

# 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

29 août 2012

## 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  $\theta_0$  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.** Dans un premier temps, nous allons vérifier formellement que l'expression de  $\theta(t, x)$  (1.1) proposée est solution de l'équation de convection diffusion (1.2). Dans un deuxième temps, nous justifierons les calculs effectués.

On pose  $G(x, t, y) = \exp\left(-\frac{(x - Vt - y)^2}{4\nu t}\right)$ . On a

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

Quitte à permuter les opérateurs de dérivation et d'intégration, on en déduit que

$$\begin{aligned} \frac{\partial}{\partial x} \int_{-\infty}^{\infty} \theta_0(y) G(x, t, y) dy &= \int_{-\infty}^{\infty} \theta_0(y) \frac{\partial G}{\partial x} dy \\ &= - \int_{-\infty}^{\infty} \theta_0(y) \frac{x - Vt - y}{2\nu t} G(x, t, y) dy. \end{aligned} \quad (1.3)$$

De manière similaire,

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

et

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

On obtient ainsi l'expression des dérivées partielles de  $\theta(t, x)$  pour tout  $t > 0$ , à savoir

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

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

$$\lim_{t \rightarrow 0} \frac{1}{\sqrt{4\pi\nu t}} \int_{-\infty}^{\infty} \theta_0(y) \exp\left(-\frac{(x - Vt - y)^2}{4\nu t}\right) dy = \theta_0(x). \quad (1.4)$$

Rappelons que,

$$\int_{-\infty}^{\infty} \exp(-x^2) dx = \sqrt{\pi}. \quad (1.5)$$

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 (1.5),  $\int \rho(x, t, y) dy = 1$  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  $\delta$  et  $\varepsilon$  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 (1.4) découle alors du fait que  $\theta_0$  est continue, uniformément bornée.

Reste à prouver que les commutations des opérateurs d'intégration et de dérivation effectuées lors du calcul des dérivées partielles de  $\theta(t, x)$  sont licites. Pour tout  $x$  de  $\mathbb{R}$  et tout  $t > 0$ , il existe des constantes  $C_1(x, t)$  et  $C_2(x, t)$  telles que si  $z$  est suffisamment proche de  $x$ ,

$$\left| \frac{z - Vt - y}{2\nu t} \right| \leq C_1(x, t)(1 + |y|)$$

et

$$(z - Vt - y)^2 \geq \frac{|y|^2}{2} + C_2(x, t).$$

En posant  $C(x, t) = C_1(x, t) \exp(-C_2(x, t)/4\nu t)$ , il vient

$$\left| \frac{\partial G}{\partial x}(z, t, y) \right| \leq C(x, t)(1 + |y|) \exp\left(-\frac{|y|^2}{8\nu t}\right).$$

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

$$\left| \theta_0(y) \frac{\partial G}{\partial x}(z, t, y) \right| \leq C(x, t)(1 + |y|) \exp\left(-\frac{|y|^2}{8\nu t}\right) \sup_s |\theta_0(s)|$$

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 l'échange des opérateurs d'intégration et de dérivation dans (1.3) est licite. On peut procéder de manière similaire pour justifier les deux autres commutations effectuées.

**Exercice 1.2.2** On suppose que la donnée initiale  $\theta_0$  est dérivable et uniformément bornée sur  $\mathbb{R}$ . Vérifier que

$$\theta(t, x) = \theta_0(x - Vt) \tag{1.6}$$

est bien une solution de

$$\begin{cases} \frac{\partial \theta}{\partial t} + V \frac{\partial \theta}{\partial x} = 0 & \text{pour } (x, t) \in \mathbb{R} \times \mathbb{R}_*^+ \\ \theta(t = 0, x) = \theta_0(x) & \text{pour } x \in \mathbb{R}. \end{cases} \tag{1.7}$$

Montrer que (1.6) est la limite de (1.1) lorsque le paramètre  $\nu$  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,  $\theta$  vérifie l'équation différentielle annoncée. De plus,  $\theta$  vérifie trivialement la condition initiale.

Par un raisonnement analogue à celui qui nous avait permis d'établir la continuité de la solution en  $t = 0$  dans l'exercice 1.2.1, 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

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = f & \text{dans } \Omega \times \mathbb{R}_*^+ \\ u = 0 & \text{sur } \partial\Omega \times \mathbb{R}_*^+ \\ u(t=0) = u_0 & \text{dans } \Omega \end{cases} \quad (1.8)$$

dans un domaine  $\Omega$  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 (1.8). 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 (1.8) 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 parties sur le terme de droite. On obtient ainsi que

$$\frac{1}{2} \frac{d}{dt} \left( \int_0^1 u^2 dx \right) = - \int_0^1 \left| \frac{\partial u}{\partial x} \right|^2 dx. \quad (1.9)$$

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

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

par l'inégalité de Cauchy-Schwarz 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  $v(x) = u(t, x)$ , on déduit de (1.9) que

$$\frac{1}{2} \frac{dE}{dt}(t) \leq -E(t)$$

où

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

Ainsi,

$$\frac{1}{2} \frac{d(Ee^{2t})}{dt} = \left( \frac{1}{2} \frac{dE}{dt} + E \right) e^{2t} \leq 0$$

et pour tout  $t \geq 0$ ,

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

**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

$$u(t, x) = \frac{1}{2} (u_0(x+t) + u_0(x-t)) + \frac{1}{2} (U_1(x+t) - U_1(x-t)), \quad (1.10)$$

est la solution unique de

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - \Delta u = f & \text{dans } \Omega \times \mathbb{R}_*^+ \\ u = 0 & \text{sur } \partial\Omega \times \mathbb{R}_*^+ \\ u(t=0) = u_0 & \text{dans } \Omega \\ \frac{\partial u}{\partial t}(t=0) = u_1 & \text{dans } \Omega \end{cases} \quad (1.11)$$

dans la classe des fonctions régulières.

**Correction.** La fonction  $u(t, x)$  définie par (1.10) est trivialement une solution de l'équation des ondes (1.11). 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 aux dérivées partielles par  $\frac{\partial u}{\partial t}$ , on obtient par intégration par parties que

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

On rappelle que pour toutes fonctions  $a$ ,  $b$  et  $g$  régulières,

$$\frac{d}{dt} \left( \int_{a(t)}^{b(t)} g(t, x) dx \right) = \int_{a(t)}^{b(t)} \frac{\partial g}{\partial t}(t, x) dx + b'(t)g(t, b(t)) - a'(t)g(t, a(t)).$$

On en déduit que

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



c'est-à-dire

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

Ainsi,

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

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

$$\int_{y_0}^{y_1} \left[ \left| \frac{\partial u}{\partial t}(x, t) \right|^2 + \left| \frac{\partial u}{\partial x}(x, t) \right|^2 \right] dx \leq \int_{x_0}^{x_1} \left[ \left| \frac{\partial u}{\partial t}(x, 0) \right|^2 + \left| \frac{\partial u}{\partial x}(x, 0) \right|^2 \right] dx = 0 \quad (1.12)$$

où  $x_0 = y_0 - t$  et  $x_1 = y_1 + t$ . On déduit de (1.12) 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 (1.10) 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 (1.11) 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  $[x - t, x + t]$ . On dit que 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 (1.11) sans utiliser la formule explicite (1.10). En une dimension d'espace, on pose  $\Omega = (0, 1)$  et on suppose  $f = 0$ . Soit  $u(t, x)$  une solution régulière de (1.11). 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 parties sur le second. Aucun terme de bord n'apparaît suite à l'intégration par parties car comme  $u(t, 0) = u(t, 1) = 0$ , on a  $\partial u/\partial t(t, 0) = \partial u/\partial t(t, 1) = 0$ . On a donc

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

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

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  $\Omega$ . Il y a une déperdition d'énergie par le bord de  $\Omega$ . 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

$$\begin{cases} i \frac{\partial u}{\partial t} + \Delta u - Vu = 0 & \text{dans } \mathbb{R}^N \times \mathbb{R}_*^+ \\ u(t=0) = u_0 & \text{dans } \mathbb{R}^N. \end{cases} \quad (1.13)$$

Soit  $u(t, x)$  une solution régulière de (1.13) 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

$$\mathcal{R} \left( \frac{\partial v}{\partial t} \bar{v} \right) = \frac{1}{2} \frac{\partial |v|^2}{\partial t}. \quad (1.14)$$

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

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

Par intégration par parties sur le second membre, on obtient

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

(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 pur,

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

D'après (1.14), 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

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

Par intégration par parties du second terme, on obtient que

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

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 pour l'équation de la chaleur

$$\frac{u_j^n - u_j^{n-1}}{\Delta t} + \nu \frac{-u_{j-1}^n + 2u_j^n - u_{j+1}^n}{(\Delta x)^2} = 0, \quad (1.15)$$

vérifie aussi le principe du maximum discret. On impose des conditions aux limites de Dirichlet, c'est-à-dire que la formule (1.15) 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  $\Delta t$  et  $\Delta x$ ).

**Correction.** Tout d'abord, montrons que le schéma implicite (1.15) 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

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

et  $c = \nu \Delta t / (\Delta x)^2$ . Il s'agit donc de prouver que la matrice  $A$  est inversible, ce qui est aisé, car  $A$  est symétrique, définie positive donc inversible. En effet, 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  $j \in \{0, \dots, J+1\}$ . 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 n'a rien à démontrer car  $0 \leq M$  par hypothèse. 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

$$|V|\Delta t \leq \Delta x \tag{1.16}$$

n'est pas satisfaite, le schéma décentré amont

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + V \frac{u_j^n - u_{j-1}^n}{\Delta x} = 0 \tag{1.17}$$

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 (1.11) 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 (1.11) 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,

$$u_j^{n+1} = -u_j^{n-1} + 2u_j^n + \left(\frac{\Delta t}{\Delta x}\right)^2 (u_{j-1}^n - 2u_j^n + u_{j+1}^n). \quad (1.18)$$

On initialise le schéma en posant

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

Au vu de l'équation (1.18), 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. Soit  $(\Delta t)_m$  et  $(\Delta x)_m$ , suites de discrétisations en temps et espace, tel que le schéma soit convergent. Dans ce cas, pour tout temps  $t$  et tout point de l'espace  $x$ , on a

$$u_j^{n+1} \xrightarrow{m \rightarrow \infty} u(t, x)$$

avec  $n = [t/(\Delta t)_m]$  et  $j = [x/(\Delta x)_m]$  (où les crochets désignent la partie entière). Comme nous venons de l'établir, la valeur de  $u_j^{n+1}$  dépend uniquement de la restriction de  $u_0$  et  $u_1$  sur l'intervalle  $[(j-n)(\Delta x)_m, (j+n)(\Delta x)_m]$ . Ainsi, par passage à la limite, on en déduit que la valeur de sa limite ne dépend que de la restriction de  $u_0$  et  $u_1$  à l'intervalle  $[x - t \liminf((\Delta x)_m/(\Delta t)_m), x + t \liminf((\Delta x)_m/(\Delta t)_m)]$ . Or  $u(t, x)$  dépend de toutes les valeurs de  $u_0$  et  $u_1$  sur l'intervalle  $[x - t, x + t]$ . Pour que le schéma soit convergent, il faut donc que

$$x - t \liminf_{m \rightarrow \infty} \frac{\Delta x_m}{\Delta t_m} \leq x - t \text{ et } x + t \leq x + t \liminf_{m \rightarrow \infty} \frac{(\Delta x)_m}{(\Delta t)_m},$$

c'est-à-dire que la condition CFL doit être asymptotiquement vérifiée

$$\liminf_{m \rightarrow \infty} \frac{\Delta x_m}{\Delta t_m} \geq 1.$$

**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  $u(x, y) = \frac{e^{-\sqrt{n}}}{n} \sin(ny) \text{sh}(nx)$  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. Rappelons que  $\text{sh}(x) = (e^x - e^{-x})/2$  et  $\text{ch}(x) = (e^x + e^{-x})/2$ . On vérifie sans mal que la solution proposée est une solution du système. D'autre part,

$$\frac{\partial^{k+l}u}{\partial x^k \partial y^l} = e^{-\sqrt{n}n^{l+k-1}}(i)^l \left( \frac{e^{iny} - (-1)^l e^{-iny}}{2i} \right) \left( \frac{e^{nx} - (-1)^k e^{-nx}}{2} \right).$$

On constate que en  $x = 0$ , la fonction  $u$  ainsi que toutes ses dérivées convergent vers 0 lorsque  $n$  tend vers  $+\infty$  (pour tous les entiers  $k$  et  $l$ ,  $e^{-\sqrt{n}n^{l+k-1}}$  converge vers 0 lorsque  $n$  tend vers l'infini). A contrario, si  $x > 0$ , ni  $u$  ni ses dérivées ne sont bornées par rapport à  $n$  (pour tous les entiers  $k$  et  $l$ ,  $e^{nx-\sqrt{n}n^{l+k-1}}$  diverge lorsque  $n$  tend vers l'infini). 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  $C^\infty$ ) 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 dans tout espace "raisonnable".

## Chapitre 2

# FORMULATION VARIATIONNELLE DES PROBLÈMES ELLIPTIQUES

**Exercice 2.1.1** Si  $f$  est une fonction continue sur  $[0, 1]$ , montrer que l'équation différentielle

$$\begin{cases} -\frac{d^2u}{dx^2} = f & \text{pour } 0 < x < 1 \\ u(0) = u(1) = 0. \end{cases} \quad (2.1)$$

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

$$u(x) = x \int_0^1 f(s)(1-s)ds - \int_0^x f(s)(x-s)ds \text{ pour } x \in [0, 1]. \quad (2.2)$$

**Correction.** Soit  $u$  défini par (2.2). 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 (2.1). Il reste à établir l'unicité de la solution de l'équation (2.1). L'équation étant linéaire, il suffit de montrer que toute solution  $v$  de l'équation (2.1) 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 2.2.1** 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  $\phi$  est une fonction scalaire de  $C^1(\overline{\Omega})$  et  $\sigma$  une fonction à valeurs vectorielles de  $C^1(\overline{\Omega})$ , à supports bornés dans le fermé  $\overline{\Omega}$ .



**Correction.** Un simple calcul à partir de la formule de Green (3.5) donne

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

**Exercice 2.2.2** En dimension  $N = 3$  on définit le rotationnel d'une fonction de  $\Omega$  dans  $\mathbb{R}^3$ ,  $\phi = (\phi_1, \phi_2, \phi_3)$ , comme la fonction de  $\Omega$  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  $\phi$  et  $\psi$ , 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.** En regroupant les termes puis en utilisant la formule de Green (3.5), on obtient

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

**Exercice 2.2.3** On considère le Laplacien avec condition aux limites de Neumann. Soit  $\Omega$  un ouvert borné régulier de  $\mathbb{R}^N$  et  $u$  une fonction de  $C^2(\overline{\Omega})$ . Montrer que  $u$  est une solution du problème aux limites

$$\begin{cases} -\Delta u = f & \text{dans } \Omega \\ \frac{\partial u}{\partial n} = 0 & \text{sur } \partial\Omega. \end{cases} \quad (2.3)$$

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

$$\int_{\Omega} \nabla u(x) \cdot \nabla v(x) dx = \int_{\Omega} f(x)v(x) dx \text{ pour toute fonction } v \in C^1(\overline{\Omega}). \quad (2.4)$$

En déduire qu'une condition nécessaire d'existence d'une solution dans  $C^2(\overline{\Omega})$  de (2.3) est que  $\int_{\Omega} f(x)dx = 0$ .

**Correction.** Supposons que  $u$  soit solution du problème aux limites de Neumann (2.3). En multipliant l'équation vérifiée par  $u$  par dans  $\Omega$  par une fonction test  $v \in C^1(\overline{\Omega})$ , on obtient, suite à une intégration par parties 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

$$\int_{\Omega} \nabla u(x) \cdot \nabla v(x) dx = \int_{\Omega} f(x)v(x) dx \text{ pour tout } v \in C^1(\overline{\Omega}). \quad (2.5)$$

Réciproquement, supposons que  $u$  soit une fonction régulière vérifiant (2.5). Par intégration par parties on a, pour tout  $v \in C^1(\overline{\Omega})$ ,

$$- \int_{\Omega} (\Delta u(x) + f(x))v(x) dx + \int_{\partial\Omega} \frac{\partial u}{\partial n}(x)v(x) ds = 0. \quad (2.6)$$

On procède alors en deux étapes. Dans un premier temps, on applique la relation (2.6) à des fonctions tests à support compact dans  $\Omega$ . Cela nous permet de "tester" l'équation vérifiée par  $u$  dans  $\Omega$  et d'établir l'équation  $-\Delta u = f$  dans  $\Omega$ . Dans un deuxième temps, on applique (2.6) 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  $\Omega$ ,

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

On conclut à la nullité de  $\Delta u + f$  grâce au Lemme **3.2.9** du cours. Une autre manière de procéder (si on connaît les propriétés de l'espace de Lebesgue  $L^2(\Omega)$ ) est de dire que  $\Delta u + f$  est nul car orthogonal à un sous-espace dense de  $L^2(\Omega)$ . L'égalité (2.6) devient alors

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

qui implique, par une variante du Lemme **3.2.9** pour des fonctions définies sur le bord  $\partial\Omega$ , que  $\partial u/\partial n = 0$ . Encore une fois, si on connaît les propriétés de  $L^2(\partial\Omega)$ , on peut obtenir le même résultat en disant que  $\partial u/\partial n$  est orthogonal (pour le produit scalaire de  $L^2(\partial\Omega)$ ) à un sous-espace dense de  $L^2(\partial\Omega)$ , trace des fonctions  $C^1(\overline{\Omega})$  sur le bord  $\partial\Omega$ .

Enfin, en choisissant la fonction  $v = 1$  comme fonction test dans la formulation variationnelle (2.5), on trouve que s'il existe une solution  $u$  régulière au problème aux limites (2.3), alors

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

**Exercice 2.2.4** Soit  $\Omega$  un ouvert borné régulier de  $\mathbb{R}^N$ . On considère l'équation des plaques

$$\begin{cases} \Delta(\Delta u) = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega \\ \frac{\partial u}{\partial n} = 0 & \text{sur } \partial\Omega \end{cases} \quad (2.7)$$

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 (2.7) si et seulement si  $u$  appartient à  $X$  et vérifie l'égalité

$$\int_{\Omega} \Delta u(x) \Delta v(x) dx = \int_{\Omega} f(x) v(x) dx \text{ pour toute fonction } v \in X. \quad (2.8)$$

**Correction.** On procède comme pour l'exercice précédent. Soit  $u$  une solution régulière de l'équation des plaques (2.7), pour tout  $v \in X$ ,

$$\int_{\Omega} \Delta(\Delta u)(x) v(x) dx = \int_{\Omega} f(x) v(x) dx. \quad (2.9)$$

Par intégration par parties,

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

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

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

puis par une nouvelle intégration par parties 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 (2.9) 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 (2.8), en effectuant deux intégrations par parties successives, on obtient

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

pour tout  $v \in X$ . Par application du Lemme 3.2.9 on en déduit que  $\Delta(\Delta u) - f = 0$ .

**Exercice 2.3.1** Le but de cet exercice est de montrer que l'espace  $V$ , défini par

$$V = \{v \in C^1(\bar{\Omega}), v = 0 \text{ sur } \partial\Omega\}, \quad (2.10)$$

muni du produit scalaire

$$\langle w, v \rangle = \int_{\Omega} \nabla w(x) \cdot \nabla v(x) dx, \quad (2.11)$$

n'est pas complet. Soit  $\Omega$  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

$$u_n(x) = \left| \log \left( \left| \frac{x}{2} \right|^2 + n^{-1} \right) \right|^\alpha - \left| \log(4^{-1} + n^{-1}) \right|^\alpha.$$

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.** Puisque (2.11) est le produit scalaire choisi pour  $V$ , dire qu'une suite  $u_n$  de l'espace  $V$  est de Cauchy est équivalent à dire que la suite  $\nabla u_n$  est de Cauchy dans  $L^2(\Omega)^N$ . Remarquons que, d'après l'inégalité de Poincaré du Lemme 3.3.6, on sait que la norme induite par (2.11) est équivalente à la norme suivante (qui sera celle de  $H^1(\Omega)$ )

$$N(v) = \left( \int_{\Omega} (|\nabla v|^2 + |v|^2) dx \right)^{1/2},$$

autrement dit, il existe  $C > 0$  tel que, pour tout  $v \in V$ ,

$$C N(v) \leq \left( \int_{\Omega} |\nabla v|^2 dx \right)^{1/2} \leq N(v).$$

Par conséquent, dire qu'une suite  $u_n$  de l'espace  $V$  est de Cauchy implique aussi que la suite  $u_n$  est de Cauchy dans  $L^2(\Omega)$ . Puisque l'espace  $L^2(\Omega)$  est un espace de Hilbert, donc complet, il est équivalent de dire que  $\nabla u_n$  est de Cauchy et de montrer que  $\nabla u_n$  converge dans  $L^2(\Omega)^N$  vers une limite  $\tau = \nabla u$  (où  $u$  est la limite de  $u_n$ ). Par application du théorème de convergence dominée de Lebesgue il est évident de montrer que les suites  $\nabla u_n$  de l'énoncé convergent dans  $L^2(\Omega)^N$ . (Formellement la limite  $u$  s'obtient à partir de  $u_n$  en faisant tendre  $n$  vers  $+\infty$  : on en déduit immédiatement la convergence presque partout de  $u_n$  et  $\nabla u_n$  vers  $u$  et  $\nabla u$ .) Pour

montrer que  $V$  n'est pas complet il suffit donc de montrer que la limite  $u$  n'appartient pas à  $V$  et, plus précisément, que  $\tau = \nabla u$  n'est pas continue dans  $\overline{\Omega}$ .

**Cas  $N = 1$ .** La suite  $\nabla u_n$  converge dans  $L^2(]-1, 1[)$  vers la fonction  $\tau$  définie par

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

La fonction  $\tau$  n'ayant pas de représentant continu,  $V$  n'est pas complet.

**Cas  $N = 2$ .** On trouve

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

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

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

et comme  $2\alpha - 1 < 0$ , l'intégrale est finie

$$\int_{\Omega} |\tau|^2 dx = 2^{2\alpha+1} \pi \alpha^2 (1 - 2\alpha)^{-1} \left[ \left| \log \frac{r}{2} \right|^{2\alpha-1} \right]_0^1 = \frac{2^{2\alpha+1} \pi \alpha^2}{1 - 2\alpha} (\log 2)^{2\alpha-1}.$$

Comme  $\tau$  n'est pas continue (ni même borné) en 0,  $V$  n'est pas complet.

**Cas  $N \geq 3$ .** On procède comme dans le cas précédent. On trouve que

$$\tau(x) = -\beta \frac{x}{|x|^{\beta+2}}.$$

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

En fait, dans les cas  $N = 2$  et  $N \geq 3$ , on peut même montrer que la limite  $u$  n'est pas borné à l'origine!

# Chapitre 3

## ESPACES DE SOBOLEV

**Exercice 3.2.1** Soit  $\Omega = (0, 1)$ . Montrer que la fonction  $x^\alpha$  est dérivable au sens faible dans  $L^2(\Omega)$  si et seulement si  $\alpha > 1/2$ .

**Correction.** Tout d'abord,  $x^\alpha$  appartient à  $L^2(0, 1)$  si et seulement si  $\alpha > -1/2$ . On se restreint donc à  $\alpha > -1/2$ . D'après la Définition 4.2.3,  $x^\alpha$  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  $\varphi$  est à support compact dans  $]0, 1[$ , il existe  $a > 0$  tel que  $\varphi(x) = 0$  pour tout  $x \in ]0, a[$ . 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,  $x^\alpha$  étant de classe  $C^\infty$  sur cet intervalle). On en déduit que  $x^\alpha$  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 3.2.2** Soit  $\Omega$  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

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

de sorte que, pour tout indice  $i$ , la restriction de  $f$  à  $\Omega_i$  (notée  $f_i$ ) 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  $\Omega$  et  $n^i$  la normale extérieure à l'ouvert  $\Omega_i$  de composantes  $(n_k^i)_{1 \leq k \leq N}$ . Soit  $\varphi \in C_c^\infty(\Omega)$ . En

appliquant la formule de Green (3.5) à chacun des ouverts  $\Omega_i$ , on obtient

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

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  $\psi_k$  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 3.2.3** Soit  $\Omega$  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  $\Omega$ , car  $\Omega \setminus \cup_i \Omega_i$  est de mesure nulle. De plus, pour tout indice  $k$ , et toute fonction test  $\varphi$ ,

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

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. Donc, une fonction  $C^1$  par morceaux mais discontinue n'est pas dérivable au sens faible.

**Exercice 3.2.4** Soit  $\Omega$  un ouvert borné constitué de deux ouverts  $\Omega_1$  et  $\Omega_2$  séparés par une surface  $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$ . Montrer qu'une fonction vectorielle de classe  $C^1$  sur chaque morceau  $\Omega_1$  et  $\Omega_2$  admet une divergence faible dans  $L^2(\Omega)$  si et seulement si sa composante normale est continue à travers la surface  $\Gamma$ .

**Correction.** Soit  $\sigma$  une fonction de  $\Omega$  à valeurs vectorielles. On note  $\sigma_1, \sigma_2$  les restrictions de  $\sigma$  à  $\Omega_1$  et  $\Omega_2$  respectivement et  $n^1, n^2$  les normales extérieures à  $\Omega_1$  et  $\Omega_2$ . Soit  $\varphi \in C_c^\infty(\Omega)$ , d'après la formule de Stokes (voir Exercice 2.2.1),

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

Ainsi, si la composante normale de  $\sigma$  est continue à l'interface, c'est-à-dire si  $\sigma_1 \cdot n^1 + \sigma_2 \cdot n^2 = 0$  sur  $\Gamma$ , on en déduit que

$$\int_{\Omega} \sigma(x) \cdot \nabla \varphi(x) dx = - \int_{\Omega} \psi(x) \varphi(x) dx,$$

avec  $\psi(x) = \operatorname{div} \sigma_i(x)$  pour tout  $x \in \Omega_i$  ( $i = 1, 2$ ). La fonction à valeurs vectorielles  $\sigma$  admet donc une divergence faible et  $\operatorname{div} \sigma(x) = \psi(x)$ .

Réciproquement, si  $\sigma$  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  $\Gamma$ .

**Exercice 3.3.1** Montrer que les fonctions continues,  $C^1$  par morceaux et à support borné dans  $\bar{\Omega}$ , appartiennent à  $H^1(\Omega)$ .

**Correction.** Soit  $f$  une fonction continue,  $C^1$  par morceaux et à support borné dans  $\bar{\Omega}$ . 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,  $\psi_k$  est continue par morceaux, à support borné et donc appartient à  $L^2(\Omega)$ . Ainsi  $f$  admet une dérivée faible dans  $L^2(\Omega)$  et appartient à  $H^1(\Omega)$ .

**Exercice 3.3.2** Soit  $B$  la boule unité ouverte de  $\mathbb{R}^N$ . Si  $N = 2$ , montrer que la fonction  $u(x) = |\log(|x|/2)|^\alpha$  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  $0 < \alpha < 1/2$  et  $u$  la fonction définie sur la boule unité de  $\mathbb{R}^2$  par

$$u(x) = |\log(|x|/2)|^\alpha.$$

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

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

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

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

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

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

En passant en coordonnées polaires, on obtient

$$\begin{aligned} \int_B |\psi|^2 dx &= 2\pi\alpha^2 \int_0^1 \frac{|\log(r/2)|^{2(\alpha-1)}}{r} dr \\ &= \frac{4\pi\alpha^2}{1-2\alpha} [|\log(r/2)|^{2\alpha-1}]_0^1 < +\infty. \end{aligned}$$

Ainsi,  $\psi$  est un élément de  $L^2(B)$ . 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  $\varphi \in C_0^\infty(B)$ , pour tout réel  $\varepsilon$  tel que  $0 < \varepsilon < 1$ ,

$$\begin{aligned} \int_B u(x) \nabla \varphi(x) dx &= \int_{\varepsilon < |x| < 1} u(x) \nabla \varphi(x) dx + \int_{|x| < \varepsilon} u(x) \nabla \varphi(x) dx \\ &= - \int_{\varepsilon < |x| < 1} \psi(x) \varphi(x) dx + \int_{|x|=\varepsilon} u(x) \varphi(x) ds + \int_{|x| < \varepsilon} u(x) \nabla \varphi(x) dx \\ &= - \int_{\varepsilon < |x| < 1} \psi(x) \varphi(x) dx + |\log(\varepsilon/2)|^\alpha \int_{|x|=\varepsilon} \varphi(x) ds + \int_{|x| < \varepsilon} u(x) \nabla \varphi(x) dx. \end{aligned}$$

Les deux derniers termes de l'expression tendent vers zéro lorsque  $\varepsilon$  tend vers zéro. En effet,

$$|\log(\varepsilon/2)|^\alpha \int_{|x=\varepsilon|} \varphi(x) ds \sim |\log(\varepsilon/2)|^\alpha 2\pi\varepsilon\varphi(0) \xrightarrow{\varepsilon \rightarrow 0} 0$$

et

$$\left| \int_{|x|<\varepsilon} u(x)\nabla\varphi dx \right| \leq \|u\|_{L^2(B)} \left( \int_{|x|<\varepsilon} |\nabla\varphi|^2 dx \right)^{1/2} \xrightarrow{\varepsilon \rightarrow 0} 0$$

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

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

La fonction  $u$  n'est pas bornée en 0. On vérifie néanmoins 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  $\psi$  est un élément de  $L^2(B)^N$ . 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 3.3.3** Le but de cet exercice est de montrer que le Théorème de trace **4.3.13** n'est pas vrai si l'ouvert  $\Omega$  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 note  $(x, y)$  les coordonnées d'un point de  $\mathbb{R}^2$ . 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  $2\alpha + r > 1$ . Soit  $\Gamma_1 = \partial\Omega \cap \{(x, y) \in \mathbb{R}^2 : y = 0\}$ ,  $\Gamma_2$  la partie du bord de  $\Omega$  paramétrée par la fonction  $(0, 1) \ni x \mapsto (x, y) = (x, x^r) \in \partial\Omega$  et  $\Gamma_3 = \partial\Omega \cap \{(x, y) \in \mathbb{R}^2 : x = 1\}$ . On a d'autre part,

$$\begin{aligned} \int_{\partial\Omega} |v|^2 ds &= \int_{\Gamma_1} |v(x)|^2 dx + \int_{\Gamma_2} |v(x)|^2 ds(x) + \int_{\Gamma_3} |v(1)|^2 dy \\ &= \int_0^1 x^{2\alpha} dx + \int_0^1 x^{2\alpha} (1 + r^2 x^{2r-2})^{1/2} dx + 1. \end{aligned}$$

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  $\alpha$  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 de trace **4.3.13** est mis en défaut dans ce cas.

On peut aussi montrer que l'application trace n'est pas continue : 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

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

On ne peut donc pas trouver une constante  $C > 0$  telle que, pour toute fonction  $v \in H^1(\Omega) \cap C(\overline{\Omega})$ ,

$$\|v\|_{L^2(\partial\Omega)} \leq C \|v\|_{H^1(\Omega)},$$

puisque, au contraire, quelque soit  $K > 0$ , pour  $n$  assez grand, on a

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

**Exercice 3.3.4** 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  $\Omega$  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  $\Omega$  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 convergence dominée de Lebesgue,  $\|u^n\|_{L^2(\Omega)} \rightarrow 0$ , et quelque soit  $K > 0$ , pour  $n$  assez grand,

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

L'opérateur trace défini de  $C(\overline{\Omega}) \cap L^2(\Omega)$  dans  $L^2(\partial\Omega)$  n'est pas continu, c'est-à-dire qu'il ne peut pas exister une constante  $C > 0$  telle que, pour toute fonction  $v \in L^2(\Omega) \cap C(\overline{\Omega})$ ,

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

A fortiori, il ne peut être prolongé en une application continue de  $L^2(\Omega)$  dans  $L^2(\partial\Omega)$ .



## Chapitre 4

# ÉTUDE MATHÉMATIQUE DES PROBLÈMES ELLIPTIQUES

**Exercice 4.2.1** A l'aide de l'approche variationnelle démontrer l'existence et l'unicité de la solution de

$$\begin{cases} -\Delta u + u = f & \text{dans } \Omega, \\ u = 0 & \text{sur } \partial\Omega, \end{cases} \quad (4.1)$$

où  $\Omega$  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  $\Omega$  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 parties, on obtient que

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

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 (4.1) 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

$$a(u, v) = \int_{\Omega} (\nabla u \cdot \nabla v + uv) \, dx \quad \text{et} \quad L(v) = \int_{\Omega} f v \, dx.$$

2eme Étape. Résolution du problème variationnel.

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

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

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 parties que lors de la première étape que  $\nabla u$  est un élément de

$H(\text{div})$  (voir cours, Section 4.4.2) et que  $-\Delta u + u = f$  en tant qu'éléments de  $L^2(\Omega)$  et donc presque partout dans  $\Omega$ . Enfin, comme  $u \in H_0^1(\Omega)$ , et que  $\Omega$  est un ouvert régulier, la trace de  $u$  est bien définie et  $u = 0$  presque partout sur  $\partial\Omega$ .

**Exercice 4.2.2** Soit  $\Omega$  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

$$\begin{cases} V \cdot \nabla u - \Delta u = f & \text{dans } \Omega, \\ u = 0 & \text{sur } \partial\Omega, \end{cases} \quad (4.2)$$

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

**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 parties, 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

$$a(u, v) = \int_{\Omega} (\nabla u \cdot \nabla v + (V \cdot \nabla u)v) dx \quad \text{et} \quad L(v) = \int_{\Omega} f v dx.$$

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  $a(\cdot, \cdot)$ .

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

La divergence de  $V$  étant nulle, on a

$$\int_{\Omega} (V \cdot \nabla u)u dx = \int_{\Omega} (\text{div}(uV)u - \text{div}(V)|u|^2) dx = \int_{\Omega} \text{div}(uV)u dx.$$

Par intégration par parties et comme  $u = 0$  sur le bord  $\partial\Omega$ , il vient

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

Ainsi,

$$\int_{\Omega} (V \cdot \nabla u)u dx = 0$$

et

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

La coercivité de  $a(\cdot, \cdot)$  se déduit alors de l'inégalité de Poincaré.

3ème Étape. Équivalence avec l'équation.

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

Ainsi, en majorant le membre de droite,

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

et  $\nabla u$  est un élément de  $H(\text{div})$ . On en déduit donc par intégration par parties que  $-\Delta u + V \cdot \nabla u = f$  en tant qu'éléments de  $L^2(\Omega)$ . Enfin, comme  $u \in H_0^1(\Omega)$ , on a  $u = 0$  presque partout sur  $\partial\Omega$ .

**Exercice 4.2.3** On reprend les notations et hypothèses de l'Exercice 5.2.2. Montrer que tout  $v \in H_0^1(\Omega)$  vérifie

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

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

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

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

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

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

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

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 4.2.4** On considère à nouveau le problème aux limites

$$\begin{cases} -\Delta u = f & \text{dans } \Omega, \\ u = 0 & \text{sur } \partial\Omega, \end{cases} \quad (4.3)$$

où  $\Omega$  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  $\Omega$  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 (4.3) a la même symétrie. Montrer que (4.3) 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.** Deux approches sont possibles. On peut raisonner soit directement sur le problème aux limites (4.3), soit sur la formulation variationnelle associée. Un raisonnement direct sur (4.3) peut se justifier rigoureusement si l'on admet que la solution  $u$  est régulière et unique, tandis que l'utilisation de la formulation variationnelle permet de s'affranchir de toute hypothèse restrictive. Dans les deux cas on utilise la notation  $x = (x', x_N) \in \mathbb{R}^N$  et on introduit la symétrie  $s$  de  $\mathbb{R}^N$  par rapport au plan  $x_N = 0$ , définie par

$$s(x', x_N) = (x', -x_N).$$

1ère approche.

Soit  $u$  la solution de (4.3) supposée régulière et unique. On vérifie aisément que

$$-\Delta(u \circ s) = -(\Delta u) \circ s = f \circ s = f \quad \text{dans } \Omega,$$

où on a utilisé l'invariance de  $\Omega$  par  $s$ . De même,  $u \circ s = 0$  sur  $\partial\Omega$ . Par conséquent,  $u \circ s$  est aussi solution de (4.3) et, par unicité, on en déduit que  $u \circ s = u$  dans  $\Omega$ . Par ailleurs,

$$\frac{\partial(u \circ s)}{\partial x_N} = -\frac{\partial u}{\partial x_N} \circ s,$$

ce qui implique, en  $x_N = 0$ ,

$$\frac{\partial u}{\partial x_N}(x', 0) = -\frac{\partial u}{\partial x_N}(x', 0) = 0.$$

Par conséquent, la restriction de  $u$  à  $\Omega^+ = \Omega \cap \{x_N > 0\}$  vérifie

$$\begin{cases} -\Delta u = f & \text{dans } \Omega^+, \\ \frac{\partial u}{\partial n} = 0 & \text{sur } \Omega \cap \{x_N = 0\}, \\ u = 0 & \text{sur } \partial\Omega \cap \partial\Omega^+. \end{cases} \quad (4.4)$$

2ème approche.

On va maintenant utiliser la formulation variationnelle de (4.3), sans aucune hypothèse de régularité sur sa solution. Il s'agit de trouver  $u \in H_0^1(\Omega)$  tel que

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx \quad \forall v \in H_0^1(\Omega). \quad (4.5)$$

On note  $S$  l'application de  $L^2(\Omega)$  à valeurs dans  $L^2(\Omega)$  qui à toute fonction  $v \in L^2(\Omega)$  associe la fonction  $S(v) = v \circ s$ . L'application  $S$  est une isométrie de  $L^2(\Omega)$ , c'est-à-dire que

$$\int_{\Omega} S(v) S(w) \, dx = \int_{\Omega} v w \, dx \quad \forall v, w \in L^2(\Omega).$$

De plus, la restriction de  $S$  à l'espace  $H^1(\Omega)$  est également une isométrie. En effet, pour toute fonction régulière  $v$ , on a

$$\nabla(S(v)) = \nabla(v \circ s) = (\nabla s)^*(\nabla v \circ s),$$

où  $A^*$  désigne la matrice adjointe (ou transposée) de  $A$ , et

$$\int_{\Omega} |\nabla(S(v))|^2 dx = \int_{\Omega} (\nabla v \circ s)^* \nabla s (\nabla s)^* (\nabla v \circ s) dx$$

Comme  $s$  est une isométrie de  $\mathbb{R}^N$ ,  $\nabla s (\nabla s)^*$  n'est autre que l'identité et

$$\int_{\Omega} |\nabla(S(v))|^2 dx = \int_{\Omega} |\nabla v \circ s|^2 dx,$$

et donc (par simple changement de variable),

$$\int_{\Omega} |\nabla(S(v))|^2 dx = \int_{\Omega} |\nabla v|^2 dx. \quad (4.6)$$

Par densité des fonctions régulières dans  $H^1(\Omega)$ , on en déduit la relation (4.6) pour tout fonction  $v \in H^1(\Omega)$  et que  $S$  est une isométrie de  $H^1(\Omega)$ . Par un raisonnement similiaire, l'ensemble des fonctions  $C_0^\infty(\Omega)$  étant stable par  $S$ , on en déduit que  $S$  est une isométrie de  $H_0^1(\Omega)$ . Enfin, la relation

$$\nabla(S(v)) = (\nabla s)^*(\nabla v \circ s),$$

valable pour toute fonction  $v$  régulière, s'étend par densité à tout élément de  $H^1(\Omega)$ .

Par changement de variable  $x = s(y)$  dans la formulation variationnelle (4.5), il vient

$$\int_{\Omega} ((\nabla u) \circ s) \cdot ((\nabla u) \circ s) dy = \int_{\Omega} (f \circ s)(v \circ s) dy$$

et, comme  $f \circ s = f$ ,

$$\int_{\Omega} (\nabla S(u)) \cdot (\nabla S(v)) dy = \int_{\Omega} f S(v) dy.$$

L'application  $S$  étant une isométrie de  $H_0^1(\Omega)$ , on en déduit que, pour tout  $w \in H_0^1(\Omega)$ , il existe  $v \in H_0^1(\Omega)$  tel que  $w = S(v)$  et

$$\int_{\Omega} \nabla S(u) \cdot \nabla w dy = \int_{\Omega} f w dy.$$

Par conséquent,  $S(u)$  est aussi solution de la formulation variationnelle (4.5) et, par unicité, on a  $S(u) = u$ .

Reste à montrer que la restriction de  $u$  à l'ouvert  $\Omega^+$  est solution de la formulation variationnelle associée à (4.4). Pour cela, on introduit l'espace

$$X = \{v \in H^1(\Omega^+) \text{ tel que } v = 0 \text{ sur } \partial\Omega^+ \cap \partial\Omega\},$$

et, pour l'instant, on note  $\tilde{u}$  la solution de (4.4). La formulation variationnelle de (4.4) est classiquement : trouver  $\tilde{u} \in X$  tel que

$$\int_{\Omega^+} \nabla \tilde{u} \cdot \nabla v \, dx = \int_{\Omega^+} f v \, dx \quad \forall v \in X. \quad (4.7)$$

On introduit l'application de prolongement  $P$  de  $L^2(\Omega^+)$  à valeur dans  $L^2(\Omega)$  définie par

$$P(u)(x) := \begin{cases} u(x) & \text{si } x \in \Omega^+, \\ u \circ s(x) & \text{si } x \in \Omega \setminus \Omega^+. \end{cases}$$

En fait, l'application  $P$ , restreinte aux fonctions de  $H^1(\Omega^+)$ , est continue de  $H^1(\Omega^+)$  à valeurs dans  $H^1(\Omega)$  car les traces de  $u$  et  $u \circ s$  coïncident sur  $\{x_N = 0\}$ . Dans la formulation variationnelle (4.5) on choisit la fonction test  $P(v)$  avec  $v \in X$

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

Par un changement de variables,

$$\begin{aligned} \int_{\Omega} \nabla u \cdot \nabla P(v) \, dx &= \int_{\Omega^+} \nabla u \cdot \nabla P(v) \, dx + \int_{\Omega^-} \nabla u \cdot \nabla P(v) \, dx \\ &= \int_{\Omega^+} \nabla u \cdot \nabla P(v) \, dx + \int_{\Omega^+} ((\nabla u) \circ s) \cdot ((\nabla P(v)) \circ s) \, dx \\ &= \int_{\Omega^+} \nabla u \cdot \nabla P(v) \, dx + \int_{\Omega^+} \nabla(u \circ s) \cdot \nabla(P(v) \circ s) \, dx. \end{aligned}$$

Or,  $u \circ s = u$  et  $P(v) \circ s = P(v)$ , ainsi

$$\int_{\Omega} \nabla u \cdot \nabla P(v) \, dx = 2 \int_{\Omega^+} \nabla u \cdot \nabla P(v) \, dx.$$

De même, on a également

$$\int_{\Omega} f P(v) \, dx = 2 \int_{\Omega^+} f P(v) \, dx.$$

Puisque  $P(v) = v$  dans  $\Omega^+$ , on a donc obtenu

$$\int_{\Omega^+} \nabla u \cdot \nabla v \, dx = \int_{\Omega^+} f v \, dx \quad \forall v \in X,$$

qui n'est rien d'autre que (4.7). Par unicité de la solution on en déduit  $\tilde{u} = u$  dans  $\Omega^+$ , c'est-à-dire que  $u$  est bien solution (faible) de (4.4).

**Exercice 4.2.5** Démontrer que l'unique solution  $u \in H^1(\Omega)$  de la formulation variationnelle

$$\int_{\Omega} (\nabla u \cdot \nabla v + uv) \, dx = \int_{\partial\Omega} g v \, ds + \int_{\Omega} f v \, dx \quad \forall v \in H^1(\Omega) \quad (4.8)$$

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 (4.8) 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 **4.3.13**, il existe donc une constante positive  $C$  (qui ne dépend que de  $\Omega$ ) telle que

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

d'où l'on déduit l'inégalité recherchée.

**Exercice 4.2.6** On suppose que  $\Omega$  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

$$\begin{cases} -\Delta u = f & \text{dans } \Omega \\ \frac{\partial u}{\partial n} + u = g & \text{sur } \partial\Omega \end{cases} \quad (4.9)$$

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 parties, 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

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v dx + \int_{\partial\Omega} uv ds \quad \text{et} \quad L(v) = \int_{\Omega} f v dx + \int_{\partial\Omega} g v ds.$$

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  $a(\cdot, \cdot)$ . 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, dans l'esprit de la démonstration de l'inégalité de Poincaré (voir cours, p. 94). 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 4.3.21, 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 sur chacune des composantes connexes de  $\Omega$ . L'application trace étant continue de  $H^1(\Omega)$  dans  $L^2(\partial\Omega)$ , la trace de  $v$  sur le bord de  $\Omega$  est égale à la limite des traces de  $v_{n'}$  sur le bord de  $\Omega$ . Or  $\lim_n \|v_{n'}\|_{L^2(\partial\Omega)} = 0$ , ainsi  $v = 0$  sur  $\partial\Omega$ . Finalement,  $v$  étant constante sur chacune de ces composantes connexes,  $v = 0$  dans tout  $\Omega$ , ce qui contredit le fait que  $\|v\|_{L^2(\Omega)} = \lim_n \|v_{n'}\|_{L^2(\Omega)} = 1$ .

