

# Chapitre 2

## Equation de Laplace

### 2.1 Modélisation

Dans  $\mathbb{R}^n$  on considère la variation d'une quantité telle que la température  $T$  sous l'effet de forces volumiques de densité  $f$  en l'absence de tout champ de vitesses (par exemple, la chaleur n'est pas mue par un champ de vitesses). Autrement dit, l'énergie du matériau considérée se réduit à son énergie potentielle de la forme  $\frac{1}{2}\|\nabla T\|^2$  dont la force associée dérive d'un potentiel, soit  $\vec{\nabla}T$  (loi de Fourier). Si  $V$  est un volume quelconque fixe, la force volumique  $f_V := \int_V f(x)d\Omega$  compense exactement le flux à la frontière de  $V$  d'origine potentielle :

$$f_V = - \int_{\partial V} \vec{\nabla}T \cdot \vec{dS}$$

où  $\vec{dS}$  est la normale extérieure à  $V$  le long de  $\partial V$ . On en déduit :

$$\begin{aligned} \int_V f(x)d\Omega &= - \int_{\partial V} \underbrace{\vec{\nabla}T \cdot \vec{dS}}_{=: \omega} = - \int_{\partial V} \omega \stackrel{Stokes}{=} - \int_V d\omega = \\ &= - \int_V \underbrace{\operatorname{div}(\vec{\nabla}T)}_{=: \Delta T} d\Omega = - \int_V \Delta T d\Omega. \end{aligned}$$

Ceci est vrai pour tout volume  $V \subset \mathbb{R}^n$  suffisamment régulier donc

$$f = -\Delta T \quad \text{dans } \mathbb{R}^n.$$

**Définition 2.1.1** (Laplacien). Dans  $\mathbb{R}^n$  on appelle Laplacien l'opérateur :

$$\Delta = \sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}.$$

dit elliptique car sa transformée de Fourier

$$\mathcal{F}(\Delta)(\xi) = \xi_1^2 + \xi_2^2$$

est le terme principal de l'équation d'une sphère et plus généralement d'une ellipse. Du fait de sa symétrie le laplacien est le modèle des opérateurs elliptiques :

$$x \mapsto \sum_{i=1}^n a_i(x) \frac{\partial^2}{\partial x_i^2}.$$

## 2.2 Le Laplacien comme opérateur non borné

Soit  $\Omega \subset \mathbb{R}^n$  un ouvert connexe, borné ou non. L'opérateur  $\Delta$  est bien défini au sens des distributions sur  $L^2(\Omega)$  :

$$\langle \Delta u, \varphi \rangle := \int_{\Omega} u \Delta \varphi dx, \quad \forall \varphi \in \mathcal{D}(\Omega)$$

mais il n'opère pas dans  $L^2(\Omega)$ . Pour remédier à cela, on utilise la notion d'opérateur non borné.

**Définition 2.2.1** (Opérateur borné). Un opérateur  $A : \mathcal{H} \rightarrow \mathcal{H}$  défini sur un espace de Hilbert  $\mathcal{H}$  est dit borné s'il est continu, i.e. s'il existe une constante  $C > 0$  t.q.

$$\|Ax\| \leq C\|x\|, \quad \forall x \in \mathcal{H}.$$

Un opérateur non borné est défini, en général, sur un sous-espace  $D(A)$  dense de  $\mathcal{H}$ . Un opérateur non borné est donc une application linéaire  $A : D(A) \rightarrow \mathcal{H}$ . Si  $A$  est un opérateur à domaine dense  $D(A)$ , on définit  $D(A^*)$  comme l'ensemble des vecteurs  $\varphi \in \mathcal{H}$  pour lesquels il existe un vecteur  $\varphi' \in \mathcal{H}$  vérifiant :

$$\forall \psi \in D(A), \quad \langle A\psi, \varphi \rangle = \langle \psi, \varphi' \rangle.$$

Pour tout  $\varphi \in D(A^*)$ , domaine de l'adjoint, on note  $A^*\varphi = \varphi'$ . Par densité de  $D(A)$ ,  $\varphi'$  est défini de façon unique. Un opérateur  $A$  sur  $\mathcal{H}$  est dit auto-adjoint si  $D(A^*) = D(A)$  et  $A = A^*$  sur  $D(A)$ . Malheureusement, dans beaucoup de cas intéressants,  $D(A^*)$  est beaucoup plus petit que  $D(A)$ , voire est réduit à  $\{0\}$ . Avec ces notations, on pose :

$$D(-\Delta) = \{u \in L^2(\Omega) \mid -\Delta u \in L^2(\Omega)\}.$$

On remarque que pour tout  $u \in D(-\Delta)$ , on peut définir la trace  $u|_{\partial\Omega}$  et la dérivée normale  $\frac{\partial u}{\partial n}$  le long de  $\partial\Omega$  en posant :  $\forall \varphi \in \mathcal{C}^\infty(\Omega)$ ,

$$\int_{\partial\Omega} u \frac{\partial \varphi}{\partial n} d\sigma - \int_{\partial\Omega} \frac{\partial u}{\partial n} \varphi d\sigma = \int_{\Omega} (-\Delta u) \varphi dx + \int_{\Omega} u \Delta \varphi dx, \quad \text{où } \frac{\partial \varphi}{\partial n} := \vec{n} \cdot \vec{\nabla} \varphi$$

L'adjoint de  $-\Delta$  est bien défini quand il est associé à une condition sur le bord de type Dirichlet, resp. Neumann,

**Proposition 2.2.1.** *Les opérateurs  $-\Delta_d$  et  $-\Delta_n$ , définis par :*

$$D(-\Delta_d) = \{u \in L^2(\Omega) \mid \Delta u \in L^2(\Omega) \text{ et } u|_{\partial\Omega} = 0\}$$

et :  $-\Delta_d u = -\Delta u$ ,  $\forall u \in D(-\Delta_d)$ , resp. :

$$D(-\Delta_n) = \{u \in L^2(\Omega) \mid \Delta u \in L^2(\Omega) \text{ et } \left. \frac{\partial u}{\partial n} \right|_{\partial\Omega} = 0\}$$

et :  $-\Delta_n u = -\Delta u$ ,  $\forall u \in D(-\Delta_n)$ , sont auto-adjoints et positifs.

*Démonstration.* On a les égalités :

$$\int_{\Omega} (-\Delta u) \varphi dx = \int_{\Omega} u (-\Delta \varphi) dx, \quad \forall u, \varphi \in D(-\Delta_d)$$

i.e.  $(-\Delta_d)^* = -\Delta_d$ , resp.

$$\int_{\Omega} (-\Delta u) \varphi dx = \int_{\Omega} u (-\Delta \varphi) dx, \quad \forall u, \varphi \in D(-\Delta_n)$$

i.e.  $(-\Delta_n)^* = -\Delta_n$ .

De plus, si  $u \in D(-\Delta)$ , on prolonge  $u \in L^2(\Omega)$  et  $-\Delta u \in L^2(\Omega)$  par 0 dans  $\mathbb{R}^n \setminus \Omega$ . On en déduit alors :

$$\int_{\Omega} (-\Delta u) u dx = \int_{\mathbb{R}^n} \overline{\mathcal{F}(-\Delta u)} \mathcal{F}(u) dx = \int_{\mathbb{R}^n} |\xi|^2 |\mathcal{F}(u)(\xi)|^2 d\xi \geq 0$$

i.e.  $-\Delta$  est un opérateur positif.  $\square$

**Corollaire 2.2.2.** *Le spectre de l'opérateur  $-\Delta_d$ , resp.  $-\Delta_n$ , est formé d'une suite de valeurs propres  $(\lambda_k^{(d)})_{k \geq 0} \in (\mathbb{R}^+)^{\mathbb{N}}$ , resp.  $(\lambda_k^{(n)})_{k \geq 0} \in (\mathbb{R}^+)^{\mathbb{N}}$ , sans point d'accumulation à distance finie, t.q.  $\lim_{k \rightarrow +\infty} \lambda_k^{(d)} = +\infty$ , resp.  $\lim_{k \rightarrow +\infty} \lambda_k^{(n)} = +\infty$ , et les vecteurs propres correspondants forment une base hilbertienne de  $L^2(\Omega)$ .*

*Démonstration.* C'est une conséquence (admise) de la théorie des opérateurs auto-adjoints positifs.  $\square$

L'exemple le plus simple de la situation décrite dans le Corollaire 2.2.2 s'obtient en dimension  $n = 1$  d'espace. On pose  $\Omega = ]0, L[$  avec  $L > 0$  et on considère l'équation :

$$-u'' - \lambda u = 0, \quad u(0) = u(L) = 0.$$

Par le calcul, on trouve directement que les valeurs propres sont les réels :

$$\lambda_k = \left(\frac{k\pi}{L}\right)^2 \geq 0, \quad k \geq 0$$

de vecteurs propres associés :

$$t \mapsto w_{\lambda_k}(t) = \sin\left(\frac{k\pi}{L}t\right), \quad k \geq 0.$$

On vérifie que la suite  $(w_k)_{k \geq 0}$  est une base hilbertienne de  $L^2(0, T)$ .

*Remarque 8.* L'analyse ci-dessus se généralise au cas du problème de Dirichlet (resp. de Neumann, et d'autres) dans un ouvert  $\Omega$  borné. Pour évaluer le comportement asymptotique des valeurs propres on introduit la fonction de comptage :

$$N(\lambda) := \#\{\lambda_k \leq \lambda\}, \quad \forall \lambda \in \mathbb{R}.$$

Ces valeurs propres apparaissent de manière fondamentale dans les équations de propagation des ondes et de la chaleur sur un ouvert  $\Omega$ .

Elles apparaissent aussi en théorie des nombres. Si  $\Omega$  est un carré de côté 1, alors les fonctions propres et les valeurs propres du Laplacien sont donnés par :

$$w_{\lambda_{n,m}}(x, y) = \sin(n\pi x) \sin(m\pi y), \quad \lambda_{n,m} = (n^2 + m^2)\pi^2, \quad n, m \geq 0.$$

L'évaluation de  $N(\lambda)$  dans ce cas est bien un problème de théorie des nombres : compter le nombre de points à coordonnées entières contenus dans un cercle de rayon  $\sqrt{\lambda}$ ,  $\lambda > 0$ . On démontre que :

$$N(\lambda) = \frac{\sqrt{\lambda}}{4\pi} + o(\sqrt{\lambda}), \quad \forall \lambda > 0.$$

## 2.3 Formulation variationnelle

### Introduction et formalisme

Soit  $\Omega \subset \mathbb{R}^n$  un ouvert borné de  $\mathbb{R}^n$  et soit  $f \in L^2(\Omega)$ . On considère le problème aux limites : trouver  $u$  solution de (2.1)

$$u - \Delta u = f \quad \text{dans } \Omega, \quad u = 0 \quad \text{sur } \partial\Omega. \quad (2.1)$$

ce qui, avec les notations de la Section 2.2, se réécrit : trouver  $u \in D(-\Delta_d)$  t.q.

$$(I - \Delta_d)u = f.$$

D'après le Corollaire 2.2.2, l'opérateur  $I - \Delta_d$  est inversible et le problème (2.1) admet  $u := (I - \Delta_d)^{-1}f$  pour unique solution dans  $D(-\Delta_d)$ .

