

Analyse variationnelle des équations aux dérivées partielles

Sonia Fliss & Philippe Moireau

Édition 2026

Alan Turing aurait déclaré que les équations aux dérivées partielles sont créées par Dieu et les conditions aux limites par le Diable! La situation a changé, le Diable a changé de place... On peut dire que les principaux défis se situent au niveau des interfaces, avec le Diable non loin d'elles...

–Jacques-Louis Lions (1928-2001)

Preface

Ce cours est une introduction aux équations aux dérivées partielles par le prisme du calcul des variations que l'on rencontre dans de nombreux problèmes physiques par le principe (quasi philosophique) de moindre action de Maupertuis, Leibniz, Euler, Lagrange, etc. Le fameux opérateur de Laplace, que nous allons rencontrer tout au long de ce cours, est en effet un formidable couteau suisse que l'on croise partout en physique, chimie, biologie, mais aussi mécanique et dans tous les champs des mathématiques. On le rencontre en analyse bien sûr, mais aussi en optimisation, en probabilités et même en statistique. Nous pouvons y voir un clin d'œil de l'histoire, car Pierre-Simon de Laplace est l'un des principaux scientifiques de la période napoléonienne qui a vu se développer les premières écoles d'ingénieurs. Parfois appelé le Newton français, il a marqué son époque à la fois par sa vision scientifique, mais aussi sa carrière politique, devenant un exemple d'un scientifique acteur de la vie publique.

Dans ce cours nous voulons présenter toute la chaîne mathématique depuis la modélisation, la formulation, l'analyse, la discrétisation, l'analyse numérique et la simulation. La modélisation traduit notre compréhension du phénomène physique, la formulation sa mise en équation. L'analyse implique de garantir par la théorie mathématique que le problème admet une solution, idéalement une seule et que cette solution est stable par rapport aux données d'entrée du problème. Ceci est indispensable pour que le modèle de la réalité étudiée, ayant simplifié son sujet d'étude, n'en soit pas pour autant simpliste. La discrétisation, elle, consiste en la reformulation du problème en une approximation qui peut être calculée numériquement. Dans ce cours, c'est la méthode des éléments finis qui est une approche variationnelle par essence qui est présentée. Simulation signifie résolution pratique de cette formulation discrète. Reste à souligner l'importance de l'étape d'analyse numérique parfois plus méconnue des autres communautés au-delà des mathématiques appliquées. Elle cherche à garantir que la discrétisation a bien une solution, mais surtout qu'elle est proche de la solution mathématique initialement prévue. Cette thématique est tout à fait passionnante, car elle fait le lien entre deux mondes, l'abstrait et le concret; elle offre des garanties *a priori* sur ce que la machine calcule; et elle présente une esthétique mathématique propre, à l'interface entre des mathématiques très concrètes des premières années de licence, notamment l'algèbre linéaire, et l'analyse beaucoup plus abstraite des problèmes aux équations partielles.

On s'amuse souvent à dire qu'il y a deux types d'ingénieurs, ceux qui proposent des solutions qui marchent sans savoir pourquoi, et ceux qui proposent des solutions qui ne marchent pas, mais savent pourquoi. Nous pensons modestement qu'on peut avoir l'exigence de se former à inventer des solutions nouvelles tout en ne négligeant pas de savoir pourquoi ces innovations fonctionnent. C'est frappant avec la méthode des éléments finis, imaginée il y a plus d'un siècle, mais toujours d'actualité, merveille d'ingéniosité dans sa formulation et dont la garantie du fonctionnement s'appuie sur un cadre abstrait extrêmement limpide, élégant et rigoureux.

Ce document est le fruit d'une collaboration pour unifier la présentation du cours d'analyse variationnelle de l'Ensta et de celui de l'École polytechnique. Les deux formations sont très proches sur ce sujet et nous sommes heureux de participer à notre niveau au rapprochement effectif des cours, des élèves et des enseignants-chercheurs de ces établissements membres de l'Institut Polytechnique de Paris.