3ème Étape. Équivalence avec le problème aux limites.

Tout d'abord, en appliquant la formulation variationnelle à des fonctions tests  $v \in C_c^\infty(\Omega)$  on établit que  $\nabla u$  est un élément de  $H(\text{div})$  (voir cours, Section 4.4.2) et par intégration par parties que

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

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

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

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 4.2.1** *En toute rigueur, l'intégrale  $\int_{\partial\Omega} (\frac{\partial u}{\partial n} + u)v \, ds$  n'est a priori pas correctement définie. Cependant, comme  $\nabla u$  est un élément de  $H(\text{div})$ , il admet une trace normale sur  $\partial\Omega$  (voir le Théorème 4.4.7). 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  $\int_{\partial\Omega} (\frac{\partial u}{\partial n} + u)v \, ds$  est justifiée a posteriori.*

**Exercice 4.2.7** On suppose que  $\Omega$  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

$$\begin{cases} -\Delta u = f & \text{dans } \Omega, \\ \frac{\partial u}{\partial n} = 0 & \text{sur } \partial\Omega_N, \\ u = 0 & \text{sur } \partial\Omega_D, \end{cases} \quad (4.10)$$

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

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx \quad \text{et} \quad L(v) = \int_{\Omega} f v \, dx.$$

L'espace  $V$  a été étudié dans la Remarque 4.3.18 du cours dont on rappelle les résultats. 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  $v \in V$ ,

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

Cette inégalité s'établit par contradiction (voir dans le cours la deuxième démonstration de l'inégalité de Poincaré, p. 94). Supposons que cette inégalité soit fautive pour toute constante  $C$ . Dans ce cas, pour tout entier  $n$ , il existe  $v_n \in V$  tel que

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

Quitte à diviser  $v_n$  par sa norme  $L^2$ , on peut supposer que  $\|v_n\|_{L^2(\Omega)} = 1$ . Ainsi,  $v_n$  est borné dans  $H^1(\Omega)$  et d'après le Théorème de Rellich 4.3.21, il existe une sous-suite  $v_{n'}$  de  $v_n$  et un élément  $v$  de  $L^2(\Omega)$  tels que  $v_{n'}$  converge vers  $v$  en norme  $L^2$ . Or  $\nabla v_n$  converge vers zéro. On en déduit que  $v_{n'}$  est une suite de Cauchy dans  $H^1(\Omega)$ . En particulier,  $v$  appartient à  $H^1(\Omega)$  et le gradient de  $v$  est égal à la limite des gradients de  $v_{n'}$ , c'est-à-dire  $\nabla v = 0$ . D'après la Proposition 4.2.5, on en déduit que  $v$  est une constante (car  $\Omega$  est connexe). Comme  $v$  appartient à  $V$ , la trace de  $v$  sur  $\partial\Omega_D$  est nulle. La mesure superficielle de  $\partial\Omega_D$  étant non nulle, on en déduit que  $v = 0$ , ce qui contredit le fait que  $\|v\|_{L^2(\Omega)} = \lim_{n'} \|v_{n'}\|_{L^2(\Omega)} = 1$ .

Enfin, si  $u$  est une solution du problème variationnel, on en déduit que  $\nabla u$  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

$$\int_{\partial\Omega} \frac{\partial u}{\partial n} v \, ds = \int_{\Omega} (\Delta u v + \nabla u \cdot \nabla v) \, dx = \int_{\Omega} (-f v + \nabla u \cdot \nabla v) \, dx = 0.$$

(On renvoie à la Remarque 4.2.1 pour une interprétation rigoureuse de l'intégrale de bord ci-dessus.) Quitte à supposer  $\Omega$  et  $\partial\Omega_N$  assez réguliers, l'ensemble des traces de fonctions de  $V$  sur le bord est dense dans l'espace  $L^2(\partial\Omega_N)$  (voir la Remarque 4.3.17). Ainsi, la trace  $\partial u/\partial n$  sur  $\partial\Omega_N$  est nulle. Enfin,  $u = 0$  presque partout sur  $\partial\Omega_D$  car  $u \in V$ . Ainsi, la solution  $u$  de la formulation variationnelle est bien solution du problème aux limites (4.10).

**Exercice 4.2.8** Démontrer l'inégalité de Poincaré-Wirtinger : si  $\Omega$  est borné, régulier et connexe, il existe une constante  $C > 0$  telle que, pour tout  $v \in H^1(\Omega)$ ,

$$\|v - m(v)\|_{L^2(\Omega)} \leq C \|\nabla v\|_{L^2(\Omega)} \quad \text{avec} \quad m(v) = \frac{\int_{\Omega} v \, dx}{\int_{\Omega} dx}. \quad (4.11)$$

**Correction.** On démontre cette inégalité par contradiction (comme dans le cours pour l'inégalité de Poincaré, p. 94). On suppose que l'inégalité de Poincaré Wirtinger est fautive. Dans ce cas, pour tout entier naturel  $n \geq 1$ , il existe un élément  $u_n$  de  $H^1(\Omega)$  tel que

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

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

$$1 = \|v_n\|_{L^2(\Omega)} > n \|\nabla v_n\|_{L^2(\Omega)}. \quad (4.12)$$

Ainsi, la suite  $v_n$  est bornée dans  $H^1(\Omega)$ . Comme  $\Omega$  est borné régulier, d'après le Théorème de Rellich 4.3.21, 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 (4.12),  $\nabla v_n$  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  $H^1(\Omega)$  vers un élément  $v$ . De plus, on a

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

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

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

Comme  $\nabla v = 0$ ,  $m(v) = 0$  et  $\Omega$  est connexe,  $v$  est une constante de moyenne nulle d'après la Proposition 4.2.5. Ainsi  $v = 0$ , ce qui contredit l'égalité  $\|v\|_{L^2(\Omega)} = 1$  et achève la démonstration de l'inégalité (4.11).

**Exercice 4.2.9** On suppose que  $\Omega$  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

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx + \left( \int_{\Omega} u \, dx \right) \left( \int_{\Omega} v \, dx \right) \quad \text{et} \quad L(v) = \int_{\Omega} f(x)v(x) \, dx.$$

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

$$a(u, v) = L(v) \quad \forall v \in H^1(\Omega). \quad (4.13)$$

Afin d'appliquer le Théorème de Lax-Milgram, la seule hypothèse non triviale à vérifier porte sur la coercivité de la forme bilinéaire  $a(\cdot, \cdot)$ . En raisonnant par l'absurde (comme dans l'Exercice 4.2.8 précédent), on établit qu'il existe  $C > 0$  tel que, pour tout  $v \in H^1(\Omega)$ ,

$$\|v\|_{L^2(\Omega)} \leq C \left( \|\nabla v\|_{L^2(\Omega)} + \left| \int_{\Omega} v \, dx \right| \right),$$

d'où l'on déduit aisément qu'il existe  $\nu > 0$  tel que, pour tout  $v \in H^1(\Omega)$ ,

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

(On utilise ici le fait que  $\Omega$  est borné connexe.) Le Théorème de Lax-Milgram nous assure alors l'existence et l'unicité de la solution de (4.13).

2. Détermination du problème aux limites

En prenant la fonction test  $v = 1$  dans la formulation variationnelle (4.13), on obtient que

$$\int_{\Omega} u \, dx = \frac{1}{|\Omega|} \int_{\Omega} f \, dx. \quad (4.14)$$

En appliquant l'inégalité de Cauchy-Schwarz à la forme linéaire  $L(v)$  et au deuxième terme de la forme bilinéaire  $a(u, v)$  on déduit de (4.13) et (4.14)

$$\left| \int_{\Omega} \nabla u \cdot \nabla v \, dx \right| \leq C \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)}.$$

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

$$-\text{div}(\nabla u) = f - |\Omega|^{-1} \int_{\Omega} f \, dx \quad \text{dans } \Omega.$$



Enfin, comme  $\nabla u \in H(\text{div})$ , la trace de  $\partial u / \partial n$  sur la frontière de  $\Omega$  est correctement définie (voir le Théorème 4.4.7) et on établit aisément que  $\partial u / \partial n = 0$ . Le problème aux limites résolu est donc

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

Dans le cas particulier  $\int_{\Omega} f \, dx = 0$ ,  $u$  est solution du problème de Neumann (5.25) avec en plus la condition (4.14) qui assure l'unicité de cette solution.

**Exercice 4.2.10** Soit  $\Omega$  un ouvert borné et  $K$  un compact connexe de  $\mathbb{R}^N$  inclus dans  $\Omega$  (on suppose que  $\Omega \setminus K$  est régulier). Soit  $f \in L^2(\Omega)$ . On considère un problème de conduction dans  $\Omega$  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

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

muni de la norme de  $H^1(\Omega \setminus K)$ . Notons que la constante dans la définition de l'espace  $X$  varie d'une fonction  $v$  à l'autre : on la désigne dans la suite par  $v(\partial K)$ . Comme sous-espace fermé de  $H^1(\Omega \setminus K)$ ,  $X$  est un espace de Hilbert.

1ère 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 parties, on en déduit que

$$\int_{\Omega \setminus K} \nabla u \cdot \nabla v \, dx - \int_{\partial K} \frac{\partial u}{\partial n} v \, ds = \int_{\Omega \setminus K} f v \, dx \quad (4.15)$$

Comme  $v(x)$  est une constante,  $v(\partial K)$ , sur  $\partial K$ , on a

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

puisque la solution doit vérifier la condition  $\int_{\partial K} \frac{\partial u}{\partial n} ds = 0$ . L'équation (4.15) vérifiée par  $u$  se simplifie en

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

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

$$a(u, v) = L(v) \quad \forall v \in X, \quad (4.16)$$

où  $a(\cdot, \cdot)$  est la forme bilinéaire et  $L(\cdot)$  la forme linéaire, définies sur  $X$  par

$$a(u, v) = \int_{\Omega \setminus K} \nabla u \cdot \nabla v \, dx \quad \text{et} \quad L(v) = \int_{\Omega \setminus K} f v \, dx.$$

2ème Étape. Existence de solution.

L'application du Théorème de Lax-Milgram est triviale grâce à l'inégalité de Poincaré pour les fonctions de  $X$  et assure l'existence et l'unicité d'une solution au problème variationnel (4.16).

3ème Étape. Équivalence avec le problème aux limites.

On déduit immédiatement de la formulation variationnelle (4.16) que  $\nabla u \in H(\text{div})$  et

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

Comme  $\nabla u \in H(\text{div})$ ,  $\partial u / \partial n$  admet une trace sur  $\partial K$  (voir le Théorème 4.4.7). En intégrant par parties dans la formulation variationnelle (4.16), il vient

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

Comme la fonction test  $v$ , et donc la constante  $v(\partial K)$ , est quelconque, on en déduit que  $\int_{\partial K} \frac{\partial u}{\partial n} ds = 0$ . Enfin, les conditions de type Dirichlet  $u = 0$  sur  $\partial\Omega$  et  $u = \text{constante}$  sur  $\partial K$  ont été incluses dans la définition de l'espace  $X$  auquel appartient  $u$ .

**Exercice 4.2.11** Soit  $\Omega$  un ouvert borné régulier de  $\mathbb{R}^N$  et  $A$  une application de  $\Omega$  dans l'ensemble des matrices symétriques  $N \times N$ . On suppose que l'application  $A$  est uniformément bornée, c'est-à-dire qu'il existe une constante  $\beta > 0$  telle que, presque partout dans  $\Omega$ ,

$$|A(x)\xi \cdot \xi| \leq \beta |\xi|^2 \text{ pour tout } \xi \in \mathbb{R}^N,$$

et qu'elle est uniformément elliptique, c'est-à-dire qu'il existe une constante  $\alpha > 0$  telle que

$$A(x)\xi \cdot \xi \geq \alpha |\xi|^2 \text{ pour tout } \xi \in \mathbb{R}^N.$$

Soit  $f \in L^2(\Omega)$ , montrer qu'il existe une unique solution faible au problème aux limites

$$\begin{cases} -\text{div}(A\nabla u) = f & \text{dans } \Omega, \\ u = 0 & \text{sur } \partial\Omega. \end{cases} \quad (4.17)$$

**Correction.** Dans un premier temps, établissons la formulation variationnelle associée. On note  $\sigma = A\nabla u$  qui est une fonction de  $\Omega$  dans  $\mathbb{R}^N$ . De (4.17) on déduit que

$$-\operatorname{div}\sigma = f \text{ dans } \Omega.$$

On multiplie alors cette équation par une fonction test  $v \in H_0^1(\Omega)$  (espace choisi à cause de la condition aux limites de Dirichlet) et on utilise la formule d'intégration par parties, dite de Stokes (voir l'Exercice 3.2.1 ou le Théorème 4.4.7),

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

Comme  $v$  s'annule sur le bord  $\partial\Omega$  on obtient, pour tout  $v \in H_0^1(\Omega)$ ,

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

On note  $a(u, v)$  la forme bilinéaire définie par le membre de gauche de cette équation et  $L(v)$  la forme linéaire définie par le membre de droite. La formulation variationnelle du problème aux limites (4.17) est donc de déterminer  $u \in H_0^1(\Omega)$  tel que

$$a(u, v) = L(v) \quad \text{pour tout } v \in H_0^1(\Omega). \quad (4.18)$$

Il faut maintenant vérifier les hypothèses du Théorème de Lax-Milgram afin d'établir l'existence et l'unicité d'une solution de (4.18). Elles sont trivialement vérifiées, exceptées la continuité et la coercivité de la forme bilinéaire  $a$  que nous détaillons.

Pour presque tout  $x \in \Omega$ , la matrice  $A(x)$  étant symétrique réelle, elle est diagonalisable dans une base orthonormée de vecteurs propres. Comme pour tout  $\xi \in \mathbb{R}^N$ ,

$$0 \leq A(x)\xi \cdot \xi \leq \beta|\xi|^2,$$

la plus grande valeur propre de  $A(x)$  est inférieure à  $\beta$  et donc  $\rho(A) = \|A\|_2 \leq \beta$  (cf. le Lemme 13.1.6). On en déduit donc que, pour tout  $\xi, \xi' \in \mathbb{R}^N$ ,

$$|A(x)\xi \cdot \xi'| \leq \|A(x)\|_2 |\xi| |\xi'| \leq \beta |\xi| |\xi'|.$$

D'après cette majoration et l'inégalité de Cauchy-Schwarz, on a

$$|a(u, v)| \leq \beta \int_{\Omega} |\nabla u| |\nabla v| dx \leq \beta \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}.$$

La forme bilinéaire  $a$  est donc continue sur  $H_0^1(\Omega)$ . De plus, d'après l'hypothèse d'uniforme ellipticité, pour tout  $u \in H_0^1(\Omega)$ ,

$$a(u, u) = \int_{\Omega} A\nabla u \cdot \nabla u dx \geq \alpha \int_{\Omega} |\nabla u|^2 dx.$$

L'inégalité de Poincaré nous permet de conclure que  $a$  est coercive. La formulation variationnelle (4.18) admet donc une solution unique. Reste à prouver que cette solution est également solution du problème aux limites (4.17). De (4.18) on déduit que la fonction  $\sigma = A\nabla u$  admet une divergence faible dans  $L^2(\Omega)$  (voir la Définition 4.2.6) et que  $-\operatorname{div}(A\nabla u) = f$  presque partout dans  $\Omega$ . De plus, comme  $u$  est un élément de  $H_0^1(\Omega)$ , on a aussi  $u = 0$  presque partout sur  $\partial\Omega$ .

**Exercice 4.2.12** Montrer l'existence et l'unicité de la solution de

$$\begin{cases} -\operatorname{div}(A\nabla u) + u = f & \text{dans } \Omega, \\ \frac{\partial u}{\partial n_A} = g & \text{sur } \partial\Omega, \end{cases}$$

avec  $f \in L^2(\Omega)$  et  $g \in L^2(\partial\Omega)$ . On rappelle que  $\partial u / \partial n_A = (A\nabla u) \cdot n$ .

**Correction.** La formulation variationnelle consiste à trouver  $u \in H^1(\Omega)$  tel que

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

avec

$$a(u, v) = \int_{\Omega} (A\nabla u \cdot \nabla v + uv) \, dx \quad \text{et} \quad L(v) = \int_{\Omega} f v \, dx + \int_{\partial\Omega} g v \, ds.$$

L'existence d'une solution à ce problème découle d'une application aisée du théorème de Lax-Milgram en suivant les mêmes étapes que dans l'Exercice 4.2.11 précédent. Pour montrer que la formulation variationnelle est bien équivalente au problème aux limites, nous allons, comme dans le cours (voir le Lemme 5.2.13), supposer que la solution est suffisamment régulière, c'est-à-dire appartient à l'espace  $H^2(\Omega)$ . (On peut effectivement généraliser le Lemme 5.2.13 et montrer que, si  $\Omega$  est régulier et  $a_{ij} \in C^1(\bar{\Omega})$  pour tout indice  $i, j$ , alors la solution  $u$  appartient à  $H^2(\Omega)$ .) 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  $\frac{\partial u}{\partial n_A} = g$  dans  $L^2(\partial\Omega)$ .

**Exercice 4.2.13** 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.** L'application  $v \rightarrow v^+$  de  $L^2(\Omega)$  dans  $L^2(\Omega)$  est évidemment continue car elle est 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  $H^1(\Omega)$ . On veut montrer que  $v_n^+$  converge vers  $v^+$  dans  $H^1(\Omega)$ . A cet effet, on va plus précisément prouver que de toute suite extraite de  $v_n^+$ , on peut extraire une sous-suite convergente dans  $H^1(\Omega)$  vers  $v^+$ , ce qui nous permettra de conclure à la convergence de toute la suite  $v_n^+$ . 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. En effet, d'après le Théorème de Rellich 4.3.21 on peut extraire une sous-suite de  $v_{n'}$  qui converge dans  $L^2(\Omega)$  et, d'après les propriétés de convergence dans  $L^2(\Omega)$ , on peut encore en extraire une sous-suite qui converge presque partout dans  $\Omega$ . D'après le Lemme 5.2.24,

$$\begin{aligned} \|\nabla v_{n''}^+ - \nabla v^+\|_{L^2(\Omega)} &= \|1_{v_{n''}>0} \nabla v_{n''} - 1_{v>0} \nabla v\|_{L^2(\Omega)} \\ &\leq \|1_{v_{n''}>0} (\nabla v_{n''} - \nabla v)\|_{L^2(\Omega)} + \|(1_{v_{n''}>0} - 1_{v>0}) \nabla v\|_{L^2(\Omega)} \\ &\leq \|\nabla v_{n''} - \nabla v\|_{L^2(\Omega)} + \|(1_{v_{n''}>0} - 1_{v>0}) \nabla v\|_{L^2(\Omega)}. \end{aligned}$$

Il est clair que le premier terme du second membre converge vers zéro. Par contre, le deuxième terme est plus délicat : expliquons pourquoi. On sait que si  $G$  est une fonction continue de  $\mathbb{R}$  dans  $\mathbb{R}$ , alors la convergence presque partout de  $v_{n''}$  vers  $v$  implique la convergence presque partout de  $G(v_{n''})$  vers  $G(v)$ . Mais malheureusement la fonction  $x \rightarrow 1_{x>0}$  n'est pas continue sur  $\mathbb{R}$  car elle est discontinue en 0. Par conséquent, il est faux de dire que  $1_{v_{n''}>0} > 0$  converge presque partout vers  $1_{v>0}$ . Comme contre-exemple (en dimension 1) prenons  $v \equiv 0$  et  $v_n(x) = n^{-1} \sin(nx)$  qui converge bien presque partout vers 0 mais par valeurs alternativement positives et négatives : la suite  $1_{v_n(x) > 0}$  oscille entre les valeurs 1 et 0 et ne converge pas presque partout.

On va néanmoins pouvoir passer à la limite dans le terme  $\|(1_{v_{n''}>0} - 1_{v>0})\nabla v\|_{L^2(\Omega)}$  en utilisant le fait que l'application  $x \rightarrow 1_{x>0}$  est continue sur  $\mathbb{R} \setminus \{0\}$  et que sur l'ensemble  $v^{-1}(0) := \{x \in \Omega : v(x) = 0\}$  on a  $\nabla v = 0$  presque partout (suivant l'indication de l'exercice, pas évidente à démontrer!). En effet, on a

$$1_{v_{n''}>0}(x) \rightarrow 1_{v>0}(x) \text{ pour presque tout } x \in \Omega \setminus v^{-1}(0).$$

Dans le deuxième membre de l'égalité

$$\|(1_{v_{n''}>0} - 1_{v>0})\nabla v\|_{L^2(\Omega)}^2 = \int_{\Omega \setminus v^{-1}(0)} (1_{v_{n''}>0} - 1_{v>0})^2 |\nabla v|^2 dx$$

on peut utiliser le théorème de convergence dominée de Lebesgue, donc

$$\|(1_{v_{n''}>0} - 1_{v>0})\nabla v\|_{L^2(\Omega)} \rightarrow 0 \text{ lorsque } n'' \rightarrow 0$$

et

$$\nabla v_{n''}^+ \rightarrow \nabla v^+ \text{ dans } L^2(\Omega).$$

On en déduit que toute la suite  $\nabla v_n^+$  converge vers  $\nabla v^+$ . 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  $L^2(\Omega)$ . L'application qui à  $v$  associe  $v^+$  est continue de  $H^1(\Omega)$  dans  $H^1(\Omega)$ .

**Exercice 4.3.1** 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

$$\begin{cases} -\operatorname{div}(2\mu e(u) + \lambda \operatorname{tr}(e(u)) \operatorname{Id}) = f & \text{dans } \Omega, \\ u = 0 & \text{sur } \partial\Omega, \end{cases} \quad (4.19)$$

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 (4.19)). On peut retrouver la continuité directement, en appliquant la formulation variationnelle à la fonction test  $v = u$ . On obtient

$$\int_{\Omega} (2\mu|e(u)|^2 + \lambda(\operatorname{div}u)^2) dx = \int_{\Omega} f \cdot u dx.$$

En combinant cette égalité à l'inégalité de Korn pour  $u \in H_0^1(\Omega)^N$

$$C \int_{\Omega} (|e(u)|^2 + (\operatorname{div}u)^2) dx \geq \|u\|_{H^1(\Omega)}^2$$

et à l'inégalité de Cauchy-Schwarz, on en déduit qu'il existe une constante  $C > 0$  telle que pour tout  $f \in L^2(\Omega)^N$ , on a

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

**Exercice 4.3.2** Soit  $\Omega$  un ouvert connexe de  $\mathbb{R}^N$ . Soit l'ensemble  $\mathcal{R}$  des "mouvements rigides" de  $\Omega$  défini par

$$\mathcal{R} = \left\{ v(x) = b + Mx \text{ avec } b \in \mathbb{R}^N, M = -M^t \text{ matrice antisymétrique} \right\}. \quad (4.20)$$

Montrer que  $v \in H^1(\Omega)^N$  vérifie  $e(v) = 0$  dans  $\Omega$  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  $\nabla v$ ,

$$W_{ij} = \frac{1}{2} \left( \frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i} \right).$$

Chaque fonction  $W_{ij}$  est un élément de  $L^2(\Omega)$ . De plus, pour toute fonction  $\varphi \in C_c^\infty(\Omega)$ , en effectuant diverses intégrations par parties, on établit que

$$\begin{aligned} \int_{\Omega} W_{ij} \frac{\partial \varphi}{\partial x_k} dx &= \frac{1}{2} \int_{\Omega} \left( \frac{\partial v_i}{\partial x_k} \frac{\partial \varphi}{\partial x_j} - \frac{\partial v_j}{\partial x_k} \frac{\partial \varphi}{\partial x_i} \right) dx \\ &= \int_{\Omega} \left( e_{ik}(v) \frac{\partial \varphi}{\partial x_j} - e_{jk}(v) \frac{\partial \varphi}{\partial x_i} \right) dx. \end{aligned}$$

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 constante  $M_{ij}$  telle que  $W_{ij}(x) = M_{ij}$  presque partout. De plus,

$W$  étant antisymétrique, la matrice  $M$  l'est également. Puisque  $e(v) = 0$ , on en déduit que

$$\nabla v = W + e(v) = M,$$

et donc

$$\nabla(v - Mx) = 0.$$

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 4.3.3** 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,

$$\int_{\Omega} 2\mu e(u) \cdot e(v) dx + \int_{\Omega} \lambda \operatorname{div} u \operatorname{div} v dx = \int_{\Omega} f \cdot v dx + \int_{\partial\Omega_N} g \cdot v ds \quad \forall v \in V, \quad (4.21)$$

si et seulement si  $u$  réalise le minimum sur  $V$  de l'énergie

$$J(v) = \frac{1}{2} \int_{\Omega} (2\mu |e(v)|^2 + \lambda |\operatorname{div} v|^2) dx - \int_{\Omega} f \cdot v dx - \int_{\partial\Omega_N} g \cdot v ds. \quad (4.22)$$

(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 symétrique

$$a(u, v) = \int_{\Omega} (2\mu e(u) \cdot e(v) + \lambda (\operatorname{div} u)(\operatorname{div} v)) dx$$

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  $V$ . Plus précisément, on a dans ce cas

$$J(u - v) = J(u) - a(u, v) + L(v) + a(v, v)/2.$$

Ainsi,  $u$  est un minimiseur de  $J$  sur  $V$ , si et seulement si pour tout  $v \in V$ ,

$$a(u, v) - L(v) \leq a(v, v)/2. \quad (4.23)$$

Dans (4.23) on remplace  $v$  par  $\alpha v$  avec  $\alpha > 0$ , et on divise par  $\alpha$  pour obtenir

$$a(u, v) - L(v) \leq \alpha a(v, v)/2 \quad \text{pour tout } \alpha > 0.$$

En faisant tendre  $\alpha$  vers 0, puis en écrivant la même inégalité pour  $-v$ , on en déduit qu'une condition nécessaire pour que (4.23) soit vérifié est

$$a(u, v) - L(v) = 0 \quad \text{pour tout } v \in V,$$

ce qui n'est rien d'autre que la formulation variationnelle (4.21). De plus comme  $a(v, v) \geq 0$ , cette condition est nécessaire et suffisante. On a donc établi que  $u$  est un minimiseur de  $J$  sur  $V$  si et seulement si  $u$  est solution de la formulation variationnelle (4.21).

**Exercice 4.3.4** Soit  $\Omega$  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)

$$\int_{\Omega} f \cdot (Mx + b) dx + \int_{\partial\Omega} g \cdot (Mx + b) ds = 0 \quad \forall b \in \mathbb{R}^N, \forall M = -M^t \in \mathbb{R}^{N \times N}$$

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.** Supposons que  $u$  soit solution du système de l'élasticité avec conditions aux bords de Neumann. En multipliant l'équation vérifiée par  $u$  dans  $\Omega$  par une fonction test  $v \in H^1(\Omega)^N$ , on obtient suite à une intégration par parties que

$$\int_{\Omega} (2\mu e(u) \cdot e(v) + \lambda(\operatorname{div}u)(\operatorname{div}v)) dx = \int_{\Omega} f \cdot v dx + \int_{\partial\Omega} g \cdot v ds.$$

En appliquant cette équation à un élément  $v$  de la forme  $v(x) = Mx + b$ , où  $b \in \mathbb{R}^N$  et  $M$  est une matrice antisymétrique, on en déduit, puisque  $e(v) = 0$  et  $\operatorname{div}v = 0$ , que

$$\int_{\Omega} f \cdot (Mx + b) dx + \int_{\partial\Omega} g \cdot (Mx + b) ds = 0,$$

qui est donc une condition nécessaire d'existence de solution. 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

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

Notons que  $a(u, v)$  et  $L(v)$  sont toutes deux correctement définies. Leurs valeurs sont indépendantes des représentants  $u$  et  $v$  choisis dans  $H^1(\Omega)^N$ . En effet, soit  $u_1$  et  $u_2$  (resp.  $v_1$  et  $v_2$ ) deux éléments de  $H^1(\Omega)^N$ , représentants de  $u$  (resp.  $v$ ) dans  $V$ . Il existe  $M$  et  $N$  matrices antisymétriques  $N \times N$ ,  $b$  et  $c$  vecteurs de  $\mathbb{R}^N$  tels que  $u_1 = u_2 + Mx + b$  et  $v_1 = v_2 + Nx + c$ . On a alors  $e(u_1) = e(u_2)$  et  $e(v_1) = e(v_2)$ . Ainsi  $a(u_1, v_1) = a(u_2, v_2)$ . De plus, d'après la condition de compatibilité,  $L(v_1) = L(v_2)$ . Afin d'appliquer le théorème de Lax-Migran, seule la coercivité de la forme bilinéaire n'est pas tout à fait évidente à établir. Il s'agit de prouver qu'il existe une constante  $C$  telle que

$$\|v\|_V^2 \leq Ca(v, v) \text{ pour tout } v \in V, \quad (4.24)$$



où

$$\|v\|_V = \inf_{M,b} \|v + Mx + b\|_{H^1(\Omega)}, \text{ avec } M \text{ matrice antisymétrique et } b \in \mathbb{R}^N.$$

On démontre cette inégalité par contradiction (comme dans le cours pour l'inégalité de Poincaré, **p. 94**). Supposons que la relation (4.24) soit fautive pour tout  $C$ . Dans ce cas, il existe une suite  $v_n$  d'éléments de  $V$  telle que

$$1 = \|v_n\|_V^2 \geq n a(v_n, v_n).$$

Rappelons (voir le cours, **p. 140**) qu'il existe  $\nu > 0$  tel que pour tout  $v \in H^1(\Omega)^N$ ,

$$a(v, v) \geq \nu \|e(v)\|_{L^2(\Omega)}^2.$$

Ainsi,

$$1 = \|v_n\|_V^2 \geq \nu n \|e(v_n)\|_{L^2(\Omega)}^2$$

Par abus de langage on note également  $v_n$  l'élément de  $H^1(\Omega)^N$  tel que  $\|v_n\|_{H^1(\Omega)} = \|v_n\|_V$  (on confond un élément de  $V$  avec un représentant particulier). D'après le théorème de Rellich, la suite  $v_n$  étant bornée dans  $H^1(\Omega)^N$ , il existe une sous-suite  $v_{n'}$  convergente dans  $L^2(\Omega)^N$ . On rappelle que d'après l'inégalité de Korn,  $v \mapsto \|v\|_{L^2(\Omega)^N} + \|e(v)\|_{L^2(\Omega)^{N \times N}}$  est une norme équivalente à la norme de  $H^1(\Omega)^N$ . Comme  $e(v_{n'})$  tend vers zéro dans  $L^2(\Omega)^{N \times N}$ ,  $e(v_{n'})$  est une suite de Cauchy dans  $L^2(\Omega)^{N \times N}$ . De plus  $v_{n'}$  est une suite de Cauchy dans  $L^2(\Omega)^N$ . Ainsi, la suite  $v_{n'}$  est de Cauchy dans  $H^1(\Omega)^N$ . En conséquence,  $v_{n'}$  converge dans  $H^1(\Omega)^N$  vers un élément  $v$  tel que  $e(v) = 0$ . D'après l'exercice précédent, il existe  $M$  matrice antisymétrique et  $b \in \mathbb{R}^N$  tels que  $v(x) = Mx + b$ . En d'autres termes,  $v = 0$  dans  $V$ . D'autre part, la convergence dans  $H^1(\Omega)^N$  implique la convergence dans  $V$ . Comme  $\|v_{n'}\|_V = 1$  on a donc  $\|v\|_V = 1$ , ce qui est contradictoire avec le fait que  $v = 0$ . La forme bilinéaire  $a$  est donc coercive sur  $V$  et la formulation variationnelle admet donc une solution unique.

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  $\sigma \cdot n$ , il serait nécessaire de montrer que  $u$  est en fait un élément de  $H^2(\Omega)^N$  (ce qu'on a admis pour le Laplacien). A défaut, on peut toujours utiliser le fait que chaque ligne de  $\sigma$  est un élément de  $H(\text{div})$  et utiliser la définition faible de la trace de la composante normale de  $\sigma$  sur le bord comme élément de  $H^{-1/2}(\partial\Omega)$  (voir le Théorème 4.4.7)

**Exercice 4.3.5** On suppose que  $\Omega$  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

$$\begin{cases} -\mu\Delta u - (\mu + \lambda)\nabla(\text{div}u) = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega. \end{cases} \quad (4.25)$$

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  $2\mu + \lambda > 0$ .

**Correction.** La formulation variationnelle consiste à trouver  $u \in H_0^1(\Omega)^N$  tel que

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

où

$$a(u, v) = \int_{\Omega} (\mu \nabla u \cdot \nabla v + (\lambda + \mu)(\operatorname{div} u)(\operatorname{div} v)) dx$$

et

$$L(v) = \int f \cdot v dx.$$

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

$$\int_{\Omega} (\operatorname{div} u)^2 dx = \int_{\Omega} \sum_{i,j=1}^N \frac{\partial u_i}{\partial x_i} \frac{\partial u_j}{\partial x_j} dx.$$

Par intégration par parties, il vient

$$\begin{aligned} \int_{\Omega} (\operatorname{div} u)^2 dx &= \int_{\Omega} \sum_{i,j=1}^N \frac{\partial u_i}{\partial x_j} \frac{\partial u_j}{\partial x_i} dx = \int_{\Omega} \nabla u \cdot (\nabla u)^t dx \\ &\leq \int_{\Omega} |\nabla u| |(\nabla u)^t| dx = \int_{\Omega} |\nabla u|^2 dx. \end{aligned}$$

Ainsi,

$$a(u, u) \geq \int_{\Omega} (\mu + \min(0, \lambda + \mu)) |\nabla u|^2 dx = \min(\mu, \lambda + 2\mu) \int_{\Omega} |\nabla u|^2 dx.$$

La forme bilinéaire  $a(\cdot, \cdot)$  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 4.3.6** Vérifier l'équivalence de (4.25) et (4.19) si  $\lambda$  et  $\mu$  sont constants. Montrer que (4.25) et (4.19) ne sont plus équivalents si  $\lambda$  et  $\mu$  sont des fonctions (régulières), même si on remplace l'équation vectorielle de (4.25) par

$$-\operatorname{div}(\mu \nabla u) - \nabla((\mu + \lambda)\operatorname{div} u) = f \text{ dans } \Omega.$$

**Correction.** Soit  $u$  la solution du problème variationnel associé à (4.19) : pour tout  $v \in C_c^\infty(\Omega)^N$ , on a

$$\int_{\Omega} \mu \sum_{i,j=1}^N \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \frac{\partial v_i}{\partial x_j} dx + \int_{\Omega} \lambda (\operatorname{div} u)(\operatorname{div} v) dx = \int_{\Omega} f \cdot v dx.$$

Or par intégration par parties,

$$\begin{aligned}
\int_{\Omega} \mu \frac{\partial u_j}{\partial x_i} \frac{\partial v_i}{\partial x_j} dx &= - \int_{\Omega} u_j \frac{\partial}{\partial x_i} \left( \mu \frac{\partial v_i}{\partial x_j} \right) dx \\
&= - \int_{\Omega} \mu u_j \frac{\partial^2 v_i}{\partial x_j \partial x_i} dx - \int_{\Omega} u_j \frac{\partial \mu}{\partial x_i} \frac{\partial v_i}{\partial x_j} dx \\
&= \int_{\Omega} \frac{\partial(\mu u_j)}{\partial x_j} \frac{\partial v_i}{\partial x_i} dx - \int_{\Omega} u_j \frac{\partial \mu}{\partial x_i} \frac{\partial v_i}{\partial x_j} dx \\
&= \int_{\Omega} \mu \frac{\partial u_j}{\partial x_j} \frac{\partial v_i}{\partial x_i} dx + \int_{\Omega} u_j \left( \frac{\partial \mu}{\partial x_j} \frac{\partial v_i}{\partial x_i} - \frac{\partial \mu}{\partial x_i} \frac{\partial v_i}{\partial x_j} \right) dx.
\end{aligned}$$

Ainsi,

$$\begin{aligned}
\int_{\Omega} \mu \sum_{i,j=1}^N \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \frac{\partial v_i}{\partial x_j} dx + \int_{\Omega} \lambda (\operatorname{div} u)(\operatorname{div} v) dx = \\
\int_{\Omega} \left( \mu \nabla u \cdot \nabla v + (\lambda + \mu)(\operatorname{div} u)(\operatorname{div} v) \right) dx + \int_{\Omega} u \cdot \left( (\operatorname{div} v) \nabla \mu - (\nabla v)^t \nabla \mu \right) dx.
\end{aligned}$$

Si  $\mu$  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  $v \in H_0^1(\Omega)^N$ ,

$$\int_{\Omega} \left( \mu \nabla u \cdot \nabla v + (\lambda + \mu)(\operatorname{div} u)(\operatorname{div} v) \right) dx = \int_{\Omega} f \cdot v dx,$$

qui est équivalent au problème aux limites consistant à trouver  $u$  tel que

$$\begin{cases} -\mu \Delta u - \nabla((\mu + \lambda) \operatorname{div} u) = f & \text{dans } \Omega, \\ u = 0 & \text{sur } \partial\Omega. \end{cases}$$

Si de plus  $\lambda$  est constant, on retrouve le problème aux limites (4.25). Enfin, si l'un des coefficient  $\mu$  ou  $\lambda$  n'est pas constant, (4.19) et (4.25) ne sont en général pas équivalents.

**Exercice 4.3.7** 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  $\Omega$  de longueur  $L > 0$  et de section  $\omega$ , où  $\omega$  est un ouvert borné connexe régulier de  $\mathbb{R}^{N-1}$  (les coefficients de Lamé  $\lambda$  et  $\mu$  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

$$\begin{cases} -\operatorname{div}(\sigma) = 0 & \text{dans } \Omega \\ \sigma n = g & \text{sur } \partial\omega \times (0, L) \\ u' = 0 & \text{sur } \omega \times \{0, L\} \\ (\sigma n) \cdot n = 0 & \text{sur } \omega \times \{0, L\} \end{cases}, \quad (4.26)$$

avec

$$\sigma = 2\mu e(u) + \lambda \operatorname{tr}(e(u)) \operatorname{Id},$$

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 (4.26) est donnée par  $u = (0, \dots, 0, u_N)$  où  $u_N(x')$  est la solution du Laplacien suivant

$$\begin{cases} -\Delta' u_N = 0 & \text{dans } \omega \\ \mu \frac{\partial u_N}{\partial n} = g_N & \text{sur } \partial\omega \end{cases} \quad (4.27)$$

où  $\Delta'$  est le Laplacien dans la variable  $x' \in \mathbb{R}^{N-1}$ .

**Correction.** Soit  $u_N$  la solution de (4.27). On pose  $u = (0, \dots, 0, u_N)$ . Pour tout  $i$  et  $j$  tels que  $1 \leq i, j < N$ , on a

$$\begin{aligned} e_{ij}(u) &= 0 \\ e_{iN}(u) &= e_{Ni}(u) = \frac{1}{2} \frac{\partial u_N}{\partial x_i} \\ e_{NN}(u) &= 0. \end{aligned}$$

En particulier,  $\text{tr}(e(u)) = 0$ . On en déduit que,

$$-\text{div} \sigma = -\text{div}(2\mu e(u) + \lambda \text{tr}(e(u)) \text{Id}) = -\mu(0, \dots, 0, \Delta' u_N) = 0.$$

De plus,

$$\sigma e_N = 2\mu e(u) e_N = \mu(\nabla' u_N, 0) \quad \text{et} \quad \sigma(u) e_N \cdot e_N = 0.$$

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

$$\sigma n = 2\mu \left( \sum_{k=1}^{N-1} e_{jk}(u) n_k \right)_{1 \leq j \leq N} = 2\mu(0, \dots, 0, 1/2 \nabla' u_N \cdot n') = (0, \dots, 0, g_N).$$

Ainsi,  $u$  est bien l'unique solution du problème aux limites (4.26).

**Exercice 4.3.8** Généraliser l'Exercice 5.3.7 au cas d'une condition aux limites latérale du type

$$u' = 0 \quad \text{et} \quad (\sigma n) \cdot e_N = g_N \quad \text{sur } \partial\omega \times (0, L).$$

**Correction.** La solution construite dans l'exercice précédent vérifie également ces conditions aux limites.

**Exercice 4.3.9** A l'aide de l'approche variationnelle démontrer l'existence et l'unicité de la solution de l'équation des plaques

$$\begin{cases} \Delta(\Delta u) = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega \\ \frac{\partial u}{\partial n} = 0 & \text{sur } \partial\Omega \end{cases} \quad (4.28)$$

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 (4.28) (voir l'Exercice 3.2.4) consiste à déterminer  $u \in H_0^2(\Omega)$  tel que

$$a(u, v) = L(v) \quad \text{pour tout } v \in H_0^2(\Omega), \quad (4.29)$$

où

$$a(u, v) = \int_{\Omega} \Delta u \Delta v dx \quad \text{et} \quad L(v) = \int_{\Omega} f v dx.$$

Afin d'appliquer le Théorème de Lax-Milgram, la seule hypothèse non trivialement vérifiée est la coercivité de la forme bilinéaire  $a(\cdot, \cdot)$ . Or pour tout  $u \in H_0^2(\Omega)$ , on établit suite à deux intégrations par parties successives que

$$a(u, u) = \sum_{i,j=1}^N \int_{\Omega} \left| \frac{\partial^2 u}{\partial x_i \partial x_j}(x) \right|^2 dx = \|\nabla^2 u\|_{L^2(\Omega)}^2.$$

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  $H_0^2(\Omega)$ ,

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

Par conséquent, il existe  $C > 0$  tel que pour tout  $u \in H_0^2(\Omega)$ ,

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

La coercivité de la forme bilinéaire  $a(\cdot, \cdot)$  est donc établie et il existe une unique solution au problème variationnel (4.29).

Reste à établir que la solution du problème variationnel est solution du problème aux limites. Soit  $\omega$  un ouvert inclus dans un compact de  $\Omega$ . Il existe  $\theta \in C_c^\infty(\Omega)$  telle que  $\theta = 1$  sur  $\omega$ . Pour toute fonction  $v \in C_c^\infty(\Omega)$  de support inclus dans  $\omega$ ,

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

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 parties successives sur le membre de gauche de l'équation précédente. On en déduit que

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

Cette équation étant vérifiée pour toute fonction  $v \in C_c^\infty(\Omega)$  de support inclus dans  $\omega$ , on en déduit que pour presque tout  $x \in \omega$ ,

$$\Delta(\Delta u)(x) = f(x).$$

Cette relation reste valable pour presque tout  $x \in \Omega$  : il suffit de considérer une suite  $\omega_n$  de compacts tels que  $\cup_n \omega_n = \Omega$ . 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 4.3.10** Soit  $V$  l'espace des champs de vitesse à divergence nulle défini par

$$V = \left\{ v \in H_0^1(\Omega)^N : \operatorname{div} v = \sum_{i=1}^n \frac{\partial v_i}{\partial x_i} = 0 \text{ p.p. dans } \Omega \right\}.$$

Soit  $J(v)$  l'énergie définie pour  $v \in V$  par

$$J(v) = \frac{1}{2} \int_{\Omega} \mu |\nabla v|^2 dx - \int_{\Omega} f \cdot v dx. \quad (4.30)$$

Soit  $u \in V$  la solution unique de la formulation variationnelle

$$\int_{\Omega} \mu \nabla u \cdot \nabla v dx = \int_{\Omega} f \cdot v dx \quad \forall v \in V. \quad (4.31)$$

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 (4.31).

**Correction.** Il suffit d'appliquer la Proposition 3.3.4 à la formulation variationnelle (4.31) pour conclure. A défaut, on peut prouver l'équivalence entre le problème de minimisation de l'énergie et la formulation variationnelle à la main. Notons que pour tout éléments  $u$  et  $v$  de  $V$ ,

$$J(u - v) = J(u) - a(u, v) + L(v) + a(v, v)/2,$$

où  $a(\cdot, \cdot)$  est la forme bilinéaire continue, définie sur  $V \times V$  par

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

et  $L$  est la forme linéaire continue sur  $V$  définie par

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

Ainsi,  $u$  est un minimiseur de  $J$  sur  $V$  si et seulement si

$$a(u, v) - L(v) \leq a(v, v)/2 \quad \text{pour tout } v \in V.$$

Dans cette inégalité on remplace  $v$  par  $\alpha v$  avec  $\alpha > 0$ , et on divise par  $\alpha$  pour obtenir

$$a(u, v) - L(v) \leq \alpha a(v, v)/2 \quad \text{pour tout } \alpha > 0.$$

En faisant tendre  $\alpha$  vers 0, puis en écrivant la même inégalité pour  $-v$ , on en déduit qu'une condition nécessaire pour que  $u$  soit un minimiseur de  $J$  sur  $V$  est

$$a(u, v) - L(v) = 0 \quad \text{pour tout } v \in V,$$

ce qui n'est rien d'autre que la formulation variationnelle (4.31). De plus comme  $a(v, v) \geq 0$ , cette condition est nécessaire et suffisante. On a donc établi que  $u$  est un minimiseur de  $J$  sur  $V$  si et seulement si  $u$  est solution de la formulation variationnelle (4.31).