Par multiplication de (2.1) par  $v \in \mathcal{D}(\Omega)$ , on obtient : (2.2)

$$\underbrace{\int_{\Omega} (u\bar{v} + \nabla u \bar{\nabla} v) dx}_{=: (u,v)} = \int_{\Omega} f\bar{v} dx, \quad \forall v \in \mathcal{D}(\Omega). \quad (2.2)$$

ce qui peut s'interpréter comme la résolution d'un problème de représentation d'une forme sesquilinéaire par le théorème de représentation de Riesz. Pour cela, on introduit :

$$H^1(\Omega) = \{v \in L^2(\Omega) \mid \vec{\nabla} v \in L^2(\Omega)^n\}$$

L'espace  $H^1(\Omega)$  muni du produit scalaire  $((\cdot, \cdot))$  est un espace de Hilbert. On remarque que la relation :

$$\int_{\Omega} \vec{\varphi} \cdot \vec{\nabla} u dx = - \int_{\Omega} \operatorname{div}(\vec{\varphi}) u dx + \int_{\partial\Omega} \underbrace{\vec{n} \cdot \vec{\varphi}}_{=: \varphi_n} u d\sigma, \quad \forall \vec{\varphi} \in \mathcal{C}^\infty(\Omega)^n$$

permet de définir les valeurs de  $u \in H^1(\Omega)$  sur  $\partial\Omega$ , ce qui permet de définir le sous-espace :

$$H_0^1(\Omega) = \{v \in H^1(\Omega) \mid v|_{\partial\Omega} = 0\}.$$

**Proposition 2.3.1.** *Le sous-espace  $H_0^1(\Omega)$  est fermé dans  $H^1(\Omega)$ .*

*Démonstration.* Soit  $(u_k)_{k \geq 0} \in H_0^1(\Omega)^{\mathbb{N}}$  t.q.  $u_k \xrightarrow[k \rightarrow +\infty]{} u$  dans  $H^1(\Omega)$ , i.e. :

$$u_k \xrightarrow[k \rightarrow +\infty]{} u \quad \text{et} \quad \vec{\nabla} u_k \xrightarrow[k \rightarrow +\infty]{} \vec{\nabla} u \quad \text{dans } L^2(\Omega).$$

Alors :  $\forall \vec{\varphi} \in \mathcal{C}^\infty(\Omega)^n$ ,

$$\int_{\Omega} \vec{\varphi} \cdot \vec{\nabla} u_k dx \xrightarrow{k \rightarrow +\infty} \int_{\Omega} \vec{\varphi} \cdot \vec{\nabla} u dx,$$

d'une part, et

$$\int_{\Omega} \vec{\varphi} \cdot \vec{\nabla} u_k dx = - \int_{\Omega} \operatorname{div}(\vec{\varphi}) u_k dx \xrightarrow{k \rightarrow +\infty} - \int_{\Omega} \operatorname{div}(\vec{\varphi}) u dx$$

d'autre part, d'où on déduit que :

$$\int_{\partial\Omega} \varphi_n u d\sigma = 0, \quad \forall \vec{\varphi} \in \mathcal{C}^\infty(\Omega)^n$$

i.e.  $u = 0$  sur  $\partial\Omega$  et  $u \in H_1^0(\Omega)$ . □

On admettra le résultat de densité :

**Proposition 2.3.2.** *Le sous-espace  $H_0^1(\Omega)$  de  $H^1(\Omega)$  coïncide avec l'adhérence de  $\mathcal{D}(\Omega)$  pour la norme  $\|\cdot\|$  de  $H^1(\Omega)$  :*

$$\overline{\mathcal{D}(\Omega)}^{H^1} = H_0^1(\Omega) \subsetneq H^1(\Omega)$$

**Proposition 2.3.3.** *Le problème (2.1) admet une unique solution  $u \in H_0^1(\Omega)$*

*Démonstration.* Compte tenu de (2.2) et de la Proposition 2.3.2, on est ramené à résoudre : trouver  $u \in H_0^1(\Omega)$  t.q. :

$$\forall v \in H_0^1(\Omega), \quad ((u, v)) = \int_{\Omega} f \bar{v} dx. \quad (2.3)$$

On remarque que  $\ell : v \mapsto \int_{\Omega} f \bar{v} dx$  est une forme antilinéaire continue sur  $H^1(\Omega)$ . D'après le Théorème de représentation de Riesz, il existe un unique  $w \in H^1(\Omega)$  t.q. :  $\ell(v) = ((w, v))$ ,  $\forall v \in H^1(\Omega)$ . Soit  $P$  la projection orthogonale :  $H^1(\Omega) \rightarrow H_0^1(\Omega)$ , bien définie d'après la Proposition 2.3.1. Alors, (2.3) se réécrit :

$$\forall v \in H^1(\Omega), \quad ((u, Pv)) = ((w, Pv)).$$

Nécessairement :  $u \in H_0^1(\Omega) \Rightarrow u = Pv$ . □

On généralise le résultat de la Proposition 2.3.3 dans le cadre fonctionnel abstrait de deux espaces de Hilbert  $H, V$  ( $H = L^2(\Omega)$  et  $V = H_0^1(\Omega)$  dans l'exemple précédent) munis de deux produits scalaires différents. On suppose que  $V \subset H$  et que cette injection est continue et dense ( $\overline{V}^H = H$ ). Alors, l'identification de  $H$  avec son dual (antidual si on considère des espaces de fonctions complexes) réalise une injection de cet espace dans  $V^*$ , ce qui donne le triplet :

$$V \subset H \simeq H^* \subset V^*.$$

Dans ce cadre mes formes sesquilinéaires continues  $(u, v) \mapsto a(u, v)$  s'identifient à des opérateurs linéaires continus  $A_a : V \rightarrow V^*$  suivant la formule :

$$\forall (u, v) \in V \times V, \quad A_a(u)(v) = a(u, v).$$

et par restriction de leur image à  $H$  définissent encore des opérateurs non bornés dans  $H$  dont les domaines et actions sont définis par :

$$D(A_a^H) = \{u \in V \mid A_a(u) \in H\} \quad \text{et} \quad A_a^H(u) = A_a(u), \quad \forall u \in D(A_a^H).$$

On en déduit, en particulier :

$$\forall u \in D(A_a^H), \quad \forall v \in V, \quad (A_a^H(u), v)_H = A_a(u)(v) = {}_{V^*}\langle A_a(u), v \rangle_V.$$

## Le Théorème de Lax-Milgram

Dans le formalisme ci-dessus, le Théorème de Lax-Milgram généralise le Théorème de représentation de Riesz.

**Théorème 2.3.4.** *Soit  $V$  un espace de Hilbert et soit  $(u, v) \mapsto a(u, v)$  une forme sesquilinéaire continue sur  $V \times V$ . On suppose qu'il existe une constante  $\alpha > 0$  t.q.*

$$|a(u, u)| \geq \alpha \|u\|_V^2, \quad \forall u \in V.$$

*Alors, l'opérateur  $A_a : V \rightarrow V^*$ ,  $u \mapsto A_a(u)(\cdot) = a(u, \cdot)$ , réalise un isomorphisme de  $V \rightarrow V^*$  et on a :*

$$\|A_a^{-1}\|_{\mathcal{L}_c(V^*, V)} \leq \frac{1}{\alpha}$$

*Démonstration.* On commence par remarquer que  $A_a$  est linéaire (immédiat) injective. En effet, soit  $u \in V$  t.q.  $A_a(u) = 0$ . Alors :

$$0 = |A_a(u)(u)| = |a(u, u)| \geq \alpha \|u\|_V^2 \Rightarrow \|u\| = 0$$

i.e.  $u = 0$ . Il reste à vérifier que  $A_a(V) = V^*$ . Pour cela on va montrer successivement que  $A_a(V)$  est fermé et dense dans  $V^*$ . Comme  $V$  est complet, il en est de même de  $V^*$  et on est ramené à montrer que  $A_a(V)$  est complet dans  $V^*$ . Soit  $(u_k)_{k \geq 0} \in V^{\mathbb{N}}$ . On suppose que la suite  $(A_a(u_k))_{k \geq 0}$  est de Cauchy. On a :  $\forall k, p \geq 0$ ,

$$\begin{aligned} \alpha \|u_k - u_p\|_V^2 &\leq |a(u_k - u_p, u_k - u_p)| = |A_a(u_k - u_p)(u_k - u_p)| \leq \\ &\leq \|A_a(u_k - u_p)\| \|u_k - u_p\|_V \end{aligned}$$

donc

$$\|u_k - u_p\|_V \leq \frac{1}{\alpha} \|A_a(u_k - u_p)\|.$$

On en déduit que la suite  $(u_k)_{k \geq 0}$  est de Cauchy dans  $V$ , donc convergente dans  $V$  qui est complet. Soit  $u \in V$  sa limite dans  $V$ . On remarque que  $A_a$  est continue sur  $V$  par continuité de  $a$ . En effet :  $\forall w, v \in V$ ,

$$\begin{aligned} |A_a(w)(v)| &= |a(w, v)| \leq \|a\| \|w\|_W \|v\|_V \\ \Rightarrow \|A_a(w)\|_{V^*} &= \sup_{v \in V \setminus \{0\}} \frac{|A_a(w)(v)|}{\|v\|_V} \leq \|a\| \|w\|_W \end{aligned}$$

On en déduit que  $A_a(u_k) \xrightarrow{k \rightarrow +\infty} A_a(u)$  dans  $V^*$ , i.e. que  $A_a(V)$  est complet.

Pour montrer que  $A_a(V)$  est dense dans  $V^*$ , on remarque que :

$$\begin{aligned} A_a(V) = \overline{A_a(V)} &= \overline{\text{Im}(A_a)} = \text{Ker}(A_a^*)^\perp \Rightarrow \overline{A_a(V)} = V^* \\ \iff \text{Ker}(A_a^*)^\perp = V^* &\iff \text{Ker}(A_a^*)^{\perp\perp} = V^{*\perp} \iff \overline{\text{Ker}(A_a^*)} = \{0\} \\ &= \text{Ker}(A_a^*) \end{aligned}$$

Par définition de l'adjoint :  $\forall u, v \in V$ ,

$$A_a(u)(v) = A_a^*(v)(u) = a(u, v).$$

Soit  $A_a^*(v) = 0$  avec  $v \in V$ . Alors :

$$0 = A_a^*(v)(v) = a(v, v) \geq \alpha \|v\|_V^2 \Rightarrow \|v\|_V = 0$$

i.e. :  $v = 0$ , ce qui achève de montrer que  $A_a(V) = V^*$ , et finalement que  $A_a$  est un isomorphisme de  $V \rightarrow V^*$ . Soit  $f \in V^*$ . On a :

$$\begin{aligned} \alpha \|A_a^{-1}(f)\|_V^2 &\leq a(A_a^{-1}(f), A_a^{-1}(f)) = f(A_a^{-1}(f)) \leq \|f\|_{V^*} \|A_a^{-1}(f)\|_V \\ \Rightarrow \alpha \|A_a^{-1}(f)\|_V &\leq \|f\|_{V^*}, \quad \forall f \in V^*. \end{aligned}$$

Il en résulte :

$$\|A_a^{-1}\|_{\mathcal{L}_c(V^*, V)} = \sup_{f \in V^* \setminus \{0\}} \frac{\|A_a^{-1}(f)\|_V}{\|f\|_{V^*}} \leq \frac{1}{\alpha}.$$

□



On va donner différents types d'applications de ce formalisme. Il convient de garder à l'esprit que ces extensions peuvent être combinées entre elles. Pour simplifier l'exposé et parce que c'est aussi le cadre de ces applications, on se limite à des fonctions à valeurs réelles.

