

2019/2020

MASTER 2 RECHERCHE FONDAMENTALE EN MATHÉMATIQUES

---

# Stabilité de schémas différences finies avec conditions de bord Lax-Wendroff Inverse

---

*Encadrant* : Benjamin BOUTIN

Rédigé par Pierre LE BARBENCHON



# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Schéma semi-discret . . . . .	2
1.2	Schéma discret . . . . .	2
1.3	Points fantômes . . . . .	3
1.4	Méthode Lax Wendroff Inverse . . . . .	4
<b>2</b>	<b>Stabilité</b>	<b>5</b>
2.1	Étude de la matrice Quasi-Toeplitz . . . . .	6
2.2	Théorie GKS . . . . .	14
2.2.1	Stabilité du problème de Cauchy sur $\mathbb{Z}$ . . . . .	14
2.2.2	Stabilité du problème avec une seule condition de bord . . . . .	16
<b>3</b>	<b>Consistance</b>	<b>17</b>
3.1	Étude de l'intérieur . . . . .	17
3.2	Étude du bord . . . . .	18
<b>4</b>	<b>Convergence</b>	<b>20</b>
<b>5</b>	<b>Conclusion</b>	<b>21</b>
<b>6</b>	<b>Annexes</b>	<b>21</b>
6.1	Programmes Python pour tracer les solutions approchées par Upwind . . . . .	21
6.2	Programmes Python pour tracer les solutions approchées par Beam-Warming . . . . .	23

## 1 Introduction

Nous allons étudier la convergence de schémas numériques discret et semi-discret avec condition de bords pour l'équation de transport

$$\begin{cases} \partial_t u + a \partial_x u = 0 \\ u(0, x) = f(x) \\ u(t, x_A) = g(t) \end{cases} \quad (1)$$

où  $t \in \mathbb{R}^+$ ,  $x \in [x_A, 1]$ ,  $a \in \mathbb{R}^+$ ,  $f : [x_A, 1] \rightarrow \mathbb{R}^d$ ,  $x_A$  proche de 0,  $g : \mathbb{R}^+ \rightarrow \mathbb{R}^d$  et on cherche la solution  $u : \mathbb{R}^+ \times [0, 1] \rightarrow \mathbb{R}^d$  qui vérifie (1).

Le but est d'approcher une solution de cette équation en discrétisant soit seulement l'espace (semi-discret), soit l'espace et le temps (discret). Les schémas que l'on considèrera seront des schémas avec un seul pas en temps. La théorie pour les schémas avec pas multiple en temps est développée dans les livres [Cou11] et [Str04].

La discrétisation en espace s'écrit en posant  $J \in \mathbb{N}^*$ ,  $\Delta x = \frac{1}{J+1}$  et  $x_j = j\Delta x$  pour  $j \in \llbracket 0; J \rrbracket$ . De plus, on écrit  $x_A = \sigma\Delta x$  avec  $\sigma \in ]-1, 1[$ .

La discrétisation en temps s'écrit en posant  $\Delta t \in ]0, 1[$  et  $t^n = n\Delta t$  pour  $n \in \mathbb{N}$ .

## 1.1 Schéma semi-discret

Les schémas numériques semi-discrets sont de la forme

$$\begin{cases} \frac{d}{dt}U_j(t) = QU_j(t), & t \geq 0, j \in \llbracket 0; J \rrbracket \\ B_0U(t) = g(t), & t \geq 0 \\ U_j(0) = f(x_j), & j \in \llbracket 0; J \rrbracket \end{cases} \quad (2)$$

où les opérateurs  $Q$  et  $B_0$  sont des combinaisons linéaires d'opérateurs de translations en espace et tels que l'opérateur  $B_0$  ne dépend pas de  $J$ .

**Exemple 1.** Un schéma Upwind avec conditions de bord

$$\begin{cases} \frac{d}{dt}U_j(t) = -a \frac{U_j(t) - U_{j-1}(t)}{\Delta x}, & t \geq 0, j \in \llbracket 0; J \rrbracket, \\ U_{-1}(t) + (1 + \sigma)(U_1(t) - U_0(t)) = g(t), & t \geq 0, \\ U_j(0) = f(x_j), & j \in \llbracket 0; J \rrbracket. \end{cases} \quad (3)$$

## 1.2 Schéma discret

Les schémas numériques discrets sont de la forme

$$\begin{cases} U_j^{n+1} = QU_j^n, & n \in \mathbb{N}, j \in \llbracket 0; J \rrbracket \\ B_0U^n = g(t^n), & t \in \mathbb{N} \\ U_j(0) = f(x_j), & j \in \llbracket 0; J \rrbracket \end{cases} \quad (4)$$

où les opérateurs  $Q$  et  $B_0$  sont des combinaisons linéaires d'opérateurs de translations en espace et tels que l'opérateur  $B_0$  ne dépend pas de  $J$ .

**Exemple A 2.** Le même schéma Upwind avec conditions de bord en utilisant un schéma Euler explicite pour discrétiser le temps

$$\begin{cases} U_j^{n+1} = U_j^n - \frac{a\Delta t}{\Delta x}(U_j^n - U_{j-1}^n), & n \in \mathbb{N}, j \in \llbracket 0; J \rrbracket, \\ U_{-1}^n + (1 + \sigma)(U_1^n - U_0^n) = g(t^n), & n \in \mathbb{N}, \\ U_j^0 = f(x_j), & j \in \llbracket 0; J \rrbracket. \end{cases} \quad (\text{Upw})$$

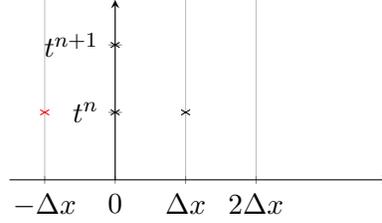
**Remarque 3.** Certains schémas discrets ne proviennent pas d'une discrétisation en temps d'un schéma semi-discret. En effet, ces schémas mêlent les deux discrétisations en temps et en espace.

**Exemple B 4.** En posant  $\lambda = \frac{a\Delta t}{\Delta x}$ , on peut définir le schéma de Beam-Warming de la manière suivante

$$\begin{cases} U_j^{n+1} = \frac{(\lambda - 1)(\lambda - 2)}{2}U_j^n + \lambda(2 - \lambda)U_{j-1}^n + \frac{\lambda(\lambda - 1)}{2}U_{j-2}^n, & n \in \mathbb{N}, j \in \llbracket 0; J \rrbracket, \\ U_{-1}^n + (1 + \sigma)(U_1^n - U_0^n) = g(t^n), & n \in \mathbb{N}, \\ U_{-2}^n + (2 + \sigma)(U_1^n - U_0^n) = g(t^n), & n \in \mathbb{N}, \\ U_j^0 = f(x_j), & j \in \llbracket 0; J \rrbracket. \end{cases} \quad (\text{BW1})$$

### 1.3 Points fantômes

On appelle *points fantômes* des points en dehors du domaine de définition auxquels on voudra assigner des valeurs.

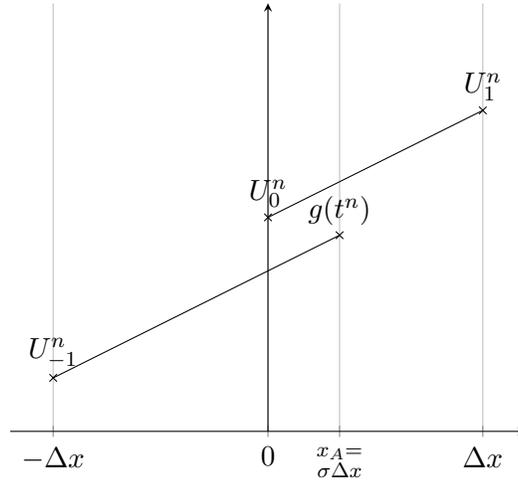


**Remarque 5.** Les points fantômes sont nécessaires afin d'appliquer un schéma jusqu'au bord sans avoir à modifier le schéma. En effet, par exemple, pour un schéma  $U_j^{n+1} = \mathcal{F}(U_{j-1}^n, U_j^n, U_{j+1}^n)$ , afin de calculer  $U_0^{n+1}$ , il est nécessaire d'avoir un point  $U_{-1}^n$  appelé *point fantôme*. C'est pourquoi dans la définition du schéma (3) (resp. (Upw)) on voit apparaître un  $U_{-1}(t)$  (resp.  $U_{-1}^n$ ).

Le but est de trouver des points fantômes qui permettent d'avoir de la stabilité et de la consistance au bord afin d'obtenir la convergence du schéma complet.

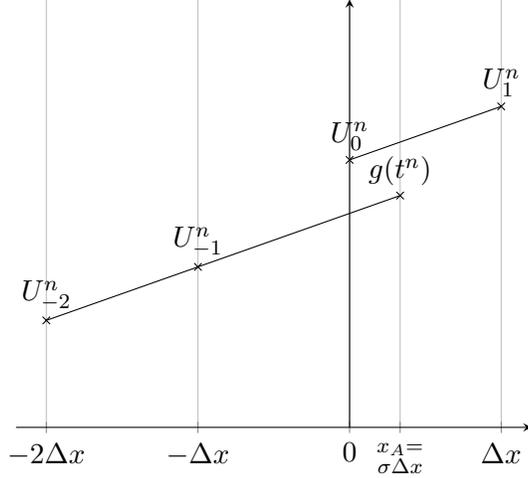
**Exemple A 6.** Pour le schéma (Upw), le point fantôme a été défini en posant

$$\frac{g(t^n) - U_{-1}^n}{(1 + \sigma)\Delta x} = \frac{U_1^n - U_0^n}{\Delta x}$$



**Exemple B 7.** Pour le schéma (BW1), les points fantômes ont été définis en posant

$$\frac{g(t^n) - U_{-1}^n}{(1 + \sigma)\Delta x} = \frac{U_1^n - U_0^n}{\Delta x} \quad \text{et} \quad \frac{g(t^n) - U_{-2}^n}{(2 + \sigma)\Delta x} = \frac{U_1^n - U_0^n}{\Delta x}$$



#### 1.4 Méthode Lax Wendroff Inverse

On peut définir des points fantômes à l'aide de la méthode Lax Wendroff Inverse (ILW) présentée dans l'article [VS15].

**Exemple C 8.** On utilise le schéma de Beam Warming (BW1) en modifiant les points fantômes en des points fantômes ILW à l'ordre 3.

$$\begin{cases} U_{-1}^n = g(t^n) + (1 + \sigma) \frac{\Delta x g'(t^n)}{a} + \frac{(1 + \sigma)^2 \Delta x^2 g''(t^n)}{2} \\ U_{-2}^n = g(t^n) + (2 + \sigma) \frac{\Delta x g'(t^n)}{a} + \frac{(2 + \sigma)^2 \Delta x^2 g''(t^n)}{2} \end{cases} \quad (\text{BW2})$$

Parfois le coût en calcul des points ILW à des ordres élevés est trop important et on préfère remplacer les derniers termes par des termes d'extrapolation. Cette méthode est appelée méthode de Lax Wendroff Inverse simplifiée.

