

## Chapitre 2

# Méthodes de différences finies et volumes finis pour les problèmes elliptiques et paraboliques

### 2.1 Principe des deux méthodes

#### 2.1.1 Cas de la dimension 1

On considère le problème unidimensionnel

$$-u''(x) = f(x), \quad \forall x \in ]0, 1[, \quad (2.1.1)$$

$$u(0) = u(1) = 0, \quad (2.1.2)$$

où  $f \in C([0, 1])$ . Les conditions aux limites (2.1.2) considérées ici sont dites de type Dirichlet homogène (le terme homogène désigne les conditions nulles). Cette équation modélise par exemple la diffusion de la chaleur dans un barreau conducteur chauffé (terme source  $f$ ) dont les deux extrémités sont plongées dans de la glace.

#### Méthode de différences finies.

Soit  $(x_k)_{k=0, \dots, N+1}$  une subdivision de  $[0, 1]$ , avec :

$$x_0 = 0 < x_1 < x_2 < \dots < x_N < x_{N+1} = 1.$$

Pour  $i = 0, \dots, N$ , on note  $h_{i+1/2} = x_{i+1} - x_i$  et on définit le "pas" du maillage par :

$$h = \max_{i=0, \dots, N} h_{i+1/2}. \quad (2.1.3)$$

Pour simplifier l'exposé, on se limitera dans un premier temps à un pas constant :

$$h_{i+1/2} = h \quad \forall i \in [0, N].$$

On écrit l'équation aux dérivées partielles (2.1.1) aux points  $x_i$

$$-u''(x_i) = f(x_i), \quad \forall i = 1, \dots, N,$$

Effectuons un développement de Taylor en  $x_i$  :

$$u(x_{i+1}) = u(x_i) + hu'(x_i) + \frac{h^2}{2}u''(x_i) + \frac{h^3}{6}u'''(x_i) + \frac{h^4}{24}u^{(4)}(\zeta_i),$$

$$u(x_{i-1}) = u(x_i) - hu'(x_i) + \frac{h^2}{2}u''(x_i) - \frac{h^3}{6}u'''(x_i) + \frac{h^4}{24}u^{(4)}(\eta_i),$$

avec  $\zeta_i \in [x_i, x_{i+1}]$ ,  $\eta_i \in [x_{i-1}, x_i]$ . En additionnant, on obtient :

$$u(x_{i+1}) + u(x_{i-1}) = 2u(x_i) + h^2u''(x_i) + O(h^2)$$

Il semble donc raisonnable d'approcher la dérivée seconde  $-u''(x_i)$  par le "quotient différentiel"

$$\frac{2u(x_i) - u(x_{i-1}) - u(x_{i+1})}{h^2}.$$

Sous des hypothèses de régularité sur  $u$ , on peut montrer (voir lemme 2.12 page 16) que cette approximation est d'ordre 2 au sens

$$R_i = u''(x_i) + \frac{2u(x_i) - u(x_{i-1}) - u(x_{i+1})}{h^2} = O(h^2)$$

On appelle erreur de consistance au point  $x_i$  la quantité  $R_i$ .

### Méthode des volumes finis.

On ne se donne plus des points mais des volumes de contrôle  $K_i$ ,  $i = 1, \dots, N$ , avec  $K_i = ]x_{i-1/2}, x_{i+1/2}[$ , et on note  $h_i = x_{i+1/2} - x_{i-1/2}$ . Pour chaque volume de contrôle  $K_i$ , on se donne un point  $x_i \in K_i = ]x_{i-1/2}, x_{i+1/2}[$ . On pourra considérer par exemple (mais ce n'est pas le seul point possible) :  $x_i = 1/2(x_{i+1/2} + x_{i-1/2})$ . On intègre l'équation  $-u'' = f$  sur  $K_i$  :

$$\int_{x_{i-1/2}}^{x_{i+1/2}} -u''(x)dx = \int_{x_{i-1/2}}^{x_{i+1/2}} f(x)dx$$

et  $f_i = \frac{1}{h_i} \int_{x_{i-1/2}}^{x_{i+1/2}} f(x)dx$ . On obtient :

$$-u'(x_{i+1/2}) + u'(x_{i-1/2}) = h_i f_i, \quad i = 1, \dots, N.$$

On cherche donc à approcher les flux  $-u'(x_{i+1/2})$  aux interfaces  $x_{i+1/2}$  des mailles. Notons que l'opérateur à approcher est ici d'ordre 1, alors qu'il était d'ordre 2 en différences finies pour la même équation.

On se donne une inconnue par maille (ou volume de contrôle  $i$ ), qu'on note  $u_i$ , et on espère approcher ainsi la valeur  $u(x_i)$  (ou  $\frac{1}{h_i} \int_{K_i} u$ ). On approche  $u'(x_{i+1/2})$  par le quotient différentiel

$$\frac{u(x_{i+1}) - u(x_i)}{h_{i+1/2}}.$$

Le schéma numérique s'écrit donc :

$$-\frac{u_{i+1} - u_i}{h_{i+1/2}} + \frac{u_i - u_{i-1}}{h_{i-1/2}} = h_i f_i \quad i = 2, \dots, N-1. \quad (2.1.4)$$

Pour la première et  $N$ -ième équations, on tient compte des conditions aux limites (2.1.2), et on  $u'(0)$  (resp.  $u'(1)$ ) par  $\frac{u(x_1)}{h_{1/2}}$  (resp.  $\frac{u(x_N)}{h_{N+1/2}}$ ), ce qui donne comme première et dernière équations du schéma numérique :

$$-\frac{u_2 - u_1}{h_{3/2}} + \frac{u_1}{h_{1/2}} = h_1 f_1, \quad (2.1.5)$$

$$-\frac{u_N - u_{N-1}}{h_{N-1/2}} - \frac{u_N}{h_{N+1/2}} = h_N f_N, \quad (2.1.6)$$

**Remarque 2.1** Si le pas du maillage est constant :  $h_i = h, \forall i = 1, \dots, N$  (on dit aussi que le maillage est régulier), on peut montrer (exercice 1 page 46) que les équations des schémas volumes finis et différences finies aux conditions de bord et au second membre près. Si le maillage n'est pas régulier, ceci n'est plus vérifié.

### 2.1.2 Cas de la dimension 2 ou 3

On considère maintenant le problème (1.3.2) en dimension 2 ou 3, sur un ouvert borné  $\Omega$  de  $\mathbb{R}^d$ ,  $d = 2$  ou 3, avec conditions aux limites de Dirichlet homogènes qui s'écrivent maintenant :

$$u(x) = 0, \forall x \in \partial\Omega, \quad (2.1.7)$$

où  $\partial\Omega$  désigne la frontière de  $\Omega$ .

#### Méthode de différences finies.

Supposons (pour simplifier) que le domaine  $\Omega$  soit un carré (c.à.d.  $d = 2$ , le cas rectangulaire se traite tout aussi facilement). On se donne un pas de maillage constant  $h$  et des points  $x_{i,j} = (ih, jh)$ ,  $i = 1, \dots, N$ ,  $j = 1, \dots, N$ . En effectuant les développements limités de Taylor (comme au paragraphe 2.1.1 page 8) dans les deux directions (voir exercice 13), on approche  $-\partial_i^2 u(x_{i,j})$  (resp.  $-\partial_j^2 u(x_{i,j})$ ) par

$$\frac{2u(x_{i,j}) - u(x_{i+1,j}) - u(x_{i-1,j})}{h^2} \text{ (resp. par } \frac{2u(x_{i,j}) - u(x_{i,j+1}) - u(x_{i,j-1})}{h^2}).$$

Ce type d'approche est limité à des géométries simples. Pour mailler des géométries compliquées, il est en général plus facile d'utiliser des triangles (tétraèdres en dimension 3), auquel cas la méthode des différences finies est plus difficile à généraliser.

#### Méthode de volumes finis.

On suppose maintenant que  $\Omega$  est un ouvert polygonal de  $\mathbb{R}^2$ , et on se donne un maillage  $\mathcal{T}$  de  $\Omega$ , c.à.d., en gros, un découpage de  $\Omega$  en volumes de contrôle polygonaux  $K$ . En intégrant l'équation (1.3.2) sur  $K$ , on obtient :

$$\int_K -\Delta u dx = \int_K f dx.$$

Par la formule de Stokes, on peut réécrire cette équation :

$$-\int_{\partial K} \nabla u(x) \cdot \mathbf{n}_K(x) d\gamma(x) = \int_K f(x) dx,$$

où  $d\gamma(x)$  désigne l'intégrale par rapport à la mesure uni-dimensionnelle sur le bord de l'ouvert  $\Omega$ , et où  $\mathbf{n}_K$  désigne le vecteur normal unitaire à  $\partial K$  extérieur à  $K$ . Comme  $K$  est polygonal, on peut décomposer  $\partial K$  en arêtes  $\sigma$  qui sont des segments de droite, et en appelant  $\mathcal{E}_K$  l'ensemble des arêtes de  $\partial K$ , on a donc :

$$-\sum_{\sigma \in \mathcal{E}_K} \int_{\sigma} \nabla u \cdot \mathbf{n}_{K,\sigma} d\gamma(x) = \int_K f(x) dx,$$

où  $\mathbf{n}_{K,\sigma}$  désigne le vecteur normal unitaire à  $\sigma$  extérieur à  $K$  (noter que ce vecteur est constant sur  $\sigma$ ). On cherche donc maintenant à approcher la dérivée normale  $\nabla u \cdot \mathbf{n}_{K,\sigma}$  de manière consistante sur chaque arête  $\sigma$ . On se donne donc des inconnues discrètes notées  $(u_K)_{K \in \mathcal{T}}$ , qui, on l'espère vont s'avérer être des approximations de  $u(x_K)$ . Pour une arête  $\sigma = K|L$  séparant les volumes de contrôle  $K$  et  $L$ , il est tentant d'approcher la dérivée normale  $\nabla u \cdot \mathbf{n}_{K,\sigma}$  par le quotient différentiel

$$\frac{u(x_L) - u(x_K)}{d_{K,L}},$$

où  $d_{K,L}$  est la distance entre les points  $x_K$  et  $x_L$ . Cependant, cette approximation ne pourra être justifiée que si la direction du vecteur défini par les deux points  $x_K$  et  $x_L$  est la même que celle de la normale  $\mathbf{n}_{K,\sigma}$ , c.à.d. si le segment de droite  $x_K x_L$  est orthogonal à l'arête  $K|L$ . Pour un maillage triangulaire à angles strictement inférieurs à  $\pi/2$ , ceci est facile à obtenir en choisissant les points  $x_K$  comme intersection des médiatrices du triangle  $K$ , voir Figure 2.1.

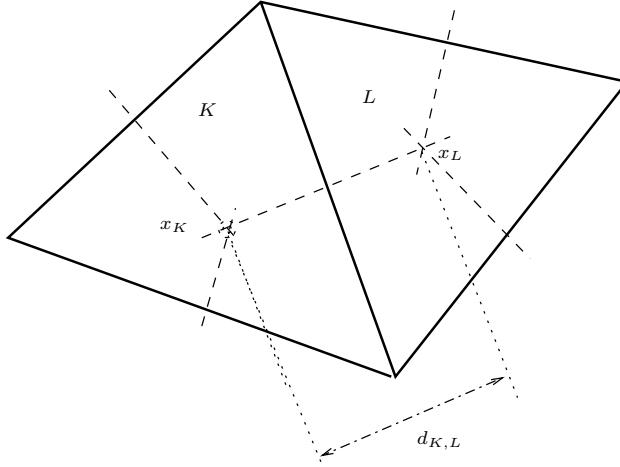


FIG. 2.1 – Exemple de volumes de contrôle pour la méthode des volumes finis en deux dimensions d'espace

On se placera ici dans ce cas, et on verra plus loin d'autres possibilités. on approche donc  $\nabla u \cdot \mathbf{n}_K|_{\sigma}$  par  $\frac{u(x_L) - u(x_K)}{d_{K,L}}$  et en notant  $|\sigma|$  la longueur de l'arête  $\sigma$ , on approche :

$$\int_{\sigma} \nabla u \cdot \mathbf{n}_K d\gamma \text{ par } F_{K,\sigma} = |\sigma| \frac{u_L - u_K}{d_{K,L}}, \text{ pour tout } \sigma \in \mathcal{E}_K \text{ et pour tout } K \in \mathcal{T}.$$

Le schéma volumes finis s'écrit donc

$$\sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma} = |K|f_K, \quad (2.1.8)$$

où  $|K|$  est la mesure de  $K$ , et  $f_K = \frac{1}{|K|} \int_K f(x)dx$ , et où les flux numériques  $F_{K,\sigma}$  sont définis (en tenant compte des conditions limites pour les arêtes du bord) par :

$$F_{K,\sigma} = \begin{cases} -|\sigma| \frac{u_L - u_K}{d_{K,L}} & \text{si } \sigma = K|L, \\ -|\sigma| \frac{u_K}{d_{K,\sigma}} & \text{si } \sigma \subset \partial\Omega \text{ et } \sigma \in \mathcal{E}_K, \end{cases} \quad (2.1.9)$$

où  $d_{K,\sigma}$  = distance entre  $x_K$  et  $\sigma$

### Comparaison des méthodes

Cette introduction aux différences finies et volumes finis nous permet de remarquer que les différences finies sont particulièrement bien adaptées dans le cas de domaines rectangulaires ou parallélépipédiques, pour lesquels on peut facilement définir des maillages structurés (cartésiens dans le cas présent) c.à.d. dont on peut indexer les mailles par un ordre  $(i, j)$  naturel.

Dans le cas de domaines plus complexes, on maille souvent à l'aide de triangles (ou tétraèdres) et dans ce cas la méthode des différences finies ne se généralise pas facilement. On a alors recours soit aux volumes finis, dont on vient de donner le principe, soit aux éléments finis, que nous aborderons ultérieurement.

### 2.1.3 Questions d'analyse numérique

Voici un certain nombre de questions, qui sont typiquement du domaine de l'analyse numérique, auxquelles nous tenterons de répondre dans la suite :

1. Le problème qu'on a obtenu en dimension finie, (avec des inconnues localisées aux noeuds du maillage dans le cas de la méthode des différences finies et dans les mailles dans le cas de la méthode des volumes finis) admet-il une (unique) solution ? On montrera que oui.
2. La solution du problème discret converge-t-elle vers la solution du problème continu lorsque le pas du maillage  $h$  tend vers 0 ? Dans le cas des différences finies en une dimension d'espace, le pas du maillage est défini par

$$h = \sup_{i=1 \dots N} |x_{i+1} - x_i|. \quad (2.1.10)$$

Dans le cas des volumes finis en une dimension d'espace, il est défini par :

$$h = \sup_{i=1 \dots N} |x_{i+1/2} - x_{i-1/2}|. \quad (2.1.11)$$

en deux dimensions d'espace, le pas  $h$  est défini par

$$h = \sup_{K \in \mathcal{T}} \text{diam}(K), \text{ avec } \text{diam}(K) = \sup_{x,y \in K} d(x,y),$$

où  $\mathcal{T}$ , le maillage, est l'ensemble des volumes de contrôle  $K$ . Notons que la réponse à cette question n'est pas évidente *a priori*. La solution discrète peut converger vers la solution continue, elle peut aussi converger mais vers autre chose que la solution du problème continu, et enfin elle peut ne pas converger du tout.

## 2.2 Etude de la méthode différences finies pour un problème elliptique unidimensionnel

On cherche à discréteriser le problème aux limites, suivant :

$$\begin{cases} -u''(x) + c(x)u(x) = f(x), & 0 < x < 1, \\ u(0) = u(1) = 0, \end{cases} \quad (2.2.12)$$

où  $c \in C([0, 1], \mathbb{R}_+)$ , et  $c \in C([0, 1], \mathbb{R})$ , qui peut modéliser par exemple un phénomène de diffusion - réaction d'une espèce chimique. On se donne un pas du maillage constant  $h = \frac{1}{N+1}$ , et une subdivision de  $]0, 1[$ , notée  $(x_k)_{k=0, \dots, N+1}$ , avec :  $x_0 = 0 < x_1 < x_2 < \dots < x_N < x_{N+1} = 1$ . Soit  $u_i$  l'inconnue discrète associée au noeud  $i$  ( $i = 1, \dots, N$ ). On pose  $u_0 = u_{N+1} = 0$ . On obtient les équations discrètes en approchant  $u''(x_i)$  par quotient différentiel par développement de Taylor, comme on l'a vu au paragraphe 2.1.1 page 8.

$$\begin{cases} \frac{1}{h^2}(2u_i - u_{i-1} - u_{i+1}) + c_i u_i = f_i, & i = 1, \dots, N, \\ u_0 = u_{N+1} = 0. \end{cases} \quad (2.2.13)$$

avec  $c_i = c(x_i)$  et  $f_i = f(x_i)$ . On peut écrire ces équations sous forme matricielle :

$$A_h U_h = b_h, \text{ avec } U_h = \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix} \text{ et } b_h = \begin{pmatrix} f_1 \\ \vdots \\ f_N \end{pmatrix} \quad (2.2.14)$$

$$\text{et } A_h = \begin{pmatrix} 2 + c_1 h^2 & -1 & 0 & \dots & 0 \\ -1 & 2 + c_2 h^2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 + c_{N-1} h^2 & -1 \\ 0 & \dots & 0 & -1 & 2 + c_N h^2 \end{pmatrix}. \quad (2.2.15)$$

Les questions suivantes surgissent alors naturellement :

1. Le système (2.2.14) admet-il un unique solution ?
2. A-t-on convergence de  $U_h$  vers  $u$  et en quel sens ?

Nous allons répondre par l'affirmative à ces deux questions. Commençons par la première.

**Proposition 2.2** Soit  $c = (c_1, \dots, c_N)^t \in \mathbb{R}^N$  tel que  $c_i \geq 0$  pour  $i = 1, \dots, N$  ; alors la matrice  $A_h$  définie par (2.2.15) est symétrique définie positive, et donc inversible.

**Démonstration :** La matrice  $A_h$  est évidemment symétrique. Montrons qu'elle est définie positive. Soit  $v = (v_1 \dots v_N)^t$ , on pose  $v_0 = v_{N+1} = 0$ . Calculons le produit scalaire  $A_h v \cdot v = v^t A_h v$ . On a :

$$A_h v \cdot v = \frac{1}{h^2} (v_1 \dots v_N) \begin{pmatrix} 2 + c_1 h^2 & -1 & 0 \\ -1 & \ddots & \ddots \\ 0 & \ddots & -1 & 2 + c_N h^2 \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix},$$

c'est-à-dire :

$$A_h v \cdot v = \frac{1}{h^2} \sum_{i=1}^N v_i (-v_{i-1} + (2 + c_i h^2) v_i - v_{i+1}).$$

On a donc, par changement d'indice :

$$A_h v \cdot v = \frac{1}{h^2} \left[ \sum_{i=1}^N (-v_{i-1} v_i) + \sum_{i=1}^N (2 + c_i h^2) v_i^2 - \sum_{j=2}^{N+1} v_{j-1} v_j \right].$$

Et comme on a posé  $v_0 = 0$  et  $v_{N+1} = 0$ , on peut écrire :

$$A_h v \cdot v = \frac{1}{h^2} \sum_{i=1}^N (2 + c_i h^2) v_i^2 + \frac{1}{h^2} \sum_{i=1}^N (-2 v_i v_{i-1}),$$

soit encore :

$$A_h v \cdot v = \sum_{i=1}^N c_i v_i^2 + \frac{1}{h^2} \sum_{i=1}^N (-2 v_i v_{i-1} + v_i^2 + v_{i-1}^2) + v_N^2.$$

On a donc finalement :

$$A_h v \cdot v = \sum_{i=1}^N c_i v_i^2 + \frac{1}{h^2} \sum_{i=1}^N (v_i - v_{i-1})^2 + v_N^2 \geq 0, \quad \forall v = (v_1, \dots, v_N) \in \mathbb{R}^N.$$

Si on suppose  $A_h v \cdot v = 0$ , on a alors

$$\sum_{i=1}^N c_i h^2 v_i^2 = 0 \text{ et } v_i - v_{i-1} = 0, \quad \forall i = 1 \dots N.$$

On a donc  $v_1 = v_2 = \dots = v_N = v_0 = v_{N+1} = 0$ . Remarquons que ces égalités sont vérifiées même si les  $c_i$  sont nuls. Ceci démontre que la matrice  $A_h$  est bien définie.

**Remarque 2.3 (Existence et unicité de la solution)** *On a montré ci-dessus que  $A_h$  est symétrique définie positive, donc inversible, ce qui entraîne l'existence et l'unicité de la solution de (2.2.14). On aurait pu aussi démontrer l'existence et l'unicité de la solution de (2.2.14) directement, en montrant que  $\text{Ker}(A_h) = 0$  (voir exercice 2 page 46). On rappelle qu'en dimension finie, toute application linéaire injective ou surjective est bijective. On en déduit ainsi l'existence de la solution du système (2.2.14).*

**Remarque 2.4 (Caractère défini et conditions limites)** *Dans la démonstration de la proposition 2.2, si  $c_i > 0$  pour tout  $i = 1, \dots, N$  le terme  $\sum_{i=1}^N c_i h^2 v_i^2 = 0$  permet de conclure que  $v_i = 0$  pour tout  $i = 1, \dots, N$ . Par contre, si  $c_i \geq 0$  (ou même  $c_i = 0$  pour tout  $i = 1, \dots, N$ , c'est grâce aux conditions au limites de Dirichlet homogènes (représentées par le fait qu'on pose  $v_0 = 0$  et  $v_{N+1} = 0$  ce qui permet d'écrire alors les équations 1 et  $N$  sous la même forme que l'équation i) qu'on peut montrer que que  $v_i = 0$ , pour tout  $i = 1, \dots, N$ , car  $v_i = v_{i-1}$ , pour tout  $i = 1, \dots, N$ , et  $v_0 = 0$ . En particulier, la matrice de discrétisation de  $-u''$  par différences finies avec conditions aux limites de Neumann homogènes :*

$$\begin{cases} -u'' = f, \\ u'(0) = u'(1) = 0. \end{cases} \quad (2.2.16)$$

*donne une matrice  $A_h$  qui est symétrique et positive, mais non définie (voir exercice 10 page 49). De fait la solution du problème continu (2.2.16) n'est pas unique, puisque les fonctions constantes sur  $[0, 1]$  sont solutions de (2.2.16).*

Nous allons maintenant nous préoccuper de la question de la convergence.

**Définition 2.5 (Matrices monotones)** Soit  $A \in \mathcal{M}_N(\mathbb{R})$ ,  $A = (a_{ij})_{i=1,\dots,N, j=1,\dots,N}$ . On dit que  $A$  est positive (ou  $A \geq 0$ ) si  $a_{ij} \geq 0$ ,  $\forall i, j = 1, \dots, N$ . On dit que  $A$  est monotone si  $A$  est inversible et  $A^{-1} \geq 0$ .

L'avantage des schémas à matrices monotones est de satisfaire la propriété de conservation de la positivité, qui peut être cruciale dans les applications physiques :

**Définition 2.6 (Conservation de la positivité)** Soit  $A \in \mathcal{M}_N(\mathbb{R})$ ,  $A = (a_{ij})_{i=1,\dots,N, j=1,\dots,N}$  ; on dit que  $A$  conserve la positivité si  $Av \geq 0$  entraîne  $v \geq 0$  (les inégalités s'entendent composante par composante).

On a en effet la proposition suivante :

**Proposition 2.7 (Monotonie et conservation de la positivité)** Soit  $A \in \mathcal{M}_N(\mathbb{R})$ . Alors  $A$  conserve la positivité si et seulement si  $A$  est monotone.

**Démonstration :** Supposons d'abord que  $A$  conserve la positivité, et montrons que  $A$  inversible et que  $A^{-1}$  a des coefficients  $\geq 0$ . Si  $x$  est tel que  $Ax = 0$ , alors  $Ax \geq 0$  et donc, par hypothèse,  $x \geq 0$ . Mais on a aussi  $Ax \leq 0$ , soit  $A(-x) \geq 0$  et donc par hypothèse,  $x \leq 0$ . On en déduit  $x = 0$ , ce qui prouve que  $A$  est inversible. La conservation de la positivité donne alors que  $y \geq 0 \Rightarrow A^{-1}y \geq 0$ . En prenant  $y = e_1$  on obtient que la première colonne de  $A^{-1}$  est positive, puis en prenant  $y = e_i$  on obtient que la  $i$ -ème colonne de  $A^{-1}$  est positive, pour  $i = 2, \dots, N$ . Donc  $A^{-1}$  a tous ses coefficients positifs.

Réciproquement, supposons maintenant que  $A$  est inversible et que  $A^{-1}$  a des coefficients positifs. Soit  $x \in \mathbb{R}^N$  tel que  $Ax = y \geq 0$ , alors  $x = A^{-1}y \geq 0$ . Donc  $A$  conserve la positivité. ■

**Remarque 2.8 (Principe du maximum)** On appelle principe du maximum continu le fait que si  $f \geq 0$  alors le minimum de la fonction  $u$  solution du problème (2.2.12) page 13 est atteint sur les bords. Cette propriété mathématique correspond à l'intuition physique qu'on peut avoir du phénomène : si on chauffe un barreau tout en maintenant ses deux extrémités à une température fixe, la température aux points intérieurs du barreau sera supérieure à celle des extrémités. Il est donc souhaitable que la solution approchée satisfasse la même propriété (voir exercice 5 page 48 à ce sujet).

**Lemme 2.9** Soit  $c = (c_1, \dots, c_N)^t \in \mathbb{R}^N$ , et  $A_h \in \mathcal{M}_N(\mathbb{R})$  définie par (2.2.15). Si  $c_i \geq 0$  pour tout  $i = 1, \dots, N$ , alors  $A_h$  est monotone.

**Démonstration :** On va montrer que si  $v \in \mathbb{R}^N$ ,  $A_h v \geq 0$  alors  $v \geq 0$ . On peut alors utiliser la proposition 2.7 pour conclure. Soit  $v = (v_1, \dots, v_N)^t \in \mathbb{R}^N$ . Posons  $v_0 = v_{N+1} = 0$ . Supposons que  $A_h v \geq 0$ . On a donc

$$-\frac{1}{h^2}v_{i-1} + \left(\frac{2}{h^2} + c_i\right)v_i - \frac{1}{h^2}v_{i+1} \geq 0, \quad i = 1, \dots, N \quad (2.2.17)$$

Soit

$$p = \min \left\{ i \in \{1, \dots, N\}; v_p = \min_{j=1, \dots, N} v_j \right\}.$$

Supposons que  $\min_{j=1, \dots, N} v_j < 0$ . On a alors  $p \geq 1$  et :

$$\frac{1}{h^2}(v_p - v_{p-1}) + c_p v_p + \frac{1}{h^2}(v_p - v_{p+1}) \geq 0.$$

On en déduit que

$$\frac{2}{h^2}c_p v_p \geq \frac{1}{h^2}(v_{p-1} - v_p) + \frac{1}{h^2}(v_{p+1} - v_p) \geq 0.$$

Si  $c_p > 0$ , on a donc  $v_p \geq 0$ , et donc  $v_i \geq 0$ ,  $\forall i = 1, \dots, N$ . Si  $c_p = 0$ , on doit alors avoir  $v_{p-1} = v_p = v_{p+1}$  ce qui est impossible car  $p$  est le plus petit indice  $j$  tel que  $v_j = \min_{i=1, \dots, N} v_i$ . Donc dans ce cas le minimum ne peut pas être atteint pour  $j = p > 1$ . On a ainsi finalement montré que  $\min_{i \in \{1, \dots, N\}} v_i \geq 0$ , on a donc  $v \geq 0$ .  $\blacksquare$

**Définition 2.10 (Erreur de consistance)** *On appelle erreur de consistance la quantité obtenue en remplaçant l'inconnue par la solution exacte dans le schéma numérique. Dans le cas du schéma (2.2.13), l'erreur de consistance au point  $x_i$  est donc définie par :*

$$R_i = \frac{1}{h^2}(2u(x_i) - u(x_{i-1}) - u(x_{i+1})) + c(x_i)u(x_i) - f(x_i). \quad (2.2.18)$$

L'erreur de consistance  $R_i$  est donc l'erreur qu'on commet en remplaçant l'opérateur  $-u''$  par le quotient différentiel

$$\frac{1}{h^2}(2u(x_i) - u(x_{i-1}) - u(x_{i+1})).$$

Cette erreur peut être évaluée si  $u$  est suffisamment régulière, en effectuant des développements de Taylor.

**Définition 2.11 (Ordre du schéma)** *On dit qu'un schéma de discréétisation à  $N$  points de discréétisation est d'ordre  $p$  s'il existe  $C \in \mathbb{R}$ , ne dépendant que de la solution exacte, tel que l'erreur de consistance satisfasse :*

$$\max_{i=1, \dots, N} (R_i) < ch^p,$$

où  $h$  est le pas du maillage défini par (2.1.3) (c.à.d. le maximum des écarts  $x_{i+1} - x_i$ ). On dit qu'un schéma de discréétisation est consistant si

$$\max_{i=1, \dots, N} (R_i) \rightarrow 0 \text{ lorsque } h \rightarrow 0,$$

où  $N$  est le nombre de points de discréétisation.

**Lemme 2.12** *Si la solution de (2.2.12) vérifie  $u \in C^4([0, 1])$ , alors le schéma (2.2.13) est consistant d'ordre 2, et on a plus précisément :*

$$|R_i| \leq \frac{h^2}{12} \sup_{[0,1]} |u^{(4)}|, \quad \forall i = 1, \dots, N. \quad (2.2.19)$$

**Démonstration :** Par développement de Taylor, on a :

$$\begin{aligned} u(x_{i+1}) &= u(x_i) + hu'(x_i) + \frac{h^2}{2}u''(x_i) + \frac{h^3}{6}u'''(x_i) + \frac{h^4}{24}u^{(4)}(\xi_i) \\ u(x_{i-1}) &= u(x_i) - hu'(x_i) + \frac{h^2}{2}u''(x_i) - \frac{h^3}{6}u'''(x_i) + \frac{h^4}{24}u^{(4)}(\eta_i) \end{aligned}$$

En additionnant ces deux égalités, on obtient que :

$$\frac{1}{h^2}(u(x_{i+1}) + u(x_i) - 2u(x_i)) = u''(x_i) + \frac{h^2}{24}(u^{(4)}(\xi_i) + u^{(4)}(\eta_i)),$$

ce qui entraîne que :

$$|R_i| \leq \frac{h^2}{12} \sup_{[0,1]} |u^{(4)}|. \quad (2.2.20)$$

■

**Remarque 2.13 (Sur l'erreur de consistance)**

1. Si on note  $\bar{U}_h : (u(x_i))_{i=1 \dots N}$  le vecteur dont les composantes sont les valeurs exactes de la solution de (2.2.12), et  $U_h = (u_1 \dots u_N)^t$  la solution de (2.2.13), on a :

$$R = A_h(U_h - \bar{U}_h). \quad (2.2.21)$$

2. On peut remarquer que si  $u^{(4)} = 0$ , les développements de Taylor effectués ci-dessus se résument à :

$$-u''(x_i) = \frac{2u(x_i) - u(x_{i-1}) - u(x_{i+1})}{h^2},$$

et on a donc  $R_i = 0$ , pour tout  $i = 1, \dots, N$ , et donc  $u_i = u(x_i)$ , pour tout  $i = 1 \dots N$ . Dans ce cas (rare !), le schéma de discréétisation donne la valeur exacte de la solution en  $x_i$ , pour tout  $i = 1, \dots, N$ . Cette remarque est bien utile lors de la phase de validation de méthodes et numériques et/ou programmes informatiques pour la résolution de l'équation (2.2.12). En effet, si on choisit  $f$  telle que la solution soit un polynôme de degré inférieur ou égal à 3, alors on doit avoir une erreur entre solution exacte et approchée inférieure à l'erreur machine.

La preuve de convergence du schéma utilise la notion de consistante, ainsi qu'une notion de stabilité, que nous introduisons maintenant :

**Proposition 2.14** On dit que le schéma (2.2.13) est stable, au sens où la matrice de discréétisation  $A_h$  satisfait :

$$\|A_h^{-1}\|_\infty \leq \frac{1}{8}. \quad (2.2.22)$$

On peut réécrire cette inégalité comme une estimation sur les solutions du système (2.2.14) :

$$\|U_h\| \leq \frac{1}{8} \|f\|_\infty. \quad (2.2.23)$$

**Démonstration :** On rappelle que par définition, si  $M \in \mathcal{M}_N(\mathbb{R})$ ,

$$\|M\|_\infty = \sup_{\substack{v \in \mathbb{R}^N \\ v \neq 0}} \frac{\|M\|_\infty}{\|v\|_\infty}, \text{ avec } \|v\|_\infty = \sup_{i=1, \dots, N} |v_i|.$$

Pour montrer que  $\|A_h^{-1}\|_\infty \leq \frac{1}{8}$ , on décompose la matrice  $A_h$  sous la forme  $A_h = A_{0h} + \text{diag}(c_i)$  où  $A_{0h}$  est la matrice de discréétisation de l'opérateur  $-u''$  avec conditions aux limites de Dirichlet homogènes, et

$$A_{0h} = \begin{bmatrix} \frac{2}{h^2} & -\frac{1}{h^2} & & & & 0 & & \\ -\frac{1}{h^2} & \ddots & & & & & & \\ & \ddots & \ddots & & & & & \\ & & & \ddots & & & & -\frac{1}{h^2} \\ 0 & & & & -\frac{1}{h^2} & & \frac{1}{h^2} & \end{bmatrix} \quad (2.2.24)$$

et  $diag(c_i)$  désigne la matrice diagonale de coefficients diagonaux  $c_i$ . Les matrices  $A_{0h}$  et  $A_h$  sont inversibles, et on a :

$$A_{0h}^{-1} - A_h^{-1} = A_{0h}^{-1} A_h A_h^{-1} - A_{0h}^{-1} A_{0h} A_h^{-1} = A_{0h}^{-1} (A_h - A_{0h}) A_h^{-1}.$$

Comme  $diag(c_i) \geq 0$ , on a  $A_h \geq A_{0h}$ , et comme  $A_{0h}$  et  $A_h$  sont monotones, on en déduit que :

$$0 \leq A_h^{-1} \leq A_{0h}^{-1}, \text{ (composante par composante).}$$

On peut maintenant remarquer que si  $B \in \mathcal{M}_N(\mathbb{R})$ , et si  $B \geq 0$  (c.à.d.  $B_{ij} \geq 0$  pour tout  $i$  et  $j$ ), on a

$$\|B\|_\infty = \sup_{\substack{v \in \mathbb{R}^N \\ \|v\|=1}} \sup_{i=1,\dots,N} |(Bv)_i| = \sup_{\substack{v \in \mathbb{R}^N \\ \|v\|=1}} \sup_{i=1,\dots,N} \left| \sum_{j=1}^N B_{ij} v_j \right| \|B\|_\infty = \sup_{i=1,\dots,N} \sum_{j=1}^N B_{ij}.$$

On a donc  $\|A_h^{-1}\| = \sup_{i=1,\dots,N} \sum_{j=1}^N (A_h^{-1})_{ij} \leq \sup_{i=1,\dots,N} \sum_{j=1}^N (A_{0h}^{-1})_{ij}$  car  $A_h^{-1} \leq A_{0h}^{-1}$  ; d'où on déduit que  $\|A_h^{-1}\|_\infty \leq \|A_{0h}^{-1}\|_\infty$ . Il ne reste plus qu'à estimer  $\|A_{0h}^{-1}\|_\infty$ . Comme  $A_{0h}^{-1} \geq 0$ , on a

$$\|A_{0h}^{-1}\|_\infty = \|A_{0h}^{-1} e\|_\infty \text{ avec } e = (1, \dots, 1)^t.$$

Soit  $d = A_{0h}^{-1} e \in \mathbb{R}^N$ . On veut calculer  $\|d\|_\infty$ , où  $d$  vérifie  $A_{0h}d = e$ . Or le système linéaire  $A_{0h}d = e$  n'est autre que la discrétisation par différences finies du problème

$$\begin{cases} -u'' = 1 \\ u(0) = u(1) = 0 \end{cases}, \quad (2.2.25)$$

dont la solution exacte est :

$$u_0(x) = \frac{x(1-x)}{2},$$

qui vérifie  $u_0^{(4)}(x) = 0$ . On en conclut, par la remarque 2.13, que

$$u_0(x_i) = d_i, \quad \forall i = 1 \dots N.$$

Donc  $\|d\|_\infty = \sup_{i=1,N} \frac{ih(ih-1)}{2}$  où  $h = \frac{1}{N+1}$  est le pas de discrétisation. Ceci entraîne que

$$\|d\|_\infty \leq \sup_{[0,1]} \left| \frac{x(x-1)}{2} \right| = \frac{1}{8}, \text{ et donc que } \|A_{0h}^{-1}\|_\infty \leq \frac{1}{8}.$$

■

**Remarque 2.15 (Sur la stabilité)** Noter que l'inégalité (2.2.23) donne une estimation sur les solutions approchées indépendantes du pas de maillage. C'est ce type d'estimation qu'on recherchera par la suite pour la discrétisation d'autres problèmes comme garant de la stabilité d'un schéma numérique.

**Définition 2.16 (Erreur de discrétisation)** On appelle erreur de discrétisation en  $x_i$ , la différence entre la solution exacte en  $x_i$  et la  $i$ -ème composante de la solution donnée par le schéma numérique

$$e_i = u(x_i) - u_i, \quad \forall i = 1, \dots, N. \quad (2.2.26)$$

**Théorème 2.17** Soit  $u$  la solution exacte de

$$\begin{cases} -u'' + cu = f, \\ u(0) = u(1) = 0. \end{cases}$$

On suppose  $u \in C^4([0, 1])$ . Soit  $u_h$  la solution de (2.2.13). Alors l'erreur de discréétisation définie par (2.2.26) satisfait

$$\max_{i=1, \dots, N} |e_i| \leq \frac{1}{96} \|u^{(4)}\|_\infty h^2.$$

Le schéma est donc convergent d'ordre 2.

**Démonstration :** Soit  $U_h = (U_1, \dots, U_n)^t$  et  $\bar{U}_h = (u(x_1), \dots, u(x_N))^t$ , on cherche à majorer  $\|\bar{U}_h - U_h\|_\infty$ . On a  $A(\bar{U}_h - U_h) = R$  où  $R$  est l'erreur de consistance (voir remarque 2.13). On a donc

$$\|\bar{U}_h - U_h\|_\infty \leq \|A_h^{-1}\|_\infty \|R\|_\infty \leq \frac{1}{8} \times \frac{1}{12} \|u^{(4)}\|_\infty = \frac{1}{96} \|u^{(4)}\|_\infty$$

■

**Remarque 2.18 (Sur la convergence)** On peut remarquer que la preuve de la convergence s'appuie sur la stabilité (elle-même déduite de la conservation de la positivité) et sur la consistance. Dans certains livres d'analyse numérique, vous trouverez la "formule" : stabilité + consistance  $\implies$  convergence. Il faut toutefois prendre garde au fait que ces notions de stabilité et convergence peuvent être variables d'un type de méthode à un autre (comme nous le verrons en étudiant la méthode des volumes finis, par exemple).

**Remarque 2.19 (Contrôle des erreurs d'arrondi)** On cherche à calculer la solution approchée de  $-u'' = f$ . Le second membre  $f$  est donc une donnée du problème. Supposons que des erreurs soient commises sur cette donnée (par exemple des erreurs d'arrondi, ou des erreurs de mesure). On obtient alors un nouveau système, qui s'écrit  $A_h \tilde{U}_h = b_h + \varepsilon_h$ , où  $\varepsilon_h$  représente la discréétisation des erreurs commises sur le second membre. Si on résout  $A_h \tilde{U}_h = b_h + \varepsilon_h$  au lieu de  $A_h U_h = b_h$ , l'erreur commise sur la solution du système s'écrit

$$E_h = \tilde{U}_h - U_h = A_h^{-1} \varepsilon_h.$$

On en déduit que

$$\|E_h\|_\infty \leq \frac{1}{8} \|\varepsilon_h\|_\infty.$$

On a donc une borne d'erreur sur l'erreur qu'on obtient sur la solution du système par rapport à l'erreur commise sur le second membre.

## 2.3 Schéma volumes finis pour un problème elliptique en une dimension d'espace

### 2.3.1 Origine du Schéma

On va étudier la discréétisation par volumes finis du problème (2.1.1)–(2.1.2), qu'on rappelle ici :

$$\begin{cases} -u_{xx} = f, & x \in ]0, 1[, \\ u(0) = u(1) = 0. \end{cases} \quad (2.3.27)$$

**Définition 2.20 (Maillage volumes finis)** On appelle maillage volumes finis de l'intervalle  $[0, 1]$ , un ensemble de  $N$  mailles  $(K_i)_{i=1,\dots,N}$ , telles que  $K_i = ]x_{i-1/2}, x_{i+1/2}[$ , avec  $x_{1/2} = 0 < x_{3/2} < x_{5/2} < \dots < x_{N+1/2} = 1$ , et on note  $K_i = x_{i+1/2} - x_{i-1/2}$ . On se donne également  $N$  points  $(x_i)_{i=1,\dots,N}$  situés dans les mailles  $K_i$ . On a donc :

$$0 = x_{1/2} < x_1 < x_{3/2} < \dots < x_{i-1/2} < x_i < x_{i+1/2} < \dots < x_{N+1/2} = 1.$$

On notera  $h_{i+1/2} = x_{i+1} - x_i$ , et  $h = \max_{i=1,\dots,N}$ , et pour des questions de notations, on posera également  $x_0 = 0$  et  $x_{N+1} = 1$ .

On intègre (2.1.1) sur  $K_i = x_{i+1/2} - x_{i-1/2}$ , et on obtient :

$$-u_x(x_i + 1/2) + u_x(x_i - 1/2) = \int_{K_i} f(x) dx. \quad (2.3.28)$$

On pose :  $f_i = \frac{1}{h_i} \int_{K_i} f(x) dx$ , et on introduit les inconnues discrètes  $(u_i)_{i=1,\dots,N}$  (une par maille) et les équations discrètes du schéma numérique :

$$F_{i+1/2} - F_{i-1/2} = h_i f_i, \quad i = 1, \dots, N, \quad (2.3.29)$$

où  $F_{i+1/2}$  est le flux numérique en  $x_{i+1/2}$  qui devrait être une approximation raisonnable de  $-u_x(x_{i+1/2})$ . On pose alors :

$$\begin{aligned} F_{i+1/2} &= -\frac{u_{i+1} - u_i}{h_{i+1/2}}, \quad i = 1, \dots, N, \\ F_{1/2} &= -\frac{u_1}{h_{1/2}}, \quad F_{N+1/2} = \frac{u_N}{h_{N+1/2}}, \end{aligned}$$

pour tenir compte des conditions aux limites de Dirichlet homogènes  $u(0) = u(1) = 0$ . On peut aussi écrire :

$$F_{i+1/2} = -\frac{u_{i+1} - u_i}{h_{i+1/2}}, \quad i = 0, \dots, N, \quad (2.3.30)$$

$$\text{en posant } u_0 = u_{N+1} = 0. \quad (2.3.31)$$

On peut écrire le système linéaire obtenu sur  $(u_1, \dots, u_N)^t$  sous la forme

$$A_h U_h = b_h, \quad (2.3.32)$$

avec

$$(A_h)_i = \frac{1}{h_i} \left[ \frac{-1}{h_{i+1/2}} (u_{i+1} - u_i) + \frac{1}{h_{i-1/2}} (u_i - u_{i-1}) \right] \text{ et } (b_h)_i = f_i.$$

**Remarque 2.21 (Non consistance au sens des différences finies)**

L'approximation de  $-u''(x_i)$  par

$$\frac{1}{h_i} \left[ \frac{-1}{h_{i+1/2}} (u(x_{i+1}) - u(x_i)) + \frac{1}{h_{i-1/2}} (u(x_i) - u(x_{i-1})) \right]$$

n'est pas consistante dans le cas général : voir exercice 7.

On peut montrer que les deux schémas différences finies et volumes sont identiques "au bord près" dans le cas d'un maillage uniforme avec  $x_i$  : centre de la maille voir exercice 1 page 46.