## Conditions aux limites et inégalité de Poincaré

Le formalisme ci-dessous dit formulation variationnelle a l'avantage de prendre en compte les conditions aux limites attachées à l'edp étudiée. En choisissant adéquatement  $V$  et la forme linéaire  $\ell$  dans le problème : trouver  $u$  solution de

$$u \in V \quad \text{et} \quad ((u, v))_V = \ell(v), \quad \forall v \in V$$

on peut traiter différentes conditions aux limites. Soit  $\Gamma \subset \partial\Omega$  une partie de mesure non nulle du bord  $\partial\Omega$ . On pose :

$$V_\Gamma := \{v \in H^1(\Omega) \mid v|_\Gamma = 0\}. \quad (2.4)$$

**Proposition 2.3.5.** *L'espace  $V_\Gamma$  défini par (2.4) est un sous-espace fermé de  $H^1(\Omega)$ .*

*Démonstration.* Par définition,  $V_\Gamma$  est un sous-espace de  $H^1(\Omega)$ .

Soit  $(u_k)_{k \geq 0} \in V_\Gamma^{\mathbb{N}}$  t.q.  $u_k \xrightarrow[k \rightarrow +\infty]{} u$  dans  $H^1(\Omega)$ . Soit  $\forall \vec{\varphi} \in \mathcal{C}^\infty(\Omega)^n$ . On a :

$$\begin{aligned} \int_{\Omega} \vec{\varphi} \cdot \vec{\nabla} u_k dx &= - \int_{\Omega} \operatorname{div}(\vec{\varphi}) u_k dx + \int_{\partial\Omega} \underbrace{\vec{n} \cdot \vec{\varphi}}_{=: \varphi_n} u_k d\sigma \\ &\stackrel{u_k \in V_\Gamma}{=} - \int_{\Omega} \operatorname{div}(\vec{\varphi}) u_k dx + \int_{\partial\Omega \setminus \Gamma} \varphi_n u_k d\sigma, \quad \forall k \geq 0, \end{aligned}$$

d'où on déduit :

$$\begin{aligned} \int_{\partial\Omega} \varphi_n u d\sigma &= \int_{\Omega} \vec{\varphi} \cdot \vec{\nabla} u dx + \int_{\Omega} \operatorname{div}(\vec{\varphi}) u dx = \lim_{k \rightarrow +\infty} \left( \int_{\Omega} \vec{\varphi} \cdot \vec{\nabla} u_k dx + \int_{\Omega} \operatorname{div}(\vec{\varphi}) u_k dx \right) \\ &= \lim_{k \rightarrow +\infty} \int_{\partial\Omega \setminus \Gamma} \varphi_n u_k d\sigma. \end{aligned}$$

En particulier, si  $\operatorname{supp}(\vec{\varphi}) \subset \Gamma$  alors :

$$\int_{\partial\Omega} \varphi_n u d\sigma = \int_{\Gamma} \varphi_n u d\sigma = 0.$$

On en déduit que  $u|_\Gamma = 0$ , i.e.  $u \in V_\Gamma$ . □

Soit  $f \in L^2(\Omega)$  et soit  $g \in L^2(\partial\Omega)$ . On introduit la forme linéaire continue sur  $V_\Gamma$  :

$$\ell : v \mapsto \int_{\Omega} f v dx + \int_{\partial\Omega} g v d\sigma$$

Soit  $\lambda > 0$ . D'après le théorème de Lax-Milgram ( ou le Théorème de représentation de Riesz appliqué à  $V_\Gamma$ ), il existe un unique  $u \in V_\Gamma$  t.q. :

$$\forall v \in V_\Gamma, \quad \lambda \int_{\Omega} u v dx + \int_{\Omega} \vec{\nabla} u \cdot \vec{\nabla} v dx = \ell(v). \quad (2.5)$$

En prenant  $v = \varphi \in \mathcal{D}(\Omega) \subset V_\Gamma$ , et en intégrant par parties, on obtient :

$$\lambda u - \Delta u = f \quad \text{dans } \mathcal{D}'(\Omega), \quad u = 0 \quad \text{sur } \Gamma. \quad (2.6)$$

On suppose que  $u$  est assez régulier pour appliquer la formule de Green. Après multiplication de (2.6) par  $v \in V_\Gamma$  :

$$\int_{\Omega} f v dx = \int_{\Omega} (\lambda u - \Delta u) v dx = \underbrace{\lambda \int_{\Omega} u v dx + \int_{\Omega} \vec{\nabla} u \cdot \vec{\nabla} v dx}_{\stackrel{(2.5)}{=} \ell(v)} - \int_{\partial\Omega \setminus \Gamma} \frac{\partial u}{\partial n} v d\sigma.$$

Par comparaison avec (2.5), on en déduit que :

$$\frac{\partial u}{\partial n} = g \quad \text{sur } \partial\Omega \setminus \Gamma.$$

et  $u$  est finalement solution du problème aux limites :

$$\lambda u - \Delta u = f \quad \text{dans } \mathcal{D}'(\Omega), \quad u = 0 \quad \text{sur } \Gamma, \quad \frac{\partial u}{\partial n} = g \quad \text{sur } \partial\Omega \setminus \Gamma;$$

## Résolution analytique par les séries de Fourier

On considère le problème avec conditions aux limites :

$$\left\{ \begin{array}{ll} \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 & \text{dans } ]0, L[ \times ]0, M[ =: Q, \\ u(x, M) = 0 & \text{dans } ]0, L[ \\ u(x, 0) = u_0(x) & \text{dans } ]0, L[ \\ u(0, y) = u(L, y) = 0 & \text{dans } ]0, M[ \end{array} \right. \quad (2.7)$$

**Théorème 2.3.6** (Existence et unicité). *Soit  $u_0 \in \mathcal{C}^2(]0, L[, \mathbb{R}) \cap \mathcal{C}([0, L], \mathbb{R})$ . Alors, il existe une unique fonction  $u \in \mathcal{C}^2(]0, L[ \times ]0, M[, \mathbb{R}) \cap \mathcal{C}([0, L] \times [0, M], \mathbb{R})$  solution de (2.7).*

**Proposition 2.3.7.** *On suppose que  $u_0 \in \mathcal{C}^2([0, L], \mathbb{R})$  et que  $u_0(0) = u_0(L) = 0$ . Alors la solution du problème (2.7) se développe en série de Fourier sous la forme :*

$$u(x, t) = \sum_{n \in \mathbb{Z}} c_n(u_0) \frac{\sinh\left(\frac{n\pi}{L}(M - y)\right)}{\sinh\left(\frac{n\pi M}{L}\right)} \sin\left(\frac{n\pi x}{L}\right) \quad (2.8)$$

où  $(c_n(u_0))_{n \in \mathbb{Z}}$  est la suite des coefficients de Fourier de  $u_0$ .

*Démonstration.* On commence par chercher  $u$  sous la forme  $u(x, t) = X(x)Y(y)$  où  $X$  et  $Y$  sont de classe  $\mathcal{C}^2$ , en accord avec le Théorème 2.3.6. Après report dans (2.7), on obtient :

$$\frac{X''}{X} = -\frac{Y''}{Y} = \lambda \in \mathbb{R}.$$

Si  $\lambda = \omega^2 > 0$  avec  $\omega > 0$ , alors

$$X(x) = ae^{\omega x} + be^{-\omega x}, \quad a, b \in \mathbb{R}$$

avec :  $X(0) = X(L) = 0$ , ce qui entraîne que  $(a, b)$  est solution du système :

$$\begin{cases} a + b & = 0, \\ ae^{\omega L} + be^{-\omega L} & = 0 \end{cases}$$

dont l'unique solution est  $a = b = 0$ , en contradiction avec  $u \neq 0$ . Donc  $\lambda = -\omega^2 < 0$  et on a :

$$X(x) = a \cos(\omega x) + b \sin(\omega x)$$

avec :  $X(0) = 0 \Rightarrow a = 0$ . Il reste :  $X(L) = b \sin(\omega L) = 0$  et donc  $\omega \in \frac{\pi}{L}\mathbb{Z}$ . On obtient donc la suite de solutions  $(u_n)_{n \in \mathbb{Z}}$  définies par :

$$u_n(x, y) = \frac{\sinh\left(\frac{n\pi}{L}(M - y)\right)}{\sinh\left(\frac{n\pi M}{L}\right)} \sin\left(\frac{n\pi x}{L}\right), \quad \forall (x, t) \in Q.$$

Il reste à vérifier la condition au bord en  $y = 0$ , ce qu'on cherche a priori pour  $u = \sum_{n \in \mathbb{Z}} \alpha_n u_n$ . Formellement, on trouve que :

$$u(x, 0) = \sum_{n \in \mathbb{Z}} \alpha_n u_n(x, 0) = \sum_{n \in \mathbb{Z}} \alpha_n \sin\left(\frac{n\pi x}{L}\right) = u_0(x).$$

Le problème admet une solution ssi  $u_0$  coïncide avec sa série de Fourier et si cette dernière est impaire. Comme  $u_0(0) = 0$ , on prolonge  $u_0$  en une fonction impaire de classe  $\mathcal{C}^1$  sur  $[-L, L]$ . La condition  $u_0(L) = 0$  permet de prolonger  $u_0$  par périodicité à  $\mathbb{R}$  en une fonction périodique de période  $2L > 0$ , continue et  $\mathcal{C}^1$  par morceaux. On en déduit que la série de Fourier de  $u_0$  est absolument convergente sur  $\mathbb{R}$  vers  $u_0$  et que la série des coefficients est dans  $\ell^1$ . Il reste à vérifier qu'on peut dériver sous le signe somme dans  $u = \sum_{n \in \mathbb{Z}} \alpha_n u_n$  lorsque  $(\alpha_n)_{n \in \mathbb{Z}} \in \ell^1$ . Soit  $T > 0$ . On a :

$$|u_n(x, y)| \leq |\alpha_n|, \quad \forall (x, t) \in \mathbb{R} \times [0, T]$$

et la série majorante est convergente par hypothèse sur  $u_0$ . On en déduit que la série de fonctions continues  $\sum \alpha_n u_n$  est uniformément convergente sur tout compact de  $[0, L] \times \mathbb{R}^+$  donc de somme continue sur  $[0, L] \times \mathbb{R}^+$ . Soit  $\delta \in ]0, M[$ . On a :  $\forall (x, y) \in [0, L] \times [\delta, M]$ ,

$$|u_n(x, y)| \leq C \frac{\sinh(\frac{n\pi}{L}(M-y))}{\sinh(\frac{n\pi M}{L})} \leq C \frac{\sinh(\frac{n\pi}{L}(M-\delta))}{\sinh(\frac{n\pi M}{L})} \underset{n \rightarrow +\infty}{\sim} C e^{-\frac{n\pi}{L}\delta}$$

ainsi que :  $\forall k, \ell > 0$ ,

$$\left| \frac{\partial^{k+\ell}}{\partial x^k \partial y^\ell} u_n(x, y) \right| \leq C \left( \frac{n\pi}{L} \right)^{k+\ell} e^{-\frac{n\pi}{L}\delta}, \quad \forall (x, y) \in [0, L] \times [\delta, M]$$

où la série majorante est convergente. On en déduit que la série des dérivées partielles  $\sum \frac{\partial^{k+\ell}}{\partial t^k \partial x^\ell} u_n$  est uniformément convergente sur tout compact de  $[0, L] \times ]0, M]$ , donc que la somme de la série  $\sum \alpha_n u_n$  est de classe  $\mathcal{C}^\infty$  sur  $[0, L] \times ]0, M]$ . On conclut par unicité de la série de Fourier de  $u_0$ .  $\square$

**Proposition 2.3.8.** *Si  $u_0 \in \mathcal{C}^2([0, L], \mathbb{R})$ , alors (2.8) est l'unique solution de (2.7).*

*Démonstration.* C'est une conséquence du principe du maximum.  $\square$

## Le principe du maximum

**Définition 2.3.1** (Principe du maximum). On appelle principe du maximum continu le fait que si  $f \geq 0$  alors le minimum de la solution  $u$  du problème (2.24) est atteint sur le bord du domaine de définition de  $u$ .

