Résolution pratique des équations aux dérivées partielles

David Manceau

Table des matières

Introduction

1	Exe	mples	classiques d'e.d.p.	9
	1.1	Modél	isation, mise en équation	9
		1.1.1	Équation de la chaleur	9
		1.1.2	Mouvement d'un fluide irrotationnel et incompressible	11
		1.1.3	Trafic routier	11
		1.1.4	Équation des ondes	12
	1.2	Classif	fication des e.d.p	13

I Méthode des différences finies

 $\mathbf{17}$

 $\mathbf{7}$

2	Pro	blèmes elliptiques	19				
	2.1	Problèmes elliptiques 1d. approche formelle	19				
	2.2	Justification des calculs : consistance, stabilité et convergence	22				
	2.3	Problèmes elliptiques en dimension supérieure à 1	27				
3	Pro	blèmes hyperboliques	31				
	3.1	Définitions et solution explicite	31				
	3.2	Méthode numérique : approche formelle	32				
	3.3	Consistance, stabilité et convergence	33				
	3.4	Étude de la stabilité par analyse de Fourier	34				
	3.5	Exemples de schémas	37				
		3.5.1 Le schéma explicite centré	37				
		3.5.2 Les schémas explicites décentrés	37				
		3.5.3 Le schéma de Lax	38				
		3.5.4 Le schéma implicite centré	39				
		3.5.5 Le schéma de Lax-Wendroff	40				
4	Problèmes paraboliques						
	4.1	Propriétés	41				
	4.2	Exemples de schémas	42				
		4.2.1 Le schéma explicite centré	42				
		4.2.2 Le schéma implicite centré	43				
		4.2.3 Le θ -schéma et le schéma de Crank-Nicholson	44				

5	Bilan sur les différences finies et exercices5.1Problèmes elliptiques	47 47
	5.2Problèmes hyperboliques5.3Problèmes paraboliques	48 51
Π	Méthode des éléments finis	53
6	Entroduction à la notion de formulation variationnelle5.1Bref rappel sur les distributions5.2Les formules de Green5.3Applications à l'équation de Laplace5.4Un première formulation variationnelle du problème de Dirichlet homogène	55 55 57 58 60
7	Espaces de Sobolev 7.1 Définitions et propriétés 7.2 L'espace $H_0^1(\Omega)$ 7.3 Traces et formules de Green 7.4 Théorème de compacité de Rellich et applications 7.5 Dualité	 63 63 66 68 70 72
8	Analyse variationnelle des problèmes elliptiques 8.1 Théorie abstraite	75 75 77 81 84
9	Approximation variationnelle0.1Approximation interne et système matriciel équivalent	87 87 88
10	Méthode des éléments finis en dimension 1.0.1 Éléments finis \mathbb{P}_1 .0.2 Convergence et estimation d'erreur pour la méthode \mathbb{P}_1 .0.3 Cas d'une condition de Neumann.0.4 Méthode des éléments finis \mathbb{P}_2	91 91 95 98 100
11	Méthode des éléments finis en dimension $d \ge 2$ 1 1.1 Maillages triangulaires 1 1.2 Éléments finis \mathbb{P}_k en dimension $d \ge 2$ 1 1.3 Calcul du second membre et assemblage de la matrice de rigidité 1 11.3.1 Calcul du second membre 1 11.3.2 Assemblage de la matrice de rigidité 1 1.4 Remarques sur le cas Ω non polyédrique 1	103 106 111 111 112 113
	1.5 Maillages rectangulaires	114

12 Théorie spectrale	117
12.1 Rappels	. 118
12.2 Application au cadre variationnel	. 119
12.3 Analyse numérique spectrale	. 121
13 Problèmes paraboliques	125
13.1 Formulation variationnelle	. 125
13.2 Résolution numérique par éléments finis	. 128
Bibliographie	131

Introduction

L'objectif de ce cours est de présenter la résolution numérique de certaines e.d.p. (équations aux dérivées partielles) linéaires d'ordre 2 (*i.e.* contenant des dérivées d'ordre au plus 2) à travers des exemples classiques intervenant dans la modélisation de nombreux problèmes de physique, biologie, économie, ... etc.

Pour donner un aperçu général succinct de ce qu'est la résolution numérique d'une e.d.p., considérons une e.d.p. écrite sous la forme générale

$$\mathcal{A}u = f \quad \text{dans } \Omega, \tag{1}$$

où Ω est un ouvert de \mathbb{R}^n , f représente les données du problème, \mathcal{A} est un opérateur (représentant les dérivées) et u est la solution du problème.

Résoudre numériquement l'e.d.p. (1) consiste à calculer une approximation de la solution u sous forme d'un vecteur $U \in \mathbb{R}^N$, $N \in \mathbb{N}$. Pour calculer cette approximation, on utilise une méthode numérique dont l'objectif est de déterminer une matrice carrée Ad'ordre N (on notera $A \in \mathbb{R}^{N \times N}$) et un vecteur $b \in \mathbb{R}^N$ tels que U soit solution du système linéaire

$$4U = b. (2)$$

Un exemple simple consiste à prendre pour vecteur U une approximation du vecteur $(u(x_1), \ldots, u(x_N))^T$, où x_1, \ldots, x_N sont N points distincts de Ω (maillage de Ω). C'est le choix de la méthode numérique employée qui détermine A et b, dans le sens où deux méthodes numériques différentes peuvent amener à des choix de A et b différentes. Le système (2) représente alors une discrétisation de l'e.d.p. (1).

Dans le cadre de ce cours, on présente deux familles de méthodes numériques : la méthode des différences finies (la plus simple à mettre en place mais limitée, notamment, à des domaines Ω simples), et la méthode des éléments finis. Ces deux méthodes comportent chacune une idée clé :

- 1. Différences finies : utilisation du développement de Taylor pour approcher les dérivées.
- 2. Élements finis : utilisation de la formulation variationnelle de l'e.d.p. puis approximation de l'espace des solutions par des espaces vectoriels de dimension finie.

Lorsque l'on a obtenu le système (2), il faut s'assurer que celui-ci admet une solution unique et que celle-ci est une "bonne" approximation de u. Ensuite, on résout le système AU = b.

En TP, on travaillera avec le logiciel de calcul Scilab qui possède une fonction dédiée à la résolution des systèmes linéaires. Ainsi, le travail le plus important consistera à calculer A et b. Dans le cas de problèmes complexes (domaine Ω , fonction f "compliqués"), il peut s'avérer nécessaire de considérer un maillage très fin (nombre N très grand). Dans ce cas, le calcul de la solution du système linéaire par la fonction préprogrammée de Scilab peut devenir très long. Pour mener à bien ce calcul, il faut alors étudier plus précisément la matrice A du système et utiliser une méthode de résolution de problème matriciel adaptée à la matrice A. Cet aspect de la résolution numérique des e.d.p. ne sera pas étudié dans ce cours (sur ce sujet, on pourra notamment consulter l'annexe de [1] et l'ouvrage [2]).

La première partie de ce cours est consacrée à la méthode des différences finies et son contenu s'inspire de [1, 3, 6, 7] et [9]. La seconde partie concerne la méthode des éléments finis et est très largement inspirée de [1] et [12].

Chapitre 1

Exemples classiques d'e.d.p.

1.1 Modélisation, mise en équation

Dans cette première partie, on décrit certains phénomènes aboutissant aux équations qui seront étudiées par la suite.

1.1.1 Équation de la chaleur

Soit Ω un domaine de \mathbb{R}^d , $d \geq 1$. On suppose le domaine Ω occupé par un matériau homogène, isotrope et conducteur de la chaleur (plaque de métal par exemple). On note xla variable d'espace (un point de Ω) et t la variable de temps. Les sources de chaleur à l'instant $t \geq 0$ et au point $x \in \Omega$ sont notées f(t, x) et la température u(t, x), toutes deux exprimées en Kelvin K.

La quantité de chaleur est c u où c est la chaleur spécifique (constante dépendant du matériau) en J.kg⁻¹.K⁻¹ (J=Joules). D'après la loi de conservation de l'énergie, pour tout volume élémentaire V, la variation en temps de la quantité de chaleur est le bilan de ce qui est produit par les sources ou rentre à travers les parois. Autrement dit, on a

$$\frac{d}{dt}\left(\int_{V} c \, u \, dx\right) = \int_{V} f \, dx - \int_{\partial V} q \cdot \nu \, d\sigma, \qquad (1.1.1)$$

où ∂V est le bord de V, ν la normale extérieure unitaire au bord ∂V , q le vecteur flux de chaleur et \cdot désigne le produit scalaire de \mathbb{R}^d . D'après le théorème de Gauss (ou théorème de la divergence, ou formule de Green), on a

$$\int_{\partial V} q \cdot \nu \, d\sigma = \int_{\Omega} \operatorname{div}(q) \, dx, \qquad (1.1.2)$$

où la **divergence** de la fonction vectorielle $q = (q_1, \ldots, q_d)^T$ est donnée par

$$\operatorname{div}(q) := \sum_{i=1}^{n} \frac{\partial q_i}{\partial x_i}.$$

D'après (1.1.1) et (1.1.2), on obtient

$$c\,\partial_t u = f - \operatorname{div}(q).$$

De plus, d'après la loi de Fourier, on a

$$q = -k\,\nabla u,$$

où k est la conductivité thermique du matériau (constante) en W.m⁻¹.K⁻¹ (m=mètres, W=Watt) et le **gradient** de la fonction scalaire u est défini par

$$\nabla u := (\partial_{x_1} u, \dots, \partial_{x_d} u)^T.$$

On en déduit

$$c \partial_t u - \operatorname{div}(k\nabla u) = f.$$

L'opérateur laplacien est défini par $\Delta := \partial_{x_1}^2 + \cdots + \partial_{x_n}^2$. Alors, puisque k est constante et div $(\nabla \cdot) = \Delta$, on obtient l'équation de la chaleur

$$c\,\partial_t u - k\Delta u = f.\tag{1.1.3}$$

Il est nécessaire d'ajouter à cette équation des conditions aux limites (*i.e.* portant sur le comportement de la solution au bord du domaine Ω). On donne ci-dessous les deux exemples les plus classiques de conditions aux limites (d'autres conditions existent).

1. Condition de Dirichlet.

Si le bord est un conducteur thermique idéal (les échanges de chaleur se font de manière instantanée), on a

$$\forall x \in \partial \Omega, \ \forall t > 0, \quad u(t, x) = g(t, x),$$

où g est la température ambiante. En général, g est constante. Dans le cas g = 0, la condition de Dirichlet est dite homogène.

2. Condition de Neumann.

Si le bord est un mauvais conducteur (par exemple l'air), en négligeant les effets de radiation, on a

$$\frac{\partial u}{\partial \nu} = 0 \quad \text{sur } \mathbb{R}^+ \times \partial \Omega,$$

où $\frac{\partial u}{\partial \nu} = \nabla u \cdot \nu$ est la dérivée normale de u.

Enfin, il faut ajouter une condition initiale

 $\forall x \in \Omega, \quad u(t=0,x) = u_0(x),$

où u_0 est la température initiale du domaine Ω .

Alors, dans le cas d'un condition de Dirichlet homogène l'équation complète s'écrit

$$\begin{cases} c \partial_t u(t,x) - k\Delta u(t,x) = f(t,x) & \text{dans } \mathbb{R}^+_* \times \Omega, \\ u(t,x) = 0 & \text{dans } \mathbb{R}^+_* \times \partial\Omega, \\ u(t=0,x) = u_0(x) & \text{dans } \Omega. \end{cases}$$
(1.1.4)

Remarque 1.1.1. Si le terme source f ne dépend pas du temps, pour t suffisamment grand, u devient aussi indépendant du temps. On dit alors que u est à l'état d'équilibre. Dans ce cas, u(t, x) = u(x) d'où $\partial_t u = 0$ et u est solution de l'équation de Poisson

$$\Delta u = f', \quad \text{dans } \Omega, \tag{1.1.5}$$

où $f' := k^{-1} f$.

1.1.2 Mouvement d'un fluide irrotationnel et incompressible

On suppose $\Omega \subset \mathbb{R}^d$ occupé par un fluide irrotationnel, incompressible à l'équilibre (*i.e.* qui ne varie pas en fonction du temps). Alors, la vitesse locale $u : \Omega \to \mathbb{R}^d$ du fluide vérifie

$$\operatorname{rot}(u) = 0$$
 et $\operatorname{div}(u) = 0$,

où rot(u) est le **rotationnel** de u défini comme étant la matrice carrée d'ordre d de coefficients

$$\operatorname{rot}(u)_{ij} := \frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}, \quad \forall \ i, j = 1, \dots, d.$$

On suppose de plus l'ouvert Ω simplement connexe. Alors, rot(u) = 0 entraîne qu'il existe une fonction $\varphi : \Omega \to \mathbb{R}$ appelée **fonction potentiel** telle que

$$u = \nabla \varphi.$$

Remarque 1.1.2. L'égalité $rot(\nabla \cdot) = 0$ a toujours lieu. Par contre, la réciproque nécessite de se placer dans un ouvert simplement connexe.

Puisque div(u) = 0 et $u = \nabla \varphi$, on en déduit que φ vérifie l'équation de Laplace

$$\Delta \varphi = 0$$
, dans Ω ,

à laquelle il faut ensuite ajouter des conditions aux limites.

1.1.3 Trafic routier

Dans le cas d'une route à une seule voie, on note u le nombre de voitures par unité de longueur de la route (on parle de densité). On se place sur un tronçon de route $\Omega := [0, 1]$ en supposant qu'il n'y a pas d'arrêts ni de sorties. La quantité M(t) de voitures sur le tronçon Ω à l'instant t est donnée par

$$M(t) := \int_0^1 u(t,x) \, dx.$$

Soit F le flux de voitures par unité de temps qui passent à l'instant t au point x, *i.e.* moyenne du nombre de voitures passant en x. Par conservation de la masse, la variation en temps de la masse totale est égale au flux rentrant moins le flux sortant, *i.e.*

$$\partial_t M(t) = F(t,0) - F(t,1) = -\int_0^1 \partial_x F(t,y) \, dy.$$

Puisque $\partial_t M(t) = \int_0^1 \partial_t u(t, x) \, dx$, on en déduit (formellement) l'équation de conservation

$$\partial_t u + \partial_x F = 0$$
 dans $[0, 1]$.

Or Le flux F est une fonction de la densité u : F(t, x) = f(u(t, x)). En supposant que chaque voiture roule à une vitesse moyenne constante v, on a F(t, x) = v u(x) et l'on obtient l'équation de transport (ou d'advection) de vitesse v:

$$\partial_t u + v \,\partial_x u = 0 \quad \text{dans } \mathbb{R}^+ \times [0, 1].$$
 (1.1.6)

1.1.4 Équation des ondes

On considère une corde tendue fixée à ses deux extrémités, représentée par l'intervalle [0, L], L > 0. Sous l'action d'une force normale d'amplitude f = f(t, x), la corde se déforme (ou vibre). On suppose la masse m de la corde répartie uniformément et la corde entièrement élastique. On cherche alors à déterminer les vibrations (déplacements) de la corde en tout point x et temps t, soit l'inconnue u = u(t, x). Pour tout $x \in [0, L]$, on



FIGURE 1.1 – Corde soumise à une force normale f.

note w(t, x) la position (dans \mathbb{R}^2) de x à l'instant t. On a alors $w(t, x) = x e_1 + u(t, x) e_2$, où (e_1, e_2) est la base canonique de \mathbb{R}^2 . On désigne par T(t, x) l'action (tension) en w(t, x)de la partie droite de la corde sur la partie gauche. En tout point, la tension est tangente à la corde et donc il existe T_1 tel que

$$T(t,x) = T_1(t,x) \,\partial_x w(t,x).$$

On suppose dans la suite T_1 constante. Soit $\Delta x > 0$ petit. Alors, la section de la corde $[x, x + \Delta x]$ est soumise à la force

$$T(t, x + \Delta x) - T(t, x) + \Delta x f(t, x) e_2 = T_1(\partial_x w(t, x + \Delta x) - \partial_x w(t, x)) + \Delta x f(t, x) e_2.$$

La densité linéaire de la corde est le réel $\rho := m/L$. Celle-ci permet de déterminer la masse de la corde dans la section $[x, x + \Delta x]$ qui est alors donnée par $\rho \Delta x$. En appliquant la loi de Newton, on obtient ainsi

$$T_1(\partial_x w(t, x + \Delta x) - \partial_x w(t, x)) + \Delta x f(t, x) e_2 = \rho \,\Delta x \,\partial_t^2 w(t, x)$$

En divisant par Δx puis en faisant $\Delta x \to 0$, on aboutit à $T_1 \partial_x^2 w - \rho \partial_t^2 w = -f e_2$. En projetant cette équation dans la direction e_2 , on obtient que u est solution de l'équation des ondes

$$\partial_t^2 u - c^2 \,\partial_x^2 u = g, \quad \text{dans } \mathbb{R}^+_* \times [0, L], \tag{1.1.7}$$

où $c^2 := \rho/T_1$ et $g := f/T_1$. La corde étant fixée au bord, on a la condition aux limites de Dirichlet homogène

L'équation étant de degré 2 en temps, il est nécessaire d'ajouter deux conditions initiales :

$$\begin{cases} u(t=0,x) = u_0(x) \text{ pour } x \in [0,L], \\ \partial_t u(t=0,x) = u_1(x) \text{ pour } x \in [0,L]. \end{cases}$$
(1.1.9)

1.2 Classification des e.d.p.

Avant de donner une classification des e.d.p. linéaires du second ordre, on définit la notion de problème bien posé due à Hadamard. Celle-ci est de grande importance dans la résolution numérique des e.d.p. pour s'assurer que la solution approchée calculée est effectivement proche de la solution exacte du problème considéré.

Pour énoncer cette définition, il est nécessaire de donner une formulation générale des e.d.p. étudiées. Pour cela, on note f les données du problèmes (second membre, données initiales, etc.), u la solution du problème et \mathcal{A} une application qui agit sur u. Le problème consiste alors à trouver u solution de

$$\mathcal{A}(u) = f. \tag{1.2.1}$$

Lorsque \mathcal{A} est une application linéaire (ce qui sera toujours le cas dans ce cours) l'e.d.p. sera dite **linéaire** (et **non linéaire** le cas échéant).

Définition 1.2.1. Le problème (1.2.1) est dit **bien posé** si pour toute donnée f il existe une solution unique u, et si cette solution dépend continûment de f.

Remarque 1.2.2. Cette définition contient trois conditions dont on donne ici l'importance pour la résolution numérique.

- 1. Existence d'une solution, sans quoi l'élaboration d'une méthode numérique n'aurait pas d'intérêt.
- 2. L'unicité de la solution, assure que si la méthode numérique converge alors celle-ci converge vers la "bonne" solution.
- 3. La continuité de la solution par rapport aux données. Cette dernière condition est très importante numériquement puisque dans le cadre d'une méthode numérique les données du problèmes sont elles même approchées. Ainsi, si la solution du problème n'est pas continue par rapport aux données, le fait d'approcher ces données peut amener à fortement perturber la solution à calculer. La continuité assure qu'une légère perturbation des données n'entraînera qu'une faible perturbation de la solution.

Définition 1.2.3. On appelle **ordre** d'une équation aux dérivées partielles l'ordre de la plus grande dérivée apparaissant dans l'équation.

On donne ci-dessous une classification des e.d.p. linéaires du second ordre portant sur des fonctions de deux variables réelles u(x, y). Une telle équation s'écrit

$$a \partial_x^2 u + b \partial_{xy}^2 u + c \partial_y^2 u + d \partial_x u + e \partial_y u + f u = g, \qquad (1.2.2)$$

où, pour simplifier, a, b, c, d, e, f, g sont supposés constants.

Définition 1.2.4. L'équation (1.2.2) est dite

- 1. elliptique si $b^2 4ac < 0$,
- 2. parabolique si $b^2 4ac = 0$,
- 3. hyperbolique si $b^2 4ac > 0$.

Ci-dessous, on donne les exemples classiques d'e.d.p. linéaires du second ordre correspondant aux trois classifications (elliptique, parabolique, hyperbolique). Dans la suite du cours, ce sont ces exemples que l'on étudiera numériquement.

1. Problèmes elliptiques.

L'équation elliptique modèle est l'équation de Poisson donnée par (1.1.5). Comme vu précédemment celle-ci modélise, entre autres, la répartition de la chaleur à l'état d'équilibre ou encore le potentiel lié à la vitesse d'un fluide irrotationnel, incompressible en équilibre.

2. Problèmes paraboliques

L'exemple modèle est celui de la conduction de la chaleur instationnaire donnée par (1.1.3). Un autre modèle amenant à une e.d.p. parabolique est le **modèle de Black et Scholes** (issu de la finance).

On note u(t, x) la valeur d'une option au temps t > 0 pour un actif $x \in \mathbb{R}$. Si on note σ la volatilité (degré de dépendance d'un titre par rapport aux fluctuations du marché) de l'action et r le taux d'intérêt, alors u vérifie

$$\partial_t u(t,x) + r \, x \, \partial_x u(t,x) + \frac{\sigma^2 x^2}{2} \, \partial_x^2 u(t,x) = r \, u(t,x), \quad (t,x) \in (0,T) \times \mathbb{R}.$$

La condition initiale est une condition finale, *i.e.* on fixe la valeur de l'option au temps T:

$$u(t = T, x) = \max(x - k, 0)$$

Ce modèle permet de déterminer la valeur initiale (au temps t = 0) à donner à l'option pour que celle-ci atteigne la valeur k au bout d'un temps T. On peut montrer au prix de changements de variables que la résolution de l'équation de Black et Scholes se ramène à celle de l'équation de la chaleur.

3. Problèmes hyperboliques

L'exemple modèle est celui de l'équation des ondes donnée par (1.1.7). Cette équation peut se ramener à une équation du type (1.1.6). Pour cela, on pose $v := \partial_t u$ et $w := \partial_x u$. Alors, on a $\partial_t v = f + c^2 \partial_x w$ et $\partial_x v = \partial_t w$. En notant $U := (v, w)^T$, on obtient

$$\partial_t U = \begin{pmatrix} f \\ 0 \end{pmatrix} + \begin{pmatrix} c^2 \partial_x w \\ \partial_x u \end{pmatrix}$$
$$= \begin{pmatrix} f \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & c^2 \\ 1 & 0 \end{pmatrix} \partial_x U.$$

En posant $A := \begin{pmatrix} 0 & c^2 \\ 1 & 0 \end{pmatrix}$ et $F := \begin{pmatrix} f \\ 0 \end{pmatrix}$, on en déduit que U vérifie l'équation d'advection vectorielle

$$\partial_t U - A \,\partial_x U = F.$$

Ainsi, il est plus commode de considérer comme modèle hyperbolique l'équation (plus simple) d'advection donnée par (1.1.6).

Première partie Méthode des différences finies

Chapitre 2

Problèmes elliptiques

2.1 Problèmes elliptiques 1d, approche formelle

On considère une barre électrique chauffée par une source de chaleur f dont les extrémités sont plongées dans la glace. Soit u(x) la température au point x de la barre. On modélise la barre par l'intervalle [0, 1]. Alors, dans le cas stationnaire, u est solution de

$$\begin{cases} -u''(x) = f(x), \quad x \in]0, 1[, \\ u(0) = u(1) = 0. \end{cases}$$
(2.1.1)

On va étudier la résolution numérique de l'e.d.p. (2.1.1) en supposant la solution u régulière (autrement dit $u \in C^2(0,1)$). On admettra dans cette partie le caractère bien posé de cette e.d.p. qui sera démontré plus loin.

Remarque 2.1.1. Pour que la solution u de (2.1.1) soit régulière, il est nécessaire que f soit continue. Dans ce cas, il est alors assez simple de déterminer u. L'intérêt de développer une méthode numérique pour résoudre l'équation (2.1.1) réside dans le fait que cette méthode s'adapte ensuite à tout problème elliptique et s'écrit simplement dans le cas de l'équation (2.1.1).

On décrit la méthode en trois parties : choix du maillage, choix du schéma numérique et détermination du problème discret.

1ère étape : Choix de la discrétisation, maillage.

On considère une subdivision

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

de l'intervalle [0, 1], où $N \in \mathbb{N}$. Pour i = 0, ..., N, on pose $h_i := x_{i+1} - x_i$. Le **pas du maillage** est défini par

$$h := \max_{i=0,\dots,N} h_i.$$

Pour la méthode des différences finies, on considère un pas de maillage constant, *i.e.*

$$h = h_i, \quad \forall \ i = 0, \dots, N.$$

On a alors $x_{i+1} = x_i + h$ pour tout i = 0, ..., N. La première étape de discrétisation consiste à remplacer le problème (2.1.1) par

$$\begin{cases} -u''(x_i) = f(x_i), \quad \forall \ i = 1, \dots, N, \\ u(x_0) = u(x_{N+1}) = 0. \end{cases}$$
(2.1.2)

2ème étape : Construction d'un schéma numérique.

u

On suppose $u \in C^2(0, 1)$. Alors u admet un développement limité sous la forme

$$u(x_{i+1}) = u(x_i + h)$$

= $u(x_i) + h u'(x_i) + \frac{h^2}{2}u''(x_i) + O(h^3),$

 et

$$(x_{i-1}) = u(x_i - h)$$

= $u(x_i) - h u'(x_i) + \frac{h^2}{2}u''(x_i) + O(h^3),$

où $|O(h^3)| \leq c h^3$ où c est une constante indépendante de h. En additionnant les deux égalités précédentes, on obtient l'expression suivante pour $u''(x_i)$:

$$u''(x_i) = \frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2} + O(h).$$
(2.1.3)

Autrement dit, l'expression

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

est une approximation de $u''(x_i)$ pour *h* suffisamment petit. Pour i = 0, ..., N + 1, on note u_i une approximation de $u(x_i)$, d'où $u_0 = u_{N+1} = 0$ (prise en compte des conditions aux limites). Alors, à priori (cela sera justifié dans la section suivante), l'expression

$$\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2},\tag{2.1.4}$$

est une approximation de $u''(x_i)$.

Remarque 2.1.2. Ce choix d'approximation n'est pas unique, il est celui qui détermine le schéma numérique considéré. Dans notre cas le schéma sera dit **centré** car il fait intervenir i - 1 et i + 1, d'où une symétrie par rapport à l'indice i.

Avec ce choix d'approximation, on peut approcher le problème (2.1.2) par le problème discret suivant : trouver $u_1, \ldots, u_N \in \mathbb{R}$ tels que

$$\begin{cases}
-\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} = f(x_i), \quad \forall \ i = 1, \dots, N, \\
u_0 = u_{N+1} = 0.
\end{cases}$$
(2.1.5)

3ème étape : Passage au problème matriciel.

On va écrire (2.1.5) sous forme d'un système matriciel. On pose $U_h = (u_1, \ldots, u_N)^T$. Il faut tout d'abord prendre en compte les conditions aux limites $u_0 = u_{N+1} = 0$ (celles-ci n'interviennent que à ce niveau). Pour i = 1, puisque $u_0 = 0$, le problème se simplifie

$$-\frac{u_2 - 2\,u_1}{h^2} = f(x_1)$$

De même, pour i = N, on a

$$-\frac{-2\,u_N+u_{N-1}}{h^2} = f(x_N)$$

Ainsi, (2.1.5) s'écrit

$$-\frac{1}{h^2} \begin{pmatrix} -2 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \dots & 0 \\ \vdots & \ddots & & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 1 & -2 \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ \vdots \\ \vdots \\ u_N \end{pmatrix} = \begin{pmatrix} f(x_1) \\ \vdots \\ \vdots \\ \vdots \\ f(x_N) \end{pmatrix}.$$

Autrement dit, le vecteur U_h est solution du système matriciel

$$A_h U_h = b_h, (2.1.6)$$

où $A_h \in \mathbb{R}^{N \times N}$ et $b_h \in \mathbb{R}^N$ sont donnés par

$$A_{h} = -\frac{1}{h^{2}} \begin{pmatrix} -2 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 1 & -2 \end{pmatrix} \quad \text{et} \quad b_{h} = \begin{pmatrix} f(x_{1}) \\ \vdots \\ \vdots \\ f(x_{N}) \end{pmatrix}.$$
(2.1.7)

Ainsi, on a la méthode suivante pour obtenir une approximation numérique de la solution u de (2.1.1):

- 1. On choisit un pas de maillage h > 0 petit (détermine la subdivision $(x_i)_{i=0,\dots,N+1}$).
- 2. On détermine une approximation de $u''(x_i)$ par les développements de Taylor,
- 3. On en déduit un système matriciel $A_h U_h = b_h$ dont la solution $U_h = (u_1, \ldots, u_N)^T$ approche le vecteur $(u(x_1), \ldots, u(x_N))^T$.
- 4. On résout le système $A_h U_h = b_h$.

Exemple 2.1.3. On donne dans la figure 2.1 les graphes obtenus pour f = 2 suivant le nombre de points de discrétisation choisi, ainsi que le graphe de la solution exacte.

Remarque 2.1.4. Le choix du pas de maillage h détermine le nombre N + 2 de points de discrétisation. Ainsi, si h est très petit, le nombre N + 2 de points considérés est très grand et l'approximation U_h est une meilleure représentation de la solution u. Dans la pratique, un pas de maillage h trop petit peut entraîner un coût de calcul prohibitif et il est donc nécessaire de déterminer un pas de maillage suffisamment petit pour obtenir une bonne approximation sans que le temps de calcul soit trop long.



FIGURE 2.1 – À gauche, approximation de la solution de (2.1.1), avec f = 2, pour N + 2 points de discrétisation avec N = 2, 7 et 20. À droite, la solution exacte.

Il reste certains points à vérifier :

- 1. que le système matriciel $A_h U_h = b_h$ admet une solution unique,
- 2. que U_h est continue par rapport aux données b_h (autrement dit qu'une légère variation de b_h n'entraîne qu'une légère variation de U_h),
- 3. que, pour $h \to 0$, le vecteur U_h solution de $A_h U_h = b_h$ converge bien vers $(u(x_1), \ldots, u(x_N))^T$ (autrement dit, que c'est effectivement une bonne approximation de u).

Remarque 2.1.5. Les deux premiers points forment une version discrète de la notion de problème bien posé d'Hadamard. Le premier point est logique du fait que l'on cherche à déterminer U_h , d'où la nécessité d'existence. L'unicité est utile pour obtenir le U_h cherché (et non une autre solution). Le deuxième point est particulièrement important car en pratique le vecteur b_h est obtenu en calculant des valeurs approchées des $f(x_i)$, ainsi on commet une légère erreur sur b_h . La continuité de U_h par rapport à b_h permet de s'assurer que la légère erreur commise sur b_h n'entraîne qu'une légère erreur sur U_h .

2.2 Justification des calculs : consistance, stabilité et convergence

Pour justifier l'approche formelle de la section précédente, on introduit trois notions que l'on définira rigoureusement plus loin :

- 1. consistance, signifie que le système matriciel $A_h U_h = b_h$ est une approximation de l'e.d.p. (2.1.1),
- 2. stabilité, signifie que la solution U_h est continue par rapport au données b_h ,
- 3. convergence, signifie que U_h converge (dans un sens à préciser) lorsque $h \to 0$ vers la solution u de (2.1.1).

On fera appel au résultat suivant :

Proposition 2.2.1 (Principe du maximum discret). Soit $b \in \mathbb{R}^N$ tel que $b_i \ge 0$ pour tout i = 1, ..., N. Si $U \in \mathbb{R}^N$ vérifie $A_h U = b$, où A_h est donnée par (2.1.7), alors $U_i \ge 0$ pour tout i = 1, ..., N.

Démonstration. Soit k le plus petit entier tel que

$$U_k = \min_{i=1,\dots,N} U_i.$$

On suppose $U_k < 0$. Trois cas sont possibles : k = 1, 1 < k < N et k = N.

Pour k = 1, $(A_h U)_1 = b_1 \ge 0$ entraîne

$$-\frac{U_2 - 2\,U_1}{h^2} \ge 0,$$

d'où

$$U_2 \le 2 U_1.$$

Or $U_1 = \min U_i \leq U_2$ donc $2U_1 \leq U_2 + U_1 < U_2$ car $U_1 < 0$. Finalement $U_2 < U_2$: impossible. Pour k = N on procède de la même manière. Pour 1 < k < N, $(AU)_k = b_k \geq 0$ entraîne

$$-\frac{U_{k+1} - 2U_k + U_{k-1}}{h^2} \ge 0,$$

d'où

$$(U_{k+1} - U_k) + (U_{k-1} - U_k) \le 0.$$

Or $U_{k+1} - U_k \ge 0$ et $U_{k-1} - U_k \ge 0$ donc $(U_{k+1} - U_k) + (U_{k-1} - U_k) = 0$. Alors $U_k = U_{k-1} = U_{k+1}$, en réitérant le même raisonnement sur k - 1 on aboutit au cas k = 1 vu précédemment. L'hypothèse $U_k < 0$ est donc fausse, d'où $U_k \ge 0$. Alors $0 \le U_k \le U_i$ pour tout $i = 1, \ldots, N$.

On peut maintenant montrer le résultat suivant :

Proposition 2.2.2. Le système matriciel (2.1.6) admet une unique solution.

Remarque 2.2.3. Le principe du maximum discret est l'ingrédient clé pour la démonstration de la Proposition 2.2.2. C'est une méthode classique pour les schémas numériques des e.d.p. elliptiques de démontrer un principe du maximum discret et d'en déduire l'existence et unicité du problème discret. Démonstration. Supposons que le système (2.1.6) admet deux solutions U et V. Alors W := U - V est solution de $A_h W = 0$ donc $A_h W \ge 0$ et $A_h W \le 0$. Par le principe du maximum discret, on en déduit $W \ge 0$ et $W \le 0$ d'où W = 0 donc U = V. En particulier on en déduit que l'application linéaire associée à la matrice A_h est injective or, en dimension finie, toute application linéaire injective est bijective donc le matrice A_h est inversible et le système (2.1.6) admet une unique solution.

Remarque 2.2.4. Une autre preuve consiste à montrer que A_h est symétrique définie positive. Cette propriété permet de plus d'utiliser la méthode de Cholesky pour la résolution numérique du système $A_h U_h = b_h$.

Définition 2.2.5. (Erreur de consistance) On appelle **erreur de consistance** R d'un schéma numérique la quantité obtenue en remplaçant, dans le schéma numérique, l'inconnue par la solution exacte u. En particulier, pour le schéma (2.1.6), on a $R = A_h U - b_h$ où $U = (u(x_1), \ldots, u(x_N))^T$.

Dans le cadre du schéma numérique (2.1.6), l'erreur de consistance r_i au point x_i (*i*ème coordonnée de R) est donnée par

$$r_i = -\frac{1}{h^2} (u(x_{i+1}) - 2u(x_i) + u(x_{i-1})) - f(x_i).$$
(2.2.1)

Définition 2.2.6. (Ordre d'un schéma) On dit qu'un schéma numérique à N points de discrétisation est d'**ordre** $p \in \mathbb{N}$ s'il existe une constante $C \in \mathbb{R}$ indépendante de la solution exacte telle que l'erreur de consistance vérifie

$$\max_{i=1,\dots,N} |r_i| \le C h^p$$

De plus, on dit que le schéma est consistant si

$$\lim_{h \to 0} \max_{i=1,\dots,N} |r_i| = 0.$$

Dans le cadre du schéma (2.1.6), on a le

Lemme 2.2.7. Si la solution exacte u de (2.1.1) vérifie $u \in C^4(0,1)$ alors le schéma (2.1.6) est consistant d'ordre 2. Précisément, on a

$$|r_i| \le \frac{h^2}{12} \max_{x \in [0,1]} |u^{(4)}(x)|, \quad \forall i = 1, \dots, N.$$
 (2.2.2)

Démonstration. Puisque $u \in C^4(0,1)$ on a le développement de Taylor à l'ordre 4 :

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

où $\xi_i \in [0, 1]$. Le dernier terme $\frac{h^4}{24} u^{(4)}(\xi_i)$ est le reste de la formule de Taylor. De même, on a

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

où $\nu_i \in [0, 1]$. On en déduit

$$r_{i} = -\frac{1}{h^{2}}(u(x_{i+1}) - 2u(x_{i}) + u(x_{i-1})) - f(x_{i})$$

$$= -\frac{1}{h^{2}}(h^{2} u''(x_{i}) + \frac{h^{4}}{24}u^{(4)}(\xi_{i}) + \frac{h^{4}}{24}u^{(4)}(\nu_{i})) - f(x_{i})$$

$$= -u''(x_{i}) - f(x_{i}) + \frac{h^{2}}{24}(u^{(4)}(\xi_{i}) + u^{(4)}(\nu_{i})).$$

Or, comme u est la solution exacte de (2.1.1) on a $-u''(x_i) = f(x_i)$ donc

$$r_i = \frac{h^2}{24} (u^{(4)}(\xi_i) + u^{(4)}(\nu_i))$$

On en déduit (2.2.2).

Remarque 2.2.8. On peut remarquer que si, de plus, u vérifie $u^{(4)} = 0$ alors $r_i = 0$ donc $u_i = u(x_i)$.

Définition 2.2.9. Un schéma numérique est dit **stable** pour la norme $|| \cdot ||$ si sa solution (quand elle existe) est continue pour la norme $|| \cdot ||$ par rapport aux données. En particulier, pour le schéma $A_h U_h = b_h$, cela signifie qu'il existe une constante C > 0 indépendante de h et b_h telle que $||U_h|| \leq C ||b_h||$.

Remarque 2.2.10. On rappelle que pour toute norme vectorielle $\|\cdot\|$ sur \mathbb{R}^N , on peut définir une norme matricielle associée (norme subordonnée), encore notée $\|\cdot\|$, définie sur $\mathbb{R}^{N \times N}$ par

$$\forall A \in \mathbb{R}^{N \times N}, \quad ||A|| := \max\{||AU||, ||U|| = 1\}.$$

En particulier pour la norme vectorielle $||\cdot||_\infty$ définie par $||U||_\infty=\max\{|U_i|\mid i=1,\ldots,N\},$ on a

$$\forall A := (a_{ij})_{1 \le i,j \le N} \in \mathbb{R}^{N \times N}, \quad ||A||_{\infty} = \max_{1 \le i \le N} \sum_{j=1}^{N} |a_{ij}|$$

Proposition 2.2.11. Le schéma numérique (2.1.6) est stable pour la norme $|| \cdot ||_{\infty}$. En particulier, on a $||A_h^{-1}||_{\infty} \leq 1/8$.

Démonstration. Soit U_h la solution de (2.1.6) alors on a $U_h = A_h^{-1} b_h$ et donc

$$||U_h||_{\infty} = ||A_h^{-1}b_h||_{\infty} \le ||A_h^{-1}||_{\infty} ||b_h||_{\infty}.$$

Il suffit donc de montrer qu'il existe une constante C indépendante de h telle que $||A_h^{-1}||_{\infty} \leq C$. On note $e = (1, \ldots, 1)^T \in \mathbb{R}^N$ et $A^{-1} = (b_{ij}) \in \mathbb{R}^{N \times N}$, on a

$$||A^{-1}e||_{\infty} = \max_{1 \le i \le N} |(Be)_i| = \max_{1 \le i \le N} |\sum_{j=1}^N b_{ij}|.$$

Soit $v \in \mathbb{R}^N$ avec $v_i \ge 0$ pour tout *i*. Comme A_h est inversible, il existe un unique $w \in \mathbb{R}^N$ tel que $v = A_h w$. Par le principe du maximum discret, on a $w_i \ge 0$ et donc $(A_h^{-1}v)_i \ge 0$.

Si on choisit $v = e^j$ le *j*-ème vecteur de la base canonique de \mathbb{R}^N on a $(A_h^{-1}v)_i = (A_h^{-1})_{ij}$ d'où $(A_h^{-1})_{ij} \ge 0$. On en déduit

$$||A_h^{-1}e||_{\infty} = \max_{1 \le i \le N} |\sum_{j=1}^N b_{ij}| = \max_{1 \le i \le N} \sum_{j=1}^N b_{ij} = ||A_h^{-1}||_{\infty}.$$

On pose $d = A_h^{-1}e$. Alors $e = A_h d$, autrement dit d est la solution du schéma numérique associé à l'équation (2.1.1) avec f = 1, à savoir

$$\begin{cases} -w''(x) = 1 & \text{dans }]0,1[, \\ w(0) = w(1) = 0. \end{cases}$$

Or w''(x) = -1 entraîne $w(x) = -\frac{x^2}{2} + \alpha x + \beta$ et w(0) = w(1) = 0 donne $\beta = 0$ et $\alpha = \frac{1}{2}$. Donc $w(x) = -\frac{x}{2}(x-1)$. En particulier, on a $w^{(4)}(x) = 0$. D'après la Remarque 2.2.8, on en déduit $w(x_i) = d_i$ d'où $d_i = -\frac{x_i}{2}(x_i - 1)$. Alors, on obtient

$$||A_h^{-1}||_{\infty} = ||A_h^{-1}e||_{\infty} = ||d||_{\infty} = \max_{i=1,\dots,N} |d_i| = \frac{1}{2} \max_{i=1,\dots,N} |x_i(x_i-1)| \le \frac{1}{2} \max_{x \in [0,1]} |x(x-1)| = \frac{1}{8}.$$

Remarque 2.2.12. le schéma (2.1.6) est dit inconditionnellement stable car il n'est pas nécessaire de fixer une condition sur le pas h pour avoir la stabilité.

Définition 2.2.13. Pour un schéma numérique $A_h U_h = b_h$ de l'équation (2.1.1), on appelle **erreur de discrétisation** (ou de convergence) le vecteur $e \in \mathbb{R}^N$ dont les coefficients sont

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

où u est la solution exacte de (2.1.1). De plus, on dit que le schéma numérique converge en norme $||\cdot||$ si l'erreur de discrétisation tend vers 0 en norme $||\cdot||$ lorsque le pas h tend vers 0.

Exemple 2.2.14. Sur la figure 2.2 on a tracé $||e||_{\infty}$ suivant le nombre de points de discrétisations N + 2 (ici N + 2 varie de 2 à 50) pour $f(x) := \pi \sin(\pi x)$.

Théorème 2.2.15. Soit u la solution exacte de (2.1.1) et U_h la solution du schéma numérique (2.1.6). On suppose $u \in C^4(0, 1)$. Alors, l'erreur de discrétisation vérifie

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

Donc le schéma numérique (2.1.6) converge en norme $|| \cdot ||_{\infty}$ et est d'ordre 2.

