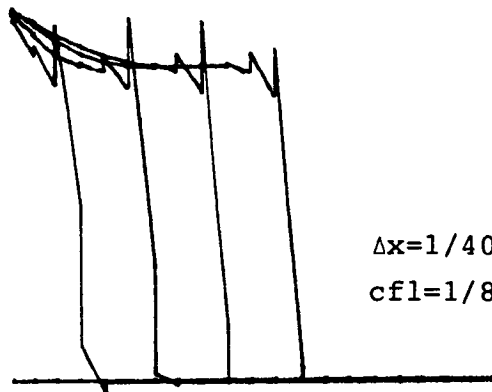
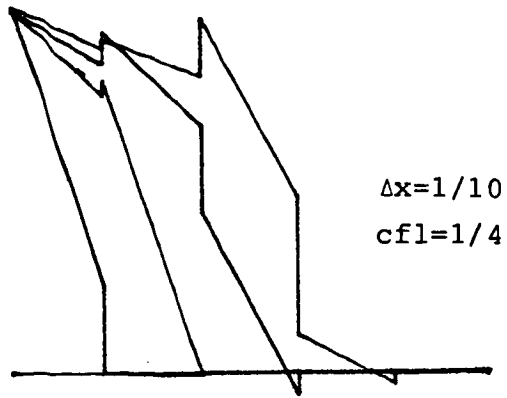
$$e(h=1/160) = + \infty$$

ce qui montre encore une fois que le schéma n'est pas convergent. On a aussi essayé le cas  $P_1 = +10$ ,  $P_2 = 30$  et on a obtenu des résultats semblables.

Peut-on conclure que cette paire de paramètres  $P_1, P_2$  n'existe pas et que le schéma n'est pas convergent? En fait, il faudrait attendre d'avoir une démonstration de cette propriété, et donc rigoureusement, on ne peut rien affirmer. Mais quand même, les expériences réalisées indiquent que pour un rang assez grand de paramètres, le schéma ne converge pas sous la condition de stabilité du cas linéaire!

Un autre indice dans cette direction peut être obtenu en comparant les figures 21 et 22, où la solution de LRG est montrée, avec la figure 26, où la solution donnée par LRGP est montrée. Deux remarques très importantes sont à faire :

- i) la région "régulière" de  $u_h$  donnée par LRG est toujours "au-dessus" de celle de  $u_h$  donné par LRGP. Ceci veut dire que la vitesse avec laquelle LRG "transporte" sa région "régulière" est inférieure à la véritable vitesse. Et ceci *indépendamment* du choix de  $P_1, P_2$ .
- ii) Par conséquent, étant donné que LRG est un schéma conservatif, ceci se traduit dans le fait que *le front approchant le choc se déplace à une vitesse aussi inférieure à la vitesse correcte*. Ainsi, le front donné par LRG est toujours "derrière" le choc!.



(explode)

$\Delta x=1/160$   
 $cfl=1/16$

Figure 22 : Cas d'un choc : schéma LRG. Raffinement du maillage sous la condition de stabilité (du cas linéaire)  $\Delta t = \text{const.} (\Delta x)^{3/2}$ . Deuxième prolongement de la fonction  $f$  (le schéma ne converge pas!)

Ceci étant donné, dans l'attente d'une démonstration rigoureuse, il est raisonnable de considérer que *dans ce cas* le schéma LRG ne converge pas.

#### 4.4.2 - Les autres schémas

##### i) Le cas CFL = 1/2

Sur les figures 23 à 27, on montre les solutions  $u_h$  données par les schémas G, MUSCL, LRGP et G-1/2, dans cet ordre. Encore une fois, on constate que tous ces schémas sont positifs et monotones (LRGP et G-1/2 sont monotones seulement sur les *moyennes*), mais dans ce cas, il est moins facile que dans le cas précédent de déterminer les schémas qui approchent le mieux la solution exacte. Tout de même, il est intéressant de remarquer que lorsque  $h \downarrow 0$ , la solution donnée par le schéma LRGP semble raidir son front plus que les autres, propriété "héritée" en quelque sorte, du schéma LRG.

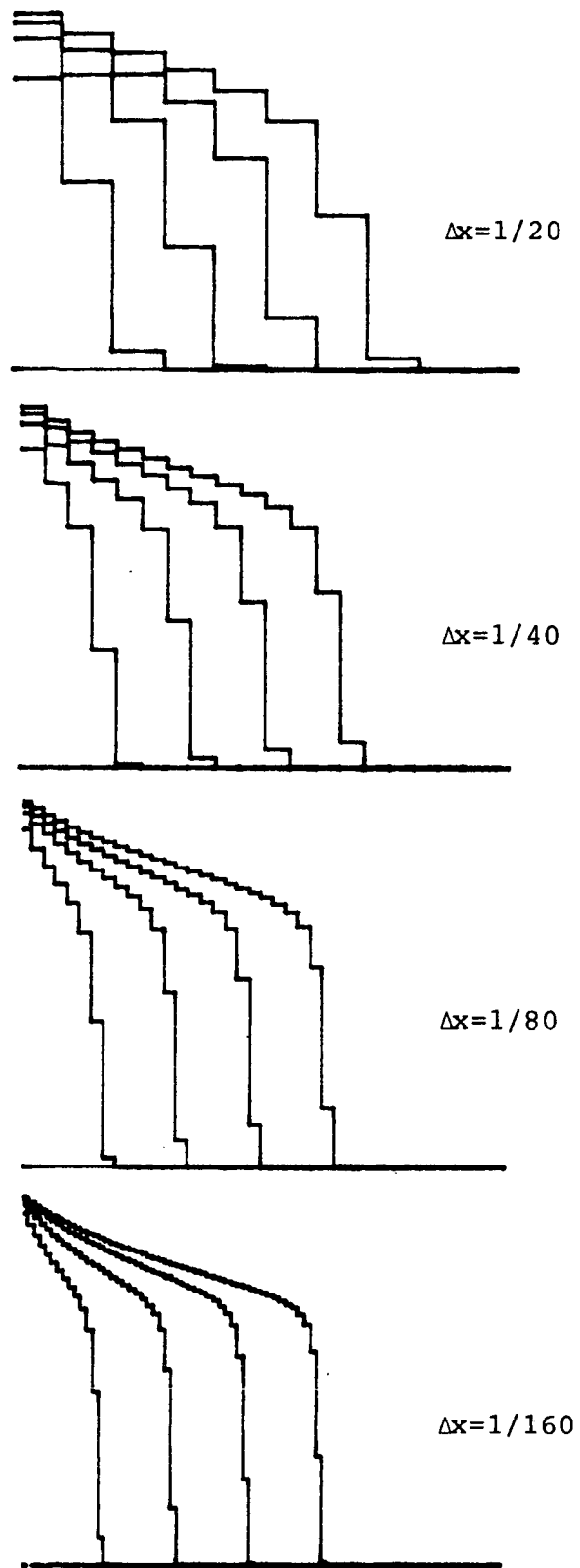


Figure 23 : Cas d'un choc : schéma de Godunov. Raffinement du maillage sous la condition CFL = 1/2.

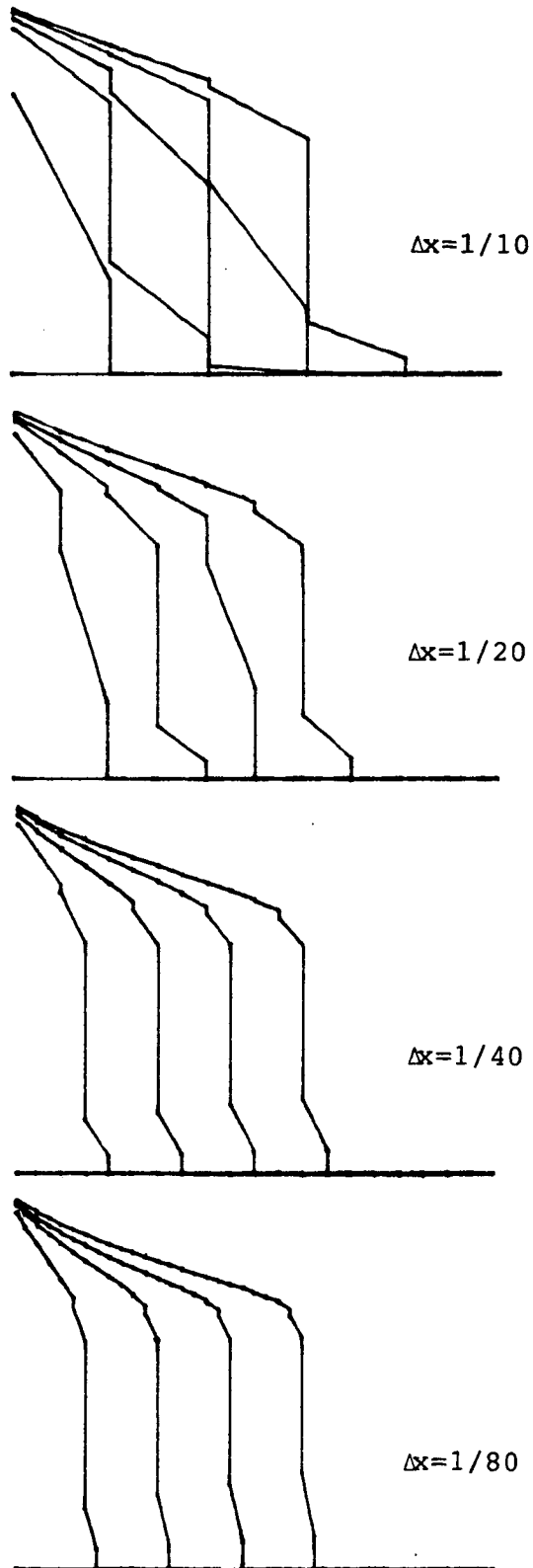


Figure 24 : Cas d'un choc : schéma MUSCL. Raffinement du maillage sous la condition CFL = 1/2.

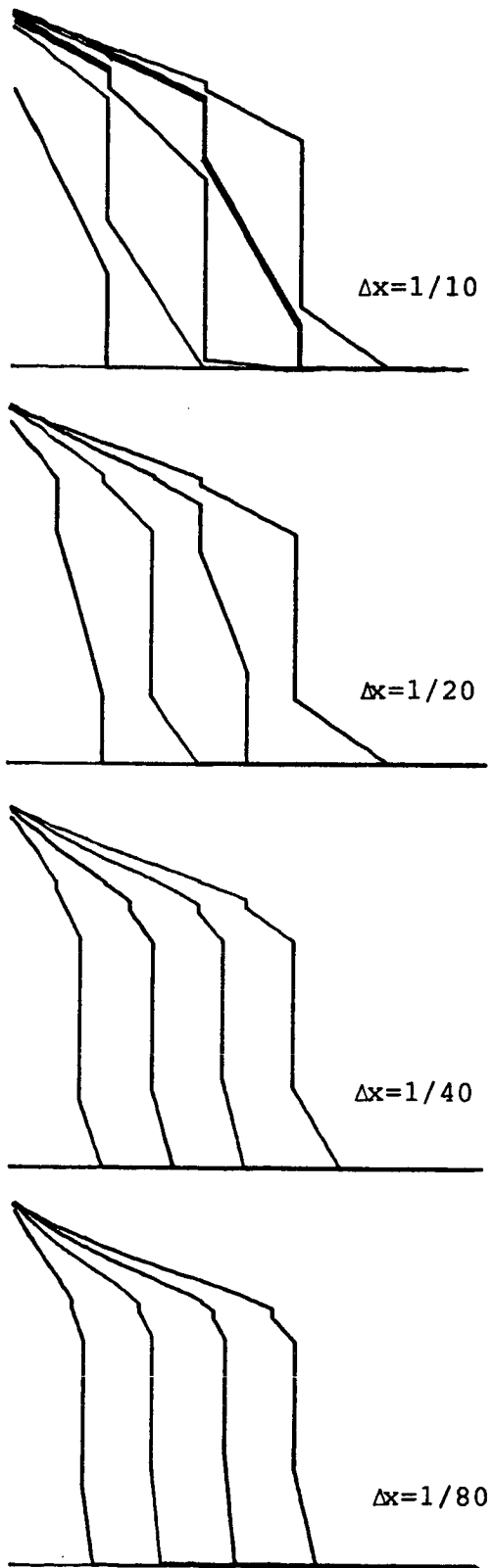


Figure 25 : Cas d'un choc : schéma BIP. Raffinement du maillage sous la condition CFL = 1/2.

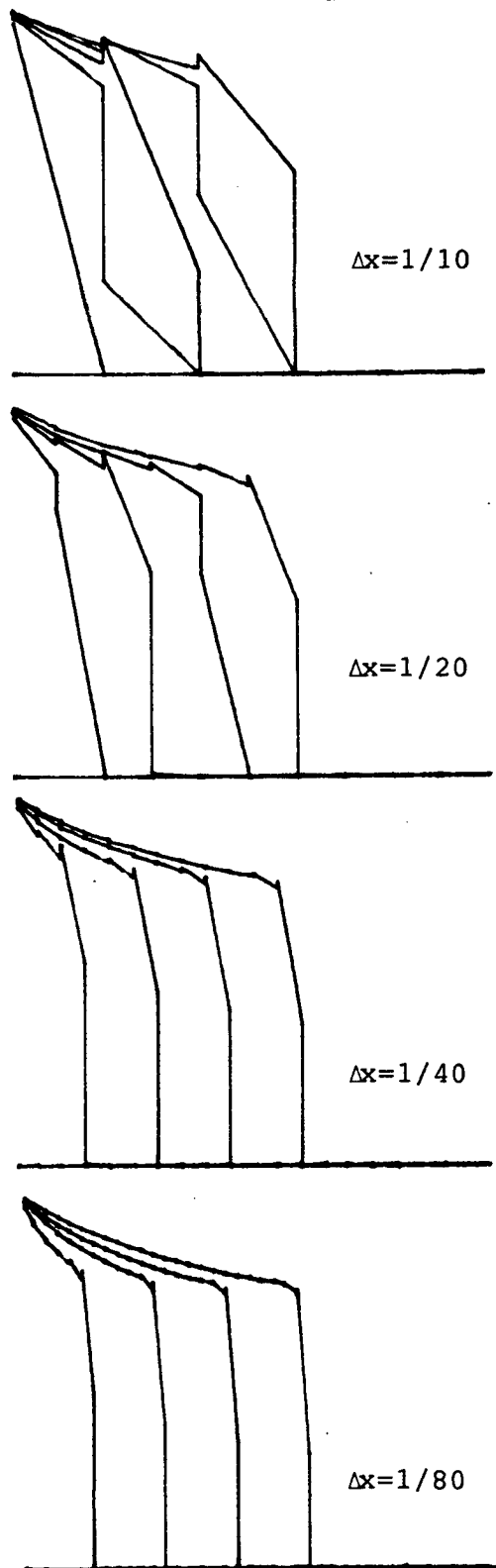


Figure 26 : Cas d'un choc : schéma LRGP. Raffinement du maillage sous la condition CFL = 1/2.



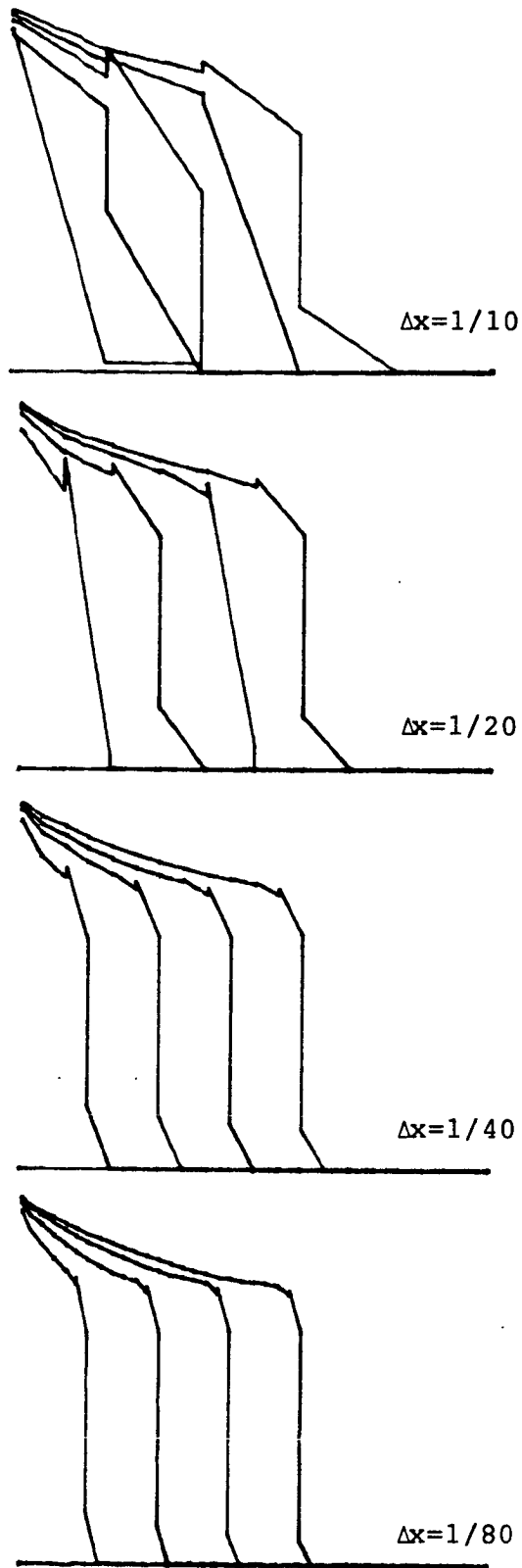


Figure 27 : Cas d'un choc : schéma G-1/2. Raffinement du maillage sous la condition CFL = 1/2.

Sur la figure 28, on montre les courbes d'erreur et d'ordre de convergence des schémas. On y constate :

- i) que si on ne considère pas la schéma LRGP, on peut ordonner les schémas du "meilleur au moins bon" de la façon suivante

$$G-1/2 \geq BIP \geq MUSCL \geq G$$

où "meilleur" signifie qu'il produit dans tous les cas *des erreurs plus petites*. Dans cette expérience, pour  $h_0 = 1/10$ , l'erreur produite par LRG est plus petite que celle produite par  $G-1/2$ , mais pour les autres valeurs de  $h$  il produit des erreurs plus grandes que celles de MUSCL et plus petites que celles de  $G$ .

- ii) que la situation est beaucoup moins claire en ce qui concerne l'ordre de convergence. Si on ne considère que l'ordre de convergence du plus petit  $h$ , on voit qu'il est

$$\begin{aligned} &\approx 4/5 \quad \text{pour BIP, G} \\ &\approx 3/4 \quad \text{pour MUSCL} \\ &\approx 3/5 \quad \text{pour G-1/2} \\ &\approx 1/2 \quad \text{pour LRGP.} \end{aligned}$$

D'autre part, il semble que lorsque  $h \rightarrow 0$ , tous les schémas semblent "augmenter" en ordre de convergence, sauf  $G-1/2$ .

Il est intéressant de voir comment les propriétés du schéma LRG se manifestent dans le schéma LRGP : au fur et à mesure que  $h \rightarrow 0$ , les fronts donnés par LRGP deviennent plus raides et le front commence à se situer derrière la position correcte du choc. C'est ce qui explique la forme de la courbe d'erreur de LRGP. Par contre, ce phénomène ne se manifeste de la même façon chez  $G-1/2$ ; en effet, les " $a_1$ " introduisent chez  $G-1/2$  la diffusion qui manque à LRGP, corrigeant ainsi ce défaut.

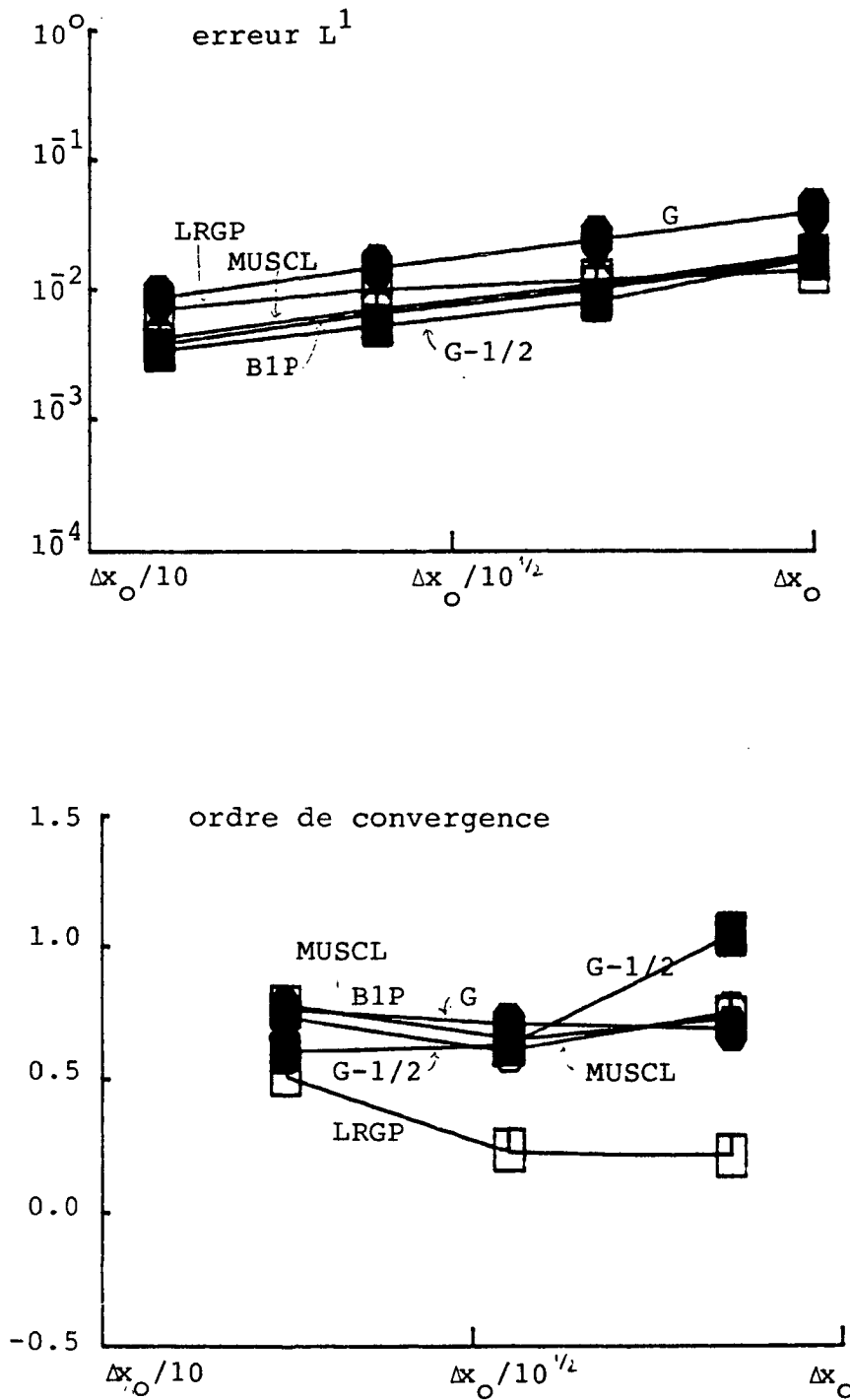


Figure 28 : Cas d'un choc : comparaison numérique des propriétés de convergence des différents schémas sous la condition CFL = 1/2. (On a pris  $\Delta x_0 = 1/10$ , sauf pour le schéma de Godunov, pour lequel  $\Delta x_0 = 1/20$ ).

ii) Le cas CFL = 1/4

Sur les figures 29 à 33, on montre la solution donnée par les schémas G, MUSCL, BIP, LRGP et G-1/2, dans cet ordre. Il est difficile de voir des différences par rapport au cas CFL = 1/2.

Considérons donc la figure 34, où l'on a dessiné les courbes d'erreur et d'ordre de convergence. On y constate

i) encore une fois le schéma LRGP mis à part, on peut ordonner du "meilleur au moins bon" les schémas comme pour CFL = 1/2

$$G-1/2 \geq BIP \geq MUSCL \geq G .$$

Le schéma LRGP produit des erreurs plus petites que G-1/2, sauf pour la dernière valeur de h;

ii) On peut aussi ordonner cette fois-ci tous les schémas, du meilleur au moins bon, de la façon suivante :

$$G-1/2 \geq BIP \geq MUSCL \geq G \geq LRGP$$

où "meilleur" signifie qu'il possède un ordre de convergence plus grand.

L'introduction de "diffusion" par le fait d'avoir diminué CFL explique le comportement de LRGP : les erreurs qu'il produit maintenant sont les plus petites et l'ordre de convergence a augmenté. Cependant, il possède l'ordre de convergence le plus bas d'entre tous les schémas considérés.

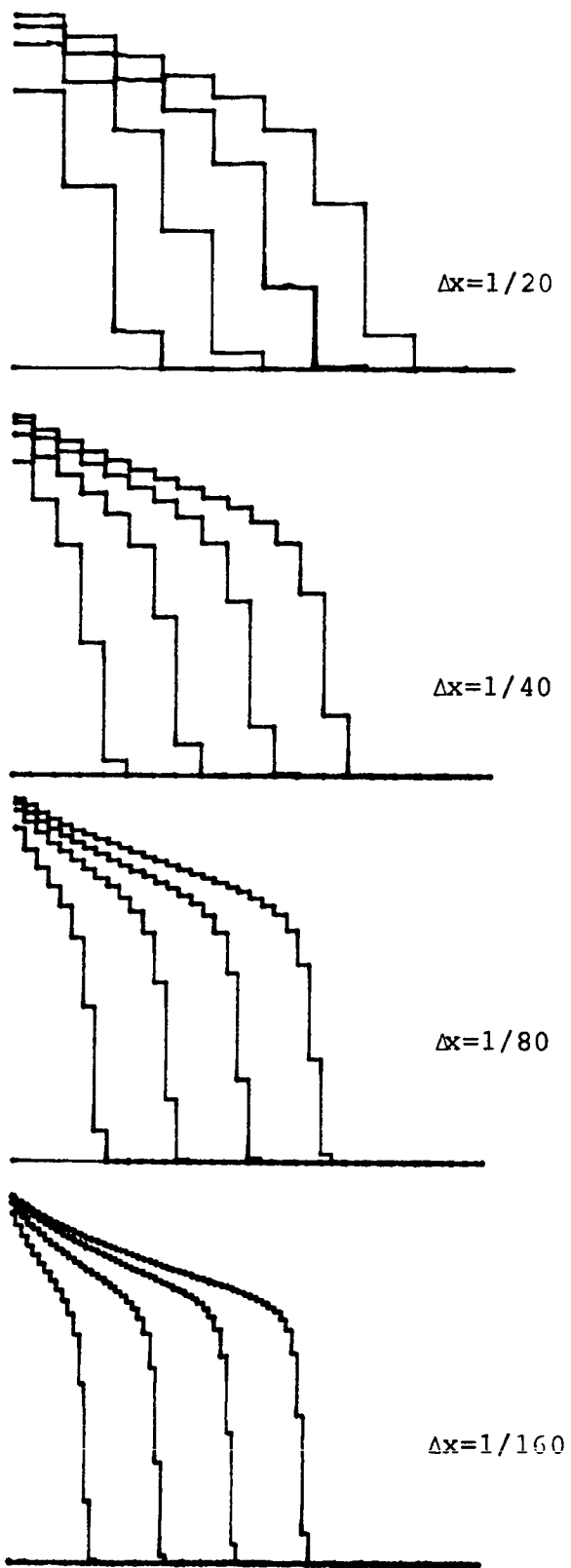


Figure 29 : Cas d'un choix : schéma de Godunov. Raffinement du maillage sous la condition  $CFL = 1/4$ .

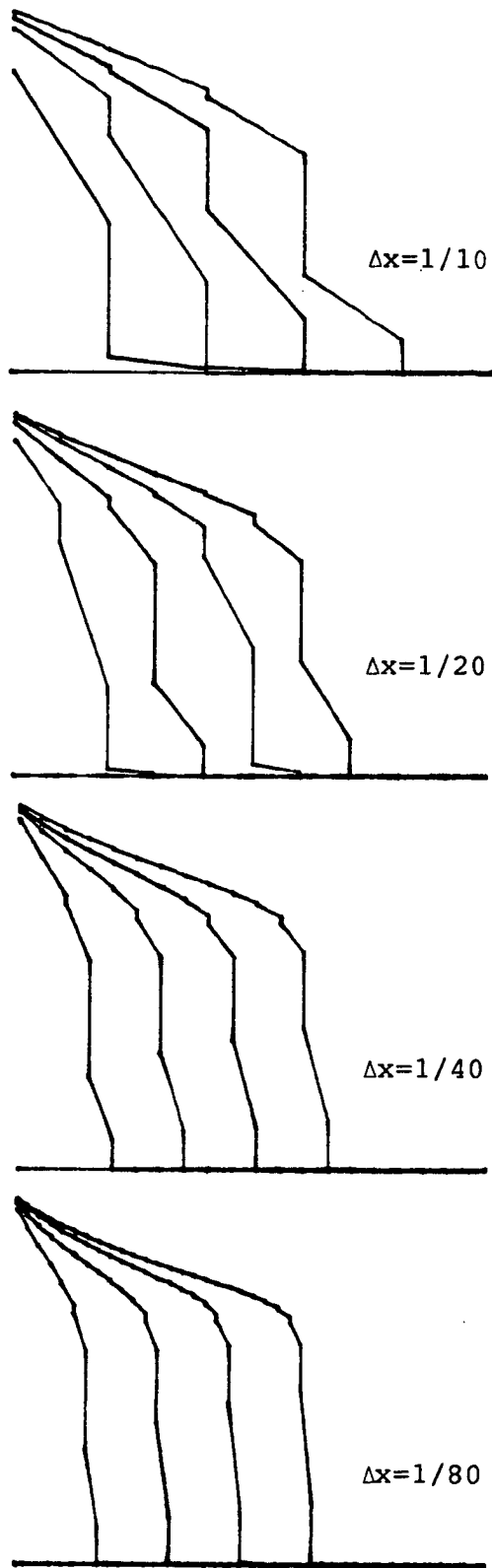


Figure 30 : Cas d'un choc : schéma MUSCL. Raffinement du maillage sous la condition CFL = 1/4.

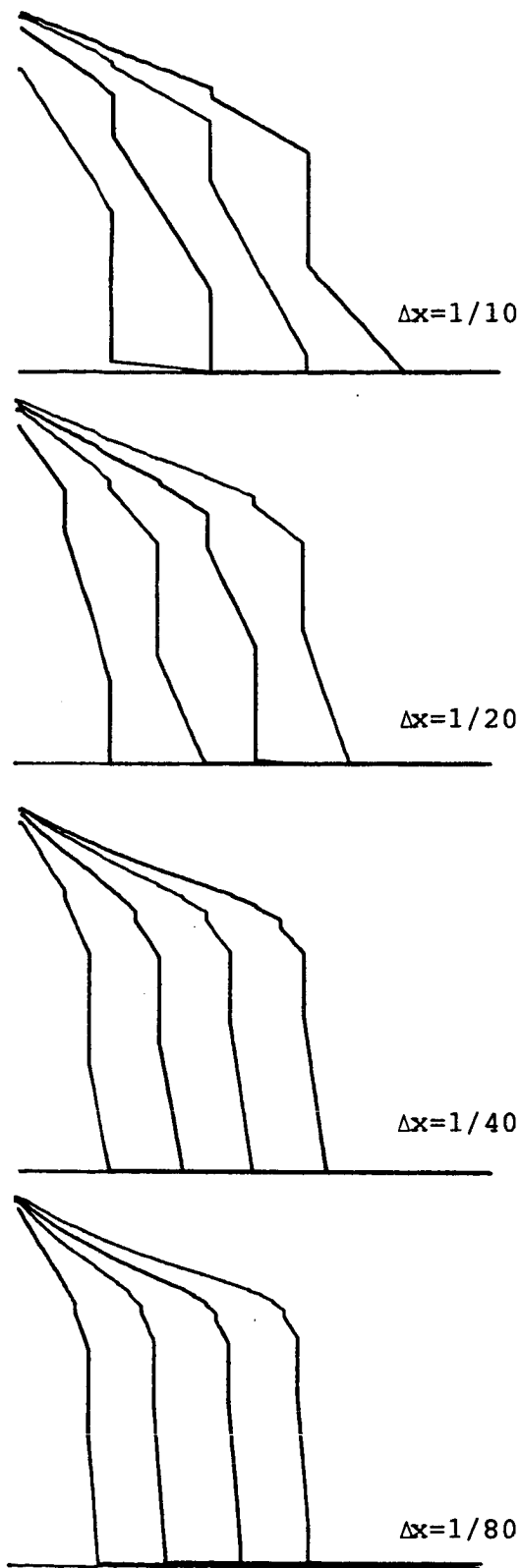


Figure 31 : Cas d'un choc : schéma BIP. Raffinement du maillage sous la condition CFL = 1/4.

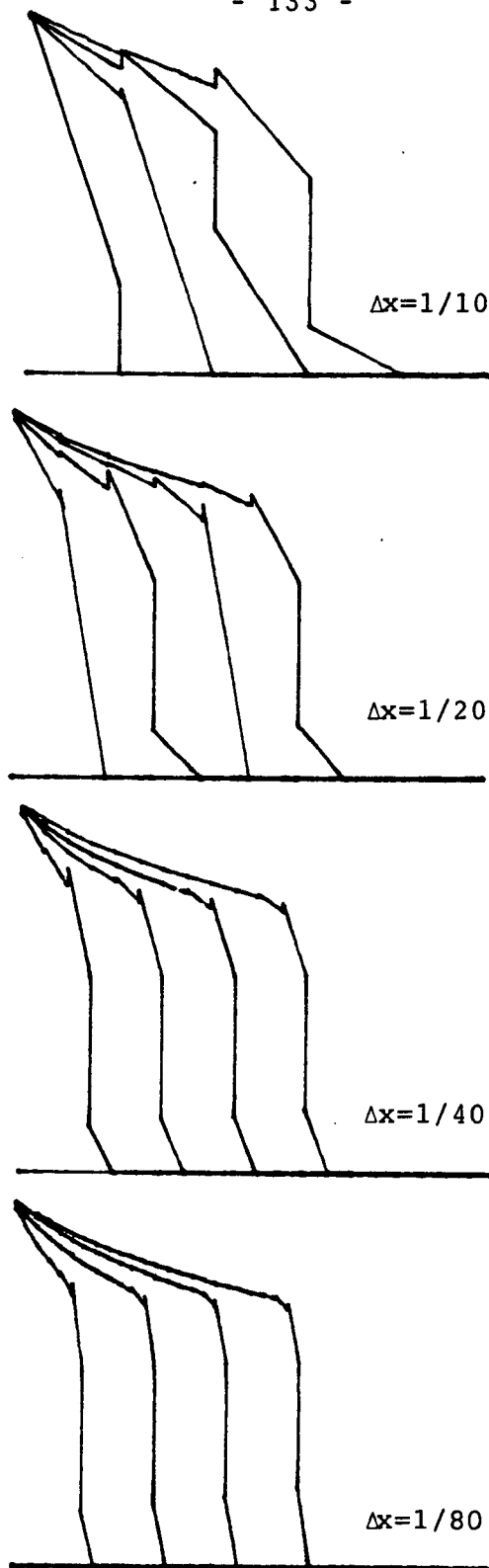


Figure 32 : Cas d'un choc : schéma LRGP. Raffinement du maillage sous la condition CFL = 1/4.