### 2.3.2 Analyse mathématique du schéma.

On va démontrer ici qu'il existe une unique solution  $(u_1, \dots, u_N)^t$  au schéma (2.3.29)–(2.3.31), et que cette solution, et que cette solution converge, en un certain sens, vers la solution de problème continu (2.3.27) lorsque le pas du maillage tend vers 0.

**Proposition 2.22 (Existence de la solution du schéma volumes finis)** *Soit  $f \in C([0, 1])$  et  $u \in C^2([0, 1])$  solution de (2.3.27). Soit  $(K_i)_{i=1, \dots, N}$  le maillage par la définition 2.20 page 20. Alors il existe une unique solution  $u_h = (u_1, \dots, u_N)^t$  de (2.3.29)–(2.3.31).*

**Démonstration :** Le schéma s'écrit

$$-\frac{u_{i+1} - u_i}{h_{i+1/2}} + \frac{u_i - u_{i-1}}{h_{i-1/2}} = h_i f_i, \quad i = 1, \dots, N.$$

(où on a posé  $u_0 = 0$  et  $u_{N+1} = 0$ ) En multipliant par  $u_i$  et en sommant de  $i = 1$  à  $N$ , on obtient donc :

$$\sum_{i=1}^N -\frac{u_{i+1} - u_i}{h_{i+1/2}} u_i + \sum_{i=1}^N \frac{u_i - u_{i-1}}{h_{i-1/2}} u_i = \sum_{i=1}^N h_i f_i u_i.$$

En effectuant un changement d'indice sur la deuxième somme, on obtient :

$$\sum_{i=1}^N -\frac{u_{i+1} - u_i}{h_{i+1/2}} u_i + \sum_{i=0}^{N-1} \frac{u_{i+1} - u_i}{h_{i+1/2}} u_{i+1} = \sum_{i=1}^N h_i f_i u_i;$$

en regroupant les sommes, on a donc :

$$\sum_{i=1}^N \frac{(u_{i+1} - u_i)^2}{h_{i+1/2}} + \frac{u_1^2}{h_{1/2}} + \frac{u_N^2}{h_{N+1/2}} = \sum_{i=1}^N h_i f_i u_i.$$

Si  $f_i = 0$  pour tout  $i = 1, \dots, N$ , on a bien alors  $u_i = 0$  pour tout  $i = 1 \dots N$ . Ceci démontre l'unicité de  $(u_i)_{i=1 \dots N}$  solution de (2.3.29)–(2.3.31), et donc son existence, puisque le système (2.3.29)–(2.3.31) est un système linéaire carré d'ordre  $N$ . (On rappelle qu'une matrice carrée d'ordre  $N$  est inversible si et seulement si son noyau est réduit à  $\{0\}$ ). ■

**Lemme 2.23 (Consistance des flux)** *Soit  $u \in C^2([0, 1])$  solution de (2.3.27). On se donne une subdivision de  $[0, 1]$ . On appelle  $\bar{F}_{i+1/2} = -u_x(x_{i+1/2})$  le flux exact en  $x_{i+1/2}$ , et  $F_{i+1/2}^* = -\frac{u(x_{i+1}) - u(x_i)}{h_{i+1/2}}$  le quotient différentiel qui approche la dérivée première  $-u_x(x_{i+1/2})$ . On dit que le flux numérique  $F_{i+1/2} = -\frac{u_{i+1} - u_i}{h_{i+1/2}}$  est consistant s'il existe  $C \in \mathbb{R}_+$  ne dépendant que de  $u$  telle que l'erreur de consistance sur le flux, définie par :*

$$R_{i+1/2} = \bar{F}_{i+1/2} - F_{i+1/2}^*,$$

vérifie

$$|R_{i+1/2}| \leq Ch. \quad (2.3.33)$$

La démonstration de ce résultat s'effectue facilement à l'aide de développements de Taylor. On peut aussi montrer (voir exercice 8 page 49) que si  $x_{i+1/2}$  est au centre de l'intervalle  $[x_i x_{i+1}]$ , l'erreur de consistance sur les flux est d'ordre 2, *i.e.* il existe  $C \in \mathbb{R}_+$  ne dépendant que de  $u$  telle que  $R_{i+1/2} \leq Ch^2$ . Notez que cette propriété de consistance est vraie sur les flux, et non pas sur l'opérateur  $-u''$  (voir remarque 2.21).

**Définition 2.24 (Conservativité)** On dit que le schéma volumes finis (2.3.29)–(2.3.31) est conservatif, au sens où, lorsqu'on considère une interface  $x_{i+1/2}$  entre deux mailles  $K_i$  et  $K_{i+1}$ , le flux numérique entrant dans une maille est égal à celui sortant de l'autre.

C'est grâce à la conservativité et à la consistance des flux qu'on va montrer la convergence du schéma volumes finis.

**Théorème 2.25 (Convergence du schéma volumes finis)** On suppose que la solution  $u$  de (2.3.27) vérifie  $u \in C^2([0, 1])$ . On pose pour  $e_i = u(x_i) - u_i$  pour  $i = 1, \dots, N$ , et  $e_0 = e_{N+1} = 0$ . Il existe  $C \geq 0$  ne dépendant que de  $u$  tel que :

$$\sum_{i=0}^N \frac{(e_{i+1} - e_i)^2}{h} \leq Ch^2, \quad (2.3.34)$$

$$\sum_{i=1}^N he_i^2 \leq Ch^2 \quad (2.3.35)$$

$$\max_{i=1 \dots N} |e_i| \leq Ch. \quad (2.3.36)$$

(On rappelle que  $h = \sup_{i=1 \dots N} h_i$ .)

**Démonstration :** Ecrivons le schéma volumes finis (2.3.29) :

$$F_{i+1/2} - F_{i-1/2} = h_i f_i,$$

l'équation exacte intégrée sur la maille  $K_i$  (2.3.28) :

$$\bar{F}_{i+1/2} - \bar{F}_{i-1/2} = h_i f_i,$$

où  $\bar{F}_{i+1/2}$  est défini dans le lemme 2.23, et soustrayons :

$$\bar{F}_{i+1/2} - F_{i+1/2} - \bar{F}_{i-1/2} + F_{i-1/2} = 0.$$

En introduisant  $R_{i+1/2} = \bar{F}_{i+1/2} - F_{i+1/2}^*$ , on obtient :

$$F_{i+1/2}^* - F_{i+1/2} - F_{i-1/2}^* + F_{i-1/2} = -R_{i+1/2} + R_{i-1/2}$$

ce qui s'écrit encore, au vu de la définition de  $e_i$ ,

$$-\frac{1}{h_{i+1/2}}(e_{i+1} - e_i) + \frac{1}{h_{i-1/2}}(e_i - e_{i-1}) = -R_{i+1/2} + R_{i-1/2}.$$

On multiplie cette dernière égalité par  $e_i$  et on somme de 1 à  $N$  :

$$\sum_{i=1}^N -\frac{1}{h_{i+1/2}}(e_{i+1} - e_i)e_i + \sum_{i=1}^N \frac{1}{h_{i-1/2}}(e_i - e_{i-1})e_i \sum_{i=1}^N -R_{i+1/2}e_i + \sum_{i=1}^N R_{i-1/2}e_i,$$

ce qui s'écrit encore :

$$\sum_{i=1}^N -\frac{1}{h_{i+1/2}}(e_{i+1} - e_i)e_i + \sum_{i=0}^{N-1} \frac{1}{h_{i+1/2}}(e_{i+1} - e_i)e_{i+1} \sum_{i=1}^N -R_{i+1/2}e_i + \sum_{i=0}^{N-1} R_{i+1/2}e_{i+1}$$

En réordonnant les termes, on obtient, en remarquant que  $e_0 = 0$  et  $e_{N+1} = 0$  :

$$\sum_{i=0}^N \frac{(e_{i+1} - e_i)^2}{h_{i+1/2}} = \sum_{i=0}^N R_{i+1/2}(e_{i+1} - e_i).$$

Or,  $R_{i+1/2} \leq C h$  (par le lemme 2.23). On a donc

$$\sum_{i=0}^N \frac{(e_{i+1} - e_i)^2}{h_{i+1/2}} \leq C h \sum_{i=0}^N \frac{|e_{i+1} - e_i|}{\sqrt{h_{i+1/2}}} \sqrt{h_{i+1/2}},$$

et, par l'inégalité de Cauchy-Schwarz :

$$\sum_{i=0}^N \frac{(e_{i+1} - e_i)^2}{h_{i+1/2}} \leq C h \left( \sum_{i=0}^N \frac{|e_{i+1} - e_i|^2}{h_{i+1/2}} \right)^{1/2} \times \left( \sum_{i=0}^N h_{i+1/2} \right)^{1/2}.$$

En remarquant que  $\sum_{i=0}^N h_{i+1/2} = 1$ , on déduit que :

$$\sum_{i=0}^N \frac{(e_{i+1} - e_i)^2}{h_{i+1/2}} \leq C h \left( \sum \frac{|e_{i+1} - e_i|^2}{h_{i+1/2}} \right)^{1/2},$$

et donc

$$\left( \sum_{i=0}^N \frac{(e_{i+1} - e_i)^2}{h_{i+1}} \right)^{1/2} \leq C h.$$

On a ainsi démontré (2.3.34). Démontrons maintenant (2.3.36). Pour obtenir une majoration de  $|e_i|$  par  $C h$ , on remarque que :

$$|e_i| = \left| \sum_{j=1}^i e_j - e_{j-1} \right| \leq \sum_{j=1}^i |e_j - e_{j-1}| \leq \sum_{j=1}^N |e_j - e_{j-1}|.$$

On en déduit, par l'inégalité de Cauchy Schwarz, que :

$$|e_i| \leq \left( \sum \frac{|e_j - e_{j-1}|^2}{h_{i+1/2}} \right)^{1/2} \left( \sum h_{i+1/2} \right)^{1/2},$$

ce qui entraîne  $\max_{i=1 \dots N} |e_i| \leq C h$ . Notons que de cette estimation, on déduit immédiatement l'estimation (2.3.35). ■

**Remarque 2.26 (Espaces fonctionnels et normes discrètes)** *On rappelle qu'une fonction  $u$  de  $L^2(]0, 1[)$  admet une dérivée faible dans  $L^2(]0, 1[)$  s'il existe  $v \in L^2(]0, 1[)$  telle que*

$$\int_{]0,1[} u(x) \varphi'(x) dx = - \int_{]0,1[} v(x) \varphi(x) dx, \quad (2.3.37)$$

*pour toute fonction  $\varphi \in C_c^1(]0, 1[)$ , où  $C_c^1(]0, 1[)$  désigne l'espace des fonctions de classe  $C^1$  à support compact dans  $]0, 1[$ . On peut montrer que  $v$  est unique, voir par exemple [1]. On notera  $v = Du$ . On*

peut remarquer que si  $u \in C^1(]0, 1[)$ , alors  $Du = u'$ , dérivée classique. On note  $H^1(]0, 1[)$  l'ensemble des fonctions de  $L^2(]0, 1[)$  qui admettent une dérivée faible dans  $L^2(]0, 1[)$  :  $H^1(]0, 1[) = \{u \in L^2(]0, 1[) ; Du \in L^2(]0, 1[)\}$ . On a  $H^1(]0, 1[) \subset C(]0, 1[)$  et on définit

$$H_0^1(]0, 1[) = \{u \in H^1(]0, 1[) ; u(0) = u(1) = 0\}.$$

Pour  $u \in H^1(]0, 1[)$ , on note :

$$\|u\|_{H_0^1} = \left( \int_0^1 (Du(x))^2 dx \right)^{1/2}.$$

C'est une norme sur  $H_0^1$  qui est équivalente à la norme  $\|\cdot\|_{H^1}$  définie par  $\|u\|_{H^1} = \left( \int u^2(x) dx + \int (Du)^2(x) dx \right)^{1/2}$ , ce qui se démontre grâce à l'inégalité de Poincaré :

$$\|u\|_{L^2(]0, 1[)} \leq \|Du\|_{L^2(]0, 1[)} \text{ pour tout } u \in H_0^1(]0, 1[). \quad (2.3.38)$$

Soit maintenant  $\mathcal{T}$  un maillage volumes finis de  $[0, 1]$  (voir définition 2.20), on note  $X(\mathcal{T})$  l'ensemble des fonctions de  $[0, 1]$  dans  $\mathbb{R}$ , constantes par maille de ce maillage. Pour  $v \in X(\mathcal{T})$ , on note  $v_i$  la valeur de  $v$  sur la maille  $i$  ; on peut écrire les normes  $L^2$  et  $L^\infty$  de  $v$  :

$$\|v\|_{L^2(]0, 1[)}^2 = \sum_{i=1}^N h_i v_i^2,$$

et

$$\|v\|_{L^\infty(]0, 1[)} = \max_{i=1}^N |v_i|.$$

Par contre, la fonction  $v$  étant constante par maille, elle n'est pas dérivable au sens classique, ni même au sens faible. On peut toutefois définir une norme  $H^1$  discrète de  $v$  de la manière suivante :

$$|v|_{1, \mathcal{T}} = \left( \sum_{i=0}^N h_{i+1/2} \left( \frac{v_{i+1} - v_i}{h_{i+1/2}} \right)^2 \right)^{1/2}$$

On peut définir une sorte de "dérivée discrète" de  $v$  par les pentes

$$p_{i+1/2} = \frac{v_{i+1} - v_i}{h_{i+1/2}}.$$

On peut alors définir une  $D_{\mathcal{T}}v$ , fonction constante par intervalle et égale à  $p_{i+1/2}$  sur l'intervalle  $x_i, x_{i+1}$ . La norme  $L^2$  de  $D_{\mathcal{T}}v$  est donc définie par :

$$\|D_{\mathcal{T}}v\|_{L^2(]0, 1[)}^2 = \sum_{i=0}^N h_{i+1/2} p_{i+1/2}^2 = \sum_{i=0}^N \sum_{i=0}^N h_{i+1/2} \frac{(v_{i+1} - v_i)^2}{h_{i+1/2}}.$$

On peut montrer (Exercice 12) que si  $u_{\mathcal{T}} : ]0, 1[ \rightarrow \mathbb{R}$  est définie par  $u_{\mathcal{T}}(x) = u_i \quad \forall x \in K_i$  où  $(u_i)_{i=1, \dots, N}$  solution de (2.3.29)–(2.3.31), alors  $|u_{\mathcal{T}}|_{1, \mathcal{T}}$  converge dans  $L^2(]0, 1[)$  lorsque  $h$  tend vers 0, vers  $\|Du\|_{L^2(]0, 1[)}$ , où  $u$  est la solution de (2.3.27).

**Remarque 2.27 (Dimensions supérieures)** En une dimension d'espace, on a obtenu une estimation d'erreur en norme " $H_0^1$  discrète" et en norme  $L^\infty$ . En dimension supérieure ou égale à 2, on aura une estimation en  $h$ , en norme  $H_0^1$  discrète, en norme  $L^2$ , mais pas en norme  $L^\infty$ . Ceci tient au fait que l'injection de Sobolev  $H^1([0, 1]) \subset C([0, 1])$  n'est vraie qu'en dimension 1. La démonstration de l'estimation d'erreur en norme  $L^2$  (2.3.35) se prouve alors directement à partir de l'estimation en norme  $H_0^1$  discrète, grâce à une "inégalité de Poincaré discrète", équivalent discret de la célèbre inégalité de Poincaré continue<sup>1</sup> (voir (2.3.38) pour la dimension 1).

### Prise en compte de discontinuités

On considère ici un barreau conducteur constitué de deux matériaux de conductivités  $\lambda_1$  et  $\lambda_2$  différentes, et dont les extrémités sont plongées dans de la glace. On suppose que le barreau est de longueur 1, que le matériau de conductivité  $\lambda_1$  (resp.  $\lambda_2$ ) occupe le domaine  $\Omega_1 = ]0, 1/2[$  (resp.  $\Omega_2 = ]1/2, 1[$ ). Le problème de conduction de la chaleur s'écrit alors :

$$\begin{cases} (-\lambda_1(x)u_x)_x = f(x) & x \in ]0, 1/2[ \\ (-\lambda_2(x)u_x)_x = f(x) & x \in ]1/2, 1[ \\ u(0) = u(1) = 0, \\ -(\lambda_1 u_x)(1/2) = -(\lambda_2 u_x)(1/2) \end{cases} \quad (2.3.39)$$

**Remarque 2.28** La dernière égalité traduit la conservation du flux de chaleur à l'interface  $x = .5$ . On peut noter que comme  $\lambda$  est discontinu en ce point, la dérivée  $u_x$  le sera forcément elle aussi.

On choisit de discréteriser le problème par volumes finis. On se donne un maillage volumes finis comme défini par la définition 2.20 page 20, en choisissant les mailles telles que la discontinuité de  $\lambda$  soit située sur un interface de deux mailles qu'on note  $K_k$  et  $K_{k+1}$ . On a donc, avec les notations du paragraphe (2.1.1)  $x_{k+1/2} = 0.5$ . La discréétisation par volumes finis s'écrit alors

$$F_{i+1/2} - F_{i-1/2} = h_i f_i, \quad i = 1, \dots, N,$$

où les flux numériques  $F_{i+1/2}$  sont donnés par

$$F_{i+1/2} = \lambda_* \frac{u_{i+1} - u_i}{h_{i+1/2}}, \quad \text{avec } \lambda_* = \begin{cases} \lambda_1 & \text{si } x_{i+1/2} > 0.5, \\ \lambda_2 & \text{si } x_{i+1/2} < 0.5. \end{cases}$$

Il ne reste donc plus qu'à calculer le flux  $F_{k+1/2}$ , approximation de  $(\lambda u_x)(x_{k+1/2})$  (avec  $x_{k+1/2} = 0.5$ ). On introduit pour cela une inconnue auxiliaire  $u_{k+1/2}$  que l'on pourra éliminer plus tard, et on écrit une discréétisation du flux de part et d'autre de l'interface.

$$\begin{aligned} F_{k+1/2} &= -\lambda_1 \frac{u_{k+1/2} - u_k}{h_k^+}, \quad \text{avec } h_k^+ = x_{k+1/2} - x_k, \\ F_{k+1/2} &= -\lambda_2 \frac{u_{k+1} - u_{k+1/2}}{h_{k+1}^-} \quad \text{avec } h_{k+1}^- = x_{k+1} - x_{k+1/2}. \end{aligned}$$

L'élimination (et le calcul) de l'inconnue se fait en écrivant la conservation du flux numérique :

$$-\lambda_1 \frac{u_{k+1/2} - u_k}{h_k^+} = -\lambda_2 \frac{u_{k+1} - u_{k+1/2}}{h_{k+1}^-}$$

<sup>1</sup>Soit  $\Omega$  un ouvert borné de  $\mathbb{R}^N$ , et  $u \in H_0^1(\Omega)$ , alors  $\|u\|_{L^2(\Omega)} \leq \text{diam}(\Omega) \|Du\|_{L^2(\Omega)}$ .

On en déduit la valeur de  $u_{k+1/2}$

$$u_{k+1/2} = \frac{\frac{\lambda_1}{h_k^+} u_k + \frac{\lambda_2}{h_{k+1}^-} u_{k+1}}{\frac{\lambda_1}{h_k^+} + \frac{\lambda_2}{h_{k+1}^-}}$$

On remplace  $u_{k+1/2}$  par cette valeur dans l'expression du flux  $F_{k+1/2}$ , et on obtient :

$$F_{k+1/2} = \frac{\lambda_1 \lambda_2}{h_k^+ \lambda_2 + h_{k+1}^- \lambda_1} (u_{k+1} - u_k).$$

Si le maillage est uniforme, on obtient

$$F_{k+1/2} = \frac{2\lambda_1 \lambda_2}{\lambda_1 + \lambda_2} \left( \frac{u_{i+1} - u_i}{h} \right).$$

Le flux est donc calculé en faisant intervenir la moyenne harmonique des conductivités  $\lambda_1$  et  $\lambda_2$ . Notons que lorsque  $\lambda_1 = \lambda_2$ , on retrouve la formule habituelle du flux.

### 2.3.3 Exemples de discréétisation par différences finies ou volumes finis des problèmes elliptiques en dimension 2.

#### Différences finies

On considère maintenant le problème de diffusion dans un ouvert  $\Omega$  de  $\mathbb{R}^2$  :

$$\begin{cases} -\Delta u = f \text{ dans } \Omega, \\ u = 0 \quad \text{sur } \partial\Omega. \end{cases} \quad (2.3.40)$$

Le problème est bien posé au sens où : Si  $f \in C^1(\Omega)$ , alors il existe une unique solution  $u \in C(\bar{\Omega}) \cap C^2(\Omega)$ , solution de (2.3.40). Si  $f \in L^2(\Omega)$  alors il existe une unique fonction  $u \in H^2(\Omega)$  au sens faible<sup>2</sup> de (2.3.40), c.à.d. qui vérifie :

$$\begin{cases} u \in H_0^1(\Omega), \\ \int_{\Omega} \nabla u(x) \nabla v(x) dx = \int_{\Omega} f(x) v(x) dx, \quad \forall v \in H_0^1(\Omega). \end{cases} \quad (2.3.41)$$

On peut montrer (voir cours Equations aux dérivées partielles) que si  $u \in C^2(\Omega)$ , alors  $u$  est solution de (2.3.40) si et seulement si  $u$  est solution faible de (2.3.40). Pour discréétiser le problème, on se donne un certain nombre de points, alignés dans les directions  $x$  et  $y$ , comme représentés sur la figure 2.2 (on prend un pas de maillage uniforme et égal à  $h$ ). Certains de ces points sont à l'intérieur du domaine  $\Omega$ , d'autres sont situés sur la frontière  $\partial\Omega$ .

Comme en une dimension d'espace, les inconnues discrètes sont associées aux noeuds du maillage. On note  $\{P_i, i \in I\}$  les points de discréétisation, et on écrit l'équation aux dérivées partielles en ces points :

$$-\Delta u(P_i) - \frac{\partial^2 u}{\partial x^2}(P_i) - \frac{\partial^2 u}{\partial y^2}(P_i) = f(P_i).$$

1er cas :

---

<sup>2</sup>Par définition,  $H^2(\Omega)$  est l'ensemble des fonctions de  $L^2(\Omega)$  qui admet des dérivées faibles jusqu'à l'ordre 2 dans  $L^2(\Omega)$ .

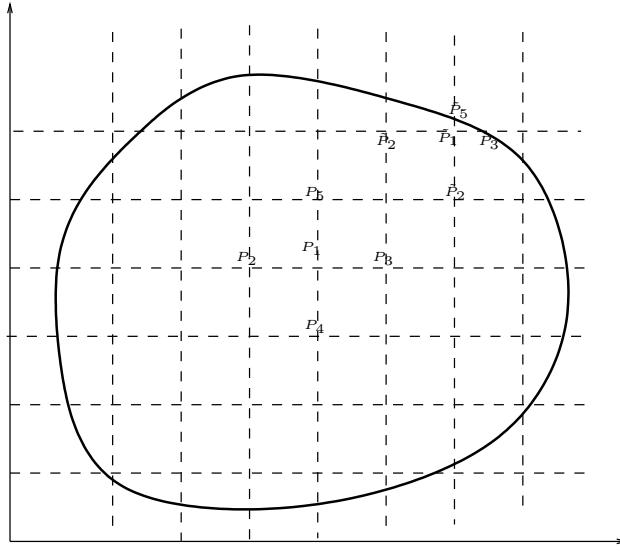


FIG. 2.2 – Discrétisation différences finies bi-dimensionnelle

Dans le cas de points points "vraiment intérieurs", tel que le point  $P_1$  sur la figure 2.2, *i.e.* dont tous les points voisins sont situés à l'intérieur de  $\Omega$ , les quotients différentiels

$$\frac{2u(P_1) - u(P_2) - u(P_3)}{h^2} \text{ et } \frac{2u(P_1) - u(P_5) - u(P_4)}{h^2}$$

sont des approximations consistantes à l'ordre 2 de  $-\partial_1^2 u(P_1)$  et  $-\partial_2^2 u(P_1)$ .

Par contre, pour un point "proche" du bord tel que le point  $\tilde{P}_1$ , les mêmes approximations (avec les points  $\tilde{P}_2$ ,  $\tilde{P}_3$ ,  $\tilde{P}_4$  et  $\tilde{P}_5$ ) ne seront que d'ordre 1 en raison des différences de distance entre les points (faire les développements de Taylor pour s'en convaincre).

Une telle discrétisation amène à un système linéaire  $A_h U_h = b_h$ , où la structure de  $A_h$  (en particulier sa "largeur de bande", c.à.d. le nombre de diagonales non nulles) dépend de la numérotation des noeuds. On peut montrer que la matrice  $A_h$  est monotone et le schéma est stable. De la consistance et la stabilité, on déduit, comme en une dimension d'espace, la convergence du schéma.

### 2.3.4 Exemple d'implémentation en dimension 2

#### Le problème modèle

On considère le problème modèle suivant (par exemple de conduction de la chaleur) :

$$-\operatorname{div}(\lambda_i \nabla u(x)) = f(x) \quad x \in \Omega_i, \quad i = 1, 2 \quad (2.3.42)$$

où  $\lambda_1 > 0$ ,  $\lambda_2 > 0$  sont les conductivités thermiques dans les domaines  $\Omega_1$  et avec  $\Omega_2$ , avec  $\Omega_1 = ]0, 1[ \times ]0, 1[$  et  $\Omega_2 = ]0, 1[ \times ]1, 2[$ . On appelle  $\Gamma_1 = ]0, 1[ \times \{0\}$ ,  $\Gamma_2 = \{1\} \times ]0, 2[$ ,  $\Gamma_3 = ]0, 1[ \times \{2\}$ , et  $\Gamma_4 = \{0\} \times ]0, 2[$  les frontières extérieures de  $\Omega$ , et on note  $I = ]0, 1[ \times \{1\}$  l'interface entre  $\Omega_1$  et  $\Omega_2$  (voir Figure 2.3). Dans la suite, on notera  $\lambda$  la conductivité thermique sur  $\Omega$ , avec  $\lambda|_{\Omega_i} = \lambda_i$ ,  $i = 1, 2$ .

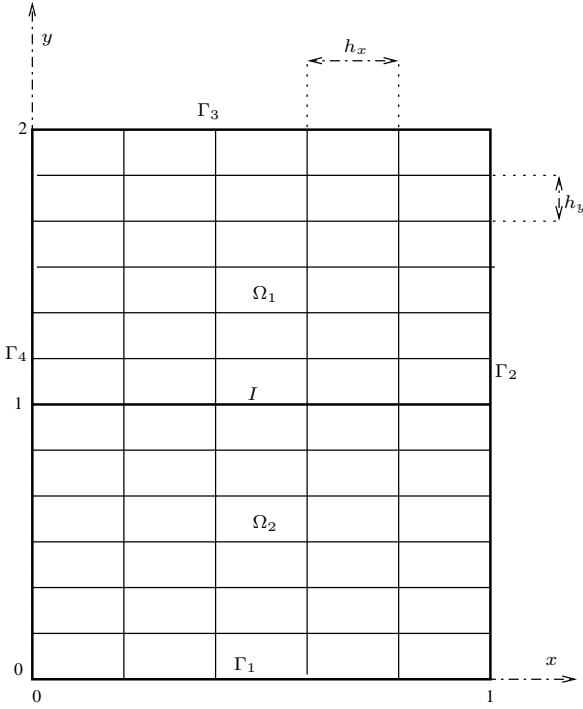


FIG. 2.3 – Domaine d'étude

On va considérer plusieurs types de conditions aux limites, en essayant d'expliquer leur sens physique. On rappelle que le flux de chaleur par diffusion est égal  $\mathbf{q}$  est donné par la loi de Fourier :  $\mathbf{q} = -\lambda \nabla u \cdot \mathbf{n}$ , où  $\mathbf{n}$  est le vecteur normal unitaire à la surface à travers laquelle on calcule le flux.

1. Conditions aux limites de type Fourier (Robin) sur  $\Gamma_1 \cup \Gamma_3$  : On suppose qu'il existe un transfert thermique entre les parois  $\Gamma_1$  et  $\Gamma_3$  et l'extérieur. Ce transfert est décrit par la condition de Fourier (Robin dans la littérature anglo-saxonne), qui exprime que le flux transféré est proportionnel à la différence de température entre l'extérieur et l'intérieur :

$$-\lambda \nabla u \cdot \mathbf{n}(x) = \alpha(u(x) - u_{ext}), \quad \forall x \in \Gamma_1 \cup \Gamma_3. \quad (2.3.43)$$

où  $\alpha > 0$  est le coefficient de transfert thermique,  $\mathbf{n}$  le vecteur unitaire normal à  $\partial\Omega$  extérieur à  $\Omega$ , et  $u_{ext}$  est la température extérieure (donnée).

2. Conditions aux limites de type Neumann sur  $\Gamma_2$  : On suppose que la paroi  $\Gamma_2$  est parfaitement isolée, et que le flux de chaleur à travers cette paroi est donc nul. Ceci se traduit par une condition dite “de Neumann homogène” :

$$-\lambda \nabla u \cdot \mathbf{n} = 0 \quad \forall x \in \Gamma_2. \quad (2.3.44)$$

3. Conditions aux limites de type Dirichlet sur  $\Gamma_4$  : Sur la paroi  $\Gamma_4$ , on suppose que la température est fixée. Ceci est une condition assez difficile à obtenir expérimentalement pour un problème de type chaleur, mais qu'on peut rencontrer dans d'autres problèmes pratiques.

$$u(x) = g(x), \quad \forall x \in \Gamma_4. \quad (2.3.45)$$

4. Conditions sur l'interface  $I$  : On suppose que l'interface  $I$  est par exemple le siège d'une réaction chimique surfacique  $\theta$  qui provoque un dégagement de chaleur surfacique. On a donc un saut du flux de chaleur au travers de l'interface  $I$ . Ceci se traduit par la condition de saut suivante :

$$-\lambda_1 \nabla u_1(x) \cdot \mathbf{n}_1 - \lambda_2 \nabla u_2(x) \cdot \mathbf{n}_2 = \theta(x), \quad x \in I. \quad (2.3.46)$$

où  $\mathbf{n}_i$  désigne le vecteur unitaire normal à  $I$  et extérieur à  $\Omega_i$ , et  $\theta$  est une fonction donnée.

### Discrétisation par volumes finis

On se donne un maillage "admissible"  $\mathcal{T}$  de  $\Omega$

$$\bar{\Omega} = \bigcup_{K \in \mathcal{T}} \bar{K}.$$

Par "admissible", on entend un maillage tel qu'il existe des points  $(x_K)_{K \in \mathcal{T}}$  situés dans les mailles, tels que chaque segment  $x_K x_L$  soit orthogonal à l'arête  $K|L$  séparant la maille  $K$  de la maille  $L$ , comme visible sur la figure 2.4.

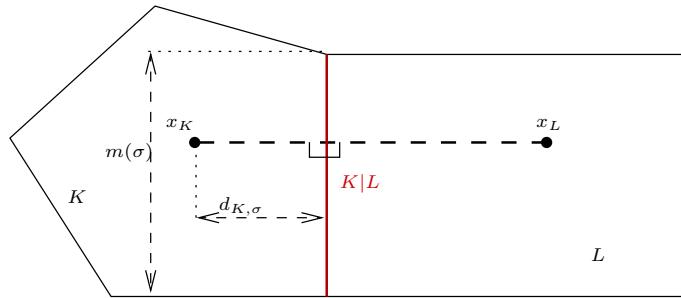


FIG. 2.4 – Condition d'orthogonalité pour un maillage volumes finis

Cette condition est nécessaire pour obtenir une approximation consistante du flux de diffusion (c'est-à-dire de la dérivée normale sur l'arête  $K|L$ ), voir remarque 2.29. Dans le cas présent, le domaine représenté sur la figure 2.3 étant rectangulaire, cette condition est particulièrement facile à vérifier en prenant un maillage rectangulaire. Par souci de simplicité, on prendra ce maillage uniforme, et on notera  $h_x = 1/n$  le pas de discrétisation dans la direction  $x$  et  $h_y = 1/p$  le pas de discrétisation dans la direction  $y$ . Le maillage est donc choisi de telle sorte que l'interface  $I$  coïncide avec un ensemble d'arêtes du maillage qu'on notera  $\mathcal{E}_I$ . On a donc

$$\bar{I} = \bigcup_{\sigma \in \mathcal{E}_I} \bar{\sigma},$$

où le signe  $\bar{\phantom{a}}$  désigne l'adhérence de l'ensemble. On se donne ensuite des inconnues discrètes  $(u_K)_{K \in \mathcal{T}}$  associées aux mailles et  $(u_\sigma)_{\sigma \in \mathcal{E}}$  associées aux arêtes.

Pour obtenir le schéma volumes finis, on commence par établir les bilans par maille en intégrant l'équation sur chaque maille  $K$  (notons que ceci est faisable en raison du fait que l'équation est sous forme conservative, c'est-à-dire sous la forme :  $-\operatorname{div}(\operatorname{flux}) = f$ ). On obtient donc :

$$\int_K -\operatorname{div}(\lambda_i \nabla u(x)) dx = \int_K f(x) dx,$$

soit encore, par la formule de Stokes,

$$\int_{\partial K} -\lambda_i \nabla u(x) \cdot \mathbf{n}(x) d\gamma(x) = m(K) f_K,$$

où  $\mathbf{n}$  est le vecteur unitaire normal à  $\partial\Omega$ , extérieur à  $\Omega$ , et  $\gamma$  désigne le symbole d'intégration sur la frontière. On décompose ensuite le bord de chaque maille  $K$  en arêtes du maillage :  $\partial K = \bigcup_{\sigma \in \mathcal{E}_K} \bar{\sigma}$  où  $\mathcal{E}_K$  représente l'ensemble des arêtes de  $K$ . On obtient alors :

$$\sum_{\sigma \in \mathcal{E}_K} \int_{\sigma} -\lambda_i \nabla u \cdot \mathbf{n}_{K,\sigma} d\gamma(x) = m(K) f_K$$

où  $\mathbf{n}_{K,\sigma}$  est le vecteur unitaire normal à  $\sigma$  extérieur à  $K$ . On écrit alors une “équation approchée” :

$$\sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma} = m(K) f_K,$$

où  $F_{K,\sigma}$  est le flux numérique à travers  $\sigma$ , qui approche le flux exact  $F_{K,\sigma}^* = \int_{\sigma} -\lambda_i \nabla u \cdot \mathbf{n}_{K,\sigma} d\gamma(x)$ . Pour obtenir le schéma numérique, il nous reste à exprimer le flux numérique  $F_{K,\sigma}$  en fonction des inconnues discrètes  $(u_K)_{K \in \mathcal{T}}$  associées aux mailles et  $(u_{\sigma})_{\sigma \in \mathcal{E}}$  associées aux arêtes (ces dernières seront ensuite éliminées) :

$$F_{K,\sigma} = -\lambda_i \frac{u_{\sigma} - u_K}{d_{K,\sigma}} m(\sigma), \quad (2.3.47)$$

où  $d_{K,\sigma}$  est la distance du point  $x_K$  à l'arête  $\sigma$  et  $m(\sigma)$  est la longueur de l'arête  $\sigma$  (voir Figure 2.4). L'équation associée à l'inconnue  $u_K$  est donc :

$$\sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma} = m(K) f_K.$$

On a ainsi obtenu autant d'équations que de mailles. Il nous reste maintenant à écrire une équation pour chaque arête, afin d'obtenir autant d'équations que d'inconnues.

En ce qui concerne les arêtes intérieures, on écrit la conservativité du flux, ce qui nous permettra d'éliminer les inconnues associées aux arêtes internes. Soit  $\sigma = K|L \subset \Omega_i$ , On a alors :

$$F_{K,\sigma} = -F_{L,\sigma}. \quad (2.3.48)$$

On vérifiera par le calcul (cf. exercice 15 page 54) que, après élimination de  $u_{\sigma}$ , ceci donne

$$F_{K,\sigma} = -F_{L,\sigma} = \lambda_i \frac{m(\sigma)}{d_{\sigma}} (u_K - u_L), \quad (2.3.49)$$

où  $d_{\sigma} = d(x_K, x_L)$ .

**Remarque 2.29 (Consistance du flux)** *On appelle erreur de consistance associée au flux (2.3.47) l'expression :*

$$R_{K,\sigma} = -\frac{1}{m(\sigma)} \int_{\sigma} \nabla u(x) \cdot \mathbf{n}_{K,\sigma} d\gamma(x) - F_{K,\sigma}^*, \text{ où } F_{K,\sigma}^* = -\lambda_i \frac{u(x_{\sigma}) - u(x_K)}{d_{K,\sigma}} m(\sigma),$$

où  $x_{\sigma}$  est l'intersection de  $\sigma$  avec l'arête  $K|L$ ,  $u$  la solution exacte.

On dit que le flux numérique donné par l'expression (2.3.47) est consistant si

$$\lim_{h(\mathcal{T}) \rightarrow 0} \max_{K \in \mathcal{T}, \sigma \in K} |R_{K,\sigma}| = 0,$$

où  $h(\mathcal{T})$  est le pas du maillage, i.e.  $h(\mathcal{T}) = \max_{K \in \mathcal{T}} \text{diam}(K)$ , avec  $\text{diam}(K) = \sup_{(x,y) \in K^2} d(x,y)$ . On vérifie facilement que si  $u$  est suffisamment régulière et si le segment  $x_K x_L$  est colinéaire au vecteur normal  $\mathbf{n}$ , alors le flux numérique est consistant. Cette propriété, alliée à la propriété de conservativité des flux, permet de démontrer la convergence du schéma, comme on l'a fait dans le cas unidimensionnel.

**Remarque 2.30 (Cas du maillage cartésien de la figure 2.3)** Dans le cas du maillage cartésien considéré pour notre problème, il est naturel de choisir les points  $x_K$  comme les centres de gravité des mailles. Comme le maillage est uniforme, on a donc  $d_{K,\sigma} = \frac{h_x}{2}$  (resp.  $\frac{h_y}{2}$ ) et  $|\sigma| = h_y$  (resp.  $|\sigma| = h_x$ ) pour une arête  $\sigma$  verticale (resp. horizontale).

Ecrivons maintenant la discrétisation des conditions aux limites et interface :

1. Condition de Neumann sur  $\Gamma_2$  Sur  $\Gamma_2$ , on a la condition de Neumann (2.3.44) :  $\lambda_i \nabla u \cdot \mathbf{n} = 0$ , qu'on discrétise par :  $\sigma \in \mathcal{E}_K$  et  $\sigma \subset \Gamma_2$ ,  $F_{K,\sigma} = 0$ .
2. Condition de Dirichlet sur  $\Gamma_4$  La discrétisation de la condition de Dirichlet (2.3.45) peut s'effectuer de la manière suivante :

$$u_\sigma = \frac{1}{m(\sigma)} \int_\sigma g(y) d\gamma(y).$$

L'expression du flux numérique est alors :

$$F_{K,\sigma} = -\lambda_i \frac{u_\sigma - u_K}{d_{K,\sigma}} m(\sigma).$$

3. Condition de Fourier sur  $\Gamma_1 \cup \Gamma_3$  Sur  $\Gamma_1 \cup \Gamma_3$  on a la condition de Fourier (2.3.43) :

$$-\lambda_i \nabla u \cdot \mathbf{n} = \alpha(u(x) - u_{ext}) \quad \forall x \in \Gamma_1 \cup \Gamma_3$$

qu'on discrétise par

$$F_{K,\sigma} = -m(\sigma) \lambda_i \frac{u_\sigma - u_K}{d_{K,\sigma}} = m(\sigma) \alpha(u_\sigma - u_{ext}) \text{ pour } \sigma \subset \Gamma_1 \cup \Gamma_3.$$

Après élimination de  $u_\sigma$  (cf. exercice 15 page 54), on obtient :

$$F_{K,\sigma} = \frac{\alpha \lambda_i m(\sigma)}{\lambda_i + \alpha d_{K,\sigma}} (u_K - u_{ext}). \quad (2.3.50)$$

4. Condition de saut pour le flux sur  $I$  Si  $\sigma = K|L \in \mathcal{E}_I$ , la discrétisation de la condition de saut (2.3.46) se discrétise facilement en écrivant :

$$F_{K,\sigma} + F_{L,\sigma} = \theta_\sigma, \text{ avec } \theta_\sigma = \frac{1}{|\sigma|} \int_\sigma \theta(x) d\gamma(x). \quad (2.3.51)$$

3.

Après élimination de l'inconnue  $u_\sigma$  (voir exercice 15 page 54), on obtient

$$F_{K,\sigma} = \frac{\lambda_1 m(\sigma)}{\lambda_1 d_{L,\sigma} + \lambda_2 d_{K,\sigma}} [\lambda_2 (u_K - u_L) + d_{L,\sigma} \theta_\sigma]. \quad (2.3.52)$$

On a ainsi éliminé toutes les inconnues  $u_\sigma$ , ce qui permet d'obtenir un système linéaire dont les inconnues sont les valeurs  $(u_K)_{K \in \mathcal{T}}$ .

**Remarque 2.31 (Implantation informatique de la méthode)** *Lors de l'implantation informatique, la matrice du système linéaire est construite “par arête” (contrairement à une matrice éléments finis, dont nous verrons plus tard la construction “par élément”), c.à.d. que pour chaque arête, on additionne la contribution du flux au coefficient de la matrice correspondant à l'équation et à l'inconnue concernées.*

## 2.4 Problèmes paraboliques.

On a vu au paragraphe 1.3.2 comme exemple type de problème parabolique l'équation de la chaleur instationnaire :

$$u_t - \Delta u = f$$

qui fait intervenir la dérivée en temps d'ordre 1,  $u_t$ , ainsi qu'un opérateur différentiel d'ordre 2 en espace. Pour que ce problème soit bien posé, il faut spécifier des conditions aux limites sur la frontière de  $\Omega$ , et une condition initiale en  $t = 0$ .

### 2.4.1 Le problème continu, et la discréétisation espace-temps

On considère maintenant le même problème en une dimension d'espace. Au temps  $t = 0$ , on se donne une condition initiale  $u_0$ , et on considère des conditions aux limites de type Dirichlet homogène. Le problème unidimensionnel s'écrit :

$$\begin{cases} u_t - u_{xx} = 0, \forall x \in ]0, 1[, \forall t \in ]0, T[ \\ u(x, 0) = u_0(x), \forall x \in ]0, 1[, \\ u(0, t) = u(1, t) = 0, \forall t \in ]0, T[, \end{cases} \quad (2.4.53)$$

où  $u(x, t)$  représente la température au point  $x$  et au temps  $t$ . On admettra le théorème d'existence et unicité suivant :

**Théorème 2.32 (Résultat d'existence et unicité)** *Si  $u_0 \in C(]0, 1[, \mathbb{R})$  alors il existe une unique fonction  $u \in C^2(]0, 1[ \times ]0, T[, \mathbb{R}) \cap C([0, 1] \times [0, T], \mathbb{R})$  qui vérifie (2.4.53).*

On a même  $u \in C^\infty(]0, 1[ \times ]0, T[, \mathbb{R})$ . Ceci est appelé, effet “régularisant” de l'équation de la chaleur.

**Proposition 2.33 (Principe du maximum)** *Sous les hypothèses du théorème 2.32, soit  $u$  la solution du problème (2.4.53) ;*

1. si  $u^0(x) \geq 0$  pour tout  $x \in [0, 1]$ , alors  $u(x, t) \geq 0$ , pour tout  $t \geq 0$  pour tout  $x \in ]0, 1[$ .
2.  $\|u\|_{L^\infty([0, 1] \times [0, T])} \leq \|u\|_{L^\infty([0, 1])}$ .