Démonstration. On pose $U := (u(x_1), \ldots, u(x_N))^T$. Alors, $||e||_{\infty} = ||U - U_h||_{\infty}$. Par définition de l'erreur de consistance R, on a

$$R = A_h U - b_h = A_h (U - U_h) \quad \text{car} \quad A_h U_h = b_h.$$

Donc $U - U_h = A_h^{-1} R$. Alors, d'après le Lemme 2.2.7 et la Proposition 2.2.11, on obtient

$$|e||_{\infty} = ||A_h^{-1}R||_{\infty} \le ||A_h^{-1}||_{\infty} ||R||_{\infty} = \frac{h^2}{96} ||u^{(4)}||_{\infty}$$



FIGURE 2.2 – Graphe de $||e||_{\infty}$ en fonction de N pour $f(x) = \pi \sin(\pi x)$.

Remarque 2.2.16.

- 1. La stratégie employée est classique : on étudie la consistance puis la stabilité du schéma et cela permet d'en déduire la convergence.
- 2. L'ordre de convergence du schéma est important pour comparer différents schémas. Le pas h étant petit (donc inférieur à 1), la quantité h^p est d'autant plus petite que p est grand. Ainsi l'ordre de convergence d'un schéma détermine sa vitesse de convergence.
- 3. La constante de stabilité permet de déterminer l'erreur commise en approchant le second membre.

En effet, supposons que l'on ne puisse pas calculer explicitement $f(x_i)$ mais en obtenir une valeur approchée. Alors on obtient un second membre

$$(c_h)_i = f(x_i) + \varepsilon_i,$$

où ε_i est l'erreur commise en approchant $f(x_i)$ et $\varepsilon := (\varepsilon_1, \ldots, \varepsilon_N)^T$. On va alors résoudre le problème $A_h \tilde{U}_h = c_h$ au lieu de $A_h U_h = b_h$. L'erreur commise en calculant \tilde{U}_h est donc

$$||\tilde{U}_h - U_h||_{\infty} = ||A_h^{-1}(c_h - b_h)||_{\infty} = ||A_h^{-1}\varepsilon||_{\infty} \le ||A_h^{-1}||_{\infty} ||\varepsilon||_{\infty}$$

Ainsi, connaissant une borne sur $||A_h^{-1}||_{\infty}$ et l'erreur commise sur les données $||\varepsilon||_{\infty}$, on détermine l'erreur commise dans le calcul de \tilde{U}_h .

2.3 Problèmes elliptiques en dimension supérieure à 1

Le principe est exactement le même que celui de la dimension 1, la seule différence réside dans l'écriture. On va étudier seulement le cas de la dimension 2, le cas de dimensions supérieures étant complètement analogue. On cherche à résoudre numériquement le problème

$$\begin{cases} -\Delta u = f \quad \text{dans } \Omega =]0, 1[^2, \\ u = 0 \quad \text{sur } \partial \Omega, \end{cases}$$
(2.3.1)

où $u = u(x, y), \ \Delta = \partial_x^2 + \partial_y^2$ et $\partial \Omega$ est le bord de Ω .

On commence par définir un maillage de Ω . On pose

 $x_i := ih$ et $y_j := jh$,

où $0\leq i,j\leq N+1,\,h:=1/(N+1)$ et $N\in\mathbb{N}.$

Remarque 2.3.1. La méthode des différences finies nécessite de considérer un domaine Ω rectangulaire du fait du type de maillage considéré (pour des extensions à des domaines non rectangulaires voir, par exemple, [3] page 242).

On va déterminer $u_{i,j}$ qui approche $u(x_i, y_j)$. Par le développement de Taylor, on a

$$\partial_x^2 u(x_i, y_j) = \frac{u(x_{i+1}, y_j) - 2u(x_i, y_j) + u(x_{i-1}, y_j)}{h^2} + O(h^3).$$

 et

$$\partial_y^2 u(x_i, y_j) = \frac{u(x_i, y_{j+1}) - 2u(x_i, y_j) + u(x_i, y_{j-1})}{h^2} + O(h^3)$$

Un schéma numérique possible est alors de considérer l'approximation suivante de $\Delta u(x_i, x_j)$

$$\Delta_h u_{i,j} = \frac{-4u_{i,j} + u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}}{h^2}$$

Avec cette notation, le problème discrétisé est : trouver $u_{i,j}$ tels que

$$\begin{cases} -\Delta_h u_{i,j} = f(x_i, y_j) & \text{pour } 1 \le i, j \le N, \\ u_{0,j} = u_{N+1,j} = u_{i,0} = u_{i,N+1} = 0 & \text{pour } 1 \le i, j \le N, \end{cases}$$
(2.3.2)

Pour écrire (2.3.2) sous forme matricielle, on pose

$$U_h = (u_{11}, \dots, u_{1N}, u_{21}, \dots, u_{2N}, \dots, u_{NN})^T.$$

Alors le problème (2.3.2) s'écrit

$$A_h U_h = b_h,$$

où $A_h \in \mathbb{R}^{N^2 \times N^2}$ et $b_h \in \mathbb{R}^{N^2}$ sont donnés par

$$A_{h} = -\frac{1}{h^{2}} \begin{pmatrix} B & C & 0 & \dots & 0 \\ C & B & C & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & C \\ 0 & \dots & \dots & C & B \end{pmatrix},$$

 et

$$b_h = (f(x_1, y_1), \dots, f(x_1, y_N), f(x_2, y_1), \dots, f(x_N, y_N))^T$$

 avec

$$B = \begin{pmatrix} -4 & 1 & 0 & \dots & 0 \\ 1 & -4 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 1 & -4 \end{pmatrix} \in \mathbb{R}^{N \times N} \text{ et } C = I_N \in \mathbb{R}^{N \times N}$$

Tout les résultats de consistance, stabilité et convergence du cas de la dimension 1 s'adaptent sans modification majeur.

Exemple 2.3.2. Soit $u(x, y) := -x(x-1)y(y-1)\exp(xy)$ et $f = -\Delta u$, de sorte que u est solution de (2.3.1). Sur la figure 2.3, on a tracé les isovaleurs de l'approximation de u donnée par (2.3.2) pour un pas d'espace h = 0.02 que l'on compare à la solution exacte u.



FIGURE 2.3 – Comparaison entre solution approchée et solution exacte pour le problème de Dirichlet en dimension 2

Remarque 2.3.3. On peut choisir un pas h en x et k en y différents. Dans ce cas $A_h \in \mathbb{R}^{NM \times NM}$, où h := 1/(N+1) et k := 1/(M+1).

Chapitre 3 Problèmes hyperboliques

Dans ce chapitre, on étudie la résolution numérique par différences finies de l'équation d'advection

$$\begin{cases} \partial_t u(t,x) + a(t,x)\partial_x u(t,x) = 0, & (t,x) \in \mathbb{R}^+_* \times \mathbb{R}, \\ u(t=0,x) = u_0(x), & x \in \mathbb{R}, \end{cases}$$
(3.0.1)

où la vitesse de transport a et la donnée initiale u_0 sont données.

3.1 Définitions et solution explicite

Définition 3.1.1. On appelle solution classique de (3.0.1) toute fonction $u \in C^1(\mathbb{R}^*_+ \times \mathbb{R})$ qui vérifie (3.0.1).

Définition 3.1.2. On appelle **courbe caractéristique** de (3.0.1) la solution X = X(t, y) de

$$\begin{cases} \partial_t X(t,y) = a\left(t, X(t,y)\right), & (t,y) \in \mathbb{R}^*_+ \times \mathbb{R}, \\ X(t=0,y) = y, & y \in \mathbb{R}. \end{cases}$$
(3.1.1)

Dans la suite, pour simplifier, on suppose a constant. La solution de (3.1.1) est alors donnée par X(t, y) = at + y.

Proposition 3.1.3. On suppose $u_0 \in C^1(\mathbb{R})$. Alors, les solutions classiques de (3.0.1) sont constantes (par rapport à la variable t) sur les courbes caractéristiques, i.e. si u est solution classique de (3.0.1), on a

$$u(t, X(t, y)) = u_0(X(0, y)) = u_0(y), \quad \forall (t, y) \in \mathbb{R}^+_* \times \mathbb{R}.$$
 (3.1.2)

Réciproquement, si u vérifie (3.1.2) alors u est solution classique de (3.0.1).

On sait que X(t, y) = at + y. Alors, sous les conditions de la Proposition 3.1.3, si u est solution classique de (3.0.1), on a $u(t, at + y) = u_0(y)$. On en déduit $u(t, y) = u_0(y - at)$. Ainsi, on obtient le résultat suivant :

Corollaire 3.1.4. Si $u_0 \in C^1(\mathbb{R})$, il existe une unique solution classique u de (3.0.1) donnée par

 $u(t,y) = u_0(y - at) \quad pour(t,x) \in \mathbb{R}^*_+ \times \mathbb{R}.$

Remarque 3.1.5. La condition $u_0 \in C^1(\mathbb{R})$ est nécessaire pour que $u \in C^1(\mathbb{R}^*_+ \times \mathbb{R})$.

Démonstration de la Proposition 3.1.3. On suppose que u est solution classique de (3.0.1) et on calcule

$$\frac{\partial}{\partial t} \Big[u(t, X(t, y)) \Big] = \partial_t u(t, X(t, y)) + \partial_t X(t, y) \partial_x u(t, X(t, y)) \\ = \partial_t u(t, X(t, y)) + a \, \partial_x u(t, X(t, y)) = 0,$$

d'où

$$u(t, X(t, y)) = u(0, X(0, y)) = u_0(y).$$

Réciproquement, soit u telle que $u(t, X(t, y)) = u_0(y)$. Alors $\partial_t[u(t, X(t, y))] = 0$, d'où (d'après le calcul ci-dessus)

$$\partial_t u(t, X(t, y)) + \partial_t X(t, y) \partial_x u(t, X(t, y)) = 0,$$

ce qui donne le résultat puisque la droite X(t, y) est une bijection de \mathbb{R} .

3.2 Méthode numérique : approche formelle

Comme dans le cas des problèmes elliptiques, on commence par discrétiser l'espace de définition de l'e.d.p. (3.0.1). On note Δx et Δt les pas d'espace et de temps, supposés positifs et petits. On se place sur un intervalle de temps borné [0, T], où $T = N\Delta t$ avec $N \in \mathbb{N}$. On pose

$$\begin{cases} x_j = j\Delta x, \quad j \in \mathbb{Z}, \\ t_n = n\Delta t, \quad n \in \{0, \dots, N\}. \end{cases}$$

Alors, on cherche une approximation u_i^n de $u(t_n, x_j)$. Les développements de Taylor donnent

$$u(t_{n+1}, x_j) = u(t_n, x_j) + \Delta t \,\partial_t u(t_n, x_j) + O(\Delta t^2),$$

 et

$$u(t_n, x_{j+1}) = u(t_n, x_j) + \Delta x \,\partial_x u(t_n, x_j) + O(\Delta x^2),$$

$$u(t_n, x_{j-1}) = u(t_n, x_j) - \Delta x \,\partial_x u(t_n, x_j) + O(\Delta x^2).$$

Donc on a

$$\partial_t u(t_n, x_j) = \frac{u(t_{n+1}, x_j) - u(t_n, x_j)}{\Delta t} + O(\Delta t),$$

$$\partial_x u(t_n, x_j) = \frac{u(t_n, x_{j+1}) - u(t_n, x_{j-1})}{2\Delta x} + O(\Delta x).$$

Un premier choix simple de schéma numérique pour l'e.d.p. (3.0.1) est alors

$$\begin{cases} \frac{u_j^{n+1} - u_j^n}{\Delta t} + a \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} = 0, \\ u_j^0 = u_0(x_j). \end{cases}$$
(3.2.1)

Ce schéma peut encore s'écrire sous la forme (en laissant de coté la condition initiale)

$$u_j^{n+1} = u_j^n + a \, \frac{\Delta t}{2\Delta x} \, (u_{j+1}^n - u_{j-1}^n). \tag{3.2.2}$$

Il est dit **explicite** en temps (le terme en n + 1 est donné en fonction de n) et **centré** en espace.

3.3 Consistance, stabilité et convergence

Pour u_j^n une approximation de $u(t_n, x_j)$, on note $u^n = (u_j^n)_{j \in \mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}$. Les schémas que l'on va considérer sont les schémas qui s'écrivent sous la forme

$$u^{n+1} = E_{\Delta} u^n, \quad \forall \ n \in \{0, \dots, N\},$$
(3.3.1)

où E_{Δ} est une application linéaire sur $\mathbb{R}^{\mathbb{Z}}$. On parle de schéma à deux niveaux.

Remarque 3.3.1. L'approximation u^n de u à l'instant t_n est ici un élément de $\mathbb{R}^{\mathbb{Z}}$ parce qu'aucune condition aux limites n'a été spécifiée. Dans la pratique, lorsque l'on ajoute une condition aux limites, u^n est un vecteur et E_{Δ} une matrice.

Définition 3.3.2. Pour un schéma numérique du type $u^{n+1} = E_{\Delta}u^n$ de l'e.d.p. (3.0.1), on définit l'**erreur de consistance** r^n au temps t_n avec $n \ge 1$ par

$$r^n := \frac{1}{\Delta t} (\tilde{u}^n - E_\Delta \tilde{u}^{n-1}),$$

où $\tilde{u}^n := (u(t_n, x_j))_{j \in \mathbb{Z}}$ avec u la solution exacte de (3.0.1).

Remarque 3.3.3. Par exemple, pour le schéma (3.2.1), à partir de l'expression (3.2.2) on a (au point x_j)

$$r_j^n = \frac{1}{\Delta t} (\tilde{u}_j^n - (E_\Delta \tilde{u}^{n-1})_j) = \frac{u(t_n, x_j) - u(t_{n-1}, x_j)}{\Delta t} + a \frac{u(t_{n-1}, x_{j+1}) - u(t_{n-1}, x_{j-1})}{2\Delta x}.$$

Autrement dit, on obtient que r_j^n est l'erreur commise en remplaçant u_j^n par $u(t_n, x_j)$, ce qui est conforme à la définition donnée dans le chapitre précédent.

So t $n \in \{0, \ldots, N\}$. Avec ces notations, on a

$$\tilde{u}^{n} - u^{n} = \tilde{u}^{n} - E_{\Delta} u^{n-1} = E_{\Delta} (\tilde{u}^{n-1} - u^{n-1}) + \tilde{u}^{n} - E_{\Delta} \tilde{u}^{n-1}$$

= $E_{\Delta} (\tilde{u}^{n-1} - u^{n-1}) + r^{n} \Delta t.$

Si n = 0, puisque $\tilde{u}^0 = u^0$, on a

$$r^0 = \frac{\tilde{u}^0 - u^0}{\Delta t} = 0,$$

alors on obtient

$$\tilde{u}^n - u^n = \Delta t \sum_{k=1}^n E_{\Delta}^{n-k} r^k.$$

On en déduit pour toute norme $|| \cdot ||$ sur $\mathbb{R}^{\mathbb{Z}}$:

$$\begin{aligned} ||\tilde{u}^n - u^n|| &\leq \Delta t \sum_{k=1}^n ||E_{\Delta}^{n-k} r^k|| \leq n \Delta t \max_{1 \leq k \leq n} ||E_{\Delta}^{n-k} r^k|| \\ &\leq T \max_{1 \leq k \leq n} ||E_{\Delta}^{n-k} r^k||. \end{aligned}$$

Si le schéma est stable pour la norme $|| \cdot ||$, alors il existe une constante c indépendante de n, Δt et Δx telle que pour toute condition initiale u^0 , on a $||E_{\Delta}^n u^0|| \leq c ||u^0||$. En particulier, on obtient

$$||E_{\Delta}^{n-k}r^k|| \le c \,||r^k||.$$

Si le schéma est consistant d'ordre p pour la norme $|| \cdot ||$ alors on a $||r^n|| \le c |\Delta x|^p$ où c est une constante indépendante de n, Δt et Δx .

Finalement, on obtient que, si le schéma est consistant d'ordre p et stable pour la norme $|| \cdot ||$, alors

$$||\tilde{u}^n - u^n|| \le c T \, |\Delta x|^p.$$

On a donc le résultat suivant :

Théorème 3.3.4 (Théorème de Lax). Si le schéma (3.3.1) est consistant d'ordre p et stable alors il est convergent et l'erreur de discrétisation tend vers 0 comme Δx^p lorsque $\Delta x \to 0$ et $\Delta t \to 0$.

On va voir dans la suite comment déterminer si un schéma du type (3.3.1) est stable pour la norme $l^2(\mathbb{Z})$.

3.4 Étude de la stabilité par analyse de Fourier

Un schéma a deux niveaux $u^{n+1} = E_{\Delta} u^n$ peut encore s'écrire

$$\sum_{m=-k}^{k} c_m u_{j+m}^{n+1} = \sum_{m=-k}^{k} d_m u_{j+m}^n, \qquad (3.4.1)$$

où $k \in \mathbb{N}$ et $c_m, d_m \in \mathbb{R}$ pour tout $m \in \{-k, \ldots, k\}$. On suppose $u^n \in l^2(\mathbb{Z})$, *i.e.* $\sum_{j \in \mathbb{Z}} |u_j^n|^2 < \infty$.

On définit la fonction $v^n : \mathbb{R} \to \mathbb{R}$ par

$$v^n(x) = u_j^n \quad \text{si } x \in [(j - 1/2)\Delta x, (j + 1/2)\Delta x].$$
 (3.4.2)

Alors, $v^n \in L^2(\mathbb{R})$ et sa transformée de Fourier est définie par

$$\mathcal{F}v^n(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} v^n(x) e^{-ix\xi} dx$$

D'après le théorème de Plancherel, on a

$$||\mathcal{F}v^{n}||_{L^{2}(\mathbb{R})} = ||v^{n}||_{L^{2}(\mathbb{R})} = \left(\int_{\mathbb{R}} |v^{n}|^{2} dx\right)^{\frac{1}{2}} = \left(\sum_{j \in \mathbb{Z}} \Delta x |u_{j}^{n}|^{2}\right)^{\frac{1}{2}} = \sqrt{\Delta x} ||u^{n}||_{l^{2}(\mathbb{Z})},$$

donc étudier la stabilité l^2 de u^n revient à étudier la continuité de l'application $v^0 \mapsto \mathcal{F}v^n$ dans $L^2(\mathbb{R})$. De plus, on a $v^n(x + m\Delta x) = u_{j+m}^n$ pour $x \in [(j - \frac{1}{2})\Delta x, (j + \frac{1}{2})\Delta x]$.

Or $\mathcal{F}[v^n(x+m\Delta x)](\xi) = e^{im\Delta x\xi} \mathcal{F}v^n(\xi)$. Ainsi, par transformée de Fourier, le schéma (3.4.1) s'écrit

$$\sum_{m=-k}^{k} c_m \mathcal{F} v^{n+1}(\xi) e^{im\Delta x\xi} = \sum_{m=-k}^{k} d_m \mathcal{F} v^n(\xi) e^{im\Delta x\xi}.$$

On en déduit

$$\mathcal{F}v^{n+1}(\xi) = g(\xi; \Delta x, \Delta t)\mathcal{F}v^n(\xi),$$

où le **coefficient d'amplification** $g(\xi; \Delta x, \Delta t)$ est donné par

$$g(\xi; \Delta x, \Delta t) = \frac{\sum_{m=-k}^{k} d_m e^{im\Delta x\xi}}{\sum_{m=-k}^{k} c_m e^{im\Delta x\xi}}.$$

On a alors le résultat suivant sur la stabilité l^2 des schémas à deux niveaux :

Théorème 3.4.1 (Critère de Von Neumann). Le schéma à deux niveaux (3.4.1) est stable en norme $l^2(\mathbb{Z})$ si et seulement si

$$\sup_{\xi \in \mathbb{R}} |g(\xi; \Delta x, \Delta t)| \le 1 + M\Delta t, \tag{3.4.3}$$

où la constante M est indépendante de Δt et Δx .

Démonstration. On a

$$||u^n||_{l^2(\mathbb{Z})} = \frac{1}{\sqrt{\Delta x}} ||\mathcal{F}v^n||_{L^2(\mathbb{R})}$$

 et

$$\begin{aligned} ||\mathcal{F}v^{n}(\xi)||_{L^{2}(\mathbb{R})} &\leq \sup_{\xi \in \mathbb{R}} |g(\xi; \Delta x, \Delta t)| \, ||\mathcal{F}v^{n-1}(\xi)||_{L^{2}(\mathbb{R})} \\ &\leq \left(\sup_{\xi \in \mathbb{R}} |g(\xi; \Delta x, \Delta t)| \right)^{n} \, ||\mathcal{F}v^{0}(\xi)||_{L^{2}(\mathbb{R})}, \end{aligned}$$

d'où

$$||u^{n}||_{l^{2}(\mathbb{Z})} \leq \left(\sup_{\xi \in \mathbb{R}} |g(\xi; \Delta x, \Delta t)|\right)^{n} ||u^{0}||_{l^{2}(\mathbb{Z})}.$$

Alors, si (3.4.3) est vérifiée, on obtient

$$||u^{n}||_{l^{2}(\mathbb{Z})} \leq (1 + M\Delta t)^{n} ||u^{0}||_{l^{2}(\mathbb{Z})}, \quad \forall \ n \in \{0, \dots, N\}.$$

Or, puisque $1 + M\Delta t > 1$, on a

 $(1 + M\Delta t)^n \le (1 + M\Delta t)^N \le e^{MN\Delta t} = e^{MT}.$

Donc

$$|u^{n}||_{l^{2}(\mathbb{Z})} \leq e^{MT} ||u^{0}||_{l^{2}(\mathbb{Z})}, \quad \forall \ n \in \{0, \dots, N\},$$

la constante e^{MT} étant indépendante de n, Δt et Δx , le schéma est stable.

Réciproquement, si le schéma est stable alors on a

$$||u^n||_{l^2(\mathbb{Z})} \le c ||u^0||_{l^2(\mathbb{Z})}, \quad \forall \ n \in \{0, \dots, N\},$$

d'où

$$\sup_{u^0 \in l^2(\mathbb{Z}) \setminus \{0\}} \frac{||u^n||_{l^2(\mathbb{Z})}}{||u^0||_{l^2(\mathbb{Z})}} \le c.$$

On admettra le résultat suivant

Lemme 3.4.2. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue et bornée et $\mathcal{P} \in \mathcal{L}(L^2(\mathbb{R}))$ l'opérateur défini par $\mathcal{F}(\mathcal{P}u)(\xi) = f(\xi)\mathcal{F}u(\xi)$. Alors $||\mathcal{P}||_{\mathcal{L}(L^2(\mathbb{R}))} = ||f||_{\infty}$, i.e.

$$\sup_{u \in L^2(\mathbb{R}) \setminus \{0\}} \frac{||\mathcal{P}u||_{L^2(\mathbb{R})}}{||u||_{L^2(\mathbb{R})}} = ||f||_{\infty}.$$

On applique le lemme pour l'opérateur \mathcal{P} défini par $\mathcal{F}(\mathcal{P}v^0) = \mathcal{F}v^n = g(\xi; \Delta x, \Delta t)^n \mathcal{F}v^0$. On en déduit

$$\sup_{\xi \in \mathbb{R}} |g(\xi; \Delta x, \Delta t)|^n = \sup_{v^0 \in L^2(\mathbb{R}) \setminus \{0\}} \frac{||v^n||_{L^2(\mathbb{R})}}{||v^0||_{L^2(\mathbb{R})}} = \sup_{u^0 \in l^2(\mathbb{Z}) \setminus \{0\}} \frac{||u^n||_{l^2(\mathbb{Z})}}{||u^0||_{l^2(\mathbb{Z})}} \le c.$$

Alors, $|g(\xi; \Delta x, \Delta t)|^n \leq c$ et donc, pour $n = N = \frac{T}{\Delta t}$, on a

$$|g(\xi; \Delta x, \Delta t)| \le c^{\frac{\Delta t}{T}} \le 1 + M\Delta t,$$

pour tout $\xi \in \mathbb{R}$.

Remarque 3.4.3. Le critère de Von Neumann est une condition nécessaire et suffisante donc celle-ci permet de montrer la stabilité d'un schéma numérique mais aussi son instabilité.

Proposition 3.4.4. Le schéma à deux niveaux (3.2.2) est inconditionnellement instable en norme $l^2(\mathbb{Z})$ (et donc, à fortiori, en norme $l^{\infty}(\mathbb{Z})$).

Démonstration. On a

$$v^{n+1}(x) = v^n(x) - a \frac{\Delta t}{2\Delta x} (v^n(x + \Delta x) - v^n(x - \Delta x)),$$

ce qui entraîne

$$\mathcal{F}v^{n+1}(\xi) = (1 - \frac{\lambda}{2}(e^{i\xi\Delta x} - e^{-i\xi\Delta x}))\mathcal{F}v^n(\xi) = (1 - i\lambda\sin(\xi\Delta x))\mathcal{F}v^n(\xi).$$

où $\lambda := a \frac{\Delta t}{\Delta x}$. On a donc pour coefficient d'amplification $g(\xi; \Delta t, \Delta x) = 1 - i\lambda \sin(\xi \Delta x)$. On obtient

$$|g(\xi; \Delta t, \Delta x)|^2 = 1 + \lambda^2 \sin^2(\xi \Delta x),$$

donc $\sup_{\xi \in \mathbb{R}} |g(\xi; \Delta t, \Delta x)| > 1$ pour tout $\Delta x, \Delta t > 0$.
3.5 Exemples de schémas

Pour toute la suite, on pose $\lambda := a \frac{\Delta t}{\Delta x}$. On note u la solution exacte de (3.0.1), que l'on supposera de classe C^2 , et on pose $\tilde{u}_j^n := u(t_n, x_j)$.

3.5.1 Le schéma explicite centré

Il est défini par

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

C'est le schéma (3.2.2) vu précédemment. Il est explicite en temps $(u^{n+1} \text{ est donné en fonction de } u^n)$ et centré en espace (dû à l'expression $u_{j+1}^n - u_{j-1}^n$). Ce schéma n'étant pas stable, il est inutilisable en pratique.

3.5.2 Les schémas explicites décentrés

Il y a le schéma explicite décentré à droite (ou décentré aval)

$$u_j^{n+1} = u_j^n - \lambda (u_{j+1}^n - u_j^n), \text{ pour } a < 0,$$

et le schéma explicite décentré à gauche (ou décentré amont)

$$u_j^{n+1} = u_j^n - \lambda (u_j^n - u_{j-1}^n), \text{ pour } a > 0.$$

<u>Consistance</u>. On va montrer la consistance du schéma explicite décentré à gauche (le cas décentré à droite se traite de la même manière). On a

$$\tilde{u}_j^{n+1} = \tilde{u}_j^n + \Delta t \,\partial_t \tilde{u}_j^n + \frac{\Delta t^2}{2} \,\partial_t^2 \tilde{u}_j^n + O(\Delta t^3),$$

 et

$$\tilde{u}_{j-1}^n = \tilde{u}_j^n - \Delta x \,\partial_x \tilde{u}_j^n + \frac{\Delta x^2}{2} \,\partial_x^2 \tilde{u}_j^n + O(\Delta x^3).$$

L'erreur de consistance au temps t_{n+1} et au point x_j est donc

$$\begin{split} r_{j}^{n+1} &= \frac{1}{\Delta t} \Big(\tilde{u}_{j}^{n} + \Delta t \, \partial_{t} \tilde{u}_{j}^{n} + \frac{\Delta t^{2}}{2} \, \partial_{t}^{2} \tilde{u}_{j}^{n} - \tilde{u}_{j}^{n} + O(\Delta t^{3}) \\ &+ \lambda (\tilde{u}_{j}^{n} - \tilde{u}_{j}^{n} + \Delta x \, \partial_{x} \tilde{u}_{j}^{n} - \frac{\Delta x^{2}}{2} \, \partial_{x}^{2} \tilde{u}_{j}^{n}) + O(\Delta x^{3}) \Big) \\ &= \partial_{t} \tilde{u}_{j}^{n} + \frac{\Delta t}{2} \, \partial_{t}^{2} \tilde{u}_{j}^{n} + O(\Delta t^{2}) + \frac{\lambda \Delta x}{\Delta t} \left(\partial_{x} \tilde{u}_{j}^{n} - \frac{\Delta x}{2} \, \partial_{x}^{2} \tilde{u}_{j}^{n} \right) + O\left(\frac{\Delta x^{3}}{\Delta t} \right) \\ &= \partial_{t} \tilde{u}_{j}^{n} + \frac{\Delta t}{2} \, \partial_{t}^{2} \tilde{u}_{j}^{n} + a \, \partial_{x} \tilde{u}_{j}^{n} - \frac{a \Delta x}{2} \, \partial_{x}^{2} \tilde{u}_{j}^{n} + O\left(\Delta t^{2} + \frac{\Delta x^{3}}{\Delta t} \right) \\ &= \frac{a^{2} \Delta t}{2} \, \partial_{x}^{2} \tilde{u}_{j}^{n} - \frac{a \Delta x}{2} \, \partial_{x}^{2} \tilde{u}_{j}^{n} + O\left(\Delta t^{2} + \frac{\Delta x^{3}}{\Delta t} \right), \end{split}$$

car $\partial_t u + a \partial_x u = 0$ et $\partial_t^2 u - a^2 \partial_x^2 u = 0$ (s'obtient en dérivant $\partial_t u + a \partial_x u = 0$ en temps d'une part et en espace d'autre part). Alors, en supposant le rapport $\frac{\Delta x}{\Delta t}$ constant, on obtient

$$r^{n+1} = \frac{a\Delta x}{2}(\lambda - 1)\partial_x^2 \tilde{u}_j^n + O(\Delta t^2 + \Delta x^2)$$

Ainsi, le schéma est consistant d'ordre 2 en temps et 1 en espace si $\lambda \neq 1$ et d'ordre au moins 2 en temps et en espace si $\lambda = 1$.

<u>Stabilité</u>. Étudions maintenant la stabilité. Par transformée de Fourier, le schéma explicite décentré à gauche s'écrit

$$\mathcal{F}v^{n+1}(\xi) = (1 - \lambda(1 - e^{-i\Delta x\xi}))\mathcal{F}v^n(\xi).$$

Le coefficient d'amplification est donc $g(\xi; \Delta x, \Delta t) = (1 - \lambda(1 - \cos(\xi \Delta x)) - i\lambda \sin(\xi \Delta x)).$ On a

$$|g(\xi; \Delta x, \Delta t)|^2 = (1 - \lambda(1 - \cos(\xi \Delta x)))^2 + \lambda^2 \sin(\xi \Delta x)$$

= 1 + 2\lambda + 2\lambda^2 - 2\lambda(\lambda - 1)\cos(\xi \Delta x))
= 1 - 2\lambda(1 - \lambda)(1 - \cos(\xi \Delta x)).

Comme $1 - \cos(\xi \Delta x) \ge 0$, en étudiant le signe de $\lambda(1 - \lambda)$ et le fait que $\lambda > 0$ (car a > 0), on obtient

$$\max_{\xi \in \mathbb{R}} |g(\xi; \Delta x, \Delta t)| = \begin{cases} 1 & \text{si } \lambda \le 1, \\ \sqrt{1 - 2\lambda(1 - \lambda)} & \text{si } \lambda > 1. \end{cases}$$

On en déduit que le schéma explicite décentré à gauche est stable en norme $l^2(\mathbb{Z})$ sous la condition dite **C.F.L.** (Courant-Friedrichs-Levy) :

$$\lambda = a \frac{\Delta t}{\Delta x} \le 1.$$

3.5.3 Le schéma de Lax

C'est le schéma explicite donné par

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

<u>Consistance</u>. On a

$$\tilde{u}_j^{n+1} = \tilde{u}_j^n + \Delta t \,\partial_t \tilde{u}_j^n + \frac{\Delta t^2}{2} \,\partial_t^2 \tilde{u}_j^n + O(\Delta t^3),$$

pour la variable en temps, et

$$\begin{split} \tilde{u}_{j-1}^n + \tilde{u}_{j+1}^n &= \tilde{u}_j^n - \Delta x \, \partial_x \tilde{u}_j^n + \frac{\Delta x^2}{2} \, \partial_x^2 \tilde{u}_j^n + \tilde{u}_j^n + \Delta x \, \partial_x \tilde{u}_j^n + \frac{\Delta x^2}{2} \, \partial_x^2 \tilde{u}_j^n + O(\Delta x^3) \\ &= 2\tilde{u}_j^n + \Delta x^2 \, \partial_x^2 \tilde{u}_j^n + O(\Delta x^3), \end{split}$$

avec

$$\begin{split} \tilde{u}_{j+1}^n - \tilde{u}_{j-1}^n &= \tilde{u}_j^n + \Delta x \, \partial_x \tilde{u}_j^n + \frac{\Delta x^2}{2} \, \partial_x^2 \tilde{u}_j^n - \tilde{u}_j^n + \Delta x \, \partial_x \tilde{u}_j^n - \frac{\Delta x^2}{2} \, \partial_x^2 \tilde{u}_j^n + O(\Delta x^3) \\ &= 2\Delta x \, \partial_x \tilde{u}_j^n + O(\Delta x^3), \end{split}$$

pour la variable d'espace. Donc l'erreur de consistance au temps t_{n+1} et au point x_j est

$$\begin{split} r_{n+1}^{j} &= \frac{1}{\Delta t} \left(\tilde{u}_{j}^{n} + \Delta t \,\partial_{t} \tilde{u}_{j}^{n} + \frac{\Delta t^{2}}{2} \,\partial_{t}^{2} \tilde{u}_{j}^{n} + O(\Delta t^{3}) - \tilde{u}_{j}^{n} - \frac{\Delta x^{2}}{2} \,\partial_{x}^{2} \tilde{u}_{j}^{n} + \lambda \Delta x \,\partial_{x} \tilde{u}_{j}^{n} + O(\Delta x^{3}) . \right) \\ &= \partial_{t} \tilde{u}_{j}^{n} + a \,\partial_{x} \tilde{u}_{j}^{n} + \frac{\Delta t}{2} \,\partial_{t}^{2} \tilde{u}_{j}^{n} - \frac{\Delta x^{2}}{2\Delta t} \,\partial_{x}^{2} \tilde{u}_{j}^{n} + O\left(\Delta t^{2} + \frac{\Delta x^{3}}{\Delta t}\right) . \end{split}$$

Or $\partial_t^2 \tilde{u}_j^n = a^2 \partial_x^2 \tilde{u}_j^n$. Ainsi, en supposant le rapport $\frac{\Delta x}{\Delta t}$ constant, on obtient

$$r_{n+1}^{j} = \left(\frac{a^{2}\Delta t}{2} - \frac{\Delta x^{2}}{2\Delta t}\right)\partial_{x}^{2}\tilde{u}_{j}^{n} + O(\Delta t^{2} + \Delta x^{2})$$
$$= \frac{a}{2}\left(a\Delta t - \frac{\Delta x^{2}}{a\Delta t}\right)\partial_{x}^{2}\tilde{u}_{j}^{n} + O(\Delta t^{2} + \Delta x^{2})$$
$$= \frac{a}{2}\left(\lambda - \frac{1}{\lambda}\right)\Delta x\,\partial_{x}^{2}\tilde{u}_{j}^{n} + O(\Delta t^{2} + \Delta x^{2})$$

On en déduit que le schéma est d'ordre 1 en temps et 2 en espace si $\lambda \neq 1$ et d'ordre 2 en temps et en espace si $\lambda = 1$.

<u>Stabilité</u>. Par transformée de Fourier, on obtient

$$\mathcal{F}v^{n+1}(\xi) = (\cos(\xi\Delta x) - i\lambda\sin(\xi\Delta x))\mathcal{F}v^n(\xi).$$

Le coefficient d'amplification est donc $g(\xi; \Delta x, \delta t) = \cos(\xi \Delta x) - i\lambda \sin(\xi \Delta x)$, d'où

$$|g(\xi; \Delta x, \delta t)|^2 = \cos^2(\xi \Delta x) + \lambda^2 \sin^2(\xi \Delta x).$$

On en déduit immédiatement que le schéma de Lax est stable en norme $l^2(\mathbb{Z})$ si et seulement si la condition C.F.L. $\lambda \leq 1$ est vérifiée.

3.5.4 Le schéma implicite centré

C'est le schéma implicite (*i.e.* u^n est donné en fonction de u^{n+1}) donné par

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

Celui-ci peut se réécrire

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

soit encore $u^n = A_{\Delta} u^{n+1}$ où A_{Δ} est un opérateur sur $\mathbb{R}^{\mathbb{Z}}$. Ainsi, en inversant A_{Δ} , on retrouve bien un schéma a deux niveaux. <u>Consistance</u>. On a

$$\begin{aligned} \frac{\tilde{u}_j^{n+1} - \tilde{u}_j^n}{\Delta t} &= \frac{1}{\Delta t} \left(\tilde{u}_j^{n+1} - \left(\tilde{u}_j^{n+1} - \Delta t \,\partial_t \tilde{u}_j^{n+1} + \frac{\Delta t^2}{2} \partial_t^2 \tilde{u}_j^{n+1} \right) + O(\Delta t^3) \right) \\ &= \partial_t \tilde{u}_j^{n+1} - \frac{\Delta t}{2} \partial_t^2 \tilde{u}_j^{n+1} + O(\Delta t^2), \end{aligned}$$

 et

$$\frac{\tilde{u}_{j+1}^{n+1} - \tilde{u}_{j-1}^{n+1}}{2\Delta x} = \partial_x \tilde{u}_j^{n+1} + O(\Delta x^2).$$

Donc l'erreur de consistance au temps t_{n+1} et au point x_j est

$$r_j^{n+1} = -\frac{\Delta t}{2} \partial_t^2 \tilde{u}_j^{n+1} + O(\Delta x^2 + \Delta t^2).$$

Le schéma implicite est donc consistant d'ordre 1 en temps et 2 en espace. <u>Stabilité</u>. Par transformée de Fourier, on obtient

$$\mathcal{F}v^{n}(\xi) = (1 + i\lambda\sin(\xi\Delta x))\mathcal{F}v^{n+1}(\xi),$$

donc le coefficient d'amplification est

$$g(\xi; \Delta x, \Delta t) = \frac{1}{1 + i\lambda\sin(\xi\Delta x)}$$

Alors, on a

$$|g(\xi; \Delta x, \Delta t)|^2 = \frac{1}{1 + \lambda^2 \sin^2(\xi \Delta x)} \le 1$$

Le schéma implicite centré est donc inconditionnellement stable en norme $l^2(\mathbb{Z})$.

3.5.5 Le schéma de Lax-Wendroff

C'est le schéma explicite donné par

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

<u>Consistance</u>. Par les mêmes calcul que précédemment, on obtient facilement que le schéma est consistant d'ordre 2 en temps et en espace sous l'hypothèse $\frac{\Delta x}{\Delta t}$ constant. <u>Stabilité</u>. Par transformée de Fourier, on obtient le coefficient d'amplification $g(\xi; \Delta x, \Delta t) = 1 - i\lambda \sin(\Delta x \xi) + \lambda^2 (\cos(\xi \Delta x) - 1)$. On a alors

$$|g(\xi;\Delta x,\Delta t)| = 1 + \lambda^2 (\lambda^2 - 1)(\xi \cos(\Delta x) - 1)^2.$$

Donc

$$\max_{\xi \in \mathbb{R}} |g(\xi; \Delta x, \Delta t)| = \begin{cases} 1 + 4\lambda^2(\lambda^2 - 1) & \text{si } \lambda > 1\\ 1 & \text{si } \lambda \le 1 \end{cases}$$

D'après le critère de Von-Neumann, le schéma est donc stable en norme $l^2(\mathbb{Z})$ sous la condition C.F.L. $\lambda \leq 1$.

Chapitre 4

Problèmes paraboliques

4.1 Propriétés

Dans ce chapitre, on étudie la résolution numérique par différences finies de l'équation de la chaleur

$$\begin{cases} \partial_t u - \partial_x^2 u = 0, & (t, x) \in]0, T[\times]0, 1[, \\ u(t = 0, x) = u_0(x), & x \in]0, 1[, \\ u(t, 0) = u(t, 1) = 0, & t \in]0, T[, \end{cases}$$

$$(4.1.1)$$

où la condition initiale u_0 est donnée.

On admettra ici le résultat d'existence et unicité suivant (voir par exemple [1]) :

Théorème 4.1.1. Si $u_0 \in C^1(]0, 1[)$, alors il existe une unique solution

 $u \in C^{2}([0, T[\times]0, 1[) \cap C^{1}([0, T] \times [0, 1])),$

de (4.1.1).

Remarque 4.1.2. On peut montrer qu'en réalité, on a $u \in C^{\infty}([0,T] \times [0,1])$ alors que la condition initiale est seulement C^1 . On parle de propriété de lissage de l'équation de la chaleur ou encore de régularisation.

On verra plus loin que pour les schémas numériques de l'équation de la chaleur on peut obtenir (comme dans le cas elliptique) de la stabilité en norme l^{∞} . Ceci est dû au fait que l'équation de la chaleur vérifie un principe du maximum (voir [1]).

Proposition 4.1.3 (Principe du maximum). On note $K := [0, T] \times [0, 1]$. Soient $u_0 \in C^1(]0, 1[)$ et $u \in C^2(]0, T[\times]0, 1[) \cap C^1([0, T] \times [0, 1])$ la solution de (4.1.1). Si $u_0 \ge 0$ dans]0, 1[, alors $u \ge 0$ dans]0, 1[.

Dans la suite, on va étudier quelques schémas numériques pour l'équation de la chaleur. On utilisera les notations suivantes :

- $t_n := n\Delta t$ et $x_j := j\Delta x$ avec $n \in \{0, \dots, N\}$, où $N\Delta t = T$ et $j \in \{0, \dots, J+1\}$ où $(J+1)\Delta x = 1, N, J \in \mathbb{N}$.
- L'approximation de $u(t_n, x_j)$ est notée u_j^n et v^n est la fonction définie par (3.4.2) associée au vecteur $u^n = (u_j^n)_{1 \le j \le J}$.
- Si u est la solution exacte de (4.1.1), on note $\tilde{u}_j^n := u(t_n, x_j)$.

4.2 Exemples de schémas

4.2.1 Le schéma explicite centré

On le définit par

$$\begin{pmatrix}
 u_{j}^{n+1} - u_{j}^{n} \\
 \Delta t - \frac{u_{j+1}^{n} - 2u_{j}^{n} + u_{j-1}^{n}}{\Delta x^{2}} = 0, & 1 \le n \le N, \ 1 \le j \le J, \\
 u_{j}^{0} = u_{0}(x_{j}), & 1 \le j \le J, \\
 u_{0}^{n} = u_{J+1}^{n} = 0, & 0 \le n \le N.$$
(4.2.1)

Remarque 4.2.1. Pour tout les schémas numériques, les conditions initiales et aux limites sont toujours les mêmes. On ne les rappellera plus pour les schémas suivants.