**Exemple D 9.** On utilise le schéma de Beam Warming (BW1) en modifiant les points fantômes en des points fantômes ILW simplifié à l'ordre 3 en n'utilisant que deux termes ILW. En utilisant les mêmes notations que [VS15], on pose  $d = 3$  et  $k_d = 2$ .

$$\begin{cases} U_{-1}^n = g(t^n) + (1 + \sigma) \frac{\Delta x g'(t^n)}{a} + \frac{(1 + \sigma)^2 \Delta x^2}{2} \frac{U_2^n - 2U_1^n + U_0^n}{\Delta x^2} \\ U_{-2}^n = g(t^n) + (2 + \sigma) \frac{\Delta x g'(t^n)}{a} + \frac{(2 + \sigma)^2 \Delta x^2}{2} \frac{U_2^n - 2U_1^n + U_0^n}{\Delta x^2} \end{cases} \quad (\text{BW3})$$

Le troisième terme provient de la dérivée seconde du polynôme d'interpolation d'ordre 2 au point 0.

## 2 Stabilité

Pour étudier la stabilité des différents schémas, on introduit la norme  $\ell_{\Delta x}^2(I)$ , pour  $I$  un ensemble d'indices, définie de la manière suivante : pour toute suite  $(u_k)_{k \in I}$ , on pose

$$\|u\|_{\Delta x} = \left( \sum_{k \in I} \Delta x |u_k|^2 \right)^{1/2}.$$

En général, on prendra  $I = \mathbb{Z}$ ,  $I = \mathbb{N}$  ou  $I = \llbracket 0; J \rrbracket$ .

Pour un schéma semi-discret, la stabilité est définie de la manière suivante.

**Définition 10** (Stabilité). On dit qu'un schéma semi-discret est *stable* s'il existe deux constantes  $K$  et  $\alpha$  telles que la solution  $U(t)$ , associée aux données initiale  $f$  et de bord  $g$ , vérifie, pour tout  $\Delta x$ , pour tout  $t \in \mathbb{R}^+$ ,

$$\|U(t)\|_{\Delta x}^2 \leq K e^{\alpha t} \left( \|f\|_{\Delta x}^2 + \int_0^t |g(\tau)|^2 d\tau \right).$$

**Définition 11** (Stabilité en temps). On dit qu'un schéma semi-discret est *stable en temps* si, pour  $g = 0$ , il existe une constante  $C$  (indépendante de  $f$ ) telle que, pour tout  $t \in \mathbb{R}^+$ , pour tout  $\Delta t$ , on a

$$\|U(t)\|_{\Delta x} \leq C \|f\|_{\Delta x}.$$

Pour un schéma discret, la stabilité est définie de la manière suivante.

**Définition 12** (Stabilité). On dit qu'un schéma discret est *stable* s'il existe deux constantes  $K$  et  $\alpha$  telles que la solution  $U_j^n$ , associée aux données initiale  $f$  et de bord  $g$ , vérifie, pour tout  $\Delta x$ , pour tout  $\Delta t$ , pour tout  $n \in \mathbb{N}$ ,

$$\|U^n\|_{\Delta x}^2 \leq K e^{\alpha n \Delta t} \left( \|f\|_{\Delta x}^2 + \Delta t \sum_{k=0}^n |g(k\Delta t)|^2 \right).$$

**Définition 13** (Stabilité en temps). On dit qu'un schéma discret est *stable en temps* si, pour  $g = 0$ , il existe une constante  $C$  (indépendante de  $f$ ) telle que, pour tout  $\Delta x$  et pour tout  $n \in \mathbb{N}$ , on a

$$\|U^n\|_{\Delta x} \leq C \|f\|_{\Delta x}.$$

**Définition 14** (Stabilité de Von Neumann). Pour un schéma discret (4), on note  $\gamma(\xi)$  son symbole<sup>1</sup> pour tout  $\xi \in \mathbb{R}$ . La condition de Von Neumann est satisfaite s'il existe une constante  $C$  tel que pour tout  $\xi \in \mathbb{R}$ , on a

$$|\gamma(\xi)| \leq 1 + C\Delta t.$$

**Remarque 15.** En pratique, on cherche à vérifier la condition plus forte

$$\forall \xi \in \mathbb{R}, \quad |\gamma(\xi)| \leq 1.$$

Cette condition induit la stabilité en temps du schéma.

---

1. Cette notion sera développée plus précisément dans la Section 2.2.1.

## 2.1 Étude de la matrice Quasi-Toeplitz

L'analyse des matrices Toeplitz et Quasi-Toeplitz est illustrée dans l'article [BW93]. Durant le séminaire encadré par Benjamin Boutin, on a compris que le spectre asymptotique de la matrice Toeplitz liée au schéma sans bord était inclus dans l'intérieur du spectre de l'opérateur sur  $\mathbb{Z}$  et que dans le spectre asymptotique de la matrice Quasi-Toeplitz (celle avec les conditions de bords), on voyait apparaître en plus un nombre fini de valeurs propres ne dépendant que des conditions de bords. De plus, le spectre asymptotique de la matrice circulante (correspondant au cas périodique) est le spectre de l'opérateur sur  $\mathbb{Z}$  qui est l'ensemble  $\gamma(\mathbb{R})$  où  $\gamma$  est le symbole du schéma.

On note  $\overset{\circ}{T}_J$  les matrices de Toeplitz circulantes, autrement dit la matrice associée au problème périodique et on note  $T_J$  les matrices Quasi-Toeplitz incluant les conditions de bords.

On écrit alors un schéma discret sous la forme suivante

$$\begin{cases} U^{n+1} = T_J U^n + b_n \\ U^0 = (f(x_j))_{j \in \llbracket 0; J \rrbracket}. \end{cases} \quad (5)$$

avec  $U^n = \begin{pmatrix} U_0^n \\ \vdots \\ U_J^n \end{pmatrix} \in \mathbb{R}^{J+1}$  pour tout  $n \in \mathbb{N}$  et  $T_J \in \mathcal{M}_{J+1}(\mathbb{R})$ . Le vecteur colonne  $b_n$  de taille  $J+1$  ne dépend que de  $g$ . La matrice Quasi-Toeplitz  $T_J$  encode le schéma intérieur et au bord.

**Exemple A 16.** Pour le schéma (Upw), on a les matrices suivantes

$$T_J = \begin{pmatrix} 1 + \lambda\sigma & -\lambda(1 + \sigma) & \cdots & 0 \\ \lambda & 1 - \lambda & 0 & \\ & \ddots & \ddots & 0 \\ 0 & & \lambda & 1 - \lambda \end{pmatrix} \quad \text{et} \quad b_n = \begin{pmatrix} \lambda g(t^n) \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

**Proposition 17.** Si  $(T_J^k)_{k \in \mathbb{N}, J \in \mathbb{N}}$  est bornée pour la norme  $\|\cdot\|_{\Delta x}$ , alors le schéma (5) est stable en temps.

**Remarque 18.** Il faut bien garder en tête que le  $\Delta x$  présent dans la norme et le  $J$  de  $T_J^k$  sont liés.

*Démonstration.* On suppose que  $g = 0$ , puisqu'on recherche la stabilité en temps. Soit  $C > 0$  telle que, pour tout  $k \in \mathbb{N}$ , pour tout  $J \in \mathbb{N}$ , on a  $\|T_J^k\| \leq C$ . Soit  $\Delta x \in ]0, 1[$ . On pose  $J$  associé à  $\Delta x$ . Pour le schéma (5), on a  $U^{n+1} = T_J U^n$  pour tout  $n \in \mathbb{N}$ . Donc, par récurrence, on a

$$U^n = T_J^n U^0.$$

On peut alors écrire

$$\|U^n\|_{\Delta x} \leq C \|U^0\|_{\Delta x} = C \|f\|_{\Delta x},$$

ce qui montre la stabilité du schéma (5). □

Si on travaille à  $\Delta x$  fixé, on a les propriétés suivantes

**Proposition 19.** Soit  $A \in \mathcal{M}_{J+1}(\mathbb{R})$ . Si  $\rho(A) < 1$ , alors  $(A^k)_{k \in \mathbb{N}^*}$  est bornée.

**Proposition 20.** Soit  $A \in \mathcal{M}_{J+1}(\mathbb{R})$ . Si  $(A^k)_{k \in \mathbb{N}^*}$  est bornée, alors  $\rho(A) \leq 1$ .

*Démonstration.* Soit une valeur propre  $\mu \in \mathbb{C}$  de  $A$ . Soit un vecteur propre  $X$  de norme 1 de la matrice  $A$  associée à  $\mu$ . On a, pour tout  $k \in \mathbb{N}$ ,

$$|\mu|^k = \|\mu^k X\| = \|A^k X\| \leq M,$$

où  $M$  est une borne des puissances de  $A$ . Donc  $|\mu| \leq 1$ . Ainsi le rayon spectral de  $A$  vérifie

$$\rho(A) \leq 1.$$

□

**Remarque 21.** Si  $\rho(A) = 1$ , alors on a équivalence entre  $(A^k)_{k \in \mathbb{N}^*}$  est bornée et les valeurs propres sont de multiplicité 1. Cela fournit une réciproque à la Propriété 20.

Pour avoir la stabilité de notre schéma, il faut que le rayon spectral soit contrôlé par 1 mais uniformément en  $\Delta x$ , d'où l'hypothèse de la propriété 17. Pour obtenir cette hypothèse, on peut utiliser le théorème suivant.

**Proposition 22.** Soit  $n_0 \in \mathbb{N}$  la largeur du stencil du schéma (5). Si le schéma (5) est Von Neumann-stable, si pour tout  $J > n_0$ ,  $\rho(T_J) \leq 1$  et si  $(\|T_J - \overset{\circ}{T}_J\|)_{J \geq 1}$  est bornée, alors  $(T_J^k)_{k \in \mathbb{N}, J \in \mathbb{N}}$  est bornée.

*Démonstration.* Ce résultat est démontré dans l'article [Sou09].

□

**Rappel 23.** La matrice  $\overset{\circ}{T}_J$  est la matrice de Toeplitz circulante. Ainsi la matrice  $\|T_J - \overset{\circ}{T}_J\|$  ne fait apparaître que la condition de bord et possède une majorité de 0.

**Exemple 24.** Pour le schéma (Upw), on a

$$\overset{\circ}{T}_J = \begin{pmatrix} 1 - \lambda & 0 & 0 & \lambda \\ \lambda & \ddots & \ddots & 0 \\ & \ddots & \ddots & 0 \\ 0 & & \lambda & 1 - \lambda \end{pmatrix} \quad \text{et} \quad T_J - \overset{\circ}{T}_J = \begin{pmatrix} \lambda(\sigma + 1) & -\lambda(1 + \sigma) & 0 & -\lambda \\ 0 & \cdots & & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & & 0 \end{pmatrix}$$

**Exemple A 25.** Pour le schéma (Upw) et pour  $J = 200$ , on obtient numériquement le domaine de stabilité présenté en Figure 2.

**Proposition 26.** Pour  $\sigma \in ]-1, 1[$ , les valeurs propres de la matrice  $T_J \in \mathcal{M}_{J+1}(\mathbb{R})$  associée au schéma (Upw) sont :

- $1 - \lambda$  de multiplicité  $J - 1$  et
- $\frac{2 - \lambda + \lambda\sigma \pm i\lambda\sqrt{(3 - \sigma)(1 + \sigma)}}{2}$  toutes deux de multiplicité 1.