Ces dernières propriétés peuvent être importantes dans le modèle physique, et il est donc souvent souhaitable que les solutions approchées les vérifient également. Pour calculer une solution approchée, on se donne une discréétisation en temps et en espace, qu'on notera  $\mathcal{D}$ . On choisit pour l'instant de discréétiser par différences finies en temps et en espace. La discréétisation consiste donc à se donner un ensemble de points  $t_n$ ,  $n = 1, \dots, M$  de l'intervalle  $]0, T[$ , et un ensemble de points  $x_i$ ,  $i = 1, \dots, N$ . Pour simplifier, on considère un pas constant en temps et en espace. Soit :  $h = \frac{1}{N+1} = \Delta x$  le pas de discréétisation en espace, et  $k = \Delta t = \frac{T}{M}$ , le pas de discréétisation en temps. On pose alors  $t_n = nk$  pour  $n = 0, \dots, M$  et  $x_i = ih$  pour  $i = 0, \dots, N+1$ . On cherche à calculer une solution approchée  $u_{\mathcal{D}}$  du problème (2.4.53) ; plus précisément, on cherche à déterminer  $u_{\mathcal{D}}(x_i, t_n)$  pour  $i = 1, \dots, N$ , et  $n = 1, \dots, M$ . Les inconnues discrètes sont notées  $u_i^{(n)}$ ,  $i = 1, \dots, N$  et  $n = 1, \dots, M$ .

## 2.4.2 Discrétisation par Euler explicite en temps.

L'approximation en temps par la méthode d'Euler explicite consiste à écrire la première équation de (2.4.53) en chaque point  $x_i$  et temps  $t_n$ , à approcher  $u_t(x_i, t_n)$  par le quotient différentiel :

$$\frac{u(x_i, t_{n+1}) - u(x_i, t_n)}{k},$$

et  $-u_{xx}(x_i, t_n)$  par

$$\frac{1}{h^2}(2u(x_i, t_n) - u(x_{i-1}, t_n) - u(x_{i+1}, t_n)).$$

**Remarque 2.34** *On a choisi une discrétisation en espace de type différences finies, mais on aurait aussi bien pu prendre un schéma de volumes finis ou d'éléments finis.*

On approche donc :

$$-u_{xx}(x_i, t_n) \text{ par } \frac{1}{h^2}(2u(x_i, t_n) - u(x_{i-1}, t_n) - u(x_{i+1}, t_n)).$$

On obtient le schéma suivant :

$$\begin{cases} \frac{u_i^{(n+1)} - u_i^{(n)}}{k} + \frac{1}{h^2}(2u_i^{(n)} - u_{i-1}^{(n)} - u_{i+1}^{(n)}) = 0, & i = 1, \dots, N, \quad n = 1, \dots, M, \\ u_i^0 = u_0(x_i), & i = 1, \dots, N, \\ u_0^{(n)} = u_{N+1}^{(n)} = 0, & \forall n = 1, \dots, M. \end{cases} \quad (2.4.54)$$

le schéma est dit explicite, car la formule ci-dessus donne  $u_i^{(n+1)}$  de manière explicite en fonction des  $(u_i^{(n)})_{i=1, \dots, N}$ . En effet on a :

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

avec  $\lambda = \frac{k}{h^2}$ .

### Consistance du schéma

Soit  $\bar{u}_i^{(n)} = u(x_i, t_n)$  la valeur exacte de la solution en  $x_i$  et  $t_n$  : L'erreur de consistance  $R_i$  en  $(x_i, t_n)$  peut s'écrire comme la somme des erreurs de consistance en temps et en espace :  $R_i^{(n)} = \tilde{R}_i^{(n)} + \hat{R}_i^{(n)}$  avec :

$$\tilde{R}_i^{(n)} = \frac{\bar{u}_i^{(n+1)} - \bar{u}_i^{(n)}}{k} - u_t(x_i, t_n) \text{ et } \hat{R}_i^{(n)} = \frac{1}{h^2} \left( 2\bar{u}_i^{(n)} - \bar{u}_{i-1}^{(n)} - \bar{u}_{i+1}^{(n)} \right) - u_{xx}(x_i, t_n).$$

**Proposition 2.35** *Le schéma (2.4.54) est consistant d'ordre 1 en temps et d'ordre 2 en espace, c'est à dire qu'il existe  $C \in \mathbb{R}_+$  ne dépendant que de  $u$  tel que :*

$$|R_i^{(n)}| \leq C(k + h^2). \quad (2.4.55)$$

**Démonstration :** On a vu lors de l'étude des problèmes elliptiques que l'erreur de consistance en espace  $\tilde{R}_i^{(n)}$  est d'ordre 2 (voir formule (2.2.19) page 16). Un développement de Taylor en temps donne facilement que  $\tilde{R}_i^{(n)}$  est d'ordre 1 en temps. ■

## Stabilité

On a vu à la proposition 2.33 page 32 que la solution exacte vérifie :

$$\|u\|_{L^\infty([0,1] \times [0, T])} \leq \|u_0\|_{L^\infty([0,1])}$$

Si on choisit correctement les pas de temps et d'espace, nous allons voir qu'on peut avoir l'équivalent discret sur la solution approchée.

**Définition 2.36** *On dit qu'un schéma est  $L^\infty$ -stable si la solution approchée est bornée dans  $L^\infty$  indépendamment du pas du maillage.*

**Proposition 2.37** *Si la condition de stabilité*

$$\lambda = \frac{k}{h^2} \leq \frac{1}{2} \quad (2.4.56)$$

*est vérifiée, alors le schéma (2.4.54) est  $L^\infty$ -stable au sens où :*

$$\sup_{\substack{i=1, \dots, N \\ n=1, \dots, M}} |u_i^{(n)}| \leq \|u_0\|_\infty$$

**Démonstration :** On peut écrire le schéma sous la forme

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

soit encore :

$$u_i^{(n+1)} = (1 - 2\lambda)u_i^{(n)} + \lambda u_{i-1}^{(n)} + \lambda u_{i+1}^{(n)}.$$

Si  $0 \leq \lambda \leq \frac{1}{2}$ , on a  $\lambda \geq 0$  et  $1 - 2\lambda \geq 0$ , et la quantité  $u_i^{(n+1)}$  est donc combinaison convexe de  $u_i^{(n)}$ ,  $u_{i-1}^{(n)}$  et  $u_{i+1}^{(n)}$ . Soit  $M^{(n)} = \max_{i=1, \dots, N} u_i^{(n)}$ , on a alors :

$$u_i^{(n+1)} \leq (1 - 2\lambda)M^{(n)} + \lambda M^{(n)} + \lambda M^{(n)}, \quad \forall i = 1, \dots, N,$$

et donc  $u_i^{(n+1)} \leq M^{(n)}$ . On en déduit en passant au maximum que :

$$M^{(n+1)} \leq M^{(n)}.$$

On montre de la même manière que

$$\min_{i=1, \dots, N} u_i^{(n+1)} \geq \min_{i=1, \dots, N} u_i^{(n)}.$$

On en déduit  $\max_{i=1, \dots, N} (u_i^{(n+1)}) \leq \max u_i^0$  et  $\min_{i=1, \dots, N} (u_i^{(n+1)}) \geq \min u_i^0$  d'où le résultat. ■

## Convergence

**Définition 2.38** Soit  $u$  la solution du problème (2.4.53) et  $(u_i^{(n)})_{\substack{i=1, \dots, N \\ n=1, \dots, M}}$  la solution de (2.4.54). On appelle erreur de discréétisation au point  $(x_i, t_n)$  la quantité  $e_i^n = u(x_i, t_n) - u_i^n$ .

**Théorème 2.39** Sous les hypothèses du théorème 2.32, et sous la condition de stabilité (2.4.56), il existe  $C \in \mathbb{R}_+$  ne dépendant que de  $u$  tel que

$$\|e_i^{(n+1)}\|_\infty \leq \|e_i^{(0)}\|_\infty + TC(k + h^2), \text{ pour tout } i = 1, \dots, N \text{ et } n = 0, \dots, M - 1.$$

Ainsi, si  $\|e_i^{(0)}\|_\infty = 0$ , alors  $\max_{i=1, \dots, N} \|e_i^{(n)}\|$  tend vers 0 lorsque  $k$  et  $h$  tendent vers 0, pour tout  $n = 1, \dots, M$ . Le schéma (2.4.54) est donc convergent.

**Démonstration :** On note  $\bar{u}_i^{(n)} = u(x_i, t_n)$ . On a donc, par définition de l'erreur de consistance,

$$\frac{\bar{u}_i^{(n+1)} - \bar{u}_i^{(n)}}{k} - \frac{1}{h^2}(2\bar{u}_i^{(n)} - \bar{u}_{i-1}^{(n)} - \bar{u}_{i+1}^{(n)}) = R_i^{(n)}. \quad (2.4.57)$$

D'autre part, le schéma numérique s'écrit :

$$\frac{u_i^{(n+1)} - u_i^{(n)}}{k} - \frac{1}{h^2}(2u_i^{(n)} - u_{i-1}^{(n)} - u_{i+1}^{(n)}) = 0. \quad (2.4.58)$$

Retranchons (2.4.58) à (2.4.57), on obtient :

$$\frac{e_i^{(n+1)} - e_i^{(n)}}{k} - \frac{1}{h^2}(2e_i^{(n)} - e_{i+1}^{(n)} - e_{i-1}^{(n)}) = R_i^{(n)},$$

soit encore :

$$e_i^{(n+1)} = (1 - 2\lambda)e_i^{(n)} + \lambda e_{i-1}^{(n)} + \lambda e_{i+1}^{(n)} + kR_i^{(n)}$$

Or  $(1 - 2\lambda)e_i^{(n)} + \lambda e_{i-1}^{(n)} + \lambda e_{i+1}^{(n)} \leq \|e^{(n)}\|_\infty$ , car  $\lambda \leq \frac{1}{2}$ , et donc comme le schéma est consistant, l'inégalité (2.4.55) entraîne que :

$$|e_i^{(n+1)}| \leq \|e^{(n)}\|_\infty + kC(k + h^2).$$

On a donc par récurrence :

$$\|e_i^{(n+1)}\|_\infty \leq \|e^{(0)}\|_\infty + MkC(k + h^2)$$

ce qui démontre le théorème. ■

Donnons maintenant un exemple où lorsque la condition (2.4.55) n'est pas vérifiée, le schéma est instable.

**Exemple de non convergence si  $\lambda > \frac{1}{2}$**

Montrons que si la condition  $\lambda \leq \frac{1}{2}$  n'est pas respectée, on peut construire une condition initiale pour lequel le schéma n'est pas stable. Soit  $u_0 \in C([-1, 1], \mathbb{R})$  qui vérifie (voir Figure (2.5)) :

$$\begin{cases} u_0(x) \geq 0 \\ u_0(x) \neq 0 \text{ si } x \in ]-1; -1 + \varepsilon[ \\ u_0(x) = 0 \text{ si } x > -1 + \varepsilon \end{cases}$$

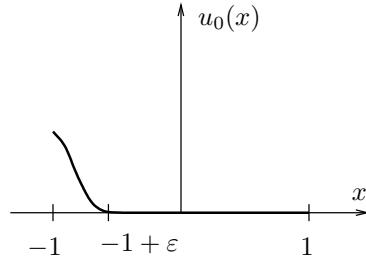


FIG. 2.5 – Condition initiale pour le contre exemple

On considère le problème :

$$\begin{cases} u_t - u_{xx} = 0, \forall x \in ]-1, 1[; \forall t > 0. \\ u(x, 0) = u_0(x), \forall x \in ]-1, 1[ \\ u(1, t) = u(-1, t) = 0, \forall t > 0. \end{cases} \quad (2.4.59)$$

On peut montrer que la solution exacte  $u$  de (2.4.59) vérifie  $u(x, t) > 0, \forall x \in ]-1, 1[, \forall t > 0$ . En particulier, pour un temps  $T > 0$  donné, on a  $u(0, T) > 0$ . Soit  $M \in \mathbb{N}$  et  $k = T/M$ . Soit  $u_i^{(n)}$  la solution approchée par (2.4.54), sensée approcher  $u(x_i, t_n)$  ( $i \in \{-N, \dots, N\}$ ,  $n \in N$ ). On va montrer que  $u_0^M = 0$  pour  $k$  et  $h$  choisis de manière non admissible ; ceci montre que le schéma ne peut pas converger. Calculons  $u_0^M$  :

$$u_0^M = (1 - 2\lambda)u_0^{M-1} + \lambda u_{-1}^{M-1} + \lambda u_1^{M-1}.$$

Donc  $u_0^M$  dépend de

$$\begin{aligned} & u^{(M-1)} \text{ sur } [-h, h] \\ & u^{(M-2)} \text{ sur } [-2h, 2h] \\ & \vdots \\ & u^{(0)} \text{ sur } [-Mh, Mh] = \left[-\frac{T}{k}h, \frac{T}{k}h\right] \end{aligned}$$

Par exemple, si on prend  $\frac{h}{k} = \frac{1}{2T}$  on obtient :  $[-\frac{T}{k}h, \frac{T}{k}h] = [-\frac{1}{2}, \frac{1}{2}]$ , et donc, si  $\varepsilon < \frac{1}{2}$ , on a  $u_0^M = 0$ . On peut donc remarquer que si  $\frac{h}{k} = \frac{1}{2T}$ , même si  $h \rightarrow 0$  et  $k \rightarrow 0$ ,

$$u_0^M \not\rightarrow u(0, T).$$

Le schéma ne converge pas ; notons que ceci n'est pas en contradiction avec le résultat de convergence 2.39 page 35, puisqu'ici, on n'a pas satisfait à la condition  $\frac{k}{h^2} \leq \frac{1}{2}$ .

### Stabilité au sens des erreurs d'arrondi

On considère le schéma d'Euler explicite pour l'équation (2.4.53). On appelle  $u$  la solution exacte de (2.4.53),  $u_D$  la solution exacte de (2.4.54),  $u_{num}$  la solution effectivement calculée. On peut écrire :

$$u - u_{num} = u - u_D + u_D - u_{num}.$$

On sait que l'erreur de discréétisation  $u - u_D$  tend vers 0 lorsque  $h$  et  $k$  tendent vers 0, sous condition de stabilité (2.4.56), c.à.d.

$$\lambda \leq \frac{1}{2}.$$

Pour contrôler l'erreur entre la solution  $u_D$  du schéma (2.4.54) et la solution numérique obtenue  $u_{num}$ , on cherche à estimer l'amplification de l'erreur commise sur la donnée initiale. Rappelons que le schéma s'écrit :

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

avec  $\lambda = \frac{k}{h^2}$ . Ce schéma se met sous la forme  $u^{(n+1)} = AU^{(n)}$ , avec

$$A = \begin{pmatrix} 1 - 2\lambda & \lambda & 0 & \dots & 0 \\ \lambda & 1 - 2\lambda & \lambda & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \lambda & 1 - 2\lambda & \lambda \\ 0 & \dots & 0 & \lambda & 1 - 2\lambda \end{pmatrix}$$

**Définition 2.40 (Stabilité au sens des erreurs d'arrondi)** *Supposons que l'on commette une erreur  $\varepsilon^0$  sur la condition initiale. La nouvelle condition initiale  $\tilde{u}^0$ , s'écrit donc  $\tilde{u}^0 = u^0 + \varepsilon^0$ . A cette nouvelle condition initiale correspond une nouvelle solution calculée  $\tilde{u}^{(n)} = u^{(n)} + \varepsilon^{(n)}$ . On dit que le schéma est stable au sens des erreurs d'arrondi s'il existe  $C > 0$  indépendant de  $n$  tel que  $\varepsilon^{(n)} \leq C\varepsilon^0$ .*

On peut trouver une condition suffisante pour que le schéma 2.4.54 soit stable au sens des erreurs d'arrondi. En effet, on va démontrer le résultat suivant :

**Proposition 2.41** *On suppose que  $\lambda = \frac{k}{h^2} < \frac{1}{2}$ . Alors le schéma 2.4.54 est stable au sens des erreurs d'arrondi.*

**Démonstration :** Soit donc une condition initiale perturbée  $\tilde{u}^0 = u^0 + \varepsilon^0$  à laquelle on associe une nouvelle solution calculée  $\tilde{u}^{(n)} = u^{(n)} + \varepsilon^{(n)}$ . On a  $\varepsilon^{(n)} = A^n \varepsilon^0$ . Comme  $A$  est symétrique,  $A$  est diagonalisable dans  $\mathbb{R}$ . Soient  $\mu_1, \dots, \mu_N$  les valeurs propres de  $A$ , et  $e_1, \dots, e_N$  les vecteurs propres associés, c'est-à-dire tels que  $Ae_i = \mu_i e_i, \forall i = 1, \dots, N$ . On décompose la perturbation  $\varepsilon^0$  sur la base des vecteurs propres :

$$\varepsilon^0 = \sum_{i=1}^N a_i e_i. \text{ On a donc } A^n \varepsilon^0 = \sum_{i=1}^N a_i \mu_i^n e_i = \varepsilon^{(n)}.$$

Si on prend par exemple :  $\varepsilon^0 = a_i e_i$ , on obtient  $\varepsilon^{(n)} = a_i \mu_i^n e_i$ . Il y a diminution de l'erreur d'arrondi sur  $\varepsilon^0$  si

$$\sup_{i=1 \dots N} |\mu_i| \leq 1$$

c'est-à-dire si  $\rho(A) \leq 1$ , où  $\rho(A)$  désigne le rayon spectral de  $A$ . Calculons  $\rho(A)$ . On écrit :  $A = I + \lambda B$  où  $B$  est la matrice symétrique définie négative, définie par :

$$B = \begin{pmatrix} -2 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1 & -2 & 1 \\ 0 & \dots & 0 & 1 & -2 \end{pmatrix} \quad (2.4.60)$$

Soit  $\mathcal{VP}(A)$  l'ensemble de valeurs propres de  $A$ . Alors  $\mathcal{VP}(A) = \{1 + \lambda\mu, \mu \in \mathcal{VP}(B)\}$ . Or  $\mathcal{VP}(B) = \{-4 \sin^2 \frac{j\pi}{2(N+1)}, j = 1, \dots, N\}$  (voir Lemme 2.42 plus loin). Pour que  $\varepsilon^{(n)} < \varepsilon^0$ , il faut donc que :

$$\sup_{j=1, \dots, N} \left| 1 - 4\lambda \sin^2 \frac{j\pi}{2(N+1)} \right| < 1,$$

c.à.d.

$$\lambda \sin^2 \frac{j\pi}{2(N+1)} < \frac{1}{2}.$$

Une condition suffisante pour avoir une diminution de l'erreur est donc que  $\lambda < \frac{1}{2}$ . ■

**Lemme 2.42 (Valeurs propres de  $B$ )** *L'ensemble  $\mathcal{VP}(B)$  des valeurs propres de la matrice  $B$  définie par (2.4.60) est donné par :*

$$\mathcal{VP}(B) = \left\{ -4 \sin^2 \frac{j\pi}{2(N+1)}, j = 1, \dots, N \right\}.$$

**Démonstration :** Les valeurs propres  $B$  peuvent se calculer à partir des valeurs propres de l'opérateur continu ; on commence donc par chercher  $u$  solution de :

$$\begin{cases} -u'' + \alpha u = 0, \\ u(0) = u(1) = 0. \end{cases}$$

Cherchons  $u(x)$  sous la forme :

$$u(x) = a \cos \sqrt{\alpha}x + b \sin \sqrt{\alpha}x$$

Comme  $u(0) = 0$ , on a :  $a = 0$ . De même,  $u(1) = B \sin \sqrt{\alpha} = 0$ , et donc  $\sqrt{\alpha} = k\pi$ . Les valeurs propres et vecteurs propres associés de l'opérateur continu sont donc :  $(k^2\pi^2, \sin k\pi x)$   $k \in \mathbb{N}^*$ . Pour  $k = 1, \dots, N$ , soit  $v^{(k)} \in \mathbb{R}^N$  tel que  $v_i^{(k)} = \sin k\pi i h$ . Calculons  $Bv^{(k)}$  :

$$(Bv^{(k)})_i = v_{i-1}^{(k)} - 2v_i^{(k)} + v_{i+1}^{(k)}$$

et donc

$$(Bv^{(k)})_i = \sin k\pi(i-1)h - 2 \sin k\pi i h + \sin k\pi(i+1)h$$

En développant, on obtient :

$$(Bv^{(k)})_i = \sin k\pi i h \cos(-k\pi h) + \cos k\pi i h \sin(-k\pi h) - 2 \sin k\pi i h + \sin k\pi i h \cos k\pi h + \cos k\pi i h \sin k\pi h.$$

Après simplifications, il vient :

$$(Bv^{(k)})_i = 2 \sin k\pi i h (-1 + \cos k\pi h).$$

Or,  $\cos k\pi h = 1 - 2 \sin^2 \frac{k\pi h}{2}$ . On a donc :

$$\begin{aligned} (Bv^{(k)})_i &= 2 \sin k\pi i h \times \left( -2 \sin^2 \frac{k\pi h}{2} \right) \\ &= -4 \sin^2 \frac{k\pi h}{2} (v^{(k)})_i, \quad \forall k = 1 \dots N. \end{aligned}$$

On a  $h = \frac{1}{N+1}$ , et donc les valeurs propres de  $B$  s'écrivent  $\mu_k = -4 \sin^2 \frac{k\pi}{2(N+1)}$ ,  $k = 1, \dots, N$ . ■

## Stabilité au sens de Von Neumann

L'analyse de stabilité au sens de Von Neumann consiste à étudier l'impact du schéma sur un mode de Fourier isolé. Pour que le mode de Fourier en question soit solution du problème continu, on remplace les conditions de Dirichlet homogènes du problème (2.4.53) par des conditions périodiques, et pour alléger les notations, on considère l'intervalle  $]0, 2\pi[$  comme intervalle d'étude en espace plutôt que l'intervalle  $]0, 1[$ .

**Problème continu avec conditions aux limites périodiques** On considère le problème avec conditions aux limites périodiques

$$\begin{cases} u_t - u_{xx} = 0, & t \in ]0, T[, x \in ]0, 2\pi[, \\ u(0, t) = u(2\pi, t), & \forall t \in ]0, T[, \\ u(x, 0) = u_0(x). \end{cases} \quad (2.4.61)$$

Le problème (2.4.61) est bien posé, au sens où  $\forall u_0 \in C([0, 2\pi])$ , il existe une unique  $u \in C^2([0, 2\pi] \times ]0, T[, \mathbb{R})$  solution de (2.4.61). On suppose que  $u_0 \in L^2([0, 2\pi])$ . On rappelle que  $L^2$  est un espace de Hilbert, et que  $\{e^{inx}, n \in \mathbb{Z}\}$  est une base hilbertienne<sup>3</sup> de  $L^2([0, 2\pi])$ . On décompose donc la condition initiale dans cette base hilbertienne :  $u_0(x) = \sum_{n \in \mathbb{Z}} c_n(0) e^{inx}$  (au sens de la convergence dans  $L^2$ ). Dans un premier temps, calculons formellement les solutions de (2.4.61) sous la forme d'un développement dans la base hilbertienne :

$$u(x, t) = \sum_{n \in \mathbb{Z}} c_n(t) e^{inx}.$$

En supposant qu'on ait le droit de dériver terme à terme, on a donc :

$$u_t(x, t) = \sum_{n \in \mathbb{Z}} c'_n(t) e^{inx} \text{ et } u_{xx}(x, t) = \sum_{n \in \mathbb{Z}} -c_n(t) n^2 e^{inx}.$$

On obtient, en remplaçant dans l'équation

$$c'_n(t) = -n^2 c_n(t)$$

c'est-à-dire  $c_n(t) = c_n(0) e^{-n^2 t}$  en tenant compte de la condition initiale. On a donc finalement :

$$u(x, t) = \sum_{n \in \mathbb{Z}} c_n(0) e^{-n^2 t} e^{inx}. \quad (2.4.62)$$

Justifions maintenant ce calcul formel. On a :

$$\sum_{n \in \mathbb{Z}} |c_n(0)|^2 = \|u^0\|_{L^2}^2 < +\infty$$

De plus, en dérivant (2.4.62) terme à terme, on obtient :

$$u_t - u_{xx} = 0.$$

La condition de périodicité est bien vérifiée par  $u$  donnée par (2.4.62). Enfin on a bien :  $u(x, t) \rightarrow u_0(t)$  lorsque  $t \rightarrow 0$ , donc la condition initiale est vérifiée. On peut remarquer qu'il y a "amortissement" des coefficients de Fourier  $c_n(0)$  lorsque  $t$  augmente, c.à.d. qu'on a :  $c_n(t) \leq c_n(0)$ ,  $\forall t > 0$ .

---

<sup>3</sup>Soit  $H$  un espace de Hilbert,  $(e_i)_{i \in \mathbb{Z}}$  est une base hilbertienne de  $H$  si :  $(e_i)_{i \in \mathbb{Z}}$  est une famille orthonormée telle que  $\forall x \in H, \exists (x_i)_{i \in \mathbb{Z}} \subset \mathbb{R} ; x = \sum_{i \in \mathbb{Z}} x_i e_i$  au sens de la convergence dans  $H$ , avec  $x_i = (x, e_i)$ , où  $(., .)$  désigne le produit scalaire sur  $H$ .

**Discrétisation du problème (2.4.61)** Si on utilise le schéma (2.4.54), pour la discrétisation de (2.4.61) on a :

$$u_j^{(n+1)} = (1 - 2\lambda)u_j^{(n)} + \lambda u_{j-1}^{(n)} + \lambda u_{j+1}^{(n)}. \quad (2.4.63)$$

On prend comme condition initiale  $u^0(x) = a_p e^{ipx}$ , pour  $p \in \mathbb{Z}$  fixé. En discrétisant, on obtient :  $u_j^0(x) = a_p e^{ipjh}$ , pour  $j = 1, \dots, N$ , avec  $h = \frac{2\pi}{N+1}$ . On a bien  $u_0^0 = u_{N+1}^0 = 0$ . Calculons :

$$u_j^{(1)} = (1 - 2\lambda)a_p e^{ipjh} + \lambda a_p e^{ip(j-1)h} + \lambda a_p e^{ip(j+1)h}$$

donc :  $u_j^{(1)} = a_p e^{ipjh} \xi_p$ . On appelle  $\xi_p$  le facteur d'amplification associé à la fonction  $e^{ipx}$  (appelé aussi “ $p$ -ième mode”). On a donc :

$$\begin{cases} u_j^{(1)} = \xi_p u_j^{(0)} \\ \vdots \\ u_j^{(n)} = (\xi_p)^n u_j^{(0)} \end{cases}$$

On dit que le schéma est “stable au sens de Von Neumann” : si :

$$|\xi_p| < 1, \quad \forall p.$$

Calculons  $\xi_p$  :

$$\begin{aligned} \xi_p &= 1 - 2\lambda + 2\lambda \cos ph \\ &= 1 - 2\lambda + 2\lambda(1 - 2 \sin^2 \frac{ph}{2}) \\ &= 1 - 4\lambda \sin^2 \left( \frac{2\pi}{N+1}, \frac{p}{2} \right). \end{aligned}$$

Pour avoir  $|\xi_p| < 1$ , il faut  $\lambda \sin^2 \left( \frac{2\pi}{N+1}, \frac{p}{2} \right) < \frac{1}{4}$ . Une condition suffisante pour que le schéma soit stable au sens de Von Neumann est que :  $\lambda < \frac{1}{2}$ . Remarquons que c'est la même condition que pour la stabilité des erreurs d'arrondis.

**Convergence du schéma avec la technique de Von Neumann** Soit  $u \in C^2([0, 2\pi] \times [0, T], \mathbb{R})$  la solution exacte de (2.4.61) on a  $u(jh, nk) = \sum_{p \in \mathbb{Z}} c_p(0) e^{-p^2 nk} e^{ipjh}$  où  $h = \frac{2\pi}{N+1}$  est le pas de discrétisation

en espace et  $k = \frac{T}{M}$  le pas de discrétisation en temps. Soit  $u_{\mathcal{D}}$  la solution de (2.4.54), et :

$$u_{\mathcal{D}}(jh, nk) = \sum_{p \in \mathbb{Z}} c_p(0) \xi_p^{(n)} e^{ipjh}.$$

On cherche à montrer la convergence de  $u_{\mathcal{D}}$  vers  $u$  au sens suivant :

**Proposition 2.43** Soit  $u_0 = \sum_{n \in \mathbb{Z}} c_n(0) e^{inx}$  et  $u$  la solution du problème (2.4.61). On note  $u_{\mathcal{D}}$  la solution approchée obtenue par le schéma d'Euler explicite (2.4.63). Alors  $\forall \varepsilon > 0$ ,  $\exists \eta \geq 0$  tel que si  $k \leq \eta$  et  $\frac{k}{h^2} \leq \frac{1}{2}$ , alors

$$|u(jh, nk) - u_{\mathcal{D}}(jh, nk)| \leq \varepsilon, \forall j = 1 \dots N, n = \frac{T}{k}.$$

**Démonstration :** On note  $(u - u_{\mathcal{D}})_j^{(n)}$  la quantité  $u(jh, nk) - u_{\mathcal{D}}(jh, nk)$ . On fera l'hypothèse supplémentaire :

$$\sum_{p \in \mathbb{Z}} |c_p(0)| < +\infty.$$

Donc pour tout  $\varepsilon \in \mathbb{R}^+$ , il existe  $A \in \mathbb{R}$  tel que  $2 \sum_{|p| \geq A} |c_p(0)| \leq \varepsilon$ . On écrit alors :

$$(u - u_{\mathcal{D}})_j^{(n)} \leq \sum_{|p| \leq A} c_p(0)(e^{-p^2 nk} - \xi_p^n) e^{ipjh} + \sum_{|p| \geq A} c_p(0)(e^{-p^2 nk} - \xi_p^n) e^{ipjh}$$

On a donc :

$$(u - u_{\mathcal{D}})_j^{(n)} \leq X + 2 \sum_{|p| \geq A} |c_p(0)|, \text{ avec } X = \sum_{|p| \leq A} |c_p(0)|(e^{-p^2 nk} - \xi_p^n)$$

et  $2 \sum_{|p| \geq A} |c_p(0)| \leq 2\varepsilon$ . Montrons maintenant que  $X \rightarrow 0$  lorsque  $h \rightarrow 0$ . Remarquons que

$$e^{-p^2 nk} - \xi_p^n = e^{-p^2 T} - \xi_p^n, \text{ et } \xi_p = 1 - 4\lambda \sin^2 \frac{ph}{2}.$$

Or,  $\sin^2 \frac{ph}{2} = \frac{p^2 h^2}{4} + O(h^4)$ , et  $\lambda = \frac{k}{h^2}$ . Donc :  $4\lambda \sin^2 \frac{ph}{2} = p^2 k + O(kh^2)$ . On en déduit :

$$(\xi_p)^n = \left(1 - 4\lambda \sin^2 \frac{ph}{2}\right)^{T/k} \text{ et donc } \ln \xi_p^n = \frac{T}{k} \ln \left(1 - 4\lambda \sin^2 \frac{ph}{2}\right) = -Tp^2 + O(h^2).$$

On en déduit que  $\xi_p^n \rightarrow e^{-p^2 T}$  lorsque  $h \rightarrow 0$ . Tous les termes de  $X$  tendent vers 0, et  $X$  est une somme finie ; on a donc montré que  $(u - u_{\mathcal{D}})_j^{(n)}$  tend vers 0 lorsque  $h$  tend vers 0. ■

**Remarque 2.44** *On peut adapter la technique de Von Neumann au cas Dirichlet homogène sur  $[0, 1]$ , en effectuant le développement de  $u$  par rapport aux fonctions propres de l'opérateur  $u''$  avec conditions aux limites de Dirichlet :*

$$u(x, t) = \sum c_n(t) \sin(n\pi x).$$

*L'avantage du développement en série de Fourier est qu'il marche pour n'importe quel opérateur linéaire à condition d'avoir pris des conditions aux limites périodiques.*

### 2.4.3 Schéma implicite et schéma de Crank-Nicolson

Commençons par un petit rappel sur les équations différentielles (voir aussi polycopié d'analyse numérique de licence, sur le site web <http://www.cmi.univ-mrs.fr/~herbin>). On considère le problème de Cauchy :

$$\begin{cases} y'(t) &= f(y(t)), t > 0. \\ y(0) &= y_0 \end{cases} \quad (2.4.64)$$

Soit  $k$  un pas (constant) de discrétisation, on rappelle que les schémas d'Euler explicite et implicite pour la discrétisation de ce problème s'écrivent respectivement :

$$\text{Euler explicite : } \frac{y^{(n+1)} - y^{(n)}}{k} = f(y^{(n)}), \quad n \geq 0 \quad (2.4.65)$$

$$\text{Euler implicite : } \frac{y^{(n+1)} - y^{(n)}}{k} = f(y^{(n+1)}), \quad n \geq 0, \quad (2.4.66)$$

avec  $y^{(n)} = y_0$ . On rappelle également que le  $\theta$ -schéma, où  $\theta$  est un paramètre de l'intervalle  $[0, 1]$  s'écrit :

$$y^{(n+1)} = y^{(n)} + k\theta f(y^{(n+1)}) + k(1-\theta)f(y^{(n)}). \quad (2.4.67)$$

Remarquons que pour  $\theta = 0$  on retrouve le schéma (2.4.65) et pour  $\theta = 1$  le schéma (2.4.66). On peut facilement adapter le  $\theta$  schéma à la résolution des équations paraboliques. Par exemple, le  $\theta$ -schéma pour la discréétisation en temps du problème (2.4.53), avec une discréétisation par différences finies en espace s'écrit :

$$\begin{aligned} \frac{u_i^{(n+1)} - u_i^{(n)}}{k} &= \frac{\theta}{h^2}(-2u_i^{(n+1)} + u_{i-1}^{(n+1)} + u_{i+1}^{(n+1)}) + \frac{1-\theta}{h^2}(-2u_i^{(n)} + u_{i-1}^{(n)} + u_{i+1}^{(n)}), \quad ; n \geq 0, \quad i = 1, \dots, N \\ u_i^{(0)} &= u_0(x_i), \quad i = 1, \dots, N. \end{aligned} \quad (2.4.68)$$

Si  $\theta = 0$ , on retrouve le schéma d'Euler explicite ; si  $\theta = 1$ , celui d'Euler implicite. Dans ce cas où  $\theta = \frac{1}{2}$  ce schéma s'appelle schéma de Crank-Nicolson. Notons que dès que  $\theta > 0$ , le schéma est implicite, au sens où on n'a pas d'expression explicite de  $u_i^{(n+1)}$  en fonction des  $u_j^{(n)}$ .

**Proposition 2.45 (Consistance du  $\theta$ -schéma)** *Le  $\theta$  schéma (2.4.68) pour la discréétisation du problème (2.4.53) est d'ordre 2 en espace. Il est d'ordre 2 en temps si  $\theta = \frac{1}{2}$ , et d'ordre 1 sinon.*

**Démonstration :** On pose  $\bar{u}_j^n = u(x_j, t_n)$ ,  $h = \frac{1}{N+1}$ ,

$$R_j^{(n)} = \frac{\bar{u}_j^{(n+1)} - \bar{u}_j^{(n)}}{k} + \frac{\theta}{h^2} \left( -2\bar{u}_i^{(n+1)} + \bar{u}_{i-1}^{(n+1)} + \bar{u}_{i+1}^{(n+1)} \right) + \frac{1-\theta}{h^2} \left( -2\bar{u}_i^{(n)} + \bar{u}_{i-1}^{(n)} + \bar{u}_{i+1}^{(n)} \right)$$

On va montrer, en effectuant des développements limités, que :  $|R_j^{(n)}| \leq C(k + h^2)$  si  $\theta \neq \frac{1}{2}$  et que  $|R_j^{(n)}| \leq C(k^2 + h^2)$  si  $\theta = \frac{1}{2}$ . En effet, on décompose

$$R_j^{(n)} = T_j^{(n,1)} + \theta T_j^{(n,2)} + (1-\theta)T_j^{(n,3)}$$

avec :

$$\begin{aligned} T_j^{(n,1)} &= \frac{\bar{u}_j^{(n+1)} - \bar{u}_j^{(n)}}{k}, \quad T_j^{(n,2)} = \frac{\theta}{h^2} \left( -2\bar{u}_i^{(n+1)} + \bar{u}_{i-1}^{(n+1)} + \bar{u}_{i+1}^{(n+1)} \right) \\ T_j^{(n,3)} &= \frac{1-\theta}{h^2} \left( -2\bar{u}_i^{(n)} + \bar{u}_{i-1}^{(n)} + \bar{u}_{i+1}^{(n)} \right) \end{aligned}$$

Effectuons un développement limité pour calculer  $T_j^{(n,1)}$  :

$$T_j^{(n,1)} = (\bar{u}_t)(x_j, t_n) + \frac{k}{2}(u_{tt})(x_j, t_n) + R_1 \quad \text{avec } |R_1| \leq Ck^2.$$

Faisons de même pour  $T_j^{(n,2)}$  :

$$T_j^{(n,2)} = \theta(\bar{u}_{xx}(x_j, t_{n+1}) + R_2) \quad \text{avec } |R_2| \leq Ch^2.$$

Or  $\bar{u}_{xx}(x_j, t_{n+1}) = \bar{u}_{xx}(x_j, t_n) + k\bar{u}_{xxt}(x_j, t_n) + R_3$  avec  $|R_3| \leq Ck^2$ , donc :

$$T_j^{(n,2)} = \theta(u_{xx}(x_j, t_n) + k u_{xxt}(x_j, t_n) + R_4) \quad \text{avec } |R_4| \leq C(h^2 + k^2).$$

De même pour  $T_j^{(n,3)}$ , on a :

$$T_j^{(n,3)} = (1 - \theta)u_{xx}(x_j, t_n) + R_5, \text{ avec } |R_5| \leq Ck^2.$$

En regroupant, on obtient que

$$R_j^{(n)} = u_t(x_j, t_n) - u_{xx}(x_j, t_n) \frac{k}{2} \frac{\partial}{\partial t} u_t(x_j, t_n) + \theta k(u_{xx})(x_j, t_n) + R$$

avec  $R = R_1 + R_4 + R_5$

- Si  $\theta = \frac{1}{2}$ , on a un schéma d'ordre 2 en temps et en espace.

En effet,  $\frac{k}{2}(\bar{u}_{tt})(x_j, t_n) - \theta k(\bar{u}_{xxt})(x_j, t_n) = \frac{\partial}{\partial t} \left( k \left[ \frac{1}{2}(\bar{u}_t)(x_j, t_n) - \theta(\bar{u}_{xx})(x_j, t_n) \right] \right)$  et  $u_t - u_{xx} = 0$ .

- Si  $\theta \neq \frac{1}{2}$  : on a un schéma d'ordre 2 en espace et d'ordre 1 en temps. ■

**Proposition 2.46 (Stabilité au sens de Von Neumann)** *Si  $\theta \geq \frac{1}{2}$  le  $\theta$ -schéma est inconditionnellement stable. En particulier, les schémas d'Euler implicite et de Crank-Nicolson sont inconditionnellement stables. Si  $\theta < \frac{1}{2}$  le schéma est stable sous condition.*

$$\lambda \leq \frac{1}{2(1 - 2\theta)}.$$

(On retrouve en particulier que le schéma d'Euler explicite n'est que si  $\lambda \leq \frac{1}{2}$ ).

**Démonstration :** On remplace les conditions aux limites de Dirichlet sur  $[0, 1]$  par des conditions périodiques sur  $[0, 2\pi]$ . La solution exacte écrit alors :

$$u = \sum_{p \in \mathbb{Z}} c_p(0) e^{-p^2 t} e^{ipx}.$$

Pronons comme condition initiale  $u_0(x) = e^{ipx}$ . On a :

$$u_j^{(n+1)} - u_j^{(n)} = \frac{k}{h^2} \left[ -\theta(2u_j^{(n+1)} - u_{j-1}^{(n+1)} - u_{j+1}^{(n+1)}) - (1 - \theta)(2u_j^{(n)} - u_{j-1}^{(n)} - u_{j+1}^{(n)}) \right],$$

ce qui écrit encore, avec :  $\lambda = \frac{k}{h^2}$  :

$$(1 + 2\lambda)u_j^{(n+1)} - \lambda\theta u_{j-1}^{(n+1)} - \lambda\theta u_{j+1}^{(n+1)} = (1 - 2\lambda(1 - \theta))u_j^{(n)} + \lambda(1 - \theta)u_{j+1}^{(n)} + \lambda(1 - \theta)u_{j-1}^{(n)}. \quad (2.4.69)$$

En discrétilisant la condition initiale (mode de Fourier) on obtient  $u_j^{(0)} = e^{ipjh}$  et on cherche le facteur d'amplification  $\xi_p$  tel que  $u_j^1 = \xi_p u_j^0 = \xi_p e^{ipjh}$  ; en appliquant le schéma ci-dessus pour  $n = 0$ , on obtient :

$$(1 + 2\lambda\theta)\xi_p - \lambda\theta\xi_p[e^{-iph} + e^{iph}] = [1 - 2\lambda(1 - \theta)] + \lambda(1 - \theta)[e^{iph} + e^{iph}]$$

et donc :

$$\xi_p = \frac{1 - 2\lambda(1 - \theta) + 2\lambda(1 - \theta)\cos ph}{(1 + 2\lambda\theta) - 2\lambda\cos ph} = \frac{1 - 4\lambda(1 - \theta)\sin^2 ph/2}{1 + 4\lambda\theta\sin^2 \frac{ph}{2}}$$

Pour que le schéma soit stable au sens de Von Neumann, il faut que :  $|\xi_p| < 1$  pour tout  $p$ , soit encore :

$$1 - 4\lambda(1 - \theta) \sin^2 \frac{ph}{2} < 1 + 4\lambda\theta \sin^2 \frac{ph}{2} \quad (2.4.70)$$

et

$$4\lambda(1 - \theta) \sin^2 \frac{ph}{2} - 1 < 1 + 4\lambda\theta \sin^2 \frac{ph}{2} \quad (2.4.71)$$

L'inégalité (2.4.70) est toujours vérifiée. En ce qui concerne l'inégalité (2.4.71), on distingue deux cas :

1. Si  $\theta \leq \frac{1}{2}$  alors  $0 \leq 1 - \theta \leq \theta$  et dans ce cas (2.4.71) est toujours vraie.
2. Si  $\theta < \frac{1}{2}$ , on veut :

$$4\lambda \left[ (1 - \theta) \sin^2 \frac{ph}{2} - \theta \sin^2 \frac{ph}{2} \right] < 2$$

Il faut donc que

$$\lambda < \frac{1}{2} \left\{ (1 - 2\theta) \sin^2 \frac{ph}{2} \right\}^{-1}$$

Une condition suffisante est donc :

$$\lambda \leq \frac{1}{2(1 - 2\theta)} \text{ si } \theta < \frac{1}{2}. \quad \blacksquare$$

### Convergence du schéma d'Euler implicite.

Prenons  $\theta = 1$  dans le  $\theta$ -schéma : on obtient le schéma d'Euler implicite :

$$(1 + 2\lambda)u_j^{(n+1)} - \lambda u_{j-1}^{(n+1)} - \lambda u_{j+1}^{(n+1)} = u_j^{(n)} \quad (2.4.72)$$

On rappelle que ce schéma est inconditionnellement stable au sens de Von Neumann. On va montrer de plus qu'il est  $L^\infty$ -stable :

**Proposition 2.47 (Stabilité  $L^\infty$  pour Euler implicite)** *Si  $(u_j^{(n)})_{j=1,\dots,N}$  est solution du schéma (2.4.72), alors :*

$$\max_{j=1,\dots,N} u_j^{(n+1)} \leq \max_{j=1,\dots,N} u_j^{(n)} \leq \max_{j=1,\dots,N} u_j^{(0)} \quad (2.4.73)$$

de même :

$$\min_{j=1,\dots,N} u_j^{(n+1)} \geq \min_{j=1,\dots,N} u_j^{(n)} \geq \min_{j=1,\dots,N} u_j^{(0)} \quad (2.4.74)$$

Le schéma (2.4.72) est donc  $L^\infty$  stable.