Ce schéma peut encore s'écrire sous la forme

$$u_{j}^{n+1} = u_{j}^{n} + \lambda (u_{j+1}^{n} - 2u_{j}^{n} + u_{j-1}^{n}), \quad \text{où} \quad \lambda := \frac{\Delta t}{\Delta x^{2}}$$

Il s'agit donc d'un schéma à deux niveaux : $u^{n+1} = Au^n$ où $A \in \mathbb{R}^{J \times J}$ est donnée par

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

Contrairement au problème d'advection, A n'est plus un opérateur sur $\mathbb{R}^{\mathbb{Z}}$ mais une matrice car on s'est placé dans l'intervalle borné [0, 1]. En pratique, pour résoudre numériquement l'équation de la chaleur (4.1.1) il suffit de calculer pour $n \in \{1, \ldots, N\}$ les itérées $u^n = A^n u^0$ connaissant $u^0 = (u_0(x_j))_{1 \le j \le J}$.

<u>Consistance</u>. Dans le cas des problèmes elliptiques on a montré que l'on a

$$\frac{1}{\Delta x^2} (\tilde{u}_{j+1}^n - 2\tilde{u}_j^n + \tilde{u}_{j-1}^n) = \partial_x^2 \tilde{u}_j^n + O(\Delta x).$$

L'erreur de consistance au temps t_{n+1} et au point x_i est donc

$$\begin{aligned} r_j^{n+1} &= \partial_t \tilde{u}_j^n + \frac{\Delta t}{2} \, \partial_t^2 \tilde{u}_j^n - \partial_x^2 \tilde{u}_j^n + O(\Delta t^2 + \Delta x^2) \\ &= \frac{\Delta t}{2} \, \partial_t^2 \tilde{u}_j^n + O(\Delta t^2 + \Delta x^2). \end{aligned}$$

Le schéma est donc consistant d'ordre 1 en temps et 2 en espace. Stabilité. On va étudier la stabilité l^{∞} du schéma. On a

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

On pose $M^{(n)} = \max_{1 \le j \le J} u_j^n$. Si $\lambda \le \frac{1}{2}$, $(1 - 2\lambda) \ge 0$ donc (puisque $\lambda > 0$)

$$u_j^{n+1} \le (1-2\lambda)M^{(n)} + \lambda M^{(n)} + \lambda M^{(n)} = M^{(n)}.$$

On en déduit

$$\max_{1 \le j \le J} u_j^{n+1} \le \max_{1 \le j \le J} u_j^n.$$

On obtient donc

$$\max_{1 \le j \le J} u_j^n \le \max_{1 \le j \le J} u_j^0.$$

De même, on a

$$\min_{1 \le j \le J} u_j^n \le \min_{1 \le j \le J} u_j^0,$$

d'où $||u^n||_{\infty} \leq ||u^0||_{\infty}$ (principe du maximum discret). Autrement dit, le schéma explicite centré est stable l^{∞} sous la condition $\lambda \leq \frac{1}{2}$.

4.2.2 Le schéma implicite centré

Il est défini par

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} - \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{\Delta x^2} = 0,$$
(4.2.2)

que l'on peut aussi écrire

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

avec $\lambda := \frac{\Delta t}{\Delta x^2}$ comme précédemment.

Le schéma est consistant d'ordre 1 en temps et 2 en espace (calcul identique à celui du schéma explicite).

<u>Stabilité</u>. Les schémas explicite et implicite sont tout deux du même ordre mais le schéma implicite est plus compliqué à mettre en place puisqu'il faut inverser la matrice A telle que $u^n = Au^{n+1}$ pour pouvoir l'utiliser. L'avantage du schéma implicite est que celui-ci, comme on va le montrer ci-dessous, est inconditionnellement stable en norme l^{∞} .

On a

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

Soit $j_0 \in \{1, \ldots, J\}$ tel que

$$u_{j_0}^{n+1} = \max_{1 \le j \le J} u_j^{n+1}$$

Alors, puisque $u_{j_0}^{n+1} \ge u_j^{n+1}$, on obtient

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

d'où $\max_{1\leq j\leq J}u_j^n\geq u_{j_0}^{n+1}=\max_{1\leq j\leq J}u_j^{n+1}.$ On en déduit

$$\max_{1 \le j \le J} u_j^0 \ge \max_{1 \le j \le J} u_j^n, \quad \forall \ n = 1, \dots, N$$

On en déduit le principe du maximum discret et donc le schéma implicite centré est inconditionnellement stable en norme l^{∞} .

4.2.3 Le θ -schéma et le schéma de Crank-Nicholson

Il s'agit d'une combinaison convexe du schéma explicite et du schéma implicite suivant un paramètre $\theta \in [0, 1]$. Autrement dit, celui-ci est donné par

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} - \theta \, \frac{u_{j-1}^{n+1} - 2u_j^{n+1} + u_{j+1}^{n+1}}{\Delta x^2} - (1 - \theta) \, \frac{u_{j-1}^n - 2u_j^n + u_{j+1}^n}{\Delta x^2} = 0 \tag{4.2.3}$$

Pour $\theta = 0$, on obtient le schéma explicite et, pour $\theta = 1$, on obtient le schéma implicite.

Proposition 4.2.2. Le θ -schéma (4.2.3) est consistant :

- 1. d'ordre 1 en temps et 2 en espace si $\theta \neq \frac{1}{2}$,
- 2. d'ordre 2 en temps et en espace si $\theta = \frac{1}{2}$, dans ce cas le θ -schéma est appelé schéma de Crank-Nicholson.

Démonstration. Partant des développements limités obtenus pour les schémas précédents, l'erreur de consistance au temps t_{n+1} et au point x_j est donnée par

$$\begin{split} r_j^{n+1} &= \partial_t \tilde{u}_j^n + \frac{\Delta t}{2} \, \partial_t^2 \tilde{u}_j^n - \theta \partial_x^2 \tilde{u}_j^{n+1} - (1-\theta) \partial_x^2 \tilde{u}_j^n + O(\Delta t^2 + \Delta x^2) \\ &= \partial_t \tilde{u}_j^n - \partial_x^2 \tilde{u}_j^n + \frac{\Delta t}{2} \, \partial_t^2 \tilde{u}_j^n - \theta (\partial_x^2 \tilde{u}_j^{n+1} - \partial_x^2 \tilde{u}_j^n) + O(\Delta t^2 + \Delta x^2) \\ &= \frac{\Delta t}{2} \, \partial_t^2 \tilde{u}_j^n - \theta (\partial_x^2 \tilde{u}_j^{n+1} - \partial_x^2 \tilde{u}_j^n) + O(\Delta t^2 + \Delta x^2). \end{split}$$

Or on a

$$\partial_x^2 \tilde{u}_j^{n+1} = \partial_x^2 \tilde{u}_j^n + \Delta t \, \partial_t \partial_x^2 \tilde{u}_j^n + O(\Delta t^2)$$

d'où

$$r_{j}^{n+1} = \frac{\Delta t}{2} \partial_{t}^{2} \tilde{u}_{j}^{n} - \theta \Delta t \partial_{t} \partial_{x}^{2} \tilde{u}_{j}^{n} + O(\Delta t^{2} + \Delta x^{2})$$
$$= \frac{\Delta t}{2} \partial_{t} \left(\partial_{t} \tilde{u}_{j}^{n} - 2\theta \partial_{x}^{2} \tilde{u}_{j}^{n} \right) + O(\Delta t^{2} + \Delta x^{2})$$

Le terme $\partial_t \tilde{u}_j^n - 2\theta \partial_x^2 \tilde{u}_j^n$ ne s'annulant (indépendamment de n et j) que si $\theta = \frac{1}{2}$, on obtient le résultat cherché.

Étant donné qu'établir un principe du maximum pour le θ -schéma peut s'avérer complexe (dû à la dépendance suivant le paramètre θ), on va étudier la stabilité l^2 du schéma. On a le résultat suivant :

Proposition 4.2.3. Soit $\lambda := \frac{\Delta t}{\Delta x^2}$. Le θ -schéma (4.2.3) est :

- 1. inconditionnellement stable en norme l^2 si $\theta \geq \frac{1}{2}$,
- 2. stable en norme l^2 sous la condition $\lambda \leq \frac{1}{2(1-2\theta)}$ si $\theta < \frac{1}{2}$.

Démonstration. Par transformée de Fourier, on obtient

$$\mathcal{F}v^{n+1}(\xi) - \mathcal{F}v^n(\xi) - 2\theta\lambda(\cos(\xi\Delta x) - 1)\mathcal{F}v^{n+1}(\xi) - 2(1-\theta)\lambda(\cos(\xi\Delta x) - 1)\mathcal{F}v^n(\xi) = 0,$$

d'où

$$(1 - 2\theta\lambda(\cos(\xi\Delta x) - 1))\mathcal{F}v^{n+1}(\xi) = (1 + 2(1 - \theta)\lambda(\cos(\xi\Delta x) - 1))\mathcal{F}v^{n}(\xi).$$

On obtient donc pour coefficient d'amplification

$$g(\xi;\Delta x,\Delta t) = \frac{1+2(1-\theta)\lambda(\cos(\xi\Delta x)-1)}{1-2\theta\lambda(\cos(\xi\Delta x)-1)} = \frac{1-4(1-\theta)\lambda\sin^2(\frac{\xi\Delta x}{2})}{1+4\theta\lambda\sin^2(\frac{\xi\Delta x}{2})}.$$

En particulier,

$$|g(\xi; \Delta x, \Delta t)| = \frac{|1 - 4(1 - \theta)\lambda \sin^2(\frac{\xi \Delta x}{2})|}{1 + 4\theta\lambda \sin^2(\frac{\xi \Delta x}{2})}.$$

Si $\theta \geq \frac{1}{2}$, alors $1 - \theta \leq \theta$ donc

$$|1 - 4(1 - \theta)\lambda\sin^2(\frac{\xi\Delta x}{2})| \le 1 + 4(1 - \theta)\lambda\sin^2(\frac{\xi\Delta x}{2}) \le 1 + 4\theta\lambda\sin^2(\frac{\xi\Delta x}{2}),$$

et comme $1 + 4\theta\lambda \sin^2(\frac{\xi\Delta x}{2}) \ge 1$ on obtient $|g(\xi; \Delta x, \Delta t)| \le 1$ pour tout λ . Si $\theta < \frac{1}{2}, |g(\xi; \Delta x, \Delta t)| \le 1$ si et seulement si

$$1 - 4(1 - \theta)\lambda \sin^2\left(\frac{\xi\Delta x}{2}\right) \le 1 + 4\theta\lambda \sin^2\left(\frac{\xi\Delta x}{2}\right)$$

 et

$$4(1-\theta)\lambda\sin^2\left(\frac{\xi\Delta x}{2}\right) - 1 \le 1 + 4\theta\lambda\sin^2\left(\frac{\xi\Delta x}{2}\right)$$

La première inégalité a toujours lieu, et, pour la seconde, on montre aisément que celle-ci équivaut à la condition $\lambda \leq \frac{1}{2(1-2\theta)}$.

Chapitre 5

Bilan sur les différences finies et exercices

5.1 Problèmes elliptiques

On a considéré le problème modèle

$$\begin{cases}
-\Delta u = f \quad \text{dans } \Omega, \\
u = 0 \quad \text{sur } \partial \Omega,
\end{cases}$$
(5.1.1)

où $\Omega :=]0, 1[^d \text{ avec } d := 1 \text{ ou } 2.$

Pour le cas d = 1, on a étudié la discrétisation donnée par

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

qui est équivalente au système matriciel $A_h U_h = b_h$ où $U_h := (u_1, \ldots, u_N)^T$, $b_h := (f(x_1), \ldots, f(x_N))^T$ et $A_h \in \mathbb{R}^{N \times N}$ est donnée par

$$A_h := -\frac{1}{h^2} \begin{pmatrix} -2 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \dots & 0 \\ \vdots & \ddots & & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 1 & -2 \end{pmatrix}$$

Exercice 5.1.1. La condition aux limites de Fourier est une situation intermédiaire entre condition de Dirichlet et condition de Neumann, celle-ci s'écrit (en dimension n quelconque) $\frac{\partial u}{\partial \nu} + \alpha u = \beta$ sur $\partial \Omega$ où $\alpha > 0, \beta \in \mathbb{R}$.

On considère le problème unidimensionnel, sujet a une condition aux limites de Fourier, suivant

$$\begin{cases} -u''(x) + c u(x) = f(x) & \text{dans }]0,1[, \\ u'(0) - \alpha(u - \beta) = 0, \\ u'(1) + \alpha(u - \beta) = 0, \end{cases}$$
(5.1.2)

où $f \in C([0,1]), c \ge 0, \alpha > 0$ et $\beta \in \mathbb{R}$. On admet l'existence d'une solution Déterminer une discrétisation de (5.1.2) par différences finies.

Exercice 5.1.2. On considère le problème

$$\begin{cases} -u''(x) + c(x) u(x) = 0 & \text{dans }]0, 1[, \\ u(0) = \alpha, \\ u(1) = \beta, \end{cases}$$
(5.1.3)

où $c \in C([0,1]), c \ge 0$ et $\alpha, \beta \in \mathbb{R}$. On admet l'existence et l'unicité d'une solution $u \in C^2([0,1[) \cap C^1([0,1]))$ de (5.1.3). Déterminer une discrétisation de (5.1.3) par différences finies et étudier celle-ci (consistance, stabilité et convergence).

Pour le cas d = 2, on a étudié la discrétisation de (5.1.1) par le schéma à cinq points donné par $-\Delta_h u_{i,j} = f(x_i, y_j)$ où $\Delta_h u_{i,j}$ est défini par

$$\Delta_h u_{i,j} = \frac{-4u_{i,j} + u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1}}{h^2}.$$

Une autre discrétisation possible est le schéma à 9 points, dans ce cas on a $-\Delta_{i,j}u_{i,j} = b_{i,j}$, où on a posé

$$\Delta_{i,j}u := \frac{u_{i-1,j-1} + u_{i+1,j-1} + u_{i-1,j+1} + u_{i+1,j+1} + 4u_{i-1,j} + 4u_{i,j-1} + 4u_{i,j+1} - 20u_{i,j}}{6h^2}$$

 et

$$b_{i,j} := f(x_i, y_j) + \frac{f(x_i, y_{j-1}) + f(x_{i-1}, y_j) - 4f(x_i, x_j) + f(x_{i+1}, y_j) + f(x_i, y_{j+1})}{12}.$$

Exercice 5.1.3. Montrer que le schéma à 9 points est d'ordre 4.

5.2 Problèmes hyperboliques

On a considéré le problème modèle de l'équation d'advection

$$\begin{cases} \partial_t u + a \,\partial_x u = 0, \quad (t, x) \in \mathbb{R}^+_* \times \mathbb{R}, \\ u(t = 0, x) = u_0(x), \quad x \in \mathbb{R}, \end{cases}$$
(5.2.1)

où la vitesse de transport a et la donnée initiale u_0 sont données. Dans le tableau 5.1, on rappelle les résultats obtenus pour les différents schémas étudiés (avec $\lambda := a \frac{\Delta t}{\Delta x}$).

Remarque 5.2.1.

1. Dans la pratique, il est nécessaire de se placer dans un domaine borné (en temps et en espace). Dans ce cas, on prend $(t, x) \in [0, T] \times [a, b]$ et il faut rajouter des conditions aux limites en espace (à juger suivant le problème étudié). On obtient, pour le cas des schémas à deux niveaux, l'égalité $U^{n+1} = AU^n$. Alors, pour calculer U^n , il suffit d'implémenter U^0 à partir de la donnée initiale u_0 et de calculer U^n par la formule $U^n = A^n U^0$.

Schéma	Ordre	Stabilité
$\frac{\text{Explicite centr}\acute{e}}{u_j^{n+1} = u_j^n - \frac{\lambda}{2}(u_{j+1}^n - u_{j-1}^n)}$	$O(\Delta t + \Delta x^2)$ si $\lambda \neq 1$ $O(\Delta t^2 + \Delta x^2)$ si $\lambda = 1$ $\frac{\Delta t}{\Delta x}$ constant	instable l^{∞} et l^2
$\frac{\text{Explicite décentré aval}}{u_j^{n+1} = u_j^n - \lambda(u_{j+1}^n - u_j^n)}$	$O(\Delta t + \Delta x^2) \text{ si } \lambda \neq 1$ $O(\Delta t^2 + \Delta x^2) \text{ si } \lambda = 1$ $\frac{\Delta t}{\Delta x} \text{ constant}$	stable l^2 si $\lambda \leq 1$
$\frac{\text{Explicite décentré amont } (a > 0)}{u_j^{n+1} = u_j^n - \lambda(u_j^n - u_{j-1}^n)}$	$O(\Delta t + \Delta x^2) \text{ si } \lambda \neq 1$ $O(\Delta t^2 + \Delta x^2) \text{ si } \lambda = 1$ $\frac{\Delta t}{\Delta x} \text{ constant}$	stable l^2 si $\lambda \leq 1$
$\frac{\underline{\text{Lax}}}{u_j^{n+1} = \frac{1}{2}(u_{j-1}^n + u_{j+1}^n) - \frac{\lambda}{2}(u_{j+1}^n - u_j^n)}$	$O(\Delta t + \Delta x^2) \text{ si } \lambda \neq 1$ $O(\Delta t^2 + \Delta x^2) \text{ si } \lambda = 1$ $\frac{\Delta t}{\Delta x} \text{ constant}$	stable l^2 si $\lambda \leq 1$
$\frac{\text{Implicite centré}}{u_j^n = u_j^{n+1} + \frac{\lambda}{2}(u_{j+1}^{n+1} - u_{j-1}^{n+1})}$	$O(\Delta t + \Delta x^2)$	stable l^2
$ \frac{\text{Lax-Wendroff}}{u_j^{n+1}} = u_j^n - \frac{\lambda}{2}(u_{j+1}^n - u_{j-1}^n) \\ + \frac{\lambda^2}{2}(u_{j+1}^n - 2u_j^n + u_{j+1}^n) $	$O(\Delta t^2 + \Delta x^2)$ $\frac{\Delta t}{\Delta x} \text{ constant}$	stable l^2 si $\lambda \leq 1$

FIGURE 5.1 – Exemples de différents schémas pour l'équation d'advection

2. Comme pour le problème de Poisson, on peut se ramener à un système matriciel. En effet, pour le problème de Dirichlet on est passé du problème continu $-\Delta u = f$ dans Ω , u = 0 sur $\partial\Omega$, au problème discret $A_h U_h = b_h$ où A_h est la discrétisation de $-\Delta$. Pour l'équation de transport (5.2.1) avec conditions aux limites de Dirichlet homogènes le problème discret s'écrit TU + a XU = B où le vecteur U contient les $u_{n,j}$ pour tout n et j, T est la discrétisation de ∂_t alors que X est la discrétisation de ∂_x et la donnée B est obtenue à partir de u_0 . En suivant cette méthode, il est alors nécessaire, comme dans le cas du problème de Poisson, de montrer que la matrice T + a X est inversible.

Exercice 5.2.2. En considérant des conditions aux limites de Dirichlet, déterminer, pour chacun des schémas du tableau 5.1, la matrice A.

Dans le cas du schéma implicite, on a $U^n = BU^{n+1}$. Il faut alors vérifier que B est inversible pour obtenir A telle que $U^{n+1} = AU^n$.

Exercice 5.2.3. Étudier la consistance et la stabilité du schéma de Lax-Friedrich :

$$\frac{2u_j^{n+1} - u_{j+1}^n - u_{j-1}^n}{2\Delta t} + a \,\frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} = 0.$$
(5.2.2)

Exercice 5.2.4. Montrer que les schémas décentré amont et de Lax-Friedrich sont stables en norme l^{∞} (on pourra s'inspirer de ce qui a été fait pour les problèmes paraboliques).

Exercice 5.2.5. On considère le schéma suivant :

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

où α, β et γ dépendent de $\lambda, \Delta x, \Delta t$. Déterminer α, β et γ de sorte que le schéma soit d'ordre 2. Quel schéma vu précédemment retrouve t-on?

Exercice 5.2.6. On considère l'équation des ondes :

$$\begin{aligned} \partial_t^2 u - \partial_x^2 u &= f & \text{dans }]0, T[\times]0, 1[, \\ u(t,0) &= u(t,1) &= 0 & \text{pour } t \in]0, T[\\ u(t=0) &= u_0 & \text{dans } [0,1], \\ \partial_t u(t=0) &= u_1 & \text{dans } [0,1]. \end{aligned}$$

Pour cette équation, on peut obtenir un θ -schéma analogue à celui de l'équation de la chaleur en discrétisant la dérivée seconde en temps par un schéma centré (voir [1] page 60). Une autre méthode consiste à écrire l'équation des ondes comme une équation d'advection en dimension 2 comme vu dans la section 1.1.4. Alors, $U = (\partial_t u, \partial_x u)^T$, est solution de

$$\partial_t U - A \,\partial_x U = F,\tag{5.2.3}$$

où $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ et $F = \begin{pmatrix} f \\ 0 \end{pmatrix}$. Déterminer pour l'équation (5.2.3) un schéma du type Lax-Wendroff et montrer que celui-ci est d'ordre 2 en temps et en espace, et stable l^2 .

Schéma	Ordre	Stabilité
$u_j^{n+1} = u_j^n + \lambda(u_{j+1}^n - 2u_j^n + u_{j-1}^n)$	$O(\Delta t + \Delta x^2)$	stable l^{∞} si $\lambda \leq \frac{1}{2}$
$u_{j}^{n} = u_{j}^{n+1} - \lambda (u_{j+1}^{n+1} - 2u_{j}^{n+1} + u_{j-1}^{n+1})$	$O(\Delta t + \Delta x^2)$	stable l^{∞}
$ \begin{array}{rcl} \theta & - schéma \\ u_{j}^{n+1} &= u_{j}^{n} + \theta \lambda (u_{j+1}^{n+1} - 2u_{j}^{n+1} + u_{j-1}^{n+1}) \\ &+ (1 - \theta) \lambda (u_{j+1}^{n} - 2u_{j}^{n} + u_{j-1}^{n}) \end{array} $	$O(\Delta t + \Delta x^2) \text{ si } \theta \neq \frac{1}{2}$ $O(\Delta t^2 + \Delta x^2) \text{ si } \theta = \frac{1}{2}$	• stable l^2 pour $\theta \le \frac{1}{2}$ • stable l^2 si $\lambda \le \frac{1}{2(1-\theta)}$ pour $\theta > \frac{1}{2}$

FIGURE 5.2 – Exemples de différents schémas pour l'équation de la chaleur

5.3 Problèmes paraboliques

On a considéré le problème modèle de l'équation de la chaleur

$$\begin{cases} \partial_t u - \partial_x^2 u = 0, & (t, x) \in]0, T[\times]0, 1[, \\ u(t = 0, x) = u_0(x), & x \in]0, 1[, \\ u(t, 0) = u(t, 1) = 0, & t \in]0, T[, \end{cases}$$

$$(5.3.1)$$

où la condition initiale u_0 est donnée. Dans le tableau 5.2, on rappelle les résultats obtenus pour les différents schémas étudiés (où on a posé $\lambda := \frac{\Delta t}{\Delta x^2}$).

Remarque 5.3.1. Les remarques faites sur l'utilisation pratique des schémas des problèmes hyperboliques sont tout aussi valables pour le cas parabolique.

Exercice 5.3.2. Pour chacun des schémas du tableau 5.2, déterminer leur équivalent en dimension 2 d'espace (se placer, par exemple, dans le cas $x \in [0,1]^2$). Dans le cas du schéma implicite, u^n est donné en fonction de u^{n+1} il faut donc inverser la matrice A telle que $u^n = Au^{n+1}$. Déterminer cette matrice A en dimension 2 d'espace.

Exercice 5.3.3. Montrer que le schéma de Crank-Nicholson est stable l^{∞} si $\lambda \leq 1$.

Deuxième partie Méthode des éléments finis

Chapitre 6

Introduction à la notion de formulation variationnelle

6.1 Bref rappel sur les distributions

Exemple 6.1.1. Considérons l'équation de transport

$$\begin{cases} \partial_t u - a \,\partial_x u = 0 & \text{dans } \mathbb{R}^+_* \times \mathbb{R}, \\ u(t = 0, x) = u_0(x) & \text{dans } \mathbb{R}. \end{cases}$$
(6.1.1)

Alors, u est le "transport" de u_0 en fonction du temps t. En particulier, si u_0 n'est pas continue, on peut s'attendre à ce que u ne le soit pas non plus et donc ne vérifie pas (6.1.1) au sens classique. Dans ce cas, il est nécessaire de pouvoir donner un sens "faible" à la notion de solution de (6.1.1) de sorte qu'une solution faible puisse ne pas être nécessairement continue.

Soit $\varphi \in C_c^1(\mathbb{R}^+_* \times \mathbb{R})$. On multiplie l'équation de transport (6.1.1) par φ et on intègre sur $\mathbb{R}^+_* \times \mathbb{R}$. Par intégration par parties, on obtient

$$\int_{\mathbb{R}^+_* \times \mathbb{R}} u(t, x) \partial_t \varphi(t, x) - a \, u(t, x) \partial_x \varphi(t, x) \, dt dx$$
$$= \int_{\mathbb{R}} [u(t, x) \varphi(t, x)]_0^\infty \, dx + \int_{\mathbb{R}^+_*} [a \, u(t, x) \varphi(t, x)]_{-\infty}^\infty \, dt$$

Le terme de droite est nul car φ est à support compact dans $\mathbb{R}^+_* \times \mathbb{R}$ donc s'annule au bord, on en déduit

$$\int_{\mathbb{R}^+_* \times \mathbb{R}} u(t,x) \partial_t \varphi(t,x) - a \, u(t,x) \partial_x \varphi(t,x) \, dt dx = 0.$$
(6.1.2)

Une fonction u vérifiant (6.1.2), pour toute $\varphi \in C_c^1(\mathbb{R}^+_* \times \mathbb{R})$, est dite solution faible de l'équation de transport (6.1.1).

Comme $\varphi \in C_c^1(\mathbb{R}^+_* \times \mathbb{R})$, φ est bornée et donc pour que l'égalité (6.1.2) ait un sens il suffit que u soit intégrable sur le support de φ . Ainsi, l'égalité (6.1.2) a un sens pour toute fonction $\varphi \in C_c^1(\mathbb{R}^+_* \times \mathbb{R})$ si u est intégrable sur tout compact, *i.e.* $u \in L^1_{\text{loc}}(\mathbb{R}^+_* \times \mathbb{R})$. La

forme linéaire T_u définie sur $C_c^1(\mathbb{R}^+_* \times \mathbb{R})$ par

$$< T_u, \varphi > := \int_{\mathbb{R}^+_* \times \mathbb{R}} u(t, x) \varphi(t, x) dt dx,$$

est appelée **distribution associée** à u. On définit $\partial_t T_u$ et $\partial_x T_u$ comme étant les formes linéaires continues définies sur $C_c^1(\mathbb{R}^+_* \times \mathbb{R})$ par

$$\langle \partial_t T_u, \varphi \rangle := -\langle T_u, \partial_t \varphi \rangle$$
 et $\langle \partial_x T_u, \varphi \rangle := -\langle T_u, \partial_x \varphi \rangle$.

L'égalité (6.1.2) est alors équivalente à

$$\forall \varphi \in C_c^1(\mathbb{R}^+_* \times \mathbb{R}), \qquad \langle \partial_t T_u + a \, \partial_x T_u, \varphi \rangle = 0. \tag{6.1.3}$$

Si u est une solution faible de l'équation de transport (6.1.1), on dit encore que u est solution de (6.1.1) au sens des distributions (au sens où T_u est solution de (6.1.3)).

Définition 6.1.2. Soit Ω un ouvert de \mathbb{R}^d , $d \geq 1$. Une **distribution** sur Ω est une forme linéaire continue sur $C_c^{\infty}(\Omega)$ muni de la topologie induite. On note $\mathcal{D}'(\Omega)$ l'espace vectoriel réel des distributions sur Ω .

Remarque 6.1.3. La définition des distributions et la notation $\mathcal{D}'(\Omega)$ sont dus à Laurent Schwartz. La notation $\mathcal{D}'(\Omega)$ vient du fait que Schwartz note $\mathcal{D}(\Omega)$ l'ensemble $C_c^{\infty}(\Omega)$ et que, pour un espace vectoriel E, l'espace dual (espace vectoriel des formes linéaires continues sur E) est noté E'.

Exemple 6.1.4. Represent l'argumentation de l'Exemple 6.1.1, pour toute fonction $f \in L^1_{loc}(\Omega)$, on définit une distribution associée T_f par

$$\forall \varphi \in C_c^{\infty}(\Omega), \qquad \langle T_f, \varphi \rangle := \int_{\Omega} f(x)\varphi(x) \ dx.$$

En général, on écrit f au lieu de T_f tant qu'il n'y a pas de confusion possible entre la distribution et la fonction.

Définition 6.1.5.

- 1. Un **multi-indice** α est un k-uplet $(\alpha_1, \ldots, \alpha_k)$, avec $k \leq d$, sa longueur est $|\alpha| = \alpha_1 + \cdots + \alpha_k$ et on note D^{α} l'opérateur différentiel défini par $D^{\alpha} = \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \ldots \partial_{x_k}^{\alpha_k}$.
- 2. Soit Ω un ouvert de \mathbb{R}^d et $T \in \mathcal{D}'(\Omega)$. Pour tout multi-indice α , $D^{\alpha}T$ désigne la distribution définie par

$$\forall \varphi \in C_c^{\infty}(\Omega), \qquad \langle D^{\alpha}T, \varphi \rangle := (-1)^{|\alpha|} \langle T, D^{\alpha}\varphi \rangle.$$
(6.1.4)

On dit que $D^{\alpha}T$ est la **dérivée** d'exposant α de T.

Exercice 6.1.6. Si $f \in L^1_{loc}(\mathbb{R})$, la dérivée f' de f au sens des distributions est donnée par

$$\langle f', \varphi \rangle = - \langle f, \varphi' \rangle, \quad \forall \varphi \in C_c^{\infty}(\mathbb{R}).$$

On a vu dans l'Exemple 6.1.1 que la formule d'intégration par parties est la base de la définition d'une distribution sur \mathbb{R} . Dans le cas de problèmes de dimension supérieure, il est alors naturel de faire appel à une formule d'intégration par parties dans \mathbb{R}^d , d > 1. Ce qui est l'objet de la section suivante.

6.2 Les formules de Green

Définition 6.2.1. Un ouvert Ω de \mathbb{R}^d , d > 1, est dit **régulier de classe** C^1 si son bord $\partial\Omega$ est une hypersurface (variété de dimension d-1) régulière et si Ω est situé d'un seul coté de sa frontière.

Dans toute la suite, ν désigne la normale extérieure au bord $\partial\Omega$ de Ω , *i.e.* $\nu(x)$ est perpendiculaire au plan tangent à $\partial\Omega$ au point $x \in \partial\Omega$, pointe vers l'extérieur de Ω et est de norme 1.

On rappelle la formule de Green (voir, par exemple, [4]) :

Théorème 6.2.2 (Formule de Green). Soit Ω un ouvert de \mathbb{R}^d régulier de classe C^1 . Alors, pour toute fonction $w \in C^1(\overline{\Omega})^d$ à support borné dans $\overline{\Omega}$, on a

$$\int_{\Omega} \operatorname{div}(w)(x) \, dx = \int_{\partial \Omega} w(x) \cdot \nu(x) \, d\sigma_x, \qquad (6.2.1)$$

où $d\sigma_x$ désigne la **mesure surfacique** sur $\partial\Omega$ (mesure de Lebesgue en dimension d-1), \cdot est le produit scalaire usuel de \mathbb{R}^d et div $(w) := \partial_{x_1}w_1 + \cdots + \partial_{x_d}w_d$ est la **divergence** $de w := (w_1, \ldots, w_d).$

En prenant w(x) := u(x) v(x), où $u \in C^1(\overline{\Omega})^d$ et $v \in C^1(\overline{\Omega})$, on obtient

$$\operatorname{div}(w) = \nabla v \cdot u + v \operatorname{div}(u).$$

On en déduit le résultat suivant :

Corollaire 6.2.3 (Formule d'intégration par parties). Soit Ω un ouvert de \mathbb{R}^d régulier de classe C^1 . Soient $u \in C^1(\overline{\Omega})^d$ et $v \in C^1(\overline{\Omega})$ à support borné dans $\overline{\Omega}$. Alors, on a

$$\int_{\Omega} \nabla v \cdot u \, dx = -\int_{\Omega} v \operatorname{div}(u) \, dx + \int_{\partial \Omega} v \, u \cdot \nu \, d\sigma_x.$$
(6.2.2)

Remarque 6.2.4. En dimension d = 1, avec $\Omega =]a, b[$ on a $\nu(a) = -1$, $\nu(b) = 1$ et la formule (6.2.2) devient alors

$$\int_{a}^{b} u'v \, dx = -\int_{a}^{b} uv' \, dx + (u(b)v(b) - u(a)v(a)).$$

Autrement dit, on retrouve bien la formule d'intégration par parties classique.

On donne une troisième formulation de la formule de Green qui sera très utile pour déterminer une formulation variationnelle pour les problèmes elliptiques.

Corollaire 6.2.5. Soit Ω un ouvert de \mathbb{R}^d régulier de classe C^1 . Soient $u \in C^2(\overline{\Omega})$ et $v \in C^1(\overline{\Omega})$ à support borné dans $\overline{\Omega}$. Alors, on a

$$\int_{\Omega} \Delta uv \ dx = -\int_{\Omega} \nabla u \cdot \nabla v \ dx + \int_{\partial \Omega} v \frac{\partial u}{\partial \nu} \ d\sigma_x, \tag{6.2.3}$$

 $o\hat{u} \frac{\partial u}{\partial \nu} := \nabla u \cdot \nu$ est la dérivée normale de u sur $\partial \Omega$.

Démonstration. Il suffit dans (6.2.2) de remplacer u par ∇u .

6.3 Applications à l'équation de Laplace

Dans cette section, on déduit des formules de Green des propriétés importantes de l'équation de Laplace. Dans toute la suite Ω désigne un ouvert borné de \mathbb{R}^d , régulier de classe C^1 .

Proposition 6.3.1 (Formule de la moyenne sphérique). Soit $u \in C^2(\overline{\Omega})$ une fonction harmonique dans Ω (i.e. $\Delta u = 0$ dans Ω). Alors, pour toute boule $B_R(y)$ de centre $y \in \Omega$ et de rayon R > 0 tel que $B_R(y) \subset \Omega$, on a

$$u(y) = \frac{1}{\omega_d R^{d-1}} \int_{\partial B_R(y)} u(x) \, d\sigma_x, \tag{6.3.1}$$

$$u(y) = \frac{d}{\omega_d R^d} \int_{B_R(y)} u(x) \, dx,$$
(6.3.2)

où $\omega_d = |\partial B_1(0)|$ est la mesure de Lebesgue de la sphère unité de \mathbb{R}^d .

Remarque 6.3.2. La formule (6.3.1) est bien une moyenne puisque $|\partial B_R(y)| = \omega_d R^{d-1}$. *Démonstration.* Soit $0 < \rho \leq R$. En prenant v = 1 dans la formule de Green (6.2.3) avec $\Omega = B_\rho(y)$, on obtient

$$0 = \int_{B_{\rho}(y)} \Delta u \, dx = \int_{\partial B_{\rho}(y)} \frac{\partial u}{\partial \nu} \, d\sigma_x.$$
(6.3.3)

On va reformuler l'intégrale du second membre à l'aide d'un changement de variable. Si $x \in B_{\rho}(y)$, alors $x = y + \rho \omega$, où $\omega \in \partial B_1(0)$ et $\nu(x) = w$.

On note $v(\omega) = u(y + \rho\omega) = u(x)$ pour $\omega \in \partial B_1(0)$ et $x \in \partial B_\rho(y)$. On a

$$\frac{\partial v}{\partial \rho}(\omega) = \omega \cdot \nabla u(y + \rho \omega) = \frac{\partial u}{\partial \nu}(x)$$

On en déduit par changement de variable $x = y + \rho \omega$ (car $d\sigma_x = \rho^{d-1} d\sigma_\omega$)

$$\int_{\partial B_{\rho}(y)} \frac{\partial u}{\partial \nu} \, d\sigma_x = \int_{\partial B_1(0)} \frac{\partial v}{\partial \rho}(\omega) \rho^{d-1} \, d\sigma_\omega$$
$$= \rho^{d-1} \frac{\partial}{\partial \rho} \int_{\partial B_1(0)} v(\omega) d\sigma_\omega$$
$$= \rho^{d-1} \frac{\partial}{\partial \rho} \int_{\partial B_1(0)} u(y + \rho\omega) \, d\sigma_\omega$$

Par le second changement de variable $\omega = (x - y)/\rho$, on obtient

$$\int_{\partial B_{\rho}(y)} \frac{\partial u}{\partial \nu} \, d\sigma_x = \rho^{d-1} \frac{\partial}{\partial \rho} \left[\rho^{1-d} \int_{\partial B_{\rho}(y)} u(x) \, d\sigma_x \right].$$

D'après l'égalité (6.3.3), on a alors

$$\frac{\partial}{\partial \rho} \left[\rho^{1-d} \int_{\partial B_{\rho}(y)} u(x) \, d\sigma_x \right] = 0$$

Donc l'application $\rho \mapsto \rho^{1-d} \int_{\partial B_{\rho}(y)} u(x) \, d\sigma_x$ est constante, en particulier on en déduit pour tout $0 < \rho \leq R$

$$\rho^{1-d} \int_{\partial B_{\rho}(y)} u(x) \ d\sigma_x = R^{1-d} \int_{\partial B_R(y)} u(x) \ d\sigma_x. \tag{6.3.4}$$

Comme $u \in C^2(\Omega)$, $u(y + \rho \omega) = u(y) + O(\rho)$ d'où

$$\lim_{\rho \to 0} \rho^{1-d} \int_{\partial B_{\rho}(y)} u(x) \, d\sigma_x = \lim_{\rho \to 0} \rho^{1-d} \int_{\partial B_{\rho}(y)} u(y) + O(\rho) \, d\sigma_x$$
$$= \lim_{\rho \to 0} \rho^{1-d} |\partial B_{\rho}(y)| \, u(y) = \omega_d \, u(y).$$

D'après (6.3.4) obtient $u(y) = \omega_d^{-1} R^{1-d} \int_{\partial B_R(y)} u(x) \, d\sigma_x$, qui donne l'égalité (6.3.1). Pour l'égalité (6.3.2), il suffit d'appliquer l'égalité (6.3.1) pour $R = \rho$ et ensuite d'intégrer la variable ρ sur]0, R[.

La formule de la moyenne sphérique permet de montrer que toute fonction harmonique vérifie un principe du maximum.

Proposition 6.3.3 (Principe du maximum). Soit $u \in C^2(\overline{\Omega})$ une fonction harmonique dans Ω supposé connexe. Alors, on a

$$\min_{x \in \partial \Omega} u(x) = \min_{x \in \overline{\Omega}} u(x) \quad et \quad \max_{x \in \partial \Omega} u(x) = \max_{x \in \overline{\Omega}} u(x).$$

Démonstration. On montre $\max_{x \in \partial \Omega} u(x) = \max_{x \in \overline{\Omega}} u(x)$ (l'autre égalité s'en déduit en considérant -u).

Supposons : il existe $y \in \Omega$ tel que $\max_{x \in \overline{\Omega}} u(x) = u(y) = M$. On pose

$$\Omega_M := \{ x \in \Omega \mid u(x) = M \}.$$

Puisque $y \in \Omega_M$, $\Omega \neq \emptyset$. De plus, $\Omega_M = u^{-1}(\{M\})$ et u est continue donc Ω_M est fermé dans Ω . Soit $z \in \Omega_M$. On applique la formule de la moyenne (6.3.2) a la fonction harmonique u - M dans la boule $B_R(z)$:

$$0 = u(z) - M = \frac{n}{\omega_d R^d} \int_{B_R(z)} u(x) - M \, dx.$$

Alors, $\int_{B_R(z)} u(x) - M \, dx = 0$. Or u - M est continue et $u - M \leq 0$ dans $B_R(z)$ donc u = M dans $B_R(z)$. On en déduit $B_R(z) \subset \Omega_M$ donc Ω_M est ouvert dans Ω . Finalement, Ω_M est un ouvert fermé non vide de Ω connexe donc $\Omega_M = \Omega$, d'où u = M dans Ω . En conclusion, on en déduit soit $\max_{x \in \partial \Omega} u(x) = \max_{x \in \overline{\Omega}} u(x)$ soit u est constante. \Box

On en déduit l'unicité d'une solution régulière du problème de Dirichlet :

Théorème 6.3.4. Si Ω est connexe et $u \in C^2(\overline{\Omega})$ est solution de

$$\begin{cases} -\Delta u = f \quad dans \ \Omega, \\ u = g \quad sur \ \partial\Omega, \end{cases}$$
(6.3.5)

alors u est l'unique solution dans $C^2(\overline{\Omega})$ de (6.3.5).

Démonstration. Soit $v \in C^2(\Omega)$ une autre solution de (6.3.5). On note w := v - u, alors $w \in C^2(\Omega)$ vérifie $\Delta w = 0$ dans Ω et w = 0 sur $\partial \Omega$. D'après le principe du maximum, on a

$$0 = \min_{\partial \Omega} w = \min_{\Omega} w \le w \le \max_{\Omega} w = \max_{\partial \Omega} w = 0 \quad \text{dans } \Omega,$$

d'où w = 0 dans Ω .

6.4 Un première formulation variationnelle du problème de Dirichlet homogène

On considère le problème de Dirichlet homogène

$$\begin{cases} -\Delta u = f \quad \text{dans } \Omega, \\ u = 0 \quad \text{sur } \partial \Omega. \end{cases}$$
(6.4.1)

On suppose que $u \in C^2(\overline{\Omega})$ est la solution de (6.4.1). Soit $\varphi \in C_c^{\infty}(\Omega)$, en multipliant l'équation (6.4.1) par φ et en intégrant sur Ω , on obtient :

$$-\int_{\Omega} \Delta u \, \varphi \, \, dx = \int_{\Omega} f \, \varphi \, \, dx.$$

D'après la formule d'intégration par parties (6.2.3), on a

$$\int_{\Omega} \nabla u \cdot \nabla \varphi \, dx - \int_{\partial \Omega} \frac{\partial u}{\partial \nu} \varphi \, d\sigma_x = \int_{\Omega} f \, \varphi \, dx.$$

Comme φ est à support compact dans Ω , le terme de bord est nul : $\int_{\partial\Omega} \frac{\partial u}{\partial \nu} \varphi \, d\sigma_x = 0$. On en déduit que u vérifie

$$\int_{\Omega} \nabla u \cdot \nabla \varphi \, dx = \int_{\Omega} f \, \varphi \, dx.$$

Proposition 6.4.1. Soit $u \in C^2(\overline{\Omega})$. On pose

$$X := \{ \varphi \in C^1(\overline{\Omega}) \mid \varphi = 0 \text{ sur } \partial \Omega \}.$$

Alors, u est solution de (6.4.1) si et seulement si $u \in X$ et vérifie

$$\int_{\Omega} \nabla u \cdot \nabla \varphi \, dx = \int_{\Omega} f \, \varphi \, dx, \qquad \forall \, \varphi \in X.$$
(6.4.2)

Remarque 6.4.2.