**Proposition 2.3.9** (Le principe du maximum). *Soit  $\Omega \subset \mathbb{R}^n$  un ouvert borné et soit  $f \in \mathcal{C}^2(\Omega)$ .*

a) Si  $\Delta f > 0$  dans  $\Omega$  alors :

$$\forall x \in \Omega, \quad f(x) < \max_{y \in \partial\Omega} f(y).$$

b) Si  $\Delta f = 0$  dans  $\Omega$  ( $f$  est dite harmonique), alors

$$\forall x \in \Omega, \quad \min_{y \in \partial\Omega} f(y) \leq f(x) \leq \max_{y \in \partial\Omega} f(y).$$

*Démonstration.* On remarque que  $\partial\Omega = \overline{\Omega} \cap \Omega^c$  est fermé comme intersection de fermés, et borné par hypothèse sur  $\Omega$ , donc c'est un compact de  $\mathbb{R}^n$  et  $f$  y atteint ses bornes.

a) Soit  $\Delta f > 0$  dans  $\Omega$ . On suppose qu'il existe  $x_0 \in \Omega$  t.q. :

$$f(x_0) \geq \max_{y \in \partial\Omega} f(y).$$

Comme  $\overline{\Omega}$  est un compact de  $\mathbb{R}^n$  (en dimension finie), on peut supposer que

$$f(x_0) = \max_{y \in \overline{\Omega}} f(x) \geq \max_{y \in \partial\Omega} f(x)$$

Comme  $\Omega$  est ouvert, on a  $\nabla f(x_0) = 0$ . Soit  $r > 0$  t.q., si  $B_r(x_0)$  est la boule ouverte de centre  $x_0$  et de rayon  $r > 0$ , on ait  $B_r(x_0) \subset \Omega$ , et soit  $x \in B_r(x_0) \setminus \{x_0\}$ . On pose :

$$\varphi(t) = f(x_0 + t(x - x_0)), \quad \forall t \in [0, 1].$$

Alors  $f \in \mathcal{C}^2(\Omega, \mathbb{R}) \Rightarrow \varphi \in \mathcal{C}^2([0, 1], \mathbb{R})$  et on a

$$\varphi(1) = f(x) = f(x_0) + \int_0^1 (1-t)f''(x_0 + t(x-x_0)) \cdot (x-x_0)^2 dt \leq f(x_0)$$

$$\Rightarrow \int_0^1 (1-t)f''(x_0 + t(x-x_0)) \cdot (x-x_0)^2 dt \leq 0.$$

Soit  $u = \frac{1}{\|x-x_0\|}(x-x_0)$ . On a, par bilinéarité de la différentielle :

$$\int_0^1 (1-t)f''(x_0 + t(x-x_0)) \cdot u^2 dt \leq 0.$$

On en déduit, par continuité de  $f''$  :

$$\lim_{x \rightarrow x_0} \int_0^1 (1-t)f''(x_0 + t(x-x_0)) \cdot u^2 dt = \int_0^1 (1-t) dt f''(x_0) \cdot u^2 = \frac{1}{2} f''(x_0) \cdot u^2 \leq 0.$$

Ceci étant vrai pour tout  $x \in B_r(x_0)$ , on en déduit, par définition des dérivées partielles d'ordre 2 :

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) \leq 0, \quad i, j = 1, \dots, n.$$

en contradiction avec  $\Delta f > 0$  dans  $\Omega$  et  $x_0 \in \Omega$ .

- b) Soit  $\Delta f = 0$  dans  $\Omega$ . Il suffit de montrer que  $f \leq \max_{y \in \partial\Omega} f$  dans  $\Omega$ , l'autre inégalité étant alors obtenue en considérant  $-f$ . Soit  $\varepsilon > 0$ . On pose :

$$g_\varepsilon(x) = f(x) + \varepsilon \|x\|^2 \quad \forall x \in \Omega.$$

Alors :

$$\Delta g_\varepsilon = \Delta f + 2n\varepsilon = 2n\varepsilon > 0 \quad \text{dans } \Omega.$$

On remarque que  $\partial\Omega$  étant borné, il existe  $M > 0$  t.q.  $\|x\| \leq M$ ,  $\forall x \in \partial\Omega$ . De a), on déduit que :  $\forall x \in \Omega$ ,

$$g_\varepsilon(x) < \max_{y \in \partial\Omega} g \leq \max_{y \in \partial\Omega} f + \varepsilon M$$

et donc :  $\forall x \in \Omega$ ,

$$f(x) \leq g_\varepsilon(x) < \max_{y \in \partial\Omega} f + \varepsilon M.$$

On en déduit :

$$\sup_{x \in \Omega} f(x) \leq \max_{y \in \partial\Omega} f + \varepsilon M.$$

Ceci étant vrai pour tout  $\varepsilon > 0$ , il en résulte que :

$$\sup_{x \in \Omega} f(x) \leq \max_{y \in \partial\Omega} f.$$

□

**Proposition 2.3.10** (Principe du maximum faible). *Soit  $\Omega \subset \mathbb{R}^N$  un ouvert borné de  $\mathbb{R}^N$  et soit  $u \in H^1(\Omega)$  solution de :*

$$\begin{cases} \Delta u \geq 0 & \text{p.p. dans } \Omega, \\ u^+ \in H_0^1(\Omega). \end{cases}$$

*Alors :  $u \leq 0$  p.p. dans  $\Omega$ .*

**Lemme 2.3.11** (Lemme préliminaire). *Soit  $f \in H^{-1}(\Omega) = (H_0^1(\Omega))'$  t.q.  $f \geq 0$  p.p. dans  $\mathcal{D}'(\Omega)$ , i.e. :*

$${}_{H^{-1}}\langle f, \varphi \rangle_{H_0^1} \geq 0, \quad \forall \varphi \in \mathcal{D}(\Omega, \mathbb{R}^+).$$

*Alors*

$$\forall v \in H_0^1(\Omega), \quad {}_{H^{-1}}\langle f, v^+ \rangle_{H_0^1} \geq 0.$$

*Démonstration.* Soit  $v \in H_0^1(\Omega)$  et soit  $(\varphi_n)_{n \geq 0} \in \mathcal{D}(\Omega)^{\mathbb{N}}$  t.q.  $\varphi_n \xrightarrow{n \rightarrow +\infty} v$  dans  $H_0^1(\Omega)$ . On a :

$$\|\nabla v^+ - \nabla \varphi_n^+\|_2 \leq \|\nabla v - \nabla \varphi_n\|_2 \xrightarrow{n \rightarrow +\infty} 0$$

i.e.  $\varphi_n^+ \xrightarrow{n \rightarrow +\infty} v^+$  dans  $H_0^1(\Omega)$ . Soit  $n \geq 0$ . On remarque que  $\varphi_n^+$  est à support compact car fermé dans le support de  $\varphi_n$  qui est fermé. Donc il existe une suite régularisante  $(\rho_\delta)_{\delta > 0}$  t.q.  $\rho_\delta \star \varphi_n^+ \in \mathcal{D}(\Omega)$  dès que  $\delta \in ]0, \delta'_n[$  est assez petit, i.e. inférieur à la distance du support de  $\varphi_n^+$  au bord  $\partial\Omega$ , et  $\rho_\delta \star \varphi_n^+ \xrightarrow{\delta \rightarrow 0} \varphi_n^+$  dans  $H_0^1(\Omega)$ . Par définition de la convolution :  $\rho_\delta \star \varphi_n^+ \geq 0$  dans  $\Omega$ ,  $\forall \delta \in ]0, \delta'_n[$ ,  $\forall n \geq 0$ . Par hypothèse sur  $f$  :

$$H^{-1}\langle f, \rho_\delta \star \varphi_n^+ \rangle_{H_0^1} \geq 0, \quad \forall \delta \in ]0, \delta'_n[.$$

On conclut après extraction d'une suite diagonale  $(\rho_{\delta_n})_{n \geq 0}$  t.q.  $\rho_{\delta_n} \star \varphi_n^+ \xrightarrow{n \rightarrow +\infty} v^+$  dans  $H_0^1(\Omega)$  :

$$H^{-1}\langle f, v^+ \rangle_{H_0^1} = \lim_{n \rightarrow +\infty} \underbrace{H^{-1}\langle f, \rho_{\delta_n} \star \varphi_n^+ \rangle_{H_0^1}}_{\geq 0} \geq 0.$$

□

*Démonstration de la Proposition 2.3.10.* Soit  $f = -\Delta u$ . Alors  $f \in H^{-1}(\Omega)$  et  $f \leq 0$  p.p. dans  $\Omega$ . Par hypothèse sur  $u^+$  :

$$H^{-1}\langle f, u^+ \rangle_{H_0^1} = \int_{\Omega} \nabla u \nabla u^+ dx = \int_{\Omega} |\nabla u^+|^2 dx \geq 0 \Rightarrow \nabla u^+ = 0 \quad \text{p.p. dans } \Omega$$

i.e.  $u^+ = 0$  p.p. dans  $\Omega$ .

□

*Remarque 9.* On a utilisé le fait que  $H_0^1(\Omega)$  est un espace de Hilbert pour le produit scalaire  $(u, v) \mapsto \int_{\Omega} \nabla u \nabla v dx$ .

**Corollaire 2.3.12.** Soit  $\Omega \subset \mathbb{R}^N$  un ouvert borné et soit  $g \in H^1(\Omega)$ . Si  $u \in H^1(\Omega)$  est solution de :

$$\begin{cases} \Delta u = 0 & \text{dans } \Omega, \\ u - g \in H_0^1(\Omega), \end{cases}$$

alors  $|u| \leq \|g\|_{\infty}$  p.p. dans  $\Omega$ .