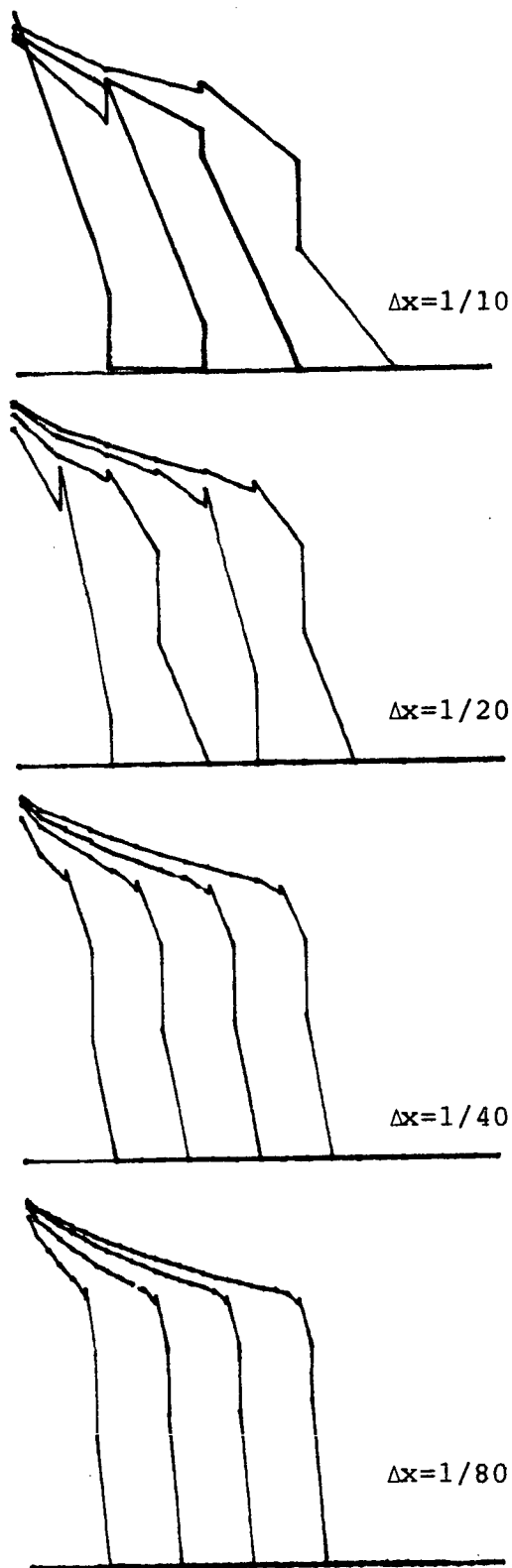
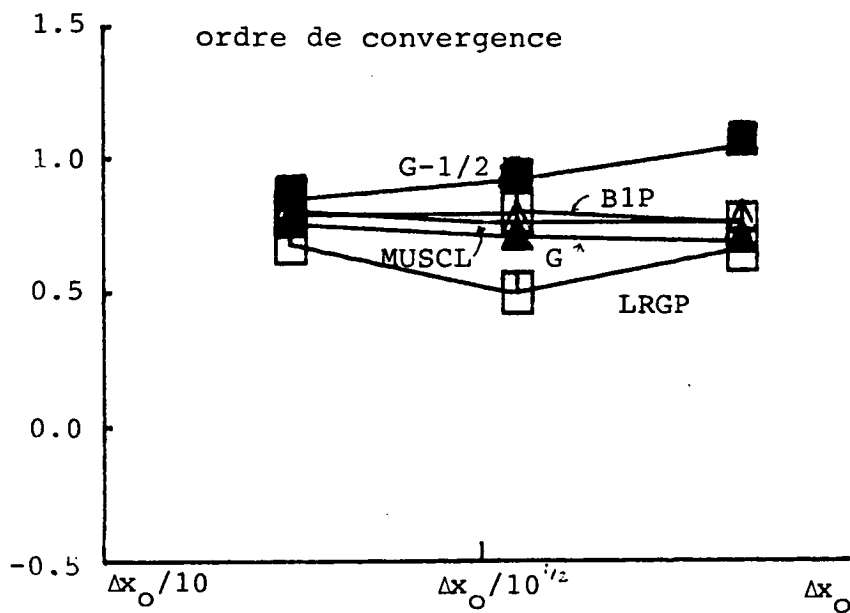
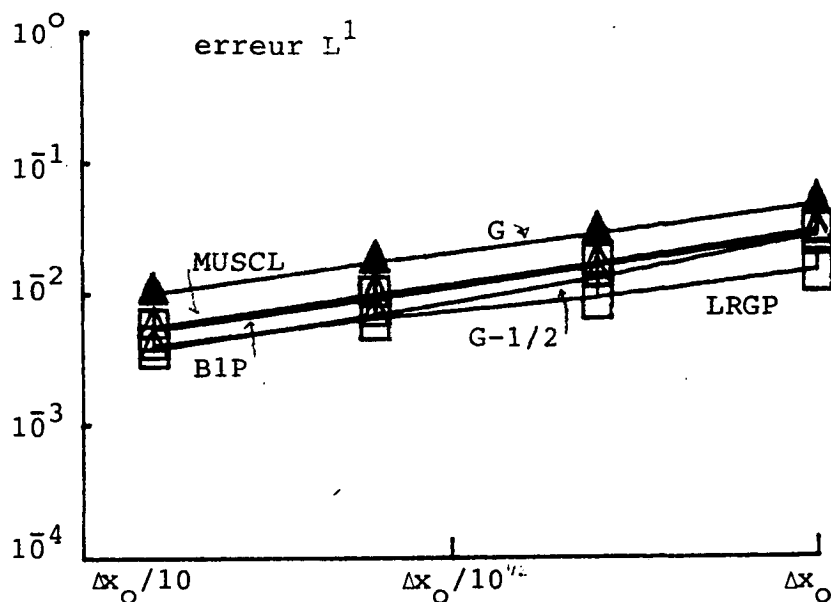


Figure 33 : Cas d'un choc : schéma G-1/2. Raffinement du maillage sous la condition CFL = 1/4.



**Figure 34** : Cas d'un choc : comparaison numérique des propriétés de convergence des différents schémas sous la condition CFL = 1/4.  
 (On a pris  $\Delta x_0 = 1/10$ , sauf pour le schéma de Godunov pour lequel  $\Delta x_0 = 1/20$ ).

iii) Le cas CFL = 1/8

Sur les figures 35 à 39, on montre les solutions  $u_h$  données par les schémas G, MUSCL, BIP, LRGP et G-1/2, respectivement. On y observe les mêmes propriétés que dans les cas CFL = 1/2 et CFL = 1/4.

Considérons la figure 40, on y constate :

- i) en ce qui concerne les erreurs, on voit que MUSCL produit des erreurs plus petites que G; BIP produit des erreurs très très légèrement plus petites que MUSCL; ensuite c'est G-1/2 qui produit les erreurs plus petites que BIP, et finalement LRGP produit les erreurs plus petites. On a ainsi :

$$\text{LRGP} \geq \text{G-1/2} \geq \text{BIP} \geq \text{MUSCL} \geq \text{G} .$$

Une comparaison rapide avec la figure 34 nous montre que la diminution de CFL n'a pas trop changé le comportement des erreurs;

- ii) en ce qui concerne l'ordre de convergence, on voit maintenant que l'ordre de convergence de G-1/2 est  $\approx 1$ . Pour MUSCL et BIP, l'ordre de convergence est probablement le même,  $\approx 0.8$ , et pour G il est  $\approx 3/4$ . L'ordre de convergence de LRGP a augmenté en proportion de la diminution de CFL, pour la valeur plus petite de  $h$ , il se situe entre celui de G et celui de BIP.

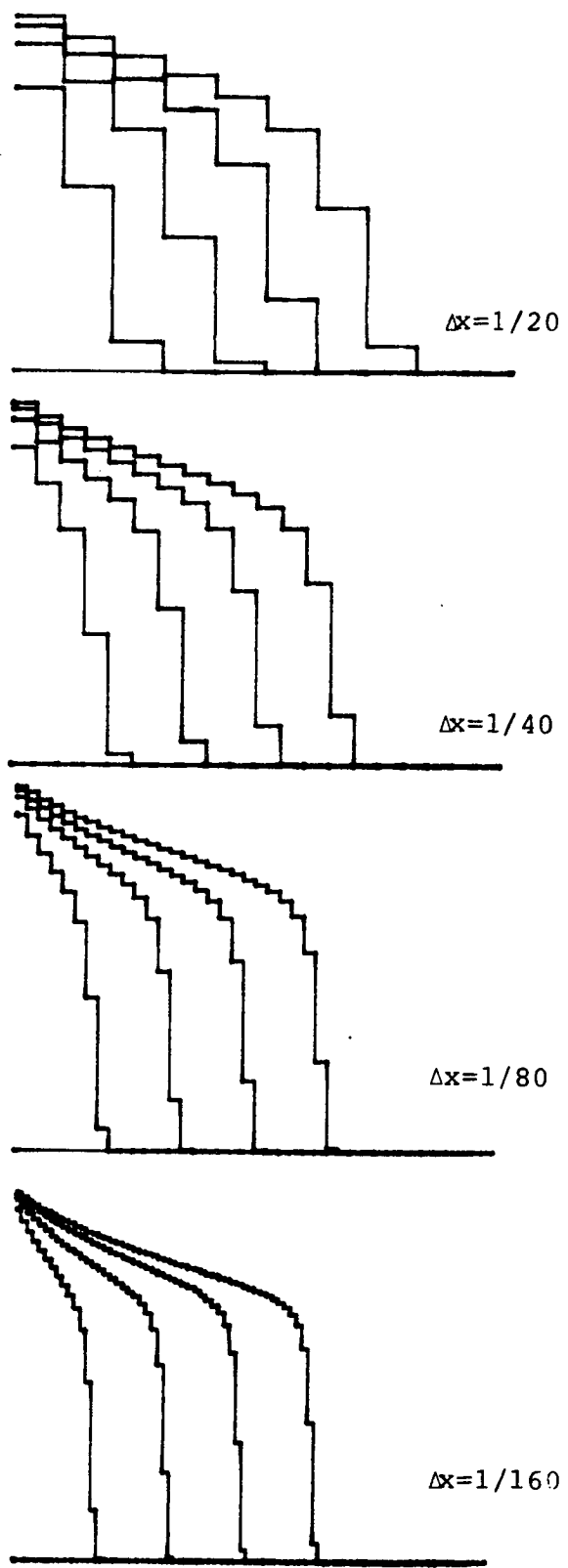


Figure 35 : Cas d'un choc : schéma de Godunov. Raffinement du maillage sous la condition CFL = 1/8.

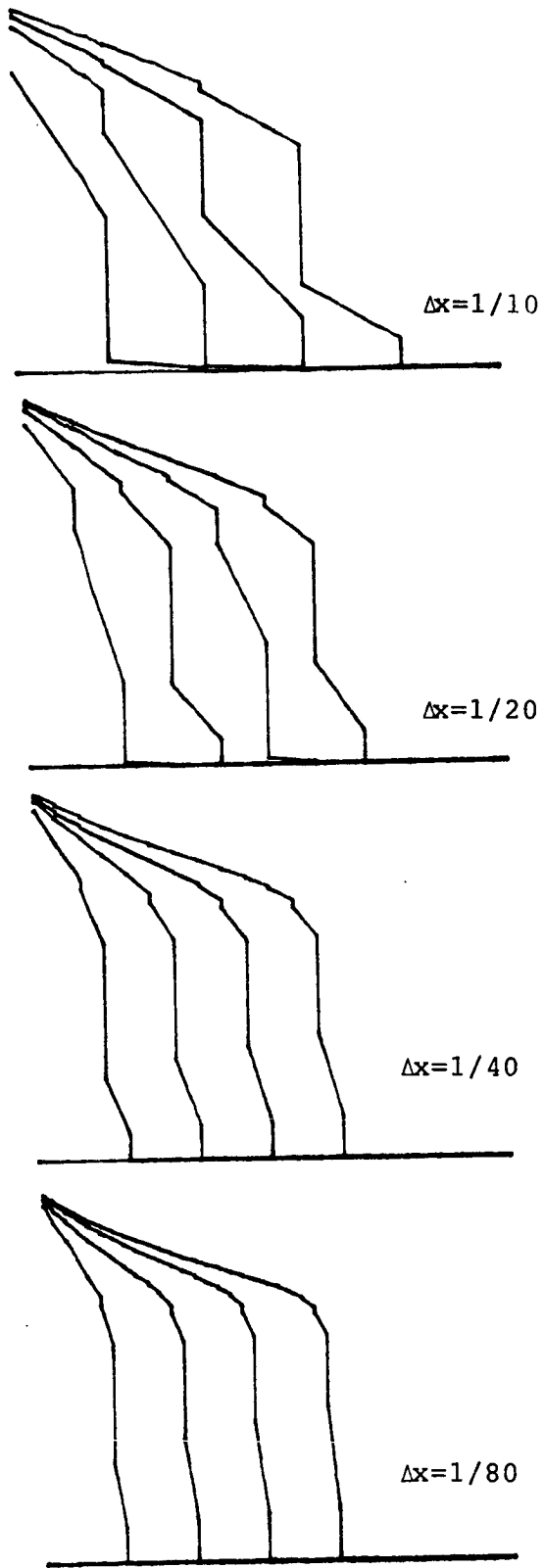


Figure 36 : Cas d'un choc : schéma MUSCL. Raffinement du maillage sous la condition CFL = 1/8.

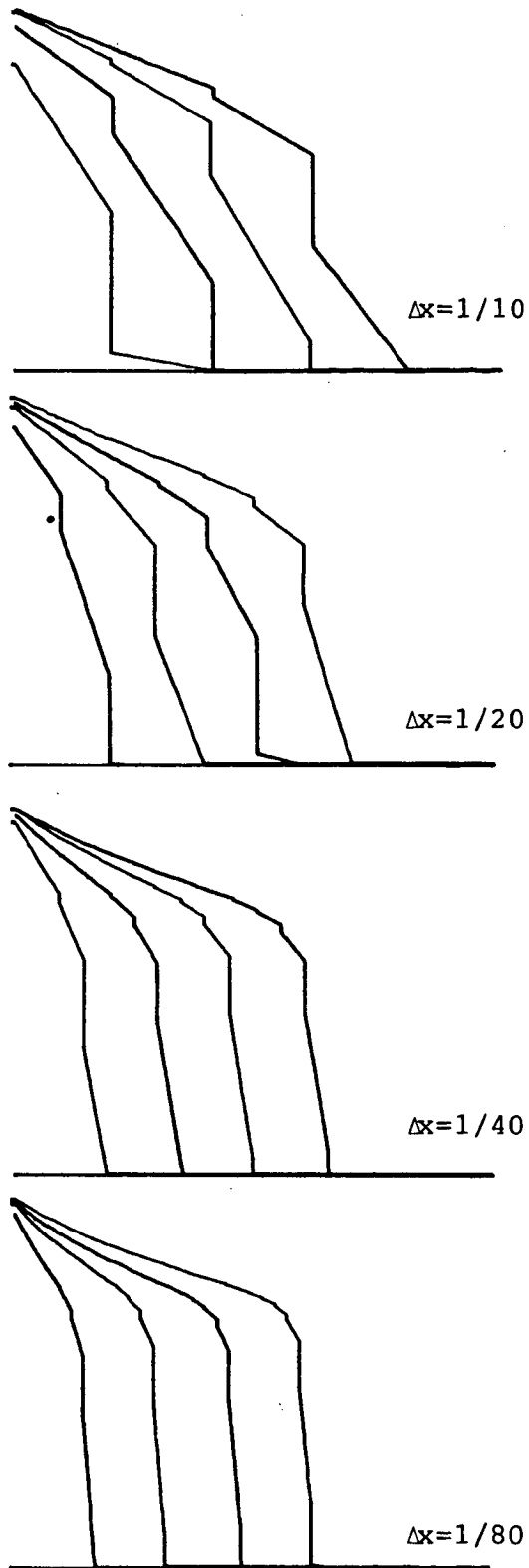


Figure 37 : Cas d'un choc : schéma BIP. Raffinement du maillage sous la condition CFL = 1/8.

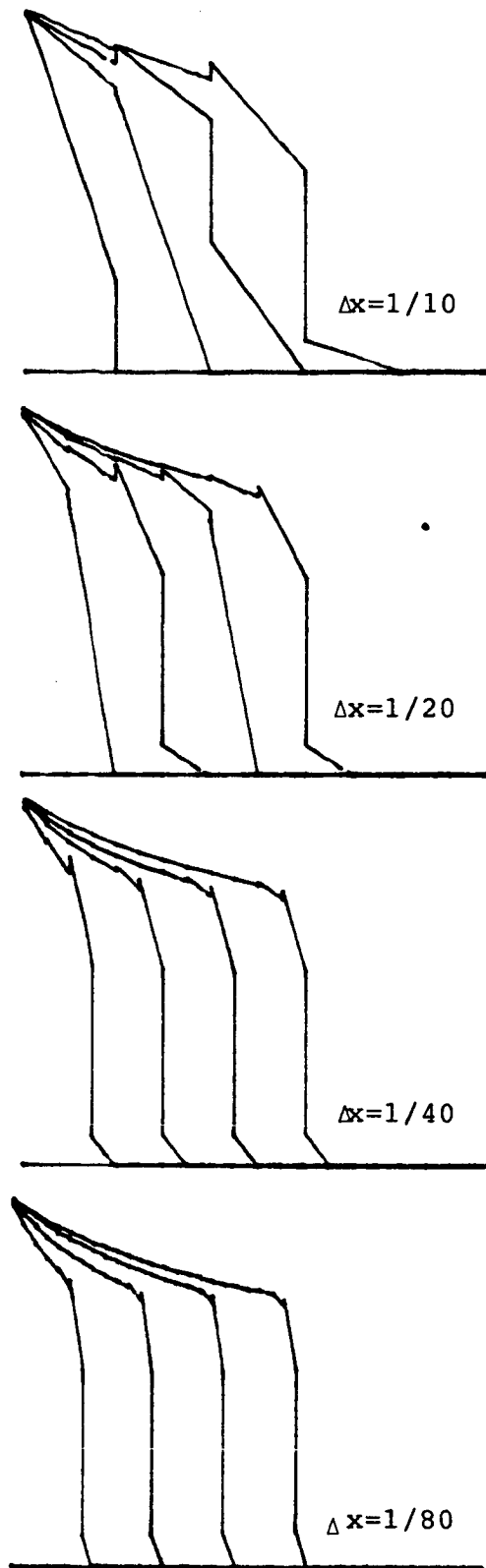


Figure 38 : Cas d'un choc : schéma LRGP. Raffinement du maillage sous la condition CFL = 1/8.

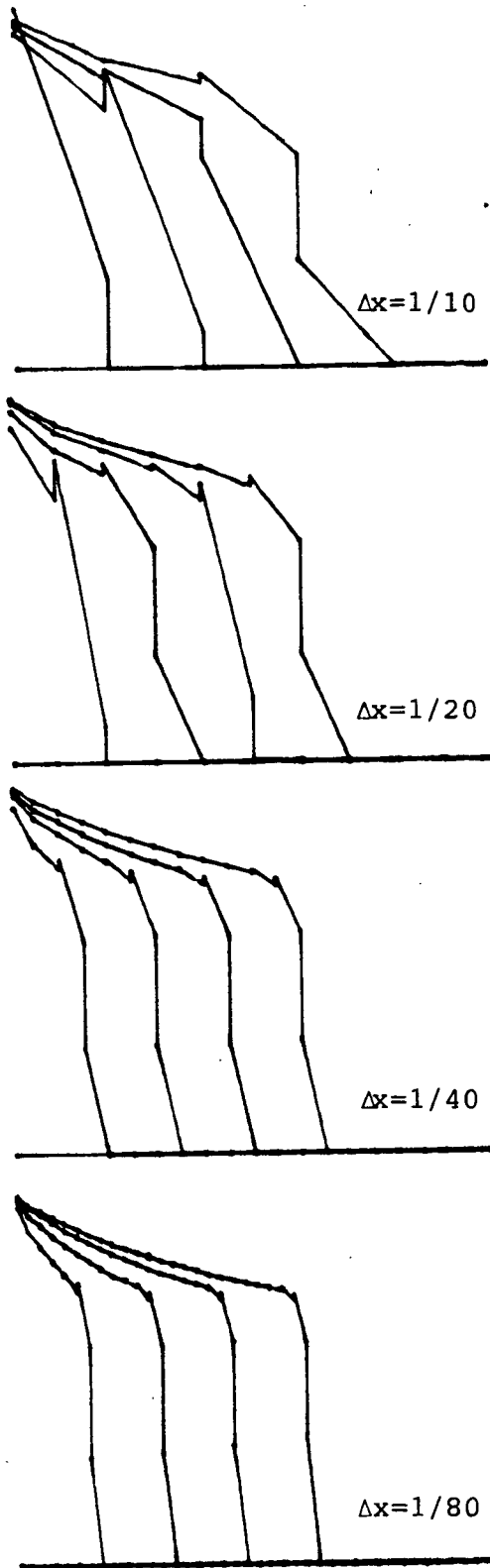


Figure 39 : Cas d'un choc : schéma G-1/2. Raffinement du maillage sous la condition CFL = 1/8.



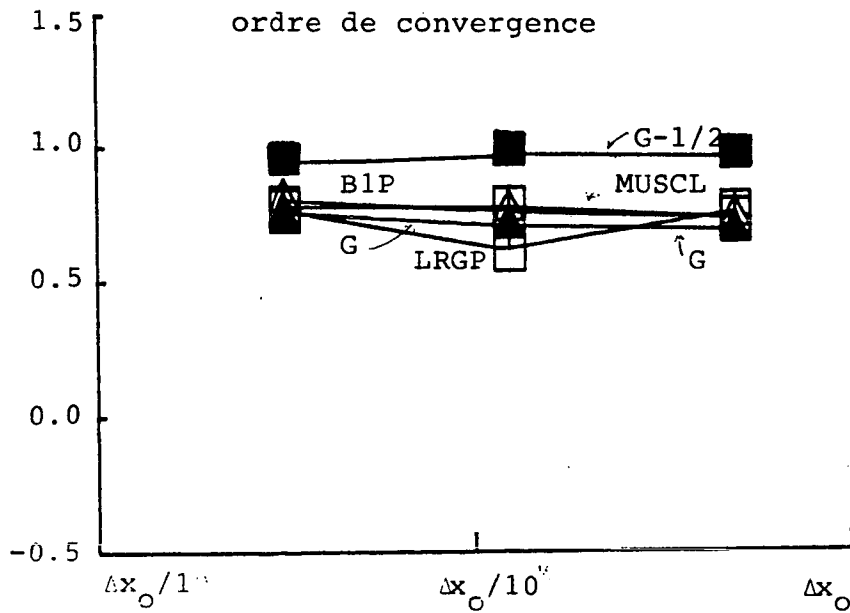
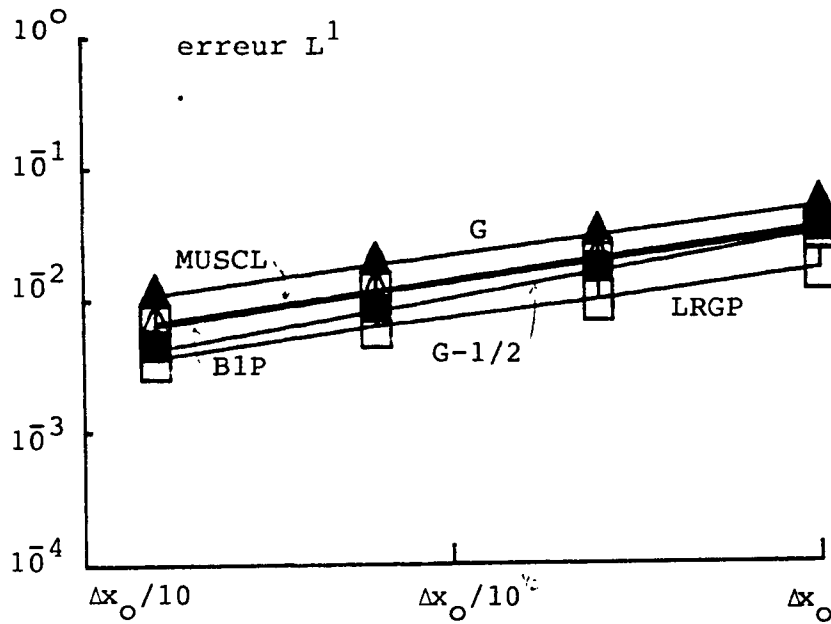


Figure 40 : Cas d'un choc : comparaison numérique des propriétés de convergence des différents schémas sous la condition CFL = 1/8.

(On a pris  $\Delta x_0 = 1/10$  sauf pour le schéma de Godunov pour lequel  $\Delta x_0 = 1/20$ )

iv) Quelques commentaires

Dans ces expériences numériques, on voit que :

- i) comme dans le premier cas, voir 4.3, tous les schémas semblent être positifs et monotones (pour LRGP et G-1/2 la monotonie est au niveau des moyennes seulement)
- ii) si on ne considère pas le schéma LRGP, on peut ordonner les schémas de la façon suivante

$$G-1/2 > BIP \tilde{\succ} MUSCL > G$$

où  $A > B$  (resp.  $\tilde{\succ}$ ) veut dire qu'*en général*, le schéma A produit des erreurs plus petites (resp. très semblables) et possède un ordre de convergence supérieur (resp. très semblable) à celui de B;

- iii) l'ordre de convergence de LRGP est inférieur *en général* à celui de G mais il varie beaucoup avec CFL. En effet, la diminution de CFL introduit de la diffusion et ceci a pour conséquence l'augmentation de l'ordre de convergence de LRGP. Une autre conséquence est la diminution de l'erreur produite par ce schéma. Pour  $CFL \geq 1/4$  et  $h \geq 1/10$ , LRGP produit l'erreur la plus petite;
- iv) lorsque  $h \rightarrow 0$ , le schéma LRGP a une forte tendance à reproduire le comportement du schéma LRG, qui ne converge pas dans ce cas. En effet, le front approchant le choc devient plus raide et se déplace à une vitesse inférieure à la vitesse correcte. Cet inconvénient est surmonté par l'introduction de diffusion, ce qui peut se faire, soit en *augmentant*  $h$ , soit en *diminuant* CFL!

4.4.3 - Résumé

La conclusion principale de ces expériences est que, pour CFL "petite" et  $h$  "grand", le meilleur choix est LRGP, autrement, le meilleur choix est G-1/2, surtout si on veut prendre  $h$  "petit". Contrairement au schéma LRG,

*qui ne converge pas* dans ce cas, tous les schémas "LRG-modifiés" semblent converger (avec un ordre de magnitude qui varie entre 0.5 et 1.0, si on ne considère pas LRGP pour CFL = 1/2). LRG donne un choc trop raide qui se déplace moins vite que le choc de la solution exacte. On observe ce comportement aussi dans LRGP, lorsque  $h \rightarrow 0$ , mais on corrige ceci en augmentant  $h$  ou en diminuant CFL.

Dans le tableau 3, on résume les différentes propriétés des schémas observées sur les expériences faites.

	G	MUSCL	BIP	LRGP	G-1/2	LRG
Cond. de stabilité	$CFL \leq 1$	$CFL \leq 1/2$	$CFL \leq 1/2$	$CFL \leq 1/2$	$CFL \leq 1/2$	$CFL \leq c h^{1/2} (?)$
Positif	oui	oui	oui	oui	oui	non
Monotone	oui	oui	oui	au niveau des moyennes	au niveau des moyennes	non
Erreur	plus petites que celle de G	très légèrement plus petites que celles de MUSCL	pour CFL petit et h "grand", il produit les plus petites erreurs	plus petites que celles de BIP et de LRGP pour h petit	ne converge pas !	
Ordre de convergence	$\approx .80$	$\approx .80$	$\approx .75$	$\approx 0.5$	$\approx 0.6$	-
	$CFL=1/2$	$CFL=1/4$	$\approx .85$	$\approx 0.65$	$\approx 0.9$	
	$CFL=1/8$			$\approx .85$	$\approx 1.0$	

schémas "LRG-modifiés"

Tableau 3 : propriétés théoriques et "expérimentales" des schémas pour le cas (b)

4.5 - Un cas non-linéaire : la fonction  $f$  est convexe et la solution  $\mathcal{C}^0 \forall t > 0$  : une étude de détente

4.5.1 - Le schéma LRG

Sur la figure 41, on montre la solution  $u_h$  donnée par LRG. On a pris, comme dans les cas précédents :

$$\begin{aligned} \text{CFL} &= 1/4 \quad , \quad h = 1/10 \\ \text{CFL} &= 1/8 \quad , \quad h = 1/40 \\ \text{CFL} &= 1/16 \quad , \quad h = 1/160 \end{aligned}$$

c'est-à-dire, on fait tendre  $h$  vers zéro de façon à vérifier la condition de stabilité du cas linéaire. On voit que  $u_h$  semble converger vers la solution entropique, en particulier les oscillations de  $u_h$  aux niveaux  $u_h = 0$  ,  $u_h = 1$  sont amorties rapidement (plus rapidement que dans le cas (a), d'une discontinuité de contact, ce qui est d'ailleurs très raisonnable), lorsque  $h \rightarrow 0$ . En effet, on peut constater que  $u_h$  converge vers la solution entropique en regardant la figure 42. L'ordre de convergence est  $\approx 1.14$  !

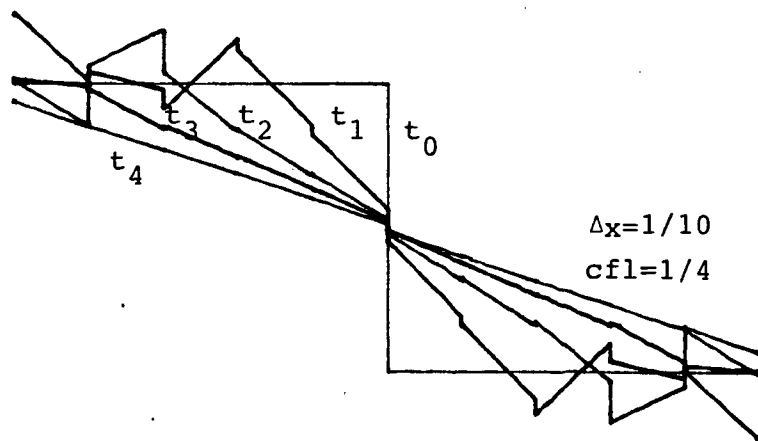
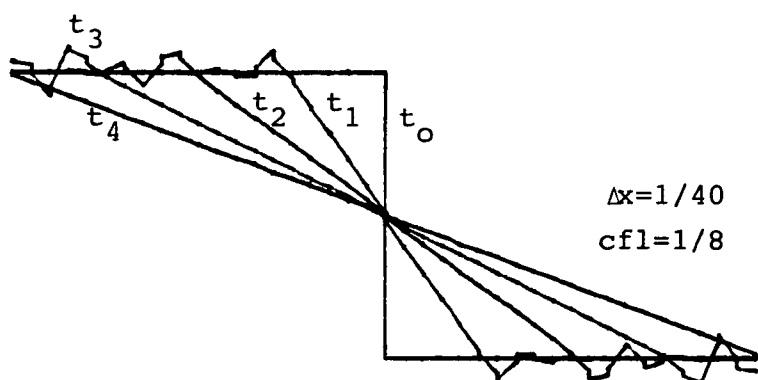
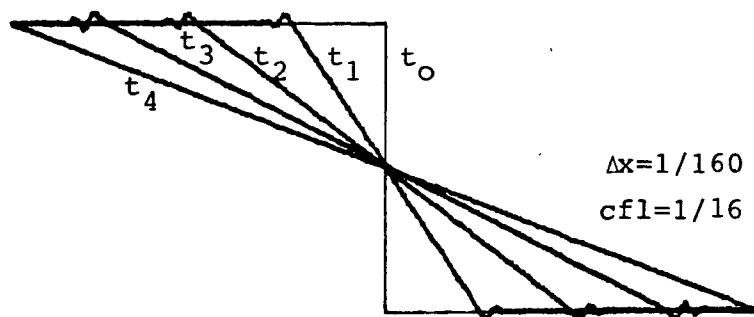


Figure 41 : Cas d'une onde de détente : schéma LRG. Raffinement du maillage sous la condition (de stabilité linéaire)  $\Delta t = \text{const}(\Delta x)^{3/2}$

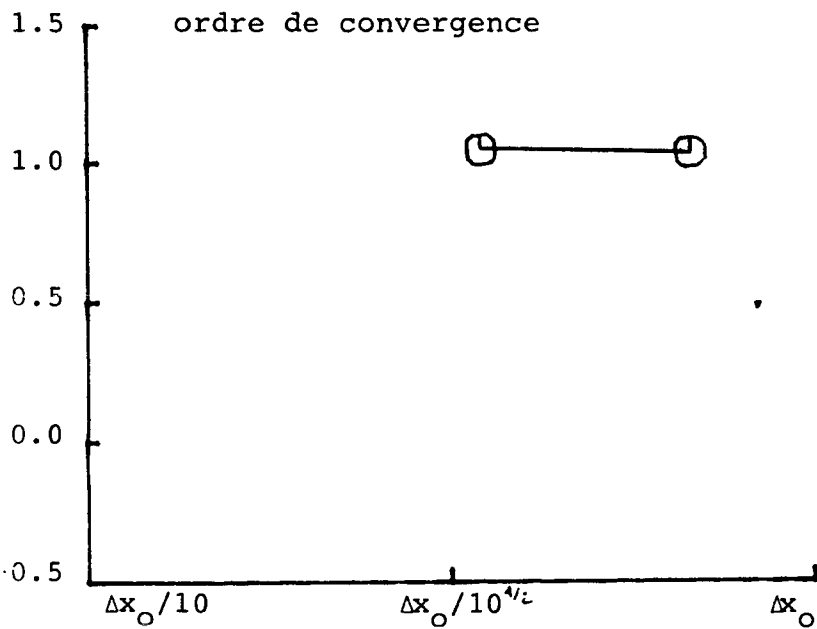
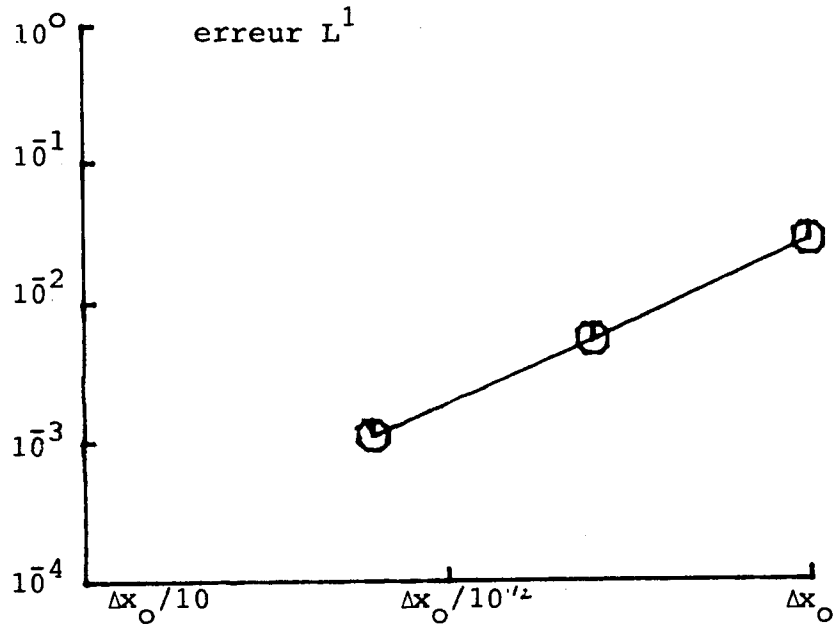


Figure 42 : Cas d'une onde de détente : schéma LRG. Raffinement du maillage sous la condition de stabilité linéaire  $\Delta t = \text{const.}(\Delta x)^{3/2}$ . (On a pris  $\Delta x_0 = 1/10$ ).

Pour voir l'effet des conditions aux bords sur l'évolution de l'erreur, on a pris l'erreur à  $t = t_4$  et aussi à  $t = t_2$ , temps pour lequel l'influence des conditions aux bords est négligeable. Les résultats sont semblables. Pour les deux cas l'ordre de convergence est à peu près de 1.14.

#### 4.5.2 - Les autres schémas

##### i) Le cas CFL = 1/2

Sur les figures 43 à 47, on montre les solutions  $u_h$  données par les schémas G, MUSCL, BIP, LRGP et G, respectivement. On y constate que :

- i) encore une fois, tous les schémas semblent être positifs et monotones (sauf pour LRGP et G-1/2 pour lesquels la monotonie est seulement pour les moyennes);
- ii) la solution exacte pour le dernier temps considéré,  $t_0$ , est la ligne droite joignant les extrémités de la polygonale qui est la condition initiale. En comparant  $u_h(t_0)$  à cette ligne, on peut obtenir beaucoup d'informations sur le schéma considéré. Ainsi, on voit que G est un peu trop diffusive :  $u_h(t_0)$  semble être au-dessus de la solution exacte pour  $x \in ]0, 1/2[$ , et en-dessous pour  $x \in ]1/2, 1[$ . Les autres schémas sont moins diffusifs, en effet,  $u_h(t)$  "touche" les lignes  $u=1$  et  $u=0$  pour  $x > 0$  et  $x < 1$  respectivement;
- iii) les "pentes" utilisées par BIP sont plus raides que celle de MUSCL, c'est-à-dire, BIP *diffuse moins que* MUSCL. On peut voir ceci en comparant les  $x$  pour lesquelles  $u_h(t, x) = 0$  ou  $= 1$ . On a toujours si

$$u_{h, \text{MUSCL}}(t, x_M) = 0 = u_{h, \text{BIP}}(t, x_M)$$

alors  $x_M < x_B$  (on obtient une propriété similaire pour  $u_h = 1$ );

- iv) ce "défaut" de BIP s'accroît beaucoup chez LRGP. Ceci est compréhensible car LRGP est un schéma qui diffuse très peu. On pourrait dire aussi, qu'il est un schéma qui *"anti-diffuse"* beaucoup. Ceci se traduit par l'apparition de chocs parasites dans  $u_h$ !
- v) ces chocs parasites apparaissent aussi chez G-1/2, mais la situation n'est pas si tragique comme pour LRGP.



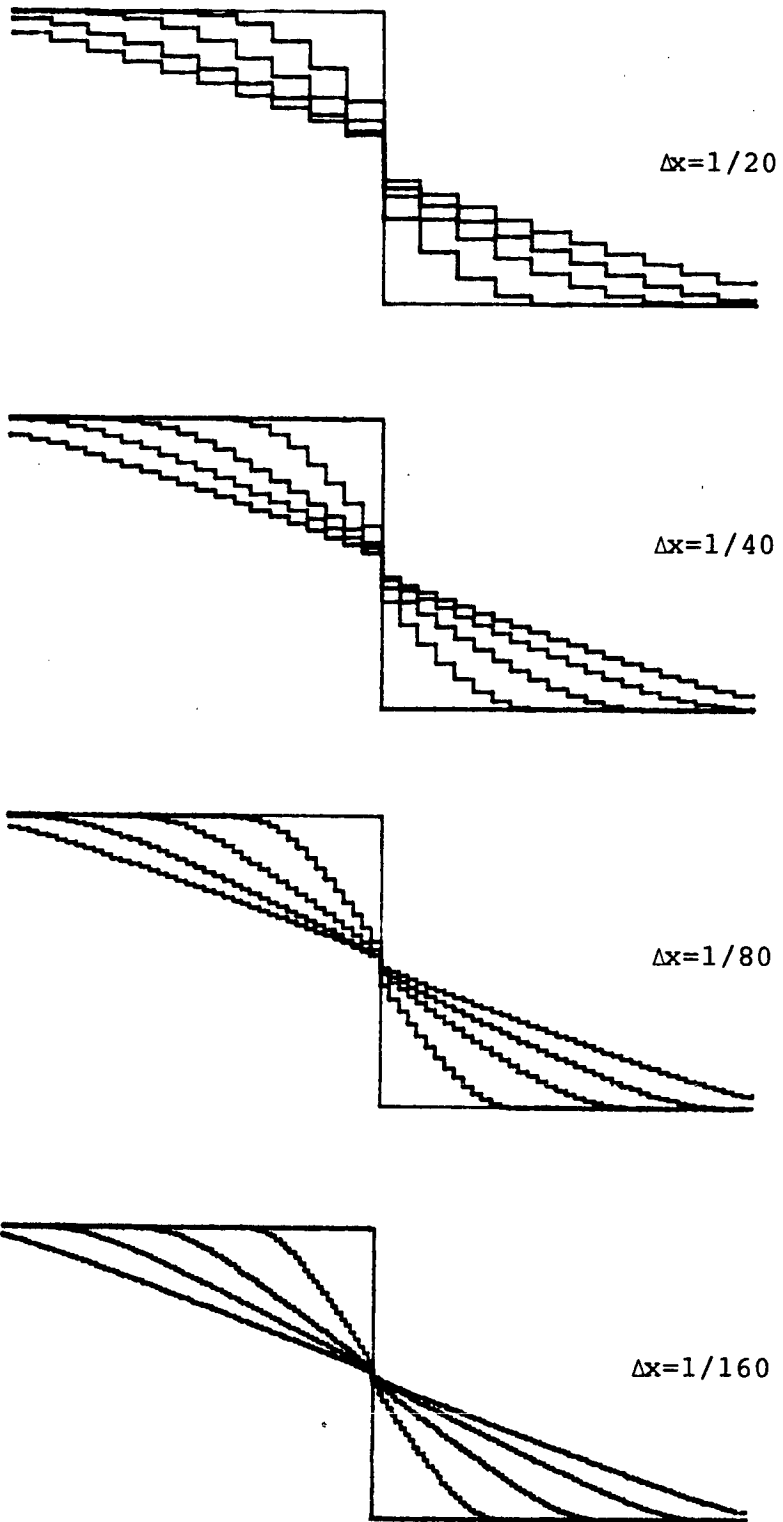


Figure 43 : Cas d'une onde de détente : schéma de Godunov. Raffinement du maillage sous la condition CFL = 1/2.