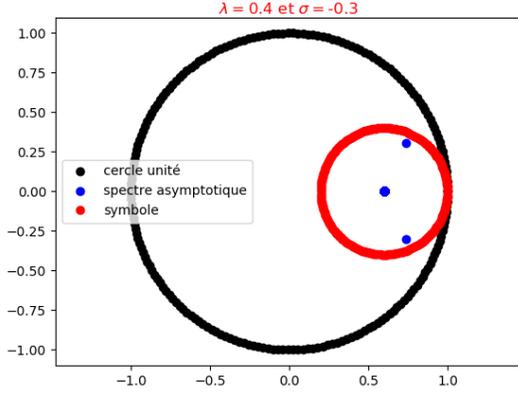


FIGURE 1 – Spectre asymptotique du schéma  $(Upw)$  pour  $(\lambda, \sigma) = (0.4, -0.3)$ .

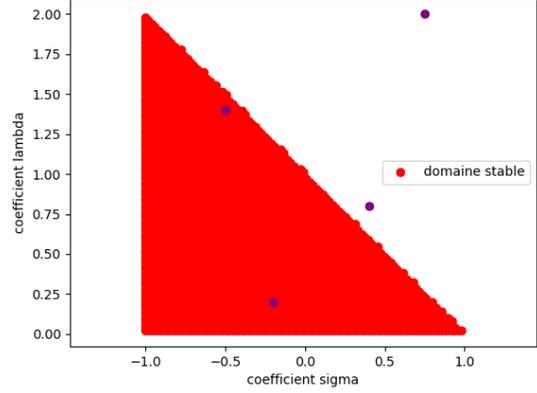


FIGURE 2 – Domaine de stabilité  $\rho(T_J) \leq 1$  pour le schéma  $(Upw)$  – les points violets sont les paramètres des courbes de la Figure 3.

*Démonstration.* On résout un système linéaire afin de trouver les valeurs propres de  $T_J$ . Soit  ${}^t(x_0, \dots, x_J)$  un vecteur propre de  $T_J$  associé à la valeur propre  $\mu$ . On obtient

$$\lambda x_{k-1} + (1 - \lambda - \mu)x_k = 0, \quad \forall k \in \llbracket 1; J \rrbracket.$$

Pour  $\mu \neq 1 - \lambda$ , on trouve

$$x_k = \left( \frac{-\lambda}{1 - \lambda - \mu} \right)^k x_0, \quad \forall k \in \llbracket 1; J \rrbracket.$$

En utilisant la première ligne de  $T_J$ , on obtient

$$\begin{aligned} (1 + \lambda\sigma - \mu)x_0 - \lambda(1 + \sigma)x_1 &= 0 \\ (1 + \lambda\sigma - \mu) - \lambda(1 + \sigma) \left( \frac{-\lambda}{1 - \lambda - \mu} \right) &= 0 \\ (1 - \lambda - \mu)(1 + \lambda\sigma - \mu) + \lambda^2(1 + \sigma) &= 0 \\ \mu^2 - \mu(2 - \lambda + \lambda\sigma) + 1 - \lambda + \lambda\sigma + \lambda^2 &= 0 \end{aligned}$$

Pour  $\sigma \in ]-1, 1[$ , on trouve

$$\mu_{1,2} = \frac{2 - \lambda + \lambda\sigma \pm i\lambda\sqrt{(3 - \sigma)(1 + \sigma)}}{2}.$$

□

Pour étudier la stabilité du schéma, on cherche des valeurs propres  $\mu$  tel que  $|\mu| < 1$ , de ce fait, pour le schéma  $(Upw)$ , on veut

$$\begin{aligned} |\mu_{1,2}|^2 &= \frac{(2 - \lambda + \lambda\sigma)^2}{4} + \lambda^2 \frac{(3 - \sigma)(1 + \sigma)}{4} < 1 \\ 4 + \lambda^2 + \lambda^2\sigma^2 - 4\lambda - 2\lambda^2\sigma + 4\lambda\sigma + \lambda^2(-\sigma^2 + 2\sigma + 3) &< 4 \\ \lambda^2 - \lambda + \lambda\sigma &< 0 \\ \lambda - 1 + \sigma &< 0 \\ \lambda &< 1 - \sigma. \end{aligned}$$

Pour le schéma  $(Upw)$ , afin d'obtenir la stabilité, il faut et il suffit d'avoir  $\lambda \in [0, 2]$  et  $\lambda < 1 - \sigma$ .

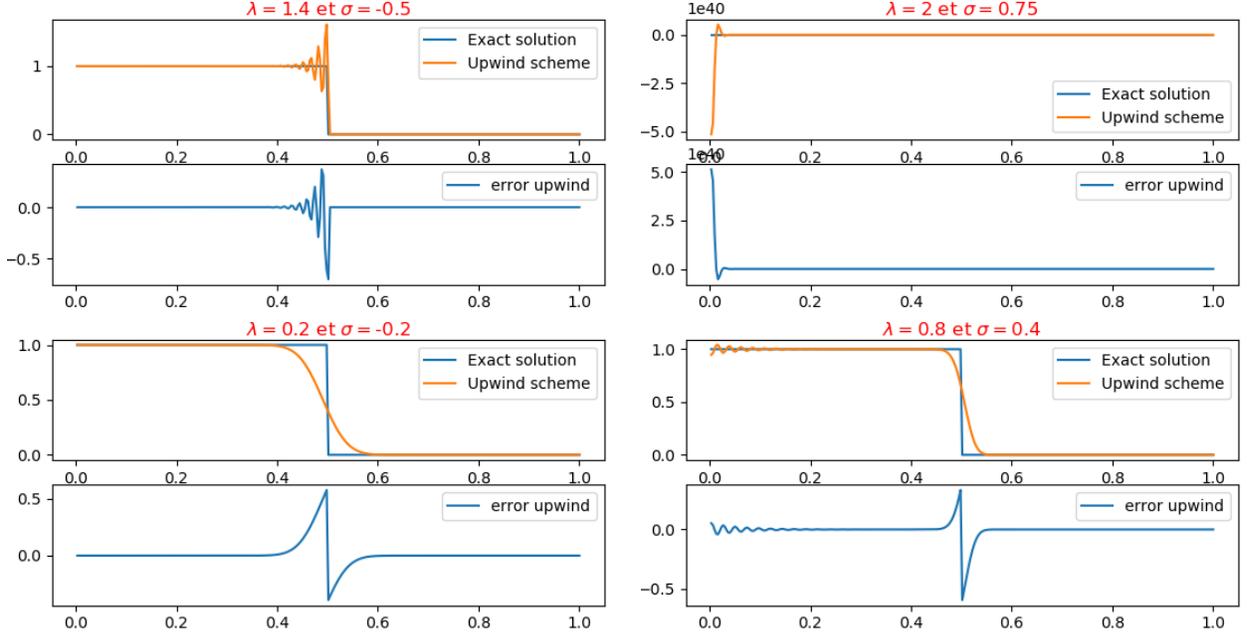


FIGURE 3 – Tracé de résolution<sup>2</sup> de (1) par le schéma (Upw), avec une condition initiale nulle et une condition de bord égale à 1, au temps  $t = 0.5$  pour différents couples  $(\sigma, \lambda)$  donnés en violet sur la Figure 2.

**Remarque 27.** Dans la Figure 3, pour l'exemple  $(\lambda, \sigma) = (1.4, -0.5)$ , on observe une oscillation, mais celle-ci n'est pas liée aux conditions de bord, mais à l'aspect  $\lambda \in ]1, 2]$ , puisque dans ce cas là, le schéma n'est pas monotone et ne vérifie pas le principe du maximum. De plus, cela montre que ce n'est pas parce que les puissances de  $T_J$  sont bornées (Propriété 19) que ses puissances sont bornées uniformément en  $J$ , ce qui justifie la Propriété 22. D'ailleurs, le schéma n'est pas Von Neumann-stable puisque le symbole est de module plus grand que 1.

**Exemple 28.** Pour les schémas de Beam-Warming (BW1), (BW2) et (BW3), on va écrire les matrices Quasi-Toeplitz  $T_J$  sous la forme

$$T_J = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & 0 & 0 \\ a_{2,1} & a_{2,2} & a_{2,3} & \vdots & \vdots \\ \frac{\lambda(\lambda-1)}{2} & \lambda(2-\lambda) & \frac{(\lambda-1)(\lambda-2)}{2} & \vdots & \vdots \\ \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & & \frac{\lambda(\lambda-1)}{2} & \lambda(2-\lambda) & \frac{(\lambda-1)(\lambda-2)}{2} \end{pmatrix} \text{ avec } A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{pmatrix}.$$

**Exemple B 29.** Pour le schéma (BW1), on a

$$\begin{aligned} U_0^{n+1} &= \left( \frac{(\lambda-1)(\lambda-2)}{2} + \lambda(2-\lambda)(1+\sigma) + \lambda \frac{(\lambda-1)(2+\sigma)}{2} \right) U_0^n \\ &\quad + \left( (\lambda-2)\lambda(1+\sigma) + \lambda \frac{(1-\lambda)(2+\sigma)}{2} \right) U_1^n + \left( \lambda(2-\lambda) + \lambda \frac{\lambda-1}{2} \right) g(t^n) \\ &= \left( 1 + \frac{\lambda^2}{2} - \frac{\lambda}{2} + \frac{3}{2}\lambda\sigma - \frac{\lambda^2\sigma}{2} \right) U_0^n + \left( \frac{\lambda^2\sigma}{2} - \lambda - \frac{3}{2}\lambda\sigma \right) U_1^n + \left( \frac{3}{2}\lambda - \frac{\lambda^2}{2} \right) g(t^n) \end{aligned}$$

2. Le code Python pour obtenir ces courbes est donné en Section 6.1.

$$U_1^{n+1} = \left( \lambda(2 - \lambda) + (1 + \sigma) \frac{\lambda(\lambda - 1)}{2} \right) U_0^n + \left( \frac{(\lambda - 1)(\lambda - 2)}{2} - (1 + \sigma) \frac{\lambda(\lambda - 1)}{2} \right) U_1^n + \frac{\lambda(\lambda - 1)}{2} g(t^n).$$

Ainsi, on peut écrire la matrice  $A$  et le vecteur colonne  $b_n$  :

$$A = \begin{pmatrix} 1 + \frac{\lambda^2}{2} - \frac{\lambda}{2} + \frac{3}{2}\lambda\sigma - \frac{\lambda^2\sigma}{2} & \frac{\lambda^2\sigma}{2} - \lambda - \frac{3}{2}\lambda\sigma & 0 \\ \lambda(2 - \lambda) + (1 + \sigma) \frac{\lambda(\lambda - 1)}{2} & \frac{(\lambda - 1)(\lambda - 2)}{2} - (1 + \sigma) \frac{\lambda(\lambda - 1)}{2} & 0 \end{pmatrix} \quad \text{et} \quad b_n = \begin{pmatrix} \left( \frac{3}{2}\lambda - \frac{\lambda^2}{2} \right) g(t^n) \\ \frac{\lambda(\lambda - 1)}{2} g(t^n) \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

En utilisant des outils numériques, à  $J$  fixé à 200, on trace le domaine de stabilité, *i.e.* le domaine où  $\rho(T_J) < 1$  et on obtient la Figure 5.

