

EXERCICES CORRIGÉS

-

ANALYSE NUMÉRIQUE ET OPTIMISATION Une introduction à la modélisation mathématique et à la simulation numérique

G. Allaire, S. Gaubert, O. Pantz

Ecole Polytechnique

MAP 431

4 octobre 2006

Introduction

Ce recueil rassemble tous les exercices proposés dans le cours de deuxième année d'introduction à l'analyse numérique et l'optimisation de Grégoire Allaire [1]. Toute référence à ce dernier se distinguera des références internes au recueil par ses caractères gras. Par exemple, **(1.1)** fait référence à la première formule du cours. Malgré notre vigilance, ce manuscrit comporte sans aucun doute (encore) de multiples erreurs de tout ordre. De nombreux exercices mériteraient un traitement plus élégant autant d'un point de vue mathématique que stylistique. Nous invitons d'ailleurs tout lecteur à participer à son amélioration. Vous pouvez nous signaler toute erreur ou approximation en envoyant un mail à l'adresse

`olivier.pantz@polytechnique.org`

Nous serons également heureux de recevoir de nouvelles solutions aux exercices proposés ou toutes autres suggestions. Bon courage.

G. Allaire, S. Gaubert, O. Pantz
Paris, Juillet 2006

Chapitre 1

INTRODUCTION A LA MODÉLISATION MATHÉMATIQUE ET A LA SIMULATION NUMÉRIQUE

Exercice 1.2.1 On suppose que la donnée initiale θ_0 est continue et uniformément bornée sur \mathbb{R} . Vérifier que

$$\theta(t, x) = \frac{1}{\sqrt{4\pi\nu t}} \int_{-\infty}^{+\infty} \theta_0(y) \exp\left(-\frac{(x - Vt - y)^2}{4\nu t}\right) dy \quad (1.1)$$

est bien une solution de

$$\begin{cases} \frac{\partial \theta}{\partial t} + V \frac{\partial \theta}{\partial x} - \nu \frac{\partial^2 \theta}{\partial x^2} = 0 & \text{pour } (x, t) \in \mathbb{R} \times \mathbb{R}_*^+ \\ \theta(t = 0, x) = \theta_0(x) & \text{pour } x \in \mathbb{R} \end{cases} \quad (1.2)$$

Correction. Afin de montrer que $\theta(t, x)$ est une fonction régulière et déterminer ses dérivées partielles, on souhaite appliquer le théorème de dérivation sous le signe somme. A cet effet, on pose $G(x, t, y) = \exp\left(-\frac{(x - Vt - y)^2}{4\nu t}\right)$. On a

$$\begin{aligned} \frac{\partial G}{\partial x} &= -\frac{x - Vt - y}{2\nu t} G(x, t, y) \\ \frac{\partial^2 G}{\partial x^2} &= \left(-\frac{1}{2\nu t} + \frac{(x - Vt - y)^2}{4\nu^2 t^2}\right) G(x, t, y) \\ \frac{\partial G}{\partial t} &= \frac{(x + Vt - y)(x - Vt - y)}{4\nu t^2} G(x, t, y). \end{aligned}$$

Pour tout x de \mathbb{R} et tout $t > 0$, il existe des constantes $C(x, t)$ et $\beta(x, t)$ positives telles que si z est suffisamment proche de x ,

$$\left| \frac{\partial G}{\partial x}(z, t, y) \right| \leq C(x, t)(1 + |y|) \exp(-\beta(x, t)y).$$

Comme $\theta_0(y)$ est uniformément bornée, on en déduit que

$$\left| \theta_0(y) \frac{\partial G}{\partial x}(z, t, y) \right| \leq C(x, t)(1 + |y|) \exp(-\beta(x, t)y) \sup_s |\theta_0(s)|$$

pour tout z appartenant à un voisinage de x . Le terme de droite est intégrable par rapport à y . Ainsi, d'après le théorème de dérivation sous le signe somme, on en déduit que

$$\begin{aligned} \frac{\partial}{\partial x} \int_{-\infty}^{\infty} \theta_0(y) G(x, t, y) dy &= \int_{-\infty}^{\infty} \theta_0(y) \frac{\partial G}{\partial x} dy \\ &= - \int_{-\infty}^{\infty} \theta_0(y) \frac{x - Vt - y}{2\nu t} G(x, t, y) dy. \end{aligned}$$

Par un raisonnement analogue, on obtient que

$$\begin{aligned} \frac{\partial^2}{\partial x^2} \int_{-\infty}^{\infty} \theta_0(y) G(x, t, y) dy &= \\ &- \int_{-\infty}^{\infty} \theta_0(y) \left(\frac{1}{2\nu t} - \frac{(x - Vt - y)^2}{4\nu^2 t^2} \right) G(x, t, y) dy \end{aligned}$$

et

$$\frac{\partial}{\partial t} \int_{-\infty}^{\infty} \theta_0(y) G(x, t, y) dy = \int_{-\infty}^{\infty} \theta_0(y) \frac{(x + Vt - y)(x - Vt - y)}{4\nu t^2} G(x, t, y) dy.$$

Ainsi, $\theta(t, x)$ est dérivable pour tout $t > 0$ et

$$\begin{aligned} \frac{\partial \theta}{\partial x} &= - \frac{1}{\sqrt{4\pi\nu t}} \int_{-\infty}^{\infty} \theta_0(y) \frac{x - Vt - y}{2\nu t} G(x, t, y) dy \\ \frac{\partial^2 \theta}{\partial x^2} &= - \frac{1}{\sqrt{4\pi\nu t}} \int_{-\infty}^{\infty} \theta_0(y) \left(\frac{1}{2\nu t} - \frac{(x - Vt - y)^2}{4\nu^2 t^2} \right) G(x, t, y) dy \\ \frac{\partial \theta}{\partial t} &= \frac{1}{\sqrt{4\pi\nu t}} \int_{-\infty}^{\infty} \theta_0(y) \left(\frac{(x + Vt - y)(x - Vt - y)}{4\nu t^2} - \frac{1}{2t} \right) G(x, t, y) dy. \end{aligned}$$

On vérifie alors aisément que

$$\frac{\partial \theta}{\partial t} + V \frac{\partial \theta}{\partial x} - \nu \frac{\partial^2 \theta}{\partial x^2} = 0$$

Il ne reste plus qu'à prouver que $\theta(t, x)$ est prolongeable en $t = 0$ et vérifie bien la condition initiale, c'est à dire que

$$\lim_{t \rightarrow 0} \frac{1}{\sqrt{4\pi\nu t}} \int_{-\infty}^{\infty} \theta_0(y) \exp\left(-\frac{(x - Vt - y)^2}{4\nu t}\right) dy = \theta_0(x). \quad (1.3)$$

Rappelons que,

$$\int_{-\infty}^{\infty} \exp(-x^2) dx = \sqrt{\pi}. \quad (1.4)$$

Pour établir cette relation, il suffit de calculer $\left(\int_{-\infty}^{\infty} e^{-x^2} dx\right)^2 = \int_{\mathbb{R}^2} e^{-|x|^2} dx$ en coordonnées polaires. On pose

$$\rho(x, t, y) = \frac{1}{\sqrt{4\pi\nu t}} \exp\left(-\frac{(x - Vt - y)^2}{4\nu t}\right).$$

D'après (1.4), $\int \rho(x, t, y) dy = 1$ pour tout x et t . Enfin, pour tout $x \in \mathbb{R}$, on constate que pour tout y différent de x , $\lim_{t \rightarrow 0} \rho(x, t, y) = 0$. Ainsi, x étant fixé, $\rho(x, t, y)$ est une fonction de y se concentrant en x lorsque t tend vers zéro. Pour être plus précis, on montre que pour tout δ et ε réels strictement positifs, il existe $t(\delta, \varepsilon)$ tel que pour tout $t < t(\delta, \varepsilon)$,

$$\left| \int_{x-\delta}^{x+\delta} \rho(x, t, y) dy - 1 \right| \leq \varepsilon.$$

et

$$\left| \int_{-\infty}^{x-\delta} \rho(x, t, y) dy + \int_{x+\delta}^{\infty} \rho(x, t, y) dy \right| \leq \varepsilon.$$

L'équation (1.3) découle alors du fait que θ_0 est continue, uniformément bornée.

Exercice 1.2.2 On suppose que la donnée initiale θ_0 est dérivable et uniformément bornée sur \mathbb{R} . Vérifier que

$$\theta(t, x) = \theta_0(x - Vt) \quad (1.5)$$

est bien une solution de

$$\begin{cases} \frac{\partial \theta}{\partial t} + V \frac{\partial \theta}{\partial x} = 0 & \text{pour } (x, t) \in \mathbb{R} \times \mathbb{R}_*^+ \\ \theta(t = 0, x) = \theta_0(x) & \text{pour } x \in \mathbb{R}. \end{cases} \quad (1.6)$$

Montrer que (1.5) est la limite de (1.1) lorsque le paramètre ν tend vers zéro.

Correction.

$$\frac{\partial \theta}{\partial t}(x, t) = -V \frac{\partial \theta_0}{\partial x}(x - Vt) = -V \frac{\partial \theta}{\partial x}(x).$$

Ainsi, θ vérifie l'équation différentielle annoncée. De plus, θ vérifie trivialement la condition initiale.

Par un raisonnement analogue à celui qui nous avait permis d'établir la continuité de la solution dans l'exercice précédent, on montre que

$$\lim_{\nu \rightarrow 0} \frac{1}{\sqrt{4\pi\nu t}} \int_{-\infty}^{+\infty} \theta_0(y) \exp\left(-\frac{(x - Vt - y)^2}{4\nu t}\right) dy = \theta_0(x - Vt) = \theta(t).$$

Exercice 1.3.1 On se propose de retrouver une propriété de décroissance exponentielle en temps (voir la formule (1.1)) de la solution de l'équation de la chaleur

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = f & \text{dans } \Omega \times \mathbb{R}_*^+ \\ u = 0 & \text{sur } \partial\Omega \times \mathbb{R}_*^+ \\ u(t = 0) = u_0 & \text{dans } \Omega \end{cases} \quad (1.7)$$

dans un domaine Ω borné. En une dimension d'espace, on pose $\Omega = (0, 1)$ et on suppose que $f = 0$. Soit $u(t, x)$ une solution régulière de (1.7). En multipliant l'équation par u et en intégrant par rapport à x , établir l'égalité

$$\frac{1}{2} \frac{d}{dt} \left(\int_0^1 u^2(t, x) dx \right) = - \int_0^1 \left| \frac{\partial u}{\partial x}(t, x) \right|^2 dx$$

Montrer que toute fonction $v(x)$ continûment dérivable sur $[0, 1]$, telle que $v(0) = 0$, vérifie l'inégalité de Poincaré

$$\int_0^1 v^2(x) dx \leq \int_0^1 \left| \frac{dv}{dx}(x) \right|^2 dx.$$

En déduire la décroissance exponentielle en temps de $\int_0^1 u^2(t, x) dx$.

Correction. En multipliant l'équation différentielle (1.7) par u on obtient par intégration que

$$\int_0^1 \frac{\partial u}{\partial t} u dx = \int_0^1 \frac{\partial^2 u}{\partial x^2} u dx.$$

Quitte à supposer u suffisamment régulière, on peut appliquer le théorème d'intégration sous le signe somme au terme de gauche et effectuer une intégration par partie sur le terme de droite. On obtient ainsi que

$$\frac{1}{2} \frac{d}{dt} \left(\int_0^1 u^2 dx \right) = - \int_0^1 \left| \frac{\partial u}{\partial x} \right|^2 dx. \quad (1.8)$$

Soit v une fonction de classe C^1 sur $[0, 1]$ telle que $v(0) = 0$. Pour tout $x \in [0, 1]$,

$$v^2(x) = \left(\int_0^x \frac{dv}{dx}(y) dy \right)^2 \leq x \int_0^x \left| \frac{dv}{dx}(y) \right|^2 dy \leq \int_0^1 \left| \frac{dv}{dx}(y) \right|^2 dy$$

d'où

$$\int_0^1 v^2(x) dx \leq \int_0^1 \left| \frac{dv}{dx}(x) \right|^2 dx.$$

En appliquant cette dernière inégalité à $v(x) = u(t, x)$ et (1.8)

$$\frac{1}{2} \frac{df}{dt}(t) \leq -f(t)$$

où

$$f(t) = \int_0^1 u^2(x, t) dx.$$

Ainsi,

$$\frac{1}{2} \frac{d(fe^{2t})}{dt} = \left(\frac{1}{2} \frac{df}{dt} + f \right) e^{2t} \leq 0$$

et pour tout $t \geq 0$,

$$f(t)e^{2t} \leq f(0).$$

Exercice 1.3.2 On se place en dimension $N = 1$ d'espace. On suppose que les données initiales u_0 et u_1 sont des fonctions régulières, et que $f = 0$ avec $\Omega = \mathbb{R}$. On note U_1 une primitive de u_1 . Vérifier que

$$u(t, x) = \frac{1}{2} (u_0(x+t) + u_0(x-t)) + \frac{1}{2} (U_1(x+t) - U_1(x-t)), \quad (1.9)$$

est la solution unique de

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - \Delta u = f & \text{dans } \Omega \times \mathbb{R}_*^+ \\ u = 0 & \text{sur } \partial\Omega \times \mathbb{R}_*^+ \\ u(t=0) = u_0 & \text{dans } \Omega \\ \frac{\partial u}{\partial t}(t=0) = u_1 & \text{dans } \Omega \end{cases} \quad (1.10)$$

dans la classe des fonctions régulières.

Correction. La fonction

$$u(t, x) = \frac{1}{2} (u_0(x+t) + u_0(x-t)) + \frac{1}{2} (U_1(x+t) - U_1(x-t))$$

où U_1 est une primitive de u_1 est trivialement une solution de l'équation des ondes (1.10). Comme l'équation est linéaire, il suffit de prouver l'unicité pour $u_0 = u_1 = 0$. Soit $x_0 < x_1$ et $2t < x_1 - x_0$. En multipliant l'équation différentielle par $\frac{\partial u}{\partial t}$, on obtient par intégration par partie que

$$\begin{aligned} 0 = \int_{x_0+t}^{x_1-t} \frac{\partial}{\partial t} \left(\left| \frac{\partial u}{\partial t}(x, t) \right|^2 \right) dx + \int_{x_0+t}^{x_1-t} \frac{\partial}{\partial t} \left(\left| \frac{\partial u}{\partial x}(x, t) \right|^2 \right) dx \\ - 2 \frac{\partial u}{\partial x} \frac{\partial u}{\partial t}(x_1 - t) + 2 \frac{\partial u}{\partial x} \frac{\partial u}{\partial t}(x_0 + t). \end{aligned}$$

Par commutation de la dérivation et de l'intégration, on en déduit que

$$\begin{aligned} 0 = \frac{d}{dt} \left(\int_{x_0+t}^{x_1-t} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 + \left| \frac{\partial u}{\partial x}(x, t) \right|^2 dx \right) \\ + \left| \frac{\partial u}{\partial t}(x_0 + t, t) \right|^2 + \left| \frac{\partial u}{\partial t}(x_1 - t, t) \right|^2 + \left| \frac{\partial u}{\partial x}(x_0 + t, t) \right|^2 + \left| \frac{\partial u}{\partial x}(x_1 - t, t) \right|^2 \\ - 2 \frac{\partial u}{\partial x} \frac{\partial u}{\partial t}(x_1 - t, t) + 2 \frac{\partial u}{\partial x} \frac{\partial u}{\partial t}(x_0 + t, t) \end{aligned}$$

c'est à dire

$$\begin{aligned} - \frac{d}{dt} \left(\int_{x_0+t}^{x_1-t} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 + \left| \frac{\partial u}{\partial x}(x, t) \right|^2 dx \right) = \\ \left| \left(\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x} \right)(x_0 + t, t) \right|^2 + \left| \left(\frac{\partial u}{\partial t} - \frac{\partial u}{\partial x} \right)(x_1 - t, t) \right|^2. \end{aligned}$$

Ainsi,

$$\frac{d}{dt} \left(\int_{x_0+t}^{x_1-t} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 + \left| \frac{\partial u}{\partial x}(x, t) \right|^2 dx \right) \leq 0.$$

Pour tout $t \geq 0$, pour tout y_0 et y_1 tels que $y_0 \leq y_1$, on a donc

$$\int_{y_0}^{y_1} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 + \left| \frac{\partial u}{\partial x}(x, t) \right|^2 dx \leq \int_{x_0}^{x_1} \left| \frac{\partial u}{\partial t}(x, 0) \right|^2 + \left| \frac{\partial u}{\partial x}(x, 0) \right|^2 dx = 0 \quad (1.11)$$

où $x_0 = y_0 - t$ et $x_1 = y_1 + t$. On déduit de (1.11) que $u(x, t) = 0$ pour tout x et $t \geq 0$, ce qui achève la démonstration.

Exercice 1.3.3 Vérifier que la solution (1.9) au point (x, t) ne dépend des données initiales u_0 et u_1 qu'à travers leurs valeurs sur le segment $[x - t, x + t]$. Vérifier aussi $u(-t, x)$ est solution de (1.10) dans $\Omega \times \mathbb{R}_*^-$, quitte à changer le signe de la vitesse initiale $u_1(x)$.

Correction. On rappelle que

$$u(t, x) = \frac{1}{2}(u_0(x + t) + u_0(x - t)) + \frac{1}{2}(U_1(x + t) - U_1(x - t)),$$

où U_1 est une primitive de u_1 . Comme

$$U_1(x + t) - U_1(x - t) = \int_{x-t}^{x+t} u_1(y) dy$$

ne dépend que de la restriction de u_1 sur l'intervalle $[x - t, x + t]$, on en déduit que $u(t, x)$ ne dépend que de u_0 et u_1 restreints à $[x - t, x + t]$. L'information se propage à vitesse finie. Enfin, on vérifie sans mal que $u(-t, x)$ est solution de la même équation sur $\Omega \times \mathbb{R}_*^-$, quitte à remplacer u_1 par $-u_1$.

Exercice 1.3.4 On se propose de démontrer un principe de conservation de l'énergie pour l'équation des ondes (1.10) sans utiliser la formule explicite (1.9). En une dimension d'espace, on pose $\Omega = (0, 1)$ et on suppose $f = 0$. Soit $u(t, x)$ une solution régulière de (1.10). En multipliant l'équation par $\frac{\partial u}{\partial t}$ et en intégrant par rapport à x , établir l'égalité d'énergie

$$\frac{d}{dt} \left(\int_0^1 \left| \frac{\partial u}{\partial t}(t, x) \right|^2 dx + \int_0^1 \left| \frac{\partial u}{\partial x}(t, x) \right|^2 dx \right) = 0.$$

Conclure et comparer à ce qui se passe pour l'équation de la chaleur.

Correction. En multipliant l'équation des ondes par $\partial u / \partial t$, on obtient par intégration

$$\int_0^1 \frac{\partial^2 u}{\partial t^2} \frac{\partial u}{\partial t} dx - \int_0^1 \frac{\partial^2 u}{\partial x^2} \frac{\partial u}{\partial t} dx = 0.$$

On applique alors le théorème de dérivation sous le signe somme au premier terme de l'équation et on effectue une intégration par partie sur le second. Il vient

$$\frac{1}{2} \frac{d}{dt} \left(\int_0^1 \left| \frac{\partial u}{\partial t} \right|^2 dx \right) + \int_0^1 \frac{\partial u}{\partial x} \frac{\partial^2 u}{\partial t \partial x} dx = 0.$$

En appliquant à nouveau le théorème de dérivation sous le signe somme (au deuxième terme cette fois), on établit la conservation de l'énergie.

Dans le cas de l'équation de la chaleur avec condition de Dirichlet, l'énergie totale décroît exponentiellement. La température tend à devenir uniformément nulle au sein de l'ouvert Ω . Il y a une déperdition d'énergie par le bord de Ω . Le comportement est très différent pour la solution de l'équation des ondes. L'énergie est conservée au cours du temps et l'onde est réfléchiée sur les bords.

Exercice 1.3.5 On se propose de démontrer des principes de conservation de l'énergie pour l'équation de Schrödinger

$$\begin{cases} i \frac{\partial u}{\partial t} + \Delta u - Vu = 0 & \text{dans } \mathbb{R}^N \times \mathbb{R}_*^+ \\ u(t=0) = u_0 & \text{dans } \mathbb{R}^N. \end{cases} \quad (1.12)$$

Soit $u(t, x)$ une solution régulière de (1.12) en une dimension d'espace qui décroît vers zéro (ainsi que $\frac{\partial u}{\partial x}$) lorsque $|x| \rightarrow +\infty$. Montrer que pour toute fonction dérivable $v(t)$ on a

$$\mathcal{R} \left(\frac{\partial v}{\partial t} \bar{v} \right) = \frac{1}{2} \frac{\partial |v|^2}{\partial t},$$

où \mathcal{R} désigne la partie réelle et \bar{v} le complexe conjugué de v . En multipliant l'équation par \bar{u} et en intégrant par rapport à x , établir l'égalité d'énergie

$$\int_{\mathbb{R}} |u(t, x)|^2 dx = \int_{\mathbb{R}} |u_0(x)|^2 dx.$$

En multipliant l'équation par $\frac{\partial \bar{u}}{\partial t}$, montrer que

$$\int_{\mathbb{R}} \left(\left| \frac{\partial u}{\partial x}(t, x) \right|^2 + V(x) |u(t, x)|^2 \right) dx = \int_{\mathbb{R}} \left(\left| \frac{\partial u_0}{\partial x}(x) \right|^2 + V(x) |u_0(x)|^2 \right) dx.$$

Correction. Soit v une fonction dérivable,

$$\mathcal{R} \left(\frac{\partial v}{\partial t} \bar{v} \right) = \frac{1}{2} \left(\frac{\partial v}{\partial t} \bar{v} + \frac{\partial \bar{v}}{\partial t} v \right) = \frac{1}{2} \frac{\partial v \bar{v}}{\partial t}$$

On a bien

$$\mathcal{R} \left(\frac{\partial v}{\partial t} \bar{v} \right) = \frac{1}{2} \frac{\partial |v|^2}{\partial t}. \quad (1.13)$$

En multipliant l'équation de Schrödinger par \bar{u} , on obtient par intégration que

$$i \int_{\mathbb{R}} \frac{\partial u}{\partial t} \bar{u} + \frac{\partial^2 u}{\partial x^2} \bar{u} - V|u|^2 dx = 0$$

Par intégration par partie sur le second membre, on obtient

$$i \int_{\mathbb{R}} \frac{\partial u}{\partial t} \bar{u} dx = \int_{\mathbb{R}} \left| \frac{\partial u}{\partial x} \right|^2 + V|u|^2 dx$$

(les hypothèses de décroissance effectuées sur u permettent d'éliminer les termes de bords à "l' infini"). Comme le second membre est réel, $\int_{\mathbb{R}} \frac{\partial u}{\partial t} \bar{u} dx$ est un imaginaire pure,

$$\mathcal{R} \left(\int_{\mathbb{R}} \frac{\partial u}{\partial t} \bar{u} dx \right) = 0.$$

D' après (1.13), on a donc

$$\int_{\mathbb{R}} \frac{\partial |u|^2}{\partial t} dx = 0.$$

Pourvu que la solution u soit suffisamment régulière, on peut commuter le signe somme et intégrale, ainsi

$$\frac{d}{dt} \int_{\mathbb{R}} |u|^2 dx = 0$$

et

$$\int_{\mathbb{R}} |u(t, x)|^2 dx = \int_{\mathbb{R}} |u_0|^2 dx.$$

En multipliant l'équation de Schrödinger par $\frac{\partial \bar{u}}{\partial t}$, il vient

$$\int_{\mathbb{R}} i \left| \frac{\partial u}{\partial t} \right|^2 + \frac{\partial^2 u}{\partial x^2} \frac{\partial \bar{u}}{\partial t} - V u \frac{\partial \bar{u}}{\partial t} dx = 0$$

Par intégration par partie du second membre, on obtient que

$$\int_{\mathbb{R}} i \left| \frac{\partial u}{\partial t} \right|^2 - \frac{\partial u}{\partial x} \frac{\partial^2 \bar{u}}{\partial t \partial x} - V u \frac{\partial \bar{u}}{\partial t} dx = 0.$$

En considérant la partie réelle de cette égalité, il vient

$$\int_{\mathbb{R}} \frac{\partial}{\partial t} \left(\left| \frac{\partial u}{\partial x} \right|^2 + V |u|^2 \right) dx = 0.$$

Il suffit d'échanger la dérivation par rapport au temps et le signe intégrale afin d'obtenir le résultat escompté

$$\int_{\mathbb{R}} \left(\left| \frac{\partial u}{\partial x} \right|^2 + V |u|^2 \right) dx = \int_{\mathbb{R}} \left(\left| \frac{\partial u_0}{\partial x} \right|^2 + V |u_0|^2 \right) dx.$$

Exercice 1.4.1 Le but de cet exercice est de montrer que le schéma implicite

$$\frac{u_j^n - u_j^{n-1}}{\Delta t} + V \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} + \nu \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} = 0, \quad (1.14)$$

avec $V = 0$, vérifie aussi le principe du maximum discret. On impose des conditions aux limites de Dirichlet, c'est-à-dire que la formule (1.14) est valable pour $1 \leq j \leq J$ et on fixe $u_0^n = u_{J+1}^n = 0$ pour tout $n \in \mathbb{N}$. Soit deux constantes $m \leq 0 \leq M$ telles que $m \leq u_j^0 \leq M$ pour $1 \leq j \leq J$. Vérifier que l'on peut bien calculer de manière unique les u_j^{n+1} en fonction des u_j^n . Montrer que pour tous les temps $n \geq 0$ on a encore les inégalités $m \leq u_j^n \leq M$ pour $1 \leq j \leq J$ (et ceci sans condition sur Δt et Δx).

Correction. Tout d'abord, montrons que le schéma implicite (1.14) est correctement défini. On pose $U^n = (u_j^n)_{1 \leq j \leq J}$. On vérifie que le schéma implicite équivaut à déterminer U^n tel que

$$AU^n = U^{n-1}.$$

où

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

et $c = \nu \Delta t / (\Delta x)^2$. Il s'agit donc de prouver que la matrice A est inversible, ce qui est aisé. En effet, A est symétrique, définie positive donc inversible : Soit $X \in \mathbb{R}^J$. Par convention, on pose $X_0 = X_{J+1} = 0$. On a

$$X^T A X = \sum_{j=0}^J \frac{X_j^2 + X_{j+1}^2}{2} + c(X_{j+1} - X_j)^2.$$

Reste à prouver que le schéma vérifie le principe du maximum. On raisonne par récurrence sur n . Supposons que $m \leq u_j^{n-1} \leq M$ pour tout $j \in \{0, \dots, J+1\}$ rappelons que d'après les conditions aux bords, $m \leq 0 \leq M$. Soit $m' = \inf_{j \in \{1, \dots, J\}} u_j^n$ et $M' = \sup_{j \in \{1, \dots, J\}} u_j^n$.

Montrons que $M' \leq M$. Si $M' = 0$, on a rien à démontrer. Dans le cas contraire, soit $k \in \{1, \dots, J\}$ tel que $M' = u_k^n$. D'après le schéma,

$$(1+2c)u_k^n = u_k^{n-1} + 2c \left(\frac{u_{k-1}^n + u_{k+1}^n}{2} \right)$$

Comme $\frac{u_{k-1}^n + u_{k+1}^n}{2} \leq u_k^n$, on en déduit que

$$(1+2c)u_k^n \leq u_k^{n-1} + 2cu_k^n,$$

d'où

$$M' = u_k^n \leq u_k^{n-1} \leq M.$$

Quitte à remplacer u par $-u$, on obtient également $m' \geq m$.

Exercice 1.4.2 Montrer que, si la condition CFL

$$|V| \Delta t \leq \Delta x. \quad (1.15)$$

n'est pas satisfaite, le schéma décentré amont

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + V \frac{u_j^n - u_{j-1}^n}{\Delta x} = 0 \quad (1.16)$$

pour l'équation d'advection est instable pour la donnée initiale $u_j^0 = (-1)^j$.

Correction. Le schéma décentré amont est défini par

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + V \frac{u_j^n - u_{j-1}^n}{\Delta x} = 0.$$

Considérons comme donnée initiale $u_j^0 = (-1)^j$. On montre par une récurrence évidente que

$$u_j^n = \left(1 - \frac{2V\Delta t}{\Delta x}\right)^n (-1)^j.$$

Ainsi, la suite u^n reste bornée si et seulement si

$$\left|1 - \frac{2V\Delta t}{\Delta x}\right| \leq 1.$$

ou encore si la condition CFL

$$\frac{|V|\Delta t}{\Delta x} \leq 1$$

est vérifiée.

Exercice 1.4.3 Écrire un schéma explicite centré en espace pour l'équation des ondes (1.10) en une dimension d'espace et sans terme source. Préciser comment démarrer les itérations en temps. Vérifier l'existence d'un cône de dépendance discret analogue à celui continu illustré par la Figure 1.3. En déduire que, si ce schéma converge, les pas de temps et d'espace doivent nécessairement satisfaire la condition (de type CFL) $\Delta t \leq \Delta x$.

Correction. Pour l'équation des ondes (1.10) sans terme source, le schéma explicite centré est

$$\frac{u_j^{n+1} - 2u_j^n + u_j^{n-1}}{(\Delta t)^2} + \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} = 0.$$

Ainsi,

$$u_j^{n+1} = -u_j^{n-1} + 2u_j^n + \left(\frac{\Delta t}{\Delta x}\right)^2 (u_{j-1}^n - 2u_j^n + u_{j+1}^n). \quad (1.17)$$

On initialise le schéma en posant

$$u_j^0 = u_0(j\Delta x) \text{ et } u_j^1 = u_j^0 + \Delta t u_1(j\Delta x).$$

Au vu de l'équation (1.17), on montre par une récurrence évidente que la valeur de u_j^{n+1} ne dépend que des valeurs des u_{j+k}^1 pour k entier, $-n \leq k \leq n$ et de u_{j+l}^0 pour l entier, $-n < l < n$.

On note $u(t, x)$ la solution de l'équation des ondes. Comme la valeur $u((n+1)\Delta t, j\Delta x)$ dépend des valeurs de u_0 et u_1 sur $[j\Delta x - (n+1)\Delta t, j\Delta x + (n+1)\Delta t]$, pour que le schéma converge, on doit avoir

$$(j - n - 1)\Delta x \leq j\Delta x - (n + 1)\Delta t \text{ et } j\Delta x + (n + 1)\Delta t \leq (j + n + 1)\Delta x,$$

conditions qui sont équivalentes à

$$\Delta t \leq \Delta x.$$

Exercice 1.5.1 Le but de cet exercice est de montrer que le problème de Cauchy pour le Laplacien est mal posé. Soit le domaine bidimensionnel $\Omega = (0, 1) \times (0, 2\pi)$. On considère le problème de Cauchy en x et le problème aux limites en y suivant

$$\begin{cases} -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 0 & \text{dans } \Omega \\ u(x, 0) = u(x, 2\pi) = 0 & \text{pour } 0 < x < 1 \\ u(0, y) = 0, \frac{\partial u}{\partial x}(0, y) = -e^{-\sqrt{n}} \sin(ny) & \text{pour } 0 < y < 2\pi \end{cases}$$

Vérifier que $u(x, y) = \frac{e^{-\sqrt{n}}}{n} \sin(ny) \operatorname{sh}(nx)$ est une solution. Montrer que la condition initiale et toutes ses dérivées en $x = 0$ convergent uniformément vers 0, tandis que, pour tout $x > 0$, la solution trouvée $u(x, y)$ et toutes ses dérivées ne sont pas bornées quand n tend vers l'infini. Conclure.

Correction. Ici, x joue le rôle du temps. On vérifie sans mal que la solution proposée est une solution du système. D'autre part,

$$\frac{\partial^k u}{\partial x^k} = \begin{cases} e^{-\sqrt{n}} n^k \sin(ny) \sinh(nx) & \text{si } k \text{ est pair} \\ e^{-\sqrt{n}} n^{k-1} \sin(ny) \cosh(nx) & \text{si } k \text{ impair} \end{cases}$$

et

$$\begin{cases} \frac{\partial^{2p} u}{\partial y^{2p}} = e^{-\sqrt{n}} (-1)^p n^{2p-1} \sin(ny) \sinh(nx) \\ \frac{\partial^{2p+1} u}{\partial y^{2p+1}} = e^{-\sqrt{n}} (-1)^p n^{2p} \cos(ny) \sinh(nx). \end{cases}$$

On constate que en $x = 0$, u ainsi que toutes ses dérivées convergent vers 0 lorsque n tend vers $+\infty$. A contrario, si $x > 0$, ni u ni ses dérivées ne sont bornées par rapport à n . Or, pour des conditions initiales (i.e. en $x = 0$) nulles, la fonction $u = 0$ est une solution triviale du système. Ainsi, des perturbations infinitésimales des conditions initiales (même pour la norme très forte C^∞) induisent de très grandes perturbations de la solution (pour n'importe quelle norme raisonnable, même faible). Le problème de Cauchy proposé est donc mal posé.

Chapitre 2

MÉTHODE DES DIFFÉRENCES FINIES

Exercice 2.2.1 Montrer que le schéma à six points

$$\begin{aligned} & \frac{u_{j+1}^{n+1} - u_{j+1}^n}{12\Delta t} + \frac{5(u_j^{n+1} - u_j^n)}{6\Delta t} + \frac{u_{j-1}^{n+1} - u_{j-1}^n}{12\Delta t} \\ & + \nu \frac{-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}}{2(\Delta x)^2} + \nu \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{2(\Delta x)^2} = 0 \end{aligned} \quad (2.1)$$

n'est rien d'autre que le θ -schéma

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \theta \nu \frac{-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}}{(\Delta x)^2} + (1 - \theta) \nu \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} = 0 \quad (2.2)$$

avec $\theta = 1/2 - (\Delta x)^2/12\nu\Delta t$

Correction. Il suffit de constater que

$$\begin{aligned} & \frac{u_{j+1}^{n+1} - u_{j+1}^n}{12\Delta t} + \frac{5(u_j^{n+1} - u_j^n)}{6\Delta t} + \frac{u_{j-1}^{n+1} - u_{j-1}^n}{12\Delta t} \\ & = \frac{(u_j^{n+1} - u_j^n)}{\Delta t} + \frac{u_{j+1}^{n+1} - u_{j+1}^n}{12\Delta t} - \frac{2(u_j^{n+1} - u_j^n)}{12\Delta t} + \frac{u_{j-1}^{n+1} - u_{j-1}^n}{12\Delta t} \\ & = \frac{(u_j^{n+1} - u_j^n)}{\Delta t} - \frac{-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}}{12\Delta t} + \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{12\Delta t} . \end{aligned}$$

En remplaçant cette expression dans le schéma à six points, on en déduit que ce dernier est équivalent à

$$\begin{aligned} & \frac{u_j^{n+1} - u_j^n}{\Delta t} + \left(\frac{\nu}{2} - \frac{(\Delta x)^2}{12\Delta t} \right) \frac{-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}}{(\Delta x)^2} \\ & + \left(\frac{\nu}{2} + \frac{(\Delta x)^2}{12\Delta t} \right) \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} = 0 \end{aligned}$$

qui n'est rien d'autre que le θ schéma avec $\theta = 1/2 - (\Delta x)^2/12\nu\Delta t$.

Exercice 2.2.2 Pour chacun des schémas de la Sous-section 2.2.1, vérifier que l'erreur de troncature est bien du type annoncé dans le Tableau 2.1. (On remarquera que tous ces schémas sont consistants sauf celui de DuFort-Frankel.)

Correction. Le calcul de l'erreur de troncature d'un schéma est souvent délicat à mener. Si on ne procède pas de manière soignée et méthodique, on peut aisément se retrouver englué dans un calcul inextricable, dont le coût croît exponentiellement en fonction de l'ordre à déterminer. Quelques règles simples permettent en général d'éviter ce travers. L'erreur de troncature se calcule en développant tous les termes du schéma au même point à l'aide des formules de Taylor. Le point choisi n'a évidemment aucune influence sur le résultat obtenu (l'ordre du schéma ne dépend pas du point considéré). Par contre, ce choix influe sur la taille du calcul qui en résulte. Il est recommandé de diviser le calcul en plusieurs étapes. Les développements calculés lors d'une étape pouvant être réutilisés à une autre. Il faut absolument utiliser l'équation vérifiée par la solution (par exemple remplacer les dérivées en temps par des dérivées en espace). Cela simplifie considérablement les calculs, et nous permet de déterminer l'ordre optimal du schéma. Enfin, il faut éviter à tout prix d'effectuer des calculs inutiles et ne pas manipuler des termes d'ordre non significatifs. Enfin, un petit truc classique consiste à utiliser les symétries du schéma, qui peuvent impliquer que les termes non nul du développement sont nécessairement soit paires, soit impaires.

Les schémas explicite, implicite et de Crank-Nicholson ne sont que des cas particuliers du θ -schéma. Ce dernier possède des termes communs avec le schéma à 6 points dont nous donnons le développement ci-dessous. Le schéma d'ordre le plus élevé étant le schéma à 6 points, d'ordre 2 en temps et 4 en espace, on peut donc négliger les termes en $o((\Delta x)^4)$ et $o((\Delta t)^2)$. On effectue nos développements au point (t_n, x_j) (un autre choix raisonnable consisterait à effectuer les développements au point $(t_n + \Delta t/2, x_j)$). Par développement de Taylor, puis en utilisant le fait que u est solution de l'équation de la chaleur, on a

$$\begin{aligned} \frac{u(t_{n+1}, x_j) - u(t_n, x_j)}{\Delta t} &= \\ &= \left(\frac{\partial u}{\partial t} + \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2} + \frac{(\Delta t)^2}{6} \frac{\partial^3 u}{\partial t^3} \right) (t_n, x_j) + o((\Delta t)^2) \\ &= \left(\nu \frac{\partial^2 u}{\partial x^2} + \nu^2 \frac{\Delta t}{2} \frac{\partial^4 u}{\partial x^4} + \nu^3 \frac{(\Delta t)^2}{6} \frac{\partial^6 u}{\partial x^6} \right) (t_n, x_j) + o((\Delta t)^2). \end{aligned}$$

De même,

$$\begin{aligned} \frac{u(t_n, x_{j-1}) - 2u(t_n, x_j) + u(t_n, x_{j+1}))}{(\Delta x)^2} &= \\ &= \left(\frac{\partial^2 u}{\partial x^2} + \frac{(\Delta x)^2}{12} \frac{\partial^4 u}{\partial x^4} + 2 \frac{(\Delta x)^4}{6!} \frac{\partial^6 u}{\partial x^6} \right) (t_n, x_j) + o((\Delta x)^4). \end{aligned}$$

En remplaçant n par $n + 1$ dans l'expression précédente, on obtient suite à un

développement en (t_n, x_j) que

$$\begin{aligned} \frac{u(t_{n+1}, x_{j-1}) - 2u(t_{n+1}, x_j) + u(t_{n+1}, x_{j+1}))}{(\Delta x)^2} = \\ \left(\frac{\partial^2 u}{\partial x^2} + \frac{(\Delta x)^2}{12} \frac{\partial^4 u}{\partial x^4} + 2 \frac{(\Delta x)^4}{6!} \frac{\partial^6 u}{\partial x^6} \right) (t_n, x_j) \\ + \Delta t \left(\frac{\partial^3 u}{\partial t \partial x^2} + \frac{(\Delta x)^2}{12} \frac{\partial^5 u}{\partial t \partial x^4} \right) (t_n, x_j) \\ + \frac{(\Delta t)^2}{2} \left(\frac{\partial^4 u}{\partial t^2 \partial x^2} \right) (t_n, x_j) + o((\Delta x)^4 + (\Delta t)^2) \end{aligned}$$

De l'équation $\partial u / \partial t = \nu \partial^2 u / \partial x^2$, il vient

$$\begin{aligned} \frac{u(t_{n+1}, x_{j-1}) - 2u(t_{n+1}, x_j) + u(t_{n+1}, x_{j+1}))}{(\Delta x)^2} = \left(\frac{\partial^2 u}{\partial x^2} + \left(\frac{(\Delta x)^2}{12} + \nu \Delta t \right) \frac{\partial^4 u}{\partial x^4} \right. \\ \left. + \left(2 \frac{(\Delta x)^4}{6!} + \frac{\nu(\Delta t)(\Delta x)^2}{12} + \frac{\nu^2(\Delta t)^2}{2} \right) \frac{\partial^6 u}{\partial x^6} \right) (t_n, x_j) + o((\Delta x)^4 + (\Delta t)^2) \end{aligned}$$

1. Consistance des schémas explicite, implicite, θ -schéma et Crank-Nicholson.
Par combinaison linéaire, des développements calculés précédemment,

$$\begin{aligned} \frac{u(t_{n+1}, x_j) - u(t_n, x_j)}{\Delta t} + \theta \nu \frac{-u(t_{n+1}, x_{j-1}) + 2u(t_{n+1}, x_j) - u(t_{n+1}, x_{j+1}))}{(\Delta x)^2} \\ + (1 - \theta) \nu \frac{-u(t_n, x_{j-1}) + 2u(t_n, x_j) - u(t_n, x_{j+1}))}{(\Delta x)^2} \\ = (1 - \theta - (1 - \theta)) \nu \frac{\partial^2 u}{\partial x^2} \\ + \left(\frac{\nu \Delta t}{2} - \theta \left(\frac{(\Delta x)^2}{12} + \nu \Delta t \right) - (1 - \theta) \frac{(\Delta x)^2}{12} \right) \nu \frac{\partial^4 u}{\partial x^4} \\ + \left(\frac{1}{6} - \frac{\theta}{2} \right) (\Delta t)^2 \nu^3 \frac{\partial^6 u}{\partial x^6} + o((\Delta t)^2 + (\Delta x)^2). \end{aligned}$$

Après simplification,

$$\begin{aligned} \frac{u(t_{n+1}, x_j) - u(t_n, x_j)}{\Delta t} + \theta \nu \frac{-u(t_{n+1}, x_{j-1}) + 2u(t_{n+1}, x_j) - u(t_{n+1}, x_{j+1}))}{(\Delta x)^2} \\ + (1 - \theta) \nu \frac{-u(t_n, x_{j-1}) + 2u(t_n, x_j) - u(t_n, x_{j+1}))}{(\Delta x)^2} \\ = \left(\left(\left(\frac{1}{2} - \theta \right) \nu \Delta t - \frac{(\Delta x)^2}{12} \right) \nu \frac{\partial^4 u}{\partial x^4} + \left(\frac{1}{6} - \frac{\theta}{2} \right) (\Delta t)^2 \nu^3 \frac{\partial^6 u}{\partial x^6} \right) (t_n, x_j) \\ + o((\Delta t)^2 + (\Delta x)^2). \end{aligned}$$

Ainsi pour le $\theta \neq 1/2$ (en particulier pour les schémas explicite et implicite), le θ -schéma est d'ordre un en temps et deux en espace, tandis que le schéma de Crank-Nicholson est d'ordre deux en temps et en espace.

2. Consistance du schéma à 6 points.

Il nous reste à considérer le terme

$$\frac{u(t_{n+1}, x_{j+1}) - u(t_n, x_{j+1})}{\Delta t} + \frac{u(t_{n+1}, x_{j-1}) - u(t_n, x_{j-1})}{\Delta t}$$

D'après le développement effectué au début de l'exercice, puis en développant le résultat obtenu en (t_n, x_j) , on a

$$\begin{aligned} & \frac{u(t_{n+1}, x_{j+1}) - u(t_n, x_{j+1})}{\Delta t} + \frac{u(t_{n+1}, x_{j-1}) - u(t_n, x_{j-1})}{\Delta t} \\ &= \nu \left(\frac{\partial^2 u}{\partial x^2} + \frac{\nu \Delta t}{2} \frac{\partial^4 u}{\partial x^4} + \frac{\nu^2 (\Delta t)^2}{6} \frac{\partial^6 u}{\partial x^6} \right) (t_n, x_{j+1}) \\ & \quad + \nu \left(\frac{\partial^2 u}{\partial x^2} + \frac{\nu \Delta t}{2} \frac{\partial^4 u}{\partial x^4} + \frac{\nu^2 (\Delta t)^2}{6} \frac{\partial^6 u}{\partial x^6} \right) (t_n, x_{j-1}) + o((\Delta t)^2) \\ &= \left(2\nu \left(\frac{\partial^2 u}{\partial x^2} + \frac{\nu \Delta t}{2} \frac{\partial^4 u}{\partial x^4} + \frac{\nu^2 (\Delta t)^2}{6} \frac{\partial^6 u}{\partial x^6} \right) + \nu (\Delta x)^2 \left(\frac{\partial^4 u}{\partial x^4} + \frac{\nu \Delta t}{2} \frac{\partial^6 u}{\partial x^6} \right) \right. \\ & \quad \left. + \frac{\nu (\Delta x)^4}{12} \frac{\partial^6 u}{\partial x^6} \right) (t_n, x_j) + o((\Delta t)^2 + (\Delta x)^4) \end{aligned}$$

Soit,

$$\begin{aligned} & \frac{u(t_{n+1}, x_{j+1}) - u(t_n, x_{j+1})}{\Delta t} + \frac{u(t_{n+1}, x_{j-1}) - u(t_n, x_{j-1})}{\Delta t} \\ &= 2\nu \frac{\partial^2 u}{\partial x^2} + (\nu^2 \Delta t + \nu (\Delta x)^2) \frac{\partial^4 u}{\partial x^4} + \left(\frac{\nu^3 (\Delta t)^2}{3} + \frac{\nu^2 \Delta t (\Delta x)^2}{2} + \frac{\nu (\Delta x)^4}{12} \right) \frac{\partial^6 u}{\partial x^6} \\ & \quad + o((\Delta t)^2 + (\Delta x)^4) \end{aligned}$$

Par combinaison linéaire avec les autres développements effectués, on obtient (après simplification)

$$\begin{aligned} & \frac{u(t_{n+1}, x_{j+1}) - u(t_n, x_{j+1})}{12\Delta t} + \frac{5(u(t_{n+1}, x_j) - u(t_n, x_j))}{6\Delta t} \\ & + \frac{u(t_{n+1}, x_{j-1}) - u(t_n, x_{j-1})}{12\Delta t} - \nu \frac{u(t_{n+1}, x_{j-1}) - 2u(t_{n+1}, x_j) + u(t_{n+1}, x_{j+1})}{2(\Delta x)^2} \\ & \quad - \nu \frac{u(t_n, x_{j-1}) - 2u(t_n, x_j) + u(t_n, x_{j+1})}{2(\Delta x)^2} \\ &= \left(\frac{3}{6!} \nu (\Delta x)^4 - \frac{\nu^3}{12} (\Delta t)^2 \right) \frac{\partial^6 u}{\partial x^6} + o((\Delta x)^4 + (\Delta t)^2). \end{aligned}$$

Le schéma à 6 points est donc d'ordre 4 en espace et 2 en temps.

3. Consistance du schéma de DuFort-Frankel (2.7)

$$\frac{u(t_{n+1}, x_j) - u(t_{n-1}, x_j)}{2\Delta t} = \frac{\partial u}{\partial t} + o\left(\left(\frac{\Delta t}{\Delta x}\right)^2\right)$$

et

$$\begin{aligned} & \frac{-u(t_n, x_{j-1}) + u(t_{n+1}, x_j) + u(t_{n-1}, x_j) - u(t_n, x_{j+1}))}{(\Delta x)^2} \\ &= -\frac{\partial^2 u}{\partial x^2} + \left(\frac{\Delta t}{\Delta x}\right)^2 \frac{\partial^2 u}{\partial t^2} - \frac{(\Delta x)^2}{12} \frac{\partial^4 u}{\partial x^4} + o\left(\left(\frac{\Delta t}{\Delta x}\right)^2 + (\Delta x)^2\right) \end{aligned}$$

En combinant ces deux expressions, on en déduit que si u est solution de l'équation de la chaleur,

$$\begin{aligned} & \frac{u(t_{n+1}, x_j) - u(t_{n-1}, x_j)}{2\Delta t} \\ &+ \nu \frac{-u(t_n, x_{j-1}) + u(t_{n+1}, x_j) + u(t_{n-1}, x_j) - u(t_n, x_{j+1}))}{(\Delta x)^2} \\ &= \left(\left(\frac{\Delta t}{\Delta x}\right)^2 - \frac{(\Delta x)^2}{12}\right) \frac{\partial^4 u}{\partial x^4} + o\left(\left(\frac{\Delta t}{\Delta x}\right)^2 + (\Delta x)^2\right) \end{aligned}$$

Le schéma est d'ordre $\mathcal{O}((\Delta t/\Delta x)^2 + (\Delta x)^2)$.

4. Consistance du schéma de Gear (2.8)

$$\begin{aligned} & 3u(t_{n+1}, x_j) - 4u(t_n, x_j) + u(t_{n-1}, x_j) \\ &= 2(\Delta t) \frac{\partial u}{\partial t} - \frac{2}{3}(\Delta t)^3 \frac{\partial^3 u}{\partial t^3}(t_{n+1}, x_j) + \mathcal{O}((\Delta t)^4) \end{aligned}$$

et

$$\begin{aligned} & -u(t_{n+1}, x_{j-1}) + 3u(t_{n+1}, x_j) - u(t_{n+1}, x_{j+1}) \\ &= -(\Delta x)^2 \frac{\partial^2 u}{\partial x^2} - \frac{(\Delta x)^4}{24} \frac{\partial^4 u}{\partial x^4} + \mathcal{O}((\Delta x)^6). \end{aligned}$$

En appliquant ces deux développements de Taylor à la solution u de l'équation des ondes, on obtient

$$\begin{aligned} & \frac{3u(t_{n+1}, x_j) - 4u(t_n, x_j) + u(t_{n-1}, x_j)}{(\Delta x)^2} \\ &+ \nu \frac{-u(t_{n+1}, x_{j-1}) + 3u(t_{n+1}, x_j) - u(t_{n+1}, x_{j+1}))}{2\Delta t} \\ &= -\frac{\nu(\Delta x)^2}{24} \frac{\partial^4 u}{\partial x^4} + \frac{(\Delta t)^2}{3} \frac{\partial^6 u}{\partial x^6} + \mathcal{O}((\Delta t)^3 + (\Delta x)^3) \end{aligned}$$

Le schéma de Gear est donc d'ordre 2 en temps et en espace.

Exercice 2.2.3 Montrer que le schéma de Crank-Nicholson (2.2) (avec $\theta = 1/2$) est stable en norme L^∞ si $\nu\Delta t \leq (\Delta x)^2$, et que le schéma de DuFort-Frankel

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t} + \nu \frac{-u_{j-1}^n + u_j^{n+1} + u_j^{n-1} - u_{j+1}^n}{(\Delta x)^2} = 0, \quad (2.3)$$

est stable en norme L^∞ si $2\nu\Delta t \leq (\Delta x)^2$ (on applique aux deux schémas des conditions aux limites de type Dirichlet homogènes).

Correction. On va montrer que sous une condition CFL appropriée, le schéma de Crank-Nicholson vérifie le principe du maximum discret.

Soit k et l tels que

$$u_k^{n+1} = M = \max_j u_j^{n+1} \text{ et } u_l^{n+1} = m = \min_j u_j^{n+1}.$$

Notons que M est positif ou nul et m négatif ou nul. On va montrer que

$$M \leq \max(0, \max_j u_j^n) \quad (2.4)$$

$$\text{et } \min(0, \min_j u_j^n) \leq m. \quad (2.5)$$

Dans un premier temps, on considère l'inégalité (2.4). Cette dernière est trivialement vérifiée si $M = 0$. On peut donc se restreindre au cas $M \neq 0$. Le maximum de u_j^{n+1} pour tout $j \in \{0, \dots, N+1\}$ est atteint en un élément $k \in \{1, \dots, N\}$ et d'après (2.2) avec $\theta = 1/2$,

$$\frac{M - u_k^n}{\Delta t} + \nu \frac{-u_{k-1}^n + 2u_k^n - u_{k+1}^n}{2(\Delta x)^2} \leq 0,$$

soit

$$M \leq \left(1 - \frac{\nu\Delta t}{(\Delta x)^2}\right) u_k^n + \frac{\nu\Delta t}{2(\Delta x)^2} (u_{k-1}^n + u_{k+1}^n).$$

Si

$$\nu\Delta t \leq (\Delta x)^2, \quad (2.6)$$

le terme de droite est une combinaison convexe des coordonnées de u^n , et le premier point de (2.4) est vérifié. La minoration de m s'en déduit en remplaçant u^n par $-u^n$ et M par $-m$. Si la condition CFL (2.6) est vérifiée, le schéma de Crank-Nicholson vérifie le principe du maximum discret. En conséquence, il est stable pour la norme L^∞ .

Le schéma de DuFort-Frankel (2.3) est défini par

$$\left(\frac{1}{2\Delta t} + \frac{\nu}{(\Delta x)^2}\right) u_j^{n+1} = \left(\frac{1}{2\Delta t} - \frac{\nu}{(\Delta x)^2}\right) u_j^{n-1} + \frac{\nu}{(\Delta x)^2} (u_{j-1}^n + u_{j+1}^n)$$

Si $2\nu\Delta t \leq (\Delta x)^2$, u_j^{n+1} est une combinaison convexe de u_j^{n-1} , u_{j-1}^n et u_{j+1}^n . Ainsi, il est stable pour la norme L^∞ , c'est à dire

$$\|u^n\|_\infty \leq \max(\|u^0\|_\infty, \|u^1\|_\infty).$$

Finalement, on peut remarquer que la différence de traitement des deux schémas est due à leur nature : implicite pour le schéma de Crank-Nicholson, explicite pour le schéma de DuFort-Frankel.

Exercice 2.2.4 Montrer que le θ -schéma (2.2) est stable en norme L^2 inconditionnellement si $1/2 \leq \theta \leq 1$, et sous la condition CFL $2(1 - 2\theta)\nu\Delta t \leq (\Delta x)^2$ si $0 \leq \theta < 1/2$.

Correction. Étudions la stabilité en norme L^2 du θ -schéma. Par application de la transformation de Fourier, il vient

$$\left(1 + \frac{2\theta\nu\Delta t}{(\Delta x)^2}(1 - \cos(2k\pi\Delta x))\right) \hat{u}^{n+1}(k) = \left(1 + \frac{2(\theta - 1)\nu\Delta t}{(\Delta x)^2}(1 - \cos(2k\pi\Delta x))\right) \hat{u}^n(k).$$

Ainsi, le schéma sera stable en norme L^2 dès que

$$\left|1 + \frac{2(\theta - 1)\nu\Delta t}{(\Delta x)^2}(1 - \cos(2k\pi\Delta x))\right| \leq \left|1 + \frac{2\theta\nu\Delta t}{(\Delta x)^2}(1 - \cos(2k\pi\Delta x))\right|$$

pour tout k , c'est à dire

$$\left|1 - \frac{4\nu\Delta t \sin^2(k\pi\Delta x)}{(\Delta x)^2 + 4\theta\nu\Delta t \sin^2(k\pi\Delta x)}\right| \leq 1$$

ou encore

$$0 \leq \frac{4\nu\Delta t \sin^2(k\pi\Delta x)}{(\Delta x)^2 + 4\theta\nu\Delta t \sin^2(k\pi\Delta x)} \leq 2.$$

Comme θ est positif, cette condition est équivalente à

$$(\Delta x)^2 \geq 2(1 - 2\theta)\nu\Delta t \sin^2(k\pi\Delta x).$$

Cette dernière relation est vérifiée pour tout k dès que $(1 - 2\theta) \leq 0$ ou $(\Delta x)^2 \geq 2(1 - 2\theta)\nu\Delta t$.

Exercice 2.2.5 Montrer que le schéma à 6 points (2.1) est inconditionnellement stable en norme L^2 .

Correction. Par transformation de Fourier appliquée au schéma à 6 points (2.1), on obtient

$$\left(\frac{\cos(2k\pi\Delta x)}{6\Delta t} + \frac{5}{6\Delta t}\right) (\hat{u}^{n+1} - \hat{u}^n) + \frac{\nu}{(\Delta x)^2}(4 - \cos(2k\pi\Delta x))(\hat{u}^{n+1} + \hat{u}^n) = 0,$$

c'est à dire

$$\left(5 + \cos(2k\pi\Delta x) + \frac{6\nu\Delta t}{(\Delta x)^2}(4 - \cos(2k\pi\Delta x))\right) \hat{u}^{n+1} = \left(5 + \cos(2k\pi\Delta x) - \frac{6\nu\Delta t}{(\Delta x)^2}(4 - \cos(2k\pi\Delta x))\right) \hat{u}^n.$$

Le schéma est donc L^2 -stable dès que

$$\begin{aligned} 5 + \cos(2k\pi\Delta x) + \frac{6\nu\Delta t}{(\Delta x)^2}(4 - \cos(2k\pi\Delta x)) \\ \geq \left| 5 + \cos(2k\pi\Delta x) - \frac{6\nu\Delta t}{(\Delta x)^2}(4 - \cos(2k\pi\Delta x)) \right|. \end{aligned}$$

Relation qui est trivialement vérifiée indépendamment de Δx et Δt .

Exercice 2.2.6 Montrer que le schéma de Gear

$$\frac{3u_j^{n+1} - 4u_j^n + u_j^{n-1}}{2\Delta t} + \nu \frac{-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}}{(\Delta x)^2} = 0 \quad (2.7)$$

est inconditionnellement stable et donc convergent en norme L^2 .

Correction. En appliquant la transformation de Fourier au schéma de Gear (2.7), on obtient

$$(3 + c \sin^2(k\pi\Delta x)) \hat{u}^{n+1} = 4\hat{u}^n - \hat{u}^{n-1}, \quad (2.8)$$

où $c = \frac{8\nu\Delta t}{(\Delta x)^2}$. On introduit les polynômes (dépendants implicitement de k , Δt et Δx)

$$P(X) = (3 + 8c \sin^2(k\pi\Delta x))X^2 - 4X + 1.$$

On note λ_1 et λ_2 les racines de P et $\Delta = (\lambda_2 - \lambda_1)^2$ son discriminant. Les solutions de (2.8) s'expriment explicitement en fonction de \hat{u}^0 et \hat{u}^1 :

$$\hat{u}^n = \begin{cases} \left(\frac{\lambda_2 \lambda_1^n - \lambda_1 \lambda_2^n}{\lambda_2 - \lambda_1} \right) \hat{u}^0 + \left(\frac{\lambda_2^n - \lambda_1^n}{\lambda_2 - \lambda_1} \right) \hat{u}^1 & \text{si } \Delta \neq 0, \\ (1-n)\lambda_1^n \hat{u}^0 + n\lambda_1^{n-1} \hat{u}^1 & \text{si } \Delta = 0. \end{cases}$$

Une condition nécessaire de stabilité est donc que $|\lambda_1|$ et $|\lambda_2|$ soient au plus égaux à un. Dans ce cas, afin que le schéma soit stable, il suffit qu'il existe deux réels δ et β tels que pour tout k , Δx et Δt ,

$$|\Delta(k, \Delta x, \Delta t)| \leq \delta \implies \max(|\lambda_1(k, \Delta x, \Delta t)|, |\lambda_2(k, \Delta x, \Delta t)|) < \beta < 1. \quad (2.9)$$

En effet, posons $C(\beta) = \max_n n\beta^{n-1}$. Comme $0 < \beta < 1$, $C(\beta) < +\infty$. De plus, si $|\Delta(k, \Delta x, \Delta t)| \geq \delta$,

$$\left| \frac{\lambda_2^n - \lambda_1^n}{\lambda_2 - \lambda_1} \right| \leq 2/\sqrt{\delta};$$

si $0 < |\Delta(k, \Delta x, \Delta t)| < \delta$,

$$\left| \frac{\lambda_2^n - \lambda_1^n}{\lambda_2 - \lambda_1} \right| = \left| \sum_{k=0}^{n-1} \lambda_1^k \lambda_2^{n-1-k} \right| \leq n \max(|\lambda_1|, |\lambda_2|)^{n-1} \leq C(\beta),$$

et si $\Delta(k, \Delta x, \Delta t) = 0$, $n|\lambda_1|^{n-1} \leq C(\beta)$. De ces trois inégalités, on en déduit que

$$|\hat{u}^n(k)| < K(|\hat{u}^0 + \hat{u}^1|)$$

où $K = 1 + 3(C(\beta) + 2/\sqrt{\delta})$. Ainsi, $\|\hat{u}^n\|_{L^2} \leq K(\|u^0\|_{L^2} + \|u^1\|_{L^2})$.

Reste à prouver que la condition de stabilité (2.9) est en effet vérifiée. Tout d'abord, on vérifie que pour tout k , Δx et Δt , $|\lambda_1| \leq 1$ et $|\lambda_2| \leq 1$. Enfin, λ_1 et λ_2 sont des fonctions continues de Δ . Or si $\Delta = 0$, $\lambda_1 = \lambda_2 = 1/2$. Il existe donc δ et β , $1/2 < \beta < 1$ tels que la condition (2.9) est vérifiée.

Exercice 2.2.7 Montrer que le schéma de DuFort-Frankel (2.3) est inconditionnellement stable en norme L^2 . Montrer que, si on fait tendre Δt et Δx vers 0 de telle manière que le rapport $\Delta t/\Delta x$ tende aussi vers 0, alors le schéma de DuFort-Frankel est convergent. (On dit qu'il est "conditionnellement" convergent.)

Correction. Par transformation de Fourier, on obtient que

$$\hat{u}^{n+1} - \hat{u}^{n-1} + \frac{2\nu\Delta t}{(\Delta x)^2}(-2\hat{u}^n \cos(2k\pi\Delta x) + \hat{u}^{n+1} + \hat{u}^{n-1}) = 0.$$

Soit encore

$$(1 + c)\hat{u}^{n+1}(k) - 2c \cos(k\pi\Delta x)\hat{u}^n(k) - (1 - c)\hat{u}^{n-1}(k) = 0,$$

où

$$c = \frac{2(\Delta t)\nu}{(\Delta x)^2}.$$

Notons que dans le cas $c \leq 1$, on a prouvé précédemment la stabilité L^∞ du schéma de DuFort-Frankel. Cette dernière impliquant la stabilité L^2 , nous n'avons plus qu'à étudier le cas $c > 1$. On procède comme pour l'exercice précédent. Considérons le polynôme

$$P(X) = (1 + c)X^2 - 2c \cos(k\pi\Delta x)X - (1 - c)$$

On vérifie sans mal que les racines de P sont de module inférieur ou égale à un. Et si Δ désigne le discriminant associé,

$$|\lambda_1|, |\lambda_2| \leq (1 + c)^{-1} (c|\cos(2k\pi\Delta x)| + |\Delta|^{1/2}/2).$$

Or $c^2 \cos^2(2k\pi\Delta x) = \Delta/4 + c^2 - 1$. Ainsi,

$$|\lambda_1|, |\lambda_2| \leq (1 + c)^{-1} (|\Delta/4 + c^2 - 1|^{1/2} + |\Delta|^{1/2}/2).$$

Le terme de gauche est continue par rapport à Δ et c . De plus, pour $\Delta = 0$, il est égale à $(\frac{c-1}{c+1})^{1/2}$. Sous la condition CFL $c < M$, il existe γ tel que $(\frac{c-1}{c+1})^{1/2} < \gamma < 1$. Comme $[1, M] \times 0$ est un compact, il existe δ et ε tel que pour tout $1 \leq c \leq M$,

$$|\Delta| \leq \delta \implies (|\Delta/4 + c^2 - 1|^{1/2} + |\Delta|^{1/2}/2) < \gamma + \varepsilon < 1.$$

La condition de stabilité (2.9) énoncée dans la correction de l'Exercice 2.2.5 est vérifiée. Le schéma est donc convergent pourvu que le rapport $\Delta t/(\Delta x)^2$ reste borné. Enfin, la stabilité combinée à la consistance implique la convergence.

Exercice 2.2.8 Montrer que le schéma explicite

$$\frac{u_{j,k}^{n+1} - u_{j,k}^n}{\Delta t} + \nu \frac{-u_{j-1,k}^n + 2u_{j,k}^n - u_{j+1,k}^n}{(\Delta x)^2} + \nu \frac{-u_{j,k-1}^n + 2u_{j,k}^n - u_{j,k+1}^n}{(\Delta y)^2} = 0 \quad (2.10)$$

est stable en norme L^∞ (et même qu'il vérifie le principe du maximum) sous la condition CFL

$$\frac{\nu \Delta t}{(\Delta x)^2} + \frac{\nu \Delta t}{(\Delta y)^2} \leq \frac{1}{2}.$$

Correction. Le schéma explicite (2.10) est défini par

$$\begin{aligned} u_{j,k}^{n+1} = & \left(1 - 2 \left(\frac{\nu \Delta t}{(\Delta x)^2} + \frac{\nu \Delta t}{(\Delta y)^2} \right) \right) u_{j,k}^n \\ & + \frac{\nu \Delta t}{(\Delta x)^2} (u_{j-1,k}^n + u_{j+1,k}^n) + \frac{\nu \Delta t}{(\Delta y)^2} (u_{j,k-1}^n + u_{j,k+1}^n) \end{aligned}$$

Si

$$\frac{\nu \Delta t}{(\Delta x)^2} + \frac{\nu \Delta t}{(\Delta y)^2} \leq 1/2,$$

$u_{j,k}^{n+1}$ est une combinaison convexe de coordonnées de u^n et

$$|u_{j,k}^{n+1}| \leq \|u^n\|_\infty.$$

Exercice 2.2.9 Montrer que le schéma de Peaceman-Rachford

$$\begin{aligned} \frac{u_{j,k}^{n+1/2} - u_{j,k}^n}{\Delta t} + \nu \frac{-u_{j-1,k}^{n+1/2} + 2u_{j,k}^{n+1/2} - u_{j+1,k}^{n+1/2}}{2(\Delta x)^2} + \nu \frac{-u_{j,k-1}^n + 2u_{j,k}^n - u_{j,k+1}^n}{2(\Delta y)^2} &= 0 \\ \frac{u_{j,k}^{n+1} - u_{j,k}^{n+1/2}}{\Delta t} + \nu \frac{-u_{j-1,k}^{n+1/2} + 2u_{j,k}^{n+1/2} - u_{j+1,k}^{n+1/2}}{2(\Delta x)^2} + \nu \frac{-u_{j,k-1}^{n+1/2} + 2u_{j,k}^{n+1/2} - u_{j,k+1}^{n+1/2}}{2(\Delta y)^2} &= 0. \end{aligned}$$

est précis d'ordre 2 en espace et temps et inconditionnellement stable en norme L^2 (pour des conditions aux limites de périodicité dans chaque direction).

Correction. 1. Consistance

En effectuant la soustraction des deux équations définissant le schéma, on obtient l'expression de $u^{n+1/2}$ en fonction de u^n et u^{n+1} .

$$u_{j,k}^{n+1/2} = \frac{u_{j,k}^{n+1} + u_{j,k}^n}{2} + \frac{\nu \Delta t}{4(\Delta y)^2} (u_{j,k-1}^n - 2u_{j,k}^n + u_{j,k+1}^n - u_{j,k-1}^{n+1} + 2u_{j,k}^{n+1} - u_{j,k+1}^{n+1}).$$

En substituant l'expression de $u^{n+1/2}$ dans l'une des équations du schéma, on détermine la relation reliant u^{n+1} à u^n . On pourrait effectuer le calcul explicite de cette

expression, puis établir la consistance. Cependant, cela constitue un calcul fastidieux qu'on peut éviter. On introduit donc la fonction intermédiaire

$$v_{\Delta t, \Delta x, \Delta y}(t, x, y) = \frac{u(t + \Delta t, x, y) + u(t, x, y)}{2} + \frac{\nu \Delta t}{4(\Delta y)^2} \left(u(t, x, y - \Delta y) - 2u(t, x, y) + u(t, x, y + \Delta y) - u(t + \Delta t, x, y - \Delta y) + 2u(t + \Delta t, x, y) - u(t + \Delta t, x, y + \Delta y) \right) \quad (2.11)$$

Pour toute solution u de l'équation de la chaleur, l'erreur de troncature est

$$E(u) = \frac{v(t, x, y) - u(t, x, y)}{\Delta t} + \nu \frac{-v(t, x - \Delta x, y) + 2v(t, x, y) - v(t, x + \Delta x, y)}{2(\Delta x)^2} + \nu \frac{-u(t, x, y - \Delta y) + 2u(t, x, y) - u(t, x, y + \Delta y)}{2(\Delta y)^2}$$

où v est définie par (2.11). Par développement de Taylor, on établit que

$$v_{\Delta t, \Delta x, \Delta y} = u + \frac{\Delta t}{2} \frac{\partial u}{\partial t} + \frac{(\Delta t)^2}{4} \left(\frac{\partial^2 u}{\partial t^2} - \nu \frac{\partial^3 u}{\partial t \partial y^2} \right) + \frac{(\Delta t)^3}{24} \left(2 \frac{\partial^3 u}{\partial t^3} - 3\nu \frac{\partial^4 u}{\partial t^2 \partial y^2} \right) + o((\Delta t)^3 + (\Delta t)(\Delta y)^2)$$

puis que

$$E(u) = \frac{v - u}{\Delta t} - \frac{\nu}{2} \left(\frac{\partial^2 v}{\partial x^2} + \frac{(\Delta x)^2}{12} \frac{\partial^4 v}{\partial x^4} + \frac{\partial^2 u}{\partial y^2} + \frac{(\Delta y)^2}{12} \frac{\partial^4 u}{\partial y^4} \right) + o((\Delta x)^2 + (\Delta y)^2) \\ = \frac{\nu^3 (\Delta t)^2}{24} \Delta \left(\frac{\partial^4 u}{\partial x^2 \partial y^2} - \Delta^2 u \right) - \frac{\nu}{24} \left((\Delta x)^2 \frac{\partial^4 u}{\partial x^4} + (\Delta y)^2 \frac{\partial^4 u}{\partial y^4} \right) + o((\Delta x)^2 + (\Delta y)^2 + (\Delta t)^2).$$

L'erreur de troncature est d'ordre 2 en espace et en temps.