- 1. La formulation (6.4.2) a l'avantage de ne nécessiter que de la régularité C^1 et est équivalente à (6.4.1) si et seulement $u \in C^2(\overline{\Omega})$. Ainsi la formulation (6.4.2) est une formulation faible du problème (6.4.1) dite **formulation variationnelle**. Les fonctions $\varphi \in X$ en sont les **fonctions tests**.
- 2. Celle-ci a un sens en mécanique : il s'agit du principe des travaux virtuels. De ce fait, chercher u en tant que solution de (6.4.2) plutôt que (6.4.1) a davantage de sens du point de vue de la mécanique.

Pour démontrer la Proposition 6.4.1, on fera appel au résultat suivant dont on donne la démonstration plus loin :

Lemme 6.4.3. Soit $g \in C(\Omega)$. Si pour toute $\phi \in C_c(\Omega)$, on a

$$\int_{\Omega} g \phi \, dx = 0,$$

alors g = 0 dans Ω .

Démonstration de la Proposition 6.4.1. Il suffit de montrer que si u est solution de (6.4.2), alors u est solution de (6.4.1). En intégrant par parties la formulation variationnelle (6.4.2), on obtient

$$\int_{\Omega} (\Delta u + f) \varphi \, dx = 0, \qquad \forall \ \varphi \in X.$$

D'après le Lemme 6.4.3, on a $\Delta u + f = 0$ dans Ω car $\Delta u + f \in C(\Omega)$. De plus, $u \in X$ donc u = 0 sur $\partial \Omega$.

Démonstration du Lemme 6.4.3. Supposons qu'il existe $x_0 \in \Omega$ tel que $g(x_0) \neq 0$, par exemple $g(x_0) > 0$. Comme g est continue dans Ω , il existe un voisinage $\omega \subset \Omega$ de x_0 tel que g > 0 dans ω . Soit $\varphi \in C_c(\Omega)$ telle que supp $(\varphi) \subset \omega, \varphi \neq 0$ sur ω et $\varphi \geq 0$ alors

$$0 = \int_{\Omega} g \varphi \, dx = \int_{\omega} g \varphi \, dx.$$

Comme $g \varphi$ est positive ou nulle sur ω et continue, on en déduit $g \varphi = 0$ sur ω . Or g > 0 sur ω et $\varphi \neq 0$ dans ω , on aboutit donc à une contradiction.

La formulation faible (6.4.2) nécessite une régularité C^1 , on va voir comment déterminer un espace E contenant l'espace X (faisant donc appel à moins de régularité) tel que pour $u \in E$ la formulation variationnelle (6.4.2) a un sens pour des fonctions tests $\varphi \in E$.

Remarquons tout d'abord que l'égalité (6.4.2) a un sens si l'on suppose seulement $\nabla u \in L^2(\Omega)^d$. De plus, en mécanique, l'énergie associée au problème (6.4.1) est donnée par

$$E(u) = \frac{1}{2} \int_{\Omega} |\nabla u|^2 \, dx - \int_{\Omega} u f \, dx < \infty.$$

En particulier si on considère un terme source f = 0, le fait que l'énergie soit finie entraîne $\nabla u \in L^2(\Omega)^d$. Ainsi, il est naturel de considérer une distribution u dont le gradient au sens des distributions est une distribution associée à une fonction de $L^2(\Omega)^d$ notée ∇u . Enfin, en supposant $f \in L^2(\Omega)$, il suffit que $u \in L^2(\Omega)$ pour que la deuxième intégrale de l'énergie soit finie, en effet d'après l'inégalité de Cauchy-Schwarz on a

$$\int_{\Omega} u f \, dx \le ||u||_{L^{2}(\Omega)} \, ||f||_{L^{2}(\Omega)} < \infty.$$

Finalement, pour $u \in L^2(\Omega)$, $\nabla u \in L^2(\Omega)^d$ avec u = 0 sur $\partial\Omega$, la formulation variationnelle (6.4.2) a un sens et la propriété d'énergie finie est vérifiée (au sens des distributions). On est alors amené à chercher une solution u de la formulation variationnelle (6.4.2) dans l'espace

$$H^{1}(\Omega) := \{ u \in L^{2}(\Omega) \mid \nabla u \in L^{2}(\Omega)^{d}, u = 0 \text{ sur } \partial \Omega \}.$$

Cet espace fait partie des espaces de Sobolev que l'on va étudier dans le chapitre suivant.

Chapitre 7

Espaces de Sobolev

7.1 Définitions et propriétés

Définition 7.1.1. Soit Ω un ouvert de \mathbb{R}^d . L'espace de Sobolev $W^{m,p}(\Omega)$, où $m, p \in \mathbb{N}^*$ est défini par

$$W^{m,p}(\Omega) := \{ u \in L^p(\Omega) \mid \forall \ \alpha \text{ multi-indice}, \ |\alpha| \le m, \ D^{\alpha}u \in L^p(\Omega) \},$$
(7.1.1)

où $D^{\alpha}u$ est la dérivée d'exposant α au sens des distributions de u. Dans le cas particulier p = 2, on note $W^{m,2}(\Omega) = H^m(\Omega)$.

Remarque 7.1.2.

- 1. L'espace $H^1(\Omega)$ donné par la Définition 7.1.1 est bien le même espace que celui défini à la fin du chapitre précédent. En effet, $\nabla u \in L^2(\Omega)^d$ si et seulement si, pour tout $i = 1, \ldots, d, \ \partial_{x_i} u \in L^2(\Omega)$, ce qui est encore équivalent à $D^{\alpha} u \in L^2(\Omega)$ pour tout multi-indice α tel que $|\alpha| = 1$.
- 2. La Définition 7.1.1 a encore un sens pour $p = +\infty$. En particulier, en dimension 1 on peut montrer que $W^{1,\infty}(I)$ est l'ensemble des fonctions Lipschitziennes sur l'intervalle I.
- 3. Il existe une classe plus générale d'espaces de Sobolev (qui ne seront pas étudiés dans ce cours). Ce sont les espaces $W^{s,p}(\Omega)$ où $s \in \mathbb{R}$. Par exemple, dans le cas $p = 2, s \ge 0$ et $\Omega = \mathbb{R}^d$, on définit l'espace de Sobolev $H^s(\mathbb{R}^d)$ par

$$H^{s}(\mathbb{R}^{d}) := \{ u \in L^{2}(\mathbb{R}^{d}) \mid \xi \mapsto (1 + |\xi|^{2})^{\frac{s}{2}} \mathcal{F}u(\xi) \in L^{2}(\mathbb{R}^{d}) \},\$$

où \mathcal{F} est la transformée de Fourier.

Pour les besoins de ce cours, on étudiera seulement les espaces de Sobolev $H^m(\Omega)$. La notation H pour ces espaces vient de la propriété suivante (dont la démonstration est laissée au lecteur) :

Proposition 7.1.3. L'espace de Sobolev $H^m(\Omega)$ est un espace de Hilbert muni du produit scalaire (\cdot, \cdot) défini par

$$(u,v) := \sum_{|\alpha| \le m} \int_{\Omega} D^{\alpha} u \, D^{\alpha} v \, dx.$$
(7.1.2)

Exemple 7.1.4. Pour m = 1, le produit scalaire (7.1.2) s'écrit simplement

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

et la norme associée est donnée par

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

où $|\cdot|$ désigne la norme euclidienne de \mathbb{R}^d .

On a le résultat d'injection suivant :

Théorème 7.1.5. Si Ω est un ouvert borné régulier de classe C^1 et si $m > \frac{d}{2}$ alors $H^m(\Omega)$ est un sous-espace de $C(\overline{\Omega})$. En particulier, en dimension d = 1 on a $H^m(\Omega) \subset C(\overline{\Omega})$ pour tout $m \in \mathbb{N}^*$.

Remarque 7.1.6.

- 1. Si m > 1 et $u \in H^m(\Omega)$ alors $D^{\alpha}u \in H^{m-1}(\Omega)$ pour tout multi-indice α tel que $|\alpha| = 1$. Donc si $m 1 > \frac{d}{2}$, $D^{\alpha}u \in C(\overline{\Omega})$ d'où $u \in C^1(\overline{\Omega})$. Ainsi, si m est très grand devant d alors les fonctions de $H^m(\Omega)$ sont des fonctions régulières.
- 2. Les fonctions de $H^m(\Omega)$ sont des fonctions qui ne sont définies que presque partout contrairement aux fonctions de $C(\overline{\Omega})$, ainsi l'inclusion $H^m(\Omega) \subset C(\overline{\Omega})$ peut paraître étrange. En fait, il faut comprendre dans cette inclusion que toute fonction $u \in H^m(\Omega)$ admet un représentant $\tilde{u} \in C(\overline{\Omega})$.

Démonstration. On donne le résultat dans le cas d = 1 (voir [5] pour le cas général). On considère $\Omega :=]0, 1[$ pour simplifier. On fera appel au résultat suivant dont on donne la démonstration plus bas.

Lemme 7.1.7. Si $u \in H^1(0,1)$ est telle que u' = 0 sur]0,1[alors u est constante p.p. sur]0,1[.

Soit $u \in H^1(0, 1)$. Si $x \in [0, 1]$, alors, par l'inégalité de Cauchy-Schwarz, on a

$$\left| \int_0^x u'(s) \, ds \right| \le \left(\int_0^x \, ds \right)^{1/2} \, \|u'\|_{L^2(0,x)} = \sqrt{x} \, \|u'\|_{L^2(0,1)} < \infty.$$

Ainsi, on peut définir sur]0,1[la fonction v par

$$v(x) := \int_0^x u'(s) \, ds, \quad \forall \ x \in]0,1[.$$
 (7.1.3)

De même, pour $x, y \in]0, 1[$, on a

$$|v(x) - v(y)| = \left| \int_x^y u'(s) \, ds \right| \le \sqrt{|x - y|} \, ||u'||_{L^2(0, 1)},$$

d'où v est continue sur [0,1]. Soit $\varphi \in C_c^{\infty}(0,1)$. D'après le théorème de Fubini, on a

$$\int_0^1 v(x)\varphi'(x) \, dx = \int_0^1 \int_0^x u'(s)\varphi'(x) \, ds \, dx = \int_0^1 u'(s) \int_s^1 \varphi'(x) \, dx \, ds = -\int_0^1 u'(s)\varphi(s) \, ds.$$

On en déduit que la dérivée au sens des distributions v' de v est u' donc, d'après le Lemme 7.1.7, il existe une constante c telle que u = v + c p.p. dans]0, 1[, donc u admet un représentant \tilde{u} qui est continu sur [0, 1].

Remarque 7.1.8. La démonstration précédente montre de plus, d'après (7.1.3), que l'on a

$$\tilde{u}(x) = \tilde{u}(y) + \int_x^y u'(s) \, ds, \quad \forall \ x, y \in]0, 1[.$$

Démonstration du Lemme 7.1.7. Soit $u \in H^1(0,1)$ telle que u' = 0 sur]0,1[. Alors, pour toute $\varphi \in C_c^1(0,1)$, on a

$$\langle u', \varphi \rangle = -\int_0^1 u \, \varphi' \, dx = 0.$$

Soit $\psi \in C_c(0,1)$. Pour toute $w \in C_c(0,1)$, la fonction $\left(w - \int_0^1 w \, ds\right) \psi$ est continue, à support compact dans (0,1) et à moyenne nulle sur (0,1) donc celle-ci admet une primitive $\varphi \in C_c^1(\Omega)$. Alors, on obtient

$$0 = \langle u', \varphi \rangle = -\int_0^1 u(x) \left(w(x) - \int_0^1 w(s) \, ds \, \psi(x) \right) \, dx$$

D'après le théorème de Fubini, on en déduit

$$\int_0^1 \left(u - \int_0^1 u\psi \, dx \right) w \, ds = 0, \quad \forall \ w \in C_c(0,1).$$

Alors, on a $u = \int_0^1 u \psi dx$ p.p. dans]0, 1[.

Notation 7.1.9. Pour $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, on note $x' := (x_1, \ldots, x_{d-1})$, de sorte que $x = (x', x_d)$. On désigne par \mathbb{R}^d_+ l'ouvert de \mathbb{R}^d défini par

$$\mathbb{R}^{d}_{+} := \{ x = (x', x_{d}) \in \mathbb{R}^{d} \mid x_{d} > 0 \}.$$
(7.1.4)

Théorème 7.1.10. Soit $m \in \mathbb{N}^*$.

- 1. L'espace $C_c^m(\mathbb{R}^d)$ est dense dans $H^m(\mathbb{R}^d)$.
- 2. Si Ω est un ouvert borné régulier de classe C^m , m > 0, ou si $\Omega = \mathbb{R}^d_+$, alors $C^m_c(\overline{\Omega})$ est dense dans $H^m(\Omega)$.

Remarque 7.1.11. Il faut prendre garde de ne pas confondre $C_c^m(\overline{\Omega})$ et $C_c^m(\Omega)$, on a seulement l'inclusion stricte $C_c^m(\Omega) \subsetneq C_c^m(\overline{\Omega})$. En effet, les fonctions de $C_c^m(\Omega)$ sont nécessairement nulles sur $\partial\Omega$ contrairement aux fonctions de $C_c^m(\overline{\Omega})$.

Schéma de preuve. La démonstration de ce résultat est assez technique. On décrit seulement les principales étapes de la démonstration.

1) Pour $u \in H^m(\mathbb{R}^d)$, on construit (procédé de régularisation) une suite $(u_n)_{n\geq 0}$ de $C^m(\mathbb{R}^d)$ en posant $u_n := u \star \varphi_n$, où $(\varphi_n)_{n\geq 0}$ est une suite de fonctions régulière et \star désigne le produit de convolution. Alors la suite $(u_n)_{n\geq 0}$ converge dans $H^m(\mathbb{R}^d)$ vers u. Ensuite, on construit (procédé de troncature) une suite $(v_n)_{n\geq 0}$ de $C_c^m(\mathbb{R}^d)$ en posant $v_n := u_n \psi_n$ où $(\psi_n)_{n\geq 0}$ est une suite de fonctions régulières ayant pour support la boule unité de rayon 2n. Enfin, on montre que la suite $(v_n)_{n\geq 0}$ converge vers u dans $H^m(\mathbb{R}^d)$.

2) Le résultat obtenu dans le cas de \mathbb{R}^d se généralise au cas de \mathbb{R}^d_+ en montrant qu'il existe un opérateur de prolongement continu de \mathbb{R}^d_+ dans \mathbb{R}^d . On obtient finalement le résultat pour tout ouvert Ω borné régulier de classe C^m en passant de Ω à \mathbb{R}^d_+ par cartes locales. C'est ce dernier point qui justifie de considérer des ouverts situés d'un seul côté de leur bord (cela vient du fait que \mathbb{R}^d_+ vérifie cette propriété).

 \square

7.2 L'espace $H_0^1(\Omega)$

Définition 7.2.1. Soit Ω un ouvert de \mathbb{R}^d . L'espace $H_0^1(\Omega)$ est défini comme étant l'adhérence de $C_c^{\infty}(\Omega)$ dans $H^1(\Omega)$.

Remarque 7.2.2.

- 1. La Définition 7.2.1 signifie que $u \in H_0^1(\Omega)$ si et seulement si il existe une suite u_n de $C_c^{\infty}(\Omega)$ telle que $\lim_{n \to +\infty} ||u_n u||_{H^1(\Omega)} = 0.$
- 2. On peut montrer que l'adhérence dans $H^1(\Omega)$ de $C_c^1(\Omega)$ et $C_c^{\infty}(\Omega)$ coïncident, autrement dit $H_0^1(\Omega)$ est aussi l'adhérence dans $H^1(\Omega)$ de $C_c^1(\Omega)$.
- 3. D'après le Théorème 7.1.10, on a $H_0^1(\mathbb{R}^d) = H^1(\mathbb{R}^d)$. Pour $\Omega = \mathbb{R}^d_+$ ou Ω un ouvert borné régulier de classe C^1 , l'espace $C_c^1(\Omega)$ est un sous-espace strict de $C_c^1(\overline{\Omega})$ donc $H_0^1(\Omega)$ est un sous-espace strict de $H^1(\Omega)$.

Par définition, l'espace $H_0^1(\Omega)$ est un sous-espace fermé de $H^1(\Omega)$, on en déduit donc le résultat suivant

Proposition 7.2.3. L'espace $H_0^1(\Omega)$ est un espace de Hilbert pour le produit scalaire de $H^1(\Omega)$.

Une propriété importante de l'espace $H_0^1(\Omega)$ est que ses éléments vérifient l'inégalité de Poincaré.

Proposition 7.2.4 (Inégalité de Poincaré). Soit Ω un ouvert de \mathbb{R}^d borné dans une direction d'espace (ou plus). Alors, il existe une constante c > 0 telle que

$$\forall v \in H_0^1(\Omega), \quad \int_{\Omega} |v|^2 \ dx \le c \ \int_{\Omega} |\nabla v|^2 \ dx. \tag{7.2.1}$$

Démonstration. Puisque Ω est borné dans une direction, il existe $k \in \{1, \ldots, d\}, m \leq M$ tels que si $x \in \Omega$ alors $m \leq x_k \leq M$. Pour simplifier et sans perdre en généralité, on peut supposer k = d. Pour $x \in \Omega$, on note $x = (x', x_d)$. Soit $\varphi \in C_c^{\infty}(\Omega)$, alors, puisque φ est nulle au bord de Ω , on a

$$\varphi(x) = \int_m^{x_d} \frac{\partial \varphi}{\partial x_d}(x', t) dt.$$

Par l'inégalité de Cauchy-Schwarz, on obtient

$$\begin{aligned} |\varphi(x)|^2 &= \left| \int_m^{x_d} \frac{\partial \varphi}{\partial x_d}(x',t) \, dt \right|^2 \\ &\leq \int_m^{x_d} dt \, \int_m^{x_d} \left| \frac{\partial \varphi}{\partial x_d}(x',t) \right|^2 \, dt \\ &\leq (M-m) \int_m^{x_d} \left| \frac{\partial \varphi}{\partial x_d}(x',t) \right|^2 \, dt \end{aligned}$$

On en déduit par le théorème de Fubini

$$\int_{\Omega} |\varphi(x)|^2 dx \leq (M-m) \int_{\Omega} \int_m^{x_d} \left| \frac{\partial \varphi}{\partial x_d}(x',t) \right|^2 dt dx$$
$$\leq (M-m) \int_m^M \int_{\Omega} \left| \frac{\partial \varphi}{\partial x_d}(x) \right|^2 dx dt$$
$$\leq (M-m)^2 \int_{\Omega} |\nabla \varphi|^2 dx,$$

ce qui donne le résultat pour toute fonction φ de $C_c^{\infty}(\Omega)$. Si $v \in H_0^1(\Omega)$, il existe une suite φ_n de $C_c^{\infty}(\Omega)$ telle que $\lim_{n \to +\infty} ||v - \varphi_n||_{H^1(\Omega)} = 0$. Autrement dit, on a

$$\lim_{n \to +\infty} \int_{\Omega} |v - \varphi_n|^2 \, dx = \lim_{n \to +\infty} \int_{\Omega} |\nabla v - \nabla \varphi_n|^2 \, dx = 0.$$

On obtient

$$\begin{aligned} \|v\|_{L^{2}(\Omega)} &\leq \|v - \varphi_{n}\|_{L^{2}(\Omega)} + \|\varphi_{n}\|_{L^{2}(\Omega)} \\ &\leq \|v - \varphi_{n}\|_{L^{2}(\Omega)} + (M - m)\|\nabla\varphi_{n}\|_{L^{2}(\Omega)^{d}} \\ &\leq \|v - \varphi_{n}\|_{L^{2}(\Omega)} + (M - m)\|\nabla\varphi_{n} - \nabla v\|_{L^{2}(\Omega)^{d}} + (M - m)\|\nabla v\|_{L^{2}(\Omega^{d})}. \end{aligned}$$

En passant à la limite, quand $n \to +\infty$, on déduit (7.2.1).

Remarque 7.2.5. L'inégalité de Poincaré est fausse dans $H^1(\Omega)$. En effet, les fonctions constantes sont dans $H^1(\Omega)$ et si u est constante alors $\nabla u = 0$ p.p. dans Ω . Si l'inégalité de Poincaré avait lieu dans $H^1(\Omega)$ on aurait alors u = 0 p.p. dans Ω . En particulier, l'inégalité de Poincaré est vraie dans $H^1_0(\Omega)$ notamment car la seule fonction constante de $H^1_0(\Omega)$ est la fonction identiquement nulle.

Corollaire 7.2.6. Soit Ω un ouvert de \mathbb{R}^d connexe et borné dans une direction d'espace (ou plus). Alors, la semi norme définie par

$$||v||_{H^1_0(\Omega)} := ||\nabla v||_{L^2(\Omega)}, \tag{7.2.2}$$

est une norme sur $H_0^1(\Omega)$ équivalente à la norme induite par celle de $H^1(\Omega)$.

Démonstration. Si $||v||_{H_0^1(\Omega)} = 0$ alors $\nabla v = 0$ p.p. dans Ω . Comme Ω est connexe, on en déduit que v est constante p.p. dans Ω . La seule constante de $H_0^1(\Omega)$ étant la fonction identiquement nulle, on en déduit v = 0 p.p. dans Ω . Ainsi, $|| \cdot ||_{H_0^1(\Omega)}$ est bien une norme sur $H_0^1(\Omega)$. Pour montrer l'équivalence, il suffit de montrer qu'il existe deux constantes $c_1 > 0$ et $c_2 > 0$ telles que pour toute $v \in H_0^1(\Omega)$, on a

$$c_1 ||v||_{H^1(\Omega)} \le ||v||_{H^1_0(\Omega)} \le c_2 ||v||_{H^1(\Omega)}$$

La deuxième inégalité est évidente avec $c_2 = 1$. Pour la première, il suffit d'appliquer l'inégalité de Poincaré, on obtient

$$||v||_{H^{1}(\Omega)}^{2} = ||v||_{L^{2}(\Omega)}^{2} + ||\nabla v||_{L^{2}(\Omega)}^{2} \le (1+c) ||\nabla v||_{L^{2}(\Omega)}^{2},$$

ce qui donne le résultat en prenant $c_1 = (1+c)^{-1/2}$.

7.3 Traces et formules de Green

En dimension $d \geq 2$, les fonctions de $H^1(\Omega)$ peuvent ne pas être continues. En particulier, on sait que celles-ci sont seulement définies presque partout sur Ω . Or le bord $\partial\Omega$ d'un ouvert Ω de \mathbb{R}^d est un ensemble de mesure nulle, il n'est donc pas clair de pouvoir définir la valeur d'une fonction de $H^1(\Omega)$ sur $\partial\Omega$. Pour cela, on introduit l'application trace.

Définition 7.3.1. Soit Ω un ouvert borné de \mathbb{R}^d . L'application trace γ_0 est l'application linéaire définie de $H^1(\Omega) \cap C(\overline{\Omega})$ dans $L^2(\partial\Omega) \cap C(\overline{\partial\Omega})$ par $\gamma_0 v := v_{|\partial\Omega}$.

Remarque 7.3.2. L'application trace considérée est définie sur $H^1(\Omega) \cap C(\overline{\Omega})$ car on va chercher à étendre cette application à $H^1(\Omega)$ tout entier. Considérer l'application trace sur $C(\overline{\Omega})$ ne serait pas suffisant car il ne s'agit pas d'un sous-espace de $H^1(\Omega)$.

Théorème 7.3.3. Soit Ω un ouvert borné régulier de classe C^1 ou $\Omega = \mathbb{R}^d_+$. Alors, l'application γ_0 se prolonge en une application linéaire continue de $H^1(\Omega)$ dans $L^2(\partial\Omega)$ notée encore γ_0 . En particulier, il existe une constante c > 0 telle que, pour toute $u \in H^1(\Omega)$, on a

$$||\gamma_0 u||_{L^2(\partial\Omega)} \le c \, ||u||_{H^1(\Omega)}. \tag{7.3.1}$$

On note $u_{|\partial\Omega} := \gamma_0 u$.

Remarque 7.3.4. Le résultat est faux si l'on se place seulement dans $L^2(\Omega)$ ou si l'ouvert Ω n'est pas régulier.

Démonstration. On montre le résultat pour le cas $\Omega = \mathbb{R}^d_+$, le cas général étant plus technique. Pour cela, il suffit de montrer que l'inégalité (7.3.1) a lieu pour des fonctions régulières, le résultat s'en déduit ensuite par densité. Soit $v \in C_c^{\infty}(\overline{\mathbb{R}^d_+})$. Alors, la restriction de v au bord $\partial \mathbb{R}^d_+$ est $\gamma_0 v(x') = v(x', 0)$, où $x' \in \mathbb{R}^{d-1}$, et vérifie

$$\begin{aligned} |v(x',0)|^2 &= -\int_0^\infty \frac{\partial}{\partial x_d} (v(x',x_d)^2) \, dx_d \\ &= -2\int_0^\infty v(x',x_d) \, \frac{\partial v}{\partial x_d} (x',x_d) \, dx_d. \end{aligned}$$

De l'inégalité $-2ab \leq a^2 + b^2$, on déduit

$$|v(x',0)|^2 \le \int_0^\infty v(x',x_d)^2 \, dx_d + \int_0^\infty \left|\frac{\partial v}{\partial x_d}(x',x_d)\right|^2 \, dx_d$$

En intégrant en x', on obtient

$$\int_{\mathbb{R}^{d-1}} |v(x',0)|^2 \, dx' \le \int_{\mathbb{R}^d_+} v(x',x_d)^2 \, dx_d + \int_{\mathbb{R}^d_+} \left| \frac{\partial v}{\partial x_d}(x',x_d) \right|^2 \, dx_d \le ||v||_{H^1(\mathbb{R}^d_+)}^2,$$

d'où $||\gamma_0 v||_{L^2(\partial \mathbb{R}^d_+)} \leq ||v||_{H^1(\mathbb{R}^d_+)}$. Par densité de $C_c^1(\overline{\mathbb{R}^d_+})$ dans $H^1(\mathbb{R}^d_+)$, on obtient le résultat (dans \mathbb{R}^d_+).

Un conséquence très importante du théorème de trace est que celui-ci permet de donner une caractérisation de $H_0^1(\Omega)$ qui s'avérera très utile dans le chapitre suivant. On admettra le résultat et on renvoie à [5, 10, 12] pour plus de détails.

Corollaire 7.3.5 (Admis). Soit Ω un ouvert borné régulier de classe C^1 . Alors, $H_0^1(\Omega)$ est donné par

$$H_0^1(\Omega) = \{ v \in H^1(\Omega) \mid v_{|\partial\Omega} = 0 \text{ sur } \partial\Omega \},$$
(7.3.2)

où on a utilisé la notation $v_{|\partial\Omega} := \gamma_0 v$.

Remarque 7.3.6.

- 1. L'égalité (7.3.2) se justifie du fait que toute fonction de $H_0^1(\Omega)$ est limite d'une suite de fonctions de $C_c^1(\Omega)$ qui sont nécessairement nulles sur $\partial\Omega$. Néanmoins cette rapide explication ne suffit pas à montrer l'égalité (7.3.2).
- 2. Cette caractérisation de $H_0^1(\Omega)$ peut encore s'écrire $H_0^1(\Omega) = \text{Ker}(\gamma_0)$ où γ_0 désigne l'application trace définie de $H^1(\Omega)$ dans $L^2(\partial\Omega)$. On peut montrer que l'on a $\text{Im}(\gamma_0) \subsetneq L^2(\partial\Omega)$. On note $H^{\frac{1}{2}}(\partial\Omega) := \text{Im}(\gamma_0)$. On peut aussi montrer que $H^{\frac{1}{2}}(\partial\Omega)$ est un sous-espace dense de $L^2(\partial\Omega)$ (voir [10] pour plus de détails).

Une autre application importante du théorème de trace est la généralisation de la formule de Green à $H^1(\Omega)$.

Corollaire 7.3.7 (Formule de Green dans $H^1(\Omega)$). Soit Ω un ouvert borné régulier de classe C^1 . Soient $u \in H^1(\Omega)^d$ et $v \in H^1(\Omega)$, alors on a

$$\int_{\Omega} u \cdot \nabla v \, dx = -\int_{\Omega} v \operatorname{div}(u) \, dx + \int_{\partial \Omega} \gamma_0 v \, \gamma_0 u \cdot \nu \, d\sigma_x.$$
(7.3.3)

Démonstration. Le résultat a déjà été vu dans le cas de fonctions dans $C^1(\overline{\Omega})$. La généralisation à $H^1(\Omega)$ s'obtient par densité grâce à la continuité de l'application trace γ_0 . Comme $C^1(\overline{\Omega})$ est dense dans $H^1(\Omega)$, il existe deux suites $(u_n)_n$ de $C^1(\overline{\Omega})^d$ et $(v_n)_n$ de $C^1(\overline{\Omega})$ qui convergent dans $H^1(\Omega)^d$ et $H^1(\Omega)$ vers u et v. Pour tout n, d'après le Corollaire 6.2.3, on a

$$\int_{\Omega} u_n \cdot \nabla v_n \, dx = -\int_{\Omega} v_n \operatorname{div}(u_n) \, dx + \int_{\partial \Omega} v_n \, u_n \cdot \nu \, d\sigma_x.$$
(7.3.4)

Comme $(u_n)_n$ converge vers u dans $H^1(\Omega)$, $(u_n)_n$ converge vers u dans $L^2(\Omega)^d$ et $(\operatorname{div}(u_n))_n$ converge vers $\operatorname{div}(u)$ dans $L^2(\Omega)$. De même, $(v_n)_n$ converge vers v dans $L^2(\Omega)$ et $(\nabla v_n)_n$ converge vers ∇v dans $L^2(\Omega)^d$. On obtient alors en passant à la limite dans (7.3.4)

$$\int_{\Omega} u \cdot \nabla v \, dx = -\int_{\Omega} v \operatorname{div}(u) \, dx + \lim_{n \to +\infty} \int_{\partial \Omega} v_n \, u_n \cdot \nu \, d\sigma_x$$

Pour passer à la limite dans la dernière intégrale, on utilise le théorème de trace. Comme l'application γ_0 est continue de $H^1(\Omega)$ dans $L^2(\partial\Omega)$, on obtient que $(\gamma_0 u_n)_n$ converge vers $\gamma_0 u$ dans $L^2(\partial\Omega)^d$ et $(\gamma_0 v_n)_n$ converge vers $\gamma_0 v$ dans $L^2(\partial\Omega)$. La normale ν étant continue, on en déduit

$$\int_{\partial\Omega} v_n \, u_n \cdot \nu \, d\sigma_x = \int_{\partial\Omega} \gamma_0 v_n \, \gamma_0 u_n \cdot \nu \, d\sigma_x \xrightarrow[n \to +\infty]{} \int_{\partial\Omega} \gamma_0 v \, \gamma_0 u \cdot \nu \, d\sigma_x,$$

ce qui donne le résultat.

On peut donner d'autres résultats de trace pour les espace $H^m(\Omega)$ avec m > 1 (par exemple définir une dérivée sur le bord). On va donner ci-dessous un résultat dans le cas m = 2.

Définition 7.3.8. Soit Ω un ouvert borné de \mathbb{R}^d . L'**application trace** γ_1 est l'application linéaire définie de $C^1(\overline{\Omega}) \cap H^2(\Omega)$ dans $C(\overline{\partial\Omega}) \cap L^2(\overline{\partial\Omega})$ par $\gamma_1 v := \frac{\partial v}{\partial v_1 \partial \Omega}$.

Théorème 7.3.9. Soit Ω un ouvert borné régulier de classe C^1 . Alors, l'application trace γ_1 se prolonge en une application linéaire continue de $H^2(\Omega)$ dans $L^2(\partial\Omega)$, encore notée γ_1 . En particulier, il existe c > 0 telle que, pour toute $v \in H^2(\Omega)$, on a

$$||\gamma_1 v||_{L^2(\partial\Omega)} \le c \, ||v||_{H^2(\Omega)}.\tag{7.3.5}$$

On note $\frac{\partial v}{\partial \nu}_{|\partial \Omega} := \gamma_1 v.$

Démonstration. Si $v \in H^2(\Omega)$, $\nabla v \in H^1(\Omega)^d$. D'après le théorème de trace dans $H^1(\Omega)$, on a

$$||\gamma_0(\nabla v)||_{L^2(\partial\Omega)} \le c \, ||\nabla v||_{H^1(\Omega)} \le c \, ||v||_{H^2(\Omega)}.$$

La normale ν étant continue sur $\partial\Omega$, on déduit $||\gamma_0(\nabla v) \cdot \nu||_{L^2(\partial\Omega)} \leq c ||v||_{H^2(\Omega)}$, d'où le résultat.

Comme dans le cas du théorème de trace dans $H^1(\Omega)$, on en déduit une généralisation de la formule de Green.

Corollaire 7.3.10. Soit Ω un ouvert borné régulier de classe C^2 . Si $u \in H^2(\Omega)$ et $v \in H^1(\Omega)$, on a

$$\int_{\Omega} \Delta u \, v \, dx = -\int_{\Omega} \nabla u \cdot \nabla v \, dx + \int_{\partial \Omega} \gamma_1 u \, \gamma_0 v \, d\sigma_x.$$

Remarque 7.3.11. La démonstration est analogue à celle de la formule de Green dans $H^1(\Omega)$. Néanmoins, il est nécessaire ici de supposer que l'ouvert Ω est de classe C^2 pour avoir le résultat de densité des fonctions de $C^2(\overline{\Omega})$ dans $H^2(\Omega)$.

7.4 Théorème de compacité de Rellich et applications

Théorème 7.4.1 (Rellich). Soit Ω un ouvert borné de \mathbb{R}^d . De toute suite bornée de $H_0^1(\Omega)$ on peut extraire une sous-suite convergente dans $L^2(\Omega)$. Autrement dit, l'injection de $H_0^1(\Omega)$ dans $L^2(\Omega)$ est compacte.

De plus, si Ω est régulier de classe C^1 , alors l'injection de $H^1(\Omega)$ dans $L^2(\Omega)$ est aussi compacte.

Remarque 7.4.2.

1. Pour le cas de l'injection de $H^1(\Omega)$ dans $L^2(\Omega)$, on renvoie à [5] pour une démonstration.

- 2. La définition d'opérateur compact est rappelée dans le chapitre 12 (Définition 12.1.2).
- 3. Le résultat peut être faux dans $H^1(\Omega)$ si l'on ne suppose pas l'ouvert régulier.
- 4. Il faut prendre garde que le résultat est une convergence dans $L^2(\Omega)$ et non dans $H^1(\Omega)$. En particulier, on n'a pas, à priori, la convergence des gradients dans $L^2(\Omega)^d$ fort.
- 5. Il existe une version plus générale de ce résultat, appelée théorème de Rellich-Kondrachov et qui s'énonce dans les espaces de Sobolev $W^{1,p}(\Omega)$ (voir [5]).

Démonstration. La démonstration que l'on donne ci-après provient de [13], page 85, et serait due à Lars Hörmander (voir aussi [12], page 29). Soit $(u_n)_n$ une suite bornée de $H_0^1(\Omega)$. Comme $L^2(\Omega)$ est séparable, d'après le théorème de Banach-Alaoglu, il existe $u \in L^2(\Omega)$ et une sous-suite $(u_{n_j})_j$ qui converge dans $L^2(\Omega)$ faible vers u. On pose $v_j := u_{n_j} - u$ et donc on veut montrer qu'il existe une sous-suite de $(v_n)_n$ qui converge vers 0 dans $L^2(\Omega)$ fort. On étend chaque v_n par 0 en dehors de Ω . On peut alors montrer (voir [5] page 173) que $v_n \in H^1(\mathbb{R}^d)$ et que les normes de v_n dans $H^1(\mathbb{R}^d)$ et $H_0^1(\Omega)$ sont les mêmes. C'est cette propriété qui rend fondamental le fait de se placer dans $H_0^1(\Omega)$. On a

$$\mathcal{F}v_n(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\Omega} v_n(x) e^{-ix\cdot\xi} dx = \frac{1}{\sqrt{2\pi}} \langle v_n, f_\xi \rangle_{L^2(\Omega)}$$

où $f_{\xi}(x) := e^{ix \cdot \xi} \in L^2(\Omega)$ (car $f_{\xi} \in L^{\infty}(\Omega) \subset L^2(\Omega)$, puisque $|\Omega| < \infty$). Alors, comme $(v_n)_n$ converge vers 0 dans $L^2(\Omega)$ faible, on en déduit que $\mathcal{F}v_n(\xi)$ converge vers 0 dans \mathbb{R} . De plus, puisque $(v_n)_n$ est bornée dans $H^1_0(\Omega)$, on a

$$|\mathcal{F}v_n(\xi)| \le ||v_n||_{L^2(\Omega)} ||f_{\xi}||_{L^2(\Omega)} \le ||v_n||_{H^1_0(\Omega)} \le c.$$

Pour tout $\rho > 0$, la constante c est intégrable sur la boule $|\xi| \leq \rho$. Alors, d'après le théorème de convergence dominée de Lebesgue, on a

$$\int_{|\xi| \le \rho} |\mathcal{F}v_n(\xi)|^2 d\xi \quad \underset{n \to +\infty}{\longrightarrow} \quad 0, \quad \forall \ \rho > 0.$$
(7.4.1)

Puisque la suite $(v_n)_n$ est bornée dans $H_0^1(\Omega)$, il existe c > 0 telle que $\|\nabla v_n\|_{L^2(\Omega)^d} \leq c$. Par le théorème de Plancherel, obtient alors

$$\|\xi \mathcal{F}v_n\|_{L^2(\mathbb{R}^d)^d} = \|\mathcal{F}(\nabla v_n)\|_{L^2(\mathbb{R}^d)^d} = \|\nabla v_n\|_{L^2(\mathbb{R}^d)^d} = \|\nabla v_n\|_{L^2(\Omega)^d} \le c,$$

d'où $\int_{\mathbb{R}^d} |\xi|^2 |\mathcal{F}v_n|^2 d\xi \leq c$. On en déduit

$$\int_{|\xi|>\rho} |\mathcal{F}v_n|^2 \, d\xi = \frac{1}{\rho^2} \int_{|\xi|>\rho} \rho^2 \, |\mathcal{F}v_n|^2 \, d\xi \le \frac{c}{\rho^2}, \quad \forall \ \rho > 0.$$
(7.4.2)

Pour tout $\rho > 0$, on a

$$\|v_n\|_{L^2(\Omega)}^2 = \|\mathcal{F}(v_n)\|_{L^2(\mathbb{R}^d)}^2 = \int_{|\xi| \le \rho} |\mathcal{F}v_n|^2 \ d\xi + \int_{|\xi| > \rho} |\mathcal{F}v_n|^2 \ d\xi$$

Alors, (7.4.1) et (7.4.2) entraînent

$$\limsup_{n} \|v_n\|_{L^2(\Omega)}^2 \le \limsup_{n} \int_{|\xi| \le \rho} |\mathcal{F}v_n|^2 \, d\xi + \limsup_{n} \int_{|\xi| > \rho} |\mathcal{F}v_n|^2 \, d\xi \le \frac{c}{\rho^2}, \quad \forall \ \rho > 0.$$

On en déduit, en faisant tendre ρ vers l'infini, $\lim_n ||v_n||^2_{L^2(\Omega)} = 0$. Donc la suite $(v_n)_n$ converge dans $L^2(\Omega)$ fort vers 0.

En application de ce résultat de compacité, on va donner une nouvelle démonstration de l'inégalité de Poincaré, dans un cadre plus restrictif que celui vu précédemment mais qui a l'avantage de s'adapter pour d'autres inégalités du même type (voir l'inégalité de Poincaré-Wirtinger plus bas).

Proposition 7.4.3 (Inégalité de Poincaré bis). Soit Ω un ouvert borné connexe de \mathbb{R}^d . Alors, il existe c > 0 telle que, pour toute $v \in H^1_0(\Omega)$, on a

$$||v||_{L^2(\Omega)} \le c \, ||\nabla v||_{L^2(\Omega)}.$$

Démonstration. On raisonne par l'absurde. Supposons : il existe une suite $(u_n)_n$ de $H_0^1(\Omega)$ telle que pour tout n, on a

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

On pose $v_n := u_n/||u_n||_{L^2(\Omega)}$. Alors, on obtient

$$1 = ||v_n||_{L^2(\Omega)} > n \, ||\nabla v_n||_{L^2(\Omega)}. \tag{7.4.3}$$

On en déduit que v_n et ∇v_n sont bornées dans $L^2(\Omega)$ donc v_n est bornée dans $H^1(\Omega)$. D'après le théorème de Rellich, il existe une sous-suite $(v_{n_k})_k$ qui converge dans $L^2(\Omega)$ vers $v \in H^1_0(\Omega)$. D'après (7.4.3), on a $\nabla v = 0$ p.p. dans Ω connexe donc v est constante p.p. dans Ω . Puisque $v \in H^1_0(\Omega)$, on en déduit v = 0 p.p. dans Ω . Or $||v||_{L^2(\Omega)} = \lim_k ||v_{n_k}||_{L^2(\Omega)} = 1$, d'où une contradiction.

Proposition 7.4.4 (Inégalité de Poincaré-Wirtinger). Soit Ω un ouvert connexe et borné régulier de classe C^1 . Pour $u \in H^1(\Omega)$, on note $\overline{u} := \frac{1}{|\Omega|} \int_{\Omega} u \, dx$, où $|\Omega|$ désigne la mesure de Lebesgue de Ω . Alors, il existe une constante c > 0 telle que

$$\forall \ u \in H^1(\Omega), \quad ||u - \overline{u}||_{L^2(\Omega)} \le c \, ||\nabla u||_{L^2(\Omega)}.$$

Démonstration. Laissée en exercice.

7.5 Dualité

On rappelle que pour un espace de Banach E, on note E' son dual, *i.e.* l'ensemble des formes linéaires continues sur E. De plus, pour $L \in E'$ et $v \in E$, on note $\langle L, v \rangle := Lv$. L'application $\langle \cdot, \cdot \rangle$ est appelée **crochet de dualité**. Celle-ci permet de mettre en avant le fait que Lv peut être vu comme un produit scalaire d'après les théorèmes de représentation de Riesz et Riesz-Fréchet (voir [5]) que l'on rappelle ci-dessous :

Théorème 7.5.1 (Théorème de représentation de Riesz). Soient Ω un ouvert de \mathbb{R}^d , $1 et <math>L \in (L^p(\Omega))'$. Alors, il existe un unique $f \in L^{p'}(\Omega)$, où $\frac{1}{p} + \frac{1}{p'} = 1$, tel que

$$\forall u \in L^p(\Omega), \quad \langle L, u \rangle = \int_{\Omega} f(x) u(x) \, dx.$$

De plus, on a

$$||f||_{L^{p'}(\Omega)} = ||L||_{(L^p(\Omega))'}.$$
Théorème 7.5.2 (Théorème de représentation de Riesz-Fréchet). Soient $(H, (\cdot, \cdot)_H)$ un espace de Hilbert sur $\mathbb{K} := \mathbb{R}$ ou \mathbb{C} et $T \in H'$. Alors, il existe un unique $v \in H$ tel que

$$\forall \ u \in H, \quad \langle T, u \rangle = (v, u)_H.$$

De plus, on a

$$||T||_{H'} = ||v||_H.$$

Remarque 7.5.3.