**Exercice 4.3.11** 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  $\omega$  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

$$\begin{cases} \nabla p - \mu \Delta u = 0 & \text{dans } \Omega \\ \operatorname{div} u = 0 & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\omega \times (0, L) \\ pn - \mu \frac{\partial u}{\partial n} = p_0 n & \text{sur } \omega \times \{0\} \\ pn - \mu \frac{\partial u}{\partial n} = p_L n & \text{sur } \omega \times \{L\} \end{cases} \quad (4.32)$$

où  $p_0$  et  $p_L$  sont deux pressions constantes. Montrer que la solution unique de (4.32) 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  $\Delta'$  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  $u_N(x')$  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 (4.32). 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, comme  $u_n$  ne dépend pas de  $x_n$ ,

$$\operatorname{div} u = \frac{\partial u_N}{\partial x_N} = 0.$$

Enfin, comme  $\partial u / \partial n = 0$  sur  $\omega \times \{0, 1\}$  et

$$\begin{cases} p = p_0 & \text{sur } \omega \times \{0\}, \\ p = p_1 & \text{sur } \omega \times \{L\}, \end{cases}$$

les conditions aux limites imposées aux extrémités du profil sont également vérifiées.

**Exercice 4.3.12** Généraliser l'Exercice **5.3.11** au cas des équations de Navier-Stokes

$$\begin{cases} (u \cdot \nabla)u + \nabla p - \mu \Delta u = f & \text{dans } \Omega \\ \operatorname{div} u = 0 & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega. \end{cases} \quad (4.33)$$

**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 5

# MÉTHODE DES ÉLÉMENTS FINIS

**Exercice 5.2.1** Appliquer la méthode des éléments finis  $\mathbb{P}_1$  au problème

$$\begin{cases} -u'' = f \text{ dans } ]0, 1[, \\ u(0) = \alpha, \quad u(1) = \beta. \end{cases}$$

Vérifier que les conditions aux limites de Dirichlet non-homogènes apparaissent dans le second membre du système linéaire qui en découle.

**Correction.** On introduit les points du maillage  $x_i = i/(n+1)$  et l'espace discret

$$V_h := \{v_h \in C^0([0, 1]; \mathbb{R}) : v_{|[x_i, x_{i+1}]} \in \mathbb{P}_1 \text{ pour tout } i \in \{0, \dots, n\}\}$$

qui est un sous-espace vectoriel de dimension finie de  $H^1(0, 1)$ . On définit un autre sous-espace vectoriel  $V_{0h} := V_h \cap H_0^1(0, 1)$  ainsi que le sous-espace affine

$$V_h^{\alpha\beta} := \{v_h \in V_h : v_h(0) = \alpha, \quad v_h(1) = \beta\}.$$

La formulation variationnelle, issue de l'utilisation des éléments finis  $\mathbb{P}_1$ , consiste à trouver  $u_h \in V_h^{\alpha\beta}$  tel que

$$\int_0^1 u_h' v_h' dx = \int_0^1 f v_h dx \text{ pour toute fonction } v_h \in V_{0h}.$$

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  $\phi_j$  comme fonction test, on obtient, à l'aide de la formulation variationnelle, que pour tout  $0 < j < n+1$ ,

$$\sum_{i=0}^{n+1} (u_h)_i \int_0^1 \phi_i' \phi_j' dx = \int_0^1 f \phi_j dx,$$

où  $(u_h)_i$  sont les coordonnées de  $u_h$  dans la base  $(\phi_i)$ . 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

$$\mathcal{K}_h = h^{-1} \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix}. \quad (5.1)$$

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 5.2.2** On reprend le problème de Neumann

$$\begin{cases} -u'' + au = f \text{ dans } ]0, 1[, \\ u'(0) = \alpha, u'(1) = \beta, \end{cases} \quad (5.2)$$

en supposant que la fonction  $a(x) = 0$  dans  $]0, 1[$ . 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.** D'après les résultats du cours suivant (**6.13**), le système linéaire obtenu en considérant  $a = 0$  est

$$\mathcal{K}_h U_h = b_h, \quad (5.3)$$

où la matrice de rigidité (de taille  $(n+2)$ , correspondant aux  $(n+2)$  points du maillage  $x_i = i/(n+1)$ ) est donnée par

$$\mathcal{K}_h = \frac{1}{h} \begin{pmatrix} 1 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 1 \end{pmatrix}$$

et  $b_h$  est défini comme dans le cas  $a \neq 0$  par

$$(b_h)_i = \int_{x_{i-1}}^{x_{i+1}} f(x) \phi_i(x) dx \quad \text{pour tout } 1 \leq i \leq n,$$

$$(b_h)_0 = \int_0^{x_1} f(x) \phi_0(x) dx - \alpha,$$

$$(b_h)_{n+1} = \int_{x_n}^1 f(x) \phi_{n+1}(x) dx + \beta.$$

La matrice  $\mathcal{K}_h$  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 (5.3) 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

$$\int_0^1 f(x) dx + \beta - \alpha = \sum_{i=0}^{n+1} \int_0^1 f(x) \phi_i(x) dx + \beta - \alpha = \sum_{i=0}^{n+1} (b_h)_i = 0.$$

On retrouve ainsi, au niveau discret, la condition de compatibilité sur les données pour obtenir l'existence d'une solution au problème (5.2), comme il est démontré au Théorème 5.2.18. Par ailleurs, si on utilise une formule de quadrature

$$\int_0^1 \psi(x) dx \approx I(\psi(x_j)),$$

qui soit linéaire par rapport à  $\psi$ , alors

$$\begin{aligned} \sum_{i=0}^{n+1} (b_h)_i &= \beta - \alpha + \sum_{i=0}^{n+1} I(f(x_j) \phi_i(x_j)) = \beta - \alpha + I\left(f(x_j) \sum_{i=0}^{n+1} \phi_i(x_j)\right) \\ &= \beta - \alpha + I(f(x_j)) \approx \beta - \alpha + \int_0^1 f(x) dx. \end{aligned}$$

Donc, modulo l'approximation de la formule de quadrature, on retrouve la même condition de compatibilité sur les données pour pouvoir résoudre le système linéaire (5.3).

**Exercice 5.2.3** Appliquer la méthode des différences finies (voir le Chapitre 2) au problème de Dirichlet

$$\begin{cases} -u'' = f & \text{dans } ]0, 1[, \\ u(0) = u(1) = 0. \end{cases} \quad (5.4)$$

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) que celle issue de la méthode des éléments finis  $\mathbb{P}_1$  mais avec un second membre  $b_h$  différent. Même question pour le problème de Neumann (5.2).

**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

$$\begin{cases} -\frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} = f(x_i) & \text{pour tout } 1 \leq i \leq n, \\ u_0 = 0, \quad u_{n+1} = 0. \end{cases}$$

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  $n$

$$\mathcal{K}_h = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix} \quad \text{et} \quad b_h = \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix}.$$

La matrice  $\mathcal{K}_h$  diffère de la matrice obtenue par la méthode des éléments 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  $(u_i)_{-1 \leq i \leq n+2}$  tel que

$$\begin{cases} -\frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} + a(x_i)u_i = f(x_i) & \text{pour tout } 0 \leq i \leq n+1, \\ \frac{u_1 - u_{-1}}{2h} = \alpha, \quad \frac{u_{n+2} - u_n}{2h} = \beta, \end{cases}$$

où les noeuds “fictifs”  $x_{-1}$  et  $x_{n+2}$  ont été introduits afin que les conditions aux limites soient discrétisées à l’ordre 2. Si on élimine du système linéaire final les degrés de liberté artificiellement introduits, on obtient les matrices suivantes (de taille  $n+2$ )

$$\mathcal{K}_h = \frac{1}{h^2} \begin{pmatrix} 1 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 1 \end{pmatrix} + \begin{pmatrix} a(x_0)/2 & 0 & \cdots & & 0 \\ 0 & a(x_1) & & & \\ \vdots & & \ddots & & \vdots \\ & & & a(x_n) & 0 \\ 0 & & \cdots & 0 & a(x_{n+1})/2 \end{pmatrix}$$

et  $b_h = (-\frac{\alpha}{h} + f(x_0)/2, f(x_1), \dots, f(x_n), \frac{\beta}{h} + f(x_{n+1})/2)^T$ . 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

$$\mathcal{K}_h^{EF} = h \mathcal{K}_h \quad \text{et} \quad b_h^{EF} = h b_h.$$

**Exercice 5.2.4** 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

si l'extrémité  $x_0$  est fixée), l'espace de minimisation est l'ensemble des  $u$  tels que  $u_0 = 0$ . Si aucune extrémité n'est fixée, l'espace de minimisation n'a pas à être contraint. Par contre, dans ce dernier cas, l'existence d'un minimiseur n'est assurée que si la condition de compatibilité

$$\sum_{j=0}^{n+1} f_j = 0$$

est vérifiée. En effet, dans le cas contraire, en notant  $\mathbb{I} = (1, \dots, 1)$ , alors

$$J(\alpha \mathbb{I}) = -\alpha \sum_{j=0}^{n+1} f_j$$

qui tend vers  $-\infty$  si  $\alpha$  tends vers l'infini avec le signe de  $\sum_{j=0}^{n+1} f_j$ .

Il y a une forte similitude entre le problème obtenu et la résolution de l'équation

$$-k\Delta u = f$$

par la méthode des 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  $U_h = (U_h^0, \dots, U_h^{n+1})$  les coordonnées de  $u_h$  dans la base classique de  $V_h$ . On a alors

$$I(u_h) = \sum_{j=0}^n \frac{k}{2} \frac{(U_h^{j+1} - U_h^j)^2}{\Delta x} - \sum_{j=0}^{n+1} \left( \int_0^1 f(x)\phi_j(x) dx \right) U_h^j.$$

Si on utilise la formule des trapèzes afin d'évaluer l'intégrale apparaissant dans la définition de  $I$ , on obtient

$$I(u_h) = \sum_{j=0}^n \frac{k}{2} \frac{(U_h^{j+1} - U_h^j)^2}{\Delta x} - \sum_{j=0}^{n+1} f(x_j)\phi_j U_h^j \Delta x.$$

En posant  $f_j = (\Delta x)^2 f(x_j)$ , on retrouve l'expression  $J$  à un coefficient  $\Delta x$  près.

**Exercice 5.2.5** Démontrer l'équivalent du Théorème 6.2.6 de convergence de la méthode des éléments finis en dimension 1 appliquée au problème de diffusion (5.2) avec conditions aux limites de type Neumann.

**Correction.** La démonstration est identique mot pour mot à celle effectuée dans le cas de conditions aux limites de Dirichlet. Plus précisément, en notant  $x_j = j/(n+1)$  les points du maillage, la formulation éléments finis consiste à déterminer

$$u_h \in V_h := \{v \in C([0, 1]) \text{ tel que } v|_{[x_j, x_{j+1}]} \in \mathbb{P}_1 \text{ pour tout } 0 \leq j \leq n\},$$

tel que, pour tout  $v_h \in V_h$ , on ait

$$A(u_h, v_h) = L(v_h), \quad (5.5)$$

où

$$A(u, v) = \int_0^1 (u'v' + auv) dx$$

et

$$L(v) = \int_0^1 fv dx - \alpha v(0) + \beta v(1).$$

Comme la fonction  $a$  vérifie  $a(x) \geq a_0 > 0$  dans  $(0, 1)$ , la forme bilinéaire  $A$  est continue et coercive sur  $H^1(0, 1)$ . La forme linéaire  $L$  est continue sur  $H^1(0, 1)$  (les fonctions  $H^1(0, 1)$  s'injectant de manière continue dans  $C([0, 1])$ , cf. Lemme 4.3.3). D'après le Lemme 6.1.1, l'approximation de Galerkin ci-dessus admet une solution unique et d'après le Lemme de Céa 6.1.2, il existe  $C > 0$  indépendant de  $h$  tel que

$$\|u - u_h\|_{H^1(0,1)} \leq C \inf_{v_h \in V_h} \|u - v_h\|_{H^1(0,1)}.$$

En choisissant  $v_h = r_h u$ , on  $r_h$  est l'opérateur d'interpolation de  $H^1(0, 1)$  dans  $V_h$  de la Définition 6.2.4, il vient

$$\|u - u_h\|_{H^1(0,1)} \leq C \|u - r_h u\|_{H^1(0,1)}.$$

D'après le Lemme 6.2.5, on a

$$\lim_{h \rightarrow 0} \|u - r_h u\|_{H^1(0,1)} = 0$$

et si  $u \in H^2(0, 1)$ , alors il existe une constante  $C$  indépendante de  $h$  telle que

$$\|u - r_h u\|_{H^1(0,1)} \leq Ch \|u''\|_{L^2(0,1)}.$$

Il s'en suit que

$$\lim_{h \rightarrow 0} \|u - u_h\|_{H^1(0,1)} = 0$$

et que si  $u \in H^2(0, 1)$ , il existe  $C$  indépendant de  $h$  tel que

$$\|u - u_h\|_{H^1(0,1)} \leq Ch \|u''\|_{L^2(0,1)}.$$

**Exercice 5.2.6** En généralisant les arguments précédents, démontrer le Théorème 6.2.14 de convergence de la méthode des éléments finis  $\mathbb{P}_2$  à la résolution du problème aux limites

$$\begin{cases} -u'' = f \text{ dans } ]0, 1[ \\ u(0) = u(1) = 0, \end{cases}$$



**Correction.** D'après le Lemme de Céa **6.1.2**, il existe une constante  $C$  indépendante de  $h$  telle que

$$\|u - u_h\|_{H^1(0,1)} \leq C \inf_{v_h \in V_{0h}} \|u - v_h\|_{H^1(0,1)}, \quad (5.6)$$

où  $V_{0h}$  est l'espace des éléments finis  $\mathbb{P}_2$  nuls aux bords. Afin de majorer le terme de droite, on introduit l'opérateur d'interpolation de  $H_0^1$  sur  $V_{0h}$  qui à  $v$  associe

$$r_h v = \sum_{j=1}^n v(x_j) \psi_j + \sum_{j=0}^n v(x_{j+1/2}) \psi_{j+1/2},$$

où  $\psi_j$  et  $\psi_{j+1/2}$  sont les fonctions de bases de l'espace des éléments finis  $\mathbb{P}_2$  (voir cours, **p. 166**). L'introduction de cette opérateur d'interpolation nous permet de majorer le terme de droite de (5.6) par la norme  $H^1$  de  $(u - r_h u)$ . Il ne nous reste donc plus qu'à estimer ce dernier terme.

Dans le cas  $h = 1$ , il existe une constante  $C_1$  telle que, pour tout  $v \in H^3(0,1)$ ,

$$\|(r_1 v - v)'\|_{L^2(0,1)} \leq C_1 \|v'''\|_{L^2(0,1)}. \quad (5.7)$$

Afin d'établir cette inégalité, on peut effectuer un raisonnement par l'absurde dans l'esprit de la démonstration de l'inégalité de Poincaré (voir cours, **p. 94**). Supposons que l'inégalité (5.7) soit violée pour toute constante  $C_1$ . Il existe une suite  $v_n$  d'éléments de  $H^3(0,1)$  telle que pour tout  $n \in \mathbb{N}$ ,

$$\|(r_1 v_n - v_n)'\|_{L^2(0,1)} > n \|v_n'''\|_{L^2(0,1)}. \quad (5.8)$$

Quitte à modifier  $v_n$  par l'ajout d'un polynôme de degré 2, on peut supposer que

$$\int_0^1 v_n dx = \int_0^1 v_n' dx = \int_0^1 v_n'' dx = 0$$

sans changer l'inégalité (5.8) car tout  $p \in \mathbb{P}_2$  vérifie  $p''' = 0$  et  $r_1 p - p = 0$ . De plus, quitte à multiplier  $v_n$  par une constante, on peut également supposer que

$$1 = \|(r_1 v_n - v_n)'\|_{L^2(0,1)}. \quad (5.9)$$

Par application successive de l'inégalité de Poincaré-Wirtinger (**5.28**), on en déduit que  $v_n''$ ,  $v_n'$  et  $v_n$  sont bornés dans  $L^2(0,1)$  et donc que  $v_n$  est borné dans  $H^3(0,1)$ . D'après le Théorème de Rellich **4.3.21** et quitte à extraire une sous-suite de  $v_n$ , on peut donc supposer que la suite  $v_n$  est convergente dans  $H^2(0,1)$ . De plus, comme  $v_n'''$  converge vers zéro dans  $L^2(0,1)$ , la suite  $v_n$  est convergente dans  $H^3(0,1)$ . Soit  $v$  la limite de  $v_n$ . Tout d'abord, on a

$$\|v'''\|_{L^2(0,1)} = \lim_{n \rightarrow +\infty} \|v_n'''\|_{L^2(0,1)} = 0,$$

donc  $v \in \mathbb{P}_2$ . D'autre part, l'opérateur  $r_1$  étant continu de  $H^1(0,1)$  à valeurs dans  $\mathbb{P}_2$  (les fonctions  $H^1$  en dimension 1 s'injectant dans l'espace des fonctions continues, cf. Lemme **4.3.3**), on peut passer à la limite dans (5.9), d'où

$$1 = \|(r_1 v - v)'\|_{L^2(0,1)}.$$

Comme  $v$  est un polynôme de degré inférieur ou égal à 2, on a

$$r_1 v = v,$$

ce qui est contradictoire avec l'égalité précédente et établit (5.7).

Nous allons généraliser l'inégalité (5.7) pour tout  $h = 1/(n+1)$ , en prenant soin d'exprimer la dépendance de la constante de majoration en fonction de  $n$ . On décompose la norme  $L^2$  de  $r_h v - v$  en une somme de termes portant sur chacune des mailles :

$$\|(r_h v - v)'\|_{L^2(0,1)}^2 = \int_0^1 |(r_h v - v)'(x)|^2 dx = \sum_{j=0}^n \int_{jh}^{(j+1)h} |(r_h v - v)'(x)|^2 dx.$$

Par le changement de variables  $x = h(j+y)$  dans chaque maille, et en posant  $v_j(y) = v(x)$ , on a

$$\|(r_h v - v)'\|_{L^2(0,1)}^2 = h^{-1} \sum_{j=0}^n \int_0^1 |(r_1 v_j - v_j)'(y)|^2 dy \leq C_1^2 h^{-1} \sum_{j=0}^n \|v_j'''\|_{L^2(0,1)}^2.$$

En effectuant ce changement de variable en sens inverse, on établit que

$$\|v_j'''\|_{L^2(0,1)}^2 = h^5 \int_{jh}^{(j+1)h} |v'''(x)|^2 dx.$$

Ainsi,

$$\|(r_h v - v)'\|_{L^2(0,1)}^2 \leq C_1^2 h^4 \|v'''\|_{L^2(0,1)}^2.$$

D'après l'inégalité de Poincaré, la norme  $L^2$  du gradient est une norme équivalente à la norme  $H^1$  sur  $H_0^1$ . Il existe donc  $C_0$  tel que pour tout  $v \in H^3(0,1) \cap H_0^1(0,1)$ ,

$$\|r_h v - v\|_{H^1(0,1)} \leq C_0 h^2 \|v'''\|_{L^2(0,1)}. \quad (5.10)$$

La convergence de la méthode des éléments finie  $\mathbb{P}_2$ , dans le cas régulier où la solution  $u$  appartient à  $H^3(0,1)$ , découle alors de la majoration (5.6) :

$$\|u - u_h\|_{H^1(0,1)} \leq C \|u - r_h u\|_{H^1(0,1)} \leq C C_0 h^2 \|u'''\|_{L^2(0,1)}^2.$$

Il reste à établir la convergence dans le cas où on n'effectue aucune hypothèse de régularité sur la solution du problème. A cet effet, on établit dans un premier temps la continuité uniforme, par rapport à  $h$ , de  $r_h$  comme application linéaire de  $H^1(0,1)$  dans  $H^1(0,1)$ . Notons que la continuité de l'opérateur  $r_h$  est évidente, les fonctions  $H^1(0,1)$  s'injectant de manière continue dans  $C([0,1])$  (voir le Lemme 4.3.3). Il nous faut donc seulement étudier la dépendance de la constante de continuité en fonction de  $h$ . Par le changement de variables  $x = h(j+y)$  et avec  $v_j(y) = v(x)$ , on a

$$\|(r_h v)'\|_{L^2(0,1)}^2 = \sum_{j=0}^n \int_{x_j}^{x_{j+1}} |(r_h v)'|^2 dx = h^{-1} \sum_{j=0}^n \int_0^1 |(r_1 v_j)'|^2 dy.$$

L'opérateur  $r_1$  étant continu, il existe une constante  $C_3$  telle que

$$\|(r_h v)'\|_{L^2(0,1)}^2 \leq h^{-1} C_3 \sum_{j=0}^n \int_0^1 |v'_j|^2 dx.$$

En effectuant le changement de variable en sens inverse, on obtient

$$\|(r_h v)'\|_{L^2(0,1)}^2 \leq C_3 \|v'\|_{L^2(0,1)}^2.$$

D'après l'inégalité de Poincaré, la norme  $L^2$  du gradient est équivalente à la norme  $H^1$  sur  $H_0^1$ . On a donc démontré la continuité uniforme de  $r_h$ , c'est-à-dire qu'il existe une constante  $C$  indépendante de  $h$  telle que, pour tout  $v \in H_0^1(0, 1)$ ,

$$\|r_h v\|_{H^1(0,1)} \leq C \|v\|_{H^1(0,1)}.$$

Comme dans le cas régulier, la convergence de la méthode s'obtient en majorant (5.6) par  $\|u - r_h u\|_{H^1(0,1)}$ , puis en approchant  $u$  par une fonction régulière  $\varphi$

$$\|u - u_h\|_{H^1(0,1)} \leq C \|u - r_h u\|_{H^1(0,1)} \leq C (\|u - \varphi\|_{H^1(0,1)} + \|\varphi - r_h \varphi\|_{H^1(0,1)} + \|r_h(\varphi - u)\|_{H^1(0,1)}).$$

Comme l'espace des fonctions de classe  $C^\infty$  à support compact est dense dans  $H_0^1(0, 1)$ , pour tout  $\varepsilon > 0$ , il existe  $\varphi \in C_c^\infty(\Omega)$  tel que

$$\|u - \varphi\|_{H^1(0,1)} \leq \varepsilon.$$

Par ailleurs, l'opérateur  $r_h$  étant uniformément continu par rapport à  $h$ , on a

$$\|r_h(\varphi - u)\|_{H^1(0,1)} \leq C \|\varphi - u\|_{H^1(0,1)},$$

d'où l'on déduit

$$\|u - u_h\|_{H^1(0,1)} \leq C (2\varepsilon + \|\varphi - r_h \varphi\|_{H^1(0,1)}).$$

Le petit paramètre  $\varepsilon$  étant fixé, par application de (5.10) à  $\varphi$ , il existe  $h_0 > 0$  tel que, pour tout  $0 < h \leq h_0$ , on a

$$\|\varphi - r_h \varphi\|_{H^1(0,1)} \leq \varepsilon,$$

Par conséquent, on obtient que, pour tout  $\varepsilon > 0$ , il existe  $h_0 > 0$  tel que, pour tout  $0 < h \leq h_0$ ,

$$\|u - u_h\|_{H^1(0,1)} \leq C\varepsilon,$$

ce qui prouve la convergence de la méthode des éléments finis  $\mathbb{P}_2$  au sens où

$$\lim_{h \rightarrow 0} \|u - u_h\|_{H^1(0,1)} = 0.$$





**Exercice 5.3.1** Soit  $\mathcal{T}_h$  un maillage de  $\bar{\Omega}$  pour  $\Omega$  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  $n_{s0}$  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  $3n_t + n_s = 2n_c + n_{s0}$ .

**Correction.** Plutôt que de vérifier les relations d'Euler, on se propose de les retrouver directement en effectuant un raisonnement constructif par récurrence sur le nombre de triangles du maillage. On cherche à déterminer les vecteurs  $L \in \mathbb{Z}^4$  et les entiers  $\alpha$  tel que, pour tout maillage borné et simplement connexe du plan,  $L \cdot \mathbf{n} + \alpha = 0$  où  $\mathbf{n} = (n_t, n_c, n_s, n_{s0})$ . Tout d'abord, la relation doit être vérifiée par le maillage trivial constitué d'un unique triangle. On a donc

$$L \cdot (1, 3, 3, 0) + \alpha = 0. \quad (5.12)$$

Considérons maintenant un maillage borné et simplement connexe comportant plusieurs triangles. On choisit un chemin à l'intérieur de ce maillage, constitué d'une succession d'arêtes reliant deux sommets distincts du bord, tel que ce chemin ne s'intersecte pas. Le domaine maillé  $\Omega$  étant simplement connexe, ce chemin le sépare en deux ouverts simplement connexes  $\Omega_1$  et  $\Omega_2$ . On note  $\mathbf{n}_1$  et  $\mathbf{n}_2$  les vecteurs composés du nombre de triangles, d'arêtes, de sommets et de sommets intérieurs de chacun de ces deux sous-maillages. On note  $\tilde{n}_c$  et  $\tilde{n}_s$  les nombres de cotés et de sommet du chemin qui vérifient évidemment que  $\tilde{n}_s = \tilde{n}_c + 1$ . On remarque que

$$\mathbf{n} = \mathbf{n}_1 + \mathbf{n}_2 + (0, -\tilde{n}_c, -\tilde{n}_s, \tilde{n}_s - 2).$$

Si les trois maillages  $\Omega$ ,  $\Omega_1$  et  $\Omega_2$  vérifient la relation affine désirée, c'est-à-dire  $L \cdot \mathbf{n} + \alpha = 0$ ,  $L \cdot \mathbf{n}_1 + \alpha = 0$  et  $L \cdot \mathbf{n}_2 + \alpha = 0$ , on en déduit par linéarité que nécessairement

$$\tilde{n}_c L \cdot (0, -1, -1, 1) + L \cdot (0, 0, -1, -1) - \alpha = 0.$$

Le nombre de cotés  $\tilde{n}_c$  du chemin pouvant prendre des valeurs quelconques, on en déduit que nécessairement

$$L \cdot (0, 0, 1, 1) = -\alpha \quad \text{et} \quad L \cdot (0, -1, -1, 1) = 0. \quad (5.13)$$

Les trois relations (5.12) et (5.13) sont donc des conditions nécessaires pour que la formule  $L \cdot x = \alpha$  soit vérifiée par tout maillage. On peut reformuler ces conditions sous la forme

$$L \in \text{Vect}((-2, 1, 0, 1); (-1, 0, 1, 1)) \quad \text{et} \quad \alpha = -L \cdot (0, 0, 1, 1).$$

Ainsi, on a uniquement deux relations d'Euler indépendantes possibles :

$$-2n_t + n_c + n_{s0} = 1 \quad \text{et} \quad -n_t + n_s + n_{s0} = 2.$$

On vérifie que ces relations sont bien équivalentes à celles proposées par l'énoncé.

Par récurrence sur le nombre de triangles du maillage, on obtient de plus que ces conditions sont suffisantes. En effet, l'équation (5.12) permet d'initialiser la récurrence. Ensuite, tout maillage borné simplement connexe peut être coupé en deux sous-maillages (strictement plus petits) par un chemin comme décrit ci-dessus. En additionnant les relations  $L \cdot \mathbf{n}_1 + \alpha = 0$  et  $L \cdot \mathbf{n}_2 + \alpha = 0$ , vérifiées par ces sous-maillages, on obtient bien  $L \cdot \mathbf{n} + \alpha = 0$  grâce à l'équation (5.13).

**Exercice 5.3.2** 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  $x = \sum_{j=1}^{N+1} \lambda_j(x) a_j$ , et que l'application qui à  $x$  associe  $p(x) - p(0)$  est linéaire, on a

$$p(x) - p(0) = \sum_{j=1}^{N+1} \lambda_j(x) (p(a_j) - p(0)).$$

Comme  $\sum_{j=1}^{N+1} \lambda_j = 1$ , on en déduit que

$$p(x) = \sum_{j=1}^{N+1} \lambda_j(x) p(a_j).$$

**Exercice 5.3.3** 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 leurs 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  $\Sigma_2$  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

$$p(x) = \sum_{j=1}^{N+1} p(a_j) \lambda_j(x) (2\lambda_j(x) - 1) + \sum_{1 \leq j < j' \leq N+1} 4p(a_{jj'}) \lambda_j(x) \lambda_{j'}(x), \quad (5.14)$$

où les  $(\lambda_j(x))_{1 \leq j \leq N+1}$  sont les coordonnées barycentriques de  $x \in \mathbb{R}^N$ .

**Correction.** Rappelons tout d'abord que le treillis  $\Sigma_2$  est défini par

$$\Sigma_2 = \left\{ x \in K \text{ tel que } \lambda_j(x) \in \left\{ 0, \frac{1}{2}, 1 \right\} \text{ pour } 1 \leq j \leq N \right\}$$

Il contient naturellement les sommets du simplexe (pour lesquels les coordonnées barycentriques sont  $\lambda_i(a_j) = \delta_{ij}$ ) et les points milieu des arêtes (pour lesquels les coordonnées barycentriques sont  $\lambda_i(a_{jj'}) = (\delta_{ij} + \delta_{ij'})/2$ ). Réciproquement, tout éléments du treillis  $\Sigma_2$  est soit un sommet, soit un point milieu d'une arête. En effet, la somme des coordonnées barycentriques d'un point étant égale à un, les coordonnées baycentriques  $\lambda_i(x)$  d'un point du treillis sont soit égale à un pour l'un des sommet, nulle pour les autres (dans ce cas  $x$  est un sommet du simplexe), soit égale à  $1/2$  pour deux sommets et nulle pour les autres (dans ce cas,  $x$  est un point milieu d'une arête; voir la Figure 6.9). Le treillis  $\Sigma_2$  est donc précisément constitué de l'ensemble des points milieu des arêtes et des sommets du simplexe.

Comme les coordonnées barycentriques sont des fonctions affines, les fonctions  $q_j$  et  $q_{jj'}$  définies, pour  $1 \leq j < j' \leq N + 1$ , par

$$\begin{aligned} q_j(x) &= \lambda_j(x) (2\lambda_j(x) - 1) \\ q_{jj'}(x) &= 4\lambda_j(x)\lambda_{j'}(x) \end{aligned}$$

sont des polynômes de degré deux. De plus, ils sont indépendants les uns des autres, car

$$q_j(a_i) = \delta_{ij}, \quad q_j(a_{ii'}) = 0, \quad q_{jj'}(a_i) = 0, \quad q_{jj'}(a_{ii'}) = \delta_{ij}\delta_{ij'}.$$

Il suffit donc de vérifier que la famille constituée par les polynômes  $(q_j)_{1 \leq j \leq N+1}$  et  $(q_{jj'})_{1 \leq j < j' \leq N+1}$  contient autant d'élément que l'espace  $\mathbb{P}_2$  pour en déduire qu'elle en est génératrice et conclure. L'espace des polynômes  $\mathbb{P}_2$ , engendré par  $N$  monômes de type  $x_i^2$ ,  $N(N-1)/2$  du type  $x_i x_j$  (avec  $i < j$ ),  $N$  monômes de type  $x_i$  et la constante unité, est de dimension  $N + N(N-1)/2 + N + 1 = N(N+1)/2 + N + 1$ . La famille constituée des  $q_j$  et  $q_{jj'}$  compte précisément  $N + 1 + (N+1)N/2$  éléments. Elle engendre donc l'espace des polynômes de degré 2 et tout polynôme  $p \in \mathbb{P}_2$  se décompose sur cette base selon (5.14).

**Exercice 5.3.4** Soit  $\mathcal{T}_h$  un maillage de  $\bar{\Omega}$  pour  $\Omega$  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.** Dans le plan, pour un treillis d'ordre  $k$ , on compte  $(k+1)(k+2)/2$  éléments ou noeuds, 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 5.3.5** Démontrer la formule (6.43) en dimension  $N = 2$ , c'est-à-dire

$$\int_K \lambda_1(x)^{\alpha_1} \lambda_2(x)^{\alpha_2} \lambda_3(x)^{\alpha_3} dx = 2\text{Aire}(K) \frac{\alpha_1! \alpha_2! \alpha_3!}{(\alpha_1 + \alpha_2 + \alpha_3 + 2)!}, \quad (5.15)$$

où  $K$  est un simplexe de  $\mathbb{R}^2$ ,  $\lambda_i(x)$  sont les coordonnées barycentriques de  $x$  et  $\alpha_i$  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

$$I = \int_S \lambda_1^{\alpha_1} \lambda_2^{\alpha_2} \lambda_3^{\alpha_3} |\det(\nabla F)| d\lambda_1 d\lambda_2,$$

avec  $\lambda_3 = 1 - (\lambda_1 + \lambda_2)$ . Or

$$|\det(\nabla F)| = \left| \frac{\partial F}{\partial \lambda_1} \wedge \frac{\partial F}{\partial \lambda_2} \right| = |(a_1 - a_3) \wedge (a_2 - a_3)| = 2\text{Aire}(K).$$

Ainsi,

$$I = 2\text{Aire}(K) \int_S \lambda_1^{\alpha_1} \lambda_2^{\alpha_2} \lambda_3^{\alpha_3} d\lambda_1 d\lambda_2. \quad (5.16)$$

Il reste 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  $\lambda_2$

$$\int_S \lambda_1^{\alpha_1} \lambda_2^{\alpha_2} \lambda_3^{\alpha_3} d\lambda_1 d\lambda_2 = \int_0^1 \lambda_1^{\alpha_1} (1 - \lambda_1)^{\alpha_2 + \alpha_3 + 1} d\lambda_1 \int_0^1 t^{\alpha_2} (1 - t)^{\alpha_3} dt.$$

Par intégration par partie 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 (5.16) nous donne (5.15).

**Exercice 5.3.6** Montrer que les formules de quadrature

$$\int_K \psi(x) dx \approx \text{Volume}(K) \psi(a_0), \quad (5.17)$$

avec  $a_0 = (N + 1)^{-1} \sum_{i=1}^{N+1} a_i$ , le barycentre de  $K$ , et

$$\int_K \psi(x) dx \approx \frac{\text{Volume}(K)}{N + 1} \sum_{i=1}^{N+1} \psi(a_i). \quad (5.18)$$

sont exactes pour  $\psi \in \mathbb{P}_1$ .

**Correction.** Soit  $p$  un polynôme de degré 1, il existe  $q$  polynôme de degré 1 en  $\lambda$  tel que  $q(\lambda(x)) = p(x)$ . Or

$$\int_K 1 dx = \text{Volume}(K) \quad \text{et} \quad \int_K \lambda_k dx = \frac{\text{Volume}(K)}{N+1} \sum_{i=1}^{N+1} \lambda_k(a_i).$$

On en déduit donc que

$$\int_K q(\lambda(x)) dx = \frac{\text{Volume}(K)}{N+1} \sum_{i=1}^{N+1} q(\lambda(a_i)),$$

et que la formule (5.18) est exacte pour les polynômes de degré 1. De plus, comme  $p$  est affine,  $p(a_0) = 1/(N+1) \sum_i p(a_i)$ , ce qui établit l'exactitude de la formule (5.17)

**Exercice 5.3.7** 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.** On procède comme pour l'Exercice 5.3.6 en vérifiant l'exactitude des formules de quadrature pour les polynômes de la forme

$$p(x) = q(\lambda_1(x), \lambda_2(x), \lambda_3(x)),$$

où  $(\lambda_i)$  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

$$\int_K q(\lambda_1(x), \lambda_2(x), \lambda_3(x)) dx = T_2(q), \quad (5.19)$$

où  $T_2$  est défini par

$$T_2(q) = \frac{\text{Aire}(K)}{3} \sum_{1 \leq i < j \leq 3} q((e_i + e_j)/2),$$

$(e_1, e_2, e_3)$  désignant la base canonique de  $\mathbb{R}^3$  et que pour tout polynôme  $q$  de trois variables et de degré trois,

$$\int_K q(\lambda_1(x), \lambda_2(x), \lambda_3(x)) dx = T_3(q), \quad (5.20)$$

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 (5.19) et (5.20) 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  $\sum_{i=1}^3 \alpha_i \leq 3$ ,

$$S(\alpha_1, \alpha_2, \alpha_3) = T_3(X_1^{\alpha_1} X_2^{\alpha_2} X_3^{\alpha_3}).$$

Pour toute permutation  $\sigma$  de  $\{1, 2, 3\}$  et pour tout multi-indice  $\alpha \in \mathbb{N}^3$ , on a

$$S(\alpha_1, \alpha_2, \alpha_3) = S(\alpha_{\sigma_1}, \alpha_{\sigma_2}, \alpha_{\sigma_3})$$

et

$$T_\beta(X_1^{\alpha_1} X_2^{\alpha_2} X_3^{\alpha_3}) = T_\beta(X_1^{\alpha_{\sigma_1}} X_2^{\alpha_{\sigma_2}} X_3^{\alpha_{\sigma_3}}) \text{ pour } \beta = 1, 2.$$

Ces considérations d'invariance, nous permettent de limiter le nombre de vérifications à effectuer. Seuls les 4 cas et les 7 cas suivants sont à considérer pour vérifier respectivement la première et deuxième formule de quadrature :

$$\begin{aligned} S(0, 0, 0) &= \text{Aire}(K) &= T_2(1) &= T_3(1), \\ S(1, 0, 0) &= \text{Aire}(K)/3 &= T_2(X_1) &= T_3(X_1), \\ S(1, 1, 0) &= \text{Aire}(K)/12 &= T_2(X_1 X_2) &= T_3(X_1 X_2), \\ S(2, 0, 0) &= \text{Aire}(K)/6 &= T_2(X_1^2) &= T_3(X_1^2), \\ S(1, 1, 1) &= \text{Aire}(K)/60 &= T_3(X_1 X_2 X_3), \\ S(2, 1, 0) &= \text{Aire}(K)/30 &= T_3(X_1^2 X_2), \\ S(3, 0, 0) &= \text{Aire}(K)/60 &= T_3(X_1^3). \end{aligned}$$

**Exercice 5.3.8** 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  $\psi$ , 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  $\psi$  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  $\psi$ , 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

$$\left| \int_K \psi dx - \int_K T_{a_0}(x - a_0) dx \right| \leq C \text{Vol}(K)h^{k+1}.$$

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  $\psi$  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 5.3.9** On considère le carré  $\Omega = ]-1, +1[^2$  maillé suivant la Figure 5.1. 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).

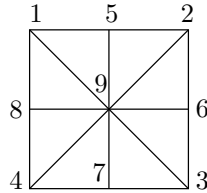


FIGURE 5.1 – Exemple de maillage et de numérotation des nœuds.

**Correction.** On note  $V_h$  l'espace des éléments finis  $\mathbb{P}_1$  associé au maillage de la Figure 5.1. L'espace  $V_h$  est de dimension 9. Pour tout  $i \in \{1, \dots, 9\}$ , on note  $\phi_i$  la fonction de base associée au  $i$ ème nœud (on utilise la numérotation des nœuds indiquée sur la figure). En d'autres termes,  $\phi_i$  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  $\phi_i$  et  $\phi_j$  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  $\phi_i$  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 5.3.10** Appliquer la méthode des éléments finis  $\mathbb{P}_1$  au problème de Dirichlet

$$\begin{cases} -\Delta u = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega, \end{cases} \quad (5.21)$$

dans le carré  $\Omega = ]0, 1[^2$  avec le maillage triangulaire uniforme de la Figure 5.2. 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.

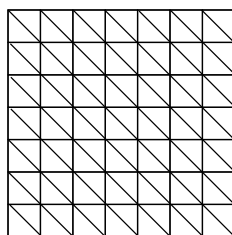


FIGURE 5.2 – Maillage triangulaire uniforme d'un carré

**Correction.** On note  $n + 2$  le nombre de noeuds du maillage 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  $1 \leq i, j \leq n$ ). On numérote les noeuds du maillage ligne par ligne. En d'autres termes, on pose  $a_{i+jn} = x_{i,j}$  pour tout

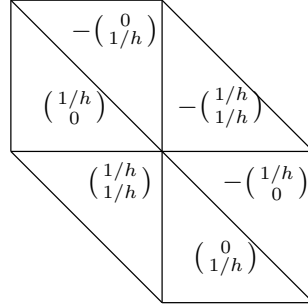


FIGURE 5.3 – Valeurs du gradient d'une fonction de base

$1 \leq i, j \leq n$ . Enfin, on note  $\phi_k$  la fonction de base  $\mathbb{P}_1$  associée au nœud  $a_k$ . La Figure 5.3 représente les valeurs du gradient d'une fonction de base  $\phi_k$  sur son support (constant par maille). La matrice de rigidité est définie par  $(\mathcal{K}_h)_{k,l} = \int_{\Omega} \nabla \phi_k \cdot \nabla \phi_l \, dx$ . Si  $k = l$ , on a

$$(\mathcal{K}_h)_{k,k} = \int_{\Omega} |\nabla \phi_k|^2 \, dx.$$

Le gradient  $\nabla \phi_k$  est nul sur tout  $\Omega$  à l'exclusion des 6 triangles contenant  $a_k$ . Sur chacun d'entre eux,  $|\nabla \phi_k|^2$  est constant, égal à  $1/h^2$  sur quatre d'entre eux,  $2/h^2$  sur les deux autres (voir Figure 5.3). Enfin, l'aire des triangles du maillage étant égale à  $h^2/2$ , on trouve

$$(\mathcal{K}_h)_{k,k} = 4.$$

Si  $a_k$  et  $a_l$  sont des nœuds voisins, c'est-à-dire si  $k = l + 1$ ,  $k = l - 1$ ,  $k = l + n - 1$ ,  $k = l + n$ ,  $k = l - n$  ou  $k = l - n + 1$ , les supports de  $\phi_k$  et  $\phi_l$  ne sont pas disjoints. Cependant, le terme  $(\mathcal{K}_h)_{k,l}$  est nul dans les cas  $k = l - n + 1$  et  $k = l + n - 1$  (les gradients des fonctions  $\phi_k$  et  $\phi_l$  sont orthogonaux). Dans les autres cas, on a

$$(\mathcal{K}_h)_{k,l} = -1.$$

Par conséquent,  $\mathcal{K}_h$  est une matrice tridiagonale par blocs (avec  $n$  lignes et  $n$  colonnes de blocs)

$$\mathcal{K}_h = \begin{pmatrix} D & E & & 0 \\ E & D & E & \\ & \cdot & \cdot & \cdot \\ 0 & & E & D & E \\ & & & E & D \end{pmatrix},$$

où les blocs  $E$  et  $D$  sont les matrices  $n \times n$  données par

$$D = \begin{pmatrix} 4 & -1 & & 0 \\ -1 & 4 & -1 & \\ & \cdot & \cdot & \cdot \\ & & -1 & 4 & -1 \\ 0 & & & -1 & 4 \end{pmatrix}, \quad E = \begin{pmatrix} -1 & 0 & & \\ 0 & -1 & 0 & \\ & \cdot & \cdot & \cdot \\ & & 0 & -1 & 0 \\ & & & 0 & -1 \end{pmatrix}.$$

On obtient donc en effet la matrice issue de la méthode des différences finies (voir cours, section **2.2.6**) multipliée par  $h^2$ . Cependant, le second membre du système linéaire obtenu diffère car, en général,

$$(b_h)_k = \int_{\Omega} f \phi_k dx \neq h^2 f(a_k).$$

**Exercice 5.3.11** On reprend les notations de l'Exercice **6.3.10**. On note  $n$  le nombre de points du maillage sur un côté du carré (supposé être le même pour chaque côté). 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  $\Omega$ ).

**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  $|k - l| > n$ ,  $(\mathcal{K}_h)_{k,l} = 0$ . 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  $\phi_k$  et  $\phi_l$  sont disjoints. Ainsi, la matrice du système obtenu à une demi-largeur de bande de l'ordre de  $n^2$  pour  $n$  grand.

**Exercice 5.3.12** 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  $X \in \mathbb{R}^N$  : on dit que  $X \geq 0$  si toutes ses coordonnées  $X_i$  sont positives ou nulles. On va tout d'abord montrer que, pour toute M-matrice  $B$ , s'il existe deux vecteurs  $X, Y \in \mathbb{R}^N$  tels que  $BX = Y \geq 0$ , alors on a aussi  $X \geq 0$ . Introduisons un 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 l'on déduit, par définition de  $i_0$  et puisque  $b_{i_0 j} \leq 0$  pour  $j \neq i_0$ ,

$$\left( \sum_{j=1}^N b_{i_0 j} \right) X_{i_0} \geq Y_{i_0} \geq 0.$$



Comme  $\sum_{j=1}^N b_{i0j} > 0$ , on en déduit que  $X_{i0} \geq 0$  et donc que  $X \geq 0$ . Cette propriété permet de montrer que  $B$  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 nécessairement positifs (prendre  $X$  égal, successivement, à chacun des vecteurs de la base canonique).