Nous tenons à remercier très chaleureusement d'un côté, Patrick Ciarlet et Eric Lunéville, et de l'autre, François Alouges et Grégoire Allaire, dont les précédentes éditions de ce cours respectivement à l'Ensta et à l'École polytechnique ont été d'une aide très précieuse pour la rédaction de ce document. Nous souhaitons aussi vivement remercier nos relecteurs E. Becache, D. Chapelle, L. Chesnel, S. Imperiale, F. Nabet. Malgré leurs relectures attentives, nous avons sans doute laissé des erreurs dans ce manuscrit et nous serions heureux de prendre en compte toutes les remarques que vous nous adresseriez à sonia.fliss@ensta.fr ou philippe.moireau@polytechnique.edu.

Sonia Fliss et Philippe Moireau,

Table des matières

Preface	v
Table des matières	vii
1 Modélisation et calcul des variations	1
1.1 Introduction au calcul des variations en dimension 1	1
1.1.1 Le principe de moindre action	1
1.1.2 Quelques rappels de calcul différentiel	2
1.1.3 Condition nécessaire d'optimalité	4
1.2 Problème modèle en dimension supérieure	6
1.2.1 Un exemple en dimension 2 d'espace	6
1.2.2 Notion de domaine en dimension supérieure	7
1.2.3 Formules de Green pour les fonctions régulières	8
1.2.4 Des formules de Green aux équations d'Euler Lagrange	10
1.3 Autres exemples de problèmes variationnels	12
1.3.1 Un problème électrostatique	12
1.3.2 Calcul des variations en traitement de signal	12
1.3.3 Équations des poutres et plaques en flexion	14
1.4 Des lois de conservations aux solutions à l'équilibre	15
1.4.1 Lois des conservations et lois de comportements	15
1.4.2 Problèmes stationnaires et principe variationnel	17
1.4.3 Les problèmes modèles de ce cours	17
1.5 Le calcul des variations d'Euler et Lagrange	18
1.5.1 Le défi de Bernoulli	19
1.5.2 Les équations d'Euler-Lagrange	20
1.5.3 Des relation de Snell-Descartes à la nécessité d'affaiblir la régularité	23
2 Espaces de Sobolev	25
2.1 Fonctions de carré intégrable	25
2.1.1 L'espace $L^2(\Omega)$	25
2.1.2 Théorème de prolongement par densité	27
2.1.3 Dérivation faible des fonctions de $L^2(\Omega)$	29
2.2 Les espaces de Sobolev	32
2.2.1 L'espace $H^1(\Omega)$ et l'espace $H_0^1(\Omega)$	32
2.2.2 Les espaces $H(\Omega, \text{div})$, $H^m(\Omega)$ et $H_0^m(\Omega)$	36
2.2.3 Théorèmes de trace	38
2.2.4 Formules de Green	44
2.2.5 Inégalités de type Poincaré	47
2.3 Lien entre la dérivée au sens faible et la dérivée au sens des distributions	51
3 Analyse hilbertienne des problèmes variationnels	57
3.1 Introduction	57
3.2 Espaces de Hilbert et propriétés élémentaires	59
3.3 Le théorème de Représentation de Riesz et ses conséquences	63
3.4 Convergence faible	65
3.5 Lax-Milgram	68
3.5.1 Le théorème de Lax-Milgram	68