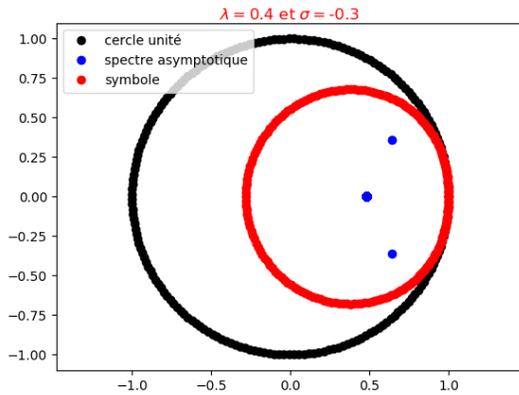


FIGURE 4 – Spectre asymptotique du schéma (BW1) pour  $(\lambda, \sigma) = (0.4, -0.3)$ .

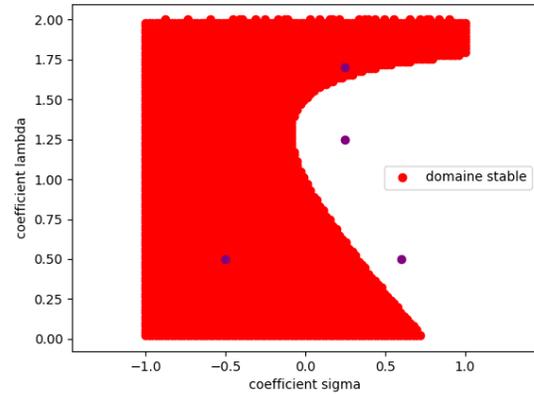


FIGURE 5 – Domaine de stabilité  $\rho(T_J) \leq 1$  pour le schéma (BW1) – les points violets sont les paramètres des courbes de la Figure 6.

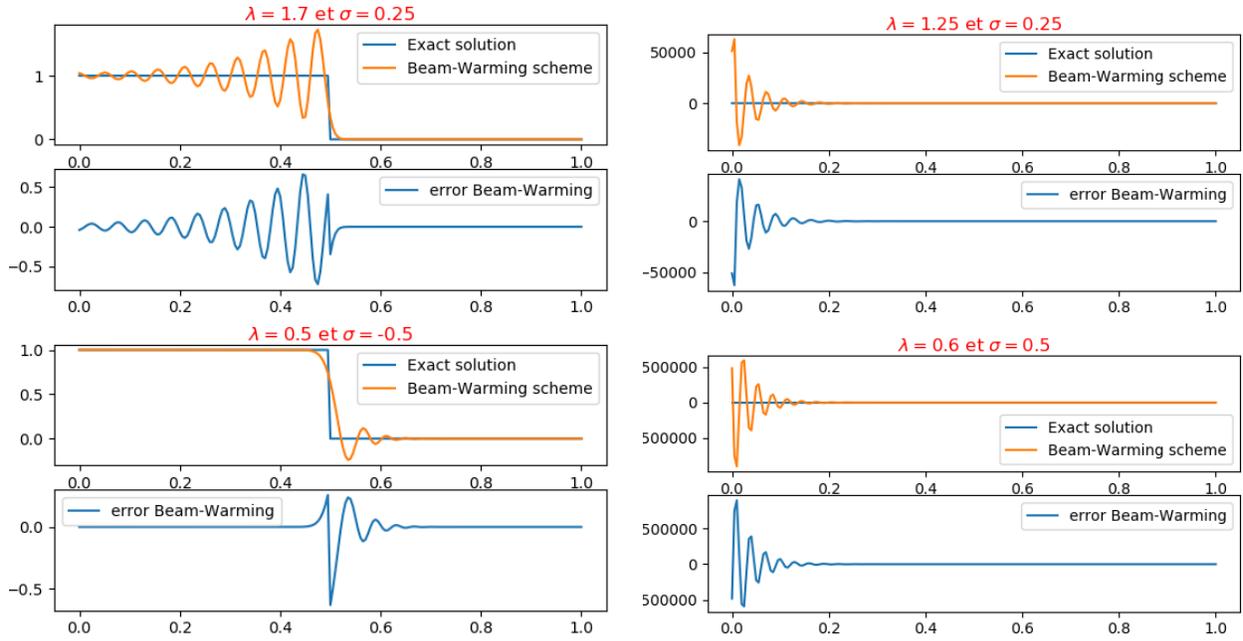


FIGURE 6 – Tracé de résolution<sup>3</sup> de (1) par le schéma (BW1), avec une condition initiale nulle et une condition de bord égale à 1, au temps  $t = 0.5$  pour différents couples  $(\sigma, \lambda)$  donnés en violet sur la Figure 5.

**Exemple C 30.** Pour le schéma (BW2), on a

$$A = \begin{pmatrix} \frac{(\lambda-1)(\lambda-2)}{2} & 0 & 0 \\ \lambda(2-\lambda) & \frac{(\lambda-1)(\lambda-2)}{2} & 0 \end{pmatrix}$$

Le vecteur  $b_n$  dépend de  $g$ ,  $g'$  et  $g''$ .

La matrice  $T_J$  est alors triangulaire inférieure avec des coefficients diagonaux tous égaux à  $\frac{(\lambda-1)(\lambda-2)}{2}$ . Pour avoir la condition  $\rho(T_J) < 1$ , il faut et il suffit d'avoir  $\lambda \in ]0, 4[$ . Cependant pour être Von Neumann-stable, il faut que  $\lambda < 2$ . On obtient donc le domaine de stabilité présenté en Figure 8.

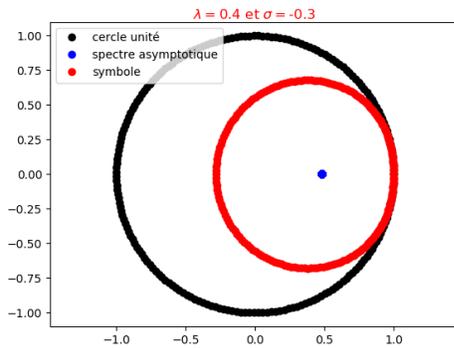


FIGURE 7 – Spectre asymptotique du schéma (BW2) pour  $(\lambda, \sigma) = (0.4, -0.3)$ .

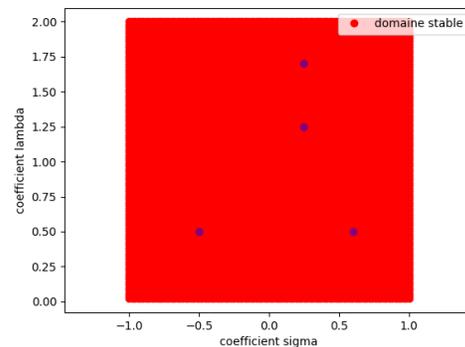


FIGURE 8 – Domaine de stabilité  $\rho(T_J) \leq 1$  pour le schéma (BW2) – les points violets sont les paramètres des courbes de la Figure 9.

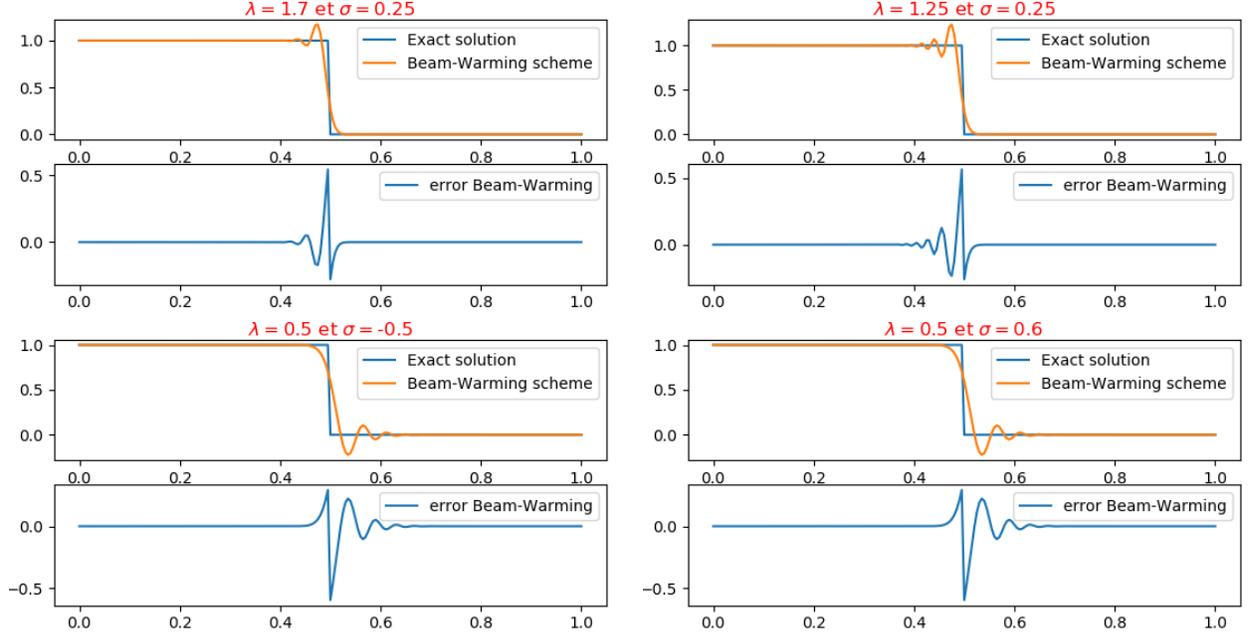


FIGURE 9 – Tracé de résolution de (1) par le schéma (BW2), avec une condition initiale nulle et une condition de bord égale à 1, au temps  $t = 0.5$  pour différents couples  $(\sigma, \lambda)$  donnés en violet sur la Figure 8.