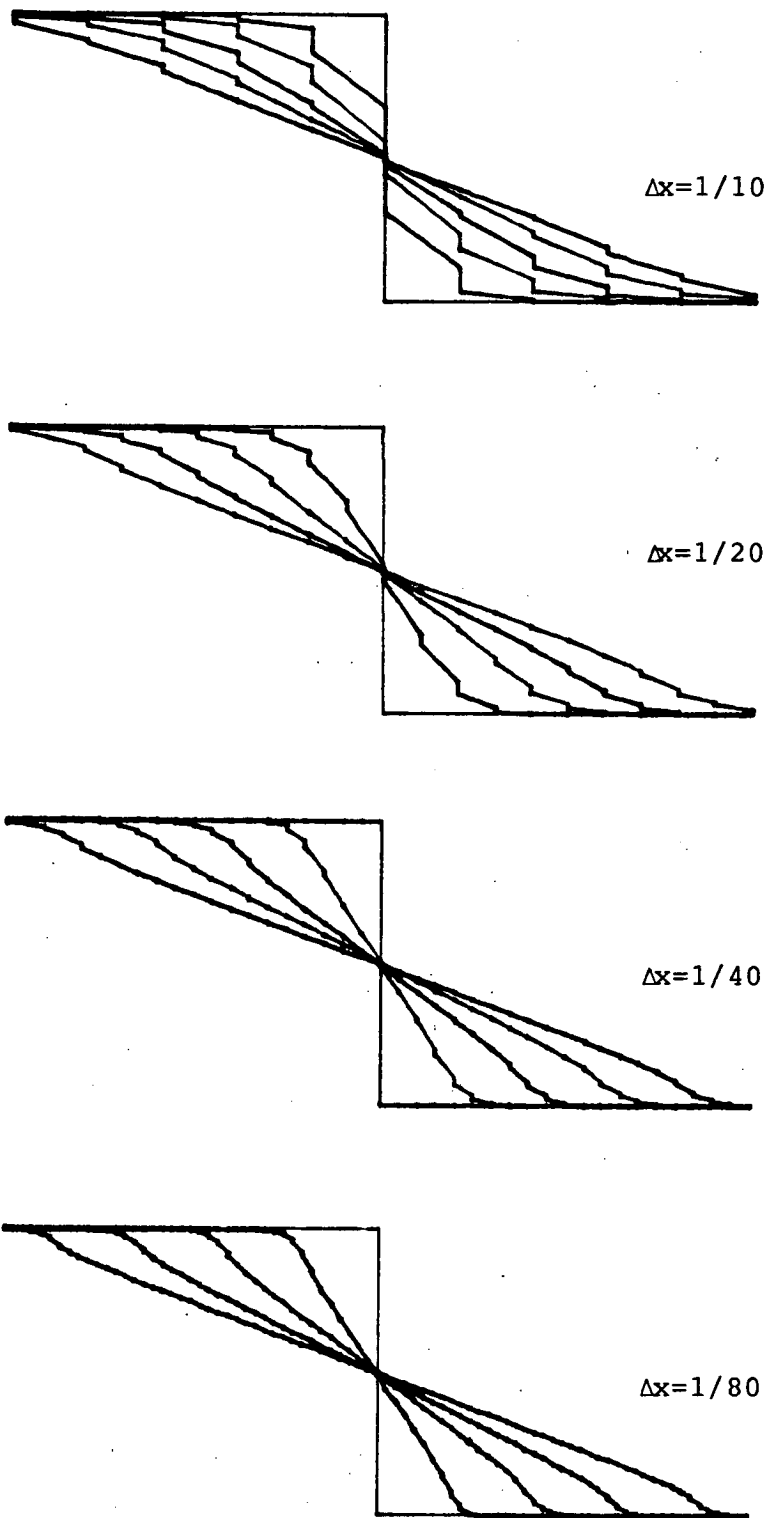


Figure 44 : Cas d'une onde de détente : schéma MUSCL. Raffinement du maillage sous la condition CFL = 1/2.

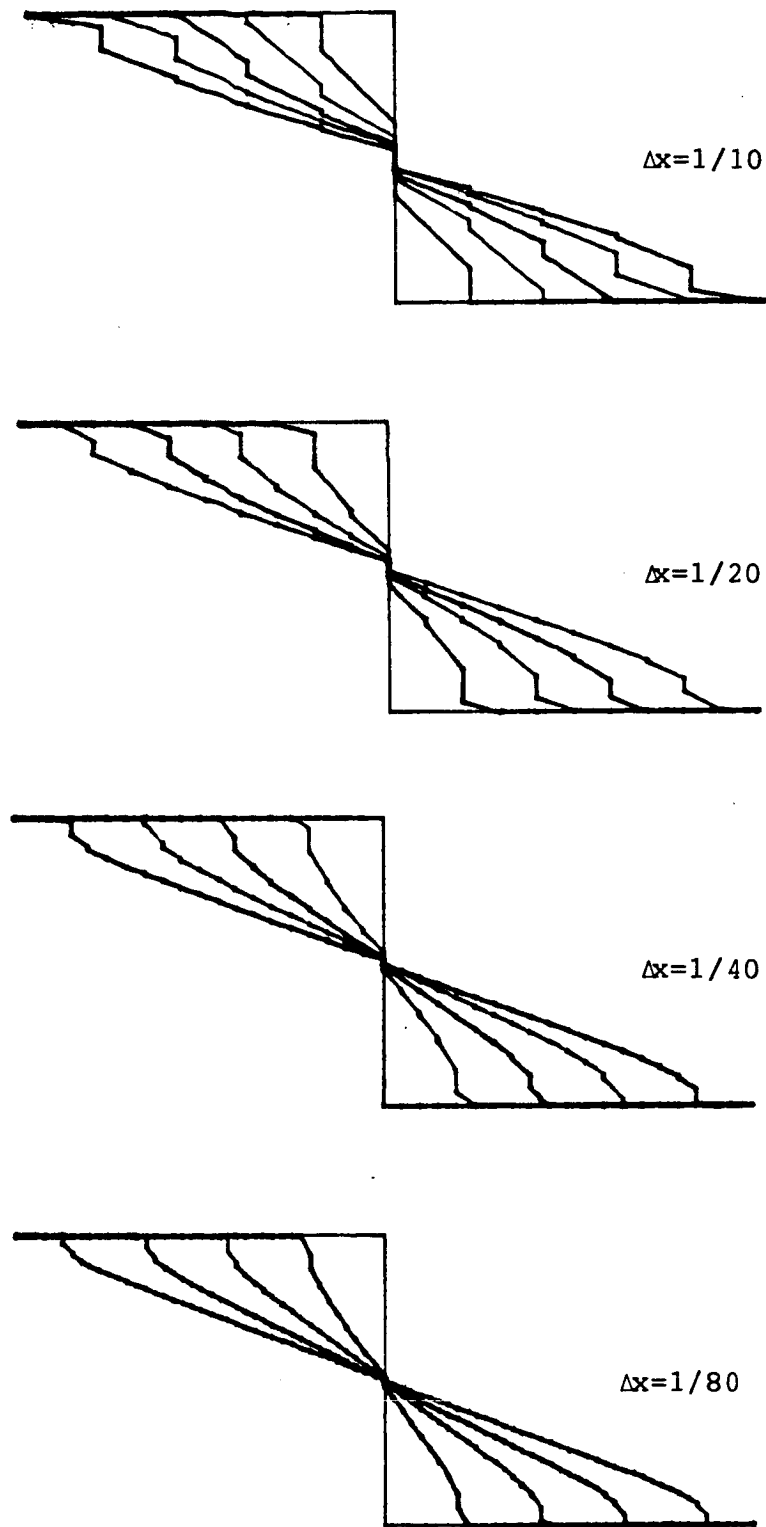


Figure 45 : Cas d'une onde de détente : schéma BIP. Raffinement du maillage sous la condition CFL = 1/2.

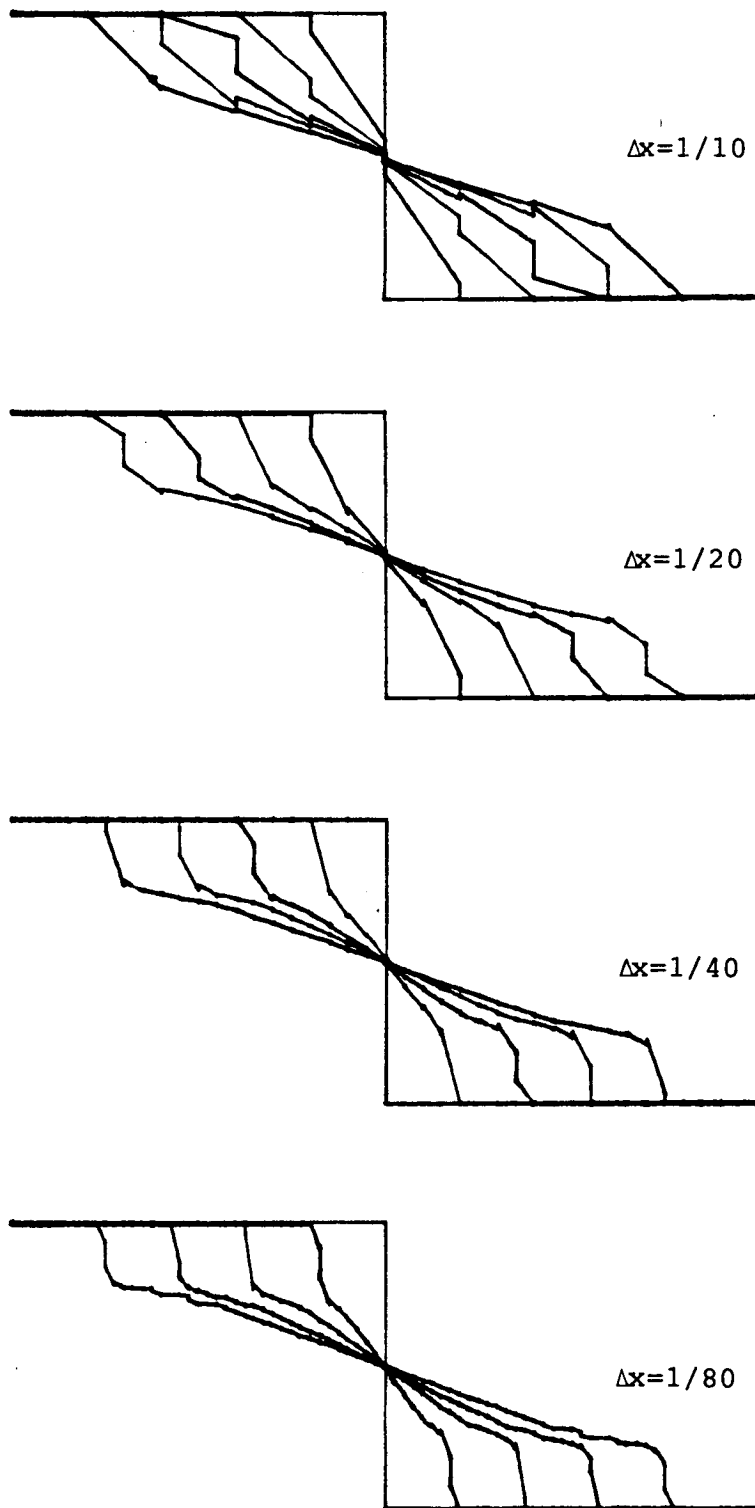


Figure 46 : Cas d'une onde de détente : schéma LRGP. Raffinement du maillage sous la condition CFL = 1/2.

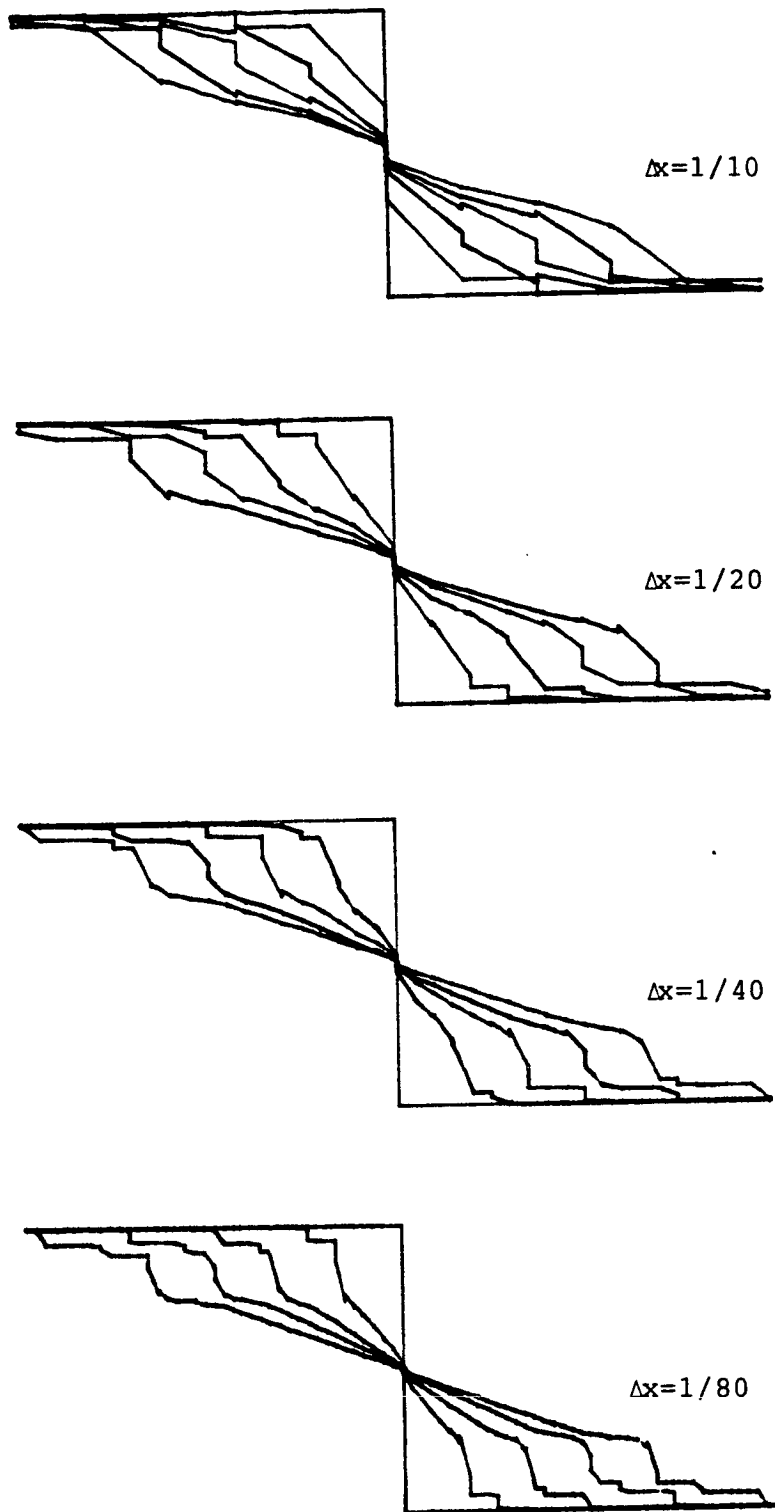


Figure 47 : Cas d'une onde de détente : schéma G-1/2. Raffinement du maillage sous la condition CFL = 1/2.

Voyons comment ceci se manifeste sur les propriétés de convergence de ces schémas. Considérons la figure 48, on y constate que :

- i) MUSCL produit les erreurs les plus petites. Il possède aussi le plus petit ordre de convergence :  $\approx 0.1$  pour la plus petite valeur de  $h$ ;
- ii) cette fois BIP se comporte d'une façon très différente de celle de MUSCL. Il produit, dans le rang de valeur de  $h$  considéré, des erreurs plus grandes que MUSCL, mais par contre, il possède un ordre de convergence nettement supérieur :  $\approx 0.5$ , pour la dernière valeur de  $h$ ;
- iii) la comportement de LRGP et de G-1/2 est semblable. Les erreurs produites par G-1/2 sont plus petites que celles produites par LRGP, mais elles sont plus grandes que celles de BIP. De même, l'ordre de convergence de G-1/2 est plus grand que celui de LRGP et plus petit que celui de BIP. *Ceci est dû à l'apparition des chocs parasites;*
- iv) finalement, le schéma G possède l'ordre de convergence le plus grand :  $\approx 0.7$  ! Il donne des erreurs plus petites que LRGP et G-1/2, et que BIP pour la plus petite valeur de  $h$ .

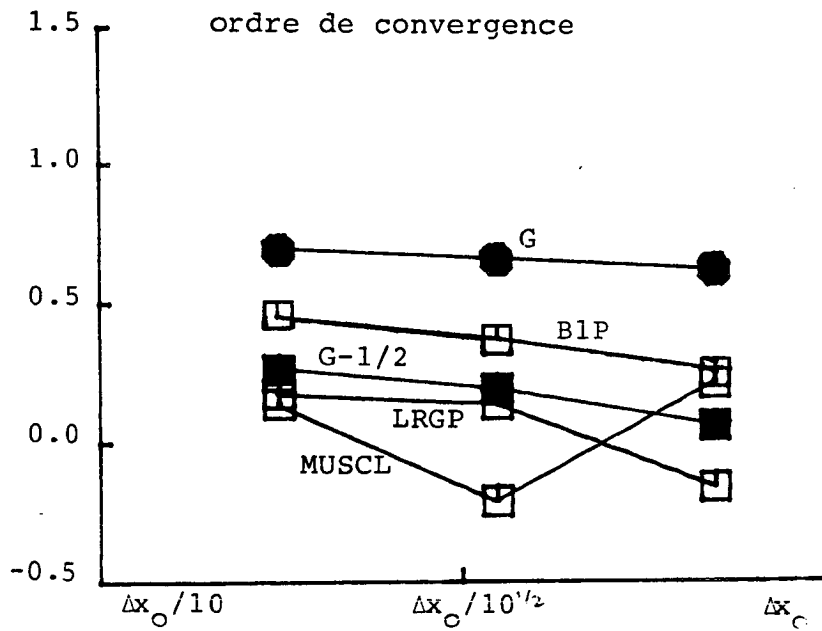
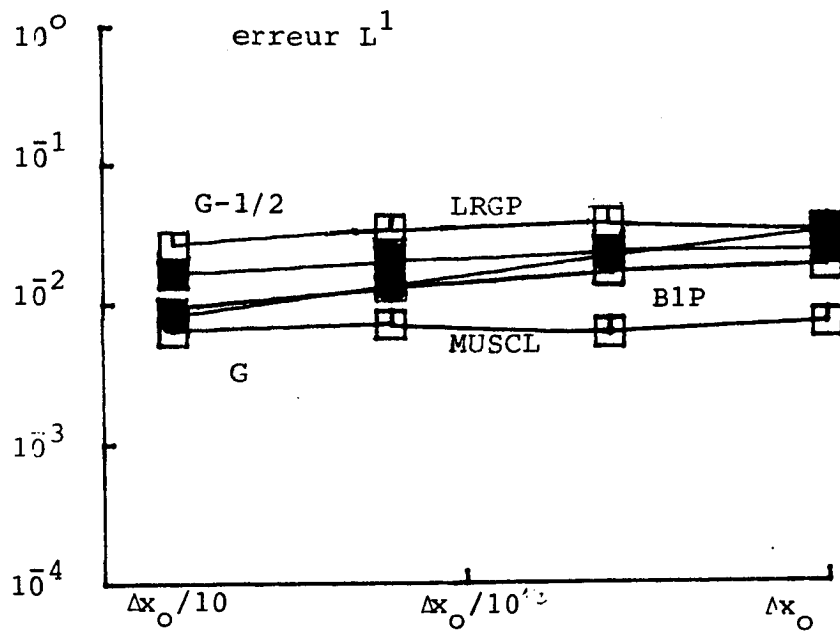


Figure 48 : Cas d'une onde de détente : comparaison numérique des propriétés de convergence des différents schémas, sous la condition CFL = 1/2. (On a pris  $\Delta x_0 = 1/10$ , sauf pour le schéma de Godunov pour lequel on a pris  $\Delta x_0 = 1/20$ ).

On a aussi rencontré deux phénomènes intéressants : d'abord, le fait que MUSCL produit des erreurs très petites, mais qu'il *ne semble pas converger trop vite vers la solution exacte*; ensuite que LRGP et G-1/2 ont un comportement très mauvais dû *aux chocs parasites qu'ils produisent!*

ii) Le cas CFL = 1/4

Sur les figures 49 à 53, on montre les solutions données par les schémas G, MUSCL, BIP, LRGP et G-1/2, dans cet ordre. On ne voit pas tellement de différence par rapport au cas précédent, mais il faut quand même remarquer que la diminution de CFL semble avoir amélioré les solutions au moins pour LRGP.

Sur la figure 54, on peut mieux voir l'effet de cette diminution,

- i) MUSCL produit maintenant, sauf pour la valeur de  $h$  la plus grande, les erreurs les plus petites, et en même temps il possède l'ordre de convergence le plus grand;
- ii) BIP possède, comme dans le cas précédent, un ordre de convergence  $\approx 0.5$ , et produit, sauf pour la valeur de  $h$  la plus grande, les erreurs les plus petites, après MUSCL;
- iii) G-1/2 et LRGP n'ont pas subi une influence positive due à la diminution du CFL. Ils se comportent, grosso modo, de la même façon que dans le cas précédent;
- iv) le schéma G produit les erreurs les plus grandes (sauf pour la plus petite valeur de  $h$ ), mais par contre, il possède, après MUSCL, le meilleur ordre de convergence  $\approx 0.7$ .



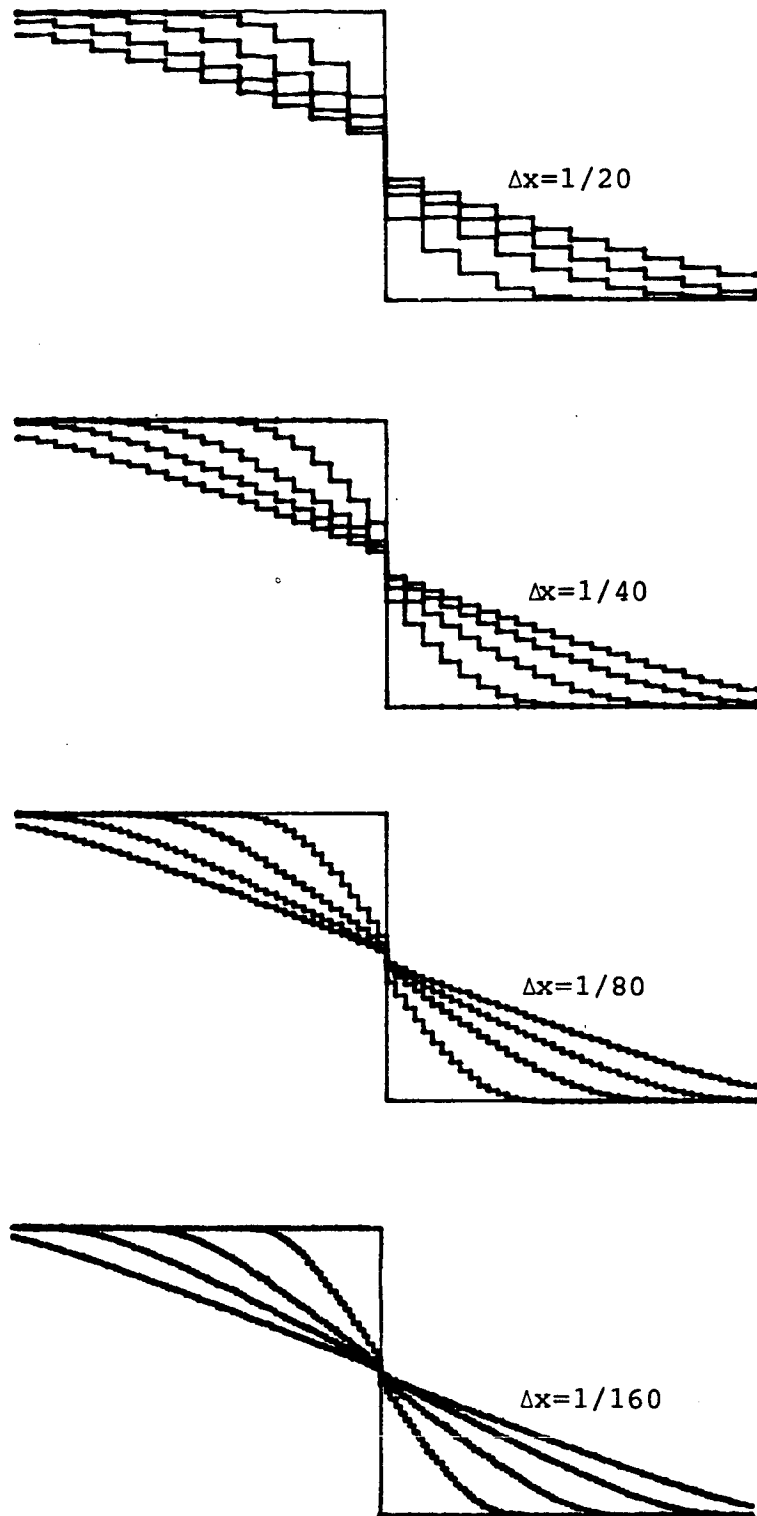


Figure 49 : Cas d'une onde de détente : schéma de Godunov. Raffinement du maillage sous la condition CFL = 1/4.

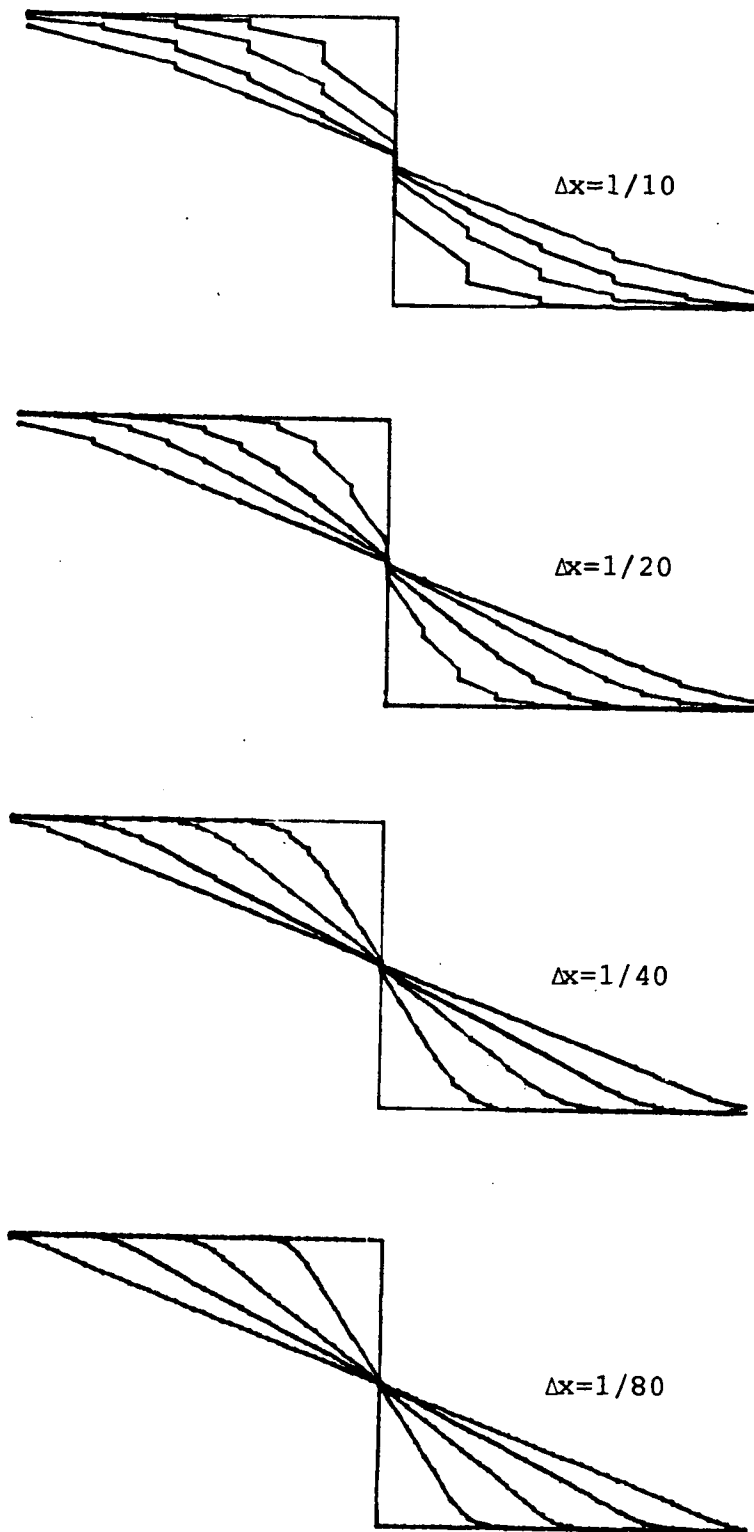


Figure 50 : Cas d'une onde de détente : schéma MUSCL. Raffinement du maillage sous la condition CFL = 1/4.

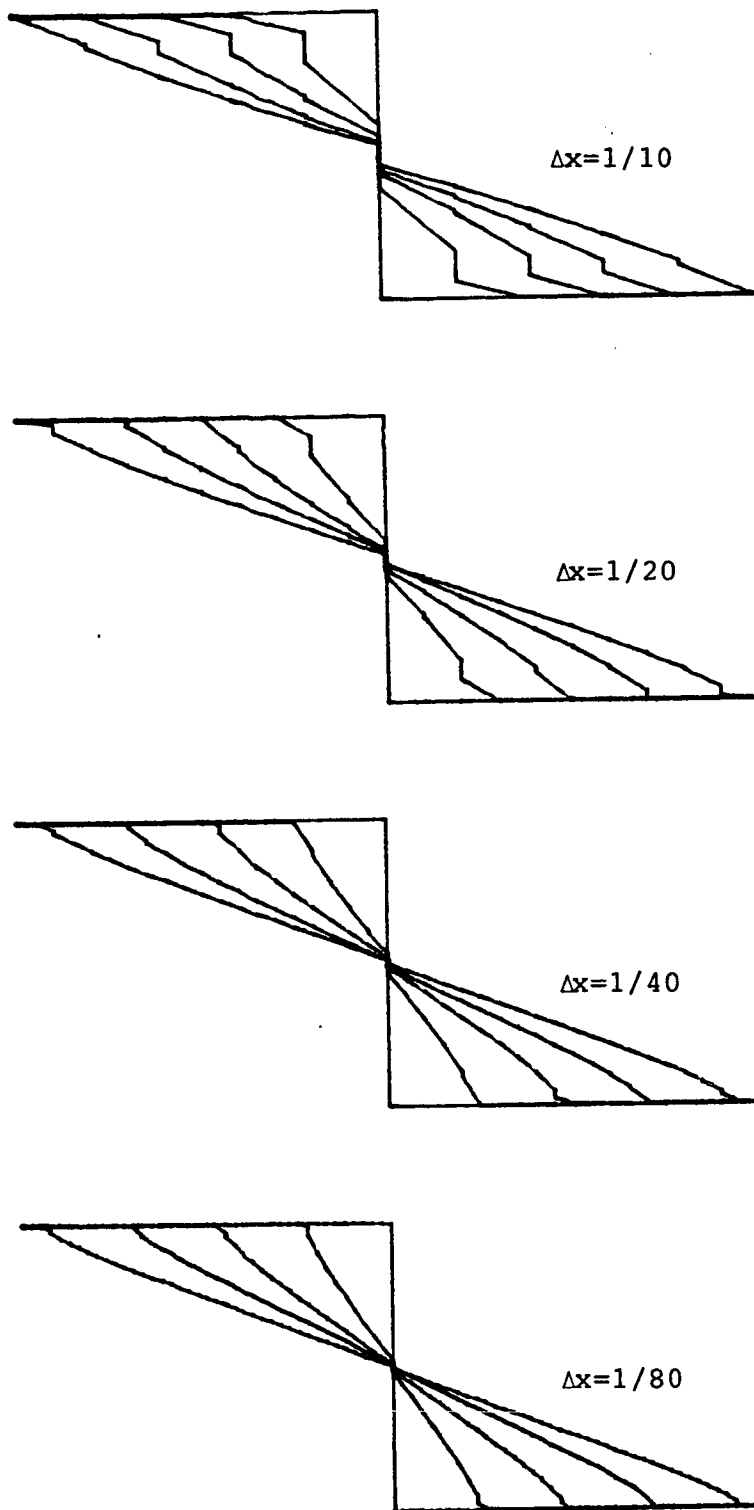


Figure 51 : Cas d'une onde de détente : schéma BIP. Raffinement du maillage sous la condition CFL = 1/4.

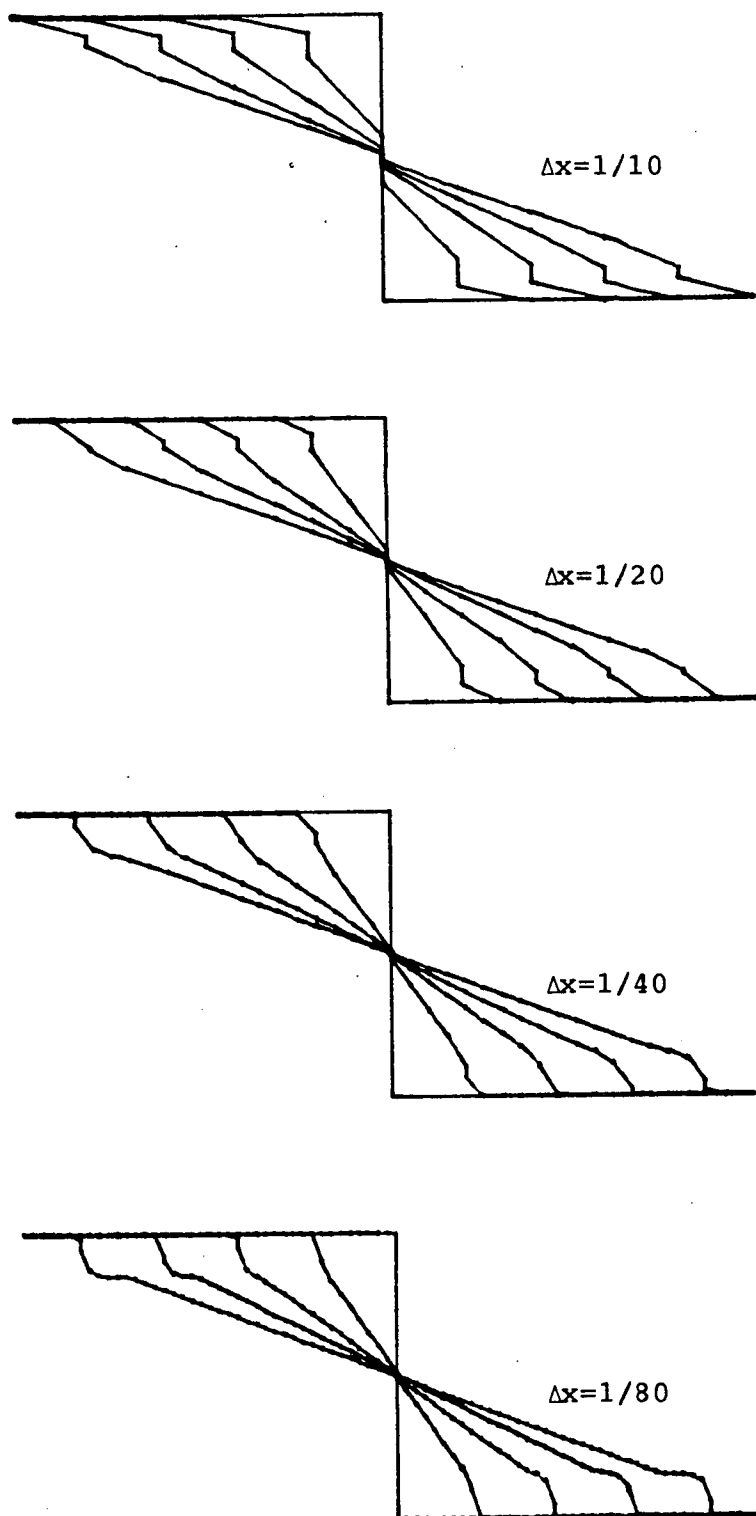


Figure 52 : Cas d'une onde de détente : schéma LRGP. Raffinement du maillage sous la condition CFL = 1/4.