**Exercice 5.3.13** On se place en dimension  $N = 2$ . Soit  $u_h$  la solution approchée du problème de Dirichlet (5.21) 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  $\Omega$  si  $f(x) \geq 0$  dans  $\Omega$ . 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  $\epsilon > 0$ ,  $\mathcal{K}_h + \epsilon \text{Id}$  est une M-matrice. En effet, dans ce cas et d'après l'exercice précédent, tous les coefficients de la matrice  $(\mathcal{K}_h + \epsilon \text{Id})^{-1}$  sont positifs. L'application qui à une matrice associe son inverse étant continue sur l'ensemble des matrices inversibles, on en déduit en faisant tendre  $\epsilon$  vers 0 que les coefficients de la matrice  $\mathcal{K}_h^{-1}$  sont positifs. On conclut alors en rappelant que  $u_h = \mathcal{K}_h^{-1}b_h$  et que  $f(x) \geq 0$  implique  $b_h \geq 0$ .

Tout d'abord, il est clair que, pour tout  $i$ ,

$$(\mathcal{K}_h)_{ii} = \int_{\Omega} |\nabla \phi_i|^2 dx > 0. \quad (5.22)$$

Considérons ensuite deux sommets distincts  $a_i$  et  $a_j$  communs à un triangle  $T_k$  du maillage.

Le gradient de  $\phi_i$  est orthogonal au coté du triangle  $T_k$  opposé à  $a_i$ , car la restriction de  $\phi_i$  à cette arête est la constante 0. Il en est de même pour  $\nabla \phi_j$ . On en déduit que l'angle formés par les vecteurs  $\nabla \phi_i$  et  $\nabla \phi_j$  est égal à  $(\pi - \alpha_{ij})$  où  $\alpha_{ij}$  est l'angle du triangle  $T_k$  à son sommet distinct de  $a_i$  et  $a_j$ . Comme on a supposé que tous les angles des triangles du maillages étaient inférieurs à  $\pi/2$ , on en déduit que  $(\pi - \alpha_{ij}) \geq \pi/2$  et donc

$$\nabla \phi_i \cdot \nabla \phi_j \leq 0 \quad \text{sur } T_k.$$

Le raisonnement étant valable sur tous les triangles du maillage, on en déduit que

$$(\mathcal{K}_h)_{ij} = \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j dx \leq 0 \quad \text{pour tout } i \neq j. \quad (5.23)$$

Soit  $n_0$  le nombre de nœuds du maillage situés à l'intérieur du domaine  $\Omega$  et  $n$  le nombre de nœuds total, on numérote les nœuds  $a_i$  du maillage de sorte que  $a_i \in \partial\Omega$  pour  $i > n_0$ . Comme

$$1 = \sum_{j=1}^n \phi_j,$$

pour tout  $i$  tel que  $1 \leq i \leq n_0$ ,

$$0 = \int_{\Omega} \nabla \phi_i \cdot \nabla 1 dx = \sum_{j=1}^n \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j dx.$$

Ainsi,

$$\sum_{j=1}^{n_0} \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j \, dx = - \sum_{j=n_0+1}^n \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j \, dx. \quad (5.24)$$

Or on a prouvé précédemment que le second membre de (5.24) est positif. On a donc montré que

$$\sum_{j=1}^{n_0} (\mathcal{K}_h)_{ij} \geq 0. \quad (5.25)$$

De (5.22), (5.23), (5.25), on déduit que  $\mathcal{K}_h + \varepsilon \text{Id}$  est une M-matrice pour tout  $\varepsilon > 0$ , ce qui achève la démonstration. Remarquons qu'il est nécessaire d'ajouter  $\varepsilon \text{Id}$  pour que  $\mathcal{K}_h + \varepsilon \text{Id}$  soit une M-matrice car l'inégalité (5.25) est large alors qu'elle doit être stricte pour une M-matrice.

**Exercice 5.3.14** Appliquer la méthode des éléments finis  $\mathbb{P}_k$  au système de l'élasticité (4.19). Montrer en particulier que la matrice de rigidité  $\mathcal{K}_h$  est dans ce cas d'ordre  $N n_{dl}$  où  $N$  est la dimension d'espace et  $n_{dl}$  est le nombre de nœuds de degrés de liberté.

**Correction.** La formulation variationnelle de l'élasticité linéarisée consiste à déterminer  $u \in H_0^1(\Omega)^N$  tel que

$$\int_{\Omega} (2\mu e(u) \cdot e(v) + \lambda(\text{div}u)(\text{div}v)) \, dx = \int_{\Omega} f \cdot v \, dx \text{ pour tout } v \in H_0^1(\Omega)^N. \quad (5.26)$$

Soit  $\mathcal{T}_h$  un maillage régulier de  $\Omega$ , on introduit les espaces discrets

$$V_h = \{u = (u_i)_{1 \leq i \leq N} \in C(\Omega; \mathbb{R})^N : u_i|_K \in \mathbb{P}_k \text{ pour tout } K \in \mathcal{T}_h, 1 \leq i \leq N\}$$

et

$$V_{0h} = \{u \in V_h : u = 0 \text{ sur } \partial\Omega\}.$$

Soit  $(\phi_i)_{i=1, \dots, n_{dl}}$  les fonctions de base associées aux degrés de liberté du treillis d'ordre  $k$  du maillage  $\mathcal{T}_h$ , intérieur à  $\Omega$ . Les fonctions  $\phi_i$  sont à valeurs dans  $\mathbb{R}$ . On introduit alors les fonctions  $(\Phi_i^k)_{1 \leq i \leq n_{dl}}^{1 \leq k \leq N}$  à valeurs dans  $\mathbb{R}^N$  définies par  $\Phi_i^k = \phi_i e_k$  où  $e_k$  est le  $k$ -ème vecteur de la base canonique de  $\mathbb{R}^N$ . Les fonctions  $\Phi_i^k$  forment une base de l'espace  $V_{0h}$  car chaque élément de  $V_{0h}$  est défini de manière unique par ses valeurs (dans  $\mathbb{R}^N$ ) aux nœuds de degrés de liberté du treillis d'ordre  $k$ . Ainsi, la dimension de l'espace  $V_{0h}$  est égale à  $N n_{dl}$  où  $n_{dl}$  est le nombre de nœuds de degrés de liberté. La méthode des éléments finis  $\mathbb{P}_k$  appliquée à la résolution de (5.26) consiste à déterminer  $u_h \in V_{0h}$  tel que

$$\int_{\Omega} (2\mu e(u_h) \cdot e(v) + \lambda(\text{div}u_h)(\text{div}v)) \, dx = \int_{\Omega} f \cdot v \, dx \text{ pour tout } v \in V_{0h},$$

c'est-à-dire à résoudre le système linéaire

$$\mathcal{K}_h U_h = b_h$$

où  $U_h$  est le vecteur de  $\mathbb{R}^{N \times n_{dl}}$  des coordonnées de  $u_h$  dans la base des  $\Phi_i^k$ ,  $\mathcal{K}_h$  est la matrice de rigidité de taille  $(Nn_{dl}) \times (Nn_{dl})$

$$(\mathcal{K}_h)_{ij}^{kl} = \int_{\Omega} (2\mu e(\Phi_i^k) \cdot e(\Phi_j^l) + \lambda(\operatorname{div}\Phi_i^k)(\operatorname{div}\Phi_j^l)) \, dx$$

et le second membre

$$(b_h)_i^k = \int_{\Omega} f \cdot \Phi_i^k \, dx.$$

L'existence d'une solution à ce problème est évidente par application du théorème de Lax-Milgram.

**Exercice 5.3.15** Expliciter la matrice de rigidité  $\mathcal{K}_h$  obtenue par application de la méthode des éléments finis  $\mathbb{P}_k$  au problème de Neumann

$$\begin{cases} -\Delta u + au = f & \text{dans } \Omega \\ \frac{\partial u}{\partial n} = g & \text{sur } \partial\Omega, \end{cases} \quad (5.27)$$

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  $\Omega$ .

**Correction.** La formulation variationnelle de (5.27) est : trouver  $u \in H^1(\Omega)$  tel que

$$\int_{\Omega} (\nabla u \cdot \nabla v + auv) \, dx = \int_{\Omega} f v \, dx + \int_{\partial\Omega} g v \, ds \quad \text{pour tout } v \in H^1(\Omega).$$

L'espace d'approximation issu de la méthode des éléments finis  $\mathbb{P}_k$ , associé au problème de Neumann (5.27) 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  $(\phi_i)_{i=1, \dots, n_{dl}}$  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

$$(b_h)_i = \int_{\Omega} f \phi_i \, dx + \int_{\partial\Omega} g \phi_i \, ds.$$

**Exercice 5.3.16** 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

$$V_{0h} = \{u \in C(\Omega; \mathbb{R})^N : u|_K \in \mathbb{P}_k \text{ pour tout } K \in \mathcal{T}_h, u = 0 \text{ sur } \partial\Omega\}.$$

Soit  $(\phi_i)_{i=1, \dots, n_{dl}}$  les fonctions de base associées aux degrés 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 car les termes  $\int_{\Omega} (V \cdot \nabla \phi_i) \phi_j dx$  sont anti-symétriques. Par exemple, en dimension  $N = 1$ ,  $\Omega = (0, 1)$ , pour un maillage de points  $x_j = j/(n+1)$  on trouve que

$$\int_0^1 (V \cdot \nabla \phi_i) \phi_j dx = \begin{cases} 0 & \text{si } i = j \\ V/2 & \text{si } i = j + 1 \\ -V/2 & \text{si } i = j - 1 \\ 0 & \text{sinon.} \end{cases}$$

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 5.3.17** On se propose de résoudre numériquement l'équation des plaques (4.28) 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

$$p(a_j), \nabla p(a_j), \nabla \nabla p(a_j), \frac{\partial p(b_j)}{\partial n} \quad j = 1, 2, 3, \quad (5.28)$$

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 (5.28) pour chaque sommet et milieu d'arête du maillage. En déduire une méthode d'éléments finis (dite d'Argyris) pour résoudre (4.28).

### Correction.

#### 1. Unisolvanace (équivalent du Lemme 6.3.3)

On considère l'application qui à un élément de  $\mathbb{P}_5$  associe les 21 valeurs (5.28) (rappelons que le gradient compte pour 2 valeurs réelles et la hessienne des dérivées secondes pour 3 valeurs réelles). Comme  $\mathbb{P}_5$  est un espace de dimension 21 (cf. la preuve du Lemme 6.3.3), 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  $a_1 = (-1, 0)$ ,  $a_2 = (1, 0)$  (voir Remarque 6.3.10 du cours). Soit  $p \in \mathbb{P}_5$  annulant toutes les valeurs (5.28). Montrons que  $p$  est le polynôme nul. On pose  $q_1(x_1) = p(x_1, 0)$  et  $q_2(x_1) = \partial p/\partial x_2(x_1, 0)$ . Par hypothèse, on a

$$q_1(\pm 1) = q_1'(\pm 1) = q_1''(\pm 1) = 0.$$

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, comme  $\partial p/\partial x_2 = \partial p/\partial n$  sur le segment  $[a_1, a_2]$ , on a par hypothèse 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  $x_2^2$ . Pour des raisons d'invariance, par changement de coté, 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 du cours) :

**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  $C^1$  sur  $K \cup K'$  si et seulement si pour  $i = 1, 2$ ,

$$\begin{aligned} p_K(a_i) &= p_{K'}(a_i), & \nabla p_K(a_i) &= \nabla p_{K'}(a_i), \\ \nabla \nabla p_K(a_i) &= \nabla \nabla p_{K'}(a_i), & \partial p_K/\partial n(b) &= \partial p_{K'}/\partial n(b), \end{aligned} \quad (5.29)$$

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  $\Gamma$  aux 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  $p_{K'}$ , de degré au plus cinq, coïncident sur  $\Gamma$  si et seulement si pour  $i = 1, 2$  les 6 conditions suivantes sont satisfaites

$$p_K(a_i) = p_{K'}(a_i), \quad \frac{\partial p_K}{\partial \tau}(a_i) = \frac{\partial p_{K'}}{\partial \tau}(a_i) \text{ et } \frac{\partial^2 p_K}{\partial \tau^2}(a_i) = \frac{\partial^2 p_{K'}}{\partial \tau^2}(a_i)$$

( $\tau$  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  $\Gamma$  sont des polynômes de degré 4 égaux si et seulement si pour  $i = 1, 2$  les 5 conditions suivantes sont vérifiées

$$\frac{\partial p_K}{\partial n}(a_i) = \frac{\partial p_{K'}}{\partial n}(a_i), \quad \frac{\partial^2 p_K}{\partial n^2}(a_i) = \frac{\partial^2 p_{K'}}{\partial n^2}(a_i)$$

et si

$$\frac{\partial p_K}{\partial n}(b) = \frac{\partial p_{K'}}{\partial n}(b).$$

On a donc prouvé que si les conditions (5.29) étaient vérifiées, alors  $v$  était  $C^1$ . Réciproquement, si  $v$  est de classe  $C^1$ , les restrictions de  $\partial p_K/\partial n$  et  $\partial p_{K'}/\partial n$  à l'arête commune  $\Gamma$  coïncident et pour  $i = 1, 2$  on a  $\partial^2 p_{K'}/\partial n \partial \tau(a_i) = \partial^2 p_K/\partial n \partial \tau(a_i)$ . Enfin, on a d'ores et déjà prouvé que les autres conditions de (5.29) étaient satisfaites, ce qui achève la preuve du Lemme.

### 3. Méthode d'Argyris

Tout d'abord, l'espace

$$V_h = \{v \in C^1(\Omega) : v|_{K_i} \in \mathbb{P}^5 \text{ pour tout } K_i \in \mathcal{T}_h\}$$

est inclus dans  $H^2(\Omega)$  (la dérivée d'un élément de  $V_h$  est continue, dérivable par morceaux et 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  $v$ ,  $\nabla v$  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  $(\psi_k)$  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  $|\beta| \leq 2$  ( $\beta$  est un multi-indice, si  $\beta = (\beta_1, \beta_2)$ ,  $\partial^\beta \varphi$  désigne la dérivée partielle  $\partial^{\beta_1 + \beta_2} \varphi / \partial^{\beta_1} x_1 \partial^{\beta_2} x_2$ ).

Afin de résoudre l'équation des plaques (4.28), 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  $(\psi_k)$  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  $\Omega$ . 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

$$\mathcal{K}_h U_h = b_h, \quad (5.30)$$

où la matrice de rigidité (de taille  $6n_{s0} + n_{c0}$ ) 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  $6 \times 6$  blocs de taille  $n_{s0} \times n_{s0}$ . La matrice  $F_h$  est un vecteur colonne constitué de 6 sous-matrices de taille  $n_{s0} \times n_{c0}$ . On pose

$$\begin{aligned} D_h &= (E_h^{ij})_{(i,j) \in \{1, \dots, 6\}^2} \\ F_h &= (G_h^i)_{i \in \{1, \dots, 6\}}. \end{aligned}$$

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  $H_h$ , de taille  $n_{c0} \times n_{c0}$ , 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

$$\begin{aligned} (c_h^i)_k &= \int_{\Omega} f_h \varphi_{k,s_i} \quad k \in \{1, \dots, n_{s0}\} \text{ et } i \in \{1, \dots, 6\} \\ (d_h)_k &= \int_{\Omega} f_h \psi_k \quad k \in \{1, \dots, n_{c0}\}. \end{aligned}$$

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 (5.30).

**Exercice 5.3.18** 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  $r_h v$ , la restriction de  $r_h v$  à un  $N$ -simplexe  $K_i$  est simplement  $r_{K_i} v$  (l'opérateur  $r_{K_i}$  est défini dans le cours par (6.56)). 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 rappelle que  $K$  est l'image par une application affine d'un simplexe de référence  $K_0$ . On utilise la même notation que dans le cours en introduisant la matrice  $B$ , composante linéaire de cette application (voir (6.62) dans le cours). 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  $K_i$  (on rappelle que  $|v|_{H^2(K)} = (\int_K |\nabla \nabla v| dx)^{1/2}$ ). Ainsi,

$$\|v - r_h v\|_{L^2(\Omega)}^2 \leq C \sum_{K_i \in \mathcal{T}_h} \|B_i\|^4 |v|_{H^2(K_i)}^2.$$

Il suffit de combiner cette estimation avec l'inégalité (voir le Lemme 6.3.20 dans le cours)

$$\|B_i\| \leq \text{diam}(K_i) / \rho(K_0) \leq Ch$$

pour conclure.

**Exercice 5.3.19** 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 tous 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

$$p_1(x) = x_3x_4 = (1 - x_1)(1 - x_2),$$

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  $\Sigma_2$  sont les sommets, les milieux des arêtes et le centre du carré  $K$ . Pour des raisons de symétrie, seules 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. Puis

$$P_5(x) = -4x_3(x_3 - 1)x_4(2x_4 - 1) = 4(1 - x_1)x_1(1 - x_2)(1 - 2x_2)$$

qui vaut 1 pour  $x = (a_1 + a_2)/2$  et zéro sur les autres nœuds du treillis. Enfin,

$$P_9(x) = 16x_1x_2x_3x_4 = 16x_1x_2(1 - x_1)(1 - x_2),$$

qui vaut 1 en  $x = (a_1 + a_2 + a_3 + a_4)/4$  et zéro sur les autres nœuds du treillis.

**Exercice 5.3.20** 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). Soit  $R_h$  et  $W_h$  les coordonnées respectives de  $r_h$  et  $w_h$  dans les bases de  $\mathcal{Q}_h$  et  $V_{0h}$ . On note  $B_h$  la matrice associée à la forme bilinéaire  $b$  définie sur  $\mathcal{Q}_h \times V_{0h}$  par  $b(r_h, w_h) = \int_{\Omega} w_h \cdot \nabla r_h \, dx$ . Ainsi,

$$W_h \cdot B_h^* R_h = B_h W_h \cdot R_h = - \int_{\Omega} \text{div}(w_h) r_h \, dx = \int_{\Omega} w_h \cdot \nabla r_h \, dx.$$

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  $\lambda_i$  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  $\lambda_k(x) > 0$  à l'intérieur de  $K_i$  et donc

$$\left( \int_{K_i} \lambda_1(x) \cdots \lambda_{N+1}(x) dx \right) > 0,$$

d'où l'on déduit que  $\nabla r_h(K_i) = 0$ . Ainsi,  $r_h$  est une fonction constante,  $R_h$  est proportionnel au vecteur  $(1, \dots, 1)$  et

$$\dim(B_h^*) = 1.$$

**Exercice 5.3.21** On considère les équations de Stokes

$$\begin{cases} \nabla p - \mu \Delta u = f & \text{dans } \Omega \\ \operatorname{div} u = 0 & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega \end{cases} \quad (5.31)$$

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.** Le schéma proposé consiste à déterminer les vecteurs  $U_h = (u_j)_{1 \leq j \leq n}$  et  $P_h = (p_j)_{0 \leq j \leq n+1}$  tels que

$$K_h \begin{pmatrix} U_h \\ P_h \end{pmatrix} = \begin{pmatrix} b_h \\ 0 \end{pmatrix},$$

où  $b_h = (f(x_j))_{1 \leq j \leq n}$  et

$$K_h = \begin{pmatrix} A & C \\ B & 0 \end{pmatrix}$$

avec  $A$  et  $B$  des matrices de taille  $n \times n$

$$A = \frac{\mu}{h^2} \begin{pmatrix} 2 & -1 & & & & \\ -1 & \ddots & \ddots & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & -1 & \\ & & & \ddots & -1 & 2 \end{pmatrix} \text{ et } B = \frac{1}{2h} \begin{pmatrix} 0 & 1 & & & & \\ -1 & \ddots & \ddots & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & 1 & \\ & & & \ddots & -1 & 0 \end{pmatrix},$$

et  $C$  une matrice rectangulaire de taille  $n \times (n + 2)$

$$C = \frac{1}{2h} \begin{pmatrix} -1 & 0 & 1 & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & -1 & 0 & 1 \end{pmatrix},$$

Remarquons tout d'abord que ce système linéaire compte  $2n$  équations pour  $2n+2$  inconnues : il y a donc soit aucune solution, soit une infinité de solutions ! Réglons tout de suite la question de l'existence d'au moins une solution en exhibant la solution particulière suivante :  $u_j = 0$  pour tout  $1 \leq j \leq n$ ,  $p_0 = p_1 = 0$ ,  $p_{2k+1} = \sum_{l=1}^k f(x_{2l})$  et  $p_{2k} = \sum_{l=1}^k f(x_{2l-1})$ .

Nous sommes donc dans le cas d'une infinité de solutions qui diffèrent deux à deux par un élément du noyau de  $K_h$  qui est de dimension au moins égale à 2. Remarquons que la pression dans le système de Stokes (5.31) n'est définie qu'à l'addition d'une constante près puisque seul son gradient apparaît dans les équations. Par conséquent, il est normal que le noyau de  $K_h$  ne soit pas réduit au vecteur nul et contienne au moins les pressions discrètes  $P_h$  proportionnelles au vecteur  $\mathbb{I} = (1, \dots, 1)$  qui est la discrétisation naturelle des constantes. Néanmoins, malgré la suppression d'un tel degré de liberté (par exemple, en fixant la moyenne ou bien une valeur de référence de la pression  $P_h$ ), le noyau de  $K_h$  sera toujours de dimension au moins égale à 1. Ce système linéaire est donc mal posé puisqu'il admet une infinité de solutions.

Il est instructif de calculer explicitement ce noyau pour se rendre compte du type d'instabilités numériques développées par ce schéma centré. Soit  $(U_h, P_h) \in \mathbb{R}^n \times \mathbb{R}^{n+2}$  qui appartient à  $\text{Ker}(K_h)$ . L'équation  $BU_h = 0$  conduit à

$$u_{j+1} = u_{j-1} \quad \text{pour tout } 1 \leq j \leq n,$$

autrement dit, toutes les valeurs paires (ou impaires) des composantes de la vitesse discrète coïncident. Considérons tout d'abord le cas où  $n$  est pair. Comme  $u_0 = 0$ , on en déduit que  $u_j = 0$  pour tout indice  $j$  pair. De même, comme  $u_{n+1} = 0$ ,  $u_j = 0$  pour tout indice  $j$  impair. Ainsi,  $U_h = 0$  et l'équation  $AU_h + CP_h = 0$  est équivalente à

$$p_{j-1} = p_{j+1} \quad \text{pour tout } 1 \leq j \leq n,$$

c'est-à-dire qu'à nouveau toutes les valeurs paires (ou impaires) des composantes de la pression discrète coïncident. On en conclut que  $(U_h, P_h) \in \text{Ker}(K_h)$  si et seulement si  $U_h = 0$  et  $P_h$  est engendré par les deux vecteurs  $\mathbb{I} = (1, \dots, 1)$  et  $(1, 0, \dots, 1, 0)$ . Le deuxième vecteur qui engendre les pressions discrètes du noyau traduit une instabilité numérique d'oscillation entre deux valeurs alternant d'une maille à l'autre.

Considérons maintenant le cas  $n$  impair (i.e.  $n = 2m + 1$  avec  $m$  entier naturel). De l'équation  $BU_h = 0$  et des conditions aux limites  $u_0 = u_{n+1} = 0$  on peut seulement déduire que  $U_h$  est un multiple de  $U_h^0 = (0, 1, \dots, 0, 1, 0)$ . Soit  $\alpha \in \mathbb{R}$  tel

que  $U_h = \alpha U_h^0$ . L'équation  $AU_h + CP_h = 0$  conduit à

$$p_{j+1} - p_{j-1} = (-1)^j \frac{4\mu\alpha}{h} \quad \text{pour tout } 1 \leq j \leq n,$$

c'est-à-dire que  $p_{2k+1} = p_1 + k \frac{4\mu\alpha}{h}$  et  $p_{2k} = p_0 - k \frac{4\mu\alpha}{h}$ . Les valeurs  $\alpha$ ,  $p_0$  et  $p_1$  étant quelconques, le noyau  $\text{Ker}(K_h)$  est de dimension 3 engendré par

$$\text{Ker } K_h = \text{Vect} \left( \begin{pmatrix} U^0 \\ P_h^0 \end{pmatrix}, \begin{pmatrix} 0 \\ P_h^1 \end{pmatrix}, \begin{pmatrix} 0 \\ P_h^2 \end{pmatrix} \right)$$

avec  $P_h^0 = \frac{4\mu}{h}(0, 0, 1, -1, 2, -2, \dots, m-1, -(m-1), m)$ ,  $P_h^1 = (0, 1, \dots, 0, 1, 0)$  et  $P_h^2 = (1, 0, \dots, 1, 0, 1)$ . Par simple addition,  $P_h^1 + P_h^2 = \mathbb{1}$ , on retrouve la discrétisation d'une pression constante. Le cas  $n$  impair est encore plus mal posé que le cas  $n$  pair puisque le noyau  $\text{Ker}(K_h)$  est encore plus grand et contient un mode d'oscillations couplées de la vitesse et de la pression. Dans tous les cas on a bien une indétermination de la pression suivant un mode d'oscillation du type annoncé  $(1, 0, 1, 0, \dots, 1, 0)$ .



## Chapitre 6

# PROBLÈMES AUX VALEURS PROPRES

**Exercice 6.1.1** Soit  $\Omega = \mathbb{R}^N$ . Montrer que  $u(x) = \exp(ik \cdot x)$  est une solution de

$$-\Delta u = \lambda u \quad (6.1)$$

si  $|k|^2 = \lambda$ . Une telle solution est appelée onde plane.

**Correction.** Soit  $u(x) = \exp(ik \cdot x)$  avec  $k \in \mathbb{R}^N$ , on a  $\nabla u(x) = i \exp(ik \cdot x)k$  et

$$\Delta u = \operatorname{div}(\nabla u) = -|k|^2 \exp(ik \cdot x).$$

Ainsi,  $u$  est solution de l'équation (6.1) dès que  $|k|^2 = \lambda$ . Sur cet exemple, on voit que le Laplacien dans un domaine non borné peut admettre une infinité non dénombrable de valeurs propres ("généralisées", car la "fonction propre"  $\exp(ik \cdot x)$  n'appartient pas à  $L^2(\mathbb{R}^N)$ ).

**Exercice 6.1.2** Soit un potentiel régulier  $V(x)$ . Montrer que, si  $\mathbf{u}(x, t) = e^{-i\omega t}u(x)$  est solution de

$$i \frac{\partial \mathbf{u}}{\partial t} + \Delta \mathbf{u} - V \mathbf{u} = 0 \quad \text{dans } \mathbb{R}^N \times \mathbb{R}_*^+, \quad (6.2)$$

alors  $u(x)$  est solution de

$$-\Delta u + Vu = \omega u \quad \text{dans } \mathbb{R}^N. \quad (6.3)$$

**Correction.** Il suffit d'effectuer le calcul. En effet,

$$\begin{aligned} i \frac{\partial \mathbf{u}}{\partial t}(x, t) &= e^{-i\omega t} \omega u(x) \\ \Delta \mathbf{u}(x, t) &= e^{-i\omega t} \Delta u(x). \end{aligned}$$

Comme  $\mathbf{u}$  est solution de l'équation de Schrödinger (6.2), on en déduit que

$$-\Delta u + Vu = \omega u.$$

**Exercice 6.1.3** 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 (6.3) si  $\omega = \text{tr}(A^{1/2})$ . Une telle solution est appelée état fondamental.

**Correction.** Rappelons tout d'abord que la matrice  $A^{1/2}$  est définie comme la matrice dont les vecteurs propres sont identiques à ceux de la matrice  $A$  et dont les valeurs propres sont les racines carrées des valeurs propres de  $A$ . Plus précisément, la matrice  $A$  étant symétrique définie positive, elle admet une base orthonormale de vecteurs propres. Il existe donc une matrice unitaire  $U$  et une matrice diagonale  $D$  telles que  $A = UDU^*$ . Les coefficients de  $D$  sont les valeurs propres (positives) de la matrice  $A$ . On a alors  $A^{1/2} = UEU^*$ , où  $E = D^{1/2}$  est la matrice diagonale définie par  $E_{ii} = D_{ii}^{1/2}$ . La matrice  $A^{1/2}$  est évidemment indépendante de la base de vecteurs propres choisie, c'est-à-dire du choix de  $U$ .

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

$$\Delta u = \text{div}(\nabla u) = -\text{div}(uA^{1/2}x).$$

On rappelle que pour toute fonction  $f$  à valeurs réelles et  $\sigma$  à valeurs vectorielles,  $\text{div}(f\sigma) = \nabla f \cdot \sigma + f(\text{div}\sigma)$ . Ainsi,

$$\Delta u = -(A^{1/2}x) \cdot \nabla u - (\text{div}(A^{1/2}x))u = (Ax \cdot x)u - \text{tr}(A^{1/2})u,$$

et  $u$  est bien solution de l'équation

$$-\Delta u + Vu = \text{tr}(A^{1/2})u.$$

**Exercice 6.2.1** Montrer que l'application identité  $\text{Id}$  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  $\text{Id}$  est évidemment la boule unité. Si l'application  $\text{Id}$  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 6.2.2** Soit l'espace de Hilbert  $\ell_2$  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  $\ell^2$ ,

$$\|Ax\|_{\ell^2}^2 = \sum_i |a_i x_i|^2 \leq \sup_i |a_i|^2 \sum_i |x_i|^2 = \sup_i |a_i|^2 \|x\|_{\ell^2}^2.$$

Ainsi,  $A$  est une application continue de  $\ell^2$  dans  $\ell^2$ .

Supposons que  $\lim a_i = 0$ . Soit  $x^n$  une suite d'éléments de la boule unité de  $\ell^2$ . 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  $\ell^2$ . Pour tout entier  $k \geq 0$ , on note  $x^{n,k}$  la sous-suite extraite de  $x^n$  telle que  $y^{n,k} = Ax^{n,k}$ . 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  $n$ ,

$$\sum_{i>l} |z_i^n|^2 = \sum_{i>l} |y_i^{n,n}|^2 = \sum_{i>l} |a_i|^2 |x_i^{n,n}|^2 \leq \varepsilon^2 \|x^{n,n}\|_{\ell^2}^2 \leq \varepsilon^2.$$

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 entier  $n$  positif, il existe  $i > n$  tel que  $|a_i| > M$ . On peut donc définir une suite  $x^n$  de  $\ell^2$  et une suite  $i_n$  d'entiers naturels croissants telles que, pour tout indice  $k$ ,

$$x_k^n = \delta_k^{i_n}, \text{ et } |a_{i_n}| > M$$

et  $i_n$  strictement croissante. Autrement dit, toutes les composantes du vecteur  $x^n$  sont nulles, sauf la  $i_n$ -ième qui est égale à 1. On pose  $y^n = Ax^n$ . La suite  $x^n$  est bornée dans  $\ell^2$ , tandis que la suite  $y^n$  d'éléments de  $\ell^2$  n'admet pas de sous-suite convergente. En effet, pour tout  $n$  et  $p$  tels que  $n \neq p$ , on a

$$\|y^n - y^p\|_{\ell^2}^2 = |a_{i_n}|^2 + |a_{i_p}|^2 > 2M^2.$$

Ainsi,  $A$  n'est pas compacte.

**Exercice 6.2.3** 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 (voir l'Exercice 6.2.1).

**Exercice 6.2.4** Soit  $V$  un espace de Hilbert réel de dimension infinie et  $A$  une application linéaire continue, définie positive, auto adjointe, compacte de  $V$  dans  $V$ . On note  $u_k$  et  $\lambda_k$  les valeurs et vecteurs propres de  $A$ . Montrer que, pour  $v \in V$ , l'équation  $Au = v$  admet une unique solution  $u \in V$  si et seulement si  $v$  vérifie

$$\sum_{k=1}^{+\infty} \frac{|\langle v, u_k \rangle|^2}{\lambda_k^2} < +\infty. \quad (6.4)$$

**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 (6.4),

$$u = \sum_k \frac{\langle v, u_k \rangle}{\lambda_k} u_k$$

appartient à  $V$  (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 6.2.5** 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  $L^2(0,1)$ ,

$$(Af, g)_{L^2(0,1)} = \int_0^1 (Af)g dx = \int_0^1 (x^2 + 1)fg dx = (f, Ag)_{L^2(0,1)}.$$

Ainsi,  $A$  est auto-adjointe. De plus,  $A$  est positive car

$$(Af, f)_{L^2(0,1)} = \int_0^1 (x^2 + 1)|f(x)|^2 dx \geq 0.$$

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  $\lambda$ . Dans ce cas, pour toute fonction  $g \in L^2(0,1)$ ,

$$\int_0^1 (x^2 + 1)f(x)g(x) dx = (Af, g)_{L^2(0,1)} = \lambda(f, g)_{L^2(0,1)} = \lambda \int_0^1 f(x)g(x) dx.$$

On en déduit que

$$((x^2 + 1)f - \lambda f, g(x))_{L^2(0,1)} = 0.$$

En choisissant  $g = (x^2 + 1 - \lambda)f$ , on en déduit que

$$\|(x^2 + 1 - \lambda)f\|_{L^2(0,1)} = 0$$

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  $\lambda$  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 6.3.1** 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  $\lambda$  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  $\lambda_k$  est une suite non bornée, croissante de réels positifs.

**Exercice 6.3.2** 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

$$\begin{cases} -\Delta u_k = \lambda_k u_k & \text{p.p. dans } \Omega \\ u_k = 0 & \text{p.p. sur } \partial\Omega. \end{cases} \quad (6.5)$$

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.** On cherche à déterminer les fonctions  $u \in H_0^1(0, 1)$  telles que

$$u'' + \lambda u = 0. \quad (6.6)$$

Tout d'abord, comme  $u \in H_0^1(0, 1)$ ,  $u$  est continue, et d'après l'équation (6.6),  $u$  est de classe  $C^2$ . Par récurrence, il s'en suit que  $u$  est en fait de classe  $C^\infty$ . Ainsi,

l'équation (6.6) est une équation différentielle classique. De plus, si elle admet une solution non nulle (avec  $u(0) = u(1) = 0$ ), on a nécessairement  $\lambda > 0$ . En effet,

$$\int_0^1 |u'|^2 dx = - \int_0^1 u'' u dx = \lambda \int_0^1 |u|^2 dx$$

et  $\lambda = \int_0^1 |u'|^2 dx / \int_0^1 |u|^2 dx > 0$ . Il est bien connu que les solutions de l'équation différentielle ordinaire (6.6) 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 non nul (car  $u(1) = 0$ ). Les vecteurs propres du Laplacien unidimensionnel avec conditions aux limites de Dirichlet sont donc les fonctions

$$u_k = \sqrt{2} \sin(k\pi x)$$

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  $a(u, v) = \int_0^1 u'v' dx$  est une forme bilinéaire symétrique, continue et coercive sur  $H_0^1(]0, 1[)$ , on peut appliquer le Théorème 7.3.2. Ainsi,  $(u_k/k\pi)_{k \geq 1}$  est une base de hilbertienne  $H^1(]0, 1[)$  et  $(u_k)_{k \geq 1}$  une base hilbertienne de  $L^2(]0, 1[)$ . On en déduit que la série

$$\sum_{k=1}^{\infty} a_k \sin(kx)$$

converge dans  $L^2(]0, 1[)$  si et seulement si  $\sum_k a_k^2 < \infty$  et dans  $H_0^1(]0, 1[)$  si et seulement si  $\sum_k k^2 a_k^2 < \infty$ .

**Exercice 6.3.3** 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 (6.5).

**Correction.** Soit  $u_k(x) = \sqrt{2} \sin(k\pi x)$  les fonctions propre du Laplacien avec conditions de Dirichlet sur  $]0, 1[$ . Pour tout indice  $1 \leq p \leq N$ , et tout  $k \in \mathbb{N}_*$ , on introduit la fonction  $u_{p,k}$  de  $]0, L_p[$  à valeurs dans  $\mathbb{R}$  définie par

$$u_{p,k}(x_p) = u_k(x_p/L_p).$$

Enfin, pour tout  $k = (k_1, \dots, k_N) \in \mathbb{N}_*^N$  et tout  $x = (x_1, \dots, x_N) \in \mathbb{R}^N$ , on pose

$$v_k(x) = \prod_{p=1}^N u_{p,k_p}(x_p).$$

On vérifie sans peine que pour  $k \in \mathbb{N}_*^N$ ,  $v_k$  est une fonction propre du Laplacien sur  $\Omega$  avec conditions aux bords de Dirichlet de valeur propre

$$\lambda_k = \sum_{p=1}^N (k_p \pi / L_p)^2.$$

Il est aussi aisé de vérifier que la famille  $v_k$  est orthonormale dans  $L^2(\Omega)$ . Pour conclure, il reste à prouver que la famille  $v_k$  forme une base de  $L^2(\Omega)$ , c'est-à-dire que si  $w \in L^2(\Omega)$  vérifie

$$\langle v_k, w \rangle_{L^2(\Omega)} = 0 \text{ pour tout } k \in \mathbb{N}_*^p, \quad (6.7)$$

alors  $w = 0$ . Procédons par récurrence sur la dimension  $N$ . Ce résultat est vrai pour  $N = 1$ . Supposons que le résultat soit établi pour  $\Omega$  de dimension  $N - 1$ . On introduit la fonction  $\tilde{w} \in L^2(]0, L_N[)$  définie par

$$\tilde{w}(x_N) = \int_{\tilde{\Omega}} w(x) \prod_{p=1}^{N-1} u_{p, k_p}(x_p) d\tilde{x},$$

où  $\tilde{\Omega} = ]0, L_1[ \times \dots \times ]0, L_{N-1}[$  et  $\tilde{x} = (x_1, \dots, x_{N-1})$ . D'après (6.7), 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  $u_{N, k}$  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

$$\int_{\tilde{\Omega}} w_{x_N}(\tilde{x}) \prod_{p=1}^{N-1} u_{k_p}(x_p) d\tilde{x} = 0,$$

et d'après l'hypothèse de récurrence,  $w_{x_N} = 0$ , ce qui achève la démonstration.

**Exercice 6.3.4** On considère à nouveau un ouvert  $\Omega$  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 aux limites de Neumann sur  $]0, 1[$  sont, pour  $k \geq 0$ , les fonctions

$$u_k(x) = \cos(k\pi x)$$

de valeurs propres  $k^2\pi^2$  (Attention : ici, la collection des valeurs propres démarre à  $k = 0$ ). En suivant le même raisonnement que lors de l'Exercice 6.3.3, on montre que les fonctions propres du Laplacien avec conditions de Neumann sur  $\Omega = ]0, L_1[ \times \dots \times ]0, L_p[$  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

$$\lambda_k = \sum_{p=1}^N (k_p \pi / L_p)^2.$$

**Exercice 6.3.5** 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  $\lambda_1$  de (6.5).

**Correction.** Soit  $(u_k)_{k \geq 1}$ , base hilbertienne de  $L^2(\Omega)$ , fonctions propres du Laplacien avec conditions aux limites de Dirichlet (6.5) et  $\lambda_k$  les valeurs propres associées (ordonnées par ordre croissant). Soit  $u$  un élément de  $H_0^1(\Omega)$ .

$$\|u\|_{L^2(\Omega)}^2 = \sum_{k \geq 1} |\langle u, u_k \rangle_{L^2(\Omega)}|^2 \leq \lambda_1^{-1} \sum_k \lambda_k |\langle u, u_k \rangle_{L^2(\Omega)}|^2 = \lambda_1^{-1} \|\nabla u\|_{L^2(\Omega)}^2.$$

Ainsi, l'inégalité de Poincaré

$$\int_{\Omega} |v(x)|^2 dx \leq C \int_{\Omega} |\nabla v(x)|^2 dx. \quad (6.8)$$

est vérifiée pour  $C = \lambda_1^{-1}$ . Cette valeur est optimale car  $\|u_1\|_{L^2(\Omega)}^2 = \lambda_1^{-1} \|\nabla u_1\|_{L^2(\Omega)}^2$ .

**Exercice 6.3.6** Soit  $\Omega$  un ouvert borné régulier et connexe. Montrer que la première valeur propre du Laplacien dans  $\Omega$  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 pour la fonction propre constante, car

$$\begin{cases} \Delta 1 = 0 & \text{dans } \Omega \\ \frac{\partial 1}{\partial n} = 0 & \text{sur } \partial\Omega. \end{cases}$$

Si  $\lambda$  est une valeur propre du Laplacien de fonction propre  $u$ , on a

$$\|\nabla u\|_{L^2(\Omega)}^2 = \lambda \|u\|_{L^2(\Omega)}^2.$$

Ainsi, les valeurs propres du Laplacien avec conditions aux limites de Neumann sont strictement positives sauf si  $\|\nabla u\|_{L^2(\Omega)} = 0$  auquel cas  $\lambda = 0$ . Comme  $\Omega$  est connexe, si  $\lambda = 0$  la seule fonction propre associée possible est  $u(x) = \text{Cte}$  dans  $\Omega$ . Ainsi, la première valeur propre du Laplacien avec condition aux limites de Neumann est 0 et elle est simple.

**Exercice 6.3.7** Soit  $\Omega$  un ouvert borné régulier connexe de classe  $C^1$  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  $(u_k)_{k \geq 1}$  de

$$H := \left\{ v \in L^2(\Omega)^N \text{ tel que pour tout } \varphi \in H_0^1(\Omega), \int_{\Omega} v \cdot \nabla \varphi dx = 0 \right\}$$

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

$$\begin{cases} \nabla p_k - \mu \Delta u_k = \lambda_k u_k & \text{p.p. dans } \Omega \\ \operatorname{div} u_k = 0 & \text{p.p. dans } \Omega \\ u_k = 0 & \text{p.p. sur } \partial\Omega. \end{cases}$$

On admettra que l'espace

$$V := \{v \in H_0^1(\Omega)^N : \operatorname{div}(v) = 0 \text{ p.p dans } \Omega\}$$

est dense dans  $H$ .

**Correction.** Notons tout d'abord que  $H$  est un espace de Hilbert, en tant que sous-espace fermé de  $L^2(\Omega)^N$ . On munit  $V$  du produit scalaire

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

D'après le théorème de Rellich, l'injection de  $V$  dans  $H$  est compacte. De plus comme l'espace  $V$  est dense dans  $H$  on peut appliquer le théorème (7.3.2) d'où l'on déduit l'existence d'une famille positive et croissante de valeurs propres  $\lambda_k$  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 le 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  $\Omega$ ). 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 6.3.8** 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)

$$-\Delta u + Vu = \lambda u \quad \text{dans } \mathbb{R}^N. \quad (6.9)$$

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 (6.9). 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,

$$\begin{aligned} \|u\|_{L^2(\mathbb{R}^N)}^2 &\leq \int_{B_1} |u|^2 dx + \int_{\mathbb{R}^N \setminus B_1} |x|^2 |u|^2 dx \\ &\leq C \left( \int_{B_2} |\nabla u|^2 dx + \int_{B_2 \setminus B_1} |u|^2 dx \right) + \int_{\mathbb{R}^N \setminus B_1} |x|^2 |u|^2 dx \\ &\leq (C + 1) \|u\|_V^2. \end{aligned}$$

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  $H^1(\mathbb{R}^N)$ .



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  $(u_n)_{n \in \mathbb{N}}$  une suite bornée de  $V$ ,  $\|u_n\|_V^2 < M$ . Nous allons construire une sous-suite de  $(u_n)_{n \in \mathbb{N}}$  dont les restrictions à tout borné sont convergentes en norme  $L^2$  par un procédé d'extraction diagonal.

Dans un premier temps, on construit par récurrence une suite  $((u_n^k)_{n \in \mathbb{N}})_{k \in \mathbb{N}}$  de sous-suites de  $(u_n)_{n \in \mathbb{N}}$  telle que pour tout  $k \in \mathbb{N}_*$ , la suite  $(u_n^k)_{n \in \mathbb{N}}$  soit convergente sur la boule  $B_k$  de rayon  $k$ , centrée en l'origine.

On pose  $(u_n^0)_{n \in \mathbb{N}} = (u_n)_{n \in \mathbb{N}}$ . Soit  $k$  un entier naturel. Supposons  $(u_n^k)_{n \in \mathbb{N}}$  soit déjà construite. On note  $v_n^k$  la restriction de  $u_n^k$  à la boule  $B_{k+1}$ . Par hypothèse,  $(u_n^k)_{n \in \mathbb{N}}$  est une suite bornée de  $V$ . On en déduit que  $v_n^k$  est borné dans  $H^1(B_{k+1})$ . D'après le Théorème de Rellich, il existe une sous-suite  $(v_{n_p}^k)_{p \in \mathbb{N}}$  de  $v_n^k$  convergente dans  $L^2(B_{k+1})$ . On définit alors  $(u_p^{k+1})_{p \in \mathbb{N}}$  comme la suite extraite de  $(u_n^k)_{n \in \mathbb{N}}$  définie par  $u_p^{k+1} = u_{n_p}^k$ .

La suite  $(u_n^n)_{n \in \mathbb{N}}$  est convergente dans  $L^2(B_k)$  pour tout entier naturel  $k$ . Dans la suite, on note  $u_n$  cette suite. Comme  $u_n$  est convergente sur toute boule bornée, elle est convergente presque partout. On note  $u$  sa limite. Notons que la restriction de  $u$  à toute boule  $B_k$  appartient à  $L^2(B_k)$  et que la restriction de  $u_n$  à  $B_k$  converge vers la restriction de  $u$  à cette même boule dans  $L^2(B_k)$ .