**Exemple D 31.** Pour le schéma (BW3), on a

$$\begin{aligned}
 U_0^{n+1} &= \left( \frac{(\lambda-1)(\lambda-2)}{2} + \lambda(2-\lambda)\frac{(1+\sigma)^2}{2} + \lambda\frac{(\lambda-1)(2+\sigma)^2}{4} \right) U_0^n \\
 &\quad - \left( \lambda(2-\lambda)(1+\sigma)^2 + \lambda\frac{\lambda-1}{2}(2+\sigma)^2 \right) U_1^n + \left( \lambda(2-\lambda)\frac{(1+\sigma)^2}{2} + \lambda\frac{(\lambda-1)(2+\sigma)^2}{4} \right) U_2^n \\
 &\quad + \lambda(2-\lambda) \left( g(t^n) + (1+\sigma)\frac{\Delta x g'(t^n)}{a} \right) + \frac{\lambda(\lambda-1)}{2} \left( g(t^n) + (2+\sigma)\frac{\Delta x g'(t^n)}{a} \right) \\
 U_1^{n+1} &= \left( \lambda(2-\lambda) + \lambda\frac{(\lambda-1)(1+\sigma)^2}{2} \right) U_0^n + \left( \frac{(\lambda-1)(\lambda-2)}{2} - (1+\sigma)^2\frac{\lambda(\lambda-1)}{2} \right) U_1^n \\
 &\quad + \frac{\lambda(\lambda-1)(1+\sigma)^2}{2} U_2^n + \frac{\lambda(\lambda-1)}{2} \left( g(t^n) + (1+\sigma)\frac{\Delta x g'(t^n)}{a} \right)
 \end{aligned}$$

Ainsi, on peut écrire la matrice  $A$  :

$$A = \begin{pmatrix} \frac{(\lambda-1)(\lambda-2)}{2} + \frac{\lambda(2-\lambda)(1+\sigma)^2}{2} + \frac{\lambda(\lambda-1)(2+\sigma)^2}{4} & -\frac{2\lambda(2-\lambda)(1+\sigma)^2}{2} - \frac{\lambda(\lambda-1)(2+\sigma)^2}{2} & \frac{\lambda(2-\lambda)(1+\sigma)^2}{2} + \frac{\lambda(\lambda-1)(2+\sigma)^2}{4} \\ \lambda(2-\lambda) + \frac{\lambda(\lambda-1)(1+\sigma)^2}{2} & \frac{(\lambda-1)(\lambda-2)}{2} - \frac{(1+\sigma)^2\lambda(\lambda-1)}{2} & \frac{\lambda(\lambda-1)(1+\sigma)^2}{2} \end{pmatrix}$$

Le vecteur  $b_n$  dépend de  $g$  et  $g'$ .

En utilisant des outils numériques, à  $J$  fixé à 200, on trace le domaine de stabilité de (BW3), *i.e.* le domaine où  $\rho(T_J) < 1$  et on obtient la Figure 11.

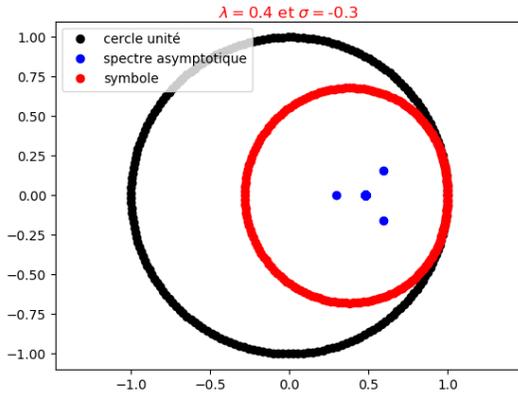


FIGURE 10 – Spectre asymptotique du schéma (BW3) pour  $(\lambda, \sigma) = (0.4, -0.3)$ .

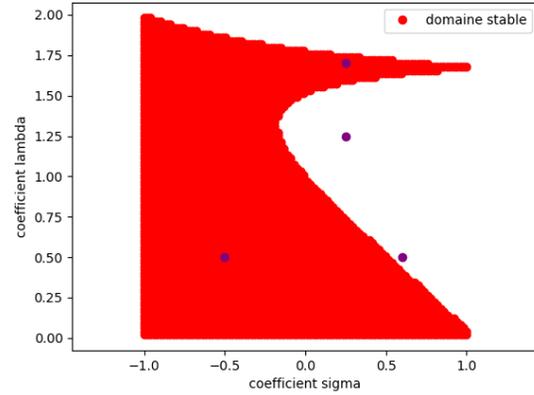


FIGURE 11 – Domaine de stabilité  $\rho(T_J) \leq 1$  pour le schéma (BW3) – les points violets sont les paramètres des courbes de la Figure 12.

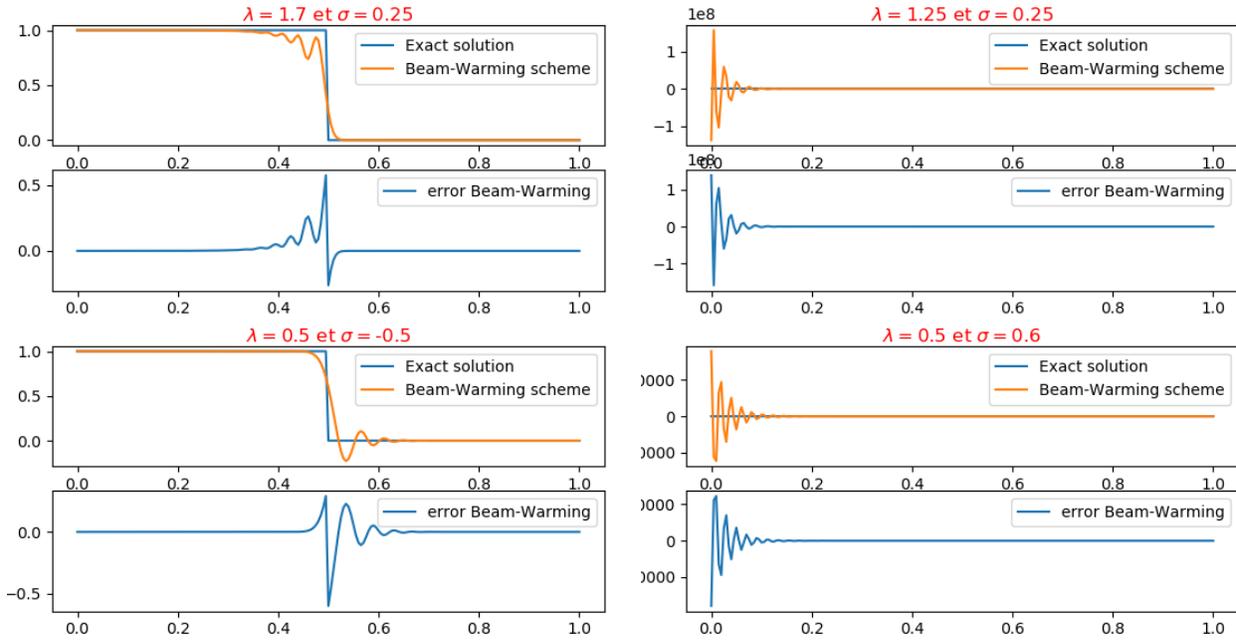


FIGURE 12 – Tracé de résolution de (1) par le schéma (BW3), avec une condition initiale nulle et une condition de bord égale à 1, au temps  $t = 0.5$  pour différents couples  $(\sigma, \lambda)$  donnés en violet sur la Figure 11.

**Remarque 32.** Les quatre exemples précédents présentent la particularité de traiter à chaque fois des matrices Toeplitz triangulaire inférieure (modulo les conditions de bords), ainsi le spectre asymptotique de la matrice de Toeplitz pure est réduit à un seul point. Pour éviter cet écueil, il suffit de considérer des schémas qui ne soient pas complètement décentrés, par exemple les schémas de Lax-Wendroff ou de Lax-Friedrichs, cependant on doit alors également traiter le bord de droite qui engendre à son tour des valeurs propres isolées.

## 2.2 Théorie GKS

La théorie GKS (Gustafsson, Kreiss et Sundström) ou analyse par mode propre est introduite dans l'article [GKS72] en 1972 et relatée dans l'article [VS15] et le livre [Gus08].

**Proposition 33.** Pour obtenir la stabilité d'un schéma semi-discret (2) ou discret (4), il suffit d'obtenir :

- (a) la stabilité du problème de Cauchy sur  $\mathbb{Z}$ ,
- (b) la stabilité du problème avec uniquement la condition de bord à droite,
- (c) la stabilité du problème avec uniquement la condition de bord à gauche.

Les points (b) et (c) se traitent de la même manière, on présentera la technique dans la Section 2.2.2.

### 2.2.1 Stabilité du problème de Cauchy sur $\mathbb{Z}$

Afin d'obtenir la stabilité du problème de Cauchy, on peut utiliser l'analyse de Fourier, notamment en étudiant le symbole du schéma. Cette méthode est parfois appelée « Analyse de stabilité de Von Neumann ». On peut s'appuyer sur le livre [Str04] pour comprendre plus en détail cette méthode.

**Exemple A 34.** On utilise la transformée de Fourier pour donner le symbole du schéma. Pour le schéma (Upw), on a

$$U_j^{n+1} = U_j^n - \lambda(U_j^n - U_{j-1}^n).$$

En passant à la transformée de Fourier, on trouve

$$\begin{aligned} \widehat{U}^{n+1}(\xi) &= \sum_{j \in \mathbb{Z}} U_j^{n+1} e^{ij\xi\Delta x} \\ &= \sum_{j \in \mathbb{Z}} \left( (1 - \lambda)U_j^n e^{ij\xi\Delta x} + \lambda U_{j-1}^n e^{ij\xi\Delta x} \right) \\ &= (1 - \lambda)\widehat{U}^n(\xi) + \lambda e^{i\xi} \widehat{U}^n(\xi) \\ &= \underbrace{(1 - \lambda + \lambda e^{i\xi})}_{\gamma(\xi)} \widehat{U}^n(\xi) \end{aligned}$$

Pour obtenir la stabilité, on veut avoir  $|\gamma(\xi)| \leq 1$  pour tout  $\xi \in \mathbb{R}$ . Ainsi, on doit avoir  $\lambda \in [0, 1]$  pour obtenir la stabilité du schéma (Upw).

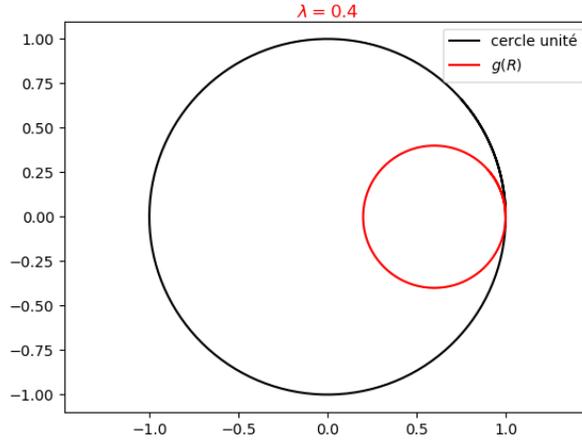


FIGURE 13 – Symbole du schéma de Upwind pour  $\lambda = 0.4$ .

**Exemple B 35.** Pour le schéma (BW1), on a

$$U_j^{n+1} = \frac{(\lambda - 1)(\lambda - 2)}{2} U_j^n + \lambda(2 - \lambda) U_{j-1}^n + \frac{\lambda(\lambda - 1)}{2} U_{j-2}^n.$$

En passant à la transformée de Fourier, on trouve

$$\begin{aligned} \widehat{U}^{n+1}(\xi) &= \sum_{j \in \mathbb{Z}} U_j^{n+1} e^{ij\xi\Delta x} \\ &= \sum_{j \in \mathbb{Z}} \left( \frac{(\lambda - 1)(\lambda - 2)}{2} U_j^n e^{ij\xi\Delta x} + \lambda(2 - \lambda) U_{j-1}^n e^{ij\xi\Delta x} + \frac{\lambda(\lambda - 1)}{2} U_{j-2}^n e^{ij\xi\Delta x} \right) \\ &= \underbrace{\left( \frac{(\lambda - 1)(\lambda - 2)}{2} + \lambda(2 - \lambda) e^{i\xi} + \frac{\lambda(\lambda - 1)}{2} e^{2i\xi} \right)}_{\gamma(\xi)} \widehat{U}^n(\xi) \end{aligned}$$

Pour obtenir la stabilité, on veut avoir  $|\gamma(\xi)| \leq 1$  pour tout  $\xi \in \mathbb{R}$ .