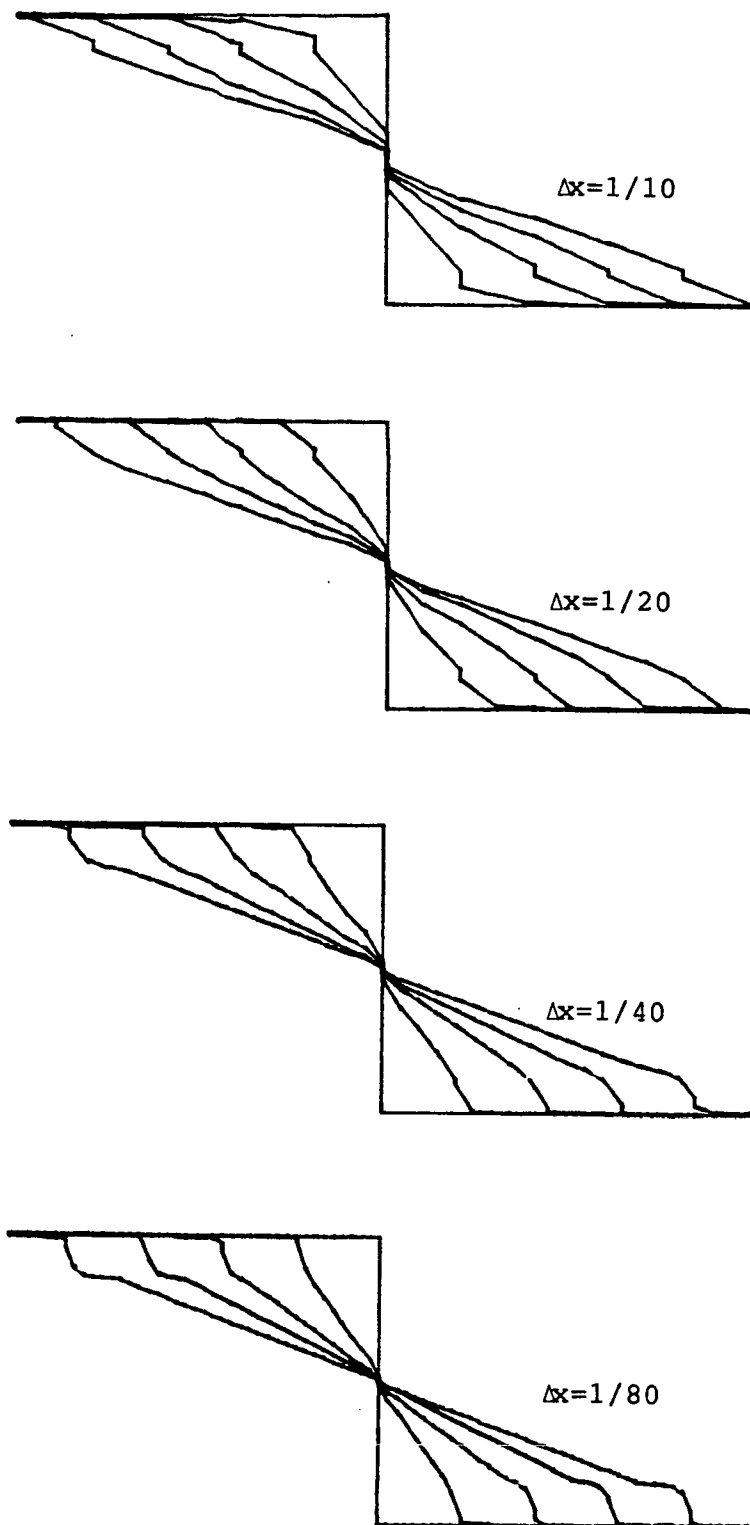


Figure 53 : Cas d'une onde de détente : schéma G-1/2. Raffinement du maillage sous la condition CFL = 1/4.

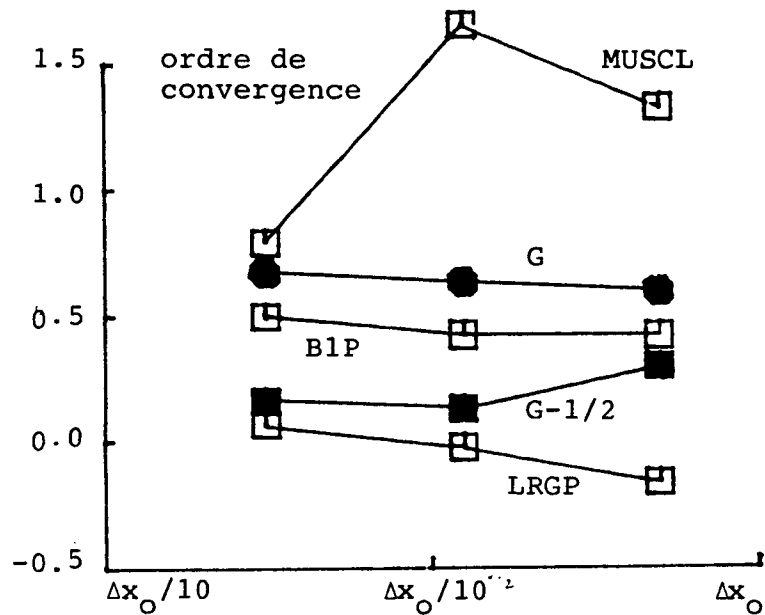
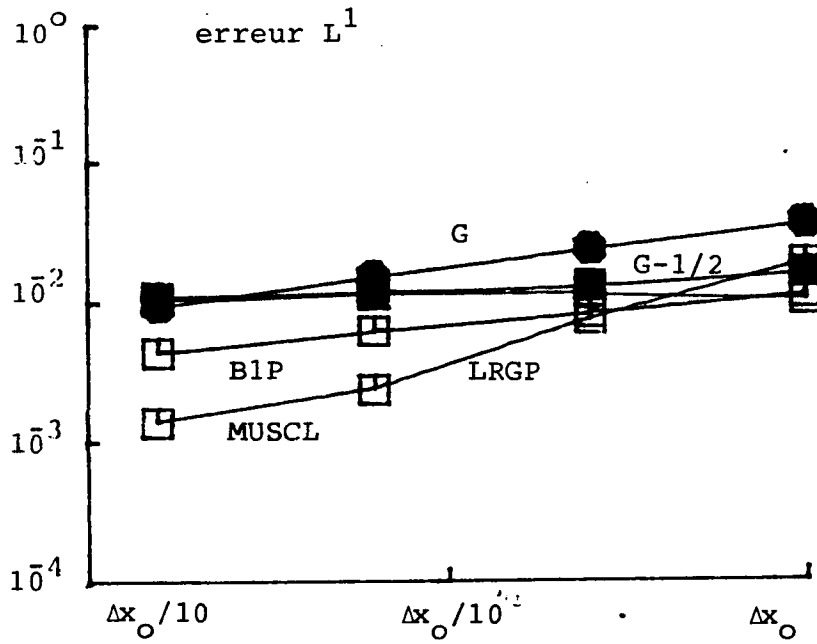


Figure 54 : Cas d'une onde de détente : comparaison numérique des propriétés de convergence des différents schémas sous la condition CFL = 1/4. (On a pris  $\Delta x_0 = 1/10$ , sauf pour le schéma de Godunov pour lequel on a pris  $\Delta x_0 = 1/20$ ).

Ainsi, la diminution de CFL se manifeste par une nette amélioration de MUSCL et une presque négligeable amélioration de BIP, G-1/2 et LRGP (*les chocs parasites ne disparaissent pas*). L'influence sur G est aussi négligeable mais il faut remarquer que maintenant G produit les erreurs les plus grandes.

iii) Le cas CFL = 1/8

Sur les figures 55 à 59, on montre les solutions données par G, MUSCL, BIP; LRGP et G-1/2 dans cet ordre. Il est difficile de voir, sur les figures, l'influence de cette nouvelle diminution de CFL.

Considérons donc la figure 60. On y constate que :

- i) comme dans le cas précédent, MUSCL possède l'ordre de convergence le plus grand :  $> 1$ . Mais cette fois-ci, il ne produit l'erreur la plus petite que pour la valeur de  $h$  la plus petite;
- ii) BIP possède, après MUSCL, l'ordre de convergence le plus élevé ( $\approx 0.6$  pour la dernière valeur de  $h$ ) et produit les erreurs les plus petites (après MUSCL pour la valeur de  $h$  la plus petite, et après LRGP pour les autres valeurs);
- iii) maintenant G-1/2 et LRGP ont subi une influence positive due à la diminution de CFL : pour G-1/2 l'ordre de convergence a augmenté beaucoup (de  $\approx 0.2$  à  $\approx 0.5$ ) (pour LRGP ceci n'est pas tellement le cas), et LRGP produit les plus petites erreurs, sauf pour la valeur la plus petite de  $h$  (on retrouve ici une propriété de LRGP déjà trouvée dans le cas d'un choc!);
- iv) finalement, le comportement de G n'a pratiquement pas changé.

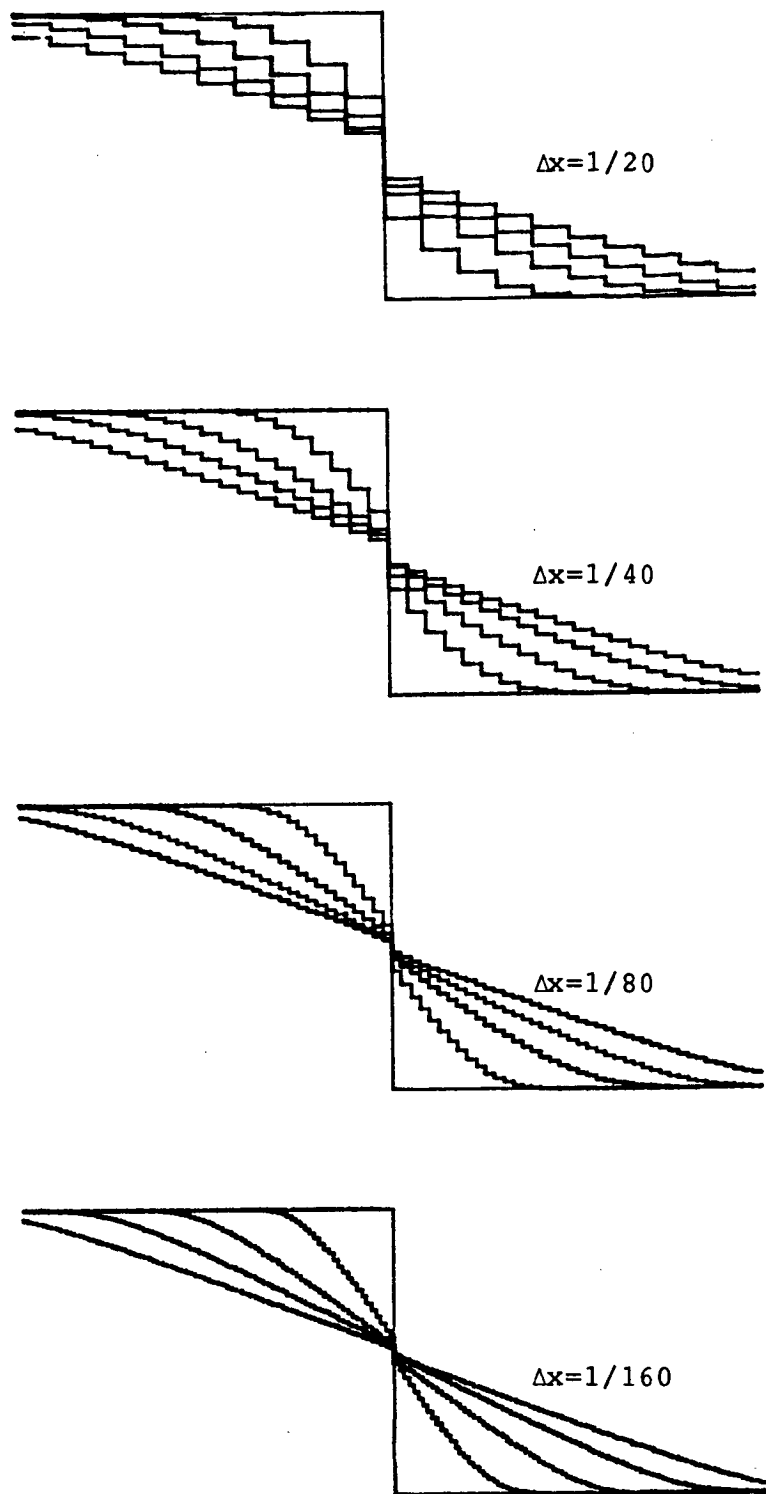


Figure 55 : Cas d'une onde de détente : schéma de Godunov. Raffinement du maillage sous la condition CFL = 1/8.



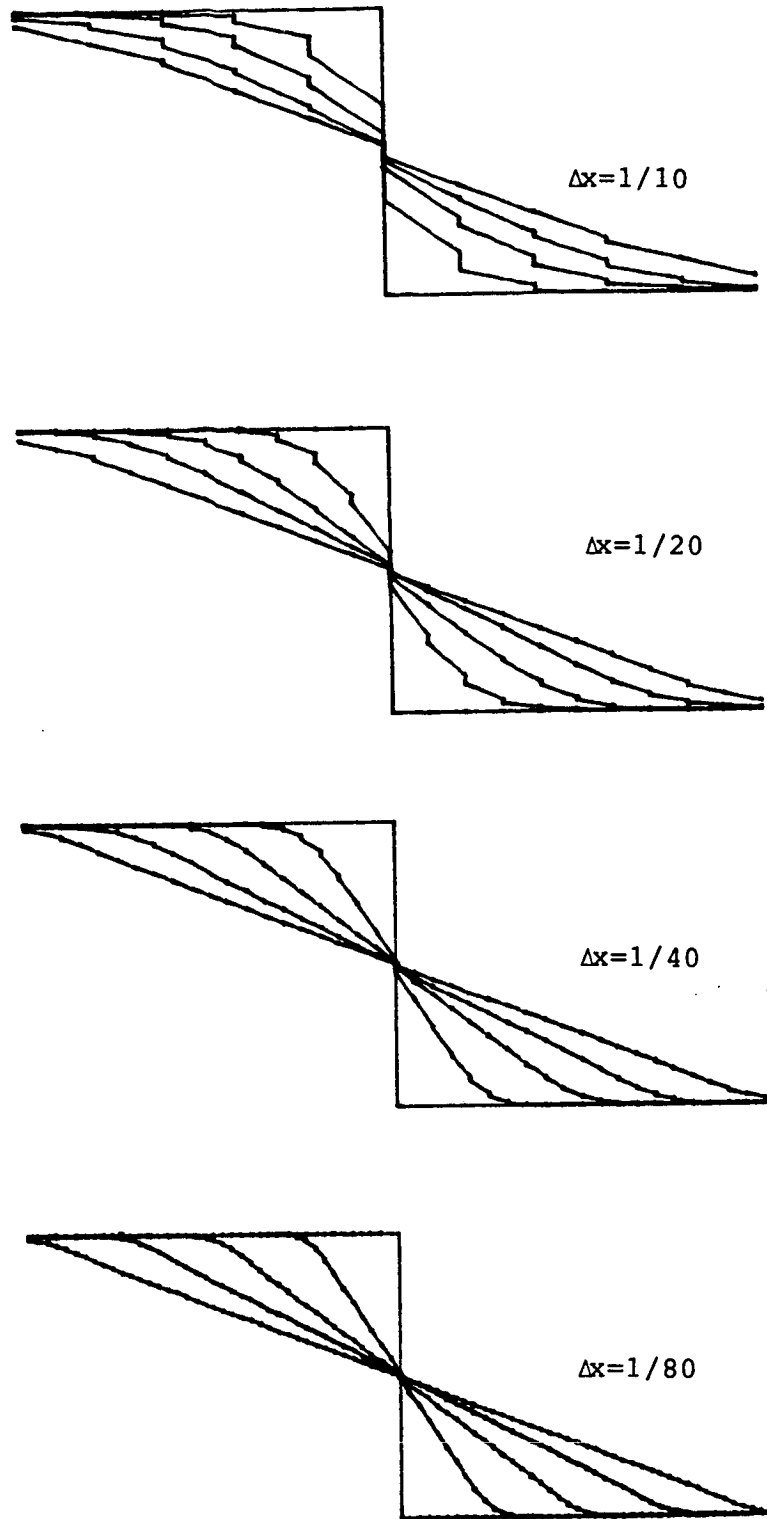


Figure 56 : Cas d'une onde de détente : schéma MUSCL. Raffinement du maillage sous la condition CFL = 1/8

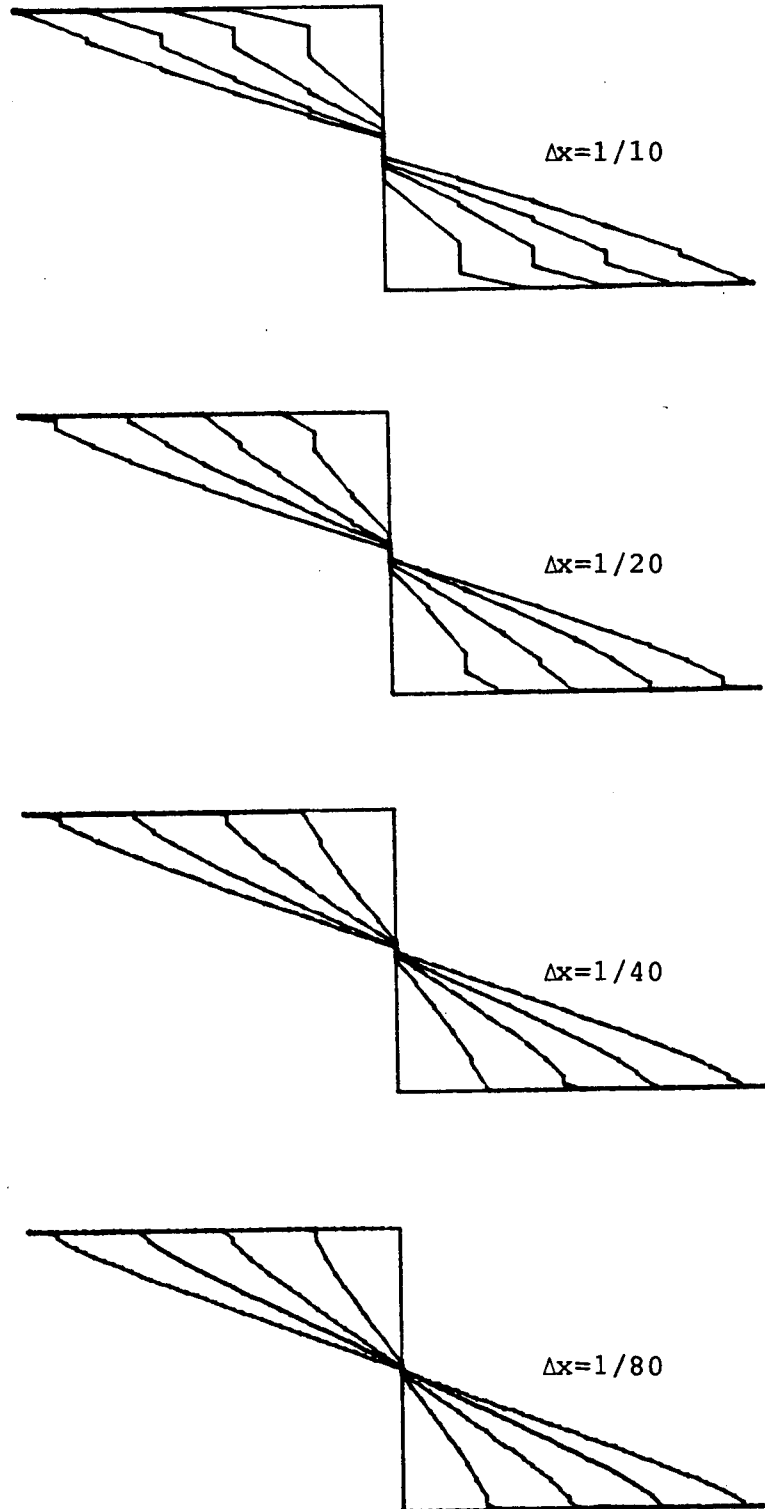


Figure 57 : Cas d'une onde de détente : schéma B1P. Raffinement du maillage sous la condition CFL = 1/8.

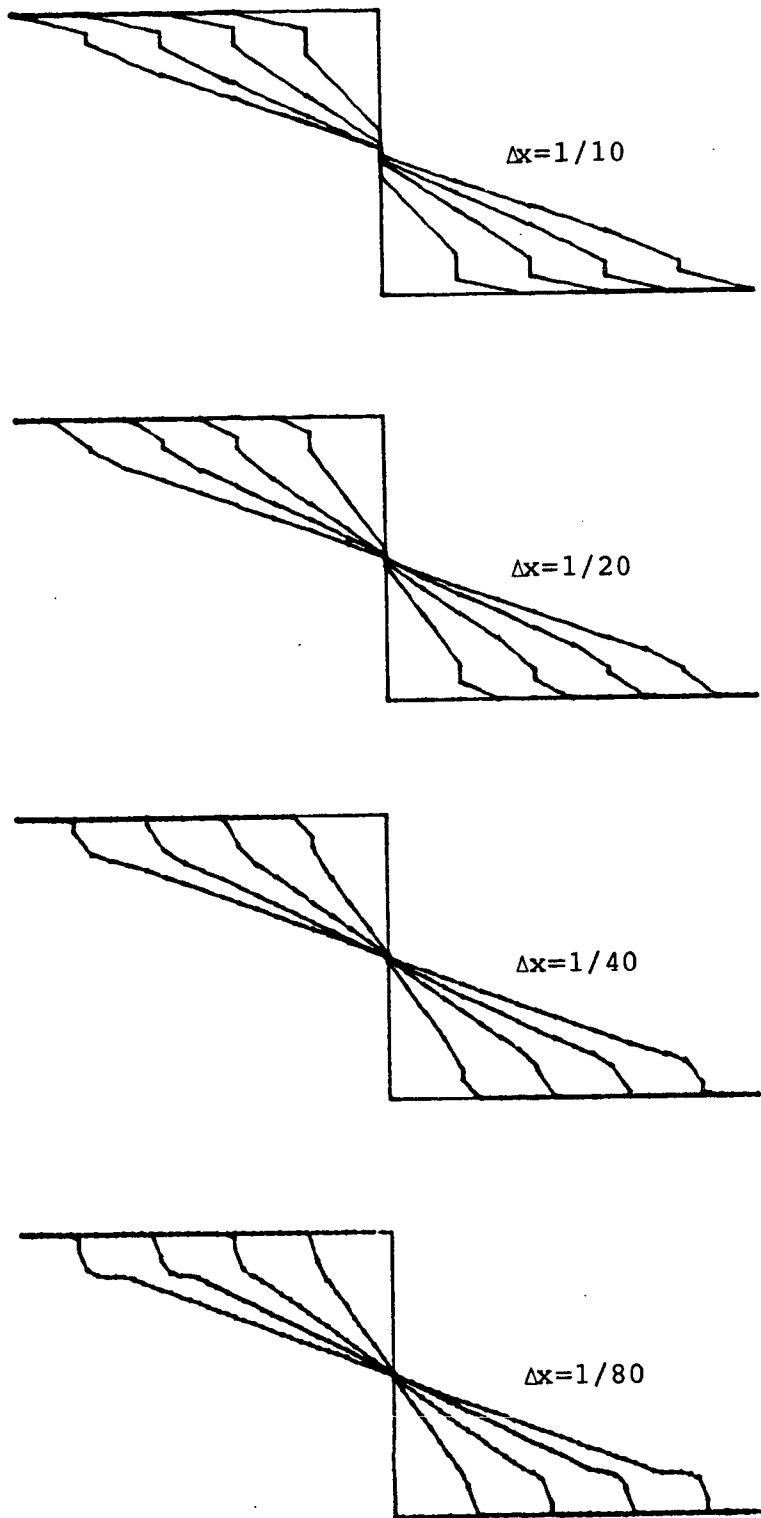


Figure 58 : Cas d'une onde de détente : schéma LRGP. Raffinement du maillage sous la condition CFL = 1/8.

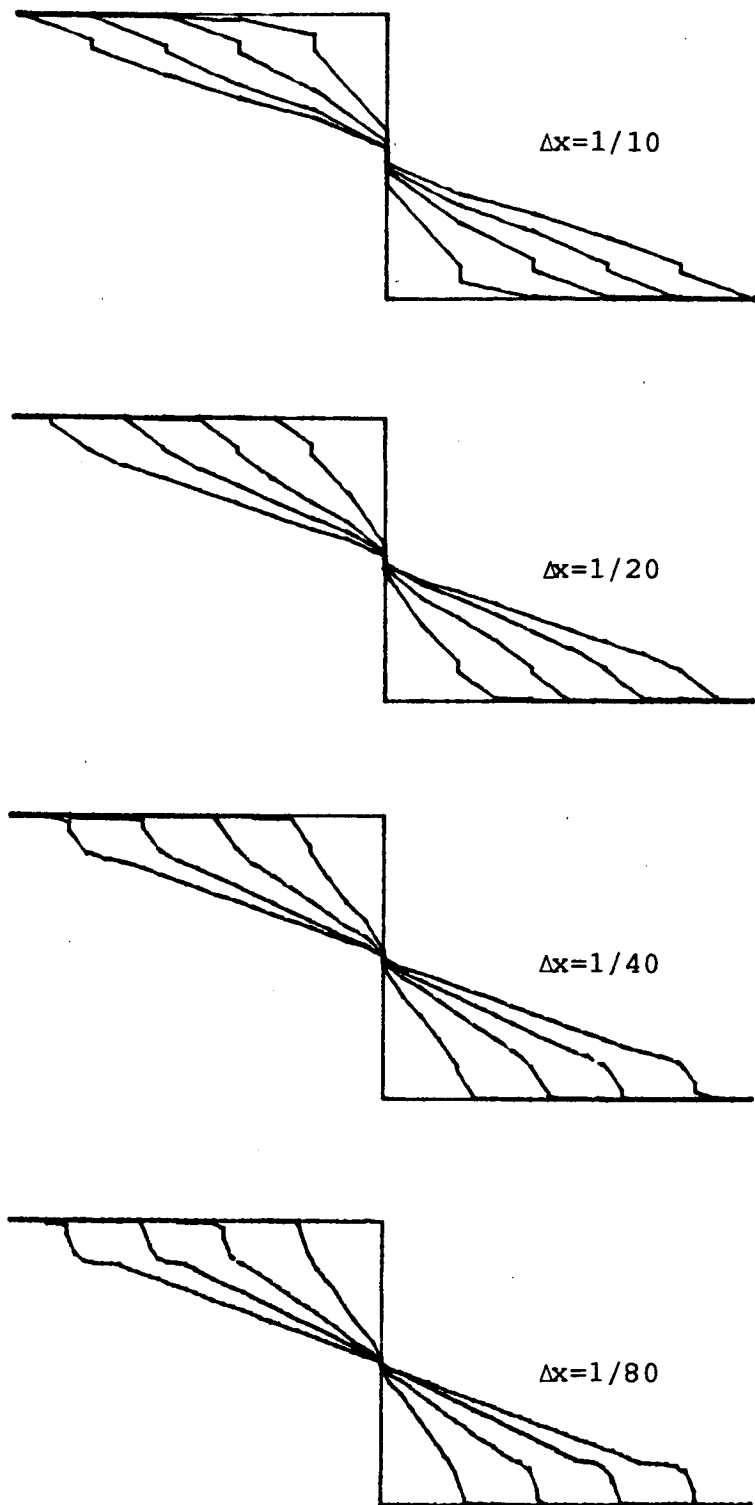


Figure 59 : Cas d'une onde de détente : schéma G-1/2. Raffinement du maillage sous la condition CFL = 1/8.

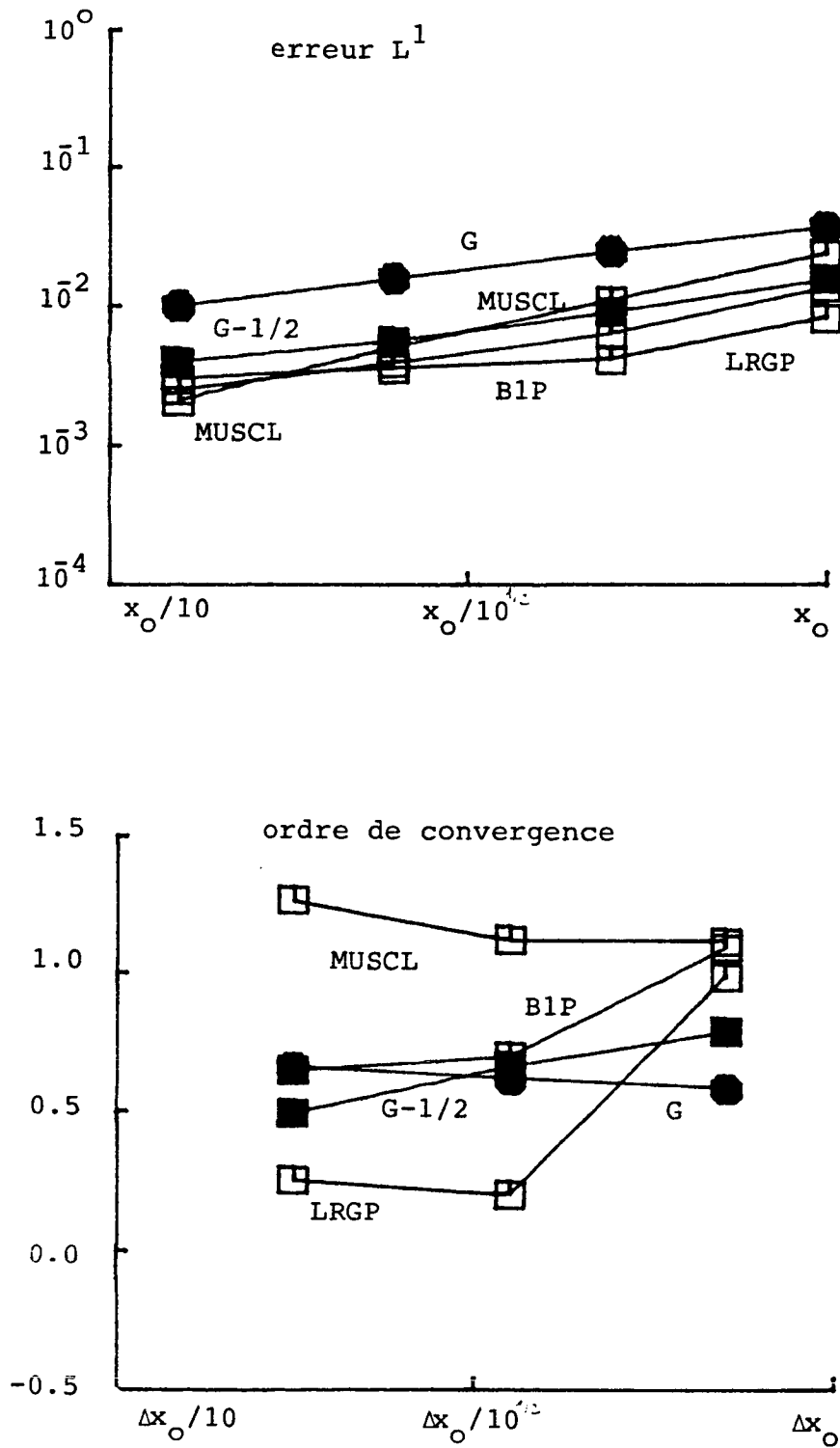


Figure 60 : Cas d'une onde de détente : comparaison numérique des propriétés de convergence des différents schémas sous la condition CFL = 1/8 (On a pris  $\Delta x_0 = 1/10$ , sauf pour le schéma de Godunov pour lequel on a pris  $\Delta x_0 = 1/20$ ).

iv) Quelques commentaires

Dans ces expériences numériques, on voit que :

- i) comme dans les deux cas précédents, tous les schémas semblent être positifs et monotones, la monotonie étant au niveau des moyennes pour les schémas LRGP et G-1/2;
- ii) on ne peut plus "ordonner" les schémas comme on l'a fait pour les deux cas précédents. En effet, pour CFL = 1/2, MUSCL produit les erreurs les plus petites mais en même temps, il possède l'ordre de convergence le plus petit aussi; pour CFL = 1/8, LRGP produit les erreurs les plus petites seulement pour h "assez grand";
- iii) dans ce cas, les schémas LRGP et aussi G-1/2 ont mis en évidence qu'ils sont des schémas "anti-diffusifs" : ils font apparaître des chocs parasites qui ne devraient pas exister. Cette tendance est fortement accentuée lorsque  $h \rightarrow 0$  où lorsque CFL croît. Il est très intéressant de voir que le schéma LRG ne possède pas du tout cette propriété. Dans le but de conserver la positivité, le schéma LRGP met à zéro les pentes de  $u_h$  dans les éléments  $]x_i, x_{i+1/2}[$  où  $u_{i+1/2} = 1$  ou  $= 0$ . Ceci permet que les pentes des éléments voisins soient trop raides, ce qui génère petit à petit les chocs parasites. Il est très intéressant aussi de se demander pourquoi les chocs ne sont pas apparus dans le cas précédent. Une réponse est la suivante : à la différence de ce cas, la "vitesse de transport" de  $u = 1$  était inférieure à celle de  $u = 1-\epsilon$ , pour  $\epsilon$  petit! On peut donc conclure les chocs parasites de LRGP et G-1/2 vont apparaître à chaque fois que la vitesse de transport de  $u = u_*$  est plus grande que celle de  $u = u_* - \epsilon$  (resp.  $= u_* + \epsilon$ ) où  $u_h \in [a, u_*]$  (resp.  $[u_*, a]$ ), pour  $\epsilon$  petit. Pour remédier à cet inconvénient il faut augmenter de la diffusion! (en faisant h grand, ou CFL petit!);

- iv) à la différence des cas précédents, MUSCL et BIP ont ici un comportement assez différent. Pour  $CFL = 1/2$ , MUSCL produit les plus petites erreurs et possède le plus petit ordre de convergence. Avec la diminution de CFL, l'ordre de convergence s'améliore considérablement ( $\approx 1$  pour  $CFL > 1/2$ ) mais par contre, les erreurs qu'il produit deviennent de plus en plus grandes en relation à celles produites par BIP (pour  $CFL = 1/8$  par exemple BIP produit des erreurs plus petites que MUSCL). BIP possède par contre un comportement qu'on pourrait qualifier de plus stable par rapport aux variations de CFL : son ordre de convergence augmente lentement avec la diminution de cette quantité, en même temps que les erreurs qu'il produit diminuent d'une façon "monotone". On peut dire que BIP est supérieur à MUSCL pour CFL petit et  $h$  grand, et que MUSCL est supérieur à BIP pour  $h$  petit!;
- v) comme dans les deux cas précédents, l'influence de la diminution de CFL sur les solutions données par G est négligeable. Il est aussi intéressant de noter qu'à la différence des cas précédents, l'ordre de convergence de G pour  $CFL > 1/8$  ( $\approx 0.6$ ) est supérieur à l'ordre de convergence de tous les schémas considérés (sauf pour  $CFL = 1/4$  où seul MUSCL possède un ordre de convergence plus grand!).

#### 4.5.3 - Résumé

La conclusion principale de ces expériences est que LRGP donne les erreurs les plus petites pourvu que  $h$  soit suffisamment grand et CFL assez petit. Pour  $h$  petit (et  $CFL > 1/2$ ), le meilleur choix est MUSCL. Après ce schéma, le meilleur choix est BIP.

On a vu apparaître un phénomène intéressant : les solutions données par LRGP et G-1/2 présentent des chocs parasites. Les chocs deviennent plus raides au fur et à mesure que CFL augmente et  $h \rightarrow 0$ . La vitesse de transport de  $u$  pour  $u \approx 0,1$  est ralentie par la projection utilisée pour conserver la positivité des schémas.

Dans [ 1 ], on a démontré que  $G-1/2$  converge vers la solution entropique sous la condition  $CFL \leq 1/2$ . Ces expériences semblent contredire ce fait, mais ceci n'est pas le cas. En effet, si on voit la définition de  $G-1/2$ , paragraphe 4.2.1, on remarque tout de suite qu'il existe deux paramètres qui contrôlent les termes d'anti-diffusion, à savoir :  $c_0$  et  $\alpha$ . Pour avoir la convergence de  $G-1/2$ , une condition suffisante est que  $\alpha > 0$ . Pour les deux premières expériences, on a vu que cette condition n'était pas nécessaire, mais ceci n'est plus le cas dans le cas que l'on considère maintenant! Au paragraphe suivant, on va donc évaluer l'influence de  $\alpha$  sur la solution donnée par  $G-1/2$ . Mais avant, on va résumer les différentes propriétés des schémas observés dans ce cas.



	G	MUSCL	BlP	LRGP	G-1/2	LRG
positif	oui	oui	oui	oui	oui	non
monotone	oui	oui	oui	au niveau des moyennes	au niveau des moyennes	non
conditions de stabilité	$CFL \leq 1$	$CFL \leq 1/2$	$CFL \leq 1/2$	$CFL \leq 1/2$	$CFL \leq 1/2$	$CFL \leq ch^{1/2}$
particularités	-	-	-	apparition des chocs parasites		oscillations pour $u \approx 1,0$
erreurs		produit les erreurs les plus petites pour $CFL=1/2$ et pour h "petit" si $CFL > 1/2$	produit des erreurs plus petites que MUSCL pour $CFL > 1/2$ et h grand	produit les erreurs les plus petites seulement pour h grand et CFL petit	produit les erreurs plus grandes que LRGP	
ordre de convergence	$\approx 0.65$	$CFL=1/2$ 0.10 $CFL=1/4$ 0.80 $CFL=1/8$ 1.25	$\approx 0.50$ $\approx 0.50$ $\approx 0.60$	$\approx .20$ $\approx .10$ $\approx .25$	$\approx .30$ $\approx .20$ $\approx .50$	$\approx 1$

Schémas "LRG-modifiés"

Tableau 4 : Propriétés théoriques et "expérimentales" des schémas pour le cas (c).

#### 4.6 - L'influence du paramètre d'anti-diffusion $\alpha$ sur la solution donnée par le schéma G-1/2

##### 4.6.1 - Introduction

A la fin du paragraphe précédent, on avait remarqué que dans [1] on a prouvé que si  $\alpha > 0$ , voir la définition de G-1/2 au paragraphe 4.2.1, alors G-1/2 est un schéma dont la solution converge sous la solution entropique. Or, dans les trois expériences considérées, on a pris  $\alpha = 0$ . Dans les deux premières, on voit que G-1/2 est un schéma convergent, mais ceci n'est pas évident pour la troisième. Ceci indique qu'il est raisonnable de penser que, pour ce cas, la condition  $\alpha > 0$  est aussi *nécessaire* pour la convergence de G-1/2.

Dans ce paragraphe, on va donc étudier l'influence de  $\alpha$  sur la solution donnée par G-1/2. Et pour faire ceci, on va faire les expériences numériques qu'on décrit à la suite :

(I) pour chacun des trois cas considérés, on va considérer les cas  
CFL = 1/2 et CFL = 1/4.

(II) pour CFL fixé, on va prendre  
 $\alpha = 0, 1/4, 1/2, 3/4, 1$ .