Soit  $\varepsilon$  un réel positif. On pose  $\alpha = (5M/\varepsilon)^{1/2}$ . On a

$$\begin{aligned} \int_{\mathbb{R}^N} |u - u_n|^2 dx &\leq \int_{|x| < \alpha} |u - u_n|^2 dx + 1/\alpha^2 \int_{|x| > \alpha} |x|^2 |u - u_n|^2 dx \\ &\leq \int_{|x| < \alpha} |u - u_n|^2 dx + 4M/\alpha^2. \end{aligned}$$

Pour  $n$  assez grand,  $\|u - u_n\|_{L^2(B_\alpha)}^2 \leq M/\alpha^2$  et

$$\int_{\mathbb{R}^N} |u - u_n|^2 dx \leq 5M/\alpha^2 = \varepsilon.$$

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  $V$ . L'injection de  $V$  dans  $L^2(\mathbb{R}^N)$  est compacte et  $V$  est dense dans  $L^2(\mathbb{R}^N)$  ( $V$  contient les fonctions  $C^\infty$  à support compact). 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 (6.9) et dont les valeurs propres associées  $\lambda_k$  sont positives et convergent vers l'infini.

Afin de déterminer l'ensemble des fonctions propres de (6.9), on considère dans un premier temps le cas unidimensionnel ( $N = 1$ ) et  $V(x) = x^2$ . Le cas général s'en déduira aisément. On s'intéresse donc au problème aux valeurs propres

$$-u'' + |x|^2 u = \lambda u. \quad (6.10)$$

Comme proposé par l'énoncé, on cherche les fonctions propres de la forme  $u_k(x) = p_k(x)e^{-x^2/2}$ , où  $p_k$  est un polynôme de degré  $k$ . Notons que s'il existe une telle fonction propre pour tout  $k$ , toutes les fonctions propres auront été exhibées, la famille ainsi obtenue étant dense dans  $V$ . On note  $\lambda_k$  la valeur propre associée à  $u_k$ . D'après (6.10), si  $u_k$  est de la forme suggérée, on vérifie après un simple calcul de dérivation de fonctions produits que le polynôme  $p_k$  est solution de l'équation différentielle

$$-p_k'' + (1+x)p_k' + p_k = \lambda_k p_k. \quad (6.11)$$

On en déduit d'ores et déjà une condition nécessaire sur  $\lambda_k$ . Si on suppose que le terme de plus haut degré de  $p_k$  est  $x^k$  (ce qui est toujours possible,  $p_k$  étant défini à une constante multiplicative près), le terme de plus haut degré du membre de gauche de l'équation (6.11) est  $(k+1)x^k$  tandis que celui de gauche est  $\lambda_k x^k$ . On a donc

$$\lambda_k = k + 1.$$

On cherche donc  $p_k$  de la forme

$$-p_k'' + (1+x)p_k' - kp_k = 0. \quad (6.12)$$

En posant  $X = x + 1$  et  $q_k(X) = p_k(x)$ , déterminer  $p_k$  équivaut à rechercher un polynôme  $q_k$  de degré  $k$  tel que

$$q_k'' - Xq_k' + kq_k = 0.$$

Il est aisé de vérifier à la main que cette équation différentielle admet une unique solution polynomiale de degré  $k$  (toujours à une constante multiplicative près). En effet, si on décompose  $q_k$  sous la forme

$$q_k(X) = \sum_{i=0}^k a_i X^i,$$

on obtient une relation de récurrence très simple entre les coefficients  $a_i$ , à savoir pour tout entier  $i$ ,

$$(i-k)a_i = a_{i+2}(i+2)(i+1),$$

qui permet de déterminer  $a_i$  en fonction de  $a_{i+2}$ , sauf (et heureusement, sinon tous les termes seraient nuls) pour  $i = k$  où on peut fixer  $a_k$  de manière arbitraire (on

choisit  $a_k = 1$ ). Notons enfin que cette relation définit bien un polynôme car elle implique toujours  $a_{-1} = 0$  et  $a_{-2} = 0$  (et de ce fait  $a_i = 0$  pour  $i < 0$ ). On a donc établi que l'équation (6.12) admet une unique solution polynomiale de degré  $k$ . Les polynômes  $q_k$  sont connus sous le nom de polynômes d'Hermite et peuvent être alternativement définis par la relation de récurrence

$$q_k = Xq_{k-1} - (k-1)q_{k-2},$$

$q_0 = 1$  et  $q_1 = X$ . Les fonctions propres de (6.10) sont donc de la forme  $q_k(x+1)e^{-x^2/2}$  et de valeurs propres  $\lambda_k = k+1$ . Cherchons à étendre ce résultat au cas général. Tout d'abord, toujours dans le cas unidimensionnel les fonctions propres de

$$-u''_\alpha + \alpha|x|^2u_\alpha = \lambda_\alpha u_\alpha. \quad (6.13)$$

se déduisent aisément du cas  $\alpha = 1$ . En effet, on vérifie (par un simple changement de variable) que si  $u$  est fonction propre de (6.10) de valeur propre  $\lambda$ , alors  $u_\alpha(x) = u(\alpha^{1/4}x)$  est fonction propre de (6.13) de valeur propre  $\lambda_\alpha = \sqrt{\alpha}\lambda$ . Ainsi, les fonctions propres de (6.13) sont les fonctions  $u_{\alpha,k} = q_k(\alpha^{1/4}x+1)e^{-\alpha^{1/2}x^2/2}$  de valeurs propres  $\lambda_{\alpha,k} = \sqrt{\alpha}(k+1)$ .

Le cas  $N > 1$  se déduit du cas unidimensionnel par diagonalisation de la matrice  $A$ . En effet, la matrice  $A$  étant symétrique, définie positive, elle admet une base de vecteurs propres. En se plaçant dans une telle base, l'équation (6.9) se réécrit sous la forme

$$\sum_{j=1}^N -\frac{\partial^2 u}{\partial x_j^2} + \alpha_j x_j^2 u = \lambda u, \quad (6.14)$$

où les  $\alpha_j$  sont les valeurs propres de la matrice  $A$ . Si on recherche  $u(x)$  sous la forme  $u(x) = v_1(x_1)v_2(x_2)\cdots v_N(x_N)$ , on obtient que  $u$  vérifie (6.14) si et seulement si chacun des  $v_j$  est une fonction propre de (6.13) avec  $\alpha = \alpha_j$ . De plus, la valeur propre  $\lambda$  est égale à la somme des valeurs propres associées aux vecteurs propres  $v_j$ . A tout multi-indice  $\sigma = (\sigma_1, \dots, \sigma_N) \in \mathbb{N}^N$ , on peut donc associer un vecteur propre  $u_\sigma$  solution de (6.14) de la forme

$$u_\sigma(x) = \prod_{j=1}^N u_{\alpha_j, \sigma_j}(x_j) = \left( \prod_{j=1}^N q_{\sigma_j}(\alpha_j^{1/4}x+1) \right) e^{-A^{1/2}x \cdot x/2}$$

de valeur propre

$$\lambda_\sigma = \sum_{j=1}^N \lambda_{\alpha_j, \sigma_j} = \sum_{j=1}^N \sqrt{\alpha_j}(\sigma_j + 1).$$

Enfin, on a bien obtenu toutes les fonctions propres de  $V$ , l'espace engendré par les  $u_\sigma$  étant dense dans  $L^2(\mathbb{R}^N)$ .

**Exercice 6.3.9** Soit  $\Omega$  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  $H_0^2(\Omega)$  (voir Exercice 4.3.9). Comme l'injection de  $H_0^2(\Omega)$  dans  $L^2(\Omega)$  est compacte et que  $H_0^2(\Omega)$  est dense dans  $L^2(\Omega)$ , la conclusion découle de l'application du Théorème 7.3.2.

**Exercice 6.4.1** 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 (5.18), 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  $\phi_i$  sont les fonctions de base des éléments finis  $\mathbb{P}_1$ . Pour tout  $i$  et  $j$  tels que  $|i - j| > 1$ , les supports de  $\phi_i$  et  $\phi_j$  sont disjoints et

$$(\mathcal{M}_h)_{ij} = 0.$$

Si  $j = i + 1$ ,

$$\begin{aligned} (\mathcal{M}_h)_{ij} &= \int_{ih}^{(i+1)h} \phi_i(x) \phi_{i+1}(x) dx = \int_{ih}^{(i+1)h} \frac{((i+1)h - x)}{h} \frac{(x - ih)}{h} dx \\ &= h^{-2} \int_0^h (h - x)x dx = h/6. \end{aligned}$$

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 bien du type annoncé.

Soit  $(U, \lambda) \in \mathbb{R}^n \times \mathbb{R}$  ( $n = h^{-1} - 1$ ) tels que

$$\mathcal{M}_h U = \lambda U \quad (6.15)$$

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

$$u_h(x) = \begin{cases} 0 & \text{si } x \in [0, h/2[ \\ U_j & \text{si } x \in [jh - h/2, jh + h/2[, 1 \leq j \leq n \\ 0 & \text{si } x \in [1 - h/2, 1[ \end{cases} \quad (6.16)$$

D'après (6.15), pour tout  $x \in \mathbb{R}$ , on a

$$h \frac{u_h(x-h) + 4u_h(x) + u_h(x+h)}{6} = \lambda u_h(x). \quad (6.17)$$

**Remarque 6.4.1** On a choisit  $u_h$  impaire de période 2 afin que l'équation (6.17) 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 (6.17), 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

$$\lambda_k = \frac{h}{3}(2 + \cos(kh\pi)), \quad \text{où } k \in \mathbb{Z}. \quad (6.18)$$

Enfin, en remarquant que

$$\{\lambda_k \text{ tel que } k \in \{0, \dots, n+1\}\} = \{\lambda_k \text{ tel que } k \in \mathbb{Z}\},$$

on peut limiter notre analyse aux entiers  $k \in \{0, \dots, n+1\}$ .

Réciproquement, pour tout  $k \in \{1, \dots, n\}$ , les fonctions  $u_h(x)$  vérifiant l'équation (6.17) 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 (6.16), il est nécessaire que  $u_h$  soit impaire, à valeur réelles. On en déduit alors qu'on a

$$\hat{u}_{k+2(n+1)j} = -\hat{u}_{-(k+2(n+1)j)},$$

et que les coefficients de Fourier  $\hat{u}_m$  sont imaginaires purs. Par conséquent, 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  $\lambda_0$  et  $\lambda_{n+1}$  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  $\lambda_k$  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 (5.18). Pour tout entier  $i$  et  $j$ , on obtient

$$(\widetilde{\mathcal{M}}_h)_{ij} = \sum_{k=1}^n h/2(\phi_i(hk)\phi_j(hk) + \phi_i(h(k+1))\phi_j(h(k+1))) = h \sum_{k=1}^n \delta_i^k \delta_j^k = h\delta_i^j.$$

Donc  $\widetilde{\mathcal{M}}_h = h \text{Id}$ . 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

$$\mathcal{K}_h = 6h \text{Id} - 6\mathcal{M}_h$$

que les valeurs propres du problème spectral sont de la forme

$$\lambda_k = h(2 - 2 \cos(kh\pi)),$$

et  $k \in \{1, \dots, n\}$ .

# Chapitre 7

## PROBLÈMES D'ÉVOLUTION

**Exercice 7.2.1** Soit  $\Omega$  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

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = f & \text{p.p. dans } \Omega \times ]0, T[ \\ u = 0 & \text{p.p. sur } \partial\Omega \times ]0, T[ \\ u(x, 0) = u_0(x) & \text{p.p. dans } \Omega. \end{cases} \quad (7.1)$$

1. En supposant que la solution  $u$  de (7.1) est assez régulière dans  $]0, T[ \times \Omega$ , montrer que, pour tout  $t \in [0, T]$ , on a l'égalité d'énergie suivante

$$\begin{aligned} \frac{1}{2} \int_{\Omega} u(x, t)^2 dx + \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds &= \frac{1}{2} \int_{\Omega} u_0(x)^2 dx \\ &+ \int_0^t \int_{\Omega} f(x, s) u(x, s) dx ds. \end{aligned} \quad (7.2)$$

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

$$z(t) \leq a e^{bt} \quad \forall t \in [0, T].$$

3. En appliquant le lemme de Gronwall avec  $z(t) = \frac{1}{2} \int_{\Omega} u(x, t)^2 dx$ , déduire de (7.2) que, pour tout  $t \in [0, T]$ ,

$$\begin{aligned} \frac{1}{2} \int_{\Omega} u(x, t)^2 dx + \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds &\leq \frac{e^t}{2} \left( \int_{\Omega} u_0(x)^2 dx \right. \\ &\left. + \int_0^T \int_{\Omega} f(x, s)^2 dx ds \right). \end{aligned} \quad (7.3)$$



**Correction.**

1. En intégrant le produit de l'équation d'évolution par  $u$  sur  $\Omega$ , on obtient

$$\int_{\Omega} \left( \frac{\partial u}{\partial t} u - \Delta u u \right) dx = \int_{\Omega} f u dx.$$

Par intégration par parties et en échangeant l'opérateur de dérivation en temps et l'intégrale, il vient

$$\frac{d}{dt} \left( \frac{1}{2} \int_{\Omega} u^2 dx \right) + \int_{\Omega} |\nabla u|^2 dx = \int_{\Omega} f u dx.$$

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

$$z(t) = \frac{1}{2} \int_{\Omega} |u(x, t)|^2 dx,$$

$$a = \frac{1}{2} \left( \int_{\Omega} |u_0(x)|^2 dx + \int_0^T \int_{\Omega} |f(x, s)|^2 dx ds \right)$$

et  $b = 1$ . D'après l'égalité d'énergie établie précédemment à la question 1 et en utilisant l'inégalité  $f u \leq (|f|^2 + |u|^2)/2$ , on a pour tout  $0 < t < T$ ,

$$\begin{aligned} z(t) &\leq z(t) + \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds \\ &\leq \frac{1}{2} \left( \int_{\Omega} |u_0(x)|^2 dx + \int_0^t \int_{\Omega} (|f(x, s)|^2 + |u(x, s)|^2) dx ds \right) \\ &\leq a + \int_0^t z(s) ds. \end{aligned}$$

D'après le lemme de Gronwall,  $z(t) \leq a e^t$ . En intégrant cette inégalité, on obtient

$$a + \int_0^t z(s) ds \leq a e^t.$$

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 7.2.2** Au vu de l'estimation

$$\int_{\Omega} u(x, t)^2 dx + \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds \leq C \left( \int_{\Omega} u_0(x)^2 dx + \int_0^t \int_{\Omega} f(x, s)^2 dx ds \right), \quad (7.4)$$

vérifiée par la solution  $u$  de (7.1), où la constante  $C$  est indépendante de  $T$ , on voit que le terme  $e^t$  n'est certainement pas optimal dans la majoration (7.3). 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 (7.2) que, pour tout  $t \in [0, T]$ ,

$$\int_{\Omega} u(x, t)^2 dx + 2 \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds \leq \left( \left( \int_{\Omega} u_0(x)^2 dx \right)^{1/2} + \int_0^t ds \left( \int_{\Omega} f(x, s)^2 dx \right)^{1/2} \right)^2. \quad (7.5)$$

**Correction.**

1. On suppose dans un premier temps que  $g$  est une fonction régulière. Soit  $\varepsilon$  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  $\varepsilon$  tend vers zéro pour obtenir l'inégalité désirée.

Si  $g$  n'est pas continue, 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 (7.2) 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 7.2.3** 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 (7.1) est assez régulière dans  $]0, T[ \times \Omega$ . Montrer que, pour tout  $t \in [0, T]$ , on a l'égalité d'énergie suivante

$$\begin{aligned} \frac{1}{2} \int_{\Omega} |\nabla u(x, t)|^2 dx + \int_0^t \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, s) \right|^2 dx ds &= \frac{1}{2} \int_{\Omega} |\nabla u_0(x)|^2 dx \\ &+ \int_0^t \int_{\Omega} f(x, s) \frac{\partial u}{\partial t}(x, s) dx ds. \end{aligned} \tag{7.6}$$

**Correction.** En multipliant l'équation (7.1) vérifiée par  $u$  par  $\frac{\partial u}{\partial t}$ , on obtient, suite à une intégration sur  $\Omega$  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 parties, 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 7.2.4** Soit  $\Omega$  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

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = f & \text{dans } \Omega \times ]0, T[ \\ \frac{\partial u}{\partial n} = 0 & \text{sur } \partial\Omega \times ]0, T[ \\ u(x, 0) = u_0(x) & \text{dans } \Omega \end{cases} \quad (7.7)$$

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 (7.7) admet donc une unique solution

$$u \in L^2(]0, T[; H^1(\Omega)) \cap C([0, T]; L^2(\Omega)).$$

**Exercice 7.2.5** Soit  $\Omega$  un ouvert borné régulier de  $\mathbb{R}^N$ . Soit  $A(x)$  une fonction de  $\Omega$  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

$$\begin{cases} \frac{\partial u}{\partial t} - \operatorname{div}(A(x) \nabla u) = f & \text{dans } \Omega \times ]0, T[ \\ u = 0 & \text{sur } \partial\Omega \times ]0, T[ \\ u(x, 0) = u_0(x) & \text{pour } x \in \Omega, \end{cases} \quad (7.8)$$

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 réelle, elle est diagonalisable dans une base orthonormée de vecteurs propres. Comme pour tout  $\xi \in \mathbb{R}^N$ ,

$$0 \leq A(x) \xi \cdot \xi \leq \beta |\xi|^2,$$

la plus grande valeur propre de  $A(x)$  est inférieure à  $\beta$  et donc  $\rho(A) = \|A\|_2 \leq \beta$  (cf. le Lemme **13.1.6**). D'après cette majoration et l'inégalité de Cauchy-Schwarz, pour tout  $u$  et  $v \in H_0^1(\Omega)$ , on a

$$|a(u, v)| \leq \beta \int_{\Omega} |\nabla u| |\nabla v| dx \leq \beta \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}.$$

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  $u \in L^2(]0, T[; H_0^1(\Omega)) \cap C([0, T]; L^2(\Omega))$  au problème aux limites (7.8).

**Exercice 7.3.1** Soit  $\Omega$  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

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - \Delta u = f & \text{p.p. dans } \Omega \times ]0, T[ \\ u = 0 & \text{p.p. sur } \partial\Omega \times ]0, T[ \\ u(x, 0) = u_0(x) & \text{p.p. dans } \Omega \\ \frac{\partial u}{\partial t}(x, 0) = u_1(x) & \text{p.p. dans } \Omega. \end{cases} \quad (7.9)$$

1. En supposant que la solution  $u$  de (7.9) 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  $\Omega$ , 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. En appliquant l'inégalité

$$2 \int_0^t \int_{\Omega} f(x, s) \frac{\partial u}{\partial t}(x, s) dx ds \leq \int_0^t \int_{\Omega} f^2(x, s) dx ds + \int_0^t \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, s) \right|^2 dx ds.$$

à celle précédemment obtenue, on obtient que

$$\begin{aligned} & \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 dx + \int_{\Omega} |\nabla u(x, t)|^2 dx \\ & \leq \int_{\Omega} |u_1(x)|^2 dx + \int_{\Omega} |\nabla u_0(x)|^2 dx + \int_0^t \int_{\Omega} f^2(x, s) dx ds + \int_0^t \int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^2 dx ds. \end{aligned}$$

D'après le Lemme de Gronwall (voir Exercice **8.2.1**) appliqué à

$$z(t) = \int_{\Omega} \left| \frac{\partial u}{\partial t}(x, t) \right|^2 dx + \int_{\Omega} |\nabla u(x, t)|^2 dx,$$

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 question de l'exercice, on déduit à l'aide de l'inégalité de Schwarz que

$$\begin{aligned} & \int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^2 dx + \int_{\Omega} |\nabla u|^2 dx \\ & \leq \int_{\Omega} (|u_1|^2 + |\nabla u_0|^2) dx + 2 \int_0^t \left( \int_{\Omega} f^2 dx \right)^{1/2} \left( \int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^2 dx \right)^{1/2} ds. \end{aligned}$$

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 l'on déduit l'estimation recherchée avec la constante  $C = 2$  en utilisant l'inégalité  $(a + b)^2 \leq 2(a^2 + b^2)$ .

**Exercice 7.3.2** 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 (7.9) 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  $\partial^{m-1}u/\partial t^{m-1}$  est elle-même solution d'une équation d'onde avec conditions de Dirichlet homogènes au bord, sans terme source.

**Exercice 7.3.3** Soit  $\Omega$  un ouvert borné régulier connexe de  $\mathbb{R}^N$ . Soit  $\rho > 0$  la densité,  $\mu > 0$  et  $\lambda$  les coefficients de Lamé d'un solide élastique tels que  $2\mu + N\lambda > 0$ . 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

$$\begin{cases} \rho \frac{\partial^2 u}{\partial t^2} - \operatorname{div} (2\mu e(u) + \lambda \operatorname{tr}(e(u)) \operatorname{Id}) = f & \text{dans } \Omega \times ]0, T[, \\ u = 0 & \text{sur } \partial\Omega \times ]0, T[, \\ u(t=0) = u_0(x) & \text{dans } \Omega, \\ \frac{\partial u}{\partial t}(t=0) = u_1(x) & \text{dans } \Omega. \end{cases} \quad (7.10)$$

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

$$\begin{cases} \rho \frac{d^2}{dt^2} \langle u(t), v \rangle_{L^2(\Omega)} + a(u(t), v) = \int_{\Omega} f \cdot v \, dx \\ u(t=0) = u_0; \quad \frac{du}{dt}(t=0) = u_1 \end{cases}$$

pour tout  $v \in H_0^1(\Omega)^N$ . La forme bilinéaire  $a$  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 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 prenant la fonction test  $v = \partial u / \partial t$  dans la formulation variationnelle, on obtient

$$\frac{d}{dt} \left( \frac{\rho}{2} \int_{\Omega} \left| \frac{\partial u}{\partial t} \right|^2 \, dx + \mu \int_{\Omega} |e(u)|^2 \, dx + \frac{\lambda}{2} \int_{\Omega} (\operatorname{div}u)^2 \, dx \right) = \int_{\Omega} f \cdot \frac{\partial u}{\partial t} \, dx.$$

L'égalité d'énergie en découle par une simple intégration en temps. En procédant comme pour l'Exercice 7.3.1 on obtient les estimations d'énergie suivantes. Tout d'abord, pour tout temps final  $T$  il existe une constante  $C(T)$ , ne dépendant pas des données autres que  $T$ , telle que

$$\int_{\Omega} \rho \left| \frac{\partial u}{\partial t} \right|^2 \, dx + \int_{\Omega} (2\mu |e(u)|^2 + \lambda(\operatorname{div}u)^2) \, dx \leq C(T) \left( \int_{\Omega} \rho |u_1|^2 \, dx + \int_{\Omega} (2\mu |e(u_0)|^2 + \lambda(\operatorname{div}u_0)^2) \, dx + \int_0^t \int_{\Omega} |f|^2 \, dx \, ds \right).$$

De plus, il existe une constante  $C$  (indépendante de toutes les données y compris  $T$ ) telle que

$$\int_{\Omega} \rho \left| \frac{\partial u}{\partial t} \right|^2 \, dx + \int_{\Omega} (2\mu |e(u)|^2 + \lambda(\operatorname{div}u)^2) \, dx \leq C \left( \int_{\Omega} \rho |u_1|^2 \, dx + \int_{\Omega} (2\mu |e(u_0)|^2 + \lambda(\operatorname{div}u_0)^2) \, dx + \left( \int_0^t \left( \int_{\Omega} |f|^2 \, dx \right)^{1/2} \, ds \right)^2 \right).$$

**Exercice 7.4.1** 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 7.4.2** Soit  $\Omega$  un ouvert borné régulier de  $\mathbb{R}^N$ . Soit  $u_0 \in L^2(\Omega)$  et  $u$  la solution du problème

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = 0 & \text{dans } ]0, +\infty[ \times \Omega \\ u(x, t) = 0 & \text{sur } ]0, +\infty[ \times \partial\Omega \\ u(x, 0) = u_0(x) & \text{dans } \Omega. \end{cases} \quad (7.11)$$

Soit  $\lambda_1$  la plus petite valeur propre du Laplacien avec conditions aux limites de Dirichlet. On note  $u_1$  la valeur propre associée normalisée en norme  $L^2$ . Montrer qu'il existe une constante positive  $C$  telle que

$$\|u(t) - \alpha_1^0 e^{-\lambda_1 t} u_1\|_{L^2(\Omega)} \leq C e^{-\lambda_2 t} \quad \forall t > 1, \quad \text{avec } \alpha_1^0 = \int_{\Omega} u_0 u_1 dx, \quad (7.12)$$

où  $\lambda_k$  désigne la  $k$ -ème valeur propre du Laplacien avec condition aux limites de Dirichlet.

**Correction.** On rappelle que la solution  $u$  de l'équation (7.11) est donnée par

$$u(t) = \sum_{k=1}^{\infty} \alpha_k^0 e^{-\lambda_k t} u_k \quad \text{et} \quad \alpha_k^0 = \int_{\Omega} u_0 u_k dx,$$

où les  $\lambda_k$  sont les valeurs propres du Laplacien avec conditions aux limites de Dirichlet et  $u_k$  est la famille de vecteurs propres associés, base orthonormale de  $L^2(\Omega)$ . 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 7.4.3** Soit  $\Omega$  un ouvert borné régulier de  $\mathbb{R}^N$ . On note  $u_1$  la première fonction propre du Laplacien dans  $\Omega$  avec condition de Dirichlet,  $\lambda_1$  la valeur propre associée. On rappelle que l'on peut choisir  $u_1 > 0$  dans  $\Omega$  (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 (7.1).

Soit  $\varepsilon > 0$ . Montrer que l'on peut trouver une constante positive  $K$  telle que

$$-Ku_1(x) \leq u(x, \varepsilon) \leq Ku_1(x) \quad \forall x \in \bar{\Omega}, \quad (7.13)$$

et en déduire qu'il existe une constante positive  $C$  telle que

$$\max_{x \in \bar{\Omega}} |u(x, t)| \leq Ce^{-\lambda_1 t} \quad \forall t > \varepsilon. \quad (7.14)$$

**Correction.** D'après la Proposition 8.4.5, pour tout  $\varepsilon > 0$ ,  $u(x, \varepsilon)$  est une fonction de classe  $C^\infty(\bar{\Omega})$ . Rappelons que  $u_1(x)$  est également une fonction régulière sur  $\bar{\Omega}$ , qu'elle est strictement positive sur  $\Omega$  et que  $\partial u_1 / \partial n > 0$  sur  $\partial\Omega$ . Par continuité de  $\partial u_1 / \partial n$  on en déduit l'existence d'une constante  $C > 0$  telle que  $\partial u_1 / \partial n \geq C > 0$  sur  $\partial\Omega$ . Soit

$$K_1 = 1 + \sup_{x \in \partial\Omega} \left| \frac{\partial u}{\partial n}(x, \varepsilon) / \frac{\partial u_1}{\partial n}(x) \right|.$$

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  $\omega$  de  $\partial\Omega$  tel que pour tout  $x \in \omega \cap \bar{\Omega}$ ,

$$v_\pm(x) \geq 0.$$

Il existe un compact  $A \subset \Omega$  tel que  $\Omega \subset A \cup \omega$ . On pose

$$K_2 = \max_{x \in A} |u(x, \varepsilon) / u_1(x)|,$$

qui est un nombre fini puisque  $u_1$  est une fonction continue strictement positive sur  $A$  (fermé), 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 (7.1) 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, pour tout  $x \in \Omega$  et tout  $t \geq \varepsilon$ ,

$$|u(x, t)| \leq \left( Ke^{\lambda_1 \varepsilon} \max_{x \in \bar{\Omega}} u_1(x) \right) e^{-\lambda_1 t}.$$

**Exercice 7.4.4** Soit  $\Omega$  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 (7.1). Montrer que

$$\|u\|_{L^\infty(\Omega \times \mathbb{R}_*^+)} \leq \|u_0\|_{L^\infty(\Omega)} + \frac{D^2}{2N} \|f\|_{L^\infty(\Omega \times \mathbb{R}_*^+)}, \quad (7.15)$$

où  $D = \sup_{x, y \in \Omega} |x - y|$  est le diamètre de  $\Omega$ . On pourra utilement introduire la fonction  $\psi \in H_0^1(\Omega)$  telle que  $-\Delta\psi = 1$  dans  $\Omega$ .

**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 (cf. Proposition 8.4.2),  $u \leq u_+$ . 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, on a déjà  $u = u_+ \geq 0$ . 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 7.4.1** 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 troncature 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 par parties sur  $\Omega$  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

$$\frac{1}{2} \int_{\Omega} |\tilde{u}(T)|^2 dx - \frac{1}{2} \int_{\Omega} |\tilde{u}(0)|^2 dx + \int_0^T \int_{\Omega} |\nabla \tilde{u}(t)|^2 dx dt = 0.$$

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  $\psi$  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  $t$ , d'après le résultat précédent on a

$$v(x, t) \leq \|u_0\|_{L^\infty(\Omega)}. \quad (7.16)$$

On a donc obtenu que  $u_+ \leq \|u_0\|_{L^\infty} + \|f\|_{L^\infty}\psi$ . Il reste à majorer  $\psi$  afin d'obtenir l'estimation souhaitée. Sans perte de généralité, on peut supposer que l'origine de  $\mathbb{R}^N$  appartient au bord de  $\Omega$ . On définit une fonction

$$\psi_+(x) = \frac{D^2 - |x|^2}{2N}.$$

Comme

$$-\Delta\psi_+ = 1 = -\Delta\psi \text{ dans } \Omega$$

et

$$\psi_+(x) \geq 0 = \psi(x) \text{ sur } \partial\Omega,$$

le principe du maximum implique que  $\psi(x) \leq \psi_+(x) \leq D^2/2N$  dans  $\Omega$ , ce qui achève la preuve.

**Exercice 7.4.5 (difficile)** Démontrer rigoureusement la Proposition **8.4.5** suivante : Soit  $\Omega$  un ouvert borné régulier de classe  $C^\infty$  de  $\mathbb{R}^N$ , et soit un temps final  $T > 0$ . Soit  $u_0 \in L^2(\Omega)$ , et  $u$  l'unique solution dans  $C([0, T]; L^2(\Omega)) \cap L^2(]0, T[; H_0^1(\Omega))$  de

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = 0 & \text{dans } ]0, T[ \times \Omega \\ u(x, t) = 0 & \text{sur } ]0, T[ \times \partial\Omega \\ u(x, 0) = u_0(x) & \text{dans } \Omega. \end{cases} \quad (7.17)$$

Alors, pour tout  $\varepsilon > 0$ ,  $u$  est de classe  $C^\infty$  en  $x$  et  $t$  dans  $\bar{\Omega} \times [\varepsilon, T]$ .

Pour cela on introduira, pour tout entier  $m \geq 0$ , l'espace

$$W^{2m}(\Omega) = \{v \in H^{2m}(\Omega), v = \Delta v = \dots = \Delta^{m-1}v = 0 \text{ sur } \partial\Omega\}, \quad (7.18)$$

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  $C^\ell([\varepsilon, T], W^{2m}(\Omega))$ .

**Correction.** On démontre d'abord l'équivalence des normes 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  $\|v\|_{H^{2(m-1)}}$  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  $u \in C^\ell([\varepsilon, T], W^{2m}(\Omega))$  pour tout entier  $\ell$ . 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  $C^\ell([\varepsilon, T], W^{2m}(\Omega))$ . Notons que toute fonction propre  $u_i$  appartient à  $W^{2m}(\Omega)$  pour tout  $m$  (car  $\Delta^m u_i = \lambda_i^m u_i = 0$  sur  $\partial\Omega$ ). 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 orthonormale de vecteurs propres 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  $\ell$ , il existe une constante  $C(\varepsilon, m, \ell)$  telle que, pour tout  $t \geq \varepsilon$  et tout indice  $j$ , on a

$$|(-\lambda_j)^{m+\ell} e^{-\lambda_j t}|^2 \leq C(\varepsilon, m, \ell).$$

Ainsi, pour tout  $t \geq \varepsilon$ , on en déduit

$$\left\| \frac{\partial^\ell (w^k - w^l)}{\partial t^\ell}(t) \right\|_{W^{2m}} \leq C(\varepsilon, m, \ell) \sum_{j=k}^l |\alpha_0^j|^2,$$

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  $u \in C^\ell([\varepsilon, T]; W^{2m}(\Omega))$ . Puisque  $\ell$  et  $m$  sont des entiers quelconques et que, d'après le Théorème **4.3.25** et la Remarque **4.3.26**,  $H^m(\Omega)$  est inclus dans  $C^k(\bar{\Omega})$  si  $m - N/2 > k$ , on conclut que  $u$  appartient à  $C^\infty(\bar{\Omega} \times [\varepsilon, T])$ .

**Exercice 7.4.6** 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

$$\begin{cases} \frac{\partial u}{\partial t} - \Delta u = f & \text{dans } ]0, +\infty[ \times \mathbb{R}^N \\ u(x, 0) = u_0(x) & \text{dans } \mathbb{R}^N. \end{cases} \quad (7.19)$$

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

$$u(t) = S(t)u_0 + \int_0^t S(t-s)f(s) ds, \quad (7.20)$$

c'est-à-dire

$$u(x, t) = \int_{\mathbb{R}^N} u_0(y) e^{-\frac{|x-y|^2}{4t}} \frac{dy}{(4\pi t)^{N/2}} + \int_0^t \int_{\mathbb{R}^N} f(y, s) e^{-\frac{|x-y|^2}{4(t-s)}} \frac{dy ds}{(4\pi(t-s))^{N/2}}.$$

**Correction.** Par mesure de commodité, on notera indifféremment la transformée de Fourier d'une fonction  $v$  par  $\hat{v}$  ou  $\mathcal{F}(v)$ . La linéarité de l'opérateur  $S(t)$  est évidente.

**Continuité de l'opérateur.** Montrons que, pour tout réel  $t$  positif ou nul,  $S(t)$  est continu de  $L^2(\mathbb{R}^N)$  dans  $L^2(\mathbb{R}^N)$ . La norme  $L^2(\mathbb{R}^N)$  de  $S(t)u_0$  étant égale à la norme  $L^2(\mathbb{R}^N)$  de sa transformée de Fourier, il suffit de vérifier la continuité de  $L^2(\mathbb{R}^N)$  dans  $L^2(\mathbb{R}^N)$  de l'application qui à  $u_0$  associe  $\mathcal{F}(S(t)u_0)$ . Notons tout d'abord que  $S(t)u_0$  est égal au produit de convolution de  $u_0$  et d'une Gaussienne

$$S(t)u_0 = \frac{1}{(4\pi t)^{N/2}} u_0 * e^{-|x|^2/4t}.$$

Or, la transformée de Fourier d'un produit de convolution est égale (à un coefficient près, du à la définition de la transformée de Fourier choisie) au produit des transformées de Fourier. Plus précisément, on a pour toutes fonctions  $v$  et  $w$  de  $L^2(\mathbb{R}^N)$

$$\mathcal{F}(v * w) = (2\pi)^{N/2} \mathcal{F}(v) \mathcal{F}(w).$$

Enfin, la transformée de Fourier d'une Gaussienne est encore une Gaussienne : pour tout réel  $a$ , on a

$$\mathcal{F}(e^{-a|x|^2})(\xi) = (2a)^{-N/2} e^{-|\xi|^2/4a}.$$

On déduit de ces deux propriétés que

$$\mathcal{F}(S(t)u_0)(\xi) = \mathcal{F}(u_0) e^{-|\xi|^2 t}.$$

Ainsi, pour tout  $u_0 \in L^2(\mathbb{R}^N)$  et tout  $t \geq 0$ ,

$$\|S(t)u_0\|_{L^2(\mathbb{R}^N)} = \|\hat{u}_0 e^{-|\xi|^2 t}\|_{L^2(\mathbb{R}^N)} \leq \|\hat{u}_0\|_{L^2(\mathbb{R}^N)} = \|u_0\|_{L^2(\mathbb{R}^N)}$$

et  $S(t)$  est un opérateur continu de  $L^2(\mathbb{R}^N)$  dans  $L^2(\mathbb{R}^N)$  de norme inférieure à 1.

**Propriété de semi-groupe.** Montrons que  $S(t)$  est un semi-groupe. Pour tout réels  $t$  et  $t'$  positifs, on a

$$\mathcal{F}(S(t+t')u_0) = \mathcal{F}(u_0)e^{-|\xi|^2 t}e^{-|\xi|^2 t'} = \mathcal{F}(S(t)u_0)e^{-|\xi|^2 t'} = \mathcal{F}(S(t')(S(t)u_0))$$

et en appliquant la transformée de Fourier inverse, on obtient que

$$S(t+t')(u_0) = S(t')(S(t)u_0).$$

Ainsi,  $S(t+t') = S(t')S(t)$ .

**Cas homogène.** On considère tout d'abord le cas homogène, c'est-à-dire  $f = 0$ . Montrons que  $u(\cdot) = S(\cdot)u_0 \in C(\mathbb{R}^+, L^2(\mathbb{R}^N))$ . Soit  $t$  et  $t'$  réels positifs ou nuls, on a

$$\begin{aligned} \|u(t') - u(t)\|_{L^2(\mathbb{R}^N)} &= \|(S(t') - S(t))u_0\|_{L^2(\mathbb{R}^N)} \\ &\leq \|u_0 - S(|t-t'|)u_0\|_{L^2(\mathbb{R}^N)} = \|(1 - e^{-|\cdot|^2|t-t'|})\hat{u}_0(\cdot)\|_{L^2(\mathbb{R}^N)}. \end{aligned}$$

Or  $|(1 - e^{-|\cdot|^2|t-t'|})\hat{u}_0|$  est borné uniformément par rapport à  $t$  et  $t'$  par  $2|\hat{u}_0|$ . Ainsi, par application du Théorème de convergence dominée de Lebesgue, on en déduit la continuité de l'application  $u$  de  $\mathbb{R}^+$  à valeurs dans  $L^2(\mathbb{R}^N)$ .

Montrons que  $u(\cdot) = S(\cdot)u_0 \in C^1(\mathbb{R}_*^+, L^2(\mathbb{R}^N))$ , ce qui est équivalent à prouver la dérivabilité de la transformée de Fourier de  $u$

$$\mathcal{F}(u(t)) = \mathcal{F}(S(t)u_0) = e^{-t|\xi|^2} \hat{u}_0.$$

Formellement, on calcule que

$$\frac{d}{dt}(\mathcal{F}(S(t)u_0)) = -|\xi|^2 e^{-t|\xi|^2} \hat{u}_0 = -|\xi|^2 \mathcal{F}(S(t)u_0).$$

Tout d'abord la fonction  $|\xi|^2 e^{-|\xi|^2 t} \hat{u}_0$  appartient effectivement à  $L^2(\mathbb{R}^N)$ . Afin de prouver qu'il s'agit bien de la dérivée de  $e^{-|\xi|^2 t} \hat{u}_0$ , il reste à établir que pour tout  $t > 0$ ,

$$\|(\delta t)^{-1}(e^{-|\xi|^2(t+\delta t)} - e^{-|\xi|^2 t} + \delta t |\xi|^2 e^{-|\xi|^2 t})\|_{L^2(\mathbb{R}^N)} = o(\delta t).$$

A cet effet, on peut utiliser la formule de Taylor avec reste intégrale suivante

$$e^{-|\xi|(t+\delta t)} = e^{-|\xi|t} - \delta t |\xi|^2 e^{-|\xi|^2 t} + (\delta t)^2 \int_0^1 |\xi|^4 e^{-|\xi|^2(t+s\delta t)} ds.$$

On en déduit que, dès que  $\delta t > -t/2$ , on a

$$\delta t^{-1}(e^{-|\xi|(t+\delta t)} - e^{-|\xi|t} + \delta t |\xi|^2 e^{-|\xi|^2 t}) = \delta t \int_0^1 |\xi|^4 e^{-|\xi|^2(t+s\delta t)} ds \leq \delta t \int_0^1 |\xi|^4 e^{-|\xi|^2 t/2} ds.$$

Or, pour  $t$  fixé, on a

$$\sup_{\xi} |\xi|^4 e^{-|\xi|^2 t/2} = C_t < \infty.$$

Il s'en suit que

$$\|(\delta t)^{-1}(e^{-|\xi|^2(t+\delta t)} - e^{-|\xi|^2 t} + \delta t|\xi|^2 e^{-|\xi|^2 t})\|_{L^2(\mathbb{R}^N)} \leq C_t \|u_0\|_{L^2(\mathbb{R}^N)} \delta t,$$

ce qui conclut la preuve de la dérivabilité de  $u$  dans le cas  $f = 0$ . En appliquant la transformée de Fourier inverse à la formule pour la dérivée de  $\mathcal{F}(S(t)u_0)$ , il vient

$$\frac{d}{dt}(S(t)u_0) = \Delta(S(t)u_0),$$

ce qui prouve que  $S(t)u_0$  est solution de l'équation de la chaleur avec  $f = 0$ .

**Cas non homogène.** D'après les résultats précédents, on a établi que dans le cas  $f = 0$ , la fonction  $u$  définie par (7.19) était solution de (7.20). D'après la linéarité de la définition (7.19) de  $u$  et de l'équation (7.20), il suffit de considérer le cas  $f \neq 0$  et  $u_0 = 0$  pour conclure. Dans ce cas, on a

$$u(t) = \int_0^t S(t-s)f(s) ds.$$

Par un simple changement de variable, on en déduit que

$$u(t) = \int_0^t S(s)f(t-s) ds.$$

Enfin, l'opérateur  $S(s)$  étant uniformément continu de  $L^2(\mathbb{R}^N)$  à valeurs dans  $L^2(\mathbb{R}^N)$  et  $f$  appartenant à  $C^1(\mathbb{R}^+, L^2(\mathbb{R}^N))$ , on en déduit que  $u$  est dérivable par rapport à  $t$  dans  $L^2(\mathbb{R}^N)$  et que

$$\frac{du}{dt} = \int_0^t S(s) \frac{df}{dt}(t-s) ds + S(t)f(0).$$

Il en découle que

$$\frac{d\hat{u}}{dt}(\xi) = \int_0^t e^{-(t-s)|\xi|^2} \frac{d\hat{f}}{dt}(s, \xi) ds + e^{-t|\xi|^2} \hat{f}(0, \xi).$$

En effectuant une intégration par parties sur le premier terme du membre de droite de cette équation, il vient

$$\begin{aligned} \frac{d\hat{u}}{dt}(\xi) &= \left[ e^{-(t-s)|\xi|^2} \hat{f}(s, \xi) \right]_0^t - |\xi|^2 \int_0^t e^{-(t-s)|\xi|^2} \hat{f}(s, \xi) ds + e^{-t|\xi|^2} \hat{f}(0, \xi) \\ &= \hat{f}(t, \xi) - |\xi|^2 \int_0^t e^{-(t-s)|\xi|^2} \hat{f}(s, \xi) ds \\ &= \hat{f}(t, \xi) - |\xi|^2 \hat{u}. \end{aligned}$$

Par transformation de Fourier inverse, on en déduit que

$$\frac{du}{dt} = \Delta u + f,$$

et on a évidemment  $u(t=0) = 0$ .



**Exercice 7.4.7 (égalité d'énergie)** Soit  $u$  la solution de l'équation de la chaleur (7.19) avec  $f = 0$ . 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  $\nabla u$  est  $i\xi\hat{u}$ . Comme la transformée de Fourier est une isométrie de  $L^2(\mathbb{R}^N)$  et  $\hat{u} = \hat{u}_0 e^{-|\xi|^2 t}$ , on a

$$\begin{aligned} \frac{1}{2} \|u(T)\|_{L^2(\mathbb{R}^N)}^2 + \int_0^T \|\nabla u(t)\|_{L^2(\mathbb{R}^N)}^2 dt &= \frac{1}{2} \|\hat{u}(T)\|_{L^2(\mathbb{R}^N)}^2 + \int_0^T \|\xi\hat{u}(t)\|_{L^2(\mathbb{R}^N)}^2 dt \\ &= \frac{1}{2} \int_{\mathbb{R}^N} |u_0(T, \xi)|^2 e^{-2|\xi|^2 T} d\xi + \int_{\mathbb{R}^N} \int_0^T |\xi|^2 |u_0(t, \xi)|^2 e^{-2|\xi|^2 t} dt d\xi \\ &= \frac{1}{2} \int_{\mathbb{R}^N} |u_0(T, \xi)|^2 e^{-2|\xi|^2 T} d\xi - \frac{1}{2} \int_{\mathbb{R}^N} \int_0^T \frac{\partial}{\partial t} (|u_0(t, \xi)|^2 e^{-2|\xi|^2 t}) dt d\xi \\ &= \frac{1}{2} \int_{\mathbb{R}^N} |u_0(\xi)|^2 d\xi = \frac{1}{2} \|u_0\|_{L^2(\mathbb{R}^N)}^2. \end{aligned}$$

**Exercice 7.4.8 (principe du maximum)** Soit  $u$  la solution de l'équation de la chaleur (7.19) avec  $f = 0$ . 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.** On majore la formule explicite pour  $u$

$$\|u(t)\|_{L^\infty(\mathbb{R}^N)} \leq \frac{\|u_0\|_{L^\infty(\mathbb{R}^N)}}{(4\pi t)^{N/2}} \int_{\mathbb{R}^N} e^{-y^2/4t} dy = \|u_0\|_{L^\infty(\mathbb{R}^N)}.$$

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 7.4.9 (effet régularisant)** Soit  $u$  la solution de l'équation de la chaleur (7.19) avec  $f = 0$ . Montrer que  $u \in C^\infty(\mathbb{R}^N \times \mathbb{R}_*^+)$ .