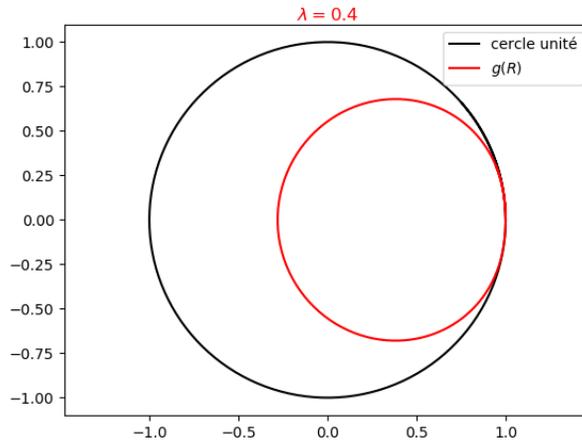


FIGURE 14 – Symbole du schéma de Beam-Warming pour  $\lambda = 0.4$ .

## 2.2.2 Stabilité du problème avec une seule condition de bord

Pour davantage de détails sur cette partie, on peut consulter l'article [VS15] et le livre [Gus08].

**Définition 36** (Valeur propre généralisée). Pour un schéma semi-discret (2), on appelle *valeur propre « généralisée »* un élément  $s \in \mathbb{C}$  associé à un vecteur propre  $v(t, x_j) = e^{st}\phi_j$  tel que :

- $(\phi_j)_{j \in \mathbb{N}}$  non nulle,
- $\frac{d}{dt}v(t, x_j) = Qv(t, x_j), t \geq 0, j \in \llbracket 0; J \rrbracket$ ,
- $B_0v(t, \cdot) = 0, t \geq 0$ ,
- $\Re(s) \geq 0$ ,
- si  $\Re(s) > 0$ ,  $v(t, x)$  est borné quand  $x$  tend vers l'infini,
- si  $\Re(s) = 0$ ,  $v(t, x)$  doit être stable au voisinage de  $s$ ,
- $\|\phi_j\|_{\Delta x} < \infty$ .

**Proposition 37.** Le schéma semi-discret (2) est stable si et seulement s'il ne possède pas de valeur propre généralisée.

**Remarque 38.** L'appellation valeur propre généralisée vient du fait qu'en injectant l'expression  $U_j(t) = e^{st}\phi_j$  dans l'équation

$$\frac{d}{dt}U_j(t) = QU_j(t),$$

on obtient

$$s\phi_j = Q\phi_j.$$

Donc  $s$  est valeur propre de l'opérateur  $Q$  et a pour vecteur propre  $(\phi_j)_j$ .

**Exemple 39.** On va étudier le schéma semi-discret (3). Supposons qu'il existe un  $s$  qui vérifie les conditions de la Définition 36. On injecte la solution  $U_j(t) = e^{st}\phi_j$  dans (3), afin d'obtenir

$$se^{st}\phi_j = -ae^{st}\frac{\phi_j - \phi_{j-1}}{\Delta x}$$

$$\frac{\Delta xs}{a}\phi_j = -(\phi_j - \phi_{j-1})$$

On pose  $\tilde{s} = \frac{\Delta xs}{a}$ . On remarque que  $\Re(s) \geq 0$  si et seulement si  $\Re(\tilde{s}) \geq 0$ . On obtient

$$(\tilde{s} + 1)\phi_j = \phi_{j-1}.$$

Autrement dit, pour tout  $j \in \mathbb{N}$ , on a

$$\phi_j = \left( \frac{1}{\tilde{s} + 1} \right)^j \phi_0.$$

On intègre cela dans la condition de bord (avec  $g = 0$ ) afin d'avoir

$$\begin{aligned} \tilde{s}\phi_0 &= -\phi_0 - (1 + \sigma)(\phi_1 - \phi_0) \\ &= \sigma\phi_0 - (1 + \sigma)\phi_1 \end{aligned}$$

$$= \left( \sigma - (1 + \sigma) \frac{1}{\tilde{s} + 1} \right) \phi_0$$

Ainsi, comme  $(\phi_j)_{j \in \mathbb{N}}$  non nulle, on a  $\phi_0 \neq 0$  et on peut écrire

$$\begin{aligned} (\tilde{s} + 1)\tilde{s} &= \sigma(\tilde{s} + 1) - (1 + \sigma) \\ \tilde{s}^2 + (1 - \sigma)\tilde{s} + 1 &= 0 \end{aligned}$$

Pour  $\sigma \in ]-1, 1[$ , on a

$$\Delta = (1 - \sigma)^2 - 4 = (\sigma - 3)(\sigma + 1) < 0.$$

D'où

$$\tilde{s}_{1,2} = \frac{-(1 - \sigma) \pm i\sqrt{(3 - \sigma)(1 + \sigma)}}{2}$$

On obtient alors  $\Re(\tilde{s}) = -\frac{1 - \sigma}{2} < 0$  pour  $\sigma \in ]-1, 1[$ , donc il n'existe pas de valeur propre généralisée et le schéma (3) est stable par la Proposition 37.

**Remarque 40.** Pour un schéma discret, on définit une valeur propre généralisée  $z \in \mathbb{C}$  associée au vecteur  $v(t^n, x_j) = z^n \phi_j$  en imposant les conditions suivantes :

- $(\phi_j)_{j \in \mathbb{N}}$  non nulle,
- $v(t^{n+1}, x_j) = Qv(t^n, x_j)$ ,  $n \in \mathbb{N}$ ,  $j \in \llbracket 0; J \rrbracket$ ,
- $B_0 v(t^n, \cdot) = 0$ ,  $n \in \mathbb{N}$ ,
- $|z| \geq 1$ ,
- si  $|z| > 1$ ,  $v(t^n, x)$  est bornée quand  $x$  tend vers l'infini,
- si  $|z| = 1$ ,  $v(t^n, x)$  doit être stable au voisinage de  $s$ ,
- $\|\phi_j\|_{\Delta x} < \infty$ .

L'étude ensuite se fait de manière similaire.

### 3 Consistance

Dans toute cette partie, on considère que la solution exacte  $u$  est suffisamment régulière.

#### 3.1 Étude de l'intérieur

**Définition 41** (Consistance à l'intérieur). On dit qu'un schéma discret est *consistant à l'intérieur* d'ordre  $p_0$  en temps et  $q_0$  en espace si la solution exacte  $u$  vérifie

$$u(t^{n+1}, x_j) = Qu(t^n, x_j) + \Delta t O(\Delta t^{p_0} + \Delta x^{q_0})$$

avec  $p_0$  et  $q_0$  maximaux.

On note  $\tau_j^n := u(t^{n+1}, x_j) - Qu(t^n, x_j)$ , appelée *erreur de troncature*.

**Remarque 42.** L'étude que l'on présente ici est formelle. Pour être rigoureux, il faut préciser la définition de  $O(\Delta t^{p_0})$  et notamment la topologie que l'on utilise lorsque  $\Delta t$  tend vers 0. Dans ce document, on a préféré mettre l'accent sur la stabilité et l'étude Quasi-Toeplitz/Théorie GKS en ne présentant qu'un aspect formel de la consistance.

**Exemple A 43.** Pour le schéma (Upw), on a, en utilisant les formules de Taylor Lagrange, pour tout  $n \in \mathbb{N}$ , pour tout  $j \in \llbracket 1; J-1 \rrbracket$ ,

$$\begin{aligned} u(t^{n+1}, x_j) - Qu(t^n, x_j) &= u(t^n + \Delta t, x_j) - (1 - \lambda)u(t^n, x_j) - \lambda u(t^n, x_{j-1}) \\ &= \Delta t \partial_t u(t^n, x_j) + O(\Delta t^2) + \lambda \Delta x \partial_x u(t^n, x_j) + O(\Delta x^2) \\ &= -a \Delta t \partial_x u(t^n, x_j) + O(\Delta t^2) + a \Delta t \partial_x u(t^n, x_j) + O(\Delta x^2) \\ &= O(\Delta t^2 + \Delta x^2). \end{aligned}$$

Sous la condition CFL, *i.e.*  $\lambda = \frac{a\Delta t}{\Delta x} \approx 1$ , on obtient alors une consistance d'ordre 1 en temps et en espace.

**Exemple B 44.** Pour le schéma (BW1), on pose

$$\begin{aligned} \tau_j^n &= u(t^{n+1}, x_j) - Qu(t^n, x_j) \\ &= u(t^{n+1}, x_j) - \frac{(\lambda - 1)(\lambda - 2)}{2} u(t^n, x_j) + \lambda(2 - \lambda)u(t^n, x_{j-1}) + \frac{\lambda(\lambda - 1)}{2} u(t^n, x_{j-2}). \end{aligned}$$

En utilisant les formules de Taylor Lagrange, on trouve

$$\begin{aligned} \tau_j^n &= \left( 1 - \frac{(\lambda - 1)(\lambda - 2)}{2} - \lambda(2 - \lambda) - \lambda \frac{\lambda - 1}{2} \right) u(t^n, x_j) \\ &\quad + (-\lambda \Delta x + \lambda(2 - \lambda)\Delta x + \lambda(\lambda - 1)\Delta x) \partial_x u(t^n, x_j) \\ &\quad + \left( \frac{\lambda^2 \Delta x^2}{2} - \lambda \frac{2 - \lambda}{2} \Delta x^2 - \lambda(\lambda - 1)\Delta x^2 \right) \partial_{xx}^2 u(t^n, x_j) \\ &\quad + O(\Delta t^3 + \lambda(2 - \lambda)\Delta x^3 + \lambda(\lambda - 1)\Delta x^3) \\ &= O(\Delta t^3 + \Delta x^2 \Delta t + \Delta x \Delta t^2). \end{aligned}$$

Sous la condition CFL, *i.e.*  $\lambda = \frac{a\Delta t}{\Delta x} \approx 1$ , on obtient alors une consistance d'ordre 2 en espace.

### 3.2 Étude du bord

**Définition 45** (Consistance au bord). On dit qu'un schéma discret est *consistant au bord* d'ordre  $q_1$  en espace si, en s'appuyant sur la deuxième équation de (4), on peut écrire

$$B_0 u(t^n, x_0) = g(t^n) + O(\Delta x^{q_1})$$

avec  $q_1$  maximal.