*Démonstration.* Si  $g \notin L^\infty(\Omega)$ , il n'y a rien à montrer. On suppose que  $g \in L^\infty(\Omega)$ . On remarque que  $(u - \|g\|_\infty)^+ \in H_0^1(\Omega)$ . D'après la Proposition 2.3.10 appliquée à  $u - \|g\|_\infty$  avec  $f = 0$ , on a :  $u \leq \|g\|_\infty$  p.p. dans  $\Omega$ . De même,  $(u + \|g\|_\infty)^- \in H_0^1(\Omega)$ . D'après la Proposition 2.3.10 appliquée à  $u + \|g\|_\infty$  avec  $f = 0$ , on a :  $u \geq -\|g\|_\infty$  p.p. dans  $\Omega$ .  $\square$

**Proposition 2.3.13.** *Soit  $\Omega \subset \mathbb{R}^N$  un ouvert borné de  $\mathbb{R}^N$  et soit  $f, c \in L^\infty(\Omega)$ . On suppose que  $c \geq \eta > 0$  p.p. dans  $\Omega$  pour une constante  $\eta > 0$ . Soit  $u \in H_0^1(\Omega)$  solution de :*

$$-\Delta u + cu = f \quad \text{dans } \mathcal{D}'(\Omega)$$

Alors

$$\|u\|_\infty \leq \frac{\|f\|_\infty}{\eta}.$$

*Démonstration.* Soit  $k \in \mathbb{R}^+$ . On remarque que  $(u - k)^+ \in H_0^1(\Omega)$ . On en déduit :

$$\int_{\Omega} \nabla u \nabla (u - k)^+ dx + \int_{\Omega} cu(u - k)^+ dx = \int_{\Omega} f(u - k)^+ dx.$$

avec

$$\begin{aligned} \int_{\Omega} \nabla u \nabla (u - k)^+ dx &= \int_{\Omega} \nabla(u - k) \nabla (u - k)^+ dx = \int_{\Omega} \nabla(u - k)^+ \nabla(u - k)^+ dx = \\ &= \int_{\Omega} |\nabla(u - k)^+|^2 dx \geq 0, \\ \int_{\Omega} cu(u - k)^+ dx &= \int_{\Omega} c(u - k)(u - k)^+ dx + k \int_{\Omega} c(u - k)^+ dx = \\ &= \int_{\Omega} c|(u - k)^+|^2 dx + k \int_{\Omega} c(u - k)^+ dx \\ &\geq \eta \int_{\Omega} |(u - k)^+|^2 dx + k \int_{\Omega} c(u - k)^+ dx. \end{aligned}$$

On en déduit :

$$0 \leq \int_{\Omega} |\nabla(u - k)^+|^2 dx + \eta \int_{\Omega} |(u - k)^+|^2 dx \leq \int_{\Omega} (f - ck)(u - k)^+ dx.$$

On pose :

$$k = \frac{\|f\|_\infty}{\eta}.$$



Alors :

$$ck \geq \|f\|_\infty \Rightarrow f - ck \leq 0 \quad \text{p.p. dans } \Omega \Rightarrow \int_{\Omega} (f - ck)(u - k)^+ dx \leq 0$$

Il en résulte :

$$\eta \int_{\Omega} |(u - k)^+|^2 dx \leq 0$$

i.e. :  $(u - k)^+ = 0$  p.p. Le même raisonnement avec  $(u + k)^-$  conduit à  $u \geq -\frac{\|f\|_\infty}{\eta}$  p.p. dans  $\Omega$ .  $\square$

## 2.4 Calcul approché par les différences finies en dimension 1

Soit  $f \in \mathcal{C}([0, 1])$ . On cherche  $u : [0, 1] \rightarrow \mathbb{R}$  solution de :

$$-u''(x) = f(x), \quad x \in ]0, 1[, \quad u(0) = u(1) = 0. \quad (2.9)$$

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. On se donne une subdivision de  $[0, 1]$  :

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

supposée régulière de pas  $h = \frac{1}{N+1} > 0$  pour simplifier, et on veut calculer des approximations  $u_i \sim u(x_i)$ ,  $i = 0, \dots, N+1$ , en chaque point  $x_i$  de la subdivision. Pour cela, on considère le problème : trouver  $u^{(N)} \in \mathbb{R}^N$ , de composantes  $u_1, \dots, u_N$ , solution du système :

$$\begin{cases} \frac{1}{h^2}(-u_{i-1} + 2u_i - u_{i+1}) = f_i, & i = 1, \dots, N, \\ u_0 = u_{N+1} = 0 \end{cases} \quad (2.10)$$

où  $f_i = f(x_i)$ ,  $i = 1, \dots, N$ , sont donnés au second membre.

**Définition 2.4.1** (Erreur de consistance). On pose :

$$\varepsilon_i^{(N)} = \frac{1}{h^2}(-u(x_{i-1}) + 2u(x_i) - u(x_{i+1})) - f(x_i), \quad i = 1, \dots, N$$

On appelle erreur de consistance du schéma (2.10) la quantité

$$\|\varepsilon^{(N)}\|_\infty = \max_{1 \leq i \leq N} |\varepsilon_i^{(N)}|, \quad N > 0.$$

Le schéma (2.10) est dit consistant si  $\lim_{N \rightarrow +\infty} \|\varepsilon^{(N)}\|_\infty = 0$ .

Par application de la formule de Taylor, on trouve :

$$\varepsilon^{(N)} = O(h) = O\left(\frac{1}{N+1}\right) \xrightarrow{N \rightarrow +\infty} 0,$$

i.e. le schéma (2.10) est consistant.

*Remarque 10.* Si  $u \in \mathcal{C}^4(]0, 1[)$  alors il existe une constante  $C > 0$  indépendante de  $u$  t.q. :

$$\|\varepsilon^{(N)}\|_\infty \leq C \|u^{(4)}\| h^2.$$

**Définition 2.4.2.** Le schéma (2.10) est dit convergent si

$$\lim_{N \rightarrow +\infty} \max_{1 \leq i \leq N} |u_i^{(N)} - u(x_i)| = 0.$$

**Proposition 2.4.1.** *Le schéma (2.10) est convergent.*

*Démonstration.* Soit  $(\mu^{(N)})_{1 \leq i \leq N} \in \mathbb{R}^N$  et soit  $z^{(N)} \in \mathbb{R}^N$  solution du schéma :

$$\begin{cases} \frac{1}{h^2}(-z_{i-1} + 2z_i - z_{i+1}) = \mu_i^{(N)}, & i = 1, \dots, N, \\ z_0 = z_{N+1} = 0 \end{cases} \quad (2.11)$$

qui se réécrit sous forme matricielle :

$$\frac{1}{h^2} A_N z^{(N)} = \mu^{(N)} \quad (2.12)$$

où  $A_N \in \mathbb{R}^{N \times N}$  est la matrice carrée d'ordre  $N$  de coefficients :

$$a_{ij}^{(N)} = \begin{cases} 2 & \text{si } i = j, \\ -1 & \text{si } |i - j| = 1, \\ 0 & \text{si } |i - j| > 1. \end{cases} \quad (2.13)$$

On remarque que  $A_N$  est symétrique réelle, donc diagonalisable dans une bon, de valeurs propres dans  $\mathbb{R}$ . Soit  $x \in \mathbb{R}^N$  un vecteur propre de  $A_N$  de valeur propre  $\lambda \in \mathbb{R}$ . L'équation  $A_N x = \lambda x$  s'écrit composante par composante :

$$-x_{i-1} + (2 - \lambda)x_i - x_{i+1} = 0, \quad i = 1, \dots, N \quad (2.14)$$

$$x_0 = x_{N+1} = 0. \quad (2.15)$$

L'équation caractéristique associée s'écrit :

$$r^2 + (\lambda - 2)r + 1 = 0. \quad (2.16)$$

On vérifie que (2.14)–(2.15) admet des solutions non nulles ssi le discriminant de (2.16) est  $\leq 0$ , i.e. ssi  $|\lambda - 2| < 2$ . On pose :  $\lambda = 2 + 2 \cos \theta$ ,  $\theta \in ]0, \pi[$ . Alors (2.16) admet les racines

$$r^{\pm} = -e^{\pm i\theta}.$$

On en déduit

$$x_k = a \cos(k\theta) + b \sin(k\theta), \quad 0 \leq k \leq N + 1$$

avec  $x_0 = a = 0$ . On en déduit :

$$x_{N+1} = b \sin((N + 1)\theta) = 0 \iff \theta \in \frac{\pi}{N + 1} \mathbb{N}.$$

Plus précisément :

$$0 < \theta < \pi \Rightarrow \theta = \frac{p\pi}{(N + 1)} = \theta_p, \quad 1 \leq p \leq N$$

On en déduit que  $A_N$  admet  $N$  valeurs propres distinctes  $> 0$  :

$$\begin{aligned} 4 > \lambda_1^{(N)} &= 4 \left( \cos \left( \frac{\pi}{N + 1} \right) \right)^2 > \dots > \lambda_p^{(N)} = 4 \left( \cos \left( \frac{p\pi}{N + 1} \right) \right)^2 > \\ &> \dots > \lambda_N^{(N)} = 4 \left( \cos \left( \frac{N\pi}{N + 1} \right) \right)^2 = 4 \left( \sin \left( \frac{\pi}{N + 1} \right) \right)^2 > 0. \end{aligned} \quad (2.17)$$

En particulier  $A_N$  est inversible et donc

$$z^{(N)} = h^2 A_N^{-1} \mu^{(N)} \Rightarrow \|z^{(N)}\|_2 \leq h^2 \underbrace{\|A_N^{-1}\|_2}_{=\rho(A_N^{-1})} \|\mu^{(N)}\|_2$$

où  $\|A_N^{-1}\|_2 = \rho(A_N^{-1})$  par symétrie de  $A_N^{-1}$ . Les valeurs propres de  $A_N^{-1}$  sont :

$$0 < \frac{1}{\lambda_1^{(N)}} < \dots < \frac{1}{\lambda_N^{(N)}} \Rightarrow \rho(A_N^{-1}) = \frac{1}{\lambda_N^{(N)}} = \frac{1}{4 \left( \sin \left( \frac{\pi}{N+1} \right) \right)^2} = \frac{1}{4(\sin(\pi h))^2}$$

et donc :

$$\|z^{(N)}\|_2 \leq \left( \frac{h}{2 \sin(\pi h)} \right)^2 \|\mu^{(N)}\|_2. \quad (2.18)$$

Soit  $X^{(N)} \in \mathbb{R}^N$  le vecteur de composantes  $X_i^{(N)} = x_i$ ,  $1 \leq i \leq N$ . Si  $z_i^{(N)} = u(x_i) - u_i$ ,  $1 \leq i \leq N$ , alors  $z^{(N)} = u(X^{(N)}) - u^{(N)}$  et  $\mu^{(N)} = \varepsilon^{(N)}$  avec

$$\|\varepsilon^{(N)}\|_2 \leq \sqrt{N} \|\varepsilon^{(N)}\|_\infty \leq C\sqrt{N}h \leq C\sqrt{h}$$

donc

$$\begin{aligned} \|u(X^{(N)}) - u^{(N)}\|_\infty &\leq \|u(X^{(N)}) - u^{(N)}\|_2 \leq C\sqrt{h} \left( \frac{h}{2\sin(\pi h)} \right)^2 \\ &\leq C\sqrt{h} \leq \frac{C}{\sqrt{N}} \xrightarrow{N \rightarrow +\infty} 0 \end{aligned}$$

□

## Autres conditions aux limites

### Conditions de Dirichlet non homogènes

On considère les conditions sur le bord non homogènes :

$$u(0) = a, \quad u(1) = b.$$

Le schéma (2.10) reste inchangé à l'exception du second membre qui devient :

$$f_i = \begin{cases} f(x_1) + a & \text{si } i = 1, \\ f(x_N) + b & \text{si } i = N, \\ f(x_i) & \text{si } 1 < i < N, \end{cases} \quad (2.19)$$

### Conditions de Neumann et de Fourier

On considère les conditions sur le bord :

- (i) de type Neumann en  $x = 0$  :  $u'(0) = a$ ,
- (ii) de type Fourier en  $x = 1$  :  $u'(1) + \alpha u(1) = b$  avec  $\alpha > 0$ .

Les premiers termes des développements en série de Taylor de  $u'(0)$  et  $u'(1)$  suggèrent de choisir :

$$u_0 = u_1 - ah, \quad u_{N+1} = \frac{u_N + bh}{1 + \alpha h}.$$

Alors le schéma (2.10) devient : trouver  $u_0, u_1, \dots, u_N, u_{N+1}$  t.q.

$$\begin{cases} \frac{1}{h}(u_0 - u_1) = -a, \\ \frac{1}{h^2}(-u_{i-1} + 2u_i - u_{i+1}) = f_i, & i = 1, \dots, N, \\ \frac{1}{h}(-u_N + (1 + \alpha h)u_{N+1}) = b. \end{cases} \quad (2.20)$$

### Schéma numérique bien posé

**Définition 2.4.3.** Le schéma numérique (2.10) est dit bien posé s'il admet une unique solution.

Le schéma (2.10) se réécrit sous forme matricielle : trouver  $U_N \in \mathbb{R}^N$  solution de :

$$A_N U_N = f \quad (2.21)$$

où  $A_N \in \mathbb{R}^{N \times N}$  est la matrice carrée introduite dans (2.12) et définie par (2.13).