2. Étude de la stabilité L^2

En appliquant la transformation de Fourier au schéma, on en déduit que

$$\hat{u}^{n+1/2} = \frac{1 - 2 \frac{\nu \Delta t}{(\Delta y)^2} \sin^2(l\pi \Delta y)}{1 + 2 \frac{\nu \Delta t}{(\Delta x)^2} \sin^2(k\pi \Delta x)} \hat{u}^n$$

et

$$\hat{u}^{n+1} = \frac{1 - 2 \frac{\nu \Delta t}{(\Delta x)^2} \sin^2(k\pi \Delta x)}{1 + 2 \frac{\nu \Delta t}{(\Delta y)^2} \sin^2(l\pi \Delta y)} \hat{u}^{n+1/2}.$$

Ainsi, $\hat{u}^{n+1}(k, l) = A(k, l) \hat{u}^n(k, l)$ où

$$A(k, l) = \frac{1 - 2 \frac{\nu \Delta t}{(\Delta y)^2} \sin^2(l\pi \Delta y)}{1 + 2 \frac{\nu \Delta t}{(\Delta y)^2} \sin^2(l\pi \Delta y)} \frac{1 - 2 \frac{\nu \Delta t}{(\Delta x)^2} \sin^2(k\pi \Delta x)}{1 + 2 \frac{\nu \Delta t}{(\Delta x)^2} \sin^2(k\pi \Delta x)}.$$

Comme pour tout $x \geq 0$, $|(1-x)/(1+x)| \leq 1$, on a $|A(k, l)| \leq 1$. Le schéma est inconditionnellement stable en norme L^2 .

Exercice 2.2.10 Montrer que le schéma de directions alternées (2.31) est précis d'ordre 2 en espace et temps et inconditionnellement stable en norme L^2 (pour des conditions aux limites de périodicité dans chaque direction).

Correction.

1. Étude de la consistance

Le schéma se décompose en deux étapes

$$\left(\frac{\text{Id}}{\Delta t} - \frac{\nu}{2} M_y \right) u^{n+1/2} - \left(\frac{\text{Id}}{\Delta t} + \frac{\nu}{2} M_y \right) u^n = 0$$

et

$$\left(\frac{\text{Id}}{\Delta t} - \frac{\nu}{2} M_x \right) u^{n+1} - \left(\frac{\text{Id}}{\Delta t} + \frac{\nu}{2} M_x \right) u^{n+1/2} = 0,$$

où

$$(M_x v)_{j,k} = \frac{v_{j+1,k} - 2v_{j,k} + v_{j-1,k}^n}{(\Delta x)^2}$$

et

$$(M_y v)_{j,k} = \frac{v_{j,k+1} - 2v_{j,k} + v_{j,k-1}^n}{(\Delta y)^2}.$$

Afin d'appliquer la définition de la consistance donnée dans le cours, il faut exhiber la relation reliant u^{n+1} à u^n . Il faut donc supprimer l'inconnue intermédiaire $u^{n+1/2}$ des équations définissant le schéma numérique. A cet effet, il suffit de multiplier la deuxième équation par $(\frac{\text{Id}}{\Delta t} - \frac{\nu}{2} M_y)$ et de constater que cette matrice commute avec $(\frac{\text{Id}}{\Delta t} + \frac{\nu}{2} M_x)$. On obtient ainsi

$$\begin{aligned} \left(\text{Id} - \frac{\nu \Delta t}{2} M_y \right) \left(\text{Id} - \frac{\nu \Delta t}{2} M_x \right) u^{n+1} = \\ \left(\text{Id} + \frac{\nu \Delta t}{2} M_x \right) \left(\text{Id} - \frac{\nu \Delta t}{2} M_y \right) u^{n+1/2}. \end{aligned}$$

D'après la première équation du schéma, il vient

$$\begin{aligned} (\Delta t)^{-1} \left(\text{Id} - \frac{\nu \Delta t}{2} M_y \right) \left(\text{Id} - \frac{\nu \Delta t}{2} M_x \right) u^{n+1} - \\ (\Delta t)^{-1} \left(\text{Id} + \frac{\nu \Delta t}{2} M_x \right) \left(\text{Id} + \frac{\nu \Delta t}{2} M_y \right) u^n = 0. \end{aligned}$$

Pour toute fonction v , on note $M_y(v)$ la fonction définie par

$$M_y(v)(t, x, y) = \frac{v(t, x, y + \Delta y) - 2v(t, x, y) + v(t, x, y - \Delta y)}{(\Delta y)^2}.$$

On définit de même la fonction $M_x(v)$ en échangeant les rôles respectifs de x et y . De plus, on note

$$\tau(v)(t, x, y) = v(t + \Delta t, x, y).$$

En effectuant un développement de Taylor en (t, x, y) , on montre que

$$M_y(v) = \frac{\partial^2 v}{\partial y^2} + \mathcal{O}((\Delta y)^2).$$

On en déduit que

$$\begin{aligned} (\Delta t)^{-1} \left(\text{Id} - \frac{\nu \Delta t}{2} M_y \right) \left(\text{Id} - \frac{\nu \Delta t}{2} M_x \right) (\tau(v)) = \\ \tau \left(\frac{v}{\Delta t} - \frac{\nu}{2} \Delta v + \frac{\nu^2 \Delta t}{4} \frac{\partial^4 v}{\partial x^2} \partial y^2 + \mathcal{O}((\Delta x)^2 + (\Delta y)^2) \right) \end{aligned}$$

De même,

$$\begin{aligned} (\Delta t)^{-1} \left(\text{Id} + \frac{\nu \Delta t}{2} M_x \right) \left(\text{Id} + \frac{\nu \Delta t}{2} M_y \right) (v) = \\ \frac{v}{\Delta t} + \frac{\nu}{2} \Delta v + \frac{\nu^2 \Delta t}{4} \frac{\partial^4 v}{\partial x^2} \partial y^2 + \mathcal{O}((\Delta x)^2 + (\Delta y)^2). \end{aligned}$$

Ainsi,

$$\begin{aligned} (\Delta t)^{-1} \left(\text{Id} - \frac{\nu \Delta t}{2} M_y \right) \left(\text{Id} - \frac{\nu \Delta t}{2} M_x \right) (\tau(v)) \\ - (\Delta t)^{-1} \left(\text{Id} + \frac{\nu \Delta t}{2} M_x \right) \left(\text{Id} + \frac{\nu \Delta t}{2} M_y \right) (v) = \\ \frac{\tau(v) - v}{\Delta t} - \nu \Delta \left(\frac{\tau(v) + v}{2} \right) + \mathcal{O}(\Delta t + (\Delta x)^2 + (\Delta y)^2) = \\ \frac{\partial v}{\partial t} - \nu \Delta v + \mathcal{O}(\Delta t + (\Delta x)^2 + (\Delta y)^2), \end{aligned}$$

d'où on déduit que le schéma est d'ordre 2 en espace et 1 en temps.

Remarque 2.2.1 *Le point essentiel sur lequel repose la démonstration de la consistance porte sur la propriété de commutation employée au début de la preuve.*

2. étude de la stabilité L^2

En appliquant la transformation de Fourier au schéma, on établit que

$$\hat{u}^{n+1/2} = \frac{1 - \frac{\nu \Delta t}{(\Delta x)^2} \sin^2(\pi k \Delta x)}{1 + \frac{\nu \Delta t}{(\Delta x)^2} \sin^2(\pi k \Delta x)} \hat{u}^n$$

et

$$\hat{u}^{n+1} = \frac{1 - \frac{\nu \Delta t}{(\Delta y)^2} \sin^2(\pi l \Delta y)}{1 + \frac{\nu \Delta t}{(\Delta y)^2} \sin^2(\pi l \Delta y)} \hat{u}^{n+1/2}.$$

Ainsi, $|\hat{u}^{n+1}| \leq |\hat{u}^{n+1/2}| \leq |\hat{u}^n|$ et le schéma est inconditionnellement stable L^2 .

Exercice 2.3.1 Montrer que le schéma implicite centré

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + V \frac{u_{j+1}^{n+1} - u_{j-1}^{n+1}}{2\Delta x} = 0 \quad (2.12)$$

est consistant avec l'équation d'advection (2.32), précis à l'ordre 1 en temps et 2 en espace, inconditionnellement stable en norme L^2 , donc convergent.

Correction. La consistance et la précision ne posent pas de problèmes. Pour la stabilité L^2 , l'analyse de Fourier conduit à

$$\hat{u}^{n+1}(k) = \left(1 + i \frac{V\Delta t}{\Delta x} \sin(2\pi k\Delta x)\right)^{-1} \hat{u}^n(k) = A(k)\hat{u}^n(k).$$

On vérifie alors que le module du facteur d'amplification est toujours plus petit que 1

$$|A(k)|^2 = \left(1 + \left(\frac{V\Delta t}{\Delta x} \sin(2\pi k\Delta x)\right)^2\right)^{-1} \leq 1.$$

Le schéma est inconditionnellement stable. La convergence s'obtient alors par le Théorème de Lax 2.2.20.

Exercice 2.3.2 Montrer que le schéma de Lax Wendroff

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + V \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} - \left(\frac{V^2\Delta t}{2}\right) \frac{u_{j-1}^n - 2u_j^n + u_{j+1}^n}{(\Delta x)^2} = 0. \quad (2.13)$$

est stable et convergent en norme L^2 si $|V|\Delta t \leq \Delta x$.

Correction. Il suffit de montrer la stabilité en norme L^2 afin d'en déduire la convergence par le théorème de Lax. En appliquant la transformation de Fourier au schéma de Lax-Wendroff (2.13), on obtient

$$\hat{u}^{n+1}(k) = A(k)\hat{u}^n(k)$$

où

$$A(k) = 1 - 2 \left(\frac{V\Delta t}{\Delta x}\right)^2 \sin^2(k\pi\Delta x) - i \frac{V\Delta t}{\Delta x} \sin(2k\pi\Delta x)$$

Le schéma est stable en norme L^2 dès que $|A(k)| \leq 1$. On montre aisément que

$$|A(k)|^2 = 1 - 4 \sin^4(k\pi\Delta x) \left(\frac{V\Delta t}{\Delta x}\right)^2 \left(1 - \left(\frac{V\Delta t}{\Delta x}\right)^2\right).$$

Ainsi, le schéma est stable et convergent dès que

$$\frac{|V|\Delta t}{\Delta x} \leq 1.$$

Exercice 2.3.3 Montrer que le schéma de Lax-Friedrichs préserve le principe du maximum discret si la condition CFL $|V|\Delta t \leq \Delta x$ est satisfaite, tandis que le schéma de Lax-Wendroff ne le préserve pas sauf si $V\Delta t/\Delta x$ vaut $-1, 0$, ou 1 .

Correction. 1. Schéma de Lax-Friedrichs

$$u_j^{n+1} = \left(\frac{1}{2} + \frac{V\Delta t}{2\Delta x}\right) u_{j+1}^n + \left(\frac{1}{2} - \frac{V\Delta t}{2\Delta x}\right) u_{j-1}^n.$$

Ainsi, u_j^{n+1} est une combinaison linéaire convexe de u_{j+1}^n et u_j^n dès que $|V|\Delta t \leq \Delta x$. Sous cette condition, le schéma vérifie le principe du maximum discret.

2. Schéma de Lax-Wendroff

$$u_j^{n+1} = \frac{V\Delta t}{2\Delta x} \left(\frac{V\Delta t}{\Delta x} - 1\right) u_{j+1}^n + \left(1 - \left(\frac{V\Delta t}{\Delta x}\right)^2\right) u_j^n + \frac{V\Delta t}{2\Delta x} \left(\frac{V\Delta t}{\Delta x} + 1\right) u_{j-1}^n.$$

Le schéma préserve le principe du maximum discret si et seulement si chacun des coefficients apparaissant dans le terme de droite est positif, c'est à dire si et seulement si $V\Delta t/\Delta x = -1, 0$ ou 1 .

Exercice 2.3.4 Montrer que le schéma de Lax-Wendroff (2.13) est le seul schéma précis à l'ordre 2 en espace et temps qui soit du type

$$u_j^{n+1} = \alpha u_{j-1}^n + \beta u_j^n + \gamma u_{j+1}^n,$$

où α, β, γ dépendent seulement de $V\Delta t/\Delta x$.

Correction. L'erreur de troncature est

$$E = (\Delta t)^{-1} (u(x_j, t_{n+1}) - \alpha u(x_{j-1}, t_n) - \beta u(t_n, x_j) - \gamma u(t_n, x_{j+1})).$$

En effectuant un développement de Taylor en (x_j, t_n) , on montre que

$$E = (\Delta t)^{-1} (1 - (\alpha + \beta + \gamma)) u + \frac{\partial u}{\partial t} + \frac{\Delta x}{\Delta t} (\alpha - \gamma) \frac{\partial u}{\partial x} + \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2} - \frac{\alpha + \gamma}{2} \frac{(\Delta x)^2}{\Delta t} \frac{\partial^2 u}{\partial x^2} + \mathcal{O}\left((\Delta t)^2 + \frac{|\alpha - \gamma|}{c} (\Delta x)^2\right).$$

Si u est solution de l'équation d'advection,

$$\partial u / \partial t = -V \partial u / \partial x \text{ et } \partial^2 u / \partial t^2 = V^2 \partial^2 u / \partial x^2.$$

Ainsi,

$$E = (\Delta t)^{-1} (1 - (\alpha + \beta + \gamma)) u - \frac{\Delta x}{\Delta t} (c - (\alpha - \gamma)) \frac{\partial u}{\partial x} + \frac{(\Delta x)^2}{2\Delta t} (c^2 - (\alpha + \gamma)) \frac{\partial^2 u}{\partial x^2} + \mathcal{O}\left((\Delta t)^2 \left(1 + \frac{|\alpha - \gamma|}{c^3}\right)\right), \quad (2.14)$$

où $c = V\Delta t/\Delta x$. Si le schéma est d'ordre 2 en temps et en espace, on doit avoir

$$\begin{aligned} 1 - (\alpha + \beta + \gamma) &= \mathcal{O}\left((\Delta t)^3 \left(1 + \frac{|\alpha - \gamma|}{c^3}\right)\right), \\ c - \alpha + \gamma &= \mathcal{O}\left(c(\Delta t)^2 \left(1 + \frac{|\alpha - \gamma|}{c^3}\right)\right), \\ c^2 - (\alpha + \gamma) &= \mathcal{O}\left(c^2(\Delta t) \left(1 + \frac{|\alpha - \gamma|}{c^3}\right)\right). \end{aligned}$$

En faisant tendre vers zéro Δt et Δx à c constant, on obtient le système linéaire suivant

$$\begin{aligned} 1 - (\alpha + \beta + \gamma) &= 0 \\ c - \alpha + \gamma &= 0 \\ c^2 - (\alpha + \gamma) &= 0. \end{aligned}$$

D'où l'on déduit que

$$\begin{aligned} \alpha &= c(1 + c)/2 \\ \beta &= 1 - c^2 \\ \gamma &= c(c - 1)/2. \end{aligned}$$

Enfin, comme $\alpha - \gamma = \mathcal{O}(c)$, d'après (2.14), le schéma est en effet au moins d'ordre 2 en espace et en temps.

Exercice 2.3.5 Montrer que le schéma explicite décentré amont

$$\begin{aligned} \frac{u_j^{n+1} - u_j^n}{\Delta t} + V \frac{u_j^n - u_{j-1}^n}{\Delta x} &= 0 \quad \text{si } V > 0 \\ \frac{u_j^{n+1} - u_j^n}{\Delta t} + V \frac{u_{j+1}^n - u_j^n}{\Delta x} &= 0 \quad \text{si } V < 0. \end{aligned} \tag{2.15}$$

est consistant avec l'équation d'advection (2.32), précis à l'ordre 1 en espace et temps, stable et convergent en norme L^2 si la condition CFL $|V|\Delta t \leq \Delta x$ est satisfaite.

Correction. La consistance d'ordre 1 en temps et en espace est aisée à établir. En effet, dans le cas $V > 0$,

$$\begin{aligned} \frac{u(t_{n+1}, x_j) - u(t_n, x_j)}{\Delta t} + V \frac{u(t_n, x_j) - u(t_n, x_{j-1})}{\Delta x} \\ = (u_t + Vu_x)(t_n, x_j) + \mathcal{O}(\Delta t + \Delta x). \end{aligned}$$

Le cas $V < 0$ est identique. Enfin, la stabilité L^2 se déduit de la stabilité L^∞ .

Exercice 2.3.6 Montrer que l'équation équivalente du schéma décentré amont (2.15) est

$$\frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} - \frac{|V|}{2} (\Delta x - |V|\Delta t) \frac{\partial^2 u}{\partial x^2} = 0.$$

Correction. Considérons le cas $V > 0$. L'erreur de troncature du schéma décentré amont (2.15) est

$$E = \frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} + \frac{\Delta t}{2} \frac{\partial^2 u}{\partial t^2} - \frac{V \Delta x}{2} \frac{\partial^2 u}{\partial x^2} + \mathcal{O}((\Delta t)^2 + (\Delta x)^2).$$

Soit u tel que l'erreur de troncature du schéma décentré soit d'ordre 2 en espace et en temps, on a

$$\frac{\partial^2 u}{\partial t^2} = V^2 \frac{\partial^2 u}{\partial x^2} + \mathcal{O}(\Delta t + \Delta x).$$

Ainsi, l'erreur de troncature pour u est égale à

$$\frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} + \frac{V}{2} (V \Delta t - \Delta x) \frac{\partial^2 u}{\partial x^2} + \mathcal{O}((\Delta t)^2 + (\Delta x)^2).$$

L'équation équivalente dans le cas $V > 0$ est donc

$$\frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} + \frac{V}{2} (V \Delta t - \Delta x) \frac{\partial^2 u}{\partial x^2} = 0.$$

Il suffit de substituer Δx par $-\Delta x$ pour obtenir l'équation équivalente dans le cas $V < 0$. Enfin, on peut résumer ces deux résultats par l'équation

$$\frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} - \frac{|V|}{2} (\Delta x - |V| \Delta t) \frac{\partial^2 u}{\partial x^2} = 0$$

valable dans les deux cas.

Exercice 2.3.7 Montrer que l'équation équivalente du schéma de Lax-Wendroff (2.13) est

$$\frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} + \frac{V(\Delta x)^2}{6} \left(1 - \frac{(V \Delta t)^2}{(\Delta x)^2} \right) \frac{\partial^3 u}{\partial x^3} = 0.$$

Correction. L'erreur de troncature dans le cas du schéma de Lax-Wendroff (2.13) est

$$E(u) = \frac{u(t_{n+1}, x_j) - u(t_n, x_j)}{\Delta t} + V \frac{u(t_n, x_{j+1}) - u(t_n, x_{j-1})}{2\Delta x} - \left(\frac{V^2 \Delta t}{2} \right) \frac{u(t_n, x_{j-1}) - 2u(t_n, x_j) + u(t_n, x_{j+1}))}{(\Delta x)^2}.$$

En effectuant un développement de Taylor en (t_n, x_j) , on montre que

$$E(u) = \frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} + \frac{\Delta t}{2} \left(\frac{\partial^2 u}{\partial t^2} - V^2 \frac{\partial^2 u}{\partial x^2} \right) + \frac{(\Delta t)^2}{6} \frac{\partial^3 u}{\partial t^3} + \frac{V(\Delta x)^2}{6} \frac{\partial^3 u}{\partial x^3} + \mathcal{O}((\Delta x)^3 + (\Delta t)^3). \quad (2.16)$$

Soit u tel que $E(u)$ soit d'ordre 3 en espace et en temps, on montre que dans ce cas

$$\frac{\partial^2 u}{\partial t^2} = V^2 \frac{\partial^2 u}{\partial x^2} + \mathcal{O}((\Delta x)^2 + (\Delta t)^2)$$

De plus, $\partial^3 u / \partial t^3 = -V^3 \partial^3 u / \partial^3 x + \mathcal{O}(\Delta t + \Delta x)$. En substituant ces expressions dans l'équation (2.16), on obtient l'équation équivalente attendue :

$$\frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} + \frac{V(\Delta x)^2}{6} \left(1 - V^2 \left(\frac{\Delta t}{\Delta x} \right)^2 \right) \frac{\partial^3 u}{\partial x^3} = \mathcal{O}((\Delta x)^3 + (\Delta t)^3).$$

Exercice 2.3.8 Soit l'équation

$$\begin{cases} \frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} - \nu \frac{\partial^2 u}{\partial x^2} - \mu \frac{\partial^3 u}{\partial x^3} = 0 \text{ pour } (x, t) \in \mathbb{R} \times \mathbb{R}_*^+ \\ u(t = 0, x) = \sin(\omega x + \phi) \text{ pour } x \in \mathbb{R}, \end{cases}$$

avec $V, \nu, \mu, \omega, \phi \in \mathbb{R}$. Montrer que sa solution est

$$u(t, x) = \exp(-\nu\omega^2 t) \sin(\omega(x - (V + \mu\omega^2)t) + \phi)$$

(on admettra son unicité). En déduire que la diffusion atténue l'amplitude de la solution, tandis que la dispersion modifie la vitesse de propagation.

Correction. On détermine les dérivées partielles de u intervenant dans l'équation donnée. On a

$$\begin{aligned} \frac{\partial u}{\partial t} &= -\nu\omega^2 u + \omega(V + \mu\omega^2) \exp(-\nu\omega^2 t) \cos(\omega(x - (V + \mu\omega^2)t) + \phi) , \\ \frac{\partial u}{\partial x} &= -\omega \exp(-\nu\omega^2 t) \cos(\omega(x - (V + \mu\omega^2)t) + \phi) , \\ \frac{\partial^2 u}{\partial x^2} &= -\omega^2 u , \end{aligned}$$

et

$$\frac{\partial^3 u}{\partial x^3} = \omega^3 \exp(-\nu\omega^2 t) \cos(\omega(x - (V + \mu\omega^2)t) + \phi) .$$

En sommant ces différents termes, on obtient

$$\frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} - \nu \frac{\partial^2 u}{\partial x^2} - \mu \frac{\partial^3 u}{\partial x^3} = 0.$$

L'atténuation de l'amplitude est $\exp(-\nu\omega^2 t)$. Elle est donc d'autant plus forte que le terme de diffusion ν est important par rapport à ω^{-2} . La vitesse de propagation de l'onde est $(V + \mu\omega^2)$ et dépend donc du terme de dispersion μ .

Exercice 2.3.9 On définit le schéma "saute-mouton" (leapfrog, en anglais)

$$\frac{u_j^{n+1} - u_j^{n-1}}{2\Delta t} + V \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} = 0.$$

Étudier la consistance et l'erreur de troncature de ce schéma. Montrer par analyse de Fourier qu'il est stable sous la condition CFL $|V|\Delta t \leq M\Delta x$ avec $M < 1$.

Correction.

1. Étude de la consistance

Par développement de Taylor en (t_n, x_j) on a

$$\frac{u(t_{n+1}, x_j) - u(t_{n-1}, x_j)}{2\Delta t} + V \frac{u(t_n, x_{j+1}) - u(t_n, x_{j-1})}{2\Delta x} = \frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} + \frac{(\Delta t)^2}{12} \frac{\partial^3 u}{\partial t^3} + V \frac{(\Delta x)^2}{12} \frac{\partial^3 u}{\partial x^3} + \mathcal{O}((\Delta t)^3 + (\Delta x)^3).$$

Si u est solution de l'équation d'advection, l'erreur de troncature est donc

$$E = \frac{1}{12} ((\Delta x)^2 - (\Delta t)^2 V^3) \frac{\partial^3 u}{\partial x^3}.$$

Ainsi, le schéma saute-mouton est consistant, d'ordre 2 en espace et en temps.

2. Stabilité L^2

Par transformation de Fourier, on obtient

$$\hat{u}^{n+1}(k) = \hat{u}^{n-1}(k) - i2c \sin(2\pi k \Delta x) \hat{u}^n(k). \quad (2.17)$$

où $c = \frac{V\Delta t}{\Delta x}$. On introduit les polynômes (dépendants implicitement de k , Δt et Δx)

$$P(X) = X^2 + i2c \sin(2\pi k \Delta x) X - 1.$$

On note λ_1 et λ_2 les racines de P et $\Delta = 4(1 - c^2 \sin^2(2\pi k \Delta x))$ son discriminant. Les solutions de (2.17) s'expriment explicitement en fonction de \hat{u}^0 et \hat{u}^1 :

$$\hat{u}^n = \begin{cases} \left(\frac{\lambda_2 \lambda_1^n - \lambda_1 \lambda_2^n}{\lambda_2 - \lambda_1} \right) \hat{u}^0 + \left(\frac{\lambda_2^n - \lambda_1^n}{\lambda_2 - \lambda_1} \right) \hat{u}^1 & \text{si } \Delta \neq 0, \\ (1-n)\lambda_1^n \hat{u}^0 + n\lambda_1^{n-1} \hat{u}^1 & \text{si } \Delta = 0. \end{cases}$$

Si $c > 1$, le module de la somme des deux racines est égale à $2c > 2$. Le module de l'une des deux racines est plus grand que un et le schéma est instable. Si $c = 1$, on peut avoir $\Delta = 0$ pour certaines valeurs de k et Δx . Dans ce cas, $\lambda_1 = \lambda_2 = i$ et

$$\hat{u}^n = (n\hat{u}^1 + i(n-1)\hat{u}^0)i^{n-1}.$$

Le schéma est instable.

Considérons le cas où c est majoré par une constante $M < 1$. Dans ce cas, les racines de P sont de même module

$$|\lambda_1| = |\lambda_2| = 1.$$

De plus, $|\lambda_1 - \lambda_2| = \sqrt{\Delta} > 2\sqrt{1 - M^2} > 0$. On déduit de l'expression explicite de \hat{u}^n en fonction de \hat{u}^0 et \hat{u}^1 que

$$|\hat{u}^n| \leq \frac{|\hat{u}^0| + |\hat{u}^1|}{\sqrt{1 - M^2}}.$$

Ainsi, sous la condition CFL $V\Delta t/\Delta x < M < 1$, le schéma saute mouton est stable L^2 .

Exercice 2.3.10 On définit le schéma de Crank-Nicholson

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + V \frac{u_{j+1}^{n+1} - u_{j-1}^{n+1}}{4\Delta x} + V \frac{u_{j+1}^n - u_{j-1}^n}{4\Delta x} = 0.$$

Étudier la consistance et l'erreur de troncature de ce schéma. Montrer par analyse de Fourier qu'il est inconditionnellement stable.

Correction.

1. Consistance

Par développement de Taylor en (t_n, x_n) , on montre que

$$\begin{aligned} \frac{u(t_{n+1}, x_j) - u(t_n, x_j)}{\Delta t} + V \frac{u(t_{n+1}, x_{j+1}) - u(t_{n+1}, x_{j-1})}{4\Delta x} \\ + V \frac{u(t_n, x_{j+1}) - u(t_n, x_{j-1})}{4\Delta x} = \frac{\partial u}{\partial t} + V \frac{\partial u}{\partial x} + \frac{\Delta t}{2} \left(\frac{\partial^2 u}{\partial t^2} + V \frac{\partial^2 u}{\partial t \partial x} \right) \\ + \frac{(\Delta t)^2}{6} \left(\frac{\partial^3 u}{\partial t^3} + \frac{V}{2} \frac{\partial^3 u}{\partial x \partial t^2} \right) \\ + \frac{(\Delta x)^2 V}{6} \frac{\partial^3 u}{\partial x^3} + \mathcal{O}((\Delta t)^3 + (\Delta x)^3). \end{aligned}$$

Ainsi, si u est solution de l'équation d'advection, l'erreur de troncature est

$$E(u) = \frac{V}{12} (2(\Delta x)^2 - V^2(\Delta t)^2) \frac{\partial^3 u}{\partial x^3} + \mathcal{O}((\Delta t)^3 + (\Delta x)^3).$$

Le schéma de Crank-Nicholson est donc d'ordre 2 en temps et en espace.

2. Stabilité L^2

Par transformation de Fourier, on établit que

$$\hat{u}^{n+1} \left(1 + \frac{iV\Delta t}{2\Delta x} \sin(2\pi k\Delta x) \right) = \hat{u}^n \left(1 - \frac{iV\Delta t}{2\Delta x} \sin(2\pi k\Delta x) \right).$$

Ainsi, $|\hat{u}^{n+1}| = |\hat{u}^n|$. Le schéma est donc inconditionnellement stable en norme L^2 .

Exercice 2.3.11 Finir la démonstration du Lemme 2.3.6 en calculant $A(k)^n$, et montrer la stabilité du schéma sous condition CFL grâce à (2.41).

Correction. On utilise l'analyse de Fourier pour obtenir

$$\hat{U}^{n+1}(k) = \begin{pmatrix} \hat{u}^{n+1}(k) \\ \hat{u}^n(k) \end{pmatrix} = \begin{pmatrix} \frac{2-(1-2\theta)\alpha(k)}{1+\theta\alpha(k)} & -1 \\ 1 & 0 \end{pmatrix} \hat{U}^n(k) = A(k)\hat{U}^n(k),$$

où

$$\alpha(k) = 4 \left(\frac{\Delta t}{\Delta x} \right)^2 \sin^2(\pi k\Delta x)$$

Ainsi, $\hat{U}^{n+1}(k) = A(k)^n \hat{U}^0(k)$. Les valeurs propres de la matrice $A(k)$ sont les racines du polynôme

$$\lambda^2 - \frac{2 - (1 - 2\theta)\alpha(k)}{1 + \theta\alpha(k)} \lambda + 1 = 0, \quad (2.18)$$

dont le discriminant est

$$\Delta(k) = -\frac{\alpha(k)}{(1 + \theta\alpha(k))^2} (4 - (1 - 4\theta)\alpha(k)).$$

Si $\alpha(k) = 0$, le polynôme (2.18) possède une racine double $\lambda = 1$. Si $\alpha(k) \neq 0$, d'après la condition CFL, $\Delta(k) < 0$ et le polynôme possède deux racines distinctes complexes, conjuguées l'une de l'autre, de module 1. Considérons le premier cas, c'est à dire $\alpha(k) = 0$. Dans ce cas, il existe p tel que $k = p(N + 1)$. On note v la vitesse initiale et v_j la vitesse discrétisée.

$$\hat{v}(k) = \sum_{j=0}^N v_j e^{i2\pi k j \Delta x} = \sum_{j=0}^N v_j e^{i2\pi p} = \sum_{j=0}^N v_j = 0,$$

d'après l'hypothèse de vitesse moyenne initiale nulle effectuée. Ainsi, $\hat{u}^1(k) = \hat{u}^0(k) + \Delta t \hat{v}(k) = \hat{u}^0(k)$. Or

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

d'où

$$\hat{U}^n(k) = A(k)^n \hat{U}^n(0) = A(k)^n \begin{pmatrix} \hat{u}^0(k) \\ \hat{u}^0(k) \end{pmatrix} = \begin{pmatrix} \hat{u}^0(k) \\ \hat{u}^0(k) \end{pmatrix} = \hat{U}^0(k).$$

On a montré que si $\alpha(k) = 0$, $\hat{u}^n(k) = \hat{u}^0(k)$ pour tout n .

Reste à considérer le cas $\alpha(k) \neq 0$. Dans ce cas, le polynôme (2.18) possède deux racines distinctes λ et $\bar{\lambda}$. La matrice $A(k)$ est diagonalisable. Plus précisément,

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

On en déduit que

$$A(k)^n = \frac{1}{\lambda - \bar{\lambda}} \begin{pmatrix} \lambda & \bar{\lambda} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda^n & 0 \\ 0 & \bar{\lambda}^n \end{pmatrix} \begin{pmatrix} 1 & -\bar{\lambda} \\ -1 & \lambda \end{pmatrix}.$$

On obtient ainsi une expression explicite de $\hat{u}^{n+1}(k)$ en fonction de $\hat{u}^0(k)$ et $\hat{v}(k)$. Plus précisément,

$$\hat{u}^{n+1}(k) = \frac{1}{\lambda - \bar{\lambda}} \left(((\lambda^{n+1} - \bar{\lambda}^{n+1}) - (\lambda^n - \bar{\lambda}^n)) \hat{u}^0(k) - \Delta t (\lambda^{n+1} - \bar{\lambda}^{n+1}) \hat{v}(k) \right),$$

ou encore

$$\hat{u}^{n+1}(k) = \frac{1}{\lambda - \bar{\lambda}} \left((\lambda^n(\lambda - 1) - \bar{\lambda}^n(\bar{\lambda} - 1)) \hat{u}^0(k) - \Delta t (\lambda^{n+1} - \bar{\lambda}^{n+1}) \hat{v}(k) \right).$$

Un calcul explicite de la racine λ du polynôme (2.18) nous donne

$$\lambda = \frac{2 - (1 - 2\theta)\alpha(k) + i\sqrt{-\Delta(k)}}{2(1 + \theta\alpha(k))}.$$

La condition CFL, stipule que $4 - (1 - 4\theta)\alpha(k)$ est minoré par une constante strictement positive. Ainsi, il existe une constante C_1 indépendante de k telle que

$$\left| \frac{\lambda - 1}{\lambda - \bar{\lambda}} \right| = (2(1 + \theta\alpha(k)))^{-1} \left| 1 + i2(1 - 2\theta) \sqrt{\frac{\alpha(k)}{4 - (1 - 4\theta)\alpha(k)}} \right| < C_1. \quad (2.19)$$

D'autre part, en utilisant à nouveau la condition CFL, on établit qu'il existe une constante C_2 , indépendante de k telle que

$$\frac{\Delta t}{|\lambda - \bar{\lambda}|} \leq C_2 \frac{\Delta t}{\sqrt{\alpha(k)}}$$

Or

$$\min_{k: \sin(k\pi\Delta x) \neq 0} \sqrt{\alpha(k)} = 2 \left(\frac{\Delta t}{\Delta x} \right) \sin(\pi\Delta x) > \left(\frac{\Delta t}{\Delta x} \right) \pi\Delta x = \pi\Delta t$$

dès que Δx est assez petit. Ainsi,

$$\frac{\Delta t}{|\lambda - \bar{\lambda}|} \leq \pi^{-1} C_2.$$

De cette dernière estimation, de l'estimation (2.19), de l'expression de $u^{n+1}(k)$ et le module de λ étant égale à 1, on déduit que

$$|\hat{u}^{n+1}(k)| \leq 2C_1 |\hat{u}^0(k)| + \pi^{-1} C_2 |\hat{v}(k)|.$$

Le schéma est donc stable pour la norme L^2 et il existe une constante C telle que

$$\|u^n\|_{L^2} \leq C (\|u^0\|_{L^2} + \|v\|_{L^2})$$

Exercice 2.3.12 On considère le cas limite du Lemme 2.3.6, c'est-à-dire $\Delta t/\Delta x = (1 - 4\theta)^{-1/2}$ avec $0 \leq \theta < 1/4$. Montrer que le θ -schéma centré

$$\begin{aligned} & \frac{u_j^{n+1} - 2u_j^n + u_j^{n-1}}{(\Delta t)^2} + \theta \frac{-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}}{(\Delta x)^2} \\ & + (1 - 2\theta) \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} + \theta \frac{-u_{j-1}^{n-1} + 2u_j^{n-1} - u_{j+1}^{n-1}}{(\Delta x)^2} = 0 \end{aligned} \quad (2.20)$$

est instable dans ce cas en vérifiant que $u_j^n = (-1)^{n+j}(2n - 1)$ est une solution (remarquez qu'il s'agit d'une instabilité "faible" puisque la croissance de u^n est linéaire et non exponentielle).

Correction. Soit $u_j^n = (-1)^{n+j}(2n - 1)$,

$$-u_{j-1}^n + 2u_j^n - u_{j+1}^n = 4(-1)^{n+j}(2n - 1)$$

et

$$u_j^{n+1} - 2u_j^n + u_j^{n-1} = 4(-1)^{n+j+1}(2n - 1).$$

En substituant ces relations dans l'expression du θ schéma et en considérant le cas $(\Delta t/\Delta x)^2 = (1 - 4\theta)^{-1}$, on en déduit que

$$\begin{aligned} & \frac{u_j^{n+1} - 2u_j^n + u_j^{n-1}}{(\Delta t)^2} + \theta \frac{-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}}{(\Delta x)^2} \\ & + (1 - 2\theta) \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} + \theta \frac{-u_{j-1}^{n-1} + 2u_j^{n-1} - u_{j+1}^{n-1}}{(\Delta x)^2} \\ & = 4(\Delta t)^{-2}(2n - 1)(-1)^{n+j}(-1 - \theta(1 - 4\theta)^{-1} \\ & \quad + (1 - 2\theta)(1 - 4\theta)^{-1} - \theta(1 - 4\theta)^{-1}) = 0. \end{aligned}$$

Exercice 2.3.13 Montrer que le θ -schéma centré (2.20) conserve l'énergie discrète, c'est-à-dire que $E^{n+1} = E^1$ pour tout $n \geq 0$, où

$$E^{n+1} = \sum_{j=0}^N \left(\frac{u_j^{n+1} - u_j^n}{\Delta t} \right)^2 + a_{\Delta x}(u^{n+1}, u^n) + \theta a_{\Delta x}(u^{n+1} - u^n, u^{n+1} - u^n)$$

avec

$$a_{\Delta x}(u, v) = \sum_{j=0}^N \left(\frac{u_{j+1} - u_j}{\Delta x} \right) \left(\frac{v_{j+1} - v_j}{\Delta x} \right).$$

Correction. On multiplie (2.20) par $u_j^{n+1} - u_j^{n-1}$ et il vient

$$\begin{aligned} & \frac{1}{(\Delta t)^2} (u_j^{n+1} - u_j^n - (u_j^n - u_j^{n-1})) (u_j^{n+1} - u_j^n + (u_j^n - u_j^{n-1})) \\ & + \frac{1}{(\Delta x)^2} (-u_{j-1}^n + 2u_j^n - u_{j+1}^n) (u_j^{n+1} - u_j^{n-1}) \\ & + \frac{\theta}{(\Delta x)^2} (-(u_{j-1}^{n+1} - u_{j-1}^n) + 2(u_j^{n+1} - u_j^n) - (u_{j+1}^{n+1} - u_{j+1}^n)) (u_j^{n+1} - u_j^{n-1}) \\ & + \frac{\theta}{(\Delta x)^2} (-(u_{j-1}^{n-1} - u_{j-1}^n) + 2(u_j^{n-1} - u_j^n) - (u_{j+1}^{n-1} - u_{j+1}^n)) (u_j^{n+1} - u_j^{n-1}) = 0 \end{aligned}$$

Si on somme par rapport à j , comme

$$\sum_{j=0}^N (-u_{j-1}^n + 2u_j^n - u_{j+1}^n) v_j = \sum_{j=0}^N (u_{j+1}^n - u_j^n) (v_{j+1} - v_j),$$

on en déduit que

$$\begin{aligned} & \sum_{j=0}^N \left(\frac{u_j^{n+1} - u_j^n}{\Delta t} \right)^2 - \sum_{j=0}^N \left(\frac{u_j^n - u_j^{n-1}}{\Delta t} \right)^2 + a_{\Delta x}(u^n, u^{n+1} - u^{n-1}) \\ & + a_{\Delta x}(u^{n+1} - u^n, u^{n+1} - u^n + (u^n - u^{n-1})) \\ & + a_{\Delta x}(u^{n-1} - u^n, u^{n+1} - u^n + (u^n - u^{n-1})) = 0, \end{aligned}$$

c'est à dire

$$\begin{aligned} & \sum_{j=0}^N \left(\frac{u_j^{n+1} - u_j^n}{\Delta t} \right)^2 + a_{\Delta x}(u^n, u^{n+1}) + \theta a_{\Delta x}(u^{n+1} - u^n, u^{n+1} - u^n) \\ &= \sum_{j=0}^N \left(\frac{u_j^n - u_j^{n-1}}{\Delta t} \right)^2 + a_{\Delta x}(u^{n-1}, u^n) + \theta a_{\Delta x}(u^n - u^{n-1}, u^n - u^{n-1}). \end{aligned}$$

Exercice 2.3.14 Montrer que le schéma de Lax-Friedrichs

$$\frac{1}{2\Delta t} \begin{pmatrix} 2v_j^{n+1} - v_{j+1}^n - v_{j-1}^n \\ 2w_j^{n+1} - w_{j+1}^n - w_{j-1}^n \end{pmatrix} - \frac{1}{2\Delta x} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} v_{j+1}^n - v_{j-1}^n \\ w_{j+1}^n - w_{j-1}^n \end{pmatrix} = 0, \quad (2.21)$$

est stable en norme L^2 sous la condition CFL $\Delta t \leq \Delta x$, et qu'il est précis à l'ordre 1 en espace et temps si le rapport $\Delta t/\Delta x$ est gardé constant lorsque Δt et Δx tendent vers zéro.

Correction.

1. Consistance

On pose $U = (v, w)$ et

$$J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

On effectue un développement de Taylor en (t_n, x_j) sur le schéma (2.21) :

$$\begin{aligned} & \frac{1}{2\Delta t} (2U(t_{n+1}, x_j) - U(t_n, x_{j+1}) - U(t_n, x_{j-1})) \\ & \quad - \frac{1}{2\Delta x} J (U(t_n, x_{j+1}) - U(t_n, x_{j-1})) \\ &= \frac{\partial U}{\partial t} - J \frac{\partial U}{\partial x} - \frac{(\Delta x)^2}{2\Delta t} \left(1 - \left(\frac{\Delta t}{\Delta x} \right)^2 \right) \frac{\partial^2 U}{\partial x^2} + \mathcal{O} \left((\Delta x)^2 + \frac{(\Delta x)^4}{\Delta t} \right) \end{aligned}$$

Le schéma est donc consistant et précis à l'ordre 1 si le rapport $\Delta t/\Delta x$ est constant.

2. Stabilité L^2

Étudions la stabilité L^2 . Par transformation de Fourier, on établit que

$$\begin{pmatrix} \hat{v}^{n+1} \\ \hat{w}^{n+1} \end{pmatrix} = \left(\cos(2k\pi\Delta x) + i \sin(2k\pi\Delta x) \frac{\Delta t}{\Delta x} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right) \begin{pmatrix} \hat{v}^n \\ \hat{w}^n \end{pmatrix}.$$

On pose $\alpha = \cos(2k\pi\Delta x)$ et $\beta = \sin(2k\pi\Delta x) \frac{\Delta t}{\Delta x}$. On diagonalise la matrice $A(k)$ et on établit que

$$A(k) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \alpha - i\beta & 0 \\ 0 & \alpha + i\beta \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Notons que

$$\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^2 = I.$$

Ainsi, le schéma est stable L^2 si et seulement si $|\alpha + i\beta| \leq 1$. Or

$$|\alpha + i\beta|^2 = (\cos(2k\pi\Delta x)^2 + \sin(2k\pi\Delta x)^2)(\Delta t/\Delta x)^2.$$

Le schéma est donc stable en norme L^2 sous la condition CFL $\Delta t \leq \Delta x$.

Exercice 2.3.15 Montrer que le schéma de Lax-Wendroff

$$\begin{aligned} & \frac{1}{\Delta t} \begin{pmatrix} v_j^{n+1} - v_j^n \\ w_j^{n+1} - w_j^n \end{pmatrix} - \frac{1}{2\Delta x} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} v_{j+1}^n - v_{j-1}^n \\ w_{j+1}^n - w_{j-1}^n \end{pmatrix} \\ & + \frac{\Delta t}{2(\Delta x)^2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2 \begin{pmatrix} -v_{j-1}^n + 2v_j^n - v_{j+1}^n \\ -w_{j-1}^n + 2w_j^n - w_{j+1}^n \end{pmatrix} = 0 \end{aligned} \quad (2.22)$$

est stable en norme L^2 sous la condition CFL $\Delta t \leq \Delta x$, et qu'il est précis à l'ordre 2 en espace et temps.

Correction.

1. Consistance

On pose $U = (v, w)$ et

$$J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

On effectue un développement de Taylor en (t_n, x_j) sur le schéma (2.22) :

$$\begin{aligned} & \frac{1}{\Delta t} (U(t_{n+1}, x_j) - U(t_n, x_j)) - \frac{1}{2\Delta x} J (U(t_n, x_{j+1}) - U(t_n, x_{j-1})) \\ & + \frac{\Delta t}{2(\Delta x)^2} (-U(t_n, x_j - 1) + 2U(t_n, x_j) - U(t_n, x_{j+1})) = \frac{\partial U}{\partial t} + \frac{\Delta t}{2} \frac{\partial^2 U}{\partial x^2} \\ & + J \frac{(\Delta t)^2}{6} \frac{\partial^3 U}{\partial t^3} - J \frac{\partial U}{\partial x} - \frac{(\Delta x)^2}{6} \frac{\partial^3 U}{\partial x^3} - \frac{\Delta t}{2} \frac{\partial^2 U}{\partial x^2} + \mathcal{O}((\Delta t)^3 + (\Delta x)^3). \end{aligned}$$

Si U est solution de l'équation (2.43), on en déduit que l'erreur de troncature est

$$E(U) = \frac{1}{6} ((\Delta t)^2 - (\Delta x)^2) J \frac{\partial^3 U}{\partial x^3} + \mathcal{O}((\Delta x)^2 + (\Delta t)^2).$$

2. Stabilité L^2 .

Établissons la stabilité L^2 sous la condition CFL $\Delta t \leq \Delta x$. Par transformation de Fourier, on établit que

$$\hat{U}^{n+1}(k) = \left(1 - 2 \left(\frac{\Delta t}{\Delta x} \right)^2 \sin^2(k\pi\Delta x) + i \frac{\Delta t}{\Delta x} \sin(2k\pi\Delta x) J \right) \hat{U}^n(k)$$

On pose $\alpha = 1 - 2 \left(\frac{\Delta t}{\Delta x} \right)^2 \sin^2(k\pi\Delta x)$ et $\beta = \frac{\Delta t}{\Delta x} \sin(2k\pi\Delta x)$ et on procède comme pour l'exercice précédent. Ainsi, le schéma est stable L^2 si et seulement si $|\alpha + i\beta| \leq 1$. Or

$$|\alpha + i\beta|^2 = 1 - 4 \sin^3(k\pi\Delta x) \left(\frac{\Delta t}{\Delta x} \right)^2 \left(1 - \left(\frac{\Delta t}{\Delta x} \right)^2 \right).$$

Ainsi, le schéma est stable L^2 dès que

$$\frac{\Delta t}{\Delta x} \leq 1.$$

Chapitre 3

FORMULATION VARIATIONNELLE DES PROBLÈMES ELLIPTIQUES

Exercice 3.1.1 Si f est une fonction continue sur $[0, 1]$, montrer que l'équation différentielle

$$\begin{cases} -\frac{d^2u}{dx^2} = f & \text{pour } 0 < x < 1 \\ u(0) = u(1) = 0. \end{cases} \quad (3.1)$$

admet une solution unique dans $C^2([0, 1])$ donnée par la formule

$$u(x) = x \int_0^1 f(s)(1-s)ds - \int_0^x f(s)(x-s)ds \text{ pour } x \in [0, 1]. \quad (3.2)$$

Correction. Soit u défini par (3.2). La continuité de la fonction f assure la dérivabilité de la fonction u . On a

$$u'(x) = \int_0^1 f(s)(1-s)ds - \int_0^x f(s)ds,$$

d'où $-u''(x) = f$. De plus, u vérifie les conditions aux limites $u(0) = u(1) = 0$. Ainsi, u est bien solution de l'équation différentielle (3.1). Il reste à établir l'unicité de la solution de l'équation (3.1). L'équation étant linéaire, il suffit de montrer que toute solution v de l'équation (3.1) avec $f = 0$ est nulle. La dérivée seconde de v étant nulle, on en déduit que v est une fonction affine. Enfin, les conditions aux limites impliquent la nullité de la fonction v .

Exercice 3.2.1 Dédurre de la formule de Green (3.5) la formule de Stokes

$$\int_{\Omega} \operatorname{div} \sigma(x) \phi(x) dx = - \int_{\Omega} \sigma(x) \cdot \nabla \phi(x) dx + \int_{\partial \Omega} \sigma(x) \cdot n(x) \phi(x) ds,$$

où ϕ est une fonction scalaire de $C^1(\overline{\Omega})$ et σ une fonction à valeurs vectorielles de $C^1(\overline{\Omega})$, à supports bornés dans le fermé $\overline{\Omega}$.

Correction.

$$\begin{aligned}
 \int_{\Omega} (\nabla \cdot \sigma(x) \phi(x) + \sigma(x) \cdot \nabla \phi(x)) dx &= \sum_{i=1}^n \int_{\Omega} \left(\frac{\partial \sigma_i}{\partial x_i}(x) \phi(x) + \sigma_i(x) \frac{\partial \phi}{\partial x_i}(x) \right) dx \\
 &= \sum_{i=1}^n \int_{\Omega} \frac{\partial \sigma_i \phi}{\partial x_i}(x) dx = \sum_{i=1}^n \int_{\partial \Omega} \sigma_i(x) \phi(x) n_i(x) ds \\
 &= \int_{\partial \Omega} \sigma(x) \cdot n(x) \phi(x) ds.
 \end{aligned}$$

Exercice 3.2.2 En dimension $N = 3$ on définit le rotationnel d'une fonction de Ω dans \mathbb{R}^3 , $\phi = (\phi_1, \phi_2, \phi_3)$, comme la fonction de Ω dans \mathbb{R}^3 définie par

$$\text{rot} \phi = \left(\frac{\partial \phi_3}{\partial x_2} - \frac{\partial \phi_2}{\partial x_3}, \frac{\partial \phi_1}{\partial x_3} - \frac{\partial \phi_3}{\partial x_1}, \frac{\partial \phi_2}{\partial x_1} - \frac{\partial \phi_1}{\partial x_2} \right).$$

Pour ϕ et ψ , fonctions à valeurs vectorielles de $C^1(\overline{\Omega})$, à supports bornés dans le fermé $\overline{\Omega}$, déduire de la formule de Green (3.5)

$$\int_{\Omega} \text{rot} \phi \cdot \psi dx - \int_{\Omega} \phi \cdot \text{rot} \psi dx = - \int_{\partial \Omega} (\phi \times n) \cdot \psi ds.$$

Correction.

$$\begin{aligned}
 &\int_{\Omega} (\text{rot} \phi \cdot \psi - \phi \cdot \text{rot} \psi) dx \\
 &= \int_{\Omega} \left[\left(\frac{\partial \phi_3}{\partial x_2} - \frac{\partial \phi_2}{\partial x_3} \right) \psi_1 + \left(\frac{\partial \phi_1}{\partial x_3} - \frac{\partial \phi_3}{\partial x_1} \right) \psi_2 + \left(\frac{\partial \phi_2}{\partial x_1} - \frac{\partial \phi_1}{\partial x_2} \right) \psi_3 \right. \\
 &\quad \left. - \left(\frac{\partial \psi_3}{\partial x_2} - \frac{\partial \psi_2}{\partial x_3} \right) \phi_1 - \left(\frac{\partial \psi_1}{\partial x_3} - \frac{\partial \psi_3}{\partial x_1} \right) \phi_2 - \left(\frac{\partial \psi_2}{\partial x_1} - \frac{\partial \psi_1}{\partial x_2} \right) \phi_3 \right] dx \\
 &= \int_{\Omega} \frac{\partial}{\partial x_1} (\phi_2 \psi_3 - \phi_3 \psi_2) + \frac{\partial}{\partial x_2} (\phi_3 \psi_1 - \phi_1 \psi_3) + \frac{\partial}{\partial x_3} (\phi_1 \psi_2 - \phi_2 \psi_1) dx \\
 &= \int_{\partial \Omega} \begin{pmatrix} \phi_2 \psi_3 - \phi_3 \psi_2 \\ \phi_3 \psi_1 - \phi_1 \psi_3 \\ \phi_1 \psi_2 - \phi_2 \psi_1 \end{pmatrix} \cdot n ds \\
 &= \int_{\partial \Omega} (\phi \times \psi) \cdot n ds.
 \end{aligned}$$

Exercice 3.2.3 On considère le Laplacien avec condition aux limites de Neumann. Soit u une fonction de $C^2(\overline{\Omega})$. Montrer que u est une solution du problème aux limites

$$\begin{cases} -\Delta u = f & \text{dans } \Omega \\ \frac{\partial u}{\partial n} = 0 & \text{sur } \partial \Omega. \end{cases} \quad (3.3)$$

si et seulement si u appartient à $C^1(\overline{\Omega})$ et vérifie l'égalité

$$\int_{\Omega} \nabla u(x) \cdot \nabla v(x) dx = \int_{\Omega} f(x) v(x) dx \text{ pour toute fonction } v \in C^1(\overline{\Omega}). \quad (3.4)$$

En déduire qu'une condition nécessaire d'existence d'une solution dans $C^2(\overline{\Omega})$ de (3.3) est que $\int_{\Omega} f(x) dx = 0$.

Correction. Supposons que u soit solution du problème aux limites de Neumann (3.3)

$$\begin{cases} -\Delta u = f & \text{dans } \Omega \\ \frac{\partial u}{\partial n} = 0 & \text{sur } \partial\Omega. \end{cases}$$

En multipliant l'équation vérifiée par u par dans Ω par une fonction test $v \in C^1(\overline{\Omega})$, on obtient, suite à une intégration par partie que

$$\int_{\Omega} \nabla u(x) \cdot \nabla v(x) dx - \int_{\partial\Omega} \frac{\partial u}{\partial n}(x) v(x) ds = \int_{\Omega} f(x) v(x) dx.$$

Comme $\partial u / \partial n = 0$ sur $\partial\Omega$, on en déduit que

$$\int_{\Omega} \nabla u(x) \cdot \nabla v(x) dx = \int_{\Omega} f(x) v(x) dx \text{ pour tout } v \in C^1(\overline{\Omega}). \quad (3.5)$$

Réciproquement, supposons que u soit une fonction régulière vérifiant (3.5). Par intégration par partie on a

$$- \int_{\Omega} (\Delta u(x) - f(x)) v(x) dx + \int_{\partial\Omega} \frac{\partial u}{\partial n}(x) v(x) ds = 0 \quad (3.6)$$

pour tout $v \in C^1(\overline{\Omega})$. On procède en deux étapes : Dans un premier temps, on applique la relation (3.6) à des fonctions tests à support compact dans Ω . Cela nous permet de “tester” l'équation vérifiée par u dans Ω et d'établir l'équation $-\Delta u = f$ dans Ω . Dans un deuxième temps, on applique (3.6) à des fonctions tests non nulles sur $\partial\Omega$, ce qui nous permet de “tester” l'équation vérifiée par u sur le bord du domaine et d'en déduire que $\partial u / \partial n = 0$ sur $\partial\Omega$. Plus précisément, pour toute fonction test v à support compact dans Ω ,

$$\int_{\Omega} (\Delta u(x) - f(x)) v(x) dx = 0.$$

On peut conclure à la nullité de $\Delta u - f$ de deux manières différentes. La première consiste à appliquer le Lemme 3.2.9 du cours. Plus simplement, $\Delta u - f$ est nulle car orthogonale à un sous espace dense de $L^2(\Omega)$. L'égalité (3.6) implique alors que $\partial u / \partial n$ est elle nulle car orthogonale (pour le produit scalaire $L^2(\partial\Omega)$) à un sous espace dense de $L^2(\Omega)$, trace des fonctions $C^1(\overline{\Omega})$ sur le bord $\partial\Omega$ du domaine Ω .

En choisissant la fonction $v = 1$ comme fonction test dans la formulation variationnelle, on en déduit que s'il existe une solution u régulière au problème aux limites (3.3),

$$\int_{\Omega} f(x) dx = 0.$$

Exercice 3.2.4 On considère l'équation des plaques

$$\begin{cases} \Delta(\Delta u) = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega \\ \frac{\partial u}{\partial n} = 0 & \text{sur } \partial\Omega \end{cases} \quad (3.7)$$

On note X l'espace des fonctions v de $C^2(\overline{\Omega})$ telles que v et $\frac{\partial v}{\partial n}$ s'annulent sur $\partial\Omega$. Soit u une fonction de $C^4(\overline{\Omega})$. Montrer que u est une solution du problème aux limites (3.7) si et seulement si u appartient à X et vérifie l'égalité

$$\int_{\Omega} \Delta u(x) \Delta v(x) dx = \int_{\Omega} f(x) v(x) dx \text{ pour toute fonction } v \in X. \quad (3.8)$$

Correction. On procède comme pour l'exercice précédent. Soit u une solution régulière de l'équation des plaques (3.7), pour tout $v \in X$,

$$\int_{\Omega} \Delta(\Delta u)(x) v(x) dx = \int_{\Omega} f(x) v(x) dx. \quad (3.9)$$

Par intégration par partie,

$$\int_{\Omega} \Delta(\Delta u)(x) v(x) dx = - \int_{\Omega} \nabla(\Delta u) \nabla v(x) dx + \int_{\partial\Omega} \frac{\partial(\Delta u)}{\partial n}(x) v(x) ds.$$

Comme $v = 0$ sur $\partial\Omega$, on en déduit que

$$\int_{\Omega} \Delta(\Delta u)(x) v(x) dx = - \int_{\Omega} \nabla(\Delta u) \nabla v(x) dx$$

puis par une nouvelle intégration par partie que

$$\int_{\Omega} \Delta(\Delta u)(x) v(x) dx = \int_{\Omega} \Delta u(x) \Delta v(x) dx - \int_{\partial\Omega} \Delta u(x) \frac{\partial v}{\partial n}(x) ds.$$

Comme $\frac{\partial v}{\partial n}(x) = 0$ sur $\partial\Omega$, le dernier terme de cette équation est nulle. Ainsi, on déduit de (3.9) que

$$\int_{\Omega} \Delta u(x) \Delta v(x) dx = \int_{\Omega} f(x) v(x) dx.$$

La réciproque s'établit comme lors de l'exercice précédent. Supposons que u soit une solution du problème variationnel (3.8), en effectuant deux intégrations par partie successives, on obtient

$$\int_{\Omega} (\Delta(\Delta u) - f) v dx = 0,$$

pour tout $v \in X$. Or X est un sous espace dense de $L^2(\Omega)$, ainsi $\Delta(\Delta u) - f = 0$.

Exercice 3.3.1 Le but de cet exercice est de montrer que l'espace V , défini par

$$V = \{v \in C^1(\overline{\Omega}), v = 0 \text{ sur } \partial\Omega\}, \quad (3.10)$$

muni du produit scalaire

$$\langle w, v \rangle = \int_{\Omega} \nabla w(x) \cdot \nabla v(x) dx, \quad (3.11)$$

n'est pas complet. Soit Ω la boule unité ouverte de \mathbb{R}^N . Si $N = 1$, on définit la suite

$$u_n(x) = \begin{cases} -x - 1 & \text{si } -1 < x < -n^{-1}, \\ (n/2)x^2 - 1 + 1/(2n) & \text{si } -n^{-1} \leq x \leq n^{-1}, \\ x - 1 & \text{si } n^{-1} < x < 1. \end{cases}$$

Si $N = 2$, pour $0 < \alpha < 1/2$, on définit la suite

$$u_n(x) = |\log(|x|^2 + n^{-1})|^{\alpha/2} - |\log(1 + n^{-1})|^{\alpha/2}.$$

Si $N \geq 3$, pour $0 < \beta < (N - 2)/2$, on définit la suite

$$u_n(x) = \frac{1}{(|x|^2 + n^{-1})^{\beta/2}} - \frac{1}{(1 + n^{-1})^{\beta/2}}.$$

Montrer que la suite u_n est de Cauchy dans V mais qu'elle ne converge pas dans V lorsque n tend vers l'infini.

Correction.

D'après l'inégalité de Poincaré, une suite u_n de l'espace V est de Cauchy, si et seulement si ∇u_n est une suite de Cauchy dans $L^2(\Omega)^N$.

L'espace $L^2(\Omega)^N$ étant complet, on en déduit que u_n est de Cauchy dans V si et seulement si ∇u_n est convergente dans $L^2(\Omega)^N$.

Ainsi, si V était un espace complet, toute suite de u_n de V telle que ∇u_n converge vers un élément τ de $L^2(\Omega)^N$ serait convergente vers un élément u de V . En particulier, $\tau = \nabla u$ serait une fonction continue. Afin de prouver que V n'est pas complet, il suffit donc dans chacun des cas de vérifier que ∇u_n converge dans $L^2(\Omega)^N$ vers une fonction discontinue.

Cas $N = 1$. La suite ∇u_n converge dans $L^2(-1, 1)$ vers la fonction τ définie par

$$\tau(x) = \begin{cases} -1 & \text{si } x < 0 \\ 1 & \text{si } x > 0 \end{cases}$$

La fonction τ n'ayant pas de représentant continu, V n'est pas complet.

Cas $N = 2$. Soit $\tau : \Omega \rightarrow \mathbb{R}^2$ la fonction définie pour tout $x \neq 0$ par

$$\tau(x) = -\frac{2\alpha x}{|x|^2} (-\log(|x|^2))^{\alpha-1}.$$

On vérifie sans mal que τ appartient à $L^2(\Omega)^2$. En effet,

$$\int_{\Omega} |\tau|^2 dx = \pi 2^{1+\alpha} \alpha \int_0^1 \frac{1}{r \log(r)^{2-2\alpha}} dr.$$

($\int_0^{1/2} \frac{1}{r(\log r)^{2(\alpha-1)}} dr < +\infty$) et que la suite ∇u_n converge dans $L^2(\Omega)^2$ vers τ . Il suffit à cet effet, d'appliquer le théorème de Lebesgue en remarquant que

$$|\nabla u_n - \tau|^2 \leq 2 \max(|\nabla u|, 2/\log(2))^2.$$

Comme τ n'est pas continue, V n'est pas complet.

Cas $N \geq 3$. On procède comme dans le cas précédent. En particulier, u_n et de Cauchy et le gradient de u_n converge dans $L^2(\Omega)^3$ vers

$$\tau = -\beta \frac{x}{|x|^{\beta+2}},$$

La fonction τ appartient bien à L^2 , car $\int_0^{1/2} r^{-2\beta+N-3} dr < +\infty$ dès que $\beta < (N-2)/2$ mais n'est pas continue en 0.

Chapitre 4

ESPACES DE SOBOLEV

Exercice 4.2.1 Soit $\Omega = (0, 1)$. Montrer que la fonction x^α est dérivable au sens faible dans $L^2(\Omega)$ si et seulement si $\alpha > 1/2$.

Correction. Tout d'abord, x^α appartient à $L^2(0, 1)$ si et seulement si $\alpha > -1/2$. Si α est un réel strictement plus grand que $-1/2$, d'après la définition 4.2.3, x^α admet une dérivée faible dans $L^2(0, 1)$ si et seulement si il existe $w \in L^2(0, 1)$ tel que pour tout $\varphi \in C_c^\infty(0, 1)$,

$$\int_0^1 x^\alpha \varphi'(x) dx = - \int_0^1 w(x) \varphi(x) dx.$$

Or comme φ est à support compact, il existe $a > 0$ tel que $\varphi([0, a]) = 0$. Ainsi,

$$\begin{aligned} \int_0^1 x^\alpha \varphi'(x) dx &= \int_a^1 x^\alpha \varphi'(x) dx \\ &= - \int_a^1 \alpha x^{\alpha-1} \varphi(x) dx = - \int_0^1 \alpha x^{\alpha-1} \varphi(x) dx \end{aligned}$$

(Les intégrations par partie sur $(a, 1)$ ne posent aucun problème, x^α étant de classe C^∞ sur cet intervalle). On en déduit que x^α admet une dérivée faible L^2 si et seulement si $\alpha x^{\alpha-1} = w \in L^2(0, 1)$, c'est à dire $\alpha - 1 > -1/2$.

Exercice 4.2.2 Soit Ω un ouvert borné. Montrer qu'une fonction continue sur $\overline{\Omega}$, et C^1 par morceaux est dérivable au sens faible dans $L^2(\Omega)$.

Correction. Soit f une fonction continue sur $\overline{\Omega}$, C^1 par morceaux. Par définition, il existe une famille d'ouverts deux à deux disjoints $(\Omega_i)_{i=1, \dots, N}$ telle que

$$\bigcup_i \overline{\Omega_i} = \overline{\Omega}$$

et $f_i = f|_{\Omega_i}$ soit de classe C^1 . On note $\Gamma_{i,j} = \overline{\Omega_i} \cap \overline{\Omega_j}$ la frontière commune entre deux sous-ouverts de Ω et n^i la normale extérieure à l'ouvert Ω_i . Soit $\varphi \in C_c^\infty(\Omega)$.

En appliquant la formule de Green (3.5) à chacun des ouverts Ω_i , on obtient

$$\begin{aligned} \int_{\Omega} f(x) \frac{\partial \varphi}{\partial x_k}(x) dx &= \sum_i \int_{\Omega_i} f_i(x) \frac{\partial \varphi}{\partial x_k}(x) dx \\ &= \sum_i \int_{\partial \Omega_i} f_i(x) \varphi(x) n_k^i ds - \int_{\Omega_i} \frac{\partial f_i}{\partial x_k} \varphi(x) dx \\ &= - \left(\sum_i \int_{\Omega_i} \frac{\partial f_i}{\partial x_k} \varphi(x) dx \right) + \sum_{\substack{i,j \\ i \neq j}} \int_{\Gamma_{i,j}} f_i(x) \varphi(x) n_k^i ds. \end{aligned}$$

Or pour tout couple (i, j) et tout point $x \in \Gamma_{i,j}$, $n_k^i(x) = -n_k^j(x)$, et comme f est continue, $f_i(x) = f_j(x)$. On en déduit que

$$\sum_{\substack{i,j \\ i \neq j}} \int_{\Gamma_{i,j}} f_i(x) \varphi(x) n_k^i ds = \sum_{\substack{i,j \\ i < j}} \int_{\Gamma_{i,j}} \varphi(x) (f_i(x) n_k^i + f_j(x) n_k^j) ds = 0$$

et

$$\int_{\Omega} f(x) \frac{\partial \varphi}{\partial x_k}(x) dx = - \sum_i \int_{\Omega_i} \frac{\partial f_i}{\partial x_k} \varphi(x) dx = - \int_{\Omega} \psi_k(x) \varphi(x) dx,$$

où $\psi_k : \Omega \rightarrow \mathbb{R}$ est défini pour tout $x \in \Omega_i$ par $\psi_k(x) = \partial f_i / \partial x_k(x)$. Enfin, la fonction ψ_k étant continue par morceaux sur un ouvert borné, elle appartient à $L^2(\Omega)$. Ainsi f admet une dérivée faible L^2 et $\partial f / \partial x_k = \psi_k$.

Exercice 4.2.3 Soit Ω un ouvert borné. Montrer qu'une fonction C^1 par morceaux mais pas continue n'est pas dérivable au sens faible dans $L^2(\Omega)$.

Correction. On utilise les mêmes notations que l'exercice précédent, on a toujours

$$\int_{\Omega} f(x) \frac{\partial \varphi}{\partial x_k}(x) dx = - \left(\sum_i \int_{\Omega_i} \frac{\partial f_i}{\partial x_k} \varphi(x) dx \right) + \sum_{\substack{i,j \\ i < j}} \int_{\Gamma_{i,j}} \varphi(x) (f_i(x) n_k^i + f_j(x) n_k^j) ds$$

d'où

$$\int_{\Omega} f(x) \frac{\partial \varphi}{\partial x_k}(x) dx = - \left(\sum_i \int_{\Omega_i} \frac{\partial f_i}{\partial x_k} \varphi(x) dx \right) + \sum_{\substack{i,j \\ i < j}} \int_{\Gamma_{i,j}} \varphi(x) (f_i(x) - f_j(x)) n_k^i ds$$

Supposons que f soit dérivable au sens faible dans L^2 . Dans ce cas, il existe une fonction $w \in L^2(\Omega)$ telle que pour tout $\varphi \in C_c^\infty(\Omega)$,

$$\sum_{\substack{i,j \\ i < j}} \int_{\Gamma_{i,j}} \varphi(x) (f_i(x) - f_j(x)) n_k^i ds = \int_{\Omega} w(x) \varphi(x) dx.$$

En particulier, pour tout $\varphi \in C_c^\infty(\Omega_i)$, on a

$$\int_{\Omega_i} w(x) \varphi(x) dx = 0.$$

Ainsi, $w = 0$ presque partout sur Ω , car $\Omega \setminus \cup_i \Omega_i$ est de mesure nulle. De plus, pour tout indice k , et tout fonction test φ ,

$$\sum_{\substack{i,j \\ i < j}} \int_{\Gamma_{i,j}} \varphi(x)(f_i(x) - f_j(x))n_k^i ds = 0.$$

On en déduit que pour tout $x \in \cup_{i,j} \Gamma_{i,j}$, $f_i(x) = f_j(x)$, c'est à dire que f est continue.

Exercice 4.2.4 Soit Ω un ouvert borné constitué de deux ouverts Ω_1 et Ω_2 séparés par une surface $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$. Montrer qu'une fonction vectorielle de classe C^1 sur chaque morceau Ω_1 et Ω_2 admet une divergence faible dans $L^2(\Omega)$ si et seulement si sa composante normale est continue à travers la surface Γ .

Correction. Soit σ une fonction de Ω à valeurs vectorielles. On note σ_1, σ_2 les restrictions de σ à Ω_1 et Ω_2 respectivement et n^1, n^2 les normales extérieures à Ω_1 et Ω_2 . Soit $\varphi \in C_c^\infty(\Omega)$, d'après la formule de Stokes (voir Exercice 3.2.1),

$$\begin{aligned} \int_{\Omega} \sigma(x) \cdot \nabla \varphi(x) dx &= \int_{\Omega_1} \sigma_1(x) \cdot \nabla \varphi(x) dx + \int_{\Omega_2} \sigma_2(x) \cdot \nabla \varphi(x) dx \\ &= \int_{\Gamma} \sigma_1(x) \cdot n^1(x) \varphi(x) ds - \int_{\Omega_1} \operatorname{div}(\sigma_1)(x) \varphi(x) dx \\ &\quad + \int_{\Gamma} \sigma_2(x) \cdot n^2(x) \varphi(x) ds - \int_{\Omega_2} \operatorname{div}(\sigma_2)(x) \varphi(x) dx. \end{aligned}$$

Ainsi, si la composante normale de σ est continue à l'interface, σ admet une divergence faible et $\operatorname{div}(\sigma)(x) = \operatorname{div}(\sigma_i)(x)$ pour tout $x \in \Omega_i$ ($i = 1, 2$).