**Démonstration :** Prouvons l'estimation (2.4.73), la preuve de (2.4.74) est similaire. Soit  $j_0$  tel que  $u_{j_0}^{(n+1)} = \max_{j=1,\dots,N} u_j^{(n+1)}$  Par définition du schéma d'Euler implicite (2.4.72), On a :

$$u_{j_0}^{(n)} = (1 + 2\lambda)u_{j_0}^{(n+1)} - \lambda u_{j_0-1}^{(n+1)} - \lambda u_{j_0+1}^{(n+1)}.$$

On en déduit :  $u_{j_0}^{(n+1)} \leq \max_{j=1,\dots,N} u_j^{(n)}$ , ce qui prouve que

$$\max_{j=1,\dots,N} u_j^{(n+1)} \leq \max_{j=1,\dots,N} u_j^{(n)}.$$

Donc le schéma (2.4.72) est  $L^\infty$  stable. ■

**Théorème 2.48** Soit  $e^{(n)}$  l'erreur de discrétisation, définie par

$$e_j^{(n)} = u(x_j, t_n) - u_j^{(n)} \text{ pour } j = 1, \dots, N.$$

Alors  $\|e^{(n+1)}\|_\infty \leq \|e^{(0)}\|_\infty + TC(k + h^2)$ . Si  $\|e^{(0)}\|_\infty = 0$ , le schéma est donc convergent (d'ordre 1 en temps et 2 en espace).

**Démonstration :** En utilisant la définition de l'erreur de consistance, on obtient :

$$(1 + 2\lambda)e_j^{(n+1)} - \lambda e_{j-1}^{(n)} - \lambda e_{j+1}^{(n)} = e_j^{(n)} + R_j^{(n)}$$

et donc :

$$\|e^{(n+1)}\|_\infty \leq \|e^{(n)}\|_\infty + kC(k + h^2)$$

On en déduit, par récurrence sur  $n$ , que :

$$\|e^{(n+1)}\|_\infty \leq \|e^{(0)}\|_\infty + TC(k + h^2)$$

d'où la convergence du schéma. ■

On peut montrer que le schéma saute-mouton (ou “Leap-frog”)

$$\frac{u_j^{(n+1)} - u_j^{(n-1)}}{2k} = \frac{1}{h^2}(u_{j-1}^{(n)} - 2u_j^{(n)} + u_{j+1}^{(n)})$$

est d'ordre 2 en espace et en temps (voir exercice 21 page 57). Malheureusement il est aussi inconditionnellement instable. On peut le modifier pour le rendre stable, en introduisant le schéma Dufort-Frankel, qui s'écrit :

$$\frac{u_j^{(n+1)} - u_j^{(n-1)}}{2k} = \frac{1}{h^2}(u_{j-1}^{(n)} - (u_j^{(n+1)} + u_j^{(n-1)}) + u_{j+1}^{(n)})$$

Ce schéma est consistant et inconditionnellement stable.

#### 2.4.4 Cas de la Dimension 2

Soit  $\Omega$  un ouvert borné de  $\mathbb{R}^2$ , on considère le problème suivant :

$$\begin{cases} u_t - \Delta u = 0 & x \in \Omega, t \in ]0, T[ \\ u(x, 0) = u_0(x) & x \in \Omega \\ u(x, t) = g(t) & x \in \partial\Omega \quad \forall t \in ]0, T[ \end{cases}$$

Si le domaine est rectangulaire, ce problème se discrétise facilement à l'aide de  $\theta$  schéma en temps et de différences finies en espace, en prenant un maillage rectangulaire. On peut montrer, comme dans le cas 1D, la consistance, la stabilité, la  $L^\infty$  stabilité, la stabilité au sens de Von Neumann

## 2.5 Exercices

**Exercice 1 (Comparaison différences finies- volumes finis)** *Suggestions en page 63, corrigé en page 66.*

On considère le problème :

$$\begin{aligned} -u''(x) &= f(x), \quad x \in ]0, 1[, \\ u(0) &= a, \quad u(1) = b, \end{aligned} \tag{2.5.75}$$

Ecrire les schémas de différences finies et volumes finis avec pas constant pour le problème (2.5.75), et comparer les schémas ainsi obtenus.

**Exercice 2 (Conditionnement “efficace”.)** *Suggestions en page 63, corrigé en page 66.*

Soit  $f \in C([0, 1])$ . Soit  $N \in \mathbb{N}^*$ ,  $N$  impair. On pose  $h = 1/(N+1)$ . Soit  $A$  la matrice définie par (2.2.24) page 17, issue d'une discréétisation par différences finies (vue en cours) du problème (2.3.27) page 19.

Pour  $u \in \mathbb{R}^N$ , on note  $u_1, \dots, u_N$  les composantes de  $u$ . Pour  $u \in \mathbb{R}^N$ , on dit que  $u \geq 0$  si  $u_i \geq 0$  pour tout  $i \in \{1, \dots, N\}$ . Pour  $u, v \in \mathbb{R}^N$ , on note  $u \cdot v = \sum_{i=1}^N u_i v_i$ .

On munit  $\mathbb{R}^N$  de la norme suivante : pour  $u \in \mathbb{R}^N$ ,  $\|u\| = \max\{|u_i|, i \in \{1, \dots, N\}\}$ . On munit alors  $\mathcal{M}_N(\mathbb{R})$  de la norme induite, également notée  $\|\cdot\|$ , c'est-à-dire  $\|B\| = \max\{\|Bu\|, u \in \mathbb{R}^N \text{ t.q. } \|u\| = 1\}$ , pour tout  $B \in \mathcal{M}_N(\mathbb{R})$ .

**Partie I** Conditionnement de la matrice et borne sur l'erreur relative

1. (Existence et positivité de  $A^{-1}$ ) Soient  $b \in \mathbb{R}^N$  et  $u \in \mathbb{R}^N$  t.q.  $Au = b$ . Remarquer que  $Au = b$  peut s'écrire :

$$\begin{cases} \frac{1}{h^2}(u_i - u_{i-1}) + \frac{1}{h^2}(u_i - u_{i+1}) = b_i, \quad \forall i \in \{1, \dots, N\}, \\ u_0 = u_{N+1} = 0. \end{cases} \tag{2.5.76}$$

Montrer que  $b \geq 0 \Rightarrow u \geq 0$ . [On pourra considérer  $p \in \{0, \dots, N+1\}$  t.q.  $u_p = \min\{u_j, j \in \{0, \dots, N+1\}\}$ .]

En déduire que  $A$  est inversible.

2. (Préliminaire...) On considère la fonction  $\varphi \in C([0, 1], \mathbb{R})$  définie par  $\varphi(x) = (1/2)x(1-x)$  pour tout  $x \in [0, 1]$ . On définit alors  $\phi \in \mathbb{R}^N$  par  $\phi_i = \phi(ih)$  pour tout  $i \in \{1, \dots, N\}$ . Montrer que  $(A\phi)_i = 1$  pour tout  $i \in \{1, \dots, N\}$ .
3. (calcul de  $\|A^{-1}\|$ ) Soient  $b \in \mathbb{R}^N$  et  $u \in \mathbb{R}^N$  t.q.  $Au = b$ . Montrer que  $\|u\| \leq (1/8)\|b\|$  [Calculer  $A(u \pm \|b\|\phi)$  avec  $\phi$  défini à la question 2 et utiliser la question 1]. En déduire que  $\|A^{-1}\| \leq 1/8$  puis montrer que  $\|A^{-1}\| = 1/8$ .
4. (calcul de  $\|A\|$ ) Montrer que  $\|A\| = \frac{4}{h^2}$ .
5. (Conditionnement pour la norme  $\|\cdot\|$ ) Calculer  $\|A^{-1}\| \|A\|$ . Soient  $b, \delta_b \in \mathbb{R}^N$ . Soient  $u, \delta_u \in \mathbb{R}^N$  t.q.  $Au = b$  et  $A(u + \delta_u) = b + \delta_b$ . Montrer que  $\frac{\|\delta_u\|}{\|u\|} \leq \|A^{-1}\| \|A\| \frac{\|\delta_b\|}{\|b\|}$ .

Montrer qu'un choix convenable de  $b$  et  $\delta_b$  donne l'égalité dans l'inégalité précédente.

**Partie II** Borne réaliste sur l'erreur relative : Conditionnement “efficace”

On se donne maintenant  $f \in C([0, 1], \mathbb{R})$  et on suppose (pour simplifier...) que  $f(x) > 0$  pour tout  $x \in ]0, 1[$ . On prend alors, dans cette partie,  $b_i = f(ih)$  pour tout  $i \in \{1, \dots, N\}$ . On considère aussi le vecteur  $\varphi$  défini à la question 2 de la partie I.

- Montrer que  $h \sum_{i=1}^N b_i \varphi_i \rightarrow \int_0^1 f(x) \phi(x) dx$  quand  $N \rightarrow \infty$  et que  $\sum_{i=1}^N b_i \varphi_i > 0$  pour tout  $N$ . En déduire qu'il existe  $\alpha > 0$ , ne dépendant que de  $f$ , t.q.  $h \sum_{i=1}^N b_i \varphi_i \geq \alpha$  pour tout  $N \in \mathbb{N}^*$ .
- Soit  $u \in \mathbb{R}^N$  t.q.  $Au = b$ . Montrer que  $N\|u\| \geq \sum_{i=1}^N u_i = u \cdot A\varphi \geq \frac{\alpha}{h}$  (avec  $\alpha$  donné à la question 1). Soit  $\delta_b \in \mathbb{R}^N$  et  $\delta_u \in \mathbb{R}^N$  t.q.  $A(u + \delta_u) = b + \delta_b$ . Montrer que  $\frac{\|\delta_u\|}{\|u\|} \leq \frac{\|f\|_{L^\infty([0,1])}}{8\alpha} \frac{\|\delta_b\|}{\|b\|}$ .
- Comparer  $\|A^{-1}\| \|A\|$  (question I.5) et  $\frac{\|f\|_{L^\infty([0,1])}}{8\alpha}$  (question II.2) quand  $N$  est “grand” (ou quand  $N \rightarrow \infty$ ).

**Exercice 3 (Conditionnement, réaction diffusion 1d.)** *Corrigé en page 69.*

On s'intéresse au conditionnement pour la norme euclidienne de la matrice issue d'une discréétisation par Différences Finies du problème aux limites suivant :

$$\begin{aligned} -u''(x) + u(x) &= f(x), \quad x \in ]0, 1[, \\ u(0) &= u(1) = 0. \end{aligned} \quad (2.5.77)$$

Soit  $N \in \mathbb{N}^*$ . On note  $U = (u_j)_{j=1 \dots N}$  une “valeur approchée” de la solution  $u$  du problème (2.5.77) aux points  $\left(\frac{j}{N+1}\right)_{j=1 \dots N}$ . On rappelle que la discréétisation par différences finies de ce problème consiste à chercher  $U$  comme solution du système linéaire  $AU = \left(f\left(\frac{j}{N+1}\right)\right)_{j=1 \dots N}$  où la matrice  $A \in M_N(\mathbb{R})$  est définie par  $A = (N+1)^2 B + Id$ ,  $Id$  désigne la matrice identité et

$$B = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix}$$

- (Valeurs propres de la matrice  $B$ .)

On rappelle que le problème aux valeurs propres

$$\begin{aligned} -u''(x) &= \lambda u(x), \quad x \in ]0, 1[, \\ u(0) &= u(1) = 0. \end{aligned} \quad (2.5.78)$$

admet la famille  $(\lambda_k, u_k)_{k \in \mathbb{N}^*}$ ,  $\lambda_k = (k\pi)^2$  et  $u_k(x) = \sin(k\pi x)$  comme solution. Montrer que les vecteurs  $U_k = \left(u_k\left(\frac{j}{N+1}\right)\right)_{j=1 \dots N}$  sont des vecteurs propres de la matrice  $B$ . En déduire toutes les valeurs propres de la matrice  $B$ .

- En déduire les valeurs propres de la matrice  $A$ .
- En déduire le conditionnement pour la norme euclidienne de la matrice  $A$ .

**Exercice 4 (Erreur de consistance)** *Suggestions en page 63, corrigé en page 69.*

On considère la discréétisation à pas constant par le schéma aux différences finies symétrique à trois points (vu en cours) du problème (2.3.27) page 19, avec  $f \in C([0, 1])$ . Soit  $N \in \mathbb{N}^*$ ,  $N$  impair. On pose  $h = 1/(N+1)$ . On note  $u$  la solution exacte,  $x_i = ih$ , pour  $i = 1, \dots, N$  les points de discréétisation, et  $(u_i)_{i=1, \dots, N}$  la solution du système discréétisé.

1. Montrer que si  $f$  est constante, alors

$$\max_{1 \leq i \leq N} |u_i - u(x_i)| = 0.$$

2. Soit  $N$  fixé, et  $\max_{1 \leq i \leq N} |u_i - u(x_i)| = 0$ . A-t-on forcément que  $f$  est constante sur  $[0, 1]$ ? (justifier la réponse.)

**Exercice 5 (Principe du maximum)** *Suggestions en page 63, corrigé en page 69*

On considère le problème :

$$\begin{cases} -u''(x) + c(x)u(x) = f(x), & 0 < x < 1, \\ u(0) = a, u(1) = b, \end{cases} \quad (2.5.79)$$

où  $c \in C([0, 1], \mathbb{R}_+)$ , et  $c \in C([0, 1], \mathbb{R})$ , et  $(a, b) \in \mathbb{R}^2$ .

1. Donner la discréétisation par différences finies de ce problème. On appelle  $U_h$  la solution approchée (c.à.d.  $U_h = (u_1, \dots, u_N)^t$ , où  $u_i$  est l'inconnue discrète en  $x_i$ ).
2. On suppose ici que  $c = 0$ . Montrer que  $u_i \geq \min(a, b)$ , pour tout  $i = 1, \dots, N$ .

**Exercice 6 (Problème elliptique 1d, discréétisation par différences finies)** <sup>4</sup> *Suggestions en page 63, corrigé en page 70.*

Soit  $f \in C^2([0, 1])$ . On s'intéresse au problème suivant :

$$\begin{aligned} -u_{xx}(x) + \frac{1}{1+x}u_x(x) &= f(x), \quad x \in ]0, 1[, \\ u(0) = a, u(1) &= b. \end{aligned} \quad (2.5.80)$$

On admet que ce problème admet une et une seule solution  $u$  et on suppose que  $u \in C^4(]0, 1[)$ . On cherche une solution approchée de (2.5.80) par la méthode des différences finies. Soit  $n \in \mathbb{N}^*$ , et  $h = \frac{1}{N+1}$ . On note  $u_i$  la valeur approchée recherchée de  $u$  au point  $ih$ , pour  $i = 0, \dots, N+1$ . On utilise les approximations centrées les plus simples de  $u_x$  et  $u_{xx}$  aux points  $ih$ ,  $i = 1, \dots, n$ . On pose  $u_h = (u_1, \dots, u_n)^t$ .

1. Montrer que  $u_h$  est solution d'un système linéaire de la forme  $A_h u_h = b_h$ ; donner  $A_h$  et  $b_h$ .
2. Montrer que le schéma numérique obtenu est consistant et donner une majoration de l'erreur de consistance (on rappelle que l'on a supposé  $u \in C^4$ ).
3. Soit  $v \in \mathbb{R}^n$ , montrer que  $A_h v \geq 0 \Rightarrow v \geq 0$  (ceci s'entend composante par composante). Cette propriété s'appelle conservation de la positivité. En déduire que  $A_h$  est monotone.
4. On définit  $\theta$  par

$$\theta(x) = -\frac{1}{2}(1+x)^2 \ln(1+x) + \frac{2}{3}(x^2 + 2x) \ln 2, \quad x \in [0, 1].$$

- 4.a. Montrer qu'il existe  $C \geq 0$ , indépendante de  $h$ , t.q.

$$\max_{1 \leq i \leq n} \left| \frac{1}{h^2}(-\theta_{i-1} + 2\theta_i - \theta_{i+1}) + \frac{1}{2h(1+ih)}(\theta_{i+1} - \theta_{i-1}) - 1 \right| \leq Ch^2,$$

<sup>4</sup>Cet exercice est tiré du livre *Exercices d'analyse numérique matricielle et d'optimisation*, de P.G. Ciarlet et J.M. Thomas, Collection Mathématiques pour la maîtrise, Masson, 1982

avec  $\theta_i = \theta(x_i)$ ,  $i = 0, \dots, n + 1$ .

4.b On pose  $\theta_h = (\theta_1, \dots, \theta_n)^t$ . Montrer que  $(A_h \theta_h)_i \geq 1 - Ch^2$ , pour  $i = 1, \dots, N$ .

4.c Montrer qu'il existe  $M \geq 0$  ne dépendant pas de  $h$  t.q.  $\|A_h^{-1}\|_\infty \leq M$ .

5. Montrer la convergence, en un sens à définir, de  $u_h$  vers  $u$ .

6. Que peut on dire si  $u \notin C^4$ , mais seulement  $u \in C^2$  ou  $C^3$  ?

7. On remplace dans (3.9.71)  $\frac{1}{1+x}$  par  $\alpha u_x(x)$ , avec  $\alpha$  donné (par exemple  $\alpha = 100$ ). On utilise pour approcher (3.9.71) le même principe que précédemment (approximations centrées de  $u_x$  et  $u_{xx}$ ). Que peut on dire sur la consistance, la stabilité, la convergence du schéma numérique ?

**Exercice 7 (Non consistance des volumes finis)** *Suggestions en page 63, corrigé en page 74*

Montrer que la discrétisation de l'opérateur  $-u''$  par le schéma volumes finis n'est pas toujours consistante au sens des différences finies, *i.e.* que l'erreur de consistance définie par (voir remarque 2.21 page 20)

$$R_i = \frac{1}{h_i} \left[ \frac{-1}{h_{i+1/2}} (u(x_{i+1}) - u(x_i)) + \frac{1}{h_{i-1/2}} (u(x_i) - u(x_{i-1})) \right] - u''(x_i)$$

ne tend pas toujours vers 0 lorsque  $h$  tend vers 0.

**Exercice 8 (Consistance des flux)** *Corrigé en page 75 Corrigé en page 75*

Montrer que le flux défini par (2.3.30) est consistant d'ordre 1 dans le cas général, et qu'il est d'ordre 2 si  $x_{i+1/2} = (x_{i+1} + x_i)/2$ .

**Exercice 9 (Conditions aux limites de Neumann)** *Suggestions en page 64, corrigé en page 75*

On considère ici l'équation le problème de diffusion réaction avec conditions aux limites de Neumann homogènes (correspondant à une condition physique de flux nul sur le bord) :

$$\begin{cases} -u''(x) + cu(x) = f(x), & x \in ]0, 1[, \\ u'(0) = u'(1) = 0, \end{cases} \quad (2.5.81)$$

avec  $c \in \mathbb{R}_+^*$ , et  $f \in C([0, 1])$ . Donner la discrétisation de ce problème par

1. différences finies,
2. volumes finis

Montrer que les matrices obtenues ne sont pas inversibles. Proposer une manière de faire en sorte que le problème soit bien posé, compatible avec ce qu'on connaît du problème continu.

**Exercice 10 (Conditions aux limites de Fourier (ou Robin))** *Suggestions en page 64, corrigé en page 76*

On considère le problème :

$$\begin{cases} -u''(x) + cu(x) = f(x), & x \in ]0, 1[, \\ u'(0) - \alpha(u - \tilde{u}) = 0, \\ u'(1) + \alpha(u - \tilde{u}) = 0, \end{cases} \quad (2.5.82)$$

avec  $c \in \mathbb{R}_+$ ,  $f \in C([0, 1])$ ,  $\alpha \in \mathbb{R}_+^*$ , et  $\tilde{u} \in \mathbb{R}$ .

Donner la discrétisation de ce problème par

1. différences finies,
2. volumes finis

Dans les deux cas, écrire le schéma sous la forme d'un système linéaire de  $N$  équations à  $N$  inconnues, en explicitant matrice et second membre ( $N$  est le nombre de noeuds internes en différences finies, de mailles en volumes finis).

**Exercice 11 (Problème elliptique 1d, discréétisation par volumes finis)** *Suggestions en page 64, corrigé en page 77*

Soient  $a, b \geq 0$ ,  $c, d \in \mathbb{R}$  et  $f \in C([0, 1], \mathbb{R})$ ; on cherche à approcher la solution  $u$  du problème suivant :

$$-u_{xx}(x) + au_x(x) + b(u(x) - f(x)) = 0, \quad x \in [0, 1], \quad (2.5.83)$$

$$u(0) = c, \quad u(1) = d. \quad (2.5.84)$$

On suppose (mais il n'est pas interdit d'expliquer pourquoi...) que (2.5.83)-(2.5.84) admet une solution unique  $u \in C^2([0, 1], \mathbb{R})$ .

Soient  $N \in \mathbb{N}^*$  et  $h_1, \dots, h_N > 0$  t.q.  $\sum_{i=1}^N h_i = 1$ . On pose  $x_{\frac{1}{2}} = 0$ ,  $x_{i+\frac{1}{2}} = x_{i-\frac{1}{2}} + h_i$ , pour  $i = 1, \dots, N$  (de sorte que  $x_{N+\frac{1}{2}} = 1$ ),  $h_{i+\frac{1}{2}} = \frac{h_{i+1} + h_i}{2}$ , pour  $i = 1, \dots, N-1$ , et  $f_i = \frac{1}{h_i} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} f(x) dx$ , pour  $i = 1, \dots, N$ .

Pour approcher la solution  $u$  de (2.5.83)-(2.5.84), on propose le schéma numérique suivant :

$$F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}} + bh_i u_i = bh_i f_i, \quad i \in \{1, \dots, N\}, \quad (2.5.85)$$

avec  $(F_{i+\frac{1}{2}})_{i \in \{0, \dots, N\}}$  donné par les expressions suivantes :

$$F_{i+\frac{1}{2}} = -\frac{u_{i+1} - u_i}{h_{i+\frac{1}{2}}} + au_i, \quad i \in \{1, \dots, N-1\}, \quad (2.5.86)$$

$$F_{\frac{1}{2}} = -\frac{u_1 - c}{\frac{h_1}{2}} + ac, \quad F_{N+\frac{1}{2}} = -\frac{d - u_N}{\frac{h_N}{2}} + au_N. \quad (2.5.87)$$

En tenant compte des expressions (2.5.86) et (2.5.87), le schéma numérique (2.5.85) donne donc un système de  $N$  équations à  $N$  inconnues (les inconnues sont  $u_1, \dots, u_N$ ).

1. Expliquer comment, à partir de (2.5.83) et (2.5.84), on obtient ce schéma numérique.
2. (Existence de la solution approchée.)
  - (a) On suppose ici que  $c = d = 0$  et  $f_i = 0$  pour tout  $i \in \{1, \dots, N\}$ . Montrer qu'il existe un unique vecteur  $U = (u_1, \dots, u_N)^t \in \mathbb{R}^N$  solution de (2.5.85). Ce vecteur est obtenu en prenant  $u_i = 0$ , pour tout  $i \in \{1, \dots, N\}$ . (On rappelle que dans (2.5.85) les termes  $F_{i+\frac{1}{2}}$  et  $F_{i-\frac{1}{2}}$  sont donnés par (2.5.86) et (2.5.87).)
  - (b) On revient maintenant au cas général (c'est à dire  $c, d \in \mathbb{R}$  et  $f \in C([0, 1], \mathbb{R})$ ). Montrer qu'il existe un unique vecteur  $U = (u_1, \dots, u_N)^t \in \mathbb{R}^N$  solution de (2.5.85). (On rappelle, encore une fois, que dans (2.5.85) les termes  $F_{i+\frac{1}{2}}$  et  $F_{i-\frac{1}{2}}$  sont donnés par (2.5.86) et (2.5.87).)

Soient  $\alpha, \beta > 0$ . On suppose, dans tout la suite de l'exercice, qu'il existe  $h > 0$  tel que  $\alpha h \leq h_i \leq \beta h$ , pour tout  $i \in \{1, \dots, N\}$ . On note  $\bar{u}_i = \frac{1}{h_i} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} u(x) dx$ , pour  $i = 1, \dots, N$ . (On rappelle que  $u$  est la solution exacte de (2.5.83)-(2.5.84).)

3. (Non consistance du schéma au sens des différences finies)

(a) Montrer que le système peut se mettre sous la forme  $AU = B$ , où  $B$  est définie par

$$B_1 = bf_1 + \frac{2c}{h_1^2} + \frac{ac}{h_1},$$

$$B_i = bf_i, \quad i = 2, \dots, N-1,$$

$$B_N = bf_n + \frac{2d}{h_N^2}.$$

(b) On pose  $\bar{R} = A\bar{U} - B$  avec  $\bar{U} = (\bar{u}_1, \dots, \bar{u}_N)^t$ . Vérifier que pour tout  $i \in \{1, \dots, N\}$ ,  $\bar{R}_i$  peut se mettre sous la forme :

$$\bar{R}_i = \bar{R}_i^1 + \bar{R}_i^2$$

où  $\sup_{i=1, \dots, N} |\bar{R}_i^1| \leq C_1$  et  $\sup_{i=1, \dots, N} |\bar{R}_i^2| \leq C_2 h$ .

(c) On se restreint dans cette question au cas où  $a = 0$ ,  $b > 0$ ,  $f = 0$ ,  $c = 1$ ,  $d = e^{\sqrt{b}}$ ,  $N = 2q$ ,  $h_i = h$  si  $i$  est pair et  $h_i = \frac{h}{2}$  si  $i$  est impair, avec  $h = \frac{2}{3N}$ .

Montrer que  $\|\bar{R}\|_\infty$  ne tend pas vers 0 avec  $h$ .

4. (Consistance des flux.) En choisissant convenablement  $(\bar{F}_{i+\frac{1}{2}})_{i \in \{0, \dots, N\}}$ , montrer que :

$$\bar{F}_{i+\frac{1}{2}} - \bar{F}_{i-\frac{1}{2}} + bh_i \bar{u}_i = bh_i f_i, \quad i \in \{1, \dots, N\}, \quad (2.5.88)$$

et que  $(\bar{F}_{i+\frac{1}{2}})_{i \in \{0, \dots, N\}}$  vérifie les égalités suivantes :

$$\bar{F}_{i+\frac{1}{2}} = -\frac{\bar{u}_{i+1} - \bar{u}_i}{h_{i+\frac{1}{2}}} + a\bar{u}_i + R_{i+\frac{1}{2}}, \quad i \in \{1, \dots, N-1\}, \quad (2.5.89)$$

$$\bar{F}_{\frac{1}{2}} = -\frac{\bar{u}_1 - c}{\frac{h_1}{2}} + ac + R_{\frac{1}{2}}, \quad \bar{F}_{N+\frac{1}{2}} = -\frac{d - \bar{u}_N}{\frac{h_N}{2}} + au_N + R_{N+\frac{1}{2}}, \quad (2.5.90)$$

avec,

$$|R_{i+\frac{1}{2}}| \leq C_1 h, \quad i \in \{0, \dots, N\}, \quad (2.5.91)$$

où  $C_1 \in \mathbb{R}$ , et  $C_1$  ne dépend que de  $\alpha, \beta$ , et  $u$ .

5. (Estimation d'erreur.) On pose  $e_i = \bar{u}_i - u_i$ , pour  $i \in \{1, \dots, N\}$  et  $E = (e_1, \dots, e_N)^t$ .

(a) Montrer que  $E$  est solution du système (de  $N$  équations) suivant :

$$G_{i+\frac{1}{2}} - G_{i-\frac{1}{2}} + bh_i e_i = 0, \quad i \in \{1, \dots, N\}, \quad (2.5.92)$$

avec  $(G_{i+\frac{1}{2}})_{i \in \{0, \dots, N\}}$  donné par les expressions suivantes :

$$G_{i+\frac{1}{2}} = -\frac{e_{i+1} - e_i}{h_{i+\frac{1}{2}}} + ae_i + R_{i+\frac{1}{2}}, \quad i \in \{1, \dots, N-1\}, \quad (2.5.93)$$

$$G_{\frac{1}{2}} = -\frac{e_1}{\frac{h_1}{2}} + R_{\frac{1}{2}}, \quad G_{N+\frac{1}{2}} = -\frac{-e_N}{\frac{h_N}{2}} + ae_N + R_{N+\frac{1}{2}}, \quad (2.5.94)$$

- (b) En multipliant (2.5.92) par  $e_i$  et en sommant sur  $i = 1, \dots, N$ , montrer qu'il existe  $C_2 \in \mathbb{R}$ , ne dépendant que de  $\alpha, \beta$ , et  $u$  tel que :

$$\sum_{i=0}^N (e_{i+1} - e_i)^2 \leq C_2 h^3, \quad (2.5.95)$$

avec  $e_0 = e_{N+1} = 0$ .

- (c) Montrer qu'il existe  $C_3 \in \mathbb{R}$ , ne dépendant que de  $\alpha, \beta$ , et  $u$  tel que :

$$|e_i| \leq C_3 h, \quad \text{pour tout } i \in \{1, \dots, N\}. \quad (2.5.96)$$

6. (Principe du maximum.) On suppose, dans cette question, que  $f(x) \leq d \leq c$ , pour tout  $x \in [0, 1]$ . Montrer que  $u_i \leq c$ , pour tout  $i \in \{1, \dots, N\}$ . (On peut aussi montrer que  $u(x) \leq c$ , pour tout  $x \in [0, 1]$ .)
7. On remplace, dans cette question, (2.5.86) et (2.5.87) par :

$$F_{i+\frac{1}{2}} = -\frac{u_{i+1} - u_i}{h_{i+\frac{1}{2}}} + au_{i+1}, \quad i \in \{1, \dots, N-1\}, \quad (2.5.97)$$

$$F_{\frac{1}{2}} = -\frac{u_1 - c}{\frac{h_1}{2}} + au_1, \quad F_{N+\frac{1}{2}} = -\frac{d - u_N}{\frac{h_N}{2}} + ad. \quad (2.5.98)$$

Analyser brièvement le nouveau schéma obtenu (existence de la solution approchée, consistance des flux, estimation d'erreur, principe du maximum).

### Exercice 12 (Convergence de la norme $H^1$ discrète)

Montrer que si  $u_{\mathcal{T}} : ]0, 1[ \rightarrow \mathbb{R}$  est définie par  $u_{\mathcal{T}}(x) = u_i \quad \forall x \in K_i$  où  $(u_i)_{i=1, \dots, N}$  solution de (2.3.29)–(2.3.31), alors  $|u_{\mathcal{T}}|_{1, \mathcal{T}}$  converge dans  $L^2(]0, 1[)$  lorsque  $h$  tend vers 0, vers  $\|Du\|_{L^2(]0, 1[)}$ , où  $u$  est la solution de (2.3.27).

### Exercice 13 (Discrétisation 2D par différences finies)

Ecrire le système linéaire obtenu lorsqu'on discrétise le problème

$$\begin{cases} -\Delta u = f \text{ dans } \Omega = ]0, 1[ \times ]0, 1[, \\ u = 0 \text{ sur } \partial\Omega. \end{cases} \quad (2.5.99)$$

par différences finies avec un pas uniforme  $h = 1/N$  dans les deux directions d'espace. Montrer l'existence et l'unicité de la solution du système linéaire obtenu.

Soit  $\Omega = ]0, 1[^2 \subset \mathbb{R}^2$ . On se propose d'étudier deux schémas numériques pour le problème suivant :

$$\begin{cases} -\Delta u(x, y) + k \frac{\partial u}{\partial x}(x, y) = f(x, y), & (x, y) \in \Omega, \\ u = 0, & \text{sur } \partial\Omega, \end{cases} \quad (2.5.100)$$

où  $k > 0$  est un réel donné et  $f \in C(\bar{\Omega})$  est donnée. On note  $u$  la solution exacte de (2.5.100) et on suppose que  $u \in C^4(\bar{\Omega})$ .

1. (Principe du maximum)

Montrer que pour tout  $\varphi \in C^1(\bar{\Omega})$  t.q.  $\varphi = 0$  sur  $\partial\Omega$ , on a :

$$\int_{\Omega} \nabla u(x) \cdot \nabla \varphi(x) \, dx + \int_{\Omega} k \frac{\partial u}{\partial x}(x) \varphi(x) \, dx = \int_{\Omega} f(x) \varphi(x) \, dx.$$

En déduire que si  $f \leq 0$  sur  $\bar{\Omega}$ , on a alors  $u \leq 0$  sur  $\bar{\Omega}$ .

Soit  $N \in \mathbb{N}$ , on pose  $h = \frac{1}{N+1}$ , et  $u_{i,j}$  est la valeur approchée recherchée de  $u(ih, jh)$ ,  $(i, j) \in \{0, \dots, N+1\}^2$ . On pose  $f_{i,j} = f(ih, jh)$ , pour tout  $(i, j) \in \{1, \dots, N\}^2$ . On s'intéresse à deux schémas de la forme :

$$\begin{cases} a_0 u_{i,j} - a_1 u_{i-1,j} - a_2 u_{i+1,j} - a_3 u_{i,j-1} - a_4 u_{i,j+1} = f_{i,j}, & \forall (i, j) \in \{1, \dots, N\}^2, \\ u_{i,j} = 0, & (i, j) \in \gamma, \end{cases} \quad (2.5.101)$$

où  $a_0, a_1, a_2, a_3, a_4$  sont données (ce sont des fonctions données de  $h$ ) et  $\gamma = \{(i, j), (ih, jh) \in \partial\Omega\}$  ( $\gamma$  dépend aussi de  $h$ ). Le premier schéma, schéma [I], correspond au choix suivant des  $a_i$  :

$$a_0 = \frac{4}{h^2}, \quad a_1 = \frac{1}{h^2} + \frac{k}{2h}, \quad a_2 = \frac{1}{h^2} - \frac{k}{2h}, \quad a_3 = a_4 = \frac{1}{h^2}.$$

Le deuxième schéma, schéma [II], correspond au choix suivant des  $a_i$  :

$$a_0 = \frac{4}{h^2} + \frac{k}{h}, \quad a_1 = \frac{1}{h^2} + \frac{k}{h}, \quad a_2 = a_3 = a_4 = \frac{1}{h^2}.$$

2. (Consistance)

Donner une majoration de l'erreur de consistance en fonction de  $k$ ,  $h$  et des dérivées de  $u$ , pour les schémas [I] et [II]. Donner l'ordre des schémas [I] et [II].

3. (Principe du maximum discret)

Dans le cas du schéma [II] montrer que si  $(w_{i,j})$  vérifie :

$$a_0 w_{i,j} - a_1 w_{i-1,j} - a_2 w_{i+1,j} - a_3 w_{i,j-1} - a_4 w_{i,j+1} \leq 0, \quad \forall (i, j) \in \{1, \dots, N\}^2,$$

on a alors

$$w_{i,j} \leq \max_{(n,m) \in \gamma} (w_{n,m}), \quad \forall (i, j) \in \{1, \dots, N\}^2.$$

Montrer que ceci est aussi vrai dans le cas du schéma [I] si  $h$  vérifie une condition à déterminer.

4. (Stabilité)

Montrer que le schéma [II] et le schéma [I] sous la condition trouvée en 3. sont stables (au sens  $\|U\|_\infty \leq C\|f\|_\infty$ , avec une constante  $C$  à déterminer explicitement, où  $U = \{u_{i,j}\}_{(i,j) \in \{0, \dots, N+1\}^2}$  est solution de (2.5.101). [On pourra utiliser la fonction  $\phi(x, y) = \frac{1}{2}y^2$ ].

En déduire que dans le cas du schéma [II] et du schéma [I] sous la condition trouvée en 3. le problème (2.5.101) admet, pour tout  $f$ , une et une seule solution.

5. (Convergence)

Les schémas [I] et [II] sont-ils convergents ? (au sens  $\max_{(i,j) \in \{0, \dots, N+1\}^2} (|u_{i,j} - u(ih, jh)|) \rightarrow 0$  quand  $h \rightarrow 0$ ). Quel est l'ordre de convergence de chacun des schémas ?

6. (Commentaires)

Quels sont, à votre avis, les avantages respectifs des schémas [I] et [II] ?

**Exercice 15 (Elimination des inconnues d'arêtes.)** *Suggestions en page 64, corrigé en page 84*

On se place ici dans le cadre des hypothèses et notations du paragraphe 2.3.4 page 27

1. Pour chaque arête interne  $\sigma = K|L$ , calculer la valeur  $u_\sigma$  en fonction de  $u_K$  et  $u_L$  et en déduire que les flux numériques  $F_{K,\sigma}$  et  $F_{L,\sigma}$  vérifient bien (2.3.49)
2. Pour chaque arête  $\sigma \subset \Gamma_1 \cup \Gamma_3$ , telle que  $\sigma \in \mathcal{E}_K$ , calculer  $u_\sigma$  en fonction de  $u_K$  et montrer que  $F_{K,\sigma}$  vérifie bien (2.3.50)
3. Pour chaque arête  $\sigma \in \mathcal{E}_I$ , avec  $\sigma = K|L \subset \Omega_1$ , calculer la valeur  $u_\sigma$  en fonction de  $u_K$  et  $u_L$  et en déduire que les flux numériques  $F_{K,\sigma}$  et  $F_{L,\sigma}$  vérifient bien (2.3.52)
4. Ecrire le système linéaire que satisfont les inconnues  $(u_K)_{K \in \mathcal{T}}$ .

**Exercice 16 (Implantation de la méthode des volumes finis pour un problème électrique.)**

On considère le problème de conduction du courant électrique

$$-\operatorname{div}(\mu_i \nabla \phi(x)) = 0 \quad x \in \Omega_i, \quad i = 1, 2 \quad (2.5.102)$$

où  $\phi$  représente le potentiel électrique,  $j = -\mu \nabla \phi(x)$  est donc le courant électrique,  $\mu_1 > 0$ ,  $\mu_2 > 0$  sont les conductivités thermiques dans les domaines  $\Omega_1$  et avec  $\Omega_2$ , avec  $\Omega_1 = ]0, 1[ \times ]0, 1[$  et  $\Omega_2 = ]0, 1[ \times ]1, 2[$ . On appelle  $\Gamma_1 = ]0, 1[ \times \{0\}$ ,  $\Gamma_2 = \{1\} \times ]0, 2[$ ,  $\Gamma_3 = ]0, 1[ \times \{2\}$ , et  $\Gamma_4 = \{0\} \times ]0, 2[$  les frontières extérieures de  $\Omega$ , et on note  $I = ]0, 1[ \times \{0\}$  l'interface entre  $\Omega_1$  et  $\Omega_2$  (voir Figure 2.3). Dans la suite, on notera  $\mu$  la conductivité électrique sur  $\Omega$ , avec  $\mu|_{\Omega_i} = \mu_i$ ,  $i = 1, 2$ .

On suppose que les frontières  $\Gamma_2$  et  $\Gamma_4$  sont parfaitement isolées. Le potentiel électrique étant défini à une constante près, on impose que sa moyenne soit nulle sur le domaine, pour que le problème soit bien posé. La conservation du courant électrique impose que

$$\int_{\Gamma_1} j \cdot \mathbf{n} + \int_{\Gamma_3} j \cdot \mathbf{n} = 0,$$

où  $\mathbf{n}$  désigne le vecteur unitaire normal à la frontière  $\partial\Omega$  et extérieure à  $\Omega$ .

Enfin, on suppose que l'interface  $I$  est le siège d'une réaction électrochimique qui induit un saut de potentiel. On a donc pour tout point de l'interface  $I$  :

$$\phi_2(x) - \phi_1(x) = \psi(x), \quad \forall x \in I,$$

où  $\phi_i$  désigne la restriction de  $\phi$  au sous domaine  $i$ . La fonction  $\phi$  est donc discontinue sur l'interface  $I$ . Notons que, par contre, le courant électrique est conservé et on a donc

$$(-\mu \nabla \phi \cdot \mathbf{n})|_2(x) + (-\mu \nabla \phi \cdot \mathbf{n})|_1(x) = 0, \forall x \in I.$$

1. Ecrire le problème complet, avec conditions aux limites.
2. Discréteriser le problème par la méthode des volumes finis, avec un maillage rectangulaire uniforme, (considérer deux inconnues discrètes pour chaque arête de l'interface) et écrire le système linéaire obtenu sur les inconnues discrètes.

**Exercice 17 (Existence de solutions “presque classiques”)** *Corrigé en page 90*

Soit  $u_0 \in L^2(\Omega)$ . On s'intéresse au problème :

$$\begin{aligned} u_t(x, t) - u_{xx}(x, t) &= 0, x \in ]0, 1[, t \in \mathbb{R}_+^*, \\ u(0, t) &= u(1, t) = 0, t \in \mathbb{R}_+^*, \\ u(x, 0) &= u_0(x), x \in ]0, 1[. \end{aligned} \tag{2.5.103}$$

1. On définit  $u : [0, 1] \times \mathbb{R}_+^* \rightarrow \mathbb{R}$  par :

$$u(x, t) = \sum_{n \in \mathbb{N}^*} e^{-n^2 \pi^2 t} a_n \sin(n\pi x), x \in [0, 1], t \in \mathbb{R}_+^*, \tag{2.5.104}$$

$$\text{avec } a_n = (\int_0^1 u_0(x) \sin(n\pi x) dx) / (\int_0^1 \sin^2(n\pi x) dx).$$

Montrer que  $u$  est bien définie de  $[0, 1] \times \mathbb{R}_+^*$  dans  $\mathbb{R}$  et est solution de (2.5.103) au sens suivant :

$$\begin{aligned} u &\in C^\infty([0, 1] \times \mathbb{R}_+^*, \mathbb{R}), \\ u_t(x, t) - u_{xx}(x, t) &= 0, \forall x \in [0, 1], \forall t \in \mathbb{R}_+^*, \\ u(0, t) &= u(1, t) = 0, \forall t \in \mathbb{R}_+^*, \\ \|u(., t) - u_0\|_{L^2([0, 1])} &\rightarrow 0, \text{ quand } t \rightarrow 0. \end{aligned} \tag{2.5.105}$$

2. Montrer qu'il existe une unique fonction  $u$  solution de (2.5.105).

**Exercice 18 (Exemple de schéma non convergent)** *Suggestions en page 65, corrigé en page 92*

Soit  $u_0 \in L^2([-4, 4])$ . On note  $u$  l'unique solution (au sens vu en cours ou en un sens inspiré de l'exercice précédent) du problème suivant :

$$\begin{aligned} u_t(x, t) - u_{xx}(x, t) &= 0, x \in ]-4, 4[, t \in ]0, 1[, \\ u(-4, t) &= u(4, t) = 0, t \in ]0, 1[, \\ u(x, 0) &= u_0(x), x \in ]-4, 4[. \end{aligned} \tag{2.5.106}$$

On sait que la solution de (2.5.106) est de classe  $C^\infty$  sur  $[-4, 4] \times ]0, 1]$  (voir l'exercice précédent). On admettra que si  $u_0 \geq 0$  p.p. sur  $[-4, 4]$  et  $u_0 \neq 0$  (dans  $L^2([-4, 4])$ ) alors  $u(x, t) > 0$  pour tout  $x \in ]-4, 4[$  et tout  $t \in ]0, 1]$ .

On suppose maintenant que  $u_0 \in C([-4, 4], \mathbb{R})$ ,  $u_0(-4) = u_0(4) = 0$ ,  $u_0 \geq 0$  sur  $[-4, 4]$ ,  $u_0$  nulle sur  $[-3, 4]$  et qu'il existe  $a \in ]-4, -3[$  t.q.  $u_0(a) > 0$ . On a donc  $u(x, t) > 0$  pour tout  $x \in ]-4, 4[$ .

Avec les notations du cours, on considère la solution de (2.5.106) donnée par le schéma d'Euler explicite (2.4.54) avec le pas de temps  $k = 1/(M+1)$  et le pas d'espace  $h = 8/(N+1)$  ( $M, N \in \mathbb{N}^*$ ,  $N$  impair). La solution approchée est définie par les valeurs  $u_i^n$  pour  $i \in \{-(N+1)/2, \dots, (N+1)/2\}$  et  $n \in \{0, \dots, M+1\}$ . La valeur  $u_i^n$  est censée être une valeur approchée de  $\bar{u}_i^n = u(ih, nk)$ .

1. Donner les équations permettant de calculer  $u_i^n$  pour  $i \in \{-(N+1)/2, \dots, (N+1)/2\}$  et  $n \in \{0, \dots, M+1\}$ .
2. On suppose maintenant que  $k = h$ . Montrer que  $u_i^n = 0$  pour  $i \geq 0$  et  $n \in \{0, \dots, M+1\}$ . En déduire que  $\max\{|u_i^{M+1} - \bar{u}_i^{M+1}|, i \in \{-(N+1)/2, \dots, (N+1)/2\}$  ne tends pas vers 0 quand  $h \rightarrow 0$  (c'est-à-dire quand  $N \rightarrow \infty$ ).