**Exemple A 46.** Pour le schéma (Upw), on a

$$U_{-1}^n + (1 + \sigma)(U_1^n - U_0^n) = g(t^n)$$

Ainsi, en injectant la solution exacte  $u$  dans l'équation précédente, on trouve

$$\begin{aligned} &u(t^n, -\Delta x) + (1 + \sigma)(u(t^n, \Delta x) - u(t^n, 0)) - g(t^n) \\ &= u(t^n, -\Delta x) + (1 + \sigma)(u(t^n, \Delta x) - u(t^n, 0)) - u(t^n, \sigma \Delta x) \\ &= u(t^n, 0) - \Delta x \partial_x u(t^n, 0) + O(\Delta x^2) + (1 + \sigma)(\Delta x \partial_x u(t^n, 0) + O(\Delta x^2)) \\ &\quad - u(t^n, 0) - \sigma \Delta x \partial_x u(t^n, 0) + O(\sigma^2 \Delta x^2) \\ &= O((1 + \sigma^2)\Delta x^2) \end{aligned}$$

**Exemple B 47.** Pour le schéma (BW1), on procède de la même façon et on trouve le même résultat puisque ce sont les mêmes formules de Taylor utilisées.

**Exemple C 48.** Pour le schéma (BW2), on a, pour tout  $p \in \{1, 2\}$ ,

$$U_{-p}^n = g(t^n) + (p + \sigma) \frac{\Delta x g'(t^n)}{a} + \frac{(p + \sigma)^2 \Delta x^2 g''(t^n)}{2 a^2}$$

Ainsi, en injectant la solution exacte  $u$  dans l'équation précédente, on trouve

$$\begin{aligned} & u(t^n, -p\Delta x) - u(t^n, \sigma\Delta x) - (p + \sigma) \frac{\Delta x \partial_t u(t^n, \sigma\Delta x)}{a} - \frac{(p + \sigma)^2 \Delta x^2}{2} \frac{\partial_{tt}^2 u(t^n, \sigma\Delta x)}{a^2} \\ &= u(t^n, -p\Delta x) - u(t^n, \sigma\Delta x) + (p + \sigma) \Delta x \partial_x u(t^n, \sigma\Delta x) - \frac{(p + \sigma)^2 \Delta x^2}{2} \partial_{xx}^2 u(t^n, \sigma\Delta x) \\ &= u(t^n, 0) - p\Delta x \partial_x u(t^n, 0) + \frac{p^2 \Delta x^2}{2} \partial_{xx}^2 u(t^n, 0) + O(\Delta x^3) \\ &\quad - u(t^n, 0) - \sigma\Delta x \partial_x u(t^n, 0) - \frac{\sigma^2 \Delta x^2}{2} \partial_{xx}^2 u(t^n, 0) + O(\sigma^3 \Delta x^3) \\ &\quad + (p + \sigma) \Delta x (\partial_x u(t^n, 0) + \sigma\Delta x \partial_{xx}^2 u(t^n, 0) + O(\sigma^2 \Delta x^2)) \\ &\quad - \frac{(p + \sigma)^2 \Delta x^2}{2} (\partial_{xx}^2 u(t^n, 0) + O(\sigma \Delta x)) \\ &= \frac{p^2 \Delta x^2}{2} \partial_{xx}^2 u(t^n, 0) - \frac{\sigma^2 \Delta x^2}{2} \partial_{xx}^2 u(t^n, 0) + (p + \sigma) \sigma \Delta x^2 \partial_{xx}^2 u(t^n, 0) - \frac{(p + \sigma)^2 \Delta x^2}{2} \partial_{xx}^2 u(t^n, 0) \\ &\quad + O((1 + \sigma + \sigma^2 + \sigma^3) \Delta x^3) \\ &= O((1 + \sigma^3) \Delta x^3) \end{aligned}$$

**Exemple D 49.** Pour le schéma (BW3), on a, pour tout  $p \in \{1, 2\}$ ,

$$U_{-p}^n = g(t^n) + (p + \sigma) \frac{\Delta x g'(t^n)}{a} + \frac{(p + \sigma)^2 \Delta x^2 U_2^n - 2U_1^n + U_0^n}{\Delta x^2}$$

Ainsi, en injectant la solution exacte  $u$  dans l'équation précédente, on trouve

$$\begin{aligned} & u(t^n, -p\Delta x) - u(t^n, \sigma\Delta x) - (p + \sigma) \frac{\Delta x \partial_t u(t^n, \sigma\Delta x)}{a} - \frac{(p + \sigma)^2 \Delta x^2}{2} \frac{u(t^n, 2\Delta x) - 2u(t^n, \Delta x) + u(t^n, 0)}{\Delta x^2} \\ &= u(t^n, -p\Delta x) - u(t^n, \sigma\Delta x) + (p + \sigma) \Delta x \partial_x u(t^n, \sigma\Delta x) - \frac{(p + \sigma)^2 \Delta x^2}{2} \frac{u(t^n, 2\Delta x) - 2u(t^n, \Delta x) + u(t^n, 0)}{\Delta x^2} \\ &= u(t^n, 0) - p\Delta x \partial_x u(t^n, 0) + \frac{p^2 \Delta x^2}{2} \partial_{xx}^2 u(t^n, 0) + O(\Delta x^3) \\ &\quad - u(t^n, 0) - \sigma\Delta x \partial_x u(t^n, 0) - \frac{\sigma^2 \Delta x^2}{2} \partial_{xx}^2 u(t^n, 0) + O(\sigma^3 \Delta x^3) \\ &\quad + (p + \sigma) \Delta x (\partial_x u(t^n, 0) + \sigma\Delta x \partial_{xx}^2 u(t^n, 0) + O(\sigma^2 \Delta x^2)) \\ &\quad - \frac{(p + \sigma)^2 \Delta x^2}{2} (2\Delta x \partial_x u(t^n, 0) + \frac{4\Delta x^2}{2} \partial_{xx}^2 u(t^n, 0) + O(\Delta x^3)) \\ &\quad \quad - 2\Delta x \partial_x u(t^n, 0) - 2\frac{\Delta x^2}{2} \partial_{xx}^2 u(t^n, 0) + O(\Delta x^3)) \\ &= \frac{p^2 \Delta x^2}{2} \partial_{xx}^2 u(t^n, 0) - \frac{\sigma^2 \Delta x^2}{2} \partial_{xx}^2 u(t^n, 0) + (p + \sigma) \sigma \Delta x^2 \partial_{xx}^2 u(t^n, 0) - \frac{(p + \sigma)^2 \Delta x^2}{2} \partial_{xx}^2 u(t^n, 0) \\ &\quad + O((1 + \sigma + \sigma^2 + \sigma^3) \Delta x^3) \\ &= O((1 + \sigma^3) \Delta x^3) \end{aligned}$$

## 4 Convergence

Dans toute cette partie, on considère que la solution exacte  $u$  est suffisamment régulière.

**Définition 50** (Convergence). On dit qu'un schéma est convergent si l'erreur de convergence  $e^n = (U_j^n)_j - (u(t^n, x_j))_j$  (où  $U_j^n$  est la solution du schéma et  $u$  la solution exacte du problème) tend vers 0 en norme  $\|\cdot\|_{\Delta x}$  quand  $\Delta x$  et  $\Delta t$  tendent vers 0.

**Théorème 51** (Lax). Un schéma est convergent si et seulement s'il est stable et consistant.

**Exemple 52.** Pour le schéma Upwind suivant (avec  $\lambda = 1$ )

$$\begin{cases} U_j^{n+1} = U_j^n - \lambda(U_j^n - U_{j-1}^n), & n \in \mathbb{N}, j \in \llbracket 0; J \rrbracket, \\ U_{-1}^n = g(t^n), & n \in \mathbb{N}, \\ U_j^0 = f(x_j), & j \in \llbracket 0; J \rrbracket, \end{cases}$$

on utilise des formules de Taylor avec reste intégral pour obtenir la consistance et en écrivant

$$e^{n+1} = T e^n + \eta^n + \Delta t \varepsilon^n,$$

où  $T$  est la matrice Quasi-Toeplitz liée au schéma,  $\eta^n$  l'erreur sur le bord liée à la condition de bord et  $\varepsilon^n$  l'erreur à l'intérieur du domaine.

On a alors

$$e^n = T^n e^0 + \sum_{k=0}^{n-1} T^{n-1-k} (\eta^k + \Delta t \varepsilon^k).$$

Par définition de  $U_j^0$ , on a  $e^0 = 0$ , donc il faut contrôler la somme en  $\eta^k$  et  $\varepsilon^k$ .

**Proposition 53.** Soit  $n \in \mathbb{N}^*$ . Pour  $\sigma$  fixé et  $M$  tel que  $M > \sigma + 2$ , on a

$$\begin{aligned} \left\| \sum_{k=0}^{n-1} T^{n-1-k} \eta^k \right\|_{\Delta x}^2 &= \sum_{j=0}^J \Delta x \left| \int_{-\Delta x}^{\sigma \Delta x} \partial_x u(t^{n-1-j}, y) dy \right|^2 \\ &\leq (1 + \sigma) \Delta x^2 M \|\partial_x u(t^n, \cdot)\|_{L^2}^2 \end{aligned}$$

**Proposition 54.** Soit  $n \in \mathbb{N}^*$ . Pour  $\sigma$  fixé et  $M$  tel que  $M > 2$ , on a

$$\varepsilon_j^n = \int_{t^n}^{t^{n+1}} (t^{n+1} - y) \partial_{tt}^2 u(y, x_j) dy - \lambda \int_{x_{j-1}}^{x_j} (y - x_{j-1}) \partial_{xx}^2 u(t^n, y) dy$$

et

$$\begin{aligned} \left\| \sum_{k=0}^{n-1} T^{n-1-k} \Delta t \varepsilon^k \right\|_{\Delta x}^2 &= \Delta t^2 \sum_{j=0}^J \Delta x |\varepsilon_j^n|^2 \\ &\leq \Delta t^2 \frac{2a^2 \Delta x^2}{3} M \|\partial_{xx}^2 u(t^n, \cdot)\|_{L^2}^2 \end{aligned}$$

Ainsi, le schéma Upwind converge et est d'ordre 1 puisque

$$\|e^n\|_{\Delta x} \leq \sqrt{(1 + \sigma) M \Delta x} \|\partial_x u(t^n, \cdot)\|_{L^2} + \Delta t a \Delta x \sqrt{\frac{2M}{3}} \|\partial_{xx}^2 u(t^n, \cdot)\|_{L^2}.$$

## 5 Conclusion

L'étude de la stabilité et de la consistance est très importante pour obtenir des propriétés de convergence. On s'est concentré sur la partie liée à la stabilité, notamment à travers deux méthodes : l'étude spectrale des matrices Quasi-Toeplitz et la théorie GKS. Il existe d'autres techniques, par exemple, la méthode d'énergie, qui consiste à trouver une norme qui fait décroître  $U(t)$  ou  $U^n$  en temps, par un calcul d'intégration par parties. Si on trouve une telle norme, on aura la stabilité, puisqu'on pourra écrire (pour le cas semi-discret)

$$\|U(t)\| \leq \|U(0)\| = \|f\|.$$

Cette théorie est expliquée dans le livre [Gus08].

On a étudié la théorie GKS pour le cas semi-discret car il est plus simple à traiter que les schémas complètement discrets, cependant les idées sont les mêmes et ne demandent pas d'introduire de nouveaux concepts.

L'introduction de points fantômes par la méthode de Lax Wendroff Inverse permet d'obtenir des propriétés de stabilité (comme on peut le voir pour l'exemple (BW3)), cependant en pratique, il faut connaître les dérivées successives de la condition de bord et parfois, on ne la possède pas, c'est pourquoi on peut l'approcher ou la remplacer par des coefficients liés à une extrapolation d'autres termes comme on l'a vu dans la méthode de Lax Wendroff Inverse simplifiée.