Réciproquement, si σ possède une divergence faible, il existe donc $w \in L^2(\Omega)$ tel que

$$\int_{\Gamma} (\sigma_1 - \sigma_2)(x) \cdot n^1(x) \varphi ds = \int_{\Omega} w(x) \varphi dx,$$

et par un raisonnement similaire à celui effectué dans l'exercice précédent, on en déduit que $(\sigma_1 - \sigma_2) \cdot n^1 = 0$ sur Γ .

Exercice 4.3.1 Montrer que les fonctions continues, C^1 par morceaux et à support borné dans $\overline{\Omega}$, appartiennent à $H^1(\Omega)$.

Correction. D'après l'Exercice 4.2.2 (on utilise les mêmes notations), pour toute fonction $\varphi \in C_c^\infty(\Omega)$,

$$\int_{\Omega} f(x) \frac{\partial \varphi}{\partial x_k}(x) dx = - \int_{\Omega} \psi_k(x) \varphi(x) dx,$$

où $\psi_k(x) = \partial f_i / \partial x_k(x)$ pour tout $x \in \Omega_i$. Le support de f étant borné, ψ_k appartient à $L^2(\Omega)$. Ainsi f admet une dérivée faible dans $L^2(\Omega)$ et appartient à $H^1(\Omega)$.

Exercice 4.3.2 Soit B la boule unité ouverte de \mathbb{R}^N . Si $N = 2$, montrer que la fonction $u(x) = |\log(|x|)|^\alpha$ appartient à $H^1(B)$ pour $0 < \alpha < 1/2$, mais n'est pas bornée au voisinage de l'origine. Si $N \geq 3$, montrer que la fonction $u(x) = |x|^{-\beta}$ appartient à $H^1(B)$ pour $0 < \beta < (N-2)/2$, mais n'est pas bornée au voisinage de l'origine.

Correction.1. Cas $N = 2$ Soit α , $0 < \alpha < 1/2$ et u la fonction définie sur la boule unité de \mathbb{R}^2 par

$$u(x) = |\log(|x|)|^\alpha.$$

Tout d'abord, on vérifie que u est un élément de $L^2(B)$. En effet,

$$\int_B |u|^2 dx = 2\pi \int_0^1 |\log(r)|^{2\alpha} r dr < +\infty.$$

Reste à prouver que u admet une dérivée faible L^2 (u n'est évidemment pas bornée au voisinage de 0). Rappelons que $|x| = \sqrt{x_1^2 + x_2^2}$ est dérivable pour tout $x \neq 0$ et que $\nabla|x| = x/|x|$. Ainsi, la fonction u , en tant que fonction composée de fonctions dérivables, est dérivable au sens classique sur $B \setminus \{0\}$ et $\nabla u = \psi$ où

$$\psi(x) = -\frac{\alpha x}{|x|^2} |\log(|x|)|^{\alpha-1}.$$

De plus, ψ appartient à $L^2(\Omega)^2$. En effet,

$$\int_B |\psi|^2 dx = \int_B \left(\frac{\alpha \log(|x|)^{\alpha-1}}{|x|} \right)^2 dx$$

En passant en coordonnées polaires, on obtient

$$\begin{aligned} \int_B |\psi|^2 dx &= 2\pi\alpha^2 \int_0^1 \frac{\log(r)^{2(\alpha-1)}}{r} dr \\ &= 2\pi\alpha^2 \int_1^\infty s^{2(\alpha-1)} ds. \end{aligned}$$

Cette dernière intégrale est finie, dès que $2(\alpha - 1) < -1$ (ce qui est le cas puisque $\alpha < 1/2$). Pour être tout à fait rigoureux, il reste à prouver que la dérivée au sens classique coïncide avec la définition de la dérivée faible. Soit $\varphi \in C^\infty(\Omega)$, pour tout réel ε tel que $0 < \varepsilon < 1$,

$$\begin{aligned} \int_B u(x) \nabla \varphi(x) dx &= \int_{\varepsilon < |x| < 1} u(x) \nabla \varphi(x) dx + \int_{|x| < \varepsilon} u(x) \nabla \varphi(x) dx \\ &= - \int_{\varepsilon < |x| < 1} \psi(x) \varphi(x) dx + \int_{|x|=\varepsilon} u(x) \varphi(x) ds + \int_{|x| < \varepsilon} u(x) \nabla \varphi(x) dx \\ &= - \int_{\varepsilon < |x| < 1} \psi(x) \varphi(x) dx + |\log(\varepsilon)|^\alpha \int_{|x|=\varepsilon} \varphi(x) ds + \int_{|x| < \varepsilon} u(x) \nabla \varphi(x) dx. \end{aligned}$$

Les deux derniers termes de l'expression tendent vers zéro lorsque ε tend vers zéro. Ainsi,

$$\int_B u(x) \nabla \varphi(x) dx = - \int_B \psi(x) \varphi(x) dx,$$

ce qui achève la démonstration.

2. Cas $N \geq 3$

Soit $0 < \beta < (N - 2)/2$. On pose

$$u(x) = |x|^{-\beta}.$$

Dans un premier temps, on vérifie que u est un élément de $L^2(B)$. Soit S_N la sphère unité,

$$\int_B |u|^2 dx = |S_N| \int_0^1 |r|^{N-1-2\beta} dr < \infty.$$

Pour tout $x \neq 0$, u est dérivable au sens classique et $\nabla u(x) = \psi(x)$ où

$$\psi(x) = -\beta x |x|^{-(\beta+2)}$$

On vérifie que la fonction ψ est un élément de $L^2(B)$ ³. En effet,

$$\begin{aligned} \int_B |\psi|^2 dx &= \beta^2 \int_B |\psi|^{-2(\beta+1)} dx \\ &= \beta^2 |S_N| \int_0^1 r^{N-1} r^{-2(\beta+1)} dr \\ &= \beta^2 |S_N| \int_0^1 r^{-2\beta+N-3} dr. \end{aligned}$$

La dernière intégrale est finie car $-2\beta + N - 3 > -1$. Enfin, en procédant comme dans le cas $N = 2$, on vérifie que les dérivées faible et classique coïncident.

Exercice 4.3.3 Le but de cet exercice est de montrer que le Théorème de trace **4.3.13** n'est pas vrai si l'ouvert Ω n'est pas régulier. Soit l'ouvert $\Omega \subset \mathbb{R}^2$ défini par $0 < x < 1$ et $0 < y < x^r$ avec $r > 2$ (voir la Figure **4.2**). Soit la fonction $v(x) = x^\alpha$. Montrer que $v \in H^1(\Omega)$ si et seulement si $2\alpha + r > 1$, tandis que $v \in L^2(\partial\Omega)$ si et seulement si $2\alpha > -1$. Conclure. (On peut aussi montrer avec ce même exemple que le Théorème **4.3.5** de densité et la Proposition **4.4.2** de prolongement ne sont pas vrais pour un tel ouvert.)

Correction. On a

$$\int_\Omega |v|^2 dx dy = \int_\Omega x^{2\alpha} dx dy = \int_0^1 \left(\int_0^{x^r} x^{2\alpha} dy \right) dx = \int_0^1 x^{2\alpha+r} dx.$$

Ainsi, $v \in L^2(\Omega)$ si et seulement si $2\alpha + r > -1$. De plus, $\partial v / \partial y = 0$ et $\partial v / \partial x = \alpha x^{\alpha-1}$. On en déduit que $v \in H^1(\Omega)$ si et seulement si $2(\alpha - 1) + r > -1$, c'est à dire $2\alpha + r > 1$. D'autre part,

$$\int_{\partial\Omega} |v(x)|^2 ds = 1 + \int_0^1 x^{2\alpha} (1 + r^2 x^{2r-2})^{1/2} dx$$

(l'intégrale du membre de droite provient de l'intégration de $v(x)$ le long de la partie du bord de Ω paramétrée par la fonction $(0, 1) \ni x \mapsto x^r \in \partial\Omega$). Comme $r > 2$,

la fonction $(1 + r^2 x^{2r-2})^{1/2}$ est bornée sur $(0, 1)$ et $v \in L^2(\partial\Omega)$ si et seulement si $2\alpha > -1$. Si r est strictement supérieur à 2, il existe α tel que $1 - r < 2\alpha < -1$. Dans ce cas, $v \in H^1(\Omega)$ et $v|_{\partial\Omega} \notin L^2(\partial\Omega)$. Le Théorème 4.3.13 est mis en défaut dans ce cas. En effet, on introduit la suite croissante de fonctions $v^n \in H^1(\Omega) \cap C(\overline{\Omega})$ définie par

$$v^n(x) = \min(v(x), n).$$

La suite v^n converge vers v dans $H^1(\Omega)$ et $v|_{\partial\Omega}^n$ converge presque partout vers $v|_{\partial\Omega}$. On a alors

$$\lim_n \|v^n\|_{H^1(\Omega)} = \|v\|_{H^1(\Omega)}$$

et

$$\lim_n \|v^n\|_{L^2(\Omega)} = \int_{\partial\Omega} |v|_{\partial\Omega}(x)|^2 ds = +\infty.$$

Quel que soit $K > 0$, pour n assez grand, on a donc

$$\|v^n\|_{L^2(\partial\Omega)} > K \|v^n\|_{H^1(\Omega)},$$

ce qui achève la démonstration.

Remarque 4.3.1 *Pour être tout à fait rigoureux, on ne peut pas conclure directement du fait que $v \in H^1(\Omega)$ et $v|_{\partial\Omega} \notin L^2(\partial\Omega)$ que le théorème de Trace 4.3.13 est erroné. En effet, on pourrait tout à fait imaginer que l'application Trace γ_0 soit prolongeable en une fonction continue de $H^1(\Omega)$ dans $L^2(\partial\Omega)$, mais telle que $\gamma_0(v)$ et $v|_{\partial\Omega}$ ne coïncident pas.*

Exercice 4.3.4 Le but de cet exercice est de montrer qu'il ne peut pas y avoir de notion de trace pour des fonctions de $L^2(\Omega)$, c'est-à-dire qu'il n'existe pas de constante $C > 0$ telle que, pour toute fonction $v \in L^2(\Omega)$, on a

$$\|v|_{\partial\Omega}\|_{L^2(\partial\Omega)} \leq C \|v\|_{L^2(\Omega)}.$$

Pour simplifier, on choisit comme ouvert Ω la boule unité. Construire une suite de fonctions régulières dans $\overline{\Omega}$ égales à 1 sur $\partial\Omega$ et dont la norme dans $L^2(\Omega)$ tend vers zéro. Conclure.

Correction.

Soit T une fonction régulière de $[0; +\infty[$ à valeurs dans \mathbb{R}^+ telle que $T(0) = 1$, $T(s) = 0$ pour $s > 1$ et $0 \leq T(s) \leq 1$ pour tout s . On définit la suite u^n de fonctions de la boule Ω à valeurs dans \mathbb{R} par

$$u^n(x) = T(n(1 - |x|)).$$

Pour tout n , quel que soit $x \in \partial\Omega$, $|u^n(x)| = 1$. D'autre part, la suite $|u^n(x)|$ est majorée par 1 pour tout $x \in \Omega$. Enfin, $u_n(x) = 0$ pour tout x appartenant à la boule de rayon $1 - 1/n$. Ainsi, d'après le théorème de Lebesgue, $\|u^n\|_{L^2(\Omega)} \rightarrow 0$, et quel que soit C , pour n assez grand,

$$\|u^n\|_{L^2(\partial\Omega)} = \|u^0\|_{L^2(\partial\Omega)} > C \|u^n\|_{L^2(\Omega)}.$$

L'opérateur trace définit de $C^0(\overline{\Omega}) \cap L^2(\Omega)$ dans $L^2(\partial\Omega)$ n'est pas continue. A fortiori, il ne peut être prolongé en une application continue de $L^2(\Omega)$ dans $L^2(\partial\Omega)$.

Chapitre 5

ÉTUDE MATHÉMATIQUE DES PROBLÈMES ELLIPTIQUES

Exercice 5.2.1 A l'aide de l'approche variationnelle démontrer l'existence et l'unicité de la solution de

$$\begin{cases} -\Delta u + u = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega \end{cases} \quad (5.1)$$

où Ω est un ouvert quelconque de l'espace \mathbb{R}^N , et $f \in L^2(\Omega)$. Montrer en particulier que l'ajout d'un terme d'ordre zéro au Laplacien permet de ne pas avoir besoin de l'hypothèse que Ω est borné.

Correction.

1er Étape. Recherche de la formulation variationnelle.

On multiplie l'équation vérifiée par u par une fonction test v nulle sur $\partial\Omega$. Par intégration par partie, on obtient que

$$\int_{\Omega} (\nabla u \cdot \nabla v + u(x)v(x)) \, dx = \int_{\Omega} f v \, dx.$$

Afin que cette expression ait un sens, il suffit de choisir u et v dans $H_0^1(\Omega)$. Le problème variationnel associé à l'équation (5.1) consiste donc à déterminer $u \in H_0^1(\Omega)$ tel que

$$a(u, v) = L(v) \text{ pour tout } v \in H_0^1(\Omega),$$

où

$$a(u, v) = \int_{\Omega} (\nabla u \cdot \nabla v + u(x)v(x)) \, dx$$

et

$$L(v) = \int_{\Omega} f v \, dx.$$

2eme Étape. Résolution du problème variationnel.

La continuité de $a(., .)$ et $L(.)$ est évidente de même que la coercivité de la forme bilinéaire $a(., .)$. En effet,

$$a(u, u) = \|u\|_{H^1(\mathbb{R}^2)}^2.$$

Les hypothèses du Théorème de Lax-Milgram sont réunies. Il existe donc une solution unique au problème variationnel. On vérifie enfin en effectuant les mêmes intégrations par partie que lors de la première étape que ∇u est un élément de $H(\text{div})$ et que $-\Delta u + u = f$ en tant qu'éléments de $L^2(\Omega)$ et donc presque partout dans Ω . Enfin, comme $u \in H_0^1(\Omega)$, et que Ω est un ouvert régulier, la trace de u est bien définie et $u = 0$ presque partout sur $\partial\Omega$.

Exercice 5.2.2 Soit Ω un ouvert borné de \mathbb{R}^N . A l'aide de l'approche variationnelle démontrer l'existence et l'unicité de la solution du problème suivant de convection-diffusion

$$\begin{cases} V \cdot \nabla u - \Delta u = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega \end{cases} \quad (5.2)$$

où $f \in L^2(\Omega)$ et V est une fonction régulière à valeurs vectorielles telle que $\text{div} V = 0$ dans Ω .

Correction.

1er Étape. Recherche de la formulation variationnelle.

On multiplie l'équation vérifiée par u par une fonction test v nulle sur $\partial\Omega$. Par intégration par partie, on obtient la formulation variationnelle suivante :
Trouver $u \in H_0^1(\Omega)$ tel que

$$a(u, v) = L(v) \text{ pour tout } v \in H_0^1(\Omega),$$

où

$$a(u, v) = \int_{\Omega} (\nabla u \cdot \nabla v + (V \cdot \nabla u)v) \, dx$$

et

$$L(v) = \int_{\Omega} f v \, dx.$$

2ème Étape. Résolution du problème variationnel.

Afin d'appliquer le Théorème de Lax-Milgram, la seule hypothèse non triviale à vérifier est la coercivité de la forme bilinéaire $a(.,.)$.

$$a(u, u) = \int_{\Omega} (\nabla u \cdot \nabla u + (V \cdot \nabla u)u) \, dx$$

Or

$$\begin{aligned} \int_{\Omega} (V \cdot \nabla u)u \, dx &= \int_{\Omega} (\text{div}(uV)u - \text{div}(V)|u|^2) \, dx \\ &= - \int_{\partial\Omega} (V \cdot \nabla u)u \, dx. \end{aligned}$$

Ainsi,

$$a(u, u) = \|\nabla u\|_{L^2(\Omega)}^2$$

et la coercivité de $a(.,.)$ se déduit de l'inégalité de Poincaré.

3ème Étape. Équivalence avec l'équation.

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} (fv - (V \cdot \nabla u)v) \, dx.$$

Ainsi,

$$\left| \int_{\Omega} \nabla u \cdot \nabla v \, dx \right| \leq (\|f\|_{L^2(\Omega)} + \|V\|_{L^\infty(\Omega)} \|u\|_{H^1(\Omega)}) \|v\|_{L^2(\Omega)},$$

et ∇u est un élément de $H(\text{div})$. On en déduit donc par intégration par partie que

$$-\Delta u + V \cdot \nabla u = f \text{ en tant qu'éléments de } L^2(\Omega).$$

Exercice 5.2.3 On reprend les notations et hypothèses de l'Exercice 5.2.2. Montrer que tout $v \in H_0^1(\Omega)$ vérifie

$$\int_{\Omega} v V \cdot \nabla v \, dx = 0.$$

Montrer que la solution de la formulation variationnelle du problème de convection diffusion ne minimise pas dans $H_0^1(\Omega)$ l'énergie

$$J(v) = \frac{1}{2} \int_{\Omega} (|\nabla v|^2 + v V \cdot \nabla v) \, dx - \int_{\Omega} f v \, dx.$$

Correction. On a d'ores et déjà prouvé dans l'exercice précédent que

$$\int_{\Omega} v V \cdot \nabla v \, dx = 0$$

pour tout $v \in H_0^1(\Omega)$. Ainsi,

$$\begin{aligned} J(v) &= 1/2 \int_{\Omega} (|\nabla v|^2 + v(V \cdot \nabla v)) \, dx - \int_{\Omega} f v \, dx \\ &= 1/2 \int_{\Omega} |\nabla v|^2 \, dx - \int_{\Omega} f v \, dx \end{aligned}$$

Or le minimiseur u sur $H_0^1(\Omega)$ de J est solution du problème aux limites

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

et n'a donc aucune raison (sauf cas exceptionnel) d'être solution du problème aux limites

$$\begin{cases} V \cdot \nabla u - \Delta u = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega, \end{cases}$$

Exercice 5.2.4 On considère à nouveau le problème aux limites

$$\begin{cases} -\Delta u = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega \end{cases} \quad (5.3)$$

où Ω est un ouvert borné de l'espace \mathbb{R}^N , et f est un second membre qui appartient à l'espace $L^2(\Omega)$. On suppose que l'ouvert Ω est symétrique par rapport à l'hyperplan $x_N = 0$ de même que la donnée f (i.e. $f(x', x_N) = f(x', -x_N)$). Montrer que la solution de (5.3) a la même symétrie. Montrer que (5.3) est équivalent à un problème aux limites posé sur $\Omega^+ = \Omega \cap \{x_N > 0\}$ avec une condition aux limites de Neumann sur $\Omega \cap \{x_N = 0\}$.

Correction. Soit u la solution de (5.3). On définit

$$v \in H^1(\Omega) \text{ par } v(x', x_n) = u(x', -x_n).$$

On a alors pour tout $(x', x_n) \in \Omega$,

$$-\Delta v(x', x_n) = -\Delta u(x', -x_n) = f(x', -x_n) = f(x', x_n).$$

De plus, pour tout élément (x', x_n) du bord de Ω , $(x', -x_n) \in \partial\Omega$ et

$$v(x', x_n) = u(x', -x_n) = 0.$$

Ainsi, v est également solution du problèmes aux limites (5.3). Comme u est l'unique solution de ce système, $u = v$ et $u(x', x_n) = u(x', -x_n)$. On note $\Gamma_N = \Omega \cap \{x_n = 0\}$ et n la normale extérieure à Ω^+ . Montrons que $\partial u / \partial n = 0$ sur Γ_N . Si on suppose que u est régulier, la nullité de la dérivée normale sur Γ_N découle directement de la relation $u(x', x_n) = u(x', -x_n)$. Sans hypothèse de régularité sur u , on peut utiliser la définition faible de la trace normale d'éléments de $H(\text{div})$. Soit $\varphi \in C_c^\infty(\Omega)$. On pose $\psi(x', x_n) = (\varphi(x', -x_n) + \varphi(x', x_n))/2$. On a

$$\begin{aligned} \left\langle \frac{\partial u}{\partial n}, \psi \right\rangle_{H^{-1/2}(\Gamma_N), H^{1/2}(\Gamma_N)} &= \int_{\Omega^+} \Delta u \psi \, dx + \int_{\Omega^+} \nabla u \cdot \nabla \psi \, dx \\ &= \int_{\Omega^+} f \psi \, dx + \int_{\Omega^+} \nabla u \cdot \nabla \psi \, dx. \end{aligned}$$

De même,

$$\begin{aligned} \left\langle \frac{\partial u}{\partial n}, \psi \right\rangle_{H^{-1/2}(\Gamma_N), H^{1/2}(\Gamma_N)} &= - \int_{\Omega^-} f \psi \, dx - \int_{\Omega^-} \nabla u \cdot \nabla \psi \, dx \\ &= - \int_{\Omega^+} f \psi \, dx - \int_{\Omega^+} \nabla u \cdot \nabla \psi \, dx \end{aligned}$$

(par changement de variable $(x', x_n) \rightarrow (x', -x_n)$).

Or $\psi|_{\Gamma_N} = \varphi|_{\Gamma_N}$, ainsi,

$$\left\langle \frac{\partial u}{\partial n}, \varphi \right\rangle_{H^{-1/2}(\Gamma_N), H^{1/2}(\Gamma_N)} = - \left\langle \frac{\partial u}{\partial n}, \varphi \right\rangle_{H^{-1/2}(\Gamma_N), H^{1/2}(\Gamma_N)}$$

et $\partial u / \partial n = 0$ sur $\Omega \cap \{x_n = 0\}$. Ainsi, u est également solution du problème aux limites

$$\begin{cases} -\Delta u = f \text{ dans } \Omega^+ \\ u = 0 \text{ sur } \partial\Omega \cap \{x_n > 0\} \\ \frac{\partial u}{\partial n} = 0 \text{ sur } \Omega \cap \{x_n = 0\}. \end{cases}$$

Exercice 5.2.5 Démontrer que l'unique solution $u \in H^1(\Omega)$ de la formulation variationnelle

$$\int_{\Omega} (\nabla u \cdot \nabla v + uv) \, dx = \int_{\partial\Omega} gv \, ds + \int_{\Omega} fv \, dx \quad \forall v \in H^1(\Omega). \quad (5.4)$$

vérifie l'estimation d'énergie suivante

$$\|u\|_{H^1(\Omega)} \leq C (\|f\|_{L^2(\Omega)} + \|g\|_{L^2(\partial\Omega)}),$$

où $C > 0$ est une constante qui ne dépend pas de u , f et g .

Correction. Il suffit d'appliquer la formulation variationnelle (5.4) à la fonction test $v = u$. On en déduit que

$$\|u\|_{H^1(\Omega)}^2 = \int_{\Omega} (|\nabla u|^2 + |u|^2) \, dx = \int_{\partial\Omega} gu \, ds + \int_{\Omega} fu \, dx.$$

En appliquant l'inégalité de Cauchy-Schwarz au deuxième membre,

$$\|u\|_{H^1(\Omega)}^2 \leq \|g\|_{L^2(\partial\Omega)} \|u\|_{L^2(\partial\Omega)} + \|f\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)}.$$

Par le Théorème de Trace, il existe donc une constante positive C telle que

$$\|u\|_{H^1(\Omega)}^2 \leq C (\|g\|_{L^2(\partial\Omega)} + \|f\|_{L^2(\Omega)}) \|u\|_{H^1(\Omega)}$$

et

$$\|u\|_{H^1(\Omega)} \leq C (\|g\|_{L^2(\partial\Omega)} + \|f\|_{L^2(\Omega)})$$

Exercice 5.2.6 On suppose que Ω est un ouvert borné régulier de classe C^1 . A l'aide de l'approche variationnelle démontrer l'existence et l'unicité de la solution du Laplacien avec une condition aux limites de Fourier

$$\begin{cases} -\Delta u = f & \text{dans } \Omega \\ \frac{\partial u}{\partial n} + u = g & \text{sur } \partial\Omega \end{cases} \quad (5.5)$$

où $f \in L^2(\Omega)$ et g est la trace sur $\partial\Omega$ d'une fonction de $H^1(\Omega)$. On démontrera l'inégalité suivante (qui généralise celle de Poincaré)

$$\|v\|_{L^2(\Omega)} \leq C (\|v\|_{L^2(\partial\Omega)} + \|\nabla v\|_{L^2(\Omega)}) \quad \forall v \in H^1(\Omega).$$

Correction.

1er Étape. Recherche de la formulation variationnelle.

On multiplie l'équation vérifiée par u par une fonction test v . Par intégration par partie, on obtient

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx - \int_{\partial\Omega} \frac{\partial u}{\partial n} v \, ds = \int_{\Omega} f v \, dx.$$

Enfin, comme $\partial u / \partial n = g - u$ sur $\partial\Omega$, on en déduit que

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx - \int_{\partial\Omega} (g - u) v \, ds = \int_{\Omega} f v \, dx.$$

La formulation variationnelle retenue consiste donc à trouver $u \in H^1(\Omega)$ tel que

$$a(u, v) = L(v) \text{ pour tout } v \in H^1(\Omega),$$

où

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx + \int_{\partial\Omega} uv \, ds$$

et

$$L(v) = \int_{\Omega} f v \, dx + \int_{\partial\Omega} g v \, ds.$$

2ème Étape. Résolution du problème variationnel.

Afin d'appliquer le théorème de Lax-Milgram, la seule hypothèse non triviale à vérifier est la coercivité de la forme bilinéaire $a(., .)$. A cet effet, on va montrer qu'il existe une constante C telle que pour tout $v \in H^1(\Omega)$,

$$\|v\|_{L^2(\Omega)} \leq C(\|v\|_{L^2(\partial\Omega)} + \|\nabla v\|_{L^2(\Omega)}).$$

La coercivité est alors évidente. Afin d'établir ce dernier résultat, on raisonne par contradiction : Supposons que pour tout n , il existe v_n tel que

$$\|v_n\|_{L^2(\Omega)} > n(\|v_n\|_{L^2(\partial\Omega)} + \|\nabla v_n\|_{L^2(\Omega)}).$$

Quitte à considérer la suite $v_n/\|v_n\|_{L^2(\Omega)}$ au lieu de v_n , on peut supposer que pour tout n , $\|v_n\|_{L^2(\Omega)} = 1$. Ainsi, la suite v_n est bornée dans $H^1(\Omega)$ et d'après le théorème de Rellich, il existe une sous suite $v_{n'}$ convergente dans $L^2(\Omega)$ vers un élément v de $H^1(\Omega)$. De plus, $\nabla v_{n'}$ converge vers zéro dans $L^2(\Omega)$. Ainsi, $v_{n'}$ est une suite de Cauchy de $H^1(\Omega)$, v appartient à $H^1(\Omega)$ et $\nabla v = 0$. D'après la Proposition 4.2.5, on en déduit que v est une constante. L'application trace étant continue de $H^1(\Omega)$ dans $L^2(\partial\Omega)$, la trace de v sur le bord de Ω est égale à la limite des traces de $v_{n'}$ sur le bord de Ω . Or $\lim_n \|v_{n'}\|_{L^2(\partial\Omega)} = 0$, ainsi $v = 0$ sur $\partial\Omega$. Finalement, v étant constante, $v = 0$ dans tout Ω , ce qui contredit le fait que $\|v\|_{L^2(\Omega)} = 1$.

3ème Étape. Équivalence avec le problème aux limites.

Tout d'abord, on établit en appliquant la formulation variationnelle à des éléments $v \in C_c^\infty(\Omega)$ que ∇u est un élément de $H(\text{div})$ et par intégration par partie que

$$-\Delta u = f \text{ dans } \Omega.$$

De plus, pour toute fonction $v \in H^1(\Omega)$,

$$\begin{aligned} \int_{\partial\Omega} \left(\frac{\partial u}{\partial n} + u \right) v \, dx &= \int_{\Omega} ((\Delta u)v + \nabla u \cdot \nabla v + uv) \, dx \\ &= \int_{\Omega} (-fv + \nabla u \cdot \nabla v + uv) \, dx = \int_{\partial\Omega} g v \, ds. \end{aligned}$$

On en déduit en particulier que $\partial u / \partial n$ est un élément de $L^2(\partial\Omega)$ et que

$$\frac{\partial u}{\partial n} + u = g \text{ presque partout sur } \partial\Omega.$$

Remarque 5.2.1 En toute rigueur, l'intégrale $\int_{\partial\Omega} (\frac{\partial u}{\partial n} + u)v \, dx$ n'est a priori pas correctement définie. Cependant, comme ∇u est un élément de $H(\text{div})$, il admet une trace normale sur $\partial\Omega$. Ainsi, le calcul précédent reste valable en toute généralité quitte à remplacer l'intégrale de bord par le crochet de dualité $\langle \frac{\partial u}{\partial n} + u, v \rangle_{H^{-1/2}, H^{1/2}}$. Enfin, comme on prouve finalement que $\partial u / \partial n$ appartient à $L^2(\partial\Omega)$, l'utilisation de l'intégrale $\int_{\partial\Omega} (\frac{\partial u}{\partial n} + u)v \, dx$ est justifiée a posteriori.

Exercice 5.2.7 On suppose que Ω est un ouvert borné connexe. A l'aide de l'approche variationnelle démontrer l'existence et l'unicité de la solution du Laplacien avec des conditions aux limites mêlées

$$\begin{cases} -\Delta u = f & \text{dans } \Omega \\ \frac{\partial u}{\partial n} = 0 & \text{sur } \partial\Omega_N \\ u = 0 & \text{sur } \partial\Omega_D \end{cases} \quad (5.6)$$

où $f \in L^2(\Omega)$, et $(\partial\Omega_N, \partial\Omega_D)$ est une partition de $\partial\Omega$ telle que les mesures superficielles de $\partial\Omega_N$ et $\partial\Omega_D$ sont non nulles (voir la Figure 4.1). (Utiliser la Remarque 4.3.18.)

Correction.

La formulation variationnelle s'établit naturellement : Il s'agit de trouver $u \in V$ tel que

$$a(u, v) = L(v) \text{ pour tout } v \in V$$

où

$$V = \{v \in H^1(\Omega) : v = 0 \text{ sur } \partial\Omega_D\}$$

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx$$

et

$$L(v) = \int_{\Omega} f v \, dx.$$

L'application trace étant continue, l'espace vectoriel V , image réciproque d'un fermé par une application continue, est un sous espace fermé de $H^1(\Omega)$. Ainsi, V est un espace de Hilbert. Les formes bilinéaire et linéaire a et L étant continues, il ne reste plus qu'à établir la coercivité de la forme bilinéaire a pour pouvoir appliquer le Théorème de Lax-Milgram et en déduire l'existence et l'unicité d'une solution au problème variationnel. Il s'agit donc d'établir l'inégalité de type Poincaré suivante : Il existe $C > 0$ tel que pour tout $v \in V$,

$$\|u\|_{L^2(\Omega)} \leq C \|\nabla u\|_{L^2(\Omega)}.$$

Cette inégalité s'établit par contradiction (voir la deuxième démonstration de l'inégalité de Poincaré 4.3.10). Supposons que cette inégalité soit fausse pour toute constante C . Dans ce cas, pour tout entier n , il existe $u_n \in V$ tel que

$$\|u_n\|_{L^2(\Omega)} > n \|\nabla u_n\|_{L^2(\Omega)}.$$

Quitte à diviser u_n par sa norme L^2 , on peut supposer que $\|u_n\|_{L^2} = 1$. Ainsi, u_n est borné dans $H^1(\Omega)$ et d'après le Théorème de Rellich, il existe une sous-suite $u_{n'}$

de u_n et un élément u de $L^2(\Omega)$ tels que $u_{n'}$ converge vers u en norme L^2 . Or ∇u_n converge vers zéro. On en déduit que u_n est de Cauchy dans $H^1(\Omega)$. En particulier, u appartient à $H^1(\Omega)$ et le gradient de u est égal à la limite des gradients de $u_{n'}$, c'est à dire $\nabla u = 0$. D'après la Proposition 4.2.5, on en déduit que u est une constante. Comme u appartient à V , la restriction de u à $\partial\Omega_D$ est nulle. La mesure superficielle de $\partial\Omega_D$ étant non nulle, on en déduit que $u = 0$, ce qui contredit le fait que $\|u\|_{L^2(\Omega)} = \lim_{n'} \|u_{n'}\|_{L^2(\Omega)} = 1$.

Enfin, si u est une solution du problème variationnel, on en déduit que ∇u appartient à $H(\text{div})$ et que $-\Delta u = f$ en tant qu'éléments de $L^2(\Omega)$. Enfin, pour tout élément v de V , on a

$$\left\langle \frac{\partial u}{\partial n}, v \right\rangle_{H^{-1/2}, H^{1/2}} = \int_{\Omega} \Delta u v + \nabla u \cdot \nabla v \, dx = 0.$$

Quitte à supposer Ω et $\partial\Omega_N$ assez réguliers, la trace des fonctions V sur le bord est égal à l'ensemble des fonctions de $H^{1/2}(\Omega)$ de support inclus dans $\partial\Omega_N$. Ainsi, la restriction de $\partial u / \partial n$ à $\partial\Omega_N$ est nulle. Enfin, $u = 0$ presque partout sur $\partial\Omega_D$ car $u \in V$. Ainsi, la solution u du problème variationnel est bien solution du problème aux limites initial.

Exercice 5.2.8 Démontrer l'inégalité de Poincaré-Wirtinger : si Ω est borné, régulier et connexe, il existe une constante $C > 0$ telle que, pour tout $v \in H^1(\Omega)$,

$$\|v - m(v)\|_{L^2(\Omega)} \leq C \|\nabla v\|_{L^2(\Omega)} \text{ avec } m(v) = \frac{\int_{\Omega} v \, dx}{\int_{\Omega} dx}. \quad (5.7)$$

Correction. On peut démontrer cette inégalité par contradiction. On suppose que l'inégalité de Poincaré Wirtinger est fausse. Dans ce cas, pour tout entier naturel $n \geq 1$, il existe un élément u_n de $H^1(\Omega)$ tel que

$$\|u_n - m(u_n)\|_{L^2(\Omega)} > n \|\nabla u_n\|_{L^2(\Omega)},$$

où m est la moyenne définie par

$$m(u_n) = \int_{\Omega} u_n \, dx / \int_{\Omega} dx.$$

On pose $v_n = (u_n - m(u_n)) / \|u_n - m(u_n)\|_{L^2(\Omega)}$. La suite v_n vérifie l'équation

$$1 = \|v_n\|_{L^2(\Omega)} > n \|\nabla v_n\|_{L^2(\Omega)}. \quad (5.8)$$

Ainsi, la suite v_n est bornée dans $H^1(\Omega)$. Comme Ω est borné régulier, d'après le Théorème de Rellich, on peut extraire de v_n une sous-suite convergente dans $L^2(\Omega)$ vers un élément v de $L^2(\Omega)$. Par commodité, on note de nouveau v_n cette suite. Comme v_n est convergente dans $L^2(\Omega)$, c'est une suite de Cauchy de $L^2(\Omega)$. De plus, d'après l'équation (5.8), ∇v_n converge vers 0 dans $L^2(\Omega)$. Ainsi, v_n est une suite de Cauchy dans $H^1(\Omega)$. Comme $H^1(\Omega)$ est un espace de Hilbert, il est complet : Toute suite de Cauchy est convergente et v_n converge dans $H^1(\Omega)$. Ainsi,

$$\|\nabla v\|_{L^2(\Omega)} = \lim_n \|\nabla v_n\|_{L^2(\Omega)} \leq \lim_n (1/n) = 0,$$

$$m(v) = \lim_n m(v_n) = 0,$$

$$\|v\|_{L^2(\Omega)} = \lim_n \|v_n\|_{L^2(\Omega)} = 1.$$

Comme $\nabla v = 0$, $m(v) = 0$ et Ω est connexe, v est une constante de moyenne nulle d'après la Proposition 4.2.5. Ainsi, $v = 0$ et $\|v\|_{L^2(\Omega)} = 1$, ce qui est absurde.

Exercice 5.2.9 On suppose que Ω est un ouvert borné connexe régulier. Soit $f \in L^2(\Omega)$. On considère la formulation variationnelle suivante : trouver $u \in H^1(\Omega)$ tel que

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx + \left(\int_{\Omega} u \, dx \right) \left(\int_{\Omega} v \, dx \right) = \int_{\Omega} f v \, dx \quad \forall v \in H^1(\Omega).$$

Démontrer l'existence et l'unicité de la solution de cette formulation variationnelle. Quel problème aux limites a-t-on ainsi résolu ? En particulier, si on suppose que $\int_{\Omega} f \, dx = 0$, quel problème déjà étudié retrouve-t-on ?

Correction.

1. Existence

Soit

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx + \left(\int_{\Omega} u \, dx \right) \left(\int_{\Omega} v \, dx \right) \quad (5.9)$$

et

$$L(v) = \int_{\Omega} f(x)v(x) \, dx.$$

Le problème variationnel posé consiste à déterminer $u \in H^1(\Omega)$ tel que

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

Afin d'appliquer le Théorème de Lax-Milgram, la seule hypothèse non triviale à vérifier porte sur la coercivité de la forme bilinéaire $a(., .)$. En raisonnant par l'absurde (comme lors de la deuxième démonstration de l'inégalité de Poincaré 4.3.10), on établit qu'il existe $C > 0$ tel que pour tout $u \in H^1(\Omega)$,

$$\|u\|_{H^1(\Omega)}^2 \leq C a(u, u)$$

(On utilise ici le fait que Ω est borné connexe). Le Théorème de Lax-Milgram nous assure alors l'existence et l'unicité de la solution de (5.9).

2. Détermination du problème aux limites

Soit $\varphi \in C_c^\infty(\Omega)$,

$$\int_{\Omega} \nabla u \cdot \nabla \varphi(x) \, dx = - \left(\int_{\Omega} u(x) \, dx \right) \left(\int_{\Omega} \varphi(x) \, dx \right) + \int_{\Omega} f(x)\varphi(x) \, dx.$$

D'après l'inégalité de Cauchy-Schwarz,

$$\left| \int_{\Omega} \nabla u \cdot \nabla \varphi(x) \, dx \right| \leq \left(|\Omega|^{1/2} \left| \int_{\Omega} u \, dx \right| + \|f\|_{L^2(\Omega)} \right) \|\varphi\|_{L^2(\Omega)}.$$

Ainsi, $\nabla u \in H(\text{div})$ et

$$-\text{div}(\nabla u) = f - \int_{\Omega} u \, dx \text{ dans } \Omega.$$

De plus, en appliquant la formulation variationnelle à $v = 1$, on obtient que

$$\int_{\Omega} u \, dx = \frac{1}{|\Omega|} \int_{\Omega} f \, dx.$$

Enfin, comme $\nabla u \in H(\text{div})$, la trace de $\partial u / \partial n$ sur la frontière de Ω est correctement définie et on établit aisément que $\partial u / \partial n = 0$. Le problème aux limites résolu consiste donc à trouver $u \in H^1(\Omega)$ tel que

$$\begin{cases} -\Delta u = f - |\Omega|^{-1} \int_{\Omega} f \, dx & \text{dans } \Omega \\ \frac{\partial u}{\partial n} = 0 & \text{sur } \partial\Omega. \end{cases}$$

Dans le cas particulier $\int_{\Omega} f \, dx = 0$, u est solution du problème de Neumann (5.25).

Exercice 5.2.10 Soit Ω un ouvert borné et K un compact connexe de \mathbb{R}^N inclus dans Ω (on suppose que $\Omega \setminus K$ est régulier). Soit $f \in L^2(\Omega)$. On considère un problème de conduction dans Ω où K est une inclusion parfaitement conductrice, c'est-à-dire que l'inconnue u (la température ou le potentiel électrique, par exemple) est constante dans K (cette constante est aussi inconnue). On suppose qu'il n'y a pas de terme source dans K . Ce problème se modélise par

$$\begin{cases} -\Delta u = f & \text{dans } \Omega \setminus K \\ u = C & \text{sur } \partial K \\ \int_{\partial K} \frac{\partial u}{\partial n} ds = 0 & \text{sur } \partial K \\ u = 0 & \text{sur } \partial\Omega, \end{cases}$$

où C est une constante inconnue à déterminer. Trouver une formulation variationnelle de ce problème aux limites et démontrer l'existence et l'unicité d'une solution (u, C) .

Correction. On introduit l'espace vectoriel

$$X = \{u \in H^1(\Omega \setminus K) : u = 0 \text{ sur } \partial\Omega ; v = \text{constante sur } \partial K\}.$$

muni de la norme de $H^1(\Omega \setminus K)$. Notons que X est un espace de Hilbert. En effet, c'est un sous espace fermé de $H^1(\Omega \setminus K)$.

1ere Étape. Détermination de la formulation variationnelle.

On multiplie l'équation vérifiée par u sur $\Omega \setminus K$ par un élément v de X . Par intégration par partie, on en déduit que

$$\int_{\Omega \setminus K} \nabla u \cdot \nabla v(x) \, dx + \int_{\partial K} \frac{\partial u}{\partial n}(x) v(x) \, ds = \int_{\Omega \setminus K} f(x) v(x) \, dx \quad (5.10)$$

Comme $v(x)$ est constante sur ∂K , on a

$$\int_{\partial K} \frac{\partial u}{\partial n}(x) v(x) ds = \left(\int_{\partial K} \frac{\partial u}{\partial n} ds \right) v(\partial K).$$

Enfin, d'après l'équation vérifiée par $\partial u / \partial n$ sur ∂K ,

$$\int_{\partial K} \frac{\partial u}{\partial n}(x) v(x) ds = 0.$$

L'équation (5.10) vérifiée par u se simplifie en

$$\int_{\Omega \setminus K} \nabla u \cdot \nabla v(x) dx = \int_{\Omega \setminus K} f(x) v(x) dx.$$

La formulation variationnelle associée au problème aux limites consiste à trouver $u \in X$ tel que

$$a(u, v) = L(v), \quad (5.11)$$

où $a(., .)$ est la forme bilinéaire définie sur X par

$$a(u, v) = \int_{\Omega \setminus K} \nabla u \cdot \nabla v(x) dx$$

et $L(.)$ la forme linéaire

$$L(v) = \int_{\Omega \setminus K} f(x) v(x) dx.$$

2eme Étape. Existence de solution.

L'application du Théorème de Lax-Milgram est triviale et nous assure l'existence et l'unicité au problème variationnel (5.11).

3eme Étape. Équivalence avec le problème aux limites.

On applique dans un premier temps la formulation variationnelle à une fonction $v \in C_c^\infty(\Omega \setminus K)$. On en déduit que $\nabla u \in H(\text{div})$ et que

$$-\Delta u = f \text{ pour presque tout } x \in \Omega \setminus K.$$

Comme $\nabla u \in H(\text{div})$, $\partial u / \partial n$ admet une trace (au moins au sens faible sur ∂K). En appliquant la formulation variationnelle à un élément quelconque v de X , on en déduit que

$$\int_{\partial K} \frac{\partial u}{\partial n}(x) dx = 0 \text{ sur } \partial K.$$

Enfin, les conditions de type Dirichlet $u = 0$ sur $\partial \Omega$ et $u = \text{constante}$ sur ∂K ont été incluses dans la définition de l'espace X auquel appartient u .

Exercice 5.2.11 Montrer que si $u_1 \in H^1(\Omega_1)$ et $u_2 \in H^1(\Omega_2)$ sont solutions de (5.33) avec (5.34) et $u_i = 0$ sur $\partial \Omega$, pour $i = 1, 2$, alors la fonction u définie comme u_i dans Ω_i , $i = 1, 2$, est l'unique solution dans $H_0^1(\Omega)$ de (5.39).

Correction.

Tout d'abord, la fonction ainsi définie est bien un élément de H_0^1 . En effet, u est continue et les restrictions de u à Ω_1 et Ω_2 appartiennent respectivement à $H^1(\Omega_1)$ et $H^1(\Omega_2)$. D'après le Lemme 4.3.19, u est donc un élément de $H^1(\Omega)$. Enfin, $u = 0$ sur $\partial\Omega$. Notons que pour presque tout $x \in \Omega$,

$$\nabla u(x) = \begin{cases} \nabla u_1(x) & \text{si } x \in \Omega_1 \\ \nabla u_2(x) & \text{si } x \in \Omega_2. \end{cases}$$

Soit v un élément de $H_0^1(\Omega)$. On introduit v_1 et v_2 les restrictions de v à Ω_1 et Ω_2 . On note n la normale extérieure à Ω_1 .

$$\int_{\Omega_1} k_1 \nabla u_1 \cdot \nabla v_1 \, dx = \int_{\Omega_1} f v_1 \, dx + \int_{\Gamma} k_1 \nabla u_1 \cdot n \, dx$$

et

$$\int_{\Omega_2} k_2 \nabla u_2 \cdot \nabla v_2 \, dx = \int_{\Omega_2} f v_2 \, dx - \int_{\Gamma} k_2 \nabla u_2 \cdot n \, dx$$

Par sommation, on obtient

$$\int_{\Omega} A \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx,$$

les deux termes de flux sur l'interface Γ se compensant. On en déduit que $A \nabla u$ est un élément de $H(\text{div})$ et que

$$-\text{div}(A \nabla u) = f \text{ en tant qu'éléments de } L^2(\Omega).$$

On a donc prouvé que u est l'unique solution de (5.39).

Exercice 5.2.12 Montrer l'existence et l'unicité de la solution de

$$\begin{cases} -\text{div}(A \nabla u) + u = f & \text{dans } \Omega \\ \frac{\partial u}{\partial n_A} = g & \text{sur } \partial\Omega \end{cases}$$

avec $f \in L^2(\Omega)$ et $g \in L^2(\partial\Omega)$.

Correction. La formulation variationnelle consiste à trouver $u \in H^1(\Omega)$ tel que

$$a(u, v) = L(v) \text{ pour tout } v \in H^1(\Omega)$$

ou

$$a(u, v) = \int_{\Omega} (A \nabla u \cdot \nabla v + uv) \, dx$$

et

$$L(v) = \int_{\Omega} f v \, dx + \int_{\partial\Omega} g v \, ds.$$

L'existence d'une solution à ce problème découle d'une application aisée du théorème de Lax-Milgram. Enfin, le Lemme 5.2.13 reste valable pour un opérateur elliptique

du deuxième ordre à coefficients variables, pourvu que A soit suffisamment régulier. En particulier, si pour tout i et j , $a_{ij} \in C^1(\overline{\Omega})$, $u \in H^2(\Omega)$. Ainsi, on obtient que

$$-\operatorname{div}(A\nabla u) = f \text{ en tant qu'éléments de } L^2(\Omega),$$

que la trace $\frac{\partial u}{\partial n_A}$ est bien définie sur $\partial\Omega$ et que

$$\frac{\partial u}{\partial n_A} = g \text{ dans } L^2(\partial\Omega).$$

Exercice 5.2.13 Montrer que l'application (non-linéaire) $v \rightarrow v^+$ est continue de $L^2(\Omega)$ dans lui-même, ainsi que de $H^1(\Omega)$ dans lui-même (utiliser le fait que $\nabla u = 0$ presque partout sur l'ensemble $u^{-1}(0)$).

Correction.

La continuité de l'application $v \rightarrow v^+$ de $L^2(\Omega)$ dans $L^2(\Omega)$ est évidente, car Lipschitzienne. En effet, pour tout $u, v \in L^2(\Omega)$, on a

$$\|v^+ - u^+\|_{L^2(\Omega)} \leq \|v - u\|_{L^2(\Omega)}.$$

La continuité de cette application de $H^1(\Omega)$ dans lui-même est un peu plus délicate. Considérons une suite v_n convergeant vers v dans $H^1(\Omega)$. Soit $v_{n'}$ une sous-suite extraite quelconque de v_n . De cette sous suite, on peut extraire une nouvelle sous-suite $v_{n''}$ convergeant presque partout. On a

$$\begin{aligned} \|\nabla v_{n''}^+ - \nabla v^+\|_{L^2(\Omega)} &= \|1_{v_{n''}>0} \nabla v_{n'} - 1_{v>0} \nabla v\|_{L^2(\Omega)} \\ &\leq \|1_{v_{n''}>0} (\nabla v_{n'} - \nabla v)\|_{L^2(\Omega)} + \|(1_{v_{n''}>0} - 1_{v>0}) \nabla v\|_{L^2(\Omega)} \\ &\leq \|\nabla v_{n''} - \nabla v\|_{L^2(\Omega)} + \|(1_{v_{n''}>0} - 1_{v>0}) \nabla v\|_{L^2(\Omega)}. \end{aligned}$$

Il est clair que le premier terme du second membre converge vers zéro. Enfin,

$$\|(1_{v_{n''}>0} - 1_{v>0}) \nabla v\|_{L^2(\Omega)}^2 = \int_{\Omega \setminus v^{-1}(0)} (1_{v_{n''}>0} - 1_{v>0})^2 |\nabla v|^2 dx.$$

car $\nabla v = 0$ presque partout sur $v^{-1}(0)$. Comme l'application $x \rightarrow 1_{x>0}$ est continue sur $\mathbb{R} \setminus \{0\}$,

$$1_{v_{n''}>0}(x) \rightarrow 1_{v>0}(x) \text{ pour presque tout } x \in \Omega \setminus v^{-1}(0).$$

Ainsi, d'après le théorème de convergence dominée de Lebesgue,

$$\|(1_{v_{n''}>0} - 1_{v>0}) \nabla v\|_{L^2(\Omega)} \rightarrow 0 \text{ lorsque } n'' \rightarrow \infty$$

et

$$\nabla v_{n''}^+ \rightarrow \nabla v^+ \text{ dans } L^2(\Omega).$$

On en déduit que toute la suite ∇v_n^+ converge vers ∇v^+ . En effet, dans le cas contraire, il existerait un réel $\varepsilon > 0$, et une sous-suite $v_{n'}$ de v_n tels que

$$\|\nabla v_{n'}^+ - \nabla v^+\|_{L^2(\Omega)} > \varepsilon,$$

ce qui contredit le fait qu'on puisse construire une sous-suite $v_{n''}$ de $v_{n'}$ telle que $\nabla v_{n''}^+ \rightarrow \nabla v^+$ dans $L^2(\Omega)$. En conclusion, on a montré que si $v_n \rightarrow v$ dans $H^1(\Omega)$, alors $v_n^+ \rightarrow v^+$ dans $L^2(\Omega)$ et $\nabla v_n^+ \rightarrow \nabla v^+$ dans $L^2(\Omega)$. En d'autres termes, $v_n^+ \rightarrow v^+$ dans $H^1(\Omega)$ et l'application qui à v associe v^+ est continue de $H^1(\Omega)$ dans $H^1(\Omega)$.

Exercice 5.3.1 Montrer que l'application de $L^2(\Omega)^N$ dans $H_0^1(\Omega)^N$ qui à f fait correspondre u , unique solution faible de

$$\begin{cases} -\operatorname{div}(2\mu e(u) + \lambda \operatorname{tr}(e(u)) \operatorname{Id}) = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega. \end{cases}, \quad (5.12)$$

est linéaire continue.

Correction.

La linéarité de cette application est évidente. La continuité est une conséquence du Théorème de Lax-Milgram (qu'on a appliqué pour démontrer l'existence et l'unicité de la solution de (5.12)). On peut retrouver la continuité directement, en appliquant la formulation variationnelle à la fonction test $v = u$. On obtient

$$\int_{\Omega} (|e(u)|^2 + (\operatorname{div} u)^2) dx = \int_{\Omega} f \cdot u dx.$$

En combinant cette égalité à l'inégalité de Korn

$$C \int_{\Omega} (|e(u)|^2 + (\operatorname{div} u)^2) dx \geq \|u\|_{H^1(\Omega)}^2$$

et à l'inégalité de Cauchy-Schwarz, on en déduit que

$$\|u\|_{H^1(\Omega)} \leq C \|f\|_{L^2(\Omega)}.$$

Exercice 5.3.2 Soit Ω un ouvert connexe de \mathbb{R}^N . Soit l'ensemble \mathcal{R} des “mouvements rigides” de Ω défini par

$$\mathcal{R} = \{v(x) = b + Mx \text{ avec } b \in \mathbb{R}^N, M = -M^t \text{ matrice antisymétrique}\}. \quad (5.13)$$

Montrer que $v \in H^1(\Omega)^N$ vérifie $e(v) = 0$ dans Ω si et seulement si $v \in \mathcal{R}$.

Correction. Tout d'abord, si v appartient à \mathcal{R} , on a évidemment $e(v) = 0$. Réciproquement, soit $v \in H^1(\Omega)^N$ telle que $e(v) = 0$. On pose $w = \frac{1}{2}(\nabla v - (\nabla v)^t)$, partie antisymétrique de ∇v ,

$$w_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i} \right).$$

La fonction w_{ij} est un élément de $L^2(\Omega)$. De plus, en effectuant diverses intégrations par partie, on peut établir que pour toute fonction $\varphi \in C_c^\infty(\Omega)$,

$$\int_{\Omega} w_{ij} \frac{\partial \varphi}{\partial x_k} dx = \int_{\Omega} e_{ik}(v) \frac{\partial \varphi}{\partial x_j} - e_{jk}(v) \frac{\partial \varphi}{\partial x_i} dx.$$

Comme $e(v) = 0$, on en déduit que pour tout k ,

$$\frac{\partial w_{ij}}{\partial x_k} = 0.$$

Ainsi, chaque w_{ij} admet une dérivée faible $L^2(\Omega)$ nulle et d'après la Proposition 4.2.5, il existe une matrice constante M telle que $w_{ij}(x) = M$ presque partout. De plus, w étant antisymétrique, M l'est également. Puisque $e(v) = 0$, on en déduit que

$$\nabla v = M.$$

Enfin,

$$\nabla(v - Mx) = 0.$$

De nouveau par application de la Proposition 4.2.5, on en déduit qu'il existe un vecteur constant b tel que

$$v(x) = b + Mx \text{ pour presque tout } x \in \Omega.$$

Exercice 5.3.3 Montrer que $u \in V = \{v \in H^1(\Omega)^N \text{ tel que } v = 0 \text{ sur } \partial\Omega_D\}$ est l'unique solution de la formulation variationnelle

$$\int_{\Omega} 2\mu e(u) \cdot e(v) dx + \int_{\Omega} \lambda \operatorname{div} u \operatorname{div} v dx = \int_{\Omega} f \cdot v dx + \int_{\partial\Omega_N} g \cdot v ds \quad \forall v \in V. \quad (5.14)$$

si et seulement si u réalise le minimum sur V de l'énergie

$$J(v) = \frac{1}{2} \int_{\Omega} (2\mu |e(v)|^2 + \lambda |\operatorname{div} v|^2) dx - \int_{\Omega} f \cdot v dx - \int_{\partial\Omega_N} g \cdot v ds. \quad (5.15)$$

(Indication : on pourra s'inspirer de la Proposition 3.3.4).

Correction. Il suffit d'appliquer la Proposition 3.3.4 à la forme bilinéaire

$$a(u, v) = \int_{\Omega} (2\mu e(u) \cdot e(v) + \lambda (\operatorname{div} u)(\operatorname{div} v)) dx$$

et à la forme linéaire

$$L(v) = \int_{\Omega} f \cdot v dx + \int_{\partial\Omega_N} g \cdot v ds,$$

sur l'espace de Hilbert V .

Exercice 5.3.4 Soit Ω un ouvert borné connexe de \mathbb{R}^N . On considère le système de l'élasticité avec la condition de Neumann (5.59) sur tout le bord $\partial\Omega$. Montrer que la condition d'équilibre (vectorielle)

$$\int_{\Omega} f dx + \int_{\partial\Omega} g ds = 0$$

est une condition nécessaire et suffisante d'existence et d'unicité d'une solution dans $H^1(\Omega)^N$ (l'unicité étant obtenue "à un mouvement de corps rigide" près, c'est-à-dire à l'addition de $Mx + b$ près avec $b \in \mathbb{R}^N$ et M une matrice antisymétrique constante).

Correction.

En intégrant l'équation sur Ω , on obtient, suite à une intégration par partie, la condition de compatibilité

$$\int_{\Omega} f \, dx + \int_{\partial\Omega} g \, ds = 0.$$

Sous cette condition, on va montrer que le problème aux limites avec condition de Neumann admet une unique solution dans l'espace V , quotient de $H^1(\Omega)^N$ par l'espace des mouvements rigides \mathcal{R} . La formulation variationnelle est aisée à établir et consiste à trouver $u \in V$ tel que

$$a(u, v) = L(v) \text{ pour tout } v \in V$$

où

$$a(u, v) = \int_{\Omega} (2\mu e(u) \cdot e(v) + \lambda(\operatorname{div} u)(\operatorname{div} v)) \, dx$$

et

$$L(v) = \int_{\Omega} f \cdot v \, dx + \int_{\partial\Omega} g \cdot v$$

Notons que $a(u, v)$ et $L(v)$ sont toutes deux correctement définies. Leurs valeurs sont indépendantes des représentants u et v choisis dans $H^1(\Omega)$. Le seul point délicat afin d'appliquer le théorème de Lax-Milgram consiste à prouver la coercivité de la forme bilinéaire, c'est à dire qu'il existe une constante C telle que

$$\|u\|_V^2 \leq Ca(u, u) \text{ pour tout } u \in V. \quad (5.16)$$

où

$$\|u\|_V = \inf_{M, b} \|u + M.x + b\|_{H^1(\Omega)}, \text{ avec } M \text{ matrice antisymétrique et } b \in \mathbb{R}^n.$$

Supposons que la relation (5.16) soit fausse pour tout C . Dans ce cas, il existe une suite u_n d'éléments de V telle que

$$1 = \|u_n\|_V^2 \geq na(u_n, u_n).$$

Rappelons qu'il existe ν tel que

$$a(u, u) \geq \nu \|e(u)\|_{L^2(\Omega)}^2.$$

Ainsi,

$$1 = \|u_n\|_V^2 \geq \nu n \|e(u_n)\|_{L^2(\Omega)}^2$$

D'après le théorème de Rellich, il existe une sous-suite $u_{n'}$ convergente dans $L^2(\Omega)$ quotienté par \mathcal{R} . De plus, comme $e(u_{n'})$ tend vers zéro, on en déduit que $u_{n'}$ converge dans V vers un élément u tel que $e(u) = 0$. D'après l'exercice précédent, il existe M matrice antisymétrique et $b \in \mathbb{R}^N$ tels que $u(x) = M.x + b$. En d'autres termes, $u = 0$ dans V , ce qui contredit le fait que $\|u\|_V = 1$. Afin de prouver que la solution

du problème variationnel est solution du problème aux limites, on procède comme pour le Laplacien. En particulier, afin de donner un sens à $\sigma.n$, il serait nécessaire de montrer que u est en fait un élément de $H^2(\Omega)$ (ce qu'on a admis pour le Laplacien). A défaut, on peut toujours utiliser le fait que σ est un élément de $H(\text{div})$ et utiliser la définition faible de la trace de la normale de σ sur le bord comme élément de $H^{-1/2}(\partial\Omega)$ (voir Théorème 4.4.7)

Exercice 5.3.5 On suppose que Ω est un ouvert borné de \mathbb{R}^N et que $f \in L^2(\Omega)^N$. Montrer l'existence et l'unicité d'une solution faible dans $H_0^1(\Omega)^N$ au système de Lamé

$$\begin{cases} -\mu\Delta u - (\mu + \lambda)\nabla(\text{div}u) = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega. \end{cases} \quad (5.17)$$

sans utiliser l'inégalité de Korn. Vérifier qu'on peut affaiblir les hypothèses de positivité sur les coefficients de Lamé en supposant seulement que $\mu > 0$ et $\mu + 2\lambda > 0$.

Correction. La formulation variationnelle consiste à trouver $u \in H_0^1(\Omega)$ tel que

$$a(u, v) = L(v) \text{ pour tout } v \in H_0^1(\Omega),$$

où

$$a(u, v) = \int_{\Omega} (\mu \nabla u \cdot \nabla v + (\lambda + \mu)(\text{div}u)(\text{div}v)) \, dx$$

et

$$L(v) = \int_{\Omega} f v \, dx.$$

Afin d'appliquer le théorème de Lax-Milgram, la seule hypothèse non triviale à vérifier est la coercivité de la forme bilinéaire $a(., .)$. Or

$$\int_{\Omega} (\text{div}u)^2 \, dx = \int_{\Omega} \sum_{i,j} \frac{\partial u_i}{\partial x_j} \frac{\partial u_j}{\partial x_i} \, dx$$

Par intégration par partie, il vient

$$\begin{aligned} \int_{\Omega} (\text{div}u)^2 \, dx &= \int_{\Omega} \sum_{i,j} \frac{\partial u_i}{\partial x_j} \frac{\partial u_j}{\partial x_i} \, dx = \int_{\Omega} \nabla u \cdot (\nabla u)^t \, dx \\ &\leq \int_{\Omega} |\nabla u| |(\nabla u)^t| \, dx = \int_{\Omega} |\nabla u|^2 \, dx. \end{aligned}$$

Ainsi,

$$a(u, u) \geq \int_{\Omega} (\mu + \min(0, \lambda + \mu)) |\nabla u|^2 \, dx$$

ou encore

$$a(u, u) \geq \int_{\Omega} \min(\mu, \lambda + 2\mu) |\nabla u|^2 \, dx.$$

La forme bilinéaire $a(., .)$ est donc coercive dès que $\mu > 0$ et $\lambda + 2\mu > 0$, ce qui établit l'existence d'une solution unique au problème variationnel. On montre que u est solution du problème aux limites en procédant comme pour le Laplacien.

Exercice 5.3.6 Vérifier l'équivalence de (5.17) et (5.12) si λ et μ sont constants. Montrer que (5.17) et (5.12) ne sont plus équivalents si λ et μ sont des fonctions (régulières), même si on remplace l'équation vectorielle de (5.17) par

$$-\operatorname{div}(\mu \nabla u) - \nabla((\mu + \lambda)\operatorname{div} u) = f \text{ dans } \Omega.$$

Correction. Soit u la solution du problème variationnel de l'élasticité linéarisée avec condition de Dirichlet, pour tout $v \in C_c^\infty(\Omega)^N$,

$$\int_{\Omega} \mu \sum_{i,j} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \frac{\partial v_i}{\partial x_j} dx + \int_{\Omega} \lambda (\operatorname{div} u)(\operatorname{div} v) dx = \int_{\Omega} f \cdot v dx.$$

Or par intégration par partie,

$$\begin{aligned} \int_{\Omega} \mu \frac{\partial u_j}{\partial x_i} \frac{\partial v_i}{\partial x_j} dx &= - \int_{\Omega} u_j \frac{\partial}{\partial x_i} \left(\mu \frac{\partial v_i}{\partial x_j} \right) dx \\ &= - \int_{\Omega} \mu u_j \frac{\partial^2 v_i}{\partial x_j \partial x_i} dx - \int_{\Omega} u_j \frac{\partial \mu}{\partial x_i} \frac{\partial v_i}{\partial x_j} dx \\ &= \int_{\Omega} \frac{\partial \mu u_j}{\partial x_j} \frac{\partial v_i}{\partial x_i} dx - \int_{\Omega} u_j \frac{\partial \mu}{\partial x_i} \frac{\partial v_i}{\partial x_j} dx \\ &= \int_{\Omega} \mu \frac{\partial u_j}{\partial x_j} \frac{\partial v_i}{\partial x_i} dx + \int_{\Omega} u_j \left(\frac{\partial \mu}{\partial x_j} \frac{\partial v_i}{\partial x_i} - \frac{\partial \mu}{\partial x_i} \frac{\partial v_i}{\partial x_j} \right) dx. \end{aligned}$$

Ainsi,

$$\begin{aligned} \int_{\Omega} \mu \sum_{i,j} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \frac{\partial v_i}{\partial x_j} dx + \int_{\Omega} \lambda (\operatorname{div} u)(\operatorname{div} v) dx = \\ \int_{\Omega} \mu \nabla u \cdot \nabla v + (\lambda + \mu)(\operatorname{div} u)(\operatorname{div} v) dx + \int_{\Omega} u \cdot ((\operatorname{div} v) \nabla \mu - (\nabla v)^t \nabla \mu) dx. \end{aligned}$$

Si μ est constant, u est donc également l'unique solution du problème variationnel consistant à trouver u dans $H_0^1(\Omega)^N$ tel que pour tout $v \in H_0^1(\Omega)^N$,

$$\int_{\Omega} (\mu \nabla u \cdot \nabla v + (\lambda + \mu)(\operatorname{div} u)(\operatorname{div} v)) dx = \int_{\Omega} f \cdot v dx,$$

qui est équivalent au problème aux limites consistant à trouver u tel que

$$\begin{cases} -\mu \Delta u - \nabla \cdot ((\mu + \lambda)\operatorname{div} u) = f & \text{dans } \Omega, \\ u = 0 & \text{sur } \partial\Omega. \end{cases}$$

Si de plus λ est constant, on retrouve le problème aux limites (5.17). Enfin, si μ n'est pas constant, on ne peut rien dire.

Exercice 5.3.7 Le but de cet exercice est de trouver une solution particulière du système de l'élasticité linéarisée dans le cas d'une force de cisaillement anti-plan. On considère un domaine cylindrique homogène Ω de longueur $L > 0$ et de section ω , où

ω est un ouvert borné connexe régulier de \mathbb{R}^{N-1} (les coefficients de Lamé λ et μ sont constants). Autrement dit, $\Omega = \omega \times (0, L)$, et pour $x \in \Omega$, on note $x = (x', x_N)$ avec $0 < x_N < L$ et $x' \in \omega$. On considère le problème aux limites suivant

$$\begin{cases} -\operatorname{div}(2\mu e(u) + \lambda \operatorname{tr}(e(u)) \operatorname{Id}) = 0 & \text{dans } \Omega \\ \sigma n = g & \text{sur } \partial\omega \times (0, L) \\ u' = 0 & \text{sur } \omega \times \{0, L\} \\ (\sigma n) \cdot n = 0 & \text{sur } \omega \times \{0, L\} \end{cases} \quad (5.18)$$

où on a utilisé la notation, pour un vecteur $v = (v_1, \dots, v_N)$, $v = (v', v_N)$ avec $v' \in \mathbb{R}^{N-1}$ et $v_N \in \mathbb{R}$. On suppose que la force surfacique g est du type “cisaillement anti-plan”, c'est-à-dire que $g' = (g_1, \dots, g_{N-1}) = 0$. Montrer que la solution unique de (5.18) est donnée par $u = (0, \dots, 0, u_N)$ où $u_N(x')$ est la solution du Laplacien suivant

$$\begin{cases} -\Delta' u_N = 0 & \text{dans } \omega \\ \mu \frac{\partial u_N}{\partial n} = g_N & \text{sur } \partial\omega \end{cases}$$

où Δ' est le Laplacien dans la variable $x' \in \mathbb{R}^{N-1}$.

Correction. Soit u_N la solution du problème de Laplace

$$\begin{cases} -\Delta' u_N = 0 & \text{sur } \omega \\ \mu \frac{\partial u_N}{\partial n} = g_N & \text{sur } \partial\omega \end{cases}$$

On pose $u = (0, \dots, 0, u_N)$. Pour tout i et j tels que $i, j < N$,

$$\begin{aligned} e_{ij}(u) &= 0 \\ e_{iN}(u) &= e_{Ni}(u) = \frac{1}{2} \frac{\partial u_N}{\partial x_i} \\ e_{NN}(u) &= 0. \end{aligned}$$

En particulier, $\operatorname{tr}(e(u)) = 0$. On en déduit que,

$$-\operatorname{div}(2\mu e(u) + \lambda \operatorname{tr}(e(u)) \operatorname{Id}) = -\mu(0, \dots, 0, \Delta' u_N) = 0.$$

De plus,

$$\sigma(u) e_N = 2\mu e(u) e_N = \mu(\nabla' u_N, 0).$$

Ainsi, pour presque tout $x \in \omega \times \{0, L\}$, $(\sigma n) \cdot n = 0$. Enfin, pour presque tout $x \in \partial\omega \times (0, L)$, $n = (n', 0)$ et

$$\sigma n = 2\mu \left(\sum_{k=1}^{N-1} e_{jk} n_k \right)_j = 2\mu(0, \dots, 0, 1/2 \nabla' u_N \cdot n') = (0, \dots, 0, g_N).$$

Ainsi, u est bien l'unique solution du problème aux limites (5.18).

Exercice 5.3.8 Généraliser l'Exercice 5.3.7 au cas d'une condition aux limites latérale du type

$$u' = 0 \text{ et } (\sigma n) \cdot e_N = g_N \text{ sur } \partial\omega \times (0, L).$$

Correction. La solution du problème de l'élasticité linéarisée n'est pas modifiée par le changement des conditions aux limites proposé sur $\partial\omega \times (0, L)$.

Exercice 5.3.9 A l'aide de l'approche variationnelle démontrer l'existence et l'unicité de la solution de l'équation des plaques

$$\begin{cases} \Delta(\Delta u) = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega \\ \frac{\partial u}{\partial n} = 0 & \text{sur } \partial\Omega \end{cases} \quad (5.19)$$

où $f \in L^2(\Omega)$. On pourra remarquer que, si $u \in H_0^2(\Omega)$, alors $\frac{\partial u}{\partial x_i} \in H_0^1(\Omega)$ et

$$\int_{\Omega} |\Delta u|^2 dx = \sum_{i,j=1}^N \int_{\Omega} \left| \frac{\partial^2 u}{\partial x_i \partial x_j} \right|^2 dx.$$

On admettra le résultat de régularité suivant : si $w \in L^2(\Omega)$ et $f \in L^2(\Omega)$ vérifient pour tout $v \in C_c^\infty(\Omega)$

$$-\int_{\Omega} w \Delta v dx = \int_{\Omega} f v dx,$$

alors $(\theta w) \in H^2(\Omega)$ quelle que soit la fonction $\theta \in C_c^\infty(\Omega)$.

Correction.