**Correction.** D'après l'expression de la transformée de Fourier de  $u$ ,

$$\hat{u}(\xi, t) = \hat{u}_0(\xi) e^{-|\xi|^2 t},$$

pour tout multi-indice  $\alpha$  et pour tout  $t > 0$ , la fonction  $(\xi, t) \mapsto |\xi|^\alpha \hat{u}(\xi, t)$  est continue en temps à valeurs dans  $L^2(\mathbb{R}^N)$ . 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  $C(\mathbb{R}_*^+, H^m(\mathbb{R}^N))$ . D'après le Théorème 4.3.25 et la Remarque 4.3.26 (ou leurs équivalents dans le cas  $\Omega = \mathbb{R}^N$ ), on en déduit que  $u(t) \in C(\mathbb{R}_*^+, C^\infty(\mathbb{R}^N))$ . 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 7.4.10 (comportement asymptotique)** Soit  $u$  la solution de l'équation de la chaleur (7.19) avec  $f = 0$ . 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  $r > 0$  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 obtient

$$|u(x, t)| \leq \frac{1}{(4\pi t)^{N/2}} \left( \|u_0\|_{L^2(\mathbb{R}^N)} \left( \int_{|x-y| \geq r} e^{-\frac{|x-y|^2}{2t}} dy \right)^{1/2} + \left\| e^{-\frac{|x|^2}{4t}} \right\|_{L^2(\mathbb{R}^N)} \left( \int_{|x-y| \leq r} |u_0(y)|^2 dy \right)^{1/2} \right).$$

On note  $B(z, r)$  la boule de rayon  $r$  et de centre  $z \in \mathbb{R}^N$ . On a

$$|u(x, t)| \leq \frac{1}{(4\pi t)^{N/2}} \left( \|u_0\|_{L^2(\mathbb{R}^N)} \left\| e^{-\frac{|x|^2}{4t}} \right\|_{L^2(\mathbb{R}^N \setminus B(0, r))} + \left\| e^{-\frac{|x|^2}{4t}} \right\|_{L^2(\mathbb{R}^N)} \|u_0\|_{L^2(B(x, r))} \right).$$

Pour tout réel  $\varepsilon > 0$ , pour  $r$  assez grand, on a  $\|u_0\|_{L^2(\mathbb{R}^N)} \|e^{-\frac{|x|^2}{4t}}\|_{L^2(\mathbb{R}^N \setminus B(0, r))} < \varepsilon$ . De plus, pour  $x$  assez grand ( $r$  étant fixé), comme  $u^0 \in L^2(\mathbb{R}^N)$ , on a

$$\left\| e^{-\frac{|x|^2}{4t}} \right\|_{L^2(\mathbb{R}^N)} \|u_0\|_{L^2(B(x, r))} < \varepsilon.$$

Ainsi, pour tout  $\varepsilon$ , pour  $x$  assez grand on a  $|u(x, t)| < \frac{2\varepsilon}{(4\pi t)^{N/2}}$ . En d'autres termes,

$$\lim_{|x| \rightarrow +\infty} u(x, t) = 0 \quad \text{pour tout } t > 0.$$

On rappelle que  $\hat{u}(\xi, t) = \hat{u}_0(\xi) e^{-|\xi|^2 t}$ . Ainsi,

$$\|u(t)\|_{L^2(\mathbb{R}^N)}^2 = \|\hat{u}(t)\|_{L^2(\mathbb{R}^N)}^2 = \int_{\mathbb{R}^N} |\hat{u}_0(\xi)|^2 e^{-2|\xi|^2 t} d\xi$$

et d'après le Théorème de convergence dominée de Lebesgue,  $\|u(t)\|_{L^2(\mathbb{R}^N)}$  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 le Théorème 4.3.25 et la Remarque 4.3.26, on en déduit que, pour tout entier  $r$ , la norme de  $u(t)$  dans  $C^r(\mathbb{R})$  tend vers zéro quand  $t \rightarrow \infty$ . En particulier,

$$\lim_{t \rightarrow +\infty} u(x, t) = 0 \quad \text{pour tout } x \in \mathbb{R}^N.$$

**Exercice 7.4.11 (vitesse de propagation infinie)** Soit  $u$  la solution de l'équation de la chaleur (7.19) avec  $f = 0$ . 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  $u_0 \geq 0$ . D'après le principe du maximum, on a  $u \geq 0$ . Par contraposée, il suffit donc de montrer que s'il existe  $(x, t) \in \mathbb{R}^N \times \mathbb{R}_*^+$  tel que  $u(x, t) = 0$ , alors  $u_0(y) = 0$  presque partout. Or, d'après l'expression explicite de  $u(x, t)$ , si  $u(x, t) = 0$ , on a  $u_0(y)e^{-\frac{|x-y|^2}{4t}} = 0$  pour presque tout  $y$  et donc  $u_0(y) = 0$  presque partout.

**Exercice 7.5.1** Soit  $\eta > 0$ . On considère l'équation des ondes amortie

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} + \eta \frac{\partial u}{\partial t} - \Delta u = f & \text{p.p. dans } \Omega \times \mathbb{R}_*^+ \\ u = 0 & \text{p.p. sur } \partial\Omega \times \mathbb{R}_*^+ \\ u(x, 0) = u_0(x) & \text{p.p. dans } x \in \Omega \\ \frac{\partial u}{\partial t}(x, 0) = u_1(x) & \text{p.p. dans } x \in \Omega. \end{cases} \quad (7.21)$$

On suppose que  $u$  est une solution suffisamment régulière de (7.21) 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.** Comme on s'intéresse uniquement à une propriété asymptotique de la solution et que  $f$  est nul pour  $t$  assez grand, on peut supposer, sans perte de généralité, que  $f = 0$ . Soit  $\alpha$  un réel strictement positif. On pose  $v = e^{\alpha t}u$ . La fonction  $v$  est solution du problème aux limites

$$\begin{cases} \frac{\partial^2 v}{\partial t^2} + (\eta - 2\alpha) \frac{\partial v}{\partial t} - \Delta v - \alpha(\eta - \alpha)v = 0 & \text{p.p. dans } \Omega \times \mathbb{R}_*^+ \\ v = 0 & \text{p.p. sur } \partial\Omega \times \mathbb{R}_*^+ \\ v(x, 0) = v_0(x) & \text{p.p. dans } x \in \Omega \\ \frac{\partial v}{\partial t}(x, 0) = v_1(x) & \text{p.p. dans } x \in \Omega. \end{cases}$$

où  $v_0 = u_0$  et  $v_1 = \alpha u_0 + u_1$ . En multipliant l'équation vérifiée par  $v$  dans  $\Omega$  par  $\partial v / \partial t$ , on obtient suite à une intégration par parties et un échange de l'intégration en espace et de la dérivation par rapport au temps que

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} \left( \left| \frac{\partial v}{\partial t} \right|^2 + |\nabla v|^2 - \alpha(\eta - \alpha)|v|^2 \right) dx = (2\alpha - \eta) \int_{\Omega} \left| \frac{\partial v}{\partial t} \right|^2 dx.$$

Ainsi, si  $\alpha \leq \eta/2$ , on en déduit que le second membre est négatif et qu'après intégration en temps il existe  $C_\alpha$  tel que

$$\int_{\Omega} \left( \left| \frac{\partial v}{\partial t} \right|^2 + |\nabla v|^2 - \alpha(\eta - \alpha)|v|^2 \right) dx < C_\alpha.$$

Enfin, d'après l'inégalité de Poincaré, il existe une constante  $C > 0$  telle que

$$\int_{\Omega} |\nabla v|^2 dx \geq C \int_{\Omega} |v|^2 dx,$$

ce qui conduit à la majoration suivante

$$\int_{\Omega} \left( \left| \frac{\partial v}{\partial t} \right|^2 + (C - \alpha(\eta - \alpha))|v|^2 \right) dx < C_{\alpha}.$$

Il existe  $\alpha > 0$  suffisamment petit tel que  $C - \alpha(\eta - \alpha)$  soit strictement positif aussi. Dans ce cas il découle de l'inégalité précédente que les normes  $L^2$  de  $\partial v/\partial t$  et de  $v$  sont bornées. Comme  $u = e^{-\alpha t}v$  et  $\partial u/\partial t = (\partial v/\partial t - \alpha v)e^{-\alpha t}$ ,  $u$  et  $\partial u/\partial t$  décroissent exponentiellement vers zéro (en norme  $L^2(\Omega)$ ) lorsque  $t$  tend vers l'infini.

**Exercice 7.5.2** Soit  $u(t, x)$  la solution, supposée suffisamment régulière, de l'équation des ondes (7.9). 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 (7.9) par  $u$  et on intégrera par parties.

**Correction.** En multipliant l'équation des ondes par  $u$ , on obtient par intégration

$$\int_0^t \int_{\Omega} \frac{\partial^2 u}{\partial t^2}(x, s) u(x, s) dx ds + \int_0^t \int_{\Omega} |\nabla u(x, s)|^2 dx ds = 0.$$

En intégrant par parties en temps le premier terme de cette équation, on trouve

$$\int_{\Omega} \frac{\partial u}{\partial t} u(x, t) dx - \int_{\Omega} u_1 u_0 dx + \int_0^t \int_{\Omega} \left( |\nabla u(x, s)|^2 - \left| \frac{\partial u}{\partial t}(x, s) \right|^2 \right) dx ds = 0.$$

Ainsi,

$$\begin{aligned} \frac{1}{t} \int_0^t \int_{\Omega} \left( |\nabla u(x, s)|^2 dx ds - \left| \frac{\partial u}{\partial t}(x, s) \right|^2 \right) dx ds \\ = \frac{1}{t} \left( \int_{\Omega} u_1 u_0 dx - \int_{\Omega} \frac{\partial u}{\partial t} u(x, t) dx \right) \xrightarrow{t \rightarrow +\infty} 0, \end{aligned}$$

car, grâce à l'inégalité de Poincaré,

$$\int_{\Omega} \frac{\partial u}{\partial t}(x, t) u(x, t) dx \leq C \left\| \frac{\partial u}{\partial t}(t) \right\|_{L^2(\Omega)} \left\| \nabla u(t) \right\|_{L^2(\Omega)}$$

est 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 donc

$$\frac{1}{t} \int_0^t \int_{\Omega} |\nabla u|^2 dx,$$

convergent vers  $E_0/2$ .

**Exercice 7.5.3** On considère l'équation des ondes dans tout l'espace  $\mathbb{R}^N$

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - \Delta u = 0 & \text{dans } \mathbb{R}^N \times \mathbb{R}_+^* \\ u(x, 0) = u_0(x) & \text{dans } x \in \mathbb{R}^N \\ \frac{\partial u}{\partial t}(x, 0) = u_1(x) & \text{dans } x \in \mathbb{R}^N, \end{cases} \quad (7.22)$$

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 chapitre 9 de [3].)

**Correction.** On procède de manière identique dans les trois cas : dans un premier temps, on vérifie que pour toute fonction  $v$ ,

$$\frac{\partial^2 Mv}{\partial t^2}(x, t) = \Delta(Mv)(x, t) \quad (7.23)$$

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

$$\frac{\partial^2 Mv}{\partial t^2} = \frac{1}{2\pi} \int_{|\xi| < 1} \frac{t(\nabla^2 v(x - t\xi)\xi) \cdot \xi - 2\nabla v(x - t\xi) \cdot \xi}{(1 - |\xi|^2)^{1/2}} d\xi$$

où  $\nabla^2 v$  est la matrice hessienne des dérivées partielles secondes de  $v$  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 (7.23), on introduit, pour tous  $x$  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

$$\frac{\partial^2 Mv}{\partial t^2} = \frac{1}{2\pi t} \int_{|\xi| < 1} \frac{(\nabla^2 w\xi) \cdot \xi + 2\nabla w \cdot \xi}{(1 - |\xi|^2)^{1/2}} d\xi$$

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 parties, on montre que

$$\int_{|\xi| < r} \frac{\Delta w}{(1 - |\xi|^2)^{1/2}} d\xi = - \int_{|\xi| < r} \frac{\nabla w \cdot \xi}{(1 - |\xi|^2)^{3/2}} d\xi + \frac{1}{r(1 - r^2)^{1/2}} \int_{|\xi|=r} (\nabla w \cdot \xi) ds.$$

De même,

$$\begin{aligned} & \int_{|\xi| < r} \frac{(\nabla^2 w\xi) \cdot \xi}{(1 - |\xi|^2)^{1/2}} d\xi = \\ & - \int_{|\xi| < r} (\nabla w \cdot \xi) \left( \frac{2}{(1 - |\xi|^2)^{1/2}} + \frac{1}{(1 - |\xi|^2)^{3/2}} \right) d\xi + \frac{r}{(1 - r^2)^{1/2}} \int_{|\xi|=r} \nabla w \cdot \xi ds. \end{aligned}$$

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

$$\int_{|\xi| < 1} \frac{\Delta w - (\nabla^2 w\xi) \cdot \xi}{(1 - |\xi|^2)^{1/2}} d\xi = 2 \int_{|\xi| < 1} \frac{\nabla w \cdot \xi}{(1 - |\xi|^2)^{1/2}} d\xi.$$

De l'expression des dérivées partielles de  $Mv$  en fonction de  $w$ , on déduit que  $Mv$  vérifie l'équation des ondes (7.23). Reste à prouver que  $Mv$  vérifie bien les conditions initiales 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 intégral, 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 parties, 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

$$\frac{\partial^2 Mv}{\partial t^2} = \frac{1}{4\pi} \int_{|\xi|=1} t(\nabla^2 v(x - t\xi)\xi) \cdot \xi - 2(\nabla v \cdot \xi) ds$$

et

$$\Delta(Mv) = \frac{1}{4\pi} \int_{|\xi|=1} t\Delta v(x - t\xi) ds .$$

Soit  $(x, t)$  fixé tel que  $t > 0$ . On introduit la fonction  $w(\xi) = v(x - t\xi)$ . On a

$$\frac{\partial^2 Mv}{\partial t^2} = \frac{1}{4\pi t} \int_{|\xi|=1} (\nabla^2 w\xi) \cdot \xi + 2(\nabla w \cdot \xi) ds$$

$$\Delta(Mv) = \frac{1}{4\pi t} \int_{|\xi|=1} \Delta w ds .$$

Il suffit donc de remarquer que

$$\int_{|\xi|=1} (\nabla^2 w\xi + 2\nabla w - \Delta w\xi) \cdot \xi ds = 0 ,$$

en tant que flux d'un champ de divergence nulle. En effet,

$$\nabla \cdot (\nabla^2 w\xi) = \nabla(\Delta w) \cdot \xi + \Delta w ,$$

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 initiales annoncées (pourvu qu'on sache que la surface de la sphère est  $4\pi$ ).

**Exercice 7.5.4** On considère l'équation des ondes (7.22) 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  $\Omega$ . 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  $\Omega$ , la solution explicite donnée par l'exercice précédent est aussi solution de l'équation des ondes dans le domaine  $\Omega$ . 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 7.5.5 (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.** En dimension 3, d'après l'expression de la solution de l'équation des ondes obtenu à l'Exercice 8.5.3, on constate (en considérant une source sonore ponctuelle) que la solution est indépendante du point l'espace considéré, à un amortissement et un décalage temporel près. Tout auditeur perçoit donc le même son (tout comme le musicien par ailleurs) et seule la puissance du signal reçu varie. De plus, le son perçu ne dépend que du son émis à un instant donné, ce qui permet en particulier au musicien d'exercer un rétro-contrôle. Ces deux propriétés ne sont plus vérifiées en dimension 2. Le son perçu (même en considérant une source ponctuelle) dépend de l'emplacement de l'auditeur et dépend en plus de tout le passé. Au lieu de se succéder au cours du temps les sons successifs s'empilent et se mélangent et l'auditeur devrait extraire d'un signal "cumulé" les ondes arrivées en dernier. Enfin, il est à noter qu'en dimension 2, il est tout simplement impossible d'obtenir un silence absolu (même si la "pollution sonore" décroît progressivement au court du temps).

**Exercice 7.6.1** La discrétisation en espace d'une équation parabolique conduit à la résolution de l'équation différentielle ordinaire

$$\begin{cases} \mathcal{M} \frac{dU}{dt}(t) + \mathcal{K}U(t) = b(t) \\ U(t=0) = U^0 \end{cases}, \quad (7.24)$$

où  $\mathcal{M}$  et  $\mathcal{K}$  sont les matrices de masse et de rigidité associées au problème et  $U(t) \in \mathbb{R}^N$  est le vecteur constitué des degrés de libertés de l'approximation spatiale choisie. Dans



l'optique de résoudre numériquement cette équation ordinaire, on introduit le  $\theta$ -schéma consistant à calculer une approximation  $U^n \in \mathbb{R}^N$  de  $U(n\Delta t)$  définie par

$$\mathcal{M} \frac{U^{n+1} - U^n}{\Delta t} + \mathcal{K} (\theta U^{n+1} + (1 - \theta) U^n) = \theta b(t_{n+1}) + (1 - \theta) b(t_n). \quad (7.25)$$

où  $\theta \in [0, 1]$ . Pour  $\theta = 1/2$ , on obtient le schéma de Crank-Nicholson. Nous allons également considérer un autre schéma, dit de Gear, défini par

$$\mathcal{M} \frac{3U^{n+1} - 4U^n + U^{n-1}}{2\Delta t} + \mathcal{K} U^{n+1} = b(t_{n+1}). \quad (7.26)$$

Montrer que le schéma de Crank-Nicholson et celui de Gear sont d'ordre 2 (en temps), tandis que le  $\theta$ -schéma pour  $\theta \neq 1/2$  est d'ordre 1.

**Correction.** Schéma de Crank-Nicholson et  $\theta$ -schéma

Soit  $U$  la solution de l'équation différentielle (7.24). L'erreur de troncature du schéma du  $\theta$ -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

$$E(U) = \left( \mathcal{M} \frac{dU}{dt} + \mathcal{K} U - b \right) + \Delta t \left( \frac{\mathcal{M}}{2} \frac{d^2 U}{dt^2} + \theta \left( \mathcal{K} \frac{dU}{dt} - \frac{db}{dt} \right) \right) + (\Delta t)^2 \left( \frac{\mathcal{M}}{6} \frac{d^3 U}{dt^3} + \frac{\theta \mathcal{K}}{2} \frac{d^2 U}{dt^2} - \frac{\theta}{2} \frac{d^2 b}{dt^2} \right) + \mathcal{O}((\Delta t)^3)$$

En exploitant l'équation vérifiée par  $U$ , on en déduit que

$$E(U) = \Delta t \frac{1 - 2\theta}{2} \left( \frac{db}{dt} - \mathcal{K} \mathcal{M}^{-1} (b - \mathcal{K} U) \right) + (\Delta t)^2 \frac{1 - 3\theta}{6} \left( (\mathcal{K} \mathcal{M}^{-1})^2 (b - \mathcal{K} U) + \mathcal{K} \mathcal{M}^{-1} \frac{db}{dt} + \frac{d^2 b}{dt^2} \right) + \mathcal{O}((\Delta t)^3).$$

Pour  $\theta \neq 1/2$ , le  $\theta$ -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

$$E(U) = \left( \mathcal{M} \frac{dU}{dt} + \mathcal{K} U - b \right) (t_{n+1}) + \frac{(\Delta t)^2}{3} \mathcal{M} \frac{d^3 U}{dt^3} (t_{n+1}) + \mathcal{O}((\Delta t)^3).$$

Si  $U$  est solution de (7.24), 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 7.6.2** On considère le  $\theta$ -schéma (7.25) 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 (7.25) 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 continu. À 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 démontrer ce lemme, puis l'appliquer au  $\theta$ -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 du 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

$$\frac{\|U^{n+1}\|_{\mathcal{M}}^2 - \|U_0\|_{\mathcal{M}}^2}{2\Delta t} + \sum_{p=0}^n \mathcal{K} \hat{U}^p \cdot \hat{U}^p \leq \sum_{p=0}^n (\theta b_{p+1} + (1 - \theta)b_p) \cdot \hat{U}^p. \quad (7.27)$$

Majorons le terme de droite dans (7.27). D'après la définition de  $b$  (voir (8.57)), on a

$$\begin{aligned} (\theta b_{p+1} + (1 - \theta)b_p) \cdot \hat{U}^p &= (\theta b_{p+1} + (1 - \theta)b_p) \cdot (\theta U^{p+1} + (1 - \theta)U^p) \\ &= \int_{\Omega} (\theta f(t_{p+1}) + (1 - \theta)f(t_p)) \cdot (\theta u_h^{p+1} + (1 - \theta)u_h^p) dx. \end{aligned}$$

On en déduit que

$$\begin{aligned} &(\theta b_{p+1} + (1 - \theta)b_p) \cdot (\theta U^{p+1} + (1 - \theta)U^p) \\ &\leq \frac{1}{2} \left( \|\theta f(t_{p+1}) + (1 - \theta)f(t_p)\|_{L^2(\Omega)}^2 + \|\theta u_h^{p+1} + (1 - \theta)u_h^p\|_{L^2(\Omega)}^2 \right). \end{aligned}$$

De la définition de  $\mathcal{M}$  et en utilisant la convexité de l'application  $x \mapsto x^2$ , il en découle que

$$\begin{aligned} &(\theta b_{p+1} + (1 - \theta)b_p) \cdot (\theta U^{p+1} + (1 - \theta)U^p) \\ &\leq \frac{1}{2} \left( \theta \|f(t_{p+1})\|_{L^2(\Omega)}^2 + (1 - \theta) \|f(t_p)\|_{L^2(\Omega)}^2 + \theta \|U^{p+1}\|_{\mathcal{M}}^2 + (1 - \theta) \|U^p\|_{\mathcal{M}}^2 \right) \end{aligned}$$

L'inégalité (7.27) implique ainsi

$$\frac{1}{2\Delta t} (\|U^{n+1}\|_{\mathcal{M}}^2 - \|U_0\|_{\mathcal{M}}^2) + \sum_{p=0}^n \mathcal{K} \hat{U}^p \cdot \hat{U}^p \leq \frac{1}{2} \left( \sum_{p=0}^{n+1} \|f(t_p)\|_{L^2(\Omega)}^2 + \sum_{p=0}^{n+1} \|U^p\|_{\mathcal{M}}^2 \right).$$

On réarrange les différents termes de l'inégalité afin d'obtenir une majoration nous permettant d'appliquer l'équivalent discret du lemme de Gronwall.

$$\begin{aligned} \|U^{n+1}\|_{\mathcal{M}}^2 + \frac{2\Delta t}{1 - \Delta t} \sum_{p=0}^n \mathcal{K} \hat{U}^p \cdot \hat{U}^p \\ \leq \frac{1}{1 - \Delta t} \|U^0\|_{\mathcal{M}}^2 + \frac{\Delta t}{1 - \Delta t} \left( \sum_{p=0}^{n+1} \|f(t_p)\|_{L^2(\Omega)}^2 + \sum_{p=0}^n \|U^p\|_{\mathcal{M}}^2 \right) \end{aligned}$$

On applique la version discrète du lemme de Gronwall à  $v_n = \|U^n\|_{\mathcal{M}}^2$  avec les paramètres

$$a = \frac{1}{1 - \Delta t} \|U^0\|_{\mathcal{M}}^2 + \frac{\Delta t}{1 - \Delta t} \sum_{p=0}^{n_0} \|f(t_p)\|_{L^2(\Omega)}^2 \quad \text{et} \quad b = \frac{\Delta t}{1 - \Delta t}.$$

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 donc

$$\|U^{n+1}\|_{\mathcal{M}}^2 + \frac{2\Delta t}{1-\Delta t} \sum_{p=0}^n \mathcal{K}\hat{U}^p \cdot \hat{U}^p \leq (1-\Delta t)^{-(n+2)} \left( \|U^0\|_{\mathcal{M}}^2 + \Delta t \sum_{p=0}^{n_0+1} \|f(t_p)\|_{L^2(\Omega)}^2 \right).$$

Notons que  $(1-\Delta t)^{-n}$  est majoré par une constante indépendante du pas de temps  $\Delta t$  (mais dépendant du temps final  $T = n_0\Delta t$ ). En effet, pour un temps  $t \geq 0$  fixé, si on pose  $n = t/(\Delta t)$ , alors  $(1-\Delta t)^{-n}$  tend vers  $e^t$  lorsque  $\Delta t$  tend vers zéro. Par ailleurs,  $\Delta t \sum_{p=0}^{n_0+1} \|f(t_p)\|_{L^2(\Omega)}^2$  est une approximation consistante de  $\|f\|_{L^2((0,T)\times\Omega)}^2$ . On retrouve ainsi l'équivalent discret de l'estimation d'énergie de l'Exercice 7.3.1.

**Exercice 7.6.3** Montrer que le schéma de Gear (7.26) est inconditionnellement stable.

**Correction.** On prouve la stabilité en établissant une estimation d'énergie du même type que celle obtenue dans l'Exercice 7.6.2. 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 (7.26) par  $U^{n+1}$ . En minorant le terme de droite, on obtient

$$\begin{aligned} \frac{1}{4} \left( \|U^{n+1}\|_{\mathcal{M}}^2 - \|U^n\|_{\mathcal{M}}^2 + \|2U^{n+1} - U^n\|_{\mathcal{M}}^2 - \|2U^n - U^{n-1}\|_{\mathcal{M}}^2 \right) \\ + \Delta t \mathcal{K}U^{n+1} \cdot U^{n+1} \leq \Delta t b(t_{n+1}) \cdot U^{n+1}. \end{aligned}$$

Or, pour tout  $\alpha \neq 0$ , on a

$$b(t_{n+1}) \cdot U^{n+1} \leq \frac{1}{2} \left( \alpha^2 \|U^{n+1}\|_{\mathcal{M}}^2 + \alpha^{-2} \|f(t_{n+1})\|_{L^2(\Omega)}^2 \right).$$

Comme les matrices  $\mathcal{K}$  et  $\mathcal{M}$  sont définies positives, on peut choisir ce paramètre  $\alpha^2 > 0$  comme la constante de minoration dans l'inégalité suivante

$$\alpha^2 \|U\|_{\mathcal{M}}^2 \leq \mathcal{K}U \cdot U \quad \forall U \in \mathbb{R}^N.$$

Par conséquent, on en déduit que

$$\|U^{n+1}\|_{\mathcal{M}}^2 + \|2U^{n+1} - U^n\|_{\mathcal{M}}^2 \leq \|U^n\|_{\mathcal{M}}^2 + \|2U^n - U^{n-1}\|_{\mathcal{M}}^2 + 2\alpha^{-2}\Delta t \|f(t_{n+1})\|_{L^2(\Omega)}^2,$$

c'est-à-dire que

$$\|U^{n+1}\|_{\mathcal{M}}^2 \leq \|U^1\|_{\mathcal{M}}^2 + \|2U^1 - U^0\|_{\mathcal{M}}^2 + 2\alpha^{-2}\Delta t \sum_{p=2}^{n+1} \|f(t_p)\|_{L^2(\Omega)}^2,$$

où la dernière somme est une approximation consistante de  $\|f\|_{L^2((0,T)\times\Omega)}^2$ . Ceci prouve la stabilité du schéma de Gear sans aucune condition sur le pas de temps.

**Exercice 7.6.4** On résout par éléments finis  $\mathbb{P}_1$  et schéma explicite en temps l'équation de la chaleur (7.1) 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 (5.1) et qu'on a calculé ses valeurs propres lors de l'Exercice 13.1.3. Montrer que dans ce cas la condition CFL

$$\max_i \lambda_i \Delta t \leq 2, \quad (7.28)$$

où les  $\lambda_i$  sont les valeurs propres de  $\mathcal{K}U = \lambda \mathcal{M}U$  est bien du type  $\Delta t \leq Ch^2$ .

**Correction.** La matrice  $\mathcal{M}$  obtenue à l'aide de la méthode de mass-lumping est

$$\mathcal{M} = h \text{Id}.$$

Ainsi, les valeurs propres  $\lambda_i$  sont égales aux valeurs propres de  $\mathcal{K}$  divisées par  $h$ , et

$$\lambda_i = 4h^{-2} \sin^2 \left( \frac{i\pi}{2(n+1)} \right).$$

On en déduit que  $\lambda_i \leq 4h^{-2}$ , on retrouve une condition CFL classique, c'est-à-dire

$$2\Delta t \leq h^2.$$

**Exercice 7.7.1** Écrire le système linéaire d'équations différentielles ordinaires obtenu par semi-discrétisation de l'équation des ondes amortie (7.21).

**Correction.** Le problème discrétisé en espace consiste à déterminer  $u(t)$  fonction de  $t$  à valeur dans  $V_{0h}$  telle que pour tout  $v_h \in V_{0h}$ ,

$$\frac{d^2}{dt^2} \langle u_h(t), v_h \rangle_{L^2(\Omega)} + \eta \frac{d}{dt} \langle u_h(t), v_h \rangle_{L^2(\Omega)} + \langle \nabla u_h(t), \nabla v(t) \rangle_{L^2(\Omega)} = \langle f, v_h \rangle_{L^2(\Omega)}$$

avec

$$u_h(t=0) = u_{0,h} \quad \text{et} \quad \frac{du_h}{dt}(t=0) = u_{1,h}.$$

Si  $\phi_i$  désigne la base de  $V_{0h}$ , si on note  $U_i(t)$  les coordonnées de  $u_h(t)$  dans cette base, on a

$$\mathcal{M} \frac{d^2 U}{dt^2}(t) + \eta \mathcal{M} \frac{dU}{dt}(t) + \mathcal{K}U(t) = b(t)$$

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 7.7.2** La discrétisation spatiale de l'équation des ondes conduit à l'équation différentielle ordinaire

$$\mathcal{M} \frac{d^2 U}{dt^2}(t) + \mathcal{K}U(t) = b(t), \quad (7.29)$$

où  $\mathcal{M}$  et  $\mathcal{K}$  sont respectivement les matrices de masse et de rigidité, tandis que  $U(t) \in \mathbb{R}^N$  est le vecteur des degrés de libertés de l'approximation spatiale de la solution exacte

et  $b(t) \in \mathbb{R}^N$  dépend du terme source. Afin d'intégrer numériquement (7.29) on considère le schéma de Newmark pour les trois suites de vecteurs  $\ddot{U}^n, \dot{U}^n, U^n$

$$\begin{cases} \mathcal{M}\ddot{U}^{n+1} + \mathcal{K}U^{n+1} = b(t_{n+1}) \\ \dot{U}^{n+1} = \dot{U}^n + \Delta t(\delta\ddot{U}^{n+1} + (1-\delta)\ddot{U}^n) \\ U^{n+1} = U^n + \Delta t\dot{U}^n + \frac{(\Delta t)^2}{2}(2\theta\ddot{U}^{n+1} + (1-2\theta)\ddot{U}^n) \end{cases} \quad (7.30)$$

avec  $0 \leq \delta \leq 1$  et  $0 \leq \theta \leq 1/2$ . 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  $\theta = 1/12$ .

**Correction.** Comme indiqué dans (8.70) on peut éliminer les suites  $\ddot{U}^n, \dot{U}^n$  et réécrire le schéma sous la forme

$$\begin{aligned} \mathcal{M}\frac{U^{n+1} - 2U^n + U^{n-1}}{(\Delta t)^2} + \mathcal{K}\left(\theta U^{n+1} + \left(\frac{1}{2} + \delta - 2\theta\right)U^n + \left(\frac{1}{2} - \delta + \theta\right)U^{n-1}\right) \\ = \theta b(t_{n+1}) + \left(\frac{1}{2} + \delta - 2\theta\right)b(t_n) + \left(\frac{1}{2} - \delta + \theta\right)b(t_{n-1}). \end{aligned}$$

On introduit l'erreur de troncature de ce schéma

$$\begin{aligned} E(U) = \mathcal{M}\frac{U(t+\Delta t) - 2U(t) + U(t-\Delta t)}{(\Delta t)^2} \\ + \mathcal{K}\left(\theta U(t+\Delta t) + \left(\frac{1}{2} + \delta - 2\theta\right)U(t) + \left(\frac{1}{2} - \delta + \theta\right)U(t-\Delta t)\right) \\ - \left(\theta b(t+\Delta t) + \left(\frac{1}{2} + \delta - 2\theta\right)b(t) + \left(\frac{1}{2} - \delta + \theta\right)b(t-\Delta t)\right). \end{aligned}$$

En effectuant un développement de Taylor en  $t = t_n$ , on établit que

$$\begin{aligned} E(U) = \mathcal{M}\frac{d^2U}{dt^2} + \mathcal{K}U - b + \Delta t\left(\delta - \frac{1}{2}\right)\left(\mathcal{K}\frac{dU}{dt} - \frac{db}{dt}\right) \\ + (\Delta t)^2\left(\frac{1}{4} - \frac{\delta}{2} + \theta\right)\left(\mathcal{K}\frac{d^2U}{dt^2} - \frac{d^2b}{dt^2}\right) + \frac{(\Delta t)^2}{12}\mathcal{M}\frac{d^4U}{dt^4} \\ + \frac{(\Delta t)^3}{6}\left(\delta - \frac{1}{2}\right)\left(\mathcal{K}\frac{d^3U}{dt^3} - \frac{d^3b}{dt^3}\right) + \mathcal{O}((\Delta t)^4). \end{aligned}$$

Si  $U$  est solution de l'équation (7.29), on a

$$\mathcal{K}\frac{d^2U}{dt^2} - \frac{d^2b}{dt^2} = -\mathcal{M}\frac{d^4U}{dt^4}.$$

Ainsi,

$$\begin{aligned} E(U) = \Delta t\left(\delta - \frac{1}{2}\right)\left(\mathcal{K}\frac{dU}{dt} - \frac{db}{dt}\right) - (\Delta t)^2\left(\frac{1}{4} - \frac{\delta}{2} + \theta - \frac{1}{12}\right)\mathcal{M}\frac{d^4U}{dt^4} \\ + \frac{(\Delta t)^3}{6}\left(\delta - \frac{1}{2}\right)\left(\mathcal{K}\frac{d^3U}{dt^3} - \frac{d^3b}{dt^3}\right) + \mathcal{O}((\Delta t)^4). \end{aligned}$$

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 7.7.3** On considère le cas limite du Lemme 8.7.1, c'est-à-dire  $\delta = 1/2$  et  $\lambda_i(\Delta t)^2 = \frac{4}{1-4\theta}$  (on utilise les mêmes notations que celles introduites dans la preuve de ce dernier). Montrer que le schéma de Newmark (7.30) est instable dans ce cas en vérifiant que

$$A_i = \begin{pmatrix} -2 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{et} \quad A_i^n = (-1)^n \begin{pmatrix} n+1 & n \\ -n & 1-n \end{pmatrix}.$$

Remarquez qu'il s'agit d'une instabilité "faible" puisque la croissance de  $A_i^n$  est linéaire et non exponentielle.

**Correction.** On rappelle que les  $\lambda_i$  désignent les valeurs propres de  $\mathcal{K}U = \lambda\mathcal{M}U$  et que  $U_i^n$  est la composante du vecteur  $U^n$  suivant le  $i$ -ème vecteur propre associé à  $\lambda_i$ . D'après la démonstration du Lemme 8.7.1, le schéma de Newmark est équivalent à

$$\frac{U_i^{n+1} - 2U_i^n + U_i^{n-1}}{(\Delta t)^2} + \lambda_i \left( \theta U_i^{n+1} + \left(\frac{1}{2} + \delta - 2\theta\right)U_i^n + \left(\frac{1}{2} - \delta + \theta\right)U_i^{n-1} \right) = b_i^n,$$

c'est-à-dire

$$\begin{pmatrix} U_i^{n+1} \\ U_i^n \end{pmatrix} = A_i \begin{pmatrix} U_i^n \\ U_i^{n-1} \end{pmatrix} + \frac{(\Delta t)^2}{1 + \theta\lambda_i(\Delta t)^2} \begin{pmatrix} b_i^n \\ 0 \end{pmatrix} \quad \text{avec} \quad A_i = \begin{pmatrix} a_{11} & a_{12} \\ 1 & 0 \end{pmatrix},$$

et

$$a_{11} = \frac{2 - \lambda_i(\Delta t)^2(\frac{1}{2} + \delta - 2\theta)}{1 + \theta\lambda_i(\Delta t)^2}, \quad a_{12} = -\frac{1 + \lambda_i(\Delta t)^2(\frac{1}{2} - \delta + \theta)}{1 + \theta\lambda_i(\Delta t)^2}.$$

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 en découle que le schéma de Newmark est instable dans ce cas : pour s'en convaincre, il suffit de considérer le cas  $b = 0$ ,  $U_i^1 = U_i^0 = 1$  et  $U_j^1 = U_j^0 = 0$  pour  $j \neq i$ .

# Chapitre 8

## INTRODUCTION A L'OPTIMISATION

**Exercice 8.1.1** 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  $J(x) = x$  sur  $]0, 1[$ .
- $J$  non continue : minimisation sur  $\mathbb{R}$  de  $J(x) = x^2$  pour  $x \neq 0$ ,  $J(0) = 1$ .
- $J$  continue, minorée mais non “infinie à l'infini” : minimisation sur  $\mathbb{R}$  de  $J(x) = e^{-x}$ .

**Exercice 8.1.2** Montrer que l'on peut remplacer la propriété “infinie à l'infini” (9.3) du Théorème d'existence 9.1.3 de minimiseur en dimension finie par la condition plus faible

$$\inf_{v \in K} J(v) < \lim_{R \rightarrow +\infty} \left( \inf_{\substack{\|v\| \geq R \\ v \in K}} J(v) \right).$$

**Correction.** Soit  $(v^n)$  une suite minimisante de  $J$  sur  $K$ . Comme

$$\inf_{v \in K} J(v) < \lim_{R \rightarrow +\infty} \left( \inf_{\substack{\|v\| \geq R \\ v \in K}} J(v) \right),$$

et que  $J(v_n)$  converge vers  $\inf_{v \in K} J(v)$ , il existe  $\delta > 0$ , suffisamment petit, et  $n_0$  tel que, pour tout  $n \geq n_0$ ,

$$J(v_n) < \lim_{R \rightarrow +\infty} \left( \inf_{\substack{\|v\| \geq R \\ v \in K}} J(v) \right) - \delta.$$

Ainsi, il existe  $R$  assez grand tel que, pour tout  $n \geq n_0$ ,

$$J(v_n) < \inf_{\substack{\|v\| \geq R \\ v \in K}} J(v).$$



De cette inégalité on déduit que, pour tout  $n \geq n_0$ , on ne peut pas avoir  $\|v_n\| \geq R$ . Par conséquent, pour  $n \geq n_0$ ,  $v_n$  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 8.1.3** Montrer que la conclusion du Théorème **9.1.3** d'existence d'un minimiseur en dimension finie reste valable si on remplace l'hypothèse de continuité de  $J$  par la condition de semi-continuité inférieure

$$\forall (u^n)_{n \geq 0} \text{ suite dans } K, \quad \lim_{n \rightarrow +\infty} u^n = u \implies \liminf_{n \rightarrow +\infty} J(u^n) \geq J(u).$$

**Correction.** Soit  $(u_n)$  une suite minimisante de  $J$  sur  $K$ . Comme  $J$  est supposée infinie à l'infini,  $u_n$  est bornée, puisque  $J(u_n)$  est une suite de réels majorée. Il existe donc une sous-suite  $(u^{n_k})$  convergeant vers un élément  $u \in \mathbb{R}^N$ . Comme  $K$  est fermé,  $u \in K$ . D'autre part, comme  $J$  est semi-continue inférieurement,

$$J(u) \leq \liminf_{k \rightarrow \infty} J(u^{n_k}) = \inf_{v \in K} J(v)$$

et donc

$$J(u) = \inf_{v \in K} J(v).$$

**Exercice 8.1.4** Montrer qu'il existe un minimum pour les Exemples **9.1.1**, **9.1.6** et **9.1.7**.

**Correction.**

**Exemple 9.1.1 : Problème de transport.** On considère le problème de minimisation de

$$J(v) = \sum_{i=1}^M \sum_{j=1}^N c_{ij} v_{ij}$$

sur

$$K = \left\{ v \in \mathbb{R}_+^{M \times N} \text{ tel que } \sum_{j=1}^N v_{ij} \leq s_i \text{ et } \sum_{i=1}^M v_{ij} = r_j, \forall 1 \leq i \leq M \text{ et } 1 \leq j \leq N \right\}$$

avec l'hypothèse que

$$\sum_{j=1}^N r_j \leq \sum_{i=1}^M s_i.$$

Tout d'abord  $K$  est clairement fermé et, d'après l'hypothèse,  $K$  est non vide. Enfin,  $J$  est continue et comme  $K$  est borné, aucune hypothèse sur le comportement de  $J$  à l'infini n'est nécessaire. On en déduit qu'il existe au moins un minimiseur de  $J$  sur  $K$ .

**Exemple 9.1.6 : Optimisation quadratique sous contraintes linéaires.** Pour une matrice carrée d'ordre  $n$ , symétrique définie positive,  $A$  et une matrice rectangulaire  $B$  de taille  $m \times n$ , on considère la minimisation de

$$J(x) = \frac{1}{2} Ax \cdot x - b \cdot x,$$

sur

$$K = \text{Ker}(B).$$

L'espace admissible  $K$  est un sous espace vectoriel de  $\mathbb{R}^n$  et est donc fermé. De plus, comme  $0 \in K$ , il est non vide. Enfin,  $A$  étant supposée symétrique définie positive, l'application  $x \mapsto Ax \cdot x$  est une norme équivalente à la norme euclidienne sur  $\mathbb{R}^n$ . Il existe donc une constante  $C$  telle que

$$Ax \cdot x \geq 2C\|x\|^2$$

et

$$J(x) \geq C\|x\|^2 - \|b\|\|x\|.$$

On en déduit que  $J$  est infinie à l'infini. Comme  $J$  est également continue, on en conclut que  $J$  admet un minimiseur sur  $K$ .

**Exemple 9.1.7 : Première valeur propre.** Pour une matrice carrée d'ordre  $n$ , symétrique,  $A$ , on considère la minimisation de

$$J(x) = Ax \cdot x$$

sur

$$K = \{x \in \mathbb{R}^n \text{ tel que } \|x\| = 1\}.$$

L'espace admissible  $K$  est fermé (car image réciproque d'un fermé par une application continue) et trivialement non vide. Enfin,  $J$  est continue et comme  $K$  est borné, aucune hypothèse sur le comportement de  $J$  à l'infini n'est à vérifier. On en conclut par application du Théorème 9.1.3 que  $J$  admet un minimiseur sur  $K$ .

**Exercice 8.1.5** 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

$$\inf_{P \in \mathcal{P}_n} \|P\| \tag{8.1}$$

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 (8.1) 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 des polynômes 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 l'on déduit l'existence d'une solution au problème de minimisation de  $\|P\|$  sur  $\mathcal{P}_n$ .

2. Soit  $P_n$  la proposition stipulant que pour tout  $0 \leq p \leq n$ ,  $T_p$  est un polynôme de degré  $p$  tel que  $T_p(\cos(\theta)) = \cos(p\theta)$ . Soit  $n \geq 1$  et supposons  $P_n$  vrai. Par définition,

$$T_{n+1}(X) = 2XT_n(X) - T_{n-1}(X).$$

D'après l'hypothèse de récurrence, on en déduit que  $T_{n+1}$  est un polynôme de degré  $n + 1$  (c'est la somme d'un polynôme de degré  $n + 1$  et d'un polynôme de degré  $n - 1$ ). De plus,

$$\begin{aligned} T_{n+1}(\cos \theta) &= 2(\cos \theta)T_n(\cos \theta) - T_{n-1}(\cos \theta) \\ &= 2(\cos \theta)(\cos n\theta) - (\cos(n-1)\theta) = \cos((1+n)\theta). \end{aligned}$$

Ainsi,  $P_n$  implique  $P_{n+1}$ . Comme  $P_1$  est vraie, on en déduit que  $P_n$  est vraie pour tout  $n \geq 1$ .

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 minimale 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  $|I| + 2|J| + |K| \geq n$  avec la notation  $|I| = \text{card}(I)$ . 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  $S'(y_j) = 0$ .