**Exercice 19 (Discrétisation d'un problème parabolique.)** *Suggestions en page 65, corrigé en page 93*

Dans cet exercice on s'intéresse à des schémas numériques pour le problème :

$$\begin{cases} u_t + u_x - \varepsilon u_{xx} = 0 & (x, t) \in \mathbb{R}^+ \times ]0, T[ \\ u(1, t) = u(0, t) = 0 & t \in ]0, T[ \\ u(x, 0) = u_0(x) & x \in ]0, 1[ \end{cases} \quad (2.5.107)$$

où  $u_0$  et  $\varepsilon$  sont donnés ( $\varepsilon > 0$ ). On reprend dans la suite de l'exercice les notations du cours.

1. Donner un schéma d'approximation de (2.5.107) différences finies à pas constant en espace et Euler explicite à pas constant en temps. Montrer que l'erreur de consistance est majorée par  $C(k + h^2)$  ; avec  $C$  dépendant de la solution exacte de (2.5.107). Sous quelle(s) condition(s) sur  $k$  et  $h$  a-t-on  $\|u^n\|_\infty \leq \|u^0\|_\infty, \forall n \leq M$  ?

Donner un résultat de convergence pour ce schéma.

2. Même question que 1. en remplaçant Euler explicite par Euler implicite.

3. En s'inspirant du schéma de Crank-Nicolson (vu en cours) construire un schéma d'ordre 2 (espace et temps). Sous quelle(s) condition(s) sur  $k$  et  $h$  a-t-on  $\|u^n\|_2 \leq \|u^0\|_2, \forall n \leq M$  ? Donner un résultat de convergence pour ce schéma.

4. Dans les schémas trouvés aux questions 1., 2. et 3. on remplace l'approximation de  $u_x$  par une approximation décentrée à gauche. Quel est l'ordre des schémas obtenus et sous quelle(s) condition(s) sur  $k$  et  $h$  a-t-on  $\|u^n\|_\infty \leq \|u^0\|_\infty$  où  $\|u\|u^0\|_2, \forall n \leq M$  ? Donner un résultat de convergence pour ces schémas.

**Exercice 20 (Equation parabolique avec terme source)** *Suggestions en page 65, corrigé 95*

Soit  $u_0$  une fonction donnée de  $[0, 1]$  dans  $\mathbb{R}$ . On s'intéresse ici à la discrétisation du problème suivant :

$$u_t(t, x) - u_{xx}(t, x) - u(t, x) = 0, \quad t \in \mathbb{R}^+, \quad x \in [0, 1], \quad (2.5.108)$$

$$u(t, 0) = u(t, 1) = 0, \quad t \in \mathbb{R}_+^*; \quad u(0, x) = u_0(x), \quad x \in [0, 1]. \quad (2.5.109)$$

On note  $u$  la solution de (2.5.108), (2.5.109), et on suppose que  $u$  est la restriction à  $\mathbb{R}_+ \times [0, 1]$  d'une fonction de classe  $C^\infty$  de  $\mathbb{R}^2$  dans  $\mathbb{R}$ .

Pour  $h = \frac{1}{N+1}$  ( $N \in \mathbb{N}^*$ ) et  $k > 0$ , on pose  $x_i = ih, i \in \{0, \dots, N+1\}$ ,  $t_n = nk, n \in \mathbb{N}$ ,  $\bar{u}_i^n = u(x_i, t_n)$ , et on note  $u_i^n$  la valeur approchée recherchée de  $\bar{u}_i^n$ .

On considère les schémas numériques (2.5.110), (2.5.112) et (2.5.111), (2.5.112) définis par les équations suivantes :

$$\frac{u_i^{n+1} - u_i^n}{k} - \frac{(u_{i+1}^{n+1} + u_{i-1}^{n+1} - 2u_i^{n+1})}{h^2} - u_i^{n+1} = 0, \quad n \in \mathbb{N}, \quad i \in \{1, \dots, N\}, \quad (2.5.110)$$

$$\frac{u_i^{n+1} - u_i^n}{k} - \frac{(u_{i+1}^{n+1} + u_{i-1}^{n+1} - 2u_i^{n+1})}{h^2} - u_i^n = 0, \quad n \in \mathbb{N}, \quad i \in \{1, \dots, N\}, \quad (2.5.111)$$

$$u_0^{n+1} = u_{N+1}^{n+1} = 0, \quad n \in \mathbb{N}; \quad u_i^0 = u_0(x_i), \quad i \in \{0, \dots, N+1\}. \quad (2.5.112)$$

Pour  $n \in N$ , on note  $u^n = (u_1^n, \dots, u_N^n)^t \in \mathbb{R}^N$ .

- (Consistance) Soit  $T > 0$ . Pour  $n \in \mathbb{N}$ , et  $i \in \{1, \dots, N\}$ , on note  $R_i^n$  l'erreur de consistance (définie en cours) du schéma numérique (2.5.110), (2.5.112) [resp. du schéma numérique (2.5.111), (2.5.112)]. Montrer qu'il existe  $C \in \mathbb{R}$ , ne dépendant que de  $u$  et  $T$ , t. q.  $|R_i^n| \leq C(k + h^2)$ , pour tout  $i \in \{1, \dots, N\}$  et tout  $n \in \mathbb{N}$ , t.q.  $kn \leq T$ .
- Montrer que le schéma (2.5.110), (2.5.112) [resp. (2.5.111), (2.5.112)] demande, à chaque pas de temps, la résolution du système linéaire  $Au^{n+1} = a$  [resp.  $Bu^{n+1} = b$ ] avec  $A, B \in \mathbb{R}^{N,N}$  et  $a, b \in \mathbb{R}^N$  à déterminer.  
Montrer que  $B$  est inversible (et même s.d.p.) pour tout  $h > 0$  et  $k > 0$ . Montrer que  $A$  est inversible (et même s.d.p.) pour tout  $h > 0$  et  $k \in ]0, 1[$ .
- (Stabilité) Pour  $n \in \mathbb{N}$ , on pose  $\|u^n\|_\infty = \sup_{i \in \{1, \dots, N\}} |u_i^n|$ . Soit  $T > 0$ . On considère le schéma (2.5.111), (2.5.112). Montrer qu'il existe  $C_1(T) \in \mathbb{R}$ , ne dépendant que de  $T$ , t.q.  $\|u^n\|_\infty \leq C_1(T)\|u_0\|_\infty$ , pour tout  $h > 0$ ,  $k > 0$ , et  $n \in \mathbb{N}$  tel que  $kn \leq T$ .  
Soit  $\alpha \in [0, 1]$ . On considère le schéma (2.5.110), (2.5.112). Montrer qu'il existe  $C_2(T, \alpha) \in \mathbb{R}$ , ne dépendant que de  $T$  et de  $\alpha$ , t.q.  $\|u^n\|_\infty \leq C_2(T, \alpha)\|u_0\|_\infty$ , pour tout  $h > 0$ ,  $k \in ]0, \alpha[$ , et  $n \in \mathbb{N}$  tel que  $kn \leq T$ .
- (Estimation d'erreur) Pour  $n \in \mathbb{N}$  et  $i \in \{1, \dots, N\}$ , on pose  $e_i^n = \bar{u}_i^n - u_i^n$ . Soit  $T > 0$ . Donner, pour  $kn \leq T$ , des majorations de  $\|e^n\|_\infty$  en fonction de  $T$ ,  $C$ ,  $C_1(T)$ ,  $C_2(T, \alpha)$  (définis dans les questions précédentes),  $k$  et  $h$  pour les deux schémas étudiés.

**Exercice 21 (Schéma “saute-mouton”)** *Corrigé en page 98*

On considère le problème suivant :

$$\begin{cases} u_t(x, t) - u_{xx}(x, t) = 0, & x \in ]0, 1[, t \in ]0, T[, \\ u(0, t) = u(1, t) = 0, & t \in ]0, T[, \\ u(x, 0) = u_0(x), & x \in ]0, 1[. \end{cases} \quad (2.5.113)$$

Pour trouver une solution approchée de ((2.5.113)), on considère le schéma “saute-mouton” :

$$\begin{cases} \frac{u_j^{n+1} - u_j^{(n-1)}}{2k} = \frac{u_{j-1}^n - 2u_j^n + u_{j+1}^n}{h^2}, & j = 1, \dots, N-1, n = 1, \dots, M-1, \\ u_0^{n+1} = u_{N+1}^{n+1} = 0, & n = 1, \dots, M-1, \end{cases} \quad (2.5.114)$$

où  $(u_j^0)_{j=1, \dots, N}$  et  $(u_j^1)_{j=1, \dots, N}$  sont supposés connus,  $h = 1/N$ ,  $k = T/M$ .

- Montrer que le schéma (2.5.114) est consistant. Quel est son ordre?
- Montrer que le schéma (2.5.114) est inconditionnellement instable au sens de Von Neumann.

On modifie “légèrement” le schéma (2.5.114) en prenant

$$\begin{cases} \frac{u_j^{n+1} - u_j^{(n-1)}}{2k} = \frac{u_{j-1}^n - (u_j^{n+1} + u_j^{(n-1)}) + u_{j+1}^n}{h^2}, & j = 1, \dots, N, n = 1, \dots, M-1, \\ u_0^{n+1} = u_{N+1}^{n+1} = 0, & n = 1, \dots, M-1, \end{cases} \quad (2.5.115)$$

(schéma de Dufort-Frankel).

3. Montrer que le schéma (2.5.115) est consistant avec (2.5.113) quand  $h, k \rightarrow 0$  sous la condition  $\frac{k}{h} \rightarrow 0$ .

4. Montrer que (2.5.115) est inconditionnellement stable.

**Exercice 22 (Schémas centré et décentré)** *Corrigé en page 100*

Soient  $\alpha > 0$ ,  $\mu > 0$ ,  $T > 0$  et  $u_0 : \mathbb{R} \rightarrow \mathbb{R}$ . On s'intéresse au problème suivant :

$$\begin{aligned} u_t(x, t) + \alpha u_x(x, t) - \mu u_{xx}(x, t) &= 0, \quad x \in ]0, 1[, \quad t \in ]0, T[, \\ u(0, t) &= u(1, t) = 0, \quad t \in ]0, T[, \\ u(x, 0) &= u_0(x), \quad x \in ]0, 1[. \end{aligned} \quad (2.5.116)$$

On rappelle que  $u_t = \frac{\partial u}{\partial t}$ ,  $u_x = \frac{\partial u}{\partial x}$  et  $u_{xx} = \frac{\partial^2 u}{\partial x^2}$ . On suppose qu'il existe  $u \in C^4([0, 1] \times [0, T])$  solution (classique) de (2.5.116) (noter que ceci implique  $u_0(0) = u_0(1) = 0$ ). On pose  $A = \min\{u_0(x), x \in [0, 1]\}$  et  $B = \max\{u_0(x), x \in [0, 1]\}$  (noter que  $A \leq 0 \leq B$ ).

On discrétise le problème (2.5.116). On reprend les notations du cours. Soient  $h = 1/(N+1)$  et  $k = T/M$  ( $N, M \in \mathbb{N}^*$ ).

1. Schéma explicite décentré. Pour approcher la solution  $u$  de (2.5.116), on considère le schéma suivant :

$$\begin{aligned} \frac{1}{k}(u_i^{n+1} - u_i^n) + \frac{\alpha}{h}(u_i^n - u_{i-1}^n) - \frac{\mu}{h^2}(u_{i+1}^n - 2u_i^n + u_{i-1}^n) &= 0, \\ i \in \{1, \dots, N\}, \quad n \in \{0, \dots, M-1\}, \\ u_0^n &= u_{N+1}^n = 0, \quad n \in \{1, \dots, M\}, \\ u_i^0 &= u_0(ih), \quad i \in \{0, \dots, N+1\}. \end{aligned} \quad (2.5.117)$$

On pose  $\bar{u}_i^n = u(ih, nk)$  pour  $i \in \{0, \dots, N+1\}$  et  $n \in \{0, \dots, M\}$ .

(a) (Consistance) Montrer que l'erreur de consistance du schéma (2.5.117) est majorée par  $C_1(k + h)$ , où  $C_1$  ne dépend que de  $u$ ,  $T$ ,  $\alpha$  et  $\mu$ .

(b) (Stabilité) Sous quelle condition sur  $k$  et  $h$  (cette condition peut dépendre de  $\alpha$  et  $\mu$ ) a-t-on  $A \leq u_i^n \leq B$  pour tout  $i \in \{0, \dots, N+1\}$  et tout  $n \in \{0, \dots, M\}$ ? Sous cette condition, en déduire  $\|u^n\|_\infty \leq \|u_0\|_{L^\infty([0, 1])}$  pour tout  $n \in \{0, \dots, M\}$  (avec  $\|u^n\|_\infty = \max\{|u_i^n|, i \in \{0, \dots, N+1\}\}$ ).

(c) (Estimation d'erreur) On pose  $e_i^n = \bar{u}_i^n - u_i^n$ .

Sous la condition sur  $k$  et  $h$  trouvée précédemment, montrer que  $|e_i^n| \leq C_2(k + h)$  pour tout  $i \in \{0, \dots, N+1\}$  et tout  $n \in \{0, \dots, M\}$  avec  $C_2$  ne dépendant que de  $u$ ,  $T$ ,  $\alpha$  et  $\mu$ .

2. Schéma explicite centré.

On change dans le schéma (2.5.117) la quantité  $(\alpha/h)(u_i^n - u_{i-1}^n)$  par  $(\alpha/2h)(u_{i+1}^n - u_{i-1}^n)$ .

(a) (Consistance) Montrer que l'erreur de consistance est maintenant majorée par  $C_3(k + h^2)$ , où  $C_3$  ne dépend que de  $u$ ,  $T$ ,  $\alpha$  et  $\mu$ .

(b) Reprendre les questions de stabilité et d'estimation d'erreur du schéma (2.5.117).

**Exercice 23 (Problème parabolique non linéaire)** *Corrigé en page 103*

On se propose, dans cet exercice, de montrer l'existence d'une solution faible au problème (2.5.118)-(2.5.120), à partir de l'existence de la solution approchée donnée par un schéma numérique. L'inconnue de ce problème est la fonction  $u$  de  $[0, 1] \times [0, T]$  dans  $\mathbb{R}$ , elle doit être solution des équations suivantes :

$$\frac{\partial u}{\partial t}(x, t) - \frac{\partial^2 \varphi(u)}{\partial x^2}(x, t) = v(x, t), \quad x \in ]0, 1[, \quad t \in ]0, T[, \quad (2.5.118)$$

$$\frac{\partial \varphi(u)}{\partial x}(0, t) = \frac{\partial \varphi(u)}{\partial x}(1, t) = 0, \quad t \in ]0, T[, \quad (2.5.119)$$

$$u(x, 0) = u_0(x), \quad x \in ]0, 1[, \quad (2.5.120)$$

où  $\varphi, v, T, u_0$  sont donnés et sont t.q.

1.  $T > 0, v \in L^\infty(]0, 1[ \times ]0, T[)$ ,
2.  $\varphi$  croissante, lipschitzienne de  $\mathbb{R}$  dans  $\mathbb{R}$ ,
3.  $u_0 \in L^\infty(]0, 1[)$  et  $\varphi(u_0)$  lipschitzienne de  $[0, 1]$  dans  $\mathbb{R}$ .

Un exemple important est donné par  $\varphi(s) = \alpha_1 s$  si  $s \leq 0$ ,  $\varphi(s) = 0$  si  $0 \leq s \leq L$  et  $\varphi(s) = \alpha_2(s - L)$  si  $s \geq L$ , avec  $\alpha_1, \alpha_2$  et  $L$  donnés dans  $\mathbb{R}_+^*$ . Noter pour cet exemple que  $\varphi' = 0$  sur  $]0, L[$ .

Les ensembles  $]0, 1[$  et  $D = ]0, 1[ \times ]0, T[$  sont munis de leur tribu borélienne et de la mesure de Lebesgue sur cette tribu.

On appelle "solution faible" de (2.5.118)-(2.5.120) une solution de :

$$u \in L^\infty(]0, 1[ \times ]0, T[), \quad (2.5.121)$$

$$\int_D (u(x, t) \frac{\partial \psi}{\partial t}(x, t) + \varphi(u(x, t)) \frac{\partial^2 \psi}{\partial x^2}(x, t) + v(x, t) \psi(x, t)) dx dt + \int_{]0, 1[} u_0(x) \psi(x, 0) dx = 0, \quad (2.5.122)$$

$$\forall \psi \in C_T^{2,1}(\mathbb{R}^2),$$

où  $\psi \in C_T^{2,1}(\mathbb{R}^2)$  signifie que  $\psi$  est une fonction de  $\mathbb{R}^2$  dans  $\mathbb{R}$  deux fois continûment dérivable par rapport à  $x$ , une fois continûment dérivable par rapport à  $t$  et t.q.  $\frac{\partial \psi}{\partial x}(0, t) = \frac{\partial \psi}{\partial x}(1, t) = 0$ , pour tout  $t \in [0, T]$  et  $\psi(x, T) = 0$  pour tout  $x \in [0, 1]$ .

**Question 1** (Solution classique versus solution faible)

On suppose, dans cette question seulement, que  $\varphi$  est de classe  $C^2$ ,  $v$  est continue sur  $[0, 1] \times [0, T]$  et  $u_0$  est continue sur  $[0, 1]$ . Soit  $u \in C^2(\mathbb{R}^2, \mathbb{R})$ . On note encore  $u$  la restriction de  $u$  à  $]0, 1[ \times ]0, T[$ . Montrer que  $u$  est solution de (2.5.121)-(2.5.122) si et seulement si  $u$  vérifie (2.5.118)-(2.5.120) au sens classique (c'est-à-dire pour tout  $(x, t) \in [0, 1] \times [0, T]$ ).

On cherche maintenant une solution approchée de (2.5.118)-(2.5.120).

Soient  $N, M \in \mathbb{N}^*$ . On pose  $h = \frac{1}{N}$  et  $k = \frac{T}{M}$ . On va construire une solution approchée de (2.5.118)-(2.5.120) à partir de la famille  $\{u_i^n, i = 1, \dots, N, n = 0, \dots, M\}$  (dont on va prouver l'existence et l'unicité) vérifiant les équations suivantes :

$$u_i^0 = \frac{1}{h} \int_{(i-1)h}^{ih} u_0(x) dx, \quad i = 1, \dots, N, \quad (2.5.123)$$

$$\frac{u_i^{n+1} - u_i^n}{k} - \frac{\varphi(u_{i-1}^{n+1}) - 2\varphi(u_i^{n+1}) + \varphi(u_{i+1}^{n+1})}{h^2} = v_i^n, \quad i = 1, \dots, N, \quad n = 0, \dots, M-1, \quad (2.5.124)$$

avec  $u_0^{n+1} = u_1^{n+1}$ ,  $u_{N+1}^{n+1} = u_N^{n+1}$ , pour tout  $n = 0, \dots, M-1$  et  $v_i^n = \frac{1}{kh} \int_{nk}^{(n+1)k} \int_{(i-1)h}^{ih} v(x, t) dx dt$ , pour tout  $i = 1, \dots, N$ , pour tout  $n = 0, \dots, M$ .

**Question 2** (Existence et unicité de la solution approchée)

Soit  $n \in \{0, \dots, M-1\}$ . On suppose connu  $\{u_i^n, i = 1, \dots, N\}$ . On va prouver dans cette question l'existence et l'unicité de  $\{u_i^{n+1}, i = 1, \dots, N\}$  vérifiant (2.5.124) (avec  $u_0^{n+1} = u_1^{n+1}$ ,  $u_{N+1}^{n+1} = u_N^{n+1}$ ).

1. Soit  $a > 0$ , Pour  $s \in \mathbb{R}$ , on pose  $g_a(s) = s + a\varphi(s)$ . Montrer que  $g_a$  est une application strictement croissante bijective de  $\mathbb{R}$  dans  $\mathbb{R}$ .
2. Soit  $\bar{w} = (\bar{w}_i)_{i=1, \dots, N} \in \mathbb{R}^N$ . On pose  $\bar{w}_0 = \bar{w}_1$  et  $\bar{w}_{N+1} = \bar{w}_N$ . Montrer qu'il existe un et un seul couple  $(u, w) \in \mathbb{R}^N \times \mathbb{R}^N$ ,  $u = (u_i)_{i=1, \dots, N}$ ,  $w = (w_i)_{i=1, \dots, N}$ , t.q. :

$$\varphi(u_i) = w_i, \quad \text{pour tout } i \in \{1, \dots, N\}, \quad (2.5.125)$$

$$u_i + \frac{2k}{h^2} w_i = \frac{k}{h^2} (\bar{w}_{i-1} + \bar{w}_{i+1}) + u_i^n + k v_i^n, \quad \text{pour tout } i = 1, \dots, N. \quad (2.5.126)$$

On peut donc définir une application  $F$  de  $\mathbb{R}^N$  dans  $\mathbb{R}^N$  par  $\bar{w} \mapsto F(\bar{w}) = w$  où  $w$  est solution de (2.5.125)–(2.5.126).

3. On munit  $\mathbb{R}^N$  de la norme usuelle  $\|\cdot\|_\infty$ . Montrer que l'application  $F$  est strictement contractante. [On pourra utiliser la monotonie de  $\varphi$  et remarquer que, si  $a = \varphi(\alpha)$  et  $b = \varphi(\beta)$ , on a  $|\alpha - \beta| \geq (1/L)|a - b|$ , où  $L$  ne dépend que de  $\varphi$ .]
4. Soit  $\{u_i^{n+1}, i = 1, \dots, N\}$  solution de (2.5.124). On pose  $w = (w_i)_{i=1, \dots, N}$ , avec  $w_i = \varphi(u_i^{n+1})$  pour  $i \in \{1, \dots, N\}$ . Montrer que  $w = F(w)$ .
5. Soit  $w = (w_i)_{i=1, \dots, N}$  t.q.  $w = F(w)$ . Montrer que pour tout  $i \in \{1, \dots, N\}$  il existe  $u_i^{n+1} \in \mathbb{R}$  t.q.  $w_i = \varphi(u_i^{n+1})$ . Montrer que  $\{u_i^{n+1}, i = 1, \dots, N\}$  est solution de (2.5.124).
6. Montrer qu'il existe une unique famille  $\{u_i^{n+1}, i = 1, \dots, N\}$  solution de (2.5.124).

**Question 3** (Estimation  $L^\infty([0, 1] \times [0, T])$  sur  $u$ )

On pose  $A = \|u_0\|_{L^\infty([0, 1])}$  et  $B = \|v\|_{L^\infty([0, 1] \times [0, T])}$ . Montrer, par récurrence sur  $n$ , que  $u_i^n \in [-A - nkB, A + nkB]$  pour tout  $i = 1, \dots, N$  et tout  $n = 0, \dots, M$ . [On pourra, par exemple, considérer (2.5.124) avec  $i$  t.q.  $u_i^{n+1} = \min\{u_j^{n+1}, j = 1, \dots, N\}$ .]

En déduire qu'il existe  $c_{u_0, v, T} \in \mathbb{R}_+$  t.q.  $\|u^n\|_{L^\infty([0, 1])} \leq c_{u_0, v, T}$ .

**Question 4** (Estimation de la dérivée p.r. à  $x$  de  $\varphi(u)$ )

Montrer qu'il existe  $C_1$  (ne dépendant que de  $T$ ,  $\varphi$ ,  $v$  et  $u_0$ ) t.q., pour tout  $n = 0, \dots, M-1$ ,

$$\sum_{n=0}^{M-1} \sum_{i=1}^{N-1} (\varphi(u_{i+1}^{n+1}) - \varphi(u_i^{n+1}))^2 \leq C_1 \frac{h}{k}. \quad (2.5.127)$$

[Multiplier (2.5.124) par  $u_i^{n+1}$  et sommer sur  $i$  et sur  $n$  et utiliser l'inégalité  $a^2 - ab \geq \frac{a^2}{2} - \frac{b^2}{2}$ .]

**Question 5** (Estimation de la dérivée p.r. à  $t$  de  $\varphi(u)$ ) A

Montrer qu'il existe  $C_2$  (ne dépendant que de  $T$ ,  $\varphi$ ,  $v$  et  $u_0$ ) t.q.

$$\sum_{n=0}^{M-1} h \sum_{i=0}^{N+1} (\varphi(u_i^{n+1}) - \varphi(u_i^n))^2 \leq C_2 k. \quad (2.5.128)$$

et

$$\sum_{i=0}^{N+1} (\varphi(u_i^{n+1}) - \varphi(u_i^n))^2 \leq C_2 h, \text{ pour tout } n \in \{0, \dots, M\}. \quad (2.5.129)$$

[indication : multiplier (2.5.124) par  $\varphi(u_i^{n+1}) - \varphi(u_i^n)$  et sommer sur  $i$  et  $n$ ]

Dans la suite de l'exercice, il s'agit de passer à la limite (quand  $N, M \rightarrow \infty$ ) pour trouver une solution de (2.5.118)-(2.5.120).

Pour  $M \in \mathbb{N}^*$  donné, on prend  $N = M^2$  (et donc  $h$  et  $k$  sont donnés et  $k = T\sqrt{h}$ ), on définit (avec les  $u_i^n$  trouvés dans les questions précédentes) une fonction,  $u_h$ , sur  $[0, 1] \times [0, T]$  en posant

$$u_h(x, t) = \frac{t - nk}{k} u_h^{(n+1)}(x) + \frac{(n+1)k - t}{k} u_h^{(n)}(x), \text{ si } t \in [nk, (n+1)k]$$

et

$$u_h^{(n)}(x) = u_i^n, \text{ si } x \in (i-1)h, ih[, i = 1, \dots, N, n = 0, \dots, M.$$

Enfin, on définit  $\varphi(u_h)$  par  $\varphi(u_h)(x, t) = \varphi(u_h(x, t))$ .

**Question 6** Montrer que les suites  $(u_h)_{M \in \mathbb{N}^*}$  et  $(\varphi(u_h))_{M \in \mathbb{N}^*}$  sont bornées dans  $L^\infty([0, 1] \times [0, T])$  (on rappelle que  $h$  est donné par  $M$ ).

**Question 7**

Montrer qu'il existe  $C$  (ne dépendant que de  $T$ ,  $\varphi$ ,  $v$  et  $u_0$ ) t.q. l'on ait, pour tout  $M \in \mathbb{N}^*$  :

1. Pour tout  $t \in [0, T]$ ,

$$\int_{\mathbb{R}} |\varphi(u_h)(x + \eta, t) - \varphi(u_h)(x, t)|^2 dx \leq C\eta,$$

pour tout  $\eta \in \mathbb{R}_+^*$ , avec  $\varphi(u_h)(\cdot, t)$  prolongée par 0 hors de  $[0, 1]$ .

2.  $\|\varphi(u_h)(\cdot, t) - \varphi(u_h)(\cdot, s)\|_{L^2([0, 1])} \leq C|t - s|$ , pour tout  $t, s \in [0, T]$ .

Une conséquence des questions 6 et 7 (que l'on admet ici est que l'on peut trouver une suite  $(h_n)_{n \in \mathbb{N}}$  et  $u \in L^\infty([0, 1] \times [0, T])$  telle que, en posant  $u_n = u_{h_n}$  (on rappelle que  $k_n = T\sqrt{h_n}$ ), l'on ait, quand  $n \rightarrow \infty$ ,

1.  $h_n \rightarrow 0$  et  $k_n \rightarrow 0$ ,

2.  $u_n \rightarrow u$  dans  $L^\infty([0,1] \times [0, T])$  pour la topologie faible- $\star$ ,
3.  $\varphi(u_n) \rightarrow \varphi(u)$  dans  $L^p([0,1] \times [0, T])$ , pour tout  $p \in [1, \infty[$ .

**Question 8** Montrer que la fonction  $u$  ainsi trouvée est solution de (2.5.121),(2.5.122).

**Remarque.** On peut aussi montrer l'unicité de la solution de (2.5.121),(2.5.122).

## 2.6 Suggestions pour les exercices

### Exercice 1 page 46 (Comparaison différences finies- volumes finis)

On rappelle que le schéma différences finies s'obtient en écrivant l'équation en chaque point de discrétilisation, et en approchant les dérivées par des quotients différentiels, alors que le schéma volumes finis s'obtient en intégrant l'équation sur chaque maille et en approchant les flux par des quotients différentiels.

### Exercice 2 page 46 (Conditionnement efficace)

#### Partie 1

1. Pour montrer que  $A$  est inversible, utiliser le théorème du rang.

2. Utiliser le fait que  $\varphi$  est un polynôme de degré 2.

3. Pour montrer que  $\|A^{-1}\| = \frac{1}{8}$ , remarquer que le maximum de  $\varphi$  est atteint en  $x = .5$ , qui correspond à un point de discrétilisation car  $N$  est impair.

#### Partie 2 Conditionnement efficace

1. Utiliser la convergence uniforme. 2. Utiliser le fait que  $A\phi = (1 \dots 1)^t$ .

### Exercice 4 page 47 (Erreur de consistance)

1. Utiliser l'erreur de consistance.

2. Trouver un contre-exemple.

### Exercice 5 page 48 (Principe du maximum)

2. Poser  $u_0 = a$ ,  $u_{N+1} = b$ . Considérer  $p = \min\{i = 0, \dots, N+1 ; u_p = \min_{j=0, \dots, N+1} u_j\}$ . Montrer que  $p = 0$  ou  $N+1$ .

### Exercice 6 page 48 (Différences finies pour un problème elliptique)

Questions 1 à 3 : application directe des méthodes de démonstration vues en cours (paragraphe 2.2 page 13).

Question 4 : La fonction  $\theta$  est introduite pour montrer une majoration de  $\|A^{-1}\|$ , puisqu'on a plus  $A\Phi = 1$ , où  $\Phi_i = \varphi(x_i)$  est la et  $\varphi$  est la fonction "miracle" dans le cas  $-u'' = f$ . Une fois qu'on a montré les bonnes propriétés de la fonction  $\theta$  (questions 4.a et 4.b), on raisonne comme dans le cours pour la question 4.c (voir démonstration de la proposition 2.14 page 17).

### Exercice 7 (Non consistance des volumes finis)

Prendre  $f$  constante et égale à 1 et prendre  $h_i = h/2$  pour  $i$  pair et  $h_i = h$  pour  $i$  impair.

### Exercice 9 page 49 (Conditions aux limites de Neumann)

1. En différences finies, écrire les équations internes de manière habituelle, et éliminez les inconnues qui apparaissent au bord  $u_0$  et  $u_{N+1}$  en discrétisant convenablement les conditions aux limites. En volumes finis, c'est encore plus simples (flux nul au bord...)

2. Remarquer les constantes sont solutions du problème continu, et chercher alors par exemple une solution à moyenne nulle.

### Exercice 10 page 49 (Conditions aux limites de Fourier (ou Robin) et Neumann)

Ecrire les équations internes de manière habituelle, et éliminez les inconnues qui apparaissent au bord  $u_0$  et  $u_{N+1}$  en discrétisant convenablement les conditions aux limites.

### Exercice 11 page 50 (Volumes finis 1D)

1. Pour justifier le schéma : écrire les bilans par maille, et approchez les flux par des quotients différentiels de manière consistante.

2 (a) On pourra, par exemple, multiplier (2.5.85) par  $u_i$  et sommer pour  $i = 1, \dots, N$ , puis conclure en remarquant, en particulier, que  $\sum_{i=1}^N (u_i - u_{i-1})u_i = \frac{1}{2} \sum_{i=1}^{N+1} (u_i - u_{i-1})^2$ , avec  $u_0 = u_{N+1} = 0$ .

2 (b) Pensez au miracle de la dimension finie...

4. Effectuer les développements de Taylor

5. (b) (c) Se débarrasser des termes de convection en remarquant qu'ils ont "le bon signe", et s'inspirer de la démonstration du théorème 2.25 page 22.

### Exercice 13 (Discrétisation 2D par différences finies)

Adapter le cas unidimensionnel, en faisant attention aux conditions limites. Pour montrer l'existence et unicité, calculer le noyau de la matrice.

### Exercice 15 (Elimination des inconnues d'arêtes)

1. Ecrire la conservativité du flux :  $F_{K,\sigma} = -F_{L,\sigma}$  et en déduire la valeur de  $u_\sigma$ .

2. Trouver la valeur de  $u_\sigma$  qui vérifie

$$-m(\sigma)\lambda_i \frac{u_\sigma - u_K}{d_{K,\sigma}} = m(\sigma)\alpha(u_\sigma - u_{ext}).$$

3. Remplacer  $F_{K,\sigma}$  et  $F_{L,\sigma}$  par leurs expressions dans (2.3.51) et en déduire la valeur de  $u_\sigma$ .

4. Adopter l'ordre lexicographique pour la numérotation des mailles, puis établir l'équation de chaque maille, en commençant par les mailles interne.

### Exercice 18 (Exemple de schéma non convergent)

1. Ecrire le schéma d'Euler explicite.
2. Démontrer par récurrence que

$$\text{Si } n \in \{0, \dots, M+1\}, i \in \left\{-\frac{N+1}{2}, \dots, \frac{N+1}{2}\right\} \text{ et } i \geq -\frac{N+1}{4} + n \text{ alors } u_i^n = 0.$$

En déduire que  $u_i^n = 0$  pour  $n \in \{0, \dots, M+1\}$  et  $i \in \{0, \dots, \frac{N+1}{2}\}$  et conclure.

### Exercice 19 (Discrétisation d'un problème parabolique)

1. Calculer l'erreur de consistance et la majorer par des développements de Taylor. Chercher ensuite les conditions pour que :

$$\|u^n\|_\infty \leq \|u^0\|_\infty.$$

Pour étudier la convergence du schéma, majorer l'erreur de discrétisation :  $e_j^n = \bar{u}_j^n - u_j^n$  où  $u_j^n$  est calculé par (2.7.146), et  $\bar{u}_j^n$  est la solution du problème (2.5.107) en  $x_j = jh$  et  $t_n = nk$ .

Même chose pour les questions suivantes...

### Exercice 20 (Problème parabolique avec terme source)

1. Effectuer des développements de Taylor...
3. Montrer par récurrence que  $\max_{j=1, \dots, N} u_j^n \leq (1+k)^n \max_{j=1, \dots, N} u_j^0$ . et que  $\min_{j=1, \dots, N} u_j^{(n)} \geq (1+k)^n \min_{j=1, \dots, N} u_j^{(0)}$ .
4. Utiliser l'équation, le schéma, et l'erreur de consistance.

### Exercice 21 (Schémas de Saute-Mouton et Dufort-Frankel)

1. Effectuer des développements de Taylor pour majorer l'erreur de consistance.
2. Montrer que le facteur d'amplification  $\xi_n$  obtenu par l'analyse de stabilité de Von Neumann satisfait :

$$\xi_{n+1} - \alpha \xi_n - \xi_{n-1} = 0, \quad n \geq 2.$$

Etudier ensuite les racines de l'équation  $r^2 - \alpha r - 1 = 0$  et montrer que l'une de ses racines est, en module, supérieure à 1.

4. Reprendre la méthode développée à la question 2, en montrant que l'équation caractéristique pour  $\xi$  est maintenant :

$$p(r) = ar^2 + br + c = 0,$$

avec

$$a = \frac{1}{2k} + \frac{1}{h^2}, \quad b = -\frac{2 \cos(ph)}{h^2} \text{ et } c = \frac{1}{h^2} - \frac{1}{2h}.$$

Etudier ensuite les racines de cette équation.

## 2.7 Corrigés des exercices

### Corrigé de l'exercice 1 page 46

Le schéma différences finies pour l'équation (2.5.75) s'écrit :

$$\begin{cases} \frac{1}{h^2} (2u_i - u_{i-1} - u_{i+1}) = f_i, & i = 1, \dots, N, \\ u_0 = a, \quad u_{N+1} = b. \end{cases}$$

Le schéma volumes finis pour la même équation s'écrit :

$$F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}} = h f_i, \quad i = 1, \dots, N \quad (2.7.130)$$

$$\text{avec } F_{i+\frac{1}{2}} = -\frac{u_{i+1} - u_i}{h}, \quad i = 1, \dots, N-1 \text{ et } F_{\frac{1}{2}} = -\frac{u_1 - a}{\frac{h}{2}},$$

$$F_{N+\frac{1}{2}} = -\frac{b - u_N}{\frac{h}{2}}$$

En remplaçant les expressions des flux dans l'équation (2.7.130). On obtient :

$$\begin{aligned} \frac{1}{h^2} (2u_i - u_{i+1} - u_{i-1}) &= f_i, \quad i = 2, \dots, N-1 \\ \frac{1}{h^2} (3u_1 - 2u_2 - a) &= 2f_1 \\ \frac{1}{h^2} (3u_N - 2u_{N-1} - b) &= 2f_N, \end{aligned}$$

La différence entre les deux schémas réside dans la première et dernière équations.

### Corrigé de l'exercice 2 page 46 (Conditionnement “efficace”)

#### Partie I

1. Soit  $u = (u_1 \dots u_N)^t$ . On a

$$Au = b \Leftrightarrow \begin{cases} \frac{1}{h^2}(u_i - u_{i-1}) + \frac{1}{h^2}(u_i - u_{i+1}) = b_i, & \forall i = 1, \dots, N, \\ u_0 = u_{N+1} = 0. \end{cases}$$

Supposons  $b_i \geq 0, \forall i = 1, \dots, N$ , et soit  $p \in \{0, \dots, N+1\}$  tel que  $u_p = \min(u_i, i = 0, \dots, N+1)$ .

Si  $p = 0$  ou  $N+1$ , alors  $u_i \geq 0 \forall i = 0, N+1$  et donc  $u \geq 0$ .

Si  $p \in \{1, \dots, N\}$ , alors

$$\frac{1}{h^2}(u_p - u_{p-1}) + \frac{1}{h^2}(u_p - u_{p+1}) \geq 0$$

et comme  $u_p - u_{p-1} < 0$  et  $u_p - u_{p+1} \leq 0$ , on aboutit à une contradiction.

Montrons maintenant que  $A$  est inversible. On vient de montrer que si  $Au \geq 0$  alors  $u \geq 0$ . On en déduit par linéarité que si  $Au \leq 0$  alors  $u \leq 0$ , et donc que si  $Au = 0$  alors  $u = 0$ . Ceci démontre que l'application linéaire représentée par la matrice  $A$  est injective donc bijective (car on est en dimension finie).

2. Soit  $\varphi \in C([0, 1], \mathbb{R})$  tel que  $\varphi(x) = 1/2x(1-x)$  et  $\phi_i = \varphi(x_i)$ ,  $i = 1, N$ , où  $x_i = ih$ .  $(A\phi)_i$  est le développement de Taylor à l'ordre 2 de  $\varphi''(x_i)$ , et comme  $\varphi$  est un polynôme de degré 2, ce développement est exact. Donc  $(A\phi)_i = \varphi''(x_i) = 1$ .

3 Soient  $b \in \mathbb{R}^N$  et  $u \in \mathbb{R}^N$  tels que  $Au = b$ . On a :

$$(A(u \pm \|b\|\varphi))_i = (Au)_i \pm \|b\|(A\phi)_i = b_i \pm \|b\|.$$

Prenons d'abord  $\tilde{b}_i = b_i + \|b\| \geq 0$ , alors par la question (1),

$$u_i + \|b\|\phi_i \geq 0 \quad \forall i = 1, \dots, N.$$

Si maintenant on prend  $\bar{b}_i = b_i - \|b\| \leq 0$ , alors

$$u_i - \|b\|\phi_i \leq 0 \quad \forall i = 1, \dots, N.$$

On a donc  $-\|b\|\phi_i \leq \|b\|\phi_i$ .

On en déduit que  $\|u\|_\infty \leq \|b\| \|\phi\|_\infty$ ; or  $\|\phi\|_\infty = \frac{1}{8}$ . D'où  $\|u\|_\infty \leq \frac{1}{8}\|b\|$ .

On peut alors écrire que pour tout  $b \in \mathbb{R}^N$ ,

$$\|A^{-1}b\|_\infty \leq \frac{1}{8}\|b\|, \text{ donc } \frac{\|A^{-1}b\|_\infty}{\|b\|_\infty} \leq \frac{1}{8}, \text{ d'où } \|A^{-1}\| \leq \frac{1}{8}.$$

On montre que  $\|A^{-1}\| = \frac{1}{8}$  en prenant le vecteur  $b$  défini par  $b(x_i) = 1$ ,  $\forall i = 1, \dots, N$ . On a en effet  $A^{-1}b = \phi$ , et comme  $N$  est impair,  $\exists i \in \{1, \dots, N\}$  tel que  $x_i = \frac{1}{2}$ ; or  $\|\phi\|_\infty = \varphi(\frac{1}{2}) = \frac{1}{8}$ .

4. Par définition, on a  $\|A\| = \sup_{\|x\|_\infty=1} \|Ax\|$ , et donc  $\|A\| = \max_{i=1, N} \sum_{j=1, N} |a_{i,j}|$ , d'où le résultat.

5. Grâce aux questions 3 et 4, on a, par définition du conditionnement pour la norme  $\|\cdot\|$ ,  $\text{cond}(A) = \|A\|\|A^{-1}\| = \frac{1}{2h^2}$ .

Comme  $A\delta_u = \delta_b$ , on a :

$$\|\delta_u\| \leq \|A^{-1}\|\delta_b\| \frac{\|b\|}{\|b\|} \leq \|A^{-1}\|\delta_b\| \frac{\|A\|\|u\|}{\|b\|},$$

d'où le résultat.

Pour obtenir l'égalité, il suffit de prendre  $b = Au$  où  $u$  est tel que  $\|u\| = 1$  et  $\|Au\| = \|A\|$ , et  $\delta_b$  tel que  $\|\delta_b\| = 1$  et  $\|A^{-1}\delta_b\| = \|A^{-1}\|$ . On obtient alors

$$\frac{\|\delta_b\|}{\|b\|} = \frac{1}{\|A\|} \text{ et } \frac{\|\delta_u\|}{\|u\|} = \|A^{-1}\|.$$

D'où l'égalité.

## Partie 2 Conditionnement efficace

1. Soient  $\varphi^{(h)}$  et  $f^{(h)}$  les fonctions constantes par morceaux définies par

$$\begin{aligned}\varphi^h(x) &= \begin{cases} \varphi(ih) = \phi_i \text{ si } x \in ]x_i - \frac{h}{2}, x_i + \frac{h}{2}[, i = 1, \dots, N, \\ 0 \text{ si } x \in [0, \frac{h}{2}] \text{ ou } x \in ]1 - \frac{h}{2}, 1]. \end{cases} \text{ et} \\ f^{(h)}(x) &= \begin{cases} f(ih) = b_i \text{ si } x \in ]x_i - \frac{h}{2}, x_i + \frac{h}{2}[, \\ f(ih) = 0 \text{ si } x \in [0, \frac{h}{2}] \text{ ou } x \in ]1 - \frac{h}{2}, 1]. \end{cases}\end{aligned}$$

Comme  $f \in C([0, 1], \mathbb{R})$  et  $\varphi \in C^2([0, 1], \mathbb{R})$ , la fonction  $f_h$  (resp.  $\varphi_h$ ) converge uniformément vers  $f$  (resp.  $\varphi$ ) lorsque  $h \rightarrow 0$ . On a donc

$$h \sum_{i=1}^N b_i \varphi_i = \int_0^1 f^{(h)}(x) \varphi^{(h)}(x) dx \rightarrow \int_0^1 f(x) \varphi(x) dx \text{ lorsque } h \rightarrow 0.$$

Comme  $b_i > 0$  et  $f_i > 0 \forall i = 1, \dots, N$ , on a évidemment

$$S_N = \sum_{i=1}^N b_i \varphi_i > 0 \text{ et } S_N \rightarrow \int_0^1 f(x) \varphi(x) dx = \beta > 0 \text{ lorsque } h \rightarrow 0.$$

Donc il existe  $N_0 \in \mathbb{N}$  tel que si  $N \geq N_0$ ,  $S_N \geq \frac{\beta}{2}$ , et donc  $S_N \geq \alpha = \min(S_0, S_1, \dots, S_{N_0}, \frac{\beta}{2}) > 0$ .