La formulation variationnelle associée à l'équation des plaques (5.19) consiste à déterminer $u \in H_0^2(\Omega)$ tel que

$$a(u, v) = L(v) \quad \text{pour tout } v \in H_0^2(\Omega) \quad (5.20)$$

où

$$a(u, v) = \int_{\Omega} \Delta u \Delta v dx \text{ et } L(v) = \int_{\Omega} f v dx$$

(voir Exercice 3.2.4). Afin d'appliquer le Théorème de Lax-Milgram, la seule hypothèse non trivialement vérifiée est la coercivité de la forme bilinéaire $a(., .)$. Or pour tout $u \in H_0^2(\Omega)$, on établit suite à deux intégrations par partie successives que

$$a(u, u) = \sum_{i,j} \int_{\Omega} \left| \frac{\partial^2 u}{\partial x_i \partial x_j}(x) \right|^2 dx = \|\nabla^2 u\|_{L^2}^2.$$

En appliquant deux fois l'inégalité de Poincaré, on obtient qu'il existe des constantes C et C' positives telles que pour tout élément u de $H_0^2(\Omega)$,

$$\|u\|_{L^2(\Omega)}^2 \leq C \|\nabla u\|_{L^2}^2 \leq C' \|\nabla^2 u\|_{L^2}^2 = C' a(u, u).$$

La coercivité de $a(., .)$ est donc établie et il existe une unique solution au problème variationnel (5.20).

Reste à établir que la solution du problème variationnel est solution du problème aux limites. Soit K un compact de Ω et $\theta \in C_c^\infty(\Omega)$ telle que $\theta = 1$ sur K . Pour toute fonction $v \in C_c^\infty(\Omega)$ à support inclus dans K ,

$$\int_{\Omega} \theta(x) \Delta u(x) \Delta v(x) dx = \int_{\Omega} \Delta u(x) \Delta v(x) dx = \int_{\Omega} f(x) v(x) dx.$$

D'après le résultat de régularité admit, $\theta \Delta u$ est un élément de $H^2(\Omega)$. Il est donc licite d'effectuer deux intégrations par partie successives sur le membre de gauche de l'équation précédente. On en déduit que

$$\int_{\Omega} \Delta(\theta(x) \Delta u(x)) v(x) dx = \int_{\Omega} f(x) v(x) dx.$$

En d'autres termes, pour presque tout $x \in K$,

$$\Delta(\Delta u)(x) = f(x).$$

Cette relation reste valable pour presque tout $x \in \Omega$: Il suffit de considérer une suite K_n de compacts tels que $\cup_n K_n = \Omega$. Enfin, comme $u \in H_0^2(\Omega)$, la solution du problème variationnel vérifie automatiquement les conditions au bord $u = \partial u / \partial n = 0$.

Exercice 5.3.10 Soit V l'espace des champs de vitesse à divergence nulle. Soit $J(v)$ l'énergie définie pour $v \in V$ par

$$J(v) = \frac{1}{2} \int_{\Omega} \mu |\nabla v|^2 dx - \int_{\Omega} f \cdot v dx. \quad (5.21)$$

Soit $u \in V$ la solution unique de la formulation variationnelle

$$\int_{\Omega} \mu \nabla u \cdot \nabla v dx = \int_{\Omega} f \cdot v dx \quad \forall v \in V. \quad (5.22)$$

Montrer que u est aussi l'unique point de minimum de l'énergie, c'est-à-dire que $J(u) = \min_{v \in V} J(v)$. Réciproquement, montrer que, si $u \in V$ est un point de minimum de l'énergie $J(v)$, alors u est la solution unique de la formulation variationnelle (5.22).

Correction. Il suffit d'appliquer la Proposition 3.3.4 à la formulation variationnelle (5.22) pour conclure.

Exercice 5.3.11 Le but de cet exercice est de trouver une solution particulière des équations de Stokes dans un canal rectiligne de section uniforme, appelée profil de Poiseuille. Soit $\Omega = \omega \times (0, L)$ où $L > 0$ est la longueur du canal et ω sa section, un ouvert borné connexe régulier de \mathbb{R}^{N-1} . Pour $x \in \Omega$, on note $x = (x', x_N)$ avec $0 < x_N < L$ et $x' \in \omega$. On considère le problème aux limites suivant

$$\begin{cases} \nabla p - \mu \Delta u = 0 & \text{dans } \Omega \\ \operatorname{div} u = 0 & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\omega \times (0, L) \\ pn - \mu \frac{\partial u}{\partial n} = p_0 n & \text{sur } \omega \times \{0\} \\ pn - \mu \frac{\partial u}{\partial n} = p_L n & \text{sur } \omega \times \{L\} \end{cases} \quad (5.23)$$

où p_0 et p_L sont deux pressions constantes. Montrer que la solution unique de (5.23) est donnée par

$$p(x) = p_0 + \frac{x_N}{L}(p_L - p_0),$$

et $u = (0, \dots, 0, u_N)$ où $u_N(x')$ est la solution du Laplacien suivant

$$\begin{cases} -\mu \Delta' u_N = -\frac{(p_L - p_0)}{L} & \text{dans } \omega \\ u_N = 0 & \text{sur } \partial\omega \end{cases}$$

où Δ' est le Laplacien dans la variable $x' \in \mathbb{R}^{N-1}$.

Correction. On pose

$$p(x) = p_0 + x_N(p_L - p_0)/L,$$

et $u = (0, \dots, 0, u_N)$ où u_N est solution du problème aux limites

$$\begin{cases} -\mu \Delta' u_N = -\frac{(p_L - p_0)}{L} & \text{dans } \omega \\ u_N = 0 & \text{sur } \partial\omega. \end{cases}$$

On va montrer que (u, p) est solution du problème aux limites (5.23). On a

$$\nabla p = (0, \dots, 0, (p_L - p_0)/L),$$

$$\Delta u = (0, \dots, 0, \Delta' u_N),$$

d'où

$$\nabla p - \mu \Delta u = (0, \dots, 0, (p_L - p_0)/L - \mu \Delta' u_N) = 0.$$

De plus,

$$\operatorname{div}(u) = \frac{\partial u_N}{\partial x_N} = 0.$$

Enfin, comme

$$\begin{cases} \frac{\partial u}{\partial n} = 0 \text{ et } p = p_0 & \text{sur } \omega \times \{0\}, \\ p = p_1 & \text{sur } \omega \times \{L\} \end{cases}$$

les conditions aux limites imposées aux extrémités du profil sont également vérifiées.

Exercice 5.3.12 Généraliser l'Exercice 5.3.11 au cas des équations de Navier-Stokes

$$\begin{cases} (u \cdot \nabla)u + \nabla p - \mu \Delta u = f & \text{dans } \Omega \\ \operatorname{div} u = 0 & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega. \end{cases} \quad (5.24)$$

Correction. Avec les mêmes notations que l'exercice précédent, on vérifie que

$$(u \cdot \nabla)u = 0,$$

ainsi, u est également solution des équations de Navier-Stokes.

Chapitre 6

MÉTHODE DES ÉLÉMENTS FINIS

Exercice 6.2.1 Appliquer la méthode des éléments finis \mathbb{P}_1 au problème

$$\begin{cases} -u'' = f \text{ dans }]0, 1[\\ u(0) = \alpha, u(1) = \beta, \end{cases}$$

Vérifier que les conditions aux limites de Dirichlet non-homogènes apparaissent dans le second membre du système linéaire qui en découle.

Correction. La formulation variationnelle, issue de l'utilisation des éléments finis \mathbb{P}_1 , consiste à déterminer

$$u_h \in V_h := \{v_h \in C^0([0, 1]; \mathbb{R}) : v|_{[x_i, x_{i+1}]} \in \mathbb{P}_1 \text{ pour tout } i \in \{0, \dots, n\}\}$$

où $x_i = i/(n+1)$ tel que

$$\int_0^1 u'_h v'_h dx = \int_0^1 f v_h dx \text{ pour toute fonction } v_h \in V_{0h} = V_h \cap H_0^1(0, 1),$$

et

$$u_h(0) = \alpha, \quad u_h(1) = \beta.$$

On note $(\phi_i)_{i=0, \dots, n+1}$ la base de V_h définie par $\phi_i(x_j) = \delta_{i,j}$. En utilisant ϕ_j comme fonction test, on obtient, à l'aide de la formulation variationnelle, que pour tout $0 < j < n+1$,

$$\sum_{i=0}^{n+1} (u_h)_i \int_0^1 \phi'_i \phi'_j dx = \int_0^1 f \phi_j dx.$$

Les conditions aux limites impliquent que $(u_h)_0 = \alpha$ et $(u_h)_{n+1} = \beta$, ainsi

$$\sum_{i=1}^n (u_h)_i \int_0^1 \phi'_i \phi'_j dx = \int_0^1 f \phi_j dx - \int_0^1 (\alpha \phi'_0 + \beta \phi'_{n+1}) \phi'_j dx.$$

Déterminer $U_h = ((u_h)_i)_{1 \leq i \leq n}$ consiste donc à résoudre le système linéaire

$$\mathcal{K}_h U_h = b_h,$$

où la matrice \mathcal{K}_h est identique à celle obtenue avec des conditions de Dirichlet homogènes, tandis que le second membre est défini par

$$\begin{aligned}(b_h)_i &= \int_{x_{i-1}}^{x_{i+1}} f \phi_i dx, & \text{pour tout } 1 < i < n, \\(b_h)_1 &= \alpha/h + \int_0^{x_2} f \phi_1 dx \\(b_h)_n &= \beta/h + \int_{x_{n-1}}^1 f \phi_n dx.\end{aligned}$$

Exercice 6.2.2 On reprend le problème de Neumann

$$\begin{cases} -u'' + au = f & \text{dans }]0, 1[\\ u'(0) = \alpha, u'(1) = \beta. \end{cases} \quad (6.1)$$

en supposant que la fonction $a(x) = 0$ dans Ω . Montrer que la matrice du système linéaire issu de la méthode des éléments finis \mathbb{P}_1 est singulière. Montrer qu'on peut néanmoins résoudre le système linéaire si les données vérifient la condition de compatibilité

$$\int_0^1 f(x) dx = \alpha - \beta,$$

et que cette condition est préservée si l'on utilise des formules de quadrature. Comparer ce résultat avec le Théorème **5.2.18**.

Correction. Le système linéaire obtenu en considérant $a = 0$ est

$$\mathcal{K}_h U_h = b_h, \quad (6.2)$$

où

$$\mathcal{K}_h = \frac{1}{h} \begin{pmatrix} 1 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 1 \end{pmatrix}$$

et b_h est défini comme dans le cas $a \neq 0$. L'application \mathcal{K}_h est auto-adjointe et

positive. En effet, pour tout $(v_i) \in \mathbb{R}^{n+2}$, on a

$$\begin{aligned}
\mathcal{K}_h v \cdot v &= h^{-1} \left((v_0 - v_1)v_0 + (v_{n+1} - v_n)v_{n+1} + \sum_{i=1}^n (-v_{i+1} + 2v_i - v_{i-1})v_i \right) \\
&= h^{-1} \left((v_0 - v_1)v_0 + (v_{n+1} - v_n)v_{n+1} + \sum_{i=1}^n (v_i - v_{i+1})v_i + (v_i - v_{i-1})v_i \right) \\
&= h^{-1} \left((v_0 - v_1)v_0 + (v_{n+1} - v_n)v_{n+1} + \sum_{i=1}^n (v_i - v_{i+1})v_i + \sum_{i=0}^{n-1} (v_{i+1} - v_i)v_{i+1} \right) \\
&= h^{-1} \left((v_0 - v_1)^2 + (v_{n+1} - v_n)^2 + \sum_{i=1}^{n-1} (v_i - v_{i+1})^2 \right) \\
&= h^{-1} \sum_{i=0}^n (v_i - v_{i+1})^2.
\end{aligned}$$

Par contre \mathcal{K}_h n'est pas définie. De l'expression précédente, on déduit que $\mathcal{K}_h v \cdot v = 0$ si et seulement si $v_i = v_{i+1}$ pour tout $i = 0, \dots, n$. Ainsi, le noyau de l'application \mathcal{K}_h est l'espace vectoriel de dimension un engendré par $(1, \dots, 1)$ et l'image de \mathcal{K}_h est exactement l'orthogonal de $(1, \dots, 1)$. Le système linéaire (6.2) admet une solution si et seulement si $b_h \in (1, \dots, 1)^\perp$, c'est à dire

$$\sum_{i=0}^{n+1} (b_h)_i = 0.$$

D'après l'expression de b_h , cette condition équivaut à

$$\int_0^1 f(x) dx + \beta - \alpha = \sum_{i=0}^{n+1} \int_0^1 f(x) \phi_i(x) dx + \beta - \alpha = \sum_{i=0}^{n+1} (b_h)_i = 0.$$

Exercice 6.2.3 Appliquer la méthode des différences finies (voir le Chapitre 2) au problème de Dirichlet

$$\begin{cases} -u'' = f & \text{dans }]0, 1[\\ u(0) = u(1) = 0. \end{cases} \quad (6.3)$$

Vérifier qu'avec un schéma centré d'ordre deux, on obtient un système linéaire à résoudre avec la même matrice \mathcal{K}_h (à un coefficient multiplicatif près) mais avec un second membre b_h différent. Même question pour le problème de Neumann (6.1).

Correction. Conditions aux limites de Dirichlet

La méthode des différences finies, basée sur un schéma centré d'ordre 2, nous conduit à résoudre, dans le cas du Laplacien avec conditions de Dirichlet, le système

$$\begin{cases} -\frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} = f(x_i) & \text{pour tout } 0 < i < n+1, \\ u_0 = 0, \\ u_{n+1} = 0. \end{cases}$$

On doit donc résoudre le système

$$\mathcal{K}_h U_h = b_h$$

où $U_h = (u_i)_{1 \leq i \leq n}$, \mathcal{K}_h est la matrice d'ordre n

$$\mathcal{K}_h = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix} \text{ et } b_h = \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix}.$$

La matrice \mathcal{K}_h diffère de la matrice obtenue par la méthode des différences finies à un facteur multiplicatif $1/h$ près. La méthode des éléments finis conduit à une expression différente du second membre

$$b_h^{EF} = \left(\int_{x_{i-1}}^{x_i} f(x) \frac{x - x_i}{h} dx + \int_{x_i}^{x_{i+1}} f(x) \frac{x_{i+1} - x_i}{h} dx \right)_{1 \leq i \leq n}.$$

En pratique, on utilise une formule de quadrature pour évaluer les intégrales définissant b_h^{EF} . Si on utilise la formule des trapèzes, on obtient

$$b_h^{EF} = h(f(x_i))_{1 \leq i \leq n}.$$

Avec un tel choix, les deux méthodes conduisent au même système linéaire.

Conditions aux limites de Neumann

Pour le problème de Neumann, le système obtenu, suite à la discrétisation par différences finies, consiste à déterminer $(u_i)_{0 \leq i \leq n+1}$ tel que

$$\begin{cases} -\frac{u_{i-1} - 2u_i + u_{i+1}}{h} + ha(x_i)u_i = hf(x_i) \text{ pour tout } 0 < i < n+1, \\ \frac{u_1 - u_0}{h} = \alpha \\ \frac{u_{n+1} - u_n}{h} = \beta. \end{cases}$$

Il s'agit donc de résoudre le système linéaire $\mathcal{K}_h U_h = b_h$ où $U_h = (u_i)_{0 \leq i \leq n+1}$, \mathcal{K}_h est la matrice d'ordre $n+2$

$$\mathcal{K}_h = \frac{1}{h^2} \begin{pmatrix} -h & h & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -h & h \end{pmatrix} + \begin{pmatrix} 0 & \dots & & 0 \\ & a(x_1) & & \\ \vdots & & \ddots & \vdots \\ & & & a(x_n) \\ 0 & \dots & & & 0 \end{pmatrix}$$

et $b_h = (\alpha, f(x_1), \dots, f(x_n), \beta)^T$. Alors que le reste du schéma est d'ordre deux, la discrétisation des conditions aux limites proposée est seulement d'ordre un. Il

en résulte une perte de précision du schéma. Afin de pallier cet inconvénient, on propose la discrétisation des conditions aux limites d'ordre deux suivante

$$\frac{u_1 - u_{-1}}{2h} = \alpha \text{ et } \frac{u_{n+2} - u_n}{2h} = \beta,$$

où x_{-1} et x_{n+2} sont des noeuds fictifs. Si on élimine du système linéaire final les degrés de liberté artificiellement introduits, on obtient les expressions suivantes

$$\mathcal{K}_h = \frac{1}{h^2} \begin{pmatrix} 2 & -2 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -2 & 2 \end{pmatrix} + \begin{pmatrix} a(x_0) & 0 & \cdots & & 0 \\ 0 & a(x_1) & & & \\ \vdots & & \ddots & & \vdots \\ & & & a(x_n) & 0 \\ 0 & \cdots & & 0 & a(x_{n+1}) \end{pmatrix}$$

et $b_h = (-\frac{2\alpha}{h} + f(x_0), f(x_1), \dots, f(x_n), \frac{2\beta}{h} + f(x_{n+1}))^T$. Le système obtenu par la méthode des éléments finis, dès lors qu'on utilise la formule des trapèzes pour évaluer les intégrales, est équivalent. Plus précisément, on a alors

$$\mathcal{K}_h \simeq \frac{1}{h} \begin{pmatrix} 1 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 1 \end{pmatrix} + h \begin{pmatrix} a(x_0) & 0 & \cdots & & 0 \\ 0 & a(x_1) & & & \\ \vdots & & \ddots & & \vdots \\ & & & a(x_n) & 0 \\ 0 & \cdots & & 0 & a(x_{n+1}) \end{pmatrix}$$

et $b_h = h(-\frac{1\alpha}{h} + \frac{f(x_0)}{2}, f(x_1), \dots, f(x_n), \frac{2\beta}{h} + \frac{f(x_{n+1})}{2})^T$.

Exercice 6.2.4 On considère $(n+2)$ masses ponctuelles (alignées) situées aux points $x_j = j/(n+1)$ pour $0 \leq j \leq n+1$ et reliées entre voisines par des ressorts de même raideur $k > 0$. On applique à chaque masse ponctuelle une force longitudinale f_j . Dans l'hypothèse de petits déplacements (longitudinaux) écrire l'énergie totale du système qu'il faut minimiser (on discutera le cas des extrémités libres ou fixées). Interpréter la recherche de la position d'équilibre du système en termes d'éléments finis.

Correction. On note u_j le déplacement de la masse j . L'allongement du ressort situé entre les masses j et $j+1$ est

$$\delta L_j = u_{j+1} - u_j.$$

Sous l'hypothèse de petits déplacements, l'énergie élastique du ressort est une fonction quadratique de l'allongement égale à $\frac{k}{2}(u_{j+1} - u_j)^2$. L'énergie totale du système est égale à la somme de l'énergie élastique de chaque ressort et de l'énergie potentielle due aux forces appliquées aux masses, soit

$$J(u) = \sum_{j=0}^n \frac{k}{2} (u_{j+1} - u_j)^2 - \sum_{j=0}^{n+1} u_j f_j.$$

Si les deux extrémités sont fixées, l'énergie est à minimiser sur l'ensemble des vecteurs u tel que $u_0 = u_{n+1} = 0$. Si uniquement l'une des extrémités (par exemple x_0), l'espace de minimisation est l'ensemble de u tels que $u_0 = 0$. Si aucune extrémité n'est fixée, l'espace de minimisation n'a pas à être contraint. Par contre, l'existence d'un minimiseur n'est assurée que si la condition de compatibilité

$$\sum_{j=0}^{n+1} f_j = 0$$

est vérifiée.

Il y a une forte similitude entre le problème obtenu est la résolution de l'équation

$$-k\Delta u = f$$

par élément éléments finis \mathbb{P}_1 , qui consiste à minimiser l'énergie

$$I(u) = \frac{k}{2} \|\nabla u\|_{L^2(0,1)}^2 - \int_0^1 f(x)u(x) dx$$

sur l'espace de discrétisation V_h . Soit u_h un élément de V_h et U_h les coordonnées de u_h dans la base classique de V_h . On a alors

$$I(u_h) = \sum_{j=0}^n \frac{k}{2} (U_h^{j+1} - U_h^j)^2 \Delta x - \sum_{j=0}^{n+1} \left(\int_0^1 f(x) \phi_j(x) dx \right) U_h^j.$$

Si on utilise la formule des trapèzes afin d'évaluer l'intégrale apparaissant dans la définition de I , on obtient

$$I(u_h) = \sum_{j=0}^n \frac{k}{2} (U_h^{j+1} - U_h^j)^2 \Delta x - \sum_{j=0}^{n+1} f(x_j) \phi_j U_h^j \Delta x.$$

En posant $f_j = (\Delta x)^2 f(x_j)$, on retrouve l'expression J à un coefficient Δx près.

Exercice 6.2.5 Démontrer l'équivalent du Théorème 6.2.6 de convergence pour le problème de Neumann (6.1).

Correction. La démonstration est identique mot pour mot à celle effectuée dans le cas de conditions aux limites de Dirichlet. L'opérateur d'interpolation r_h utilisé est identique. Dans le cas de conditions aux limites de Dirichlet, on utilise en fait sa restriction à $H_0^1(\Omega)$ qui est à valeurs dans $H_0^1(\Omega) \cap V_h$, ce qui constitue l'unique différence.

Exercice 6.2.6 En généralisant les arguments précédents, démontrer le Théorème 6.2.14.

Correction. D'après le Lemme de C  a **6.1.2**, il existe une constante C ind  pendante de h telle que

$$\|u - u_h\|_{H^1(0,1)} \leq C \inf_{v_h \in V_{0h}} \|u - v_h\|_{H^1(0,1)},$$

o   V_{0h} est l'espace des   l  ments finis \mathbb{P}_2 nuls aux bords. Afin de majorer le terme de droite, on introduit l'op  rateur d'interpolation de l'espace des fonctions r  guli  res dans V_{0h} qui    v associe

$$r_h v = \sum_{j=1}^n v(x_j) \psi_j + \sum_{j=0}^n v(x_{j+1/2}) \psi_{j+1/2}.$$

Dans le cas $h = 1$, il existe des constantes C_0 et C_1 telles que

$$\|r_1 v - v\|_{L^2(0,1)} \leq C_0 \|v'''\|_{L^2(0,1)}$$

et

$$\|(r_1 v)' - v'\|_{L^2(0,1)} \leq C_1 \|v'''\|_{L^2(0,1)}.$$

Soit $h = 1/(n+1)$, on a

$$\|r_h v - v\|_{L^2(0,1)}^2 = \int_0^1 |(r_h v - v)(x)|^2 dx = \sum_{j=0}^n \int_{jh}^{(j+1)h} |(r_h v - v)(x)|^2 dx.$$

On pose $v_j(x) = v(h(j+x))$. Par changement de variable, on a

$$\|r_h v - v\|_{L^2(0,1)}^2 = h \sum_{j=0}^n \int_0^1 |(r_1 v_j - v_j)(x)|^2 dx \leq C_0^2 h \sum_{j=0}^n \|v_j'''\|_{L^2(0,1)}^2.$$

En effectuant de nouveau un changement de variable, on   tablit que

$$\|v_j'''\|_{L^2(0,1)}^2 = h^5 \int_{hj}^{(j+1)h} |v'''(x)|^2 dx.$$

Ainsi,

$$\|r_h v - v\|_{L^2(0,1)} \leq C_0 h^3 \|v'''\|_{L^2(0,1)}.$$

On proc  de de m  me pour   tablir que

$$\|(r_h v - v)'\|_{L^2(0,1)} \leq C_1 h^2 \|v'''\|_{L^2(0,1)}.$$

En rassemblant ces deux r  sultats, on en d  duit qu'il existe une constante C_2 telle que

$$\|r_h v - v\|_{H^1(0,1)} \leq C_2 h^2 \|v'''\|_{L^2(0,1)},$$

et d'apr  s le Lemme de C  a qu'il existe une constante C_3 telle que

$$\|u - u_h\|_{H^1(0,1)} \leq C_3 h^2 \|u'''\|_{L^2(0,1)}.$$

Exercice 6.2.7 Calculer explicitement la matrice de rigidité \mathcal{K}_h associée au problème consistant à trouver

$$u_h \in V_{0h} := \{v \in C^1([0, 1]) \text{ tel que } v_{[x_j, x_{j+1}]} \in \mathbb{P}_3 \text{ pour tout } 0 \leq j \leq n\} \cap H_0^2(\Omega)$$

tel que

$$\int_0^1 u_h''(x) v_h''(x) dx = \int_0^1 f(x) v_h(x) dx \quad \forall v_h \in V_{0h}. \quad (6.4)$$

Correction. La matrice de rigidité \mathcal{K}_h associée au problème (6.4) peut-être décomposée en $n \times n$ matrices blocs 2×2 , $A_{i,j}$, de sorte que si $U_h = (u_j, u_j')_{1 \leq j \leq n}$ et $V_h = (v_j, v_j')_{1 \leq j \leq n}$, on ait

$$V_h \cdot \mathcal{K}_h U_h = \sum_{i,j=1}^n (v_i, v_i') \cdot A_{i,j} (u_j, u_j').$$

Chaque matrice $A_{i,j}$ est définie par

$$A_{i,j} = \begin{pmatrix} \int_0^1 \phi_i(x) \phi_j(x) dx & \int_0^1 \phi_i(x) \psi_j(x) dx \\ \int_0^1 \psi_i(x) \phi_j(x) dx & \int_0^1 \psi_i(x) \psi_j(x) dx \end{pmatrix}$$

En comparant les supports des fonctions de bases, on constate que $A_{i,j} = 0$ dès que $|i - j| > 1$. Il suffit donc de déterminer les matrices $A_{i,j}$ pour $|i - j| \leq 1$. Il est donc nécessaire de déterminer les matrices $A_{i-1,i}$, $A_{i,i}$ et $A_{i+1,i}$. La forme bilinéaire de la formulation variationnelle étant symétrique, la matrice \mathcal{K}_h est elle même symétrique. On en déduit que la matrice $A_{i,i}$ est symétrique et que $A_{i-1,i} = A_{i,i-1}^T = A_{i+1,i}^T$. Nous n'avons donc que 7 coefficients à déterminer (4 pour la matrice $A_{i+1,i}$ et 3 pour la matrice $A_{i,i}$), soit

$$A_{i+1,i} = \begin{pmatrix} \int_0^1 \phi_{i+1}'' \phi_i'' dx & \int_0^1 \phi_{i+1}'' \psi_i'' dx \\ \int_0^1 \psi_{i+1}'' \phi_i'' dx & \int_0^1 \psi_{i+1}'' \psi_i'' dx \end{pmatrix}$$

et

$$A_{i,i} = \begin{pmatrix} \int_0^1 |\phi_i''|^2 dx & \int_0^1 \phi_i'' \psi_i'' dx \\ \int_0^1 \phi_i'' \psi_i'' dx & \int_0^1 |\psi_i''|^2 dx \end{pmatrix}$$

Afin de déterminer ces intégrales, on effectue le changement de variable $X = (x - x_i)/h$. On obtient en utilisant la parité des fonctions de base

$$A_{i+1,i} = h^{-3} \begin{pmatrix} \int_0^1 \phi''(X-1) \phi''(X) dX & \int_0^1 \phi''(X-1) \psi''(X) dX \\ \int_0^1 \psi''(X-1) \phi''(X) dX & \int_0^1 \psi''(X-1) \psi''(X) dX \end{pmatrix}$$

et

$$A_{i,i} = h^{-3} \begin{pmatrix} \int_0^1 |\phi''(X)|^2 dx & 0 \\ 0 & \int_0^1 |\psi''(X)|^2 dx \end{pmatrix}$$

Enfin, sur $[0, 1]$ on a

$$\phi''(X) = 12X - 6; \phi''(X - 1) = -12X + 6$$

et

$$\psi''(X) = 6X - 4; \psi''(X - 1) = 6X - 2.$$

Finalement, suite à un calcul de primitive élémentaire, il vient

$$A_{i+1,i} = h^{-3} \begin{pmatrix} -3 & -6 \\ 6 & 2 \end{pmatrix}$$

et

$$A_{i,i} = h^{-3} \begin{pmatrix} 6 & 0 \\ 0 & 8 \end{pmatrix}.$$

Ainsi,

$$\mathcal{K}_h = h^{-3} \begin{pmatrix} 6 & 0 & -3 & 6 & & & \\ 0 & 8 & -6 & 2 & & & \\ -3 & -6 & 6 & 0 & -3 & 6 & \\ 6 & 2 & 0 & 8 & -6 & 2 & \\ & & -3 & -6 & 6 & 0 & \ddots \\ & & 6 & 2 & 0 & 8 & \ddots \\ & & & & & \ddots & \ddots \end{pmatrix}$$

Exercice 6.3.1 Soit \mathcal{T}_h un maillage de $\bar{\Omega}$ pour Ω ouvert simplement connexe polygonal de \mathbb{R}^2 . On note n_t le nombre de triangles de \mathcal{T}_h , n_c le nombre de faces ou cotés des triangles (un coté commun à deux triangles n'est compté qu'une seule fois), n_s le nombre de sommets du maillage, et n_{0s} le nombre de sommets intérieurs du maillage (qui ne sont pas sur $\partial\Omega$). Démontrer les relations, dites d'Euler, $n_t + n_s = n_c + 1$ et $3n_t + n_s = 2n_c + n_{0s}$.

Correction. Plutôt que de vérifier les relations d'Euler, on se propose de les retrouver directement en effectuant un raisonnement par récurrence. On cherche à déterminer s'il existe un (ou plusieurs) vecteur $L \in \mathbb{Z}^4$ et un entier α tel que pour tout maillage d'un ouvert simplement connexe, $L \cdot x + \alpha = 0$ où $x = (n_t, n_c, n_s, n_{0s})$. Tout d'abord, la relation doit être vérifiée par le maillage trivial constitué d'un unique triangle. On a donc

$$L \cdot (1, 3, 3, 0) + \alpha = 0.$$

Considérons un maillage de Ω comportant plusieurs triangles. On choisit un chemin à l'intérieur de Ω constitué d'une succession d'arêtes reliant deux sommets distincts du bord. L'ouvert Ω étant simplement connexe, ce chemin sépare Ω en deux ouverts simplement connexes Ω_1 et Ω_2 . On note x_1 et x_2 les vecteurs composés du nombre de triangles, d'arêtes, de sommets et de sommets intérieurs de chacun des maillages. On note \tilde{n}_c et \tilde{n}_s les nombres de cotés et de sommet du chemin. On vérifie que

$$x = x_1 + x_2 + (0, -\tilde{n}_c, -\tilde{n}_s, \tilde{n}_s - 2).$$

De plus, $\tilde{n}_s = \tilde{n}_c + 1$. Si $L \cdot x_1 + \alpha = 0$ et $L \cdot x_2 + \alpha = 0$, on a $L \cdot x + \alpha = 0$ si et seulement si

$$\tilde{n}_c L \cdot (0, -1, -1, 1) + L \cdot (0, 0, -1, -1) - \alpha = 0.$$

Comme \tilde{n}_c est quelconque, on en déduit que les conditions nécessaires et suffisantes pour que la relation $L \cdot x + \alpha = 0$ soit vérifiée pour tout maillage sont

$$L \cdot (1, 3, 3, 0) = L \cdot (0, 0, 1, 1) = -\alpha \text{ et } L \cdot (0, -1, -1, 1) = 0,$$

ou encore $L \in \text{Vect}((-2, 1, 0, 1); (-1, 0, 1, 1))$ et $\alpha = -L \cdot (0, 0, 1, 1)$. Ainsi, on a uniquement deux relations d'Euler indépendantes :

$$-2n_t + n_c + n_{s0} = 1 \text{ et } -n_t + n_s + n_{s0} = 2.$$

On vérifie enfin que ces relations sont équivalentes à celles proposées par l'énoncé.

Exercice 6.3.2 Soit K un N -simplexe de sommets $(a_j)_{1 \leq j \leq N+1}$. Montrer que tout polynôme $p \in \mathbb{P}_1$ se met sous la forme

$$p(x) = \sum_{j=1}^{N+1} p(a_j) \lambda_j(x),$$

où les $(\lambda_j(x))_{1 \leq j \leq N+1}$ sont les coordonnées barycentriques de $x \in \mathbb{R}^N$.

Correction. Soit p un polynôme de degré un et K un N -simplexe de sommets $(a_j)_{1 \leq j \leq N+1}$. Comme $x = \sum_j \lambda_j(x) a_j$, et que l'application qui à x associe $p(x) - p(0)$ est linéaire, on a

$$p(x) - p(0) = \left(\sum_{j=1}^{N+1} \lambda_j(x) p(a_j) \right) - \left(\sum_{j=1}^{N+1} \lambda_j(x) \right) p(0).$$

Comme $\sum_j \lambda_j = 1$, on en déduit que

$$p(x) = \sum_{j=1}^{N+1} \lambda_j(x) p(a_j).$$

Exercice 6.3.3 Soit K un N -simplexe de sommets $(a_j)_{1 \leq j \leq N+1}$.

Soit $(a_{jj'})_{1 \leq j < j' \leq N+1}$ les points milieux des arêtes de K définis par leur coordonnées barycentriques

$$\lambda_j(a_{jj'}) = \lambda_{j'}(a_{jj'}) = \frac{1}{2}, \quad \lambda_l(a_{jj'}) = 0 \text{ pour } l \neq j, j'.$$

Vérifier que Σ_2 est précisément constitué des sommets et des points milieux des arêtes et que tout polynôme $p \in \mathbb{P}_2$ se met sous la forme

$$p(x) = \sum_{j=1}^{N+1} p(a_j) \lambda_j(x) (2\lambda_j(x) - 1) + \sum_{1 \leq j < j' \leq N+1} 4p(a_{jj'}) \lambda_j(x) \lambda_{j'}(x),$$

où les $(\lambda_j(x))_{1 \leq j \leq N+1}$ sont les coordonnées barycentriques de $x \in \mathbb{R}^N$.

Correction. On note $\mathbb{P}_{n,p}$ l'ensemble des polynômes de degré n de p variables. On introduit $Q \in \mathbb{P}_{2,N+1}$ le polynôme de \mathbb{R}^{N+1} , de degré 2 défini par

$$Q(X_1, \dots, X_{N+1}) = p \left(\sum_{j=1}^{N+1} X_j a_j \right).$$

Soit q_j et $q_{jj'}$ les éléments de $\mathbb{P}_{2,N+1}$ définis par

$$\begin{aligned} q_j(X_1, \dots, X_{N+1}) &= X_j(2X_j - 1) && \text{pour tout } 1 \leq j \leq N+1 \\ \text{et } q_{jj'}(X_1, \dots, X_{N+1}) &= 4X_j X_{j'} && \text{pour tout } 1 \leq j < j' \leq N+1 \end{aligned}$$

La famille $(q_j, q_{jj'})$ constituée de l'ensemble de ces polynômes est une famille libre. En effet, si (e_j) est la base canonique de \mathbb{R}^{N+1} , on a pour tout $1 \leq k \leq N+1$

$$q_j(e_k) = \delta_j^k \text{ et } q_{jj'}(e_k) = 0,$$

et pour tout couple (k, l) tel que $1 \leq k < l \leq N+1$,

$$q_{jj'}((e_k + e_l)/2) = \delta_j^k \delta_{j'}^l \text{ et } q_j(e_k) = 0.$$

On note R l'espace engendré par $(q_j, q_{jj'})$. On déduit également des relations précédentes que l'espace R est en somme directe avec l'ensemble des polynômes divisibles par

$$q_0(X_1, \dots, X_{N+1}) = 1 - \sum_{j=1}^{N+1} X_j,$$

soit

$$R \oplus q_0 \mathbb{P}_{1,N+1} \subset \mathbb{P}_{2,N+1}.$$

Enfin, notons que

$$\begin{aligned} \dim(\mathbb{P}_{2,N+1}) &= N+2 + (N+1)(N+2)/2, & \dim(R) &= N+1 + N(N+1)/2, \\ \text{et } \dim(\mathbb{P}_{1,N+1}) &= N+2. \end{aligned}$$

Ainsi,

$$\dim(R) + \dim(\mathbb{P}_{1,N+1}) = (N+1)(N+2)/2 + N+2 = \dim(\mathbb{P}_{2,N+1})$$

et

$$R \oplus q_0 \mathbb{P}_{1,N+1} = \mathbb{P}_{2,N+1}.$$

Il existe donc un unique couple $Q_1 \in R$ et $Q_2 \in q_0 \mathbb{P}_{1,N+1}$ tel que $Q = Q_1 + Q_2$. On peut aisément déterminer la décomposition de Q_1 dans la base $(q_j, q_{jj'})$. En effet,

$$Q_1 = \sum_{j=1}^{N+1} Q(e_j) q_j + \sum_{1 \leq j < j' \leq N+1} Q((e_j + e_{j'})/2) q_{jj'},$$

c'est à dire

$$Q_1 = \sum_{j=1}^{N+1} p(a_j)q_j + \sum_{1 \leq j < j' \leq N+1} p(a_{jj'})q_{j,j'},$$

Enfin, comme pour tout $x \in \mathbb{R}^N$, $\sum_j \lambda_j(x) = 1$, on a $p(x) = Q(\lambda_j(x)) = Q_1(\lambda_j(x))$, d'où

$$p(x) = \sum_{j=1}^{N+1} p(a_j)\lambda_j(x)(2\lambda_j(x) - 1) + \sum_{1 \leq j < j' \leq N+1} 4p(a_{jj'})\lambda_j(x)\lambda_{j'}(x).$$

Exercice 6.3.4 Soit \mathcal{T}_h un maillage de $\bar{\Omega}$ pour Ω ouvert simplement connexe polygonal de \mathbb{R}^2 . On note n_t le nombre de triangles de \mathcal{T}_h , n_c le nombre de faces ou cotés des triangles (un coté commun à deux triangles n'est compté qu'une seule fois), n_s le nombre de sommets du maillage, et n_{0s} le nombre de sommets intérieurs du maillage. Montrer que les dimensions de l'espace V_h d'éléments finis de Lagrange d'ordre k et de son sous espace V_{0h} des fonctions s'annulant sur le bord du domaine sont

$$\dim V_h = \frac{k(k-1)}{2}n_t + kn_s - k + 1, \quad \dim V_{0h} = \frac{k(k+1)}{2}n_t - kn_s + k + 1.$$

Correction. Pour un treillis d'ordre k , on compte $(k+1)(k+2)/2$ éléments, dont $3k$ sur le bord du triangle. En particulier, un treillis d'ordre k compte $(k+1)(k+2)/2 - 3k = (k-1)(k-2)/2$ points "internes", $3(k-1)$ points situés à l'intérieur des arêtes et 3 aux sommets.

La dimension de V_h est égale au nombre total de degrés de liberté. A l'intérieur de chaque triangle, on compte $(k-1)(k-2)/2$ degrés de liberté soit $n_t(k-1)(k-2)/2$, auxquels il faut ajouter les degrés de liberté situés à l'intérieur des arêtes, soit $(k-1)n_c$ degrés de liberté et les n_s sommets du maillage. Au total,

$$\dim(V_h) = \frac{(k-1)(k-2)}{2}n_t + (k-1)n_c + n_s$$

D'après la première formule d'Euler (voir Exercice 6.3.1), $n_c = n_t + n_s - 1$. Ainsi,

$$\dim(V_h) = \frac{(k-1)(k-2)}{2}n_t + (k-1)n_t + kn_s + (1-k) = \frac{(k-1)k}{2}n_t + kn_s + 1 - k.$$

Le nombre de degrés de liberté de V_{0h} est égal à celui de V_h , auquel il faut soustraire les degrés de liberté situés sur le bord $\partial\Omega$ du domaine qui en compte $k(n_s - n_{0s})$. On a donc

$$\dim(V_{0h}) = \frac{(k-1)k}{2}n_t + kn_s + 1 - k - k(n_s - n_{0s}) = \frac{(k-1)k}{2}n_t + kn_{0s} + 1 - k.$$

D'après les formules d'Euler,

$$\begin{aligned} n_{0s} &= 3n_t + n_s - 2n_c \\ &= 3n_t + n_s - 2(n_t + n_s - 1) \\ &= n_t - n_s + 2. \end{aligned}$$

Ainsi

$$\dim(V_{0h}) = \frac{(k-1)k}{2}n_t + kn_t - kn_s + 1 + k = \frac{(k+1)k}{2}n_t - kn_s + 1 + k.$$

Exercice 6.3.5 Démontrer la formule (6.43) en dimension $N = 2$, c'est à dire

$$\int_K \lambda_1(x)^{\alpha_1} \lambda_2(x)^{\alpha_2} \lambda_3(x)^{\alpha_3} dx = 2\text{Aire}(K) \frac{\alpha_1! \alpha_2! \alpha_3!}{(\alpha_1 + \alpha_2 + \alpha_3 + 2)!}, \quad (6.5)$$

où K est un simplexe de \mathbb{R}^2 , $\lambda_i(x)$ sont les coordonnées barycentriques de x et α_i des entiers naturels.

Correction. On pose

$$I = \int_K \lambda_1^{\alpha_1}(x) \lambda_2^{\alpha_2}(x) \lambda_3^{\alpha_3}(x) dx.$$

Soit a_i les sommets de K , et F l'application de

$$S = \{(\lambda_1, \lambda_2) \in \mathbb{R}_+^2 : \lambda_1 + \lambda_2 \leq 1\}$$

à valeurs dans K définie par

$$F(\lambda_1, \lambda_2) = \lambda_1 a_1 + \lambda_2 a_2 + (1 - \lambda_1 - \lambda_2) a_3.$$

L'application F est un difféomorphisme de S dans K . En effectuant le changement de variables $x = F(\lambda_1, \lambda_2)$ dans l'expression de I , on obtient

$$I = 2\text{Aire}(K) \int_S \lambda_1^{\alpha_1} \lambda_2^{\alpha_2} \lambda_3^{\alpha_3} d\lambda_1 d\lambda_2, \quad (6.6)$$

avec $\lambda_3 = 1 - (\lambda_1 + \lambda_2)$. Il reste donc à calculer l'intégrale figurant dans le terme de droite.

$$\int_S \lambda_1^{\alpha_1} \lambda_2^{\alpha_2} \lambda_3^{\alpha_3} d\lambda_1 d\lambda_2 = \int_0^1 \lambda_1^{\alpha_1} \left(\int_0^{1-\lambda_1} \lambda_2^{\alpha_2} (1 - \lambda_1 - \lambda_2)^{\alpha_3} d\lambda_2 \right) d\lambda_1.$$

On effectue le changement de variable $\lambda_2 = (1 - \lambda_1)t$ dans l'intégrale selon λ_2 .

$$\int_S \lambda_1^{\alpha_1} \lambda_2^{\alpha_2} \lambda_3^{\alpha_3} d\lambda_1 d\lambda_2 = \int_0^1 \lambda_1^{\alpha_1} (1 - \lambda_1)^{\alpha_2 + \alpha_3 + 1} d\lambda_1 \int_0^1 t^{\alpha_2} (1 - t)^{\alpha_3} dt$$

Par intégration par parties successives, on montre que

$$\int_0^1 t^n (1 - t)^m dt = \frac{n!m!}{(n + m + 1)!}.$$

Ainsi,

$$\int_S \lambda_1^{\alpha_1} \lambda_2^{\alpha_2} \lambda_3^{\alpha_3} d\lambda_1 d\lambda_2 = \frac{\alpha_1! (\alpha_2 + \alpha_3 + 1)!}{(\alpha_1 + \alpha_2 + \alpha_3 + 2)!} \frac{\alpha_2! \alpha_3!}{(\alpha_2 + \alpha_3 + 1)!} = \frac{\alpha_1! \alpha_2! \alpha_3!}{(\alpha_1 + \alpha_2 + \alpha_3 + 2)!}.$$

qui combinée avec (6.6) nous donne (6.5).

Exercice 6.3.6 Montrer que les formules de quadrature

$$\int_K \psi(x) dx \approx \text{Volume}(K) \psi(a_0), \quad (6.7)$$

avec $a_0 = (N+1)^{-1} \sum_{i=1}^{N+1} a_i$, le barycentre de K , et

$$\int_K \psi(x) dx \approx \frac{\text{Volume}(K)}{N+1} \sum_{i=1}^{N+1} \psi(a_i). \quad (6.8)$$

sont exactes pour $\psi \in \mathbb{P}_1$.

Correction. Soit p un polynôme de degré 1, il existe q polynôme de degré 1 en λ tel que $q(\lambda(x)) = p(x)$. Or $\int_K 1 dx = \text{Volume}(K)$ et $\int_K \lambda_k dx = \frac{\text{Volume}(K)}{N+1} \sum_i \lambda_k(a_i)$. On en déduit donc que

$$\int_K q(\lambda(x)) = \frac{\text{Volume}(K)}{N+1} \sum_i q(\lambda(a_i)),$$

et que la formule (6.8) est exacte pour les polynôme de degré 1. De plus, comme p est linéaire, $p(a_0) = 1/(N+1) \sum_i p(a_i)$, ce qui établit l'exactitude de la formule (6.7)

Exercice 6.3.7 Soit K un triangle de \mathbb{R}^2 de sommets $(a_i)_{1 \leq i \leq 3}$ et de barycentre a_0 . Soit $(a_{ij})_{1 \leq i < j \leq 3}$ les points milieux des segments d'extrémités a_i, a_j . Montrer que la formule de quadrature

$$\int_K \psi(x) dx \approx \frac{\text{Aire}(K)}{3} \sum_{1 \leq i < j \leq 3} \psi(a_{ij})$$

est exacte pour $\psi \in \mathbb{P}_2$, tandis que la formule

$$\int_K \psi(x) dx \approx \frac{\text{Aire}(K)}{60} \left(3 \sum_{i=1}^3 \psi(a_i) + 8 \sum_{1 \leq i < j \leq 3} \psi(a_{ij}) + 27 \psi(a_0) \right)$$

est exacte pour $\psi \in \mathbb{P}_3$.

Correction. Comme précédemment, il suffit de vérifier l'exactitude des formules pour les polynômes de la forme

$$p(x) = q(\lambda_1(x), \lambda_2(x), \lambda_3(x)),$$

où (λ_i) sont les coordonnées barycentriques de x et q est un polynôme de trois variables de degré 2 ou 3. En d'autres termes, il s'agit de vérifier que pour tout polynôme q de trois variables et de degré deux

$$\int_K q(\lambda_1(x), \lambda_2(x), \lambda_3(x)) dx = T_2(q), \quad (6.9)$$

où

$$T_2(q) = \frac{\text{Aire}(K)}{3} \sum_{1 \leq i < j \leq 3} q((e_i + e_j)/2)$$

et que pour tout polynôme q de trois variables et de degré trois,

$$\int_K q(\lambda_1(x), \lambda_2(x), \lambda_3(x)) dx = T_3(q), \quad (6.10)$$

où

$$T_3(q) = \frac{\text{Aire}(K)}{60} \left(3 \sum_{i=1}^3 q(e_i) + 8 \sum_{1 \leq i < j \leq 3} q((e_i + e_j)/2) + 27q((e_1 + e_2 + e_3)/3) \right).$$

On note

$$S(\alpha_1, \alpha_2, \alpha_3) = \int_K \lambda_1^{\alpha_1} \lambda_2^{\alpha_2} \lambda_3^{\alpha_3} dx = 2\text{Aire}(K) \frac{\alpha_1! \alpha_2! \alpha_3!}{(\alpha_1 + \alpha_2 + \alpha_3 + 2)!}$$

Les équations (6.9) et (6.10) sont linéaires par rapport au polynôme q . Il suffit donc de les établir pour une base de l'ensemble des polynômes de trois variables de degré deux et trois respectivement. On peut par exemple vérifier que, pour tout $\alpha_i \in \mathbb{N}$ tel que $\sum_{i=1}^3 \alpha_i \leq 2$,

$$S(\alpha_1, \alpha_2, \alpha_3) = T_2(X_1^{\alpha_1} X_2^{\alpha_2} X_3^{\alpha_3})$$

et pour tout $\alpha_i \in \mathbb{N}$ tel que $\sum_{i=1}^3 \alpha_i \leq 3$,

$$S(\alpha_1, \alpha_2, \alpha_3) = T_3(\lambda_1^{X_1} X_2^{\alpha_2} X_3^{\alpha_3}).$$

Par des considérations d'invariance, le nombre de vérifications se limite à 4 cas pour la première formule et 7 cas pour la seconde, et on a

$$\begin{aligned} S(0, 0, 0) &= \text{Aire}(K) &= T_2(1) &= T_3(1), \\ S(1, 0, 0) &= \text{Aire}(K)/3 &= T_2(X_1) &= T_3(X_1), \\ S(1, 1, 0) &= \text{Aire}(K)/12 &= T_2(X_1 X_2) &= T_3(X_1 X_2), \\ S(2, 0, 0) &= \text{Aire}(K)/6 &= T_2(X_1^2) &= T_3(X_1^2), \\ S(1, 1, 1) &= \text{Aire}(K)/60 &= T_3(X_1 X_2 X_3), \\ S(2, 1, 0) &= \text{Aire}(K)/30 &= T_3(X_1^2 X_2), \\ S(3, 0, 0) &= \text{Aire}(K)/60 &= T_3(X_1^3). \end{aligned}$$

Exercice 6.3.8 Soit $(b_i)_{1 \leq i \leq I}$ des points d'un N -simplexe K et $(\omega_i)_{1 \leq i \leq I}$ des poids réels. Soit une formule de quadrature

$$\int_K \psi(x) dx \approx \text{Volume}(K) \sum_{i=1}^I \omega_i \psi(b_i)$$

qui soit exacte pour $\psi \in \mathbb{P}_k$. Montrer que, pour une fonction régulière ψ , on a

$$\frac{1}{\text{Volume}(K)} \int_K \psi(x) dx = \sum_{i=1}^I \omega_i \psi(b_i) + \mathcal{O}(h^{k+1}),$$

où h est le diamètre de K .

Correction. Soit ψ une fonction de classe C^{k+1} . En effectuant un développement de Taylor, il existe une constante C telle que pour tout élément a du domaine (borné) considéré, il existe un polynôme T_a dépendant de ψ , de degré au plus k tel que

$$|\psi(a+u) - T_a(u)| \leq C|u|^{k+1}.$$

Considérons un simplexe K de centre de gravité a_0 , par intégration de la formule précédente sur les éléments u tels que $a_0 + u \in K$ (en particulier, $|u| < h$), on obtient que

$$\left| \int_K \psi dx - \int_K T_{a_0}(u) dx \right| \leq C \text{Vol}(K) h^{k+1}.$$

La formule de quadrature étant exacte pour les polynômes de degré inférieur ou égal à k , on a donc

$$\left| \int_K \psi dx - \text{Vol}(K) \sum_i \omega_i T_{a_0}(b_i - a_0) \right| \leq C \text{Vol}(K) h^{k+1}.$$

En utilisant à nouveau le développement de Taylor de ψ en a_0 , on en déduit que

$$\left| \int_K \psi dx - \text{Vol}(K) \sum_i \omega_i \psi(b_i) \right| \leq C' \text{Vol}(K) h^{k+1}$$

où C' est une constante indépendante de h , ce qui achève la démonstration.

Exercice 6.3.9 On considère le carré $\Omega =]-1, +1[^2$ maillé suivant la Figure 6.1. Calculer la matrice de rigidité \mathcal{K}_h des éléments finis \mathbb{P}_1 appliqués au Laplacien avec condition aux limites de Neumann (on utilisera les symétries du maillage).

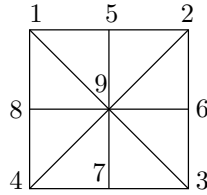


FIG. 6.1 – Exemple de maillage et de numérotation des nœuds.

Correction. On note V_h l'espace des éléments finis \mathbb{P}_1 associé au maillage 6.1. L'espace V_h est de dimension 9. Pour tout $i \in \{1, \dots, 9\}$, on note ϕ_i la fonction de base associée au i ème nœud (on utilise la numérotation des nœuds indiquée sur la figure). En d'autres termes, ϕ_i est l'unique élément de V_h tel que $\phi_i(x_j) = \delta_{ij}$ pour tout indice $j \in \{1, \dots, 9\}$. La matrice de rigidité associée à la résolution du Laplacien est définie pour tout couple d'indices i et j par

$$(\mathcal{K}_h)_{i,j} = \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j dx.$$

On a donc 81 coefficients à déterminer ! Cependant, dès que ϕ_i et ϕ_j sont à support disjoint, $(\mathcal{K}_h)_{i,j} = 0$. Enfin, en utilisant les symétries du maillage, on constate qu'il suffit de calculer six coefficients de la matrice de rigidité, les autres s'en déduisant aisément. En l'occurrence, on doit calculer $(\mathcal{K}_h)_{1,1}$, $(\mathcal{K}_h)_{1,5}$, $(\mathcal{K}_h)_{1,9}$, $(\mathcal{K}_h)_{5,5}$, $(\mathcal{K}_h)_{5,9}$ et $(\mathcal{K}_h)_{9,9}$. Le gradient des fonctions de base ϕ_i est constant sur chaque maille, qui sont toutes de même aire $1/2$. Le calcul de nos 9 coefficients est donc aisé et

$$(\mathcal{K}_h)_{1,1} = 1, (\mathcal{K}_h)_{1,5} = -1/2, (\mathcal{K}_h)_{1,9} = 0, (\mathcal{K}_h)_{5,5} = 2, (\mathcal{K}_h)_{5,9} = -1, (\mathcal{K}_h)_{9,9} = 4.$$

En rassemblant ces résultats, on obtient

$$\mathcal{K}_h = \begin{pmatrix} 1 & 0 & 0 & 0 & -1/2 & 0 & 0 & -1/2 & 0 \\ 0 & 1 & 0 & 0 & -1/2 & -1/2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1/2 & -1/2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1/2 & -1/2 & 0 \\ -1/2 & -1/2 & 0 & 0 & 2 & 0 & 0 & 0 & -1 \\ 0 & -1/2 & -1/2 & 0 & 0 & 2 & 0 & 0 & -1 \\ 0 & 0 & -1/2 & -1/2 & 0 & 0 & 2 & 0 & -1 \\ -1/2 & 0 & 0 & -1/2 & 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & 4 \end{pmatrix}$$

Exercice 6.3.10 Appliquer la méthode des éléments finis \mathbb{P}_1 au problème de Dirichlet

$$\begin{cases} -\Delta u = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega, \end{cases} \quad (6.11)$$

dans le carré $\Omega =]0, 1[^2$ avec le maillage triangulaire uniforme de la Figure 6.12. Montrer que la matrice de rigidité \mathcal{K}_h est la même matrice que celle que l'on obtiendrait par application de la méthode des différences finies (à un facteur multiplicatif h^2 près), mais que le second membre b_h est différent.

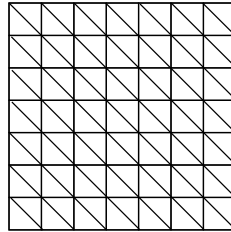
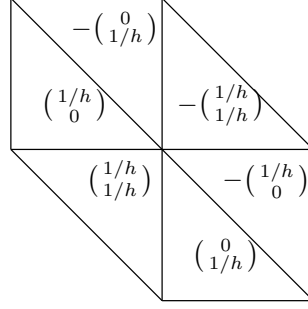


FIG. 6.2 – Maillage triangulaire uniforme d'un carré

Correction. On note n le nombre de mailles situées sur l'un des bords du domaine. Soit $h = 1/(n + 1)$, la taille d'une maille. On note $x_{i,j} = (x_i, x_j)$ les sommets du maillage où $x_i = ih$ (on a $0 < i, j < n$). On numérote les nœuds du maillage ligne par ligne. En d'autres termes, on pose $a_{i+jn} = x_{i,j}$ pour tout $0 < i, j < n$. Enfin, on note ϕ_k la fonction de base \mathbb{P}_1 associée au nœud a_k . La figure ci-dessous représente le gradient d'une fonction de base ϕ_k (constant sur chaque triangle).



On cherche à calculer $\mathcal{A}_{hk,l} = \int_{\Omega} \nabla \phi_k \cdot \nabla \phi_l dx$. Si $k = l$,

$$\mathcal{A}_{hk,k} = \int_{\Omega} |\nabla \phi_k|^2 dx.$$

Le gradient $\nabla \phi_k$ est nul sur tout Ω à l'exclusion des 6 triangles contenant a_k . Sur chacun d'entre eux, $|\nabla \phi_k|^2$ est constant, égale à $1/h^2$ sur quatre d'entre eux, $2/h^2$ sur les deux autres. Enfin, l'aire des triangles du maillage étant égale à $h^2/2$,

$$\mathcal{A}_{hk,k} = 4.$$

Si a_k et a_l sont des nœuds voisins, c'est à dire si $k = l + 1$, $k = l - 1$, $k = l + n - 1$, $k = l + n$, $k = l + n - 1$, $k = l - N$ ou $k = l - n + 1$, les supports de ϕ_k et ϕ_l ne sont pas disjoints. Cependant, le terme $\mathcal{A}_{hk,l}$ est nul dans les cas $k = l - n + 1$ et $k = l + n - 1$ (les gradients des fonctions ϕ_k et ϕ_l sont orthogonaux). Dans les autres cas, on a

$$\mathcal{A}_{hk,l} = -1.$$

En d'autres termes, on a

$$\mathcal{A}_h = \begin{pmatrix} D & E & & 0 \\ E & D & E & \\ & \cdot & \cdot & \cdot \\ & & E & D & E \\ 0 & & & E & D \end{pmatrix}$$

où E et D sont les matrices $(n-1) \times (n-1)$

$$D = \begin{pmatrix} 4 & -1 & & 0 \\ -1 & 4 & -1 & \\ & \cdot & \cdot & \cdot \\ & & -1 & 4 & -1 \\ 0 & & & -1 & 4 \end{pmatrix} \quad E = \begin{pmatrix} -1 & 0 & & \\ 0 & -1 & 0 & \\ & \cdot & \cdot & \cdot \\ & & 0 & -1 & 0 \\ & & & 0 & -1 \end{pmatrix}.$$

On obtient donc en effet la matrice issue de la méthode des différences finies multipliée par h^2 . Cependant, le second membre du système linéaire obtenu diffère, car

$$(b_h)_k = \int_{\Omega} f \phi_k dx \neq h^2 f(a_k).$$

Exercice 6.3.11 On reprend les notations de l'Exercice 6.3.10. On note n le nombre de points du maillage sur un coté du carré (supposé être le même pour chaque coté). On numérote "ligne par ligne" les nœuds du maillage (ou les degrés de liberté). Montrer que la matrice de rigidité \mathcal{K}_h des éléments finis \mathbb{P}_1 est de taille de l'ordre de n^2 et de largeur de bande de l'ordre de $2n$ (pour n grand).

Montrer que la même méthode et le même type de maillage pour le cube $\Omega =]0, 1[^3$ conduisent à une matrice de taille de l'ordre de n^3 et de largeur de bande de l'ordre de $2n^2$ (où n est le nombre de nœuds le long d'une arête du cube Ω).

Correction. La taille de la matrice \mathcal{K}_h est exactement n^2 , tandis que sa demi-largeur de bande est n , en effet, dès que $|k - l| > n$, $\mathcal{K}_{hk,l} = 0$. Dans le cas du cube, on note $a_{i+jn+kn^2} = (x_i, x_j, x_k)$ les nœuds du maillage, où $x_i = i/(n+1)$. Le nombre de degré de liberté est donc égal à n^3 . Enfin, si $|k - l| > n^2 + n$, le support des fonctions test ϕ_k et ϕ_l sont disjoints. Ainsi, la matrice du système obtenu à une demi-largeur de bande de l'ordre de n^2 pour n grand.

Exercice 6.3.12 On dit qu'une matrice carrée réelle $B = (b_{ij})_{1 \leq i,j \leq n}$ est une M-matrice si, pour tout i ,

$$b_{ii} > 0, \quad \sum_{k=1}^n b_{ik} > 0, \quad b_{ij} \leq 0 \quad \forall j \neq i.$$

Montrer que toute M-matrice est inversible et que tous les coefficients de son inverse sont positifs ou nuls.

Correction. Soit B une M-matrice et $X \in \mathbb{R}^N$ tel que $BX = Y \geq 0$. Introduisons l'indice i_0 tel que

$$X_{i_0} = \min_{1 \leq i \leq N} X_i.$$

On a alors

$$B_{i_0 i_0} X_{i_0} + \sum_{j \neq i_0} B_{i_0 j} X_j = Y_{i_0} \geq 0,$$

d'où

$$\left(\sum_{j=1}^N B_{i_0 j} \right) X_{i_0} \geq \sum_{j \neq i_0} B_{i_0 j} (X_{i_0} - X_j).$$

D'après la définition de i_0 , le second membre de cette inégalité est positif ou nul. Comme $\sum_{j=1}^N B_{i_0 j} > 0$, on en déduit que $X_{i_0} \geq 0$ et donc que $X \geq 0$. Enfin, B est inversible car injective. En effet, si $BX = 0$, $BX \geq 0$ et $B(-X) \geq 0$, d'où $X \geq 0$ et $-X \geq 0$, c'est à dire $X = 0$. Comme $BX \geq 0$ implique $X \geq 0$, les coefficients de la matrice B^{-1} sont positifs.

Exercice 6.3.13 On se place en dimension $N = 2$. Soit u_h la solution approchée du problème de Dirichlet (6.11) obtenue par la méthode des éléments finis \mathbb{P}_1 . On suppose que tous les angles des triangles $K_i \in \mathcal{T}_h$ sont inférieurs ou égaux à $\pi/2$. Montrer que $u_h(x) \geq 0$ dans Ω si $f(x) \geq 0$ dans Ω . Indication : on montrera que, pour tout $\epsilon > 0$, $\mathcal{K}_h + \epsilon \text{Id}$ est une M-matrice, où \mathcal{K}_h est la matrice de rigidité.

Correction. Soit \mathcal{K}_h la matrice du système issu de la méthode des éléments finis, avec conditions de Dirichlet. Il suffit de prouver que pour tout $\varepsilon > 0$, $\mathcal{K}_h + \varepsilon I$ est une M-matrice. En effet, dans ce cas et d'après l'exercice précédent,

$$(\mathcal{K}_h + \varepsilon I)^{-1} \geq 0.$$

L'application qui à une matrice associe son inverse étant continue sur l'ensemble des matrices inversibles, on en déduit que

$$\mathcal{K}_h^{-1} \geq 0.$$

Tout d'abord, il est clair que

$$(\mathcal{K}_h)_{ii} > 0 \tag{6.12}$$

pour tout i . Considérons ensuite deux sommets distincts S_i et S_j communs à un triangle T_k du maillage. Le gradient de ϕ_i est orthogonal au coté du triangle T_k opposé à S_i . Il en est de même pour $\nabla \phi_j$. Il découle alors des hypothèses effectuées sur le maillage que

$$\nabla \phi_i \cdot \nabla \phi_j \leq 0$$

sur T_k . Le raisonnement étant valable sur tous les triangles du maillage, on en déduit que

$$(\mathcal{K}_h)_{ij} = \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j \, dx \leq 0 \text{ pour tout } i \neq j. \tag{6.13}$$

Soit n_0 le nombre de nœuds du maillage situés à l'intérieur du domaine Ω et n le nombre de nœuds total, on numérote les nœuds S_i du maillage de sorte que $S_i \in \partial\Omega$ pour $i > n_0$. Comme

$$1 = \sum_{j=1}^n \phi_j,$$

pour tout i , $0 < i \leq n_0$,

$$0 = \int_{\Omega} \nabla \phi_i \cdot \nabla 1 \, dx = \sum_{j=1}^n \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j \, dx.$$

Ainsi,

$$\sum_{j=1}^{n_0} \nabla \phi_i \cdot \nabla \phi_j \, dx = - \sum_{j=n_0+1}^n \nabla \phi_i \cdot \nabla \phi_j \, dx.$$

Or on a prouvé précédemment que le second membre est positif. On a donc montré que

$$\sum_{i=1}^{n_0} (\mathcal{K}_h)_{ij} \geq 0. \tag{6.14}$$

De (6.12), (6.13), (6.14), on déduit que $\mathcal{K}_h + \varepsilon I$ est une M-matrice pour tout $\varepsilon > 0$, ce qui achève la démonstration.

Exercice 6.3.14 Appliquer la méthode des éléments finis \mathbb{P}_k au système de l'élasticité (5.56). Montrer en particulier que la matrice de rigidité \mathcal{K}_h est dans ce cas d'ordre Nn_{dl} où N est la dimension d'espace et n_{dl} est le nombre de nœuds de degrés de liberté.

Correction. La formulation faible de l'élasticité linéarisée consiste à déterminer $u \in H_0^1(\Omega)^N$ tel que

$$\int_{\Omega} (2\mu e(u) \cdot e(v) + \lambda(\operatorname{div} u)(\operatorname{div} v)) \, dx = \int_{\Omega} f \cdot v \, dx \text{ pour tout } v \in H_0^1(\Omega)^N.$$

Soit \mathcal{T}_h un maillage régulier de Ω , on introduit les espaces discrets

$$V_h = \{u \in C(\Omega; \mathbb{R})^N : u|_K \in \mathbb{P}_k \text{ pour tout } K \in \mathcal{T}_h\}$$

et

$$V_{0h} = \{u \in V_h : u = 0 \text{ sur } \partial\Omega\}.$$

Soit $(\phi_i)_{i=1, n_{dl}}$ les fonctions de base associées au degré de liberté du treillis d'ordre k du maillage \mathcal{T}_h . L'approximation variationnelle du problème (5.56) par la méthode des éléments finis \mathbb{P}_k consiste à déterminer $u \in V_{0h}$ tel que

$$\int_{\Omega} (2\mu e(u) \cdot e(v) + \lambda(\operatorname{div} u)(\operatorname{div} v)) \, dx = \int_{\Omega} f \cdot v \, dx \text{ pour tout } v \in V_{0h},$$

c'est à dire à résoudre le système

$$\mathcal{K}_h U_h = b_h$$

où

$$(\mathcal{K}_h)_{ij} = \int_{\Omega} (2\mu e(\phi_i) \cdot e(\phi_j) + \lambda(\operatorname{div} \phi_i)(\operatorname{div} \phi_j)) \, dx$$

et

$$(b_h)_i = \int_{\Omega} f \cdot \phi_i \, dx.$$

L'existence d'une solution à ce problème est évidente par application du théorème de Lax-Milgram. Enfin, la dimension de l'espace V_{0h} est égale à Nn_{dl} où n_{dl} est le nombre de nœuds de degrés de liberté.

Exercice 6.3.15 Expliciter la matrice de rigidité \mathcal{K}_h obtenue par application de la méthode des éléments finis \mathbb{P}_k au problème de Neumann

$$\begin{cases} -\Delta u + au = f & \text{dans } \Omega \\ \frac{\partial u}{\partial n} = g & \text{sur } \partial\Omega, \end{cases} \quad (6.15)$$

avec $f \in L^2(\Omega)$, $g \in L^2(\partial\Omega)$, et $a \in L^\infty(\Omega)$ tel que $a(x) \geq a_0 > 0$ p.p. dans Ω .

Correction. L'espace d'approximation issu de la méthode des éléments finis \mathbb{P}_k , associé au problème de Neumann (6.15) est basée sur l'espace discret

$$V_h = \{u \in C(\Omega; \mathbb{R}) : u|_K \in \mathbb{P}_k \text{ pour tout } K \in \mathcal{T}_h\}.$$

Soit $(\phi_i)_{i=1,nd}$ les fonctions de base associées au degré de liberté du treillis d'ordre k du maillage \mathcal{T}_h . L'approximation variationnelle consiste à résoudre le système

$$\mathcal{K}_h U_h = b_h,$$

où

$$(\mathcal{K}_h)_{ij} = \int_{\Omega} (\nabla \phi_i \cdot \nabla \phi_j + a \phi_i \phi_j) \, dx$$

et

$$(b_h)_i = \int_{\Omega} f \phi_i \, dx.$$

Exercice 6.3.16 Montrer que la matrice de rigidité \mathcal{K}_h obtenue par application de la méthode des éléments finis \mathbb{P}_k au problème de convection-diffusion de l'Exercice 5.2.2 est inversible mais pas symétrique.

Correction. L'espace d'approximation variationnelle du problème de convection diffusion de l'Exercice 5.2.2 est

$$V_{0h} = \{u \in C(\Omega; \mathbb{R})^N : u_i|_K \in \mathbb{P}_k \text{ pour tout } K \in \mathcal{T}_h, u = 0 \text{ sur } \partial\Omega\}.$$

Soit $(\phi_i)_{i=1,nd}$ les fonctions de base associées au degré de liberté du treillis d'ordre k du maillage \mathcal{T}_h . L'approximation variationnelle consiste à résoudre le système

$$\mathcal{K}_h U_h = b_h,$$

où

$$(\mathcal{K}_h)_{ij} = \int_{\Omega} (\nabla \phi_i \cdot \nabla \phi_j + (V \cdot \nabla \phi_i) \phi_j) \, dx$$

et

$$(b_h)_i = \int_{\Omega} f \phi_i \, dx.$$

On rappelle que la divergence de V est supposée nulle. Ainsi, pour tout u_h et v_h appartenant à V_{0h} ,

$$\int_{\Omega} (V \cdot \nabla u_h) v_h \, dx = - \int_{\Omega} ((\operatorname{div} V) v_h u_h + (V \cdot \nabla v_h) u_h) \, dx = - \int_{\Omega} (V \cdot \nabla v_h) u_h \, dx.$$

En particulier, la matrice \mathcal{K}_h est en général non symétrique, sauf si tous les termes $\int_{\Omega} (V \cdot \nabla \phi_i) \phi_k \, dx$ sont nuls. Enfin, la matrice \mathcal{K}_h est inversible car injective, en effet,

$$\langle \mathcal{K}_h U_h, U_h \rangle = \int_{\Omega} \nabla u_h \cdot \nabla u_h + (V \cdot \nabla u_h) u_h \, dx = \int_{\Omega} \nabla u_h \cdot \nabla u_h \, dx$$

et $\langle \mathcal{K}_h U_h, U_h \rangle > 0$ si $U_h \neq 0$.