Pour tout  $i \in I$ , comme  $\|P\| \geq \|R\|$ , le signe de  $S(y_i) = P(y_i) - R(y_i)$  est égal 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 8.2.1** Modifier la construction de l'Exemple 9.2.2 pour montrer qu'il n'existe pas non plus de minimum de

$$J_h(v) = \int_0^1 \left( (|v'(x)| - h)^2 + v(x)^2 \right) dx .$$

sur  $C^1[0, 1]$  pour  $h \neq 0$ .

**Correction.** Soit  $a \in [0, 1]$ . On note  $P_a$  la fonction de  $C^1(\mathbb{R}; \mathbb{R})$  paire, périodique de période 2, définie sur  $[0, 1]$  par

$$P_a(x) = \begin{cases} x^2/2a + (a-1)/2 & \text{si } 0 \leq x \leq a \\ x - 1/2 & \text{si } a \leq x \leq 1-a, \\ -(x-1)^2/(2a) + (1-a)/2 & \text{si } 1-a \leq x \leq 1 \end{cases}$$

On note  $u^n \in C^1(\mathbb{R}; \mathbb{R})$  la fonction  $2/n$ -périodique, définie par

$$u^n(x) = n^{-1}hP_{n^{-1}}(nx).$$

On vérifie que  $u^n(x) \rightarrow 0$  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 8.2.2** Soient  $J_1$  et  $J_2$  deux fonctions convexes sur  $V$ ,  $\lambda > 0$ , et  $\varphi$  une fonction convexe croissante sur un intervalle de  $\mathbb{R}$  contenant l'ensemble  $J_1(V)$ . Montrer que  $J_1 + J_2$ ,  $\max(J_1, J_2)$ ,  $\lambda J_1$  et  $\varphi \circ J_1$  sont convexes.

**Correction.** La convexité de  $J_1 + J_2$  comme de  $\lambda J_1$  est triviale à établir. Par ailleurs, rappelons que, selon la Remarque 9.2.2, une fonction  $J : V \rightarrow \mathbb{R}$  est convexe si et seulement si son épigraphe

$$Epi(J) = \{(\lambda, v) \in \mathbb{R} \times V, \lambda \geq J(v)\}$$

est convexe. Or,

$$\begin{aligned} Epi(\max(J_1, J_2)) &= \{(\lambda, v) \in \mathbb{R} \times V : \lambda \geq J_1(v) \text{ et } \lambda \geq J_2(v)\} \\ &= Epi(J_1) \cap Epi(J_2). \end{aligned}$$

L'intersection de deux convexes étant convexe,  $Epi(\max(J_1, J_2))$  est convexe et  $\max(J_1, J_2)$  est donc convexe.

D'autre part, si  $J$  est convexe et  $\varphi$  croissante,

$$\varphi \circ J(\theta x + (1-\theta)y) \leq \varphi(\theta J(x) + (1-\theta)J(y))$$

enfin comme  $\varphi$  est convexe il vient,

$$\varphi \circ J(\theta x + (1-\theta)y) \leq \theta \varphi \circ J(x) + (1-\theta) \varphi \circ J(y).$$

La convexité de  $\varphi \circ J$  est ainsi établie.

**Exercice 8.2.3** 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. Ainsi, si  $J = \sup_{i \in I} J_i$ , où  $J_i$  sont des fonctions convexes, on a

$$\text{Epi}(J) = \{(\lambda, v) \in \mathbb{R} \times V, \lambda \geq J_i(v) \text{ pour tout } i \in I\} = \bigcap_i \text{Epi}(J_i).$$

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  $(\lambda_0, v_0)$  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  $\beta$  est non nul, car  $(\lambda_0, v_0) \notin \text{Epi}(J)$  et, de plus,  $\beta$  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  $(v_0, \lambda_0)$  tel que

$$J(v_0) > \lambda_0,$$

il existe une fonction affine  $L$  telle que

$$J(v_0) > L(v_0) = \lambda_0$$

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 8.2.4** Si  $J$  est continue et  $\alpha$ -convexe, montrer que, pour tout  $\theta \in [0, 1]$ ,

$$J(\theta u + (1 - \theta)v) \leq \theta J(u) + (1 - \theta)J(v) - \frac{\alpha\theta(1 - \theta)}{2} \|u - v\|^2. \quad (8.2)$$

**Correction.** Pour tout  $n$ , on note  $K_n = \{x \in [0, 1] : 2^n x \in \mathbb{N}\}$ . Supposons que l'inégalité (8.2) 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  $\alpha$ -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 (8.2) 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 8.2.5** 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  $\alpha$ .

**Correction.**

$$\begin{aligned} J((x + y)/2) &= A(x + y) \cdot (x + y)/8 - (b \cdot x + b \cdot y)/2 \\ &= \frac{Ax \cdot x/2 - b \cdot x + Ay \cdot y/2 - b \cdot y}{2} - A(x - y) \cdot (x - y)/8 \\ &= (J(x) + J(y))/2 - A(x - y) \cdot (x - y)/8. \end{aligned}$$

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  $\alpha$  est la plus petite valeur propre de  $A$ .

**Exercice 8.2.6** Soit  $\Omega$  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  $\Omega$  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.** Soit  $u$  et  $v \in H^1(\Omega)$ , on a

$$\begin{aligned} J\left(\frac{u+v}{2}\right) &= \frac{1}{2} \left\| \frac{u+v}{2} \right\|_{H^1(\Omega)}^2 - \frac{1}{2} \int_{\Omega} f(x)(u(x) + v(x)) dx \\ &= \frac{1}{4} (\|u\|_{H^1(\Omega)}^2 + \|v\|_{H^1(\Omega)}^2) - \frac{1}{2} \left\| \frac{u-v}{2} \right\|_{H^1(\Omega)}^2 - \frac{1}{2} \int_{\Omega} f(x)(u(x) + v(x)) dx \\ &= \frac{J(u) + J(v)}{2} - \|u-v\|_{H^1(\Omega)}^2/8. \end{aligned}$$

La fonction  $J$  est donc fortement convexe sur  $H^1(\Omega)$ .

**Exercice 8.2.7** 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 8.2.8** Montrer que le Théorème 9.2.6 d'existence d'un minimiseur dans le cas fortement convexe (en dimension infini) s'applique à l'Exemple 9.1.10 du problème de minimisation de

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

sur  $H_0^1(\Omega)$  où  $\Omega$  est un ouvert borné régulier de  $\mathbb{R}^N$  et  $f \in L^2(\Omega)$  (utiliser l'inégalité de Poincaré).

**Correction.** En procédant comme lors de l'Exercice **9.2.6**, on montre que

$$J\left(\frac{u+v}{2}\right) = \frac{J(u) + J(v)}{2} - \frac{1}{8} \int_{\Omega} |\nabla(u-v)|^2 dx.$$

Comme  $\Omega$  est un ouvert borné régulier, il existe, d'après l'inégalité de Poincaré, une constante  $C$  telle que pour tout  $u \in H_0^1(\Omega)$ ,

$$\int_{\Omega} |\nabla v|^2 dx \geq C \|u\|_{H^1(\Omega)}^2.$$

Ainsi,

$$J\left(\frac{u+v}{2}\right) \leq \frac{J(u) + J(v)}{2} - \frac{C}{8} \|u-v\|_{H^1(\Omega)}^2$$

et  $J$  est fortement convexe. La fonction  $J$  étant d'autre part continue sur  $H_0^1(\Omega)$ , le Théorème **9.2.6** s'applique et  $J$  admet donc un unique minimiseur sur  $H_0^1(\Omega)$ .

**Exercice 8.2.9** 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.** Dans tout ce qui suit,  $\Omega$  désigne un ouvert borné régulier de  $\mathbb{R}^N$ .  
**Laplacien avec conditions aux limites de Neumann.** L'énergie associée au problème est

$$J(v) = \frac{1}{2} \int_{\Omega} (|\nabla v|^2 + |v|^2) dx - \int_{\Omega} f v dx - \int_{\partial\Omega} g v ds, \quad (8.3)$$

avec  $f \in L^2(\Omega)$  et  $g \in L^2(\partial\Omega)$ . La fonctionnelle  $J$  est fortement convexe, car

$$J\left(\frac{u+v}{2}\right) = \frac{J(u) + J(v)}{2} - \|u-v\|_{H^1(\Omega)}^2/8.$$

De plus  $J$  est continue sur  $H^1(\Omega)$ . Ainsi  $J$  admet un unique minimiseur sur  $H^1(\Omega)$ .

**Elasticité.** On suppose que la frontière du domaine se décompose en deux parties  $\partial\Omega_D$  et  $\partial\Omega_N$  de mesures surfacielles non nulles. L'énergie associée au système de l'élasticité est définie pour tout

$$u \in V := \{v \in H^1(\Omega)^N : v = 0 \text{ sur } \Gamma_D\}$$

par

$$J(v) = \frac{1}{2} \int_{\Omega} (2\mu|e(v)|^2 + \lambda|\operatorname{div}v|^2) dx - \int_{\Omega} f \cdot v dx - \int_{\partial\Omega_N} g \cdot v ds,$$

où  $\lambda$  et  $\mu$  sont les coefficients de Lamé du solide tels que

$$\mu > 0 \text{ et } 2\mu + N\lambda > 0.$$



La fonctionnelle  $J$  est continue sur  $H^1(\Omega)^N$  et  $V$  est un espace de Hilbert. D'autres part, pour tout  $u$  et  $v$  dans  $V$ , on a

$$J\left(\frac{u+v}{2}\right) = \frac{J(u) + J(v)}{2} - \frac{1}{8} \int_{\Omega} (2\mu|e(u-v)|^2 + \lambda|\operatorname{div}(u-v)|^2) dx.$$

L'inégalité algébrique établie au cours de la démonstration du Théorème 5.3.1 implique que

$$\int_{\Omega} 2\mu|e(u-v)|^2 dx + \int_{\Omega} \lambda|\operatorname{div}(u-v)|^2 dx \geq \nu \int_{\Omega} |e(u-v)|^2 dx$$

avec  $\nu = \min(2\mu, 2\mu + N\lambda)$ . De plus, d'après l'inégalité de Korn (5.65)

$$\|v\|_{H^1(\Omega)} \leq C\|e(v)\|_{L^2(\Omega)}. \quad (8.4)$$

Ainsi,

$$J\left(\frac{u+v}{2}\right) = \frac{J(u) + J(v)}{2} - \frac{C}{8}\|u-v\|_{H^1(\Omega)}^2.$$

La fonction  $J$  est donc fortement convexe et admet donc un minimiseur sur  $V$  d'après le Théorème 9.2.6.

**Stokes.** L'énergie associée au système de Stokes est définie pour tout

$$u \in V := \{v \in H_0^1(\Omega)^N \text{ tel que } \operatorname{div} v = 0\}$$

par

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

La fonctionnelle  $J$  est évidemment continue,  $V$  est un espace de Hilbert et on peut établir que  $J$  est fortement convexe à l'aide de l'inégalité de Poincaré (comme pour l'Exercice 9.2.8). D'après le Théorème 9.2.6,  $J$  admet donc un minimiseur unique sur  $V$ .

## Chapitre 9

# CONDITIONS D'OPTIMALITÉ ET ALGORITHMES

**Exercice 9.1.1** Montrer que la dérivabilité de  $J$  en  $u$  implique la continuité de  $J$  en  $u$ . Montrer aussi que, si  $L_1, L_2$  vérifient

$$\begin{cases} J(u+w) \geq J(u) + L_1(w) + o(w) , \\ J(u+w) \leq J(u) + L_2(w) + o(w) , \end{cases} \quad (9.1)$$

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 (9.1). 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  $\alpha > 0$ , en appliquant l'inégalité précédente à  $\alpha w$  et en divisant par  $\alpha$ , on obtient

$$0 \geq (L_1 - L_2)(w) + \frac{o(\alpha w)}{\alpha}.$$

En faisant tendre  $\alpha$  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 9.1.2 (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,  $a(w, w)$  est égale à  $o(w)$ . La fonction  $J$  est donc dérivable et

$$\langle J'(u), w \rangle = a(u, w) - L(w).$$

**Exercice 9.1.3** 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.** C'est un cas particulier de l'Exercice précédent. On a

$$J(x + y) = J(x) + (Ax - b) \cdot y + Ay \cdot y/2.$$

Ainsi,  $J$  est dérivable et si on identifie  $\mathbb{R}^N$  et son dual à l'aide du produit scalaire euclidien, on obtient

$$J'(x) = Ax - b.$$

**Exercice 9.1.4** On reprend l'Exercice **10.1.2** avec  $V = L^2(\Omega)$  ( $\Omega$  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 9.1.2,

$$\langle J'(u), w \rangle = a(u, w) - L(w),$$

d'où

$$\langle J'(u), w \rangle = \int_{\Omega} (uw - fw) \, dx = \langle u - f, w \rangle_{L^2(\Omega)}.$$

En identifiant  $L^2(\Omega)$  et son dual à l'aide du produit scalaire  $L^2(\Omega)$ , on obtient  $J'(u) = u - f$ .

**Exercice 9.1.5** On reprend l'Exercice **10.1.2** avec  $V = H_0^1(\Omega)$  ( $\Omega$  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  $L(u) = \int_{\Omega} f u \, dx$  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.** D'après le résultat établi à l'Exercice 10.1.2, 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 à  $L^2(\Omega)$ . En effet, une intégration par partie conduit à

$$\langle J'(u), w \rangle = - \int_{\Omega} (\Delta u + f) w \, dx$$

d'où l'on déduit  $J'(u) = -\Delta u - f$  si l'on identifie le produit de dualité  $\langle J'(u), w \rangle$  au produit scalaire dans  $L^2(\Omega)$ ,  $\int_{\Omega} J'(u) w \, dx$ . Si on utilise le produit scalaire  $H^1(\Omega)$  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 de son dual à l'aide du produit scalaire  $H^1(\Omega)$ . 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, dans le cadre de cette identification de  $H_0^1(\Omega)$  avec son dual, on obtient  $J'(u) = u_0$ .

**Exercice 9.1.6** Soit  $\Omega$  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

$$\frac{d}{dx} \left( \frac{\partial L}{\partial p}(u'(x), u(x), x) \right) - \frac{\partial L}{\partial t}(u'(x), u(x), x) = 0, \quad (9.2)$$

presque partout dans l'intervalle  $]0, 1[$ .

3. Si  $L$  ne dépend pas de  $x$  (i.e.  $L = L(p, t)$ ) et si  $u$  est une solution de classe  $C^2(]0, 1[)$  de l'équation différentielle (9.2), 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

$$\left| L(p+q, t+s, x) - L(p, t, x) - \frac{\partial L}{\partial p}(p, t, x) \cdot q - \frac{\partial L}{\partial t}(p, t, x) s \right| \leq \frac{K}{2}(|q|^2 + |s|^2). \quad (9.3)$$

En particulier,

$$L(p, t, x) \leq C(1 + |p|^2 + t^2),$$

et  $J$  est correctement défini. On vérifie également que

$$\langle M(u), w \rangle = \int_{\Omega} \left( \frac{\partial L}{\partial p}(\nabla u, u, x) \cdot \nabla w + \frac{\partial L}{\partial t}(\nabla u, u, x) w \right) dx$$

est une forme linéaire continue sur  $H^1(\Omega)$ . Enfin, d'après l'inégalité (9.3)

$$|J(u+w) - J(u) - \langle M(u), w \rangle| \leq \frac{K}{2} \|w\|_{H^1}^2.$$

La fonction  $J$  est donc dérivable en  $u$  de dérivée  $J'(u) = M(u)$ .

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  $\partial L / \partial p(u', u, x)$  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 :

$$\begin{aligned} \frac{d}{dx} \left( L(u', u) - u' \frac{\partial L}{\partial p}(u', u) \right) &= \frac{d(L(u', u))}{dx} - u'' \frac{\partial L}{\partial p} - u' \frac{d}{dx} \left( \frac{\partial L}{\partial p}(u', u) \right) \\ &= u' \left( \frac{\partial L}{\partial t}(u', u) - \frac{d}{dx} \left( \frac{\partial L}{\partial p}(u', u) \right) \right) = 0. \end{aligned}$$

et  $L(u', u) - u' \partial L / \partial p(u', u)$  est constant sur l'intervalle  $[0, 1]$ .

**Exercice 9.1.7** 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  $u \neq v$ . Par convexité 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

$$2J\left(\frac{u+v}{2}\right) > J(u) + J(v),$$

ce qui prouve que  $J$  est strictement convexe. Reste à prouver l'équivalence entre la stricte convexité et la deuxième inégalité de l'énoncé.

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

$$J(u) > J(v) + \langle J'(v), u - v \rangle.$$

Par sommation, 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 9.1.8** 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 :

$$J'(u+w) = a(u, \cdot) + L + a(w, \cdot) = J'(u) + a(w, \cdot).$$

Ainsi,  $J''(u)w = a(w, \cdot)$  ou encore  $J''(u)(v, w) = a(w, v)$ .

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 9.1.4 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 9.1.5 est deux fois dérivable dans  $H_0^1(\Omega)$  et  $J''(u)(v, w) = \int_{\Omega} (\nabla v \cdot \nabla w + vw) \, dx$ .

**Exercice 9.1.9** 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

$$J''(u)(w, w) \geq 0 \quad \text{et} \quad J''(u)(w, w) \geq \alpha \|w\|^2 \quad \forall u, w \in V. \quad (9.4)$$

**Correction.** Rappelons que la Proposition **10.1.5** énonce que les assertions suivantes sont équivalentes :

$$J \text{ est } \alpha\text{-convexe sur } V,$$

$$J(v) \geq J(u) + \langle J'(u), v - u \rangle + \frac{\alpha}{2} \|v - u\|^2 \quad \forall u, v \in V,$$

$$\langle J'(u) - J'(v), u - v \rangle \geq \alpha \|u - v\|^2 \quad \forall u, v \in V.$$

Montrons que pour tout  $\alpha \geq 0$ , les conditions ci-dessus 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,

$$J(v) = J(u + w) = J(u) + \langle J'(u), w \rangle + \frac{1}{2} J''(u)(w, w) + o(\|w\|^2),$$

où  $w = v - u$ . Ainsi, pour tout  $w$ ,

$$J''(u)(w, w) + o(\|w\|^2) \geq \alpha \|w\|^2.$$

Soit  $\lambda$  un réel non nul. Remplaçant  $w$  par  $\lambda w$  dans l'équation précédente, il vient

$$\lambda^2 J''(u)(w, w) + o(\lambda^2 \|w\|^2) \geq \alpha \lambda^2 \|w\|^2$$

et

$$J''(u)(w, w) + o(\lambda^2 \|w\|^2) / \lambda^2 \geq \alpha \|w\|^2.$$

En faisant tendre  $\lambda$  vers zéro, on obtient  $J''(u)(w, w) \geq \alpha \|w\|^2$ . Réciproquement, si  $J''(u)(w, w) \geq \alpha \|w\|^2$ , on pose  $f(t) = J(u + t(v - u))$ . La fonction  $f$  est deux fois dérivable avec

$$f'(t) = \langle J'(u + t(v - u)), v - u \rangle$$

et

$$f''(t) = J''(u + t(v - u))(v - u, v - u) \geq \alpha \|v - u\|^2.$$

Ainsi,

$$f'(1) - f'(0) = \int_0^1 f''(t) dt \geq \alpha \|v - u\|^2$$

c'est-à-dire

$$\langle J'(v) - J'(u), v - u \rangle \geq \alpha \|v - u\|^2,$$

ce qui n'est rien d'autre que la troisième condition de la Proposition **10.1.5**.

**Exercice 9.2.1** Soit  $K$  un convexe fermé non vide d'un espace de Hilbert  $V$ . Pour  $x \in V$ , on cherche la projection  $x_K \in K$  de  $x$  sur  $K$  (voir le Théorème **12.1.10**)

$$\|x - x_K\|^2 = \min_{y \in K} \{J(y) := \|x - y\|^2\}.$$



Montrer que la condition nécessaire et suffisante

$$\langle J'(x_K), y - x_K \rangle \geq 0 \quad \forall y \in K \quad (9.5)$$

du Théorème 10.2.1 se ramène exactement à

$$\langle x_K - x, x_K - y \rangle \leq 0, \quad \forall y \in K. \quad (9.6)$$

**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  $x_K$  (9.5) 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 (9.6).

**Exercice 9.2.2** 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 introduit le réel  $\alpha$  défini par

$$\alpha = \inf_{u \in K, \|u\|=1} \|Au\|^2.$$

Comme la sphère unité de  $K$  est un fermé compact, l'infimum est atteint en un élément  $u$  de cette dernière. De plus,  $u$  ne peut appartenir à la fois au noyau de  $A$  et à son orthogonal, car dans ce cas  $u \cdot u = 0$  d'une part et  $\|u\|^2 = 1$  d'autre part. On en déduit que  $\alpha$  est strictement positif. Par conséquent,  $J$  est  $\alpha$ -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 9.2.3** 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 fonction  $J$  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  $y \in \text{Ker } B$ ,  $(A\bar{x} - b) \cdot y = 0$ , 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 9.2.4** 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.** On développe

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

ce qui prouve que  $J$  est dérivable en tout point  $u$  de  $H_0^1(\Omega)$  avec

$$\langle J'(u), v \rangle = \int_{\Omega} (\nabla u \cdot \nabla v - f v) dx.$$

Au point de minimum, noté  $u$ , de  $J$ , l'équation d'Euler dit que  $\langle J'(u), v \rangle = 0$  pour tout  $v \in H_0^1(\Omega)$ , ce qui est la formulation variationnelle annoncée.

**Exercice 9.2.5** 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

$$\langle J'(u), v - u \rangle \geq 0, \quad \text{pour tout } v \in K,$$

c'est-à-dire

$$a(u, v - u) \geq L(v - u), \quad \forall v \in K.$$

**Exercice 9.2.6** 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

$$\langle J'_1(u), v - u \rangle + J_2(v) - J_2(u) \geq 0 \quad \text{pour tout } v \in K$$

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,

$$J_1(v) - J_1(u) + J_2(v) - J_2(u) \geq 0 \quad \text{pour tout } v \in K$$

et  $u$  est un minimiseur de  $J_1 + J_2$  sur  $K$ .

**Exercice 9.2.7** Soit  $K$  un sous-ensemble d'un espace de Hilbert  $V$ . Montrer que pour tout  $v \in K$ ,

$$K(v) = \left\{ w \in V, \exists (v^n) \in K^{\mathbb{N}}, \exists (\varepsilon^n) \in (\mathbb{R}_+^*)^{\mathbb{N}}, \right. \\ \left. \lim_{n \rightarrow +\infty} v^n = v, \lim_{n \rightarrow +\infty} \varepsilon^n = 0, \lim_{n \rightarrow +\infty} \frac{v^n - v}{\varepsilon^n} = w \right\}$$

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  $\alpha$  un réel strictement positif. D'après la définition de  $K(v)$ , il existe une suite  $v_n$  d'éléments de  $K$ , une suite  $\varepsilon_n$  de réels positifs tels que  $v_n$  converge vers  $v$ ,  $\varepsilon_n$  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  $w$  un élément de  $V$  et  $w_m$  une suite d'éléments de  $K(v)$  tels 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  $\varepsilon_\delta = \varepsilon^{n,m} \in \mathbb{R}_+^*$  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  $v$  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  $w \in V$ , en posant  $v^n = v + \frac{rw}{n\|w\|}$  et  $\varepsilon^n = \frac{r}{n\|w\|}$ , on a

$$w = \lim_{n \rightarrow 0} \frac{v^n - v}{\varepsilon^n} \in K(v).$$

En d'autres termes,  $V \subset K(v)$ , d'où  $K(v) = V$ . Enfin, pour  $K = \{0\}$ ,  $K(0) = \{0\}$ .

**Exercice 9.2.8** Soit  $A$  une matrice 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  $B\bar{x} = c$ , on peut déterminer  $p$  à partir de la relation

$$BA^{-1}B^*p = c - BA^{-1}b.$$

Si  $B$  est de rang maximal  $m$ , alors  $BA^{-1}B^*$  est inversible et on a simplement

$$p = (BA^{-1}B^*)^{-1}(c - BA^{-1}b) \text{ et } x = A^{-1}b + A^{-1}B^*(BA^{-1}B^*)^{-1}(c - BA^{-1}b).$$

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 car l'ensemble de minimisation est vide! Si les contraintes sont redondantes, on peut éliminer les contraintes surnuméraires et se ramener au cas précédent. Autrement dit, il existe au moins un vecteur  $p \in \mathbb{R}^m$  tel que  $BA^{-1}B^*p = c - BA^{-1}b$  mais il n'est unique qu'à l'addition près d'un élément de  $\text{Ker } B^*$ . Par contre,  $\bar{x}$  est toujours défini de manière unique par la relation  $\bar{x} = A^{-1}(b + B^*p)$ .

**Exercice 9.2.9** 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

$$K = \{x \in \mathbb{R}^n : |x| = 1\},$$

où  $F(x) = 1 - |x|^2$ . Les fonctions  $J$  et  $F$  sont toutes deux dérivables avec

$$J'(x) = 2Ax \quad F'(x) = -2x.$$

Comme  $F'(x) \neq 0$  pour  $x \in K$ , d'après le Théorème 10.2.8, si  $\bar{x}$  est un point de minimum de  $J$  sur la sphère unité, il existe un multiplicateur de Lagrange  $\lambda$  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  $\lambda$ . 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  $\mu$ , on a

$$J(x) = \mu.$$

Le minimum de  $J$  est donc atteint pour les vecteurs propres de plus petite valeur propre.

**Exercice 9.2.10** Rappelons que le problème de Didon (Exemple 9.1.11) consiste à déterminer  $\xi$  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

$$L(y) = \int_0^\xi \sqrt{1 + |y'|^2} dx - l = 0.$$

En utilisant les résultats précédents et ceux de l'Exercice 10.1.6, montrer que la solution du problème de Didon est nécessairement un arc de cercle.

**Correction.** Soit  $y$  et  $\xi$  solution du problème de Didon. En particulier,  $y$  est solution du même problème pour  $\xi$  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  $\lambda$  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, pour tout  $v \in H_0^1(]0, \xi[)$ ,

$$\int_0^\xi \left( v + \frac{\lambda}{\sqrt{1 + |y'|^2}} y' v' \right) dx = 0.$$

En intégrant par partie le second membre de cette équation, comme la fonction test  $v$  est quelconque, on en déduit que

$$\lambda \left( \frac{y'}{\sqrt{1 + |y'|^2}} \right)' = 1$$

et donc qu'il existe une constante  $C$  telle que

$$\frac{y'}{\sqrt{1+|y'|^2}} = \lambda^{-1}x + C. \quad (9.7)$$

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 (9.7), 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  $\lambda$  associé à la contrainte sur la longueur n'est autre que le rayon du cercle obtenu.

**Exercice 9.2.11** On étudie la première valeur propre du Laplacien dans un domaine borné  $\Omega$  (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,

$$\left| \frac{u_n - u_p}{2} \right|_{H_0^1}^2 = \frac{|u_n|_{H_0^1(\Omega)}^2 + |u_p|_{H_0^1(\Omega)}^2}{2} - \left| \frac{u_n + u_p}{2} \right|_{H_0^1}^2. \quad (9.8)$$

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 (9.8),

$$\left| \frac{u_n - u_p}{2} \right|_{H_0^1}^2 = \frac{|u_n|_{H_0^1(\Omega)}^2 + |u_p|_{H_0^1(\Omega)}^2}{2} - \alpha_{n,p}^2 \left| \frac{u_n + u_p}{2\alpha_{n,p}} \right|_{H_0^1}^2.$$

Comme  $\frac{u_n + u_p}{2\alpha_{n,p}} \in K$ , on a donc

$$\left| \frac{u_n - u_p}{2} \right|_{H_0^1}^2 \leq \frac{|u_n|_{H_0^1(\Omega)}^2 + |u_p|_{H_0^1(\Omega)}^2}{2} - \alpha_{n,p}^2 \mu$$

dont le membre de droite tend vers 0 car  $\alpha_{n,p}$  tend vers 1 tandis que  $|u_n|_{H_0^1(\Omega)}^2$  et  $|u_p|_{H_0^1(\Omega)}^2$  tendent vers  $\mu$  par définition d'une suite minimisante. 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  $\lambda$  tel que

$$J'(u) + \lambda F'(u) = 0,$$



c'est-à-dire tel que pour tout  $v \in H_0^1(\Omega)$ ,

$$\int_{\Omega} \nabla u \cdot \nabla v dx = \lambda \int_{\Omega} uv dx.$$

Ainsi,  $u$  est un vecteur propre de valeur propre  $\lambda$ . En choisissant  $v = u$  dans l'expression précédente, on en déduit de plus que  $\lambda = \mu$ . Enfin, on vérifie sans peine que  $\lambda$  est nécessairement la plus petite valeur propre du Laplacien avec conditions aux bords de Dirichlet.

**Exercice 9.2.12** 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  $x \in \mathbb{R}^n$ . Dédire de la question précédente que, si  $A \geq B$ , alors  $B^{-1} \geq A^{-1}$ .

**Correction.** 1. Tout d'abord, les deux problèmes admettent tous deux une solution en tant que problème de maximisation d'une fonction continue sur un compact non vide. 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  $b$  supposé non nul, 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'optimalité du premier ordre, il existe  $\lambda$  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  $\lambda$  pour définir  $\bar{x}$  de manière unique. Comme  $A\bar{x} \cdot \bar{x} = 1$ , on en déduit que

$$\lambda^2 = (A^{-1}b \cdot b)^{-1}.$$

On en déduit ( $\lambda$  est nécessairement positif) que

$$\lambda = (A^{-1}b \cdot b)^{-1/2},$$

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

$$(A^{-1}b \cdot b)^{1/2} = \sup_{Ax \cdot x \leq 1} b \cdot x \leq \sup_{Bx \cdot x \leq 1} b \cdot x = (B^{-1}b \cdot b)^{1/2},$$

car l'ensemble  $\{Ax \cdot x \leq 1\}$  est inclus dans l'ensemble  $\{Bx \cdot x \leq 1\}$ , d'où l'on déduit  $B^{-1} \geq A^{-1}$ .

**Exercice 9.2.13** 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  $\rho$ , sa vitesse  $u$ , et sa température  $T$ , se retrouvent grâce aux moments de la fonction  $f(v)$

$$\rho = \int_{\mathbb{R}^N} f(v) dv, \quad \rho u = \int_{\mathbb{R}^N} v f(v) dv, \quad \frac{1}{2}\rho u^2 + \frac{N}{2}\rho T = \frac{1}{2} \int_{\mathbb{R}^N} |v|^2 f(v) dv. \quad (9.9)$$

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 telles que  $H(f) < +\infty$ . On minimise  $H$  sur cet espace sous les contraintes de moment (9.9), 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  $\theta$  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  $\lambda_1$  et  $\lambda_3$  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

$$f(v) = \exp(-1 - \lambda_1 - \lambda_2 \cdot v - \lambda_3 |v|^2).$$

Il est plus simple de réécrire la fonction  $f(v)$  avec trois autres paramètres  $\mu_1 \in \mathbb{R}$ ,  $\mu_2 \in \mathbb{R}^N$ ,  $\mu_3 \in \mathbb{R}$  sous la forme (équivalente)

$$f(v) = \mu_1 \exp\left(-\frac{|v - \mu_2|^2}{\mu_3}\right).$$

Il reste à calculer ces paramètres grâce aux contraintes (9.9). Par imparité on remarque tout d'abord que

$$\int_{\mathbb{R}^N} (v - \mu_2) \mu_1 \exp\left(-\frac{|v - \mu_2|^2}{\mu_3}\right) dv = 0,$$

c'est-à-dire que  $\rho(u - \mu_2) = 0$  et, comme  $\rho > 0$ , on a  $\mu_2 = u$ . De même,

$$\int_{\mathbb{R}^N} |v - \mu_2|^2 f(v) dv = \int_{\mathbb{R}^N} |v|^2 f(v) dv - 2\mu_2 \cdot \int_{\mathbb{R}^N} v f(v) dv + |\mu_2|^2 \int_{\mathbb{R}^N} f(v) dv$$

c'est-à-dire

$$\int_{\mathbb{R}^N} |v - \mu_2|^2 \mu_1 \exp\left(-\frac{|v - \mu_2|^2}{\mu_3}\right) dv = \rho u^2 + N\rho T - 2u \cdot \rho u + \rho |u|^2 = N\rho T.$$

Or, en posant  $w = (v - \mu_2)/\sqrt{\mu_3}$  il vient

$$\int_{\mathbb{R}^N} |v - \mu_2|^2 \mu_1 \exp\left(-\frac{|v - \mu_2|^2}{\mu_3}\right) dv = \mu_1 \mu_3^{1+N/2} \int_{\mathbb{R}^N} |w|^2 \exp(-|w|^2) dw$$

et

$$\int_{\mathbb{R}^N} |w|^2 \exp(-|w|^2) dw = \int_{\mathbb{R}^N} (-w/2) \cdot \nabla_w \exp(-|w|^2) dw = \frac{N}{2} \int_{\mathbb{R}^N} \exp(-|w|^2) dw.$$

Un calcul classique (utilisant des coordonnées radiales en dimension  $N = 2$ ) montre que

$$\int_{\mathbb{R}^N} \exp(-|w|^2) dw = \left( \int_{\mathbb{R}} \exp(-z^2) dz \right)^N = \pi^{N/2}.$$

Par conséquent,

$$N\rho T = \mu_1 \mu_3^{1+N/2} \frac{N}{2} \pi^{N/2},$$

tandis que

$$\rho = \mu_1 \mu_3^{N/2} \pi^{N/2},$$

d'où l'on déduit  $\mu_3 = 2T$  et  $\mu_1 = \rho/(2\pi T)^{N/2}$ , c'est-à-dire que  $f(v) = M(v)$ .

**Exercice 9.2.14** Calculer la condition nécessaire d'optimalité du second ordre pour chacun des problèmes d'optimisation suivants

1. Optimisation quadratique à contraintes linéaires (Exemple 9.1.6)

$$\inf_{x \in \text{Ker} B} \left\{ J(x) = \frac{1}{2} Ax \cdot x - b \cdot x \right\},$$

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  $\mathbb{R}^m$ .

2. Première valeur propre (Exemple 9.1.7)

$$\inf_{x \in \mathbb{R}^n, \|x\|=1} \{ J(x) = Ax \cdot x \},$$

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

$$Au - \lambda u = 0.$$

Comme

$$J''(u)(v, v) = 2Av \cdot v$$

et  $F''(u)(v, v) = 2v \cdot v$ , la condition d'optimalité d'ordre deux est donc

$$Av \cdot v - \lambda v \cdot v \geq 0$$

pour tout  $v$  tel que  $v \cdot u = 0$ . Autrement dit, comme la condition d'optimalité d'ordre un dit que  $\lambda$  est une valeur propre, la condition d'optimalité d'ordre deux dit en plus qu'il faut que ça soit la plus petite valeur propre de  $A$ .

**Exercice 9.2.15** 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  $F_i(x) = B_i \cdot x - c_i$  avec  $B_i$  le  $i$ -ème vecteur ligne de  $B$ . Les fonctions  $F_i$  sont dérivables et  $\langle F'_i(x), y \rangle = B_i \cdot y$ . De même, la fonction objectif

$$J(x) = \frac{1}{2}Ax \cdot x - b \cdot x$$

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

$$J'(x) + \sum_{i=1}^m p_i F'_i(x) = 0, \quad p_i \geq 0, \quad p_i F'_i = 0,$$

c'est-à-dire, en remarquant que la  $i$ -ème ligne  $B_i$  de  $B$  n'est rien d'autre que la  $i$ -ème colonne  $(B_i)^*$  de la matrice transposée  $B^*$ ,

$$A\bar{x} - b + \sum_{i=1}^m p_i (B_i)^* = 0, \quad p_i \geq 0, \quad p_i (B_i \cdot \bar{x} - c_i) = 0.$$

ou, sous une forme plus compacte, en sommant les contraintes de complémentarité

$$A\bar{x} - b + B^*p = 0, \quad p \geq 0, \quad p \cdot (B\bar{x} - c) = 0.$$

Notons que  $K$  étant convexe et  $J$  fortement convexe, il existe un unique minimiseur au problème considéré.

**Exercice 9.2.16** Soit  $f \in L^2(\Omega)$  une fonction définie sur un ouvert borné  $\Omega$ . 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  $u_\epsilon$ . Montrer que, soit  $u_\epsilon = 0$ , soit il existe  $\lambda > 0$  tel que  $u_\epsilon$  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  $u_\epsilon$  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  $F'(u_\epsilon) \neq 0$ . Les contraintes sont donc nécessairement qualifiées et 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  $u_\epsilon$  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 (cas  $\epsilon \geq \|f\|_{L^2}$ ), on a  $\lambda = 0$  et on trouve que  $u_\epsilon = 0$ .

**Exercice 9.3.1** On considère le problème d'optimisation, dit perturbé

$$\inf_{F_i(v) \leq u_i, 1 \leq i \leq m} J(v), \quad (9.10)$$

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 (9.10).

1. Montrer que si  $p$  est le multiplicateur de Lagrange pour le problème non perturbé (c'est-à-dire (9.10) avec  $u = 0$ ), alors

$$m^*(u) \geq m^*(0) - p \cdot u. \quad (9.11)$$

2. Dédire de (9.11) 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 (9.10) non perturbé est telle qu'il existe  $p_i \geq 0$  tel que

$$J'(\bar{v}) + p_i F'_i(\bar{v}) = 0, \quad p_i F_i(\bar{v}) = 0. \quad (9.12)$$

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 (9.12), 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  $F(v) \leq u$  et  $p \geq 0$  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  $m^*(u)$  en fonction de  $u$ . Plus  $p$  est important, plus une petite variation de  $u$  par rapport à zéro entraînera une forte variation de  $m^*$ . L'exemple **9.1.8** modélise les choix d'un ménage en matière de consommation entre différents produits pour un budget donné. Le ménage cherche à maximiser sa "fonction d'utilité" sous sa contrainte budgétaire. Le multiplicateur associé à la contrainte budgétaire n'est autre que l'utilité marginale, c'est-à-dire correspond à l'augmentation de la fonction d'utilité du ménage par rapport à l'augmentation de leur budget.

**Exercice 9.3.2** Donner un exemple de Lagrangien pour lequel l'inégalité

$$\inf_{v \in U} \left( \sup_{q \in P} \mathcal{L}(v, q) \right) \geq \sup_{q \in P} \left( \inf_{v \in U} \mathcal{L}(v, q) \right). \quad (9.13)$$

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

$$\inf_{v \in U} \left( \sup_{q \in P} \mathcal{L}(v, q) \right) = \sup_{\mathbb{R}} F > \inf_{\mathbb{R}} F = \sup_{q \in P} \left( \inf_{v \in U} \mathcal{L}(v, q) \right).$$

**Exercice 9.3.3** 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

$$F(q) = \mathcal{L}(\varphi(q), q) = \min_{v \in U} \mathcal{L}(v, q). \quad (9.14)$$

L'application  $F$  est l'infimum d'une famille de fonctions concaves continues. Elle est donc elle-même concave et, comme elle ne prend pas de valeurs infinies, elle est continue d'après l'Exercice 8.2.7. Comme  $P$  est compact et que  $F$  est continue,  $F$  admet au moins un maximum sur  $P$  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  $q \in P$ ,

$$\mathcal{L}(v^*, q) \leq \mathcal{L}(v^*, q^*). \quad (9.15)$$

Pour tout  $t \in [0, 1]$  et tout  $q \in P$ , on pose

$$v_t = \varphi((1-t)q^* + tq)$$

D'après la concavité de  $\mathcal{L}(v, \cdot)$ , on a pour tout  $v \in U$

$$\mathcal{L}(v, (1-t)q^* + tq) \geq (1-t)\mathcal{L}(v, q^*) + t\mathcal{L}(v, q), \quad (9.16)$$

d'où l'on déduit pour  $v = v_t$ , puisque  $q^*$  maximise  $F$  sur  $P$ , que

$$\begin{aligned} F(q^*) &\geq F((1-t)q^* + tq) = \mathcal{L}(v_t, (1-t)q^* + tq) \\ &\geq (1-t)\mathcal{L}(v_t, q^*) + t\mathcal{L}(v_t, q). \end{aligned}$$



Or, par définition (9.14) de  $F$ , on a  $\mathcal{L}(v_t, q^*) \geq F(q^*)$ , ce qui conduit à

$$F(q^*) \geq (1-t)F(q^*) + t\mathcal{L}(v_t, q),$$

ce qui donne en fin de compte que, pour tout  $q \in P$  et tout  $t \neq 0$ ,

$$F(q^*) \geq \mathcal{L}(v_t, q).$$

On ne peut malheureusement pas prendre  $t = 0$  dans l'inégalité ci-dessus et conclure, puisque  $v_0 = v^*$ , que  $\mathcal{L}(v^*, q) \leq F(q^*) = \mathcal{L}(v^*, q^*)$ . Il faut donc utiliser un argument de "passage à la limite" quand  $t$  tend vers zéro. 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

$$F(q^*) = \mathcal{L}(v^*, q^*) \geq \lim_{n \rightarrow +\infty} \mathcal{L}(v_{t_n}, q) = \mathcal{L}(\tilde{v}, q).$$

Pour conclure, il suffit donc de prouver que  $\tilde{v} = v^*$  et ainsi obtenir l'inégalité (9.15). La concavité (9.16) de  $\mathcal{L}(v_{t_n}, \cdot)$  et le fait que  $v_{t_n}$  minimise  $\mathcal{L}(v, (1-t_n)q^* + t_nq)$  donnent

$$(1-t_n)\mathcal{L}(v_{t_n}, q^*) + t_n\mathcal{L}(v_{t_n}, q) \leq \mathcal{L}(v_{t_n}, (1-t_n)q^* + t_nq) \leq \mathcal{L}(v, (1-t_n)q^* + t_nq).$$

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  $\tilde{v} = \varphi(q^*) = v^*$ , ce qui prouve (9.15).

**Exercice 9.3.4** 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 9.4.1** On considère le problème de commande optimal (10.72) avec  $K = \mathbb{R}^M$ ,  $f = 0$ ,  $z = 0$ , et  $z_T = 0$ , c'est-à-dire

$$\inf_{v(t) \in L^2([0, T]; \mathbb{R}^M)} J(v)$$

avec

$$J(v) = \frac{1}{2} \int_0^T Rv(t) \cdot v(t) dt + \frac{1}{2} \int_0^T Qy(t) \cdot y(t) dt + \frac{1}{2} Dy(T) \cdot y(T),$$

et  $y(t)$  la solution à valeurs dans  $\mathbb{R}^N$  du système différentiel linéaire

$$\begin{cases} \frac{dy}{dt} = Ay + Bv \text{ pour } 0 \leq t \leq T, \\ y(0) = y_0 \in \mathbb{R}^N. \end{cases}$$

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  $u$  le contrôle optimal du problème (10.72),  $y$  l'état du système et  $p$  l'état adjoint correspondant, défini par

$$\begin{cases} \frac{dp}{dt} = -A^* p - Qy \text{ pour } 0 \leq t \leq T, \\ p(T) = Dy(T). \end{cases}$$

On rappelle que la commande optimale est  $u = -R^{-1} B^* p$ . Ainsi, d'après l'équation différentielle ordinaire (10.71) vérifiée par  $y$ ,

$$\frac{dy}{dt} = Ay - BR^{-1} B^* p.$$

On en déduit que

$$\begin{aligned}\frac{d}{dt}(p \cdot y) &= -Qy \cdot y - A^*p \cdot y + p \cdot Ay - B^*p \cdot R^{-1}B^*p \\ &= -Qy \cdot y - R^{-1}B^*p \cdot B^*p.\end{aligned}$$

Par intégration, il vient

$$\begin{aligned}p(t) \cdot y(t) &= p(T) \cdot y(T) + \int_t^T (Qy \cdot y + R^{-1}B^*p \cdot B^*p) dt \\ &= Dy(T) \cdot y(T) + \int_t^T (Qy \cdot y + R^{-1}B^*p \cdot B^*p) dt.\end{aligned}$$

S'il existe  $t_0 \in [0, T]$  tel que  $y(t_0) = 0$ , on a  $p(t_0) \cdot y(t_0) = 0$ . 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 différentiel 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, la trajectoire optimale ne consiste pas à atteindre exactement la cible. Autrement dit, il n'est jamais "rentable" d'atteindre exactement la cible. Le coût du contrôle (terme en  $Rv \cdot v$ ) nécessaire pour s'approcher de l'état cible devient plus important que le gain réalisé dans les termes du type  $Qy \cdot y$  et  $Dy(T) \cdot y(T)$ .