- 1. Le théorème de Riesz exprime que toute forme linéaire $L \in (L^p(\Omega))'$ peut être représentée par une fonction $f \in L^{p'}$. L'application $L \mapsto f$ est un opérateur linéaire isométrique et surjectif qui permet d'identifier le dual de $L^p(\Omega)$ avec $L^{p'}(\Omega)$, *i.e.* $(L^p(\Omega))' = L^{p'}(\Omega)$.
- 2. Le théorème de Riesz-Fréchet exprime que tout espace de Hilbert peut être identifié avec son dual. Autrement dit, toute forme linéaire continue sur H peut être considérée comme un élément de H, *i.e.* H' = H.
- 3. Pour $H = L^2(\Omega)$, les deux résultats coïncident.

Définition 7.5.4. Le dual de l'espace de Sobolev $H_0^1(\Omega)$ est noté $H^{-1}(\Omega)$ et le crochet de dualité correspondant est noté $\langle \cdot, \cdot \rangle_{H^{-1}, H_0^1}$.

Remarque 7.5.5. D'après le théorème de représentation de Riesz-Fréchet, on peut identifier $H_0^1(\Omega)$ et $H^{-1}(\Omega)$ mais dans la pratique on ne fait jamais cette identification. Par contre on identifie $L^2(\Omega)$ avec son dual. La raison de ce choix est justifié par le Corollaire 7.5.7 ci-dessous.

Proposition 7.5.6. L'espace $H^{-1}(\Omega)$ admet la caractérisation suivante :

$$H^{-1}(\Omega) = \left\{ f \in \mathcal{D}'(\Omega) \mid f = g_0 + \sum_{j=1}^d \frac{\partial g_j}{\partial x_j}, \text{ où } g_0, \dots, g_d \in L^2(\Omega) \right\}.$$
 (7.5.1)

Démonstration. On définit l'opérateur T de $H^1_0(\Omega)$ dans $L^2(\Omega)^{d+1}$ par

$$Tu = \left(u, \frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_d}\right).$$

Pour toute $u \in H_0^1(\Omega)$, on a

$$||Tu||_{L^2(\Omega)^{d+1}} = ||u||_{H^1(\Omega)}.$$

Puisque $H_0^1(\Omega)$ est fermé dans $H^1(\Omega)$, on en déduit aisément que $\operatorname{Im}(T)$ est fermé dans $L^2(\Omega)^{d+1}$. De plus, on obtient immédiatement que T est injective et continue donc est bicontinue (*i.e.* d'inverse continu) de $H_0^1(\Omega)$ sur $\operatorname{Im}(T)$. Soit L une forme linéaire continue sur $H_0^1(\Omega)$. Alors LT^{-1} est linéaire continue sur $\operatorname{Im}(T)$. Puisque $\operatorname{Im}(T)$ est fermé dans $L^2(\Omega)^{d+1}$, d'après le théorème de Hahn-Banach, LT^{-1} se prolonge en une forme linéaire continue sur $L^2(\Omega)^{d+1}$. Alors, d'après le théorème de représentation de Riesz, il existe $g_0, \ldots, g_d \in L^2(\Omega)$ telles que

$$LT^{-1}u = \int_{\Omega} u_0 g_0 \, dx + \sum_{i=1}^d \int_{\Omega} u_i g_i \, dx, \quad \forall \ u \in L^2(\Omega)^{d+1}$$

Pour $u \in H_0^1(\Omega)$, on obtient

$$Lu = LT^{-1}Tu = \int_{\Omega} u g_0 \, dx + \sum_{i=1}^d \frac{\partial u}{\partial x_i} g_i \, dx$$

Donc, au sens des distributions, on a $L = g_0 - \sum_{i=1}^d \frac{\partial g_i}{\partial x_i}$.

Corollaire 7.5.7. On a les inclusions suivantes :

$$H_0^1(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega).$$
(7.5.2)

Chapitre 8

Analyse variationnelle des problèmes elliptiques

8.1 Théorie abstraite

Dans cette section, on se place dans le cadre abstrait des espaces de Hilbert. On rappelle le théorème de Lax-Milgram d'existence et unicité pour des formulations variationnelles dans les espaces de Hilbert. De plus, on rappelle aussi la version du théorème de Lax-Milgram en tant que minimisation d'énergie.

On note V un espace de Hilbert réel muni du produit scalaire $(\cdot, \cdot)_V$ et de la norme induite $|| \cdot ||_V$.

Définition 8.1.1. Soit $a: V \times V \to \mathbb{R}$ une forme bilinéaire sur V.

1. On dit que $a(\cdot, \cdot)$ est **continue** s'il existe une constante M > 0 telle que

$$\forall u, v \in V, |a(u, v)| \le M ||u||_V ||v||_V$$

2. On dit que $a(\cdot, \cdot)$ est **coercive** s'il existe une constante m > 0 telle que

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

Pour $L \in V'$ et $a(\cdot, \cdot)$ une forme bilinéaire sur V. On considère la formulation variationnelle

$$\begin{cases} \text{Trouver } u \in V \text{ telle que :} \\ \forall v \in V, \quad a(u,v) = \langle L, v \rangle . \end{cases}$$

$$(8.1.1)$$

Théorème 8.1.2 (Lax-Milgram). Soit $L \in V'$ et $a(\cdot, \cdot)$ une forme bilinéaire sur V continue et coercive. Alors, la formulation variationnelle (8.1.1) admet une solution unique.

Démonstration. Puisque $L \in V'$, d'après le théorème de Riesz-Frécher 7.5.2, il existe un unique $f \in V$ tel que, pour tout $v \in V$, on a $\langle L, v \rangle = (f, v)_V$. De même, pour $u \in V$ fixé, l'application $v \mapsto a(u, v)$ est une forme linéaire continue sur V car $a(\cdot, \cdot)$ est continue. Ainsi, pour $u \in V$ fixé, il existe un unique $\mathcal{A}u \in V$ tel que $a(u, v) = (\mathcal{A}u, v)_V$. L'application \mathcal{A} ainsi définie est linéaire de V dans V et continue car $a(\cdot, \cdot)$ est continue. Le problème (8.1.1) s'écrit alors : trouver $u \in V$ tel que $(\mathcal{A}u, v)_V = (f, v)_V$ pour tout $v \in V$. Soit encore : trouver $u \in V$ tel que $\mathcal{A}u = f$. Autrement dit, montrer l'existence et l'unicité d'une solution de (8.1.1) est équivalent à montrer que l'opérateur \mathcal{A} est bijectif.

Pour montrer l'injectivité, il suffit de montrer que $\mathcal{A}u = 0$ entraîne u = 0. D'après la coercivité de $a(\cdot, \cdot)$, on a $(\mathcal{A}u, u)_V \ge m ||u||_V^2$ dont on déduit immédiatement l'injectivité.

Pour montrer la surjectivité, on montre que l'image $\text{Im}(\mathcal{A})$ de \mathcal{A} est fermée et que $\text{Im}(\mathcal{A})^{\perp} = \{0\}$. Soit $(v_n)_n$ une suite de $\text{Im}(\mathcal{A})$ qui converge dans V vers v. Alors $v_n = \mathcal{A}u_n$, où $(u_n)_n$ est une suite de V. On a

$$\|\mathcal{A}u_n - \mathcal{A}u_m\|_V \|u_n - u_m\|_V \ge (\mathcal{A}u_n - \mathcal{A}u_m, u_n - u_m)_V \ge m \|u_n - u_m\|_V^2$$

donc $||\mathcal{A}u_n - \mathcal{A}u_m||_V \ge m ||u_n - u_m||_V$. La suite $(\mathcal{A}u_n)_n$ étant de Cauchy, on en déduit que la suite $(u_n)_n$ est de Cauchy dans V qui est complet donc $(u_n)_n$ converge vers $u \in V$. Comme \mathcal{A} est continue, on en déduit $v = \mathcal{A}u \in \text{Im}(\mathcal{A})$ donc $\text{Im}(\mathcal{A})$ est fermée. Soit $v \in \text{Im}(\mathcal{A})^{\perp} = \{v \in V \mid (u, v) = 0, \forall u \in \text{Im}(\mathcal{A})\}$. Alors on a $0 = (\mathcal{A}v, v)_V \ge m ||v||_V$ donc v = 0, d'où $\text{Im}(\mathcal{A})^{\perp} = \{0\}$. On en déduit

$$\operatorname{Im}(\mathcal{A}) = \overline{\operatorname{Im}(\mathcal{A})} = (\operatorname{Im}(\mathcal{A})^{\perp})^{\perp} = \{0\}^{\perp} = V,$$

d'où \mathcal{A} est surjective.

Lorsque la forme bilinéaire $a(\cdot, \cdot)$ est de plus symétrique, *i.e.* a(u, v) = a(v, u), il existe une autre formulation du théorème de Lax-Milgram comme minimisation d'une énergie.

Proposition 8.1.3. On se place sous les hypothèses du théorème 8.1.2 de Lax-Milgram. On note J(v) l'énergie définie pour $v \in V$ par

$$J(v) := \frac{1}{2} a(v, v) - \langle L, v \rangle.$$
 (8.1.2)

Si $a(\cdot, \cdot)$ est symétrique, alors l'unique solution $u \in V$ de la formulation variationnelle (8.1.1) vérifie

$$J(u) = \min_{v \in V} J(v).$$
 (8.1.3)

Réciproquement, si u réalise le minimum de J sur V (i.e. u vérifie (8.1.3)), alors u est l'unique solution de (8.1.1).

Remarque 8.1.4. On dit que la formulation variationnelle (8.1.1) est l'équation d'Euler associée au problème de minimisation (8.1.3).

Démonstration. Si $u \in V$ est la solution de la formulation variationnelle (8.1.1), puisque $a(\cdot, \cdot)$ est symétrique, on a pour tout $v \in V$

$$J(u+v) = \frac{1}{2}a(u+v,u+v) - \langle L,u+v \rangle$$

= $\frac{1}{2}(a(u,u) + 2a(u,v) + a(v,v)) - \langle L,u \rangle - \langle L,v \rangle$
= $J(u) + \frac{1}{2}a(v,v) \ge J(u).$

Donc pour tout $v \in V$, $J(v) \ge J(u)$ d'où u vérifie (8.1.3).

Réciproquement, si $u \in V$ vérifie (8.1.3). Soit $v \in V$ fixé, on définie $j : \mathbb{R} \to \mathbb{R}$ par j(t) := J(u + tv). Alors $j(t) \ge j(0)$ pour tout $t \in \mathbb{R}$ et on a

$$j(t) = \frac{t^2}{2} a(u, v) + t(a(u, v) - \langle L, v \rangle) + J(u).$$

On en déduit que j est dérivable et, comme 0 est le minimum de j, j'(0) = 0. Alors on obtient $a(u, v) - \langle L, v \rangle = j'(0) = 0$ donc u est solution de la formulation variation-nelle (8.1.1).

8.2 Problème de Dirichlet

Soit Ω un ouvert borné de \mathbb{R}^d . Dans un premier temps, on considère le problème de Dirichlet homogène

$$\begin{cases}
-\Delta u = f \quad \text{dans } \Omega, \\
u = 0 \quad \text{sur } \partial\Omega,
\end{cases}$$
(8.2.1)

où $f \in L^2(\Omega)$. On a vu à la fin du chapitre 3 qu'une formulation variationnelle du problème de Dirichlet est : trouver $u \in H^1(\Omega)$ avec u = 0 sur $\partial\Omega$ telle que pour toute $v \in H^1(\Omega)$ avec v = 0 sur $\partial\Omega$ on a

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

On suppose que Ω est régulier de classe C^1 . Alors, $H_0^1(\Omega) = \{u \in H^1(\Omega) \mid u = 0 \text{ sur } \partial\Omega\}$. On obtient ainsi la formulation variationnelle suivante pour le problème de Dirichlet homogène

$$\begin{cases} \text{Trouver } u \in H_0^1(\Omega) \text{ telle que} \\ \forall v \in H_0^1(\Omega), \quad \int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx. \end{cases}$$
(8.2.2)

Proposition 8.2.1. Soient Ω un ouvert borné régulier de classe C^1 de \mathbb{R}^d et $f \in C(\Omega)$. Si $u \in C^2(\overline{\Omega})$ est solution classique de (8.2.1) alors u est solution de (8.2.2). Réciproquement, si u est solution de (8.2.2) et $u \in C^2(\overline{\Omega})$ alors u est solution classique de (8.2.1).

Remarque 8.2.2.

- 1. La Proposition 8.2.1 montre que la formulation variationnelle (8.2.2) définit bien une notion de solution faible de l'équation (8.2.1).
- 2. On peut montrer le résultat plus précis suivant (voir [1]) : si $f \in L^2(\Omega)$ et u est solution de la formulation variationnelle (8.2.2) alors u est solution de (8.2.1) au sens où $-\Delta u = f$ p.p. dans Ω et u = 0 p.p. sur $\partial \Omega$ (en particulier, la distributions Δu est dans $L^2(\Omega)$).

Démonstration. Par "construction" de la formulation variationnelle, il est clair que si u est solution classique de (8.2.1) alors u est solution de la formulation variationnelle (8.2.2). Réciproquement, soit $u \in C^2(\overline{\Omega}) \cap H_0^1(\Omega)$ telle que

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

Soit $v \in C_c^1(\Omega) \subset H_0^1(\Omega)$. Par application de la formule de Green, on obtient

$$\int_{\Omega} (\Delta u + f) v \, dx = \int_{\partial \Omega} \frac{\partial u}{\partial \nu} v \, d\sigma_x = 0.$$

D'après le Lemme 6.4.3, on en déduit $\Delta u(x) + f(x) = 0$ pour tout $x \in \Omega$. De plus, comme $u \in H_0^1(\Omega), u = 0$ sur $\partial \Omega$

On pose :

$$\forall u, v \in H_0^1(\Omega), \quad a(u, v) := \int_{\Omega} \nabla u \cdot \nabla v \, dx, \qquad (8.2.3)$$

 et

$$\forall v \in H_0^1(\Omega), \quad \langle L, v \rangle := \int_{\Omega} f v \, dx. \tag{8.2.4}$$

Proposition 8.2.3. Soit Ω est un ouvert borné régulier de classe C^1 .

- 1. La forme bilinéaire $a(\cdot, \cdot)$ définie sur $H_0^1(\Omega)$ par (8.2.3) est coercive et continue.
- 2. La forme linéaire L définie sur $H_0^1(\Omega)$ par (8.2.4) est continue (i.e. $L \in H^{-1}(\Omega)$).

Démonstration.

1. D'après l'inégalité de Cauchy-Schwarz et le Corollaire 7.2.6, on a

$$|a(u,v)| \le ||\nabla u||_{L^2(\Omega)^d} \, ||\nabla v||_{L^2(\Omega)^d} = ||u||_{H^1_0(\Omega)} \, ||v||_{H^1_0(\Omega)}, \quad \forall \ u,v \in H^1_0(\Omega),$$

donc $a(\cdot, \cdot)$ est continue sur $H_0^1(\Omega)$. De plus, on a

$$a(u, u) = \int_{\Omega} |\nabla u|^2 \, dx = ||u||^2_{H^1_0(\Omega)}, \quad \forall \ u \in H^1_0(\Omega),$$

donc $a(\cdot, \cdot)$ est coercive sur $H_0^1(\Omega)$.

2. En appliquant l'inégalité de Cauchy-Schwarz puis l'inégalité de Poincaré, on obtient qu'il existe une constante c > 0 telle que

$$|\langle L, v \rangle| = \left| \int_{\Omega} f v \, dx \right| \le ||f||_{L^{2}(\Omega)} \, ||v||_{L^{2}(\Omega)} \le c \, ||f||_{L^{2}(\Omega)} \, ||v||_{H^{1}_{0}(\Omega)}, \quad \forall \, v \in H^{1}_{0}(\Omega),$$

donc L est continue sur $H_0^1(\Omega)$.

De la Proposition 8.2.3, du théorème de Lax-Milgram 8.1.2 et de la Proposition 8.1.3, on déduit le résultat suivant :

Corollaire 8.2.4. Soit Ω est un ouvert borné régulier de classe C^1 . Alors, la formulation variationnelle (8.2.2) admet une unique solution $u \in H_0^1(\Omega)$. De plus, u réalise le minimum dans $H_0^1(\Omega)$ de l'énergie J définie par

$$J(v) := \frac{1}{2} \int_{\Omega} \nabla v \cdot \nabla v \, dx - \int_{\Omega} f v \, dx.$$
(8.2.5)

Remarque 8.2.5. On rappelle que le second membre f considéré est, par hypothèse, dans $L^2(\Omega)$. Néanmoins, le Corollaire 8.2.4 est encore vrai en supposant seulement $f \in H^{-1}(\Omega)$ puisque, par définition, f est alors une forme linéaire continue sur $H_0^1(\Omega)$. Par contre, dans ce cas, la formulation variationnelle (8.2.2) s'écrit

$$\begin{cases} \text{Trouver } u \in H_0^1(\Omega) \text{ telle que} \\ \forall v \in H_0^1(\Omega), \quad \int_{\Omega} \nabla u \cdot \nabla v \, dx = \langle f, v \rangle_{H^{-1}, H_0^1}. \end{cases}$$

$$(8.2.6)$$

Corollaire 8.2.6. L'application linéaire qui à $f \in H^{-1}(\Omega)$ fait correspondre l'unique solution $u \in H_0^1(\Omega)$ de la formulation variationnelle (8.2.6) est continue. En particulier, on a l'inégalité

$$\forall f \in H^{-1}(\Omega), \quad ||u||_{H^{1}_{0}(\Omega)} \le ||f||_{H^{-1}(\Omega)}.$$
(8.2.7)

Démonstration. On note T l'opérateur qui à $f \in H^{-1}(\Omega)$ fait correspondre l'unique solution $u \in H^1_0(\Omega)$ de la formulation variationnelle (8.2.6). Alors, en prenant u comme fonction test dans (8.2.6), on obtient

$$||Tf||_{H_0^1(\Omega)}^2 = ||u||_{H_0^1(\Omega)}^2 = \int_{\Omega} \nabla u \cdot \nabla u \, dx = \langle f, u \rangle_{H^{-1}, H_0^1} \le ||f||_{H^{-1}(\Omega)} \, ||u||_{H_0^1(\Omega)} = \int_{\Omega} ||u||_$$

ce qui entraîne l'inégalité (8.2.7) et la continuité de T.

Remarque 8.2.7.

- 1. On a vu dans la démonstration précédente que l'inégalité (8.2.7) s'obtient en prenant *u* comme fonction test dans la formulation variationnelle (8.2.6). Une inégalité obtenue suivant cette méthode es appelée **estimation à priori** ou encore **estimation d'énergie**. Ce type d'inégalités est fréquemment utilisé dans l'analyse des e.d.p..
- 2. D'apres les Corollaires 8.2.4 et 8.2.6, on en déduit que le problème variationnel (8.2.6) est bien posé au sens d'Hadamard.

On a vu dans le chapitre 3 que le problème de Dirichlet vérifie un principe du maximum s'il admet une solution u régulière. Ci-dessous on montre que cela est encore le cas pour une solution faible.

Théorème 8.2.8 (Principe du maximum). Soient Ω un ouvert borné régulier de classe C^1 , $f \in L^2(\Omega)$ et $u \in H^1_0(\Omega)$ la solution de (8.2.1). Si $f \ge 0$ p.p. dans Ω , alors $u \ge 0$ p.p. dans Ω .

Démonstration. On admet le résultat suivant (voir par exemple [1] ou [5]).

Lemme 8.2.9. Si $v \in H_0^1(\Omega)$ alors $v^- := \min(v, 0) \in H_0^1(\Omega)$ et $\nabla v^- = \chi_{v<0} \nabla v$ p.p. dans Ω , où $\chi_{v<0}$ désigne la fonction caractéristique de $\{x \in \Omega \mid v(x) < 0\}$.

On prend $v = u^{-}$ comme fonction test dans la formulation variationnelle (8.2.2). On obtient

$$\int_{\Omega} f u^- dx = \int_{\Omega} \nabla u \cdot \nabla u^- dx = \int_{\Omega} \chi_{u<0} |\nabla u|^2 dx = \int_{\Omega} |\nabla u^-|^2 dx \ge 0.$$

Or $u^- \leq 0$ et $f \geq 0$ p.p. dans Ω , donc le membre de gauche de l'égalité précédente est négatif ou nul. On en déduit $||u^-||_{H_0^1(\Omega)} = 0$ donc $u^- = 0$ p.p. dans Ω car $u^- \in H_0^1(\Omega)$ et $|| \cdot ||_{H_0^1(\Omega)}$ est une norme sur $H_0^1(\Omega)$. Par définition de u^- on obtient le résultat. \square

Une propriété remarquable et très importante des équations elliptiques est que si les données (terme source et terme de bord) sont suffisamment régulières alors la solution faible gagne elle-même en régularité, c'est ce qu'on appelle la régularité elliptique.

Théorème 8.2.10 (Régularité elliptique). Soit $m \in \mathbb{N}^*$, Ω un ouvert borné de \mathbb{R}^d de classe C^{m+2} et $f \in H^m(\Omega)$. Alors, l'unique solution $u \in H_0^1(\Omega)$ de (8.2.1) appartient à $H^{m+2}(\Omega)$. De plus, l'application qui à f associe u est linéaire continue de $H^m(\Omega)$ dans $H^{m+2}(\Omega)$.

Remarque 8.2.11.

- 1. On admet ici ce résultat dont la démonstration dépasse de loin le cadre de ce cours. Celle-ci se fait en trois étapes en considérant tout d'abord le cas $\Omega = \mathbb{R}^d$ puis le cas $\Omega = \mathbb{R}^d_+$ et enfin le cas Ω de classe C^{m+2} pour lequel on applique les deux précédentes étapes via des cartes locales. La méthode employée est due à Niremberg, on renvoie à [1] et [5] pour plus de détails.
- 2. En dimension 1, le résultat est évident. En effet, le problème s'écrit alors -u'' = fdans]0,1[(par exemple), u(0) = u(1) = 0. Si $f \in H^1(\Omega)$, la solution $u \in H_0^1(0,1)$ vérifie $u'' = -f \in H^1(\Omega)$ donc $u \in H^3(0,1)$. En dimension supérieure, le même raisonnement permet seulement d'obtenir $\Delta u \in L^2(\Omega)$ ce qui n'entraîne pas à priori $u \in H^3(\Omega)$.

On rappelle que, d'après le Théorème 7.1.5, $H^m(\Omega)$ est un sous-espace de $C(\overline{\Omega})$ dès que $m > \frac{d}{2}$. En corollaire de ce résultat et de la régularité elliptique on obtient

Corollaire 8.2.12. Soient $m \in \mathbb{N}^*$, Ω un ouvert borné de \mathbb{R}^d de classe C^{m+2} et $f \in H^m(\Omega)$. Si $m > \frac{d}{2}$, la solution faible $u \in H^1_0(\Omega)$ de (8.2.1) est solution forte de (8.2.1) car $u \in C^2(\overline{\Omega})$. En particulier, si Ω est de classe C^{∞} et $f \in C^{\infty}(\overline{\Omega})$ alors $u \in C^{\infty}(\overline{\Omega})$.

Considérons maintenant le problème de Dirichlet non-homogène (où Ω est un ouvert borné régulier de classe C^1)

$$\begin{cases} -\Delta u = f \quad \text{dans } \Omega, \\ u = g \quad \text{sur } \partial \Omega, \end{cases}$$
(8.2.8)

où $f \in H^{-1}(\Omega)$. Pour qu'il existe une solution $u \in H^1(\Omega)$ au problème (8.2.8), il faut que la donnée au bord g soit la trace sur $\partial\Omega$ d'une fonction de $H^1(\Omega)$ d'après le théorème de trace (7.3.3), *i.e.* $g = \gamma_0 h$ où $h \in H^1(\Omega)$. Autrement dit $g \in H^{\frac{1}{2}}(\partial\Omega)$. On dit que l'application h est un **relèvement** de la condition aux limites.

Pour obtenir une formulation variationnelle de (8.2.8) et en déduire, notamment, un résultat d'existence et d'unicité on pose $\tilde{u} := u - h$. De sorte que la solution de (8.2.8) (au sens des distributions) est donnée par $u = \tilde{u} + h$. On obtient alors que \tilde{u} vérifie le problème de Dirichlet homogène

$$\begin{cases} -\Delta \tilde{u} = \tilde{f} \quad \text{dans } \Omega, \\ \tilde{u} = 0 \quad \text{sur } \partial \Omega, \end{cases}$$
(8.2.9)

où $\tilde{f} := f + \Delta h$. Or $h \in H^1(\Omega)$ donc ses dérivées sont des éléments de $L^2(\Omega)$, on en déduit que Δh s'écrit comme une somme de dérivées (au sens des distributions) de fonctions de $L^2(\Omega)$. D'après la caractérisation (7.5.1), on a alors $\Delta h \in H^{-1}(\Omega)$. Donc $\tilde{f} \in H^{-1}(\Omega)$, et, d'après le Corollaire 8.2.4, $\tilde{u} \in H_0^1(\Omega)$ est l'unique solution du problème variationnel

$$\begin{cases} \text{Trouver } \tilde{u} \in H_0^1(\Omega) \text{ telle que} \\ \forall v \in H_0^1(\Omega), \quad \int_{\Omega} \nabla \tilde{u} \cdot \nabla v \, dx = \left\langle \tilde{f}, v \right\rangle_{H^{-1}, H_0^1}. \end{cases}$$
(8.2.10)

Autrement dit, on a montré le résultat suivant :

Théorème 8.2.13. Soient Ω un ouvert borné régulier de classe C^1 de \mathbb{R}^d , $f \in H^{-1}(\Omega)$ et $g \in H^{\frac{1}{2}}(\partial\Omega)$. Alors, il existe une unique solution faible $u \in H^1(\Omega)$ au problème (8.2.8) (i.e. solution au sens des distributions). De plus, $u = \tilde{u} + h$ où $h \in H^1(\Omega)$ est un relèvement de g et \tilde{u} est l'unique solution de la formulation variationnelle (8.2.10).

Remarque 8.2.14.

- 1. Le relèvement h n'est pas unique. En effet, si on note h_1 et h_2 les solutions de (8.2.8) pour deux termes sources différents f_1 et f_2 alors h_1 et h_2 sont tout deux des relèvements de g.
- 2. L'unicité de la solution faible $u \in H^1(\Omega)$ du problème (8.2.8) n'est pas immédiate puisque le choix du relèvement h n'est pas unique. Celle-ci a tout de même bien lieu car le choix de h conditionne le choix de \tilde{u} . Pour montrer l'unicité, il suffit de choisir deux solutions faibles u_1 et u_2 de (8.2.8). On obtient alors que $u_1 - u_2$ est solution faible de (8.2.1) avec f = 0. Or l'unique solution de ce problème est u = 0donc $u_1 = u_2$ p.p. dans Ω .

8.3 Équation de Laplace avec conditions aux limites de Neumann

Dans un premier temps, on considère le problème suivant :

$$\begin{cases} -\Delta u + u = f \quad \text{dans } \Omega, \\ \frac{\partial u}{\partial \nu} = g \quad \text{sur } \partial \Omega, \end{cases}$$
(8.3.1)

où Ω est un ouvert borné régulier de classe C^1 , $f \in L^2(\Omega)$ et $g \in L^2(\partial \Omega)$.

Remarque 8.3.1. D'après le théorème de trace 7.3.9, il serait naturel de considérer $g \in L^2(\partial\Omega)$ telle que $g = \gamma_1 h$, où $h \in H^2(\Omega)$. En fait, on va justifier ci-dessous qu'il suffit d'avoir $g \in L^2(\partial\Omega)$ pour obtenir une formulation variationnelle de (8.3.1).

On reprend la méthode employée dans le cas du problème de Dirichlet qui consiste à considérer qu'il existe une solution $u \in C^2(\overline{\Omega})$ de (8.3.1). Soit $\varphi \in C^1(\overline{\Omega})$. En multipliant (8.3.1) par φ et on intégrant par parties, on obtient

$$\int_{\Omega} \nabla u \cdot \nabla \varphi \, dx - \int_{\partial \Omega} \frac{\partial u}{\partial \nu} \varphi \, d\sigma_x + \int_{\Omega} u \, \varphi \, dx = \int_{\Omega} f \, \varphi \, dx,$$

d'où

$$\int_{\Omega} \nabla u \cdot \nabla \varphi \, dx + \int_{\Omega} u \, \varphi \, dx = \int_{\Omega} f \, \varphi \, dx + \int_{\partial \Omega} g \, \varphi \, d\sigma_x.$$

On en déduit la formulation variationnelle suivante :

$$\begin{cases} \text{Trouver } u \in H^1(\Omega) \text{ telle que} \\ \forall v \in H^1(\Omega), \quad \int_{\Omega} \nabla u \cdot \nabla v + u \, v \, dx = \int_{\Omega} f \, v \, dx + \int_{\partial \Omega} g \, v \, d\sigma_x. \end{cases}$$
(8.3.2)

Remarque 8.3.2.

- 1. Dans l'intégrale de bord, v est la trace $\gamma_0 v \operatorname{sur} \partial \Omega$ (qui est bien définie car $v \in H^1(\Omega)$). De plus, cette intégrale a bien un sens pour $g \in L^2(\partial \Omega)$ sans nécessairement que g soit la trace γ_1 d'une fonction de $H^2(\Omega)$.
- 2. La formulation variationnelle (8.3.2) ne fait pas intervenir la dérivée normale de u. C'est ce qui justifie de chercher u seulement dans $H^1(\Omega)$ et non dans $H^2(\Omega)$. La grande différence (au sens variationnel) entre condition de Dirichlet et condition de Neumann réside dans le fait que la condition de Dirichlet est contenue dans l'espace des fonctions test considérées alors que la condition de Neumann est contenue dans la formulation variationnelle. On dit que la condition de Dirichlet est **essentielle** (ou **explicite**) alors que la condition de Neumann est dite **naturelle** (ou **implicite**).

Comme dans le cas du problème de Dirichlet, il faut s'assurer qu'une solution de la formulation variationnelle (8.3.2) définie bien une solution faible de l'équation (8.3.1).

Proposition 8.3.3. Soient Ω un ouvert borné régulier de classe C^1 de \mathbb{R}^d , $f \in C(\Omega)$ et $g \in C(\partial\Omega)$. Si $u \in C^2(\overline{\Omega})$ est solution classique de (8.3.1), alors u est solution de (8.3.2). Réciproquement, si u est solution de (8.3.2) et $u \in C^2(\overline{\Omega})$ alors u est solution classique de (8.3.1).

Démonstration. Il suffit de montrer la réciproque. Soit $u \in C^2(\overline{\Omega})$ solution de (8.3.2). On prend $v \in C^1(\overline{\Omega})$ comme fonction test. Par intégration par parties, on obtient

$$\int_{\Omega} (-\Delta u + u - f) v \, dx = \int_{\partial \Omega} \left(g - \frac{\partial u}{\partial \nu} \right) v \, d\sigma_x$$

En particulier, on a pour toute $v \in C_c^1(\Omega)$, $\int_{\Omega} (-\Delta u + u - f)v \, dx = 0$. Alors, d'après le Lemme 6.4.3, $-\Delta u + u - f = 0$ dans Ω . On en déduit

$$\forall v \in C^1(\overline{\Omega}), \quad \int_{\partial\Omega} \left(g - \frac{\partial u}{\partial\nu}\right) v \ d\sigma_x = 0.$$

De même qu'avec le Lemme 6.4.3, on en déduit $g - \frac{\partial u}{\partial \nu} = 0$.

Théorème 8.3.4. Soient Ω un ouvert borné régulier de classe C^1 , $f \in L^2(\Omega)$ et $g \in L^2(\partial \Omega)$. Alors, la formulation variationnelle (8.3.2) admet une unique solution $u \in H^1(\Omega)$. De plus, u réalise le minimum dans $H^1(\Omega)$ de l'énergie J définie par

$$J(v) := \frac{1}{2} \int_{\Omega} |\nabla v|^2 + |v|^2 \, dx - \int_{\Omega} f \, v \, dx - \int_{\partial \Omega} g \, v \, d\sigma_x.$$
(8.3.3)

Démonstration. Il suffit d'appliquer le théorème de Lax-Milgram. Pour cela on définit l'application bilinéaire $a(\cdot, \cdot)$ sur $H^1(\Omega)$ par

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

et la forme linéaire L sur $H^1(\Omega)$ par

$$\langle L, v \rangle := \int_{\Omega} f v \, dx + \int_{\partial \Omega} g v \, d\sigma_x.$$

Alors $a(u, u) = ||u||^2_{H^1(\Omega)}$ donc $a(\cdot, \cdot)$ est bien coercive. Par l'inégalité de Cauchy-Schwarz, on a

$$|a(u,v)| \le ||\nabla u||_{L^{2}(\Omega)} ||\nabla v||_{L^{2}(\Omega)} + ||u||_{L^{2}(\Omega)} ||v||_{L^{2}(\Omega)} \le 2||u||_{H^{1}(\Omega)} ||v||_{H^{1}(\Omega)},$$

donc $a(\cdot, \cdot)$ est continue. Pour $v \in H^1(\Omega)$, d'après l'inégalité de Cauchy-Schwarz et le théorème de trace 7.3.3, on obtient qu'il existe une constante c > 0 indépendante de v telle que

$$\begin{aligned} \langle L, v \rangle \,| &\leq ||f||_{L^{2}(\Omega)} \,||v||_{L^{2}(\Omega)} + ||g||_{L^{2}(\partial\Omega)} \,||v||_{L^{2}(\partial\Omega)} \\ &\leq ||f||_{L^{2}(\Omega)} \,||v||_{H^{1}(\Omega)} + c \,||g||_{L^{2}(\partial\Omega)} \,||v||_{H^{1}(\Omega)}, \end{aligned}$$

donc L est continue sur $H^1(\Omega)$.

On considère maintenant l'équation de Laplace avec la condition aux limites de Neumann

$$\begin{cases} -\Delta u = f \quad \text{dans } \Omega, \\ \frac{\partial u}{\partial \nu} = g \quad \text{sur } \partial \Omega, \end{cases}$$
(8.3.4)

où Ω est un ouvert borné régulier de classe C^1 , $f \in L^2(\Omega)$ et $g \in L^2(\partial\Omega)$. La difficulté supplémentaire par rapport au problème (8.3.1) vient du fait que l'équivalent de la forme bilinéaire $a(\cdot, \cdot)$ ne contiendra plus de terme en u et donc celle-ci n'est plus coercive sur $H^1(\Omega)$.

Il faut remarquer que le problème (8.3.4) est, pour l'instant, mal posé. Il est nécessaire d'ajouter une condition dite de compatibilité reliant f et g. En effet, supposons que u est une solution régulière de (8.3.4), alors on obtient

$$\int_{\Omega} f \, dx = -\int_{\Omega} \Delta u \, dx = -\int_{\partial \Omega} \frac{\partial u}{\partial \nu} \, d\sigma_x = -\int_{\partial \Omega} g \, d\sigma_x.$$

Il faut donc ajouter la condition de compatibilité

$$\int_{\Omega} f \, dx + \int_{\partial \Omega} g \, d\sigma_x = 0. \tag{8.3.5}$$

De plus, le problème (8.3.4) ne fait intervenir que les dérivées de u et donc la solution est définie, si l'ouvert Ω est supposé connexe, à une constante additive près (*i.e.* si u est solution alors u + c est solution pour toute constante c). Pour avoir unicité de la solution, on peut, par exemple, fixer la moyenne de u sur Ω :

$$\int_{\Omega} u \, dx = 0. \tag{8.3.6}$$

L'espace fonctionnel naturel à considérer est alors

$$V(\Omega) := \left\{ v \in H^1(\Omega) \mid \int_{\Omega} v \, dx = 0 \right\}, \tag{8.3.7}$$

qui est un espace de Hilbert pour le norme de H^1 (car fermé). Par les mêmes arguments que pour le problème précédent, on aboutit à la formulation variationnelle

$$\begin{cases} \text{Trouver } u \in V(\Omega) \text{ telle que} \\ \forall v \in V(\Omega), \quad \int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f \, v \, dx + \int_{\partial \Omega} g \, v \, d\sigma_x. \end{cases}$$
(8.3.8)

Par le même raisonnement que précédemment on obtient le résultat suivant

Proposition 8.3.5. Soient Ω un ouvert borné connexe régulier de classe C^1 de \mathbb{R}^d , $f \in C(\Omega)$ et $g \in C(\partial\Omega)$ tels que la relation de compatibilité (8.3.5) est vérifiée. Si $u \in C^2(\overline{\Omega})$ est solution classique de (8.3.4) alors u est solution de (8.3.8). Réciproquement, si u est solution de (8.3.8) et $u \in C^2(\overline{\Omega})$ alors u est solution classique de (8.3.4).

Théorème 8.3.6. Soient Ω un ouvert borné connexe régulier de classe C^1 de \mathbb{R}^d , $f \in L^2(\Omega)$ et $g \in L^2(\partial \Omega)$ tels que la relation de compatibilité (8.3.5) est vérifiée. Alors, il existe une unique solution $u \in V(\Omega)$ de (8.3.8). De plus, u réalise le minimum dans $V(\Omega)$ de l'énergie J définie par

$$J(v) := \frac{1}{2} \int_{\Omega} \nabla v \cdot \nabla v \, dx - \int_{\Omega} f v \, dx - \int_{\partial \Omega} g v \, d\sigma_x.$$

Remarque 8.3.7. Comme dans le cas du problème de Dirichlet, on a des résultats de régularité elliptique (voir [5]).

Exercice 8.3.8. Montrer le Théorème 8.3.6 en utilisant l'inégalité de Poincaré-Wirtinger (Proposition 7.4.4).

8.4 Problèmes elliptiques sous forme divergence

Les problèmes elliptiques sous forme divergence sont une généralisation de l'équation de Laplace. Par exemple, l'équation de Laplace permet de modéliser le problème stationnaire de diffusion de la chaleur dans un matériau homogène isotrope. Dès lors qu'on ne suppose plus le matériau isotrope ou homogène le problème ne peut plus s'écrire sous forme d'un Laplacien.

Dans le cas d'un matériau **anisotrope** (autrement dit la chaleur se diffuse suivant des directions privilégiées) et homogène le problème s'écrit

$$-\operatorname{div}(A\nabla u) = f,$$

où $A \in \mathbb{R}^{d \times d}$. Par exemple, en dimension 2, si on suppose que la chaleur se diffuse deux fois plus vite dans la direction x_2 par rapport à la direction x_1 alors A est de la forme

$$A = \alpha \left(\begin{array}{cc} 1 & 0 \\ 0 & 2 \end{array} \right),$$

où α est le conductivité dans la direction x_1 . Si le matériau n'est plus homogène alors A = A(x), autrement dit la conductivité dépend du point x considéré. Alors la forme générale d'un problème elliptique (avec condition de Dirichlet homogène) sous forme divergente est la suivante

$$\begin{cases} -\operatorname{div}(A\nabla u) = f \quad \operatorname{dans} \Omega, \\ u = 0 \quad \operatorname{sur} \partial\Omega, \end{cases}$$
(8.4.1)

où Ω est un ouvert borné de \mathbb{R}^d , $f \in H^{-1}(\Omega)$ et $A : \Omega \to \mathbb{R}^{d \times d}$.

Pour avoir existence et unicité d'une solution faible de (8.4.1), on suppose que la fonction A vérifie :

1. A est coercive, *i.e.* il existe $\alpha > 0$ telle que pour tout $\xi \in \mathbb{R}^d$,

$$A(x)\xi \cdot \xi \ge \alpha |\xi|^2 \quad \text{p.p. dans } \Omega, \tag{8.4.2}$$

2. A est bornée, *i.e.* il existe $\beta > 0$ telle que pour tout $\xi \in \mathbb{R}^d$,

$$|A(x)\xi| \le \beta |\xi| \quad \text{p.p. dans } \Omega. \tag{8.4.3}$$

Par application du théorème de Lax-Milgram, on a

Théorème 8.4.1. Soient Ω un ouvert borné de \mathbb{R}^d de classe C^1 , $f \in L^2(\Omega)$ et $A : \Omega \to \mathbb{R}^{d \times d}$ coercive et bornée. Alors, il existe une unique solution faible $u \in H_0^1(\Omega)$ de (8.4.1), *i.e. une solution de la formulation variationnelle*

$$\begin{cases} Trouver \ u \in H_0^1(\Omega) \ telle \ que \\ \forall \ v \in H_0^1(\Omega), \quad \int_\Omega A \nabla u \cdot \nabla v \ dx = \int_\Omega f \ v \ dx. \end{cases}$$
(8.4.4)

Exercice 8.4.2. Montrer le Théorème 8.4.1.

Un cas particulier intéressant est celui où A s'écrit $A(x) = \alpha(x) I_d$ avec $\alpha(x)$ prenant deux valeurs distinctes. On considère Ω_1 et Ω_2 deux sous-domaines de Ω (*i.e* Ω_1 et Ω_2 sont deux ouverts bornés et connexe) formant une partition de Ω , donc $\Omega = \Omega_1 \cup \Omega_2$. On suppose que α est constante dans Ω_1 et Ω_2 prenant deux valeurs différentes :

$$\alpha(x) = \begin{cases} a & \text{dans } \Omega_1, \\ b & \text{dans } \Omega_2, \end{cases}$$
(8.4.5)

où a, b > 0. Dans ce cas, le problème (8.4.1) s'écrit

$$\begin{cases} -\operatorname{div}(\alpha(x)\nabla u) = f & \operatorname{dans} \Omega, \\ u &= 0 & \operatorname{sur} \partial\Omega, \end{cases}$$
(8.4.6)

où on a écrit $\alpha(x)$ au lieu de α pour préciser la dépendance en la variable d'espace. Le problème (8.4.6) modélise, par exemple, la répartition de la chaleur pour un milieu Ω composé de deux matériaux homogènes et isotropes Ω_1 et Ω_2 de conductivité a et b. Si α était constant le problème se réduirait simplement au problème de Dirichlet classique, le fait que α prenne des valeurs différentes dans Ω_1 et Ω_2 entraîne le résultat suivant :



FIGURE 8.1 – Décomposition de Ω en Ω_1 et Ω_2

Proposition 8.4.3. Soient Ω un ouvert borné connexe de \mathbb{R}^d de classe C^1 et $f \in L^2(\Omega)$. Soient Ω_2 un ouvert borné connexe régulier de classe C^1 de Ω , $\Omega_1 := \Omega \setminus \Omega_2$ et $\Gamma := \partial \Omega_2$. Soit α donnée par (8.4.5). Alors, le problème (8.4.6) est équivalent au **problème de transmission**

$$-a\Delta u_{1} = f \qquad dans \ \Omega_{1},$$

$$-b\Delta u_{2} = f \qquad dans \ \Omega_{2},$$

$$u_{1} = u_{2} \qquad sur \ \Gamma,$$

$$a\nabla u_{1} \cdot \nu = b\nabla u_{2} \cdot \nu \quad sur \ \Gamma,$$

$$u_{1} = 0 \qquad sur \ \partial\Omega,$$

(8.4.7)

où u_i est la restriction de la solution u à Ω_i . Les conditions sur Γ sont appelées conditions aux limites de transmission à l'interface Γ .