2. On a  $N\|u\| = N \sup_{i=1, N} |u_i| \geq \sum_{i=1}^N u_i$ . D'autre part,  $A\varphi = (1 \dots 1)^t$  donc  $u \cdot A\varphi = \sum_{i=1}^N u_i$ ; or

$u \cdot A\varphi = A^t u \cdot \varphi = Au \cdot \varphi$  car  $A$  est symétrique. Donc  $u \cdot A\varphi = \sum_{i=1}^N b_i \varphi_i \geq \frac{\alpha}{h}$  d'après la question 1. Comme

$\delta_u = A^{-1}\delta_b$ , on a donc  $\|\delta_u\| \leq \|A^{-1}\| \|\delta_b\|$ ; et comme  $N\|u\| \geq \frac{\alpha}{h}$ , on obtient :  $\frac{\|\delta_u\|}{\|u\|} \leq \frac{1}{8} \frac{hN}{\alpha} \|\delta_b\| \frac{\|f\|_\infty}{\|b\|}$ .

Or  $hN = 1$  et on a donc bien :

$$\frac{\|\delta_u\|}{\|u\|} \leq \frac{\|f\|_\infty}{8\alpha} \frac{\|\delta_b\|}{\|b\|}.$$

3. Le conditionnement  $\text{cond}(A)$  calculé dans la partie 1 est d'ordre  $1/h^2$ , et donc tend vers l'infini lorsque le pas du maillage tend vers 0, alors qu'on vient de montrer dans la partie 2 que la variation relative  $\frac{\|\delta_u\|}{\|u\|}$  est inférieure à une constante multipliée par la variation relative de  $\frac{\|\delta_b\|}{\|b\|}$ . Cette dernière information est nettement plus utile et réjouissante pour la résolution effective du système linéaire.

### Corrigé de l'exercice 3 page 47 (Conditionnement, réaction diffusion 1d)

1. Pour  $k = 1$  à  $N$ , calculons  $BU_k$  :

$$(BU_k)_j = -\sin k\pi(j-1)h + 2 \sin k\pi(jh) - \sin k\pi(j+1)h, \text{ où } h = \frac{1}{N+1}.$$

En utilisant le fait que  $\sin(a+b) = \sin a \cos b + \cos a \sin b$  pour développer  $\sin k\pi(1-j)h$  et  $\sin k\pi(j+1)h$ , on obtient (après calculs) :

$$(BU_k)_j = \lambda_k(U_k)_j, \quad j = 1, \dots, N,$$

où  $\lambda_k = 2(1 - \cos k\pi h) = 2(1 - \cos \frac{k\pi}{N+1})$ . On peut remarquer que pour  $k = 1, \dots, N$ , les valeurs  $\lambda_k$  sont distinctes.

On a donc trouvé les  $N$  valeurs propres  $\lambda_1 \dots \lambda_N$  de  $B$  associées aux vecteurs propres  $U_1, \dots, U_N$  de  $\mathbb{R}^N$  tels que  $(U_k)_j = \sin \frac{k\pi j}{N+1}$ ,  $j = 1, \dots, N$ .

2. Comme  $A = Id + \frac{1}{h^2}B$ , les valeurs propres de la matrice  $A$  sont les valeurs  $\mu_i = 1 + \frac{1}{h^2}\lambda_i$ .
3. Comme  $A$  est symétrique, le conditionnement de  $A$  est donné par

$$cond_2(A) = \frac{\mu_N}{\mu_1} = \frac{1 + \frac{2}{h^2}(1 - \cos \frac{N\pi}{N+1})}{1 + \frac{2}{h^2}(1 - \cos \frac{\pi}{N+1})}.$$

### Corrigé de l'exercice 4 page 47 (Erreur de consistance)

1. Si  $f$  est constante, alors  $-u''$  est constante, et donc les dérivées d'ordre supérieur de  $u$  sont nulles. Donc par l'estimation (2.2.19) page 16 sur l'erreur de consistance, on a  $R_i = 0$  pour tout  $i = 1, \dots, N$ . Si on appelle  $U$  le vecteur de composantes  $u_i$  et  $\bar{U}$  le vecteur de  $\mathbb{R}^N$  de composantes  $u(x_i)$ , on peut remarquer facilement que  $U - \bar{U} = A^{-1}R$ , où  $R$  est le vecteur de composantes  $R_i$ . On a donc  $U - \bar{U} = 0$ , c.q.f.d.
2. Il est facile de voir que  $f$  n'est pas forcément constante, en prenant  $f(x) = \sin 2\pi x$ , et  $h = 1/2$ , on n'a donc qu'une seule inconnue  $u_1$  qui vérifie  $u_1 = 0$ , et on a également  $u(1/2) = \sin \pi = 0$ .

### Corrigé de l'exercice 5 page 48

1. On se donne une discrétisation  $(x_i)_{i=1, \dots, N}$  de  $[0, 1]$ , avec un pas constant  $h$ . On approche  $u''(x_i)$  par le quotient différentiel  $\frac{1}{h^2}(2u(x_i) - u(x_{i-1}) - u(x_{i+1}))$ .

Le schéma résultant s'écrit donc :

$$\begin{cases} \frac{1}{h^2}(2u_i - u_{i-1} - u_{i+1}) + c_i u_i = f_i, & i = 1, \dots, N. \\ u_0 = a, u_{N+1} = b. \end{cases}$$

2. Soit  $p = \min\{i; u_i = \min_{j=0, N+1} u_j\}$ .

Supposons que  $1 < p < N$ , alors on a

$$\frac{1}{h^2}(u_p - u_{p-1}) + \frac{1}{h^2}(u_p - u_{p+1}) = f_p.$$

Or  $u_p = \min u_j$ , on a donc  $u_p - u_{p-1} < 0$ , puisque  $p = \min\{i; u_i = \min_{j=0, \dots, N+1} u_j\}$ , et  $u_p - u_{p+1} \leq 0$ . Comme  $f_p \geq 0$ , on aboutit à une contradiction. On en déduit que  $p = 0$  ou  $N+1$ , ce qui prouve que  $u_i \geq \min(a, b)$ .

### Corrigé de l'exercice 6 page 48

1. On se donne un pas constant  $h = x_{i+1} - x_i = \frac{1}{N+1}$ . Comme  $u \in C^4([0, 1])$ , un développement de Taylor à l'ordre 4 aux points  $x_{i+1}$  et  $x_i$  donne :

$$\begin{cases} u(x_{i+1}) = u(x_i) + hu'(x_i) + \frac{h^2}{2} u''(x_i) + \frac{h^3}{6} u^{(3)}(x_i) + \frac{h^4}{24} u^{(4)}(\eta_i) \\ u(x_{i-1}) = u(x_i) - hu'(x_i) + \frac{h^2}{2} u''(x_i) - \frac{h^3}{6} u^{(3)}(x_i) + \frac{h^4}{24} u^{(4)}(\zeta_i), \end{cases}$$

où  $\eta_i$  (resp.  $\zeta_i$ ) appartient à l'intervalle  $[x_i, x_{i+1}]$  (resp.  $[x_{i-1}, x_i]$ ). En effectuant la somme et la différence de ces deux lignes, et on trouve :

$$\begin{cases} -u''(x_i) = \frac{2u(x_i) - u(x_{i-1}) - u(x_{i+1})}{h^2} + \frac{h^2}{24}(u^{(4)}(\eta_i) + u^{(4)}(\zeta_i)) \\ u'(x_i) = \frac{u(x_{i+1}) - u(x_{i-1})}{2h} + \frac{h^3}{3} u^{(3)}(x_i) + \frac{h^2}{24}(u^{(4)}(\eta_i) - u^{(4)}(\zeta_i)). \end{cases} \quad (2.7.131)$$

On prend alors en compte les conditions limites du système de type Dirichlet. On introduit les valeurs  $u_0 = a$  et  $u_{N+1} = b$ , et on obtient alors le schéma suivant :

$$\begin{cases} \frac{2u_i - u_{i-1} - u_{i+1}}{h^2} + \left(\frac{1}{1+ih}\right) \frac{u_{i+1} - u_{i-1}}{2h} = f_i, \quad i = 1, \dots, N, \\ u_0 = a, \quad u_{N+1} = b. \end{cases}$$

On en déduit la forme matricielle du système  $A_h u_h = b_h$

$$A_h = \begin{pmatrix} \frac{2}{h^2} & -\frac{1}{h^2} + \frac{1}{2h(1+h)} & 0 & \dots & 0 \\ -\frac{1}{h^2} + \frac{1}{2h(1+h)} & \frac{2}{h^2} & \dots & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & \vdots \\ \vdots & & \ddots & \ddots & -\frac{1}{h^2} + \frac{1}{2h(1+(N-1)h)} \\ 0 & \dots & 0 & -\frac{1}{h^2} + \frac{1}{2h(1+Nh)} & \frac{2}{h^2} \end{pmatrix},$$

soit encore

$$A_h = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & 0 & \dots & \dots & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & & \\ \vdots & & \ddots & \ddots & \ddots & -1 \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix} + \frac{1}{2h} \begin{pmatrix} 0 & \frac{1}{1+h} & 0 & \dots & \dots & 0 \\ -\frac{1}{1+2h} & \ddots & \frac{1}{1+2h} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & & \\ \vdots & & & & & \frac{1}{1+(N-1)h} \\ 0 & \dots & 0 & 0 & \frac{1}{1+Nh} & 0 \end{pmatrix}$$

avec

$$b_h = \begin{pmatrix} f_1 + a \left( \frac{1}{h^2} + \frac{1}{2h(1+h)} \right) \\ f_2 \\ \vdots \\ \vdots \\ f_{N-1} \\ f_N + b \left( \frac{1}{h^2} - \frac{1}{2h(1+Nh)} \right) \end{pmatrix} \text{ et } u_h = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_{N-1} \\ u_N \end{pmatrix}.$$

2. Majorons l'erreur de consistance :

$$R_1 = \frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2} - u''(x_i)$$

$$R_2 = \frac{u(x_{i+1}) - u(x_{i-1})}{2h} - u'(x_i)$$

La première égalité de (2.7.131) entraîne que :

$$\|R^1\|_\infty \leq \frac{h^2}{12} \sup_{[0;1]} |u^{(4)}|.$$

En effectuant alors les développements limités jusqu'à l'ordre 3, on obtient de la même manière que :

$$\|R^2\|_\infty \leq \frac{h^2}{6} \sup_{[0;1]} \|u^{(3)}\|$$

D'où on déduit finalement que :

$$\|R\|_\infty \leq \frac{h^2}{12} \left( \sup_{[0,1]} |u^{(4)}| + 2 \sup_{[0,1]} |u^{(3)}| \right). \quad (2.7.132)$$

3. Par hypothèse on suppose  $A_h v \geq 0$ , et on va montrer que  $v \geq 0$  pour  $v \in \mathbb{R}^N$ . L'hypothèse sur la matrice  $A_h$  nous permet d'écrire pour la  $i$ -ème ligne l'inéquation suivante :

$$\left( \frac{-1}{h^2} - \frac{1}{2h(1+ih)} \right) u_{i-1} + \frac{2}{h^2} u_i + \left( \frac{-1}{h^2} + \frac{1}{2h(1+ih)} \right) u_{i+1} \geq 0 \quad (2.7.133)$$

Soit

$$p = \min \left\{ i \in \{1, \dots, N\}; v_p = \min_{j=1, \dots, N} v_j \right\}.$$

Supposons d'abord que  $p = 1$ . On a alors :

$$v_1 \leq v_j, \forall j = 1, \dots, N.$$

Mais l'inéquation (2.7.133) pour  $p = 1$  s'écrit encore :

$$\frac{1}{h^2}(v_1 - v_2) + \left( \frac{1}{h^2} + c_1 \right) v_1 \geq 0,$$

On a donc

$$v_1 \geq \frac{1}{1 + c_1 h^2} (v_2 - v_1) \geq 0,$$

ce qui montre que  $\min_{j=1, \dots, N} v_j = v_1 \geq 0$  si  $p = 1$ . Un raisonnement similaire permet de montrer que si  $p = N$ , on a  $v_N \geq 0$  et donc  $\min_{j=1, \dots, N} v_j \geq 0$ .

Supposons enfin que  $p \in \{2, \dots, N-1\}$  on a alors, par la deuxième inéquation de (2.2.17) :

$$\frac{1}{h^2}(v_p - v_{p-1}) + \frac{1}{h^2}(v_p - v_{p+1}) + \frac{1}{2h(1+ph)}(v_{p+1} - v_p + v_p - v_{p-1}) \geq 0.$$

On en déduit que

$$\left( \frac{1}{h^2} - \frac{1}{2h(1+ph)} \right) (v_p - v_{p-1}) + \left( \frac{1}{h^2} + \frac{1}{2h(1+ph)} \right) (v_p - v_{p+1}) \geq 0.$$

Or  $h < 1$ , donc  $\frac{1}{h^2} - \frac{1}{2h(1+ph)} > 0$ . Les deux termes du membre de gauche de l'équation ci-dessus sont donc négatifs ou nuls. On doit donc avoir :  $v_{p-1} = v_p = v_{p+1}$  ce qui est impossible car  $p$  est le plus petit indice  $j$  tel que  $v_j = \min_{i=1, \dots, N} v_i$ . Donc dans ce cas le minimum ne peut pas être atteint pour  $j = p > 1$ . On a ainsi finalement montré que  $\min_{i \in \{1, \dots, N\}} v_i \geq 0$ , on a donc  $v \geq 0$ .

La matrice  $A_h$  est monotone si et seulement si elle est inversible et que  $A_h^{-1} \geq 0$ . Pour vérifier que  $A_h$  est inversible il suffit de remarquer que si  $A_h v = 0$ , alors  $A_h v \geq 0$  et  $A_h(-v) \geq 0$ , ce qui entraîne que  $v = 0$ , par conservation de la positivité. On en déduit que  $\text{Ker}(A_h) = \{0\}$ , et donc  $A_h$  est inversible. Pour montrer que  $A_h$  est monotone, il nous rester à prouver que  $A_h^{-1} \geq 0$ .

On se donne pour cela  $b \in \mathbb{R}^n$  tel que  $A_h b = e_i$ , où  $e_i$  est le vecteur de la base canonique de  $\mathbb{R}^n$ , ainsi  $e_i \geq 0$ . D'après la conservation de la positivité, on a  $b \geq 0$ . Cependant  $b$  est le  $i$ ème colonne de  $A_h^{-1}$ , donc  $b$  est solution de  $A_h^{-1} e_i = b$  tel que  $b \geq 0$  donc  $(A_h^{-1}) \geq 0$ .

4.a. On calcule les dérivées successives de  $\theta$  :

$$\begin{cases} \theta'(x) = -\frac{1}{2}(1+x) - (1+x)\ln(1+x) + \frac{4}{3}(1+x)\ln 2, \\ \theta''(x) = -\frac{1}{2} - 1 - \ln(1+x) + \frac{4}{3}\ln 2, \\ \theta^{(3)}(x) = \frac{-1}{1+x}, \\ \theta^{(4)}(x) = \frac{1}{(1+x)^2}. \end{cases} \quad (2.7.134)$$

On vérifie alors que  $\theta(x)$  est solution de l'équation :

$$-\theta_{xx}(x) + \frac{1}{1+x}\theta_x(x) = f(x),$$

avec  $f(x) = \frac{1}{2} + 1 + \ln(1+x) - \frac{4}{3}\ln 2 - \frac{1}{2} - \ln(1+x) + \frac{4}{3}\ln 2 = 1$ . De plus  $\theta(0) = 0$  et  $\theta(1) = 0$ , donc  $A_h\theta_h = b_h$ . On a donc :

$$\max_{1 \leq i \leq n} \left| \frac{1}{h^2}(-\theta_{i-1} + 2\theta_i - \theta_{i+1}) + \frac{1}{2h(1+ih)}(\theta_{i+1} - \theta_{i-1}) - 1 \right| \leq \frac{h^2}{12} \left( \sup_{[0;1]} |\theta^{(4)}| + 2 \sup_{[0;1]} |\theta^{(3)}| \right)$$

et comme  $\sup_{[0;1]} |\theta^{(4)}| = 1$  et  $\sup_{[0;1]} |\theta^{(3)}| = 1$ , on a

$$\max_{1 \leq i \leq n} \left| \frac{1}{h^2}(-\theta_{i-1} + 2\theta_i - \theta_{i+1}) + \frac{1}{2h(1+ih)}(\theta_{i+1} - \theta_{i-1}) - 1 \right| \leq \frac{h^2}{4}.$$

4.b On a d'après la question précédente que :  $\max_{1 \leq i \leq n} |(A_h\theta_h) - f_i| \leq \frac{h^2}{4}$ , avec  $f_i = 1$ , et donc :

$$\frac{h^2}{4} \leq (A_h\theta_h)_i - 1 \leq \frac{h^2}{4}$$

ce qui entraîne

$$\Rightarrow (A_h\theta_h)_i \geq 1 - \frac{h^2}{4}.$$

4.c Par définition de la norme on a :

$$\|B\|_\infty = \sup_{v \neq 0} \frac{\|Bv\|}{\|v\|} = \sup_{i \in (1; \cdot; N)} \sum_{j=1}^N |B_{i,j}|,$$

et comme  $A_h^{-1}$  est une matrice positive, on a :

$$\|A_h^{-1}\|_\infty = \sup_{i \in (1; \cdot; N)} \sum_{j=1}^N (A_h^{-1})_{i,j}.$$

On se donne

$$v = (1 - \frac{1}{4}h^2; \dots; 1 - \frac{1}{4}h^2),$$

et on note  $d = A_h^{-1}v$  on a donc  $A_h d = v$ . D'après la question 4b on a :

$$(A_h\theta_h)_i \geq v_i, \forall i \in (1, \dots, N)$$

ce qui peut encore s'écrire :

$$A_h(\theta_h - A_h^{-1}v) \geq 0.$$

Par conservation de la positivité (question 3), on en déduit que

$$\theta_h - A_h^{-1}v \geq 0,$$

soit encore

$$(\theta_h)_i \geq (1 - \frac{1}{4}h^2)(A_h^{-1}e)_i \text{ avec } e = (1, \dots, 1)^t$$

Or  $e \geq 0$ ,  $A_h^{-1}e \geq 0$ , et  $1 - \frac{1}{4}h^2 > 0$ . On en déduit que  $(A_h^{-1}e)_i \leq \frac{1}{1 - \frac{1}{4}h^2}(\theta_h)_i$ , soit encore :

$$\|A_h^{-1}e\|_\infty = \|A_h^{-1}\|_\infty \leq \frac{1}{1 - \frac{1}{4}h^2} \|\theta_h\|_\infty.$$

La fonction  $\theta$  est continue et bornée sur  $[0; 1]$ ; il existe donc  $K$  tel que  $|\theta(x)| \leq K$ . De plus  $\frac{1}{1 - \frac{1}{4}h^2} \leq \frac{4}{3}$ . On en déduit que

$$\|A_h^{-1}\|_\infty \leq M, \quad (2.7.135)$$

avec  $M = \frac{4}{3}K$ .

5. Dans la question 2, on a montré que le schéma est consistant d'ordre 2 avec la majoration suivante :

$$\|R^h\|_\infty \leq \frac{h^2}{12} \left( \sup_{[0;1]} \|u^{(4)}\| + 2 \cdot \sup_{[0;1]} \|u^{(3)}\| \right)$$

Soit  $\bar{u} = (u(x_1), \dots, u(x_N))^t$  le vecteur dont les composantes sont les valeurs de la solution exacte aux points de discrétisation. Par définition, on a :

$$A_h u_h - A_h \bar{u} = b_h - (b_h + R^h),$$

et donc l'erreur de discrétisation vérifie :

$$e_h = u - u_h = A_h^{-1} R^h. \quad (2.7.136)$$

Pour montrer la convergence du schéma il suffit de montrer que  $e_h$  tend vers 0 avec  $h$ . Ceci se déduit de la stabilité du schéma ainsi que la consistance. De (2.7.136), on déduit :

$$\|e_h\|_\infty = \|A_h^{-1} R^h\|_\infty \leq \|A_h^{-1}\|_\infty \|R^h\|_\infty,$$

et donc, grâce à la stabilité (2.7.135) et la consistance (2.7.132), on obtient :  $\|e_h\|_\infty \leq K h^2$ , ce qui prouve la convergence.

### Corrigé de l'exercice 7 page 49 (Non consistance des volumes finis)

Par développement de Taylor, pour  $i = 1, \dots, N$ , il existe  $\xi_i \in [x_i, x_{i+1}]$  tel que :

$$u(x_{i+1}) = u(x_i) + h_{i+\frac{1}{2}} u_x(x_i) + \frac{1}{2} h_{i+\frac{1}{2}}^2 u_{xx}(x_i) + \frac{1}{6} h_{i+\frac{1}{2}}^3 u_{xxx}(\xi_i),$$

et donc

$$R_i = -\frac{1}{h_i} \frac{h_{i+\frac{1}{2}} + h_{i-\frac{1}{2}}}{2} u_{xx}(x_i) + u_{xx}(x_i) + \rho_i, \quad i = 1, \dots, N, \quad (2.7.137)$$

où  $|\rho_i| \leq Ch$ ,  $C$  ne dépendant que de la dérivée troisième de  $u$ . Il est facile de voir que, en général,  $R_i$ , ne tend pas vers 0 lorsque  $h$  tend vers 0 (sauf dans des cas particuliers). En effet, prenons par exemple  $f \equiv 1$ ,  $h_i = h$  pour  $i$  pair,  $h_i = h/2$  pour  $i$  impair, and  $x_i = (x_{i+1/2} + x_{i-1/2})/2$ , pour  $i = 1, \dots, N$ . On a dans ce cas  $u'' \equiv -1$ ,  $u''' \equiv 0$ , et donc :

$$R_i = -\frac{1}{4} \text{ si } i \text{ est pair, et } R_i = +\frac{1}{2} \text{ si } i \text{ est impair.}$$

On en conclut que  $\sup\{|R_i|, i = 1, \dots, N\} \not\rightarrow 0$  as  $h \rightarrow 0$ .

## Corrigé de l'exercice 8 page 49

Par développement de Taylor, pour  $i = 1, \dots, N$ , il existe  $\xi_i \in x_{i+\frac{1}{2}}, x_{i+1}$   $\eta_i \in [x_i, x_{i+\frac{1}{2}}]$  tels que :

$$u(x_{i+1}) = u(x_{i+\frac{1}{2}}) + (x_{i+1} - x_{i+\frac{1}{2}})u'(x_{i+\frac{1}{2}}) + \frac{1}{2}(x_{i+1} - x_{i+\frac{1}{2}})^2u''(x_{i+\frac{1}{2}}) + \frac{1}{6}(x_{i+1} - x_{i+\frac{1}{2}})^3u'''(\eta_i),$$

$$u(x_i) = u(x_{i+\frac{1}{2}}) + (x_i - x_{i+\frac{1}{2}})u'(x_{i+\frac{1}{2}}) + \frac{1}{2}(x_i - x_{i+\frac{1}{2}})^2u''(x_{i+\frac{1}{2}}) + \frac{1}{6}(x_i - x_{i+\frac{1}{2}})^3u'''(\eta_i),$$

Par soustraction, on en déduit que :

$$u(x_{i+1}) - u(x_i) = (x_{i+1} - x_i)u'(x_{i+\frac{1}{2}}) + \frac{1}{2}(x_{i+1} - x_i)(x_{i+1} + x_i - 2x_{i+\frac{1}{2}})u''(x_{i+\frac{1}{2}}) + \tilde{\rho}_{i+\frac{1}{2}},$$

où

$$\tilde{\rho}_{i+\frac{1}{2}} = \frac{1}{6}((x_{i+1} - x_{i+\frac{1}{2}})^3u'''(\xi_i) - (x_i - x_{i+\frac{1}{2}})^3u'''(\eta_i))$$

et donc, en posant  $F_{i+\frac{1}{2}}^* = -\frac{u(x_{i+1}) - u(x_i)}{h_{i+\frac{1}{2}}}$ , on obtient après simplifications :

$$F_{i+\frac{1}{2}}^* = -u'(x_{i+\frac{1}{2}}) - \frac{1}{2} \left[ x_{i+1} + x_i - 2x_{i+\frac{1}{2}} \right] u''(x_{i+\frac{1}{2}}) + \rho_{i+\frac{1}{2}}$$

avec  $|\rho_{i+\frac{1}{2}}| \leq Ch^2$ , où  $h = \max_{i=1, \dots, N} h_i$  et  $C$  ne dépend que de  $u'''$ .

Dans le cas où  $x_{i+\frac{1}{2}} = \frac{x_{i+1} + x_i}{2}$ , on a donc  $|F_{i+\frac{1}{2}}^* + u'(x_{i+\frac{1}{2}})| \leq \rho_{i+\frac{1}{2}}$ , et donc le flux est consistant d'ordre 2.

Dans le cas général, on peut seulement majorer  $\frac{1}{2} \left( x_{i+1} + x_i - 2x_{i+\frac{1}{2}} \right)$  par  $h$ , on a donc un flux consistant d'ordre 1.

## Corrigé de l'exercice 9 page 49

1. On se donne une discréétisation  $(x_i)_{i=1, \dots, N}$  de l'intervalle  $[0, 1]$ , de pas constant et égal à  $h$ . On écrit l'équation en chaque point  $x_i$ , et on remplace  $-u''(x_i)$  par le quotient différentiel habituel. En appelant  $u_1, \dots, u_N$  les inconnues localisées aux points  $x_1, \dots, x_N$ , et  $u_0, u_{N+1}$  les inconnues auxiliaires localisées en  $x = 0$  et  $x = 1$ , on obtient les équations discrètes associées aux inconnues  $i = 1, \dots, N$ .

$$\frac{1}{h^2}(2u_i - u_{i-1} - u_{i+1}) + c_i u_i = f_i,$$

avec  $c_i = c(x_i)$  et  $f_i = f(x_i)$ . Il reste à déterminer  $u_0$  et  $u_{N+1}$ . Ceci se fait en approchant la dérivée  $u'(0)$  ( resp.  $u'(1)$  ) par  $\frac{1}{h}(u(x_1) - u(0))$  ( resp.  $\frac{1}{h}(u(1) - u(x_N))$  ).

Comme  $u'(0) = 0$  et  $u'(1) = 0$ , on obtient donc que  $u_0 = u_1$  et  $u_{N+1} = u_N$ . Le schéma différences finies s'écrit donc :

$$\begin{cases} \frac{1}{h^2}(u_1 - u_2) + c_1 u_1 = f_1 \\ \frac{1}{h^2}(2u_i - u_{i-1} - u_{i+1}) + c_i u_i = f_i, \quad i = 2, \dots, N-1, \\ \frac{1}{h^2}(u_N - u_{N-1}) + c_N u_N = f_N. \end{cases}$$

2. On se donne un maillage volumes finis, et on intègre l'équation sur chaque maille, ce qui donne le schéma

$$F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}} = h_i f_i, i = 1, \dots, N,$$

où  $F_{i+\frac{1}{2}}$  est le flux numérique à l'interface  $x_{i+\frac{1}{2}}$ . Pour  $i = 1, \dots, N-1$ , ce flux numérique est donné par

$$F_{i+\frac{1}{2}} = \frac{u_{i+1} - u_i}{h_i \frac{1}{2}}.$$

Pour  $i = 0$  et  $i = N+1$ , on se sert des conditions de Neumann, qui imposent un flux nul. On écrit donc :

$$F_{\frac{1}{2}} = F_{N+\frac{1}{2}} = 0.$$

### Corrigé de l'exercice 10 page 49

1. La discréétisation par différences finies donne comme  $i$ -ème équation (voir par exemple exercice 9 page 49) :

$$\frac{1}{h^2}(2u_i - u_{i-1} - u_{i+1}) + c_i u_i = f_i, i = 1, \dots, N.$$

Il reste donc à déterminer les inconnues  $u_0$  et  $u_{N+1}$  à l'aide de la discréétisation des conditions aux limites, qu'on approche par :

$$\frac{u_1 - u_0}{h} + \alpha(u_0 - \tilde{u}) = 0,$$

$$\frac{u_{N+1} - u_N}{h} + \alpha(u_{N+1} - \tilde{u}) = 0$$

où  $u_0$  et  $u_{N+1}$  sont les valeurs approchées en  $x_0$  et  $x_{N+1}$ , on a donc par élimination :

$$u_0 = \frac{1}{\alpha - \frac{1}{h}} \left( \alpha \tilde{u} - \frac{u_1}{h} \right) \text{ et } u_{N+1} = \frac{1}{\alpha + \frac{1}{h}} \left( \alpha \tilde{u} + \frac{u_N}{h} \right).$$

Ce qui termine la définition du schéma.

2. Par volumes finis, la discréétisation de l'équation s'écrit

$$F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}} = h_i f_i, i = 1, \dots, N,$$

et les seuls flux "nouveaux" sont encore  $F_{1/2}$  et  $F_{N+1/2}$ , qu'on obtient à partir de la discréétisation des conditions aux limites. Ceci peut se faire de plusieurs manières.

On peut, par exemple, discréétiser la condition aux limites en 0 par :

$$F_{1/2} + \alpha(u_0 - \tilde{u}) = 0, \text{ avec } F_{1/2} = \frac{u_1 - u_0}{\frac{h_1}{2}}.$$

On a dans ce cas :  $-\alpha(u_0 - \tilde{u}) \times \frac{h_1}{2} = -u_1 + u_0$ , d'où on déduit que  $u_0 = \frac{\alpha \tilde{u} h_1 + 2u_1}{\alpha h_1 + 2}$ , et qui conduit à l'expression suivante pour  $F_{1/2}$  :

$$F_{1/2} = \frac{\alpha}{\alpha h_1 + 2} (2(u_1 - \tilde{u}) - \alpha h_1 \tilde{u}).$$

Le calcul est semblable pour  $F_{N+1/2}$

## Corrigé de l'exercice 11 page 50

1. On intègre (2.5.83) sur une maille  $[x_{i-1/2}, x_{i+1/2}]$  et on obtient :

$$-u'(x_{i+1/2}) + u'(x_{i-1/2}) + a[u(x_{i+1/2}) - u(x_{i-1/2})] + b \int_{x_{i-1/2}}^{x_{i+1/2}} u(x) dx = bh_i f_i. \quad (2.7.138)$$

Pour justifier le schéma numérique proposé on remarque que :

$$u(x_{i+1}) = u(x_{i+1/2}) + (x_{i+1} - x_{i+1/2})u'(x_{i+1/2}) + \frac{1}{2}(x_{i+1} - x_{i+1/2})^2 u''(\xi_i), \text{ avec } \xi_i \in [x_{i+1/2}, x_{i+1}],$$

et de même

$$u(x_i) = u(x_{i+1/2}) + (x_i - x_{i+1/2})u'(x_{i+1/2}) + \frac{1}{2}(x_i - x_{i+1/2})^2 u''(\gamma_i), \text{ avec } \gamma_i \in [x_{i-1/2}, x_i],$$

dont on déduit :

$$u(x_{i+1}) - u(x_i) = h_{i+1/2} u'(x_{i+1/2}) + \frac{1}{8}(h_{i+1}^2 u''(\xi_i) - h_i^2 u''(\gamma_i)).$$

De plus en utilisant le fait que  $x_i$  est le milieu de  $[x_{i-1/2}, x_{i+1/2}]$  on a (voir démonstration plus loin)

$$\int_{x_{i-1/2}}^{x_{i+1/2}} u dx = u(x_i)h_i + \frac{1}{24}u''(\alpha_i)h_i^3 \quad (2.7.139)$$

D'où le schéma numérique.

Démontrons la formule (2.7.139). Pour cela il suffit (par changement de variable) de démontrer que si  $u \in C^2(\mathbb{R})$ , alors pour tout  $\alpha \geq 0$ , on a :

$$\int_{-\alpha}^{\alpha} u dx = 2\alpha u(0) + \frac{1}{3}u''(\alpha) \alpha^3. \quad (2.7.140)$$

Pour cela, on utilise une formule de type Taylor avec reste intégral, qu'on obtient en remarquant que si on pose  $\varphi(t) = u(tx)$ , alors  $\varphi'(t) = xu'(tx)$ , et  $\varphi''(t) = x^2u''(tx)$ . Or  $\varphi(1) - \varphi(0) = \int_0^1 \varphi'(t) dt$ , et par intégration par parties, on obtient donc :

$$\varphi(1) = \varphi(0) + \varphi'(0) + \int_0^1 \varphi''(t)(1-t) dt.$$

On en déduit alors que

$$u(x) = u(0) + xu'(0) + \int_0^1 x^2 u''(tx)(1-t) dt.$$

En intégrant entre  $-\alpha$  et  $\alpha$ , on obtient alors :

$$\int_{-\alpha}^{\alpha} u(x) dx = 2\alpha u(0) + A, \text{ avec } A = \int_0^1 x^2 u''(tx)(1-t) dt.$$

Comme la fonction  $u''$  est continue elle est minorée et majorée sur  $[-\alpha, \alpha]$ . Soient donc  $m = \min_{[-\alpha, \alpha]} u''$  et  $M = \max_{[-\alpha, \alpha]} u''$ . Ces deux valeurs sont atteintes par  $u''$  puisqu'elle est continue. On a donc

$u''([-\alpha, \alpha]) = [m, M]$ . De plus, la fonction  $(x, t) \mapsto x^2(1-t)$  est positive ou nulle sur  $[-\alpha, \alpha] \times [0, 1]$ . On peut donc minorer et majorer  $A$  de la manière suivante

$$m \int_0^1 x^2(1-t)dt dx \leq A \leq M \int_0^1 x^2(1-t)dt dx.$$

Or  $\int_0^1 x^2(1-t)dt dx = \frac{1}{3}\alpha^3$ . On en déduit que  $\frac{1}{3}\alpha^3m \leq A \leq \frac{1}{3}\alpha^3M$ , et donc que  $A = \frac{1}{3}\alpha^3\gamma$ , avec  $\gamma \in [m, M]$ ; mais comme  $u''$  est continue, elle prend toutes les valeurs entre  $m$  et  $M$ , il existe donc  $\beta \in [-\alpha, \alpha]$  tel que  $\gamma = u''(\beta)$ , ce qui termine la preuve de (2.7.140).

2 (a). On multiplie (2.5.85) par  $u_i$  et on somme pour  $i = 1, \dots, N$ . On obtient après changement d'indice que

$$\sum_{i=0}^{i=N} \frac{(u_{i+1} - u_i)^2}{h_{i+1/2}} + \frac{a}{2} \sum_{i=0}^{i=N} (u_{i+1} - u_i)^2 + b \sum_{i=0}^{i=N} u_i^2 h_i = 0.$$

Ce qui donne  $u_i = 0$  pour tout  $i = 1 \dots N$ , d'où en mettant le schéma sous la forme matricielle  $AU = B$  on déduit que l'application linéaire représentée par la matrice  $A$  est injective donc bijective (grâce au fait qu'on est en dimension finie) et donc que (2.5.85) admet une unique solution.

3 (a). Evident.

(b). On pose  $\bar{R} = A\bar{U} - B$ . On a donc  $R_i = R_i^{(1)} + R_i^{(2)}$ , avec

$$R_i^{(1)} = -\frac{1}{h_i} \left[ \left( \frac{\bar{u}_{i+1} - \bar{u}_i}{h_{i+1/2}} - u'(x_{i+1/2}) \right) - \left( \frac{\bar{u}_i - \bar{u}_{i-1}}{h_{i-1/2}} - u'(x_{i-1/2}) \right) \right],$$

$$R_i^{(2)} = \frac{a}{h_i} [(\bar{u}_i - u(x_{i+1/2})) - (\bar{u}_{i-1} - u(x_{i-1/2}))].$$

De plus on remarque que

$$\bar{u}_i = \frac{1}{h_i} \int_{x_{i-1/2}}^{x_{i+1/2}} u dx = u(x_{i+1/2}) - \frac{1}{2}u'(x_{i+1/2})h_i + \frac{1}{6}u''(x_{i+1/2})h_i^2 - \frac{1}{24}u^{(3)}(d_i)h_i^3 \text{ avec } d_i \in [x_{i-1/2}, x_{i+1/2}],$$

$$\bar{u}_{i+1} = \frac{1}{h_{i+1}} \int_{x_{i+1/2}}^{x_{i+3/2}} u dx = u(x_{i+1/2}) + \frac{1}{2}u'(x_{i+1/2})h_{i+1} + \frac{1}{6}u''(x_{i+1/2})h_{i+1}^2 - \frac{1}{24}u^{(3)}(\delta_i)h_{i+1}^3 \text{ avec } \delta_i \in [x_{i+1/2}, x_{i+3/2}].$$

Ce qui implique que :

$$\frac{\bar{u}_{i+1} - \bar{u}_i}{h_{i+1/2}} = u'(x_{i+1/2}) + \frac{1}{3}u''(x_{i+1/2})(h_{i+1} - h_i) + \frac{1}{24} \frac{1}{h_{i+1/2}} [u^{(3)}(\delta_i)h_{i+1}^3 + u^{(3)}(d_i)h_i^3]$$

et donc

$$\frac{1}{h_i} \left[ \frac{\bar{u}_{i+1} - \bar{u}_i}{h_{i+1/2}} - u'(x_{i+1/2}) \right] = S_i + K_i,$$

avec

$$|S_i| = \left| \frac{1}{3}u''(x_{i+1/2}) \left( \frac{h_{i+1}}{h_i} - 1 \right) + \frac{1}{24} \right| \leq Ch \text{ et } |K_i| = \left| \frac{1}{h_i h_{i+1/2}} [u^{(3)}(\delta_i)h_{i+1}^3 + u^{(3)}(d_i)h_i^3] \right| \leq Ch,$$

où  $C$  ne dépend que de  $u$ . De plus si on pose :  $L_i = \frac{1}{h_i}(\bar{u}_i - u(x_{i+1/2}))$ , par développement de Taylor, il existe  $\tilde{C}$  ne dépendant que de  $u$  telle que  $|L_i| \leq Ch$ . Finalement on conclut que  $|R_i^{(1)}| = |-S_i + S_{i+1}| \leq C_1$  et  $|R_i^{(2)}| = |-K_i + K_{i-1} + a(L_i - L_{i-1})| \leq C_2h$ .

4. Reprendre les résultats précédents... Pour  $|R_{i+1/2}| \leq Ch$  reprendre calcul du 3  $|R_{i+1/2}| = |h_i(-S_i - K_i + L_i)|$ .

5 (a). On pose  $e_i = \bar{u}_i - u_i$ . Cette définition implique que  $e_i$  est solution du système (2.5.92)-(2.5.94). (b). Un calcul similaire à celui de la question 2. donne que

$$b \sum_1^N h_i e_i + \sum_{i=0}^{i=N} \frac{(e_{i+1} - e_i)^2}{h_{i+1/2}} + \frac{a}{2} \sum_{i=0}^{i=N} (e_{i+1} - e_i)^2 = \sum_{i=0}^{i=N} R_{i+1/2} (e_{i+1} - e_i)$$

D'où en utilisant le fait que

$\alpha h \leq h_i \leq \beta h$  et l'inégalité de Cauchy-Schwarz on déduit que

$$\frac{1}{\beta h} \sum_{i=0}^{i=N} (e_{i+1} - e_i)^2 \leq \left( \sum_{i=0}^{i=N} (e_{i+1} - e_i)^2 \right)^{1/2} \left( \sum_{i=0}^{i=N} R_{i+1/2}^2 \right)^{1/2}$$

et en utilisant (2.5.91), et le fait que  $\sum_{i=0}^{i=N} h_i = 1$  entraîne  $N \leq \frac{1}{\alpha h}$ , on déduit :

$$\sum_{i=0}^{i=N} (e_{i+1} - e_i)^2 \leq C_1 \frac{\beta}{\alpha} h^3.$$

En remarquant que  $e_i = \sum_{j=0}^{j=i-1} (e_{j+1} - e_j)$  on a pour tout  $0 < i \leq N$  que

$$|e_i| \leq \left( \sum_{j=0}^{j=i} (e_{j+1} - e_j)^2 \right)^{1/2} i^{1/2} \leq \left( C_1 \frac{\beta}{\alpha} h^3 \right)^{1/2} N^{1/2}$$

et donc  $|e_i| \leq \frac{\sqrt{C_1 \beta}}{\alpha} h$ , pour tout  $0 < i \leq N$ .

### Corrigé de l'exercice 14 page 53

On note  $(x, y)$  les coordonnées d'un point de  $\mathbb{R}^2$ .

1. En multipliant la première équation de (2.5.100) par  $\varphi$  et en intégrant par parties, on trouve, pour tout  $\varphi \in C^1(\bar{\Omega})$  t.q.  $\varphi = 0$  sur  $\partial\Omega$  :

$$\int_{\Omega} \nabla u(x, y) \cdot \nabla \varphi(x, y) \, dx dy + \int_{\Omega} k \frac{\partial u(x, y)}{\partial x} \varphi(x, y) \, dx = \int_{\Omega} f(x, y) \varphi(x, y) \, dx dy. \quad (2.7.141)$$

On suppose maintenant que  $f \leq 0$  sur  $\Omega$ . On se donne une fonction  $\psi \in C^1(\mathbb{R}, \mathbb{R})$  t.q. :

$$\begin{aligned} \psi(s) &= 0, & \text{si } s \leq 0, \\ \psi(s) &> 0, & \text{si } s > 0. \end{aligned}$$

(On peut choisir, par exemple,  $\psi(s) = s^2$  pour  $s > 0$  et  $\psi(s) = 0$  pour  $s \leq 0$ ) et on prend dans (2.7.141)  $\varphi = \psi \circ u$ . On obtient ainsi :

$$\int_{\Omega} \psi'(u(x, y)) |\nabla u(x, y)|^2 \, dx dy + \int_{\Omega} k \frac{\partial u}{\partial x}(x, y) \psi(u(x, y)) dx = \int_{\Omega} f(x, y) \psi(u(x, y)) \, dx dy \leq 0. \quad (2.7.142)$$

En notant  $G$  la primitive de  $\psi$  s'annulant en 0, on a :  $\frac{\partial}{\partial x} G(u(x, y)) = \psi(u(x, y)) \frac{\partial u}{\partial x}(x, y)$ . Comme  $u = 0$  sur  $\partial\Omega$ , on obtient donc :

$$\int_{\Omega} k \frac{\partial u}{\partial x}(x, y) \psi(u(x, y)) \, dx dy = \int_{\Omega} k \frac{\partial}{\partial x} G(u(x, y)) \, dx dy = \int_{\partial\Omega} k G(u(x, y)) n_x \, d\gamma(x, y) = 0,$$

où  $n_x$  désigne la première composante du vecteur normal  $n$  à  $\partial\Omega$  extérieur à  $\Omega$ , et  $d\gamma(x, y)$  le symbole d'intégration par rapport à la mesure de Lebesgue unidimensionnelle sur  $\partial\Omega$ . De (2.7.142) on déduit alors :

$$\int_{\Omega} \psi'(u(x, y)) |\nabla u(x, y)|^2 \, dx dy \leq 0,$$

et donc, comme  $\psi' \geq 0$  et que la fonction  $(x, y) \mapsto \psi'(u(x, y)) |\nabla u(x, y)|^2$  est continue :

$$\psi'(u(x, y)) |\nabla u(x, y)|^2 = 0, \forall (x, y) \in \bar{\Omega}$$

Ceci donne aussi

$$\nabla \psi(u(x, y)) = 0, \forall (x, y) \in \bar{\Omega}.$$

La fonction  $\psi \circ u$  est donc constante sur  $\bar{\Omega}$ , comme elle est nulle sur  $\partial\Omega$ , elle est nulle sur  $\bar{\Omega}$ , ce qui donne

$$u \leq 0 \text{ sur } \bar{\Omega}$$

2. On s'intéresse ici à la consistance au sens des différences finies. On pose donc

$$\bar{u}_{i,j} = u(ih, jh) \text{ pour } i, j \in \{0, \dots, N+1\}^2.$$

On a bien  $\bar{u}_{i,j} = 0$  pour  $(i, j) \in \gamma$ , et pour  $(i, j) \in \{1, \dots, N\}^2$ , on pose :

$$R_{ij} = a_0 \bar{u}_{i,j} - a_1 \bar{u}_{i-1,j} - a_2 \bar{u}_{i+1,j} - a_3 \bar{u}_{i,j-1} - a_4 \bar{u}_{i,j+1} - f_{i,j}.$$

On rappelle que  $u$  est solution de (2.5.84),  $R_j$  est donc l'erreur de consistance. Dans le cas du schéma [I] on a :

$$R_{ij} = \frac{2\bar{u}_{ij} - \bar{u}_{i+1,j} - \bar{u}_{i-1,j}}{h^2} + \frac{2\bar{u}_{ij} - \bar{u}_{i,j+1} - \bar{u}_{i,j-1}}{h^2} + k \frac{\bar{u}_{i+1,j} - \bar{u}_{i-1,j}}{2h} - f_{ij}.$$

Comme  $u \in C^4(\bar{\Omega})$ , il existe  $\xi_{ij} \in ]0, 1[$  et  $\eta_{ij} \in ]0, 1[$  t.q.

$$\bar{u}_{i+1,j} = \bar{u}_{i,j} + h \frac{\partial u}{\partial x}(ih, jh) + \frac{h^2}{2} \frac{\partial^2 u}{\partial x^2}(ih, jh) + \frac{h^3}{6} \frac{\partial^3 u}{\partial x^3}(ih, jh) + \frac{h^4}{24} \frac{\partial^4 u}{\partial x^4}(ih + \xi_{ij}h, jh)$$

$$\bar{u}_{i-1,j} = \bar{u}_{i,j} - h \frac{\partial u}{\partial x}(ih, jh) + \frac{h^2}{2} \frac{\partial^2 u}{\partial x^2}(ih, jh) - \frac{h^3}{6} \frac{\partial^3 u}{\partial x^3}(ih, jh) + \frac{h^4}{24} \frac{\partial^4 u}{\partial x^4}(ih + \eta_{ij}h, jh).$$

On obtient des formules analogues pour  $\bar{u}_{i,j+1}$  et  $\bar{u}_{i,j-1}$ , et on en déduit

$$|R_{ij}| \leq \frac{h^2}{12} \|u_{xxxx}\|_{\infty} + \frac{h^2}{12} \|u_{yyyy}\|_{\infty} + k \frac{h^2}{6} \|u_{xxx}\|_{\infty}$$

où  $\|u_{xxxx}\|_\infty$  désigne la norme uniforme sur  $\bar{\Omega}$  de la dérivée  $h^2$  de  $u$  par rapport à  $x$  (notations analogues pour  $\|u_{yyyy}\|_\infty$  et  $\|u_{xxx}\|_\infty$ ). On obtient finalement

$$|R_{ij}| \leq C_1 h^2,$$

où  $C_1$  ne dépend que de  $u$  et  $k$ . Comme pour  $h$  petit, on a  $h^2 \leq h$ , on en déduit que schéma [I] est donc d'ordre 2.