Exercice 6.3.17 On se propose de résoudre numériquement l'équation des plaques (5.19) par une méthode d'éléments finis (de type Hermite) en dimension $N = 2$. Pour un maillage triangulaire \mathcal{T}_h on introduit l'espace discret

$$V_h = \{v \in C^1(\overline{\Omega}) \text{ tel que } v|_{K_i} \in \mathbb{P}_5 \text{ pour tout } K_i \in \mathcal{T}_h\}.$$

Montrer que tout polynôme $p \in \mathbb{P}_5$ est caractérisé de manière unique sur un triangle K par les 21 valeurs réelles suivantes

$$v(a_j), \nabla v(a_j), \nabla \nabla v(a_j), \frac{\partial p(b_j)}{\partial n} \quad j = 1, 2, 3, \quad (6.16)$$

où (a_1, a_2, a_3) sont les sommets de K , (b_1, b_2, b_3) les milieux des cotés de K , tandis que $\partial p(b_j)/\partial n$ désigne la dérivée normale au coté de b_j . Montrer que V_h est un sous-espace de $H^2(\Omega)$ dont les éléments v sont caractérisés de manière unique par les valeurs (6.16) pour chaque sommet et milieu d'arête du maillage. En déduire une méthode d'éléments finis (dite d'Argyris) pour résoudre (5.19).

Correction.

1. Unisolvance (équivalent du Lemme 6.3.3)

On considère l'application qui à un élément de \mathbb{P}_5 associe les 21 valeurs (6.16). Comme \mathbb{P}_5 est un espace de dimension 21, il suffit de montrer que cette application est injective afin de prouver qu'elle est bijective. Enfin, quitte à effectuer un changement de variables par une application affine, on peut se contenter de considérer le cas d'un triangle équilatéral tel que $a_1 = (-1, 0)$, $a_2 = (1, 0)$. Soit $p \in \mathbb{P}_5$ annulant toutes les valeurs (6.16). Montrons que p est le polynôme nul. On pose $q_1(x_1) = p(x_1, 0)$ et $q_2(x_1) = \partial p / \partial x_2(x_1, 0)$. On vérifie que

$$q_1(\pm 1) = q_1'(\pm 1) = q_1''(\pm 1) = 0.$$

Comme q_1 est un polynôme de degré au plus 5, on en déduit que $q_1 = 0$. Ainsi, p est divisible par x_2 : il existe un polynôme $q(x_1, x_2)$ tel que

$$p(x_1, x_2) = x_2 q(x_1, x_2).$$

De même, on vérifie que

$$q_2(\pm 1) = q_2'(\pm 1) = q_2(0) = 0.$$

Comme q_2 est un polynôme de degré au plus 4, on a donc $q_2 = 0$. Or $q_2(x_1) = q(x_1, 0)$, ainsi q est divisible par x_2 . On a donc prouvé que p est divisible par x_2^2 . Le polynôme q et ses dérivées s'annulent le long de la droite (a_1, a_2) . Pour des raisons d'invariance, il en est de même le long des droites (a_1, a_3) et (a_2, a_3) . On en déduit que p est également divisible par $(1 + x_1 - x_2/\sqrt{3})^2$ et $(1 - x_1 - x_2/\sqrt{3})^2$. Ainsi, p est un polynôme de degré au plus 5 divisible par un polynôme de degré 6 et $p = 0$.

2. Raccordement au niveau des mailles

Afin de résoudre le problème, il nous faut prouver le Lemme suivant (équivalent du Lemme 6.3.4) :

Lemme. Soit K et K' deux triangles ayant une arête commune $\Gamma = (a_1, a_2)$. Soit p_K et $p_{K'}$ deux éléments de \mathbb{P}_5 , alors la fonction v définie sur $K \cup K'$ par

$$v(x) = \begin{cases} p_K(x) & \text{si } x \in K \\ p_{K'}(x) & \text{si } x \in K' \end{cases}$$

est de classe C^1 si et seulement si

$$\begin{aligned} p_K(a_i) &= p_{K'}(a_i), & \nabla p_K(a_i) &= \nabla p_{K'}(a_i), \\ \nabla \nabla p_K(a_i) &= \nabla \nabla p_{K'}(a_i), & \partial p_K / \partial n(b) &= -\partial p_{K'} / \partial n(b), \end{aligned} \quad (6.17)$$

où n désigne la normale extérieure à K et b le milieu du segment $[a_1, a_2]$.

Démonstration. L'application v est de classe C^1 si et seulement si les restrictions de p_K et $p_{K'}$ coïncident sur l'arête commune Γ au deux triangles et s'il en est de même pour les polynômes $\partial p_K / \partial n$ et $\partial p_{K'} / \partial n$. Or les polynômes p_K et $p_{K'}$ coïncident sur Γ si et seulement si pour $i = 1, 2$ on a

$$p_K(a_i) = p_{K'}(a_i), \quad \frac{\partial p_K}{\partial \tau}(a_i) = \frac{\partial p_{K'}}{\partial \tau}(a_i) \text{ et } \frac{\partial^2 p_K}{\partial \tau^2}(a_i) = \frac{\partial^2 p_{K'}}{\partial \tau^2}(a_i)$$

(τ désigne le vecteur unitaire tangent à l'arête). D'autre part, les restrictions de $\partial p_K / \partial n$ et de $\partial p_{K'} / \partial n$ à Γ sont des polynômes de degré 4 égaux si et seulement si pour $i = 1, 2$ on a

$$\frac{\partial p_K}{\partial n}(a_i) = \frac{\partial p_{K'}}{\partial n}(a_i), \quad \frac{\partial^2 p_K}{\partial n^2}(a_i) = \frac{\partial^2 p_{K'}}{\partial n^2}(a_i)$$

et si

$$\frac{\partial p_K}{\partial n}(b) = \frac{\partial p_{K'}}{\partial n}(b).$$

Ce qui achève la preuve du Lemme.

3. Méthode d'Argyris

Tout d'abord, l'espace

$$V_h = \{v \in C^1(\Omega) : v|_{K_i} \in P^5 \text{ pour tout } K_i \in \mathcal{T}_h\}$$

est inclus dans $H^2(\Omega)$ (la dérivée d'un élément de V_h appartient à $H^1(\Omega)$ d'après le Lemme 4.3.19). D'après le point précédent, un élément v de V_h est entièrement déterminé par les valeurs de v , ∇v et $\nabla \nabla v$ aux nœuds du maillage ainsi que par les flux $\partial v / \partial n(b_k)$, b_k parcourant les milieux des arêtes k du maillage (on oriente de manière arbitraire chacune des arêtes). On peut donc construire une base de V_h formée des éléments $(\varphi_{i,\alpha})_{(i,\alpha)}$ et (ψ_k) où $i \in \{1, \dots, n_s\}$, $\alpha \in \mathbb{N}^2$, $|\alpha| = \alpha_1 + \alpha_2 \leq 2$ et $k \in \{1, \dots, n_c\}$ définis par

$$\begin{aligned} \partial^\beta \varphi_{i,\alpha}(a_j) &= \delta_j^i \delta_\alpha^\beta & \frac{\partial \varphi_{i,\alpha}}{\partial n}(b_l) &= 0; \\ \partial^\beta \psi_k(a_j) &= 0; & \frac{\partial \psi_k}{\partial n}(b_l) &= \delta_l^k. \end{aligned}$$

pour tout $j \in \{1, \dots, n_s\}$, $l \in \{1, \dots, n_c\}$ et $\beta \in \mathbb{N}^2$ tel que $|\beta| \leq 2$.

Afin de résoudre l'équation des plaques (5.19), on introduit le sous espace de V_h

$$V_{0h} = V_h \cap H_0^2(\Omega).$$

L'espace V_{0h} est l'ensemble des fonctions de V_h , qui s'annulent ainsi que leurs dérivées partielles sur $\partial\Omega$. Il est engendré par les éléments $(\varphi_{i,\alpha})$ et (ψ_k) où $i \in \{1, \dots, n_{s0}\}$

et $k \in \{1, \dots, n_{c0}\}$ parcourent respectivement sommets et arêtes n'appartenant pas au bord de Ω . L'approximation variationnelle consiste à trouver $u_h \in V_{0h}$ tel que

$$\int_{\Omega} \Delta u_h \Delta v_h dx = \int_{\Omega} f v_h dx \text{ pour tout } v_h \in V_{0h}.$$

D'après le Théorème de Lax-Milgram, ce problème admet une solution unique. Enfin, il équivaut à résoudre le système

$$\mathcal{K}_h U_h = b_h, \quad (6.18)$$

où la matrice de rigidité est définie par

$$\mathcal{K}_h = \begin{pmatrix} D_h & F_h \\ F_h^T & H_h \end{pmatrix},$$

où D_h et F_h sont des matrices définies par blocs. La matrice D_h est constituée de 6×6 blocs, la matrice F_h est un vecteur colonne constitué de 6 sous-matrices :

$$\begin{aligned} D_h &= (E_h^{ij})_{(i,j) \in \{1, \dots, 6\}^2} \\ F_h &= (G_h^i)_{i \in \{1, \dots, 6\}}. \end{aligned}$$

Les sous-matrices E_h^{ij} et G_h^i sont définies par

$$\begin{aligned} (E_h^{ij})_{kl} &= \int_{\Omega} \Delta \varphi_{k,s_i} \Delta \varphi_{l,s_j} dx, \quad \text{où } (k,l) \in \{1, \dots, n_{s0}\}^2 \\ (G_h^i)_{kl} &= \int_{\Omega} \Delta \varphi_{k,s_i} \Delta \psi_l dx \quad \text{où } (k,l) \in \{1, \dots, n_{s0}\} \times \{1, \dots, n_{c0}\} \end{aligned}$$

où s_i parcourt les multi-indices \mathbb{N}^2 de degré inférieur ou égal à 2 (ensemble qui contient 6 éléments). La matrice H_h est définie par

$$(H_h)_{kl} = \int_{\Omega} \Delta \psi_k \Delta \psi_l dx$$

où $(k,l) \in \{1, \dots, n_{c0}\}^2$. Enfin, Le vecteur b_h compte $6n_{s0} + n_{c0}$ composantes et est défini par

$$b_h = (c_h^1, \dots, c_h^6, d_h)$$

où $c_h^i \in \mathbb{R}^{n_{s0}}$ et $d_h \in \mathbb{R}^{n_{c0}}$ sont les vecteurs

$$\begin{aligned} (c_h^i)_k &= \int_{\Omega} f_h \varphi_{k,s_i} \quad k \in \{1, \dots, n_{s0}\} \text{ et } i \in \{1, \dots, 6\} \\ (d_h)_k &= \int_{\Omega} f_h \psi_k \quad k \in \{1, \dots, n_{c0}\}. \end{aligned}$$

Enfin, la solution u_h de l'approximation variationnelle est telle que

$$u_h = \sum_{i=0}^5 \sum_{k=1}^{n_{s0}} U_h^{in_{s0}+k} \varphi_{k,s_{i+1}} + \sum_{k=1}^{n_{c0}} U_h^{6n_{s0}+k} \psi_k,$$

où U_h est solution du système (6.18).

Exercice 6.3.18 Montrer que pour une suite de maillages réguliers, et pour des éléments finis \mathbb{P}_1 , l'opérateur d'interpolation r_h vérifie en dimension $N = 2$ ou 3

$$\|v - r_h v\|_{L^2(\Omega)} \leq Ch^2 \|v\|_{H^2(\Omega)}.$$

Correction. Par construction de $r_h u$, la restriction de $r_h u$ à un N-simplexe K_i est simplement $r_{K_i} u$. Par conséquent,

$$\|v - r_h v\|_{L^2(\Omega)}^2 = \sum_{K_i \in \mathcal{T}_h} \|v - r_{K_i} v\|_{L^2(K_i)}^2.$$

On applique la majoration (Lemme 6.3.20 avec $k = 1$)

$$\|v - r_K v\|_{L^2(K)} \leq C \|B\|^2 |v|_{H^2(K)}$$

à chacun des N-simplexe K_i . Ainsi,

$$\|v - r_h v\|_{L^2(\Omega)}^2 \leq C \sum_{K_i \in \mathcal{T}_h} \|B_i\|^4 |v|_{H^2(K)}^2.$$

Il suffit de combiner cette estimation avec l'inégalité

$$\|B_i\| \leq \text{diam}(K_i) / \rho(K_0) \leq Ch$$

pour conclure.

Exercice 6.3.19 Soit $K = [0, 1]^2$ le cube unité en dimension $N = 2$ de sommets $a^1 = (0, 0)$, $a^2 = (1, 0)$, $a^3 = (1, 1)$, $a^4 = (0, 1)$. On définit $x_3 = 1 - x_1$, $x_4 = 1 - x_2$, et \bar{i} comme la valeur de i modulo 4. Grâce à ses notations, chaque sommet a^i est défini par $x_{\bar{i}} = x_{\bar{i}+1} = 0$. Vérifier que les fonctions de base de \mathbb{Q}_1 sont

$$p_i(x) = x_{\bar{i}+2} x_{\bar{i}+3} \quad \text{pour } 1 \leq i \leq 4,$$

et que celles de \mathbb{Q}_2 sont

$$\begin{aligned} P_i(x) &= x_{\bar{i}+2}(2x_{\bar{i}+2} - 1)x_{\bar{i}+3}(2x_{\bar{i}+3} - 1) & \text{pour } 1 \leq i \leq 4 \\ P_i(x) &= -4x_{\bar{i}+2}(x_{\bar{i}+2} - 1)x_{\bar{i}+3}(2x_{\bar{i}+3} - 1) & \text{pour } 5 \leq i \leq 8 \\ P_9(x) &= 16x_1 x_2 x_3 x_4. \end{aligned}$$

Correction.

Les éléments de \mathbb{Q}_k définis sur K sont les polynômes de degré au plus k par rapport à chacune des variables x_1 et x_2 . D'après le Lemme 6.3.22, il suffit de vérifier que les fonctions proposées s'annulent sur tout les points du treillis correspondant excepté un point, différent pour chacune d'entre elles, où elles prennent la valeur un.

1. Fonctions de base \mathbb{Q}_1 .

Les points du treillis sont a^i , $i = 1, \dots, 4$. Pour des raisons de périodicité des formules, il suffit de vérifier la forme de la fonction de base p_1 . Or

$$p_1(x) = x_3 x_4 = (1 - x_1)(1 - x_2),$$

qui vaut en effet 1 pour $x = a_1$ et zéro sur les autres sommets du carré.

2. Fonctions de base \mathbb{Q}_2 .

Les points du treillis Σ_2 sont les sommets, les milieux des arêtes et le centre du carré K . Pour des raisons de symétrie, seul trois fonctions de bases sont à étudier. On a

$$P_1(x) = x_3(2x_3 - 1)x_4(2x_4 - 1) = (1 - x_1)(1 - 2x_1)(1 - x_2)(1 - 2x_2)$$

qui vaut 1 pour $x = a_1$ et zéro sur les autres nœuds du treillis.

$$P_5(x) = -4x_3(x_3 - 1)x_4(2x_4 - 1) = 4(1 - x_1)x_1(1 - x_2)(1 - 2x_2)$$

qui vaut 1 pour $x = (a_1 + a_2)/2$ et zéro sur les autres nœuds du treillis. Enfin,

$$P_9(x) = 16x_1x_2x_3x_4,$$

qui vaut 1 en $x = (a_1 + a_2 + a_3 + a_4)/4$ et zéro sur les autres nœuds du treillis.

Exercice 6.3.20 Montrer que pour la méthode des éléments finis \mathbb{P}_1 /bulle pour la vitesse et \mathbb{P}_1 pour la pression on a $\dim(\text{Ker } B_h^*) = 1$.

Correction. Soit $r_h \in \mathcal{Q}_h$ et $w_h \in V_{0h}$ (\mathcal{Q}_h et V_{0h} étant les espaces issus respectivement de la discrétisation \mathbb{P}_1 de la pression et \mathbb{P}_1 /bulle de la vitesse). Par définition,

$$W_h \cdot B_h^* R_h = B_h W_h \cdot R_h = - \int_{\Omega} \text{div}(w_h) r_h \, dx = \int_{\Omega} w_h \cdot \nabla r_h \, dx.$$

Si R_h appartient au noyau de B_h^* , on a

$$\int_{\Omega} w_h \cdot \nabla r_h \, dx = 0$$

pour tout élément $w_h \in V_{0h}$. En particulier, si on applique l'égalité précédente à la fonction bulle $\lambda_1(x) \cdots \lambda_{N+1}(x) e_k$ de la maille K_i (les λ_i sont les coordonnées barycentriques de x dans la maille K_i et $k \in \{1, \dots, N\}$), on obtient

$$\nabla r_h(K_i) \cdot e_k \left(\int_{K_i} \lambda_1(x) \cdots \lambda_{N+1}(x) \, dx \right) = 0.$$

Or

$$\left(\int_{K_i} \lambda_1(x) \cdots \lambda_{N+1}(x) \, dx \right) > 0,$$

d'où on en déduit que $\nabla r_h(K_i) = 0$. Ainsi, r_h est une fonction constante et

$$\dim(B_h^*) = 1.$$

Exercice 6.3.21 On considère les équations de Stokes

$$\begin{cases} \nabla p - \mu \Delta u = f & \text{dans } \Omega \\ \operatorname{div} u = 0 & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega \end{cases} \quad (6.19)$$

où $\mu > 0$ est la viscosité du fluide en dimension $N = 1$ (ce modèle n'a aucun intérêt puisque sa solution explicite est $u = 0$ et p une primitive de f , mais il permet de bien comprendre les problèmes de discrétisation). Pour $\Omega = (0, 1)$, on considère le maillage de points $x_j = jh$ avec $h = 1/(n+1)$ et $0 \leq j \leq n+1$. On définit la méthode de différences finies centrées (d'ordre 2) suivante

$$\begin{cases} \mu \frac{-u_{j+1} + 2u_j - u_{j-1}}{h^2} + \frac{p_{j+1} - p_{j-1}}{2h} = f(x_j) \text{ pour } 1 \leq j \leq n \\ \frac{u_{j+1} - u_{j-1}}{2h} = 0 \text{ pour } 1 \leq j \leq n \\ u_0 = u_{n+1} = 0. \end{cases}$$

Montrer que ce système d'équations algébriques est mal posé, et en particulier que la pression (p_j) est définie à l'addition d'une constante près ou d'un multiple d'une pression définie par ses composantes $(1, 0, 1, 0, \dots, 1, 0)$.

Correction. On vérifie sans mal que la pression est définie à l'addition d'une constante près ou d'un multiple de $(1, 0, \dots, 1, 0)$. Lorsque n est impair, la situation est particulièrement critique, car u_j n'est, dans ce cas, pas non plus déterminé de manière unique. Dans tous les cas, le système est mal posé.

Chapitre 7

PROBLÈMES AUX VALEURS PROPRES

Exercice 7.1.1 Soit $\Omega = \mathbb{R}^N$. Montrer que $u(x) = \exp(ik \cdot x)$ est une solution de

$$-\Delta u = \lambda u \quad (7.1)$$

si $|k|^2 = \lambda$. Une telle solution est appelée onde plane.

Correction. Soit $u(x) = \exp(ik \cdot x)$, on a $\nabla u(x) = i \exp(ik \cdot x)k$ et

$$\Delta u = \nabla \cdot \nabla u = -|k|^2 \exp(ik \cdot x).$$

Ainsi, u est solution de l'équation (7.1) dès que $|k|^2 = \omega^2$.

Exercice 7.1.2 Soit un potentiel régulier $V(x)$. Montrer que, si $\mathbf{u}(x, t) = e^{-i\omega t}u(x)$ est solution de

$$i \frac{\partial \mathbf{u}}{\partial t} + \Delta \mathbf{u} - V \mathbf{u} = 0 \quad \text{dans } \mathbb{R}^N \times \mathbb{R}_*^+, \quad (7.2)$$

alors $u(x)$ est solution de

$$-\Delta u + Vu = \omega u \quad \text{dans } \mathbb{R}^N. \quad (7.3)$$

Correction. Il suffit d'effectuer le calcul. En effet,

$$i \frac{\partial \mathbf{u}}{\partial t}(x, t) = e^{-i\omega t} \omega u(x) \quad \Delta \mathbf{u}(x, t) = e^{-i\omega t} \Delta u(x).$$

Comme \mathbf{u} est solution de l'équation de Schrödinger (7.2), on en déduit que

$$-\Delta u + Vu = \omega u.$$

Exercice 7.1.3 Soit $V(x) = Ax \cdot x$ avec A matrice symétrique réelle définie positive. Montrer que $u(x) = \exp(-A^{1/2}x \cdot x/2)$ est une solution de (7.3) si $\omega = \text{tr}(A^{1/2})$. Une telle solution est appelée état fondamental.

Correction.

Soit $u(x) = \exp(-A^{1/2}x \cdot x/2)$. On a

$$\nabla u = -\exp(-A^{1/2}x \cdot x/2)A^{1/2}x = -uA^{1/2}x$$

et

$$\Delta u = \operatorname{div}(\nabla u) = -\operatorname{div}(uA^{1/2}x)$$

On rappelle que pour toute fonction f à valeurs réelles et σ à valeur vectorielle, $\operatorname{div}(f\sigma) = \nabla f \cdot \sigma + f(\operatorname{div}\sigma)$. Ainsi,

$$\Delta u = -(A^{1/2}x) \cdot \nabla u - (\operatorname{div}(A^{1/2}x))u = (Ax \cdot x)u - \operatorname{tr}(A^{1/2})u,$$

et u est bien solution de l'équation

$$-\Delta u + Vu = \operatorname{tr}(A^{1/2})u.$$

Exercice 7.2.1 Montrer que l'application identité Id dans un espace de Hilbert V de dimension infinie n'est jamais compacte (utiliser le Lemme 7.2.6).

Correction.

L'image de la boule unité par l'application Id est évidemment la boule unité. Si l'application Id était compacte, la boule unité serait relativement compacte et donc compacte (la boule unité est fermée), ce qui est impossible d'après le Lemme 7.2.6 qui stipule que la boule unité d'un espace de Hilbert de dimension infinie n'est jamais compacte.

Exercice 7.2.2 Soit l'espace de Hilbert ℓ_2 des suites réelles $x = (x_i)_{i \geq 1}$ telles que $\sum_{i \geq 1} |x_i|^2 < +\infty$, muni du produit scalaire $\langle x, y \rangle = \sum_{i \geq 1} x_i y_i$. Soit $(a_i)_{i \geq 1}$ une suite de réels bornés, $|a_i| \leq C < +\infty$ pour tout $i \geq 1$. On définit l'application linéaire A par $Ax = (a_i x_i)_{i \geq 1}$. Vérifier que A est continue. Montrer que A est compacte si et seulement si $\lim_{i \rightarrow +\infty} a_i = 0$.

Correction.

Soit x un élément de ℓ^2 ,

$$\|Ax\|_{\ell^2}^2 = \sum_i |a_i x_i|^2 \leq \sup_i |a_i|^2 \sum_i |x_i|^2 = \sup_i |a_i|^2 \|x\|_{\ell^2}^2.$$

Ainsi, A est une application continue de ℓ^2 dans ℓ^2 .

Supposons que $\lim a_i = 0$. Soit x_n une suite d'éléments de la boule unité de ℓ^2 . On pose $y^n = Ax^n$. Afin de prouver que l'opérateur A est compact, on va construire une sous-suite de y^n convergente. On commence par construire une suite de sous-suite par récurrence : On pose $y^{n,0} = y^n$. Pour tout k , $y^{n,k}$ est une suite extraite de $y^{n,k-1}$ telle que $y_k^{n,k}$ soit convergente (c'est toujours possible puisque pour tout $k \geq 1$, $y_k^{n,k}$ est borné dans \mathbb{R}). Enfin, on procède à l'extraction d'une sous-suite diagonale en définissant la suite $z^n = y^{n,n}$. Reste à prouver que la suite z^n est de

Cauchy dans ℓ^2 . Soit $\varepsilon > 0$, comme a_i converge vers 0, il existe l tel que pour tout $i > l$, $|a_i| < \varepsilon$. On en déduit que pour tout indice n ,

$$\sum_{i>l} |z_i^n|^2 = \sum_{i>l} |a_i|^2 |y_l^{n,n}|^2 \leq \varepsilon^2 \|y^{n,n}\|_{\ell^2}^2 \leq \varepsilon^2.$$

Notons que pour tout k , la suite z_k^n est simplement convergente. Ainsi, pour n et p assez grand, on a

$$\sum_{i \leq l} |z_i^n - z_i^p|^2 \leq \varepsilon^2.$$

En combinant ces deux résultats, on en déduit que pour n et p assez grand,

$$\sum_i |z_i^n - z_i^p|^2 \leq \sum_{i \leq l} |z_i^n - z_i^p|^2 + 2 \sum_{i>l} (|z_i^n|^2 + |z_i^p|^2) \leq 5\varepsilon^2,$$

et que $\|z^n - z^p\|_{\ell^2} \rightarrow 0$ lorsque n et p convergent vers l'infini. Ainsi, A est compacte.

Reste à établir la réciproque. Supposons que la suite a_i ne converge pas vers zéro. Il existe une constante $M > 0$ telle que pour tout N , il existe $i > N$ tel que $|a_i| > M$. On peut donc définir les suites x^n de ℓ^2 et i_n de \mathbb{N} telles que

$$\text{Pour tout indice } k, x_k^n = \delta_k^{i_n},$$

$$|a_{i_n}| > M$$

et i_n strictement croissante. On pose $y^n = Ax^n$. La suite x^n est bornée dans ℓ^2 , tandis que la suite y^n d'éléments de ℓ^2 n'admet pas de sous-suite convergente. En effet, pour tout n et p on a

$$\|y^n - y^p\|_{\ell^2} > \sqrt{2}M.$$

Ainsi, A n'est pas compacte.

Exercice 7.2.3 Soit U , V et W trois espaces de Hilbert de dimension infinie, A une application linéaire continue de V dans W , et B une application linéaire continue de U dans V . Montrer que l'application AB est compacte dès que A ou B est compacte. En déduire qu'une application linéaire continue compacte n'est jamais inversible d'inverse continu en dimension infinie.

Correction.

Considérons le cas A compacte et B continue. Comme B est continue, il existe un réel M tel que l'image de la boule unité de U par B soit incluse dans la boule de V , centrée à l'origine et de rayon M . Comme A est compacte, l'image de la boule de rayon M par A est relativement compacte. Or tout sous-ensemble d'un ensemble relativement compact est relativement compact. L'image de la boule unité de U par l'application AB est donc relativement compacte : l'application AB est compacte.

Considérons le cas A continue et B compacte. L'image de la boule unité de U par B est relativement compacte dans V . Or l'image par une application continue

d'un ensemble relativement compact est relativement compact. L'image de la boule unité de U par l'application AB est relativement compacte.

Enfin, considérons une application linéaire compacte inversible A . L'application inverse A^{-1} (qui est linéaire) ne peut être continue. En effet, dans ce cas l'application identité AA^{-1} serait compacte, ce qui n'est jamais le cas en dimension infinie.

Exercice 7.2.4 On reprend les notations et les hypothèses du Théorème 7.2.8. Montrer que, pour $v \in V$, l'équation $Au = v$ admet une unique solution $u \in V$ si et seulement si v vérifie

$$\sum_{k=1}^{+\infty} \frac{|\langle v, u_k \rangle|^2}{\lambda_k^2} < +\infty.$$

Correction. Supposons qu'il existe u tel que $Au = v$. Pour tout k , $\langle Au, u_k \rangle = \langle v, u_k \rangle$ et donc

$$\langle u, u_k \rangle = \frac{\langle u, Au_k \rangle}{\lambda_k} = \frac{\langle Au, u_k \rangle}{\lambda_k} = \frac{\langle v, u_k \rangle}{\lambda_k}.$$

La famille (u_k) formant une base orthonormale,

$$\sum_k \frac{\langle v, u_k \rangle^2}{\lambda_k^2} = \sum_k \langle u, u_k \rangle^2 = \|u\|^2 < +\infty.$$

Réciproquement, si v vérifie la relation précédente,

$$u = \sum_k \frac{\langle v, u_k \rangle}{\lambda_k} u_k$$

appartient à U (la série est convergente) et $Au = v$. Enfin, le système $Au = v$ ne peut admettre plus d'une solution. En effet, l'application A étant définie positive, elle est injective.

Exercice 7.2.5 Soit $V = L^2(0, 1)$ et A l'application linéaire de V dans V définie par $(Af)(x) = (x^2 + 1)f(x)$. Vérifier que A est continue, définie positive, auto-adjointe mais pas compacte. Montrer que A n'a pas de valeurs propres. On pourra vérifier aussi que $(A - \lambda \text{Id})$ est inversible d'inverse continu si et seulement si $\lambda \notin [1, 2]$.

Correction. Continuité

$$\|Af\|_{L^2(0,1)}^2 = \int_0^1 (x^2 + 1)^2 |f(x)|^2 dx \leq \left(\max_{x \in (0,1)} (x^2 + 1)^2 \right) \int_0^1 |f(x)|^2 dx.$$

Ainsi, $\|Af\|_{L^2(0,1)} \leq 2\|f\|_{L^2(0,1)}$ et A est continue.

Positivité et symétrie

Soit f et g éléments de $L^2(0, 1)$,

$$(Af, g)_{L^2} = \int_0^1 (Af)g dx = \int_0^1 (x^2 + 1)fg dx = (f, Ag)_{L^2}.$$

Ainsi, A est auto-adjointe. De plus, A est positive car

$$(Af, f)_{L^2} = \int_0^1 (x^2 + 1)|f(x)|^2 dx \geq 0.$$

Enfin, A est définie. En effet, si $(Af, f) = 0$, la fonction $(x^2 + 1)|f(x)|^2$ est nulle presque partout, donc $f = 0$ (en tant qu'élément de $L^2(0, 1)$).

Valeurs propres

Supposons que f soit un vecteur propre de A de valeur propre λ . Dans ce cas, pour toute fonction $g \in L^2(0, 1)$,

$$\int_0^1 (x^2 + 1)f(x)g(x) dx = (Af, g)_{L^2} = \lambda(f, g)_{L^2} = \lambda \int_0^1 f(x)g(x) dx.$$

On en déduit que

$$((x^2 + 1)f - \lambda f, g(x))_{L^2} = 0.$$

En choisissant $g = (x^2 + 1 - \lambda)f$, on en déduit que

$$\|(x^2 + 1 - \lambda)f\|_{L^2} = 0$$

et que $(x^2 + 1 - \lambda)f(x) = 0$ presque partout et donc $f(x) = 0$ presque partout. L'application A n'admet pas de vecteur propre non nul.

Inversibilité de $(A - \lambda \text{Id})$ Soit $g \in L^2(0, 1)$, on cherche f tel que $(A - \lambda \text{Id})f = g$, c'est à dire tel que

$$(x^2 + 1 - \lambda)f(x) = g(x)$$

presque partout. Si $(A - \lambda \text{Id})$ est inversible, $f = (A - \lambda \text{Id})^{-1}g$ est défini par

$$f(x) = (x^2 + 1 - \lambda)^{-1}g(x)$$

pour presque tout $x \in]0, 1[$. L'inverse de $(x^2 + 1 - \lambda)$ étant défini, sauf en au plus deux points, $f(x)$ est correctement défini presque partout.

Si λ n'appartient pas à l'intervalle $[1, 2]$, il existe $C(\lambda)$ tel que $(x^2 + 1 - \lambda) > C(\lambda) > 0$. On en déduit que l'opérateur $(A - \lambda \text{Id})$ est bien inversible de $L^2(0, 1)$ dans $L^2(0, 1)$, d'inverse continue. En effet,

$$\|(A - \lambda \text{Id})^{-1}g\|_{L^2(0,1)} \leq C(\lambda)^{-1}\|g\|_{L^2(0,1)}.$$

Si $\lambda \in [1, 2]$, on constate que si $(A - \lambda \text{Id})$ était inversible, $(x^2 + 1 - \lambda)^{-1}$ serait un élément de $L^2(0, 1)$ (prendre $g = 1$). Ceci n'est pas le cas. En effet le polynôme $(x^2 + 1 - \lambda)$ admet une racine dans l'intervalle $[1, 2]$. Ainsi, $(x^2 + 1 - \lambda)^{-1}$ présente une singularité (du type $1/x$ ou $1/x^2$) dont le carré n'est pas d'intégrale finie :

$$\int_0^1 (x^2 + 1 - \lambda)^{-2} dx = +\infty.$$

Exercice 7.3.1 Démontrer une variante du Théorème 7.3.2 où l'on remplace l'hypothèse de coercivité de la forme bilinéaire $a(\cdot, \cdot)$ par l'hypothèse plus faible qu'il existe deux constantes positives $\eta > 0$ et $\nu > 0$ telles que

$$a(v, v) + \eta \|v\|_H^2 \geq \nu \|v\|_V^2 \text{ pour tout } v \in V.$$

(Dans ce cas les valeurs propres $(\lambda_k)_{k \geq 1}$ ne sont pas forcément positives, mais vérifient seulement $\lambda_k + \eta > 0$.)

Correction. Un réel λ est valeur propre de (7.12) de vecteur propre u , si et seulement si

$$a(u, v) + \eta \langle u, v \rangle_H = (\lambda + \eta) \langle u, v \rangle_H \text{ pour tout } v \in V,$$

c'est à dire si u est un vecteur propre associé à la forme bilinéaire $a(\cdot, \cdot) + \eta \langle \cdot, \cdot \rangle_H$ de valeur propre $\lambda + \eta$. Comme la forme bilinéaire $a(\cdot, \cdot) + \eta \langle \cdot, \cdot \rangle_H$ vérifie les hypothèses du Théorème 7.3.2, il existe une base hilbertienne de H de vecteurs propres u_k de (7.12) de valeurs propres $\lambda_k - \eta$ où λ_k est une suite non bornée, croissante de réels positifs.

Exercice 7.3.2 En dimension $N = 1$, on considère $\Omega =]0, 1[$. Calculer explicitement toutes les valeurs propres et les fonctions propres du Laplacien avec conditions aux limites de Dirichlet

$$\begin{cases} -\Delta u_k = \lambda_k u_k & \text{p.p. dans } \Omega \\ u_k = 0 & \text{p.p. sur } \partial\Omega. \end{cases} \quad (7.4)$$

A l'aide de la décomposition spectrale de ce problème (voir la Remarque 7.2.9), montrer que la série

$$\sum_{k=1}^{+\infty} a_k \sin(k\pi x)$$

converge dans $L^2(0, 1)$ si et seulement si $\sum_{k=1}^{+\infty} a_k^2 < +\infty$, et dans $H^1(0, 1)$ si et seulement si $\sum_{k=1}^{+\infty} k^2 a_k^2 < +\infty$.

Correction. Tout d'abord, toute fonction propre du Laplacien en dimension 1 est de classe C^∞ . Ainsi,

$$u'' + \lambda u = 0$$

est une équation différentielle classique dont les solutions sont de la forme

$$u = A \sin(\sqrt{\lambda}x) + B \cos(\sqrt{\lambda}x).$$

Les conditions aux limites de Dirichlet impliquent que $B = 0$ (car $u(0) = 0$) et $\sqrt{\lambda} = k\pi$ où k est un entier naturel (car $u(1) = 0$). Les vecteurs propres du Laplacien unidimensionnel avec conditions aux limites de Dirichlet sont donc les fonctions

$$u_k = 2 \sin(k\pi x)$$

de valeurs propres $\lambda_k = k^2\pi^2$. Comme l'injection de $H_0^1(0, 1)$ dans $L^2(0, 1)$ est compacte et que $a(u, v) = \int_0^1 \nabla u \cdot \nabla v \, dx$ est une forme bilinéaire symétrique, continue

et coercive sur $H_0^1(]0, 1[)$, on peut appliquer le Théorème 7.3.2. Ainsi, $(u_k/k\pi)$ est une base de hilbertienne $H^1(]0, 1[)$ et (u_k) une base hilbertienne de $L^2(]0, 1[)$. On en déduit que la série

$$\sum_{k=1}^{\infty} a_k \sin(kx)$$

converge dans L^2 si et seulement si $\sum_k a_k^2 < \infty$ et dans H_0^1 si et seulement si $\sum_k k^2 a_k^2 < \infty$.

Exercice 7.3.3 On considère un parallélépipède

$$\Omega =]0, L_1[\times]0, L_2[\times \cdots \times]0, L_N[,$$

où les $(L_i > 0)_{1 \leq i \leq N}$ sont des constantes positives. Calculer explicitement toutes les valeurs propres et les fonctions propres du Laplacien avec conditions aux limites de Dirichlet (7.4).

Correction.

Soit $u_k(x) = 2 \sin(k\pi x)$ les fonctions propre du Laplacien avec conditions de Dirichlet sur $]0, 1[$. Pour tout entier $0 < p < N + 1$, et tout $k \in \mathbb{N}_*$, on pose

$$u_{p,k}(x) = u_k(x/L_p).$$

Enfin, pour tout $k = (k_1, \dots, k_N) \in \mathbb{N}_*^N$, on pose

$$v_k = \prod_{p=1}^N u_{p,k_p}.$$

On vérifie que pour k , v_k est une valeur propre du Laplacien sur Ω avec conditions aux bords de Dirichlet de valeur propre

$$\lambda_k = \left(\prod_{p=1}^N k_p \pi / L_p \right)^2.$$

Pour conclure, il reste à prouver que la famille $v_k / \sqrt{\lambda_k}$ forme une base de $L^2(\Omega)$, c'est à dire que pour tout $w \in L^2(\Omega)$ tel que $\forall k \in \mathbb{N}_*^p$,

$$\langle v_k, w \rangle_{L^2(\Omega)} = 0, \quad (7.5)$$

on a $w = 0$. Supposons que le résultat soit établi pour Ω de dimension $N - 1$. On introduit la fonction $\tilde{w} \in L^2(]0, L_N[)$ définie par

$$\tilde{w}(x_N) = \int_{\tilde{\Omega}} w(x) \prod_{p < N} u_{k_p}(x_p) d\tilde{x},$$

où $\tilde{\Omega} =]0, L_1[\times \dots \times]0, L_{N-1}[$ et $\tilde{x} = (x_1, \dots, x_{N-1})$. D'après (7.5), pour tout $k \in \mathbb{N}_*$, on a

$$\int_0^{L_N} \tilde{w}(x_N) u_{N,k}(x_N) dx_N = 0.$$

Comme la famille $u_{N,k}/\sqrt{\pi/L_p}$ forme une base de $L^2(]0, L_N[)$, on en déduit que $\tilde{w}(x_N) = 0$ pour presque tout x_N . Ainsi, pour presque tout $x_N \in]0, L_N[$, la fonction $w_{x_N}(\tilde{x}) = w(\tilde{x}, x_N) \in L^2(\tilde{\Omega})$ est telle que

$$\int_{\tilde{\Omega}} w_{x_N}(\tilde{x}) \prod_{p < N} u_{k_p}(x_p) d\tilde{x} = 0,$$

et d'après l'hypothèse de récurrence, $w_{x_N} = 0$, ce qui achève la démonstration.

Exercice 7.3.4 On considère à nouveau un ouvert Ω parallélépipédique comme dans l'Exercice 7.3.3. Calculer explicitement toutes les valeurs propres et les fonctions propres du Laplacien avec conditions aux limites de Neumann sur tout le bord $\partial\Omega$.

Correction.

Les fonctions propres du Laplacien 1D avec conditions de Neumann sur $]0, 1[$ sont les fonctions

$$u_k(x) = \cos(k\pi x)$$

de valeurs propres $k^2\pi^2$. En suivant le même raisonnement que lors de l'Exercice 7.3.3, on montre que les fonctions propres du Laplacien avec conditions de Neumann sur $\Omega =]0, L_1[\times \cdots \times]0, L_p[$ sont de la forme

$$u_k(x) = \prod_{p=1}^N \cos(k_p \pi x_p / L_p)$$

où $k \in \mathbb{N}^N$. La valeur propre associée à u_k étant

$$\lambda_k = \left(\prod_{p=1}^N k_p \pi / L_p \right)^2.$$

Exercice 7.3.5 On reprend les notations et les hypothèses du Théorème 7.3.5. Montrer que la meilleure (i.e. la plus petite) constante C dans l'inégalité de Poincaré (voir la Proposition 4.3.10) est précisément la première valeur propre λ_1 de (7.4).

Correction.

Soit u_k , base hilbertienne de $L^2(\Omega)$, fonctions propres du Laplacien avec conditions aux limites de Dirichlet (7.4) et λ_k les valeurs propres associées (ordonnées par ordre croissant). Soit u un élément de $H_0^1(\Omega)$.

$$\|u\|_{L^2}^2 = \sum_k |\langle u, u_k \rangle_{L^2}|^2 \leq \lambda_1^{-1} \sum_k \lambda_k |\langle u, u_k \rangle_{L^2}|^2 = \lambda_1^{-1} \|\nabla u\|_{L^2}^2.$$

Ainsi, l'inégalité de Poincaré

$$\int_{\Omega} |v(x)|^2 dx \leq C \int_{\Omega} |\nabla v(x)|^2 dx. \quad (7.6)$$

est vérifiée pour $C = \lambda_1^{-1}$. Cette valeur est optimale car $\|u_1\|_{L^2}^2 = \lambda_1^{-1} \|\nabla u_1\|_{L^2}^2$.

Exercice 7.3.6 Soit Ω un ouvert borné régulier et connexe. Montrer que la première valeur propre du Laplacien dans Ω avec condition aux limites de Neumann est nulle et qu'elle est simple.

Correction.

Tout d'abord, zéro est valeur propre du Laplacien avec conditions aux limites de Neumann car

$$\begin{cases} \Delta 1 = 0 & \text{dans } \Omega \\ \frac{\partial 1}{\partial n} = 0 & \text{sur } \partial\Omega. \end{cases}$$

Si λ est une valeur propre du Laplacien de fonction propre u , on a

$$\|\nabla u\|_{L^2}^2 = \lambda \|u\|_{L^2}^2.$$

Ainsi, les valeurs propres du Laplacien avec conditions aux limites de Neumann sont strictement positives sauf si $\|\nabla u\|_{L^2} = 0$ auquel cas $\lambda = 0$. Comme Ω est connexe, si $\lambda = 0$ la fonction u est constante. Ainsi, la première valeur propre du Laplacien avec condition aux limites de Neumann est 0 et elle est simple.

Exercice 7.3.7 Soit Ω un ouvert borné régulier connexe de classe \mathcal{C}^1 de \mathbb{R}^N . Montrer qu'il existe une suite croissante $(\lambda_k)_{k \geq 1}$ de réels positifs qui tend vers l'infini, et une base hilbertienne de $L^2(\Omega)^N$ $(u_k)_{k \geq 1}$, telle que chaque u_k appartient à $H_0^1(\Omega)^N$, et il existe une famille de pressions $p_k \in L^2(\Omega)$ qui vérifient

$$\begin{cases} \nabla p_k - \mu \Delta u_k = \lambda_k u_k & \text{p.p. dans } \Omega \\ \operatorname{div} u_k = 0 & \text{p.p. dans } \Omega \\ u_k = 0 & \text{p.p. sur } \partial\Omega. \end{cases}$$

Correction. On introduit l'espace de Hilbert

$$V = \{v \in H_0^1(\Omega)^N : \operatorname{div} v = 0 \text{ p.p. dans } \Omega\}.$$

On munit V du produit scalaire

$$a(u, v) = \mu \int_{\Omega} \nabla u \cdot \nabla v \, dx.$$

L'injection de V dans $L^2(\Omega)^N$ étant compacte (d'après le Théorème de Rellich), il existe une famille positive et croissante de valeurs propres λ_k et $u_k \in V$ une base de $L^2(\Omega)^N$ tels que

$$a(u_k, v) = \lambda_k \int_{\Omega} u_k \cdot v \, dx \text{ pour tout } v \in V.$$

Pour tout k , on définit la forme linéaire continue L_k sur $H_0^1(\Omega)^N$ par

$$L_k(v) = \lambda_k \int_{\Omega} u_k \cdot v \, dx - a(u_k, v)$$

La forme linéaire L_k s'annule sur V et d'après de Théorème de de Rahm 5.3.9, il existe $p_k \in L^2(\Omega)$ tel que

$$L_k(v) = \int_{\Omega} p_k \operatorname{div} v \, dx \text{ pour tout } v \in H_0^1(\Omega)^N.$$

On en déduit en procédant comme lors de la résolution du problème de Stokes que

$$-\mu \Delta u + \nabla p_k = \lambda_k u_k \text{ dans } \Omega$$

(Attention, dans cette expression, la somme $-\mu \Delta u + \nabla p_k$ appartient à $L^2(\Omega)$, ce qui n'est pas forcément le cas de chacun des termes sans hypothèses supplémentaires sur la régularité de Ω). Par définition, comme les éléments u_k appartiennent à V ,

$$\begin{aligned} \operatorname{div}(u_k) &= 0 && \text{dans } \Omega \\ \text{et } u_k &= 0 && \text{sur } \partial\Omega. \end{aligned}$$

Exercice 7.3.8 On considère le problème aux valeurs propres pour l'équation de Schrödinger avec un potentiel quadratique $V(x) = Ax \cdot x$ où A est une matrice symétrique définie positive (modèle de l'oscillateur harmonique)

$$-\Delta u + Vu = \lambda u \text{ dans } \mathbb{R}^N. \quad (7.7)$$

On définit les espaces $H = L^2(\mathbb{R}^N)$ et

$$V = \{v \in H^1(\mathbb{R}^N) \text{ tel que } |x|v(x) \in L^2(\mathbb{R}^N)\}.$$

Montrer que V est un espace de Hilbert pour le produit scalaire

$$\langle u, v \rangle_V = \int_{\mathbb{R}^N} \nabla u(x) \cdot \nabla v(x) \, dx + \int_{\mathbb{R}^N} |x|^2 u(x) v(x) \, dx,$$

et que l'injection de V dans H est compacte. En déduire qu'il existe une suite croissante $(\lambda_k)_{k \geq 1}$ de réels positifs qui tend vers l'infini et une base hilbertienne de $L^2(\mathbb{R}^N)$ $(u_k)_{k \geq 1}$ qui sont les valeurs propres et les fonctions propres de (7.7). Calculer explicitement ses valeurs et fonctions propres (on cherchera u_k sous la forme $p_k(x) \exp(-Ax \cdot x/2)$ où p_k est un polynôme de degré $k-1$). Interpréter physiquement les résultats.

Correction.

1. V est un Hilbert

Tout d'abord, il est évident que $\langle \cdot, \cdot \rangle_V$ définit bien un produit scalaire sur V . Reste à montrer que V muni de la norme associée est complet pour prouver que V est un espace de Hilbert. Soit B_1 la boule unité de \mathbb{R}^N et B_2 la boule de rayon 2. Par un raisonnement par l'absurde, on montre aisément qu'il existe une constante $C \geq 1$ telle que

$$\int_{B_2} |u|^2 dx \leq C \left(\int_{B_2} |\nabla u|^2 dx + \int_{B_2 \setminus B_1} |u|^2 dx \right).$$

On en déduit que pour $u \in V$,

$$\|u\|_{H^1(\mathbb{R}^N)} \leq C\|u\|_V.$$

En effet,

$$\begin{aligned} \|u\|_{L^2(\mathbb{R})}^2 &\leq \int_{B_1} |u|^2 dx + \int_{\mathbb{R}^N \setminus B_1} |x|^2 |u|^2 dx \\ &\leq C \left(\int_{B_2} |\nabla u|^2 dx + \int_{B_2 \setminus B_1} |u|^2 dx \right) + \int_{\mathbb{R}^N \setminus B_1} |x|^2 |u|^2 dx \\ &\leq (C+1)\|u\|_V^2. \end{aligned}$$

Ainsi, si u_n est une suite de Cauchy de V , elle est également une suite de Cauchy de $H^1(\mathbb{R}^N)$. Il existe donc $u \in H^1(\mathbb{R}^N)$ telle que u_n converge vers u dans $H^1(\mathbb{R}^N)$ fort. La suite $|x|u_n$ étant elle-même de Cauchy dans $L^2(\mathbb{R}^N)$, elle converge dans $L^2(\mathbb{R}^N)$ vers une limite v de $L^2(\mathbb{R}^N)$. Enfin, pour tout $\phi \in C_c^\infty(\mathbb{R}^N)$,

$$\begin{aligned} \lim_{n \rightarrow +\infty} \int_{\mathbb{R}^N} |x|u_n(x)\phi(x)dx &= \int_{\mathbb{R}^N} |x|u(x)\phi(x)dx \\ &= \int_{\mathbb{R}^N} v(x)\phi(x)dx. \end{aligned}$$

On en déduit que $v = |x|u$ et que u_n converge vers u dans V .

2. Compacité

Soit u_n une suite bornée de V , $\|u_n\|_V < M$, il existe une sous-suite et u telle que u_n converge vers u dans $L^2(B_R)$ sur toute boule B_R centrée en l'origine de rayon R . Pour tout réel $A > 0$,

$$\begin{aligned} \int_{\mathbb{R}^N} |u - u_n|^2 dx &\leq \int_{|x| < A} |u - u_n|^2 dx + 1/A^2 \int_{|x| > A} |x|^2 |u - u_n|^2 dx \\ &\leq \int_{|x| < A} |u - u_n|^2 dx + 2M/A^2. \end{aligned}$$

Pour n assez grand, $\|u - u_n\|_{L^2(B_A)}^2 \leq M/A^2$ et

$$\int_{\mathbb{R}^N} |u - u_n|^2 dx \leq 3M/A^2.$$

On en déduit que u_n converge vers u dans $L^2(\mathbb{R}^N)$ fort. Ainsi, l'injection de V dans H est compacte.

3. Fonctions propres

La forme bilinéaire

$$a(u, v) = \int_{\mathbb{R}^N} (\nabla u(x) \cdot \nabla v(x) + (Ax \cdot x)u(x)v(x)) dx$$

est symétrique, continue et coercive sur V . On déduit donc du Théorème **7.3.2** qu'il existe une base hilbertienne de $L^2(\mathbb{R}^N)$ formée de vecteurs propres u_k de (7.7) et dont les valeurs propres associées λ_k sont positives et convergent vers l'infini.

Exercice 7.3.9 Soit Ω un ouvert borné régulier de \mathbb{R}^N . On considère le problème de vibrations pour l'équation des plaques avec condition aux limites d'encastrement

$$\begin{cases} \Delta(\Delta u) = \lambda u & \text{dans } \Omega \\ \frac{\partial u}{\partial n} = u = 0 & \text{sur } \partial\Omega. \end{cases}$$

Montrer qu'il existe une suite croissante $(\lambda_k)_{k \geq 1}$ de valeurs propres positives qui tend vers l'infini et une base hilbertienne dans $L^2(\Omega)$ de fonctions propres $(u_k)_{k \geq 1}$ qui appartiennent à $H_0^2(\Omega)$.

Correction.

On introduit la forme bilinéaire

$$a(u, v) = \int_{\Omega} \Delta u \Delta v dx$$

qui est symétrique, continue et coercive sur $H_0^2(\Omega)$. Comme l'injection de $H_0^2(\Omega)$ dans $L^2(\Omega)$ est compacte, la conclusion découle de l'application du Théorème 7.3.2.

Exercice 7.4.1 On considère le problème aux valeurs propres en dimension $N = 1$

$$\begin{cases} -u_k'' = \lambda_k u_k & \text{pour } 0 < x < 1 \\ u_k(0) = u_k(1) = 0. \end{cases}$$

On se propose de calculer la matrice de masse pour la méthode des éléments finis \mathbb{P}_1 . On reprend les notations de la Section 6.2. Montrer que la matrice de masse \mathcal{M}_h est donnée par

$$\mathcal{M}_h = h \begin{pmatrix} 2/3 & 1/6 & & & 0 \\ 1/6 & 2/3 & 1/6 & & \\ & \ddots & \ddots & \ddots & \\ & & 1/6 & 2/3 & 1/6 \\ 0 & & & 1/6 & 2/3 \end{pmatrix},$$

et que ses valeurs propres sont

$$\lambda_k(\mathcal{M}_h) = \frac{h}{3} (2 + \cos(k\pi h)) \text{ pour } 1 \leq k \leq n.$$

Montrer que, si on utilise la formule de quadrature (6.8), alors on trouve que $\mathcal{M}_h = h \text{Id}$. Dans ce dernier cas, calculer les valeurs propres du problème spectral discret.

Correction. La matrice de masse \mathcal{M}_h est définie par

$$(\mathcal{M}_h)_{ij} = \int_0^1 \phi_i(x) \phi_j(x) dx,$$

où ϕ_i sont les fonctions de base des éléments finis \mathbb{P}_1 . Pour tout i et j tels que $|i - j| > 1$, les supports de ϕ_i et ϕ_j sont disjoints et

$$(\mathcal{M}_h)_{ij} = 0.$$

Si $j = i + 1$,

$$\begin{aligned} (\mathcal{M})_{ij} &= \int_{ih}^{(i+1)h} \phi_i(x) \phi_{i+1}(x) dx = \int_{ih}^{(i+1)h} \frac{((i+1)h - x)}{h} \frac{x - ih}{h} dx \\ &= h^{-2} \int_0^h (h - x)x dx = h/6. \end{aligned}$$

Enfin, si $i = j$,

$$\begin{aligned} (\mathcal{M}_h)_{ij} &= \int_{(i-1)h}^{(i+1)h} |\phi_i(x)|^2 dx = 2 \int_{ih}^{(i+1)h} \left| \frac{(i+1)h - x}{h} \right|^2 dx \\ &= 2h^{-2} \int_0^h |h - x|^2 dx = 2h/3. \end{aligned}$$

On a donc montré que la matrice de masse obtenue par la méthode des éléments finis \mathbb{P}_1 est

$$\mathcal{M}_h = h \begin{pmatrix} 2/3 & 1/6 & & & 0 \\ 1/6 & 2/3 & 1/6 & & \\ & \ddots & \ddots & \ddots & \\ & & 1/6 & 2/3 & 1/6 \\ 0 & & & 1/6 & 2/3 \end{pmatrix}$$

Soit $(U, \lambda) \in \mathbb{R}^n \times \mathbb{R}$ ($n = h^{-1} - 1$) tels que

$$\mathcal{M}_h U = \lambda U \quad (7.8)$$

et $U \neq 0$. Afin de calculer les valeurs propres de la matrice de masse \mathcal{M}_h , on effectue une analyse de type Fourier. On introduit la fonction u_h périodique de période 2, impaire, définie sur $[0, 1]$ par

$$u_h(x) = \begin{cases} 0 & \text{si } x \in [0, h/2[\\ U_k & \text{si } x \in [kh - h/2, kh + h/2[\\ 0 & \text{si } x \in [1 - h/2, 1[\end{cases} \quad (7.9)$$

D'après (7.8), pour tout $x \in \mathcal{R}$, on a

$$h \frac{u_h(x - h) + 4u_h(x) + u_h(x + h)}{6} = \lambda u_h(x). \quad (7.10)$$

Remarque 7.4.1 On a choisit u_h impaire de période 2 afin que l'équation (7.10) soit vérifiée pour tout x et en particulier pour tout $x \in [0, h/2] \cup [1 - h/2, 1]$.

Comme u_h est périodique de période 2, il existe \hat{u}_k tel que

$$u_h(x) = \sum_{k=-\infty}^{\infty} \hat{u}_k e^{ik\pi x}.$$

En appliquant la transformée de Fourier à (7.10), on obtient

$$h \frac{e^{-ihk\pi} \hat{u}_k + 4\hat{u}_k + e^{ihk\pi} \hat{u}_k}{6} = \lambda \hat{u}_k,$$

c'est à dire

$$(\cos(k\pi h) + 2 - 3\lambda/h) \hat{u}_k = 0.$$

Comme $U \neq 0$, il existe au moins un k tel que

$$\cos(k\pi h) + 2 - 3\lambda/h = 0$$

ou encore tel que

$$\lambda = \frac{h}{3}(2 + \cos(kh\pi)).$$

Ainsi, toute valeur propre de \mathcal{M}_h est de la forme

$$\lambda_k = \frac{h}{3}(2 + \cos(kh\pi)), \quad \text{où } k \in \{0, \dots, n+1\}. \quad (7.11)$$

Réciproquement, pour tout $k \in \{1, \dots, n\}$, les fonctions $u_h(x)$ vérifiant l'équation (7.10) avec $\lambda = \lambda_k$ sont de la forme de la forme

$$u_h(x) = \sum_j \hat{u}_{k+2(n+1)j} e^{i(k+2(n+1)j)\pi x} + \hat{u}_{-(k+2(n+1)j)} e^{i(k+2(n+1)j)\pi x}.$$

Afin que u_h soit définie à partir d'un vecteur $U \in \mathbb{R}^n$ par (7.9), il est nécessaire que u_h soit impaire, à valeur réelles. On en déduit qu'on a alors

$$\hat{u}_{k+2(n+1)j} = -\hat{u}_{-(k+2(n+1)j)},$$

et que les coefficients de Fourier \hat{u}_m sont imaginaires pures. On en déduit qu'il existe une suite a_j de réels telle que

$$u_h(x) = \sum_j a_j \sin((k+2(n+1)j)\pi x).$$

Ainsi, si $\mathcal{M}_h U = \lambda_k U$, on a

$$U_p = u_h(x = hp) = \sum_j a_j \sin((k+2j(n+1))\pi ph) = \left(\sum_j a_j \right) \sin(khp\pi).$$

Un calcul similaire appliqué au cas $k = 0$ ou $k = n+1$, nous montre que λ_0 et λ_{n+1} ne sont pas des valeurs propres de \mathcal{M}_h . Finalement, comme \mathcal{M}_h est symétrique, définie positive, elle admet une base de vecteurs propres. Les seules valeurs propres possibles sont les n valeurs de λ_k pour $k \in \{1, \dots, n\}$. A chacune de ces valeurs propres, on peut associer au plus un vecteur propre. Ainsi, il ne peut y avoir de valeur propre double. On a donc prouvé que pour tout $k \in \{1, \dots, n\}$,

$$\mathcal{M}_h U^k = \lambda_k U^k,$$

où $U^k = (\sin(khp\pi))_p$.

On note $\widetilde{\mathcal{M}}_h$ la matrice de masse obtenue par la formule de quadrature (6.8). Pour tout entier i et j , on obtient

$$(\widetilde{\mathcal{M}}_h)_{ij} = \sum_{k=1}^n h/2(\phi_i(hk)\phi_j(hk) + \phi_i(h(k+1))\phi_j(h(k+1))) = h \sum_{k=1}^n \delta_i^k \delta_j^k = h\delta_i^j.$$

En utilisant la matrice de masse ainsi obtenue, les valeurs propres et vecteur propres du problème spectral discret vérifient

$$\mathcal{K}_h U_h = h\lambda_h U_h,$$

où

$$\mathcal{K}_h = h \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix}$$

On déduit des valeurs propres de \mathcal{M}_h et de la relation

$$\mathcal{K}_h = 5h \text{Id} - 6\mathcal{M}_h$$

que les valeurs propres du problème spectral sont de la forme

$$\lambda_k = 1 - 2 \cos(kh\pi),$$

et $k \in \{1, \dots, n\}$

Chapitre 8

PROBLÈMES D'ÉVOLUTION

Exercice 8.2.1 Soit Ω un ouvert borné régulier de \mathbb{R}^N . Soit un temps final $T > 0$, une donnée initiale $u_0 \in L^2(\Omega)$, et un terme source $f \in L^2(]0, T[; L^2(\Omega))$. On considère la solution u de l'équation

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = f & \text{p.p. dans } \Omega \times]0, T[\\ u = 0 & \text{p.p. sur } \partial\Omega \times]0, T[\\ u(x, 0) = u_0(x) & \text{p.p. dans } \Omega. \end{cases} \quad (8.1)$$

1. En supposant que la solution u de (8.1) est assez régulière dans $]0, T[\times \Omega$, montrer que, pour tout $t \in [0, T]$, on a l'égalité d'énergie suivante

$$\begin{aligned} \frac{1}{2} \int_{\Omega} u(x, t)^2 dx + \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds &= \frac{1}{2} \int_{\Omega} u_0(x)^2 dx \\ &+ \int_0^t \int_{\Omega} f(x, s) u(x, s) dx ds. \end{aligned} \quad (8.2)$$

2. Démontrer la propriété suivante, appelée "lemme de Gronwall" : si z est une fonction continue de $[0, T]$ dans \mathbb{R}^+ telle que

$$z(t) \leq a + b \int_0^t z(s) ds \quad \forall t \in [0, T],$$

où a, b sont deux constantes positives ou nulles, alors

$$z(t) \leq ae^{bt} \quad \forall t \in [0, T].$$

3. En appliquant le lemme de Gronwall avec $z(t) = \frac{1}{2} \int_{\Omega} u(x, t)^2 dx$, déduire de (8.2) que, pour tout $t \in [0, T]$,

$$\begin{aligned} \frac{1}{2} \int_{\Omega} u(x, t)^2 dx + \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds &\leq \frac{e^t}{2} \left(\int_{\Omega} u_0(x)^2 dx \right. \\ &\left. + \int_0^T \int_{\Omega} f(x, s)^2 dx ds \right). \end{aligned} \quad (8.3)$$

Correction.

1. En intégrant le produit de l'équation d'évolution par u sur Ω , on obtient

$$\int_{\Omega} \frac{\partial u}{\partial t} u - \Delta u u dx = \int_{\Omega} f u dx.$$

Par intégration par partie et en échangeant l'opérateur de dérivation en temps et intégrale, il vient

$$\frac{d}{dt} \frac{1}{2} \int_{\Omega} u^2 dx + \int_{\Omega} |\nabla u|^2 dx = \int_{\Omega} f u dx.$$

Il suffit alors d'effectuer une intégration en temps pour obtenir l'égalité désirée.

2. Soit $v(t) = a + b \int_0^t z(s) ds$. La fonction v est de classe C^1 et

$$v'(t) = bz(t) \leq bv(t).$$

Ainsi,

$$(v(t) \exp(-bt))' = \exp(-bt)(v'(t) - bv(t)) \leq 0$$

et $v(t) \exp(-bt) \leq v(0) = a$. Comme $z(t) \leq v(t)$, on a montré que

$$z(t) \leq a \exp(bt).$$

3. On pose

$$z(t) = \frac{1}{2} \int_{\Omega} |u(x, t)|^2 dx,$$

$$a = \frac{1}{2} \left(\int_{\Omega} |u_0(x)|^2 dx + \int_0^T \int_{\Omega} f^2 dx ds \right)$$

et $b = 1$. D'après l'égalité d'énergie établie précédemment, pour tout $0 < t < T$,

$$\begin{aligned} & z(t) + \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds \\ & \leq \frac{1}{2} \left(\int_{\Omega} |u_0(x)|^2 dx + \int_0^t \int_{\Omega} |f(x, s)|^2 + |u(x, s)|^2 dx ds \right) \\ & \leq a + \int_0^t z(s) ds. \end{aligned}$$

D'après le lemme de Gronwall, $z(t) \leq ae^t$. En intégrant cette inégalité, on obtient

$$a + \int_0^t z(s) ds \leq ae^t.$$

Cette dernière, combinée à la précédente, implique que

$$\begin{aligned} & \frac{1}{2} \int_{\Omega} |u(x, t)|^2 dx + \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds \\ & \leq \frac{1}{2} \left(\int_{\Omega} |u_0(x)|^2 dx + \int_0^T \int_{\Omega} |f(x, s)|^2 dx ds \right) e^t. \end{aligned}$$

Exercice 8.2.2 Au vu de l'estimation

$$\int_{\Omega} u(x, t)^2 dx + \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds \leq C \left(\int_{\Omega} u_0(x)^2 dx + \int_0^t \int_{\Omega} f(x, s)^2 dx ds \right), \quad (8.4)$$

vérifiée par la solution u de (8.1), où la constante C est indépendante de T , on voit que le terme e^t n'est certainement pas optimal dans la majoration (8.3). Cette estimation peut être améliorée en raisonnant de la façon suivante, avec une variante du lemme de Gronwall.

1. Soit $a \in \mathbb{R}^+$ et $g \in L^2(]0, T[)$ tel que $g \geq 0$. Montrer que, si $z(t)$ est continue de $[0, T]$ dans \mathbb{R}^+ et vérifie

$$z(t) \leq a + 2 \int_0^t g(s) \sqrt{z(s)} ds \quad \forall t \in [0, T],$$

alors

$$z(t) \leq \left(\sqrt{a} + \int_0^t g(s) ds \right)^2 \quad \forall t \in [0, T].$$

2. Dédurre de (8.2) que, pour tout $t \in [0, T]$,

$$\begin{aligned} \int_{\Omega} u(x, t)^2 dx + 2 \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds &\leq \left(\left(\int_{\Omega} u_0(x)^2 dx \right)^{1/2} \right. \\ &\quad \left. + \int_0^t ds \left(\int_{\Omega} f(x, s)^2 dx \right)^{1/2} \right)^2. \end{aligned} \quad (8.5)$$

Correction.

1. On suppose dans un premier temps que g est une fonction régulière. Soit ε un réel strictement positif. On pose

$$v(t) = \varepsilon + a + 2 \int_0^t g(s) \sqrt{z(s)} ds.$$

Comme $g(s)\sqrt{z(s)}$ est une fonction continue, la fonction v est dérivable et $v'(t) = 2g(t)\sqrt{z(t)}$. Comme $z(t) \leq v(t)$ et que g est une fonction positive,

$$v'(t) \leq 2g(t)\sqrt{v(t)}.$$

Enfin, $v(t) > 0$, ainsi d'après l'inégalité précédente, $v'(t)/2\sqrt{v(t)} \leq g(t)$ et par intégration, on obtient

$$\sqrt{v(t)} - \sqrt{v(0)} \leq \int_0^t g(s) ds.$$

Ainsi, pour tout $\varepsilon > 0$,

$$z(t) \leq v(t) \leq \left(\sqrt{a + \varepsilon} + \int_0^t g(s) ds \right)^2.$$

Il suffit de passer à la limite lorsque ε tend vers zéro pour obtenir l'inégalité désirée.

Dans le cas général, on raisonne par densité. Soit $g \in L^2(]0; T[)$ tel que $g \geq 0$ presque partout. Il existe une suite de fonctions régulières g_n positives, convergeant vers g dans $L^2(]0; T[)$. Pour tout n , on a pour tout $t \in [0; T]$,

$$z(t) \leq a + \|g_n - g\|_{L^2(]0; T[)} \|z\|_{L^1(]0; T[)}^{1/2} + 2 \int_0^t g_n(s) \sqrt{z(s)} ds.$$

D'après ce qui précède,

$$\begin{aligned} z(t) &\leq a + \|g_n - g\|_{L^2(]0; T[)} \|z\|_{L^1(]0; T[)}^{1/2} + 2 \int_0^t g_n(s) \sqrt{z(s)} ds \\ &\leq \left(\sqrt{a + \|g_n - g\|_{L^2(]0; T[)} \|z\|_{L^1(]0; T[)}^{1/2}} + \int_0^t g_n(s) ds \right)^2. \end{aligned}$$

Il suffit alors de passer à la limite lorsque n tend vers l'infini pour conclure.

2. D'après l'égalité d'énergie (8.2) et l'inégalité de Cauchy-Schwarz,

$$\begin{aligned} &\int_{\Omega} u(x, t)^2 dx + 2 \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds \\ &\leq \int_{\Omega} u_0(x)^2 dx + 2 \int_0^t \left(\int_{\Omega} f(x, s)^2 dx \right)^{1/2} \left(\int_{\Omega} u(x, s)^2 dx \right)^{1/2} ds \end{aligned}$$

On applique la variante du Lemme de Gronwall à

$$\begin{aligned} z(t) &= \int_{\Omega} u(x, t)^2 dx + 2 \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds \\ g(s) &= \left(\int_{\Omega} f(x, s)^2 dx \right)^{1/2} \\ a &= \int_{\Omega} u_0(x)^2 dx. \end{aligned}$$

Ainsi,

$$\begin{aligned} &\int_{\Omega} u(x, t)^2 dx + 2 \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds \\ &\leq \left(\left(\int_{\Omega} u_0(x)^2 dx \right)^{1/2} + \int_0^t \left(\int_{\Omega} f(x, s)^2 dx \right)^{1/2} ds \right)^2. \end{aligned}$$

Exercice 8.2.3 On suppose que les hypothèses du Théorème 8.2.7 sont vérifiées, que $u_0 \in H_0^1(\Omega)$, et que la solution u de (8.1) est assez régulière dans $]0, T[\times \Omega$. Montrer que, pour tout $t \in [0, T]$, on a l'égalité d'énergie suivante

$$\begin{aligned} \frac{1}{2} \int_{\Omega} |\nabla u(x, t)|^2 dx + \int_0^t \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, s) \right|^2 dx ds &= \frac{1}{2} \int_{\Omega} |\nabla u_0(x)|^2 dx \\ &+ \int_0^t \int_{\Omega} f(x, s) \frac{\partial u}{\partial t}(x, s) dx ds. \end{aligned} \tag{8.6}$$

Correction. En multipliant l'équation (8.1) vérifiée par u par $\frac{\partial u}{\partial t}$, on obtient, suite à une intégration sur Ω que

$$\int_{\Omega} -\Delta u(x, t) \frac{\partial u}{\partial t}(x, t) dx + \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 dx = \int_{\Omega} f(x, t) \frac{\partial u}{\partial t}(x, t) dx.$$

Par intégration par partie, il vient

$$\int_{\Omega} \nabla u(x, t) \cdot \frac{\partial \nabla u}{\partial t}(x, t) dx + \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 dx = \int_{\Omega} f(x, t) \frac{\partial u}{\partial t}(x, t) dx,$$

ou encore en échangeant les signes dérivation et intégrale,

$$\frac{d}{dt} \left(\frac{1}{2} \int_{\Omega} |\nabla u|^2 dx \right) + \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 dx = \int_{\Omega} f(x, t) \frac{\partial u}{\partial t}(x, t) dx.$$

Il suffit d'intégrer cette dernière équation suivant t pour obtenir l'égalité recherchée.