(III) pour CFL et  $\alpha$  fixés, on prendra  
 $h = 1/10, 1/20, 1/40, 1/80$ .

On va procéder comme dans les trois paragraphes précédents. On prendra  $c_0(1/10)^\alpha = 1$  (voir la définition de  $c_0$  et  $\alpha$  dans (G-1/2) au paragraphe 4.2.1).

Le premier cas sera traité dans 4.6.2, le deuxième dans 4.6.3 et le dernier dans 4.6.4. Au paragraphe 4.6.5, on trouve les conclusions de cette étude.

#### 4.6.2 - Le cas linéaire : une discontinuité de contact

Sur les figures 61 à 65 (resp. 67 à 71), on montre les solutions données pour  $G=1/2$  avec  $CFL = 1/2$  (resp.  $CFL = 1/4$ ) pour  $\alpha = 0, 1/4, 1/2, 3/4$  et 1 dans cet ordre.

On y constate que, au fur et à mesure que  $\alpha$  augmente, la diffusion du schéma augmente aussi, comme on l'attendait. Ceci fait que les erreurs augmentent aussi. Par conséquent, dans ce cas le meilleur choix de  $\alpha$  est  $\alpha = 0$ .

Ceci est mis en évidence sur la figure 66 (resp. 72) sur laquelle on voit que c'est pour  $\alpha = 0$  que  $G=1/2$  produit les plus petites erreurs et possède le plus grand ordre de convergence.

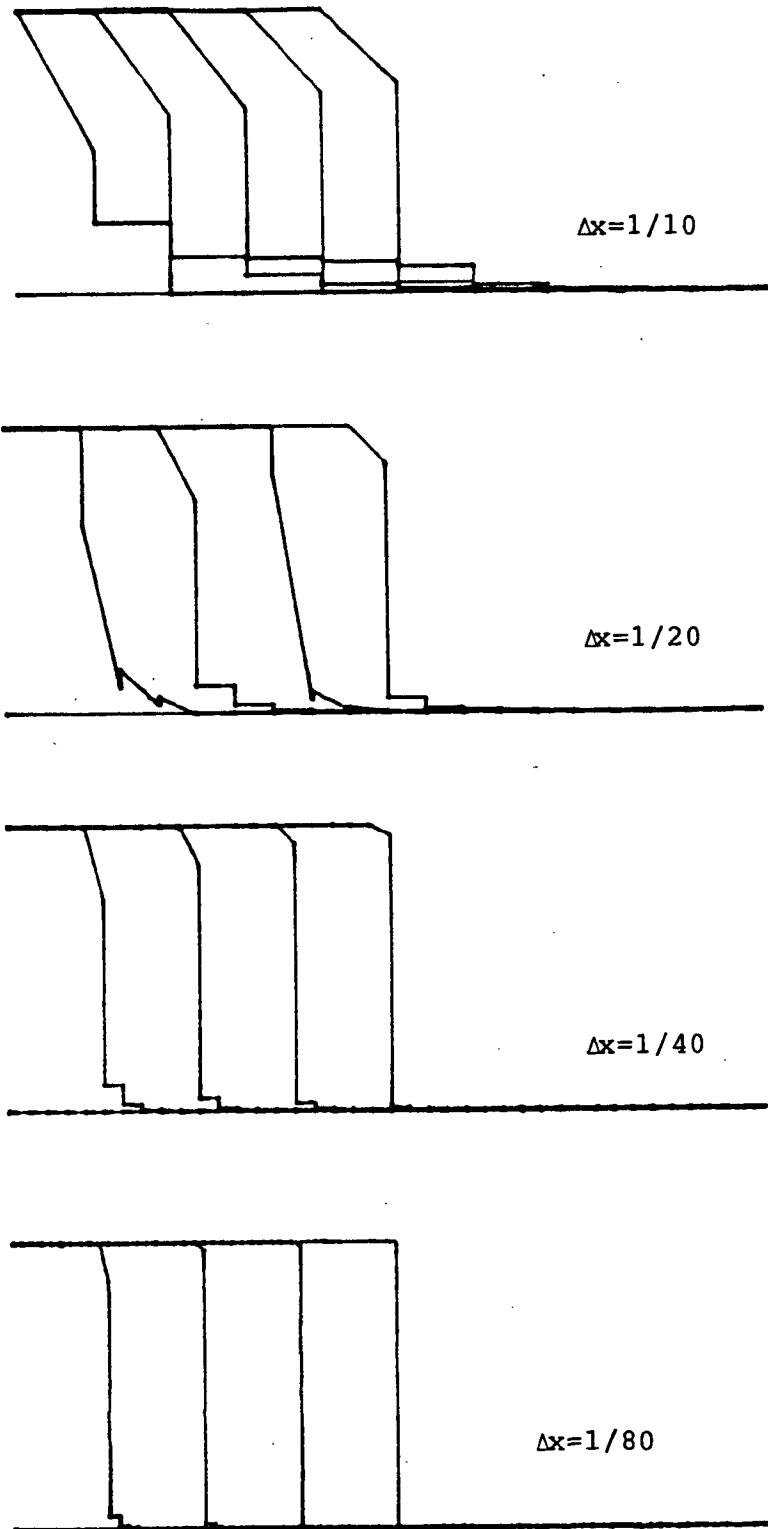


Figure 61 : Cas d'une discontinuité de contact ( $CFL = 1/2$ ). Exposant de contrôle d'anti-diffusion = 0.

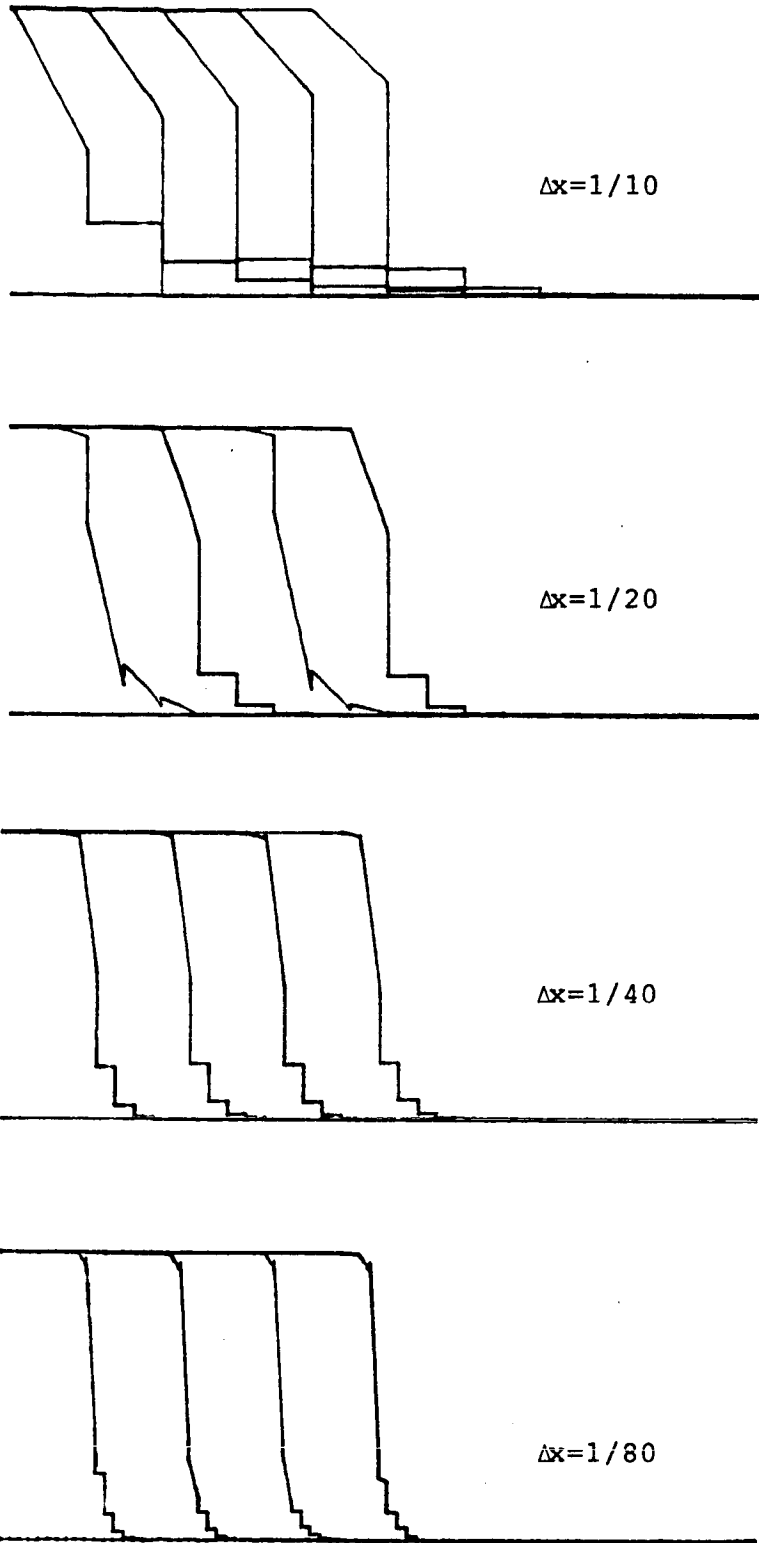


Figure 62 : Cas d'une discontinuité de contact (CFL = 1/2). Exposant de contrôle d'antidiffusion = 1/4.

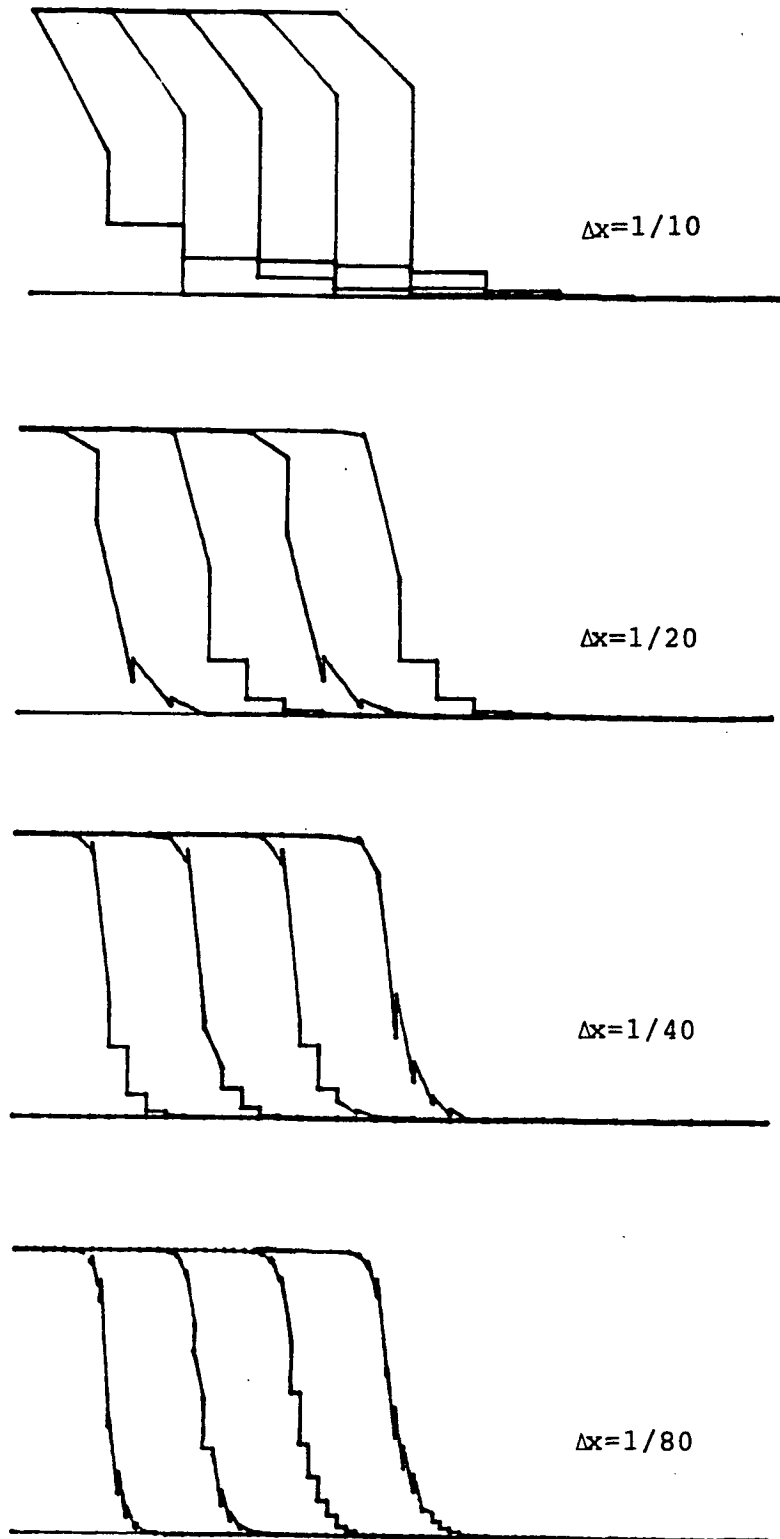


Figure 63 : Cas d'une discontinuité de contact ( $CFL = 1/2$ ). Exposant de contrôle d'antidiffusion =  $1/2$ .

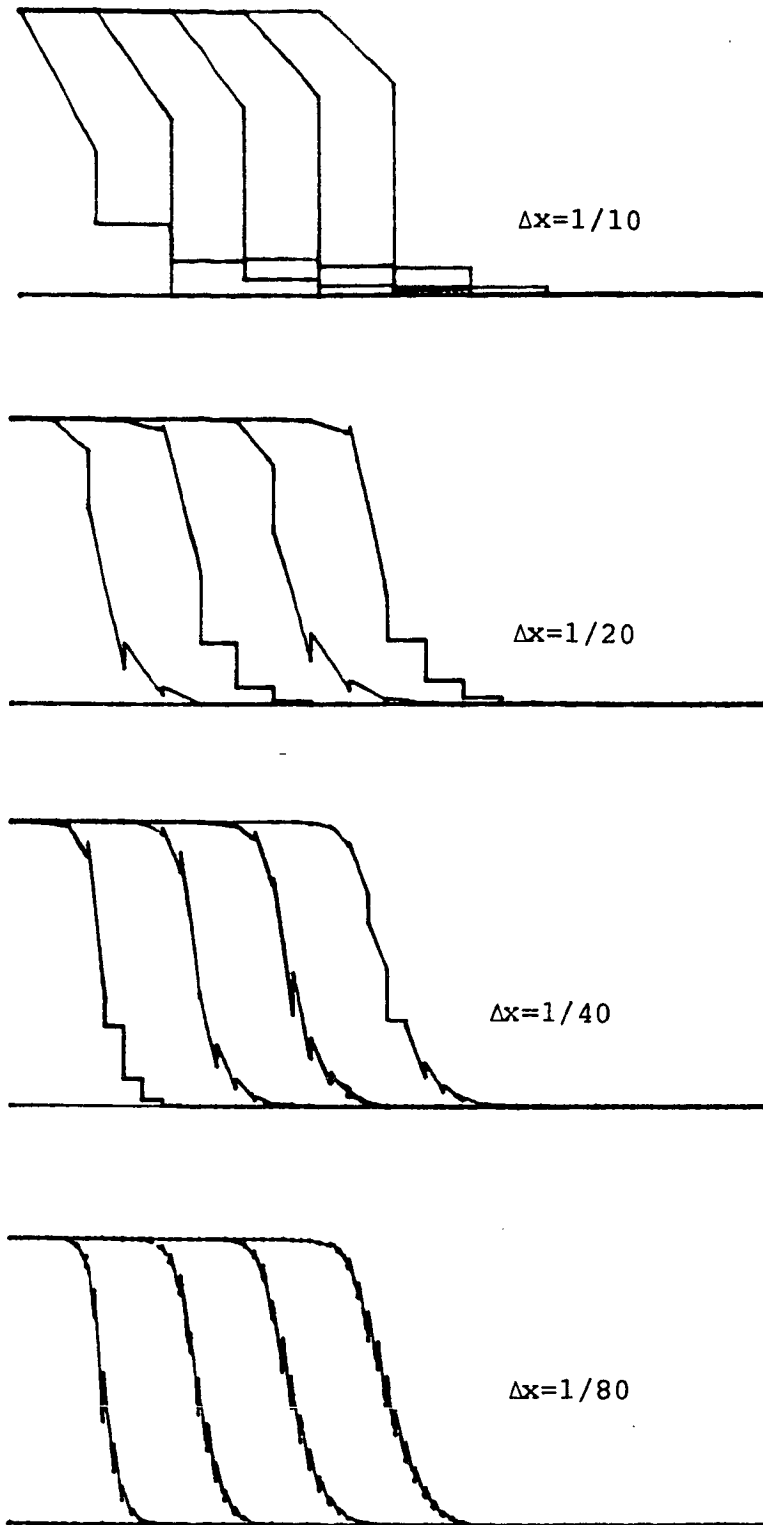


Figure 64 : Cas d'une discontinuité de contact (CFL = 1/2). Exposant de contrôle d'antidiffusion = 3/4.

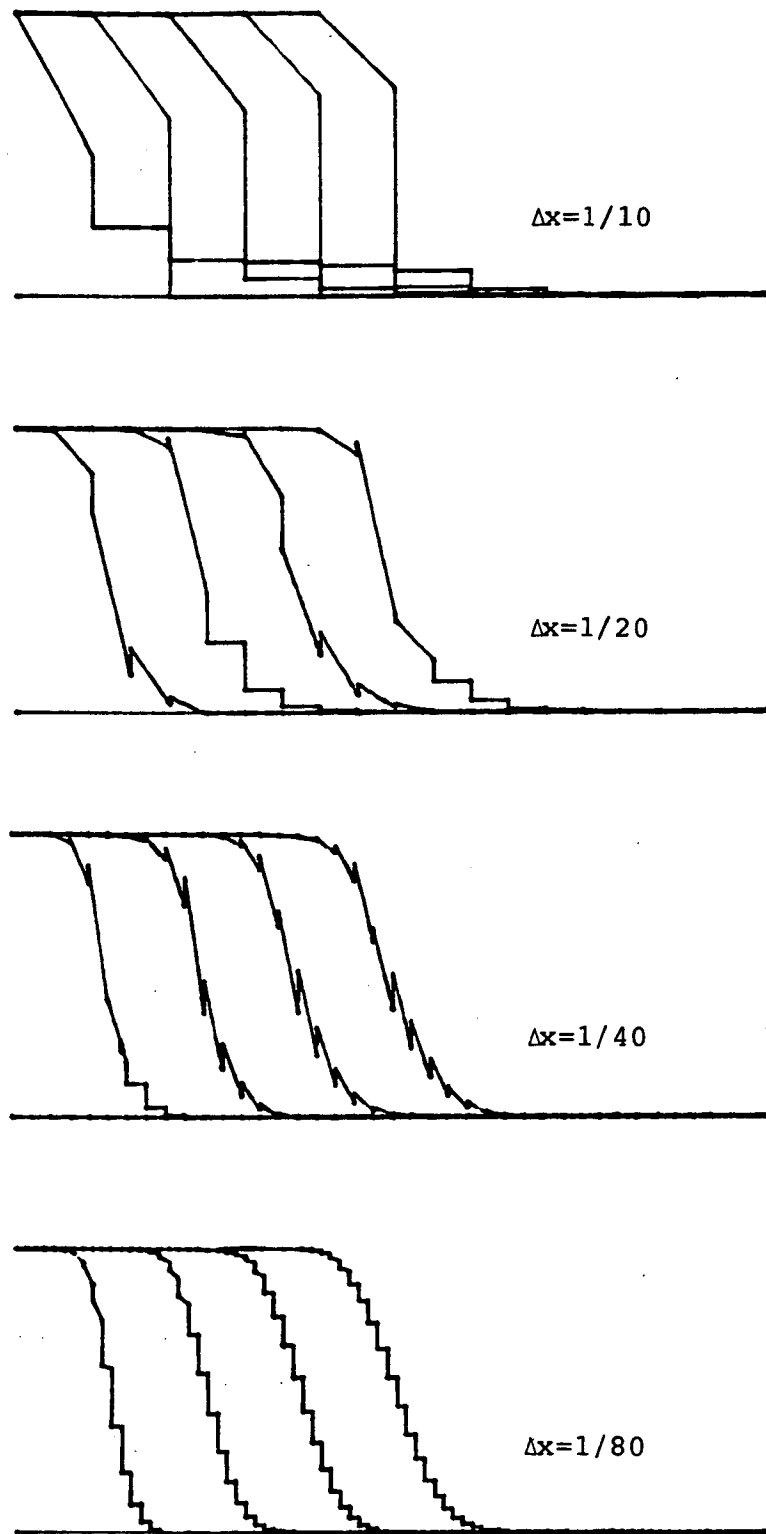


Figure 65 : Cas d'une discontinuité de contact (CFL = 1/2). Exposant de contrôle d'antidiffusion = 1.



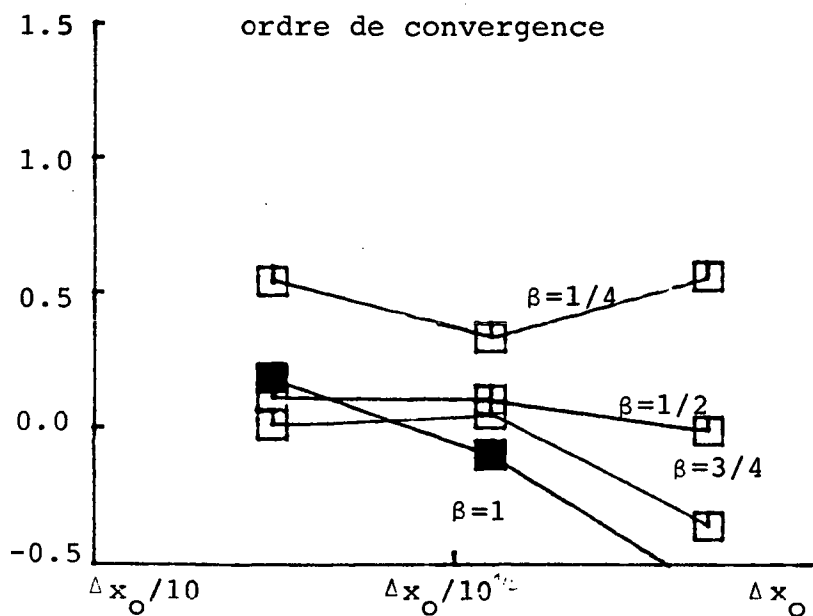
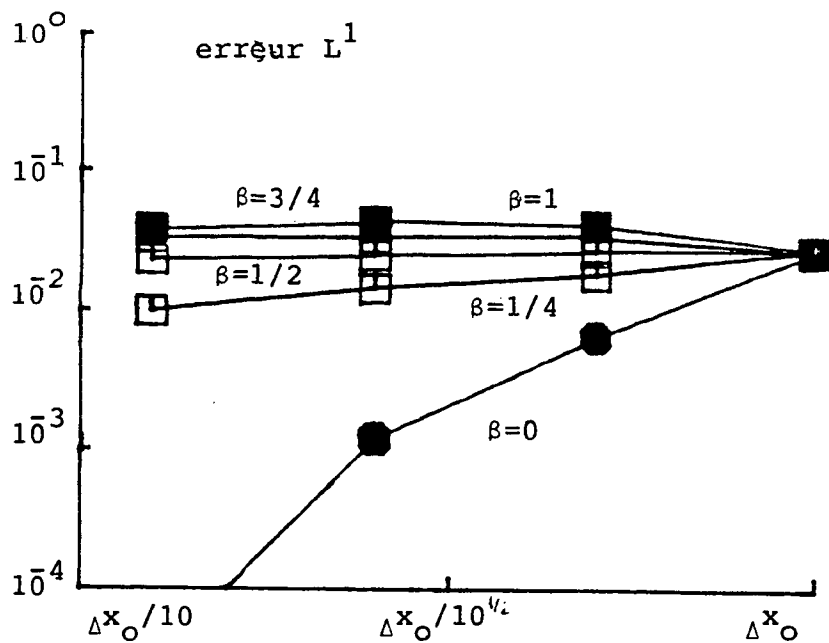


Figure 66 : Cas d'une discontinuité de contact (CFL = 1/2). Comparaison des propriétés de convergence du schéma G-1/2 pour les différentes valeurs de l'exposant de contrôle d'antidiffusion,  $\beta$ . Les valeurs de  $\alpha$ , pour  $\beta = 0$ , n'apparaissent pas sur ce dessin, elles sont respectivement (du plus grand  $\Delta x$  au plus petit) :

2.056 , 2.967 , 4.855 .

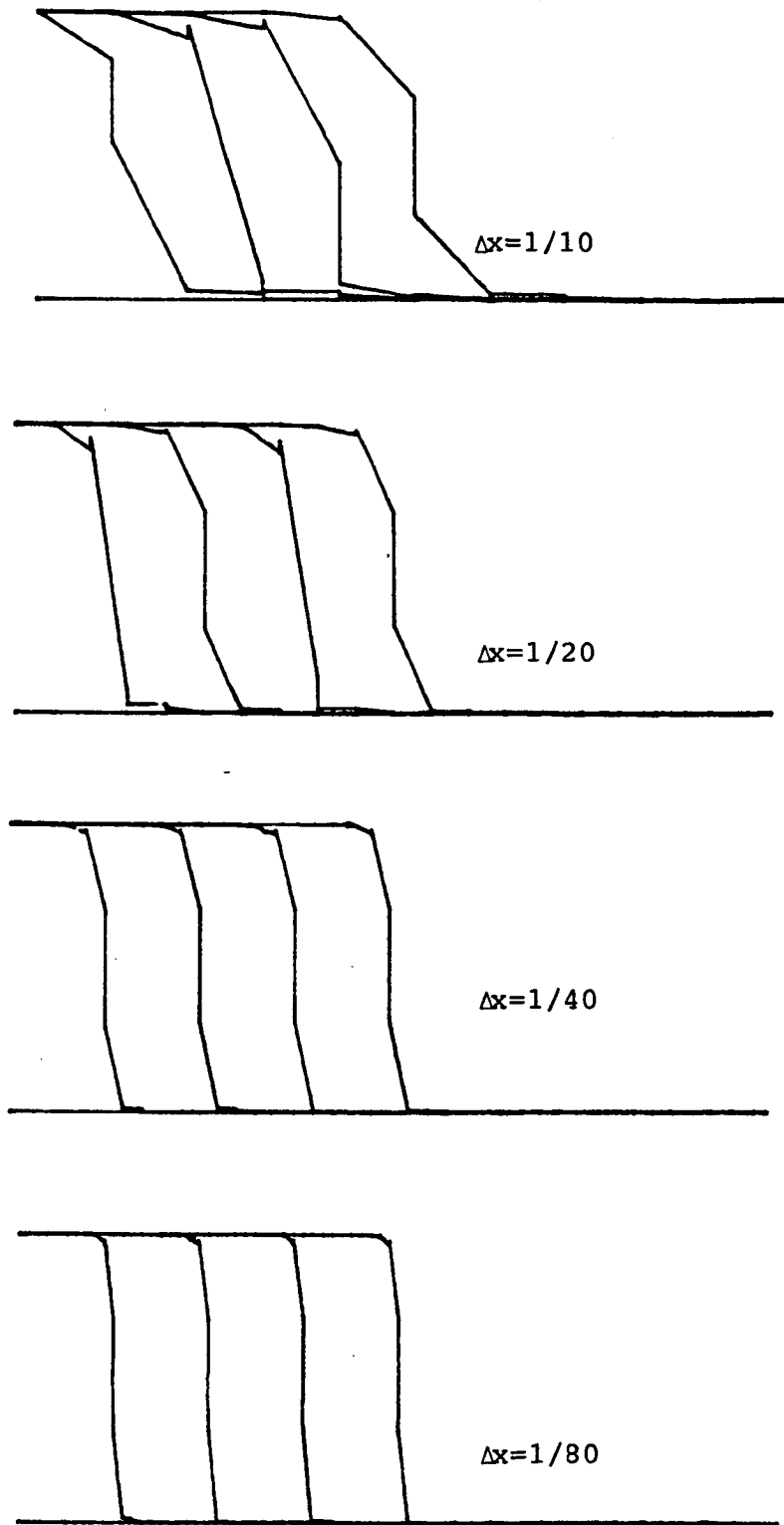


Figure 67 : Cas d'une discontinuité de contact (CFL = 1/4). Exposant de contrôle d'antidiffusion = 0.

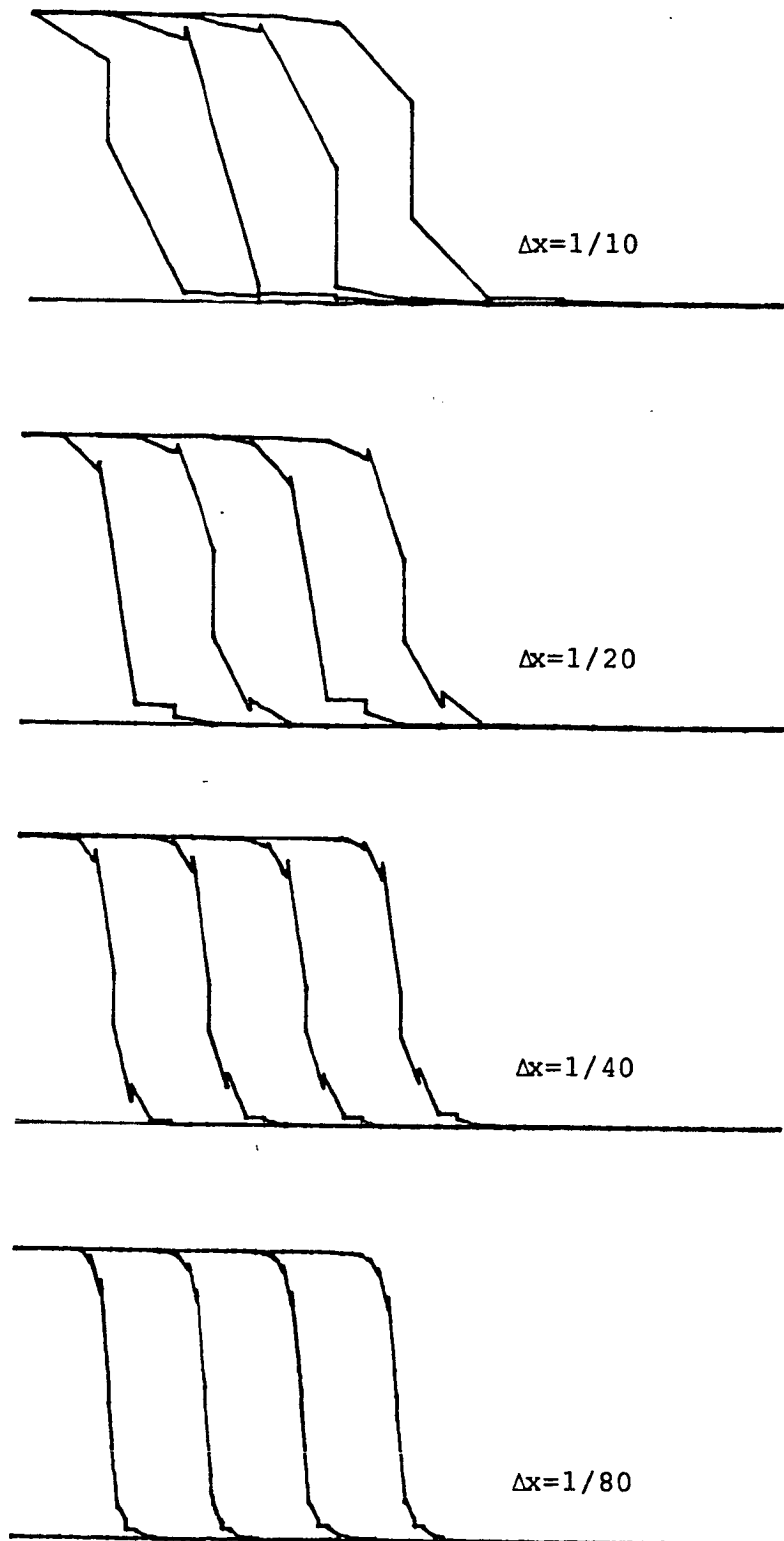


Figure 68 : Cas d'une discontinuité de contact ( $CFL = 1/4$ ). Exposant de contrôle d'antidiffusion =  $1/4$ .

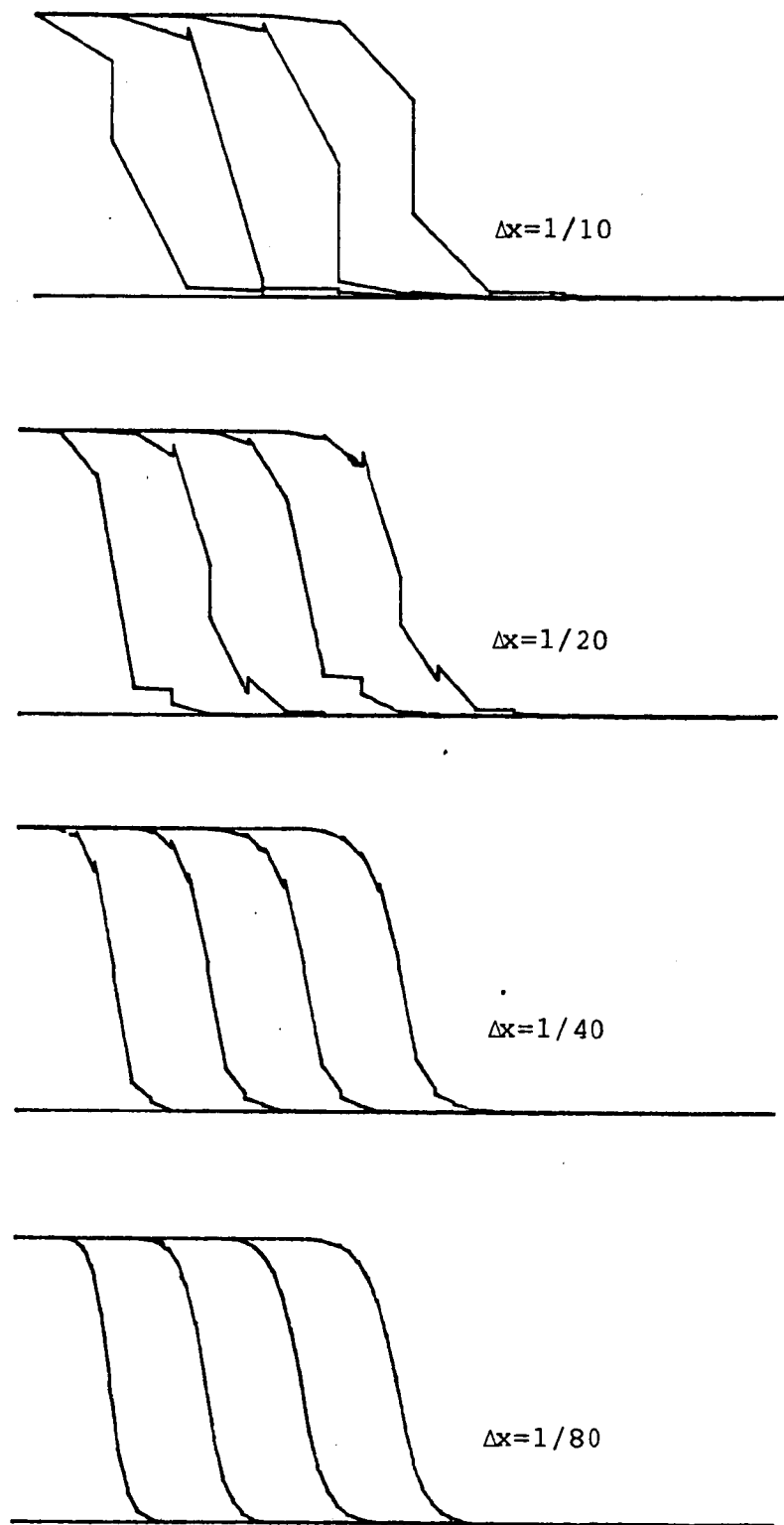


Figure 69 : Cas d'une discontinuité de contact (CFL = 1/4). Exposant de contrôle d'antidiffusion = 1/2.

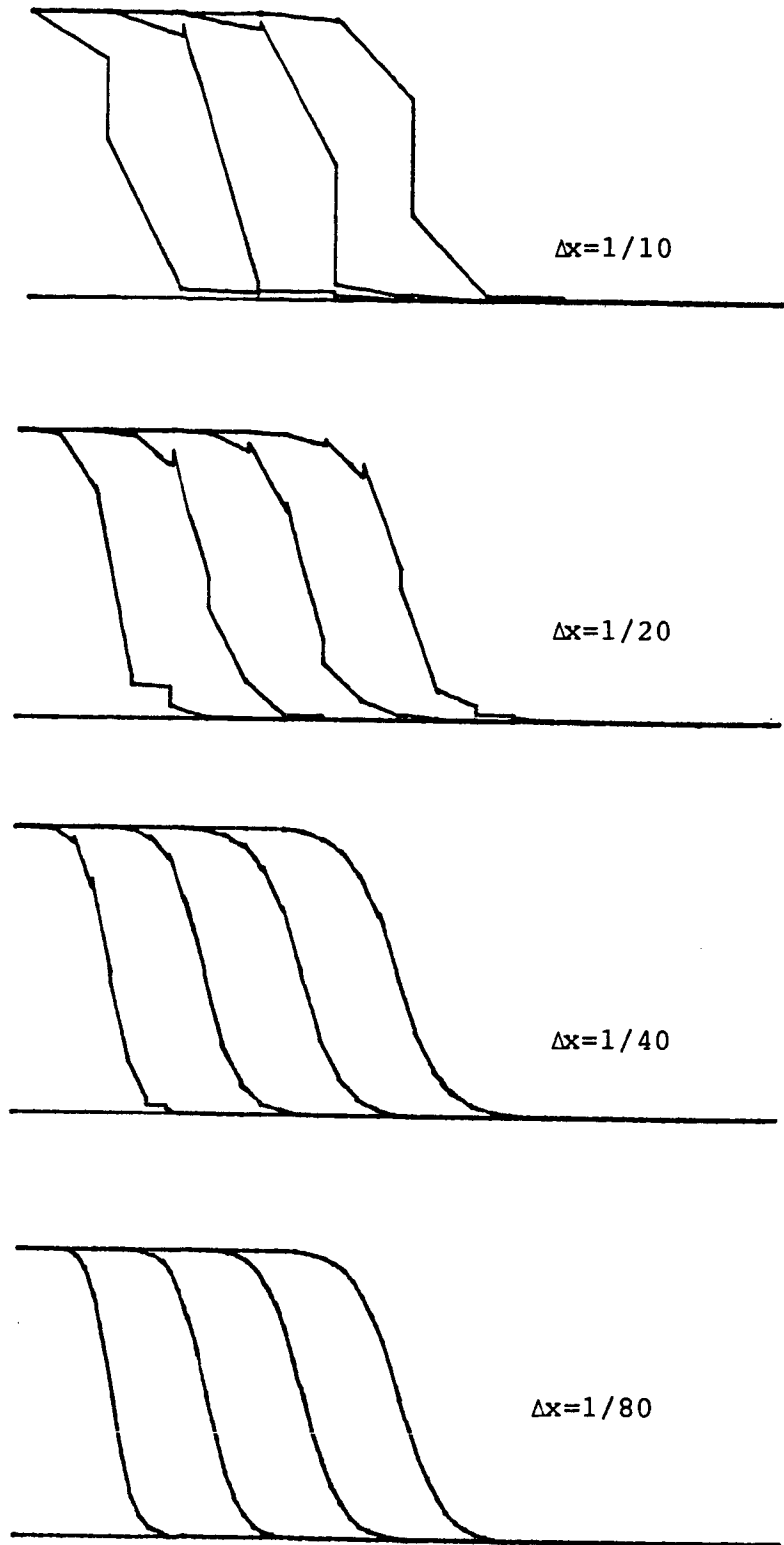


Figure 70 : Cas d'une discontinuité de contact ( $CFL = 1/4$ ). Exposant de contrôle d'antidiffusion =  $3/4$ .