Pour le schéma [II], on a :

$$R_{ij} = \frac{2\bar{u}_{ij} - \bar{u}_{i+1,j} - \bar{u}_{i-1,j}}{h^2} + \frac{2\bar{u}_{ij} - \bar{u}_{i,j+1} - \bar{u}_{i,j-1}}{h^2} + k \frac{\bar{u}_{ij} - \bar{u}_{i-1,j}}{h} - f_{ij}.$$

D'où l'on déduit

$$|R_{ij}| \leq \frac{h^2}{12} \|u_{xxxx}\|_\infty + \frac{h^2}{12} \|u_{yyyy}\|_\infty + \frac{kh}{2} \|u_{xx}\|_\infty,$$

et donc

$$|R_{ij}| \leq C_2 h$$

où  $C_2$  ne dépend que de  $u$  et  $k$ . Le schéma [II] est donc d'ordre 1.

3. Dans le cas du schéma [II], la famille des  $w_{ij}, i, j \in \{0, \dots, N+1\}^2$  vérifie :

$$\frac{1}{h^2} (w_{ij} - w_{i+1,j}) + \left( \frac{1}{h^2} + \frac{k}{h} \right) (w_{ij} - w_{i-1,j}) + \frac{1}{h^2} (w_{ij} - w_{i,j+1}) + \frac{1}{h^2} (w_{ij} - w_{i,j-1}) \leq 0, \quad \forall i, j \in \{1, \dots, N\}$$

On pose  $M = \max\{w_{ij}, (i, j) \in \{0, \dots, N+1\}^2\}$  et  $m = \max\{w_{ij}, (i, j) \in \gamma\}$ . Noter que  $\gamma = \{0, \dots, N+1\}^2 \setminus \{1, \dots, N\}^2$ . On a bien sûr  $m \leq M$  et il reste donc à montrer que  $M \leq m$ . Soit  $A = \{(i, j) \in \{0, \dots, N+1\}^2, w_{ij} = M\}$  et soit  $(\bar{i}, \bar{j}) \in A$  tel que  $\bar{i} = \max\{i, (i, j) \in A\}$  et  $\bar{j} = \max\{j, (i, j) \in A\}$ . On distingue deux cas :

1. Si  $\bar{i} \in \{0, N+1\}$  ou  $\bar{j} \in \{0, \dots, N+1\}$ , on a alors  $(\bar{i}, \bar{j}) \in \gamma$  et donc  $M = w_{\bar{i}, \bar{j}} \leq m$ .

2. Sinon, on a  $\bar{i} \notin \{0, N+1\}$  et  $\bar{j} \notin \{0, N+1\}$ , et donc  $(\bar{i}, \bar{j}) \in \{1, \dots, N\}^2$ . On en déduit que :

$$\begin{aligned} \frac{1}{h^2} (w_{\bar{i}, \bar{j}} - w_{\bar{i}+1, \bar{j}}) + \left( \frac{1}{h^2} + \frac{k}{h} \right) (w_{\bar{i}, \bar{j}} - w_{\bar{i}-1, \bar{j}}) + \\ \frac{1}{h^2} (w_{\bar{i}, \bar{j}} - w_{\bar{i}, \bar{j}+1}) + \frac{1}{h^2} (w_{\bar{i}, \bar{j}} - w_{\bar{i}, \bar{j}-1}) \leq 0, \end{aligned}$$

ce qui est impossible car  $w_{\bar{i}, \bar{j}} = M$  et donc

$$\begin{aligned} w_{\bar{i}, \bar{j}} - w_{\bar{i}, \bar{j}-1} &\geq 0, \\ w_{\bar{i}, \bar{j}} - w_{\bar{i}, \bar{j}+1} &\geq 0, \\ w_{\bar{i}, \bar{j}} - w_{\bar{i}-1, \bar{j}} &\geq 0, \\ w_{\bar{i}, \bar{j}} - w_{\bar{i}+1, \bar{j}} &> 0, \end{aligned}$$

noter que la dernière inégalité est bien stricte car  $(\bar{i}+1, \bar{j}) \notin A$  (c'est l'intérêt du choix de  $\bar{i}$ ). On a donc bien montré que  $M \leq m$ .

Dans le cas du schéma [II], si on a  $\bar{i} \notin \{0, N+1\}$  et  $\bar{j} \notin \{0, N+1\}$ , et donc  $(\bar{i}, \bar{j}) \in \{1, \dots, N\}^2$  le même raisonnement que celui du schéma 1 donne :

$$\begin{aligned} \left( \frac{1}{h^2} - \frac{k}{2h} \right) (u_{\bar{i}, \bar{j}} - u_{\bar{i}+1, \bar{j}}) + \left( \frac{1}{h^2} + \frac{k}{2h} \right) (u_{\bar{i}, \bar{j}} - u_{\bar{i}-1, \bar{j}}), \\ + \frac{1}{h^2} (u_{\bar{i}, \bar{j}} - u_{\bar{i}, \bar{j}+1}) + \frac{1}{h^2} (u_{\bar{i}, \bar{j}} - u_{\bar{i}, \bar{j}-1}) \leq 0. \end{aligned}$$

On ne peut conclure à une contradiction que si  $\frac{1}{h^2} - \frac{k}{2h} \geq 0$ . Le schéma [II] vérifie

$$w_{i,j} \leq \max_{(k,\ell) \in \gamma} (w_{k,\ell}) \quad \forall i, j \in \{1, \dots, N\}^2$$

lorsque  $h$  vérifie la condition (dite Condition de stabilité) :

$$h \leq \frac{2}{k} \quad (2.7.143)$$

4. La fonction  $\phi$  vérifie

$$\begin{aligned} \phi_x &= 0 \\ \phi_y &= y, \quad \phi_{yy} = 1, \end{aligned}$$

et donc  $-\Delta\phi + k\frac{\partial\phi}{\partial x} = -1$ . On pose maintenant  $\phi_{i,j} = \phi(ih, jh)$  pour  $i, j \in \{a, \dots, N+1\}^2$  (Noter que  $\phi$  ne vérifie pas la condition  $\phi_{i,j} = 0$  si  $(i, j) \in \gamma$ ) Comme  $\phi_{xx} = \phi_{xxx} = \phi_{xxxx} = \phi_{yyyy} = 0$ , les calculs de la question 2 montrent que pour les schémas [I] et [II],

$$a_0\phi_{i,j} - a_1\phi_{i-1,j} - a_2\phi_{i+1,j} - a_3\phi_{i,j-1} - a_4\phi_{i,j+1} = -1$$

pour  $i, j \in \{1, \dots, N\}^2$ .

En posant  $w_{i,j} = u_{i,j} + C\phi_{i,j}$  pour  $(i, j) \in \{0, \dots, N+1\}^2$  (et  $U$  solution de (2.5.101)) on a donc

$$a_0w_{i,j} - a_1w_{i-1,j} - a_2w_{i+1,j} - a_3w_{i,j-1} - a_4w_{i,j+1} = f_{i,j} - C \quad \forall i, j \in \{1, \dots, N\}$$

On prend  $C = \|f\|_\infty$ , de sorte que  $f_{i,j} - C \leq 0$  pour tout  $(i, j)$  pour le schéma [II] et pour le schéma [I] avec  $h \leq 2/k$ , la question 3 donne alors pour  $(i, j) \in \{1, \dots, N\}^2$ ,

$$w_{i,j} \leq \max\{w_{k,\ell}, (k,\ell) \in \gamma\} \leq \frac{C}{2},$$

car  $u_{i,j} = 0$  si  $(i, j) \in \gamma$  et  $-\max_{\Omega} \phi = \frac{1}{2}$ . On en déduit pour  $(i, j) \in \{1, \dots, N\}^2$ ,

$$w_{i,j} \leq \frac{C}{2} = \frac{1}{2}\|f\|_\infty.$$

Pour montrer que  $-w_{i,j} \leq \frac{1}{2}\|f\|_\infty$ , on prend maintenant  $w_{i,j} = C\phi_{i,j} - u_{i,j}$  pour  $(i, j) \in \{0, \dots, N+1\}^2$ , avec  $C = \|f\|_\infty$ . On a donc

$$a_0w_{i,j} - a_1w_{i-1,j} - a_2w_{i+1,j} - a_3w_{i,j-1} - a_4w_{i,j+1} = -C - f_{i,j} \leq 0, \forall i, j \in \{1, \dots, N\}.$$

Ici encore, pour le schéma [II] ou le schéma [I] avec la condition  $h \leq \frac{2}{k}$ , la question 3 donne

$$w_{i,j} \leq \max\{w_{k,\ell}, (k,\ell) \in \gamma\} = \frac{C}{2}$$

donc  $u_{i,j} \geq -\frac{C}{2} = -\frac{\|f\|_\infty}{2}$  pour tout  $(i, j) \in \{1, \dots, N\}^2$ . Pour le schéma [II] ou le schéma [I] avec la condition  $h \leq \frac{2}{k}$ , on a donc :

$$\|U\|_\infty \leq \frac{1}{2}\|f\|_\infty. \quad (2.7.144)$$

Le système (2.5.101) peut être vu comme un système linéaire de  $N^2$  équation, à  $N^2$  inconnues (qui sont les  $u_{i,j}$  pour  $(i,j) \in \{1, \dots, N\}^2$ ). Si le second membre de ce système linéaire est nul, l'inégalité (2.7.144)(I) prouve que la solution est nulle. Le système (2.5.101) admet donc, pour tout  $f$ , au plus une solution. Ceci est suffisant pour affirmer qu'il admet, pour tout  $f$ , une et une seule solution.

5. pour  $(i,j) \in \{0, \dots, N+1\}^2$  on pose

$$e_{ij} = u(ih, jh) - u_{i,j}.$$

On a donc, pour les schémas [I] et [II], avec les notations de la question 2 :

$$a_0 e_{ij} - a_1 e_{i-1,j} - a_2 e_{i+1,j} - a_3 e_{i,j-1} - a_4 e_{i,j+1} = R_{ij}, \quad \forall i, j \in \{1, \dots, N\}^2.$$

avec les questions 2 et 4, on a donc, pour le schéma [I], si  $h \leq \frac{2}{k}$  :

$$\max\{|e_{ij}|, (i,j) \in \{1, \dots, N\}^2\} \leq \frac{1}{2} C_1 h^2,$$

où  $C_1$  et  $C_2$  ne dépendent que de  $u$  et  $k$  (et sont données à la question 2). Les 2 schémas sont convergents. Le schéma [I] converge en " $h^2$ " et le schéma [II] en " $h$ ".

6. Le schéma [I] converge plus vite mais a une condition de stabilité  $k \leq \frac{2}{h}$ . Le schéma [II] est inconditionnellement stable.

## Corrigé de l'exercice 15 page 54

1. On a vu au paragraphe 2.3.4 page 27 que si  $\sigma$  est une arête du volume de contrôle  $K$ , alors le flux numérique  $F_{K,\sigma}$  s'écrit :

$$F_{K,\sigma} = \lambda_i \frac{u_\sigma - u_K}{d_{K,\sigma}} m(\sigma).$$

On cherche à éliminer les inconnues auxiliaires  $u_\sigma$ . Pour cela, si  $\sigma$  est une arête interne,  $\sigma = K|L$ , on écrit la conservativité du flux numérique :

$$F_{K,\sigma} = -F_{L,\sigma},$$

Ce qui entraîne, si  $\sigma$  n'est pas une arête de l'interface  $I$ , que :

$$-\lambda_i \frac{u_\sigma - u_K}{d_{K,\sigma}} m(\sigma) = \lambda_i \frac{u_\sigma - u_L}{d_{L,\sigma}} m(\sigma)$$

On en déduit que

$$u_\sigma \left( \frac{1}{d_{K,\sigma}} + \frac{1}{d_{L,\sigma}} \right) = \frac{u_K}{d_{K,\sigma}} + \frac{u_L}{d_{L,\sigma}},$$

soit encore que

$$u_\sigma = \frac{d_{K,\sigma} d_{L,\sigma}}{d_\sigma} \left( \frac{u_K}{d_{K,\sigma}} + \frac{u_L}{d_{L,\sigma}} \right).$$

Remplaçons alors dans (2.7). On obtient :

$$\begin{aligned} F_{K,\sigma} &= \lambda_i \left( \frac{d_{L,\sigma}}{d_\sigma} \left( \frac{u_K}{d_{K,\sigma}} + \frac{u_L}{d_{L,\sigma}} \right) - \frac{u_K}{d_{K,\sigma}} \right) \\ &= -\frac{\lambda_i}{d_\sigma} \left( \frac{d_{L,\sigma}}{d_{K,\sigma}} u_K + u_L - u_K - \frac{d_{L,\sigma}}{d_{K,\sigma}} u_K \right) \end{aligned}$$

On obtient donc finalement bien la formule (2.3.49).

2. Considérons maintenant le cas d'une arête  $\sigma \subset \Gamma_1 \cup \Gamma_3$ , où l'on a une condition de Fourier, qu'on a discrétisée par :

$$F_{K,\sigma} = -m(\sigma) \lambda_i \frac{u_\sigma - u_K}{d_{K,\sigma}} = m(\sigma) \alpha (u_\sigma - u_{ext}).$$

On a donc

$$u_\sigma = \frac{1}{\alpha + \frac{\lambda_i}{d_{K,\sigma}}} \left( \frac{\lambda_i u_K}{d_{K,\sigma}} + \alpha u_{ext} \right)$$

On remplace cette expression dans l'égalité précédente. Il vient :

$$F_{K,\sigma} = \frac{m(\sigma) \alpha}{\alpha + \frac{\lambda_i}{d_{K,\sigma}}} \left( \frac{\lambda_i}{d_{K,\sigma}} u_K + \alpha u_{ext} - \alpha u_{ext} - \frac{\lambda_i}{d_{K,\sigma}} u_{ext} \right),$$

Ce qui, après simplifications, donne exactement (2.3.50).

3. Considérons maintenant une arête  $\sigma = K|L$  appartenant à l'interface  $I$ . La discréétisation de la condition de saut de flux sur  $I$ . S'écrit :

$$F_{K,\sigma} + F_{L,\sigma} = \int_{\sigma} \theta(x) d\gamma(x) = m(\sigma) \theta_{\sigma}$$

Supposons que  $K$  (resp.  $L$ ) soit situé dans le milieu de conductivité (resp.  $\lambda_2$ ). En remplaçant  $F_{K,\sigma}$  et  $F_{L,\sigma}$  par leurs expressions, on obtient :

$$-\lambda_i m(\sigma) \frac{u_{\sigma} - u_K}{d_{K,\sigma}} - \lambda_2 m(\sigma) \frac{u_{\sigma} - u_L}{d_{L,\sigma}} = m(\sigma) \theta_{\sigma}.$$

On en déduit que

$$u_{\sigma} \left( \frac{\lambda_1}{d_{K,\sigma}} + \frac{\lambda_2}{d_{L,\sigma}} \right) = \left( \frac{\lambda_1 u_K}{d_{K,\sigma}} + \frac{\lambda_2 u_L}{d_{L,\sigma}} - \theta_{\sigma} \right).$$

En remplaçant  $u_{\sigma}$  dans l'expression de  $F_{K,\sigma}$ , on obtient :

$$F_{K,\sigma} = -\frac{m(\sigma)}{d_{K,\sigma}} \lambda_1 \frac{1}{\frac{\lambda_1}{d_{K,\sigma}} + \frac{\lambda_2}{d_{L,\sigma}}} \left( \frac{\lambda_1 u_K}{d_{K,\sigma}} + \frac{\lambda_2 u_L}{d_{L,\sigma}} - \theta_{\sigma} - \frac{\lambda_1 u_K}{d_{K,\sigma}} - \frac{\lambda_2 u_K}{d_{L,\sigma}} \right).$$

En simplifiant, on obtient :

$$F_{K,\sigma} = -\frac{m(\sigma) \lambda_1}{\lambda_1 d_{L,\sigma} + \lambda_2 d_{K,\sigma}} (\lambda_2 u_L - \lambda_2 u_K - d_{L,\sigma} \theta_{\sigma}),$$

ce qui est exactement (2.3.52). On obtient alors l'expression de  $F_{L,\sigma}$  :

$$F_{L,\sigma} = m(\sigma) \theta_{\sigma} - F_{K,\sigma},$$

ce qui donne, après simplifications :

$$F_{L,\sigma} = \frac{\lambda_2 m(\sigma)}{\lambda_1 d_{L,\sigma} + \lambda_2 d_{K,\sigma}} [\lambda_1 (u_L - u_K) + d_{K,\sigma} \theta_{\sigma}].$$

On vérifie bien que  $F_{K,\sigma} + F_{L,\sigma} = m(\sigma) \theta_{\sigma}$ .

4. Le système linéaire que satisfont les inconnues  $(u_K)_{K \in \mathcal{M}}$  s'écrit

$$AU = b$$

avec  $U = (u_K)_{K \in \mathcal{T}}$ . Pour construire les matrices  $A$  et  $b$ , il faut se donner une numérotation des mailles. On suppose qu'on a  $n \times 2p$  mailles ; on considère un maillage uniforme du type de celui décrit sur la figure 2.3 page 28 et on note  $h_x = \frac{1}{n}$  (resp.  $h_y = \frac{1}{p}$ ) la longueur de la maille dans la direction  $x$  (resp.  $y$ ). Comme le maillage est cartésien, il est facile de numérotter les mailles dans l'ordre “lexicographique” ; c'est-à-dire que la  $k$ -ième maille a comme centre le point  $x_{i,j} = ((i - \frac{1}{2})h_x, (j - \frac{1}{2})h_y)$ , avec  $k = n(j-1) + i$ . On peut donc déterminer le numéro de la maille (et de l'inconnue associée)  $k$  à partir de la numérotation cartésienne  $(i, j)$  de la maille.

$$k = n(j-1) + i$$

Remarquons que, comme on a choisi un maillage uniforme, on a pour tout  $K \in \mathcal{T}$  :  $m(K) = h_x h_y$ , pour toute arête intérieure verticale  $\sigma$  :  $d_{\sigma} = h_x m(\sigma) = h_y$  et pour toute arête intérieure horizontale,

$d_\sigma = h_y$  et  $m(\sigma) = h_x$ . Pour chaque numéro de maille, nous allons maintenant construire l'équation correspondante.

Mailles internes  $i = 2, \dots, n-1; j = 2, \dots, p-1, p+1, \dots, 2p-1$ .

L'équation associée à une maille interne  $K$  s'écrit

$$\sum_{\sigma \in \mathcal{E}_K} F_{K,\sigma} = m(K)f_K.$$

Avec l'expression de  $F_{K,\sigma}$  donnée par (2.3.49), ceci amène à :

$$2\lambda_m \left( \frac{h_x}{h_y} + \frac{h_y}{h_x} \right) u_k - \lambda_m \frac{h_x}{h_y} (u_{k-n} + u_{k+n}) - \lambda_m \frac{h_y}{h_x} (u_{k+1} + u_{k-1}) = h_x h_y f_k,$$

avec  $m = 1$  si  $j \leq p-1$  et  $m = 2$  si  $j \geq p+1$ .

Mailles du bord  $\Gamma_2$  Les mailles du bord  $\Gamma_2$  sont repérées par les indices  $(n, j), j = 2 \text{ à } p-1, j = p+1 \text{ à } 2p-1$ , (on exclut pour l'instant les coins).

L'équation des flux est la même que pour les mailles internes, mais le flux sur la frontière  $\Gamma_2$  est nul. Ceci donne :

$$\lambda_m \left( 2 \frac{h_x}{h_y} + \frac{h_y}{h_x} \right) u_k - \lambda_m \frac{h_x}{h_y} (u_{k-n} + u_{k+n}) - \lambda_m \frac{h_y}{h_x} u_{k-1} = h_x h_y f_k,$$

avec  $k = n(j-1) + n, j = 2 \text{ à } p-1, j = p+1 \text{ à } 2p-1$  et  $m = 1$  si  $j \leq p-1, m = 2$  si  $j \geq p+1$ .

Mailles de bord  $\Gamma_4$  Les mailles du bord  $\Gamma_4$  sont repérées par les indices  $(1, j), j = 2 \text{ à } p-1, j = p+1 \text{ à } 2p-1$ . Pour ces mailles, il faut tenir compte du fait que sur une arête de  $\Gamma_4$ , le flux  $F_{K,\sigma}$  est donné par :

$$F_{K,\sigma} = -\lambda_m \frac{g_\sigma - u_K}{d_{K,\sigma}} m(\sigma)$$

$$\text{avec } g_\sigma = \frac{1}{m(\sigma)} \int g(y) d\gamma(y).$$

D'où on tire l'équation relative à la maille  $k = n(j-1) + 1, j = 2, \dots, p-1, p+1, \dots, 2p-1$  :

$$\lambda_m \left( 2 \frac{h_x}{h_y} + 3 \frac{h_y}{h_x} \right) u_k - \lambda_m \frac{h_x}{h_y} (u_{k-n} + u_{k+n}) - \lambda_m \frac{h_y}{h_x} u_{k+1} = h_x h_y f_k + 2 \frac{h_y}{h_x} \lambda_m g_j,$$

avec  $g_j = g_{\sigma_j}$  et  $m = 1$  si  $j \leq p-1, m = 2$  si  $j \geq p+1$ .

Mailles du bord  $\Gamma_1 \cup \Gamma_3$  Pour  $j = 1$ , où  $j = 2p, i = 2 \dots n-1$ . On tient compte ici de la condition de Fourier sur la maille qui appartient au bord, pour laquelle l'expression du flux est :

$$F_{K,\sigma} = \frac{\alpha \lambda_m m(\sigma)}{\lambda_m + \alpha d_{K,\sigma}} (u_K - u_{ext}).$$

Pour une arête  $\sigma$  horizontale, on note :  $C_{F,\sigma} = \frac{\alpha m(\sigma)}{\lambda_m + \alpha d_{K,\sigma}}$ . Notons que  $C_{F,\sigma}$  est égal à

$$C_F = \frac{2\alpha h_x}{2\lambda_m + \alpha h_y}.$$

Notons que ce coefficient ne dépend pas de  $\sigma$ .

Les équations s'écrivent donc :

$$\lambda_1 \left( 2 \frac{h_y}{h_x} + \frac{h_x}{h_y} + C_F \right) u_k - \lambda_1 \frac{h_x}{h_y} u_{k+n} - \lambda_1 \frac{h_y}{h_x} (u_{k+1} + u_{k-1}) = h_x h_y f_k + \lambda_1 C_F u_{ext}, \\ k = 2, \dots, n-1,$$

$$\lambda_2 \left( \frac{2h_y}{h_x} + \frac{h_x}{h_y} + C_F \right) u_k - \lambda_2 \frac{h_x}{h_y} u_{k-n} - \lambda_2 \frac{h_y}{h_x} (u_{k+1} + u_{k-1}) = h_x h_y f_k + \lambda_2 C_F u_{ext}, \\ k = 2n(p-1) + 2, \dots, 2np-1,$$

Mailles des coins extérieurs : Il suffit de synthétiser les calculs déjà faits :

- coin sud-est :  $i = 1, j = 1, k = 1$  ; un bord Dirichlet, un bord Fourier :

$$\lambda_1 \left( \frac{3h_y}{h_x} + \frac{h_x}{h_y} + C_F \right) u_1 - \lambda_1 \frac{h_y}{h_x} u_2 - \lambda_1 \frac{h_x}{h_y} u_{n+1} = h_x h_y f_1 + \lambda_1 C_F u_{ext} + \frac{2h_y}{h_x} \lambda_1 g_1$$

- coin sud-ouest :  $i = 1n, j = 1, k = n$  ; un bord Fourier, un bord Neumann :

$$\lambda_1 \left( \frac{h_y}{h_x} + \frac{h_x}{h_y} + C_F \right) u_1 - \lambda_1 \frac{h_y}{h_x} u_{n-1} - \lambda_1 \frac{h_x}{h_y} u_{2n} = h_x h_y f_n + \lambda_1 C_F u_{ext}$$

- coin nord-ouest :  $i = 2n, j = 2p, k = 2np$ .

On a encore un bord Fourier, un bord Neumann, et l'équation s'écrit :

$$\lambda_2 \left( \frac{h_y}{h_x} + \frac{h_x}{h_y} + C_F \right) u_{2np} - \lambda_2 \frac{h_y}{h_x} u_{2np-1} - \lambda_2 \frac{h_x}{h_y} u_{2n(p-1)} = h_x h_y f_{2np} + \lambda_2 C_F u_{ext}$$

- coin nord-est :  $i = 1, j = 2p, k = n(2p-1) + 1$  un bord Dirichlet, un bord Fourier :

$$\lambda_2 \left( \frac{3h_y}{h_x} + \frac{h_x}{h_y} + C_F \right) u_k - \lambda_2 \frac{h_y}{h_x} u_{k+1} - \lambda_2 \frac{h_x}{h_y} (u_{k-n} = h_x h_y f_k + \lambda_2 C_F u_{ext}, + \frac{2h_y}{h_x} \lambda_2 g_k.$$

Interface L'expression du flux sur une arête de l'interface est donnée par (2.3.52). On pose, pour chaque arête  $\sigma$  de l'interface,

$$s_{I,\sigma} = \frac{m(\sigma)}{\lambda_1 d_{L,\sigma} + \lambda_2 d_{K,\sigma}}.$$

Notons que dans le cas du maillage uniforme considéré, ce coefficient est égal à :

$$s_I = \frac{2h_x}{(\lambda_1 + \lambda_2)h_y},$$

et qu'il est indépendant de l'arête  $\sigma$ . Tenant compte de ce flux, on obtient, pour  $k = n(p-1) + i, i = 2, \dots, N-1$

$$\lambda_1 \left( \frac{2h_y}{h_x} + \frac{h_x}{h_y} + \lambda_2 S_I \right) u_k - \lambda_1 \frac{h_y}{h_x} u_{k+1} - \lambda_1 \frac{h_y}{h_x} u_{k-1} - \lambda_1 \frac{h_x}{h_y} u_{k-n} - \lambda_1 S_I u_{k+n} = h_x h_y f_k + \lambda_1 S_I \frac{h_y}{2} \theta_i,$$

avec

$$\theta_i = \int_{\sigma_i} \theta(x) d\gamma(x).$$

Et de même, pour  $k = np + i, i = 2, \dots, N-1$ ,

$$\lambda_1 \left( \frac{2h_y}{h_x} + \frac{h_x}{h_y} + \lambda_1 S_I \right) u_k - \lambda_2 \frac{h_y}{h_x} u_{k+1} - \lambda_2 \frac{h_y}{h_x} u_{k-1} - \lambda_2 \frac{h_x}{h_y} u_{k+n} - \lambda_2 S_I u_{k-n} = h_x h_y f_k + \lambda_2 S_I \frac{h_y}{2} \theta_i.$$

Il ne reste plus qu'à traiter les coins des interfaces.

- $i = 1, j = p \quad k = n(p-1) + 1$ . Dirichlet sous l'interface

$$\lambda_1 \left( \frac{3h_y}{h_x} + \frac{h_x}{h_y} + \lambda_2 S_I \right) u_k - \lambda_1 \frac{h_y}{h_x} u_{k+1} - \lambda_1 \frac{h_x}{h_y} u_{k+n} - \lambda_1 S_I u_{k+n} = h_x h_y f_k + \lambda_1 S_I \frac{h_y}{2} \theta_i + \frac{2h_y}{h_x} \lambda_1 g_j$$

- $i = 1, j = p+1 \quad k = np+1$ , Dirichlet, dessus de l'interface

$$\lambda_2 \left( \frac{3h_y}{h_x} + \frac{h_x}{h_y} + \lambda_1 S_I \right) u_k - \lambda_2 \frac{h_y}{h_x} u_{k+1} - \lambda_2 \frac{h_x}{h_y} u_{k+n} - \lambda_2 S_I u_{k-n} = h_x h_y f_k + \lambda_2 S_I \frac{h_y}{2} \theta_i + \frac{2h_y}{h_x} \lambda_2 g_j$$

- $i = n, j = p, k = n(p-1) + n$ . Neumann, sous l'interface.

$$\lambda_1 \left( \frac{h_y}{h_x} + \frac{h_x}{h_y} + \lambda_2 S_I \right) u_k - \lambda_1 \frac{h_y}{h_x} u_{k-1} - \lambda_1 \frac{h_x}{h_y} u_{k-n} - \lambda_1 S_I u_{k+n} = h_x h_y f_k + \lambda_1 S_I \frac{h_y}{2} \theta_i$$

- $i = n, j = p+1, k = np+n$ , Neuman, dessus de l'interface

$$\lambda_2 \left( \frac{h_y}{h_x} + \frac{h_x}{h_y} + \lambda_1 S_I \right) u_k - \lambda_2 \frac{h_y}{h_x} u_{k-1} - \lambda_2 \frac{h_x}{h_y} u_{k+n} - \lambda_2 S_I u_{k-n} = h_x h_y f_k + \lambda_2 S_I \frac{h_y}{2} \theta_i.$$

On a ainsi obtenu  $2np$  équations à  $2np$  inconnues. Notons que chaque équation fait intervenir au plus 5 inconnues.

## Corrigé de l'exercice 16 page 54

1. Le problème complet s'écrit :

$$\begin{cases} -\operatorname{div}(\mu_i \nabla \phi)(x) = 0, & x \in \Omega_i, \quad i = 1, 2 \\ \nabla \phi(x) \cdot \mathbf{n}(x) = 0, & x \in \Gamma_2 \cup \Gamma_4, \\ \int_{\Gamma_1} \mu_1 \nabla \phi(x) \cdot \mathbf{n}(x) d\gamma(x) + \int_{\Gamma_3} \mu_2 \nabla \phi(x) \cdot \mathbf{n}(x) d\gamma(x) = 0, \\ \phi_2(x) - \phi_1(x) = 0, & \forall x \in I, \\ -(\mu \nabla \phi \cdot \mathbf{n})|_2(x) - (\mu \nabla \phi \cdot \mathbf{n})|_1(x) = 0, & \forall x \in I. \end{cases}$$

2. On se donne le même maillage rectangulaire uniforme qu'à l'exercice précédent. On note  $\phi_K$  l'inconnue associée à la maille  $K$  (ou  $\phi_k$  si on la référence la maille  $K$  par son numéro  $k = n(j-1)+i$ , où  $i \in \{1, \dots, n\}$  et  $j \in \{1, \dots, 2p\}$ ). Pour une maille intérieure, l'équation obtenue est la même que (2.7) en remplaçant  $\lambda_m$  par  $\mu_m$ .

Etudions maintenant le cas d'une maille proche de l'interface. Comme indiqué, on va considérer deux inconnues discrètes par arête de l'interface. Soient  $K$  et  $L$  ayant en commun l'arête  $\tilde{\sigma} \subset I$ ,  $K$  est située au dessous de  $L$ . Les équations associées à  $K$  et  $L$  s'écrivent alors

$$\sum_{\sigma \in \xi_K} F_{K,\sigma} = 0 \text{ et } \sum_{\sigma \in \xi_L} F_{L,\sigma} = 0.$$

Pour les arêtes  $\sigma \in \xi_K$  autres que  $\tilde{\sigma}$ , le flux s'écrit de manière habituelle

$$F_{K,\sigma} = \mu_1 \frac{\phi_K - \phi_M}{d_\sigma}, \text{ avec } \sigma = K|M.$$

Pour l'arête  $\sigma = \tilde{\sigma}$ , on a  $F_{K,\sigma} = \mu_1 \frac{\phi_K - \phi_\sigma}{d_\sigma} m(\sigma)$  et  $F_{L,\sigma} = \mu_2 \frac{\phi_L - \phi_\sigma^+}{d_\sigma} m(\sigma)$ , où les deux inconnues discrètes  $\phi_\sigma^+$  et  $\phi_\sigma^-$  sont reliées par les relations :

$$\phi_\sigma^+ - \phi_\sigma^- = \psi_\sigma \left( = \frac{1}{m(\sigma)} \int_\sigma \psi(x) d\gamma(x) \right)$$

$$F_{K,\sigma} + F_{L,\sigma} = 0.$$

On peut alors éléminer  $\phi_\sigma^+$  et  $\phi_\sigma^-$ ; en utilisant par exemple  $\phi_\sigma^+ = \psi_\sigma + \phi_\sigma^-$  et en remplaçant dans la deuxième équation, on obtient :

$$-\mu_1 \frac{\phi_\sigma^- - \phi_K}{d_{K,\sigma}} + \mu_2 \frac{\phi_\sigma^- + \psi_\sigma - \phi_L}{d_{L,\sigma}} = 0,$$

ce qui donne :

$$\phi_{\sigma^-} = \frac{1}{\frac{\mu_1}{d_{K,\sigma}} + \frac{\mu_2}{d_{L,\sigma}}} \left( \frac{\mu_1}{d_{K,\sigma}} \phi_K + \frac{\mu_2}{d_{L,\sigma}} \phi_L - \frac{\mu_2}{d_{L,\sigma}} \psi_\sigma \right).$$

En remplaçant cette expression dans les flux, on obtient :

$$F_{K,\sigma} = -F_{L,\sigma} = m(\sigma) \frac{\mu_1 \mu_2}{\mu_1 d_{L,\sigma} + \mu_2 d_{K,\sigma}} (\phi_K - \phi_L + \psi_\sigma)$$

On peut alors écrire l'équation discrète associée à une maille de numéro  $k$  située sous l'interface (avec  $k = n(p-1) + i$ ,  $i = 2, \dots, n-1$ ). On pose :

$$\frac{\mu_1 \mu_2}{\mu_1 d_{L,\sigma} + \mu_2 d_{K,\sigma}} = \frac{\mu_I}{d_\sigma}$$

( $\mu_I$  est donc la moyenne harmonique pondérée entre  $\mu_1$  et  $\mu_2$ ). Notons que pour une arête de  $I$ ,  $d_\sigma = h_y$ , et  $m(\sigma) = h_x$ . L'équation associée à la maille  $k$  s'écrit donc :

$$\left( 2\mu_1 \frac{h_y}{h_x} + \mu_1 \frac{h_x}{h_y} + \frac{\mu_I h_x}{h_y} \right) u_k - \mu_1 \frac{h_y}{h_x} (u_{k-1} + u_{k+1}) - \mu_1 \frac{h_x}{h_y} u_{k-n} - \mu_I \frac{h_x}{h_y} u_{k+n} = -\mu_I \frac{h_x}{h_y} \psi_i,$$

où  $\psi_i$  est le saut de potentiel à travers l'arête  $\sigma_i$  de l'interface considérée ici. De même, l'équation associée à une maille  $k$  avec  $k = np + i$ ,  $i = 2, \dots, n-1$ , située au dessus de l'interface s'écrit :

$$\left( 2\mu_1 \frac{h_y}{h_x} + \mu_1 \frac{h_x}{h_y} + \mu_I \frac{h_x}{h_y} \right) u_k - \mu_1 \frac{h_y}{h_x} (u_{k-1} + u_{k+1}) - \mu_1 \frac{h_x}{h_y} u_{k+n} - \mu_I \frac{h_x}{h_y} u_{k-n} = +\mu_I \frac{h_x}{h_y} \psi_i.$$

La discrétisation des conditions aux limites de Neumann sur  $\Gamma_2$  et  $\Gamma_4$  est effectuée de la même manière que pour le cas du problème thermique, voir exercice 15.

Il ne reste plus qu'à discrétiser la troisième équation du problème (2.7), qui relie les flux sur la frontière  $\Gamma_1$  avec les flux sur la frontière  $\Gamma_3$ . En écrivant la même condition avec les flux discrets, on obtient :

$$\mu_1 \sum_{i=1}^n \frac{2h_x}{h_y} (u_i - u_{B,i}) + \mu_2 \sum_{i=1}^n \frac{2h_x}{h_y} (u_{H,i} - u_{k(i)}) = 0,$$

où :  $\mu_{B,i}$  représente l'inconnue discrète sur la  $i$ -ème arête de  $\Gamma_1$  et  $\mu_{H,i}$  l'inconnue discrète sur la  $i$ -ème arête de  $\Gamma_3$ , et  $k(i) = n(p-1) + i$  est le numéro de la maille jouxtant la  $i$ -ème arête de  $\Gamma_3$ . Remarquons que tel qu'il est posé, le système n'est pas inversible : on n'a pas assez d'équations pour éliminer les inconnues  $u_{B,i}$  et  $u_{H,i}, i = 1 \dots N$ . On peut par exemple pour les obtenir considérer une différence de potentiel fixée entre  $\Gamma_1$  et  $\Gamma_3$ , et se donner un potentiel fixé sur  $\Gamma_1$ .

## Corrigé de l'exercice 17 page 55

On note  $\|\cdot\|_2 = \|\cdot\|_{L^2([0,1])}$ .

1) Pour  $n \in \mathbb{N}^*$ , on a

$$\int_0^1 \sin^2(n\pi x) dx = \int_0^1 \frac{1 - \cos(2n\pi x)}{2} dx = \frac{1}{2},$$

et

$$\int_0^1 |u_0(x) \sin(n\pi x)| dx \leq \|u_0\|_2 \left( \int_0^1 \sin^2(n\pi x) dx \right)^{1/2} = \frac{r_2}{2} \|u_0\|_2.$$

La quantité  $a_n$  est donc bien définie et

$$|a_n| \leq r_2 \|u_0\|_2$$

Pour tout  $t > 0$  et  $x \in [0, 1]$ , on a

$$|e^{-n^2\pi^2t^2} a_n \sin(n\pi x)| \leq r_2 \|u_0\|_2 e^{-n^2\pi^2t^2} \quad \forall n \in \mathbb{N}^*.$$

Ceci montre que la série  $\sum_{n>0} e^{-n^2\pi^2t^2} a_n \sin(n\pi x)$  est absolument convergente et donc que  $u$  est bien définie

pour tout  $t > 0$  et tout  $x \in [0, 1]$  et même pour tout  $x \in \mathbb{R}$ .

On remarque ensuite que  $u$  est de classe  $C^\infty$  sur  $\mathbb{R} \times \mathbb{R}_+^*$ , en appliquant les théorèmes classiques de dérivation terme à terme d'une série. En effet, soit  $\varepsilon > 0$ , pour tout  $x \in \mathbb{R}$  et  $t > \varepsilon$  on a

$$\left| e^{-n^2\pi^2t^2} a_n \sin(n\pi x) \right| \leq r_2 \|u_0\|_2 e^{-n^2\pi^2\varepsilon^2}, \forall n \in \mathbb{N}^*$$

Comme  $(x, t) \rightarrow e^{-n^2\pi^2t^2} a_n \sin(n\pi t)$  est continue (pour tout  $n \in \mathbb{N}^*$ ), on en déduit que  $u$  est continue sur  $\mathbb{R} \times ]\varepsilon, \infty[$ , et finalement sur  $\mathbb{R} \times ]0, \infty[$  car  $\varepsilon > 0$  est arbitraire.

Pour dériver terme à terme la série définissant  $u$ , il suffit également d'obtenir sur  $\varepsilon, \infty[ \times \mathbb{R}$  (pour tout  $\varepsilon > 0$ ) une majoration du terme général de la série des dérivées par le terme général d'une série convergente (indépendant de  $(x, t) \in \mathbb{R} \times ]\varepsilon, \infty[$ ). On obtient cette majoration en remarquant que, pour  $(x, t) \in \mathbb{R} \times ]\varepsilon, \infty[$ ,

$$|-n^2\pi^2e^{-n^2\pi^2t^2} a_n \sin(n\pi x)| \leq n^2\pi^2e^{-n^2\pi^2\varepsilon^2} r_2 \|u_0\|_2$$

On montre ainsi finalement que  $u$  est de classe  $C^1$  par rapport à  $t$  et que

$$u_t(x, t) = \sum_{n>0} -n^2\pi^2e^{-n^2\pi^2t^2} a_n \sin(n\pi x), x \in \mathbb{R}, t > 0.$$

En itérant ce raisonnement on montre que  $u$  est de classe  $C^\infty$  par rapport à  $t$  sur  $\mathbb{R} \times \mathbb{R}_+^*$ .

Un raisonnement similaire montre que  $u$  est de classe  $C^\infty$  par rapport à  $x$  sur  $\mathbb{R} \times \mathbb{R}_+^*$  et que l'on peut dériver terme à terme la série définissant  $u$ . On obtient donc aussi

$$u_{xx}(xt) = \sum_{n>0} -n^2\pi^2e^{-n^2\pi^2t^2} a_n \sin(n\pi x), x \in \mathbb{R}, t > 0,$$

et ceci donne  $u_t = u_{xx}$  sur  $\mathbb{R} \times \mathbb{R}_t^*$  et donc aussi un  $[0, 1] \times \mathbb{R}_t^*$ . Le fait que  $u(0, t) = u(1, t)$  pour tout  $t > 0$  est immédiat car  $\sin n\pi t = \sin 0 = 0$ , pour tout  $n \in \mathbb{N}^*$ .

Il reste à montrer que  $u(., t) \rightarrow u_0$  dans  $L^2([0, 1])$  quand  $t \rightarrow 0$ .

On définit  $e_n \in L^2([0, 1])$  par  $e_n(x) = \sqrt{2} \sin(n\pi x)$ . La famille  $\{e_n, n \in \mathbb{N}^*\}$  est une base hilbertienne de  $L^2([0, 1])$ .