Exercice 8.2.4 Soit Ω un ouvert borné régulier de \mathbb{R}^N . Soit un temps final $T > 0$, une donnée initiale $u_0 \in L^2(\Omega)$, et un terme source $f \in L^2(]0, T[; L^2(\Omega))$. Montrer que l'équation de la chaleur avec condition aux limites de Neumann

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = f & \text{dans } \Omega \times]0, T[\\ \frac{\partial u}{\partial n} = 0 & \text{sur } \partial\Omega \times]0, T[\\ u(x, 0) = u_0(x) & \text{dans } x \in \Omega \end{cases} \quad (8.7)$$

admet une unique solution $u \in L^2(]0, T[; H^1(\Omega)) \cap C([0, T]; L^2(\Omega))$.

Correction. On applique le Théorème 8.2.3 à $V = H^1(\Omega)$, $H = L^2(\Omega)$ et à la forme bilinéaire symétrique, continue sur V

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v dx.$$

La forme bilinéaire $a(., .)$ n'est pas coercive, mais $a(u, v) + \langle u, v \rangle_{L^2}$ étant coercive sur V , les conclusions du théorème restent valables d'après la remarque 8.2.5. Le problème (8.7) admet donc une unique solution

$$u \in L^2(]0, T[; H^1(\Omega)) \cap C([0, T]; L^2(\Omega)).$$

Exercice 8.2.5 Soit Ω un ouvert borné régulier de \mathbb{R}^N . Soit $A(x)$ une fonction de Ω dans l'ensemble des matrices symétriques réelles telles qu'il existe deux constantes $\beta \geq \alpha > 0$ vérifiant

$$\beta |\xi|^2 \geq A(x) \xi \cdot \xi \geq \alpha |\xi|^2 \quad \forall \xi \in \mathbb{R}^N, \text{ p.p. } x \in \Omega.$$

Soit un temps final $T > 0$, une donnée initiale $u_0 \in L^2(\Omega)$, et un terme source $f \in L^2(]0, T[; L^2(\Omega))$. Montrer que le problème aux limites

$$\begin{cases} \frac{\partial u}{\partial t} - \operatorname{div}(A(x) \nabla u) = f & \text{dans } \Omega \times]0, T[\\ u = 0 & \text{sur } \partial\Omega \times]0, T[\\ u(x, 0) = u_0(x) & \text{pour } x \in \Omega, \end{cases} \quad (8.8)$$

admet une unique solution $u \in L^2(]0, T[; H^1(\Omega)) \cap C([0, T]; L^2(\Omega))$.

Correction. On introduit la forme bilinéaire $a(\cdot, \cdot)$ symétrique définie pour tout u et v de $H_0^1(\Omega)$ par

$$a(u, v) = \int_{\Omega} A(x) \nabla u \cdot \nabla v dx.$$

Pour presque tout $x \in \Omega$, la matrice $A(x)$ étant symétrique, définie positive, elle admet une base de vecteurs propres. Comme pour tout $\xi \in \mathbb{R}^N$,

$$A(x)\xi \cdot \xi \leq \beta |\xi|^2,$$

la plus grande valeur propre de $A(x)$ est inférieure à β et $\|A\|_2 \leq \beta$. D'après cette majoration et l'inégalité de Cauchy-Schwarz, pour tout u et $v \in H_0^1(\Omega)$, on a

$$a(u, v) \leq \beta \int_{\Omega} \nabla u \cdot \nabla v dx \leq \beta \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}.$$

La forme bilinéaire a est donc continue sur $H_0^1(\Omega)$. De plus, pour tout $u \in H_0^1(\Omega)$, d'après l'inégalité de Poincaré,

$$a(u, u) \geq \alpha \int_{\Omega} |\nabla u|^2 dx \geq \alpha C \|u\|_{H_0^1(\Omega)}^2.$$

La forme bilinéaire a est donc coercive. D'après le Théorème 8.2.3 appliqué à la forme bilinéaire a avec $H = L^2(\Omega)$ et $V = H_0^1(\Omega)$, il existe une unique solution $u \in L^2(]0, T[; H^1(\Omega)) \cap C([0, T]; L^2(\Omega))$ au problème aux limites (8.8).

Exercice 8.3.1 Soit Ω un ouvert borné régulier de \mathbb{R}^N , et un temps final $T > 0$. On considère une donnée initiale $(u_0, u_1) \in H_0^1(\Omega) \times L^2(\Omega)$ et un terme source $f \in L^2(]0, T[; L^2(\Omega))$. On considère la solution u de l'équation des ondes

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - \Delta u = f & \text{p.p. dans } \Omega \times]0, T[\\ u = 0 & \text{p.p. sur } \partial\Omega \times]0, T[\\ u(x, 0) = u_0(x) & \text{p.p. dans } \Omega \\ \frac{\partial u}{\partial t}(x, 0) = u_1(x) & \text{p.p. dans } \Omega. \end{cases} \quad (8.9)$$

1. En supposant que la solution u de (8.9) est assez régulière dans $]0, T[\times \Omega$, montrer que, pour tout $t \in [0, T]$, on a l'égalité d'énergie suivante

$$\begin{aligned} \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 dx + \int_{\Omega} |\nabla u(x, t)|^2 dx &= \int_{\Omega} u_1(x)^2 dx + \int_{\Omega} |\nabla u_0(x)|^2 dx \\ &\quad + 2 \int_0^t \int_{\Omega} f(x, s) \frac{\partial u}{\partial t}(x, s) dx ds. \end{aligned}$$

2. En déduire qu'il existe une constante $C(T)$ (indépendante des données autre que T) telle que

$$\begin{aligned} \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 dx + \int_{\Omega} |\nabla u(x, t)|^2 dx &\leq \\ C(T) \left(\int_{\Omega} u_1(x)^2 dx + \int_{\Omega} |\nabla u_0(x)|^2 dx + \int_0^t \int_{\Omega} f(x, s)^2 dx ds \right). \end{aligned}$$

3. Montrer qu'il existe une constante C (indépendante de toutes les données y compris T) telle que

$$\int_{\Omega} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 dx + \int_{\Omega} |\nabla u(x, t)|^2 dx \leq C \left(\int_{\Omega} u_1(x)^2 dx + \int_{\Omega} |\nabla u_0(x)|^2 dx + \left(\int_0^t \left(\int_{\Omega} f(x, s)^2 dx \right)^{1/2} ds \right)^2 \right).$$

Correction.

1. Supposons que u soit une solution suffisamment régulière, de l'équation des ondes. On a

$$\frac{\partial u}{\partial t} \frac{\partial^2 u}{\partial t^2} - \frac{\partial u}{\partial t} \Delta u = f \frac{\partial u}{\partial t}.$$

Par intégration sur le domaine Ω , il vient en échangeant les opérateurs d'intégration en espace et de dérivation en temps (ce qui est licite pour u régulière) que

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^2 dx + \frac{1}{2} \frac{d}{dt} \int_{\Omega} |\nabla u|^2 dx = \int_{\Omega} f \frac{\partial u}{\partial t} dx.$$

Par intégration en temps, on obtient l'égalité voulue.

2. D'après l'inégalité de Schwarz,

$$\begin{aligned} \int_0^t \int_{\Omega} f(x, s) \frac{\partial u}{\partial t}(x, s) dx ds &\leq \left(\int_0^t \int_{\Omega} f(x, s)^2 dx ds \right)^{1/2} \left(\int_0^t \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, s) \right|^2 dx ds \right)^{1/2} \\ &\leq \frac{1}{2} \left(\int_0^t \int_{\Omega} f^2(x, s) dx ds + \int_0^t \int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^2 dx ds \right). \end{aligned}$$

En appliquant cette inégalité à l'égalité précédemment obtenue, on obtient que

$$\begin{aligned} \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 dx + \int_{\Omega} |\nabla u(x, t)|^2 dx &\leq \int_{\Omega} |u_1(x)|^2 dx + \int_{\Omega} |\nabla u_0|^2 dx + \int_0^t \int_{\Omega} f^2(x, s) dx ds + \int_0^t \int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^2 dx ds \end{aligned}$$

D'après le Lemme de Gronwall (voir Exercice 8.2.1), on en déduit que pour tout $t \leq T$,

$$\begin{aligned} \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 dx + \int_{\Omega} |\nabla u(x, t)|^2 dx &\leq e^t \left(\int_{\Omega} |u_1(x)|^2 dx + \int_{\Omega} |\nabla u_0(x)|^2 dx + \int_0^t \int_{\Omega} f^2(x, s) dx ds \right) \end{aligned}$$

3. De l'égalité obtenue dans la première partie de l'exercice, on déduit à l'aide de l'inégalité de Schwarz que

$$\begin{aligned} & \int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^2 dx + \int_{\Omega} |\nabla u(x, t)|^2 dx \\ & \leq \int_{\Omega} |u_1(x)|^2 + |\nabla u_0|^2 dx + 2 \int_0^t \left(\int_{\Omega} f^2(x, s) dx \right)^{1/2} \left(\int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^2 dx \right)^{1/2} ds. \end{aligned}$$

D'après la variante du Lemme de Gronwall (voir Exercice 8.2.2),

$$\begin{aligned} & \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 dx + \int_{\Omega} |\nabla u(x, t)|^2 dx \\ & \leq \left(\left(\int_{\Omega} |u_1(x)|^2 + |\nabla u_0(x)|^2 dx \right)^{1/2} + \int_0^t \left(\int_{\Omega} f^2(x, s) dx \right)^{1/2} ds \right)^2 \end{aligned}$$

d'où on déduit l'estimation recherchée avec $C = 2$ (il suffit d'utiliser l'inégalité $(a + b)^2 \leq 2(a^2 + b^2)$).

Exercice 8.3.2 On suppose que les hypothèses du Théorème 8.3.4 sont vérifiées, que le terme source est nul $f = 0$ et que la solution u de (8.9) est régulière dans $[0, T] \times \Omega$. Montrer que, pour tout entier $m \geq 1$, on a

$$\frac{d}{dt} \int_{\Omega} \left(\left| \frac{\partial^m u}{\partial t^m} \right|^2 + \left| \nabla \frac{\partial^{m-1} u}{\partial t^{m-1}} \right|^2 \right) dx = 0.$$

Correction. Il suffit de remarquer que la fonction $\partial^m u / \partial t^m$ est elle-même solution d'une équation d'onde avec conditions de Dirichlet homogènes au bord, sans terme force.

Exercice 8.3.3 Soit Ω un ouvert borné régulier connexe de \mathbb{R}^N . Soit une donnée initiale $(u_0, u_1) \in H_0^1(\Omega)^N \times L^2(\Omega)^N$, et un terme source $f \in L^2([0, T]; L^2(\Omega))^N$. Montrer qu'il existe une unique solution $u \in C([0, T]; H_0^1(\Omega))^N \cap C^1([0, T]; L^2(\Omega))^N$ de

$$\begin{cases} \rho \frac{\partial^2 u}{\partial t^2} - \operatorname{div} (2\mu e(u) + \lambda \operatorname{tr}(e(u)) \operatorname{Id}) = f & \text{dans } \Omega \times]0, T[, \\ u = 0 & \text{sur } \partial\Omega \times]0, T[, \\ u(t = 0) = u_0(x) & \text{dans } \Omega, \\ \frac{\partial u}{\partial t}(t = 0) = u_1(x) & \text{dans } \Omega. \end{cases} \quad (8.10)$$

En supposant que la solution u est assez régulière, montrer que, pour tout $t \in [0, T]$, on a l'égalité d'énergie suivante

$$\begin{aligned} & \frac{\rho}{2} \int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^2 dx + \mu \int_{\Omega} |e(u)|^2 dx + \frac{\lambda}{2} \int_{\Omega} (\operatorname{div} u)^2 dx = \frac{\rho}{2} \int_{\Omega} |u_1|^2 dx \\ & + \mu \int_{\Omega} |e(u_0)|^2 dx + \frac{\lambda}{2} \int_{\Omega} (\operatorname{div} u_0)^2 dx + \int_0^t \int_{\Omega} f \cdot \frac{\partial u}{\partial t} dx ds. \end{aligned}$$

En déduire une estimation d'énergie.

Correction. On introduit la forme bilinéaire $a(\cdot, \cdot)$ définie pour tout u et $v \in H_0^1(\Omega)^N$ par

$$a(u, v) = \int_{\Omega} 2\mu e(u) \cdot e(v) + \lambda(\operatorname{div} u)(\operatorname{div} v) dx.$$

La formulation variationnelle associée au système d'équations aux dérivées partielles consiste à déterminer $u \in C([0, T]; H_0^1(\Omega)^N) \cap C^1([0, T]; L^2(\Omega)^N)$ tel que

$$\begin{cases} \frac{d^2 \langle u(t), v \rangle_{L^2}}{dt^2} + a(u, v) = \int_{\Omega} f \cdot v dx \\ u(t=0) = u_0; \quad \frac{du}{dt}(t=0) = u_1 \end{cases}$$

pour tout $v \in H_0^1(\Omega)^N$. La forme bilinéaire bilinéaire a est continue et coercive sur $H_0^1(\Omega)^N$. La coercivité de la forme bilinéaire a est établie dans les preuves du Théorèmes 5.3.1 et 5.3.4 et découle du le Lemme de Korn 5.3.3 ou de sa version simplifiée 5.3.2. Les hypothèses du Théorème 8.3.1 sont vérifiées, il existe une unique solution à la formulation variationnelle. En appliquant la fonction test $v = \partial u / \partial t$ à la formulation variationnelle, on en déduit que

$$\frac{d}{dt} \left(\frac{\rho}{2} \int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^2 dx + \mu \int_{\Omega} |e(u)|^2 dx + \frac{\lambda}{2} \int_{\Omega} (\operatorname{div} u)^2 dx \right) = \int_{\Omega} f \cdot \frac{\partial u}{\partial t} dx.$$

L'égalité d'énergie en découle par une simple intégration. En procédant comme pour l'Exercice 8.3.1 on obtient les estimations d'énergie suivantes. Tout d'abord, pour tout T il existe une constante $C(T)$ ne dépendant que du temps final T telle que

$$\begin{aligned} & \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 dx + \int_{\Omega} 2\mu |e(u)|^2 + \lambda(\operatorname{div} u)^2 dx \leq \\ & C(T) \left(\int_{\Omega} |u_1(x)|^2 dx + \int_{\Omega} 2\mu |e(u_0)|^2 + \lambda(\operatorname{div} u_0)^2 dx + \int_0^t \int_{\Omega} |f(x, s)|^2 dx ds \right). \end{aligned}$$

De plus, il existe une constante C (indépendante de toutes les données y compris T) telle que

$$\begin{aligned} & \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 dx + \int_{\Omega} 2\mu |e(u)|^2 + \lambda(\operatorname{div} u)^2 dx \leq C \left(\int_{\Omega} |u_1|^2 dx + \right. \\ & \left. \int_{\Omega} 2\mu |e(u_0)|^2 + \lambda(\operatorname{div} u_0)^2 dx + \left(\int_0^t \left(\int_{\Omega} |f(x, s)|^2 dx \right)^{1/2} ds \right)^2 \right). \end{aligned}$$

Exercice 8.4.1 On reprend les hypothèses de la Proposition 8.4.1. Soit $f(x) \in L^2(\Omega)$ et u la solution de

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = f & \text{dans }]0, +\infty[\times \Omega \\ u(x, t) = 0 & \text{sur }]0, +\infty[\times \partial\Omega \\ u(x, 0) = u_0(x) & \text{dans } \Omega. \end{cases}$$

Soit $v(x) \in H_0^1(\Omega)$ la solution de

$$\begin{cases} -\Delta v = f & \text{dans } \Omega \\ v = 0 & \text{sur } \partial\Omega. \end{cases}$$

Montrer que $\lim_{t \rightarrow +\infty} \|u(x, t) - v(x)\|_{L^2(\Omega)} = 0$.

Correction.

On pose $\tilde{u}(x, t) = u(x, t) - v(x)$. La fonction \tilde{u} est solution de l'équation de la chaleur avec conditions de Dirichlet homogènes et condition initiale $\tilde{u}(x, 0) = u_0(x) - v(x)$. Ainsi, d'après la Proposition 8.4.1,

$$\lim_{t \rightarrow +\infty} \|u(x, t) - v(x)\|_{L^2(\Omega)} = 0.$$

Exercice 8.4.2 Soit Ω un ouvert borné régulier de \mathbb{R}^N . Soit $u_0 \in L^2(\Omega)$ et u la solution du problème

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = 0 & \text{dans }]0, +\infty[\times \Omega \\ u(x, t) = 0 & \text{sur }]0, +\infty[\times \partial\Omega \\ u(x, 0) = u_0(x) & \text{dans } \Omega. \end{cases} \quad (8.11)$$

Montrer qu'il existe une constante positive C telle que

$$\|u(t) - \alpha_1^0 e^{-\lambda_1 t} u_1\|_{L^2(\Omega)} \leq C e^{-\lambda_2 t} \quad \forall t > 1, \quad \text{avec } \alpha_1^0 = \int_{\Omega} u_0 u_1 dx, \quad (8.12)$$

où λ_k désigne la k -ème valeur propre du Laplacien avec condition aux limites de Dirichlet.

Correction.

On rappelle que la solution u de l'équation (8.11) est donnée par

$$u(t) = \sum_{k=1}^{\infty} \alpha_k^0 e^{-\lambda_k t} u_k.$$

Ainsi,

$$u(t) - \alpha_1^0 e^{-\lambda_1 t} u_1 = \sum_{k=2}^{\infty} \alpha_k^0 e^{-\lambda_k t} u_k.$$

et

$$\|u(t) - \alpha_1^0 e^{-\lambda_1 t} u_1\|_{L^2(\Omega)} = e^{-\lambda_2 t} \left(\sum_{k=2}^{\infty} |\alpha_k^0|^2 e^{-2(\lambda_k - \lambda_2)t} \right)^{1/2}.$$

Comme $\lambda_k - \lambda_2 \geq 0$, on en déduit que

$$\|u(t) - \alpha_1^0 e^{-\lambda_1 t} u_1\|_{L^2(\Omega)} \leq e^{-\lambda_2 t} \left(\sum_{k=2}^{\infty} |\alpha_k^0|^2 \right)^{1/2} \leq e^{-\lambda_2 t} \|u_0\|_{L^2(\Omega)}$$

Exercice 8.4.3 Soit Ω un ouvert borné régulier de \mathbb{R}^N . On note u_1 la première fonction propre du Laplacien dans Ω avec condition de Dirichlet, λ_1 la valeur propre associée. On rappelle que l'on peut choisir $u_1 > 0$ dans Ω (voir le Théorème de Krein-Rutman 7.3.10) et on admettra que l'on a aussi $\partial u_1 / \partial n > 0$ sur $\partial\Omega$. Soit $f = 0$, $u_0 \in L^2(\Omega)$ et u l'unique solution (supposée régulière) de (8.1).

Soit $\epsilon > 0$. Montrer que l'on peut trouver une constante positive K telle que

$$-Ku_1(x) \leq u(x, \epsilon) \leq Ku_1(x) \quad \forall x \in \overline{\Omega}, \quad (8.13)$$

et en déduire qu'il existe une constante positive C telle que

$$\max_{x \in \overline{\Omega}} |u(x, t)| \leq Ce^{-\lambda_1 t} \quad \forall t > \epsilon. \quad (8.14)$$

Correction. Pour tout $\varepsilon > 0$, $u(x, \varepsilon)$ est une fonction de classe $C^\infty(\overline{\Omega})$. Rappelons que $u_1(x)$ est également une fonction régulière sur $\overline{\Omega}$, qu'elle est strictement positive sur Ω et que $\partial u_1 / \partial n > 0$ sur $\partial\Omega$.

Soit

$$K_1 = 1 + \sup_{x \in \partial\Omega} \left| \frac{\partial u}{\partial n}(x, \varepsilon) \frac{\partial u_1^{-1}}{\partial n}(x) \right|.$$

On introduit les fonctions v_+ et v_- définies par

$$\begin{aligned} v_+(x) &= K_1 u_1(x) - u(x, \varepsilon) \\ v_-(x) &= K_1 u_1(x) + u(x, \varepsilon) \end{aligned}$$

On vérifie sans mal que $\partial v_\pm / \partial n > 0$ sur $\partial\Omega$. Il existe donc un voisinage ω de $\partial\Omega$ tel que pour tout $x \in \omega$,

$$v_\pm(x) \geq 0.$$

Il existe un compact $A \subset \Omega$ tel que $A \cup \omega = \Omega$. On pose

$$K_2 = \max_{x \in A} |u(x, \varepsilon) / u_1(x)|$$

et $K = \max(K_1, K_2)$. On vérifie sans peine que

$$-Ku_1(x) \leq u(x, \varepsilon) \leq Ku_1(x).$$

La fonction $\tilde{u}(x, t) = Ke^{-\lambda_1(t-\varepsilon)}u_1$ est une solution de l'équation de la chaleur (8.1) sur $t > \varepsilon$ avec $f = 0$ et $\tilde{u}(x, \varepsilon) = Ku_1(x)$ comme condition initiale. Enfin, comme

$$-\tilde{u}(x, \varepsilon) \leq u(x, \varepsilon) \leq \tilde{u}(x, \varepsilon),$$

on déduit du principe du maximum de la Proposition 8.4.2 que

$$-\tilde{u}(x, t) \leq u(x, t) \leq \tilde{u}(x, t)$$

pour tout $t \geq \varepsilon$. On a donc montré que

$$|u(x, t)| \leq \left(Ke^{\lambda_1 \varepsilon} \max_{x \in \Omega} u_1(x) \right) e^{-\lambda_1 t}$$

Exercice 8.4.4 Soit Ω un ouvert borné régulier de \mathbb{R}^N . Soit $u_0 \in L^\infty(\Omega)$, $f \in L^\infty(\Omega \times \mathbb{R}_*^+)$, et $u \in C([0, T]; L^2(\Omega)) \cap L^2(]0, T[; H_0^1(\Omega))$ l'unique solution de (8.1). Montrer que

$$\|u\|_{L^\infty(\Omega \times \mathbb{R}_*^+)} \leq \|u_0\|_{L^\infty(\Omega)} + \frac{D^2}{2N} \|f\|_{L^\infty(\Omega \times \mathbb{R}_*^+)}, \quad (8.15)$$

où $D = \sup_{x, y \in \Omega} |x - y|$ est le diamètre de Ω . On pourra utilement introduire la fonction $\psi \in H_0^1(\Omega)$ telle que $-\Delta\psi = 1$ dans Ω .

Correction. Remarquons tout d'abord qu'il suffit de montrer que pour tout $(t, x) \in \mathbb{R}_*^+ \times \Omega$,

$$u(x, t) \leq \|u_0\|_{L^\infty(\Omega)} + \frac{D^2}{2N} \|f\|_{L^\infty(\Omega \times \mathbb{R}_*^+)}.$$

En appliquant ce résultat à $-u$ au lieu de u et en combinant les deux estimations obtenues, on prouve l'estimation souhaitée.

Introduisons la fonction u_+ solution de

$$\begin{cases} \frac{\partial u_+}{\partial t} - \Delta u_+ = \|f\|_{L^\infty(\Omega \times \mathbb{R}_*^+)} & \text{dans }]0; T[\times \Omega \\ u_+(x, t) = 0 & \text{sur }]0; T[\times \partial\Omega \\ u_+(x, 0) = \|u_0\|_{L^\infty(\Omega)} & \text{dans } \Omega. \end{cases}$$

D'après le principe du maximum, $u \leq u_+$. Il suffit donc de prouver le résultat pour $u = u_+$. Dans un premier temps, on considère le cas $f = 0$. Il s'agit de prouver que pour presque tout $(t, x) \in]0, T[\times \Omega$, on a $|u(x, t)| \leq \|u_0\|_{L^\infty(\Omega)}$. D'après le principe du maximum de la Proposition 8.4.2, on a déjà $u \geq 0$. Reste à prouver que $u \leq \|u_0\|_{L^\infty(\Omega)}$. A cet effet, on procède comme lors de la preuve de la Proposition 8.4.2. On introduit la fonction $\tilde{u} = \max(u - \|u_0\|_{L^\infty}, 0)$. En vertu du Lemme 5.2.24, $\tilde{u} \in L^2(]0, T[; H_0^1(\Omega))$ et

$$\int_{\Omega} \nabla u \cdot \nabla \tilde{u} dx = \int_{\Omega} |\nabla \tilde{u}|^2 dx.$$

De même, si u est suffisamment régulière (ce que nous admettrons par la suite),

$$\int_{\Omega} \frac{\partial u}{\partial t} \tilde{u} dx = \frac{1}{2} \frac{d}{dt} \left(\int_{\Omega} |\tilde{u}|^2 dx \right),$$

Remarque 8.4.1 D'après la Proposition 8.4.6, pour tout $\delta > 0$, la fonction u appartient à $H^1(]0, T[; L^2(\Omega))$. Elle est donc assez régulière pour que le Lemme de tronature 5.2.24 s'applique de sorte à justifier l'équation précédente.

Par conséquent, en multipliant l'équation vérifiée par u par \tilde{u} , on obtient par intégration sur Ω , puis par intégration par partie que

$$\frac{1}{2} \frac{d}{dt} \left(\int_{\Omega} |\tilde{u}|^2 dx \right) + \int_{\Omega} |\nabla \tilde{u}|^2 dx = 0.$$

En intégrant cette équation en temps, il vient

$$\frac{1}{2} \int_{\Omega} |\tilde{u}|^2 dx - \frac{1}{2} \int_{\Omega} |\tilde{u}(t=0)|^2 dx + \int_0^T \int_{\Omega} |\nabla \tilde{u}|^2 dx dt = 0.$$

Comme $\tilde{u}(t=0) = 0$, on en déduit que $\tilde{u} = 0$, c'est à dire $u \leq \|u_0\|_{L^\infty}$. On se place dorénavant dans le cas général (f non nécessairement nul). Soit ψ la solution du problème aux limites $\psi \in H_0^1(\Omega)$, $-\Delta\psi = 1$. On pose $v = u_+ - \|f\|_{L^\infty(\Omega \times \mathbb{R}_+^*)}\psi$. La fonction v est solution du problème

$$\begin{cases} \frac{\partial v}{\partial t} - \Delta v = 0 & \text{dans }]0; T] \\ v(x, t) = 0 & \text{sur }]0; T[\times \partial\Omega \\ v(x, 0) = \|u_0\|_{L^\infty(\Omega)} - \|f\|_{L^\infty(\Omega \times \mathbb{R}_+^*)}\psi & \text{dans } \Omega. \end{cases}$$

Comme $\psi \geq 0$, pour tout $x \in \Omega$ on a $v(x, 0) \leq \|u_0\|_{L^\infty(\Omega)}$. Ainsi, pour tout t ,

$$v(x, t) \leq \|u_0\|_{L^\infty(\Omega)}. \quad (8.16)$$

On a donc obtenu que $u_+ \leq \|u_0\|_{L^\infty} + \|f\|_{L^\infty}\psi$. Il reste à majorer ψ afin d'obtenir l'estimation souhaitée. Sans perte de généralité, on peut supposer que l'origine de \mathbb{R}^N appartient au bord de Ω . Comme

$$-\Delta|x|^2/(2N) = 1 = -\Delta\psi \text{ dans } \Omega$$

et

$$|x|^2/(2N) \geq 0 = \psi(x) \text{ sur } \partial\Omega,$$

le principe du maximum implique que $|x|^2/2N \geq \psi(x)$. Or $|x|$ est majoré sur Ω par de diamètre de Ω , ainsi $\|\psi\|_{L^\infty(\Omega)} \leq D^2/2N$, ce qui achève la preuve.

Exercice 8.4.5 (difficile) Démontrer rigoureusement la Proposition 8.4.5. Pour cela on introduira, pour tout entier $m \geq 0$, l'espace

$$W^{2m}(\Omega) = \{v \in H^{2m}(\Omega), v = \Delta v = \dots \Delta^{m-1}v = 0 \text{ sur } \partial\Omega\}, \quad (8.17)$$

que l'on munit de la norme $\|v\|_{W^{2m}(\Omega)}^2 = \int_\Omega |(\Delta)^m v|^2 dx$, dont on montrera qu'elle est équivalente à la norme de $H^{2m}(\Omega)$. On reprendra la démonstration du Théorème 8.2.3 en montrant que la suite (w_k) des sommes partielles est de Cauchy dans $C^\ell([\epsilon, T], W^{2m}(\Omega))$.

Correction.

La démonstration se fait par récurrence sur m . Pour $m = 1$, en posant $f = \Delta v$, le Théorème 5.2.26 de régularité nous dit exactement que, si $f \in L^2(\Omega)$ et $v \in H_0^1(\Omega)$ alors $v \in H^2(\Omega)$, c'est à dire que

$$\|v\|_{H^2(\Omega)} \leq C\|\Delta v\|_{L^2(\Omega)} \text{ pour tout } v \in H_0^1(\Omega).$$

L'inégalité inverse est évidente, d'où l'équivalence des normes dans le cas $m = 1$. Supposons que $\|v\|_{W^{2(m-1)}(\Omega)}$ est une norme équivalente à $\|v\|_{H^{2(m-1)}(\Omega)}$ pour les fonctions de $W^{2(m-1)}(\Omega)$. Le Théorème de régularité 5.2.26 nous dit aussi que

$$\|v\|_{H^{2m}(\Omega)} \leq C\|\Delta v\|_{H^{2(m-1)}(\Omega)} \text{ pour tout } v \in H_0^1(\Omega),$$

c'est à dire, en utilisant l'hypothèse de récurrence pour $v \in W^{2m}(\Omega)$,

$$\|v\|_{H^{2m}(\Omega)} \leq C\|\Delta v\|_{W^{2(m-1)}(\Omega)} = C\|\Delta^{m-1}(\Delta v)\|_{L^2(\Omega)} = C\|v\|_{W^{2m}(\Omega)},$$

ce qui prouve que $\|v\|_{W^{2m}(\Omega)}$ est une norme équivalente à $\|v\|_{H^{2m}(\Omega)}$ pour les fonctions de $W^{2m}(\Omega)$ (l'inégalité inverse est évidente).

On se propose de montrer que $u \in C^\ell([\varepsilon, T], W^{2m}(\Omega))$. A cet effet, il suffit de prouver que la suite w_k des sommes partielles introduites dans la preuve du Théorème 8.2.3 est de Cauchy pour la norme $C^\ell([\varepsilon, T], W^{2m}(\Omega))$. On rappelle que

$$w^k(t) = \sum_{j=1}^k \alpha_0^j e^{-\lambda_j t} u_j.$$

Ainsi, soit l et k deux entiers naturels non nuls,

$$\left\| \frac{\partial^\ell (w^k - w^l)}{\partial t^\ell}(t) \right\|_{W^{2m}(\Omega)}^2 = \int_{\Omega} \left| \sum_{j=k}^l \alpha_0^j (-\lambda_j)^\ell e^{-\lambda_j t} (\Delta)^m u_j \right|^2 dx.$$

Comme u_j est une base de vecteur propres orthonormale du Laplacien, $(\Delta)^m u_j = (-\lambda_j)^m u_j$ et

$$\left\| \frac{\partial^\ell (w^k - w^l)}{\partial t^\ell}(t) \right\|_{W^{2m}(\Omega)}^2 = \sum_{j=k}^l (\alpha_0^j (-\lambda_j)^{m+\ell} e^{-\lambda_j t})^2.$$

Or pour tout $\varepsilon > 0$, pour tout m et ℓ , il existe une constante $C(\varepsilon, m, \ell)$ telle que

$$|(-\lambda_j)^{m+\ell} e^{-\lambda_j t}|^2 \leq C(\varepsilon, m, \ell)$$

pour tout $t \geq \varepsilon$ et tout indice j . Ainsi, pour tout $t \geq \varepsilon$, on a

$$\left\| \frac{\partial^\ell (w^k - w^l)}{\partial t^\ell}(t) \right\|_{W^{2m}} \leq C(\varepsilon, m, \ell) \sum_{j=k}^l |\alpha_0^j|^2,$$

où le second membre tend vers zéro lorsque k et l tendent vers l'infini. La suite w^k est donc de Cauchy dans $C^\ell([\varepsilon, T]; W^{2m}(\Omega))$. Elle est donc convergente dans cet espace et $u \in C^\ell([\varepsilon, T]; W^{2m}(\Omega))$.

Exercice 8.4.6 Pour $u_0 \in L^2(\mathbb{R}^N)$ et $t > 0$, on pose

$$S(t)u_0 = \frac{1}{(4\pi t)^{N/2}} \int_{\mathbb{R}^N} u_0(y) e^{-\frac{|x-y|^2}{4t}} dy.$$

Vérifier que $S(t)$ est un opérateur linéaire continu de $L^2(\mathbb{R}^N)$ dans $L^2(\mathbb{R}^N)$. En posant $S(0) = \text{Id}$ (l'identité de $L^2(\mathbb{R}^N)$), vérifier que $(S(t))_{t \geq 0}$ est un semi-groupe d'opérateurs qui dépendent continûment de t , c'est-à-dire qu'ils vérifient $S(t+t') = S(t)S(t')$ pour $t, t' \geq 0$. Soit $f \in C^1(\mathbb{R}^+; L^2(\mathbb{R}^N))$. Montrer que le problème

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = f & \text{dans }]0, +\infty[\times \mathbb{R}^N \\ u(x, 0) = u_0(x) & \text{dans } \mathbb{R}^N. \end{cases}$$

admet une unique solution $u \in C(\mathbb{R}^+; L^2(\mathbb{R}^N)) \cap C^1(\mathbb{R}_*^+; L^2(\mathbb{R}^N))$, donnée par

$$u(t) = S(t)u_0 + \int_0^t S(t-s)f(s) ds,$$

c'est-à-dire

$$u(x, t) = \int_{\mathbb{R}^N} u_0(y) e^{-\frac{|x-y|^2}{4t}} \frac{dy}{(2\pi t)^{N/2}} + \int_0^t \int_{\mathbb{R}^N} f(y, s) e^{-\frac{|x-y|^2}{4(t-s)}} \frac{dy ds}{(2\pi(t-s))^{N/2}}.$$

Correction.

1. Étude de l'opérateur $S(t)$

On a

$$S(t)u_0 = \frac{1}{(2\pi t)^{N/2}} \int_{\mathbb{R}^N} u_0(y) e^{-\frac{|x-y|^2}{4t}} dy.$$

La linéarité de l'opérateur $S(t)$ est évidente. De plus, la norme $L^2(\mathbb{R}^N)$ de $S(t)u_0$ est égale à la norme $L^2(\mathbb{R}^N)$ de sa transformée de Fourier. Ainsi, pour tout $u_0 \in L^2(\mathbb{R}^N)$,

$$\|S(t)u_0\|_{L^2(\mathbb{R}^N)} = \|\hat{u}_0 e^{-|k|^2 t}\|_{L^2(\mathbb{R}^N)} \leq \|\hat{u}_0\|_{L^2(\mathbb{R}^N)} = \|u_0\|_{L^2(\mathbb{R}^N)}$$

et $S(t)$ est un opérateur continu de $L^2(\mathbb{R}^N)$ dans $L^2(\mathbb{R}^N)$.

Pour tout t , on note $\hat{S}(t)u_0$ la transformée de Fourier de $S(t)u_0$. On a $\hat{S}(t)u_0 = \hat{u}_0 e^{-|k|^2 t}$. Ainsi,

$$\hat{S}(t+t')u_0 = \hat{u}_0 e^{-|k|^2 t} \cdot e^{-|k|^2 t'} = \hat{S}(t)u_0 e^{-|k|^2 t'} = \hat{S}(t')(\hat{S}(t)u_0),$$

en appliquant la transformée de Fourier inverse, on obtient que

$$S(t+t')(u_0) = S(t')(S(t)u_0).$$

Ainsi, $S(t+t') = S(t')S(t)$. Reste à montrer que les opérateurs $S(t)$ dépendent continûment de t . Comme $(S(t))_{t \geq 0}$ est un semi groupe, il suffit de vérifier cette propriété en $t = 0$, c'est à dire que pour tout $\varepsilon > 0$, il existe $T > 0$ tel que pour tout $t < T$ et tout $u_0 \in L^2(\mathbb{R}^N)$,

$$\|S(t)u_0 - u_0\|_{L^2(\mathbb{R}^N)} \leq \varepsilon \|u_0\|_{L^2(\mathbb{R}^N)}.$$

A nouveau, il est beaucoup plus simple de raisonner sur les transformées de Fourier, la relation ci-dessus étant équivalente à

$$\|\hat{u}_0(e^{-|k|^2 t} - 1)\|_{L^2(\mathbb{R}^N)} \leq \varepsilon \|\hat{u}_0\|_{L^2(\mathbb{R}^N)},$$

qui est vérifiée dès que $t < T = -\frac{\ln(1-\varepsilon)}{|k|^2}$.

2. Équation de la chaleur non homogène

On pose

$$u(t) = S(t)u_0 + \int_0^t S(t-s)f(s)ds.$$

Montrons que u est dérivable par rapport au temps. Le premier terme de l'expression de u est dérivable, d'après la définition même de S et

$$\frac{\partial S(t)u_0}{\partial t} = -\Delta(S(t)u_0).$$

En effectuant un changement de variable sur le deuxième terme, il reste à prouver que

$$\int_0^t S(s)f(t-s)ds$$

est dérivable par rapport à t . Comme $f \in C(\mathbb{R}^+; L^2(\mathbb{R}^N))$, on en déduit que

$$\frac{\partial}{\partial t} \left(\int_0^t S(s)f(t-s)ds \right) = S(t)f(0) + \int_0^t S(s) \frac{\partial f}{\partial t}(t-s)ds.$$

On effectue à nouveau un changement de variable (dans l'autre sens cette fois), pour en déduire que

$$\frac{\partial}{\partial t} \left(\int_0^t S(s)f(t-s)ds \right) = S(t)f(0) + \int_0^t S(t-s) \frac{\partial f}{\partial s}(s)ds.$$

Remarquons enfin que

$$S(t-s) \frac{\partial f}{\partial s}(s) = \frac{\partial S(t-s)f(s)}{\partial t} + \frac{\partial S(t-s)f(s)}{\partial s}.$$

Ainsi, on déduit que

$$\frac{\partial}{\partial t} \int_0^t S(t-s)f(s)ds = S(0)f(t) - \int_0^t \Delta(S(t-s)f(s))ds.$$

On a donc montré que u était dérivable sur \mathbb{R}_*^+ et que

$$\frac{\partial u}{\partial t} = f(t) - \Delta u.$$

Enfin, l'unicité est évidente, l'équation étant linéaire et de solution unique lorsque $f = 0$.

Exercice 8.4.7 (égalité d'énergie) Montrer que, pour tout $T > 0$,

$$\frac{1}{2} \int_{\mathbb{R}^N} u(x, T)^2 dx + \int_0^T \int_{\mathbb{R}^N} |\nabla u(x, t)|^2 dx dt = \frac{1}{2} \int_{\mathbb{R}^N} u_0(x)^2 dx.$$

Correction. On rappelle que la transformée de Fourier de ∇u est $ik\hat{u}$. Comme la transformée de Fourier est une isométrie de L^2 , on a donc

$$\begin{aligned} \frac{1}{2}\|u\|_{L^2}^2 + \int_0^T \|\nabla u\|_{L^2}^2 &= \frac{1}{2}\|\hat{u}\|_{L^2}^2 + \int_0^T \|k\hat{u}\|_{L^2}^2 \\ &= \frac{1}{2} \int_{\mathbb{R}^N} |u_0(k)|^2 e^{-2|k|^2 T} dk + \int_{\mathbb{R}^N} \int_0^T |k|^2 |u_0(k)|^2 e^{-2|k|^2 t} dt dk \\ &= \frac{1}{2} \int_{\mathbb{R}^N} |u_0(k)|^2 e^{-2|k|^2 T} dk - \frac{1}{2} \int_{\mathbb{R}^N} \int_0^T \frac{\partial}{\partial t} |u_0(k)|^2 e^{-2|k|^2 t} \\ &= \frac{1}{2} \int_{\mathbb{R}^N} |u_0(k)|^2 dk = \frac{1}{2} \|u_0\|_{L^2(\mathbb{R}^N)}^2 \end{aligned}$$

Exercice 8.4.8 (principe du maximum) Montrer que, si $u_0 \in L^\infty(\mathbb{R}^N)$, alors $u(t) \in L^\infty(\mathbb{R}^N)$ et

$$\|u(t)\|_{L^\infty(\mathbb{R}^N)} \leq \|u_0\|_{L^\infty(\mathbb{R}^N)} \quad \forall t > 0.$$

Montrer que, si $u_0 \geq 0$ presque partout dans \mathbb{R}^N , alors $u \geq 0$ dans $\mathbb{R}^N \times \mathbb{R}_*^+$.

Correction.

$$\|u(t)\|_{L^\infty} \leq \frac{\|u_0\|_\infty}{(4\pi t)^{N/2}} \int_{\mathbb{R}^N} e^{-y^2/4t} dy = \|u_0\|_\infty.$$

Enfin, d'après l'expression explicite de u en fonction de u_0 , il est évident que si $u_0 \geq 0$ presque partout, $u \geq 0$ presque partout.

Exercice 8.4.9 (effet régularisant) Montrer que $u \in C^\infty(\mathbb{R}^N \times \mathbb{R}_*^+)$.

Correction.

D'après l'expression de la transformée de Fourier de u ,

$$\hat{u}(k, t) = \hat{u}_0(k) e^{-|k|^2 t}.$$

Pour tout multi-indice α , $|k|^\alpha \hat{u}$ est un élément de $C(\mathbb{R}_*^+, L^2(\mathbb{R}^N))$. Ainsi, par transformation de Fourier inverse, $\partial^\alpha u$ est un élément de $C(\mathbb{R}_*^+, L^2(\mathbb{R}^N))$. En d'autres termes, pour tout entier m , u appartient à $C^0(\mathbb{R}_*^+, H^m(\mathbb{R}^N))$. D'après les injections de Sobolev, on en déduit que $u(t) \in C^0(\mathbb{R}_*^+, C^\infty(\mathbb{R}^N))$. En effectuant une analyse similaire sur $\frac{\partial^n u}{\partial t^n}$, on en déduit que $u \in C^\infty(\mathbb{R}_*^+, C^\infty(\mathbb{R}^N)) = C^\infty(\mathbb{R}^N \times \mathbb{R}_*^+)$.

Exercice 8.4.10 (comportement asymptotique) Montrer que

$$\lim_{|x| \rightarrow +\infty} u(x, t) = 0 \quad \forall t > 0, \quad \text{et} \quad \lim_{t \rightarrow +\infty} u(x, t) = 0 \quad \forall x \in \mathbb{R}^N.$$

Correction.

Soit r un réel positif, on décompose l'intégrale définissant $u(x, t)$ en deux intégrales

$$u(x, t) = \frac{1}{(4\pi t)^{N/2}} \left(\int_{|x-y| \geq r} u_0(y) e^{-\frac{|x-y|^2}{4t}} dy + \int_{|x-y| \leq r} u_0(y) e^{-\frac{|x-y|^2}{4t}} dy \right).$$

En appliquant l'inégalité de Cauchy-Schwarz à chacun des termes, on en déduit que

$$|u(x, t)| \leq \frac{1}{(4\pi t)^{N/2}} \left(\|u_0\|_{L^2(\mathbb{R}^N)} \left(\int_{|x-y| \geq r} e^{-\frac{|x-y|^2}{2t}} dy \right)^{1/2} \right. \\ \left. + \|e^{-\frac{x^2}{4t}}\|_{L^2(\mathbb{R}^N)} \left(\int_{|x-y| \leq r} |u_0(y)|^2 dy \right)^{1/2} \right).$$

On note B_r la boule de rayon r centrée en l'origine. On a

$$|u(x, t)| \leq \frac{1}{(4\pi t)^{N/2}} \left(\|u_0\|_{L^2(\mathbb{R}^N)} \|e^{-\frac{x^2}{4t}}\|_{L^2(\mathbb{R}^N \setminus B_r)} \right. \\ \left. + \|e^{-\frac{x^2}{4t}}\|_{L^2(\mathbb{R}^N)} \left(\int_{|x-y| \leq r} |u_0(y)|^2 dy \right)^{1/2} \right)$$

Pour tout réel $\varepsilon > 0$, pour r assez grand, on a $\|u_0\|_{L^2(\mathbb{R}^N)} \|e^{-\frac{x^2}{4t}}\|_{L^2(\mathbb{R}^N \setminus B_r)} < \varepsilon$. De plus, pour x assez grand (r étant fixé),

$$\|e^{-\frac{x^2}{4t}}\|_{L^2(\mathbb{R}^N)} \left(\int_{|x-y| \leq r} |u_0(y)|^2 dy \right)^{1/2} < \varepsilon.$$

Ainsi, pour tout ε , pour x assez grand on a $|u(x, t)| < \frac{1}{(2\pi t)^{N/2}}$. En d'autres termes,

$$\lim_{|x| \rightarrow +\infty} u(x, t) = 0 \text{ pour tout } t > 0.$$

On rappelle que $\hat{u}(k, t) = \hat{u}_0(k) e^{-|k|^2 t}$. Ainsi,

$$\|u(t)\|_{L^2(\mathbb{R}^N)}^2 = \|\hat{u}(t)\|_{L^2(\mathbb{R}^N)}^2 = \int_{\mathbb{R}^N} |\hat{u}_0(k)|^2 e^{-2|k|^2 t} dk$$

et d'après le Théorème de convergence dominé de Lebesgue, $\|u\|_{L^2(\mathbb{R}^N)}$ converge vers zéro lorsque t tend vers l'infini. Le même raisonnement appliqué aux dérivées partielles de u d'ordre quelconque nous donne que pour tout entier m , la norme H^m de $u(t)$ converge vers zéro lorsque t tend vers l'infini. D'après les injections de Sobolev, on en déduit que, pour tout entier r , la norme de $u(t)$ dans $C^r(\mathbb{R})$ tend vers zéro. En particulier,

$$\lim_{t \rightarrow +\infty} u(x, t) = 0 \text{ pour tout } x \in \mathbb{R}^N.$$

Exercice 8.4.11 (vitesse de propagation infinie) Montrer que, si $u_0 \geq 0$ et $u_0 \not\equiv 0$, alors $u(x, t) > 0$ dans $\mathbb{R}^N \times \mathbb{R}_*^+$.

Correction.

Soit $u_0 \geq 0$. Si il existe $(x, t) \in \mathbb{R}^N \times \mathbb{R}_*^+$ tel que $u(x, t) = 0$, alors $u_0(y) e^{-\frac{|x-y|^2}{4t}} = 0$ pour presque tout y et $u_0(y) = 0$ presque partout.

Exercice 8.5.1 Soit $\eta > 0$. On considère l'équation des ondes amortie

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} + \eta \frac{\partial u}{\partial t} - \Delta u = f & \text{p.p. dans } \Omega \times \mathbb{R}_+^* \\ u = 0 & \text{p.p. sur } \partial\Omega \times \mathbb{R}_+^* \\ u(x, 0) = u_0(x) & \text{p.p. dans } x \in \Omega \\ \frac{\partial u}{\partial t}(x, 0) = u_1(x) & \text{p.p. dans } x \in \Omega. \end{cases} \quad (8.18)$$

On suppose que u est une solution suffisamment régulière de (8.18) et que f est nul après un temps fini. Montrer, à l'aide d'un lemme de Gronwall (voir l'Exercice 8.2.1), que u et $\frac{\partial u}{\partial t}$ décroissent exponentiellement vers zéro lorsque le temps t tend vers l'infini.

Correction.

On pose $v = e^{\eta t} u$. Si on suppose que u est régulière, v est également solution d'une équation des ondes. On vérifie en effet que

$$\begin{cases} \frac{\partial^2 v}{\partial t^2} - \eta \Delta v = \eta e^{\eta t} f & \text{dans } \Omega \times \mathbb{R}_+^* \\ v = 0 & \text{sur } \partial\Omega \times \mathbb{R}_+^* \\ v(x, 0) = v_0 \text{ dans } \Omega \\ \frac{\partial v}{\partial t} = v_1 \text{ dans } \Omega \end{cases}$$

avec $v_0 = u_0$ et $v_1 = \eta u_1$. D'après la dernière estimation de l'Exercice 8.3.1 appliquées à v , on obtient que

$$\begin{aligned} & \int_{\Omega} \left| \frac{\partial v}{\partial t} \right|^2 + \eta |\nabla v|^2 dx \\ & \leq C \left(\int_{\Omega} v_1^2 + |\nabla v_0|^2 dx + \left(\int_0^t \left(\int_{\Omega} (\eta f(x, s) e^{\eta t})^2 dx \right)^{1/2} ds \right)^2 \right). \end{aligned}$$

Comme f est nul pour t assez grand, le second membre est borné uniformément en temps. On en déduit que v et $\frac{\partial v}{\partial t}$ sont bornées dans $L^2(\Omega)$, d'où on déduit que les normes L^2 de u et $\frac{\partial u}{\partial t}$ décroissent en $e^{-\eta t}$.

Exercice 8.5.2 Soit $u(t, x)$ la solution, supposée suffisamment régulière, de l'équation des ondes (8.9). En l'absence de terme source, montrer que

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t \int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^2 dx = \lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t \int_{\Omega} |\nabla u|^2 dx = \frac{1}{2} E_0,$$

avec E_0 l'énergie initiale

$$E_0 = \int_{\Omega} |u_1(x, t)|^2 dx + \int_{\Omega} |\nabla u_0(x, t)|^2 dx.$$

Pour cela on multipliera l'équation (8.9) par u et on intégrera par partie.

Correction.

En multipliant l'équation des ondes par u , on obtient par intégration sur $\Omega \times]0, t[$ que

$$\int_0^t \int_{\Omega} \frac{\partial^2 u}{\partial t^2} u(x, s) dx ds + \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds = 0.$$

En intégrant par partie en temps le premier terme de cette équation, on en déduit que

$$\int_{\Omega} \frac{\partial u}{\partial t} u(x, t) dx - \int_{\Omega} u_1 u_0 dx - \int_0^t \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, s) \right|^2 dx ds + \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds = 0.$$

Ainsi,

$$\begin{aligned} \frac{1}{t} \left(\int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds - \int_0^t \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, s) \right|^2 dx ds \right) \\ = \frac{1}{t} \left(\int_{\Omega} u_1 u_0 dx - \int_{\Omega} \frac{\partial u}{\partial t} u(x, t) dx \right) \xrightarrow{t \rightarrow +\infty} 0 \end{aligned}$$

(En effet, $\int_{\Omega} \frac{\partial u}{\partial t} u(x, t) dx$ uniformément borné en temps). D'autre part, l'équation de conservation de l'énergie implique que

$$\frac{1}{t} \left(\int_0^t \int_{\Omega} |\nabla u|^2 dx ds + \int_0^t \int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^2 dx ds \right) = E_0.$$

En sommant ces deux équations on obtient que

$$\frac{1}{t} \int_0^t \int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^2 dx$$

et

$$\frac{1}{t} \int_0^t \int_{\Omega} |\nabla u|^2 dx$$

convergent vers $E_0/2$.

Exercice 8.5.3 On considère l'équation des ondes dans tout l'espace \mathbb{R}^N

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - \Delta u = 0 & \text{dans } \mathbb{R}^N \times \mathbb{R}_+^* \\ u(x, 0) = u_0(x) & \text{dans } x \in \mathbb{R}^N \\ \frac{\partial u}{\partial t}(x, 0) = u_1(x) & \text{dans } x \in \mathbb{R}^N, \end{cases} \quad (8.19)$$

avec une donnée initiale (u_0, u_1) régulière et à support compact. Montrer que la solution $u(t, x)$ peut se mettre sous la forme

$$u(x, t) = (Mu_1)(x, t) + \left(\frac{\partial(Mu_0)}{\partial t} \right)(x, t),$$

où M est un opérateur de moyenne défini par

$$\text{si } N = 1, \quad (Mv)(x, t) = \frac{1}{2} \int_{-t}^{+t} v(x - \xi) d\xi,$$

$$\text{si } N = 2, \quad (Mv)(x, t) = \frac{1}{2\pi} \int_{|\xi| < t} \frac{v(x - \xi)}{\sqrt{t^2 - |\xi|^2}} d\xi,$$

$$\text{si } N = 3, \quad (Mv)(x, t) = \frac{1}{4\pi t} \int_{|\xi|=t} v(x - \xi) ds(\xi),$$

où $ds(\xi)$ est la mesure surfacique de la sphère. En déduire que la solution u en (t, x) ne dépend que des valeurs des données initiales u_0 et u_1 sur la boule $|x| \leq t$. (Pour savoir comment on trouve les expressions ci-dessus de l'opérateur M , nous renvoyons au cours de majeure [4].)

Correction.

On procède de manière identique dans les trois cas : dans un premier temps, on vérifie que pour toute fonction v ,

$$\frac{\partial^2 Mv}{\partial t^2}(x, t) = \Delta(Mv)(x, t) \quad (8.20)$$

Pour tout couple (x, t) tel que $t > 0$. On en déduit que la fonction u définie à l'aide de Mu_1 et Mu_0 vérifie bien l'équation des ondes. Il reste à montrer qu'elle vérifie les conditions aux limites, c'est à dire que

$$\begin{aligned} Mv(x, 0) &= 0 \\ \frac{\partial Mv}{\partial t}(x, 0) &= v(x) \\ \frac{\partial^2 Mv}{\partial t^2}(x, 0) &= 0 \end{aligned}$$

Le cas $N = 1$ est essentiellement élémentaire. Étudions directement les cas $N = 2$ ou 3.

Cas N=2. Tout d'abord, on effectue un changement de variable afin de définir Mv à l'aide d'une intégrale dont le domaine est indépendant du temps. On a

$$Mv = \frac{1}{2\pi} \int_{|\xi| < 1} \frac{v(x - t\xi)t}{(1 - |\xi|^2)^{1/2}} d\xi.$$

Si on suppose que v est assez régulière, on peut échanger les opérateur d'intégration et de dérivation lors du calcul des dérivées partielles. On obtient

$$\frac{\partial^2 Mv}{\partial t^2} = \frac{1}{2\pi} \int_{|\xi| < 1} \frac{t(D^2v(x - t\xi)\xi) \cdot \xi - 2\nabla v(x - t\xi) \cdot \xi}{(1 - |\xi|^2)^{1/2}} d\xi$$

et

$$\Delta(Mv) = \frac{1}{2\pi} \int_{|\xi| < 1} \frac{\Delta v(x - t\xi)}{(1 - |\xi|^2)^{1/2}} t d\xi.$$

Afin de vérifier (8.20), on introduit, pour tout x et $t > 0$ fixés, la fonction $w(\xi) = v(x - t\xi)$. Des expressions de $\partial^2 Mv / \partial t^2$ et de $\Delta(Mv)$, on déduit que

$$\frac{\partial^2 Mv}{\partial t^2} = \frac{1}{2\pi t} \int_{|\xi| < 1} \frac{(D^2 w \xi) \cdot \xi + 2\nabla w \cdot \xi}{(1 - |\xi|^2)^{1/2}} d\xi$$

et que

$$\Delta(Mv) = \frac{1}{2\pi t} \int_{|\xi| < 1} \frac{\Delta w}{(1 - |\xi|^2)^{1/2}} d\xi.$$

Soit r un réel strictement positif tel que $r < 1$. Par intégration par partie, on montre que

$$\int_{|\xi| < r} \frac{\Delta w}{(1 - |\xi|^2)^{1/2}} d\xi = - \int_{|\xi| < r} \frac{\nabla w \cdot \nabla \xi}{(1 - |\xi|^2)^{1/2}} d\xi + \frac{1}{r(1 - r^2)^{1/2}} \int_{|\xi|=r} (\nabla w \cdot \xi) ds.$$

De même,

$$\begin{aligned} & \int_{|\xi| < r} \frac{(D^2 w \xi) \cdot \xi}{(1 - |\xi|^2)^{1/2}} d\xi = \\ & - \int_{|\xi| < r} (\nabla w \cdot \xi) \left(\frac{2}{(1 - |\xi|^2)^{1/2}} + \frac{1}{(1 - |\xi|^2)^{3/2}} \right) d\xi + \frac{r}{(1 - r^2)^{1/2}} \int_{|\xi|=r} \nabla w \cdot \xi ds. \end{aligned}$$

On effectue la soustraction de ces deux expressions, puis on fait tendre r vers 1. Les termes de bords tendent vers zéro, ce qui établit que

$$\int_{|\xi| < 1} \frac{\Delta w - (D^2 w \xi) \cdot \xi}{(1 - |\xi|^2)^{1/2}} d\xi = 2 \int_{|\xi| < 1} \frac{\nabla w \cdot \xi}{(1 - |\xi|^2)^{1/2}} d\xi.$$

De l'expression des dérivées partielles de Mv en fonction de w , on en déduit que Mv vérifie l'équation des ondes. Reste à prouver que Mv vérifie bien les conditions aux limites annoncées en $t = 0$.

On a évidemment $Mv(t = 0) = 0$. De plus,

$$\frac{\partial Mv}{\partial t} = \frac{1}{2\pi} \int_{|\xi| < 1} \frac{t \nabla v(x - t\xi) \cdot \xi + v(x - t\xi)}{(1 - |\xi|^2)^{1/2}} d\xi.$$

Ainsi,

$$\frac{\partial Mv}{\partial t}(x, t = 0) = v(x) \frac{1}{2\pi} \int_{|\xi| < 1} \frac{1}{(1 - |\xi|^2)^{1/2}} d\xi.$$

En passant en coordonnées polaires afin de calculer le terme intégrale, il vient

$$\frac{\partial Mv}{\partial t}(t = 0) = v.$$

Enfin,

$$\frac{\partial^2 Mv}{\partial t^2}(t = 0) = -\frac{1}{\pi} \int_{|\xi| < 1} \frac{\nabla v(x) \cdot \xi}{(1 - |\xi|^2)^{1/2}} d\xi = -\frac{1}{\pi} \int_{|\xi| < 1} \nabla_\xi \cdot (\nabla v(x) (1 - |\xi|^2)^{1/2}) d\xi.$$

Par intégration par partie, on obtient que

$$\frac{\partial^2 Mv}{\partial t^2}(t=0) = \frac{1}{\pi} \int_{|\xi|=1} (\nabla v(x) \cdot \xi)(1 - |\xi|^2)^{1/2} d\xi = 0.$$

Cas N=3. On procède au calcul des dérivées partielles de Mv comme précédemment. Il vient

$$\frac{\partial^2 Mv}{\partial t^2} = \frac{1}{4\pi} \int_{|\xi|=1} t(D^2 v(x - t\xi)\xi) \cdot \xi - 2(\nabla v \cdot \xi) ds$$

et

$$\Delta(Mv) = \frac{1}{4\pi} \int_{|\xi|=1} t\Delta v(x - t\xi) ds.$$

Soit (x, t) fixée tel que $t > 0$. On introduit la fonction $w(\xi) = v(x - t\xi)$. On a

$$\frac{\partial^2 Mv}{\partial t^2} = \frac{1}{4\pi t} \int_{|\xi|=1} (D^2 w\xi) \cdot \xi + 2(\nabla w \cdot \xi) ds$$

$$\Delta(Mv) = \frac{1}{4\pi t} \int_{|\xi|=1} \Delta w ds.$$

Il suffit donc de remarquer que

$$\int_{|\xi|=1} (D^2 w\xi + 2\nabla w - \Delta w\xi) \cdot \xi ds = 0,$$

en tant que flux d'un champ de divergence nulle. En effet,

$$\nabla \cdot (D^2 w\xi) = \nabla(\Delta w) \cdot \xi + \Delta w,$$

et (en dimension 3),

$$\nabla \cdot (\Delta w\xi) = 3\Delta w + \nabla(\Delta w) \cdot \xi.$$

Pour finir, il est aisé de vérifier que Mv vérifie bien les conditions aux limites annoncées (pourvu qu'on sache que la surface de la sphère est 4π).

Exercice 8.5.4 On considère l'équation des ondes (8.19) dans un domaine $\Omega \subset \mathbb{R}^N$ avec des conditions aux limites indéterminées mais homogènes, et une donnée initiale (u_0, u_1) régulière et à support compact dans Ω . Vérifier qu'il existe un temps $T > 0$ tel que sur l'intervalle $[0, T]$ la solution est encore donnée par les formules de l'Exercice 8.5.3.

Correction.

Soit K l'union des supports de u_0 et u_1 . Si T est inférieur à la distance de K à la frontière de Ω , la solution explicite donnée par l'exercice précédent est aussi solution de l'équation des ondes dans le domaine Ω . En effet, les conditions aux limites sont vérifiées, car $u(x, t)$ est nul dès que la distance de x à K est supérieure à t .

Exercice 8.5.5 (application musicale) En admettant que le son se propage selon l'équation des ondes, montrer qu'il n'est pas possible d'écouter de la musique (audible) dans un monde de dimension spatiale $N = 2$, alors que c'est (fort heureusement) possible en dimension $N = 3$.

Correction.

Il est impossible d'écouter une musique audible dans un monde à deux dimensions. En effet, toutes les ondes émises sont entendues en même temps par l'auditeur (et pas seulement celles émises à un instant donné).

Exercice 8.6.1 Montrer que le schéma de Crank-Nicholson et celui de Gear sont d'ordre 2 (en temps), tandis que le θ -schéma pour $\theta \neq 1/2$ est d'ordre 1.

Correction.

Schéma de Crank-Nicholson et θ -schéma

Soit U la solution de l'équation différentielle (8.58). L'erreur de troncature du schéma du θ -schéma est

$$E(U) = \mathcal{M} \frac{U(t_{n+1}) - U(t_n)}{\Delta t} + \mathcal{K}(\theta U(t_{n+1}) + (1 - \theta)U(t_n)) - (\theta b(t_{n+1}) + (1 - \theta)b(t_n)).$$

En effectuant un développement de Taylor en $t = t_n$, on obtient

$$E(U) = \left(\mathcal{M} \frac{dU}{dt} + \mathcal{K}U - b \right) + \Delta t \left(\frac{\mathcal{M}}{2} \frac{d^2U}{dt^2} + \theta \left(\mathcal{K} \frac{dU}{dt} - \frac{db}{dt} \right) \right) + (\Delta t)^2 \left(\frac{\mathcal{M}}{6} \frac{d^3U}{dt^3} + \frac{\theta \mathcal{K}}{2} \frac{d^2U}{dt^2} - \frac{\theta}{2} \frac{d^2b}{dt^2} \right) + \mathcal{O}((\Delta t)^3)$$

En exploitant l'équation vérifiée par U , on en déduit que

$$E(U) = \Delta t \frac{1 - 2\theta}{2} \left(\frac{db}{dt} - \mathcal{K} \mathcal{M}^{-1}(b - \mathcal{K}U) \right) + (\Delta t)^2 \frac{1 - 3\theta}{6} \left((\mathcal{K} \mathcal{M}^{-1})^2(b - \mathcal{K}U) + \mathcal{K} \mathcal{M}^{-1} \frac{db}{dt} + \frac{d^2b}{dt^2} \right) + \mathcal{O}((\Delta t)^3).$$

Pour $\theta \neq 1/2$, le θ -schéma est d'ordre 1 en temps tandis que le schéma de Crank-Nicholson (qui correspond au cas $\theta = 1/2$) est d'ordre 2 en temps.

Schéma de Gear

Dans le cas du schéma de Gear, l'erreur de troncature est

$$E(U) = \mathcal{M} \frac{2U(t_{n+1}) - 4U(t_n) + U(t_{n-1}))}{2\Delta t} + \mathcal{K}U(t_{n+1}) - b(t_{n+1}).$$

En effectuant un développement de Taylor en $t = t_{n+1}$, on obtient

$$E(U) = \left(\mathcal{M} \frac{dU}{dt} + \mathcal{K}U - b \right) (t_{n+1}) + \frac{(\Delta t)^2}{3} \mathcal{M} \frac{d^3U}{dt^3} (t_{n+1}) + \mathcal{O}((\Delta t)^3).$$

Si U est solution de (8.58), on a donc

$$E(U) = \frac{(\Delta t)^2}{3} \mathcal{M} \frac{d^3 U}{dt^3}(t_{n+1}) + \mathcal{O}((\Delta t)^3).$$

Le schéma de Gear est donc d'ordre 2 en temps.

Exercice 8.6.2 On considère le θ -schéma (8.59) avec $1/2 \leq \theta \leq 1$. On note $\|U\|_{\mathcal{M}} = \sqrt{\mathcal{M}U \cdot U}$. Démontrer l'équivalent discret suivant de l'inégalité d'énergie (8.17)

$$\|U^{n_0}\|_{\mathcal{M}}^2 + \sum_{n=0}^{n_0} \Delta t \mathcal{K} \hat{U}^n \cdot \hat{U}^n \leq C \left(\|U^0\|_{\mathcal{M}}^2 + \int_0^T \|f(t)\|_{L^2(\Omega)}^2 dt + \mathcal{O}(1) \right).$$

où $n_0 = T/\Delta t$. Pour cela, on prendra le produit scalaire de (8.59) avec $\hat{U}^n = \theta U^{n+1} + (1 - \theta)U^n$.

Correction. Afin d'établir l'inégalité d'énergie demandée, on procède comme dans le cas continue. A cet effet, on utilise une version discrète du lemme de Gronwall : Si v^n est une suite de réels positifs tels que pour $a \geq v^0 \geq 0$ et $b \geq 0$, on a

$$v^{n+1} \leq a + b \sum_{p=0}^n v^p,$$

alors pour tout n , on a

$$v^n \leq a(1 + b)^n.$$

Dans un premier temps, nous allons donc démontrer ce lemme, puis l'appliquer au θ -schéma afin d'obtenir l'estimation d'énergie souhaitée. On introduit la suite w_n définie par

$$w^{n+1} = a + b \sum_{p=0}^n w^p,$$

$w^0 = a$. On vérifie que $w^n = a(1 + b)^n$ et que $v^n \leq w^n$, ce qui prouve la version discrète de lemme de Gronwall. Nous allons maintenant appliquer ce lemme afin d'obtenir l'estimation voulue.

Notons que

$$2\mathcal{M}(U^{n+1} - U^n) \cdot (\theta U^{n+1} + (1 - \theta)U^n) = \|U^{n+1}\|_{\mathcal{M}}^2 - \|U^n\|_{\mathcal{M}}^2 + (2\theta - 1)\|U^{n+1} - U^n\|_{\mathcal{M}}^2.$$

En effectuant le produit scalaire de (8.59) avec $\hat{U}^n = \theta U^{n+1} + (1 - \theta)U^n$, on obtient

$$\frac{\|U^{n+1}\|_{\mathcal{M}}^2 - \|U^n\|_{\mathcal{M}}^2}{2\Delta t} + \frac{2\theta - 1}{2\Delta t} \|U^{n+1} - U^n\|_{\mathcal{M}}^2 + \mathcal{K} \hat{U}^n \cdot \hat{U}^n = (\theta b_{n+1} + (1 - \theta)b_n) \cdot \hat{U}^n.$$

Comme $\theta \geq 1/2$, par sommation de la relation précédente, il vient

$$\frac{\|U^{n+1}\|_{\mathcal{M}}^2 - \|U_0\|_{\mathcal{M}}^2}{2\Delta t} + \sum_{p=0}^n \mathcal{K} \hat{U}^p \cdot \hat{U}^p \leq \sum_{p=0}^n (\theta b_{p+1} - (1 - \theta)b_p) \cdot (\theta U^{p+1} - (1 - \theta)U^p). \quad (8.21)$$

Majorons le terme de droite. D'après la définition de b , on a

$$\begin{aligned} & (\theta b_{p+1} - (1 - \theta)b_p) \cdot (\theta U^{p+1} - (1 - \theta)U^p) \\ &= \int_{\Omega} (\theta f(t_{p+1}) + (1 - \theta)f(t_p)) \cdot (\theta u_h^{p+1} + (1 - \theta)u_h^p) dx. \end{aligned}$$

On en déduit que

$$\begin{aligned} & (\theta b_{p+1} - (1 - \theta)b_p) \cdot (\theta U^{p+1} - (1 - \theta)U^p) \\ & \leq \frac{1}{2} (\|\theta f(t_{p+1}) + (1 - \theta)f(t_p)\|_{L^2}^2 + \|\theta u_h^{p+1} + (1 - \theta)u_h^p\|_{L^2}^2). \end{aligned}$$

De la définition de \mathcal{M} et en utilisant la convexité de l'application $x \mapsto x^2$, il en découle que

$$\begin{aligned} & (\theta b_{p+1} - (1 - \theta)b_p) \cdot (\theta U^{p+1} - (1 - \theta)U^p) \\ & \leq \frac{1}{2} (\theta \|f(t_{p+1})\|_{L^2}^2 + (1 - \theta)\|f(t_p)\|_{L^2}^2 + \theta \|U^{p+1}\|_{\mathcal{M}}^2 + (1 - \theta)\|U^p\|_{\mathcal{M}}^2). \end{aligned}$$

L'inégalité (8.21) implique ainsi

$$\frac{1}{2\Delta t} (\|U^{n+1}\|_{\mathcal{M}}^2 - \|U_0\|_{\mathcal{M}}^2) + \sum_{p=0}^n \mathcal{K} \hat{U}^p \cdot \hat{U}^p \leq \frac{1}{2} \left(\sum_{p=0}^{n+1} \|f(t_p)\|_{L^2}^2 + \sum_{p=0}^{n+1} \|U^p\|_{\mathcal{M}}^2 \right).$$

On réarrange les différents termes de l'inégalité afin d'obtenir une majoration nous permettant d'appliquer l'équivalent discret du lemme de Gronwall.

$$\begin{aligned} & \|U^{n+1}\|_{\mathcal{M}}^2 + \frac{2\Delta t}{1 - \Delta t} \sum_{p=0}^n \mathcal{K} \hat{U}^p \cdot \hat{U}^p \\ & \leq \frac{1}{1 - \Delta t} \|U^0\|_{\mathcal{M}}^2 + \frac{\Delta t}{1 - \Delta t} \left(\sum_{p=0}^{n+1} \|f(t_p)\|_{L^2}^2 + \sum_{p=0}^n \|U^p\|_{\mathcal{M}}^2 \right) \end{aligned}$$

On applique la version discrète du lemme de Gronwall à

$$v_n = \|U^n\|_{\mathcal{M}}^2,$$

$$a = \frac{1}{1 - \Delta t} \|U^0\|_{\mathcal{M}}^2 + \frac{1}{1 - \Delta t} \sum_{p=0}^{n_0} \|f(t_p)\|_{L^2}^2.$$

Pour tout $n \leq n_0$, on a $v_n \leq a(1 + b)^n$. En particulier,

$$a + b \sum_{p=0}^n v^p \leq a(1 + b)^{n+1},$$

et

$$\begin{aligned} \frac{1}{2} \|U^{n+1}\|_{\mathcal{M}}^2 + (1 - \Delta t)^{-1} \sum_{p=0}^n \mathcal{K} \hat{U}^p \cdot \hat{U}^p \Delta t \\ \leq (1 - \Delta t)^{-(n+2)} \left(\sum_0^{n_0+1} \|f(t_p)\|_{L^2}^2 \Delta t + \|U^0\|_{\mathcal{M}}^2 \right). \end{aligned}$$

Notons que $(1 - \Delta t)^{-n}$ est majoré par une constante indépendante du pas de temps Δt (mais dépendant du temps final $T = n_0 \Delta t$). En effet, $(1 - \Delta t)^{-n} \rightarrow e^t$ lorsque Δt tend vers zéro (avec $n = t/(\Delta t)$). On retrouve ainsi l'équivalent discret de l'estimation d'énergie de l'Exercice 8.3.1.