**Proposition 2.4.2.** La matrice (2.13) du système (2.21) est définie positive.

*Démonstration.* Soit  $x \in \mathbb{R}^N$ . On a :

$$Ax \cdot x = 2 \sum_{i=1}^N x_i^2 - 2 \sum_{i=1}^{N-1} x_i x_{i+1} = \sum_{i=1}^{N-1} (x_i - x_{i+1})^2 + x_1^2 + x_N^2 \geq 0.$$

Si  $Ax \cdot x = 0$ , alors

$$x_1 = x_N = 0 \quad \text{et} \quad x_i = x_{i+1}, \quad i = 1 \dots N \Rightarrow x = 0.$$

□

**Corollaire 2.4.3.** La matrice (2.13) du système (2.21) est inversible.

*Démonstration.* C'est une conséquence directe de la Proposition 2.4.2. □

On en déduit que le schéma numérique (2.10) est bien posé.

**Exemple 1.** Le problème (2.20) se réécrit sous forme matricielle : trouver  $U_h \in \mathbb{R}^{N+2}$ . solution de

$$\frac{1}{h^2} A_h U_h = f_h$$

où  $A_h = (a_{ij}^{(h)})_{0 \leq i, j \leq N+1}$  est la matrice carrée d'ordre  $N + 2$  définie par :

$$a_{ij}^{(h)} = \begin{cases} 1 & \text{si } i = j = 0, \\ -1 & \text{si } (i, j) \in \{(1, 0), (N + 1, N)\} \\ 2 & \text{si } 1 \leq i = j \leq N \\ -1 & \text{si } |i - j| = 1, 1 \leq i, j \leq N \\ 1 + \alpha h & \text{si } i = j = N + 1, \\ 0 & \text{si } |i - j| > 1, \end{cases}$$

et où le second membre  $f_h = (f_i^{(h)})_{0 \leq i \leq N+1} \in \mathbb{R}^{N+2}$  est défini par :

$$f_i^{(h)} = \begin{cases} -\frac{a}{h} & \text{si } i = 0, \\ f(x_i) & \text{si } 1 \leq i \leq N, \\ \frac{b}{h} & \text{si } i = N + 1. \end{cases}$$

Soit  $x \in \mathbb{R}^{N+2}$ . On a :

$$A_h x \cdot x = \sum_{i=0}^N (x_i - x_{i+1})^2 + \alpha h x_{N+1}^2 \geq 0$$

car  $\alpha h > 0$ . Si  $Ax \cdot x = 0$ , alors :

$$x_0 = x_1 = \dots = x_{N+1} = 0$$

i.e.  $x = 0$  et  $A_h$  est définie positive, donc inversible.

**Exemple 2.** On considère le problème aux limites avec conditions de Neumann :

$$-u'' = f \quad \text{dans } ]0, 1[, \quad u'(0) = u'(1) = 0. \quad (2.22)$$

Les approximations

$$u'(0) \sim \frac{1}{h}(u_1 - u_0), \quad u'(1) \sim \frac{1}{h}(u_{N+1} - u_N)$$

conduisent au schéma numérique :

$$\begin{cases} \frac{1}{h^2}(-u_{i-1} + 2u_i - u_{i+1}) = f_i, & i = 1, \dots, N, \\ u_0 = u_1, \\ u_{N+1} = u_N. \end{cases} \quad (2.23)$$

Sous forme matricielle, le schéma (2.23) se réécrit : trouver  $U_h \in \mathbb{R}^{N+2}$  solution de

$$\frac{1}{h^2} A_h U_h = f_h$$

où  $A_h = (a_{ij}^{(h)})_{0 \leq i, j \leq N+1}$  est la matrice carrée d'ordre  $N + 2$  définie par :

$$a_{ij}^{(h)} = \begin{cases} 1 & \text{si } i = j = 0, \\ 2 & \text{si } 1 \leq i = j \leq N, \\ -1 & \text{si } |i - j| = 1, \\ 1 & \text{si } i = j = N + 1, \\ 0 & \text{si } |i - j| > 1 \end{cases}$$

et où le second membre  $f_h = (f_i^{(h)})_{0 \leq i \leq N+1} \in \mathbb{R}^{N+2}$  est le vecteur de composantes :

$$f_i^{(h)} = \begin{cases} 0 & \text{si } i = 0, \\ f(x_i) & \text{si } 1 \leq i \leq N, \\ 0 & \text{si } i = N + 1. \end{cases}$$

Soit  $x \in \mathbb{R}^{N+2}$ . On a :

$$A_h x \cdot x = \sum_{i=0}^N (x_i - x_{i+1})^2 \geq 0$$

et

$$A_h x \cdot x = 0 \Rightarrow x_0 = \dots = x_{N+1}.$$

On vérifie directement que

$$A_h \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = 0$$

i.e.  $A_h$  est semi-définitive positive et non inversible. De fait, le problème (2.22) admet les constantes pour solutions.

**Définition 2.4.4** (Matrice monotone). Une matrice réelle est dite monotone si elle est inversible, d'inverse à coefficients  $\geq 0$ .

**Proposition 2.4.4** (Caractérisation des matrices monotones). Une matrice réelle  $A \in \mathbb{R}^{N \times N}$  est monotone ssi :  $\forall v \in \mathbb{R}^N$ ,

$$Av \geq 0 \Rightarrow v \geq 0.$$

(Les inégalités s'entendent composante par composante)

*Démonstration.* Soit  $A \in \mathbb{R}^{N \times N}$ .  $\Rightarrow$  Soit  $A \in \mathbb{R}^{N \times N}$  monotone et soit  $v \in \mathbb{R}^N$ . On suppose que  $Av \geq 0$ . Alors :

$$v_i = \underbrace{A_{ik}^{-1}}_{\geq 0} \underbrace{(Av)_k}_{\geq 0} \geq 0, \quad \forall i \in [[1, N]].$$

$\Leftarrow$  Inversement, on suppose que :  $\forall v \in \mathbb{R}^N$ ,

$$Av \geq 0 \Rightarrow v \geq 0.$$

Soit  $v \in \mathbb{R}^N$  t.q.  $Av = 0$ . Alors :

$$Av = 0 \Rightarrow v \geq 0 \quad \text{et} \quad -Av = 0 \Rightarrow -v \geq 0$$

donc  $v = 0$ , i.e.  $A$  est inversible. Soit  $i \in [[1, N]]$ . On note  $e_i$  le  $i$ ème vecteur de la base canonique de  $\mathbb{R}^N$ . Par hypothèse sur  $A$  :

$$e_i = A(A^{-1}e_i) \geq 0 \Rightarrow A^{-1}e_i \geq 0.$$

On en déduit :  $(A^{-1}e_i)_j = A_{ij}^{-1} \geq 0, \forall j \in [[1, N]]$ . Ceci étant vrai pour tout  $i \in [[1, N]]$ , il en résulte :  $A^{-1} \geq 0$ , i.e.  $A$  est monotone.  $\square$

## Principe du maximum discret

On considère le problème : trouver  $u$  solution de

$$\begin{cases} -u''(x) + c(x)u(x) = f(x) & \text{si } 0 < x < 1, \\ u(0) = 0, \\ u(1) = 0, \end{cases} \quad (2.24)$$

où  $c \in \mathcal{C}([0, 1], \mathbb{R}^+)$  et  $f \in \mathcal{C}([0, 1], \mathbb{R})$ . L'analogie du problème discrétisé (2.10) s'écrit :

$$\begin{cases} \frac{1}{h^2}(-u_{i-1} + 2u_i - u_{i+1}) + c_i u_i = f_i, & i = 1, \dots, N, \\ u_0 = u_{N+1} = 0 \end{cases} \quad (2.25)$$



où  $c_i = c(x_i)$ ,  $f_i = f(x_i)$ ,  $i = 1, \dots, N$ , sont donnés. Sous forme matricielle, le problème (2.25) se réécrit :

$$\frac{1}{h^2} A_h U_h = f_h$$

où  $A_h = (a_{ij}^{(h)})_{1 \leq i, j \leq N} \in \mathbb{R}^{N \times N}$  est la matrice carrée d'ordre  $N$  définie par :

$$a_{ij}^{(h)} = \begin{cases} 2 + c_i h^2 & \text{si } i = j, \\ -1 & \text{si } |i - j| = 1, \\ 0 & \text{si } |i - j| > 1, \end{cases} \quad (2.26)$$

**Proposition 2.4.5.** *Soit  $c = (c_1, \dots, c_N)^T \in \mathbb{R}^N$  t.q.  $c_i \geq 0, \forall i \in [[1, N]]$ . Alors la matrice  $A_h$  définie par (2.26) est symétrique, définie positive et donc inversible.*

*Démonstration.* La matrice  $A_h$  est symétrique par définition. Soit  $x \in \mathbb{R}^N$ . On a :

$$A_h x \cdot x = \sum_{i=1}^{N-1} (x_i - x_{i+1})^2 + x_1^2 + x_N^2 + h^2 \sum_{i=1}^N c_i x_i^2 \underset{c \geq 0}{\geq} 0$$

On suppose que  $A_h x \cdot x = 0$ . Alors :

$$x_1 = x_N = 0 \quad \text{et} \quad x_i = x_{i+1}, \quad i = 1, \dots, N \Rightarrow x_1 = \dots = x_N = 0.$$

□

*Remarque 11.* Si la solution  $u$  de (2.24) vérifie le principe du maximum, on souhaite qu'il en soit de même pour la solution approchée.

**Lemme 2.4.6.** *Soit  $c = (c_1, \dots, c_N)^T \in \mathbb{R}^N$  t.q.  $c_i \geq 0, \forall i \in [[1, N]]$ . Alors la matrice  $A_h$  définie par (2.26) est monotone.*

*Démonstration.* On commence par remarquer que la matrice  $A_h$  est inversible d'après la Proposition 2.4.5. Soit  $v \in \mathbb{R}^N$  t.q.  $A_h v \geq 0$ . Soit  $i_0 \in [[1, N]]$  t.q.  $v_{i_0} = \min_{1 \leq j \leq N} v_j$ . Si  $i_0 = 1$ , alors :

$$(2 + h^2 c_1) v_1 \geq v_2 \geq v_1 \Rightarrow \underbrace{(1 + h^2 c_1) v_1}_{>0} \geq 0 \Rightarrow v_1 \geq 0.$$

Si  $i_0 = N$ , alors :

$$(2 + h^2 c_N) v_N \geq v_{N-1} \geq v_N \Rightarrow \underbrace{(1 + h^2 c_N) v_N}_{>0} \geq 0 \Rightarrow v_N \geq 0.$$

On suppose que  $1 < i_0 < N$ . On a :

$$(2 + h^2 c_{i_0})v_{i_0} \geq v_{i_0-1} + v_{i_0+1} \geq 2v_{i_0} \Rightarrow h^2 c_{i_0} v_{i_0} \geq 0.$$

Si  $c_{i_0} > 0$ , alors  $v_{i_0} \geq 0$ . Sinon, si  $v_{i_0} = \min_{1 \leq j \leq N} v_j \Rightarrow c_{i_0} = 0$ , alors

$$v_{i_0} - v_{i_0-1} \geq v_{i_0+1} - v_{i_0} \geq 0 \Rightarrow v_{i_0} = v_{i_0-1} = \min_j v_j.$$

Par récurrence sur  $j \in [[1, i_0]]$ , on en déduit :

$$v_{i_0} = v_{i_0-1} = \dots = v_1 = \min_j v_j \geq 0.$$

□

**Définition 2.4.5** (Ordre du schéma). On dit que le schéma (2.25), resp. (2.10), est d'ordre  $p$  s'il existe  $C > 0$  t.q.

$$\|\varepsilon^{(N)}\|_\infty \leq Ch^p$$

où  $\varepsilon^{(N)} = (\varepsilon_i^{(N)})_{1 \leq i \leq N}$  est l'erreur de consistance du schéma (2.25) resp. (2.10), définie par :  $\forall i \in [[1, N]]$ ,

$$\varepsilon_i^{(N)} = \frac{1}{h^2}(-u(x_{i-1}) + 2u(x_i) - u(x_{i+1})) + c(x_i)u(x_i) - f(x_i). \quad (2.27)$$

resp. par la Définition 1.3.2.