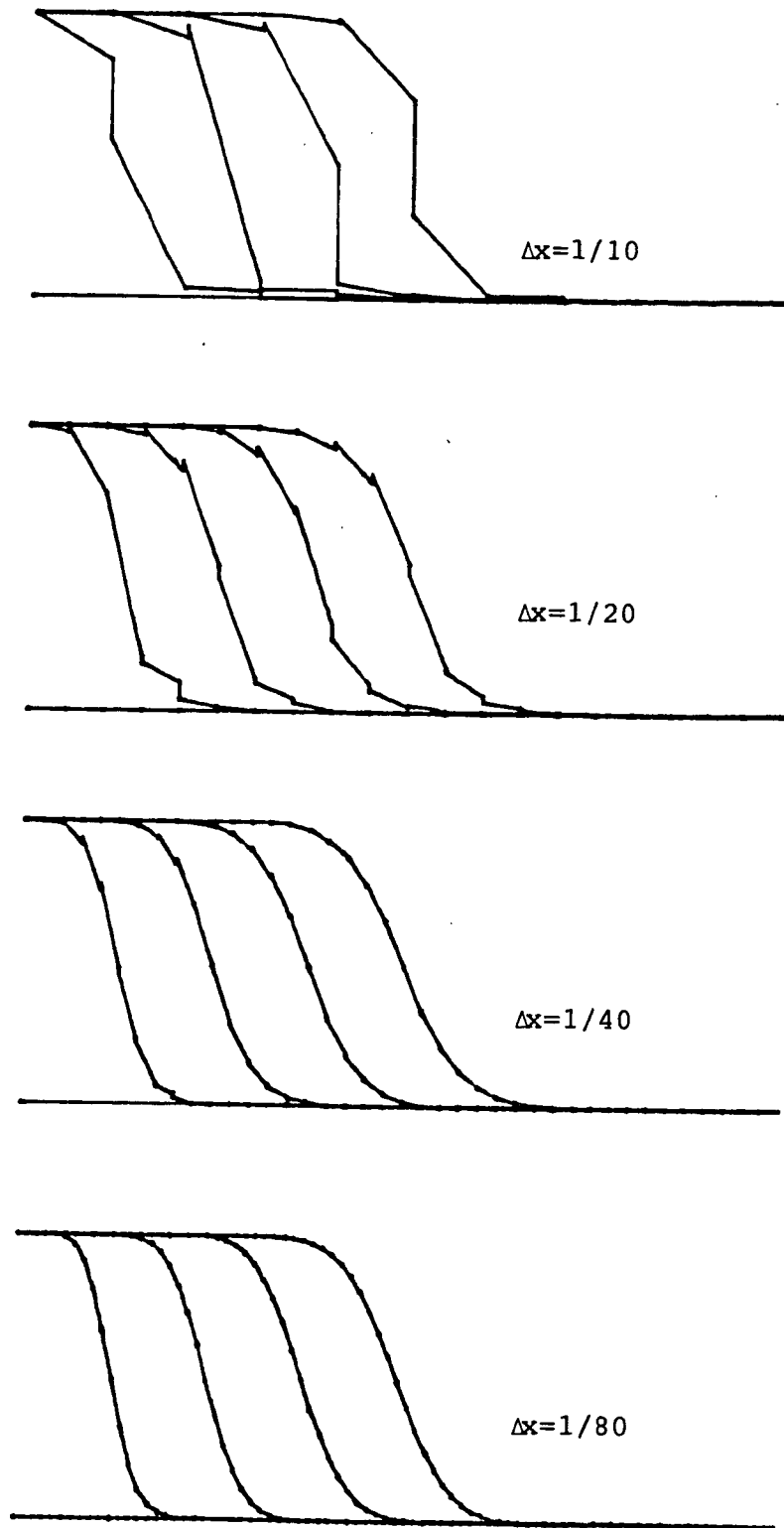


Figure 71 : Cas d'une discontinuité de contact ( $CFL = 1/4$ ). Exposant de contrôle d'antidiffusion = 1.

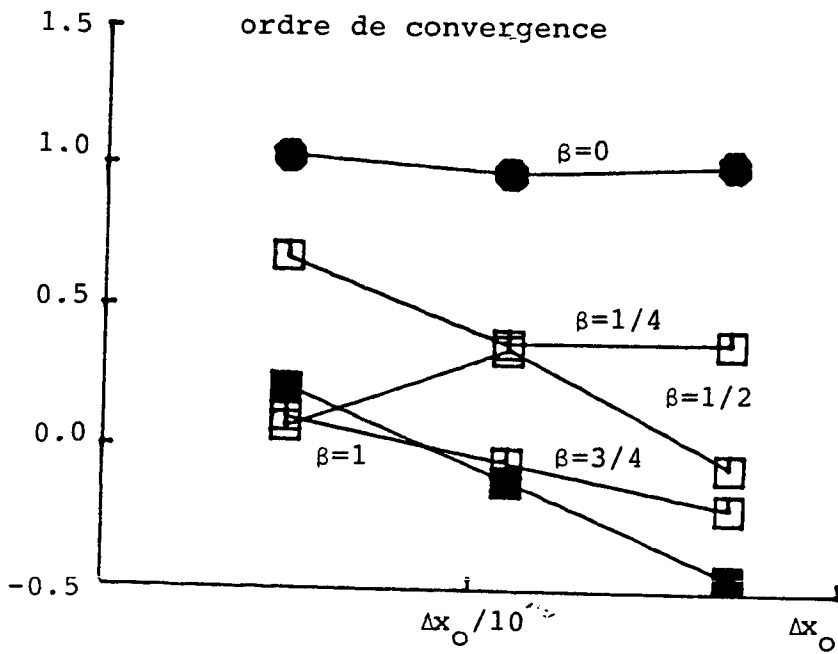
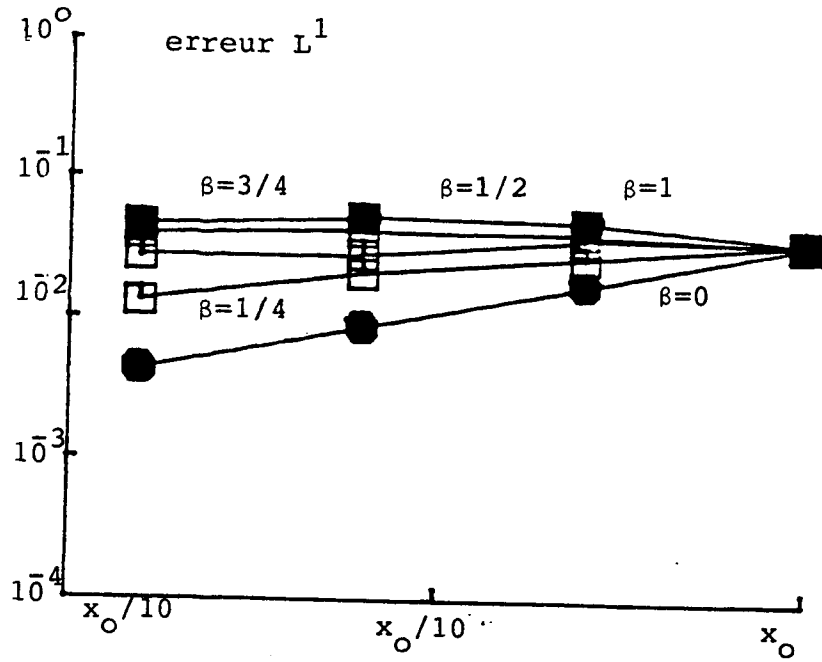


Figure 72 : Cas d'une discontinuité de contact ( $CFL = 1/4$ ). Comparaison numérique des propriétés de convergence du schéma G-1/2 pour les différentes valeurs de l'exposant de contrôle d'antidiffusion,  $\beta$ .

#### 4.6.3 - L'équation de Buckeley-Leverett : un choc

Sur les figures 73 à 77 (resp. 79 à 83), on montre les solutions données par G-1/2 pour CFL = 1/2 (resp. CFL = 1/4), pour  $\alpha = 0, 1/4, 1/2, 3/4$  et 1 sans cet ordre.

On y constate le même phénomène que dans le cas précédent.

Sur la figure 78 (resp. 81), on montre les erreurs et les ordres de convergence de chacun des cas considérés. On y voit que pour  $\alpha = 0$ , le schéma G-1/2 produit les erreurs les plus petites et possède l'ordre de convergence le plus grand.



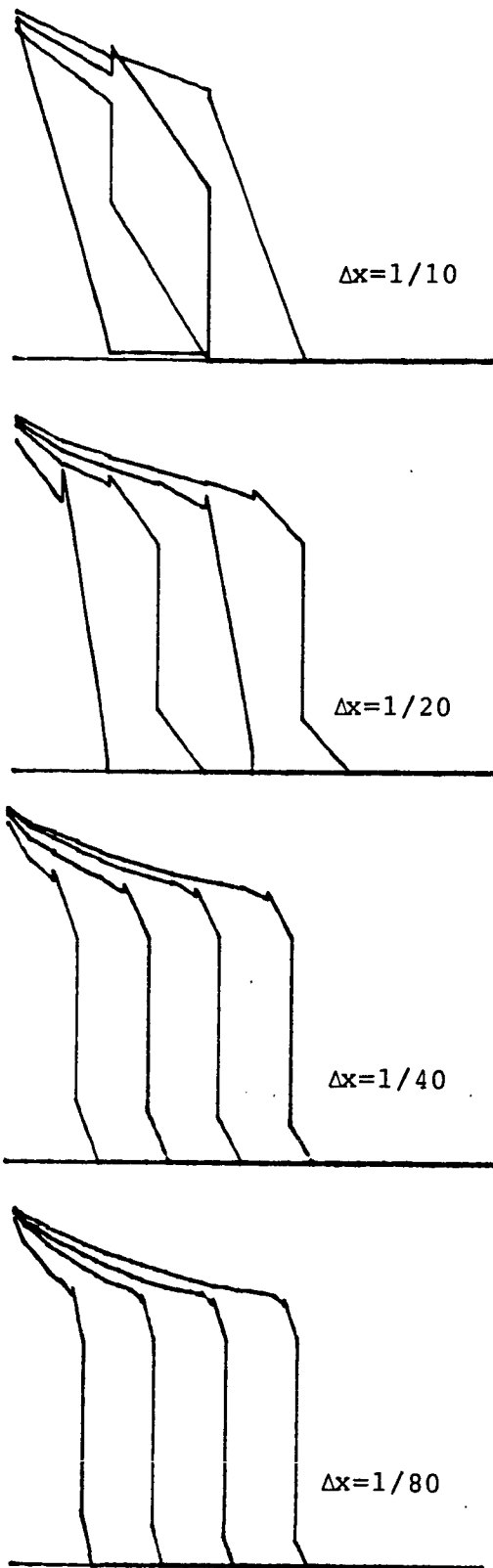


Figure 73 : Cas d'un choc : schéma G-1/2 (CFL = 1/2). Exposant du paramètre de contrôle d'antidiffusion = 0.

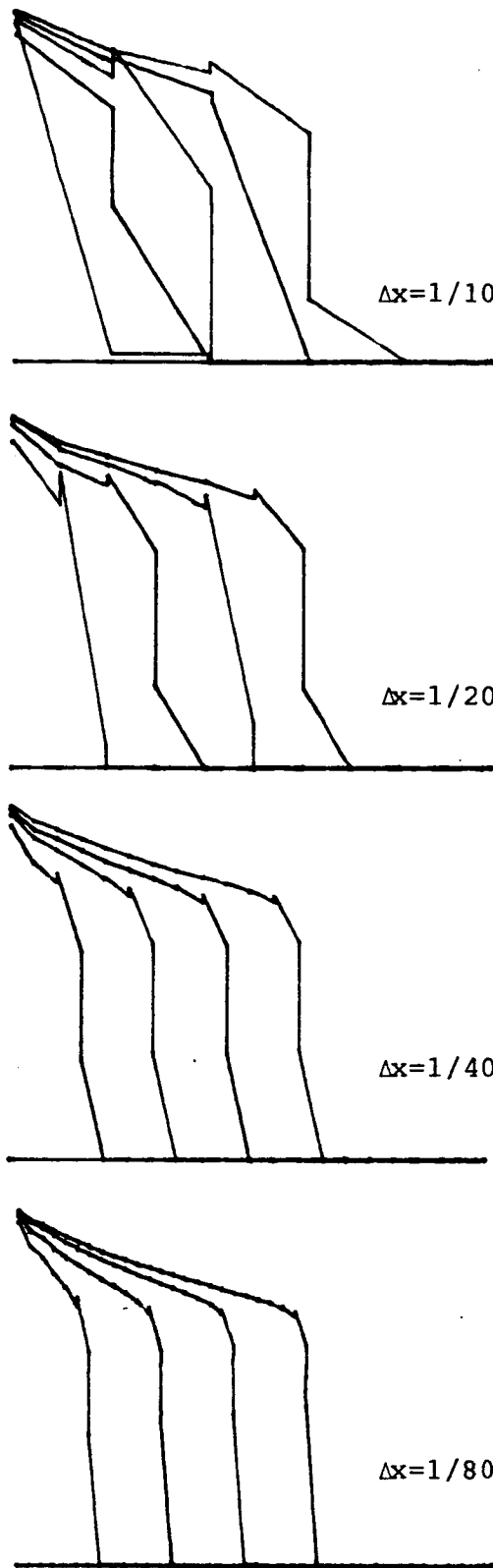


Figure 74 : Cas d'un choc : schéma G-1/2 (CFL = 1/2). Exposant du paramètre de contrôle d'antidiffusion = 1/4.

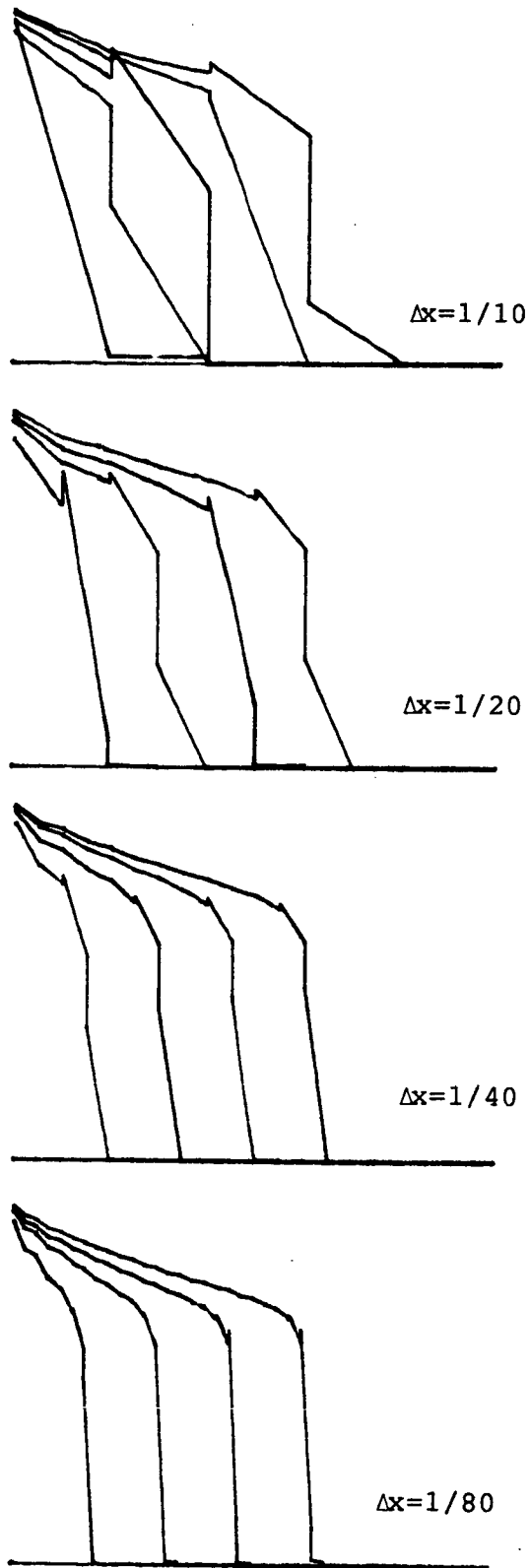


Figure 75 : Cas d'un choc : schéma G-1/2 (CFL = 1/2). Exposant du paramètre de contrôle d'antidiffusion = 1/2.

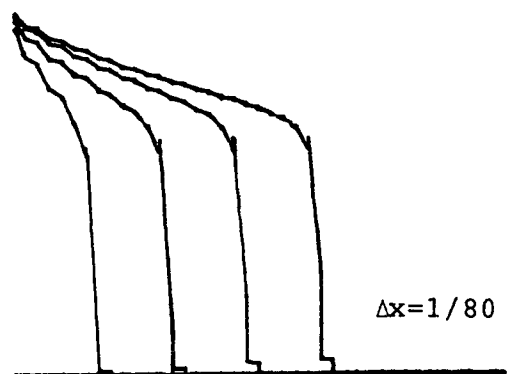
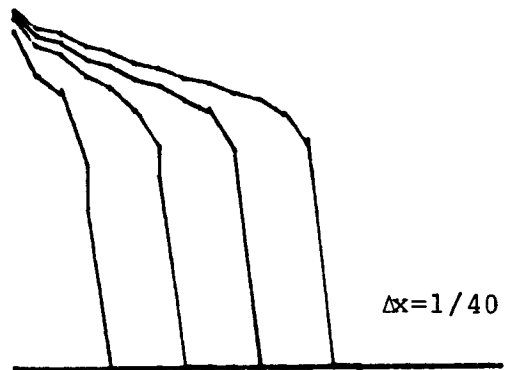
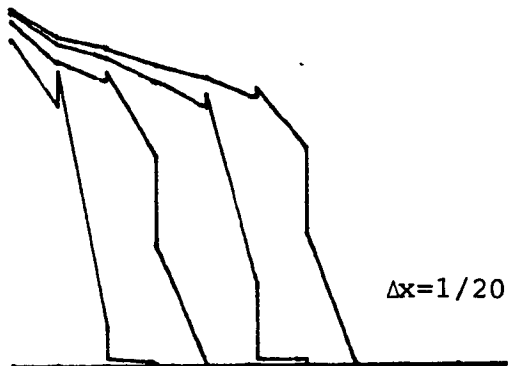
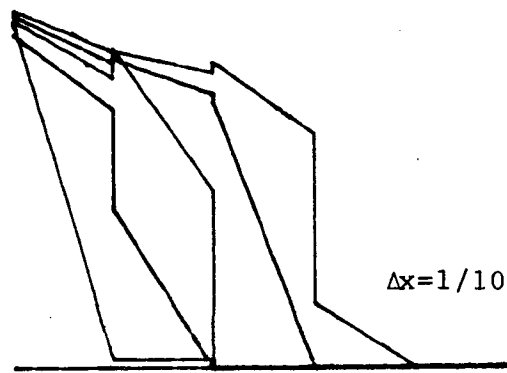


Figure 76 : Cas d'un choc : schéma G-1/2 (CFL = 1/2). Exposant du paramètre de contrôle d'antidiffusion = 3/4.

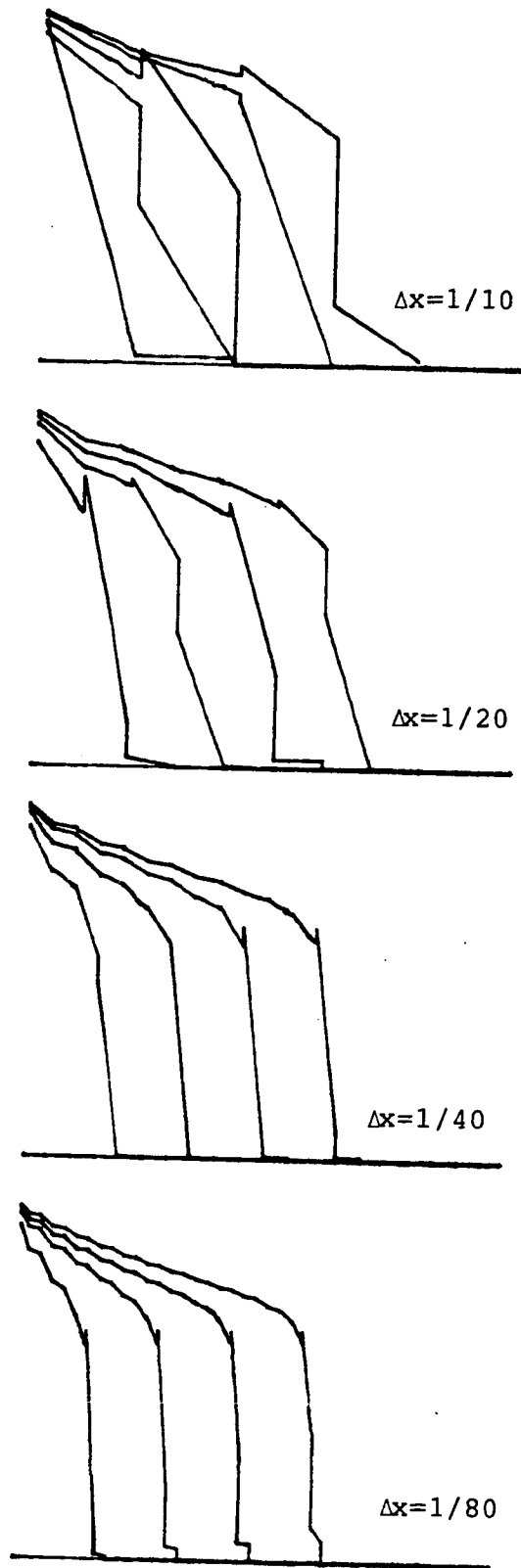


Figure 77 : Cas d'un choc : schéma G-1/2 (CFL = 1/2). Exposant du paramètre de contrôle d'antidiffusion = 1.

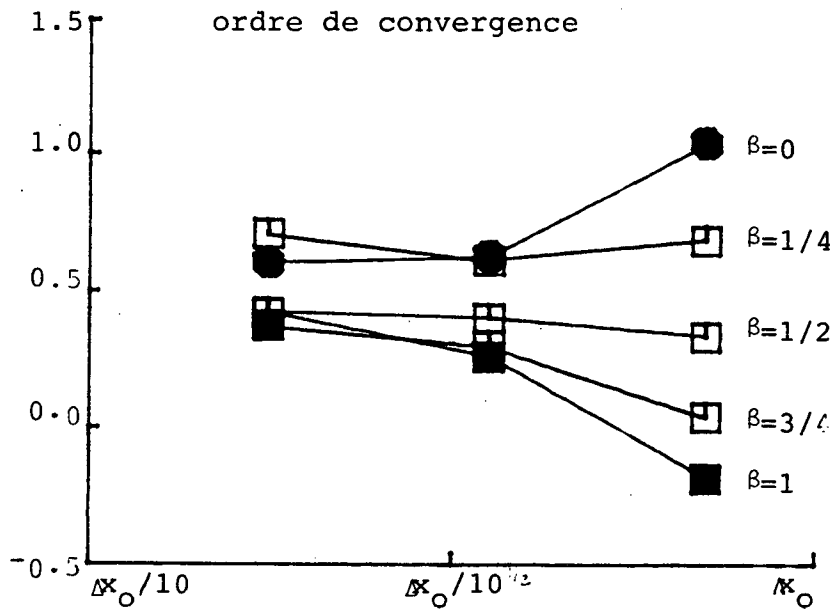
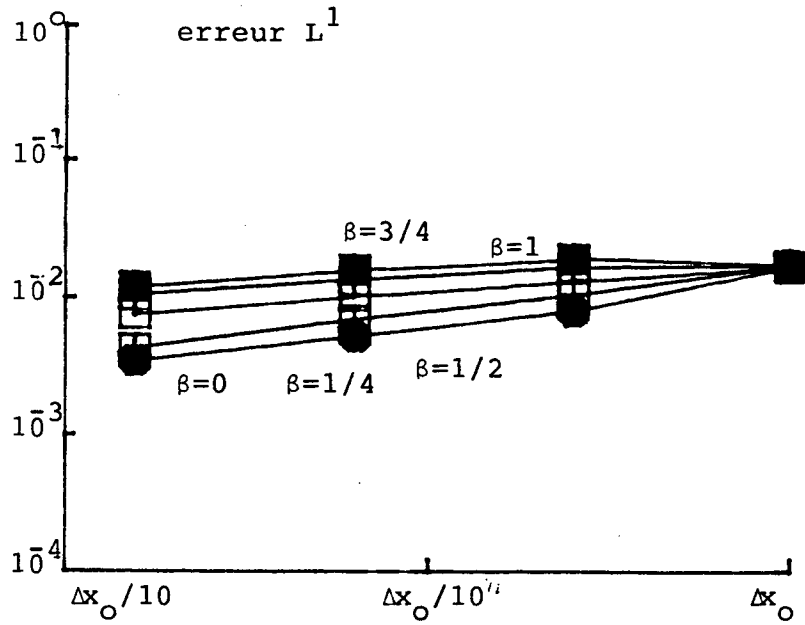


Figure 78 : Cas d'un choc : schéma G-1/2 (CFL = 1/2). Comparaison des propriétés de convergence du schéma G-1/2 pour les différentes valeurs de l'exposant de contrôle d'antidiffusion,  $\beta$ .

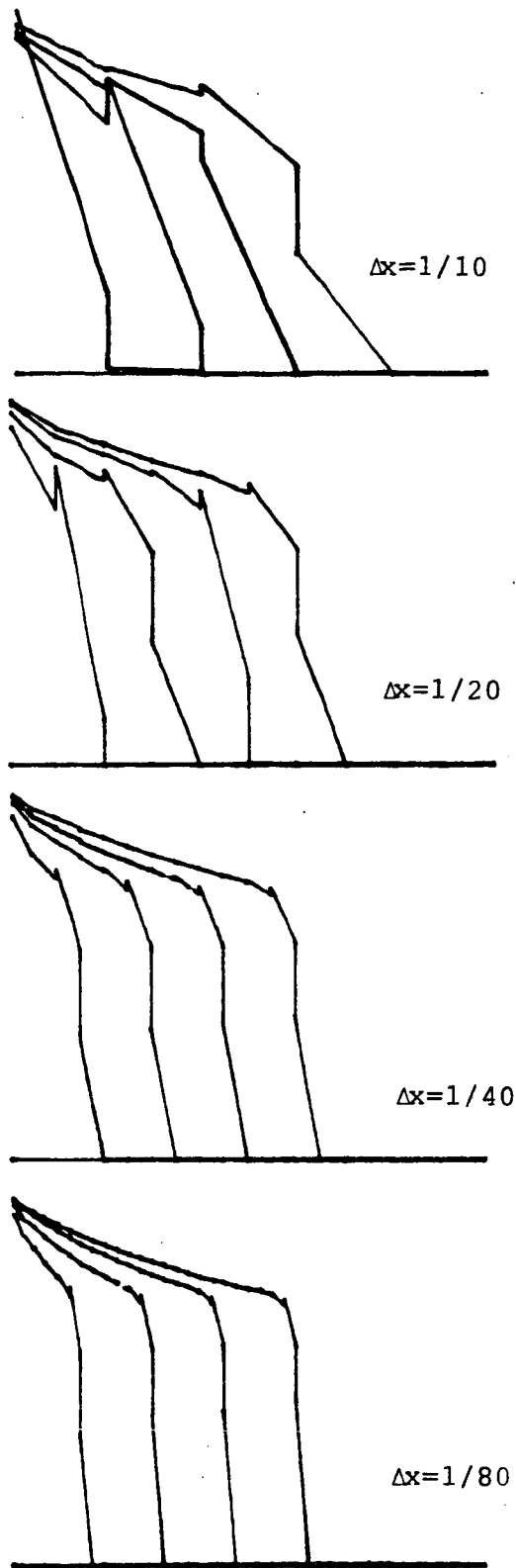


Figure 79 : Cas d'un choc : schéma G-1/2 (CFL = 1/4). Exposant du paramètre de contrôle d'antidiffusion = 0.

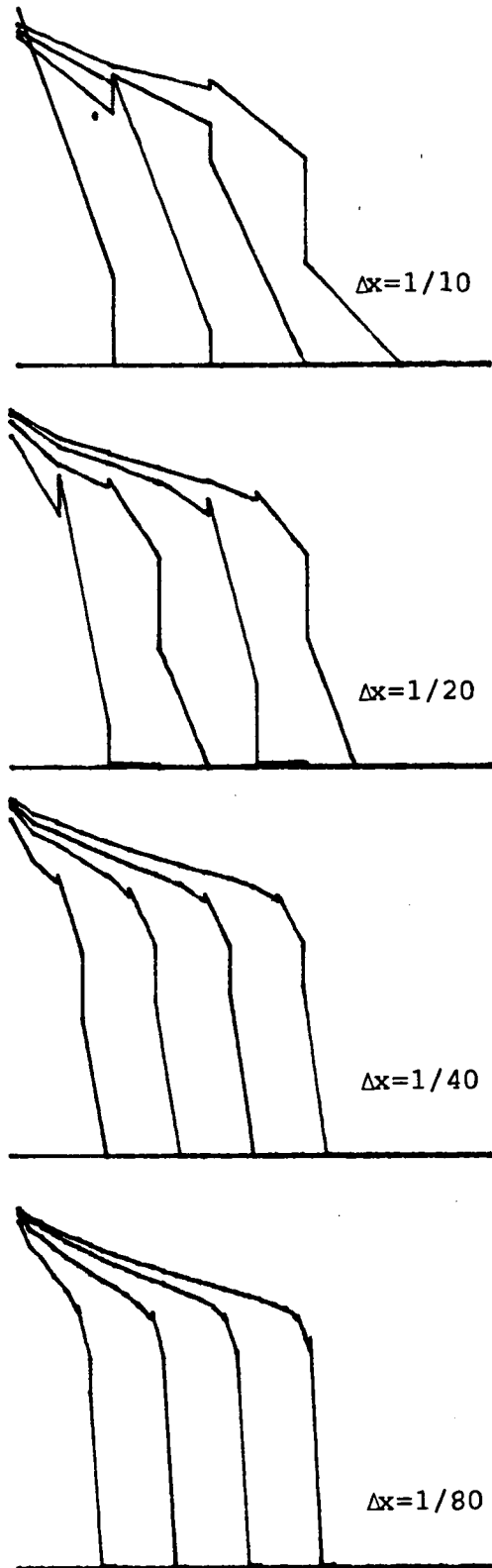


Figure 80 : Cas d'un choc : schéma G-1/2 (CFL = 1/4). Exposant de contrôle d'antidiffusion = 1/4.



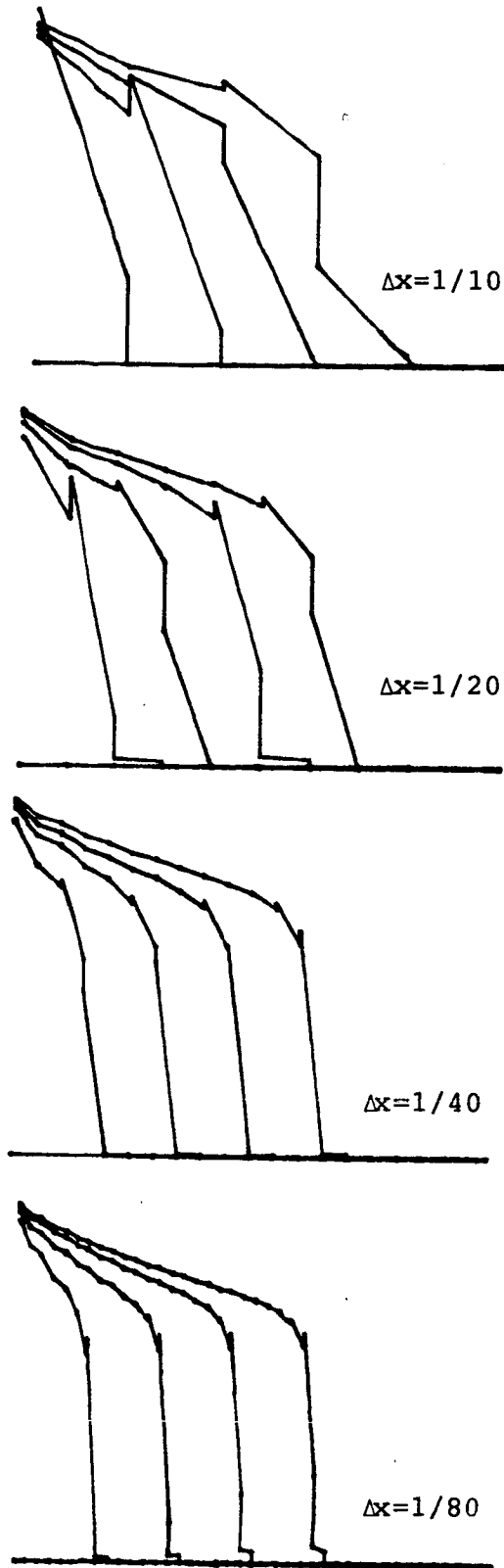


Figure 81 : Cas d'un choc : schéma G-1/2 (CFL = 1/4). Exposant de contrôle d'antidiffusion = 1/2.

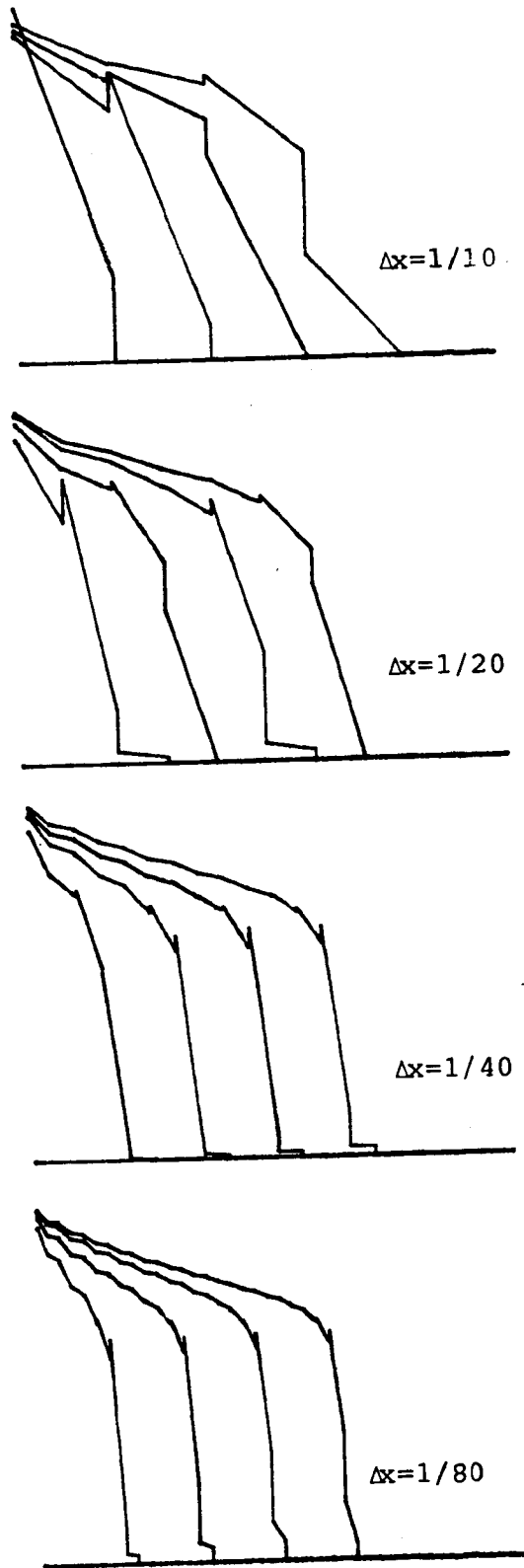


Figure 82 : Cas d'un choc : schéma G-1/2 (CFL = 1/4). Exposant de contrôle d'antidiffusion = 3/4.