Remarque 8.4.4.

- 1. En pratique, en modélisation, on obtient plutôt le problème (8.4.7). Le point important de la Proposition 8.4.3 est l'équivalence entre celui-ci et le problème (8.4.6) qui est plus simple à traiter et dont on a montré (Théorème 8.4.1) qu'il admet une solution faible unique.
- 2. On peut aussi considérer des conditions aux limites d'un autre type sur $\partial \Omega$ (Dirichlet non-homogène, Neumann, mixte...etc).
- 3. Les résultats de régularité elliptique sont encore vrais pour les problèmes elliptiques sous forme divergente (voir [8]) en supposant une certaine régularité de la fonction A.

Exercice 8.4.5. Montrer la Proposition 8.4.3.

Chapitre 9 Approximation variationnelle

Dans ce (court) chapitre, on décrit la méthode générale d'approximation d'une formulation variationnelle définie dans un espace de Hilbert qui servira de base pour l'élaboration de la méthode des éléments finis.

9.1 Approximation interne et système matriciel équivalent

Soit V un espace de Hilbert réel muni du produit scalaire $(\cdot, \cdot)_V$ et de la norme associée $||\cdot||_V$. Soient $a(\cdot, \cdot)$ une forme bilinéaire continue et coercive sur V et $f \in V$. On considère la formulation variationnelle générale

$$\begin{cases} \text{Trouver } u \in V \text{ tel que :} \\ \forall v \in V, \quad a(u,v) = (f,v)_V. \end{cases}$$
(9.1.1)

L'existence et l'unicité de la solution $u \in V$ de (9.1.1) est assurée par le théorème de Lax-Milgram. Dans cette section, on cherche à déterminer une suite $(u_h)_h$ de V telle que $||u - u_h||_V$ tend vers 0 quand h tend vers 0.

L'approximation interne consiste à considérer une suite V_h de sous-espaces fermés de V de dimension finie (qui sont des espaces de Hilbert car fermés dans V). On s'intéresse alors aux problèmes approchés :

$$\begin{cases} \text{Trouver } u_h \in V_h \text{ tel que :} \\ \forall v_h \in V_h, \quad a(u_h, v_h) = (f, v_h)_V. \end{cases}$$
(9.1.2)

La forme bilinéaire $a(\cdot, \cdot)$ étant encore coercive et continue sur les sous-espaces V_h , par le théorème de Lax-Milgram, on a existence et unicité de la solution $u_h \in V_h$ de (9.1.2). Soit $\{\phi_1, \ldots, \phi_N\}$ une base de V_h . Alors, il existe $u_1^h, \ldots, u_N^h \in \mathbb{R}$ tels que la solution $u_h \in V_h$ de (9.1.2) s'écrit

$$u_h = \sum_{j=1}^N u_j^h \phi_j.$$

Pour que l'égalité $a(u_h, v_h) = (f, v_h)_V$ ait lieu pour tout $v_h \in V_h$, il faut et il suffit qu'elle ait lieu pour tous les vecteurs de base ϕ_1, \ldots, ϕ_N . En utilisant la bilinéarité de $a(\cdot, \cdot)$, le problème (9.1.2) s'écrit alors

$$\begin{cases} \text{Trouver } u_1^h, \dots, u_N^h \in \mathbb{R} \text{ tels que }: \\ \forall i = 1, \dots, N, \quad \sum_{j=1}^N u_j^h a(\phi_j, \phi_i) = (f, \phi_i)_V. \end{cases}$$

En posant $U_h := (u_1^h, \ldots, u_N^h)^T \in \mathbb{R}^N$, on obtient que le problème (9.1.2) est équivalent au problème matriciel $\mathcal{K}_h U_h = b_h$, où $\mathcal{K}_h \in \mathbb{R}^{N \times N}$ et $b_h \in \mathbb{R}^N$ sont définis par

$$\mathcal{K}_h := (a(\phi_j, \phi_i))_{1 \le i, j \le N} \quad \text{et} \quad b_h := ((f, \phi_i)_V)_{1 \le i \le N}.$$
 (9.1.3)

Notation 9.1.1. Pour la suite, on note m et M les constantes de coercivité et de continuité de $a(\cdot, \cdot)$, *i.e.*

$$\forall u \in V, \quad a(u, u) \ge m ||u||_V^2,$$
(9.1.4)

 et

$$\forall u, v \in V, \quad |a(u, v)| \le M \, ||u||_V \, ||v||_V. \tag{9.1.5}$$

Proposition 9.1.2. On suppose de plus que $a(\cdot, \cdot)$ est symétrique. Alors, la matrice \mathcal{K}_h définie par (9.1.3) est définie positive. En particulier, \mathcal{K}_h est inversible et donc le système $\mathcal{K}_h U_h = b_h$ admet une solution unique $U_h \in \mathbb{R}^N$.

Démonstration. Par définition de \mathcal{K}_h , il est clair que si $a(\cdot, \cdot)$ est symétrique alors \mathcal{K}_h aussi. Soit $\xi \in \mathbb{R}^N \setminus \{0\}, \ \xi := (\xi_1, \ldots, \xi_N)^T$. On pose $\tilde{\xi} := \xi_1 \phi_1 + \cdots + \xi_N \phi_N \in V_h$. Puisque $a(\cdot, \cdot)$ est bilinéaire et coercive, on a

$$\mathcal{K}_h \xi \cdot \xi = \sum_{i,j=1}^N a(\phi_i, \phi_j) \xi_i \, \xi_j = \sum_{i,j=1}^N a(\xi_i \phi_i, \xi_j \phi_j) = a(\tilde{\xi}, \tilde{\xi}) \ge m \, ||\tilde{\xi}||_V^2 > 0,$$

donc \mathcal{K}_h est définie positive.

9.2 Convergence de la méthode

Il reste à montrer que la solution $u_h \in V_h$ de (9.1.2) est bien une approximation de u. Pour cela, on utilise le résultat suivant :

Lemme 9.2.1 (Lemme de Céa). Soit $u \in V$ la solution de (9.1.1) et $u_h \in V_h$ la solution de (9.1.2), alors on a

$$||u - u_h||_V \le \frac{M}{m} \inf_{v_h \in V_h} ||u - v_h||_V$$

où m et M sont données par (9.1.4) et (9.1.5).

Démonstration. Si $w_h \in V_h$, en prenant w_h comme fonction test dans (9.1.1) et (9.1.2), on obtient

$$a(u, w_h) = (f, w_h)_V = a(u_h, w_h),$$

d'où, par bilinéarité de $a(\cdot, \cdot)$, $a(u - u_h, w_h) = 0$. Soit $v_h \in V_h$, alors $w_h := v_h - u_h \in V_h$ donc $a(u - u_h, v_h - u_h) = 0$. On en déduit

$$m ||u - u_h||_V^2 \leq a(u - u_h, u - u_h) = a(u - u_h, u - v_h) + a(u - u_h, v_h - u_h)$$

= $a(u - u_h, u - v_h) \leq M ||u - u_h||_V ||u - v_h||_V,$

donc on a

$$\forall v_h \in V_h, \quad ||u - u_h||_V \le \frac{M}{m} ||u - v_h||_V$$

ce qui donne le résultat.

Théorème 9.2.2 (Théorème de convergence). On suppose qu'il existe un sous-espace \mathcal{V} de V dense dans V tel qu'il existe une application linéaire r_h de \mathcal{V} dans V_h vérifiant

$$\forall v \in \mathcal{V}, \quad \lim_{h \to 0} ||v - r_h v||_V = 0.$$
(9.2.1)

L'application r_h est applée opérateur d'interpolation de \mathcal{V} sur V_h . Alors, la solution $u_h \in V_h$ de (9.1.2) converge vers la solution $u \in V$ de (9.1.1), au sens où on a

$$\lim_{h \to 0} ||u - u_h||_V = 0 \tag{9.2.2}$$

Remarque 9.2.3. Le point important du Théorème de convergence est l'existence d'un opérateur d'interpolation. Celle-ci n'est pas immédiate et suivant le problème considéré, en particulier suivant les espaces de Hilbert V et V_h considérés, il est nécessaire de montrer l'existence de cet opérateur. On appellera les résultats de ce type des **lemmes d'interpo-lation**. Le théorème de convergence est l'analogue du théorème de Lax pour la méthode des différences finies. Ici la notion de stabilité correspond à la coercivité de $a(\cdot, \cdot)$ et la notion de consistance correspond au lemme d'interpolation.

Démonstration. Soit $\varepsilon > 0$. Puisque \mathcal{V} est dense, il existe $v \in \mathcal{V}$ tel que $||u - v||_V \leq \varepsilon$. De plus, l'existence de l'opérateur d'interpolation vérifiant (9.2.1) entraîne qu'il existe $h_0 > 0$ tel que si $h \leq h_0$ alors $||v - r_h v||_V \leq \varepsilon$. Puisque $r_h v_h \in V_h$, on a d'après le lemme de Céa :

$$\begin{aligned} ||u - u_h||_V &\leq \frac{M}{m} \inf_{v_h \in V_h} ||u - v_h||_V \leq \frac{M}{m} ||u - r_h v||_V \\ &\leq \frac{M}{m} ||u - v||_V + \frac{M}{m} ||v - r_h v||_V \leq \frac{2M}{m} \varepsilon, \end{aligned}$$

résultat

ce qui donne le résultat

En résumé l'approximation variationnelle consiste à construire des sous-espace V_h de Vdont on détermine une base $\{\phi_1, \ldots, \phi_N\}$ pour aboutir au système matriciel $\mathcal{K}_h U_h = b_h$. Pour avoir convergence de la méthode et aboutir à un système matriciel simple, il faut que l'espace V_h vérifie

- 1. qu'il existe un sous-espace dense \mathcal{V} sur lequel est défini un opérateur d'interpolation r_h vérifiant (9.2.1),
- 2. qu'il existe une base $\{\phi_1, \ldots, \phi_N\}$ telle que la résolution du système matriciel $\mathcal{K}_h U_h = b_h$ soit économique (typiquement que la matrice \mathcal{K}_h soit creuse).

La méthode des éléments finis repose sur le choix d'espaces V_h constitués de fonctions continues localement polynomiales et qui vérifient les deux points précédents.

Chapitre 10 Méthode des éléments finis en dimension 1

Dans ce chapitre, on décrit la méthode des éléments finis en dimension 1 pour le problème modèle de Dirichlet homogène :

$$\begin{cases} -u'' = f & \text{dans }]0,1[, \\ u(0) = u(1) & = 0, \end{cases}$$
(10.0.1)

où $f \in L^2(0, 1)$. La première étape de discrétisation consiste à choisir un maillage de [0, 1]: soit $(x_j)_{j=0,\dots,N+1}$ une subdivision de [0, 1] telle que

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

Pour simplifier, on suppose le pas d'espace uniforme donné par $h := \frac{1}{N+1} = x_{j+1} - x_j$, où j = 1, ..., N. La formulation variationnelle de (10.0.1) est

$$\begin{cases} \text{Trouver } u \in H_0^1(0,1) \text{ telle que :} \\ \forall v \in H_0^1(0,1), \quad \int_0^1 u' v' \, dx = \int_0^1 f v \, dx. \end{cases}$$
(10.0.2)

Notation 10.0.4. On note \mathbb{P}_k l'espace des polynômes sur \mathbb{R} à coefficients réels de degré inférieur ou égal à k.

La méthode des éléments finis consiste à définir comme espace d'approximation de $H^1(0, 1)$:

$$V_h := \{ v \in C(0,1) \mid v_{|[x_j, x_{j+1}]} \in \mathbb{P}_k, \forall j = 0, \dots, N \}.$$

Suivant le choix de $k \in \mathbb{N}^*$, on parle de méthode des éléments finis \mathbb{P}_k .

10.1 Éléments finis \mathbb{P}_1

On pose

$$V_h := \{ v \in C(0,1) \mid v_{|[x_j, x_{j+1}]} \in \mathbb{P}_1, \forall j = 0, \dots, N \},$$
(10.1.1)

 et

$$V_{0h} := \{ v \in V_h \mid v(0) = v(1) = 0 \}.$$
(10.1.2)

L'espace V_h est l'espace d'approximation de $H^1(0, 1)$ par la méthode des **éléments finis** \mathbb{P}_1 tandis que V_{0h} est l'espace d'approximation de $H_0^1(0, 1)$. D'après le chapitre précédent, il faut montrer que ces sont des sous-espaces de, respectivement, $H^1(0, 1)$ et $H_0^1(0, 1)$, en déterminer une base et montrer l'existence d'un opérateur d'interpolation défini sur un sous-espace dense de, respectivement, $H^1(0, 1)$ et $H_0^1(0, 1)$, et à valeurs dans, respectivement, V_h et V_{0h} .

On pose

$$\phi(x) := \begin{cases} 1 - |x| & \text{si } |x| \le 1, \\ 0 & \text{sinon.} \end{cases}$$
(10.1.3)

Puis on définit, pour j = 0, ..., N + 1, les fonctions ϕ_j par

$$\phi_j(x) := \phi\left(\frac{x-x_j}{h}\right). \tag{10.1.4}$$

Les fonctions ϕ_j sont des fonctions "chapeau" (voir Figure 10.1), elles vérifient $\phi_j(x_i) = \delta_{ij}$ et $\phi_j \in V_h$.



FIGURE 10.1 – Graphe de la fonction ϕ_i

De plus, $\operatorname{supp}(\phi_j) =]x_{j-1}, x_{j+1}[$ et on peut encore écrire ϕ_j sur son support par

$$\phi_j(x) = \begin{cases} \frac{x - x_{j-1}}{h} & \text{si } x \in [x_{j-1}, x_j], \\ \frac{x_{j+1} - x}{h} & \text{si } x \in [x_j, x_{j+1}]. \end{cases}$$

Remarque 10.1.1. Une définition équivalente des fonctions ϕ_i est que ce sont les seules fonctions de V_h telles que $\phi_i(x_j) = \delta_{ij}$, pour tout i, j = 0, ..., N.

Proposition 10.1.2. L'espace vectoriel V_h défini par (10.1.1) est un sous-espace de $H^1(0, 1)$ de dimension N + 2 et la famille $\{\phi_0, \ldots, \phi_{N+1}\}$ en est une base. En particulier, pour toute $v \in V_h$, on a

$$\forall x \in [0,1], \quad v(x) = \sum_{j=0}^{N+1} v(x_j)\phi_j(x).$$
 (10.1.5)

De même, l'espace vectoriel V_{0h} défini par (10.1.2) est un sous-espace de $H_0^1(0,1)$ de dimension N et la famille $\{\phi_1, \ldots, \phi_N\}$ en est une base. En particulier, pour toute $v \in V_{0h}$, on a

$$\forall x \in [0,1], \quad v(x) = \sum_{j=1}^{N} v(x_j)\phi_j(x).$$
 (10.1.6)

Remarque 10.1.3. On obtient en particulier que toute fonction de V_h et V_{0h} est définie de façon unique par ses valeurs aux noeuds x_j .

Démonstration. On montre d'abord que V_h est un sous-espace de $H^1(0,1)$. Si $v \in V_h$ alors $v \in C([0,1]) \subset L^2(0,1)$, il suffit donc de montrer que $v' \in L^2(0,1)$. Soit $\varphi \in C_c^{\infty}(0,1)$, on a

$$\langle v', \varphi \rangle = -\int_0^1 v \, \varphi' \, dx = -\sum_{j=0}^N \int_{x_j}^{x_{j+1}} v_{|[x_j, x_{j+1}]} \, \varphi' \, dx$$

Or $v_{|[x_i,x_{i+1}]} \in \mathbb{P}_1 \subset C([0,1])$ donc, d'après la formule de Green, on obtient

$$\langle v', \varphi \rangle = \sum_{j=0}^{N} \int_{x_{j}}^{x_{j+1}} v'_{[x_{j}, x_{j+1}]} \varphi \, dx - \sum_{j=0}^{N} (v(x_{j})\varphi(x_{j}) - v(x_{j+1})\varphi(x_{j+1})).$$

D'autre part, on a

$$\sum_{j=0}^{N} (v(x_j)\varphi(x_j) - v(x_{j+1})\varphi(x_{j+1})) = \sum_{j=0}^{N} v(x_j)\varphi(x_j) - \sum_{j=0}^{N} v(x_{j+1})\varphi(x_{j+1})$$
$$= \sum_{j=0}^{N} v(x_j)\varphi(x_j) - \sum_{j=1}^{N+1} v(x_{j+1})\varphi(x_{j+1})$$
$$= v(x_0)\varphi(x_0) - v(x_{N+1})\varphi(x_{N+1}) = 0,$$

 $\operatorname{car} \varphi(1) = \varphi(0) = 0$. On en déduit

$$\langle v', \varphi \rangle = \sum_{j=0}^{N} \int_{x_j}^{x_{j+1}} v'_{|[x_j, x_{j+1}]} \varphi \, dx,$$

d'où $v' = \sum_{j=0}^{N+1} v'_{[x_j, x_{j+1}]} \chi_{[x_j, x_{j+1}]} \in L^2(0, 1).$

Il reste à montrer que V_h a pour base $\{\phi_0, \ldots, \phi_{N+1}\}$ et l'égalité (10.1.5). Soit $j \in \{0, \ldots, N+1\}$, alors $\operatorname{supp}(\phi_j) \cap \operatorname{supp}(\phi_{j+1}) = [x_j, x_{j+1}]$. De plus, $\{\phi_j, \phi_{j+1}\}$ est une base de \mathbb{P}_1 sur $[x_j, x_{j+1}]$. En effet, \mathbb{P}_1 est de dimension 2 et si $\alpha, \beta \in \mathbb{R}$ vérifient $\alpha \phi_j(x) + \beta \phi_{j+1}(x) = 0$ pour $x \in [x_j, x_{j+1}]$, en prenant $x = x_j$ on obtient $\alpha = 0$ et avec $x = x_{j+1}$ on obtient $\beta = 0$.

Soit $v \in V_h$, alors pour tout $j = 0, \ldots, N+1$, $v_{\mid [x_j, x_{j+1}]} \in \mathbb{P}_1$. Donc $v_{\mid [x_j, x_{j+1}]} = \alpha \phi_j(x) + \beta \phi_{j+1}(x)$. En prenant $x = x_j$ puis $x = x_{j+1}$ on obtient $v_{\mid [x_j, x_{j+1}]} = v(x_j)\phi_j(x) + v(x_{j+1})\phi_{j+1}(x)$. D'autre part, pour $x \in [x_j, x_{j+1}]$, $\phi_i(x) = 0$ si $i \neq j$ et j + 1, donc

$$\forall x \in [x_j, x_{j+1}], \quad \sum_{i=0}^{N+1} v(x_i)\phi_i(x) = v(x_j)\phi_j(x) + v(x_{j+1})\phi_{j+1}(x) = v_{|[x_j, x_{j+1}]},$$

d'où le résultat.

L'approximation de la formulation variationnelle (10.0.2) est

$$(Trouver $u_h \in V_{0h} \text{ telle que }:$

$$(\forall v_h \in V_{0h}, \quad \int_0^1 u'_h v'_h \, dx = \int_0^1 f v_h \, dx.$$

$$(10.1.7)$$$$

Alors $u_h = u_h(x_1)\phi_1 + \dots + u_h(x_N)\phi_N$ et (10.1.7) s'écrit

$$\begin{cases} \text{Trouver } u_h(x_1), \dots, u_h(x_N) \in \mathbb{R} \text{ tels que :} \\ \forall i = 1, \dots, N, \quad \sum_{j=1}^N u_h(x_j) \int_0^1 \phi'_i \, \phi'_j \, dx = \int_0^1 f \, \phi_i \, dx \end{cases}$$

En posant $U_h := (u_h(x_1), \ldots, u_h(x_N))^T \in \mathbb{R}^N$, on obtient $\mathcal{K}_h U_h = b_h$, où $\mathcal{K}_h \in \mathbb{R}^{N \times N}$ et $b_h \in \mathbb{R}^N$ sont donnés par

$$\mathcal{K}_h := \left(\int_0^1 \phi'_i \, \phi'_j \, dx \right)_{1 \le i, j \le N} \quad \text{et} \quad b_h := \left(\int_0^1 f \, \phi_i \, dx \right)_{1 \le i \le N}$$

La matrice \mathcal{K}_h est appelée **matrice de rigidité** du système. Puisque $\operatorname{supp}(\phi_i) \cap \operatorname{supp}(\phi_j) = \emptyset$ si |i - j| > 1, la matrice \mathcal{K}_h est creuse. En particulier, on a $(\mathcal{K}_h)_{ij} = 0$ si |i - j| > 1 et pour $|i - j| \leq 1$, on a

$$(\mathcal{K}_h)_{ii} = \int_{x_{i-1}}^{x_{i+1}} \phi'_i \, \phi'_i \, dx = \int_{x_{i-1}}^{x_i} \frac{1}{h^2} \, dx + \int_{x_i}^{x_{i+1}} \frac{1}{h^2} \, dx = \frac{2}{h}$$

 et

$$(\mathcal{K}_h)_{ii+1} = \int_{x_i}^{x_{i+1}} \phi'_i \, \phi'_{i+1} \, dx = \int_{x_i}^{x_{i+1}} \left(-\frac{1}{h}\right) \frac{1}{h} \, dx = -\frac{1}{h} = (\mathcal{K}_h)_{ii-1}.$$

Autrement dit, \mathcal{K}_h est la matrice tridiagonale suivante

$$\mathcal{K}_{h} = \frac{1}{h} \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ \vdots & \ddots & & \ddots & \vdots \\ \vdots & & \ddots & \ddots & -1 \\ 0 & \dots & \dots & -1 & 2 \end{pmatrix},$$

qui est la matrice de discrétisation du Laplacien (au coefficient h^{-1} près) obtenue par la méthode des différences finies du schéma à 5 points (cela est dû au choix d'un pas de maillage uniforme).

Pour calculer le second membre b_h , lorsque la fonction f est compliquée, il faut utiliser une formule de quadrature (appelée aussi intégration numérique) dont on donne quelques exemples :

– Formule du rectangle

$$\int_{a}^{b} \psi(x) \, dx \simeq (b-a)\psi(a) \simeq (b-a)\psi(b).$$

- Formule du point milieu

$$\int_{a}^{b} \psi(x) \, dx \simeq (b-a)\psi\left(\frac{a+b}{2}\right).$$

- Formule du trapèze

$$\int_a^b \psi(x) \, dx \simeq \frac{1}{2} \, (b-a)(\psi(a) + \psi(b)).$$

– Formule de Simpson

$$\int_{a}^{b} \psi(x) \, dx \simeq \frac{1}{6} \left(b - a \right) \left[\psi(a) + \psi\left(\frac{a+b}{2}\right) + \psi(b) \right].$$

Les deux premières formules sont exactes pour les fonctions ψ affines, et la troisième est exacte pour les polynômes de second degré. Pour les fonctions régulières, ces formules sont approchées avec un reste d'ordre O(h), $O(h^2)$ et $O(h^3)$ respectivement

Remarque 10.1.4. Par hypothèse on sait seulement que $f \in L^2(0, 1)$ donc f n'est défini que presque partout. Mais, dans la pratique, f est une donnée donc connue en tout point, ce qui justifie l'emploi des formules de quadrature ci-dessus.

10.2 Convergence et estimation d'erreur pour la méthode \mathbb{P}_1

D'après le Théorème 9.2.2, pour obtenir la convergence de la méthode \mathbb{P}_1 , il suffit qu'il existe un sous-espace dense \mathcal{V} de $H_0^1(0, 1)$ sur lequel est défini un opérateur d'interpolation à valeur dans V_h . En dimension 1, le fait que $H^1(0, 1)$ est un sous-espace de C(0, 1) permet de prendre pour \mathcal{V} l'espace $H_0^1(0, 1)$ tout entier.

Définition 10.2.1. L'opérateur d'interpolation \mathbb{P}_1 est l'application r_h définie de $H^1(0, 1)$ dans V_h par :

$$\forall v \in H^1(0,1), \quad r_h v(x) := \sum_{j=0}^{N+1} v(x_j) \phi_j(x).$$

où les fonctions ϕ_j sont données par (10.1.3)-(10.1.4). En particulier, sur $H_0^1(0,1)$ l'opérateur r_h vérifie

$$\forall v \in H_0^1(0,1), \qquad r_h v(x) = \sum_{j=1}^N v(x_j) \phi_j(x).$$

Remarque 10.2.2.

- 1. D'après la Proposition 10.1.2, il est clair que $r_h v \in V_h$ et, si $v \in V_h$, alors $r_h v = v$.
- 2. La définition a un sens car $H^1(0,1)$ est un sous-espace de C(0,1) et donc toute fonction de $H^1(0,1)$ est définie en tout point de]0,1[. En dimension supérieure, les fonctions H^1 ne sont pas nécessairement continues et donc définies seulement presque partout et la définition précédente n'a plus de sens. Pour définir un opérateur d'interpolation, il sera nécessaire de considérer un sous-espace \mathcal{V} dense dans H^1 et constitué de fonctions régulières.

Lemme 10.2.3 (Lemme d'interpolation \mathbb{P}_1).

1. Pour toute $v \in H^2(0,1)$, il existe une constante c > 0 indépendante de h telle que

$$||v - r_h v||_{H^1(0,1)} \le c h ||v''||_{L^2(0,1)}.$$
(10.2.1)

2. Pour toute $v \in H^1(0,1)$, on a

$$\lim_{h \to 0} ||v - r_h v||_{H^1(0,1)} = 0.$$
(10.2.2)

Démonstration.

1) On montre (10.2.1) pour $v \in C^2([0,1])$, par densité on en déduit (10.2.1) pour $v \in H^2(0,1)$. Il faut estimer $||v - r_h v||_{L^2(0,1)}$ et $||v' - (r_h v)'||_{L^2(0,1)}$ en fonction de *h*. Soit $x \in [x_j, x_{j+1}]$. Comme $r_h v \in V_h$, on a $r_h v(x) = \alpha x + \beta$. De plus, $r_h v(x_j) = v(x_j)$ et $r_h v(x_{j+1}) = v(x_{j+1})$ donc

$$r_h v(x) = v(x_j) + \frac{v(x_{j+1}) - v(x_j)}{h} (x - x_j).$$

On obtient

$$v(x) - r_h v(x) = v(x) - v(x_j) - \frac{v(x_{j+1}) - v(x_j)}{h} (x - x_j)$$
$$= \int_{x_j}^x v'(t) dt - \frac{x - x_j}{h} \int_{x_j}^{x_{j+1}} v'(t) dt.$$

D'après le théorème des accroissements finis, il existe $y \in [x_j, x]$ et $z \in [x_j, x_{j+1}]$ tels que

$$v(x) - r_h v(x) = (x - x_j)v'(y) - (x - x_j)v'(z) = (x - x_j)\int_z^y v''(t) dt$$

Alors, d'après l'inégalité de Cauchy-Schwarz, on a

$$|v(x) - r_h v(x)|^2 \leq h^2 \left(\int_z^y v''(t) \, dt \right)^2 \leq h^2 \left(\int_{x_j}^{x_{j+1}} |v''(t)| \, dt \right)^2$$
$$\leq h^2 \left(\left(\int_{x_j}^{x_{j+1}} \, dt \right)^{1/2} \left(\int_{x_j}^{x_{j+1}} |v''(t)|^2 \, dt \right)^{1/2} \right)^2$$
$$\leq h^3 \int_{x_j}^{x_{j+1}} |v''(t)|^2 \, dt.$$

On intègre x sur $[x_j, x_{j+1}]$

$$\int_{x_j}^{x_{j+1}} |v(x) - r_h v(x)|^2 \, dx \le h^4 \int_{x_j}^{x_{j+1}} |v''(t)|^2 \, dt.$$

En sommant sur $j = 0, \ldots, N$, on obtient

$$||v - r_h v||_{L^2(0,1)} \le h^2 ||v''||_{L^2(0,1)}$$

Il reste à obtenir une estimation sur les dérivées. Pour $x \in [x_j, x_{j+1}]$, on a

$$v'(x) - (r_h v)'(x) = v'(x) - \frac{v(x_{j+1}) - v(x_j)}{h} = v'(x) - \frac{1}{h} \int_{x_j}^{x_{j+1}} v'(t) dt$$
$$= \frac{1}{h} \int_{x_j}^{x_{j+1}} v'(x) - v'(t) dt = \frac{1}{h} \int_{x_j}^{x_{j+1}} \int_t^x v''(y) dy dt.$$

Par l'inégalité de Cauchy-Schwarz, on obtient donc

$$|v'(x) - (r_h v)'(x)|^2 \leq \frac{1}{h^2} \left(\int_{x_j}^{x_{j+1}} \int_t^x v''(y) \, dy \, dt \right)^2 \leq \frac{1}{h^2} \left(\int_{x_j}^{x_{j+1}} \int_{x_j}^{x_{j+1}} |v''(y)| \, dy \, dt \right)^2$$
$$\leq h \int_{x_j}^{x_{j+1}} |v''(y)|^2 \, dy.$$

En intégrant x sur $[x_j, x_{j+1}]$, on a

$$\int_{x_j}^{x_{j+1}} |v'(x) - (r_h v)'(x)|^2 \, dx \le h^2 \int_{x_j}^{x_{j+1}} |v''(y)|^2 \, dy.$$

On en déduit $||v' - (r_h v)'||_{L^2(0,1)} \le h ||v''||_{L^2(0,1)}$, d'où

$$\begin{aligned} ||v - r_h v||^2_{H^1(0,1)} &= ||v - r_h v||^2_{L^2(0,1)} + ||v' - (r_h v)'||^2_{L^2(0,1)} \\ &\leq h^4 \, ||v''||^2_{L^2(0,1)} + h^2 \, ||v''||^2_{L^2(0,1)} \leq 2h^2 \, ||v''||^2_{L^2(0,1)}, \end{aligned}$$

pour h < 1, ce qui donne (10.2.1).

2) On montre (10.2.2) pour $v \in C^1([0,1])$. Le raisonnement est le même que pour 1). On montre tout d'abord que $||v - r_h v||_{L^2(0,1)}$ converge vers 0. Soit $x \in [x_j, x_{j+1}]$, on a

$$v(x) - r_h v(x) = \int_{x_j}^x v'(t) \, dt - \frac{x - x_j}{h} \, \int_{x_j}^{x_{j+1}} v'(t) \, dt,$$

d'où

$$v(x) - r_h v(x) \leq \int_{x_j}^{x_{j+1}} |v'(t)| \, dt + \int_{x_j}^{x_{j+1}} |v'(t)| \, dt = 2 \int_{x_j}^{x_{j+1}} |v'(t)| \, dt.$$

D'après l'inégalité de Cauchy-Schwarz, on obtient

$$\int_{x_j}^{x_{j+1}} |v(x) - r_h v(x)|^2 \, dx \le 2h^2 \int_{x_j}^{x_{j+1}} |v'(t)|^2 \, dt$$

donc $||v - r_h v||_{L^2(0,1)} \leq \sqrt{2} h ||v'||_{L^2(0,1)} \xrightarrow[h \to 0]{} 0.$ Il reste à montrer que $||v' - (r_h v)'||_{L^2(0,1)} \xrightarrow[h \to 0]{} 0.$ Soit $\varepsilon > 0$. Puisque $C^2([0,1])$ est dense dans $C^1([0,1])$, il existe $\phi \in C^2([0,1])$ telle que

$$||v' - \phi'||_{L^2(0,1)} \le \varepsilon.$$

Pour $u \in C^1([0, 1])$, on a

$$\int_{x_j}^{x_{j+1}} |(r_h u)'(x)|^2 dx = \frac{1}{h} (u(x_{j+1}) - u(x_j))^2 = \frac{1}{h} \left(\int_{x_j}^{x_{j+1}} u'(t) dt \right)^2$$
$$\leq \int_{x_j}^{x_{j+1}} |u'(t)|^2 dt,$$

donc $||(r_h u)'||_{L^2(0,1)} \le ||u'||_{L^2(0,1)}$. Alors, on en déduit

$$||(r_hv)' - (r_h\phi)'||_{L^2(0,1)} = ||(r_h(v-\phi))'||_{L^2(0,1)} \le ||v'-\phi'||_{L^2(0,1)} \le \varepsilon$$

En appliquant l'inégalité (10.2.1) à ϕ , on obtient

$$||\phi' - (r_h \phi)'||_{L^2(0,1)} \le c \, h \, ||\phi''||_{L^2(0,1)} \le \varepsilon,$$

pour h suffisamment petit, d'où

$$\begin{aligned} ||v' - (r_h v)'||_{L^2(0,1)} &\leq ||v' - \phi'||_{L^2(0,1)} + ||\phi' - (r_h \phi)'||_{L^2(0,1)} + ||(r_h \phi)' - (r_h v)'||_{L^2(0,1)} \\ &\leq c \varepsilon \xrightarrow[\varepsilon \to 0]{} 0, \end{aligned}$$

ce qui donne le résultat.

Théorème 10.2.4 (Convergence de la méthode \mathbb{P}_1). Soient $u \in H_0^1(0,1)$ la solution de (10.0.2) et $u_h \in V_h$ la solution de (10.1.7). Alors, on a

$$\lim_{h \to 0} ||u - u_h||_{H^1(0,1)} = 0.$$
(10.2.3)

Autrement dit, la méthodes des éléments finis \mathbb{P}_1 converge. De plus, si $u \in H^2(0,1)$ alors il existe une constante c > 0 telle que

$$||u - u_h||_{H^1(0,1)} \le c h ||f||_{L^2(0,1)}.$$
(10.2.4)

On dit que la convergence est linéaire.

Démonstration. La convergence (10.2.3) est la conséquence du Lemme 10.2.3 et du Théorème 9.2.2. L'estimation (10.2.4) s'obtient à partir du lemme de Céa et de (10.2.1) :

$$\begin{aligned} ||u - u_h||_{H^1(0,1)} &\leq c \inf_{v_h \in V_h} ||u - v_h||_{H^1(0,1)} \leq c ||u - r_h u||_{H^1(0,1)} \\ &\leq c h ||u''||_{L^2(0,1)} = c h ||f||_{L^2(0,1)}, \end{aligned}$$

car -u'' = f p.p. dans]0, 1[.

10.3 Cas d'une condition de Neumann

Dans cette section, on étudie la méthode des éléments finis \mathbb{P}_1 pour le problème de Neumann :

$$\begin{cases} -u'' + u = f & \text{dans }]0,1[, \\ u'(0) = \alpha, & u'(1) = \beta, \end{cases}$$
(10.3.1)

où $f \in L^2(0,1)$. La formulation variationnelle de (10.3.1) est

$$\begin{cases} \text{Trouver } u \in H^1(0,1) \text{ telle que :} \\ \forall v \in H^1(0,1), \quad \int_0^1 u' v' + u v \ dx = \int_0^1 f v \ dx + \beta v(1) - \alpha v(0). \end{cases}$$
(10.3.2)

Avec les notations de la section précédente, le problème approché est alors

$$\begin{cases} \text{Trouver } u_h \in V_h \text{ telle que :} \\ \forall v_h \in V_h, \quad \int_0^1 u'_h v'_h + u_h v_h \, dx = \int_0^1 f v_h \, dx + \beta v_h(1) - \alpha v_h(0), \end{cases}$$
(10.3.3)

qui est équivalent à

$$\begin{cases} \text{Trouver } u_h(x_0), \dots, u_h(x_{N+1}) \in \mathbb{R} \text{ tels que :} \\ \forall i = 0, \dots, N+1, \quad \sum_{j=0}^{N+1} u_h(x_j) \int_0^1 \phi'_i \, \phi'_j + \phi_i \, \phi_j \, dx = \int_0^1 f \, \phi_i \, dx + \beta \, \phi_i(1) - \alpha \, \phi_i(0). \end{cases}$$

On remarque que l'on a

$$\phi_i(1) = \phi_i(x_{N+1}) = \begin{cases} 1 & \text{si } i = N+1, \\ 0 & \text{sinon,} \end{cases}$$

 et

$$\phi_i(0) = \phi_i(x_0) = \begin{cases} 1 & \text{si } i = 0, \\ 0 & \text{sinon.} \end{cases}$$

Alors, on définit $b_h \in \mathbb{R}^{N+2}$ par :

$$(b_h)_i := \begin{cases} \int_{x_{i-1}}^{x_{i+1}} f \phi_i \, dx & \text{si } 1 \le i \le N, \\ \int_0^{x_1} f \phi_0 \, dx - \alpha & \text{si } i = 0, \\ \int_{x_N}^{x_{N+1}} f \phi_{N+1} \, dx + \beta & \text{si } i = N+1. \end{cases}$$

Ainsi, si on pose $U_h := (u_h(x_0), \dots, u_h(x_{N+1}))^T \in \mathbb{R}^{N+2}$, on obtient que U_h est solution de $\mathcal{K}_h U_h = b_h$, où la matrice de rigidité $\mathcal{K}_h \in \mathbb{R}^{(N+2) \times (N+2)}$ est définie par

$$\mathcal{K}_h := \left(\int_0^1 \phi'_i \, \phi'_j + \phi_i \, \phi_j \, dx \right)_{0 \le i,j \le N+1}$$

Exercice 10.3.1. En reprenant les calculs de la section précédente, calculer explicitement \mathcal{K}_h .

Remarque 10.3.2.

- 1. Le théorème de convergence de la méthode \mathbb{P}_1 reste vrai pour la condition de Neumann.
- 2. Si l'on considère un problème plus général du type

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

l'approximation par la méthode \mathbb{P}_1 aboutit encore à un système $\mathcal{K}_h U_h = b_h$. Mais, dans ce cas, \mathcal{K}_h est donnée par

$$\mathcal{K}_h := \left(\int_0^1 \nu(x) \, \phi_i'(x) \, \phi_j'(x) + c(x) \, \phi_i(x) \, \phi_j(x) \, dx \right)_{0 \le i, j \le N+1}.$$

Pour calculer \mathcal{K}_h on peut avoir besoin d'utiliser des formules de quadrature.

10.4 Méthode des éléments finis \mathbb{P}_2

On termine ce chapitre par un second exemple de méthode \mathbb{P}_k (en dimension d = 2 ou 3, on traitera le cas général). Les espaces d'approximation considérés sont

$$V_h := \{ v \in C(0,1) \mid v_{|[x_j, x_{j+1}]} \in \mathbb{P}_2, \forall j = 0, \dots, N \},$$
(10.4.1)

 et

$$V_{0h} := \{ v \in V_h \mid v(0) = v(1) = 0 \}.$$
(10.4.2)

Comme dans le cas de la méthode \mathbb{P}_1 , il faut déterminer une base de V_h et un opérateur d'interpolation.

Dans la méthode des éléments finis \mathbb{P}_1 , toute fonction v de V_h était linéaire sur $[x_j, x_{j+1}]$, il suffisait donc de connaître les valeurs de v en x_j et x_{j+1} pour déterminer v sur $[x_j, x_{j+1}]$. Pour la méthode des éléments finis \mathbb{P}_2 , il est nécessaire de connaître les valeurs de v en trois points de $[x_j, x_{j+1}]$. Pour cela, on définit le point milieu :

$$x_{j+1/2} := x_j + \frac{h}{2}, \quad \forall \ j = 0, \dots, N.$$

Par analogie avec la méthode \mathbb{P}_1 , on va prendre pour fonctions de base les fonctions ψ_j et $\psi_{j+1/2}$ appartenant à V_h telles que

$$\psi_j(x_i) = \delta_{ij}$$
 et $\psi_{j+1/2}(x_{i+1/2}) = \delta_{ij}$.

Remarque 10.4.1. Ce choix de fonctions de base qui est analogue à celui fait pour la méthode \mathbb{P}_1 n'est pas unique. On parle d'**éléments de Lagrange** pour le choix d'une base telle que $\psi_k(x_l) = \delta_{kl}$. D'autres choix sont possibles tels que les éléments de Hermite, où on considère en plus les valeurs de la fonction dérivée (voir [1] page 172).

On peut montrer que les fonctions ψ_j et $\psi_{j+1/2}$ sont données par

$$\psi_j(x) = \phi\left(\frac{x-x_j}{h}\right) \quad \text{où } 0 \le j \le N+1,$$

$$\psi_{j+1/2}(x) = \psi\left(\frac{x-x_{j+1/2}}{h}\right) \quad \text{où } 0 \le j \le N,$$
(10.4.3)

avec

$$\phi(x) := \begin{cases} (1+x)(1+2x) & \text{si } -1 \le x \le 0, \\ (1-x)(1-2x) & \text{si } 0 \le x \le 1, \\ 0 & \text{sinon}, \end{cases}$$
(10.4.4)

 et

$$\psi(x) := \begin{cases} (1 - 4x^2)(1 + 2x) & \text{si } |x| \le \frac{1}{2}, \\ 0 & \text{sinon.} \end{cases}$$
(10.4.5)

Comme dans le cas des éléments finis \mathbb{P}_1 , on obtient le résultat suivant :



FIGURE 10.2 – Graphes des fonctions ψ_j et $\psi_{j+1/2}$

Proposition 10.4.2. L'espace V_h défini par (10.4.1) est un sous-espace de $H^1(0,1)$ de dimension 2N + 3 et pour toute $v_h \in V_h$, on a

$$v_h = \sum_{j=0}^{N+1} v_h(x_j)\psi_j + \sum_{j=0}^{N} v_h(x_{j+1/2})\psi_{j+1/2}$$
$$= \sum_{j=0}^{2N+2} v_h(x_{j/2})\psi_{j/2}.$$

De même, l'espace V_{0h} défini par (10.4.2) est un sous-espace de $H_0^1(0,1)$ de dimension 2N + 1 et pour toute $v_h \in V_{0h}$, on a

$$v_h = \sum_{j=1}^N v_h(x_j)\psi_j + \sum_{j=0}^N v_h(x_{j+1/2})\psi_{j+1/2}$$
$$= \sum_{j=1}^{2N+1} v_h(x_{j/2})\psi_{j/2}.$$

De même que pour la méthode des éléments finis \mathbb{P}_1 , on en déduit la définition suivante d'opérateur d'interpolation :

Définition 10.4.3. L'opérateur d'interpolation \mathbb{P}_2 est l'application r_h définie de $H^1(0, 1)$ dans V_h par :

$$\forall v \in H^1(0,1), \quad r_h v := \sum_{j=0}^{2N+2} v_h(x_{j/2})\psi_{j/2}.$$

où les fonctions ψ_j sont données par (10.4.3)-(10.4.4)-(10.4.5).