**Proposition 2.4.7.** Si la solution  $u$  de (2.24), resp. de (2.9), est de classe  $C^4$ , alors :

$$\|\varepsilon^{(N)}\|_\infty \leq \frac{h^2}{12} \sup_{0 \leq x \leq 1} |u^{(4)}|. \quad (2.28)$$

*Démonstration.* D'après la formule de Taylor avec reste intégral :  $\forall x \in [0, 1]$ ,

$$\begin{aligned} u(x+h) + u(x-h) - 2u(x) &= hu'(x) + h^2 u''(x) + \int_x^{x+h} \frac{(x+h-t)^3}{6} u^{(4)}(t) dt + \\ &\quad + \int_x^{x-h} \frac{(x-h-t)^3}{6} u^{(4)}(t) dt \\ &\Rightarrow \left| \frac{1}{h^2}(-u(x-h) + 2u(x) - u(x+h)) + c(x)u(x) - f(x) \right| = \\ &= \frac{1}{h^2} \left| \int_x^{x+h} \frac{(x+h-t)^3}{6} u^{(4)}(t) dt + \int_x^{x-h} \frac{(x-h-t)^3}{6} u^{(4)}(t) dt \right| \\ &\leq \frac{2}{h^2} \left( \frac{h^4}{24} \|u^{(4)}\|_\infty \right) = \frac{h^2}{12} \|u^{(4)}\|_\infty. \end{aligned}$$

□

**Définition 2.4.6** (Stabilité). Le schéma (2.25), resp. (2.10), est stable si, pour tout  $\mu \in \mathbb{R}^N$ , la solution  $z^{(h)}$  du schéma :

$$\frac{1}{h^2} A_h z^{(h)} = \mu \quad (2.29)$$

où  $A_h \in \mathbb{R}^{N \times N}$  est la matrice carrée définie par (2.26), resp. par (2.13), vérifie :

$$\|z^{(h)}\|_\infty \leq C \|\mu\|_\infty,$$

pour une constante  $C > 0$  indépendante de  $h > 0$ .

**Proposition 2.4.8.** *Le schéma (2.25), resp. (2.10), est stable.*

*Démonstration.* La stabilité du schéma (2.10) résulte de (2.18) et de l'équivalence des normes sur  $\mathbb{R}^N$ . On note  $A_{0h}$  la matrice définie par (2.13). Alors  $A_h = A_{0h} + h^2 C_h$  où  $C_h$  est la matrice diagonale définie par :

$$(C_h)_{ii} = c_i, \quad i = 1, \dots, N.$$

On a :

$$A_{0h}^{-1} - A_h^{-1} = A_{0h}^{-1} A_h A_h^{-1} - A_{0h}^{-1} A_{0h} A_h^{-1} = A_{0h}^{-1} \underbrace{(A_h - A_{0h})}_{=h^2 C_h \geq 0} A_h^{-1}$$

Soit  $v \geq 0$ . Alors :

$$A_h - A_{0h} \geq 0 \Rightarrow (A_h - A_{0h})v \geq 0$$

D'après le Lemme 2.4.6, les matrices  $A_h$  et  $A_{0h}$  sont monotones, donc

$$(A_h - A_{0h})v \geq 0 \Rightarrow A_h^{-1}(A_h - A_{0h})v = v - A_h^{-1}A_{0h}v \geq 0.$$

Ceci est également vrai pour  $A_{0h}^{-1}v \geq 0$ , donc  $A_{0h}^{-1}v \geq A_h^{-1}v$ . Finalement :

$$A_{0h}^{-1} \geq A_h^{-1} \geq 0.$$

On remarque que si  $B$  est une matrice positive, soit  $B \geq 0$ , alors

$$\|B\|_\infty = \max_i \sum_j |B_{ij}| = \max_i \sum_j B_{ij}.$$

On en déduit :

$$\|A_h^{-1}\|_\infty = \max_i \sum_j (A_h^{-1})_{ij} \leq \max_i \sum_j (A_{0h}^{-1})_{ij} = \|A_{0h}^{-1}\|_\infty.$$

Il reste donc à estimer  $\|A_{0h}^{-1}\|_\infty$ . On a aussi :

$$\|A_{0h}^{-1}\|_\infty = \|A_{0h}^{-1}e\|_\infty \quad \text{où } e = (1, \dots, 1)^T \in \mathbb{R}^N$$

Soit  $d_h = h^2 A_{0h}^{-1}e \in \mathbb{R}^N$ . De façon équivalente,  $d_h$  est solution du système  $h^{-2}A_{0h}d_h = e$  résultant de la discrétisation de

$$-u'' = 1 \quad \text{dans } ]0, 1[, \quad u(0) = u(1) = 0$$

dont la solution exacte est

$$u_0(x) = \frac{1}{2}x(1-x), \quad \forall x \in [0, 1].$$

Soit  $x^{(h)} \in \mathbb{R}^N$  le vecteur de composantes  $x_1, \dots, x_N$ , et soit  $u_0(x^{(h)})$  le vecteur de composantes  $u_0(x_1), \dots, u_0(x_N)$ . Comme  $u_0$  est polynomiale de degré 2, on a :

$$\frac{1}{h^2}A_{0h}u_0(x^{(h)}) = e.$$

Il en résulte :

$$h^2\|A_{0h}^{-1}e\|_\infty = \|u_0(x^{(h)})\|_\infty \leq \sup_{[0,1]} |u_0| = \frac{1}{8}$$

i.e. :

$$\|A_h^{-1}\|_\infty \leq \|A_{0h}^{-1}\|_\infty = \|A_{0h}^{-1}e\|_\infty \leq \frac{1}{8h^2}.$$

Soit  $\mu \in \mathbb{R}^N$  et soit  $z^{(h)} \in \mathbb{R}^N$  solution de (2.29). On a :

$$\|z^{(h)}\|_\infty = h^2\|A_n^{-1}\mu\|_\infty \leq h^2\|A_n^{-1}\|_\infty h^2\|\mu\|_\infty \leq \frac{1}{8}\|\mu\|_\infty, \quad (2.30)$$

i.e. le schéma (2.25) est stable.  $\square$

**Définition 2.4.7** (Erreur de discrétisation). On appelle erreur de discrétisation au point  $x_i$  la quantité :

$$e_i^{(h)} = u(x_i) - u_i, \quad i = 1, \dots, N.$$

**Théorème 2.4.9.** Soit  $u$  la solution du problème (2.24). On suppose que  $u \in \mathcal{C}^4([0, 1])$ . Alors, l'erreur de discrétisation  $e^{(h)}$  du schéma (2.25) vérifie :

$$\|e^{(h)}\|_\infty \leq \frac{h^2}{96}\|u^{(4)}\|_\infty$$

Le schéma (2.25) est donc convergent d'ordre 2.

*Démonstration.* Le choix  $z^{(h)} = e^{(h)}$  dans (2.29) conduit à  $\mu = \varepsilon^{(N)}$  défini par (2.27). De (2.30) et (2.28) on déduit :

$$\|e^{(h)}\|_\infty \stackrel{(2.30)}{\leq} \frac{1}{8} \|\varepsilon^{(N)}\|_\infty \stackrel{(2.28)}{\leq} \frac{h^2}{96} \sup_{x \in [0,1]} |u^{(4)}(x)|$$

□

**Proposition 2.4.10** (Le principe du maximum discret). *La solution du problème (2.24) avec  $f = 0$  et la condition de Dirichlet non homogène  $u_0 = a$ ,  $u_{N+1} = b$ ,  $a, b \in \mathbb{R}$  donnés, vérifie :*

$$\min(a, b) \leq u_i \leq \max(a, b), \quad i = 1, \dots, N.$$

*Démonstration.* Soit  $U^{(h)} \in \mathbb{R}^N$  le vecteur de composantes  $U_i^{(h)} = u_i$ ,  $i = 1, \dots, N$ . Par définition, et compte tenu de (2.19) :

$$\frac{1}{h^2} A_h U^{(h)} = \begin{pmatrix} a \\ 0 \\ \vdots \\ 0 \\ b \end{pmatrix}.$$

Soit  $\lambda \in \mathbb{R}$ . Le calcul direct donne :

$$\frac{1}{h^2} A_h (U^{(h)} + \lambda e) = \begin{pmatrix} a + \lambda \\ 0 \\ \vdots \\ 0 \\ b + \lambda \end{pmatrix}, \quad e := \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Si  $\lambda = -\max(a, b)$  alors  $A_h(U^{(h)} + \lambda e) \leq 0 \Rightarrow U^{(h)} + \lambda e \leq 0$ , i.e.  $u_i \leq \max(a, b)$ ,  $i = 1, \dots, N$ . Si  $\lambda = -\min(a, b)$  alors  $A_h(U^{(h)} + \lambda e) \geq 0 \Rightarrow U^{(h)} + \lambda e \geq 0$ , i.e.  $u_i \geq \min(a, b)$ ,  $i = 1, \dots, N$ . □

## 2.5 Diffusion bidimensionnelle

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

$$\begin{cases} -\Delta u = f & \text{dans } \Omega, \\ u = 0 & \text{sur } \partial\Omega. \end{cases} \quad (2.31)$$

Le problème est bien posé au sens où si  $f \in \mathcal{C}^1(\overline{\Omega})$ , alors il existe une unique solution  $u \in \mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^2(\Omega)$  de (2.31). Si  $f \in L^2(\Omega)$  et si  $\Omega$  est convexe (ou à bord régulier), alors il existe une unique solution faible  $u \in H^2(\Omega)$  de (2.31); i.e. vérifiant :

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

On peut montrer que si  $u \in \mathcal{C}^2(\overline{\Omega})$ , alors  $u$  est solution de (2.31) ssi  $u$  est solution faible de (2.31). Pour discrétiser le problème on se donne un nombre fini de points alignés dans les directions des axes  $0x$  et  $0y$  comme représentés dans la Figure 2.1 (on choisit un maillage régulier de pas  $\Delta x$  et  $\Delta y$  dans les directions de  $0x$  et  $0y$  resp.) Certains points sont à l'intérieur de  $\Omega$ , d'autres sur le bord. Comme dans le cas de la dimension 1, les inconnues discrètes  $u_{ij}$  sont associées aux points  $P(x_i, y_j)$  du maillage de sorte que  $u_{ij} \sim u(P(x_i, y_j))$ . On note  $\{P_i, i \in I\}$  l'ensemble des points de discrétisation. Dans le cas de points vraiment intérieurs, tels que le point  $P_1$  sur la Figure 2.1, i.e. pour lesquels les points voisins dans le schéma sont aussi intérieurs, on a :

$$\begin{aligned} -\Delta u(P_1) &= \frac{-u(P_2) + 2u(P_1) - u(P_3)}{(\Delta x)^2} + \frac{-u(P_4) + 2u(P_1) - u(P_5)}{(\Delta y)^2} + \\ &\quad + O((\Delta x)^2 + (\Delta y)^2) \end{aligned}$$

Pour des points proches du bord, i.e. pour lesquels l'un des points voisins est hors de  $\Omega$ , on doit prendre en compte les conditions sur le bord, ce qui dégrade l'approximation qui devient de l'ordre de  $O((\Delta x) + (\Delta y))$ . Le schéma numérique peut se réécrire sous forme matricielle  $A_{\Delta x, \Delta y} U_{\Delta x, \Delta y} = F_{\Delta x, \Delta y}$  où la matrice  $A_{\Delta x, \Delta y}$  est tridiagonale par bandes et dont la largeur de bande dépend du système de numérotation des noeuds choisi. On peut montrer que la matrice  $A_{\Delta x, \Delta y}$  est inversible et monotone, et que le schéma est stable. De la stabilité et de la consistance, on déduit comme en dimension 1 la convergence du schéma.