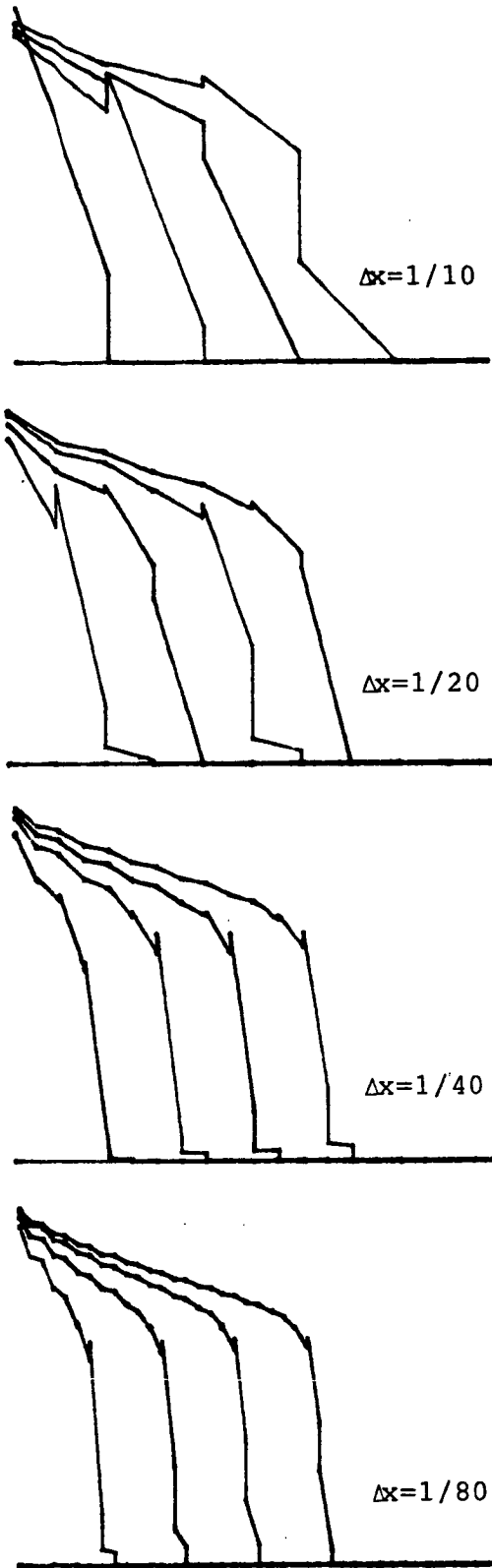


Figure 83 : Cas d'un choc : schéma G-1/2 (CFL = 1/4). Exposant de contrôle d'antidiffusion = 1.

#### 4.6.4 - Le cas d'une onde de détente

Sur les figures 85 à 89 (resp. 91 à 95), on montre les solutions données par G-1/2 pour CFL = 1/2 (resp. CFL = 1/4) pour  $\alpha = 0, 1/4, 1/2, 3/4, 1$  dans cet ordre.

Les chocs parasites qui apparaissent sur la figure 85 (resp. la figure 91) disparaissent pratiquement pour  $\alpha = 1/4$ , voir la figure 86 (resp. 92). Au fur et à mesure que  $\alpha$  augmente, on voit clairement augmenter la diffusion du schéma. Le meilleur choix du paramètre  $\alpha$  semble être  $\alpha = 1/4$ .

Sur la figure 90 (resp. 96), on montre les courbes d'erreur et d'ordre de convergence. On y constate que :

- i) pour h "petit", G-1/2 produit les plus petites erreurs et possède l'ordre de convergence le plus grand pour  $\alpha = 1/4$ ;
- ii) les erreurs produites par G-1/2 avec  $\alpha = 1/4$  et  $\alpha = 1/2$  sont à peu près les mêmes;
- iii) l'ordre de convergence de G-1/2 avec  $\alpha = 0$  est pour CFL = 1/2 (resp. CFL = 1/4)  $\approx 0.25$  (resp.  $\approx 0.20$ ). Avec  $\alpha = 1/4$ , il est  $\approx 0.75$  (resp.  $\approx 0.75$ ). Ceci montre qu'il est très important d'estimer correctement ce paramètre.

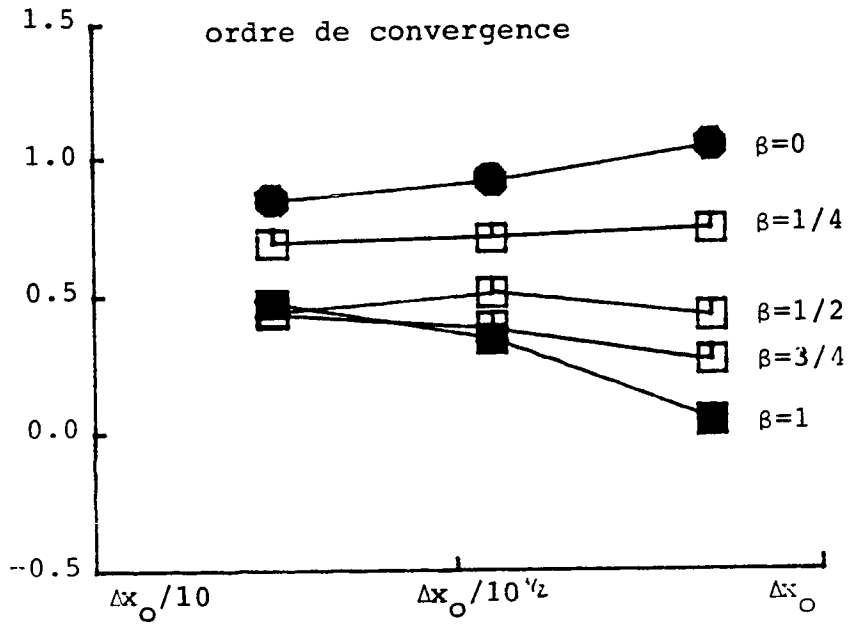
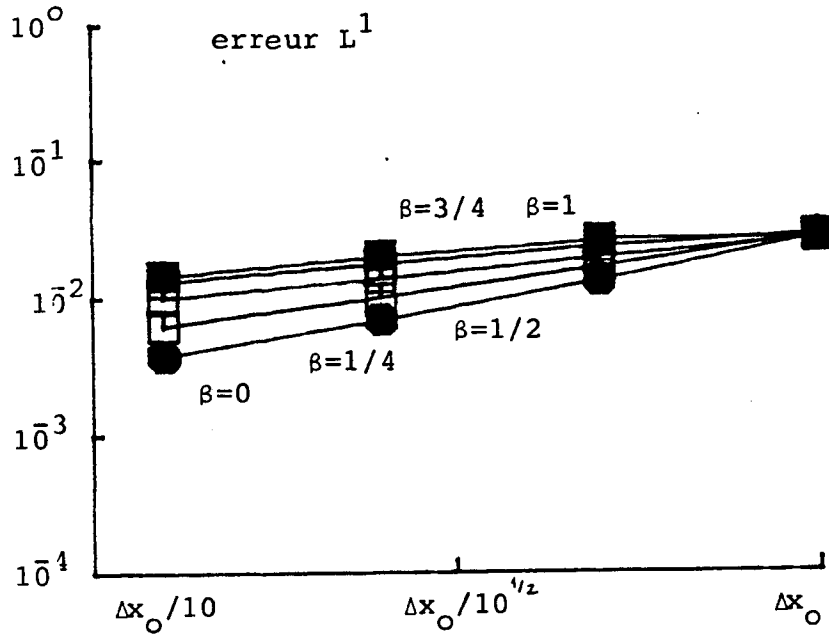


Figure 84 : Cas d'un choc : schéma G-1/2 (CFL = 1/4). Comparaison numérique des propriétés de convergence du schéma G-1/2 pour les différentes valeurs de l'exposant de contrôle d'antidiffusion,  $\beta$ .

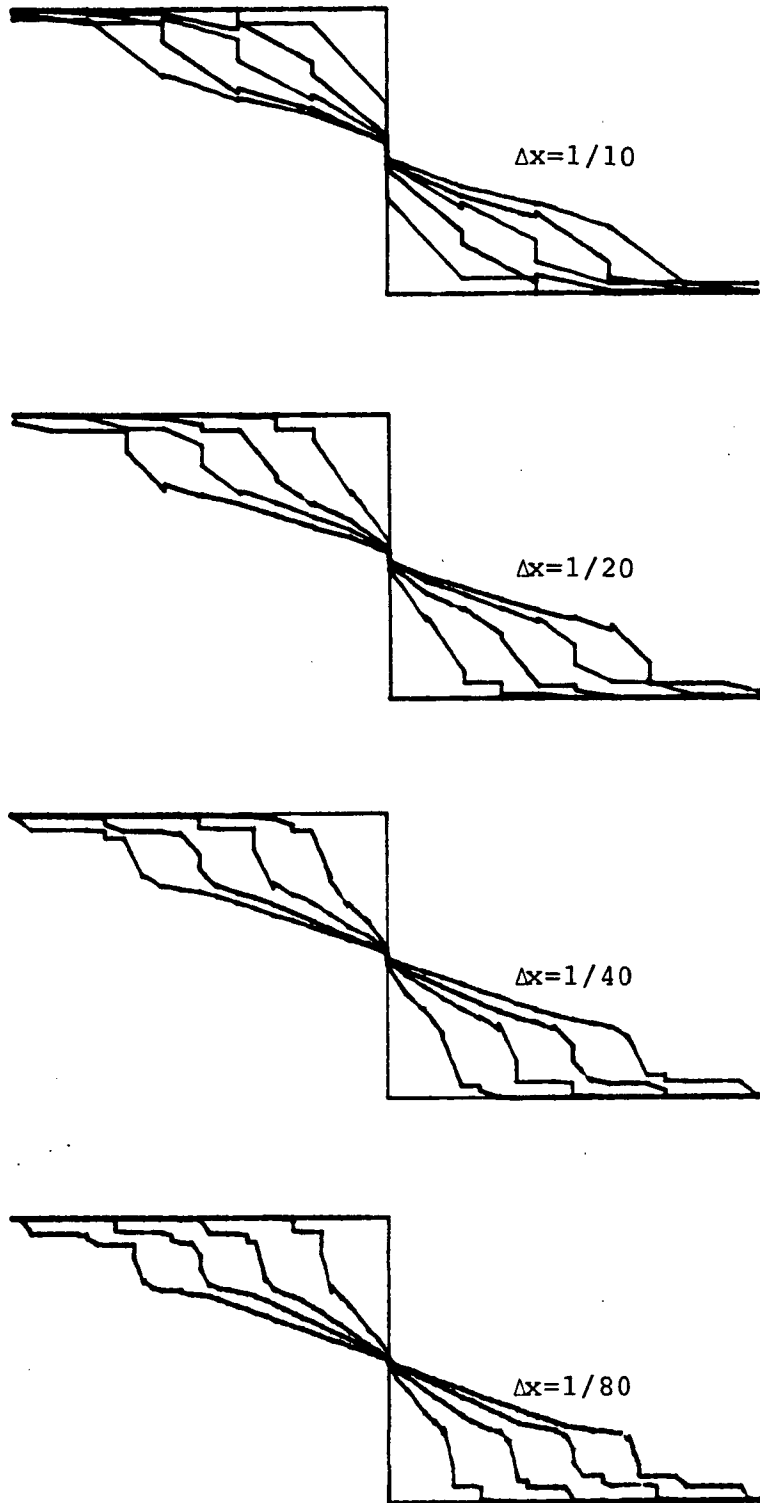


Figure 85 : Cas d'une onde de détente ( $CFL = 1/2$ ). Exposant de contrôle d'antidiffusion = 0.

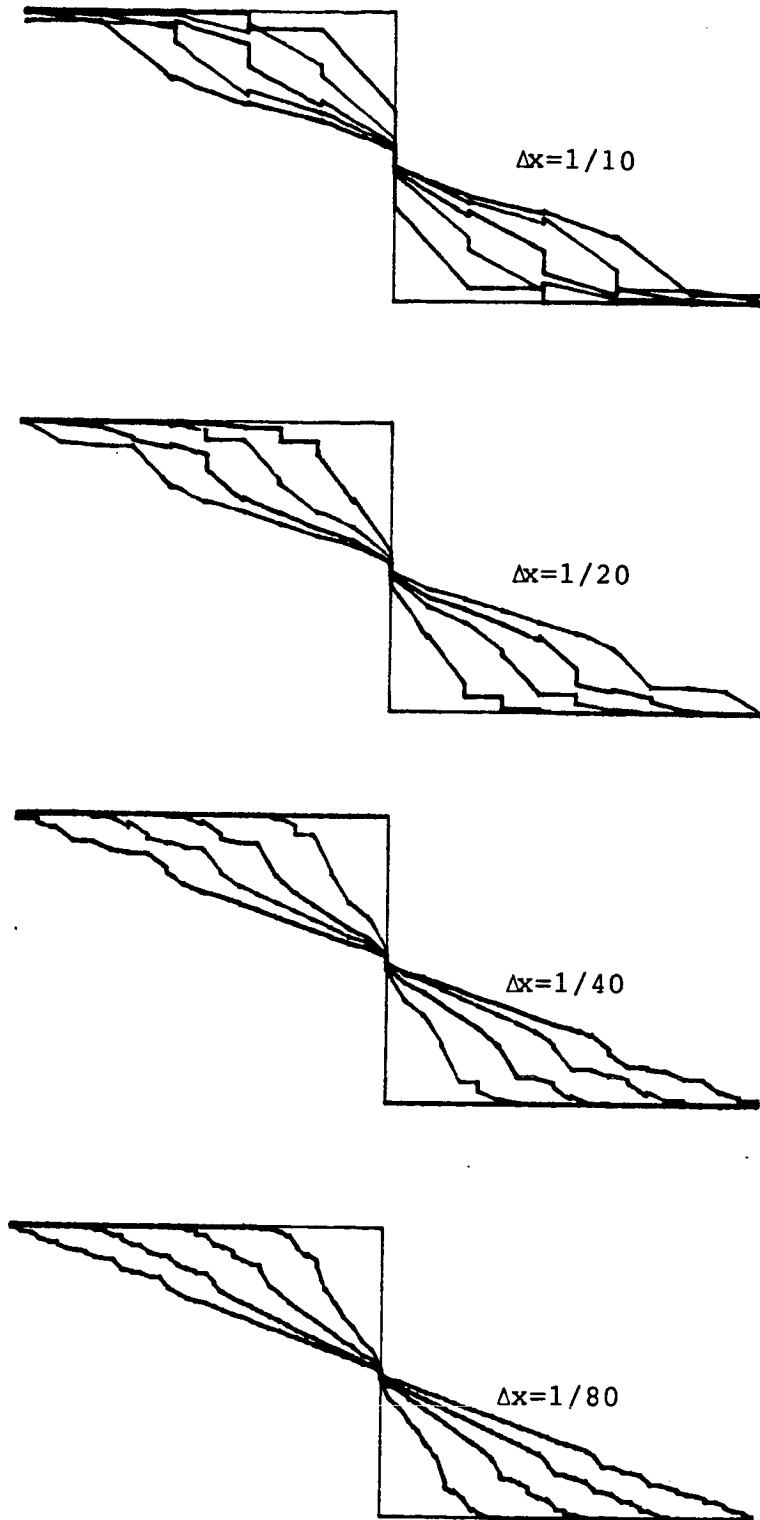


Figure 86 : Cas d'une onde de détente ( $CFL = 1/2$ ). Exposant de contrôle d'antidiffusion =  $1/4$ .

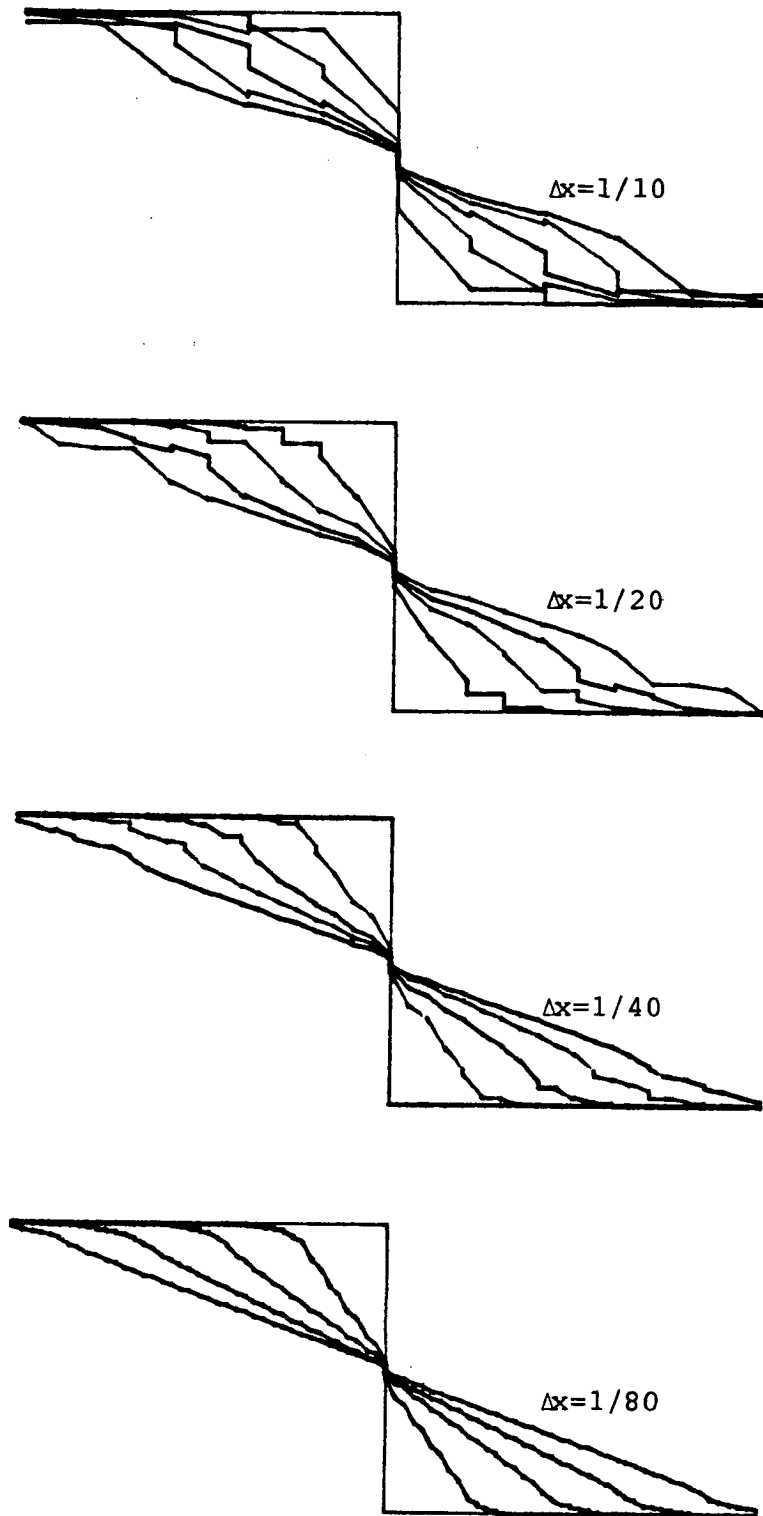


Figure 87 : Cas d'une onde de détente (CFL = 1/2). Exposant de contrôle d'antidiffusion = 1/2.



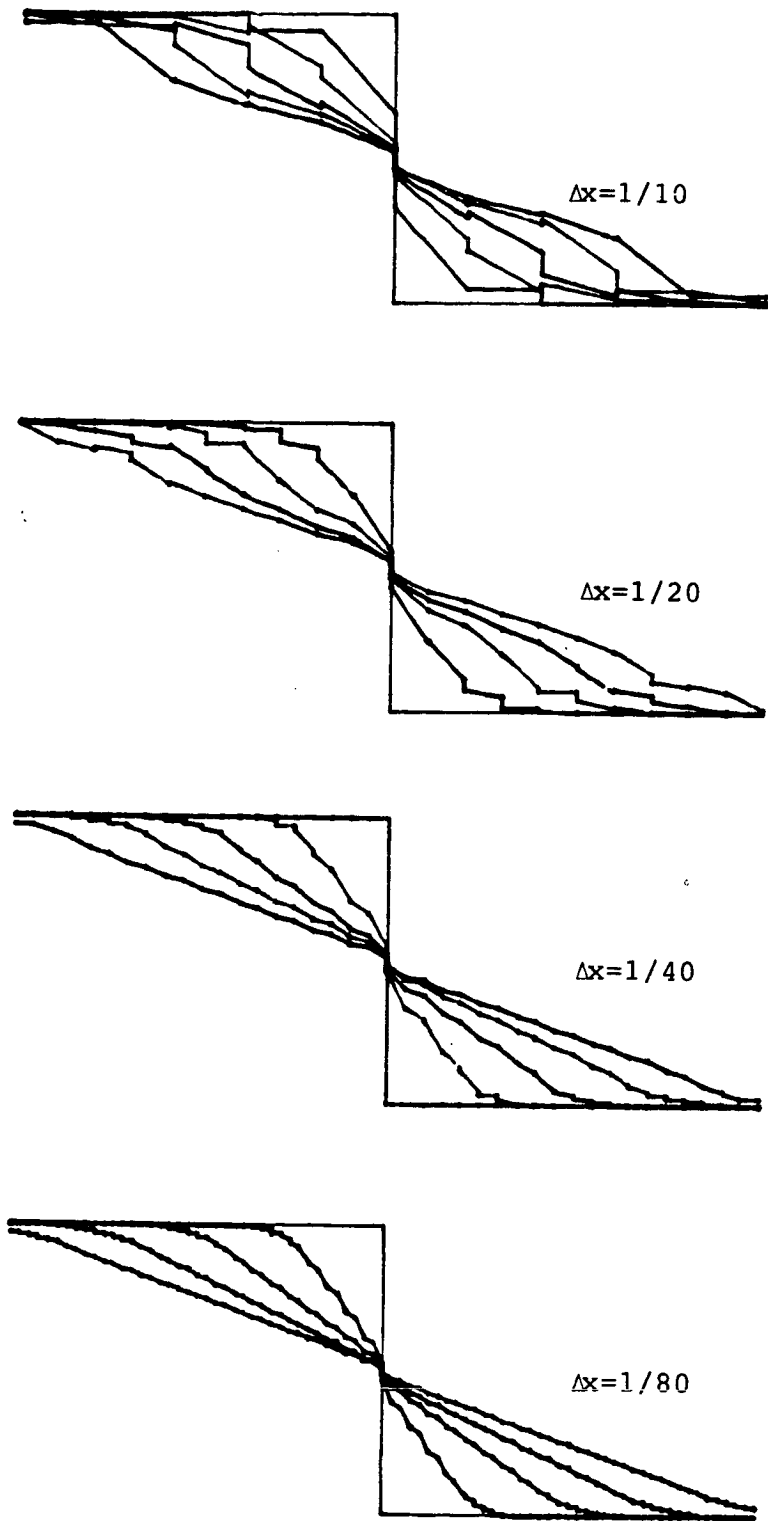


Figure 88 : Cas d'une onde de détente ( $CFL = 1/2$ ). Exposant de contrôle d'antidiffusion =  $3/4$ .

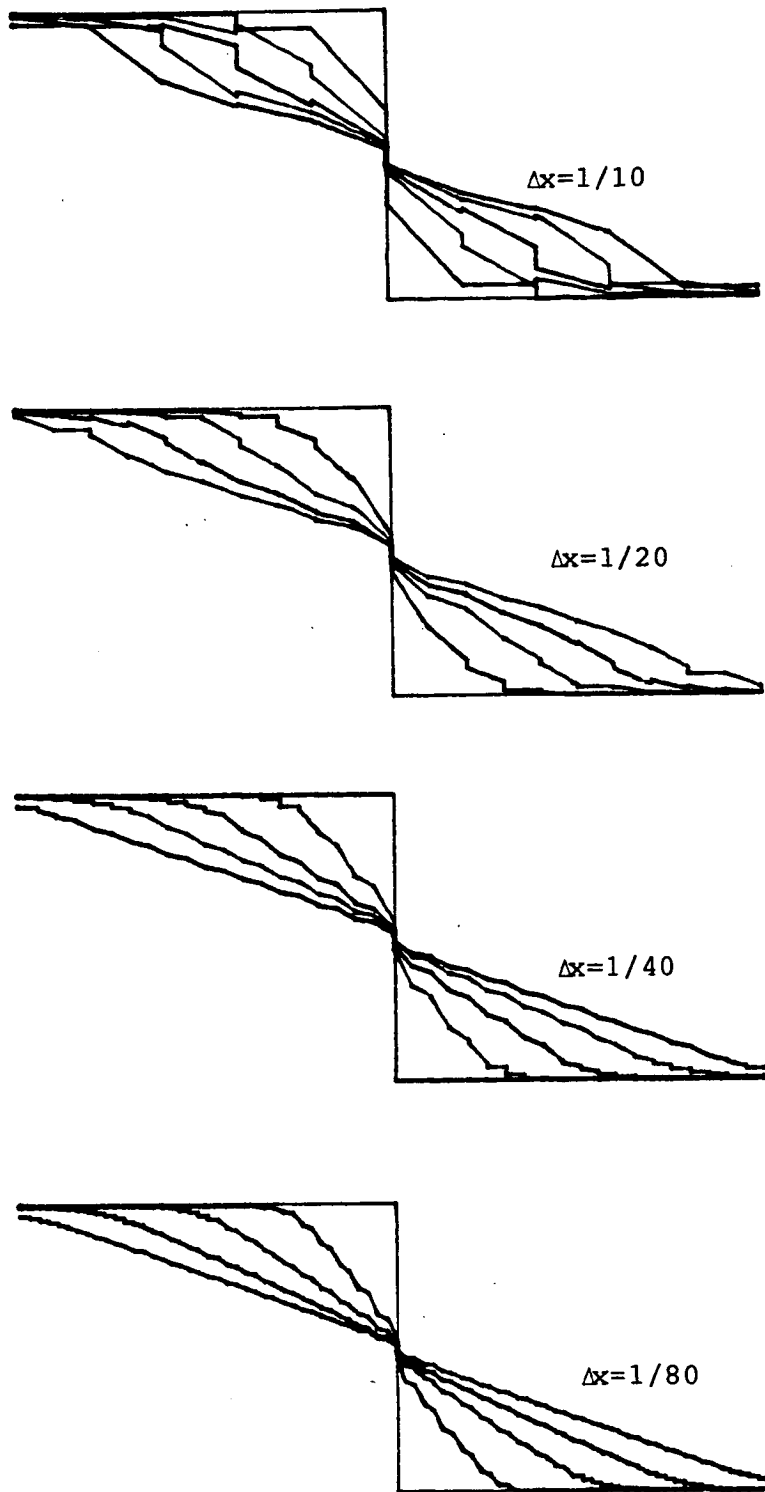


Figure 89 : Cas d'une onde de détente ( $CFL = 1/2$ ). Exposant de contrôle d'antidiffusion = 1.

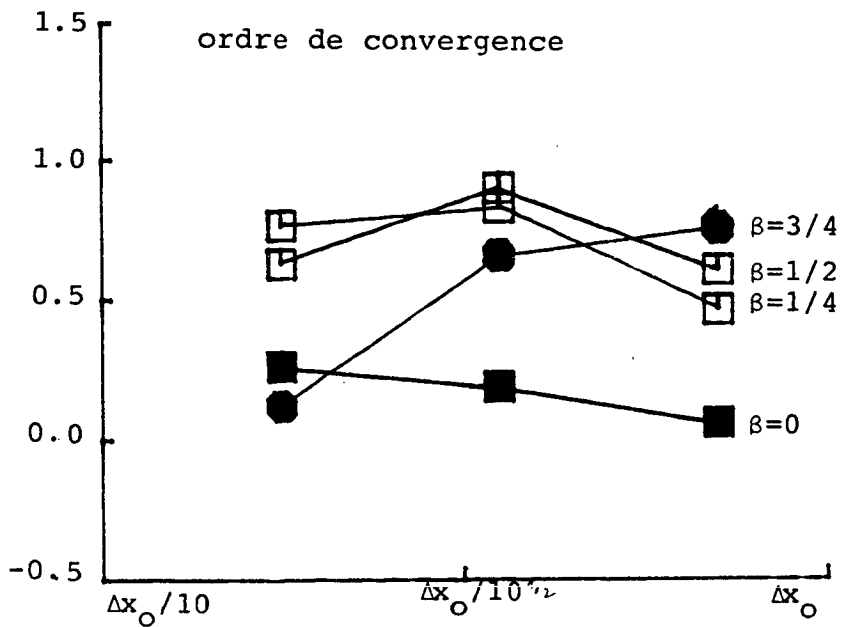
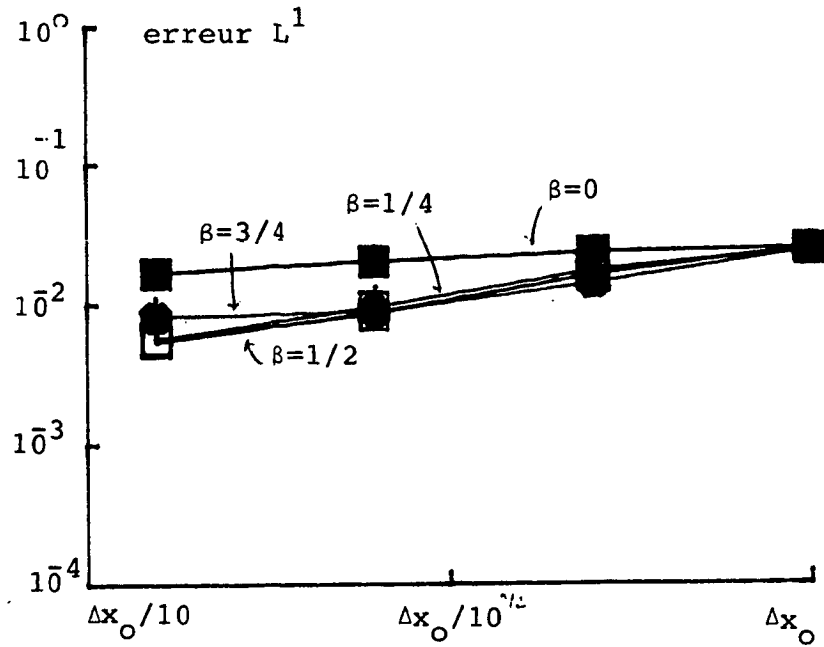


Figure 90 : Cas d'une onde de détente ( $CFL = 1/2$ ). Comparaison numérique des propriétés de convergence du schéma G-1/2 pour les différentes valeurs de l'exposant de contrôle d'antidiffusion,  $\beta$ .

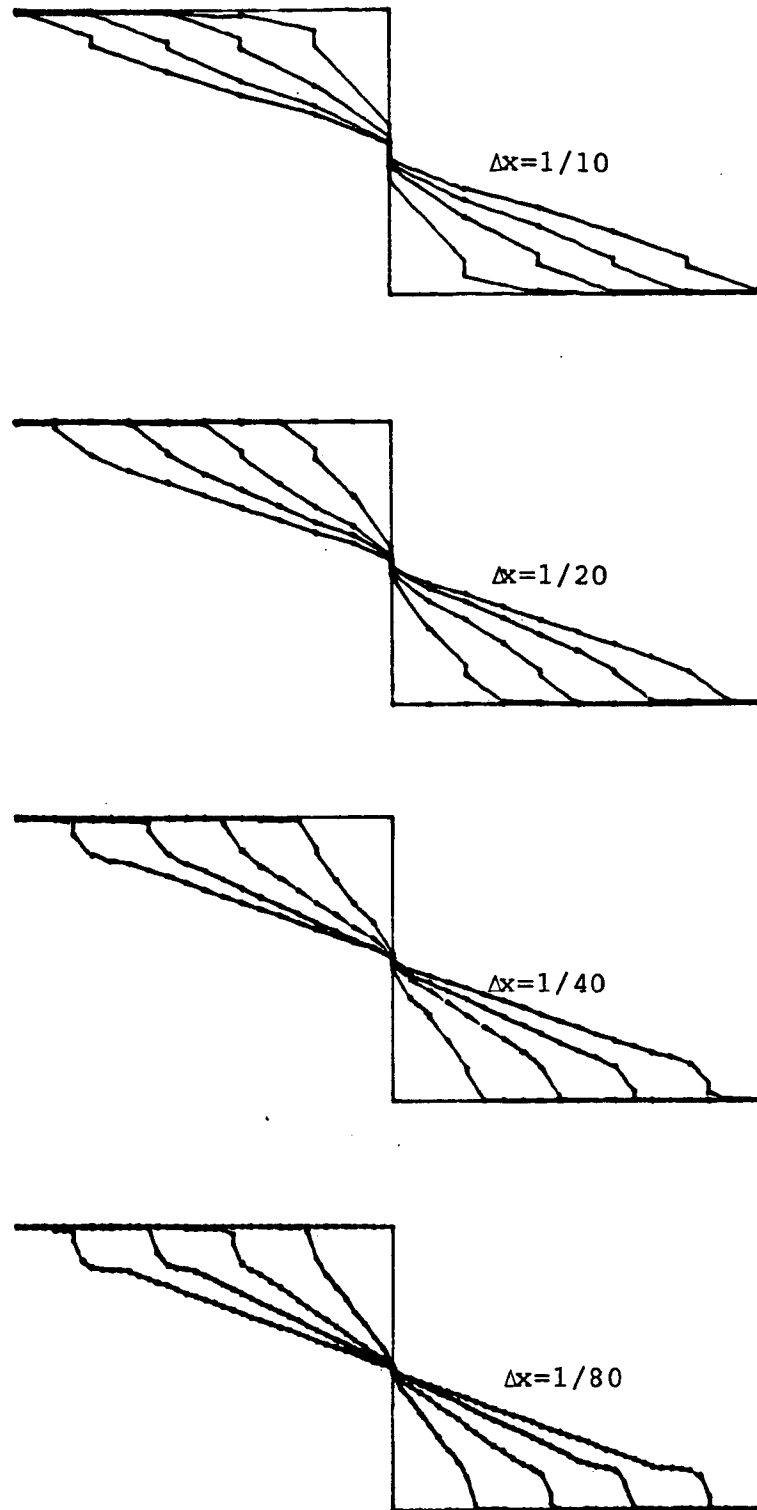


Figure 91 : Cas d'une onde de détente (CFL = 1/4). Exposant de contrôle d'antidiffusion = 0.

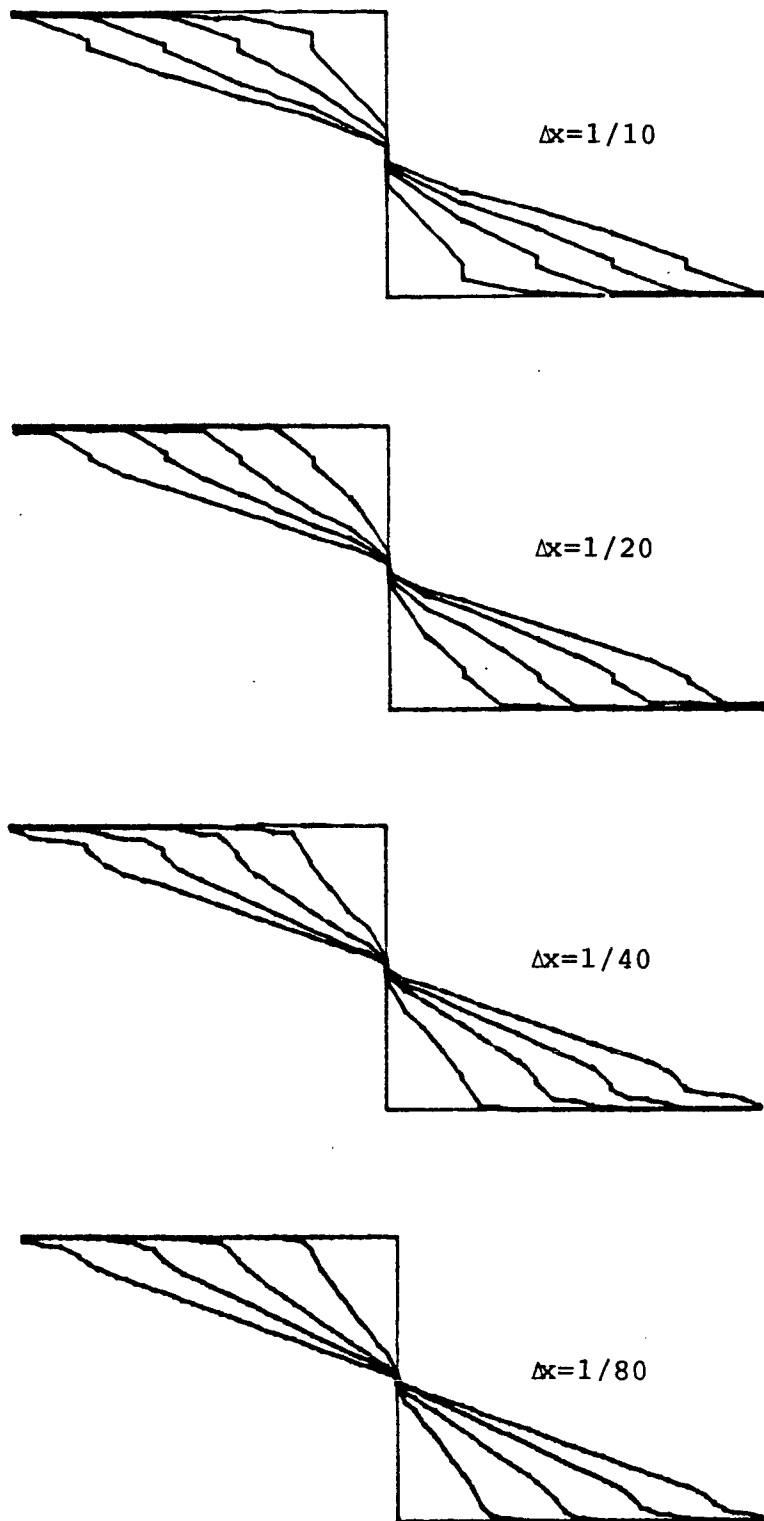


Figure 92 : Cas d'une onde de détente ( $CFL = 1/4$ ). Exposant de contrôle d'antidiffusion =  $1/4$ .