Exercice 8.6.3 Montrer que le schéma de Gear (8.61) est inconditionnellement stable.

Correction. On prouve la stabilité en établissant une estimation d'énergie du même type que celle obtenue dans l'Exercice 8.6.2. On note tout d'abord que

$$\begin{aligned} \mathcal{M}(3U^{n+1} - 4U^n + U^{n-1}) \cdot U^{n+1} &= \frac{1}{2} \left(\|U^{n+1}\|_{\mathcal{M}}^2 - \|U^n\|_{\mathcal{M}}^2 \right. \\ &\quad \left. + \|2U^{n+1} - U^n\|_{\mathcal{M}}^2 - \|2U^n - U^{n-1}\|_{\mathcal{M}}^2 + \|U^{n+1} - 2U^n + U^{n-1}\|_{\mathcal{M}}^2 \right). \end{aligned}$$

On effectue le produit scalaire du schéma de Gear (8.61) par U^{n+1} . En majorant le second terme, on obtient

$$\begin{aligned} \frac{1}{4} (\|U^{n+1}\|_{\mathcal{M}}^2 - \|U^n\|_{\mathcal{M}}^2 + \|2U^{n+1} - U^n\|_{\mathcal{M}}^2 - \|2U^n - U^{n-1}\|_{\mathcal{M}}^2) \\ + \Delta t \mathcal{K} U^{n+1} \cdot U^{n+1} \leq \Delta t \left(\frac{1}{2} \|U^{n+1}\|_{\mathcal{M}}^2 + \frac{1}{2} \|f(t_{n+1})\|_{L^2(\Omega)}^2 \right). \end{aligned}$$

Par sommation, on en déduit que

$$\begin{aligned} (1 - 2\Delta t) \|U^{n+1}\|_{\mathcal{M}}^2 + \Delta t \sum_{p=1}^{n+1} \mathcal{K} U^p \cdot U^p \\ \leq \|2U^1 - U^0\|_{\mathcal{M}}^2 + \|U^1\|_{\mathcal{M}}^2 + 2(\Delta t) \sum_{p=1}^{n+1} \|f(t_p)\|_{L^2}^2 + 2(\Delta t) \sum_{p=1}^n \|U^p\|_{\mathcal{M}}^2. \end{aligned}$$

En appliquant la version discrète du Lemme de Gronwall, on obtient l'estimation d'énergie

$$\|U^n\|_{\mathcal{M}}^2 \leq (1 - 2\Delta t)^{-(n+1)} \left(\|2U^1 - U^0\|_{\mathcal{M}}^2 + \|U^1\|_{\mathcal{M}}^2 + 2 \sum_{p=1}^{n_0+1} \|f(t_p)\|_{L^2}^2 \Delta t \right),$$

pour tout $n \leq n_0$. Comme $(1 - 2\Delta t)^{-(n+1)}$ est borné indépendamment de Δt pour un temps final donné, le schéma est stable.

Exercice 8.6.4 On résout par éléments finis \mathbb{P}_1 et schéma explicite en temps l'équation de la chaleur (8.12) en dimension $N = 1$. On utilise une formule de quadrature qui rend la matrice \mathcal{M} diagonale (voir la Remarque 7.4.3 et l'Exercice 7.4.1). On rappelle que la matrice \mathcal{K} est donnée par (6.12) et qu'on a calculé ses valeurs propres lors de l'Exercice 13.1.3. Montrer que dans ce cas la condition CFL (8.62) est bien du type $\Delta t \leq Ch^2$.

Correction. La condition CFL (8.62) est toujours valable, même si \mathcal{M} n'est pas la matrice de masse exacte. Ainsi, le schéma est stable sous la condition CFL (on a $\theta = 0$)

$$\max_k \lambda_k \Delta t \leq 2,$$

où λ_k sont les valeurs propres de \mathcal{K} , c'est à dire

$$\lambda_k = 4h^{-2} \sin^2 \left(\frac{k\pi}{2(n+1)} \right).$$

Comme $\lambda_k \leq 4h^{-2}$, on retrouve une condition CFL classique, c'est à dire

$$2\Delta t \leq h^2.$$

Exercice 8.6.5 Écrire le système linéaire d'équations différentielles ordinaires obtenu par semi-discrétisation de l'équation des ondes amortie (8.53).

Correction. Le problème discrétisé en espace consiste à déterminer $u(t)$ fonction de t à valeur dans V_{0h} tel que pour tout $v_h \in V_{0h}$,

$$\frac{d^2}{dt^2} \langle u_h(t), v_h \rangle_{L^2(\Omega)} + \eta \frac{d}{dt} \langle u_h(t), v_h \rangle_{L^2(\Omega)} + \langle \nabla u_h(t), \nabla v(t) \rangle = \langle f, v_h \rangle_{L^2(\Omega)}$$

tel que

$$u_h(t=0) = u_{0,h} \text{ et } \frac{du_h}{dt}(t=0) = u_{1,h}.$$

Si ϕ_i désigne la base de V_{0h} , si on note $U_i(t)$ les coordonnées de $u_h(t)$ dans cette base, on a

$$\frac{d^2}{dt^2} \mathcal{M}U(t) + \eta \frac{d}{dt} \mathcal{M}U(t) + \mathcal{K}U(t) = b(t)$$

où \mathcal{M} est la matrice de masse $\mathcal{M} = \langle \phi_i, \phi_j \rangle$, \mathcal{K} la matrice de rigidité $\langle \nabla \phi_i, \nabla \phi_j \rangle$ et b le terme source $\langle f, \phi_j \rangle$.

Exercice 8.7.1 Montrer que le schéma de Newmark est d'ordre 1 (en temps) pour $\delta \neq 1/2$, d'ordre 2 pour $\delta = 1/2$ et $\theta \neq 1/12$, et d'ordre 4 si $\delta = 1/2$ et $\theta = 1/12$ (on se limitera à l'équation sans amortissement).

Correction.

On introduit l'erreur de troncature

$$\begin{aligned} E(U) = & \mathcal{M} \frac{U(t+\Delta t) - 2U(t) + U(t-\Delta t)}{(\Delta t)^2} \\ & + \mathcal{K} \left(\theta U(t+\Delta t) + \left(\frac{1}{2} + \delta - 2\theta \right) U(t) + \left(\frac{1}{2} - \delta + \theta \right) U(t-\Delta t) \right) \\ & - \left(\theta b(t+\Delta t) + \left(\frac{1}{2} + \delta - 2\theta \right) b(t) + \left(\frac{1}{2} - \delta + \theta \right) b(t-\Delta t) \right). \end{aligned}$$

En effectuant un développement de Taylor en $t = t_n$, on établit que

$$\begin{aligned} E(U) = & \mathcal{M}U'' + \mathcal{K}U - b + \Delta t \left(\delta - \frac{1}{2} \right) (\mathcal{K}U' - b') \\ & + (\Delta t)^2 \left(\frac{1}{4} - \frac{\delta}{2} + \theta \right) (\mathcal{K}U'' - b'') + \frac{(\Delta t)^2}{12} \mathcal{M}U^{(4)} \\ & + \frac{(\Delta t)^3}{6} \left(\delta - \frac{1}{2} \right) (\mathcal{K}U^{(3)} - b^{(3)}) + \mathcal{O}((\Delta t)^4). \end{aligned}$$

Si U est solution de l'équation (8.67), on a

$$\mathcal{M}U'' + \mathcal{K}U - b = 0$$

et

$$\mathcal{K}U'' - b'' = -\mathcal{M}U^{(4)}.$$

Ainsi,

$$\begin{aligned} E(U) = & \Delta t \left(\delta - \frac{1}{2} \right) (\mathcal{K}U' - b') - (\Delta t)^2 \left(\frac{1}{4} - \frac{\delta}{2} + \theta - \frac{1}{12} \right) \mathcal{M}U^{(4)} \\ & + \frac{(\Delta t)^3}{6} \left(\delta - \frac{1}{2} \right) (\mathcal{K}U^{(3)} - b^{(3)}) + \mathcal{O}((\Delta t)^4). \end{aligned}$$

On vérifie aisément sur l'expression de $E(U)$ que le schéma de Newmark est d'ordre 1 pour $\delta \neq 1/2$, d'ordre 2 pour $\delta = 1/2$ et $\theta \neq 1/12$ et d'ordre (au moins) 4 si $\delta = 1/2$ et $\theta = 1/12$.

Exercice 8.7.2 On considère le cas limite du Lemme 8.7.1, c'est-à-dire $\delta = 1/2$ et $\lambda_i (\Delta t)^2 = \frac{4}{1-4\theta}$. Montrer que le schéma de Newmark est instable dans ce cas en vérifiant que

$$A_i = \begin{pmatrix} -2 & -1 \\ 1 & 0 \end{pmatrix}, \text{ et } A_i^n = (-1)^n \begin{pmatrix} n+1 & n \\ -n & 1-n \end{pmatrix}.$$

Remarquez qu'il s'agit d'une instabilité "faible" puisque la croissance de A_i^n est linéaire et non exponentielle.

Correction. D'après la démonstration du Lemme 8.7.1, on a

$$A_i = \begin{pmatrix} a_{11} & a_{12} \\ 1 & 0 \end{pmatrix},$$

$$a_{11} = \frac{2 - \lambda_i (\Delta t)^2 (\frac{1}{2} + \delta - 2\theta)}{1 + \theta \lambda_i (\Delta t)^2}, \quad a_{12} = -\frac{1 + \lambda_i (\Delta t)^2 (\frac{1}{2} - \delta + \theta)}{1 + \theta \lambda_i (\Delta t)^2}.$$

On vérifie sans mal que pour $\delta = 1/2$ et $\lambda_i (\Delta t)^2 = 4\theta/(1-\theta)$, $a_{11} = -2$ et $a_{12} = -1$. Ainsi,

$$A_i = \begin{pmatrix} -2 & -1 \\ 1 & 0 \end{pmatrix}.$$

Par récurrence on établit alors que

$$A_i^n = (-1)^n \begin{pmatrix} n+1 & n \\ -n & 1-n \end{pmatrix}$$

Il s'en suit que le schéma de Newmark est instable dans ce cas (pour s'en convaincre, il suffit par exemple de considérer le cas $b = 0$, $U_i^1 = U_i^0 = 1$)

Chapitre 9

INTRODUCTION A L'OPTIMISATION

Exercice 9.1.1 Montrer par des exemples que le fait que K est fermé ou que J est continue est en général nécessaire pour l'existence d'un minimum. Donner un exemple de fonction continue et minorée de \mathbb{R} dans \mathbb{R} n'admettant pas de minimum sur \mathbb{R} .

Correction. Exemples de non-existence de minimum

- K non fermé : minimisation de $J(x) = x^2$ sur $]0, 1[$.
- J non continue : minimisation sur \mathbb{R} de $J(x) = x^2$ pour $x \neq 0$, $J(0) = 1$.
- J non coercive : minimisation sur \mathbb{R} de $J(x) = e^{-x}$.

Exercice 9.1.2 Montrer que l'on peut remplacer la propriété “infinie à l'infini” (9.3) par la condition plus faible

$$\inf_{v \in K} J(v) < \lim_{R \rightarrow +\infty} \left(\inf_{\|v\| \geq R} J(v) \right).$$

Correction. Soit (u^n) une suite minimisante de J sur K . Comme

$$\inf_{v \in K} J(v) < \lim_{R \rightarrow +\infty} \left(\inf_{\|v\| \geq R} J(v) \right),$$

et que $J(v_n)$ converge vers $\inf_{v \in K} J(v)$, il existe $\delta > 0$ tel que pour n assez grand,

$$J(v_n) < \lim_{R \rightarrow +\infty} \left(\inf_{\|v\| \geq R} J(v) \right) - \delta.$$

Ainsi, il existe R tel que pour n assez grand,

$$J(v_n) < \inf_{\|v\| \geq R} J(v).$$

On en déduit que pour n assez grand, v appartient à la boule de rayon R . Autrement dit, la suite v_n reste bornée. La suite de la démonstration est alors identique à la démonstration initiale.

Exercice 9.1.3 Montrer que l'on peut remplacer la continuité de J par la semi-continuité inférieure de J définie par

$$\forall (u^n)_{n \geq 0} \text{ suite dans } K, \lim_{n \rightarrow +\infty} u^n = u \implies \liminf_{n \rightarrow +\infty} J(u^n) \geq J(u).$$

Correction. Seul la fin de la démonstration est modifiée. La suite minimisante (u^{n_k}) converge vers u , mais cette fois on a seulement

$$J(u) \leq \liminf_{k \rightarrow \infty} J(u^{n_k}) = \inf_{v \in K} J(v).$$

Or comme $u \in K$, $\inf_{v \in K} J(v) \leq J(u)$, d'où

$$J(u) = \inf_{v \in K} J(v).$$

Exercice 9.1.4 Montrer qu'il existe un minimum pour les Exemples 9.1.1, 9.1.6 et 9.1.7.

Correction. Les conditions du Théorème 9.1.3 sont trivialement satisfaites.

Exercice 9.1.5 Soit a et b deux réels avec $0 < a < b$, et pour $n \in \mathbb{N}^*$, soit \mathcal{P}_n l'ensemble des polynômes P de degré inférieur ou égal à n tels que $P(0) = 1$. Pour $P \in \mathcal{P}_n$, on note $\|P\| = \max_{x \in [a,b]} |P(x)|$.

1. Montrer que le problème

$$\inf_{P \in \mathcal{P}_n} \|P\| \tag{9.1}$$

a une solution.

2. On rappelle que les polynômes de Tchebycheff $T_n(X)$ sont définis par les relations

$$T_0(X) = 1, \quad T_1(X) = X, \quad T_{n+1}(X) = 2XT_n(X) - T_{n-1}(X).$$

Montrer que le degré de T_n est égal à n et que pour tout $\theta \in \mathbb{R}$, $T_n(\cos \theta) = \cos(n\theta)$. En déduire l'existence de $n+1$ réels

$$\xi_0^n = 1 > \xi_1^n > \xi_2^n > \dots > \xi_n^n = -1$$

tels que $T_n(\xi_k^n) = (-1)^k$ pour $0 \leq k \leq n$ et que $\max_{-1 \leq x \leq 1} |T_n(x)| = 1$.

3. Montrer que l'unique solution de (9.1) est le polynôme

$$P(X) = \frac{1}{T_n\left(\frac{b+a}{b-a}\right)} T_n\left(\frac{\frac{b+a}{2} - X}{\frac{b-a}{2}}\right).$$

Correction.

1. L'ensemble des polynômes de degré inférieur ou égal à n tel que $P(0) = 1$ est un sous espace affine (et fermé) de l'ensemble de polynôme de degré inférieur ou égal à n muni de la norme $\max_{x \in [a,b]} |P(x)|$. Toutes les hypothèses du Théorème 9.1.3 sont

satisfaites d'où on déduit l'existence d'une solution au problème de minimisation de $\|P\|$ sur \mathcal{P}_n .

2. Par une récurrence facile, on montre que T_n est un polynôme de degré n et que $T_n(\cos(\theta)) = \cos(n\theta)$. Pour tout $0 \leq k \leq n$, on pose $\xi_k^n = \cos(k\pi/n)$. On a $\xi_0^n = 1 > \xi_1^n > \dots > \xi_n^n = -1$ et $T_n(\xi_k^n) = \cos(k\pi) = (-1)^k$. Enfin,

$$\max_{-1 \leq x \leq 1} |T_n(x)| = \max_{\theta \in \mathbb{R}} |T_n(\cos(\theta))| = \max_{\theta \in \mathbb{R}} |\cos(n\theta)| = 1.$$

3. Soit R un polynôme de norme minimal appartenant à \mathcal{P}_n . On considère le polynôme $S = P - R$ où

$$P(X) = \frac{1}{T_n\left(\frac{b+a}{b-a}\right)} T_n\left(\frac{\frac{b+a}{2} - X}{\frac{b-a}{2}}\right).$$

On veut montrer que $S = 0$. Pour tout $k = 0, \dots, n$, on pose $y_k = \frac{a+b}{2} - \left(\frac{a-b}{2}\right) \xi_k$. D'après la question précédente, $P(y_k) = (-1)^k \|P\|$. On définit les ensembles d'indices

$$\begin{aligned} I &= \{i \in \{0, \dots, n-1\} : S(y_i) \neq 0 \text{ et } S(y_{i+1}) \neq 0\} \\ J &= \{j \in \{1, \dots, n-1\} : S(y_j) = 0\} \\ K &= \{k \in \{0, n\} : S(y_k) = 0\}. \end{aligned}$$

On vérifie que $|I| + 2|J| + |K| \geq n$. Pour tout $j \in J$, on a $|R(y_j)| = \|P\| \geq \|R\|$, d'où $\|R\| = |R(y_j)|$ et $R'(y_j) = 0$. De plus, $P'(y_j) = 0$, d'où $S'(y_j) = 0$.

De plus, pour tout $i \in I$, comme $\|P\| \geq \|R\|$, le signe de $S(y_i) = P(y_i) - R(y_i)$ est égale au signe de $P(y_i) = \|P\|(-1)^i$. De manière similaire, le signe de $S(y_{i+1})$ est $(-1)^{i+1}$. Comme $S(y_i)$ et $S(y_{i+1})$ sont de signes opposés, le polynôme S s'annule sur l'intervalle $[y_i, y_{i+1}]$ au moins une fois.

Ainsi, pour tout $j \in J$, $S(y_j) = S'(y_j) = 0$ et y_j est une racine double, pour tout $i \in I$, il existe $x_i \in]y_i, y_{i+1}[$ tel que $S(x_i) = 0$ et pour tout $k \in K$, $S(y_k) = 0$. De plus $S(0) = 0$. Ainsi, S admet au moins $|I| + 2|J| + |K| + 1$ racines (multiples). Comme S est de degré au plus $n \leq |I| + 2|J| + |K|$, on a $S = 0$.

Exercice 9.2.1 Modifier la construction de l'Exemple 9.2.2 pour montrer qu'il n'existe pas non plus de minimum de J sur $C^1[0, 1]$.

Correction. Soit $a \in [0, 1]$. On note P_a la fonction de $C^1(\mathbb{R}; \mathbb{R})$ paire, 2 périodique définie sur $[0, 1]$ par

$$P_a(x) = \begin{cases} x^2/2a + (a-1)/2 & \text{si } 0 \leq x \leq a \\ x - 1/2 & \text{si } a \leq x \leq 1-a, \\ -(x-1)^2/2(1-a) + (1-a)/2 & \text{si } 1-a \leq x \leq 1 \end{cases}$$

On note $u^n \in C^1(\mathbb{R}; \mathbb{R})$ la fonction $2/n$ -périodique, définie par

$$u^n(x) = n^{-1} h P_{n-1}(nx).$$

On vérifie de $u^n(x) \rightarrow 0$ presque partout et que $|(u^n)'(x)| \rightarrow h$ presque partout. Ainsi, l'infimum de J_h sur $C^1([0, 1])$ est nul et ne peut être atteint si $h > 0$.

Exercice 9.2.2 Soient J_1 et J_2 deux fonctions convexes sur V , $\lambda > 0$, et φ une fonction convexe croissante sur un intervalle de \mathbb{R} contenant l'ensemble $J_1(V)$. Montrer que $J_1 + J_2$, $\max(J_1, J_2)$, λJ_1 et $\varphi \circ J_1$ sont convexes.

Correction. La convexité de $J_1 + J_2$ comme de λJ_1 est triviale à établir.

$$\begin{aligned} \text{Epi}(\max(J_1, J_2)) &= \{(\lambda, v) \in \mathbb{R} \times V : \lambda \geq J_1(v) \text{ et } \lambda \geq J_2(v)\} \\ &= \text{Epi}(J_1) \cap \text{Epi}(J_2). \end{aligned}$$

L'intersection de deux convexes étant convexe, $\text{Epi}(\max(J_1, J_2))$ est convexe et $\max(J_1, J_2)$ est convexe.

Comme J est convexe et φ croissante,

$$\varphi \circ J(\theta x + (1 - \theta)y) \leq \varphi(\theta J(x) + (1 - \theta)J(y)).$$

La convexité de φ nous permet d'en déduire la convexité de $\varphi \circ J$.

Exercice 9.2.3 Soit $(L_i)_{i \in I}$ une famille (éventuellement infinie) de fonctions affines sur V . Montrer que $\sup_{i \in I} L_i$ est convexe sur V . Réciproquement, soit J une fonction convexe continue sur V . Montrer que J est égale au $\sup_{L_i \leq J} L_i$ où les fonctions L_i sont affines.

Correction. Le sup de fonction convexe est une fonction convexe. En effet, une fonction $J : V \rightarrow \mathbb{R}$ est convexe si et seulement si son épigraphe

$$\text{Epi}(J) = \{(\lambda, v) \in \mathbb{R} \times V, \lambda \geq J(v)\}$$

est convexe. Ainsi, si $J = \sup_{i \in I} J_i$, où J_i sont des fonctions convexes, on a

$$\begin{aligned} \text{Epi}(J) &= \{(\lambda, v) \in \mathbb{R} \times V, \lambda \geq J_i(v) \text{ pour tout } i \in I\} \\ &= \bigcap_i \text{Epi}(J_i). \end{aligned}$$

Une intersection de convexes étant convexe, l'épigraphe de J est convexe. La fonction J est donc convexe.

Réciproquement, supposons que J soit convexe. Soit $v_0 \in V$ et $\lambda_0 \in \mathbb{R}$ tel que $\lambda_0 < J(v_0)$, c'est à dire tel que (λ_0, v_0) n'appartienne pas à $\text{Epi}(J)$. Notons que l'ensemble $\text{Epi}(J)$ est un convexe fermé (fermé car J est continue et convexe car J est convexe). Puisque $(\lambda_0, v_0) \notin \text{Epi}(J)$, nous déduisons du Théorème 12.1.19 de séparation d'un point et d'un convexe l'existence de $\alpha, \beta \in \mathbb{R}$ et d'une forme linéaire continue $T \in V'$ tels que

$$\beta \lambda + T(v) > \alpha > \beta \lambda_0 + T(v_0) \quad \forall (\lambda, v) \in \text{Epi}(J).$$

Ainsi,

$$\beta J(v) + T(v) > \alpha > \beta \lambda_0 + T(v_0) \quad \forall v \in V$$

et

$$\beta J(v) > \beta \lambda_0 + T(v_0) - T(v) \quad \forall v \in V.$$

En appliquant l'inégalité précédente à $v = v_0$, on en déduit que β est non nul. De plus, β est nécessairement positif. On a donc

$$J(v) > \lambda_0 + \beta^{-1}(T(v_0) - T(v)) \quad \forall v \in V.$$

On pose $L(v) = \lambda_0 + \beta^{-1}(T(v_0) - T(v))$. On a prouvé que pour tout (v_0, λ_0) tel que

$$J(v_0) > \lambda_0,$$

il existe une fonction affine L telle que p

$$J(v_0) \geq L(v_0) = \lambda_0$$

et $J(v) \geq L(v)$ pour tout $v \in V$. On en déduit que

$$J = \sup_{L_i \leq J} L_i,$$

où les L_i sont des fonctions affines.

Exercice 9.2.4 Si J est continue et α -convexe, montrer que, pour tout $\theta \in [0, 1]$,

$$J(\theta u + (1 - \theta)v) \leq \theta J(u) + (1 - \theta)J(v) - \frac{\alpha\theta(1 - \theta)}{2} \|u - v\|^2. \quad (9.2)$$

Correction. Pour tout n , on note $K_n = \{x \in [0, 1] : 2^n x \in \mathbb{N}\}$. Supposons que l'inégalité (9.2) soit vérifiée pour tout $\theta \in K_n$. Soit $\theta \in K_{n+1} \setminus K_n$, il existe $\theta_1, \theta_2 \in K_n$ tels que $\theta_1 < \theta_2$ et $\theta = (\theta_1 + \theta_2)/2$. Comme J est α -convexe,

$$\begin{aligned} J(\theta u + (1 - \theta)v) &= J\left(\frac{(\theta_1 u + (1 - \theta_1)v) + (\theta_2 u + (1 - \theta_2)v)}{2}\right) \\ &\leq \frac{J(\theta_1 u + (1 - \theta_1)v) + J(\theta_2 u + (1 - \theta_2)v)}{2} + \frac{\alpha}{8}(\theta_2 - \theta_1)^2 \|u - v\|^2 \end{aligned}$$

L'inégalité (9.2) ayant été supposée exacte sur K_n , on a donc

$$\begin{aligned} J(\theta u + (1 - \theta)v) &\leq \frac{\theta_1 J(u) + (1 - \theta_1)J(v) + \theta_2 J(u) + (1 - \theta_2)J(v)}{2} \\ &\quad + \frac{\alpha\theta_1(1 - \theta_1) + \alpha(\theta_2(1 - \theta_2))}{4} \|u - v\|^2 + \frac{\alpha}{8}(\theta_2 - \theta_1)^2 \|u - v\|^2. \end{aligned}$$

et

$$J(\theta u + (1 - \theta)v) \leq \frac{\theta J(u) + (1 - \theta)J(v)}{2} + \frac{\alpha(\theta_1 + \theta_2)(2 - (\theta_1 + \theta_2))}{8} \|u - v\|^2,$$

ce qui prouve que l'inégalité est alors valable pour tout élément de K^{n+1} . On en déduit par récurrence que l'inégalité est valable pour $\theta \in \bigcup_n K^n$. Comme J est continue, l'inégalité reste valable sur l'adhérence de l'union des K_n , c'est à dire sur $[0, 1]$.

Exercice 9.2.5 Soit A une matrice symétrique d'ordre N et $b \in \mathbb{R}^N$. Pour $x \in \mathbb{R}^N$, on pose $J(x) = \frac{1}{2}Ax \cdot x - b \cdot x$. Montrer que J est convexe si et seulement si A est semi-définie positive, et que J est strictement convexe si et seulement si A est définie positive. Dans ce dernier cas, montrer que J est aussi fortement convexe et trouver la meilleure constante α .

Correction.

$$\begin{aligned} J((x+y)/2) &= A(x+y) \cdot (x+y)/8 - (b \cdot x + b \cdot y)/2 \\ &= \frac{Ax \cdot x - b \cdot x + Ay \cdot y - b \cdot y}{2} - A(x-y) \cdot (x-y)/8 \\ &= (J(x) + J(y))/2 - A(x-y) \cdot (x-y)/8. \end{aligned}$$

L'application J est donc convexe si et seulement si la matrice A est positive. Elle est strictement convexe si et seulement si A est définie positive. Dans ce cas, elle est fortement convexe et la meilleure constante α est la plus petite valeur propre de A .

Exercice 9.2.6 Soit Ω un ouvert de \mathbb{R}^N et $H^1(\Omega)$ l'espace de Sobolev associé (voir la Définition 4.3.1). Soit la fonction J définie sur Ω par

$$J(v) = \frac{1}{2} \int_{\Omega} (|\nabla v(x)|^2 + v(x)^2) dx - \int_{\Omega} f(x)v(x) dx ,$$

avec $f \in L^2(\Omega)$. Montrer que J est fortement convexe sur $H^1(\Omega)$.

Correction.

$$J((u+v)/2) = \frac{J(u) + J(v)}{2} - \|u - v\|_{H^1}^2/8.$$

(Les calculs sont identiques à ceux effectués lors de l'Exercice 9.2.5)

Exercice 9.2.7 Soit $v_0 \in V$ et J une fonction convexe majorée sur une boule de centre v_0 . Montrer que J est minorée et continue sur cette boule.

Correction. Sans perte de généralité, on peut supposer que $v_0 = 0$ et $J(0) = 0$ et que J est majorée sur une boule de rayon unité. Soit M un majorant de J sur la boule. Soit v tel que $\|v\| < 1$, on a

$$J(v) = J\left(\|v\| \frac{v}{\|v\|} + (1 - \|v\|)0\right) \leq \|v\| J\left(\frac{v}{\|v\|}\right) + (1 - \|v\|)J(0) \leq \|v\| M.$$

De plus,

$$\begin{aligned} 0 &= J(0) = J\left(\frac{1}{1 + \|v\|}v + \frac{\|v\|}{1 + \|v\|}\left(-\frac{v}{\|v\|}\right)\right) \\ &\leq \frac{1}{1 + \|v\|}J(v) + \frac{\|v\|}{1 + \|v\|}J\left(-\frac{v}{\|v\|}\right) \\ &\leq \frac{1}{1 + \|v\|}J(v) + \frac{\|v\|}{1 + \|v\|}M. \end{aligned}$$

Il découle de ces deux inégalités que

$$|J(v)| \leq M\|v\|.$$

Ainsi, J est minorée sur la boule unité et continue en zéro. Enfin, on peut appliquer ce résultat à tout point appartenant à la boule unité ouverte pour conclure que J est continue sur cette dernière.

Exercice 9.2.8 Montrer que le Théorème 9.2.6 s'applique à l'Exemple 9.1.10 (utiliser l'inégalité de Poincaré dans $H_0^1(\Omega)$).

Correction. D'après l'inégalité de Poincaré,

$$\|v\|_{H_0^1(\Omega)} = \int_{\Omega} |\nabla v|^2 dx$$

défini une norme sur H_0^1 . Ainsi, J est fortement convexe et le Théorème 9.2.6 s'applique.

Exercice 9.2.9 Généraliser l'Exercice 9.2.8 aux différents modèles rencontrés au Chapitre 5 : Laplacien avec conditions aux limites de Neumann (voir la Proposition 5.2.16), élasticité (voir l'Exercice 5.3.3), Stokes (voir l'Exercice 5.3.10).

Correction. Pas de Pb.

Chapitre 10

CONDITIONS D'OPTIMALITÉ ET ALGORITHMES

Exercice 10.1.1 Montrer que la dérivabilité de J en u implique la continuité de J en u . Montrer aussi que, si L_1, L_2 vérifient

$$\begin{cases} J(u+w) \geq J(u) + L_1(w) + o(w) , \\ J(u+w) \leq J(u) + L_2(w) + o(w) , \end{cases} \quad (10.1)$$

alors J est dérivable et $L_1 = L_2 = J'(u)$.

Correction. Si J est dérivable au sens de Fréchet en u , il existe une forme linéaire continue L telle que

$$J(u+w) = J(u) + L(w) + o(w).$$

Ainsi,

$$|J(u+w) - J(u)| \leq \|L\| \|w\| + |o(w)|.$$

Le terme de droite convergeant vers zéro lorsque w tend vers zéro, J est continue en u .

Considérons une fonction J vérifiant (10.1). De

$$J(u+w) \geq J(u) + L_1(w) + o(w)$$

et

$$-J(u+w) \geq -J(u) - L_2(w) + o(w),$$

on déduit que

$$0 \geq (L_1 - L_2)(w) + o(w).$$

Ainsi, pour tout réel $\alpha > 0$,

$$0 \geq (L_1 - L_2)(w) + \|w\| \frac{o(\alpha w)}{\alpha \|w\|}$$

(on applique l'inégalité précédente à αw et on divise par α). En faisant tendre α vers zéro, on obtient que pour tout w ,

$$0 \geq (L_1 - L_2)(w).$$

Cette inégalité appliquée $-w$, nous donne l'inégalité inverse et finalement l'égalité $L_1(w) = L_2(w)$. Il en découle que J est dérivable au sens de Fréchet et que $J' = L_1 = L_2$.

Exercice 10.1.2 (essentiel !) Soit a une forme bilinéaire symétrique continue sur $V \times V$. Soit L une forme linéaire continue sur V . On pose $J(u) = \frac{1}{2}a(u, u) - L(u)$. Montrer que J est dérivable sur V et que $\langle J'(u), w \rangle = a(u, w) - L(w)$ pour tout $u, w \in V$.

Correction. Il suffit de développer l'expression $J(u + w)$. On obtient

$$J(u + w) = J(u) + a(u, w) - L(w) + a(w, w)/2.$$

La forme bilinéaire a étant continue, $a(w, w)/\|w\|$ converge vers zéro lorsque w tend vers zéro. La fonction J est donc dérivable et

$$\langle J'(u), w \rangle = a(u, w) - L(w).$$

Exercice 10.1.3 Soit A une matrice symétrique $N \times N$ et $b \in \mathbb{R}^N$. Pour $x \in \mathbb{R}^N$, on pose $J(x) = \frac{1}{2}Ax \cdot x - b \cdot x$. Montrer que J est dérivable et que $J'(x) = Ax - b$ pour tout $x \in \mathbb{R}^N$.

Correction. Même résultat que l'exercice précédent (mais en dimension finie).

Exercice 10.1.4 On reprend l'Exercice 10.1.2 avec $V = L^2(\Omega)$ (Ω étant un ouvert de \mathbb{R}^N), $a(u, v) = \int_{\Omega} uv \, dx$, et $L(u) = \int_{\Omega} fu \, dx$ avec $f \in L^2(\Omega)$. En identifiant V et V' , montrer que $J'(u) = u - f$.

Correction. D'après l'Exercice 10.1.2,

$$\langle J'(u), w \rangle = a(u, w) - L(w),$$

d'où

$$\langle J'(u), w \rangle = \int_{\Omega} uw - fw \, dx = \langle u - f, w \rangle_{L^2}.$$

En identifiant L^2 et son dual à l'aide du produit scalaire L^2 , on obtient $J'(u) = u - f$.

Exercice 10.1.5 On reprend l'Exercice 10.1.2 avec $V = H_0^1(\Omega)$ (Ω étant un ouvert de \mathbb{R}^N) que l'on munit du produit scalaire

$$\langle u, v \rangle = \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, dx.$$

On pose $a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx$, et $L(u) = \int_{\Omega} fu \, dx$ avec $f \in L^2(\Omega)$. Montrer (au moins formellement) que $J'(u) = -\Delta u - f$ dans $V' = H^{-1}(\Omega)$. Montrer que, si on identifie V et V' , alors $J'(u) = u_0$ où u_0 est l'unique solution dans $H_0^1(\Omega)$ de

$$\begin{cases} -\Delta u_0 + u_0 = -\Delta u - f & \text{dans } \Omega \\ u_0 = 0 & \text{sur } \partial\Omega \end{cases}$$

Correction. La fonction J est dérivable et pour tout $w \in H_0^1(\Omega)$ on a

$$\langle J'(u), w \rangle = \int_{\Omega} \nabla u \cdot \nabla w - f w dx.$$

Si u appartient à $H^2(\Omega)$ alors $J'(u)$ appartient au dual de $L^2(\Omega)$. Suite à une intégration par partie, on obtient

$$\langle J'(u), w \rangle = - \int_{\Omega} (\Delta u + f) w dx.$$

Aussi, si on identifie $L^2(\Omega)$ et son dual à l'aide du produit scalaire L^2 , on obtient $J'(u) = -\Delta u - f$. Si on utilise le produit scalaire H^1 pour associer une fonction à $J'(u)$, on obtient évidemment un autre résultat. Soit v l'élément de $H_0^1(\Omega)$ associé à $J'(u)$ par identification de $H_0^1(\Omega)$ et son dual à l'aide du produit scalaire H^1 . En d'autres termes, v est l'unique élément de $H_0^1(\Omega)$ tel que pour tout $w \in H_0^1(\Omega)$,

$$\int_{\Omega} \nabla v \cdot \nabla w + v w dx = \langle J'(u), w \rangle = \int_{\Omega} \nabla u \cdot \nabla w + f w dx.$$

Par intégration par partie, on en déduit que v est solution du problème aux limites vérifié par u_0 . Ainsi $v = u_0$ et, si on identifie $H_0^1(\Omega)$ et son dual à l'aide du produit scalaire H^1 , $J'(u) = u_0$.

Exercice 10.1.6 Soit Ω un ouvert borné de \mathbb{R}^N (on pourra se restreindre au cas où $N = 1$ avec $\Omega =]0, 1[$). Soit $L = L(p, t, x)$ une fonction continue sur $\mathbb{R}^N \times \mathbb{R} \times \overline{\Omega}$, dérivable par rapport à p et t sur cet ensemble, de dérivées partielles $\frac{\partial L}{\partial p}$ et $\frac{\partial L}{\partial t}$ Lipschitziennes sur cet ensemble. On pose $V = H_0^1(\Omega)$ et $J(v) = \int_{\Omega} L(\nabla v(x), v(x), x) dx$.

1. Montrer que J est dérivable sur $H_0^1(\Omega)$ et que

$$\langle J'(u), w \rangle = \int_{\Omega} \left(\frac{\partial L}{\partial p}(\nabla u(x), u(x), x) \cdot \nabla w(x) + \frac{\partial L}{\partial t}(\nabla u(x), u(x), x) w(x) \right) dx.$$

2. Si $N = 1$ et $\Omega =]0, 1[$, montrer que, si $u \in H_0^1(0, 1)$ satisfait $J'(u) = 0$, alors u vérifie

$$\frac{d}{dx} \left(\frac{\partial L}{\partial p}(u'(x), u(x), x) \right) - \frac{\partial L}{\partial t}(u'(x), u(x), x) = 0, \quad (10.2)$$

presque partout dans l'intervalle $]0, 1[$.

3. Si L ne dépend pas de x (i.e. $L = L(p, t)$) et si $u \in C^2([0, 1])$ est une solution de classe de l'équation différentielle (10.2), montrer que la quantité

$$L(u'(x), u(x)) - u'(x) \frac{\partial L}{\partial p}(u'(x), u(x))$$

est constante sur l'intervalle $[0, 1]$.

Correction.

1. Tout d'abord, comme L est dérivable par rapport à p et t , de dérivées Lipschitziennes, on a

$$|L(p+q, t+s, x) - L(p, t, x) - \frac{\partial L}{\partial p}(p, t, x) \cdot q - \frac{\partial L}{\partial t}(p, t, x)s| \leq \frac{K}{2}(|q|^2 + |s|^2).$$

En particulier,

$$L(p, t, x) \leq C(1 + |p|^2 + t^2),$$

et J est correctement défini. On vérifie également que

$$M(u) \cdot w = \int_{\Omega} \left(\frac{\partial L}{\partial p}(\nabla u, u, x) \cdot \nabla w + \frac{\partial L}{\partial t}(\nabla u, u, x)w \right) dx$$

est une forme linéaire continue sur $H^1(\mathbb{R}^N \times \mathbb{R} \times \overline{\Omega})$. Enfin, on a

$$|J(u+w) - J(u) - M(u) \cdot w| \leq \frac{K}{2} \|w\|_{H^1}^2.$$

La fonction J est donc dérivable en u de dérivée $M(u)$.

2. Si $J'(u) = 0$, on a pour tout $w \in H_0^1(0, 1)$,

$$\int_0^1 \left(\frac{\partial L}{\partial p}(u', u, x) \cdot w' + \frac{\partial L}{\partial t}(u', u, x)w \right) dx = 0.$$

On en déduit que $\partial L / \partial p(\nabla u, u, x)$ appartient à $H^1(0, 1)$ et que

$$\frac{d}{dx} \left(\frac{\partial L}{\partial p}(u', u, x) \right) - \frac{\partial L}{\partial t}(u', u, x) = 0$$

presque partout.

3. Comme u est de classe C^2 , les calculs suivants sont licites :

$$\begin{aligned} \frac{d}{dx} \left(L(u', u) - u' \frac{\partial L}{\partial p}(u', u) \right) &= \frac{d(L(u', u))}{dx} - u'' \frac{\partial L}{\partial p} - u' \frac{d}{dx} \left(\frac{\partial L}{\partial p}(u', u) \right) \\ &= u' \left(\frac{\partial L}{\partial t}(u', u) - \frac{d}{dx} \left(\frac{\partial L}{\partial p}(u', u) \right) \right) = 0. \end{aligned}$$

Exercice 10.1.7 Montrer qu'une fonction J dérivable sur V est strictement convexe si et seulement si

$$J(v) > J(u) + \langle J'(u), v - u \rangle \quad \forall u, v \in V \quad \text{avec} \quad u \neq v,$$

ou encore

$$\langle J'(u) - J'(v), u - v \rangle > 0 \quad \forall u, v \in V \quad \text{avec} \quad u \neq v.$$

Correction. Notons tout d'abord, que ces équivalences ont été établies dans le cours dans le cas convexe avec des inégalités larges.

Soit J une fonction dérivable. Prouvons tout d'abord que J est strictement convexe si et seulement si

$$J(v) > J(u) + \langle J'(u), v - u \rangle \quad \forall u, v \in V \quad \text{avec} \quad u \neq v.$$

Soit J une fonction strictement convexe, u et $v \in V$ tels que $u \neq v$. On a

$$J\left(\frac{u+v}{2}\right) \geq J(u) + \left\langle J'(u), \frac{v-u}{2} \right\rangle.$$

De plus

$$J\left(\frac{u+v}{2}\right) < \frac{J(u) + J(v)}{2}.$$

Ainsi,

$$J(v) > J(u) + \langle J'(u), v - u \rangle.$$

Réciproquement, si J vérifie cette dernière inégalité, pour tout couple (u, v) , J est convexe. Ainsi, pour tout u et v , non seulement l'inégalité précédente est vérifiée, mais on a

$$2J\left(\frac{u+v}{2}\right) \geq 2J(v) + \langle J'(u), u - v \rangle.$$

En sommant ces deux inégalités, on obtient

$$2J\left(\frac{u+v}{2}\right) > J(u) + J(v).$$

Reste à prouver l'équivalence entre la stricte convexité et la deuxième inégalité.

Si J est une fonction strictement convexe, on vient de prouver que

$$J(v) > J(u) + \langle J'(u), v - u \rangle.$$

En commutant u et v dans cette inégalité, on obtient

$$J(u) > J(v) + \langle J'(v), u - v \rangle.$$

En sommant ces deux inégalités, on en déduit que

$$0 > \langle J'(v) - J'(u), u - v \rangle.$$

Réciproquement, si une fonction J vérifie cette inégalité pour tout couple (u, v) , elle est convexe. Ainsi,

$$J\left(\frac{u+v}{2}\right) \geq J(u) + \left\langle J'(u), \frac{u-v}{2} \right\rangle$$

et

$$J\left(\frac{u+v}{2}\right) \geq J(v) + \left\langle J'(v), \frac{v-u}{2} \right\rangle,$$

d'où

$$\begin{aligned} J\left(\frac{u+v}{2}\right) &\geq \frac{J(u) + J(v)}{2} + \frac{1}{4} \langle J'(u) - J'(v), u - v \rangle \\ &> \frac{J(u) + J(v)}{2} \end{aligned}$$

et J est strictement convexe.

Exercice 10.1.8 Soit a une forme bilinéaire symétrique continue sur $V \times V$. Soit L une forme linéaire continue sur V . On pose $J(u) = \frac{1}{2}a(u, u) - L(u)$. Montrer que J est deux fois dérivable sur V et que $J''(u)(v, w) = a(v, w)$ pour tout $u, v, w \in V$. Appliquer ce résultat aux exemples des Exercices **10.1.3**, **10.1.4**, **10.1.5**.

Correction. Tout d'abord, on montre que J est dérivable. En effet,

$$J(u + v) = J(u) + a(u, v) + L(v) + \frac{1}{2}a(v, v)$$

et comme a est continue, $a(v, v) = o(v)$. On a donc $J'(u) = a(u, \cdot) + L$. Montrons que J' est lui même dérivable au sens de Fréchet :

$$J'(u + w) = a(u + w, \cdot) + L = a(u, \cdot) + L + a(w, \cdot) = J'(u) + a(w, \cdot).$$

Ainsi, $J''(u)w = a(w, \cdot)$ ou encore $J''(u)(v, w) = a(v, w)$.

La fonctionnelle $J(x) = \frac{1}{2}Ax \cdot x - b \cdot x$ de l'Exercice 10.1.3 est deux fois dérivable dans \mathbb{R}^N et $J''(x)(X, Y) = AX \cdot Y$.

La fonctionnelle $J(u) = \frac{1}{2} \int_{\Omega} uv \, dx - \int_{\Omega} fu \, dx$ de l'Exercice 10.1.4 est deux fois dérivable dans $L^2(\Omega)$ et $J''(u)(v, w) = \int_{\Omega} vw \, dx$.

La fonctionnelle $J(u) = \frac{1}{2} \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, dx - \int_{\Omega} fu \, dx$ de l'Exercice 10.1.5 est deux fois dérivable dans $H_0^1(\Omega)$ et $J''(u)(v, w) = \int_{\Omega} (\nabla v \cdot \nabla w + vw) \, dx$.

Exercice 10.1.9 Montrer que si J est deux fois dérivable sur V les conditions des Propositions **10.1.4** et **10.1.5** sont respectivement équivalentes à

$$J''(u)(w, w) \geq 0 \quad \text{et} \quad J''(u)(w, w) \geq \alpha \|w\|^2 \quad \forall u, w \in V. \quad (10.3)$$

Correction. Montrons que pour tout $\alpha \geq 0$, les conditions de la proposition **10.1.5** sont équivalentes à

$$J''(u)(w, w) \geq \alpha \|w\|^2, \quad \forall u, w \in V$$

(l'équivalence avec les conditions de la Proposition **10.1.4** est obtenue en choisissant $\alpha = 0$). Supposons que pour tout u et v ,

$$J(v) \geq J(u) + \langle J'(u), v - u \rangle + \frac{\alpha}{2} \|u - v\|^2.$$

Comme J est deux fois différentiable,

$$\begin{aligned} J(v) &= J(u + w) \\ &= J(u) + \langle J'(u), w \rangle + \frac{1}{2} J''(u)(w, w) + o(\|w\|^2), \end{aligned}$$

où $w = v - u$. Ainsi, pour tout w ,

$$J''(u)(w, w) + o(\|w\|^2) \geq \alpha \|w\|^2.$$

Ainsi, pour tout $\lambda \neq 0$ et $w \neq 0$,

$$J''\left(\frac{w}{\|w\|}, \frac{w}{\|w\|}\right) + \frac{o(\|\lambda w\|^2)}{\|\lambda w\|^2} \geq \alpha.$$

En faisant tendre λ vers zéro, on obtient

$$J''\left(\frac{w}{\|w\|}, \frac{w}{\|w\|}\right) \geq \alpha$$

et $J''(w, w) \geq \alpha\|w\|^2$. Réciproquement, si $J''(w, w) \geq \alpha\|w\|^2$, On pose $f(t) = J(u + t(v - u))$. La fonction f est deux fois dérivable,

$$f'(t) = J'(u + t(v - u)) \cdot (v - u)$$

et

$$f''(t) = J''(u + t(v - u))(v - u, v - u) \geq \alpha\|v - u\|^2.$$

Ainsi,

$$f'(1) - f'(0) = \int_0^1 f''(t)dt \geq \alpha\|u - v\|^2$$

c'est à dire

$$\langle J'(u) - J'(v), u - v \rangle \geq \alpha\|u - v\|^2.$$

Exercice 10.2.1 Soit K un convexe fermé non vide de V . Pour $x \in V$, on cherche la projection $x_K \in K$ de x sur K (voir le Théorème 12.1.10)

$$\|x - x_K\| = \min_{y \in K} \|x - y\|.$$

Montrer que la condition nécessaire et suffisante

$$\langle J'(x_K), y - x_K \rangle \geq 0 \quad \forall y \in K \tag{10.4}$$

du Théorème 10.2.1 se ramène exactement à

$$\langle x_K - x, x_K - y \rangle \leq 0, \quad \forall y \in K. \tag{10.5}$$

Correction. Soit

$$J(y) = \|x - y\|^2.$$

La fonction J est dérivable de plus, pour tous éléments x_K et y de V , $\langle J'(x_K), y - x_K \rangle = 2\langle x - x_K, x_K - y \rangle$. La condition d'optimalité de x_K (10.4) est

$$\langle J'(x_K), y - x_K \rangle \geq 0 \text{ pour tout } y \in K,$$

c'est à dire

$$\langle x - x_K, x_K - y \rangle \geq 0 \text{ pour tout } y \in K,$$

qui n'est rien d'autre que (10.5).

Exercice 10.2.2 Soit A une matrice réelle d'ordre $p \times n$ et $b \in \mathbb{R}^p$. On considère le problème “aux moindres carrés”

$$\inf_{x \in \mathbb{R}^n} \|Ax - b\|^2.$$

Montrer que ce problème admet toujours une solution et écrire l'équation d'Euler correspondante.

Correction. On pose

$$J(x) = \|Ax - b\|^2.$$

Soit K l'orthogonal du noyau de A . On pose $\alpha = \inf_{u \in K \setminus \{0\}} \|Au\|^2 / \|u\|^2 > 0$. On constate que J est α -convexe sur K convexe. Elle admet donc un unique minimum sur K qui est un minimum sur \mathbb{R}^n , car $J(x + y) = J(x)$ pour tout élément y du noyau de A . Comme

$$\langle J'(x), y \rangle = 2(Ax - b) \cdot Ay,$$

l'équation d'Euler correspondante $J'(x) = 0$ est

$$A^*Ax = A^*b.$$

Exercice 10.2.3 On reprend l'Exemple 9.1.6

$$\inf_{x \in \text{Ker } B} \left\{ J(x) = \frac{1}{2} Ax \cdot x - b \cdot x \right\}$$

avec A matrice symétrique carrée d'ordre n , et B de taille $m \times n$ ($m \leq n$). Montrer qu'il existe une solution si A est positive et qu'elle est unique si A est définie positive. Montrer que tout point de minimum $\bar{x} \in \mathbb{R}^n$ vérifie

$$A\bar{x} - b = B^*p \text{ avec } p \in \mathbb{R}^m.$$

Correction. La fonctionnelle J est dérivable et $J'(x) = Ax - b$. Ainsi, un élément \bar{x} de $\text{Ker } B$ est un minimiseur de J sur $\text{Ker } B$ si et seulement si, pour tout $y \in \text{Ker } B$, $(Ax - b) \cdot y = 0$, c'est à dire $A\bar{x} - b \in (\text{Ker } B)^\perp$. Enfin,

$$\begin{aligned} (\text{Ker } B)^\perp &= \{x \in \mathbb{R}^n : Bx \cdot y = 0, \forall y \in \mathbb{R}^m\}^\perp \\ &= \{x \in \mathbb{R}^n : x \cdot B^*y = 0, \forall y \in \mathbb{R}^m\}^\perp \\ &= ((\text{Im } B^*)^\perp)^\perp \\ &= \text{Im } B^*. \end{aligned}$$

Il existe donc $p \in \mathbb{R}^m$ tel que $A\bar{x} - b = B^*p$.

Exercice 10.2.4 On reprend l'Exemple 9.1.10. Montrer que l'équation d'Euler vérifiée par le point de minimum $u \in H_0^1(\Omega)$ de

$$\inf_{v \in H_0^1(\Omega)} \left\{ J(v) = \frac{1}{2} \int_{\Omega} |\nabla v|^2 dx - \int_{\Omega} f v dx \right\}$$

est précisément la formulation variationnelle

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

(On retrouve ainsi un résultat de la Proposition 5.2.7.)

Correction.

$$J(u + v) = J(u) + \int_{\Omega} (\nabla u \cdot \nabla v - f v) \, dx + \frac{1}{2} \int_{\Omega} |\nabla v|^2 \, dx.$$

Ainsi, J est dérivable en tout point u de $H_0^1(\Omega)$ et

$$\langle J'(u), v \rangle = \int_{\Omega} (\nabla u \cdot \nabla v - f v) \, dx,$$

Au point de minimum de J , $J'(u) = 0$, c'est à dire

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx \text{ pour tout } v \in H_0^1(\Omega)$$

Exercice 10.2.5 Soit K un convexe fermé non vide de V , soit a une forme bilinéaire symétrique continue coercive sur V , et soit L une forme linéaire continue sur V . Montrer que $J(v) = \frac{1}{2}a(v, v) - L(v)$ admet un unique point de minimum dans K , noté u . Montrer que u est aussi l'unique solution du problème (appelé inéquation variationnelle)

$$u \in K \quad \text{et} \quad a(u, v - u) \geq L(v - u) \quad \forall v \in K.$$

Correction. La forme bilinéaire $a(.,.)$ étant coercive, la fonction J est fortement convexe. Elle admet donc un unique minimum u sur le convexe fermé non vide K . De plus, J étant symétrique,

$$\langle J'(u), w \rangle = a(u, w) - L(w).$$

Un élément u de K est un minimiseur de J sur K si et seulement si

$$\langle J'(u), v - u \rangle \geq 0, \text{ pour tout } v \in K,$$

c'est à dire

$$a(u, v - u) \geq L(v - u), \quad \forall v \in K.$$

Exercice 10.2.6 Soit J_1 et J_2 deux fonctions convexes continues sur une partie convexe fermée non vide $K \subset V$. On suppose que J_1 seulement est dérivable. Montrer que $u \in K$ est un minimum de $J_1 + J_2$ si et seulement si

$$\langle J_1'(u), v - u \rangle + J_2(v) - J_2(u) \geq 0 \quad \forall v \in K.$$

Correction. Soit u minimum de $J_1 + J_2$ sur K , alors pour tout $v \in K$ et $h \in]0, 1[$, $u + h(v - u) \in K$ et

$$\frac{J_1(u + h(v - u)) - J_1(u)}{h} + \frac{J_2(u + h(v - u)) - J_2(u)}{h} \geq 0$$

De plus,

$$J_2(u + h(v - u)) = J_2((1 - h)u + hv) \leq (1 - h)J_2(u) + hJ_2(v)$$

d'où

$$\frac{J_1(u + h(v - u)) - J_1(u)}{h} + J_2(v) - J_2(u) \geq 0.$$

En passant à la limite en $h \rightarrow 0$, on obtient

$$\langle J'_1(u), v - u \rangle + J_2(v) - J_2(u) \geq 0 \text{ pour tout } v \in K$$

La réciproque découle de (10.7). Si J_1 et J_2 vérifient l'équation précédente, J_1 étant convexe, on a

$$J_1(v) \geq J_1(u) + \langle J'_1(u), v - u \rangle.$$

Ainsi,

$$J_1(v) - J_1(u) + J_2(v) - J_2(u) \geq 0 \text{ pour tout } v \in K$$

et u est un minimiseur de $J_1 + J_2$ sur K .

Exercice 10.2.7 Soit K un sous-ensemble d'un espace de Hilbert V . Montrer que pour tout $v \in K$,

$$K(v) = \left\{ w \in V, \exists (v^n) \in K^{\mathbb{N}}, \exists (\varepsilon^n) \in (\mathbb{R}_+^*)^{\mathbb{N}}, \lim_{n \rightarrow +\infty} v^n = v, \lim_{n \rightarrow +\infty} \varepsilon^n = 0, \lim_{n \rightarrow +\infty} \frac{v^n - v}{\varepsilon^n} = w \right\}$$

est un cône fermé et que $K(v) = V$ si v est intérieur à K . Donner un exemple où $K(v)$ est réduit à $\{0\}$.

Correction. Montrons que $K(v)$ est un cône. Tout d'abord, 0 appartient toujours à $K(v)$ (il suffit de choisir $v_n = v$). Soit w un élément de $K(v)$ et α un réel strictement positif. D'après la définition de $K(v)$, il existe une suite v_n d'éléments de K , une suite ε_n de réels positifs tels que v_n converge vers v , ε_n converge vers zéro et

$$\frac{v_n - v}{\varepsilon_n} \rightarrow w.$$

On pose $\tilde{\varepsilon}_n = \alpha^{-1}\varepsilon_n$. On a

$$\frac{v_n - v}{\tilde{\varepsilon}_n} \rightarrow \alpha w,$$

d'où $\alpha w \in K(v)$ et $K(v)$ est un cône.

Montrons que $K(v)$ est fermé. Soit $w_m \in K(v)$ tel que $w_m \rightarrow w$. On note $v^{n,m}$ et $\varepsilon^{n,m}$ les suites telles que

$$\lim_{n \rightarrow +\infty} \frac{v^{n,m} - v}{\varepsilon^{n,m}} = w_m.$$

Pour tout $\delta > 0$, il existe m tel que

$$\|w_m - w\| \leq \delta/2$$

Comme $(v^{n,m} - v)/\varepsilon^{n,m}$ converge vers w_m lorsque n tend vers l'infini et, il existe n tel que

$$\left\| \frac{v^{n,m} - v}{\varepsilon^{n,m}} - w_m \right\| \leq \delta/2 \quad \text{et} \quad \|v^{n,m} - v\| \leq \delta.$$

On a montré que, pour tout $\delta > 0$, il existe $v_\delta = v^{n,m} \in K$ et $\varepsilon_\delta = \varepsilon^{n,m}$ tels que

$$\|\varepsilon_\delta\| \leq \delta, \quad \left\| \frac{v_\delta - v}{\varepsilon_\delta} - w \right\| \leq \delta \quad \text{et} \quad \|v_\delta - v\| \leq \delta.$$

Ainsi, w appartient à $K(v)$.

Si $K(v)$ est à l'intérieur de K , il existe un réel r strictement positif tel que la boule de rayon r centrée en v soit incluse dans K . Pour tout élément $w \in V$,

$$w = \lim_{n \rightarrow 0} \frac{v^n - v}{\varepsilon^n} \in K(v),$$

où $v^n = v + \frac{rw}{n\|w\|}$ et $\varepsilon^n = \frac{r}{n\|w\|}$. En d'autres termes, $V \subset K(v)$, d'où $K(v) = V$. Enfin, pour $K = 0$, $K(0) = \{0\}$.

Exercice 10.2.8 Soit A une matrice est symétrique définie positive d'ordre n , et B une matrice de taille $m \times n$ avec $m \leq n$. A l'aide des conditions d'optimalité du Théorème 10.2.8, déterminer une expression explicite de la solution \bar{x} du problème d'optimisation

$$\min_{Bx=c} \left\{ J(x) = \frac{1}{2}Ax \cdot x - b \cdot x \right\},$$

où $c \in \mathbb{R}^m$ est un vecteur donné.

Correction. Les conditions d'optimalité s'écrivent à nouveau

$$A\bar{x} - b = B^*p.$$

Ainsi, $\bar{x} = A^{-1}(b + B^*p)$ et comme $B\bar{x} = c$. Si B est de rang m , $BA^{-1}B^*$ est inversible et

$$p = (BA^{-1}B^*)^{-1}(c - BA^{-1}b) \text{ et } x = A^{-1}b + A^{-1}B^*(BA^{-1}B^*)^{-1}(c - BA^{-1}b).$$

Si B n'est pas de rang maximal, les contraintes sont soit redondantes, soit contradictoires. Si elles sont contradictoires, il n'y a pas d'optimum (l'ensemble de minimisation est vide). Si les contraintes sont redondantes, il existe $p \in \mathbb{R}^m$ tel que

$$BA^{-1}B^*p = c - BA^{-1}b,$$

et p est défini à l'addition d'un élément de $\text{Ker } B^*$ près. Par contre, \bar{x} est défini de manière unique par la relation $\bar{x} = A^{-1}(b + B^*p)$.

Exercice 10.2.9 On reprend l'Exemple 9.1.7. Soit A une matrice symétrique d'ordre n et $J(x) = Ax \cdot x$. A l'aide du Théorème 10.2.8, montrer que les points de minimum de J sur la sphère unité sont des vecteurs propres de A associés à la plus petite valeur propre.

Correction. On note K la sphère unité, définie par

$$K = \{x \in \mathbb{R}^n : F(x) = 0\},$$

où $F(x) = 1 - |x|^2$. Les fonctions J et F sont toutes deux dérivables et

$$J'(x) = 2Ax \quad F'(x) = -2x.$$

Ainsi, d'après le Théorème 10.2.8, si \bar{x} est un point de minimum de J sur la sphère unité, il existe λ tel que

$$J'(\bar{x}) + \lambda F'(\bar{x}) = 0,$$

c'est à dire

$$A\bar{x} - \lambda\bar{x} = 0.$$

Toute solution optimale \bar{x} est un vecteur propre de A de valeur propre λ . Notons que l'existence d'un minimiseur est évidente, K étant compact et J continue. Le problème de minimisation de J sur K est donc équivalent au problème de minimisation de J sur l'ensemble des vecteurs propres de A de norme un. Or pour tout vecteur propre x de A (tel que $\|x\| = 1$) de valeur propre μ , on a

$$J(x) = \mu.$$

Le minimum de J est donc atteint pour les vecteurs propres de plus petite valeur propre.

Exercice 10.2.10 En utilisant les résultats précédents et ceux de l'Exercice 10.1.6, montrer que la solution du problème de Didon (Exemple 9.1.11) est nécessairement un arc de cercle.

Correction. Tout d'abord, rappelons (en termes un peu simplifiés) le problème de Didon tel qu'il est posé dans l'Exemple 9.1.11. Il s'agit de déterminer ξ et $y : [0, \xi] \rightarrow \mathbb{R}$ tel que $y(0) = y(\xi) = 0$, maximisant

$$J(y) = \int_0^\xi y(x) dx,$$

sous la contrainte

$$L(y) = \int_0^\xi \sqrt{1 + |y'|^2} dx - l = 0.$$

On suppose que y, ξ est solution de ce problème. En particulier, y est solution du même problème pour ξ fixé. On souhaite prouver que toute solution y à ce dernier problème est un arc de cercle.

D'après l'exercice **10.1.6**, la fonctionnelle L est dérivable et pour toute fonction $v \in H_0^1([0, \xi])$, on a

$$\langle L'(y), v \rangle = \int_0^\xi \frac{1}{\sqrt{1 + |y'|^2}} y' v' dx.$$

La fonctionnelle J est également dérivable car linéaire. Ainsi, les conditions d'optimalité d'ordre un (Théorème **10.2.8**) impliquent que si y est une solution, il existe λ tel que

$$J'(y) + \lambda L'(y) = 0$$

pourvu que $L'(y) \neq 0$. Le cas $L'(y) = 0$ se traite de manière triviale et conduit à la solution $y = 0$. On a donc

$$\int_0^\xi v + \frac{\lambda}{\sqrt{1 + |y'|^2}} y' v' dx = 0$$

pour tout $v \in H_0^1([0, \xi])$. En intégrant par partie le second membre de cette équation, on en déduit que

$$\lambda \left(\frac{y'}{\sqrt{1 + |y'|^2}} \right)' = 1$$

et qu'il existe une constante C telle que

$$\frac{y'}{\sqrt{1 + |y'|^2}} = \lambda^{-1} x + C. \quad (10.6)$$

Dans un premier temps, on élève cette équation au carré afin de déterminer $|y'|^2$ en fonction de x . On obtient

$$\frac{1}{1 + |y'|^2} = 1 - (\lambda^{-1} x + C)^2.$$

En substituant cette expression dans l'équation (10.6), on en déduit que

$$y' = \frac{\lambda^{-1} x + C}{\sqrt{1 - (\lambda x + C)^2}}.$$

Par intégration, il existe une constante D telle que

$$y = -\lambda \sqrt{1 - (\lambda^{-1} x + C)^2} + D.$$

Pour conclure, il suffit de constater que

$$(y - D)^2 + (x + \lambda C)^2 = \lambda^2.$$

Ainsi, $(x, y(x))$ est bien un arc de cercle. Remarquons que le multiplicateur de Lagrange λ associé à la contrainte sur la longueur n'est autre que le rayon du cercle obtenu.

Exercice 10.2.11 On étudie la première valeur propre du Laplacien dans un domaine borné Ω (voir la Section 7.3). Pour cela on introduit le problème de minimisation sur $K = \{v \in H_0^1(\Omega), \int_{\Omega} v^2 dx = 1\}$

$$\min_{v \in K} \left\{ J(v) = \int_{\Omega} |\nabla v|^2 dx \right\}.$$

Montrer que ce problème admet un minimum (on montrera que K est compact pour les suites minimisantes à l'aide du Théorème de Rellich 4.3.21). Écrire l'équation d'Euler de ce problème et en déduire que la valeur du minimum est bien la première valeur propre et que les points de minimum sont des vecteurs propres associés.

Correction. Pour tout $v \in H_0^1(\Omega)$, on note

$$|v|_{H_0^1(\Omega)} = \left(\int_{\Omega} |\nabla v|^2 dx \right)^{1/2}.$$

D'après l'inégalité de Poincaré, $|\cdot|_{H_0^1(\Omega)}$ est une norme équivalente à la norme usuelle de $H_0^1(\Omega)$. Soit u_n une suite minimisante de J sur K . D'après le Théorème de Rellich, il existe une sous suite de u_n (que nous noterons également u_n) et un élément $u \in H_0^1(\Omega)$ tel que u_n converge vers u dans $L^2(\Omega)$. Montrons que (u_n) est une suite convergente dans $H_0^1(\Omega)$. Tout d'abord,

$$\left| \frac{u_n - u_p}{2} \right|_{H_0^1}^2 = \frac{|u_n|_{H_0^1(\Omega)}^2 + |u_p|_{H_0^1(\Omega)}^2}{2} - \left| \frac{u_n + u_p}{2} \right|_{H_0^1}^2. \quad (10.7)$$

On note

$$\mu = \inf_{v \in K} J(v)$$

et

$$\alpha_{n,p} = \left\| \frac{u_n + u_p}{2} \right\|_{L^2(\Omega)}.$$

Comme u_n converge vers u dans $L^2(\Omega)$, $\|u\|_{L^2(\Omega)} = 1$ et $u \in K$. De plus, $\alpha_{n,p}$ converge vers 1 lorsque n et p tendent vers l'infini. D'après l'équation (10.7),

$$\left| \frac{u_n - u_p}{2} \right|_{H_0^1}^2 = \frac{|u_n|_{H_0^1(\Omega)}^2 + |u_p|_{H_0^1(\Omega)}^2}{2} - \alpha_{n,p}^2 \left| \frac{u_n + u_p}{2\alpha_{n,p}} \right|_{H_0^1}^2.$$

Comme $\frac{u_n + u_p}{2\alpha_{n,p}} \in K$, on a donc

$$\left| \frac{u_n - u_p}{2} \right|_{H_0^1}^2 \leq \frac{|u_n|_{H_0^1(\Omega)}^2 + |u_p|_{H_0^1(\Omega)}^2}{2} - \alpha_{n,p}^2 \mu.$$

Ainsi, $|u_n - u_p|_{H_0^1} \rightarrow 0$ lorsque n et p tendent vers l'infini et u_n est une suite de Cauchy dans $H_0^1(\Omega)$. Ainsi, u_n converge dans $H_0^1(\Omega)$ vers u et $J(u) = \mu$.

Soit $F(v) = 1 - \int_{\Omega} |v|^2 dx$. L'ensemble de minimisation K est donné par

$$K = \{v \in H_0^1(\Omega) : F(v) = 0\}.$$

De plus, F est dérivable et pour tout $v, w \in H_0^1(\Omega)$, on a

$$\langle F'(v), w \rangle = -2 \int_{\Omega} v w dx.$$

de même, J est dérivable et

$$\langle J'(v), w \rangle = 2 \int_{\Omega} \nabla v \cdot \nabla w dx.$$

D'après le Théorème **10.2.8**, comme F' est non nul pour tout élément de K (et donc en particulier pour u), il existe λ tel que

$$J'(u) + \lambda F'(u) = 0,$$

c'est à dire tel que pour tout $v \in H_0^1(\Omega)$,

$$\int_{\Omega} \nabla u \cdot \nabla v dx = \lambda \int_{\Omega} u v dx.$$

Ainsi, u est un vecteur propre de valeur propre λ . En choisissant $v = u$ dans l'expression précédente, on en déduit de plus que $\lambda = \mu$. Enfin, on vérifie sans peine que λ est nécessairement la plus petite valeur propre du Laplacien avec conditions aux bords de Dirichlet.

Exercice 10.2.12 Soit A une matrice $n \times n$ symétrique définie positive et $b \in \mathbb{R}^n$ non nul.

1. Montrer que les problèmes

$$\sup_{Ax \cdot x \leq 1} b \cdot x \quad \text{et} \quad \sup_{Ax \cdot x = 1} b \cdot x$$

sont équivalents et qu'ils ont une solution. Utiliser le Théorème **10.2.8** pour calculer cette solution et montrer qu'elle est unique.

2. On introduit un ordre partiel dans l'ensemble des matrices symétriques définies positives d'ordre n en disant que $A \geq B$ si et seulement si $Ax \cdot x \geq Bx \cdot x$ pour tout $x \in \mathbb{R}^n$. Dédurre de la question précédente que, si $A \geq B$, alors $B^{-1} \geq A^{-1}$.