Le problème approché par la méthode des éléments finis \mathbb{P}_2 du problème de Dirichlet (10.0.2) est

Trouver
$$u_h \in V_{0h}$$
 telle que :
 $\forall v_h \in V_{0h}, \quad \int_0^1 u'_h v'_h \, dx = \int_0^1 f v_h \, dx,$
(10.4.6)

qui est équivalent à

$$\begin{cases} \text{Trouver } u_h(x_{j/2}) \in \mathbb{R}, \ j = 1, \dots, 2N+1 \text{ tels que :} \\ \forall \ i = 1, \dots, 2N+1, \quad \sum_{j=1}^{2N+1} u_h(x_j) \int_0^1 \psi'_{i/2} \ \psi'_{j/2} \ dx = \int_0^1 f \ \psi_{i/2} \ dx. \end{cases}$$

On pose $U_h := (u_h(x_{1/2}), u_h(x_1), \dots, u_h(x_N), u_h(x_{N+1/2}))^T \in \mathbb{R}^{2N+1}$. Alors, U_h est solution de $\mathcal{K}_h U_h = b_h$, où $\mathcal{K}_h \in \mathbb{R}^{(2N+1) \times (2N+1)}$ et $b_h \in \mathbb{R}^{2N+1}$ sont donnés par

$$\mathcal{K}_h := \left(\int_0^1 \psi'_{i/2} \, \psi'_{j/2} \, dx \right)_{1 \le i, j \le 2N+1} \quad \text{et} \quad b_h := \left(\int_0^1 f \, \psi_{i/2} \, dx \right)_{1 \le i \le 2N+1}$$

Exercice 10.4.4. Vérifier que l'on a

$$\mathcal{K}_{h} = \frac{1}{h} \begin{pmatrix}
16/3 & -8/3 & 0 & & & \\
-8/3 & 14/3 & -8/3 & 1/3 & 0 & & & \\
0 & -8/3 & 16/3 & -8/3 & 0 & & & \\
0 & 1/3 & -8/3 & 14/3 & -8/3 & 1/3 & 0 & & \\
& & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\
& & 0 & -8/3 & 16/3 & -8/3 & 0 & \\
& & 0 & 1/3 & -8/3 & 14/3 & -8/3 & \\
& & 0 & 0 & 0 & -8/3 & 16/3 &
\end{pmatrix}.$$
(10.4.7)

En adaptant les arguments de la méthode \mathbb{P}_1 , on peut montrer le résultat de convergence suivant :

Théorème 10.4.5 (Convergence de la méthode \mathbb{P}_2). Soit $u \in H_0^1(0, 1)$ la solution de (10.0.2) et $u_h \in V_h$ la solution de (10.4.6). Alors, on a

$$\lim_{h \to 0} ||u - u_h||_{H^1(0,1)} = 0.$$
(10.4.8)

Autrement dit, la méthode des éléments finis \mathbb{P}_2 converge. De plus, si $u \in H^3(0,1)$ alors il existe une constante c > 0 indépendante de h telle que

$$||u - u_h||_{H^1(0,1)} \le c h^2 ||u'''||_{L^2(0,1)}.$$
(10.4.9)

On dit que la convergence est quadratique.

Remarque 10.4.6. D'après le théorème de convergence, on voit que l'avantage de la méthode \mathbb{P}_2 est une accélération de la convergence (quadratique au lieu de linéaire) lorsque la solution est régulière ($u \in H^3(0,1)$). Le désavantage est que le système $\mathcal{K}_h U_h = b_h$ est plus couteux à résoudre car la matrice \mathcal{K}_h n'est plus tridiagonale mais pentadiagonale. En particulier, on remarque que si $u \notin H^3(0,1)$, il n'y a aucun intérêt à employer la méthode \mathbb{P}_2 .

Chapitre 11

Méthode des éléments finis en dimension $d \ge 2$

11.1 Maillages triangulaires

En dimension 2 (resp. dimension 3) une **triangulation** d'un domaine Ω est une subdivision de Ω en triangles (resp. tétraèdre). Dans ce cas, le maillage ne peut remplir entièrement Ω que si $\overline{\Omega}$ est une réunion finie de polyèdres, on dit que Ω est **polyédrique**. On donnera à la fin de ce chapitre une rapide description de la méthode lorsque le domaine Ω n'est pas polyédrique.

Dans la suite, on supposera toujours que Ω est un domaine polyédrique de \mathbb{R}^d , d = 2, 3. Afin d'énoncer les résultats en dimension quelconque d = 2 ou 3, on introduit la notion de *d*-simplexe non dégénéré qui correspond à un triangle en dimension 2 et à un tétraèdre en dimension 3 (non dégénérés au sens non vides).

Définition 11.1.1. Soient d + 1 points $a_j := (a_{ij})_{i=1}^d$, $1 \le j \le d+1$, non situés dans un même hyperplan de \mathbb{R}^d , c'est à dire tels que la matrice d'ordre d+1

$$A := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1d+1} \\ a_{21} & a_{22} & \dots & a_{2d+1} \\ \vdots & \vdots & \dots & \vdots \\ a_{d1} & a_{d2} & \dots & a_{dd+1} \\ 1 & 1 & \dots & 1 \end{pmatrix}$$
(11.1.1)

est inversible. On appelle *d*-simplexe non dégénéré K de sommets $(a_j)_{1 \le j \le d+1}$ l'enveloppe convexe des points a_j , $1 \le j \le d+1$.

Définition 11.1.2. Une triangulation admissible de $\overline{\Omega}$ est une famille $\mathcal{T}_h := \{K_i\}_{1 \le i \le N}$ constituée de *d*-simplexes non dégénérés tels que

- 1. $K_i \subset \Omega$ et $\overline{\Omega} = \bigcup_{i=1}^N K_i$,
- 2. l'intersection $K_i \cap K_j$ est soit vide, soit un *m*-simplexe, avec $0 \le m \le d-1$, dont les sommets sont des sommets de K_i et K_j

Le paramètre h est défini par

$$h := \max_{1 \le i \le N} \operatorname{diam}(K_i), \quad \text{où} \quad \operatorname{diam}(K_i) := \sup\{|x - y| \mid x, y \in K_i\},\$$

où $|\cdot|$ désigne la norme euclidienne. Les **sommets** ou **nœuds** de la triangulation \mathcal{T}_h sont les sommets des *d*-simplexes K_i .

Remarque 11.1.3. Précisons la signification du deuxième point de la définition dans les cas d = 2 et d = 3:

- 1. En dimension 2, 2. signifie que l'intersection de deux triangles non égaux est soit vide, soit réduite à un sommet commun, soit une arête commune entière.
- 2. En dimension 3, 2. signifie que l'intersection de deux tétraèdres non égaux est soit vide, soit réduite à un sommet commun, soit une arête commune entière, soit une face commune entière.



FIGURE 11.1 – Exemples de maillages non admissible à gauche et admissible à droite, en dimension 2

Définition 11.1.4. Soient K un d-simplexe non dégénéré de \mathbb{R}^d de sommets $(a_j)_{1 \leq j \leq d+1}$ et $x \in \mathbb{R}^d$. Alors, x est caractérisé par ses **coordonnées barycentriques** $\lambda_j^K(x) \in \mathbb{R}$, où $1 \leq j \leq d+1$, par rapport à K, définies comme solutions du système linéaire

$$\sum_{j=1}^{d+1} \lambda_j^K(x) = 1 \quad \text{et} \quad x = \sum_{j=1}^{d+1} \lambda_j^K(x) \, a_j.$$
(11.1.2)

Remarque 11.1.5. Le système défini par (11.1.2) s'écrit encore $A\lambda = X$, où $\lambda \in \mathbb{R}^{d+1}$ a pour coefficients les $\lambda_j^K(x)$, $X = (1, x_1, \dots, x_d)^T \in \mathbb{R}^{d+1}$ et A est donnée par (11.1.1). L'inversibilité de A assure l'existence et l'unicité de λ donc des coordonnées barycentriques $\lambda_j^K(x)$.

Les coordonnées bary centriques permettent de donner une caractérisation simple du d-simplexe K :

$$K = \{ x \in \mathbb{R}^d \mid 0 \le \lambda_i^K(x) \le 1, \ \forall \ i = 1, \dots, d+1 \}.$$
(11.1.3)

De plus, on remarque l'on a $\lambda_i^K(a_j) = \delta_{ij}$.

Définition 11.1.6. Soit K un d-simplexe non dégénéré de \mathbb{R}^d de coordonnées barycentriques associées $\lambda_j^K \in \mathbb{R}$, où $1 \le j \le d+1$. On appelle **treillis d'ordre** $k \in \mathbb{N}^*$ de Kl'ensemble

$$\Sigma_k := \left\{ x \in K \mid \lambda_j(x) \in \left\{ 0, \frac{1}{k}, \dots, \frac{k-1}{k}, 1 \right\}, \ \forall \ j = 1, \dots, d-1 \right\}.$$
(11.1.4)

Remarque 11.1.7. La définition de treillis permet de définir l'équivalent en dimension $d \ge 1$ de la notion de point milieu d'un intervalle. Comme vu dans le cas de la dimension 1, celle-ci sera nécessaire dans l'étude la méthode des éléments finis \mathbb{P}_k pour k > 1. On peut aussi définir Σ_0 comme étant le singleton réduit au barycentre de K.

Exemple 11.1.8.

1. Pour k = 1, on a

$$\Sigma_1 := \left\{ x \in K \mid \lambda_j(x) \in \{0, 1\}, \forall j = 1, \dots, d-1 \right\},\$$

donc Σ_1 est composé des sommets de K.

2. Pour $k = 2, \Sigma_2$ est l'ensemble des sommets et des points milieux (*i.e.* les centres des arêtes).



FIGURE 11.2 – Treillis d'ordre 2 pour un triangle à gauche et pour un tétraèdre à droite

Notation 11.1.9. On note \mathbb{P}_k , l'espace des polynômes à coefficients réels de \mathbb{R}^d de degré inférieur ou égal à k, *i.e.* tout polynôme p de \mathbb{P}_k est de la forme

$$\forall x \in \mathbb{R}^d, \quad p(x) = \sum_{\substack{i_1 \ge 0, \dots, i_d \ge 0\\i_1 + \dots + i_d \le k}} \alpha_{i_1, \dots, i_d} x_1^{i_1} \dots x_d^{i_d}, \tag{11.1.5}$$

où $\alpha_{i_1,\ldots,i_d} \in \mathbb{R}$.

Remarque 11.1.10. On peut montrer que l'on a

$$\operatorname{Card}(\Sigma_k) = \dim(\mathbb{P}_k).$$

Lemme 11.1.11 (admis, voir [1] page 177). Soient K un d-simplexe non dégénéré de \mathbb{R}^d et, pour $k \in \mathbb{N}^*$, Σ_k le treillis d'ordre k de K. On désigne par $(\sigma_j)_{1 \le j \le N_k}$ les points de Σ_k . Alors, tout polynôme de \mathbb{P}_k est déterminé de manière unique par ses valeurs aux points $(\sigma_j)_{1 \le j \le N_k}$. Plus précisément, il existe une base $(\psi_j)_{1 \le j \le N_k}$ de \mathbb{P}_k telle que

$$\psi_j(\sigma_i) = \delta_{ij} \quad pour \ tout \ 1 \le i, j \le N_k. \tag{11.1.6}$$

Dans la suite, ce lemme a essentiellement deux applications importantes. Tout d'abord, celui-ci permettra de construire une base de l'espace d'approximation. Ensuite, celui-ci permet de montrer le résultat suivant :

Lemme 11.1.12. Soient K et K' deux d-simplexes non dégénérés de \mathbb{R}^d ayant une face commune $\Gamma := \partial K \cap \partial K'$. Soit un entier $k \geq 1$. Alors, leurs treillis d'ordre k, Σ_k et Σ'_k coïncident sur cette face Γ . De plus, étant donné p_K et $p_{K'}$ deux polynômes de \mathbb{P}_k , la fonction v définie par

$$v(x) := \begin{cases} p_K(x) & si \ x \in K, \\ p_{K'}(x) & si \ x \in K', \end{cases}$$

est continue sur $K \cup K'$ si et seulement si les valeurs de p_K et $p_{K'}$ coïncident aux points des treillis sur la face commune Γ .

Remarque 11.1.13. Soit $\{\sigma_1, \ldots, \sigma_N\} = \Gamma \cap \Sigma_k$. On note $(\psi_j^K)_{1 \le j \le N_k}$ la base de \mathbb{P}_k associée à K et $(\psi_j^{K'})_{1 \le j \le N_k}$ celle associée à K', définies par (11.1.6). Pour $j \in \{1, \ldots, N\}$ fixé, si on définit p_K et $p_{K'}$ par

$$p_K = \psi_j^K \quad \text{et} \quad p_{K'} = \psi_j^{K'},$$

alors la fonction v du Lemme 11.1.12 est continue sur $K \cup K'$ puisque pour tout $i \in \{1, \ldots, N\}$, on a $p_K(\sigma_i) = \delta_{ij} = p_{K'}(\sigma_i)$.

Démonstration. Si v est continue sur $K \cup K'$ alors $p_K = p_{K'}$ sur Γ . Réciproquement, supposons que p_K et $p_{K'}$ coïncident aux points des treillis sur Γ . D'après le Lemme 11.1.11, p_K et $p_{K'}$ sont uniquement déterminés sur Γ par leurs valeurs sur $\Sigma_k \cap \Gamma$ donc, si celles-ci coïncident, on a nécessairement $p_K = p_{K'}$ sur Γ donc v est continue sur $K \cup K'$. \Box

11.2 Éléments finis \mathbb{P}_k en dimension $d \geq 2$

Dans toute la suite Ω est un domaine polyédrique. On donne tout d'abord la définition des espaces d'approximation de $H^1(\Omega)$ et $H^1_0(\Omega)$ associés à une triangulation \mathcal{T}_h de Ω .

Définition 11.2.1. Soit \mathcal{T}_h une triangulation de Ω .

1. La méthode des éléments finis triangulaire de Lagrange d'ordre $k \in \mathbb{N}^*$ est définie comme étant la méthode pour laquelle l'espace $H^1(\Omega)$ est approché par l'espace

$$V_h := \{ v \in C(\overline{\Omega}) \mid v_{|K} \in \mathbb{P}_k, \ \forall \ K \in \mathcal{T}_h \}.$$

$$(11.2.1)$$

2. Les nœuds de degrés de liberté sont les points σ_i , $1 \leq i \leq N_{dl}$, des treillis d'ordre k de chaque K de \mathcal{T}_h . Le nombre de degrés de liberté N_{dl} ne compte qu'une fois les points communs de deux treillis.

- 3. Les **degrés de liberté** d'une fonction $v \in V_h$ sont les la valeurs $v(\sigma_i)$ de v aux points σ_i .
- 4. L'espace $H_0^1(\Omega)$ est approché par l'espace

$$V_{0h} := \{ v \in V_h \mid v = 0 \text{ sur } \partial \Omega \}.$$
 (11.2.2)

Notation 11.2.2. Dans la suite, on note N_{bord} le nombre de nœuds de degrés de liberté appartenant à $\partial\Omega$ et $N := N_{dl} - N_{bord}$. De plus, on suppose les nœuds σ_i , $1 \le i \le N_{dl}$, rangés de sorte que les σ_i , pour $i = 1, \ldots, N$, sont des points intérieurs de Ω tandis que les σ_i pour $i = N + 1, \ldots, N_{dl}$, sont sur $\partial\Omega$.

Proposition 11.2.3. L'espace V_h défini par (11.2.1) est un sous-espace de $H^1(\Omega)$ de dimension finie N_{dl} . De plus, il existe une base $(\phi_i)_{1 \le i \le N_{dl}}$ de V_h définie par

$$\phi_i(\sigma_j) = \delta_{ij} \quad pour \ tout \ 1 \le i, j \le N_{dl}, \tag{11.2.3}$$

où les σ_j sont les nœuds de degrés de liberté, et pour toute $v \in V_h$, on a

$$v = \sum_{i=1}^{N_{dl}} v(\sigma_i) \,\phi_i.$$
(11.2.4)

De même, l'espace V_{0h} défini par (11.2.2) est un sous-espace de $H_0^1(\Omega)$ de dimension finie $N = N_{dl} - N_{bord}$, où N_{bord} est le nombre nœuds sur $\partial\Omega$. De plus, pour toute $v \in V_{0h}$, on a

$$v = \sum_{i=1}^{N} v(\sigma_i) \phi_i.$$
 (11.2.5)

Remarque 11.2.4. L'égalité (11.2.5) est une conséquence directe de l'égalité (11.2.4) et de la Notation 11.2.2.

Démonstration. Soit $v_h \in V_h$. Comme v_h est continue sur Ω borné, $v_h \in L^2(\Omega)$. Ainsi, pour montrer que $v_h \in H^1(\Omega)$, il suffit de montrer que, pour tout $i = 1, \ldots, d, \ \partial_{x_i} v_h \in L^2(\Omega)$. Soient $i \in \{1, \ldots, d\}$ et $\varphi \in C_c^{\infty}(\Omega)$. On a

$$\langle \partial_{x_i} v_h, \varphi \rangle = -\int_{\Omega} v_h \frac{\partial \varphi}{\partial x_i} \, dx = -\sum_{K \in \mathcal{T}_h} \int_K v_{h|K} \frac{\partial \varphi}{\partial x_i} \, dx$$

Puisque $v_h \in V_h$, $v_{h|K} \in \mathbb{P}_1$ pour tout $K \in \mathcal{T}_h$, d'après la formule de Green on en déduit

$$\langle \partial_{x_i} v_h, \varphi \rangle = \sum_{K \in \mathcal{T}_h} \int_K \frac{\partial v_{h|K}}{\partial x_i} \varphi \, dx - \sum_{K \in \mathcal{T}_h} \int_{\partial K} v_{h|K} \varphi \, \nu_i^K \, dx, \qquad (11.2.6)$$

où $\nu^{K} := (\nu_{1}^{K}, \dots, \nu_{d}^{K})^{T}$ est la normale extérieure unitaire de ∂K . Or, si $K_{m}, K_{n} \in \mathcal{T}_{h}$ sont tels que $\Gamma := \partial K_{m} \cap \partial K_{n}$ est une face commune alors $\nu^{K_{n}} = -\nu^{K_{m}}$ sur Γ . Ainsi, comme v_{h} et φ sont continues dans $\overline{\Omega}$, on obtient

$$\int_{\partial K_n} v_{h|K} \varphi \,\nu_i^{K_n} \, dx + \int_{\partial K_m} v_{h|K} \varphi \,\nu_i^{K_m} \, dx = 0.$$

On en déduit que la somme des intégrales de bord de (11.2.6) se réduit à une intégrale sur le bord de Ω qui s'annule car $\varphi \in C_c^{\infty}(\Omega)$. Finalement, on a

$$\langle \partial_{x_i} v_h, \varphi \rangle = \sum_{K \in \mathcal{T}_h} \int_K \frac{\partial v_{h|K}}{\partial x_i} \varphi \, dx,$$

d'où $\partial_{x_i} v_h = \sum_{K \in \mathcal{T}_h} \partial_{x_i} v_{h|K} \in L^2(\Omega)$, donc $v_h \in H^1(\Omega)$.

Le Lemme 11.1.12 et la Remarque 11.1.13 montrent que les éléments de V_h sont obtenus localement sur chaque $K \in \mathcal{T}_h$ par des polynômes qui coïncident sur les degrés de liberté des faces. La base de V_h est obtenue en assemblant les bases $(\psi_j^K)_{1 \le j \le N_k}$ définies dans le Lemme 11.1.11 de chaque $K \in \mathcal{T}_h$ (voir l'exemple ci-dessous sur l'assemblage des ψ_j^K).

Exemple 11.2.5. On se place en dimension 2, on prend un maillage simple $\mathcal{T} = \{K_1, K_2\}$, où K_1 et K_2 ont une arête commune et on considère le cas k = 2. Alors, les treillis de K_1 et K_2 sont formés de 6 points et ont 3 points en commun, il y a donc 2 * 6 - 3 = 9 nœuds de degré de liberté. Soient $\sigma_1, \ldots, \sigma_9$ les nœuds du maillage répartis dans le sens trigonométrique (voir figure 11.3). Les nœuds communs à K_1 et K_2 sont alors σ_1, σ_3 et σ_9 . On note $\{\psi_j^i\}_{j\in J_i}$ la base de \mathbb{P}_2 associée à K_i , avec $J_1 = \{1, 3, 4, 5, 8, 9\}$ et $J_2 = \{1, 2, 3, 6, 7, 9\}$. On va assembler les fonctions ψ_j^i pour obtenir la base $\{\phi_k\}_{1\leq k\leq 9}$ de V_h de



FIGURE 11.3 – Représentation des triangles K_1 , K_2 et des treillis d'ordre 2 associés.

la façon suivante : Si $k \in \{1, 3, 9\}$, on pose :

$$\phi_k = \begin{cases} \psi_k^1 & \text{dans } K_1, \\ \psi_k^2 & \text{dans } K_2. \end{cases}$$
Si $k \in \{4, 5, 8\} \subset J_1$, on pose :

$$\phi_k = \begin{cases} \psi_k^1 & \text{dans } K_1, \\ 0 & \text{dans } K_2. \end{cases}$$

Enfin si $k \in \{2, 6, 7\} \subset J_2$, on pose :

$$\phi_k = \begin{cases} 0 & \text{dans } K_1, \\ \psi_k^2 & \text{dans } K_2. \end{cases}$$

D'après la Remarque 11.1.13, les fonctions ϕ_k sont bien continues sur $K_1 \cup K_2$.

Considérons maintenant la formulation variation nelle du problème de Dirichlet homogène sur Ω :

$$\begin{cases} \text{Trouver } u \in H_0^1(\Omega) \text{ telle que} \\ \forall v \in H_0^1(\Omega), \quad \int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx. \end{cases}$$
(11.2.7)

où $f \in L^2(\Omega)$. On approche (11.2.7) par la formulation variationnelle

$$\begin{cases} \text{Trouver } u_h(\sigma_1), \dots, u_h(\sigma_N) \in \mathbb{R} \text{ tels que} \\ \forall i = 1, \dots, N, \quad \sum_{j=1}^N u_h(\sigma_j) \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j \, dx = \int_{\Omega} f \, \phi_i \, dx, \end{cases}$$
(11.2.8)

où les σ_j sont les nœuds de degrés de liberté intérieurs (*i.e.* non sur le bord $\partial\Omega$), N est le nombre de nœuds intérieurs et ϕ_i est donnée par (11.2.3).

En posant $U_h := (u_h(\sigma_1), \dots, u_h(\sigma_N))^T \in \mathbb{R}^N$, la formulation variationnelle (11.2.8) s'écrit encore $\mathcal{K}_h U_h = b_h$ avec $\mathcal{K}_h \in \mathbb{R}^{N \times N}$ et $b_h \in \mathbb{R}^N$ donnés par

$$\mathcal{K}_h := \left(\int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j \ dx \right)_{1 \le i, j \le N}, \tag{11.2.9}$$

 et

$$b_h := \left(\int_{\Omega} f \phi_i \, dx \right)_{1 \le i \le N}. \tag{11.2.10}$$

Dans la section suivante, on verra comment calculer explicitement la matrice de rigidité \mathcal{K}_h et le second membre b_h . On termine tout d'abord cette section par une étude de la convergence de la méthode \mathbb{P}_k .

Définition 11.2.6. Pour tout *d*-simplexe non dégénéré K de \mathbb{R}^d , on note

diam(K) :=
$$\max_{x,y \in K} |x - y|$$
 et $\rho(K)$:= $\sup_{B_r \subset K} (2r)$,

en particulier la **rondeur** $\rho(K)$ désigne le diamètre du plus grand disque contenu dans K. Soit $(\mathcal{T}_h)_{h>0}$ une suite de maillages admissibles de Ω . On dit qu'il s'agit d'une suite de **maillages réguliers** si :

1. la suite $h := \max_{K \in \mathcal{T}_h} \operatorname{diam}(K)$ tend vers 0,

2. il existe une constante c > 0 telle que, pour tout h > 0 et tout $K \in \mathcal{T}_h$, on a

$$\frac{\operatorname{diam}(K)}{\rho(K)} \le c. \tag{11.2.11}$$

Remarque 11.2.7. La condition (11.2.11) signifie que les angles des *d*-simplexes K de \mathcal{T}_h ne sont pas trop aplatis.

Définition 11.2.8. On définit r_h l'opérateur d'interpolation $\mathbb{P}_k \det C(\overline{\Omega})$ dans V_h par :

$$r_h v := \sum_{j=1}^{N_{dl}} v(\sigma_i) \phi_i,$$
 (11.2.12)

où N_{dl} est le nombre de degrés de liberté et les ϕ_i sont les fonctions de base de V_h données par la Proposition 11.2.3.

Proposition 11.2.9 (Interpolation \mathbb{P}_k). Soit $(\mathcal{T}_h)_h$ une suite de maillages réguliers de Ω . On suppose $k + 1 > \frac{d}{2}$. Alors, pour toute $v \in H^{k+1}(\Omega)$, l'interpolée $r_h v$ est bien définie et il existe c > 0, indépendante de h et v, telle que

$$||v - r_h v||_{H^1(\Omega)} \le c h^k ||v||_{H^{k+1}(\Omega)}.$$
(11.2.13)

Remarque 11.2.10. Si $k + 1 > \frac{d}{2}$, alors $H^{k+1}(\Omega) \subset C(\overline{\Omega})$ donc les fonctions de $H^{k+1}(\Omega)$ sont bien définies en tout point de $\overline{\Omega}$ et $r_h v$ est bien définie.

Démonstration. La démonstration donnée ici a essentiellement pour but de justifier la nécessité de considérer des maillages réguliers et, en particulier, la majoration uniforme (11.2.11). On admettra le résultat technique donné plus bas (voir [1] et [12]) et essentiel pour la suite. On définit tout d'abord un opérateur d'interpolation local sur chaque d-simplexe du maillage. Soient K un d-simplexe, $k \in \mathbb{N}^*$ et Σ_k le treillis d'ordre k de K. On définit l'opérateur d'interpolation r_K , pour toute fonction continue v, par

$$r_K v := p \in \mathbb{P}_k$$
 où $p(x) = v(x), \forall x \in \Sigma_k.$ (11.2.14)

Lemme 11.2.11 (admis, voir [12]). On suppose que k + 1 > d/2 et que diam $(K) \le 1$. Alors, il existe une constante c > 0 indépendante de K telle que, pour toute $v \in H^{k+1}(K)$, on a

$$||v - r_K v||_{H^1(K)} \le c \ \frac{(\operatorname{diam}(K))^{k+1}}{\rho(K)} |v|_{H^{k+1}(K)}, \tag{11.2.15}$$

 $o\dot{u} \mid \cdot \mid_{H^{k+1}(K)}$ est la semi-norme sur $H^{k+1}(K)$ définie par

$$|v|_{H^{k+1}(K)}^2 := ||v||_{H^{k+1}(K)}^2 - ||v||_{H^k(K)}^2$$

Si $v \in H^{k+1}(\Omega)$, son interpolée $r_h v$ restreinte au *d*-simplexe K est $r_K v$. On en déduit

$$||v - r_h v||_{H^1(\Omega)}^2 = \sum_{K \in \mathcal{T}_h} ||v - r_K v||_{H^1(K)}^2.$$

Pour h > 0 suffisamment petit, on peut supposer diam $(K) \leq 1$ pour tout $K \in \mathcal{T}_h$. Alors, d'après le Lemme 11.2.11 et la majoration (11.2.11), on obtient

$$||v - r_h v||_{H^1(\Omega)}^2 \le c \, h^{2k} \sum_{K \in \mathcal{T}_h} |v|_{H^{k+1}(K)}^2 \le c \, h^{2k} ||v||_{H^{k+1}(\Omega)}^2,$$

d'où le résultat.

Théorème 11.2.12. Soit $(\mathcal{T}_h)_{h>0}$ une suite de maillages réguliers de Ω . Soient $u \in H_0^1(\Omega)$ la solution de (11.2.7) et $u_h \in V_{0h}$ la solution de (11.2.8). Alors, la méthode des éléments finis \mathbb{P}_k converge :

$$\lim_{h \to 0} ||u - u_h||_{H^1(\Omega)} = 0.$$
(11.2.16)

De plus, si $u \in H^{k+1}(\Omega)$ et si $k+1 > \frac{d}{2}$ alors il existe c > 0 telle que

$$||u - u_h||_{H^1(\Omega)} \le c h^k ||u||_{H^{k+1}(\Omega)}.$$
(11.2.17)

Démonstration. On applique le Théorème 9.2.2 avec $\mathcal{V} = C_c^{\infty}(\Omega)$ qui est dense dans $H_0^1(\Omega)$. Comme $C_c^{\infty}(\Omega) \subset H^{k+1}(\Omega)$, l'estimation (11.2.13) entraîne que les hypothèses du Théorème 9.2.2 sont vérifiées, ce qui donne (11.2.16).

L'estimation d'erreur (11.2.17) s'obtient à partir du Lemme de Céa

$$||u - u_h||_{H^1(\Omega)} \le c \inf_{v_h \in V_{0h}} ||u - v_h||_{H^1(\Omega)} \le c ||u - r_h u||_{H^1(\Omega)}.$$

En appliquant ensuite l'estimation (11.2.13), on obtient le résultat.

11.3 Calcul du second membre et assemblage de la matrice de rigidité

Dans cette section, on présente sur l'exemple du problème de Dirichlet homogène, le calcul pratique des coefficients du système matriciel obtenu par la méthode des éléments finis \mathbb{P}_1 en dimension $d \geq 2$.

11.3.1 Calcul du second membre

D'après (11.2.10), le second membre est le vecteur $b_h \in \mathbb{R}^N$ de coefficients

$$(b_h)_i := \int_{\Omega} f \phi_i \, dx = \sum_{K \in \mathcal{T}_h} \int_K f \phi_i \, dx, \quad \forall \ i = 1, \dots, N,$$

où $f \in L^2(\Omega)$ est le terme source et les ϕ_i sont les fonctions de base de V_h . Comme dans le cas 1d, b_h se calcule par des formules de quadratures sur chaque $K \in \mathcal{T}_h$. On donne ci-dessous l'équivalent des formules en dimension 1 du point milieu et du trapèze.

Soit K un d-simplexe de sommets a_i , i = 1, ..., d + 1. On note $a_0 := (1+d)^{-1} \sum_{i=1}^{d+1} a_i$ le barycentre de K. Alors, pour toute fonction ψ intégrable, on a

$$\int_{K} \psi(x) \, dx \simeq |K| \, \psi(a_0), \tag{11.3.1}$$

 et

$$\int_{K} \psi(x) \, dx \simeq \frac{|K|}{d+1} \, \sum_{i=1}^{d+1} \psi(a_i). \tag{11.3.2}$$

Ces deux formules sont exactes pour $\psi \in \mathbb{P}_1$. En particulier, on peut montrer alors qu'elles sont approchées à l'ordre 2 en h pour des fonctions régulières.

Remarque 11.3.1. On peut montrer que le résultat de convergence reste vrai en remplaçant le second membre b_h par son approximation via une formule de quadrature (voir [12]).

11.3.2 Assemblage de la matrice de rigidité

D'après (11.2.9), la matrice de rigidité $\mathcal{K}_h \in \mathbb{R}^{N \times N}$ a pour coefficients

$$(\mathcal{K}_h)_{ij} := \int_{\Omega} \nabla \phi_i \cdot \nabla \phi_j \, dx, \quad \forall \ i, j = 1, \dots, N$$

Pour calculer ces coefficients, il est nécessaire de pouvoir déterminer les fonctions de base ϕ_i . Celles-ci dépendant du maillage considéré. Ci-dessous, on étudie un cas simple. On suppose $\Omega := [-1, 1]^2$ et \mathcal{T}_h le maillage donné dans la figure 11.4, où les numéros sont ceux des nœuds du maillages (par exemple $\sigma_1 := (-1, 1)$ et $\sigma_7 := (0, -1)$).



FIGURE 11.4 – Maillage de $[-1, 1]^2$ avec numérotation des nœuds.

En utilisant les propriétés de symétrie on va voir que le calcul de \mathcal{K} se ramène simplement au calcul de quelques coefficients. Considérons d'abord les termes diagonaux. Si $i \neq 9$, le nœud σ_i n'est commun qu'à deux mailles et $\nabla \phi_i$ est constant sur ces deux mailles. De plus, on a

$$\phi_1(x,y) = \phi_2(-x,y) = \phi_3(-x,-y) = \phi_4(x,-y),$$

dont on déduit

$$(\mathcal{K}_h)_{11} = (\mathcal{K}_h)_{22} = (\mathcal{K}_h)_{33} = (\mathcal{K}_h)_{44}.$$

Déterminons $(\mathcal{K}_h)_{11}$. Comme $\sup(\phi_1) \subset K_1 \cup K_2$, on a

$$(\mathcal{K}_h)_{11} = \int_{K_1} |\nabla \phi_1|^2 \, dx + \int_{K_2} |\nabla \phi_1|^2 \, dx$$

Sur $K_1, \phi_1 \in \mathbb{P}_1, \phi_1(-1, 1) = 1, \phi_1(0, 1) = 0$ et $\phi_1(0, 0) = 0$ donc $\phi_1(x, y) = -x$. Sur K_2 , on obtient de même $\phi_1(x, y) = y$. Alors, on a $(\mathcal{K}_h)_{11} = |K_1| + |K_2| = 1$.

De même, on montre facilement que $(\mathcal{K}_h)_{55} = (\mathcal{K}_h)_{77} = 2$ et $(\mathcal{K}_h)_{66} = (\mathcal{K}_h)_{88} = 2$. Pour le dernier coefficient de la diagonale, on remarque que le calcul se ramène simplement à un calcul sur $K_1 \cup K_2$. En effet, on a

$$(\mathcal{K}_h)_{99} = \sum_{i=1}^{9} \int_{K_i} |\nabla \phi_9|^2 \, dx = 4 \left(\int_{K_1} |\nabla \phi_9|^2 \, dx + \int_{K_2} |\nabla \phi_9|^2 \, dx \right).$$

En procédant comme pour le calcul de ϕ_1 , on obtient alors $(\mathcal{K}_h)_{99} = 4$.

Il reste à calculer les termes non diagonaux. Comme $\operatorname{supp}(\phi_1) \subset K_1 \cup K_2$, on a $(\mathcal{K}_h)_{12} = (\mathcal{K}_h)_{13} = (\mathcal{K}_h)_{14} = (\mathcal{K}_h)_{16} = (\mathcal{K}_h)_{17} = 0$. De plus,

$$(\mathcal{K}_h)_{15} = \int_{K_1} \nabla \phi_1 \cdot \nabla \phi_5 \ dx = \frac{1}{2} \left(\nabla \phi_1 \cdot \nabla \phi_5 \right),$$

puisque $\nabla \phi_1$ et $\nabla \phi_5$ sont constants sur K_1 . Sur K_1 , $\phi_5(x, y) = \alpha + \beta x + \gamma y$ et $\phi_5(0, 0) = 0$, $\phi_5(0, 1) = 1$ et $\phi_5(-1, 1) = 0$ d'où $\phi_5(x, y) = x + y$. On en déduit $(\mathcal{K}_h)_{15} = -1/2$. De même, on obtient $(\mathcal{K}_h)_{18} = 1/2$ et $(\mathcal{K}_h)_{19} = 0$. Le seul terme restant à calculer est $(\mathcal{K}_h)_{59}$ et, comme précédemment, on obtient $(\mathcal{K}_h)_{59} = -1$. En utilisant les propriétés de symétrie, on en déduit les autres coefficients de \mathcal{K}_h . Finalement, on obtient

$$\mathcal{K}_{h} = \begin{pmatrix} 1 & 0 & 0 & 0 & -1/2 & 0 & 0 & -1/2 & 0 \\ 0 & 1 & 0 & 0 & -1/2 & -1/2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1/2 & -1/2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1/2 & -1/2 & 0 \\ -1/2 & -1/2 & 0 & 0 & 2 & 0 & 0 & -1 \\ 0 & -1/2 & -1/2 & 0 & 0 & 2 & 0 & 0 & -1 \\ 0 & 0 & -1/2 & -1/2 & 0 & 0 & 2 & 0 & -1 \\ -1/2 & 0 & 0 & -1/2 & 0 & 0 & 2 & 0 & -1 \\ -1/2 & 0 & 0 & -1/2 & 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & -1 & -1 & -1 & 4 \end{pmatrix}$$

Exercice 11.3.2. Calculer la matrice de rigidité \mathcal{K}_h dans le cas du maillage de la figure 11.5.



FIGURE 11.5 – Autre maillage de $[-1, 1]^2$ avec numérotation des nœuds.

11.4 Remarques sur le cas Ω non polyédrique

Si Ω n'est pas polyédrique, alors on approche Ω par un ouvert polyédrique Ω_h dont les sommets sont sur la frontière de Ω . Ensuite, on maille Ω_h par une triangulation \mathcal{T}_h . L'ouvert Ω_h peut, par exemple, être choisi tel qu'il existe une constante c > 0 vérifiant

$$\operatorname{dist}(\partial\Omega, \partial\Omega_h) \le c h^2.$$

On obtient alors, sous de bonnes conditions (voir [12]), le même résultat de convergence pour la méthode \mathbb{P}_1 . Pour la méthode \mathbb{P}_k avec $k \geq 2$, la convergence de la solution approchée vers la solution exacte est d'ordre moins élevé que dans le cas d'une ouvert Ω polyédrique.

Pour obtenir une convergence plus rapide, *i.e.* d'ordre plus élevé, on peut faire appel à des éléments finis isoparamétriques. Il s'agit de mailler la partie de Ω près du bord par des mailles à bords courbes obtenus par déformation de *d*-simplexes. Pour des compléments sur ces différents points, on renvoie à [12].

11.5 Maillages rectangulaires

On termine ce chapitre par une rapide présentation du cas d'un maillage de Ω par des rectangles. Dans cette section on suppose l'ouvert Ω rectangulaire. On définit un *d*-rectangle K de \mathbb{R}^d comme étant le pavé (non dégénéré) $\prod_{i=1}^d [l_i, L_i]$ où $0 < l_i < L_i < +\infty$. On note $(a_j)_{1 < j < 2^d}$ les sommets de K.

Définition 11.5.1. Un maillage rectangulaire de $\overline{\Omega}$ est un ensemble \mathcal{T}_h de *d*-rectangles (non dégénérés) $(K_i)_{1 \le i \le n}$ qui vérifient

- 1. $K_i \subset \overline{\Omega}$ et $\overline{\Omega} = \bigcup_{i=1}^n K_i$,
- 2. l'intersection $K_i \cap K_j$ de deux *d*-rectangles distincts est *m*-rectangle, avec $0 \le m \le d-1$, dont tous les sommets sont aussi des sommets de K_i et K_j .

Le paramètre h est défini par

$$h := \max_{1 \le i \le n} \operatorname{diam}(K_i), \quad \text{où} \quad \operatorname{diam}(K_i) := \sup\{|x - y| \mid x, y \in K_i\},$$

où $|\cdot|$ désigne la norme euclidienne. Les **sommets** (ou **nœuds**) du maillage \mathcal{T}_h sont les sommets des *d*-rectangles K_i .

Notation 11.5.2. On note \mathbb{Q}_k , l'espace des polynômes à coefficients réels de \mathbb{R}^d de degré inférieur ou égal à k par rapport à chaque variable, *i.e.* tout polynôme p de \mathbb{Q}_k est de la forme

$$\forall x \in \mathbb{R}^{d}, \quad p(x) = \sum_{0 \le i_1 \le k, \dots, 0 \le i_d \le k} \alpha_{i_1, \dots, i_d} \, x_1^{i_1} \dots \, x_d^{i_d}, \tag{11.5.1}$$

où $\alpha_{i_1,\ldots,i_d} \in \mathbb{R}$.

Définition 11.5.3. Soit K un d-rectangle non dégénéré de \mathbb{R}^d . On appelle **treillis d'ordre** $k \in \mathbb{N}^*$ de K l'ensemble

$$\Sigma_k := \left\{ x \in K \mid \frac{x_j - l_j}{L_j - l_j} \in \left\{ 0, \frac{1}{k}, \dots, \frac{k - 1}{k}, 1 \right\}, \ \forall \ j = 1, \dots, d \right\}.$$
(11.5.2)

Définition 11.5.4. Soit \mathcal{T}_h un maillage rectangulaire de Ω . La méthode des éléments finis \mathbb{Q}_k est définie comme étant la méthode pour laquelle l'espace $H^1(\Omega)$ est approché par l'espace

$$V_h := \{ v \in C(\overline{\Omega}) \mid v_K \in \mathbb{Q}_k, \ \forall \ K \in \mathcal{T}_h \}.$$
(11.5.3)

Avec ses définitions, on obtient les mêmes résultats que dans le cas des maillages triangulaires.

Chapitre 12

Théorie spectrale : analyse théorique et numérique

Ce chapitre rappelle certains résultats de théorie spectrale (voir le cours de master 1, [11]) et servira de transition entre problèmes stationnaires et problèmes instationnaires. On donne tout d'abord deux exemples précisant comment la théorie spectrale intervient dans la résolution de problèmes instationnaires.

Exemple 12.0.5. (Système différentiel en dimension finie) Considérons le problème suivant : trouver $u \in C^1(\mathbb{R}_+; \mathbb{R}^d)$ solution de

$$\begin{cases} \partial_t u + Au = 0, \quad \forall \ t > 0, \\ u(0) = u_0, \end{cases}$$
(12.0.1)

où $A \in \mathbb{R}^{d \times d}$ et $u_0 \in \mathbb{R}^d$ sont fixés. Si A est symétrique définie positive, on sait que le problème (12.0.1) admet une unique solution $u \in C^1(\mathbb{R}_+; \mathbb{R}^d)$. Plus précisément, soient $\lambda_1, \ldots, \lambda_d$ les valeurs propres de A (qui est diagonalisable puisque symétrique définie positive) et v_1, \ldots, v_d une base de vecteurs propres associés. Alors, on a

$$u_0 = \sum_{k=1}^d u_k^0 v_k$$
, et $u(t) = \sum_{k=1}^d u_k^0 e^{-\lambda_k t} v_k$.