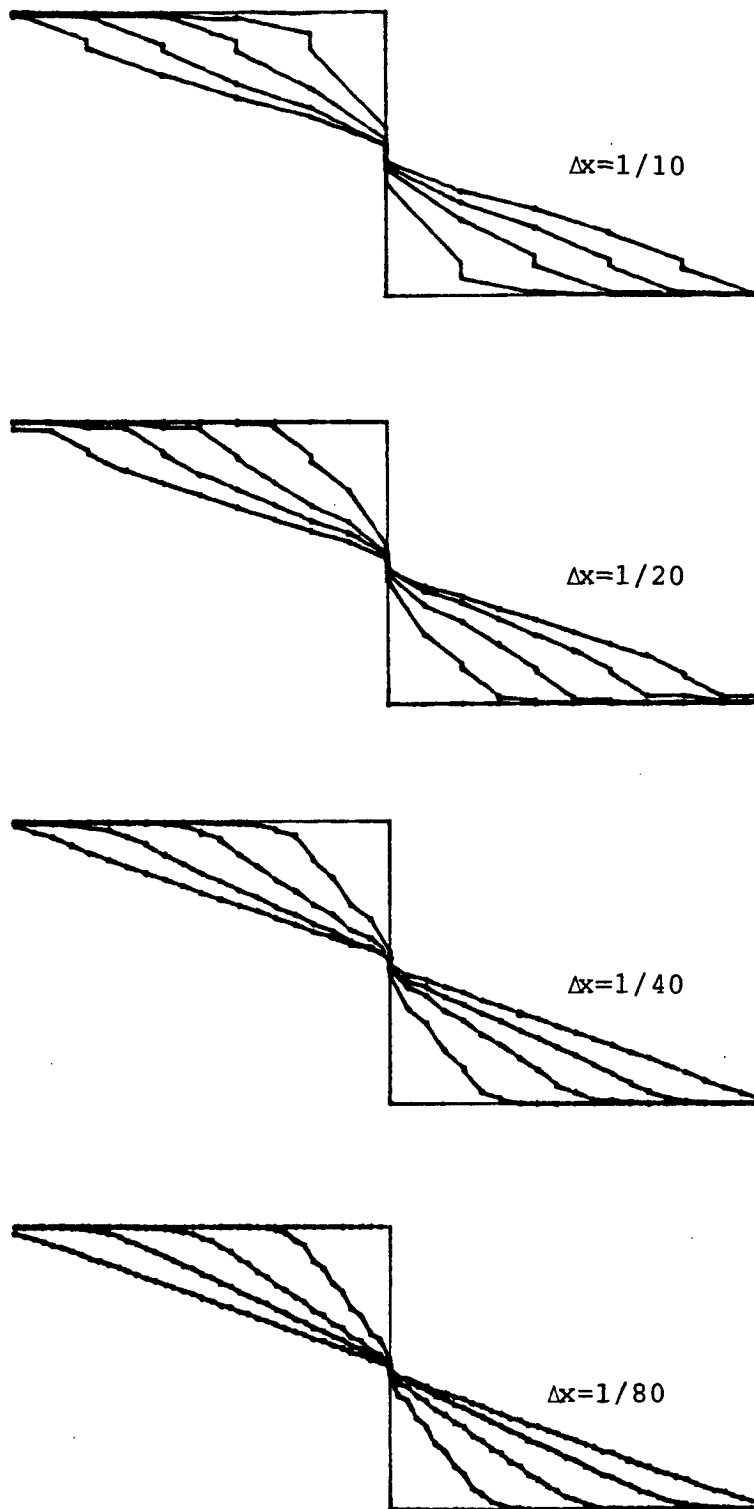


Figure 93 : Cas d'une onde de détente ( $CFL = 1/4$ ). Exposant de contrôle d'antidiffusion =  $1/2$ .

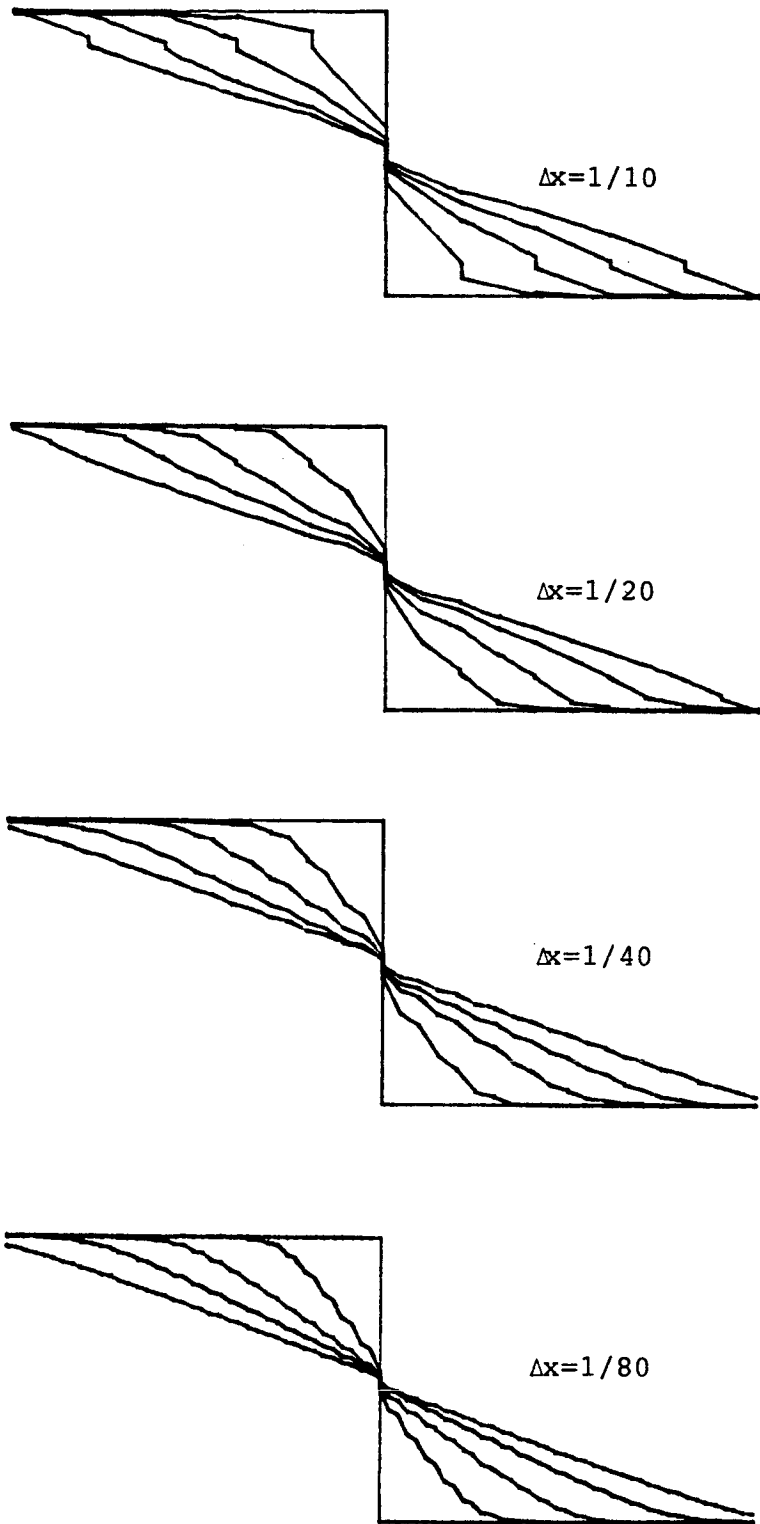


Figure 94 : Cas d'une onde de détente ( $CFL = 1/4$ ). Exposant de contrôle d'antidiffusion =  $3/4$ .

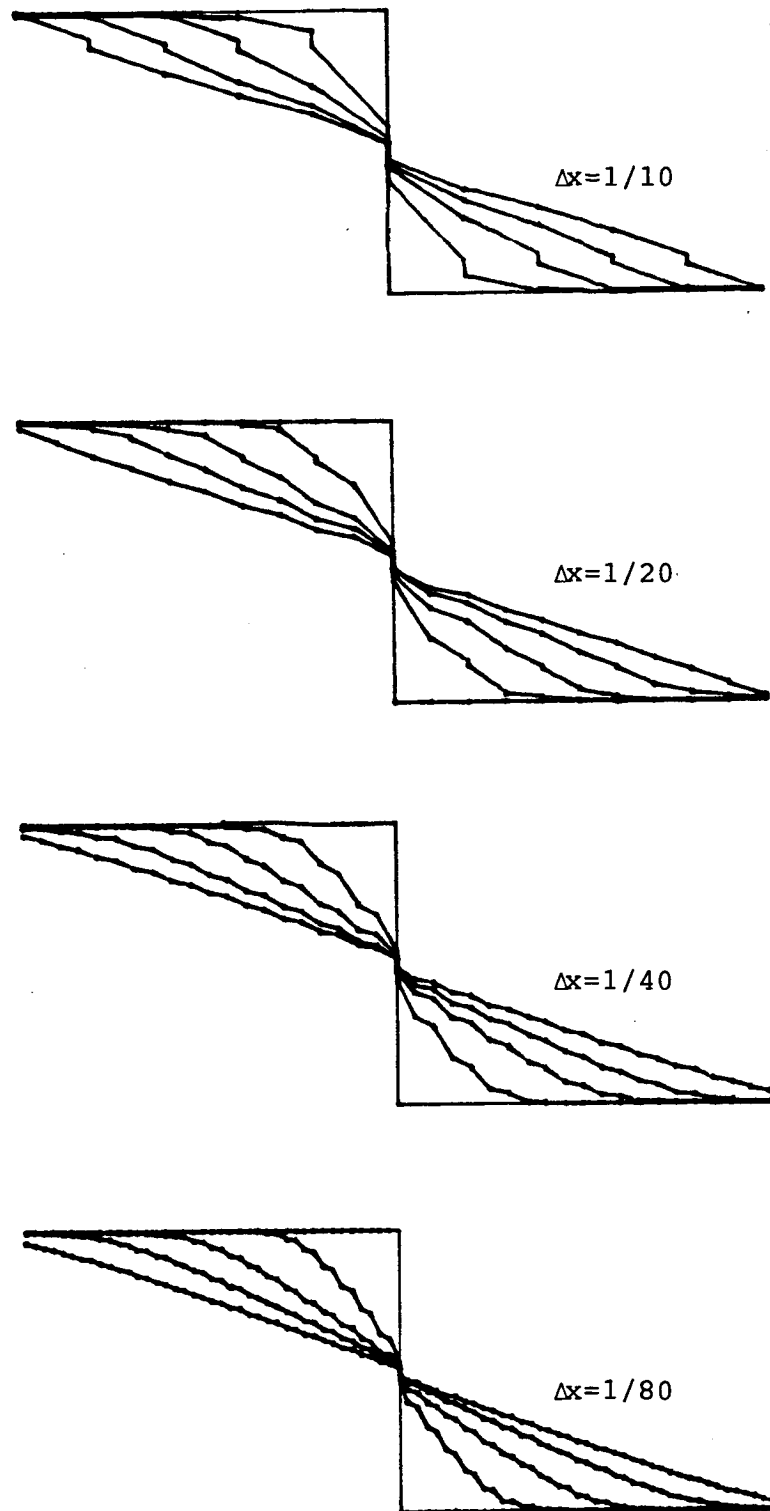


Figure 95 : Cas d'une onde de détente (CFL = 1/4). Exposant de contrôle d'antidiffusion = 1.



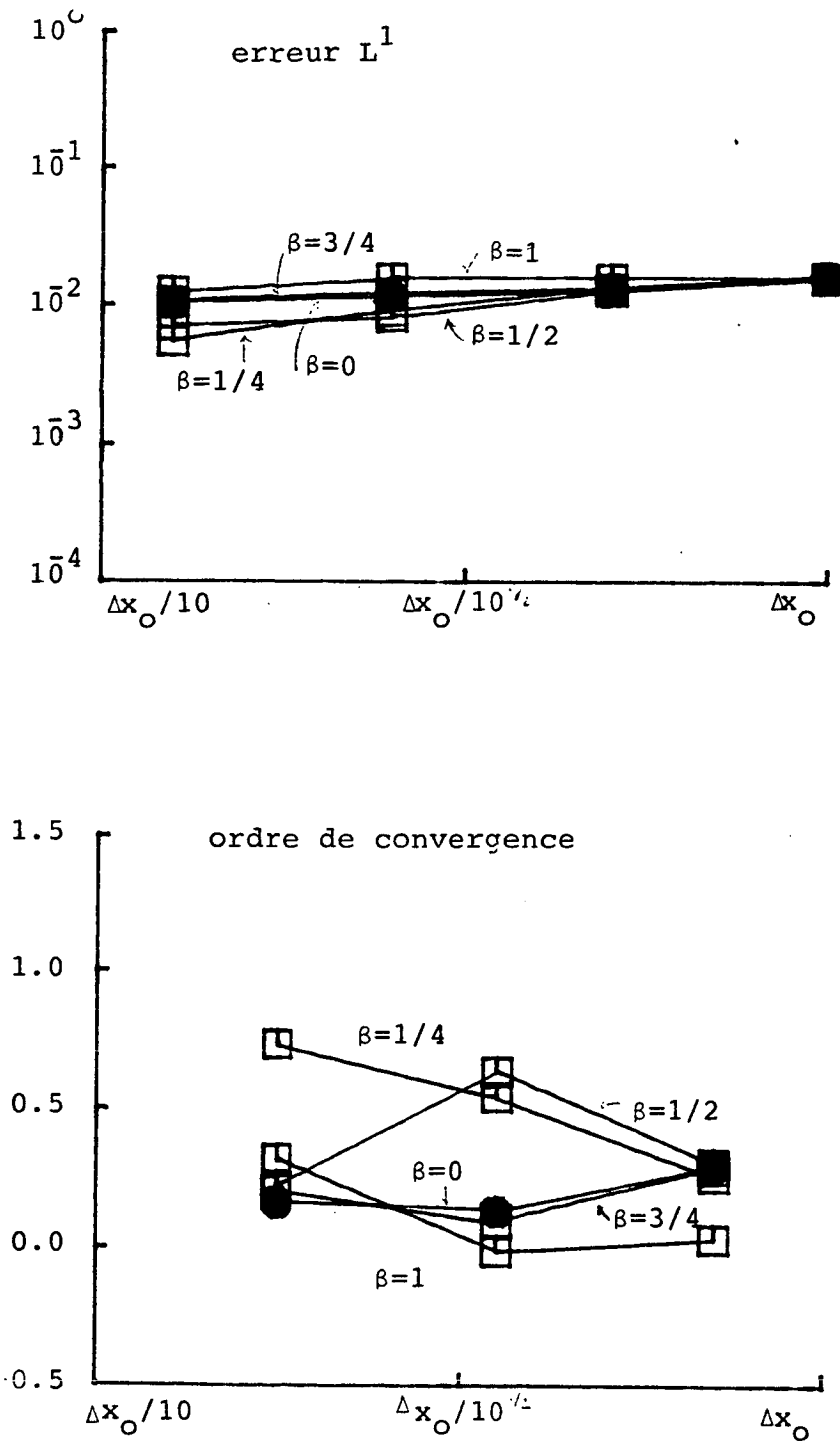


Figure 96 : Cas d'une onde de détente (CFL = 1/4). Comparaison numérique des propriétés de convergence du schéma G-1/2 pour les différentes valeurs de l'exposant de contrôle d'antidiffusion,  $\beta$ .

#### 4.6.5 - Conclusion

On a donc vu que pour le cas d'une discontinuité de contact et pour le cas d'un choc, le meilleur choix de  $\alpha$  est  $\alpha = 0$ .

Par contre, pour l'onde de détente, le meilleur choix est  $\alpha = 1/4$  et les résultats sont, avec ce choix, bien meilleurs que pour  $\alpha = 0$ .

Une comparaison des figures 48 et 90 et 54 et 96, montre qu'avec  $\alpha = 1/4$ , G-1/2 produit pratiquement les mêmes erreurs que BIP (l'ordre de convergence pour G-1/2 est supérieur pour  $h$  petit).

Il faut remarquer qu'on ne sait pas encore estimer le paramètre  $\alpha$  d'une façon optimale. En fait, on peut essayer de remplacer le terme  $c_0 h^\alpha$  par une fonction (non-linéaire!) de  $u_h$  avec laquelle G-1/2 produit de meilleures solutions. On est en train de travailler dans cette direction.

#### 4.7 - Conclusion

i) Dans ce chapitre, on a donc étudié les propriétés de convergence (et d'autres propriétés aussi) de quelques schémas numériques pour résoudre la loi de conservation scalaire non-linéaire.

Les schémas étudiés peuvent être considérés à mi-chemin entre le schéma de Godunov (linéaire, positif, monotone, d'ordre 1 et convergent sous la condition  $CFL \leq 1$ ) et le schéma LRG que l'on peut considérer comme la version d'ordre 2 du schéma de Godunov (il est aussi linéaire, mais il n'est pas positif ni monotone et *quand il converge*, il converge sous la condition  $CFL \leq c h^{1/2}$ ).

Il est bien connu qu'il n'existe pas de schémas linéaires d'ordre plus grand que 1 qui soient positifs et monotones. C'est pour ceci qu'on *doit* considérer des schémas vraiment non-linéaires (un schéma est vraiment non-linéaire quand la relation entre  $u_h^{n+1}$  et  $u_h^n$  n'est pas linéaire quand le schéma est appliqué au cas linéaire :  $f(u) = c u$ ,  $c \in \mathbb{R}$ ).

ii) Les schémas MUSCL, BIP, LRGP et G-1/2 ont été construits de façon à essayer d'améliorer les propriétés du schéma LRG et c'est pour cela qu'on les a appelés les schémas "LRG-modifiés". Ils sont tous des schémas *vraiment non-linéaires*, et les expériences numériques indiquent qu'ils ont les propriétés suivantes :

- ils sont tous positifs
- ils sont tous monotones (sauf LRGP et G-1/2 pour lesquels la monotonie est au niveau des moyennes!)
- ils sont stables sous la condition  $CFL \leq 1/2$ .  
(pour G-1/2 on peut démontrer ceci, voir [ 1 ]).

On voit donc que les quatre schémas ont des propriétés de positivité de monotonie et de stabilité bien meilleures que celles de LRG.

Il faut remarquer que du point de vue pratique c'est la condition de stabilité iii), celle qui fait que ces schémas sont préférables au schéma LRG. En effet, le nombre d'opérations nécessaire pour obtenir une approximation de la solution  $u(t)$  est :

$$w = O(N_x N_t)$$

où

$$N_t = t/\Delta t$$

$$N_x = L/h$$

avec une condition de stabilité du type  $CFL \leq c_0$  on a :

$$w = O(h^{-2})$$

mais, avec une condition de stabilité du type  $CFL \leq c_0 h^{1/2}$ , on a :

$$w = O(h^{-5/2})$$

le coût du schéma LRG est donc *nettement supérieur* à celui des autres schémas.

C'est clair, d'autre part, que le coût d'un schéma n'est pas une mesure de son efficacité. En effet, on doit plutôt regarder la fonction :

$$w \cdot e$$

où  $w$  est le coût et  $e$  est l'erreur produite par le schéma. Plus petite est  $w \cdot e$ , plus efficace est le schéma. D'après les expériences numériques, on a *presque toujours*

$$w \cdot e = O(h^{-1-\epsilon}) \quad 0 \leq \epsilon \leq 1/2$$

pour G, MUSCL, B1P, LRGP et G-1/2, et

$$w \cdot e = O(h^{-3/2})$$

en fait, pour le meilleur des schémas LRG-modifiés, on a  $\epsilon \approx 0$ , et par conséquent, il est donc préférable au schéma LRG !

iii) Les schémas MUSCL et B1P sont des schémas qu'on pourrait qualifier de "*diffusifs*" dans le sens suivant : *la diffusion qu'ils produisent augmente avec le temps.*

Par contre, les schémas LRGP et G-1/2 sont des schémas "*anti-diffusifs*" : *la diffusion qu'ils produisent diminue avec le temps.*

Ceci fait que MUSCL et B1P sont préférables lorsqu'on veut approximer une solution avec une région de régularité très grande, et que LRGP ou G-1/2 sont préférables pour approcher des solutions avec des fortes discontinuités.

Si, par "meilleur schéma", on comprend le schéma qui produit les erreurs les plus petites, on peut conclure d'après les expériences numériques faites que,

- i) pour CFL petit et  $h$  grand, le meilleur choix est le schéma LRGP. Ceci est très intéressant car du point de vue pratique, on n'est pas intéressés de faire  $h$  petit! D'autre part, ceci montre que LRGP est idéal pour être utilisé quand on perturbe notre équation par un terme de deuxième ordre : dans ce cas, le CFL doit être petit! (si le schéma est explicite);
- ii) pour  $h$  petit, le meilleur choix est G-1/2 quand il s'agit d'approximer des solutions avec de fortes discontinuités;
- iii) pour approcher des solutions avec des grandes régions de régularité, le meilleur choix est MUSCL pour CFL "grand" ou  $h$  petit. Dans les autres cas, le meilleur choix est entre B1P et G-1/2 (avec  $\alpha$  adéquatement choisi).

ANNEXE 1

Démonstration de la proposition 4

1 - Si l'on pose :

$$\psi_1(\theta, \mu) = -\frac{1}{2}((1 + 3\mu) + (3\mu - 1)e^{-i\theta})$$

$$\psi_2(\theta, \mu) = -6\mu(1 - e^{-i\theta})$$

alors, on peut écrire

$$\lambda_+(\theta, \mu) = 1 + (\psi_1(\theta, \mu) + (\psi_1^2(\theta, \mu) + \psi_2(\theta, \mu))^{1/2})\gamma$$

Si  $\mu = 0$ , il est facile de vérifier que  $\lambda_+(\theta) = 1$  (ceci implique que  $\alpha = 1$  dans ce cas!).

2 - Considérons maintenant le cas  $\mu \neq 0$ . Au voisinage de  $\theta = 0$ , on peut écrire :

$$\lambda_1(\theta, \mu) = a_0 + a_1\theta + a_2\theta^2 + a_3\theta^3 = a_4\theta^4 + o(\theta^5)$$

$$\lambda_2(\theta, \mu) = b_1\theta + b_2\theta^2 + b_3\theta^3 + b_4\theta^4 + o(\theta^5)$$

où

$$a_0 = -3\mu$$

$$a_1 = \frac{1}{2}(3\mu - 1)i$$

$$a_2 = \frac{1}{4}(3\mu - 1)$$

$$a_3 = -\frac{1}{12}(3\mu - 1)i$$

$$a_4 = -\frac{1}{48}(3\mu - 1)$$

$$b_1 = -6\mu i$$

$$b_2 = -3\mu$$

$$b_3 = \mu i$$

$$b_4 = \frac{1}{4}\mu.$$

Etant donné que  $\mu \neq 0$ ,  $a_0 \neq 0$ , on peut donc écrire :

$$\begin{aligned}\lambda_+ &= 1 + \gamma a_0 [(\psi_1/a_0) - ((\psi_1/a_0)^2 + (\psi_2/a_0^2))^{1/2}] \\ &= 1 + \gamma a_0 \Theta\end{aligned}$$

et  $(\psi_1/a_0) = 1 + \alpha_1\theta + \alpha_2\theta^2 + \alpha_3\theta^3 + \alpha_4\theta^4 + o(\theta^5)$

$$(\psi_2/a_0) = \beta_1\theta + \beta_2\theta^2 + \beta_3\theta^3 + \beta_4\theta^4 + o(\theta^5)$$

où  $\alpha_1 = (\frac{1}{6\mu} - \frac{1}{2})i$   $\beta_1 = -\frac{2}{3\mu}i$

$$\alpha_2 = (\frac{1}{12\mu} - \frac{1}{4})$$

$$\beta_2 = -\frac{1}{3\mu}$$

$$\alpha_3 = (-\frac{1}{36\mu} + \frac{1}{12})i$$

$$\beta_3 = \frac{i}{9\mu}$$

$$\alpha_4 = (-\frac{1}{144\mu} + \frac{1}{48})$$

$$\beta_4 = \frac{1}{36\mu}$$

3 - On va maintenant évaluer  $\Theta$

$$\begin{aligned}(\psi_1/a_0)^2 &= 1 + (2\alpha_1)\theta + (2\alpha_2 + \alpha_1^2)\theta^2 + (2\alpha_3 + 2\alpha_1\alpha_2)\theta^3 \\ &\quad + (2\alpha_4 + 2\alpha_1\alpha_3 + \alpha_2^2)\theta^4 + o(\theta^5)\end{aligned}$$

et donc

$$\begin{aligned}(\psi_1/a_0)^2 + (\psi_2/a_0^2) &= 1 + (2\alpha_1 + \beta_1)\theta + (2\alpha_2 + \alpha_1^2 + \beta_2)\theta^2 \\ &\quad + (2\alpha_3 + 2\alpha_1\alpha_2 + \beta_3)\theta^3 \\ &\quad + (2\alpha_4 + 2\alpha_1\alpha_3 + \alpha_2^2 + \beta_4)\theta^4 + o(\theta^5) \\ &= 1 + \gamma_1\theta + \gamma_2\theta^2 + \gamma_3\theta^3 + \gamma_4\theta^4 + o(\theta^5) \\ &= 1 + \epsilon.\end{aligned}$$

D'autre part, on sait que l'on a :

$$(1+\epsilon)^{1/2} = 1 + \frac{1}{2}\epsilon - \frac{1}{8}\epsilon^2 + \frac{1}{16}\epsilon^3 - \frac{5}{128}\epsilon^4 + o(\epsilon^5)$$

par conséquent,

$$\begin{aligned} & ((\psi_1/a_0)^2 + (\psi_2/a_0^2))^{1/2} = \\ & = 1 + \frac{1}{2}(\gamma_1\theta + \gamma_2\theta^2 + \gamma_3\theta^3 + \gamma_4\theta^4) \\ & - \frac{1}{8}(\gamma_1^2\theta^2 + 2\gamma_1\gamma_2\theta^3 + (2\gamma_1\gamma_3 + \gamma_2^2)\theta^4) + \\ & + \frac{1}{16}(\gamma_1^3\theta^3 + 3\gamma_1^2\gamma_2\theta^4) \\ & - \frac{5}{128}(\gamma_1^4\theta^4) + o(\theta^5) \\ & = 1 + \theta(\frac{1}{2}\gamma_1) + \theta^2(\frac{1}{2}\gamma_2 - \frac{1}{8}\gamma_1^2) + \theta^3(\frac{1}{2}\gamma_3 - \frac{1}{4}\gamma_1\gamma_2 + \frac{1}{16}\gamma_1^3) \\ & + \theta^4(\frac{1}{2}\gamma_4 - \frac{1}{8}(2\gamma_1\gamma_3 + \gamma_2^2) + \frac{3}{16}\gamma_1^2\gamma_2 - \frac{5}{128}\gamma_1^4) + o(\theta^5) \end{aligned}$$

et finalement

$$\begin{aligned} \theta & = \theta(\alpha_1 - \frac{1}{2}\gamma_1) + \theta^2(\alpha_2 - \frac{1}{2}\gamma_2 + \frac{1}{8}\gamma_1^2) \\ & + \theta^3(\alpha_3 - \frac{1}{2}\gamma_3 + \frac{1}{4}\gamma_1\gamma_2 - \frac{1}{16}\gamma_1^3) \\ & + \theta^4(\alpha_4 - \frac{1}{2}\gamma_4 + \frac{1}{8}(2\gamma_1\gamma_3 + \gamma_2^2) - \frac{3}{16}\gamma_1^2\gamma_2 + \frac{5}{128}\gamma_1^4) + o(\theta^5) \\ & = \delta_1\theta + \delta_2\theta^2 + \delta_3\theta^3 + \delta_4\theta^4 + o(\theta^5) . \end{aligned}$$

Ensuite,



$$\delta_1 = \alpha_1 - \frac{1}{2} \gamma_1$$

$$= \alpha_1 - \frac{1}{2}(2\alpha_1 + \beta_1)$$

$$= -\frac{1}{2} \beta_1$$

$$\delta_2 = \alpha_2 - \frac{1}{2} \gamma_2 + \frac{1}{8} \gamma_1^2$$

$$= \alpha_2 - \frac{1}{2}(2\alpha_2 + \alpha_1^2 + \beta_2) + \frac{1}{8}(4\alpha_1^2 + 4\alpha_1\beta_1 + \beta_1^2)$$

$$= -\frac{1}{2} \beta_2 + \frac{1}{2} \alpha_1\beta_1 + \frac{1}{8} \beta_1^2$$

$$\delta_3 = \alpha_3 - \frac{1}{2}(2\alpha_3 + 2\alpha_1\alpha_2 + \beta_3) + \frac{1}{4}(2\alpha_2 + \alpha_1^2 + \beta_2)(2\alpha_1 + \beta_1)$$

$$- \frac{1}{16}(8\alpha_1^3 + 12\alpha_1^2\beta_1 + 6\alpha_1\beta_1^2 + \beta_1^3)$$

$$= -\alpha_1\alpha_2 - \frac{1}{2} \beta_3 + \alpha_2\alpha_1 + \frac{1}{2} \alpha_2\beta_1 + \frac{1}{2} \alpha_1^3 + \frac{1}{4} \alpha_1^2\beta_1$$

$$+ \frac{1}{2} \beta_2\alpha_1 + \frac{1}{4} \beta_1\beta_2 - \frac{1}{2} \alpha_1^3 - \frac{3}{4} \alpha_1^2\beta_1 - \frac{3}{8} \alpha_1\beta_1^2 - \frac{1}{16} \beta_1^3$$

$$= -\frac{1}{2} \beta_3 + \frac{1}{2} \alpha_2\beta_1 + \frac{1}{2} \alpha_1\beta_2 + \frac{1}{4} \beta_1\beta_2$$

$$- \frac{1}{2} \alpha_1^2\beta_1 - \frac{3}{8} \alpha_1\beta_1^2 - \frac{1}{16} \beta_1^3$$

$$\delta_4 = \alpha_4 - \frac{1}{2}(2\alpha_4 + 2\alpha_1\alpha_3 + \alpha_2^2 + \beta_4) + \frac{1}{4}(2\alpha_1 + \beta_1)(2\alpha_3 + 2\alpha_1\alpha_2 + \beta_3)$$

$$+ \frac{1}{8}(4\alpha_2^2 + \alpha_1^4 + \beta_2^2 + 4\alpha_2\alpha_1^2 + 4\alpha_2\beta_2 + 2\alpha_1^2\beta_2)$$

$$- \frac{3}{16}(4\alpha_1^2 + 4\alpha_1\beta_1 + \beta_1^2)(2\alpha_2 + \alpha_1^2 + \beta_2)$$

$$+ \frac{5}{128}(16\alpha_1^4 + 32\alpha_1^3\beta_1 + 24\alpha_1^2\beta_1^2 + 8\alpha_1\beta_1^3 + \beta_1^4)$$

$$\begin{aligned}
 &= -\alpha_1\alpha_3 - \frac{1}{2}\alpha_2^2 - \frac{1}{2}\beta_4 + \alpha_1\alpha_3 + \alpha_1^2\alpha_2 + \frac{1}{2}\alpha_1\beta_3 + \\
 &+ \frac{1}{2}\beta_1\alpha_3 + \frac{1}{2}\beta_1\alpha_1\alpha_2 + \frac{1}{4}\beta_1\beta_3 + \frac{1}{2}\alpha_2^2 + \frac{1}{8}\alpha_1^4 + \\
 &+ \frac{1}{8}\beta_2^2 + \frac{1}{2}\alpha_2\alpha_1^2 + \frac{1}{2}\alpha_2\beta_2 + \frac{1}{4}\alpha_1^2\beta_2 - \frac{3}{2}\alpha_1^2\alpha_2 - \frac{3}{4}\alpha_1^4 - \\
 &- \frac{3}{4}\alpha_1^2\beta_2 - \frac{3}{2}\alpha_1\alpha_2\beta_1 - \frac{3}{4}\alpha_1^3\beta_1 - \frac{3}{4}\alpha_1\beta_1\beta_2 - \frac{3}{8}\beta_1^2\alpha_2 - \\
 &- \frac{3}{16}\beta_1^2\alpha_1^2 - \frac{3}{16}\beta_1^2\beta_2 + \frac{5}{8}\alpha_1^4 + \frac{5}{4}\alpha_1^3\beta_1 + \frac{15}{16}\alpha_1^2\beta_1^2 + \\
 &+ \frac{5}{16}\alpha_1\beta_1^3 + \frac{5}{128}\beta_1^4 \\
 &= -\frac{1}{2}\beta_4 + \frac{1}{2}\alpha_1\beta_3 + \frac{1}{2}\alpha_2\beta_2 + \frac{1}{2}\alpha_3\beta_1 + \frac{1}{4}\beta_1\beta_3 + \frac{1}{8}\beta_2^2 - \\
 &- \beta_1\alpha_1\alpha_2 - \frac{1}{2}\beta_2\alpha_1^2 - \frac{3}{4}\beta_1\beta_2\alpha_1 - \frac{3}{8}\beta_1^2\alpha_2 - \frac{3}{16}\beta_1^2\beta_2 \\
 &+ \frac{1}{2}\alpha_1^3\beta_1 + \frac{3}{4}\alpha_1^2\beta_1^2 + \frac{5}{16}\alpha_1\beta_1^3 + \frac{5}{128}\beta_1^4.
 \end{aligned}$$

En reportant les expressions des  $\alpha_i$  et des  $\beta_i$  dans les expressions donnant les  $\delta_i$  on obtient :

$$\delta_1 = -\frac{1}{2}\left(-\frac{2}{3\mu}\right)i = \frac{1}{3\mu} i$$

$$\begin{aligned}
 \delta_2 &= \frac{1}{2}\left(\frac{1}{3\mu}\right) + \frac{1}{2}\left(\frac{1}{6\mu} - \frac{1}{2}\right)i \cdot \left(-\frac{2i}{3\mu}\right) + \frac{1}{8}\left(\frac{2}{3\mu} i\right)^2 \\
 &= \frac{1}{6\mu} + \frac{1}{18\mu^2} - \frac{1}{6\mu} - \frac{1}{18\mu^2} = 0
 \end{aligned}$$

$$\text{Re } \delta_3 = 0$$

$$\begin{aligned}
 \delta_4 &= -\frac{1}{2}\left(\frac{1}{36\mu}\right) + \frac{1}{2}\left(\frac{1}{6\mu} - \frac{1}{2}\right)i \cdot \left(\frac{i}{9\mu}\right) \\
 &+ \frac{1}{2}\left(\frac{1}{12\mu} - \frac{1}{4}\right)\left(\frac{-1}{3\mu}\right) + \frac{1}{2}\left(-\frac{1}{36\mu} + \frac{1}{12}\right)i \cdot \left(-\frac{2i}{3\mu}\right) + \\
 &+ \frac{1}{4}\left(-\frac{2i}{3\mu}\right)\left(\frac{i}{9\mu}\right) + \frac{1}{8}\left(\frac{1}{9\mu^2} - \left(\frac{-2i}{3\mu}\right)\left(\frac{1}{6\mu} - \frac{1}{2}\right)i\right. \\
 &\quad \cdot \left.\left(\frac{1}{12\mu} - \frac{1}{4}\right) - \frac{1}{2}\left(\frac{-1}{3\mu}\right)\left(-\frac{1}{36\mu^2} + \frac{1}{6\mu} - \frac{1}{4}\right)\right. \\
 &\quad \left.- \frac{3}{4}\left(\frac{-2i}{3\mu}\right)\left(\frac{-1}{3\mu}\right)\left(\frac{1}{6\mu} - \frac{1}{2}\right)i - \frac{3}{8}\left(\frac{-4}{9\mu^2}\right)\left(\frac{1}{12\mu} - \frac{1}{4}\right)\right. \\
 &\quad \left.- \frac{3}{16}\left(\frac{-4}{9\mu^2}\right)\left(\frac{-1}{3\mu}\right) + \frac{1}{2}(-i)\left(\frac{1}{216\mu^3} - \frac{1}{24\mu^2} + \frac{1}{8\mu} - \frac{1}{8}\right)\right. \\
 &\quad \left.\cdot \left(\frac{-2i}{3\mu}\right) + \frac{3}{4}\left(-\frac{1}{36\mu^2} + \frac{1}{6\mu} - \frac{1}{4}\right)\left(\frac{-4}{9\mu^2}\right)\right) \\
 &+ \frac{5}{16}\left(\frac{1}{6\mu} - \frac{1}{2}\right)i \left(\frac{8i}{27\mu^3}\right) + \frac{5}{128}\left(\frac{16}{81\mu^2}\right) \\
 \delta_4 &= \frac{1}{\mu} \left\{ -\frac{1}{72} + \frac{1}{36} + \frac{1}{24} + \frac{1}{36} - \frac{1}{12} - \frac{1}{24} + \frac{1}{24} \right\} \\
 &+ \frac{1}{\mu^2} \left\{ -\frac{1}{108} - \frac{1}{72} - \frac{1}{108} + \frac{1}{54} + \frac{1}{72} + \frac{1}{18} + \frac{1}{36} - \right. \\
 &\quad \left. - \frac{1}{12} - \frac{1}{24} - \frac{1}{24} + \frac{1}{12} \right\} \\
 &+ \frac{1}{\mu^3} \left\{ -\frac{1}{108} - \frac{1}{216} + \frac{1}{36} + \frac{1}{72} - \frac{1}{36} + \frac{1}{72} - \frac{1}{18} + \frac{5}{108} \right\} + \\
 &+ \frac{1}{\mu^4} \left\{ -\frac{1}{648} + \frac{1}{108} - \frac{5}{324} + \frac{5}{648} \right\} \\
 &= \frac{1}{216} \mu^{-3} .
 \end{aligned}$$

Alors,  $\operatorname{Re}\theta = \frac{1}{216} \mu^{-3} \theta^4 + o(\theta^6)$

$$\operatorname{Im}\theta = \frac{1}{3\mu} \theta + o(\theta^3)$$

Et finalement, en revenant à l'expression de  $\lambda_+$ ,

$$\operatorname{Re} \lambda_+(\theta, \mu) = 1 - \frac{1}{72\mu^2} \gamma \theta^4 + o(\theta^6)$$

$$\operatorname{Im} \lambda_+(\theta, \mu) = -\gamma \theta + o(\theta^3)$$

d'où le résultat.

REFERENCES

- [1] B. Cockburn, "Le schéma  $G-k/2$  pour les lois de conservation scalaires", Congrès National d'Analyse Numérique, Bombannes, 21-25 mai 1984, pp. 53-56; INRIA rapport n° 354, 1985.
- [2] G. Chavent and G. Salzano, "A finite element method for the 1D water flooding problem with gravity", J. Comp. Physics, 45, 1982, pp. 307-344.
- [3] T. Geveci, "The significance of the stability of difference schemes in different  $l^p$ -spaces", SIAM Review, 24, n° 4, october 1982, pp. 413-426.
- [4] S. Osher, "Convergence of generalized MUSCL schemes", SIAM Journal of Numer. Anal., 22, 1985, pp. 947-861.
- [5] R.D. Richtmyer and K.W. Morton, "Difference methods for Initial value problems", John Wiley, New York, 1967.

Imprimé en France

par

l'Institut National de Recherche en Informatique et en Automatique