Pour simplifier la suite de l'exposé, on suppose que  $\Omega = ]0, a[ \times ]0, b[$  est un rectangle. Soit  $M > 0$  et soit  $N > 0$ . On définit les pas du maillage :

$$h_x := \frac{a}{N+1}, \quad h_y := \frac{b}{M+1}$$

et on considère les subdivisions

$$x_i = ih_x, \quad y_j = jh_y, \quad i = 0, \dots, N+1, \quad j = 0, \dots, M+1.$$

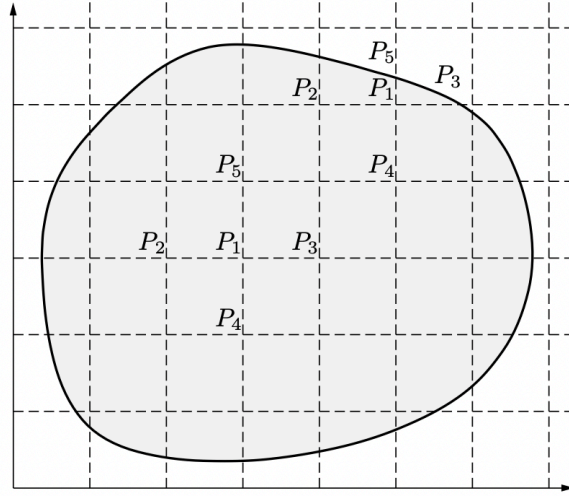


FIGURE 2.1 – Différences finies : discrétisation bidimensionnelle

de  $[0, a]$  et  $[0, b]$  resp.

L'analogie des problèmes discrétisés (2.10) et (2.25) s'écrit :

$$\left\{ \begin{array}{l} \frac{1}{h_x^2}(-u_{i-1,j} + 2u_{i,j} - u_{i+1,j}) + \frac{1}{h_y^2}(-u_{i,j-1} + 2u_{i,j} - u_{i,j+1}) = f_{ij}, \\ i = 1, \dots, N, \quad j = 1, \dots, M \\ u_{0,j} = u_{N+1,j} = 0, \quad j = 0, \dots, M + 1, \\ u_{i,0} = u_{i,N+1} = 0, \quad i = 0, \dots, N + 1 \end{array} \right. \quad (2.32)$$

où  $f_{i,j} = f(x_i, y_j)$ ,  $i = 1, \dots, N$ ,  $j = 1, \dots, M$ , sont donnés. Sous forme matricielle, le problème (2.25) se réécrit :

$$A_h U_h = f_h \quad (2.33)$$

où  $U_h \in \mathbb{R}^{NM}$  est le vecteur

$$U_h = \begin{pmatrix} U_1^{(h)} \\ \vdots \\ U_N^{(h)} \end{pmatrix}, \quad \text{avec} \quad U_j^{(h)} = \begin{pmatrix} u_{1j} \\ \vdots \\ u_{Nj} \end{pmatrix}, \quad j = 1, \dots, M,$$

où  $A_h = (A_{ij}^{(h)})_{1 \leq i,j \leq M} \in \mathbb{R}^{NM \times NM}$  est la matrice carrée d'ordre  $NM$  formée de  $M^2$  blocs de taille  $N \times N$  :

$$A_{ij}^{(h)} = \begin{cases} B_h & \text{si } i = j, \\ C_h & \text{si } |i - j| = 1, \\ 0_h & \text{si } |i - j| > 1, \end{cases}$$

où  $B_h = (b_{ij}^{(h)})_{1 \leq i, j \leq N} \in \mathbb{R}^{N \times N}$ , est la matrice carrée d'ordre  $N$  définie par ses coefficients :

$$b_{ij}^{(h)} = \begin{cases} \frac{2}{h_x^2} + \frac{2}{h_y^2} & \text{si } i = j, \\ -\frac{1}{h_x^2} & \text{si } |i - j| = 1, \\ 0 & \text{si } |i - j| > 1, \end{cases} \quad (2.34)$$

$C_h = (c_{ij}^{(h)})_{1 \leq i, j \leq N} \in \mathbb{R}^{N \times N}$ , est la matrice carrée d'ordre  $N$  définie par ses coefficients :

$$c_{ij}^{(h)} = \begin{cases} -\frac{1}{h_y^2} & \text{si } i = j, \\ 0 & \text{si } i \neq j \end{cases} \quad (2.35)$$

$0_h \in \mathbb{R}^{N \times N}$ , est la matrice carrée nulle d'ordre  $N$ .

**Proposition 2.5.1.** *Si la solution  $u$  de (2.31) est de classe  $C^4$ , alors : l'erreur de consistance  $\varepsilon^{(N,M)}$  vérifie :*

$$\|\varepsilon^{(N,M)}\|_\infty \leq C(h_x^2 + h_y^2)\|u^{(4)}\|_\infty. \quad (2.36)$$

où la constante  $C > 0$  est indépendante de  $u$  et des pas  $h_x, h_y$ .

**Proposition 2.5.2.** *Le schéma (2.32) est stable.*

**Théorème 2.5.3.** *Soit  $u$  la solution du problème (2.31). On suppose que  $u \in C^4(\bar{\Omega})$ . Alors, l'erreur de discrétisation  $e^{(h)}$  du schéma (2.32) vérifie :*

$$\|e^{(h)}\|_\infty \leq C(h_x^2 + h_y^2)\|u^{(4)}\|_\infty$$

*Le schéma (2.32) est donc convergent d'ordre 2.*

## La méthode de Gauss-Seidel

On se propose de résoudre le système (2.33) par la méthode de Gauss-Seidel (version Jacobi), ce qui donne le schéma point par point :



$$\left\{ \begin{array}{l} \frac{1}{h_x^2}(-u_{i-1,j}^n + 2u_{i,j}^{n+1} - u_{i+1,j}^n) + \frac{1}{h_y^2}(-u_{i,j-1}^n + 2u_{i,j}^{n+1} - u_{i,j+1}^n) = f_{ij}, \\ u_{ij}^0 = 0, \\ u_{0,j}^n = u_{N+1,j}^n = 0, \\ u_{i,0}^n = u_{i,N+1}^n = 0, \\ i = 1, \dots, N, \quad j = 1, \dots, M \quad n \geq 0. \end{array} \right. \quad (2.37)$$

**Proposition 2.5.4.** *Soit  $T > 0$  et soit  $h = \max(h_x, h_y)$ . Si la solution  $u$  de (2.31) est de classe  $C^4$ , alors, le schéma (2.37) est convergent d'ordre 2 :*

$$\max_{nh \leq T} \|e^n\|_\infty \leq C_T h^2 \|u^{(4)}\|_\infty.$$

où  $C_T > 0$  ne dépend que de  $T > 0$ .

*Démonstration.* On commence par remarquer que l'erreur de consistance est donnée par :

$$\varepsilon_{ij}^n = \varepsilon_{i,j}^{(N,M)}, \quad \forall (i, j) \in [[1, N]] \times [[1, M]].$$

et vérifie donc

$$\|\varepsilon^n\|_\infty \leq C(h_x^2 + h_y^2) \|u^{(4)}\|_\infty. \quad (2.38)$$

où la constante  $C > 0$  est indépendante de  $u$  et des pas  $h_x, h_y$ . Soit  $(\mu_{ij}^n)_{n \geq 0}$  une suite de réels et soit  $(z_{ij}^n)$  la suite définie par le schéma :

$$\begin{aligned} z_{ij}^n &= \frac{h_y^2}{2(h_x^2 + h_y^2)} (z_{i-1,j}^n + z_{i+1,j}^n) + \frac{h_x^2}{2(h_x^2 + h_y^2)} (z_{i,j-1}^n + z_{i,j+1}^n) + \\ &\quad + \frac{h_x^2 h_y^2}{2(h_x^2 + h_y^2)} \mu_{ij}^n, \quad i = 1, \dots, N, \quad j = 1, \dots, M \end{aligned}$$

$$z_{0j}^n = z_{N+1,j}^n = z_{i,0}^n = z_{i,M+1}^n = 0, \quad i = 1, \dots, N, \quad j = 1, \dots, M.$$

Soit  $n \geq 0$ . On a :

$$|z_{ij}^{n+1}| \leq \frac{1}{2(h_x^2 + h_y^2)} (2h_x^2 \|z^n\|_\infty + 2h_y^2 \|z^n\|_\infty) + \frac{h_x^2 h_y^2}{2(h_x^2 + h_y^2)} \|\mu^n\|_\infty$$

$$\leq \|z^n\|_\infty + \frac{h_x^2 h_y^2}{2(h_x^2 + h_y^2)} \|\mu^n\|_\infty.$$

On pose  $\mu_{ij}^n = \varepsilon_{ij}$ . alors  $z_{ij}^n = u(x_i, y_j) - u_{ij}^n = e_{ij}^n$  coïncide avec l'erreur de convergence et on a, tant que  $nh \leq T$  :

$$\begin{aligned} \|e^n\|_\infty &\leq \|e^0\|_\infty + Cnh^4 \|u^{(4)}\|_\infty \leq Ch^2 \|u^{(4)}\|_\infty + CTh^3 \|u^{(4)}\|_\infty \\ &\leq C(1 + hT)h^2 \|u^{(4)}\|_\infty \end{aligned}$$

□

# Bibliographie

- [1] Thierry Gallouët, Raphaèle Herbin. Analyse numérique des équations aux dérivées partielles. Master. Marseille, France. 2011. (<https://cel.hal.science/cel-00637008v2>) *Chapitres 1.2 et 1.4.*
- [2] Xavier Gourdon. Les Maths en Tête. Analyse. Ellipses Marketing, Paris, 2020, *p. 338–339*
- [3] Lionel Sainsaulieu. Calcul Scientifique. Cours et exercices corrigés pour le second cycle et les écoles d'ingénieurs, Masson, Paris 1996. *Chapitres 1.2, 2.1, 2.2, 3.5.*
- [4] Brigitte Lucquin, Olivier Pironneau. Introduction au calcul scientifique. Masson, Paris, 1996. *Chapitre VII.1*
- [5] Alfio Quarteroni, Riccardo Sacco, Fausto Saleri Numerical Mathematics. Springer, Berlin, 2007. *Chapitre 12.1*
- [6] Alfio Quarteroni, Riccardo Sacco, Paola Gervasio Calcul scientifique. Springer, Milan, 2010. *Chapitres 8.2.1 à 8.2.5*