**Exercice 9.4.2** 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 affine 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

$$\langle J'(v), w \rangle = 2 \left( \int_0^T \int_{\Omega} v w \, dx \, dt + \int_0^T \int_{\Omega} (y - z) y_w \, dx \, dt + \int_{\Omega} (y(T) - z_T) y_w(T) \, dx \right) \quad (9.17)$$

où  $y_w$  est solution du problème parabolique

$$\begin{cases} \frac{\partial y_w}{\partial t} - \Delta y_w = w & \text{dans } ]0, T[ \times \Omega \\ y_w = 0 & \text{sur } ]0, T[ \times \partial\Omega \\ y_w(0) = 0 & \text{dans } \Omega. \end{cases}$$

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

$$\begin{cases} -\frac{\partial p}{\partial t} - \Delta p = y - z & \text{dans } ]0, T[ \times \Omega \\ p = 0 & \text{sur } ]0, T[ \times \partial\Omega \\ p(T) = y(T) - z_T & \text{dans } \Omega. \end{cases}$$

Notons que, pour trouver l'état adjoint, on peut introduire un Lagrangien comme dans le cas de la dimension finie. On vérifie sans mal que

$$\langle J'(v), w \rangle = \int_0^T \int_{\Omega} (v + p) w \, dx.$$

La commande optimale est donc  $v = -p$ . L'état adjoint pour  $p$  est rétrograde en temps (puisque l'on impose une condition finale) mais le signe de la dérivée en temps est aussi inversé. Par le changement de variables  $\tilde{p}(t) = p(T - t)$  le lecteur se convaincra sans peine que le problème parabolique pour  $\tilde{p}$ , donc  $p$ , est bien posé.

**Exercice 9.4.3** Généraliser l'exercice précédent à l'équation des ondes.

**Correction.** Il s'agit d'étudier le problème hyperbolique

$$\begin{cases} \frac{\partial^2 y}{\partial t^2} - \Delta y = v + f & \text{dans } ]0, T[ \times \Omega \\ y = 0 & \text{sur } ]0, T[ \times \partial\Omega \\ y(0) = y_0 & \text{dans } \Omega \\ \frac{\partial y}{\partial t}(0) = y_1 & \text{dans } \Omega \end{cases}$$

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 (9.17) que précédemment. Cependant,  $y_w$  est, dans ce cas, solution du problème hyperbolique

$$\begin{cases} \frac{\partial^2 y_w}{\partial t^2} - \Delta y_w = w & \text{dans } ]0, T[ \times \Omega \\ y_w = 0 & \text{sur } ]0, T[ \times \partial\Omega \\ y_w(0) = 0 & \text{dans } \Omega \\ \frac{\partial y_w}{\partial t}(0) = 0 & \text{dans } \Omega. \end{cases}$$

A nouveau, on peut introduire un état adjoint afin de déterminer explicitement  $J'$ . L'équation vérifiée par l'état adjoint est

$$\begin{cases} \frac{\partial^2 p}{\partial t^2} - \Delta p = y - z & \text{dans } ]0, T[ \times \Omega \\ p = 0 & \text{sur } ]0, T[ \times \partial\Omega \\ p(T) = y(T) - z_T & \text{dans } \Omega \\ \frac{\partial p}{\partial t}(T) = 0 & \text{dans } \Omega \end{cases}$$

et un calcul simple conduit à

$$\langle J'(v), w \rangle = \int_0^T \int_{\Omega} (v + p)w \, dx \, dt.$$

Notons que pour trouver l'état adjoint, on peut introduire un Lagrangien comme dans le cas de la dimension finie. L'état adjoint pour  $p$  est rétrograde en temps (puisque l'on impose des conditions finales). Par le changement de variables  $\tilde{p}(t) = p(T - t)$  le lecteur se convaincra sans peine que le problème parabolique pour  $\tilde{p}$ , donc  $p$ , est bien posé.

**Exercice 9.5.1** 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  $\mu$  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 si 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 optimale, 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  $\alpha(rx, ry) = \alpha(x, y)$  puisque le gradient  $J'(x, y)$  est linéaire en  $(x, y)$ . On a donc  $\alpha(x_{n+2}, y_{n+2}) = \alpha(x_n, y_n)$  et, par récurrence,  $\alpha(x_{2p}, y_{2p}) = \alpha(x_0, y_0)$ . Ainsi,

$$(x_{2p}, y_{2p}) = \alpha(x_0, y_0)^p(x_0, y_0).$$

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  $\mu$ . 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  $\beta'$  de la suite dans le cas de l'algorithme à pas optimal. A titre indicatif, on obtient

$$\beta' = |a - b||x_0||y_0|\sqrt{ab} \left( (ax_0^2 + by_0^2)(a^3x_0^2 + b^3y_0^2) \right)^{-1/2}.$$

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 9.5.2** 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

$$u_{n+1} = P_K(u_n - \mu J'(u_n)) \tag{9.18}$$

définissant l'algorithme de gradient projeté à pas fixe en terme de multiplicateur de Lagrange.

**Correction.** L'opérateur de projection sur  $K$  est défini par

$$P_K(u) = u - (1 - u \cdot \mathbb{1})\mathbb{1}/N,$$

où  $\mathbb{1} = \sum_{i=1}^N e_i$ . L'algorithme du gradient projeté peut donc s'écrire sous la forme

$$\begin{aligned} u_{n+1} &= P_K(u_n - \mu J'(u_n)) = u_n - \mu(J'(u_n) - N^{-1}(J'(u_n) \cdot \mathbb{1})\mathbb{1}) \\ &= u_n - \mu(J'(u_n) - \lambda_n \mathbb{1}) \end{aligned}$$

avec

$$\lambda_n = N^{-1}(J'(u_n) \cdot \mathbb{I}).$$

Si un point fixe est atteint, on obtient qu'il existe un réel  $\lambda$  tel que

$$J'(u) - \lambda \mathbb{I} = 0,$$

et on retrouve la condition d'optimalité du premier ordre associée au problème de minimisation de  $J$  sur  $K$ .

**Exercice 9.5.3** Appliquer l'algorithme d'Uzawa au problème

$$\min_{v \in \mathbb{R}^N, F(v)=Bv-c \leq 0} \left\{ J(v) = \frac{1}{2}Av \cdot v - b \cdot v \right\}, \quad (9.19)$$

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

$$\mathcal{L}(u^n, p^n) = \min_v \mathcal{L}(v, p^n). \quad (9.20)$$

On rappelle que  $p^{n+1}$  est déterminé à l'aide de  $p^n$  par

$$p^{n+1} = P_{\mathbb{R}_+^M}(p^n + \mu F(u^n)), \quad (9.21)$$

où  $\mu$  est le pas de l'algorithme, choisit suffisamment petit. La matrice  $A$  étant symétrique définie positive, le problème (9.20) admet comme unique solution

$$u^n = A^{-1}(b - B^*p^n).$$

En explicitant la définition (9.21) 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 strictement contractante. Comme  $B$  est de rang  $M$ , la matrice  $BA^{-1}B^*$  est définie positive. Pour  $\mu$  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 la suite  $p^n$  converge. On note  $p$  sa limite. La suite  $u^n$  est également convergente et sa limite  $u$  est telle que

$$Au - b + B^*p = 0. \quad (9.22)$$

Enfin, comme  $p = P_{\mathbb{R}_+^M}(p + \mu F(u))$ , par définition de l'opérateur de projection, 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

$$F(u) \leq 0 \quad (9.23)$$

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,

$$F(u) \cdot p = 0. \quad (9.24)$$

De (9.22), (9.23) et (9.24), on conclut que  $u$  est solution du problème de minimisation étudié.

**Exercice 9.5.4** 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  $\varepsilon$  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,  $J_\varepsilon$  est dérivable et

$$J'_\varepsilon(v) = J'(v) + 2\varepsilon^{-1} \sum_{i=1}^M \max(F_i(v), 0) F'_i(v) .$$



Comme  $u_\varepsilon$  minimise  $J_\varepsilon$ , on a  $J'_\varepsilon(u_\varepsilon) = 0$  et

$$J'(u_\varepsilon) = -2\varepsilon^{-1} \sum_{i=1}^M \max(F_i(u_\varepsilon), 0) F'_i(u_\varepsilon). \quad (9.25)$$

De plus  $u_\varepsilon$  converge vers  $u$  pour lequel

$$J'(u) = - \sum_{i \in I(u)} \lambda_i F'_i(u). \quad (9.26)$$

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  $\mathbb{R}^N$  (appelée base duale) telle que

$$\langle F'_i(u), a_j \rangle = \delta_{ij}$$

pour tout  $i$  et  $j \in I(u)$ . Comme  $F'_i(u_\varepsilon)$  converge vers  $F'_i(u)$ , pour  $\varepsilon$  assez petit, la famille  $(F'_i(u_\varepsilon))_{i \in I(u)}$  est indépendante et il existe une famille  $(a_i^\varepsilon)_{i \in I(u)} \in \text{Vect}((a_i)_{i \in I(u)})$  telle que

$$\langle F'_i(u_\varepsilon), a_j^\varepsilon \rangle = \delta_{ij}$$

pour tout  $i$  et  $j \in I(u)$ . De plus, pour tout  $i \in I(u)$ ,  $a_i^\varepsilon$  converge vers  $a_i$ . Enfin, pour tout  $i \in I(u)$ ,

$$-2\varepsilon^{-1} \max(F_i(u_\varepsilon), 0) = \left\langle -2\varepsilon^{-1} \sum_{j \in I(u)} \max(F_j(u_\varepsilon), 0) F'_j(u_\varepsilon), a_i^\varepsilon \right\rangle.$$

Comme  $\varepsilon^{-1} \max(F_i(u_\varepsilon), 0)$  converge vers zéro pour tout  $i \notin I(u)$ ,

$$\begin{aligned} \lim_{\varepsilon} -2\varepsilon^{-1} \max(F_i(u_\varepsilon), 0) &= \lim_{\varepsilon} -2\varepsilon^{-1} \sum_{j=1}^M \max(F_j(u_\varepsilon), 0) \langle F'_j(u_\varepsilon), a_i^\varepsilon \rangle \\ &= \lim_{\varepsilon} \langle J'(u_\varepsilon), a_i^\varepsilon \rangle \\ &= \langle J'(u), a_i \rangle = \lambda_i. \end{aligned}$$

# Chapitre 10

## MÉTHODES DE LA RECHERCHE OPÉRATIONNELLE

Exercice 10.1.1 Résoudre le programme linéaire suivant

$$\max_{x_1 \geq 0, x_2 \geq 0} x_1 + 2x_2$$

sous les contraintes

$$\begin{cases} -3x_1 + 2x_2 \leq 2 \\ -x_1 + 2x_2 \leq 4 \\ x_1 + x_2 \leq 5 \end{cases}$$

**Correction.**

**Exercice 10.1.2** Montrer que l'on peut choisir la matrice  $A$  de taille  $m \times n$  et le vecteur  $b \in \mathbb{R}^m$  de telle façon que  $X_{ad}$  soit le cube unité  $[0, 1]^{n-m}$  dans le sous-espace affine de dimension  $n - m$  défini par  $Ax = b$ . En déduire que le nombre de sommets de  $X_{ad}$  est alors  $2^{n-m}$ .

**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  $x \in \mathbb{R}^N$ .

Exercice 10.1.4 Résoudre par l'algorithme du simplexe le programme linéaire

$$\min_{x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, x_5 \geq 0} x_1 + 2x_2$$

sous les contraintes

$$\begin{cases} -3x_1 + 2x_2 + x_3 = 2 \\ -x_1 + 2x_2 + x_4 = 4 \\ x_1 + x_2 + x_5 = 5 \end{cases}$$

**Exercice 10.1.5** Résoudre par l'algorithme du simplexe le programme linéaire

$$\min_{x_1 \geq 0, x_2 \geq 0} 2x_1 - x_2$$

sous les contraintes  $x_1 + x_2 \leq 1$  et  $x_2 - x_1 \leq 1/2$  (on pourra s'aider d'un dessin et introduire des variables d'écart).

**Exercice 10.1.6** Résoudre par l'algorithme du simplexe le programme linéaire

$$\min_{x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0} 3x_3 - x_4$$

sous les contraintes

$$\begin{cases} x_1 - 3x_3 + 3x_4 = 6 \\ x_2 - 8x_3 + 4x_4 = 4 \end{cases}$$

**Exercice 10.1.7** Montrer que, si  $X_{ad}$  est borné non vide, (??) admet une unique solution optimale  $x^\mu$ . Ecrire les conditions d'optimalité et en déduire que, si (??) admet une unique solution optimale  $x^0$ , alors  $x^\mu$  converge vers  $x^0$  lorsque  $\mu$  tend vers zéro.

**Exercice 10.1.8** Utiliser la dualité pour résoudre "à la main" (et sans calculs!) le programme linéaire

$$\min_{x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0} 8x_1 + 9x_2 + 4x_3 + 6x_4$$

sous les contraintes

$$\begin{cases} 4x_1 + x_2 + x_3 + 2x_4 \geq 1 \\ x_1 + 3x_2 + 2x_3 + x_4 \geq 1 \end{cases}$$

**Exercice 10.1.9** Trouver le problème dual de (??) lorsqu'on dualise aussi la contrainte  $x \geq 0$ , c'est-à-dire qu'on introduit le Lagrangien

$$L(x, p, q) = c \cdot x + p \cdot (b - Ax) - q \cdot x$$

avec  $q \in \mathbb{R}^n$  tel que  $q \geq 0$ . Comparer avec (??) et interpréter la nouvelle variable duale  $q$ . En déduire qu'il n'y a pas d'intérêt à "dualiser" aussi la contrainte  $x \geq 0$ .

**Exercice 10.1.10** Vérifier que le problème dual de (??) est à nouveau (??).

**Exercice 10.1.11** Soit  $v \in \mathbb{R}^N$ ,  $c \in \mathbb{R}^N$ ,  $A$  une matrice  $M \times N$  et  $b \in \mathbb{R}^M$ . On considère le programme linéaire

$$\inf_{\substack{v \geq 0 \\ Av \leq b}} c \cdot v. \quad (10.1)$$

Montrer que le problème dual peut se mettre sous la forme suivante, avec  $q \in \mathbb{R}^M$

$$\sup_{\substack{q \geq 0 \\ A^* q \leq c}} b \cdot q. \quad (10.2)$$

Soient  $v$  et  $q$  des solutions admissibles de (10.1) et (10.2), respectivement. Montrer que  $v$  et  $q$  sont des solutions optimales si, et seulement si,

$$(c - A^*q) \cdot v = 0 \quad \text{et} \quad (b - Ac) \cdot q = 0. \quad (10.3)$$

Les deux égalités de (10.3) sont appelées **conditions des écarts complémentaires** (primales et duales, respectivement). Généraliser au cas où le problème primal comprend en outre des contraintes égalités.

**Exercice 10.1.12** Montrer réciproquement que si  $x$  est un point de  $P$  vérifiant  $A'x = b'$ , avec  $A'$  et  $b'$  comme dans le Lemme ??, alors  $x$  est un point extrémal.

**Exercice 10.1.13** Il s'agit d'établir une réciproque au Corollaire ??. Commençons par examiner le cas spécial du polyèdre  $X_{ad} = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$  des solutions admissibles du programme linéaire standard, avec  $A \in \mathbb{Z}^{m \times n}$  de rang  $m$ . Montrer que les deux propriétés suivantes sont équivalentes :

1. pour tout  $b \in \mathbb{Z}^m$ , les points extrémaux de  $X_{ad}$  sont entiers ;
2. toutes les sous-matrices  $m \times m$  de  $A$  sont de déterminant  $\pm 1$  ou 0.

Soit maintenant  $D \in \mathbb{Z}^{m \times n}$ , et considérons  $Q = \{x \in \mathbb{R}^n \mid Dx \leq b, x \geq 0\}$ . Dédurre de l'équivalence qui précède l'équivalence des deux propriétés suivantes (théorème de Hoffman et Kruskal) :

3. pour tout  $b \in \mathbb{Z}^m$ , les points extrémaux de  $Q$  sont entiers ;
4.  $D$  est totalement unimodulaire.

**Exercice 10.1.14 (Problème de couverture)** Un centre d'appel téléphonique a une courbe de charge :  $c_t$  est le nombre de clients devant être servis à l'instant discret  $t \in \{1, \dots, T\}$ . Un certain nombre de conseillers de clientèle répondent aux appels. On simplifie le problème en supposant que tous les appels sont de même type. On supposera qu'il y a  $k$  horaires de travail possibles, l'horaire  $i$  étant caractérisé par un intervalle  $[\alpha_i, \beta_i]$ , avec  $1 \leq \alpha_i \leq \beta_i \leq T$ , ce qui revient à faire abstraction des pauses. On note  $S_i$  le salaire à verser à un conseiller de clientèle travaillant de l'instant  $\alpha_i$  à l'instant  $\beta_i$ . On pose  $u_{it} = 1$  si  $\alpha_i \leq t \leq \beta_i$ , et  $u_{it} = 0$  sinon. Justifier le problème

$$\inf_{\substack{x \in \mathbb{N}^k \\ \sum_{1 \leq i \leq k} x_i u_{it} \geq c_t, \forall 1 \leq t \leq T}} \sum_{1 \leq i \leq k} x_i S_i. \quad (10.4)$$

Montrer que l'ensemble des solutions admissibles de ce problème peut s'écrire comme l'ensemble des points entiers d'un polyèdre de la forme (??), où la matrice  $D$  est une matrice d'intervalles, c'est-à-dire une matrice à coefficients 0, 1 telle que les 1 apparaissent consécutivement sur une colonne. Montrer qu'une matrice d'intervalles est totalement unimodulaire. Conclure.

**Exercice 10.1.15** Expliciter le problème de flot à coût minimum correspondant au problème de transport particulier de l'Exemple ??. (On dessinera le graphe.)

**Exercice 10.1.16** Montrer que le problème du flot maximal est effectivement un cas particulier de problème de flot à coût minimal. (Indication : on pourra rajouter un arc au graphe intervenant dans la définition du problème du flot maximal.)

**Exercice 10.1.17** Montrer que le problème de couverture (Exercice 10.1.14) peut se modéliser par un problème de flot à coût minimum, et retrouver ainsi la conclusion de l'Exercice 10.1.14.

**Exercice 10.1.18** A un bal, il y a  $n$  garçons et  $n$  filles, chaque garçon ayant été présenté à  $r$  filles, et chaque fille ayant été présentée à  $r$  garçons. On suppose  $r \geq 1$ . Dédurre du Théorème de Birkhoff ?? qu'il est possible de former  $n$  couples de danseurs, de sorte que les danseurs d'un même couple aient déjà été présentés l'un à l'autre.

**Exercice 10.1.19** Nous dirons que le graphe est *coaccessible* pour  $\phi$  si pour tout nœud  $i$  du graphe, il existe un chemin de  $i$  à un nœud  $j$  tel que  $\phi_j \neq +\infty$ . On va montrer que si le graphe n'a pas de circuit de poids négatif et est co-accessible pour  $\phi$ , alors l'équation de Bellman (??) possède une unique solution finie, égale à  $v$ . Pour cela, on considère  $v' \in \mathbb{R}^{\mathcal{N}}$  une solution de l'équation de Bellman (??). On introduit

$$C = \{i \in \mathcal{N} \mid \phi_i \geq \min_{k \in \mathcal{N}, (i,k) \in \mathcal{A}} (c_{i,k} + v'_k)\} ,$$

et l'on choisit pour chaque  $i \in C$  un nœud  $\pi(i)$  tel que

$$v'_i = c_{i,\pi(i)} + v'_{\pi(i)} .$$

On définit ainsi une application  $\pi : C \rightarrow \mathcal{N}$ . Montrer que quel que soit  $i \in C$ , il existe un entier  $k$  tel que le  $k$ -ième itéré  $\pi^k(i)$ , n'appartienne pas à  $C$ . Conclure que  $v' \geq v$ .

**Exercice 10.1.20 (Convergence en temps fini)** Montrer que si  $\mathcal{G}$  n'a pas de circuits de poids strictement négatif, alors  $f^{|\mathcal{N}|-1}(\phi) = v$ . Si au contraire  $\mathcal{G}$  a un circuit de poids strictement négatif, et si  $\mathcal{G}$  est coaccessible pour  $\phi$  (voir l'Exercice 10.1.19 pour cette notion) alors  $f^{|\mathcal{N}|}(\phi) < f^{|\mathcal{N}|-1}(\phi)$ .

**Exercice 10.1.21** Soit  $\mathcal{G} = (\mathcal{N}, \mathcal{A})$  un graphe orienté, muni d'un poids  $c \in \mathbb{R}^{\mathcal{A}}$ . Montrer que les deux assertions suivantes sont équivalentes :

1.  $\mathcal{G}$  est sans circuits de poids strictement négatif,
2. il existe une fonction  $y \in \mathbb{R}^{\mathcal{N}}$  telle que  $-y_i + c_{i,j} + y_j \geq 0$ , pour tout  $(i, j) \in \mathcal{A}$ .

Indication pour  $1 \Rightarrow 2$ . On pourra commencer par supposer le graphe fortement connexe et exhiber un  $y$  "évident".

**Exercice 10.1.22** Comment peut-on compléter l'algorithme d'itération sur les valeurs pour construire un circuit de poids négatif ?

**Exercice 10.1.23 (Chemin de coût minimum avec contrainte de temps)** Soit  $\mathcal{G} = (\mathcal{N}, \mathcal{A})$  un graphe orienté, muni de deux valuations,  $c \in \mathbb{R}^{\mathcal{A}}$ , un coût, et  $\tau \in \mathbb{N}^{\mathcal{A}}$ , un temps, et  $\phi \in (\mathbb{R} \cup \{+\infty\})^{\mathcal{N}}$  une pénalité finale. On fixe un nœud initial  $s$

et une date limite  $T \in \mathbb{N}$ , et l'on cherche à trouver un chemin  $(\ell_0, \dots, \ell_m)$  de longueur arbitraire, partant de  $s$  (i.e.  $\ell_0 = s$ ), tel que le gain total  $c_{\ell_0, \ell_1} + \dots + c_{\ell_{m-1}, \ell_m} + \phi_{\ell_m}$  soit minimal, sous la contrainte de respecter la date limite  $T$ , i.e. sous la contrainte  $\tau_{\ell_0, \ell_1} + \dots + \tau_{\ell_{m-1}, \ell_m} \leq T$ . On supposera qu'il n'y a pas de circuit dont tous les arcs ont des temps nuls. Formuler un algorithme de programmation dynamique pour résoudre ce problème. Application : trouver par programmation dynamique le chemin de coût minimum du nœud 1 au nœud 6, en temps au plus 10, pour l'exemple de graphe (??) qui sera traité plus loin par relaxation Lagrangienne.

**Exercice 10.1.24** Considérons un amateur de théâtre, qui se rend en Juillet pour une journée voir des pièces du festival off d'Avignon. Le festival off comprend plusieurs centaines de pièces. Chaque pièce est caractérisée par un unique lieu de la ville (un théâtre), et une unique plage horaire dans la journée (par exemple, 16h00-17h30). On connaît en outre les temps nécessaires pour aller d'un théâtre à l'autre. En lisant le programme du off avant d'aller au festival, notre spectateur affecte à chaque pièce un plaisir espéré, mesuré sur une échelle de 0 à 5. Son but est de voir dans la journée une suite de pièces du off, de manière à maximiser la somme des plaisirs espérés pour les différentes pièces choisies. Montrer que ce problème peut se ramener à un chemin à coût minimal dans un graphe sans circuits, et qu'il peut se résoudre en un temps quadratique en le nombre de pièces du festival off (on négligera le temps des repas).

**Exercice 10.1.25** On peut imaginer que sur la planète Mars<sup>1</sup>, des extraterrestres apprennent aux enfants à compter avec l'addition  $(a, b) \mapsto a \oplus b = \min(a, b)$ , et la multiplication  $(a, b) \mapsto a \otimes b = a + b$ . La structure algébrique correspondante,  $(\mathbb{R} \cup \{+\infty\}, \oplus, \otimes)$  est appelée *semi-anneau min-plus*. Elle vérifie les mêmes axiomes que les anneaux, sauf que l'addition, au lieu d'être une loi de groupe, vérifie  $a \oplus a = a$ . Pour nos martiens, l'équation de Bellman associé au problème en horizon fini (??) n'est autre qu'un produit de matrices min-plus

$$v_i^T = \bigoplus_{k \in \mathcal{N}} M_{i,k} \otimes v_k^{T-1} ,$$

avec  $M_{i,k} = c_{i,k}$  si  $(i, k) \in \mathcal{A}$ , et  $M_{i,k} = +\infty$  sinon. Les martiens, qui utilisent les mêmes notations matricielles que nous, mais dans le semi-anneau min-plus, écrivent tout simplement,

$$v^t = Mv_{t-1}, \quad \text{et} \quad v^T = M^T \phi .$$

Quant à l'équation de Bellman (??), ils l'écrivent

$$v = Mv \oplus \phi . \tag{10.5}$$

Que devient le Théorème de la matrice de transfert ?? pour une matrice à coefficients dans le semi-anneau min-plus? Montrer que la solution maximale de (10.5) est égale à  $v = M^* \phi$ , où  $M^* = M^0 \oplus M \oplus M^2 \oplus \dots$ . Retrouver ainsi le Théorème ???. Montrer que si  $\mathcal{G}$  n'a pas de circuit de poids négatif,  $\lim_{T \rightarrow \infty} M^T = +\infty$  (la matrice dont tous les coefficients sont égaux à  $+\infty$ ). Retrouver ainsi le résultat d'unicité de l'Exercice 10.1.19.

1. Nous empruntons cette plaisanterie à V.P. Maslov, *Méthodes Opératorielles*, MIR, 1973

**Exercice 10.1.26** Appelons jeu Markovien fini un jeu à deux joueurs, “blanc”, et “noir”, qui déplacent à tour de rôle un jeton sur un graphe fini, à partir d’une position initiale (c’est blanc qui commence). Certains sommets du graphe sont finaux : lorsque le jeton est dans ce sommet, on sait que la partie est gagnante pour blanc, ou gagnante pour noir, ou bien nulle. Peut-on modéliser les échecs ou les dames par un tel jeu ? On suppose que la partie termine toujours. Écrire une équation de type programmation dynamique exprimant la valeur d’une position pour blanc. (Indication : cette équation fera intervenir à la fois les lois  $\min$  et  $\max$ .) Dédurre que des trois assertions suivantes, une seule est vraie : - les blancs peuvent toujours gagner ; - les noirs peuvent toujours gagner ; - les blancs et les noirs peuvent toujours forcer la partie nulle.

**Exercice 10.1.27 (Décomposition des flots)** Montrer que tout flot de  $s$  à  $p$  dans le graphe  $\mathcal{G} = (\mathcal{N}, \mathcal{A})$  peut s’écrire comme somme d’au plus  $|\mathcal{A}|$  flots associés soit à des chemins élémentaires de  $s$  à  $p$  soit à des circuits élémentaires.

**Exercice 10.1.28** Utiliser l’exercice qui précède pour formuler le problème du plus court chemin de  $s$  à  $p$  comme un problème de flot à coût minimum.

**Exercice 10.1.29** Nous dirons qu’un flot est **sans circuits** s’il peut s’écrire comme somme de flots associés à des chemins élémentaires, dans la décomposition de l’Exercice 10.1.27. Montrer que pour tout flot admissible, il existe un flot sans circuit de même valeur. Conclure la preuve du Théorème ?? dans le cas où les capacités peuvent être infinies.

**Exercice 10.1.30** Utiliser le théorème de dualité en programmation linéaire pour donner une autre preuve du Théorème de Ford et Fulkerson ?. [ATTENTION, BESOIN D’INDICATIONS, DIFFICILE]

**Exercice 10.1.31** Montrer que la valeur des flots admissibles de  $s$  à  $p$  n’est pas supérieurement bornée si, et seulement si, il existe un chemin de  $s$  à  $p$  formé d’arcs de capacité infinie.

**Exercice 10.1.32** Montrer que dans le problème  $(VC)_{\text{rel}}$ , on peut remplacer les contraintes (??) par les contraintes :

$$\text{pour tout } \mathcal{S} \subset \mathcal{V}, \text{ tel que } \mathcal{S} \neq \emptyset \text{ et } \mathcal{S} \neq \mathcal{V}, \quad \sum_{k \in \mathcal{V}, m \in \mathcal{S} \setminus \mathcal{V}, \{k, m\} \in \mathcal{E}} x_{km} \geq 2. \quad (10.6)$$

Donner un algorithme polynômial qui vérifie si les contraintes de type (10.6) sont satisfaites par un vecteur  $x \in [0, 1]^{\mathcal{E}}$ , et retourne le cas échéant l’ensemble  $\mathcal{S}$  associé à une contrainte non-satisfaite.

**Exercice 10.1.33** Calculer la fonction duale  $\mathcal{G}(\lambda)$  pour le problème (??), et retrouver ainsi la valeur de  $\max_{\lambda \in \mathbb{R}_+} \mathcal{G}(\lambda)$ .

**Exercice 10.1.34** Proposer une relaxation Lagrangienne fournissant une borne supérieure pour le problème du sac-à-dos (??), et l’appliquer à l’exemple (??).

**Exercice 10.1.35** La version *orientée* du problème du voyageur de commerce consiste à considérer un graphe orienté  $\mathcal{G} = (\mathcal{N}, \mathcal{A})$ , muni d'une fonction temps  $t : \mathcal{A} \rightarrow \mathbb{R}_+$ , et à chercher un circuit orienté passant une et une seule fois par chaque nœud. On demande de proposer une relaxation Lagrangienne, en faisant en sorte que pour chaque vecteur  $\lambda$  de multiplicateurs de Lagrange, le calcul de la fonction duale  $\mathcal{G}(\lambda)$  revienne à résoudre un problème d'affectation. Expliquer pourquoi cette relaxation est moins intéressante que la borne de Held et Karp, dans le cas particulier du problème du voyageur de commerce non-orienté.

**Exercice 10.1.36** Nous allons prouver la formule

$$\partial J(x) = \text{co} \left( \bigcup_{\substack{i \in I \\ J_i(x) = J(x)}} \partial J_i(x) \right) \quad (10.7)$$

dans le cas où tous les  $J_i$  sont des fonctions convexes de  $\mathbb{R}^n$  dans  $\mathbb{R}$ . (L'exclusion de la valeur  $+\infty$  n'est qu'une commodité, qui nous permettra d'appliquer directement les résultats du Chapitre 9, prouvés dans le cas de fonctions convexes à valeurs finies). Notons  $C$  le second membre de (10.7).

1. Montrer que  $C \subset \partial J(x)$ .
2. On suppose que  $0 \in \partial J(x)$ . En considérant le programme convexe

$$\min \quad t \quad \text{sous les contraintes} \quad y \in \mathbb{R}^n, \quad t \in \mathbb{R}, \quad J_i(y) \leq t, \quad \forall i \in I,$$

montrer que  $0 \in C$ .

3. Conclure que de manière générale,  $C = \partial J(x)$ .





## 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. On note  $\langle x, y \rangle = \sum_{i=1}^n x_i \bar{y}_i$  le produit hermitien de deux vecteurs  $x, y \in \mathbb{C}^n$ . Tout d'abord, on rappelle que les valeurs singulières de  $A$  sont les racines carrées des valeurs propres de la matrice auto-adjointe (hermitienne)  $A^*A$  (dont on vérifie sans peine qu'elle n'admet que des valeurs propres réelles positives). Par définition, on a

$$\|A\|_2 = \left( \max_{x \in \mathbb{C}^n, x \neq 0} \frac{\langle Ax, Ax \rangle}{\langle x, x \rangle} \right)^{1/2}.$$

Ainsi,

$$\|A\|_2 = \left( \max_{x \in \mathbb{C}^n, x \neq 0} \frac{\langle A^*Ax, x \rangle}{\langle x, x \rangle} \right)^{1/2}$$

est bien le maximum des valeurs singulières de  $A$ . Par ailleurs, pour tout  $x \in \mathbb{C}^n$ ,

$$\|x\|_2 = \sup_{y \in \mathbb{C}^n, \|y\|_2 \leq 1} |\langle x, y \rangle|.$$

Ainsi,

$$\|Ax\|_2 = \sup_{y \in \mathbb{C}^n, \|y\|_2 \leq 1} |\langle Ax, y \rangle| = \sup_{y \in \mathbb{C}^n, \|y\|_2 \leq 1} |\langle x, A^*y \rangle| \leq \|x\|_2 \|A^*\|_2.$$

On en déduit que  $\|A\|_2 \leq \|A^*\|_2$  et finalement  $\|A\|_2 = \|A^*\|_2$ .

2. On a

$$\|A\|_1 = \max_{x \in \mathbb{C}^n, x \neq 0} \frac{\|Ax\|_1}{\|x\|_1} = \max_{x \in \mathbb{C}^n, x \neq 0} \frac{\sum_{i=1}^n \left| \sum_{k=1}^n a_{ik} x_k \right|}{\sum_{k=1}^n |x_k|}.$$

Pour tout indice  $j$ , en choisissant  $x_k = \delta_{kj}$ , on obtient

$$\|A\|_1 \geq \sum_{i=1}^n |a_{ij}|.$$

De plus,

$$\begin{aligned}\|A\|_1 &= \max_{x \in \mathbb{C}, x \neq 0} \frac{\sum_{i=1}^n \left| \sum_{j=1}^n a_{ij} x_j \right|}{\sum_{j=1}^n |x_j|} \leq \max_{x \in \mathbb{C}, x \neq 0} \frac{\sum_{i,j=1}^n |a_{ij}| |x_j|}{\sum_{j=1}^n |x_j|} \\ &= \max_{x \in \mathbb{C}, x \neq 0} \frac{\sum_{j=1}^n \left( \sum_{i=1}^n |a_{ij}| \right) |x_j|}{\sum_{j=1}^n |x_j|} \leq \max_j \sum_{i=1}^n |a_{ij}|,\end{aligned}$$

d'où l'égalité demandée.

3. On a

$$\|A\|_\infty = \max_{x \in \mathbb{C}, x \neq 0} \left( \frac{\max_k \left| \sum_{j=1}^n a_{kj} x_j \right|}{\max_k |x_k|} \right).$$

Soit  $i \in \{1, \dots, n\}$  et  $x \in \mathbb{C}^n$  tel que, pour tout indice  $j$ ,  $x_j$  soit égal au signe de  $a_{ij}$ . On déduit de l'expression précédente que

$$\|A\|_\infty \geq \max_k \left( \sum_{j=1}^n |a_{kj}| \right).$$

Réciproquement,

$$\|A\|_\infty \leq \max_{x \in \mathbb{C}, x \neq 0} \left( \frac{\max_k \sum_{j=1}^n |a_{kj}| |x_j|}{\max_k |x_k|} \right) \leq \max_k \sum_{j=1}^n |a_{kj}|.$$

**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. On a

$$\text{cond}(A) = \|A\| \|A^{-1}\| = \|A^{-1}\| \|A\| = \text{cond}(A^{-1}).$$

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 dans une base orthonormée), 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) ou par (13.2) 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

$$\begin{aligned} (\mathcal{K}_h u^k)_j &= h^{-1} (-u_{j-1}^k + 2u_j^k - u_{j+1}^k) \\ &= h^{-1} \left( \sin \left( \frac{(j-1)k\pi}{n+1} \right) + 2 \sin \left( \frac{jk\pi}{n+1} \right) - \sin \left( \frac{(j+1)k\pi}{n+1} \right) \right) \\ &= (2ih)^{-1} \left( -e^{\frac{i(j-1)k\pi}{n+1}} + 2e^{\frac{ijk\pi}{n+1}} - e^{\frac{i(j+1)k\pi}{n+1}} - e^{-\frac{i(j-1)k\pi}{n+1}} + 2e^{-\frac{ijk\pi}{n+1}} - e^{-\frac{i(j+1)k\pi}{n+1}} \right) \\ &= (2ih)^{-1} \left( e^{\frac{ijk\pi}{n+1}} - e^{-\frac{ijk\pi}{n+1}} \right) \left( -e^{\frac{ik\pi}{n+1}} + 2 - e^{-\frac{ik\pi}{n+1}} \right) \\ &= 4h^{-1} \sin \left( \frac{jk\pi}{n+1} \right) \sin^2 \left( \frac{k\pi}{2(n+1)} \right) \\ &= 4h^{-1} \sin^2 \left( \frac{k\pi}{2(n+1)} \right) u_j^k. \end{aligned}$$

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 demie largeur de bande  $p$ . On raisonne par récurrence afin de prouver que les matrices  $L$  et  $U$  sont également des matrices bande de demie largeur de bande  $p$ . On lit par ordre croissant les colonnes de  $A$  ce qui permet de déterminer les colonnes correspondantes de  $L$  et  $U$ . La première colonne de  $A$  donne

$$a_{11} = l_{11}u_{11} \Rightarrow u_{11} = a_{11}$$

et, pour  $i \geq 2$ ,

$$a_{i1} = l_{i1}u_{11} \Rightarrow l_{i1} = a_{i1}/u_{11}$$

ce qui montre que les premières colonnes de  $L$  et  $U$  ont bien la même demie largeur de bande  $p$ . Supposons que les  $j-1$  premières colonnes de  $L$  et  $U$  sont aussi de demie largeur de bande  $p$ . Les composantes de la  $j$  ème colonne de  $U$  sont définies pour  $1 \leq i \leq j$  par

$$u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik}u_{kj}.$$

La matrice  $A$  étant une matrice bande de demi-largeur  $p$ , on a  $a_{ij} = 0$  pour tout  $i$  tels que  $j > i + p$ . Par une récurrence évidente sur l'indice  $i$  croissant, on en déduit que  $u_{ij} = 0$  pour tout  $i$  tel que  $j > i + p$ . Ainsi, la  $j$ ème colonne de  $U$  est celle d'une matrice bande de demie largeur de bande  $p$ . La  $j$ ème colonne de  $L$  est ensuite déterminée, pour  $j+1 \leq i \leq n$ , par

$$l_{ij} = \frac{a_{ij} - \sum_{k=1}^{j-1} l_{ik}u_{kj}}{u_{jj}}.$$

D'après l'hypothèse de récurrence sur la structure bande des  $j$  premières colonnes de  $L$ , on a  $l_{ik} = 0$  dès que  $i - k > p$ . Ainsi, chaque terme de la somme apparaissant dans l'expression de  $l_{ij}$  est nul dès que  $i - (j-1) > p$  et en particulier dès que  $i - j > p$ . Ainsi,  $l_{ij} = 0$  dès que  $i - j > p$  et les  $j$  premières colonnes de  $L$  ont une structure de matrice bande de demie 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  $\mathcal{O}(np^2)$  et celui de la factorisation de Cholesky est  $\mathcal{O}(np^2/2)$ .

**Correction.** On suppose que  $n \gg p \gg 1$  et on peut donc négliger les "effets de bord" pour les indices extrêmes. Puisque  $L$  et  $U$  ont la même structure bande que  $A$ , la factorisation LU conduit aux opérations suivantes

$$a_{ij} = \sum_{k=\max(i-p, j-p)}^{\min(i, j)} l_{ik} u_{kj}.$$

Raisonnons à  $i$  fixé : pour  $j$  allant de 1 à  $i - p$  il n'y a aucun calcul à faire, pour  $j$  allant de  $i - p$  à  $i$  on fait de 1 à  $p$  multiplications, de même pour  $j$  allant de  $i$  à  $i + p$  on fait de  $p$  à 1 multiplications, et enfin aucun calcul pour  $j$  allant de  $i + p$  à  $n$ . Pour la  $i$ ème ligne on fait donc  $p^2$  opérations, ce qui conduit au total à de l'ordre de  $(np^2)$  opérations.

Pour la factorisation de Cholesky on fait moitié moins de calculs puisque la matrice est symétrique.

**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 déterminant de  $M^{-1}N$  vaut

$$\det(M^{-1}N) = \det N / \det M = \det\left(\frac{1 - \omega}{\omega} D + F\right) / \det\left(\frac{1}{\omega} D - E\right) = (1 - \omega)^n.$$

On en déduit que

$$\rho(M^{-1}N)^n \geq \prod_{i=1}^n |\lambda_i(M^{-1}N)| = |\det(M^{-1}N)| = |1 - \omega|^n,$$

donc  $\rho(M^{-1}N) \geq |1 - \omega|$  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 défini 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  $K_{k-1} = [r_0, \dots, r_{k-1}]$  (c'est vrai pour  $k = 1$ ). 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

$$K_k = K_{k-1} = [r_0, \dots, r_{k-1}] = [r_0, \dots, r_k].$$

Dans tous les cas, on a donc bien  $K_k = [r_0, \dots, r_k]$ . Montrons maintenant que  $K_k = [d_0, \dots, d_k]$ . 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  $K_{k-1} = [d_0, \dots, d_{k-1}]$  (c'est vrai pour  $k = 1$ ). 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, c'est-à-dire si  $y_{k+1}$  appartient à  $K_{k-1}$ , alors on a à la fois

$$r_{k+1} = r_0 - Ay_{k+1} \perp K_{k-1} \subset K_k \text{ avec } y_{k+1} \in K_{k-1}$$

et

$$r_k = r_0 - Ay_k \perp K_{k-1} \text{ avec } y_k \in K_{k-1}.$$

On en déduit par unicité de  $y_k$  que  $y_k = y_{k+1}$  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 note le produit scalaire  $\langle x, y \rangle_A = \langle Ax, y \rangle$ . 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 itération, on effectue de l'ordre de  $n^2$  opérations, l'essentiel du temps étant consacré au calcul du produit matrice-vecteur  $Ap_k$ . Dans le cas le plus défavorable, l'algorithme converge au bout de  $n$  itérations. Dans ce cas, le nombre d'opérations 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), c'est-à-dire

$$\tilde{A}\tilde{x} = \tilde{b}, \quad \text{où } \tilde{A} = B^{-1}AB^{-*}, \tilde{b} = B^{-1}b \text{ et } \tilde{x} = B^*x.$$

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

$$\begin{array}{l} \text{initialisation} \\ \text{itérations } k \geq 1 \end{array} \left\{ \begin{array}{l} \text{choix initial } x_0 \\ r_0 = b - Ax_0 \\ p_0 = z_0 = C^{-1}r_0 \\ \\ \alpha_{k-1} = \frac{z_{k-1} \cdot r_{k-1}}{Ap_{k-1} \cdot p_{k-1}} \\ x_k = x_{k-1} + \alpha_{k-1}p_{k-1} \\ r_k = r_{k-1} - \alpha_{k-1}Ap_{k-1} \\ z_k = C^{-1}r_k \\ \beta_{k-1} = \frac{z_k \cdot r_k}{z_{k-1} \cdot r_{k-1}} \\ p_k = z_k + \beta_{k-1}p_{k-1} \end{array} \right.$$

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  $h = 1/(n + 1)$

$$A = h^{-1} \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}. \quad (13.2)$$

Montrer que pour la valeur optimale

$$\omega_{opt} = \frac{2}{1 + 2 \sin \frac{\pi}{2(n+1)}} \simeq 2\left(1 - \frac{\pi}{n}\right)$$

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 rappelle qu'en notant  $D = \text{diag}(A)$  et  $-E$  la partie strictement inférieure de  $A$ , de telle façon que  $A = D - E - E^*$ , pour  $\omega \in ]0, 2[$ , on définit

$$C_\omega = \frac{\omega}{2 - \omega} \left( \frac{D}{\omega} - E \right) D^{-1} \left( \frac{D}{\omega} - E^* \right).$$

On note  $B_\omega$  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  $\tilde{A}_\omega$  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

$$\|\tilde{A}_\omega\|_2 = \max_{y \neq 0} \frac{\langle Ay, y \rangle}{\langle B_\omega^T y, B_\omega^T y \rangle} = \max_{y \neq 0} \frac{\langle Ay, y \rangle}{\langle B_\omega B_\omega^T y, y \rangle} = \max_{y \neq 0} \frac{\langle Ay, y \rangle}{\langle C_\omega y, y \rangle}.$$

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  $C_\omega$  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  $\beta = 1$  convient.

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  $D/4 = D^{-1}$ , un calcul simple montre que  $\langle Gx, x \rangle = -\frac{\|x_1\|^2}{2h}$  et donc

$$(2 - \omega) \frac{\langle C_\omega x, x \rangle}{\langle Ax, x \rangle} \leq 1 + a \frac{\langle Dx, x \rangle}{\langle Ax, x \rangle} = 1 + \frac{2a}{h} \frac{\|x\|^2}{\langle Ax, x \rangle} \leq 1 + \frac{2a}{h\lambda_{\min}(A)},$$

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

$$\begin{aligned} \alpha &= (2 - \omega) \left( 1 + \frac{a}{2 \sin^2 \frac{\pi}{2(n+1)}} \right)^{-1} \\ &= \left( \frac{1}{2 - \omega} + \frac{2 - \omega}{8\omega \sin^2 \frac{\pi}{2(n+1)}} \right)^{-1}. \end{aligned}$$

et

$$\text{cond}_2(C_\omega^{-1}A) \leq \frac{1}{2-\omega} + \frac{2-\omega}{8\omega \sin^2 \frac{\pi}{2(n+1)}}.$$

La minimisation du terme de droite par rapport à  $\omega$  conduit à la valeur optimale

$$\omega_{opt} = \frac{2}{1 + 2 \sin \frac{\pi}{2(n+1)}} \simeq 2\left(1 - \frac{\pi}{n}\right)$$

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] BONY J.-M., *Cours d'analyse. Théorie des distributions et analyse de Fourier*, Éditions de l'École Polytechnique, Palaiseau (2001).
- [4] CIARLET P.G., *Introduction à l'analyse numérique matricielle et à l'optimisation*, Masson, Paris (1982).