On a donc :

$$\sum_{n=1}^N a_n \sin n\pi x \rightarrow u_0, \text{ dans } L^2([0, 1]), \text{ quand } n \rightarrow \infty,$$

et

$$\sum_{n=1}^{\infty} a_n^2 = 2\|u_0\|_2^2.$$

On remarque maintenant que

$$u(x, t) - u_0(x) = u(x, t) - u^{(N)}(x, t) + u^{(N)}(x, t) - u_0^{(N)}(x) - u_0^{(N)}(x) + u_0(x),$$

avec

$$\begin{aligned} u^{(N)}(x, t) &= \sum_{n=1}^N a_n e^{-n^2 \pi^2 t^2} \sin(n\pi x) \\ u_0^{(N)}(x) &= \sum_{n=1}^N a_n \sin(n\pi x). \end{aligned}$$

Il est clair que, pour tout  $N \in \mathbb{N}^*$ , on a  $u^{(N)}(., t) \rightarrow u_0^{(N)}$  uniformément sur  $\mathbb{R}$ , quand  $N \rightarrow \infty$ , et donc  $u^{(N)}(., t) \rightarrow u_0^{(N)}$  dans  $L^2([0, 1])$ .

Comme

$$\|u(., t) - u^{(N)}(., t)\|_2^2 = \sum_{n=N+1}^{\infty} a_n^2 \frac{1}{2} e^{-2n^2 \pi^2 t^2} \leq \sum_{n=N+1}^{\infty} a_n^2 \frac{1}{2} = \|u_0^{(N)} - u_0\|_2^2 \rightarrow 0$$

quand  $N \rightarrow \infty$ , on en déduit que  $u(., t) \rightarrow u_0$ ,  $qdt \rightarrow 0$ , dans  $L^2([0, 1])$ .

2) On note  $w$  la différence de 2 solutions de (2.5.105). On a donc

$$\begin{aligned} w &\in C^\infty([0, 1] \times \mathbb{R}_+^*, \mathbb{R}) \\ w_t - w_{xx} &= 0 \text{ sur } [0, 1] \times \mathbb{R}_+^* \\ w(0, t) &= w(1, t) = 0 \text{ pour } t > 0 \\ w(., t) &\rightarrow 0, \text{ dans } L^2([0, 1]), \text{ quand } t \rightarrow 0 \end{aligned}$$

Soit  $0 < \varepsilon < T < \infty$ . On intègre l'équation  $ww_t - ww_{xx} = 0$  sur  $[0, 1] \times [\varepsilon, T]$ . En utilisant une intégration par parties (noter que  $w \in C^\infty([0, 1] \times [\varepsilon, T])$ ), on obtient :

$$\frac{1}{2} \int_0^1 w^2(x, T) dx - \frac{1}{2} \int_0^1 w^2(x, \varepsilon) dx + \int_0^1 \int_\varepsilon^T w_x^2(x, t) dx dt = 0.$$

D'où l'on déduit  $\|w(., T)\|_2 \leq \|w(., \varepsilon)\|_2$ . Quand  $\varepsilon \rightarrow 0$ , on a  $\|w(., \varepsilon)\|_2 \rightarrow 0$ , on a donc  $\|w(., T)\|_2 = 0$  et donc, comme  $w(., t)$  est contenue sur  $[0, 1]$ ,  $w \in [0, 1]$ . Comme  $T > 0$  est arbitraire, on a finalement

$$w(x, t) = 0 \quad \forall t \in [0, 1], \forall t > 0$$

Ce qui montre bien l'unicité de la solution de (2.5.105).

## Corrigé de l'exercice 18 page 55

1) La formule pour calculer  $u_i^0$  est :

$$u_1^0 = u_0(ih, 0), \quad i = -\frac{N+1}{2}, \dots, \frac{N+1}{2}$$

Soit maintenant  $n \in \{0, \dots, M\}$ . On a :

$$\begin{aligned} u_i^{n+1} &= 0 \quad \text{pour} \quad i = -\frac{N+1}{2} \quad \text{et} \quad i = \frac{N+1}{2} \\ u_i^{n+1} &= u_i^n + \frac{k}{h^2} (u_{i+1}^n + u_{i-1}^n - 2u_i^n), \quad i = -\frac{N+1}{2} + 1, \dots, \frac{N+1}{2} - 1. \end{aligned}$$

2) On va montrer, par récurrence (finie) sur  $n$ , que :

$$\text{Si } n \in \{0, \dots, M+1\}, i \in \left\{-\frac{N+1}{2}, \dots, \frac{N+1}{2}\right\} \text{ et } i \geq -\frac{N+1}{4} + n \text{ alors } u_i^n = 0. \quad (2.7.145)$$

Pour initialiser la récurrence, on suppose que  $n = 0$  et  $i \geq -\frac{N+1}{4}$ . On a alors

$$ih \geq -\frac{N+1}{4} \quad \frac{8}{N+1} = -2 > -3$$

et donc  $u_i^0 = 0$ .

Soit maintenant  $n \in \{0, \dots, M\}$ . On suppose que l'hypothèse de récurrence est vérifiée jusqu'au rang  $n$ , et on démontre la propriété au rang  $n+1$ . Soit donc  $i \in \left\{-\frac{N+1}{2}, \dots, \frac{N+1}{2}\right\}$  tel que  $i \geq -\frac{N+1}{4} + (n+1)$ .

Alors :

- Si  $i = \frac{N+1}{2}$  on a bien  $u_i^{N+1} = 0$ .
- Si  $i < \frac{N+1}{2}$ , les indices  $i-1, i$  et  $i+1$  sont tous supérieurs ou égaux à  $-\frac{N+1}{4} + n$ , et donc par hypothèse de récurrence,

$$u_i^{n+1} = u_i^n \left(1 - \frac{2k}{h^2}\right) + \frac{k}{h^2} u_{i+1}^n + \frac{k}{h^2} u_{i-1}^n = 0.$$

On a donc bien démontré (2.7.145). On utilise maintenant l'hypothèse  $k = h$ , c'est-à-dire  $\frac{1}{M+1} = \frac{8}{N+1}$ .  
On a alors

$$-\frac{N+1}{4} + M + 1 = -2(M+1) + M + 1 = -(M+1) < 0.$$

On en déduit que si  $n \in \{0, \dots, M+1\}$  et  $i \geq 0$ , alors  $i \geq -\frac{N+1}{4} + n$ . On en déduit que  $u_i^n = 0$  pour  $n \in \{0, \dots, M+1\}$  et  $i \in \left\{0, \dots, \frac{N+1}{2}\right\}$ . On remarque alors que

$$\begin{aligned} \max \left\{ |u_i^{M+1} - \bar{u}_i^{M+1}|, i \in \left\{-\frac{N+1}{2}, \dots, \frac{N+1}{2}\right\} \right\} &\geq !la \max \left\{ |\bar{u}_i^{M+1}|, i \in \left\{0, \dots, \frac{N+1}{2}\right\} \right\} \\ &\geq \inf_{[0,4]} u(x, 1) > 0, \end{aligned}$$

et donc ne tend pas vers 0 quand  $h \rightarrow 0$ .

## Corrigé de l'exercice 19 page 56

1. On admettra que la solution de (2.5.107) existe et qu'elle est assez régulière. Soient  $M \in \mathbb{N}^*$  et  $N \in \mathbb{N}^*$ , et soient  $k$  le pas de temps, choisi tel que  $Mk = T$  et  $h$  le pas espace, choisi tel que  $Nh = 1$ . On applique un schéma d'Euler explicite en temps, et un schéma de différences finies centré en espace, on obtient donc :

$$u_j^{n+1} = k \left[ \frac{1}{k} u_j^n - \frac{1}{2h} (u_{j+1}^n - u_{j-1}^n) + \frac{\varepsilon}{h^2} (u_{j+1}^n + u_{j-1}^n - 2u_j^n) \right] \quad (2.7.146)$$

On tient compte des conditions aux limites et des conditions initiales en posant :

$$\begin{cases} u_0^n = u_{N+1}^n = 0, \\ u_j^0 = u_0(jh). \end{cases}$$

On a, par développement de Taylor :

$$u(x+h, t) = u(x, t) + hu_x(x, t) + \frac{h^2}{2}u_{xx}(x, t) + \frac{h^3}{6}u^{(3)}(x, t) + \frac{h^4}{24}u^{(4)}(\alpha, t),$$

$$u(x-h, t) = u(x, t) - hu_x(x, t) + \frac{h^2}{2}u_{xx}(x, t) - \frac{h^3}{6}u''''(x, t) + \frac{h^4}{24}u^{(4)}(\beta, t)$$

et

$$u(x, t+k) = u(x, t) + ku_t(x, t) + \frac{k^2}{2}u_{tt}(x, \tau_k), \quad \tau_k \in [t, t+k].$$

De ces développements de Taylor, il ressort que l'erreur de consistance vérifie  $|R| \leq C(k + h^2)$ , où  $C$  ne dépend que de  $u$ . Le schéma est donc explicite d'ordre 1 en temps et 2 en espace.

Cherchons alors les conditions pour que :

$$\|u^n\|_\infty \leq \|u^0\|_\infty.$$

Par définition,

$$\|u^n\|_\infty = \max_{j=1, \dots, N} |u_j^n|.$$

On essaye d'abord de vérifier que :  $\|u^{n+1}\|_\infty \leq \|u^n\|_\infty$ , c'est-à-dire :

$$\max_{j=1, \dots, N} |u_j^{n+1}| \leq \max_{j=1, \dots, N} |u_j^n|,$$

On veut donc montrer que

$$\begin{cases} \max_{j=1, \dots, N} u_j^{n+1} \leq \max_{j=1, \dots, N} u_j^n, \\ \min_{j=1, \dots, N} u_j^{n+1} \geq \min_{j=1, \dots, N} u_j^n. \end{cases}$$

On peut réécrire le schéma (2.7.146) :

$$u_j^{n+1} = u_j^n \left( 1 - \frac{2\varepsilon k}{h^2} \right) + u_{j+1}^n \left( -\frac{k}{2h} + \frac{k\varepsilon}{h^2} \right) + u_{j-1}^n \left( \frac{\varepsilon k}{h^2} + \frac{k}{2h} \right).$$

Posons :

$$M^n = \max_{j=1, \dots, N} u_j^n$$

Supposons que  $k$  et  $h$  vérifient :

$$1 \geq \frac{2\varepsilon k}{h^2} \text{ et } \frac{k\varepsilon}{h^2} - \frac{k}{2h} \geq 0,$$

ce qui s'écrit encore :

$$\begin{cases} \frac{k}{h^2} \leq \frac{1}{2\varepsilon} \\ k \leq \frac{2\varepsilon}{h}, \end{cases} \quad (2.7.147)$$

on a alors :

$$u_j^{n+1} \leq M^n \left( 1 - \frac{2\varepsilon k}{h^2} \right) + M^n \left( -\frac{k}{2h} + \frac{k\varepsilon}{h^2} \right) + M^n \left( \frac{\varepsilon k}{h^2} + \frac{k}{2h} \right) \quad \forall j = 1, \dots, N,$$

et donc :

$$M^{n+1} \leq M^n.$$

Posons maintenant :

$$m^n = \min_{j=1 \dots N} u_j^n.$$

Si  $k$  et  $h$  satisfont les conditions (2.7.147), on obtient de la même manière

$$m^{n+1} \geq m^n$$

On a ainsi montré que :

$$\|u^{n+1}\|_\infty \leq \|u^n\|_\infty.$$

On a de même :

$$\|u^n\|_\infty \leq \|u^{n-1}\|_\infty.$$

⋮

$$\|u^1\|_\infty \leq \|u^0\|_\infty.$$

En sommant ces inégalités, on obtient :

$$\|u^n\|_\infty \leq \|u^0\|_\infty.$$

Donc, sous les conditions (2.7.147), on a  $\|u^{n+1}\|_\infty \leq \|u^n\|_\infty$  et donc  $\|u^n\|_\infty \leq \|u^0\|_\infty$ , pour tout  $n = 1, \dots, N$ .

Pour étudier la convergence du schéma, on va tenter de majorer l'erreur de discréétisation :

$$e_j^n = \bar{u}_j^n - u_j^n,$$

où  $u_j^n$  est calculé par (2.7.146), et  $\bar{u}_j^n$  est la solution du problème (2.5.107) en  $x_j = jh$  et  $t_n = nk$ .  
On a donc, par définition de l'erreur de consistance,

$$\frac{1}{k}(\bar{u}_j^{n+1} - \bar{u}_j^n) + \frac{1}{2h}(\bar{u}_{j+1}^n - \bar{u}_{j-1}^n) - \frac{\varepsilon}{h^2}(-2\bar{u}_j^n + \bar{u}_{j+1}^n + \bar{u}_{j-1}^n) = R_j^n$$

où  $|R_j^n| \leq C(k + h^2)$

ce qui entraîne :

$$\frac{1}{k}(e_j^{n+1} - e_j^n) + \frac{1}{2h}(e_j^n - e_{j-1}^n) - \frac{\varepsilon}{h^2}(-2e_j^n + e_{j+1}^n + e_{j-1}^n) = R_j^n$$

soit encore :

$$e_j^{n+1} = \left(1 - \frac{2\varepsilon k}{h^2}\right)e_j^n + \left(-\frac{k}{2h} + \frac{k\varepsilon}{h^2}\right)e_j^{j+1} + \left(\frac{\varepsilon k}{h^2} + \frac{k}{2h}\right)e_{j-1}^n + kR_j^n.$$

de même que précédemment, on obtient sous les conditions (2.7.147)

$$\begin{aligned} |e_j^{n+1}| &\leq \|e^n\|_\infty + C(k + h^2)k \\ &\vdots \\ |e_j^n| &\leq \|e^{n-1}\|_\infty + C(k + h^2)k \\ &\vdots \\ |e_j^1| &\leq \|e^0\|_\infty + C(k + h^2)k. \end{aligned}$$

Et donc en sommant ces inégalités :

$$\|e^n\|_\infty \leq \|e^0\|_\infty + nCk(k + h^2)$$

Si à  $t = 0$  on a  $\|e^0\|_\infty = 0$ , alors :

$$\|e^{n+1}\|_\infty \leq CMk(k + h^2) = T(k + h^2).$$

Et donc sous les conditions (2.7.147) on a  $\|e^n\|_\infty$  qui tend vers 0 lorsque  $k, h \rightarrow 0$ , ce qui prouve que le schéma est convergent.

## Corrigé de l'exercice 20 page 56

1. Notons  $R_i^{(n)}$  l'erreur de consistance en  $(x_i, t_n)$ . Pour le schéma (2.5.110), on a donc par définition :

$$\begin{aligned} R_i^{(n)} &= \frac{\bar{u}_i^{(n+1)} - \bar{u}_i^{(n)}}{k} + \frac{1}{h^2}(2\bar{u}_i^{(n+1)} - \bar{u}_{i-1}^{(n+1)} - \bar{u}_{i+1}^{(n+1)}) - \bar{u}_i^{n+1} \\ &= \tilde{R}_i^{(n)} + \hat{R}_i^n, \end{aligned}$$

où

$$\tilde{R}_i^n = \frac{\bar{u}_i^{n+1} - \bar{u}_i^n}{k} - u_t(x_i, t_n) \text{ est l'erreur de consistance en temps}$$

et

$$\hat{R}_i^n = \frac{1}{h^2}(2\bar{u}_i^{n+1} - \bar{u}_{i-1}^{n+1} - \bar{u}_{i+1}^{n+1}) - (u_{xx}(x_i, t_n)) \text{ est l'erreur de consistance en espace.}$$

On a vu (voir (2.2.19)) que

$$|\hat{R}_i^n| \leq \frac{h^2}{12} \sup_{[0,1]} \left| \frac{\partial^4 u}{\partial x^4}(\cdot, t_n) \right|, \quad \forall i \in \{1, \dots, N\}$$

Effectuons maintenant un développement de Taylor en fonction du temps d'ordre 2 :

$$u(x_i, t_{n+1}) = u(x_i, t_n) + ku_t + \frac{k^2}{2}u_{tt}(x_i, \xi_n)$$

avec  $\xi_n \in [t_n, t_{n+1}]$ . Donc  $\frac{u(x_i, t_{n+1}) - u(x_i, t_n)}{k} - u_t = \frac{k}{2} u_{tt}(x_i, \xi_n)$ . Comme  $\xi_n \in [0, T]$ , et  $u_{tt}$  admet un maximum (à  $x_i$  fixé) dans  $[0, T]$  (qui est compact), on a donc

$$|\tilde{R}_i^n| \leq \frac{k}{2} \max_{[0, T]} |u_{tt}(x_i, \cdot)|.$$

Par conséquent,

$$|R_i^n| = |\tilde{R}_i^n + \hat{R}_i^n| \leq |\tilde{R}_i^n| + |\hat{R}_i^n| \leq \frac{k}{2} \max_{[0, T]} |u_{tt}(x_i, \cdot)| + \frac{h^2}{12} \max_{[0, 1]} \left| \frac{\partial^4 u}{\partial x^4}(\cdot, t_{n+1}) \right|.$$

Donc  $|R_i^n| \leq C(k + h^2)$  avec

$$C = \frac{1}{2} \max \left( \|u_{tt}\|_{L^\infty([0, 1] \times [0, T])}, \frac{1}{6} \left\| \frac{\partial^4 u}{\partial x^4} \right\|_{L^\infty([0, 1] \times [0, T])} \right).$$

Le calcul de l'erreur de consistance pour le schéma (2.5.111) s'effectue de manière semblable.

2. Le schéma (2.5.110) est complètement implicite alors que le schéma (2.5.111) ne l'est que partiellement, puisque le terme de réaction est pris à l'instant  $n$ . Le schéma (2.5.110) s'écrit :  $AU^{n+1} = U^n$  avec  $U^{n+1} = (U_1^{n+1}, \dots, U_N^{n+1})^t$ ,  $U^n = (U_1^n, \dots, U_N^n)^t$ , et

$$A = \begin{pmatrix} 1 + 2\lambda - k & -\lambda & 0 & \dots & 0 \\ -\lambda & 1 + 2\lambda - k & -\lambda & \ddots & 0 \\ 0 & & & & \\ \vdots & & & & 0 \\ 0 & 0 & -\lambda & 1 + 2\lambda - k & \end{pmatrix}$$

où  $\lambda = \frac{k}{h^2}$ . Notons que par définition,  $A$  est symétrique. De même, le schéma (2.5.111) s'écrit :  $BU^{n+1} = U^n$  avec

$$B = \frac{1}{1+k} \begin{pmatrix} 1 + 2\lambda & -1 & 0 & \dots & 0 \\ -1 & 1 + 2\lambda & & \ddots & 0 \\ 0 & & & & -1 \\ 0 & 0 & -1 & & 1 + 2\lambda \end{pmatrix}$$

On a donc  $A = \lambda A_h$ , où  $A_h$  est définie en (2.2.15) page 13, avec  $c_i = \frac{1-k}{\lambda}$ , et  $B = \frac{\lambda}{k+1} A_h$  avec  $c_i = \frac{1}{\lambda}$ . Dans les deux cas, les matrices sont donc s.d.p. en vertu de la proposition 2.2 page 13. Notons que l'hypothèse  $k \in ]0, 1[$  est nécessaire dans le cas du premier schéma, pour assurer la positivité de  $c_i$ .

3. Le schéma (2.5.111) s'écrit

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

On montre facilement par récurrence que  $\max_{j=1, \dots, N} u_j^n \leq (1+k)^n \max_{j=1, \dots, N} u_j^0$ , (voir preuve de la stabilité  $L^\infty$  d'Euler implicite page 44) et que  $\min_{j=1, \dots, N} u_j^{(n)} \geq (1+k)^n \min_{j=1, \dots, N} u_j^{(0)}$ . On en déduit que

$$\|u^{(n)}\|_\infty \leq (1+k)^n \|u_0\|_\infty$$

Or  $(1+k)^n \leq (1+k)^{T/k}$  car  $kn \leq T$ . Or

$$\begin{aligned}(1+k)^{T/k} &= \exp\left(\frac{T}{k}\ln(1+k)\right) \\ &\leq \exp\left(\frac{T}{k}k\right) = e^T\end{aligned}$$

On en déduit le résultat, avec  $C_1(T) = e^T$ . De même, pour le schéma (2.5.110), on montre par récurrence que :

$$\|u^{(n)}\|_\infty \leq \frac{1}{(1-k)^n} \|u^{(0)}\|.$$

Mais pour  $k \in ]0, \alpha[$ , avec  $\alpha \in ]0, 1[$ , on a :

$$\frac{1}{(1-k)} \leq 1 + \beta k, \quad \text{avec } \beta = \frac{1}{(1-\alpha)}.$$

On en déduit par un calcul similaire au précédent que

$$(1-k)^{T/k} \leq e^{\beta T},$$

d'où le résultat avec  $C_2(T, \alpha) = e^{\beta T}$ .

4. Par définition de l'erreur de consistance, on a pour le schéma (2.5.110)

$$\frac{\bar{u}_j^{(n+1)} - \bar{u}_j^n}{k} - \frac{\bar{u}_{j+1}^{(n+1)} + \bar{u}_{j-1}^{n+1} - 2\bar{u}_j^{(n+1)}}{h^2} - \bar{u}_j^{n+1} = R_i^{(n,1)}$$

et donc, en notant  $e_j^{(n)} = \bar{u}_j^n - u_j^{(n)}$  l'erreur de discréttisation en  $(x_j, t_n)$ , on a :

$$e_j^{(n+1)}(1 + 2\lambda - k) - \lambda e_{j-1}^{(n+1)} - \lambda e_{j+1}^{(n+1)} = e_j^{(n)} + kR_j^{(n,1)}$$

On obtient donc, de manière similaire à la question 3 :

(en considérant  $e_{j_0}^{(n+1)} = \max e_j^{(n+1)}$  puis  $e_{j_0}^{(n+1)} = \min e_j^{(n+1)}$ )

$$\frac{1}{1-k} \|e^{(n+1)}\| \leq \|e^{(n)}\| + kC(k + h^2).$$

Par récurrence sur  $n$ , on obtient alors

$$\|e^{(n)}\|_\infty \leq \left(\frac{1}{1-k}\right)^n [kC(k + h^2) + \|e^0\|_\infty]$$

d'où

$$\|e^{(n)}\|_\infty \leq C_2(T, \alpha)(TC(k + h^2) + \|e^0\|_\infty).$$

De même, pour le schéma (2.5.111), on écrit l'erreur de consistance :

$$\frac{\bar{u}_j^{(n+1)} - \bar{u}_j^n}{k} - \frac{\bar{u}_{j+1}^{(n+1)} + \bar{u}_{j-1}^{n+1} - 2\bar{u}_j^{(n+1)}}{h^2} - \bar{u}_j^n = R_j^{(n,2)}$$

et donc :

$$e_j^{(n+1)}(1 + 2\lambda) - \lambda e_{j-1}^{(n+1)} - \lambda e_{j+1}^{(n+1)} = e_j^{(n)}(1 + k) + kR_j^{(n,2)}.$$

Par des raisonnements similaires à ceux de la question 3 on obtient alors :

$$\|e_j^{(n)}\| \leq (1+k)^n (\|e^{(0)}\| + kC(k + h^2))$$

d'où

$$\|e^{(n)}\|_\infty \leq C_1(T) (\|e^{(0)}\| + kC(k + h^2)).$$

## Corrigé de l'exercice 21 page 57

1. On s'intéresse ici à l'ordre du schéma au sens des différences finies. On suppose que  $u \in C^4([0, 1] \times [0, T])$  est solution de (2.5.113) et on pose

$$\bar{u}_j^n = u(jh, nk), \quad j = 0, \dots, N, \quad k = 0, \dots, M.$$

L'erreur de consistance est définie par :

$$R_j^n = \frac{\bar{u}_j^{n+1} - \bar{u}_j^{n-1}}{2k} - \frac{\bar{u}_{j-1}^n - 2\bar{u}_j^n + \bar{u}_{j+1}^n}{h^2}, \quad j = 1, \dots, N-1, \quad k = 1, \dots, M-1.$$

On cherche une majoration de  $R_j^n$  en utilisant des développements de Taylor. Soit  $j \in \{1, \dots, N-1\}$ ,  $k \in \{1, \dots, M-1\}$ . Il existe  $(\xi_i, t_i) \in [0, 1] \times [0, T]$ ,  $i = 1, \dots, 4$ , t.q. :

$$\begin{aligned} \bar{u}_j^{n+1} &= \bar{u}_j^n + ku_t(jh, nk) + \frac{k^2}{2}u_{tt}(jh, nk) + \frac{k^3}{6}u_{ttt}(\xi_1, t_1), \\ \bar{u}_j^{n-1} &= \bar{u}_j^n - ku_t(jh, nk) + \frac{k^2}{2}u_{tt}(jh, nk) - \frac{k^3}{6}u_{ttt}(\xi_2, t_2), \\ \bar{u}_{j-1}^n &= \bar{u}_j^n - hu_x(jh, nk) + \frac{h^2}{2}u_{xx}(jh, nk) - \frac{h^3}{6}u_{xxx}(jh, nk) - \frac{h^4}{24}u_{xxxx}(\xi_3, t_3), \\ \bar{u}_{j+1}^n &= \bar{u}_j^n + hu_x(jh, nk) + \frac{h^2}{2}u_{xx}(jh, nk) + \frac{h^3}{6}u_{xxx}(jh, nk) + \frac{h^4}{24}u_{xxxx}(\xi_4, t_4). \end{aligned}$$

On en déduit :

$$R_j^n = u_t(jh, nk) + \frac{k^2}{12}(u_{ttt}(\xi_1, t_1) + u_{ttt}(\xi_2, t_2)) - u_{xx}(jh, nk) - \frac{h^2}{24}(u_{xxxx}(\xi_3, t_3) + u_{xxxx}(\xi_4, t_4)),$$

et donc, comme  $u$  est solution de (2.5.113),

$$|R_j^n| \leq C_1(k^2 + h^2),$$

où  $C_1$  ne dépend que de  $u$ . Le schéma (2.5.114) est donc consistant d'ordre 2.

2. Pour étudier la stabilité au sens de Von Neumann, on "oublie" les conditions aux limites dans (2.5.113). Plus précisément, on s'intéresse à (2.5.113) avec  $x \in \mathbb{R}$  (au lieu de  $x \in ]0, 1[$ ) et on remplace les conditions aux limites par des conditions de périodicité (exactement comme on l'a vu au paragraphe 2.4.2 page 39). Enfin, on prend une condition initiale de type "mode de Fourier", avec  $p \in \mathbb{R}$  arbitraire, et  $u_0$  défini par :

$$u_0(x) = e^{ipx}, \quad x \in \mathbb{R}.$$

La solution exacte est alors :

$$u(x, t) = e^{-p^2 t} e^{ipx}, \quad x \in \mathbb{R}, \quad t \in \mathbb{R}_+,$$

c'est-à-dire

$$u(\cdot, t) = e^{-p^2 t} u_0, \quad t \in \mathbb{R}_+.$$

Le facteur d'amplification est donc, pour tout  $t \in \mathbb{R}_+$ , le nombre  $e^{-p^2 t}$ . Ce facteur est toujours, en module, inférieur à 1. On va maintenant chercher la solution du schéma numérique sous la forme :

$$u_j^n = \xi_n e^{ipjh}, j \in \mathbb{Z}, n \in \mathbb{N}, \quad (2.7.148)$$

où  $\xi_0$  et  $\xi_1 \in \mathbb{R}$  sont donnés (ils donnent  $u_j^0$  et  $u_j^1$  pour tout  $j \in \mathbb{Z}$ ) et  $\xi_n \in \mathbb{R}$  est à déterminer de manière à ce que la première équation de (2.5.114) soit satisfaite.

Ce facteur  $\xi_n$  va dépendre de  $k, h$  et  $p$ . Pour  $k$  et  $h$  donnés, le schéma est stable au sens de Von Neumann si, pour tout  $p \in \mathbb{R}$ , la suite  $(\xi_n)_{n \in \mathbb{N}}$  est bornée. Dans le cas contraire, le schéma est (pour ces valeurs de  $k$  et  $h$ ) dit instable au sens de Von Neumann.

Un calcul immédiat donne que la famille des  $u_j^n$ , définie par (2.7.148), est solution de la première équation si et seulement si la suite  $(\xi_n)_{n \in \mathbb{N}}$  vérifie (on rappelle que  $\xi_0$  et  $\xi_1$  sont donnés) :

$$\frac{\xi_{n+1} - \xi_{n-1}}{2k} = \frac{2}{h^2} (\cos ph - 1) \xi_n, \quad n \geq 2,$$

ou encore, en posant

$$\alpha = \frac{4k}{h^2} (\cos ph - 1) \quad (\leq 0),$$

$$\xi_{n+1} - \alpha \xi_n - \xi_{n-1} = 0, \quad n \geq 2 \quad (2.7.149)$$

En excluant le cas  $\alpha = -2$  (qui correspond, pour  $k$  et  $h$  donnés, à des valeurs de  $p$  très particulières), la solution de (2.7.149) est

$$\xi_n = Ar_1^n + Br_2^n, \quad A \geq 0, \quad (2.7.150)$$

où  $A$  et  $B$  sont déterminés par  $\xi_0$  et  $\xi_1$  (de sorte que  $\xi_0 = A + B, \xi_1 = Ar_1 + Br_2$ ) et  $r_1, r_2$  sont les deux racines distinctes de :

$$r^2 - \alpha r - 1 = 0. \quad (2.7.151)$$

Les nombres  $r_1$  et  $r_2$  sont réels et comme  $r_1 r_2 = 1$ , l'un de ces nombres est, en module, supérieur à 1. Ceci montre que  $(\xi_n)_n$  est une suite non bornée (sauf pour des choix très particulier de  $\xi_0$  et  $\xi_1$ , ceux pour lesquelles  $\xi_1 = \xi_0 r_2$  où  $r_2$  est la racine de (2.7.151) de module inférieur à 1). Ce schéma est donc instable au sens de Von Neumann, pour tout  $k > 0$  et  $h > 0$ .

3. On reprend les notations de la question 1. On s'intéresse maintenant à la quantité  $S_j^n$  (qui est toujours l'erreur de consistance) :

$$S_j^n = \frac{\bar{u}_j^{n+1} - u_j^{n-1}}{2k} - \frac{\bar{u}_{j-1}^n - (\bar{u}_j^{n+1} + \bar{u}_j^{n-1}) + \bar{u}_{j+1}}{h^2}, \quad j = 1, \dots, N-1, \quad k = 0, \dots, M-1.$$

En reprenant la technique de la question 1, il existe  $(\xi_i, t_i)$ ,  $i = 1, \dots, 6$       *t.q.*

$$S_j^n = \frac{h^2}{12} (u_{ttt}(\xi_1, t_1) + u_{ttt}(\xi_2, t_2)) - \frac{h^2}{24} (u_{xxxx}(\xi_3, t_3) - u_{xxxx}(\xi_4, t_4)) + \frac{k^2}{2h^2} h_{tt}(\xi_5, t_5) + \frac{k^2}{2h^2} u_{tt}(\xi_6, t_6).$$

Ce qui donne, avec  $C_2$  ne dépendant que de  $u$ ,

$$|S_j^n| \leq C_2 \left( h^2 + k^2 + \frac{k^2}{h^2} \right), \quad j = 1, \dots, N-1, \quad k = 0, \dots, M-1.$$

Le schéma est donc consistant quand  $h \rightarrow 0$  avec  $\frac{k}{h} \rightarrow 0$ .

4. On reprend la méthode développée à la question 2, la suite  $(\xi_n)_n$  doit maintenant vérifier la relation suivante (avec  $\xi_0, \xi_1$  donnés).

$$\frac{\xi_{n+1} - \xi_{n-1}}{2k} = \frac{2 \cos(ph)}{h^2} \xi_n - \frac{\xi_{n-1} + \xi_{n+1}}{h^2}, \quad n \geq 2$$

c'est à dire :

$$\xi_{n+1} \left( \frac{1}{2k} + \frac{1}{h^2} \right) - \frac{2 \cos(ph)}{h^2} \xi_n + \xi_{n-1} \left( \frac{1}{h^2} - \frac{1}{2k} \right) = 0, \quad n \geq 2.$$

L'équation caractéristique est maintenant :

$$p(r) = ar^2 + br + c = 0,$$

avec

$$a = \frac{1}{2k} + \frac{1}{h^2}, \quad b = -\frac{2 \cos(ph)}{h^2} \quad \text{et} \quad c = \frac{1}{h^2} - \frac{1}{2k}.$$

Pour montrer la stabilité au sens de Von Neumann, il suffit d'après (2.7.150) de montrer que les deux racines du polynôme  $p$  sont de module inférieur ou égal à 1. On note  $r_1$  et  $r_2$  ces deux racines (qui peuvent être confondues) et on distingue 2 cas :

1. 1er cas : Les racines de  $p$  ne sont pas réelles. Dans ce cas, on a  $|r_1| = |r_2| = \gamma$  et

$$\gamma = \left| \frac{c}{a} \right| < 1,$$

car  $k > 0$ .

2. 2ème cas : Les racines de  $p$  sont réelles. Dans ce cas, on remarque que

$$r_1 r_2 = \frac{c}{a} < 1,$$

et l'une des racines, au moins, est donc entre  $-1$  et  $1$  (strictement). De plus on a  $p(1) = \frac{2}{h^2} - \frac{2 \cos ph}{h^2} \geq 0$  et  $p(-1) = \frac{2}{h^2} + \frac{2 \cos ph}{h^2} \geq 0$ , l'autre racine est donc aussi entre  $-1$  et  $1$  (au sens large).

On en déduit que le schéma (2.5.115) est stable au sens de Von Neumann.

## Corrigé de l'exercice 22 page 58 : schémas centré et décentré

1. Schéma explicite décentré

(a) Par définition, l'erreur de consistance en  $(x_i, t_n)$  s'écrit : On s'intéresse ici à l'ordre du schéma au sens des différences finies. On suppose que  $u \in C^4([0, 1] \times [0, T])$  est solution de (2.5.116) et on pose

$$\bar{u}_i^n = u(ih, nk), \quad i = 0, \dots, N, \quad k = 0, \dots, M.$$

Pour  $i = 1, \dots, N-1$  et  $k = 1, \dots, M-1$ , l'erreur de consistance en  $(x_i, t_k)$  est définie par :

$$R_i^n = \frac{1}{k}(\bar{u}_i^{n+1} - \bar{u}_i^n) - \frac{\alpha}{h}(\bar{u}_i^n - \bar{u}_{i-1}^n) - \frac{\mu}{h^2}(\bar{u}_{i-1}^n - 2\bar{u}_i^n + \bar{u}_{i+1}^n). \quad (2.7.152)$$

Soit  $i \in \{1, \dots, N-1\}$ ,  $k \in \{1, \dots, M-1\}$ . On cherche une majoration de  $R_i^n$  en utilisant des développements de Taylor. En utilisant ces développements, on obtient qu'il existe  $(\xi_\ell, t_\ell) \in [0, 1] \times [0, T]$ ,  $\ell = 1, \dots, 4$ , t.q. :

$$\bar{u}_i^{n+1} = \bar{u}_i^n + ku_t(ih, nk) + \frac{k^2}{2}u_{tt}(\xi_1, t_1), \quad (2.7.153)$$

$$\bar{u}_{i-1}^n = \bar{u}_i^n - hu_x(ih, nk) + \frac{h^2}{2}u_{xx}(\xi_2, t_2), \quad (2.7.154)$$

$$\bar{u}_{i-1}^n = \bar{u}_i^n - hu_x(ih, nk) + \frac{h^2}{2}u_{xx}(ih, nk) - \frac{h^3}{6}u_{xxx}(ih, nk) - \frac{h^4}{24}u_{xxxx}(\xi_3, t_3), \quad (2.7.155)$$

$$\bar{u}_{i+1}^n = \bar{u}_i^n + hu_x(ih, nk) + \frac{h^2}{2}u_{xx}(ih, nk) + \frac{h^3}{6}u_{xxx}(ih, nk) + \frac{h^4}{24}u_{xxxx}(\xi_4, t_4). \quad (2.7.156)$$

On en déduit :

$$\begin{aligned} R_i^n &= u_t(ih, nk) + \frac{k}{2}u_{tt}(\xi_1, t_1) + \alpha u_x(ih, nk) + \alpha \frac{h}{2}u_{xx}(\xi_2, t_2) \\ &\quad - \mu u_{xx}(ih, nk) - \mu \frac{h^2}{24}(u_{xxxx}(\xi_3, t_3) + \mu u_{xxxx}(\xi_4, t_4)), \end{aligned}$$

et donc, comme  $u$  est solution de (2.5.116), pour  $h$  assez petit, on a :

$$|R_i^n| \leq C_1(h + k),$$

où  $C_1$  ne dépend que de  $u$ . Le schéma (2.5.117) est donc consistant d'ordre 1 en temps et en espace.

- (b) Cherchons les conditions pour que  $u_i^{n+1}$  s'écrive comme combinaison convexe de  $u_i^n, u_{i-1}^n$  et  $u_{i+1}^n$ . On peut réécrire le schéma (2.5.117) :

$$u_i^{n+1} = au_i^n + bu_{i+1}^n + cu_{i-1}^n, \text{ avec } a = 1 - \frac{\alpha k}{h} - \frac{2\mu k}{h^2}, b = \frac{\mu k}{h^2} \text{ et } c = \frac{\alpha k}{h} + \frac{\mu k}{h^2}.$$

Il est facile de voir que  $a + b + c = 1$ , et que  $b \geq 0$ ,  $c \geq 0$ . Il reste à vérifier que  $a \geq 0$ ; pour cela, il faut et il suffit que  $\frac{\alpha k}{h} + \frac{2\mu k}{h^2} \leq 1$ . Cette condition écrit encore :

$$k \leq \frac{h^2}{\alpha h + 2\mu}. \quad (2.7.157)$$

Si  $h$  et  $k$  vérifient la condition (2.7.157), on pose :  $M^n = \max_{i=1 \dots N} u_i^n$  (resp.  $m^n = \min_{i=1 \dots N} u_i^n$ ). Comme  $u_i^{n+1}$  est une combinaison convexe de  $u_i^n, u_{i-1}^n$  et  $u_{i+1}^n$ , on a alors :  $u_i^{n+1} \leq M^n \quad \forall i = 1, \dots, N$  (resp.  $u_i^{n+1} \geq m^n \quad \forall i = 1, \dots, N$ ) et donc :  $M^{n+1} \leq M^n$  (resp.  $m^{n+1} \geq m^n$ ). On a ainsi montré que :

$$\|u^{n+1}\|_\infty \leq \|u^n\|_\infty.$$

On a de même :

$$\|u^n\|_\infty \leq \|u^{n-1}\|_\infty.$$

⋮

$$\|u^1\|_\infty \leq \|u^0\|_\infty.$$

En sommant ces inégalités, on obtient :

$$\|u^n\|_\infty \leq \|u^0\|_\infty.$$

Donc, sous la condition (2.7.157), on a  $\|u^{n+1}\|_\infty \leq \|u^n\|_\infty$  et donc  $\|u^n\|_\infty \leq \|u^0\|_\infty$ , pour tout  $n = 1, \dots, N$ .

(c) En retranchant l'égalité (2.7.152) au schéma (2.5.117), on obtient l'équation suivante sur  $e_i^n$  :

$$\frac{1}{k}(e_i^{n+1} - e_i^n) + \frac{\alpha}{h}(e_i^n - e_{i-1}^n) - \frac{\mu}{h^2}(e_{i-1}^n - 2e_i^n + e_{i+1}^n) = R_i^n.$$

ce qu'on peut encore écrire :

$$e_i^{n+1} = (1 - \frac{k\alpha}{h} - 2\frac{k\mu}{h^2})e_i^n + e_{i-1}^n \frac{k\mu}{h^2} + kR_i^n. \quad (2.7.158)$$

Sous la condition de stabilité (2.7.157), on obtient donc :

$$\begin{aligned} |e_i^{n+1}| &\leq \|e^{n+1}\|_\infty + C_1(k+h)k, \\ |e_i^n| &\leq \|e^{n-1}\|_\infty + C_1(k+h)k, \\ \vdots &\leq \vdots + \vdots \\ |e_i^1| &\leq \|e^0\|_\infty + C_1(k+h)k, \end{aligned}$$

Si à  $t = 0$ , on a  $\|e^0\| = 0$ , alors on éduit des inégalités précédentes que  $|e_i^n| \leq C1T(k+h)$  pour tout  $n \in \mathbb{N}$ . Le schéma est donc convergent d'ordre 1.

2. Schéma explicite centré.

(a) (Consistance) En utilisant les développements de Taylor (2.7.153) (2.7.155) et (2.7.156), et les développements suivants :

$$\begin{aligned} \bar{u}_{i-1}^n &= \bar{u}_i^n - hu_x(ih, nk) + \frac{h^2}{2}u_{xx}(ih, nk) - \frac{h^3}{6}u_{xxx}(\xi_5, t_5), \\ \bar{u}_{i+1}^n &= \bar{u}_i^n + hu_x(ih, nk) + \frac{h^2}{2}u_{xx}(ih, nk) + \frac{h^3}{6}u_{xxx}(\xi_6, t_6), \end{aligned}$$

on obtient maintenant :

$$\begin{aligned} R_i^n &= u_t(ih, nk) + \frac{k}{2}u_{tt}(\xi_1, t_1) + \alpha u_x(ih, nk) + \alpha \frac{h^2}{12}(u_{xxx}(\xi_5, t_5) + \mu u_{xxx}(\xi_6, t_6)) \\ &\quad - \mu u_{xx}(ih, nk) - \mu \frac{h^2}{24}(u_{xxxx}(\xi_3, t_3) + \mu u_{xxxx}(\xi_4, t_4)), \end{aligned}$$

On en déduit que

$$|R_i^n| \leq C_3(k+h^2),$$

où  $C_3 = \max(\frac{1}{2}\|u_{tt}\|_\infty, \frac{1}{6}\|u_{xxx}\|_\infty, \frac{1}{12}\|u_{xxxx}\|_\infty)$ .

(b) Le schéma s'écrit maintenant :

$$u_i^{n+1} = \tilde{a}u_i^n + \tilde{b}u_{i+1}^n + \tilde{c}u_{i-1}^n, \text{ avec } \tilde{a} = 1 - \frac{2\mu k}{h^2}, \tilde{b} = \frac{\mu k}{h^2} - \frac{\alpha k}{h} \text{ et } \tilde{c} = \frac{\mu k}{h^2} + \frac{\alpha k}{h}.$$

Remarquons que l'on a bien :  $\tilde{a} + \tilde{b} + \tilde{c} = 1$ . Pour que  $u_i^{n+1}$  soit combinaison convexe de  $u_i^n$ ,  $u_{i+1}^n$  et  $u_{i-1}^n$ , il faut et il suffit donc que  $\tilde{a} \geq 0$ ,  $\tilde{b} \geq 0$ , et  $\tilde{c} \leq 0$ . L'inégalité  $\tilde{c} \geq 0$  est toujours vérifiée. Les deux conditions qui doivent être vérifiées par  $h$  et  $k$  s'écrivent donc :

i.  $\tilde{a} \geq 0$ , i.e.  $1 - \frac{2\mu k}{h^2} \geq 0$ , soit encore

$$k \leq \frac{h^2}{2\mu}.$$

ii.  $\tilde{b} \geq 0$  i.e.  $\frac{\mu k}{h^2} - \frac{\alpha k}{h} \geq 0$ , soit encore

$$h \leq \frac{\mu}{2\alpha}.$$

Le schéma centré est donc stable sous les deux conditions suivantes :

$$h \leq \frac{\mu}{2\alpha} \text{ et } k \leq \frac{1}{2\mu}h^2. \quad (2.7.159)$$

Pour obtenir une borne d'erreur, on procède comme pour le schéma (2.5.117) : on soustrait la définition de l'erreur de consistance au schéma numérique, et on obtient :

$$e_i^{n+1} = \tilde{a}e_i^n + \tilde{b}e_{i+1}^n + \tilde{c}e_{i-1}^n + kR_i^n. \quad (2.7.160)$$

Par le même raisonnement que pour le schéma décentré, on obtient donc que si  $e_i^0 = 0$ , on a  $|e_i^n| \leq C_4(k + h^2)$ , avec  $C_4 = TC_3$ .

### Corrigé de l'exercice 23 page 59 : Discrétisation d'un problème parabolique non linéaire

Corrigé en cours de rédaction