Dans nos quatre exemples, on a des schémas entièrement décentrés à gauche. Si on avait des schémas plus centrés, il faudrait rajouter une condition de bord à droite. De plus, le spectre asymptotique de la matrice Toeplitz ne serait plus (en général) réduit à un seul point.

## 6 Annexes

### 6.1 Programmes Python pour tracer les solutions approchées par Upwind

```
import numpy as np
import scipy as sp
import matplotlib.pyplot as plt
import numba

##initial and boundary data
def initial(x):
    return 0
def inflow(t):
    g = 1
    return g

##Exact solution
def exact_pt(t,x):
    if x<t :
        return inflow(t-x/a)
    else :
        return initial(x-a*t)
```

```

exact = np.vectorize(exact_pt)

##Upwind scheme
def Upw(omega,U):
    return U[1:] - omega*(U[1:]-U[0:-1])

def SolveUpw(L,x,U,T,cfl,sigma):
    t = 0
    dx = x[1]-x[0]
    dt = np.min([cfl*dx/a,dx])
    Utmp = U #Initial condition
    U = Upw(cfl,Utmp)
    t = t + dt

    while t<T :
        # Adjustment of time step
        dt = np.min([dt,T-t])

        # Boundary conditions
        Ui0 = (1+cfl*sigma)*Utmp[0] -cfl*(1+sigma)*Utmp[1]+ cfl*inflow(t)
        Utmp = np.concatenate(([Ui0],U))
        U = Upw(dt/dx*a,Utmp)
        t = t + dt
    Ui0 = (1+cfl*sigma)*Utmp[0] -cfl*(1+sigma)*Utmp[1]+ cfl*inflow(t)
    return [x,np.concatenate(([Ui0],U))]

##Tests
a = 1.
L = 1
N = 300
x = np.linspace(0, L, N)
cfl = 0.8
sigma = 0.4
T = 0.5
U0 = np.array([initial(y) for y in x])
[x,Uupw] = SolveUpw(L,x,U0,T,cfl,sigma)
Ue = exact(T,x)

plt.subplot(3,1,1)
plt.plot(x, Ue)
plt.plot(x, Uupw)
plt.legend(['Exact solution','Upwind scheme'])
chaine= '$\lambda ='+str(cfl)+' et $\sigma ='+str(sigma)
plt.title(chaine, horizontalalignment = 'center', verticalalignment = 'center',
color='red')
plt.subplot(3,1,2)
plt.plot(x, Ue-Uupw)

```

```
plt.legend(['error upwind'])
plt.show()
```

## 6.2 Programmes Python pour tracer les solutions approchées par Beam-Warming

```
import numpy as np
import scipy as sp
import matplotlib.pyplot as plt
import numba

##initial and boundary data
def initial(x):
    return 0
def inflow(t):
    g = 1
    return g
def dinflow(t):
    g = 0
    return g
def d2inflow(t):
    g = 0
    return g

##Exact solution

def exact_pt(t,x):
    if x<t :
        return inflow(t-x/a)
    else :
        return initial(x-a*t)

exact = np.vectorize(exact_pt)

##Beam Warming scheme
def BW(omega,U):
    return (omega-1)*(omega-2)/2*U[2:] + omega*(2-omega)*U[1:-1] + omega*(omega-1)*U[0:-2]/2

def SolveBW(L,x,U,T,cfl,sigma):
    t = 0
    dx = x[1]-x[0]
    dt = cfl*dx/a
    Utmp = U #Initial condition
    U = BW(cfl,Utmp)
    t = t + dt

    while t<T :
        # Adjustment of time step
```

```

dt = np.min([dt,T-t])

# Boundary conditions
#Beam Warming Example B
Ui0 = (cfl-1)*(cfl-2)/2*Utmp[0] + cfl*(2-cfl)*(inflow(t)
- (1+sigma)*(Utmp[1]-Utmp[0])) + cfl*(cfl-1)/2*(inflow(t)-(2+sigma)*(Utmp[1]-Utmp[0]))
Ui1 = (cfl-1)*(cfl-2)/2*Utmp[1] + cfl*(2-cfl)*Utmp[0]
+ cfl*(cfl-1)/2*(inflow(t)-(1+sigma)*(Utmp[1]-Utmp[0]))

#Beam Warming Example C
# Ui0 = (cfl-1)*(cfl-2)/2*Utmp[0] + cfl*(2-cfl)*(inflow(t) +
(1+sigma)*dx*dinflow(t)/a + (1+sigma)**2/2*dx*dx*d2inflow(t)/(a*a))+cfl*(cfl-1)/2*(inflow(t)
+ (2+sigma)*dx*dinflow(t)/a + (2+sigma)**2/2*dx*dx*d2inflow(t)/(a*a))
# Ui1 = (cfl-1)*(cfl-2)/2*Utmp[1] + cfl*(2-cfl)*Utmp[0] + cfl*(cfl-1)/2*(inflow(t)
+ (1+sigma)*dx*dinflow(t)/a + (1+sigma)**2/2*dx*dx*d2inflow(t)/(a*a))

#Beam Warming Example D
# Ui0 = (cfl-1)*(cfl-2)/2*Utmp[0] + cfl*(2-cfl)*(inflow(t)
+ (1+sigma)*dx*dinflow(t)/a + (1+sigma)**2/2*(Utmp[2]-2*Utmp[1]+Utmp[0]))
+cfl*(cfl-1)/2*(inflow(t) + (2+sigma)*dx*dinflow(t)/a
+ (2+sigma)**2/2*(Utmp[2]-2*Utmp[1]+Utmp[0]))
# Ui1 = (cfl-1)*(cfl-2)/2*Utmp[1] + cfl*(2-cfl)*Utmp[0]
+ cfl*(cfl-1)/2*(inflow(t) + (1+sigma)*dx*dinflow(t)/a
+ (1+sigma)**2/2*(Utmp[2]-2*Utmp[1]+Utmp[0]))

Utmp = np.concatenate(([Ui0,Ui1],U))
U = BW(dt/dx*a,Utmp)
t = t + dt

#Beam Warming Example B
Ui0 = (cfl-1)*(cfl-2)/2*Utmp[0] + cfl*(2-cfl)*(inflow(t)
- (1+sigma)*(Utmp[1]-Utmp[0])) + cfl*(cfl-1)/2*(inflow(t)-(2+sigma)*(Utmp[1]-Utmp[0]))
Ui1 = (cfl-1)*(cfl-2)/2*Utmp[1] + cfl*(2-cfl)*Utmp[0]
+ cfl*(cfl-1)/2*(inflow(t)-(1+sigma)*(Utmp[1]-Utmp[0]))

#Beam Warming Example C
# Ui0 = (cfl-1)*(cfl-2)/2*Utmp[0] + cfl*(2-cfl)*(inflow(t) + (1+sigma)*dx*dinflow(t)/a
+ (1+sigma)**2/2*dx*dx*d2inflow(t)/(a*a))+cfl*(cfl-1)/2*(inflow(t)
+ (2+sigma)*dx*dinflow(t)/a + (2+sigma)**2/2*dx*dx*d2inflow(t)/(a*a))
# Ui1 = (cfl-1)*(cfl-2)/2*Utmp[1] + cfl*(2-cfl)*Utmp[0] + cfl*(cfl-1)/2*(inflow(t)
+ (1+sigma)*dx*dinflow(t)/a + (1+sigma)**2/2*dx*dx*d2inflow(t)/(a*a))

#Beam Warming Example D
# Ui0 = (cfl-1)*(cfl-2)/2*Utmp[0] + cfl*(2-cfl)*(inflow(t)
+ (1+sigma)*dx*dinflow(t)/a + (1+sigma)**2/2*(Utmp[2]-2*Utmp[1]+Utmp[0]))
+cfl*(cfl-1)/2*(inflow(t) + (2+sigma)*dx*dinflow(t)/a
+ (2+sigma)**2/2*(Utmp[2]-2*Utmp[1]+Utmp[0]))
# Ui1 = (cfl-1)*(cfl-2)/2*Utmp[1] + cfl*(2-cfl)*Utmp[0]
+ cfl*(cfl-1)/2*(inflow(t) + (1+sigma)*dx*dinflow(t)/a

```

```

+ (1+sigma)**2/2*(Utmp[2]-2*Utmp[1]+Utmp[0]))
    return [x,np.concatenate(([Ui0,Ui1],U))]

##Tests
a = 1.
L = 1
N = 201
x = np.linspace(0, L, N)
cfl = 0.5
sigma = 0.6
T = 0.5
U0 = np.array([initial(y) for y in x])
[x,UBW] = SolveBW(L,x,U0,T,cfl,sigma)
Ue = exact(T,x)

plt.subplot(3,1,1)
plt.plot(x, Ue)
plt.plot(x, UBW)
plt.legend(['Exact solution','Beam-Warming scheme'])
chaine= '$\lambda ='+str(cfl)+' et $\sigma ='+str(sigma)
plt.title(chaine, horizontalalignment = 'center',
verticalalignment = 'center', color='red')
plt.subplot(3,1,2)
plt.plot(x, Ue-UBW)
plt.legend(['error Beam-Warming'])
plt.show()

```

## Références

- [BW93] Richard Beam and Robert Warming. The asymptotic spectra of banded toeplitz and quasi-toeplitz matrices. *SIAM Journal on Scientific Computing*, pages 971–1006, July 1993.
- [Cou11] Jean-François Coulombel. *Stability of Finite Difference Schemes for Hyperbolic Initial Boundary value Problems*. 2011.
- [DDS18] Gautier Dakin, Bruno Després, and Jaouen Stéphane. Inverse lax-wendroff boundary treatment for compressible lagrange-remap hydrodynamics on cartesian grids. *Journal of Computational Physics*, 353 :228–257, 2018.
- [Des09] Bruno Després. Uniform asymptotic stability of strang’s explicit compact schemes for linear advection. *SIAM*, 47(5) :3956–3976, 2009.
- [GKS72] B. Gustafsson, H.-O. Kreiss, and A. Sundström. Stability theory of difference approximations for mixed initial boundary value problem. *Math. Comput*, 26 :649–686, 1972.
- [Gus08] Bertil Gustafsson. *High Order Difference Methods for Time Dependent PDE*. Springer, 2008.
- [LSZ16] Tingting Li, Chi-Wang Shu, and Mengping Zhang. Stability analysis of the inverse lax-wendroff boundary treatment for high order upwind-biased finite difference schemes. *Journal of Computational and Applied Mathematics*, 299 :140–158, 2016.
- [Sou09] Ercilia Sousa. On the edge of stability analysis. *Applied Numerical Mathematics*, 59 :1322–1336, 2009.
- [Str04] John C. Strikwerda. *Finite Difference Schemes and Partial Differential Equations*. SIAM, 2004.
- [VS15] François Vilar and Chi-Wang Shu. Development and stability analysis of the inverse lax-wendroff boundary treatment for central compact schemes. *ESAIM : Mathematical Modelling and Numerical Analysis*, 49 :39–67, 2015.