Autrement dit, la résolution du problème d'évolution (12.0.1) passe par la diagonalisation de la matrice A.

Exemple 12.0.6. (Équation de la chaleur) On rappelle que l'équation de la chaleur est donnée (sans condition initiale ni condition aux limites) par : $\partial_t u - \Delta u = 0$ où $t \in \mathbb{R}^+_*$ et $x \in \Omega$. Comparant avec le problème (12.0.1), on peut voir que l'on a "échangé" la matrice A par l'opérateur $-\Delta$. Autrement dit, on est passé d'un problème en dimension finie (A représente une application linéaire sur un espace vectoriel de dimension finie) à un problème en dimension infinie ($-\Delta$ est une application linéaire en dimension infinie). Par analogie avec le problème (12.0.1), on détermine une solution de l'équation de la chaleur en diagonalisant l'opérateur $-\Delta$.

On peut retrouver ce raisonnement en cherchant une solution de l'équation de la chaleur à variables séparables. Supposons donc $u(t, x) = \phi(t)v(x)$. Alors, on obtient

$$\phi'(t)v(x) - \phi(t)\Delta v(x) = 0,$$

d'où

$$\frac{\phi'(t)}{\phi(t)} = \frac{\Delta v(x)}{v(x)}.$$

Dans l'égalité précédente, le terme de gauche ne dépend que de la variable t et celui de droite ne dépend que de la variable x donc les deux sont constants. On en déduit

$$\frac{\phi'(t)}{\phi(t)} = \frac{\Delta v(x)}{v(x)} = -\lambda,$$

où $\lambda \in \mathbb{R}$. Finalement, on obtient

$$\phi(t) = c e^{-\lambda t}$$
 et $-\Delta v = \lambda v$,

donc λ est une valeur propre de Δ .

12.1 Rappels

Dans cette section, on rappelle (sans démonstrations) les résultats de théorie spectrale pour les opérateurs auto-adjoints compacts et on renvoie à [11] pour plus de détails. Dans toute la suite, $(V, (\cdot, \cdot)_V)$ est un espace de Hilbert réel.

Définition 12.1.1. Soit $T \in \mathcal{L}(V)$.

- 1. Un réel $\lambda \in \mathbb{R}$ est appelé valeur propre de T s'il existe $u \in V$ non nul tel que $Au = \lambda u$. On dit que u est un vecteur propre de T associé à la valeur propre λ .
- 2. On appelle **adjoint** de T l'unique $T^* \in \mathcal{L}(V)$ tel que

$$(Tx, y)_V = (x, T^*y)_V, \quad \forall x, y \in V.$$

De plus, on dit que T est **auto-adjoint** si $T = T^*$.

3. On dit que T est **défini positif** si

$$(Tx, x)_V > 0, \quad \forall \ x \in V \setminus \{0\}.$$

On rappelle aussi la notion d'opérateur compact.

Définition 12.1.2. Soient E, F deux espaces de Banach sur $\mathbb{K} := \mathbb{R}$ ou \mathbb{C} et $T \in \mathcal{L}(E, F)$. On dit que T est **compact** si l'image par T de la boule unité fermée de E est relativement compacte dans F. On note $\mathcal{K}(E, F)$ l'espace des opérateurs compacts de E dans F.

Remarque 12.1.3.

- 1. On rappelle qu'une partie K de F est **relativement compacte** si son adhérence est compacte dans F. Autrement dit, K est relativement compacte dans F si et seulement si de toute suite de K on peut en extraire une sous-suite convergente dans F.
- 2. Le théorème de Riesz affirme que la boule unité d'un espace vectoriel normé *E* est relativement compacte si et seulement si *E* est de dimension finie. On en déduit en particulier que l'opérateur identité sur *E* est un opérateur compact si et seulement si *E* est de dimension finie.

Théorème 12.1.4 (Théorème spectral des opérateurs auto-ajoints compacts). Soient $(V, (\cdot, \cdot)_V)$ un espace de Hilbert réel de dimension infinie et $T \in \mathcal{K}(V)$ défini positif et auto-adjoint. Alors, les valeurs propres de T forment une suite $(\lambda_k)_{k\geq 1}$ de \mathbb{R}^*_+ qui tend vers 0 et il existe une base hilbertienne $(u_k)_{k\geq 1}$ de V formée de vecteur propres, i.e. qui vérifient

$$u_k \in V$$
 et $Tu_k = \lambda_k u_k, \quad \forall \ k \ge 1.$

Remarque 12.1.5. On rappelle qu'une **base hilbertienne** $(e_n)_{n\geq 1}$ de V est une famille dénombrable orthonormale de V engendrant un sous-espace vectoriel dense dans V. De plus, si $(u_k)_{k\geq 1}$ est une base hilbertienne de V, on a alors

$$v = \sum_{k \ge 1} (v, u_k)_V u_k$$
 et $||v||^2 = \sum_{k \ge 1} |(v, u_k)_V|^2$, $\forall v \in V$.

12.2 Application au cadre variationnel

On va appliquer le théorème spectral des opérateurs compact au cadre variationnel. Soient $(V, (\cdot, \cdot)_V)$ et $(H, (\cdot, \cdot)_H)$ deux espaces de Hilbert réels tels que

$$V \subset H$$
 avec injection compacte,
 V est dense dans H . (12.2.1)

Soit $a(\cdot, \cdot)$ une forme bilinéaire symétrique, continue et coercive sur V. On considère le problème spectral variationnel suivant :

$$\begin{cases} \text{Trouver } \lambda \in \mathbb{R} \text{ et } u \in V \setminus \{0\} \text{ tels que} \\ a(u,v) = \lambda (u,v)_H, \quad \forall v \in V. \end{cases}$$
(12.2.2)

Définition 12.2.1. Si $\lambda \in \mathbb{R}$ et $u \in V \setminus \{0\}$ vérifient (12.2.2), on dit que λ est une valeur propre de la formulation variationnelle (12.2.2) et que u est un vecteur propre associé.

Théorème 12.2.2. Les valeurs propres de (12.2.2) forment une suite croissante $(\lambda_k)_{k\geq 1}$ de réels positifs et il existe une base hilbertienne de H formée de vecteur propres $(u_k)_{k\geq 1}$, i.e. qui vérifient

$$u_k \in V \setminus \{0\}$$
 et $a(u_k, v) = \lambda_k (u_k, v)_H$, $\forall v \in V$.

De plus, $(u_k/\sqrt{\lambda_k})_{k\geq 1}$ est une base hilbertienne de V pour le produit scalaire $a(\cdot, \cdot)$.

Démonstration. Pour $f \in H$, on considère le problème :

$$\begin{cases} \text{Trouver } u \in V \text{ tel que} \\ a(u,v) = (f,v)_H, \quad \forall v \in V. \end{cases}$$
(12.2.3)

D'après le théorème de Lax-Milgram, le problème (12.2.3) admet une solution unique $u \in V$. On définit l'opérateur \mathcal{A} de H dans V par $\mathcal{A}f := u$. Autrement dit, \mathcal{A} est l'opérateur qui à $f \in H$ associe la solution $u \in V$ de (12.2.3). L'injection \mathcal{I} de V dans H est

continue donc $||v||_H \leq c ||v||_V$ pour tout $v \in V$. En prenant $v = \mathcal{A}f$ comme fonction test dans (12.2.3), on obtient

$$m ||\mathcal{A}f||_{V}^{2} \leq a(\mathcal{A}f, \mathcal{A}f) = (f, \mathcal{A}f)_{H} \leq ||f||_{H} ||\mathcal{A}f||_{H} \leq c ||f||_{H} ||\mathcal{A}f||_{V}$$

On en déduit $\mathcal{A} \in \mathcal{L}(H, V)$. On pose $T := \mathcal{I}\mathcal{A} \in \mathcal{L}(H)$. Un produit d'opérateur est compact si l'un des deux opérateurs est compact, or l'injection \mathcal{I} de V dans H est compact donc $T \in \mathcal{K}(H)$. Soient $f, g \in H$. En prenant $v = \mathcal{A}g$ comme fonction test dans (12.2.3), on obtient

$$(f,Tg)_H = (f,\mathcal{A}g)_H = a(\mathcal{A}f,\mathcal{A}g) = a(\mathcal{A}g,\mathcal{A}f) = (g,\mathcal{A}f)_H = (g,Tf)_H,$$

donc T est auto-adjoint et, de plus, défini positif dans H car $a(\cdot, \cdot)$ est coercive. Le théorème spectral appliqué à T entraı̂ne qu'il existe une suite décroissante $(\mu_k)_{k\geq 1}$ de réels positifs qui tend vers 0 et une base hilbertienne $(u_k)_{k\geq 1}$ de H formée de vecteurs propres de T, *i.e.*

$$u_k \in V$$
 et $Tu_k = \mu_k u_k, \quad \forall \ k \ge 1.$

Le problème (12.2.2) s'écrit encore

$$a(u, v) = \lambda (u, v)_H = \lambda a(\mathcal{A}u, v), \quad \forall v \in V.$$

Ce qui équivaut à $a(u - \lambda Au, v) = 0$ pour tout $v \in V$. Comme $a(\cdot, \cdot)$ est un produit scalaire sur V, on en déduit $u = \lambda Au = \lambda Tu$. Ainsi, les valeurs propres de (12.2.2) sont les inverses des valeurs propres de T et les vecteurs propres sont les mêmes. Pour $k \ge 1$, on pose

$$\lambda_k := \frac{1}{\mu_k}$$
 et $v_k := \frac{u_k}{\sqrt{\lambda_k}}$.

Il reste à vérifier que $(v_k)_{k\geq 1}$ est une base hilbertienne de V pour le produit scalaire $a(\cdot, \cdot)$. Pour $k, j \geq 1$, on a

$$a(v_k, v_j) = \frac{a(u_k, u_j)}{\sqrt{\lambda_k \lambda_j}} = \lambda_k \frac{(u_k, u_j)_H}{\sqrt{\lambda_k \lambda_j}} = \delta_{kj},$$

car $(u_k)_{k\geq 1}$ est une base hilbertienne de H. Enfin, on a le résultat en remarquant que l'orthogonal de $(v_k)_{k\geq 1}$ dans V est contenu dans l'orthogonal de $(u_k)_{k\geq 1}$ dans H qui est réduit à $\{0\}$.

On va appliquer le résultat précédent pour l'opérateur laplacien. La formulation faible du problème $-\Delta u = \lambda u$ dans Ω avec u = 0 sur $\partial \Omega$ est donnée par

$$\begin{cases} \text{Trouver } u \in H_0^1(\Omega) \text{ tel que} \\ \int_{\Omega} \nabla u \cdot \nabla v \ dx = \lambda \int_{\Omega} u v \ dx, \quad \forall \ v \in H_0^1(\Omega). \end{cases}$$

On obtient bien un problème du type (12.2.2) avec $H := L^2(\Omega), V := H_0^1(\Omega)$ et $a(\cdot, \cdot)$ la forme bilinéaire symétrique définie sur V par

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

L'injection de V dans H est compacte d'après le théorème de Rellich. De plus l'espace $C_c^{\infty}(\Omega)$ étant dense dans $L^2(\Omega)$ et $H_0^1(\Omega)$, $H_0^1(\Omega)$ est dense dans $L^2(\Omega)$. On est donc bien dans les conditions du Théorème 12.2.2 et on en déduit le résultat suivant :

Corollaire 12.2.3. Soit Ω un ouvert borné régulier de classe C^1 de \mathbb{R}^d . Alors, il existe une suite croissante $(\lambda_k)_{k\geq 1}$ de réels positifs et une base hilbertienne $(u_k)_{k\geq 1}$ de $L^2(\Omega)$ telle que

$$u_k \in H_0^1(\Omega)$$
 et $-\Delta u_k = \lambda_k u_k$ p.p. dans Ω . (12.2.4)

On termine cette section par un résultat de régularité des fonctions propres du laplacien.

Proposition 12.2.4. Soit Ω un ouvert borné régulier de classe C^{∞} . Alors, les fonctions propres du laplacien vérifiant (12.2.4) appartiennent à $C^{\infty}(\overline{\Omega})$.

Démonstration. Si $u_k \in H_0^1(\Omega)$ vérifie (12.2.4), en particulier on a $\lambda_k u_k \in H^1(\Omega)$. D'après le théorème de régularité elliptique, on en déduit $u_k \in H^3(\Omega)$. Alors, $\lambda_k u_k \in H^3(\Omega)$ et, par régularité elliptique, $u_k \in H^5(\Omega)$. Ainsi, par récurrence on déduit $u_k \in H^m(\Omega)$, pour tout $m \in \mathbb{N}^*$. Enfin, puisque Ω est C^{∞} , d'après le Théorème 7.1.5 on a $u_k \in C^{\infty}(\overline{\Omega})$. \Box

12.3 Analyse numérique spectrale

Comme dans le cadre de la méthode des éléments finis des problèmes elliptiques, on donne tout d'abord un résultat général dans le cadre des espaces de Hilbert puis son application pour le laplacien.

Soient $(V, (\cdot, \cdot)_V)$, $(H, (\cdot, \cdot)_H)$ deux espaces de Hilbert vérifiant (12.2.1) et V_h un sousespace vectoriel de V de dimension finie. Soit $a(\cdot, \cdot)$ une forme bilinéaire symétrique, continue et coercive sur V. On considère le problème spectral variationnel suivant :

$$\begin{cases} \text{Trouver } \lambda_h \in \mathbb{R} \text{ et } u_h \in V_h \setminus \{0\} \text{ tels que} \\ a(u_h, v_h) = \lambda_h (u_h, v_h)_H, \quad \forall v_h \in V_h. \end{cases}$$
(12.3.1)

Proposition 12.3.1. Les valeurs propres de (12.3.1) forment une suite croissante finie :

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{N_{dl}}, \quad o\hat{u} \; N_{dl} := \dim(V_h),$$

et il existe une base de V_h , orthonormale dans H, $(u_{k,h})_{1 \le k \le N_{dl}}$ de vecteurs propres associés, i.e. qui vérifient

$$u_{k,h} \in V_h$$
 et $a(u_{h,k}, v_h) = \lambda_k (u_{h,k}, v_h)_H$, $\forall v_h \in V_h$.

Démonstration. Soit $\{\phi_i\}_{i=1,\dots,N_{dl}}$ une base de V_h . Alors, si $u_h \in V_h$, on a

3.7

$$u_h = \sum_{j=1}^{N_{dl}} U_j^h \phi_j, \quad \text{où } U_j^h \in \mathbb{R}, \ \forall \ j = 1, \dots, N_{dl}.$$

Pour que le couple (λ_h, u_h) soit solution de (12.3.1), il suffit qu'il vérifie (12.3.1) pour tout $v = \phi_i, i = 1, ..., N_{dl}$, et donc (12.3.1) est équivalent à

$$\begin{cases} \text{Trouver } \lambda_h \in \mathbb{R} \text{ et } U_1^h, \dots, U_{N_{dl}}^h \in \mathbb{R} \text{ tels que} \\ \sum_{j=1}^{N_{dl}} U_j^h a(\phi_j, \phi_i) = \lambda_h \sum_{j=1}^{N_{dl}} U_j^h (\phi_j, \phi_i)_H, \quad \forall i = 1, \dots, N_{dl} \end{cases}$$

On définit la **matrice de masse** $\mathcal{M}_h \in \mathbb{R}^{N_{dl} \times N_{dl}}$ par

$$(\mathcal{M}_h)_{ij} := (\phi_i, \phi_j)_H,$$

et la matrice de rigidité $\mathcal{K}_h \in \mathbb{R}^{N_{dl} \times N_{dl}}$ par

$$(\mathcal{K}_h)_{ij} := a(\phi_i, \phi_j).$$

Alors, le problème (12.3.1) est équivalent à

$$\mathcal{K}_h U_h = \lambda_h \, \mathcal{M}_h U_h,$$

où $U_h := (U_1^h, \ldots, U_{N_{dl}}^h)$. De plus, les matrices \mathcal{K}_h et \mathcal{M}_h sont symétriques définies positives. Alors, en appliquant le théorème de réduction simultanée (théorème 2.3.6 de [2]), on obtient qu'il existe $P_h \in \mathbb{R}^{N_{dl} \times N_{dl}}$ inversible telle que

$$\mathcal{M}_h = P_h P_h^T$$
 et $\mathcal{K}_h = P_h D P_h^T$,

où $D := \text{diag}(\lambda_1, \ldots, \lambda_{N_{dl}})$ et les λ_i sont les valeurs propres de \mathcal{K}_h . Alors, le problème (12.3.1) s'écrit encore

$$D\tilde{U}_h = \lambda_h \tilde{U}_h, \quad \text{où} \quad \tilde{U}_h := P_h^T U_h$$

Autrement dit, les valeurs propres de (12.3.1) sont les valeurs propres de \mathcal{K}_h et les vecteurs propres associés sont les $u_{h,k}$ donnés par

$$u_{h,k} = \sum_{j=1}^{N_{dl}} U_j^h \phi_j \quad \text{où} \quad U_{h,k} = (P_h^T)^{-1} \tilde{U}_{h,k} = (P_h^T)^{-1} e^k,$$

 $(D \text{ étant diagonale, ses vecteurs propres sont les vecteurs de base } e^k)$. On en déduit que la famille de vecteur propres $\{u_{h,k}\}_{k=1,\ldots,N_{dl}}$ forme bien une base de V_h . \Box

Remarque 12.3.2. En pratique, on résout le problème $\mathcal{K}_h U_h = \lambda_h \mathcal{M}_h U_h$ pour calculer l'approximation u_h . Cette résolution passe par la réduction simultanée de \mathcal{K}_h et U_h pour laquelle des algorithmes existent (voir [2]).

Dans le cas du problème de Dirichlet, il suffit d'appliquer le résultat précédent au cas du problème (12.2.4) en prenant pour V_{0h} le sous-espace de $H_0^1(\Omega)$ défini par la méthode des éléments finis \mathbb{P}_k . On donne ci-dessous un exemple en dimension 1 avec la méthode \mathbb{P}_1 .

On considère le problème de Dirichlet homogène sur l'intervalle [0, 1]:

$$\begin{cases} -u_k'' = \lambda_k u_k & \text{dans }]0,1[, \\ u_k(0) = u_k(1) & = 0. \end{cases}$$

Soit $x_0 = 0 < x_1 < \cdots < x_N < x_{N+1} = 1$ une discrétisation uniforme de l'intervalle [0, 1] de pas h > 0. On pose

 $V_{0h} := \{ v \in C([0,1]) \mid v_{||x_i, x_{i+1}|} \in \mathbb{P}_1, \ \forall \ i = 0, \dots, N, \ \text{et} \ v(0) = v(1) = 0 \}.$

On a vu que $\mathcal{K}_h \in \mathbb{R}^{N \times N}$ est donnée par

$$\mathcal{K}_{h} = \frac{1}{h} \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ \vdots & \ddots & & \ddots & \vdots \\ \vdots & & \ddots & \ddots & -1 \\ 0 & \dots & \dots & -1 & 2 \end{pmatrix}.$$

On se propose de calculer la matrice de masse \mathcal{M}_h correspondante. Soient $\{\phi_1, \ldots, \phi_N\}$ la base de V_{0h} . Alors, on a

$$\forall i = 1, \dots, N, \quad \phi_i(x) = \begin{cases} \frac{x - x_i}{h} & \text{si } x \in [x_{i-1}, x_i], \\ \frac{x_{i+1} - x}{h} & \text{si } x \in [x_i, x_{i+1}], \\ 0 & \text{sinon.} \end{cases}$$

La matrice de masse $\mathcal{M}_h \in \mathbb{R}^{N \times N}$ a pour coefficients

$$(\mathcal{M}_h)_{ij} = (\phi_i, \phi_j)_{L^2(0,1)}.$$

D'une part, puisque $\operatorname{supp}(\phi) \subset [x_{i-1}, x_{i+1}]$, on a

$$(\mathcal{M}_h)_{ij} = 0 \quad \text{si } |i-j| > 1.$$

D'autre part, on a

$$(\mathcal{M}_h)_{ii} = ||\phi_i||_{L^2(0,1)}^2 = \int_{x_{i-1}}^{x_i} \frac{(x-x_i)^2}{h^2} \, dx + \int_{x_i}^{x_{i+1}} \frac{(x_{i+1}-x)^2}{h^2} \, dx = \frac{h^3}{3h^2} + \frac{h^3}{3h^2} = \frac{2h}{3},$$

et, par changement de variable,

$$(\mathcal{M}_h)_{i(i+1)} = \int_{x_i}^{x_{i+1}} \frac{(x_{i+1} - x)(x - x_i)}{h^2} \, dx = \frac{1}{h^2} \int_0^h (h - y)y \, dx = \frac{h}{6} = (\mathcal{M}_h)_{i(i-1)}.$$

Finalement, la matrice de masse $\mathcal{M}_h \in \mathbb{R}^{N \times N}$ est donnée par

$$\mathcal{M}_{h} = h \begin{pmatrix} 2/3 & 1/6 & 0 & \dots & 0 \\ 1/6 & 2/3 & 1/6 & \dots & 0 \\ \vdots & \ddots & & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 1/6 \\ 0 & \dots & \dots & 1/6 & 2/3 \end{pmatrix}.$$

On donne ci-dessous un résultat de convergence de la méthode. Il faut prendre garde que seules les premières valeurs propres $\lambda_{k,h}$ sont de bonnes approximations des valeurs propres exactes. Pour obtenir un plus grand nombre d'approximations correctes, il est nécessaire de raffiner le maillage considéré. **Théorème 12.3.3.** Soient Ω un ouvert polyédrique de \mathbb{R}^d et $(\mathcal{T}_h)_{h>0}$ une suite de maillages triangulaires réguliers de Ω . Soit V_{0h} le sous-espace vectoriel de $H_0^1(\Omega)$ défini par la méthode des éléments finis \mathbb{P}_k , de dimension N. Soient $(\lambda_i, u_i) \in \mathbb{R}^*_+ \times H_0^1(\Omega)$, $i \geq 1$, les valeurs propres et vecteurs propres (orthogonaux dans $L^2(\Omega)$) du problème de Dirichlet rangées par ordre croissant

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_i \leq \cdots \leq \cdots$$

Soient les valeurs propres

$$0 < \lambda_{1,h} \leq \lambda_{2,h} \leq \cdots \leq \lambda_{N,h},$$

de l'approximation variationnelle (12.3.1) pour V_{0h} . Alors, pour tout $i \ge 1$, on a

$$\lim_{h \to 0} |\lambda_i - \lambda_{i,h}| = 0,$$

et il existe une famille de vecteurs propres $(u_{i,h})_{1 \le i \le N}$ de (12.3.1) dans V_{0h} telle que

$$\lim_{h \to 0} ||u_i - u_{i,h}||_{H^1(\Omega)} = 0.$$

De plus, si Vect $\{u_1, \ldots, u_i\} \subset H^{k+1}(\Omega)$ et k+1 > d/2, on a

$$|\lambda_i - \lambda_{i,h}| \le C_i h^{2k},$$

 $où C_i$ ne dépend pas de h et

$$||u_i - u_{i,h}||_{H^1(\Omega)} \le C_i h^k.$$

Remarque 12.3.4. la constante C_i tend vers $+\infty$ lorsque *i* tend vers $+\infty$. Il n'y a donc aucune garantie que l'approximation soit correcte pour les grandes valeurs propres.

Chapitre 13

Problèmes paraboliques

13.1 Formulation variationnelle

Dans tout ce chapitre on va considérer la résolution numérique de l'équation de la chaleur

$$\begin{cases} \partial_t u - \Delta u = f & \text{dans } \mathbb{R}^*_+ \times \Omega, \\ u = 0 & \text{sur } \mathbb{R}^*_+ \times \partial \Omega, \\ u(t = 0, x) = u_0(x) & \text{dans } \Omega, \end{cases}$$
(13.1.1)

où $f(t, \cdot) \in L^2(\Omega)$ est le terme source et $u_0 \in H_0^1(\Omega)$ la donnée initiale.

Soit $v \in C_c^{\infty}(\Omega)$. On multiplie (13.1.1) par v puis on intègre sur Ω . En effectuant une intégration par parties, on aboutit à

$$\frac{d}{dt} \int_{\Omega} u(t,x)v(x) \, dx + \int_{\Omega} \nabla u(t,x) \cdot \nabla v(x) \, dx = \int_{\Omega} f(t,x)v(x) \, dx,$$

qui a un sens pour toute fonction test $v \in H_0^1(\Omega)$. On note u(t) := u(t, x) de sorte que u est une fonction définie sur \mathbb{R}_+ à valeurs dans $H_0^1(\Omega)$. Avec cette notation, on obtient la formulation variationnelle suivante pour (13.1.1).

$$\begin{cases} \text{Trouver } u: t \in \mathbb{R}_+ \to H_0^1(\Omega) \text{ telle que} \\ \frac{d}{dt}(u(t), v)_{L^2(\Omega)} + a(u(t), v) = (f(t), v)_{L^2(\Omega)}, \quad \forall t \in \mathbb{R}_+^*, \ \forall v \in H_0^1(\Omega), \\ u(t=0) = u_0, \end{cases}$$

$$(13.1.2)$$

où $a(\cdot, \cdot)$ est la forme bilinéaire définie sur $H^1(\Omega)$ par

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

Il est nécessaire de préciser la régularité de la solution en temps, pour cela on introduit des espaces de solutions prenant en compte la variable de temps et la variable d'espace.

Notation 13.1.1. Pour $(X, || \cdot ||)$ un espace de Banach on note $C^k([0, T]; X)$ l'espace des fonctions k-fois continûment dérivables de [0, T] à valeurs dans X. L'espace $C^k([0, T]; X)$

est un espace de Banach muni de la norme

$$\|v\|_{C^k([0,T];X)} := \sum_{i=0}^k \left(\sup_{0 \le t \le T} \left\| \frac{d^i v}{dt^i}(t) \right\|_X \right).$$

On note $L^2([0,T];X)$ l'espace des fonctions v telles que l'application $t \mapsto ||v(t)||_X$ est de carré intégrable sur [0,T], *i.e.*

$$\|v\|_{L^2([0,T];X)} := \left(\int_0^T \|v(t)\|_X^2 dt\right)^{1/2} < \infty$$

Muni de la norme $\|\cdot\|_{L^2([0,T];X)}$ ainsi définie, l'espace $L^2([0,T];X)$ est un espace de Banach. De plus, si $(X, (\cdot, \cdot)_X)$ est un espace de Hilbert, alors $L^2([0,T];X)$ est un espace de Hilbert muni du produit scalaire

$$(u,v)_{L^2([0,T];X)} := \int_0^T (u(t),v(t))_X dt$$

En particulier, si $X = L^2(\Omega)$ on a l'identification $L^2([0,T]; L^2(\Omega)) = L^2([0,T] \times \Omega)$.

Comme dans le cas des problèmes elliptiques, on va considérer un problème variationnel plus général que (13.1.2). Pour cela, soient $(V, (\cdot, \cdot)_V)$ et $(H, (\cdot, \cdot)_H)$ deux espaces de Hilbert réel tels que $V \subset H$. Soient $a(\cdot, \cdot)$ une forme bilinéaire sur $V, T > 0, u_0 \in H$ et $f \in L^2(]0, T[; H)$. Alors, on considère la formulation variationnelle

$$\begin{cases} \text{Trouver } u: t \in \mathbb{R}_+ \to V \text{ telle que} \\ \frac{d}{dt}(u(t), v)_H + a(u(t), v) = (f(t), v)_H, \quad \forall t \in]0, T[, \forall v \in V, \quad (13.1.3) \\ u(t=0) = u_0. \end{cases}$$

Théorème 13.1.2 (Existence et unicité). Soient $(V, (\cdot, \cdot)_V)$ et $(H, (\cdot, \cdot)_H)$ deux espaces de Hilbert réels tels que

$$\begin{cases} V \subset H \text{ avec injection compacte,} \\ V \text{ est dense dans } H. \end{cases}$$

Soit $a(\cdot, \cdot)$ une forme bilinéaire continue, coercive et symétrique sur V. Soient T > 0, $u_0 \in H$ et $f \in L^2(]0, T[; H)$. Alors, la formulation variationnelle (13.1.3) admet une unique solution $u \in L^2(]0, T[; V) \cap C([0, T]; H)$. De plus, il existe une constante c > 0telle que

$$||u||_{L^{2}(]0,T[;V)} + ||u||_{C([0,T];H)} \le c \left(||u_{0}||_{H} + ||f||_{L^{2}(]0,T[;H)} \right)$$

Autrement dit, le problème (13.1.3) est bien posé au sens d'Hadamard.

Démonstration. On donne seulement une idée de la démonstration, pour une démonstration complète on renvoie à [1].

D'après le Théorème 12.2.2, il existe une base hilbertienne $(u_k)_{k\geq 1}$ de H telle que, pour tout $k\geq 1$,

$$u_k \in V$$
 et $a(u_k, v) = \lambda_k (u_k, v)_H$, $\forall v \in V$. (13.1.4)

Pour tout $k \geq 1$, on pose

$$\alpha_k(t) := (u(t), u_k)_H \in C([0, T]), \qquad \alpha_k^0 = (u_0, u_k)_H,$$
(13.1.5)

 et

$$\beta_k(t) := (f(t), u_k)_H \in L^2(]0, T[).$$
 (13.1.6)

D'après la Remarque 12.1.5 et la notation (13.1.5), si $u(t) \in H$ pour tout $t \in [0, T]$, alors u(t) admet pour décomposition

$$\forall t \in [0,T], \quad u(t) = \sum_{k \ge 1} \alpha_k(t) \, u_k.$$
 (13.1.7)

Supposons que u est solution de (13.1.3). On prend pour fonction test $v = u_k$, où $k \ge 1$. Alors, on obtient

$$\frac{d}{dt}(u(t), u_k)_H + a(u(t), u_k) = (f(t), u_k)_H$$

D'après (13.1.5) et (13.1.6), on en déduit

$$\alpha'_k(t) + \lambda_k \,\alpha_k(t) = \beta_k(t).$$

De plus, la condition initiale de (13.1.3) et la notation (13.1.5) entraînent

$$\alpha_k(t=0) = \alpha_k^0.$$

Autrement dit, chaque α_k est solution d'une e.d.o. du premier ordre. On en déduit l'expression exacte de α_k :

$$\forall t \in [0,T], \quad \alpha_k(t) = \alpha_k^0 e^{-\lambda_k t} + \int_0^t \beta_k(s) e^{-\lambda_k(t-s)} ds.$$
(13.1.8)

Ainsi, si u est solution de (13.1.3) alors u est donnée par (13.1.7) avec α_k donné par (13.1.8) donc, en particulier, est unique. La démonstration complète consiste à justifier que la série (13.1.7) converge.

En appliquant le Théorème 13.1.2 avec $H := L^2(\Omega), V := H_0^1(\Omega)$ et $a(\cdot, \cdot)$ la forme bilinéaire définie sur V par

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

on obtient le résultat d'existence et d'unicité d'une solution faible pour l'équation de la chaleur :

Théorème 13.1.3. Soient Ω un ouvert borné régulier de classe C^1 de \mathbb{R}^d . Soient T > 0, $u_0 \in L^2(\Omega)$ et $f \in L^2(]0, T[; L^2(\Omega))$. Alors, l'équation de la chaleur (13.1.1) admet une unique solution faible $u \in L^2(]0, T[; H_0^1(\Omega)) \cap C([0, T]; L^2(\Omega))$. De plus, il existe une constante c > 0 telle que

$$\int_{\Omega} |u(t,x)|^2 \, dx + \int_0^t \int_{\Omega} |\nabla u(s,x)|^2 \, dx ds \le c \left(\int_{\Omega} |u_0(x)|^2 \, dx + \int_0^t \int_{\Omega} |f(s,x)|^2 \, dx ds \right).$$

On donne, sans démonstration, un résultat qualitatif sur la régularité de la solution de l'équation de la chaleur (voir [1]).

Proposition 13.1.4 (Effet régularisant). Soit Ω un ouvert borné régulier de classe C^{∞} de \mathbb{R}^d . Soient T > 0, $u_0 \in L^2(\Omega)$ et f = 0. Alors, pour tout $\varepsilon > 0$, l'unique solution faible de (13.1.1) est de classe C^{∞} en temps et en espace dans $[\varepsilon, T] \times \overline{\Omega}$.

13.2 Résolution numérique par éléments finis

La résolution numérique des problèmes paraboliques sous la forme (13.1.3) se fait en deux étapes. On effectue d'abord une discrétisation par éléments finis en la variable d'espace. On aboutit ainsi à une e.d.o. en temps que l'on discrétise par différences finies.

Avec les notations de la section précédente, soit V_h un sous-espace de dimension finie N de V. On s'intéresse à la discrétisation du problème variationnel (13.1.3), *i.e.*

 $\begin{cases} \text{Trouver } u: t \in \mathbb{R}_+ \to V \text{ telle que} \\ & \frac{d}{dt}(u(t), v)_H + a(u(t), v) = (f(t), v)_H, \quad \forall t \in]0, T[, \forall v \in V, \\ & u(t = 0) = u_0. \end{cases}$

En considérant l'espace discret V_h , on obtient le problème variationnel discret suivant :

Trouver
$$u_h : t \in \mathbb{R}_+ \to V_h$$
 telle que

$$\frac{d}{dt}(u_h(t), v_h)_H + a(u_h(t), v_h) = (f(t), v_h)_H, \quad \forall t \in]0, T[, \forall v_h \in V_h,$$

$$u_h(t = 0) = u_{0,h},$$
(13.2.1)

où $u_{0,h} \in V_h$ est une approximation de u_0 . Soit $\{\phi_i\}_{i=1,\dots,N}$ une base de V_h . Alors, pour tout $t \in [0, T[, u_h(t) \text{ et } u_{0,h} \text{ admettent pour développements}$

$$u_h(t) = \sum_{j=1}^N U_h^j(t) \phi_j \quad \text{et} \quad u_{0,h} = \sum_{j=1}^N U_{0,h}^j \phi_j, \qquad (13.2.2)$$

avec $U_h^j(t), U_{0,h}^j \in \mathbb{R}$ pour tout $j = 1, \dots, N$. Avec ces développements, (13.2.1) s'écrit

$$\begin{cases} \text{Trouver } U_h^j : t \in \mathbb{R}_+ \to \mathbb{R} \quad \text{et} \quad U_{0,h}^j \in \mathbb{R}, \ j = 1, \dots, N \text{ tels que} \\ \sum_{j=1}^N (\phi_j, \phi_i)_H \frac{dU_h^j}{dt}(t) + \sum_{j=1}^N a(\phi_j, \phi_i) U_h^j(t) = (f(t), \phi_i)_H, \quad \forall t \in]0, T[, \ \forall i = 1, \dots, N, \\ U_h^i(t=0) = U_{0,h}^i, \quad \forall i = 1, \dots, N. \end{cases}$$

On pose $U_h(t) := (U_h^1(t), \dots, U_h^N(t))^T \in \mathbb{R}^N$ et $U_{0,h} := (U_{0,h}^1, \dots, U_{0,h}^N)^T \in \mathbb{R}^N$. Alors (13.2.3) est équivalent au système d'e.d.o.

$$\begin{cases} \mathcal{M}_{h}U_{h}'(t) + \mathcal{K}_{h}U_{h}(t) = b_{h}(t), & \forall t \in]0, T[, \\ U_{h}(t=0) = U_{0,h}, \end{cases}$$
(13.2.4)

où $\mathcal{M}_h, \mathcal{K}_h \in \mathbb{R}^{N \times N}$ et $b_h \in \mathbb{R}^N$ sont donnés par

$$(\mathcal{M}_h)_{ij} := (\phi_i, \phi_j)_H, \quad (\mathcal{K}_h)_{ij} := a(\phi_i, \phi_j) \quad \text{et} \quad (b_h(t))_i := (f(t), \phi_i)_H.$$

Autrement dit, \mathcal{K}_h et \mathcal{M}_h correspondent aux matrices de rigidité et de masse rencontrées dans la résolution numérique des problèmes elliptiques dans le cadre général et dans le cadre spectral. On obtient un premier résultat de convergence pour la semi-discrétisation en espace de l'équation de la chaleur (13.1.1) : **Proposition 13.2.1.** Soient Ω un ouvert polyédrique de \mathbb{R}^d et $(\mathcal{T}_h)_{h>0}$ une suite de maillages triangulaires réguliers de Ω . Soit V_{0h} l'espace de discrétisation de $H_0^1(\Omega)$ défini par la méthode des éléments finis \mathbb{P}_k , $k \geq 1$, de dimension N. Soient $f \in L^2(]0, 1[; L^2(\Omega))$, $u_0 \in H_0^1(\Omega)$ et $u \in L^2(]0, T[; H_0^1(\Omega)) \cap C([0, T]; L^2(\Omega))$ l'unique solution faible de (13.1.1). Soit u_h donné par (13.2.2) avec U_h la solution de (13.2.4). On suppose

$$\lim_{h \to 0} \|u_{0,h} - u_0\|_{L^2(\Omega)} = 0,$$

alors on a

$$\lim_{h \to 0} \|u - u_h\|_{L^2([0,T[;H_0^1(\Omega))]} = \lim_{h \to 0} \|u - u_h\|_{C([0,T];L^2(\Omega))} = 0.$$

Il reste maintenant à discrétiser (13.2.4) pour la variable en temps par un schéma aux différences finies. On suppose dans la suite $b_h \in C([0,T])$. Soient $N_T \in \mathbb{N}$ et $\Delta t := T/N_T$. On pose

$$\forall n \in \{0, \dots, N_T\}, \quad t_n := n\Delta t.$$

On considère l'approximation U_h^n de $U_h(t_n)$ définie par le θ -schéma :

$$\mathcal{M}_h \frac{U_h^{n+1} - U_h^n}{\Delta t} + \mathcal{K}_h \left(\theta U_h^{n+1} + (1-\theta) U_h^n \right) = \theta b(t_{n+1}) + (1-\theta) b(t_n),$$

ce qui donne

$$\left(\mathcal{M}_{h}+\theta\,\Delta t\,\mathcal{K}_{h}\right)U_{h}^{n+1}=\left(\mathcal{M}_{h}-\left(1-\theta\right)\Delta t\,\mathcal{K}_{h}\right)U_{h}^{n}+\Delta t\left(\theta b(t_{n+1})+\left(1-\theta\right)b(t_{n})\right).$$
 (13.2.5)

Définition 13.2.2. Un schéma aux différences finies de (13.2.4) est dit **stable** s'il existe une constante c > 0 indépendante de Δt et h telle que

$$\forall n \in \{0, \dots, N_T\}, \quad \mathcal{M}_h U_h^n \cdot U_h^n \le c.$$

Lemme 13.2.3 (Stabilité du schéma). Si $1/2 \le \theta \le 1$, le θ -schéma (13.2.5) est inconditionnellement stable. Si $0 \le \theta < 1/2$, le θ -schéma (13.2.5) est stable sous la condition CFL

$$\max_{1 \le i \le N} (\lambda_i \,\Delta t) \le \frac{2}{1 - 2\theta},\tag{13.2.6}$$

où les λ_i , $1 \leq i \leq N$, sont les valeurs propres de $\mathcal{K}U = \lambda \mathcal{M}U$.

Démonstration. Voir [1].

On a alors le résultat de convergence suivant (voir [12]) :

Théorème 13.2.4. Soient Ω un ouvert polyédrique de \mathbb{R}^d et $(\mathcal{T}_h)_{h>0}$ une suite de maillages triangulaires réguliers de Ω . Soit V_{0h} l'espace de discrétisation de $H_0^1(\Omega)$ défini par la méthode des éléments finis \mathbb{P}_k , $k \geq 1$, de dimension N. Soient T > 0, $f \in L^2(]0, 1[; L^2(\Omega))$, $u_0 \in H_0^1(\Omega)$ et u l'unique solution faible de (13.1.1) supposée "suffisamment régulière". Soit u_h la solution de (13.2.5). On suppose

$$\lim_{h \to 0} \|u_{0,h} - u_0\|_{L^2(\Omega)} = 0,$$

et que h et Δt tendent vers 0 en vérifiant (13.2.6) si $0 \leq \theta < 1/2$. Alors, on a

$$\lim_{h \to 0, \Delta t \to 0} \max_{0 \le n \le N_T} \| u(t_n) - u_h^n \|_{L^2(\Omega)} = 0$$

Bibliographie

- G. ALLAIRE, Analyse numérique et optimisation. Éditions Ellipses, Paris (2006). Version sans images et avec bandeau pour usage personnel et non reproductible disponible à l'adresse http://www.cmap.polytechnique.fr/~allaire/livre2.html
- [2] G. ALLAIRE & S.M. KABER, Algèbre linéaire numérique. Cours et exercices. Édition ellipses, Paris (2002).
- [3] A.-S. BONNET-BENDHIA & E. LUNEVILLE, Résolution numérique des équations aux dérivées partielles. Ensta (1993).
- [4] J.M. BONY, Cours d'analyse. Théorie des distributions et analyse de Fourier. Editions de l'école Polytechnique, Palaiseau (2001).
- [5] H. BREZIS, Analyse Fonctionnelle : Théorie et Applications. Masson, Paris (1983).
- [6] C. DAVEAU, Méthodes d'approximation des équations aux dérivées partielles. Cours de Master 2 de l'université de Cergy-Pontoise. Polycopié disponible à l'adresse http://www.u-cergy.fr/daveau/M2-EDP-APPROX.pdf
- [7] D. EUVRARD, Résolution numérique des équations aux dérivées partielles de la physique, de la mécanique et des sciences de l'ingénieur : différences finies, éléments finis, problèmes en domaines non bornés. Masson, Paris (1990).
- [8] D. GILBARG & N.S. TRUDINGER, Elliptic Partial Differential Equations of Second Order. 2nd ed., Springer-Verlag, Berlin, 1983.
- [9] R. HERBIN, Analyse numérique des EDP. Cours de Master de mathématiques de l'université Aix Marseille 1. Polycopié disponible à l'adresse http://www.cmi.univ-mrs.fr/~herbin/anedp.html
- [10] J.L. LIONS & E. MAGENES, Problèmes aux limites non homogènes et applications. Dunod, Paris (1970).
- [11] S. MAINGOT & D. MANCEAU, *Théorie spectrale*. Cours de Master 1 Mathématiques de l'université du Havre. Polycopié disponible à l'adresse http://d.p.manceau.free.fr/ThSpect/ThSpect.pdf
- [12] P.-A. RAVIART & J. M. THOMAS, Introduction à l'analyse numérique des équations aux dérivées partielles. Masson, Paris (1983).
- [13] L. TARTAR, An introduction to Navier-Stokes equation and oceanography. Lectures notes of Unione Matematica Italiana, Vol. 1, Springer, Berlin (2008).