Correction. 1. Si $b = 0$, le résultat est évident. On peut donc supposer par la suite $b \neq 0$. On a pose $J(x) = b \cdot x$. Soit \bar{x} la solution du problème de maximisation de J sur

$$K = \{x \in \mathbb{R}^n \text{ tel que } Ax \cdot x \leq 1\}.$$

Comme la dérivée de J est égale à b et n'est jamais nulle, le maximum de J sur K ne peut être atteint dans l'intérieur de K . Il est donc atteint sur le bord, d'où

$$\sup_{Ax \cdot x \leq 1} Ax \cdot x = \sup_{Ax \cdot x = 1} Ax \cdot x.$$

Les deux problèmes sont équivalents. Reste à déterminer la solution de ces problèmes. D'après les conditions d'optimalités du premier ordre, il existe λ tel que

$$A\bar{x} - \lambda b = 0.$$

Ainsi,

$$\bar{x} = \lambda A^{-1}b.$$

Il ne reste plus qu'à déterminer le multiplicateur de Lagrange λ pour définir \bar{x} de manière unique. Comme $A\bar{x} \cdot \bar{x} = 1$, on en déduit que

$$\lambda^2 = (A^{-1}b \cdot b)^{-1}$$

Comme λ est positif, on a

$$\lambda = (A^{-1}b \cdot b)^{-1/2},$$

ce qui détermine \bar{x} de manière unique.

2. Soit A et B deux matrices symétriques définies positives telles que $A \geq B$. Pour tout b non nul, on a

$$(A^{-1}b \cdot b)^{1/2} = \sup_{Ax \cdot x \leq 1} b \cdot x \leq \sup_{Bx \cdot x \leq 1} b \cdot x = (B^{-1}b \cdot b)^{1/2}.$$

d'où $B^{-1} \geq A^{-1}$.

Exercice 10.2.13 En théorie cinétique des gaz les molécules de gaz sont représentées en tout point de l'espace par une fonction de répartition $f(v)$ dépendant de la vitesse microscopique $v \in \mathbb{R}^N$. Les quantités macroscopiques, comme la densité du gaz ρ , sa vitesse u , et sa température T , se retrouvent grâce aux moments de la fonction $f(v)$

$$\rho = \int_{\mathbb{R}^N} f(v) dv, \quad \rho u = \int_{\mathbb{R}^N} v f(v) dv, \quad \frac{1}{2}\rho u^2 + \frac{N}{2}\rho T = \frac{1}{2} \int_{\mathbb{R}^N} |v|^2 f(v) dv. \quad (10.8)$$

Boltzmann a introduit l'entropie cinétique $H(f)$ définie par

$$H(f) = \int_{\mathbb{R}^N} f(v) \log(f(v)) dv.$$

Montrer que H est strictement convexe sur l'espace des fonctions $f(v) > 0$ mesurables telle que $H(f) < +\infty$. On minimise H sur cet espace sous les contraintes de moment (10.8), et on admettra qu'il existe un unique point de minimum $M(v)$. Montrer que ce point de minimum est une Maxwellienne définie par

$$M(v) = \frac{\rho}{(2\pi T)^{N/2}} \exp\left(-\frac{|v - u|^2}{2T}\right).$$

Correction. La fonction $\varphi(t) = t \log(t)$ est strictement convexe sur $\mathbb{R}^+ \setminus \{0\}$, en effet, $\varphi''(t) = 1/t > 0$. On en déduit que

$$\begin{aligned} H(\theta f + (1 - \theta)g) &= \int_{\mathbb{R}^N} \varphi(\theta f + (1 - \theta)g) dv \\ &\leq \int_{\mathbb{R}^N} \theta \varphi(f) + (1 - \theta) \varphi(g) dv \\ &= \theta H(f) + (1 - \theta) H(g). \end{aligned}$$

Ainsi, H est convexe. De plus, l'inégalité est une égalité si et seulement si

$$\varphi(\theta f + (1 - \theta)g) = \theta\varphi(f) + (1 - \theta)\varphi(g)$$

presque partout. En particulier, si θ est différent de 0 et 1, on en déduit que $f = g$ presque partout. La fonction H est donc strictement convexe (quitte à identifier les fonctions égales presque partout). On a

$$\langle H'(f), g \rangle = \int_{\mathbb{R}^N} ((\log f(v)) + 1)g(v) dv.$$

Les contraintes sont linéaires et les conditions d'optimalité du premier ordre impliquent qu'il existe λ_1 et λ_3 réels, $\lambda_2 \in \mathbb{R}^N$ tels que

$$\int_{\mathbb{R}^N} ((\log f(v)) + 1 + \lambda_1 + \lambda_2 \cdot v + |v|^2 \lambda_3)g(v) dv = 0$$

pour tout g . En d'autres termes,

$$(\log f(v)) + 1 + \lambda_1 + \lambda_2 \cdot v + |v|^2 \lambda_3 = 0$$

presque partout ou encore

$$f(v) = \exp(-1 - \lambda_1 - \lambda_2 \cdot v - \lambda_3 |v|^2).$$

Reste à déterminer les multiplicateurs de Lagrange λ_1 , λ_2 et λ_3 . Un calcul un peu fastidieux permet de montrer que

$$\int_{\mathbb{R}^N} \exp(-1 - \lambda_1 - \lambda_2 \cdot v - \lambda_3 |v|^2) dv = \sqrt{\pi}^N \frac{e^{-(1+\lambda_1)} e^{|\lambda_2|^2/4\lambda_3}}{\sqrt{\lambda_3}^N},$$

$$\int_{\mathbb{R}^N} v \exp(-1 - \lambda_1 - \lambda_2 \cdot v - \lambda_3 |v|^2) dv = -\sqrt{\pi}^N \lambda_2 \frac{e^{-(1+\lambda_1)} e^{|\lambda_2|^2/4\lambda_3}}{2\sqrt{\lambda_3}^{N+2}}$$

et

$$\begin{aligned} \int_{\mathbb{R}^N} |v|^2 \exp(-1 - \lambda_1 - \lambda_2 \cdot v - \lambda_3 |v|^2) dv = \\ \frac{e^{-(1+\lambda_1)} e^{|\lambda_2|^2/4\lambda_3}}{|\lambda_2|^3 \sqrt{\lambda_3}^{N-1}} \left(\frac{N}{2} \sqrt{\pi}^N \left(\frac{|\lambda_2|}{\sqrt{\lambda_3}} \right)^3 + \frac{\sqrt{\pi}^N}{4} \left(\frac{|\lambda_2|}{\sqrt{\lambda_3}} \right)^5 \right). \end{aligned}$$

Les contraintes vérifiées par v nous permettent de déterminer les multiplicateurs de Lagrange. On obtient $\lambda_2 = -u/T$, $\lambda_3 = (2T)^{-1}$ et $e^{-(1+\lambda_1)} = \sqrt{2\pi} T^{-N} e^{-|u|^2/2T} \rho$, d'où on conclut que $f = M$.

Exercice 10.2.14 Calculer la condition nécessaire d'optimalité du second ordre pour les Exemples 9.1.6 et 9.1.7

1.

$$\inf_{x \in \text{Ker } B} \left\{ J(x) = \frac{1}{2} Ax \cdot x - b \cdot x \right\},$$

où A est une matrice carrée d'ordre n , symétrique définie positive, B une matrice rectangulaire de taille $m \times n$ et b un vecteur de \mathbb{R}^n .

2.

$$\inf_{x \in \mathbb{R}^n, \|x\|=1} \{ J(x) = Ax \cdot x \},$$

où A est une matrice carrée d'ordre n , symétrique définie.

Correction.

1. On note F la fonction contrainte $F(x) = Bx$. On a

$$J''(u)(v, v) = Av \cdot v$$

De plus, $F'' = 0$. La condition d'optimalité d'ordre deux est donc

$$Av \cdot v \geq 0$$

pour tout $v \in \text{Ker } B$. Comme A est définie positive, cette condition est toujours vérifiée.

2. On note F la fonction de contrainte $F(x) = x \cdot x - 1$. D'après la condition d'optimalité du premier ordre, si u est une solution du problème de minimisation, il existe $\lambda \in \mathbb{R}$ tel que

$$2Au + \lambda u = 0.$$

Comme

$$J''(u)(v, v) = 2Av \cdot v$$

et $F''(u)(v, v) = v \cdot v$, la condition d'optimalité d'ordre deux est donc

$$2Av \cdot v + \lambda v \cdot v \geq 0$$

pour tout v tel que $v \cdot u = 0$.

Exercice 10.2.15 Soit A une matrice symétrique définie positive d'ordre n , et B une matrice de taille $m \times n$ avec $m \leq n$ et de rang m . On considère le problème de minimisation

$$\min_{x \in \mathbb{R}^n, Bx \leq c} \left\{ J(x) = \frac{1}{2} Ax \cdot x - b \cdot x \right\},$$

Appliquer le Théorème 10.2.15 pour obtenir l'existence d'un multiplicateur de Lagrange $p \in \mathbb{R}^m$ tel qu'un point de minimum \bar{x} vérifie

$$A\bar{x} - b + B^*p = 0, \quad p \geq 0, \quad p \cdot (B\bar{x} - c) = 0.$$

Correction. L'ensemble des solutions admissibles est défini par

$$K = \{x \in \mathbb{R}^n : F_i(x) \leq 0 \text{ pour tout } i = 1, \dots, m\},$$

où $F_i(x) = B_i x - c_i$. Les fonctions F_i sont dérivables et $\langle F'(x), y \rangle = B_i y = (B_i)^* \cdot y$. De même, la fonction objectif

$$J(x) = \frac{1}{2} A x \cdot x - b \cdot x$$

est dérivable et

$$J'(x) = Ax - b.$$

Comme les contraintes sont affines, elles sont automatiquement qualifiées. On peut appliquer le Théorème 10.2.15. Si \bar{x} est la solution du problème de minimisation de J sur K , il existe donc $p \in \mathbb{R}^m$ tel que

$$J'(x) + \sum_i p_i F'_i(x) = 0, \quad p_i \geq 0, \quad p_i F'_i = 0,$$

c'est à dire

$$A\bar{x} - b + \sum_i p_i (B_i)^* = 0, \quad p_i \geq 0, \quad p_i (B_i \bar{x} - c_i) = 0.$$

ou, sous une forme plus compacte,

$$A\bar{x} - b + B^* p = 0, \quad p \geq 0, \quad p \cdot (B\bar{x} - c) = 0.$$

Notons que K étant convexe et J fortement convexe, il existe un unique minimiseur au problème considéré.

Exercice 10.2.16 Soit $f \in L^2(\Omega)$ une fonction définie sur un ouvert borné Ω . Pour $\epsilon > 0$ on considère le problème de régularisation suivant

$$\min_{u \in H_0^1(\Omega), \|u-f\|_{L^2(\Omega)} \leq \epsilon} \int_{\Omega} |\nabla u|^2 dx.$$

Montrer que ce problème admet une unique solution u_ϵ . Montrer que, soit $u_\epsilon = f$, soit $u_\epsilon = 0$, soit il existe $\lambda > 0$ tel que u_ϵ est solution de

$$\begin{cases} -\Delta u_\epsilon + \lambda(u_\epsilon - f) = 0 & \text{dans } \Omega, \\ u_\epsilon = 0 & \text{sur } \partial\Omega. \end{cases}$$

Correction. On note J la fonction objectif

$$J(u) = \int_{\Omega} |\nabla u|^2 dx$$

et K l'ensemble des solutions admissibles, c'est à dire

$$K = \{v \in H_0^1(\Omega) : F(v) \leq 0\},$$

où $F(v) = \|v - f\|_{L^2(\Omega)}^2 - \epsilon^2$. L'ensemble K est un convexe fermé tandis que la fonctionnelle J est fortement convexe. Il existe donc une unique solution u_ϵ au problème de minimisation de J sur K . Les fonctionnelles J et F sont toutes deux dérivables et, pour tout $v \in H_0^1(\Omega)$, on a

$$\langle J'(u_\epsilon), v \rangle = 2 \int_{\Omega} \nabla u_\epsilon \cdot \nabla v dx$$

et

$$\langle F'(u_\epsilon), v \rangle = 2 \int_{\Omega} (u_\epsilon - f)v dx.$$

Si la contrainte est active, c'est à dire si $F(u_\epsilon) = 0$, on a $F'(u_\epsilon) = 0$. D'après le Théorème 10.2.15, il existe un réel $\lambda \geq 0$ tel que

$$J'(u_\epsilon) + \lambda F'(u_\epsilon) = 0, \quad \lambda F(u_\epsilon) = 0,$$

c'est à dire tel que pour tout $v \in H_0^1(\Omega)$,

$$\int_{\Omega} \nabla u_\epsilon \cdot \nabla v + \lambda(u_\epsilon - f)v dx = 0, \quad \lambda(\|u_\epsilon - f\|_{L^2}^2 - \epsilon) = 0.$$

On déduit de la première équation que u_ϵ est solution du problème aux limites

$$\begin{cases} -\Delta u_\epsilon + \lambda(u_\epsilon - f) = 0 & \text{dans } \Omega, \\ u_\epsilon = 0 & \text{sur } \partial\Omega. \end{cases}$$

Si la contrainte n'est pas active, $\lambda = 0$ et $u = 0$ (cas $\epsilon \geq \|f\|_{L^2}$).

Exercice 10.3.1 On considère le problème d'optimisation, dit perturbé

$$\inf_{F_i(v) \leq u_i, 1 \leq i \leq m} J(v), \quad (10.9)$$

avec $u_1, \dots, u_m \in \mathbb{R}$.

On se place sous les hypothèses du Théorème 10.3.4 de Kuhn et Tucker. On note $m^*(u)$ la valeur minimale du problème perturbé (10.9).

1. Montrer que si p est le multiplicateur de Lagrange pour le problème non perturbé (c'est-à-dire (10.9) avec $u = 0$), alors

$$m^*(u) \geq m^*(0) - pu. \quad (10.10)$$

2. Dédire de (10.10) que si $u \mapsto m^*(u)$ est dérivable, alors

$$p_i = -\frac{\partial m^*}{\partial u_i}(0).$$

Interpréter ce résultat (cf. l'Exemple 9.1.8 en économie).

Correction.

1. D'après le Théorème du **10.2.15**, la solution \bar{v} du problème (10.9) non perturbé est telle qu'il existe $p_i \geq 0$ tel que

$$J'(\bar{v}) + p_i F'_i(\bar{v}) = 0, \quad p_i F_i(\bar{v}) = 0. \quad (10.11)$$

Comme les fonctions J et F_i sont supposées convexes, pour tout v , on a

$$J(v) + p \cdot F(v) - J(\bar{v}) - p \cdot F(\bar{v}) \geq \langle J'(\bar{v}) + p \cdot F'(\bar{v}), v - \bar{v} \rangle.$$

D'après l'équation (10.11), on a donc

$$J(v) + p \cdot F(v) - J(\bar{v}) \geq 0.$$

Enfin, si v est la solution du problème perturbé, on en déduit comme $F(v) \leq u$ que

$$m^*(u) + p \cdot u - m^*(0) \geq 0.$$

2. Supposons que l'application $u \mapsto m^*(u)$ soit dérivable. Dans ce cas,

$$m^*(u) = m^*(0) + \frac{\partial m^*}{\partial u}(0) \cdot u + o(u).$$

Ainsi, d'après la question précédente,

$$\left(\frac{\partial m^*}{\partial u}(0) + p \right) \cdot u + o(u) \geq 0$$

pour tout u . En divisant cette équation par la norme de u , on obtient que pour tout élément u de norme unité,

$$\left(\frac{\partial m^*}{\partial u}(0) + p \right) \cdot u \geq 0.$$

En appliquant cette inégalité à $-u$ au lieu de u , on en déduit que

$$\frac{\partial m^*}{\partial u}(0) + p = 0.$$

Lorsque u augmente, l'ensemble des solutions admissibles croît. Ainsi, la valeur de $m^*(u)$, solution du problème de minimisation, ne peut que décroître. Grâce au multiplicateur de Lagrange p , on a une information supplémentaire : il nous permet de déterminer le taux de décroissance de $m^*(u)$ est fonction de u . Plus p est important, plus une petite variation de u par rapport à zéro entraînera une forte variation de m^* .

Exercice 10.3.2 Donner un exemple de Lagrangien pour lequel l'inégalité (10.60) est stricte avec ses deux membres finis.

Correction. On pose $U = \mathbb{R}$, $P = \mathbb{R}$ et

$$\mathcal{L}(v, q) = F(v + q),$$

où F est une fonction bornée non constante. On a alors

$$\inf_{v \in U} \left(\sup_{q \in P} \mathcal{L}(v, q) \right) = \sup_{\mathbb{R}} F > \inf_{\mathbb{R}} F = \sup_{q \in P} \left(\inf_{v \in U} \mathcal{L}(v, q) \right). \quad (10.12)$$

Exercice 10.3.3 Soit U (respectivement P) un convexe compact non vide de V (respectivement Q). On suppose que le Lagrangien est tel que $v \rightarrow \mathcal{L}(v, q)$ est strictement convexe continue sur U pour tout $q \in P$, et $q \rightarrow \mathcal{L}(v, q)$ est concave continue sur P pour tout $v \in U$. Montrer alors l'existence d'un point selle de \mathcal{L} sur $U \times P$.

Correction. Pour tout $q \in P$, on note $\varphi(q)$ l'unique minimiseur sur U de l'application $v \mapsto \mathcal{L}(v, q)$ (l'existence est assurée par la compacité de U et la continuité de \mathcal{L} , l'unicité par la stricte convexité de $v \mapsto \mathcal{L}(v, q)$). De plus, on pose

$$F(q) = \mathcal{L}(\varphi(q), q) = \min_{v \in U} \mathcal{L}(v, q).$$

L'application F est l'infimum d'une famille de fonctions concaves, semi-continues supérieurement. Elle est donc elle-même concave et semi-continue supérieurement. Comme U est compact et que F est semi-continue supérieurement, F admet au moins un maximum sur V noté q^* . On pose de plus $v^* = \varphi(q^*)$. On va montrer que (v^*, q^*) est un point selle de \mathcal{L} sur $U \times P$, c'est à dire que

$$\mathcal{L}(v^*, q) \leq \mathcal{L}(v^*, q^*) \leq \mathcal{L}(v, q^*)$$

pour tout couple $(v, q) \in U \times P$. La deuxième inégalité est évidente et découle simplement de la définition de $v^* = \varphi(q^*)$. Il reste à prouver que pour tout $q \in V$,

$$\mathcal{L}(v^*, q) \leq \mathcal{L}(v^*, q^*). \quad (10.13)$$

Pour tout $t \in [0, 1]$ et tout $q \in V$, on pose

$$v_t = \varphi((1 - t)q^* + tq)$$

D'après la concavité de $\mathcal{L}(v, \cdot)$, on a pour tout $v \in U$

$$\mathcal{L}(v, (1 - t)q^* + tq) \geq (1 - t)\mathcal{L}(v, q^*) + t\mathcal{L}(v, q),$$

d'où on déduit (puisque q^* maximise F sur P et $\mathcal{L}(v_t, q^*) \geq F(q^*)$), que

$$\begin{aligned} F(q^*) &\geq F((1 - t)q^* + tq) = \mathcal{L}(v_t, (1 - t)q^* + tq) \\ &\geq (1 - t)\mathcal{L}(v_t, q^*) + t\mathcal{L}(v_t, q) \\ &\geq (1 - t)F(q^*) + t\mathcal{L}(v_t, q), \end{aligned}$$

ce qui donne en fin de compte que pour tout $q \in V$ et tout $t \neq 0$,

$$F(q^*) \geq \mathcal{L}(v_t, q).$$

Comme U est compact, il existe une suite t_n convergent vers zéro tel que v_{t_n} soit convergente. Soit \tilde{v} la limite de v_{t_n} . D'après l'inégalité précédente, on a

$$F(q^*) = \mathcal{L}(v^*, q^*) \geq \lim \mathcal{L}(v_{t_n}, q) = \mathcal{L}(\tilde{v}, q).$$

Pour conclure, il suffit donc de prouver que $\tilde{v} = v^*$ et ainsi obtenir l'inégalité (10.13). Or, pour tout n on a

$$\begin{aligned} (1 - t_n)\mathcal{L}(v_{t_n}, q^*) + t_n\mathcal{L}(v_{t_n}, q) &\leq \mathcal{L}(v_{t_n}, (1 - t_n)q^* + t_nq) \\ &\leq \mathcal{L}(v, (1 - t_n)q^* + t_nq). \end{aligned}$$

En passant à la limite, on en déduit que pour tout $v \in U$,

$$\mathcal{L}(\tilde{v}, q^*) \leq \mathcal{L}(v, q^*).$$

Ainsi, \tilde{v} est un minimiseur de $v \mapsto \mathcal{L}(v, q^*)$. Comme cette dernière application est strictement convexe, elle admet au plus un minimiseur et $\tilde{v} = \varphi(q^*) = v^*$.

Exercice 10.3.4 Soit une matrice rectangulaire

$$A = \begin{pmatrix} 1 & 0 & 4 & 2 & 3 & 5 \\ -3 & 2 & -1 & 2 & -5 & 2 \\ -4 & 2 & -2 & 0 & -1 & 2 \\ -2 & 4 & -1 & 6 & -2 & 2 \\ -1 & 2 & -6 & 3 & -1 & 1 \end{pmatrix}.$$

On suppose que deux joueurs choisissent l'un une ligne i , l'autre une colonne j , sans qu'ils ne connaissent le choix de l'autre. Une fois révélé leurs choix, le gain (ou la perte, selon le signe) du premier joueur est déterminé par le coefficient a_{ij} de la matrice A (l'autre joueur recevant ou payant $-a_{ij}$). Montrer que la stratégie optimale de minimisation du risque conduit à un problème de min-max que l'on résoudra. Le jeu est-il équitable avec cette matrice A ?

Correction. Le premier joueur cherche à maximiser son gain quelque soit le choix du deuxième joueur, il choisit donc la ligne i tel que $\min_j a_{i,j}$ soit maximal. En adoptant cette stratégie, son gain minimal est alors

$$G_1 = \max_i \min_j a_{ij}.$$

Le deuxième joueur tient un raisonnement identique. Son gain minimal est donc

$$G_2 = - \min_j \max_i a_{ij}.$$

On résout aisément ces deux problèmes. La solution au premier problème pour le premier joueur consiste à jouer la première ligne ce qui lui assure un gain au moins nul (il ne peut pas perdre). La stratégie minimisant les risques pour le deuxième joueur consiste à jouer la première colonne ce qui lui assure au moins un gain de -1 , c'est à dire au pire une perte de 1. Le jeu n'est pas équitable. Si les deux joueurs adoptent cette stratégie, le premier joueur gagne 1 tandis que le deuxième perd 1.

Exercice 10.4.1 On considère le problème de commande optimal (10.72). On suppose que $K = \mathbb{R}^M$, $f = 0$, $z = 0$, et $z_T = 0$. Montrer que, pour tout $t \in [0, T]$,

$$p(t) \cdot y(t) = Dy(T) \cdot y(T) + \int_t^T Qy(s) \cdot y(s) ds + \int_t^T R^{-1}B^*p(s) \cdot B^*p(s) ds.$$

En déduire que s'il existe $t_0 \in [0, T]$ tel que $y(t_0) = 0$, alors $y(t) = p(t) = 0$ pour tout $t \in [0, T]$. Interpréter ce résultat.

Correction. Soit u la solution optimale au problème (10.72), y la solution au problème (10.71) associé et p la solution au problème adjoint (10.76). Un simple calcul nous donne

$$\begin{aligned} \frac{d}{dt}(p \cdot y) &= -Qy \cdot y - A^*p \cdot y + p \cdot Ay - B^*p \cdot R^{-1}B^*p \\ &= -Qy \cdot y - R^{-1}B^*p \cdot B^*p. \end{aligned}$$

Par intégration, on en déduit que

$$\begin{aligned} p \cdot y(t) &= p \cdot y(T) + \int_t^T Qy \cdot y + R^{-1}B^*p \cdot B^*p dt \\ &= Dy(T) \cdot y(T) + \int_t^T Qy \cdot y + R^{-1}B^*p \cdot B^*p dt. \end{aligned}$$

S'il existe $t_0 \in [0, T]$ tel que $y(t_0) = 0$, on a $p \cdot y(t_0) = 0$. Comme tous les termes du second membre de la formule précédente sont positifs ou nuls et de somme nulle, ils sont tous nuls. En particulier, si $t \in [t_0, T]$, $R^{-1}B^*p \cdot B^*p(t) = 0$. Comme R est symétrique, définie positive, on en déduit que $u(t) = R^{-1}B^*p(t) = 0$. La commande est donc nulle pour tout $t \in [t_0, T]$, et $y(t) = \exp(A(t-t_0))y(t_0) = 0$ pour $t \in [t_0, T]$. De même, on obtient la nullité de p sur $[t_0, T]$. Ce résultat n'est pas étonnant. Il signifie que si on cherche à annuler y alors que y est déjà nul, la commande optimale consiste simplement à ne rien faire. Reste à prouver la nullité de y , u et p sur l'intervalle $[0, t_0]$. Il suffit de constater que le couple (y, p) est solution d'un système différentielle linéaire de condition initiale $(y, p)(t_0) = (0, 0)$ (la flèche du temps est inversée). Ce système admet une solution unique : la solution nulle.

Ce résultat stipule que, si l'état initial n'est pas l'état cible, il n'est jamais rentable d'atteindre exactement ce dernier. Le coût nécessaire pour s'approcher de l'état cible devient plus important que le gain réalisé.

Exercice 10.4.2 Obtenir l'équivalent de la Proposition 10.4.4 et du Théorème 10.4.6 pour le système parabolique

$$\begin{cases} \frac{\partial y}{\partial t} - \Delta y = v + f & \text{dans }]0, T[\times \Omega \\ y = 0 & \text{sur }]0, T[\times \partial\Omega \\ y(0) = y_0 & \text{dans } \Omega \end{cases}$$

où $y_0 \in L^2(\Omega)$, $f \in L^2(]0, T[\times \Omega)$, $v \in L^2(]0, T[\times \Omega)$ est la commande, et on minimise

$$\inf_{v \in L^2(]0, T[\times \Omega)} J(v) = \int_0^T \int_{\Omega} v^2 dt dx + \int_0^T \int_{\Omega} |y - z|^2 dt dx + \int_{\Omega} |y(T) - z_T|^2 dx,$$

où $z \in L^2(]0, T[\times \Omega)$ et $z_T \in L^2(\Omega)$.

Correction. L'application qui à v associe y est linéaire continue de $L^2(]0, T[\times \Omega)$ dans $C^0([0, T]; L^2(\Omega))$. On en déduit que J est continue. De plus, J est fortement convexe et admet donc un unique minimiseur. Combinaison de fonctions différentiables, J est elle même différentiable (l'application qui à v associe y est dérivable car affine continue!) et

$$\begin{aligned} \langle J'(v), w \rangle = \\ 2 \left(\int_0^T \int_{\Omega} vw dx dt + \int_0^T \int_{\Omega} (y - z)y_w dx dt + \int_{\Omega} (y(T) - z_T)y_w(T) dx \right) \end{aligned} \quad (10.14)$$

où y_w est solution du problème parabolique

$$\begin{cases} \frac{\partial y_w}{\partial t} - \Delta y_w = w & \text{dans }]0, T[\times \Omega \\ y_w = 0 & \text{sur }]0, T[\times \partial\Omega \\ y_w(0) = 0 \end{cases}$$

La condition d'optimalité nécessaire et suffisante est $J'(y) = 0$. Comme dans le cas présenté dans le cours, la formule précédente permettant de calculer la dérivée de J est inexploitable : elle nécessite pour chaque fonction test w la résolution d'un système parabolique. On peut obtenir une expression explicite de J' en fonction d'un état adjoint p solution du système

$$\begin{cases} -\frac{\partial p}{\partial t} - \Delta p = y - z & \text{dans }]0, T[\times \Omega \\ p = 0 & \text{sur }]0, T[\times \partial\Omega \\ p(T) = y(T) - z_T. \end{cases}$$

On vérifie sans mal que

$$\langle J'(v), w \rangle = \int_0^T \int_{\Omega} (v + p)w dx.$$

Exercice 10.4.3 Généraliser l'exercice précédent à l'équation des ondes.

Correction. Il s'agit d'étudier le problème hyperbolique

$$\begin{cases} \frac{\partial^2 y}{\partial t^2} - \Delta y = v + f & \text{dans }]0, T[\times \Omega \\ y = 0 & \text{sur }]0, T[\times \partial\Omega \\ y(0) = y_0 & \text{dans } \Omega \\ \frac{\partial y}{\partial t}(0) = y_0 & \text{dans } \Omega \end{cases}$$

où $y_0 \in H_0^1(\Omega)$, $y_1 \in L^2(\Omega)$, $f \in L^2(]0, T[\times \Omega)$ et $v \in L^2(]0, T[\times \Omega)$ est la commande. On minimise

$$\inf_{v \in L^2(]0, T[\times \Omega)} J(v) = \int_0^T \int_{\Omega} v^2 dt dx + \int_0^T \int_{\Omega} |y - z|^2 dt dx + \int_{\Omega} |y(T) - z_T|^2 dx,$$

où $z \in L^2(]0, T[\times \Omega)$ et $z_T \in L^2(\Omega)$. A nouveau, J est dérivable, fortement convexe et admet donc un unique minimiseur. De plus, la dérivée de J possède la même expression (10.14) que précédemment. Cependant, y_w est dans ce cas solution du problème hyperbolique

$$\begin{cases} \frac{\partial^2 y_w}{\partial t^2} - \Delta y_w = w & \text{dans }]0, T[\times \Omega \\ y_w = 0 & \text{sur }]0, T[\times \partial\Omega \\ y_w(0) = 0 \\ \frac{\partial y_w}{\partial t} = 0 \end{cases}$$

A nouveau, on peut introduire un état adjoint afin de déterminer explicitement J' . L'équation vérifiée par l'état adjoint est

$$\begin{cases} \frac{\partial^2 p}{\partial t^2} - \Delta p = w & \text{dans }]0, T[\times \Omega \\ p = 0 & \text{sur }]0, T[\times \partial\Omega \\ p(0) = 0 \\ \frac{\partial p}{\partial t} = z_T - y(T) \end{cases}$$

et

$$\langle J'(v), w \rangle = \int_0^T \int_{\Omega} (v + p) w dx dt.$$

Notons que pour trouver l'état adjoint, on peut introduire un Lagrangien comme dans le cas de la dimension finie.

Exercice 10.5.1 Pour $V = \mathbb{R}^2$ et $J(x, y) = ax^2 + by^2$ avec $a, b > 0$, montrer que l'algorithme de gradient à pas optimal converge en une seule itération si $a = b$ ou si $x^0 y^0 = 0$, et que la convergence est géométrique dans les autres cas. Étudier aussi la convergence de l'algorithme de gradient à pas fixe : pour quelles valeurs du paramètre μ la convergence se produit-elle, pour quelle valeur est-elle la plus rapide ?

Correction. L'algorithme de gradient à pas optimal converge en une unique itération si et seulement si le minimiseur de J (en l'occurrence 0) appartient à la droite paramétrée par la fonction $t \mapsto tJ'(x, y) + (x, y)$, c'est à dire si et seulement si (x, y) et $J'(x, y)$ sont colinéaires. Comme $J'(x, y) = 2(ax, by)$, l'algorithme converge en une itération si et seulement le produit vectoriel entre (x_0, y_0) et (ax_0, by_0) est nul, c'est à dire si $a = b$ ou $x_0 y_0 = 0$. Dans le cas contraire, considérons (x_n, y_n) la solution obtenue au bout de n itérations du gradient à pas optimal. Comme le pas est choisi de manière optimal, le gradient de J en (x_{n+1}, y_{n+1}) est orthogonal au gradient de J en (x_n, y_n) . Ainsi, le gradient de J en (x_{n+2}, y_{n+2}) est colinéaire au gradient de J en (x_n, y_n) . On en déduit que (x_n, y_n) et (x_{n+2}, y_{n+2}) sont colinéaires. Il existe donc $\alpha(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$ tel que

$$(x_{n+2}, y_{n+2}) = \alpha(x_n, y_n)(x_n, y_n).$$

Enfin, pour tout réel r , on a $\alpha(rx, ry) = \alpha(x, y)$. On a donc $\alpha(x_{n+2}, y_{n+2}) = \alpha(x_n, y_n)$ et $\alpha(x_{2p}, y_{2p}) = \alpha(x_0, y_0)$. Ainsi,

$$J(x_{2p}, y_{2p}) = \alpha(x_0, y_0)^p J(x_0, y_0).$$

La convergence est donc géométrique.

Considérons l'algorithme de gradient à pas fixe. D'après l'expression de la dérivée de J ,

$$x_{n+1} = (1 - 2\mu a)x_n \text{ et } y_{n+1} = (1 - 2\mu b)y_n.$$

Par récurrence évidente, on en déduit une formule explicite de (x_n, y_n) :

$$x_n = (1 - 2\mu a)^n x_0 \text{ et } y_n = (1 - 2\mu b)^n y_0.$$

La convergence a lieu lorsque $\max(|1 - 2\mu a|, |1 - 2\mu b|) < 1$, c'est à dire

$$\mu < \min(a^{-1}, b^{-1}).$$

Le pas optimal est obtenu en minimisant $\beta = \max(|1 - 2\mu a|, |1 - 2\mu b|)$ par rapport à μ . Par une étude graphique rapide, on obtient que le pas optimal est

$$\mu_{opt} = (a + b)^{-1}.$$

La raison de la suite géométrique est alors

$$\beta = |a - b|/(a + b).$$

Pour terminer, notons qu'on peut également calculer explicitement la raison β' de la suite dans le cas de l'algorithme à pas optimal. A titre indicatif, on obtient

$$\beta' = |a - b||x_0||y_0|\sqrt{ab}((ax_0^2 + by_0^2)(a^3x_0^2 + b^3y_0^2))^{-1/2}.$$

L'algorithme du gradient à pas optimal converge au moins aussi rapidement que l'algorithme à pas fixe optimal. La convergence des deux algorithmes est identique si $a = b$ ou $a|x_0| = b|y_0|$.

Exercice 10.5.2 Soit $V = \mathbb{R}^N$ et $K = \{x \in \mathbb{R}^N \text{ tel que } \sum_{i=1}^N x_i = 1\}$. Expliciter l'opérateur de projection orthogonale P_K et interpréter dans ce cas la formule

$$u_{n+1} = P_K(u_n - \mu J'(u_n)) \quad (10.15)$$

définissant l'algorithme de gradient projeté à pas fixe en terme de multiplicateur de Lagrange.

Correction. Soit $n = \sqrt{N}^{-1} \sum_{i=1}^N e_i$ le vecteur normal à K . L'opérateur de projection sur K est

$$P_K(u) = u + (1 - u \cdot n)n.$$

La formule (10.106) implique que si u minimise J sur K

$$u = P_K(u - \mu J'(u)) = u - \mu J'(u) + (1 - u \cdot n + \mu J'(u) \cdot n)n.$$

Comme $u \cdot n = 1$, on en déduit que

$$J'(u) + \lambda n = 0.$$

où $\lambda = -J'(u) \cdot n$. On retrouve la condition d'optimalité du Théorème 10.2.8, vérifiée par les minimiseurs de J sur K , où λ est le multiplicateur de Lagrange associé à la contrainte $u \in K$.

Exercice 10.5.3 Appliquer l'algorithme d'Uzawa au problème

$$\min_{v \in \mathbb{R}^N, F(v)=Bv-c \leq 0} \left\{ J(v) = \frac{1}{2}Av \cdot v - b \cdot v \right\}, \quad (10.16)$$

où A est une matrice $N \times N$ symétrique définie positive, $b \in \mathbb{R}^N$, B une matrice $M \times N$ et $c \in \mathbb{R}^M$. Si la matrice B est de rang M , ce qui assure l'unicité de p d'après la Remarque 10.3.12, montrer que la suite p^n converge vers p .

Correction. Le Lagrangien associé à ce problème est

$$\mathcal{L}(v, q) = \frac{1}{2}Av \cdot v - b \cdot v + q \cdot (Bv - c)$$

avec $q \in \mathbb{R}_+^M$. Soit p^n la suite de multiplicateurs obtenus par l'algorithme d'Uzawa et u^n la suite d'éléments de \mathbb{R}^N définie par

$$\mathcal{L}(u^n, p^n) = \min_v \mathcal{L}(v, p^n). \quad (10.17)$$

On rappelle que p^{n+1} est déterminé à l'aide de p^n par

$$p^{n+1} = P_{\mathbb{R}_+^M} (p^n + \mu F(u^n)), \quad (10.18)$$

où μ est le pas de l'algorithme, choisit suffisamment petit. La matrice A étant symétrique définie positive, le problème (10.17) admet comme unique solution

$$u^n = A^{-1}(b - B^*p).$$

En explicitant la définition (10.18) de p^{n+1} en fonction de p^n , on obtient

$$p^{n+1} = P_{\mathbb{R}_+^M} ((\text{Id} - \mu BA^{-1}B^*)p^n + \mu(BA^{-1}b - c)).$$

Afin de prouver la convergence de la suite p^n , il suffit de montrer que l'application qui à p^{n+1} associe p^n est strictement contractante. Comme la projection $P_{\mathbb{R}_+^M}$ est contractante, il suffit de prouver que l'application

$$q \mapsto (\text{Id} - \mu BA^{-1}B^*)q + \mu(BA^{-1}b - c)$$

est contractante. Comme B est de rang M , la matrice $BA^{-1}B^*$ est définie positive. Pour μ suffisamment petit, la matrice $\text{Id} - \mu BA^{-1}B^*$ est symétrique, définie positive de valeurs propres strictement plus petites que l'identité. L'application précédente

est donc strictement contractante et l'algorithme convergent. On note p sa limite. La suite u^n est également convergente et sa limite u est telle que

$$Au - b + B^*p = 0. \quad (10.19)$$

Enfin, comme $p = P_{\mathbb{R}_+^M}(p + \mu F(u))$, pour tout $q \in \mathbb{R}_+^M$, on a

$$(p - (p + \mu F(u))) \cdot (q - p) \geq 0,$$

c'est à dire $F(u) \cdot p \geq F(u) \cdot q$. On en déduit que

$$F(u) \leq 0 \quad (10.20)$$

et que $F(u) \cdot p \geq 0$. Or comme $F(u) \leq 0$ et $p \geq 0$, on a également $F(u) \cdot p \leq 0$. Ainsi,

$$F(u) \cdot p = 0. \quad (10.21)$$

De (10.19), (10.20) et (10.21), on conclut que u est solution du problème de minimisation étudié.

Exercice 10.5.4 En plus des hypothèses de la Proposition **10.5.10**, on suppose que les fonctions J et F_1, \dots, F_M sont continûment différentiables. On note de nouveau $I(u)$ l'ensemble des contraintes actives en u , et on suppose que les contraintes sont qualifiées en u au sens de la Définition **10.2.13**. Enfin, on suppose que les vecteurs $(F'_i(u))_{i \in I(u)}$ sont linéairement indépendants, ce qui assure l'unicité des multiplicateurs de Lagrange $\lambda_1, \dots, \lambda_M$ tels que $J'(u) + \sum_{i=1}^M \lambda_i F'_i(u) = 0$, avec $\lambda_i = 0$ si $i \notin I(u)$. Montrer alors que, pour tout indice $i \in \{1, \dots, M\}$

$$\lim_{\varepsilon \rightarrow 0} \left[\frac{2}{\varepsilon} \max(F_i(u_\varepsilon), 0) \right] = \lambda_i.$$

Correction. Pour tout $i \notin I(u)$, on a $F_i(u) < 0$. Ainsi, pour ε assez petit, on a $F_i(u_\varepsilon) < 0$ et $\max(F_i(u_\varepsilon), 0) = 0$. En particulier, pour tout $i \notin I(u)$, on a bien

$$\lim_{\varepsilon \rightarrow 0} \left[\frac{2}{\varepsilon} \max(F_i(u_\varepsilon), 0) \right] = 0 = \lambda_i.$$

On pose

$$J_\varepsilon(v) = J(v) + \varepsilon^{-1} \sum_{i=1}^M [\max(F_i(v), 0)]^2.$$

Les fonction F_i étant supposées continûment dérivables, J_ε est dérivable et

$$J'_\varepsilon(v) = J'(v) + 2\varepsilon^{-1} \sum_{i=1}^M \max(F_i(v), 0) F'_i(v).$$

Comme u_ε minimise J_ε , on a $J'_\varepsilon(u_\varepsilon) = 0$ et

$$J'(u_\varepsilon) = -2\varepsilon^{-1} \sum_{i=1}^M \max(F_i(u_\varepsilon), 0) F'_i(u_\varepsilon). \quad (10.22)$$

De plus u_ϵ converge vers u pour lequel

$$J'(u) = - \sum_{i \in I(u)} \lambda_i F'_i(u). \quad (10.23)$$

Comme les applications linéaires $(F'_i(u))_{i \in I(u)}$ sont indépendantes, il existe une famille $(a_i)_{i \in I(u)}$ d'éléments de \mathbb{R}^N telle que

$$\langle F'_i(u), a_j \rangle = \delta_i^j$$

pour tout i et $j \in I(u)$. Comme $F'_i(u_\epsilon)$ converge vers $F'_i(u)$, pour ϵ assez petit, la famille $(F'_i(u_\epsilon))_{i \in I(u)}$ est indépendante et il existe une famille $(a_i^\epsilon)_{i \in I(u)} \in \text{Vect}((a_i)_{i \in I(u)})$ telle que

$$\langle F'_i(u_\epsilon), a_j^\epsilon \rangle = \delta_i^j$$

pour tout i et $j \in I(u)$. De plus, pour tout $i \in I(u)$, a_i^ϵ converge vers a_i . Enfin, pour tout $i \in I(u)$,

$$-2\epsilon^{-1} \max(F_i(u), 0) = \left\langle -2\epsilon^{-1} \sum_{j \in I(u)} \max(F_j(u_\epsilon), 0) F'_j(u_\epsilon), a_i^\epsilon \right\rangle.$$

Comme $\epsilon^{-1} \max(F_i(u), 0)$ converge vers zéro pour tout $i \notin I(u)$,

$$\begin{aligned} \lim_{\epsilon} -2\epsilon^{-1} \max(F_i(u), 0) &= \lim_{\epsilon} -2\epsilon^{-1} \sum_{j=1}^M \max(F_j(u_\epsilon), 0) \langle F'_j(u_\epsilon), a_i^\epsilon \rangle \\ &= \lim_{\epsilon} \langle J'(u_\epsilon), a_i^\epsilon \rangle \\ &= \langle J'(u), a_i \rangle = \lambda_i. \end{aligned}$$

ANALYSE NUMÉRIQUE MATRICIELLE

Exercice 13.1.1 Montrer que

1. $\|A\|_2 = \|A^*\|_2 = \text{maximum des valeurs singulières de } A$,
2. $\|A\|_1 = \max_{1 \leq j \leq n} \left(\sum_{i=1}^n |a_{ij}| \right)$,
3. $\|A\|_\infty = \max_{1 \leq i \leq n} \left(\sum_{j=1}^n |a_{ij}| \right)$.

Correction.

1. Tout d'abord, on rappelle que les valeurs singulières de A sont les racines carrées des valeurs propres de la matrice symétrique A^*A . Par définition, on a

$$\|A\|_2 = \left(\max_{x \in \mathbb{C}^n, x \neq 0} \frac{(Ax)^* \cdot Ax}{x^* \cdot x} \right)^{1/2}.$$

Ainsi,

$$\|A\|_2 = \left(\max_{x \in \mathbb{C}^n, x \neq 0} \frac{(A^*Ax)^* \cdot x}{x^* \cdot x} \right)^{1/2}$$

est bien le maximum des valeurs singulières de A (la matrice A^*A est symétrique, positive et diagonalisable).

On a pour tout $x \in \mathbb{C}^n$,

$$\|x\|_2 = \sup_{y \in \mathbb{C}^n, \|y\|_2 \leq 1} |x \cdot y|.$$

Ainsi,

$$\|Ax\|_2 = \sup_{y \in \mathbb{C}^n, \|y\|_2 \leq 1} |Ax \cdot y| = \sup_{y \in \mathbb{C}^n, \|y\|_2 \leq 1} |x \cdot A^*y| \leq \|x\|_2 \|A^*\|_2.$$

On en déduit que $\|A\|_2 \leq \|A^*\|_2$ et finalement $\|A\|_2 = \|A^*\|_2$.

2.

$$\|A\|_1 = \max_{x \in \mathbb{C}, x \neq 0} \frac{\|Ax\|_1}{\|x\|_1} = \max_{x \in \mathbb{C}, x \neq 0} \frac{\sum_i \left| \sum_k a_{ik} x_k \right|}{\sum_k |x_k|}.$$

Pour tout indice j , en choisissant $x_k = \delta_j^k$, on obtient

$$\|A\|_1 \geq \sum_i |a_{ij}|.$$

De plus,

$$\begin{aligned} \|A\|_1 &= \max_{x \in \mathbb{C}, x \neq 0} \frac{\sum_i \left| \sum_j a_{ij} x_j \right|}{\sum_j |x_j|} \leq \max_{x \in \mathbb{C}, x \neq 0} \frac{\sum_{i,j} |a_{ij}| |x_j|}{\sum_j |x_j|} \\ &= \max_{x \in \mathbb{C}, x \neq 0} \frac{\sum_j \left(\sum_i |a_{ij}| \right) |x_j|}{\sum_j |x_j|} \leq \max_j \sum_i |a_{ij}|. \end{aligned}$$

3. On a

$$\|A\|_\infty = \max_{x \in \mathbb{C}, x \neq 0} \left(\frac{\max_k \left| \sum_j a_{k,j} x_j \right|}{\max_k |x_k|} \right).$$

Soit $i \in \{1, \dots, n\}$ et $x \in \mathbb{C}^n$ telle que pour tout indice j , x_j soit égal au signe de $a_{i,j}$. On déduit de l'expression précédente que

$$\|A\|_\infty \geq \max_i \left(\sum_j |a_{i,j}| \right).$$

Réciproquement,

$$\|A\|_\infty \leq \max_{x \in \mathbb{C}, x \neq 0} \left(\frac{\max_i \sum_j |a_{i,j}| |x_j|}{\max_i |x_i|} \right) \leq \max_i \sum_j |a_{i,j}|.$$

Exercice 13.1.2 Soit une matrice $A \in \mathcal{M}_n(\mathbb{C})$. Vérifier que

1. $\text{cond}(A) = \text{cond}(A^{-1}) \geq 1$, $\text{cond}(\alpha A) = \text{cond}(A) \forall \alpha \neq 0$,
2. pour une matrice quelconque, $\text{cond}_2(A) = \frac{\mu_n(A)}{\mu_1(A)}$, où $\mu_1(A), \mu_n(A)$ sont respectivement la plus petite et la plus grande valeur singulière de A ,
3. pour une matrice normale, $\text{cond}_2(A) = \frac{|\lambda_n(A)|}{|\lambda_1(A)|}$, où $|\lambda_1(A)|, |\lambda_n(A)|$ sont respectivement la plus petite et la plus grande valeur propre en module de A ,
4. pour toute matrice unitaire U , $\text{cond}_2(U) = 1$,
5. pour toute matrice unitaire U , $\text{cond}_2(AU) = \text{cond}_2(UA) = \text{cond}_2(A)$.

Correction.

1.

$$\text{cond}(A) = \|A\| \|A^{-1}\| = \|A^{-1}\| \|A\| = \text{cond}(A^{-1}).$$

De plus d'après les propriétés élémentaires des normes subordonnées,

$$\text{cond}(A) = \|A\| \|A^{-1}\| \geq \|AA^{-1}\| = \|\text{Id}\| = 1.$$

Enfin, $\text{cond}(\alpha A) = \|\alpha A\| \|(\alpha A)^{-1}\| = |\alpha| |\alpha|^{-1} \|A\| \|A^{-1}\| = \text{cond}(A)$.

2. D'après l'Exercice 13.1.1, $\|A\|_2$ est la plus grande valeur singulière de A . Comme les valeurs singulières de A^{-1} sont les inverses des valeurs singulières de A , on en déduit que $\text{cond}_2(A) = \frac{\mu_n(A)}{\mu_1(A)}$.
3. Pour une matrice normale (donc diagonalisable), les valeurs singulières sont les modules des valeurs propres. Ainsi, d'après le point précédent, pour toute matrice normale on a encore $\text{cond}_2(A) = \frac{|\lambda_n(A)|}{|\lambda_1(A)|}$.
4. Pour une matrice unitaire, $\|U\|_2 = \|U^{-1}\|_2 = 1$. Ainsi, $\text{cond}_2(U) = 1$.
5. Si U est une matrice unitaire, on a

$$(AU)(AU)^* = AUU^*A^* = AA^* \text{ et } (UA)^*(UA) = A^*A.$$

Ainsi, la plus grande valeur singulière de A est égale à la plus grande valeur singulière de UA tandis que la plus grande valeur singulière de A^* est égale à la plus grande valeur singulière de $(AU)^*$. On a donc

$$\|AU\|_2 = \|(AU)^*\|_2 = \|A^*\|_2 = \|A\|_2 = \|UA\|_2.$$

De plus, comme $(AU)^{-1}$ et $(UA)^{-1}$ sont le produit (à gauche et à droite) de A^{-1} avec la matrice unitaire U^* , on a également

$$\|(AU)^{-1}\|_2 = \|A^{-1}\|_2 = \|A\|_2 = \|(UA)^{-1}\|_2.$$

On en déduit que $\text{cond}_2(AU) = \text{cond}_2(UA) = \text{cond}_2(A)$.

Exercice 13.1.3 Montrer que le conditionnement de la matrice de rigidité \mathcal{K}_h , donnée par (6.12) pour la méthode des éléments finis P_1 appliquée au Laplacien, est

$$\text{cond}_2(\mathcal{K}_h) \approx \frac{4}{\pi^2 h^2}. \quad (13.1)$$

On montrera que les valeurs propres de \mathcal{K}_h sont

$$\lambda_k = 4h^{-1} \sin^2 \left(\frac{k\pi}{2(n+1)} \right) \quad 1 \leq k \leq n,$$

pour des vecteurs propres u^k donnés par leurs composantes

$$u_j^k = \sin \left(\frac{jk\pi}{n+1} \right) \quad 1 \leq j, k \leq n.$$

Correction. Dans un premier temps, on vérifie que les vecteurs u^k sont les vecteurs propres de \mathcal{K}_h . On a

$$\begin{aligned} (\mathcal{K}_h u^k)_j &= h^{-1} (-u_{j-1}^k + 2u_j^k - u_{j+1}^k) \\ &= h^{-1} \left(\sin \left(\frac{(j-1)k\pi}{n+1} \right) + 2 \sin \left(\frac{jk\pi}{n+1} \right) - \sin \left(\frac{(j+1)k\pi}{n+1} \right) \right) \\ &= (2ih)^{-1} \left(-e^{\frac{i(j-1)k\pi}{n+1}} + 2e^{\frac{ijk\pi}{n+1}} - e^{\frac{i(j+1)k\pi}{n+1}} - e^{-\frac{i(j-1)k\pi}{n+1}} + 2e^{-\frac{ijk\pi}{n+1}} - e^{-\frac{i(j+1)k\pi}{n+1}} \right) \\ &= (2ih)^{-1} \left(e^{\frac{ijk\pi}{n+1}} - e^{-\frac{ijk\pi}{n+1}} \right) \left(-e^{\frac{ijk\pi}{n+1}} + 2 - e^{-\frac{ijk\pi}{n+1}} \right) \\ &= 4h^{-1} \sin \left(\frac{jk\pi}{n+1} \right) \sin^2 \left(\frac{k\pi}{2(n+1)} \right) \\ &= 4h^{-1} \sin^2 \left(\frac{k\pi}{2(n+1)} \right) u_j^k. \end{aligned}$$

La matrice \mathcal{K}_h étant normale,

$$\text{cond}_2(\mathcal{K}_h) = |\lambda_n(\mathcal{K}_h)| / |\lambda_1(\mathcal{K}_h)|.$$

La plus grande valeur propre de \mathcal{K}_h est $4h^{-1} \sin^2(n\pi/2(n+1)) \approx 4h^{-1}$ et la plus petite $4h^{-1} \sin^2(\pi/2(n+1)) \approx 4h^{-1} (\pi/2(n+1))^2 = h\pi^2$. La matrice \mathcal{K}_h étant normale, le conditionnement de \mathcal{K}_h est

$$\text{cond}_2(\mathcal{K}_h) \approx \frac{4h^{-1}}{h\pi^2} = \frac{4}{\pi^2 h^2}.$$

Exercice 13.1.4 Montrer que les factorisations LU et de Cholesky conservent la structure bande des matrices.

Correction. Considérons le cas de la factorisation LU. Soit A une matrice bande de demi largeur de bande p . On raisonne par récurrence afin de prouver que les matrices L et U sont également des matrices bande de demi largeur de bande p . Les composantes des matrices L et U sont déterminées en fonction des composantes de A colonnes par colonnes. Supposons que les $j-1$ premières colonnes de L et U soit de demi largeur de bande p . Les composantes de la j ème colonne de U sont définies pour $1 \leq i \leq j$ par

$$u_{i,j} = a_{i,j} - \sum_{k=1}^{i-1} l_{i,k} u_{k,j}.$$

La matrice A étant une matrice bande de demi-largeur p , on a $a_{i,j} = 0$ pour tout i tels que $j > i + p$. Par une (nouvelle) récurrence (sur i cette fois), on en déduit que $u_{i,j} = 0$ pour tout i tel que $j > i + p$. Ainsi, la j ème colonne de U est celle d'une matrice bande creuse de demi largeur de bande p . La j ème colonne de L est déterminée pour $j+1 \leq i \leq n$ par

$$l_{i,j} = \frac{a_{i,j} - \sum_{k=1}^{j-1} l_{i,k} u_{k,j}}{u_{j,j}}.$$

D'après l'hypothèse de récurrence sur la structure bande des j premières colonnes de L , on a $l_{i,k} = 0$ dès que $i - k > p$. Ainsi, le terme de somme apparaissant dans l'expression de $l_{i,j}$ est nul dès que $i - (j-1) > p$ et en particulier dès que $i - j > p$. Ainsi, $l_{i,j} = 0$ dès que $i - j > p$ et les j premières colonnes de L ont une structure de matrice bande de demi largeur p . Ceci achève la récurrence et prouve que la structure bande est conservée par la factorisation LU. La matrice B issue de la factorisation de Cholesky n'est autre que le produit de L par une matrice diagonale, si L est une matrice bande, B l'est également. La factorisation de Cholesky conserve donc également la structure bande.

Exercice 13.1.5 Montrer que, pour une matrice bande d'ordre n et de demie largeur de bande p , le compte d'opérations de la factorisation LU est $\mathcal{O}(np^2/3)$ et celui de la factorisation de Cholesky est $\mathcal{O}(np^2/6)$.

Correction. Voir les remarques du répertoire... N'y aurait-il pas une erreur d'énoncé?

Exercice 13.1.6 Soit A une matrice hermitienne définie positive. Montrer que pour tout $\omega \in]0, 2[$, la méthode de relaxation converge.

Correction. La matrice A étant hermitienne définie positive, sa diagonale D est constituée de réels strictement positifs. La matrice $M = D/\omega - E$ est donc inversible et la méthode de relaxation correctement définie. De plus, d'après les Lemmes 13.1.26 et 13.1.27, la méthode de relaxation est convergente dès que $M^* + N$ est définie positive. Or

$$M^* + N = \frac{2 - \omega}{\omega} D,$$

qui est définie positive pour tout $\omega \in]0, 2[$.

Exercice 13.1.7 Montrer que, pour la méthode de relaxation, on a toujours

$$\rho(M^{-1}N) \geq |1 - \omega|, \quad \forall \omega \neq 0,$$

et donc qu'elle ne peut converger que si $0 < \omega < 2$.

Correction. Le vecteur $e_1 = (1, 0, \dots, 0)$ est un vecteur propre de $M^{-1}N$ de valeur propre $1 - \omega$. Ainsi, $\rho(M^{-1}N) \geq |1 - \omega|$ et la méthode de relaxation ne peut converger que pour $\omega \in]0, 2[$.

Exercice 13.1.8 Soit A une matrice symétrique définie positive. Soit $(x_k)_{0 \leq k \leq n}$ la suite de solutions approchées obtenues par la méthode du gradient conjugué. On pose $r_k = b - Ax_k$ et $d_k = x_{k+1} - x_k$. Montrer que

(i) l'espace de Krylov K_k est aussi égal à

$$K_k = [r_0, \dots, r_k] = [d_0, \dots, d_k],$$

(ii) la suite $(r_k)_{0 \leq k \leq n-1}$ est orthogonale

$$r_k \cdot r_l = 0 \text{ pour tout } 0 \leq l < k \leq n-1,$$

(iii) la suite $(d_k)_{0 \leq k \leq n-1}$ est conjuguée par rapport à A

$$Ad_k \cdot d_l = 0 \text{ pour tout } 0 \leq l < k \leq n-1.$$

Correction.

(i) On rappelle que r_k est définie par $r_k = r_0 - Ay_k \perp K_{k-1}$, où $y_k \in K_{k-1}$. On a $Ay_k \in K_k$. Ainsi, $r_k \in K_k$ et (r_0, \dots, r_k) est une famille de K_k . Reste à montrer que cette famille est génératrice. On raisonne par récurrence. Supposons que $K_{k-1} = [r_0, \dots, r_{k-1}]$. Si r_k n'appartient pas à K_{k-1} , on a

$$\dim([r_0, \dots, r_k]) = \dim([r_0, \dots, r_{k-1}]) + 1 = \dim(K_{k-1}) + 1 \geq \dim(K_k).$$

L'espace $[r_0, \dots, r_k]$ étant inclus dans K_k et de même dimension, ils sont égaux. Reste à considérer le cas où r_k appartient à K_{k-1} . Comme r_k est orthogonal à K_{k-1} , on a dans ce cas $r_k = 0$ et $r_0 = Ay_k$. Or $y_k \in [r_0, \dots, A^{k-1}r_0]$.

Ainsi, $r_0 \in [Ar_0, \dots, A^k r_0]$. La famille $(r_0, \dots, A^k r_0)$ n'est pas libre et K_k est un espace de dimension strictement inférieure à k . Dans ce cas, on a

$$K_k = K_{k-1} = [r_0, \dots, r_{k-1}] = [r_0, \dots, r_k].$$

Comme y_k appartient à K_{k-1} , le vecteur $d_k = y_{k+1} - y_k$ appartient à K_k . Ainsi, $[d_0, \dots, d_k]$ est un sous espace de K_k . Supposons que pour un k donné, on ait $K_{k-1} = [d_0, \dots, d_{k-1}]$. Si y_{k+1} n'appartient pas à K_{k-1} , d_k n'appartient pas à K_{k-1} et $K_k = [d_0, \dots, d_k]$. Dans le cas contraire (y_{k+1} appartient à K_{k-1}), on a $y_k = y_{k+1}$ et $r_{k+1} = r_k$. En particulier, r_{k+1} appartient à K_k et est orthogonal à K_k . On a donc $r_{k+1} = 0$. On en déduit que r_k est nul et que $K_k = K_{k-1}$. On a donc a nouveau $K_k = [d_0, \dots, d_k]$.

(ii) Le vecteur r_k est orthogonal à $K_{k-1} = [r_0, \dots, r_{k-1}]$.

(iii) On a $\langle A^{-1}r_0 - y_k, y \rangle_A = 0$ pour tout $y \in K_{k-1}$. Ainsi, $\langle y_{k+1} - y_k, y \rangle_A = 0$ pour tout $y \in K_{k-1}$. En d'autres termes,

$$\langle d_k, y \rangle_A = 0, \quad \forall y \in [d_0, \dots, d_{k-1}].$$

Exercice 13.1.9 Si on considère la méthode du gradient conjugué comme une méthode directe, montrer que dans le cas le plus défavorable, $k_0 = n - 1$, le nombre d'opérations (multiplications seulement) pour résoudre un système linéaire est $N_{op} = n^3(1 + o(1))$.

Correction. A chaque itérations, on effectue de l'ordre de n^2 opérations, l'essentiel du temps étant consacré au calcul de Ap_k . Dans le cas le plus défavorable, l'algorithme converge au bout de n itérations. Dans ce cas, le nombre d'itérations est de l'ordre de n^3 .

Exercice 13.1.10 On note avec un tilde $\tilde{\cdot}$ toutes les quantités associées à l'algorithme du gradient conjugué appliqué au système linéaire (13.12). Soit $x_k = B^{-*}\tilde{x}_k$, $r_k = B\tilde{r}_k = b - Ax_k$, et $p_k = B^{-*}\tilde{p}_k$. Montrer que l'algorithme du gradient conjugué pour (13.12) peut aussi s'écrire sous la forme

$$\begin{aligned} \text{initialisation} & \quad \left\{ \begin{array}{l} \text{choix initial } x_0 \\ r_0 = b - Ax_0 \\ p_0 = z_0 = C^{-1}r_0 \end{array} \right. \\ \text{itérations } k \geq 1 & \quad \left\{ \begin{array}{l} \alpha_{k-1} = \frac{z_{k-1} \cdot r_{k-1}}{Ap_{k-1} \cdot p_{k-1}} \\ x_k = x_{k-1} + \alpha_{k-1}p_{k-1} \\ r_k = r_{k-1} - \alpha_{k-1}Ap_{k-1} \\ z_k = C^{-1}r_k \\ \beta_{k-1} = \frac{z_k \cdot r_k}{z_{k-1} \cdot r_{k-1}} \\ p_k = z_k + \beta_{k-1}p_{k-1} \end{array} \right. \end{aligned}$$

où $C = BB^*$.

Correction. L'algorithme du gradient conjugué associé à (13.12) consiste à calculer itérativement

$$\begin{aligned}\alpha_{k-1} &= \frac{\|\tilde{r}_{k-1}\|^2}{\tilde{A}\tilde{p}_{k-1} \cdot \tilde{p}_{k-1}} \\ \tilde{x}_k &= \tilde{x}_{k-1} + \alpha_{k-1}\tilde{p}_{k-1} \\ \tilde{r}_k &= \tilde{r}_{k-1} - \alpha_{k-1}\tilde{A}\tilde{p}_{k-1} \\ \beta_{k-1} &= \frac{\|\tilde{r}_k\|^2}{\|\tilde{r}_{k-1}\|^2} \\ \tilde{p}_k &= \tilde{r}_k + \beta_{k-1}\tilde{p}_{k-1}.\end{aligned}$$

En utilisant les expressions de x_k , r_k et p_k en fonction de \tilde{x}_k , \tilde{r}_k et \tilde{p}_k , on obtient

$$\begin{aligned}\alpha_{k-1} &= \frac{\|B^{-1}r_{k-1}\|^2}{Ap_{k-1} \cdot p_{k-1}} = \frac{C^{-1}r_{k-1} \cdot r_{k-1}}{Ap_{k-1} \cdot p_{k-1}} \\ x_k &= B^{-*}\tilde{x}_k = x_{k-1} + \alpha_{k-1}p_{k-1} \\ r_k &= B\tilde{r}_k = r_{k-1} - \alpha_{k-1}Ap_{k-1} \\ \beta_{k-1} &= \frac{\|B^{-1}r_k\|^2}{\|B^{-1}r_{k-1}\|^2} = \frac{C^{-1}r_k \cdot r_k}{C^{-1}r_{k-1} \cdot r_{k-1}} \\ p_k &= B^{-*}\tilde{p}_k = C^{-1}r_k + \beta_{k-1}p_{k-1}.\end{aligned}$$

L'algorithme du gradient préconditionné s'écrit donc bien sous la forme annoncée.

Exercice 13.1.11 Soit A la matrice d'ordre n issue de la discrétisation du Laplacien en dimension $N = 1$ avec un pas d'espace constant $h = 1/(n+1)$

$$A = h^{-1} \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}.$$

Montrer que pour la valeur optimale

$$\omega_{opt} = \frac{2}{1 + 2 \sin \frac{\pi}{2n}} \simeq 2\left(1 - \frac{\pi}{n+1}\right)$$

le conditionnement de la matrice $C_\omega^{-1}A$ est majoré par

$$\text{cond}_2(C_\omega^{-1}A) \leq \frac{1}{2} + \frac{1}{2 \sin \frac{\pi}{2(n+1)}},$$

et donc que, pour n grand, on gagne un ordre en n dans la vitesse de convergence.

Correction. On note B_ω la matrice définie par

$$B_\omega = \sqrt{\frac{\omega}{2-\omega}} \left(\frac{D}{\omega} - E \right) D^{-1/2}.$$

On a $C_\omega = B_\omega B_\omega^T$. Ainsi,

$$C_\omega^{-1}A = B_\omega^{-T}B_\omega^{-1}A = B_\omega^{-T}(B_\omega^{-1}AB_\omega^{-T})B_\omega^T = B_\omega^{-T}\tilde{A}_\omega B_\omega^T,$$

où $\tilde{A}_\omega = B_\omega^{-1}AB_\omega^{-T}$. Les matrices $C_\omega^{-1}A$ et \tilde{A}_ω étant semblables, elles ont les mêmes valeurs propres et

$$\text{cond}_2(C_\omega^{-1}A) = \text{cond}_2(\tilde{A}_\omega) = \|\tilde{A}_\omega\|_2 \|\tilde{A}_\omega^{-1}\|_2.$$

Afin de déterminer une majoration du conditionnement, il suffit de majorer $\|\tilde{A}_\omega\|_2$ et $\|\tilde{A}_\omega^{-1}\|_2$. On a

$$\|\tilde{A}_\omega\|_2 = \max_{x \neq 0} \frac{\langle \tilde{A}_\omega x, x \rangle}{\langle x, x \rangle} = \max_{x \neq 0} \frac{\langle B_\omega^{-1}AB_\omega^{-T}x, x \rangle}{\langle x, x \rangle} = \max_{x \neq 0} \frac{\langle AB_\omega^{-T}x, B_\omega^{-T}x \rangle}{\langle x, x \rangle}.$$

En posant $y = B_\omega^{-T}x$, on en déduit que

$$\|\tilde{A}_\omega\|_2 = \max_{y \neq 0} \frac{\langle Ay, y \rangle}{\langle B_\omega^T y, B_\omega^T y \rangle} = \max_{y \neq 0} \frac{\langle Ay, y \rangle}{\langle B_\omega^{-T} B_\omega^T y, y \rangle} = \max_{y \neq 0} \frac{\langle Ay, y \rangle}{\langle C_\omega y, y \rangle}.$$

De même, on a

$$\|\tilde{A}_\omega^{-1}\|_2 = \max_{y \neq 0} \frac{\langle C_\omega y, y \rangle}{\|\tilde{A}_\omega\|_2}.$$

Ainsi,

$$\text{cond}_2(C_\omega^{-1}A) = \max_{x \neq 0} \frac{\langle Ax, x \rangle}{\langle C_\omega x, x \rangle} \left(\min_{x \neq 0} \frac{\langle Ax, x \rangle}{\langle C_\omega x, x \rangle} \right)^{-1}.$$

Il reste à déterminer un encadrement

$$0 < \alpha \leq \frac{\langle Ax, x \rangle}{\langle C_\omega x, x \rangle} \leq \beta.$$

Majoration . On décompose C_ω sous la forme

$$C_\omega = A + \frac{\omega}{2 - \omega} F_\omega D^{-1} F_\omega^T,$$

avec $F_\omega = \frac{\omega - 1}{\omega} D - E$. Pour tout $x \neq 0$, on a

$$\frac{2 - \omega}{\omega} \langle (A_\omega - C)x, x \rangle = -\langle F_\omega D^{-1} F_\omega^T x, x \rangle = -\langle D^{-1} F_\omega^T x, F_\omega^T x \rangle \leq 0,$$

puisque la matrice D^{-1} est définie positive. Il en découle que $\beta = 1$.

Minoration . On écrit cette fois $(2 - \omega)C_\omega = A + aD + \omega G$ avec

$$G = ED^{-1}E^T - \frac{D}{4} \quad \text{et} \quad a = \frac{(2 - \omega)^2}{4\omega}.$$

Pour $x \neq 0$, on calcule le rapport

$$(2 - \omega) \frac{\langle C_\omega x, x \rangle}{\langle Ax, x \rangle} = 1 + a \frac{\langle Dx, x \rangle}{\langle Ax, x \rangle} + \omega \frac{\langle Gx, x \rangle}{\langle Ax, x \rangle}.$$

Puisque $\langle Gx, x \rangle = -\frac{|x_1|^2}{2h}$, on a

$$(2 - \omega) \frac{\langle C_\omega x, x \rangle}{\langle Ax, x \rangle} \leq 1 + a \frac{\langle Dx, x \rangle}{\langle Ax, x \rangle} = 1 + \frac{2a}{h} \frac{\|x\|^2}{\langle Ax, x \rangle} \leq 1 + \frac{2a}{h\lambda_{\min}(A)},$$

où $\lambda_{\min}(A) = 4h^{-1} \sin^2 \frac{\pi}{2(n+1)}$ est la plus petite valeur propre de A . On peut donc prendre

$$\begin{aligned} \alpha &= (2 - \omega) \left(1 + \frac{a}{2 \sin^2 \frac{\pi}{2(n+1)}} \right)^{-1} \\ &= \left(\frac{1}{2 - \omega} + \frac{2 - \omega}{2\omega \sin^2 \frac{\pi}{2(n+1)}} \right)^{-1}. \end{aligned}$$

et

$$\text{cond}_2(C_\omega^{-1}A) \leq \frac{1}{2 - \omega} + \frac{2 - \omega}{2\omega \sin^2 \frac{\pi}{2(n+1)}}.$$

La minimisation du terme de droite par rapport à ω conduit à la valeur optimale

$$\omega_{opt} = \frac{2}{1 + 2 \sin \frac{\pi}{2n}} \simeq 2 \left(1 - \frac{\pi}{n+1} \right)$$

et à la majoration

$$\text{cond}_2(C_\omega^{-1}A) \leq \frac{1}{2} + \frac{1}{2 \sin \frac{\pi}{2(n+1)}},$$

Bibliographie

- [1] ALLAIRE G., *Analyse numérique et optimisation* Éditions de l'École Polytechnique, Palaiseau (2005).
- [2] ALLAIRE G., KABER S. M., *Algèbre linéaire numérique. Cours et exercices*, Éditions Ellipses, Paris (2002).
- [3] CIARLET P.G., *Introduction à l'analyse numérique matricielle et à l'optimisation*, Masson, Paris (1982).
- [4] JOLY P., *Analyse et approximation de modèles de propagation d'ondes*, Cours de majeure SeISM, Ecole Polytechnique (2001).