3.6	Approximation interne générale	71
3.6.1	Caractère bien posé et propriétés du système linéaire	71
3.6.2	Lemme de Céa	73
4	Étude mathématique des problèmes aux limites	75
4.1	Introduction	75
4.2	Étude d'un problème elliptique scalaire	76
4.2.1	Conditions de Dirichlet homogènes	76
4.2.2	Conditions de Dirichlet non homogènes	81
4.2.3	Conditions de Neumann non homogènes	83
4.3	La méthode directe du calcul des variations	89
4.4	Propriétés des solutions faibles	91
4.4.1	Principe du maximum	91
4.4.2	Résultats de régularité	93
5	La méthode des éléments finis en 1D	97
5.1	Introduction	97
5.2	Élément fini de Lagrange $\mathbb{P}_1$ en dimension 1	97
5.2.1	Maillage et fonctions de forme	97
5.2.2	Matrices de raideur et de masse – cas constant et conditions de Dirichlet	99
5.2.3	Calcul du second membre.	101
5.2.4	Procédure d'assemblage.	102
5.2.5	Prise en compte des conditions aux limites après la procédure d'assemblage.	105
5.2.6	Passage à l'élément de référence.	106
5.2.7	Cas des coefficients variables et formules de quadrature	107
5.3	Autres exemples d'éléments finis en dimension 1	110
5.3.1	Éléments finis de Lagrange $\mathbb{P}_k$	110
5.3.2	Élément fini de Hermite pour des problèmes de type : "poutre"	113
5.4	Méthode spectrale	116
6	La méthode des éléments finis en 2D	119
6.1	Introduction	119
6.2	Élément fini de Lagrange $\mathbb{P}_1$ en dimension 2	119
6.2.1	Notion de maillage triangulaire en dimension 2	119
6.2.2	Définition de l'espace d'approximation $\mathbb{P}_1$ et fonctions de forme	121
6.2.3	Matrices de masse et raideur élémentaires.	124
6.2.4	Passage à l'élément de référence.	128
6.2.5	Formules de quadrature en dimension 2	129
6.2.6	Assemblage.	130
6.2.7	Calcul du second membre	130
6.2.8	Prise en compte des conditions aux limites de Dirichlet	131
6.3	Éléments finis de Lagrange généraux	135
6.3.1	Éléments finis de Lagrange $\mathbb{P}_k$	135
6.3.2	Éléments finis de Lagrange $\mathbb{Q}_k$	138
6.4	Éléments finis: cadre général	141
7	Analyse numérique de la méthode des éléments finis	143
7.1	Introduction	143
7.2	Vitesse de convergence de la méthode des éléments finis	144
7.2.1	Une estimation d'erreur en dimension 1	144
7.2.2	Convergence et estimation d'erreur en dimension $d \geq 2$	149
7.2.3	Prise en compte de la régularité des données	158

7.2.4	Lemme de Strang	160
7.2.5	Estimations de l'erreur en norme L^2	162
7.3	Principe du maximum discret	164
8	Analyse spectrale des problèmes aux limites	169
8.1	Introduction	169
8.2	Un problème aux valeurs propres modèle	169
8.3	Problème aux valeurs propres sous forme opérateur	170
8.4	Analyse spectrale des opérateurs auto-adjoints compacts	172
8.4.1	Opérateurs bornés	172
8.4.2	Opérateurs auto-adjoints	173
8.4.3	Opérateurs compacts	173
8.4.4	Principe du Min-Max	178
8.5	Approximation numérique des problèmes spectraux	180
8.5.1	Existence des éléments propres discrets	180
8.5.2	Analyse d'erreur	182
8.5.3	Calcul de valeurs et vecteurs propres en pratique	184
8.6	Analyse modale en mécanique	187
8.6.1	Modes de vibration d'une structure élastique	187
8.6.2	Modes de vibration de la plaque	187
	Bibliographie	191

Modélisation et calcul des variations

1

Ce premier chapitre constitue une introduction générale au calcul des variations. L'objectif est de montrer comment la modélisation physique s'est construite de concert avec la formalisation mathématique. En partant de quelques exemples emblématiques, nous introduisons les premiers outils de formalisation mathématique. On verra notamment que, les fonctions d'abord supposées régulières afin de mener les calculs vont progressivement montrer leurs limites et nous obliger à introduire de nouveaux espaces pour définir les solutions mathématiques des problèmes physiques modélisés. Ainsi, choix de modélisation et régularité mathématique seront deux facettes d'une même démarche qui se sont construites historiquement ensemble dans un dialogue permanent.

1.1 Introduction au calcul des variations en dimension 1

1.1.1 Le principe de moindre action

En 1744, Maupertuis définit¹ le principe de la moindre quantité d'action sous la forme « *Lorsqu'il arrive quelque changement dans la nature, la quantité d'action, nécessaire pour ce changement, est la plus petite qui soit possible* ». Se développe alors à partir des outils du calcul des variations, une vision de la physique où les équations fondamentales sont formulées à partir du principe de moindre action.

Considérons par exemple une corde, ou chaînette, tendue ayant une certaine densité de masse linéique $\rho \geq 0$ et fixée à ses extrémités en $x = 0$ et $x = L$.

Le principe de moindre action stipule que l'équation de la corde correspond à la minimisation d'une énergie de déformation W et de son énergie potentielle. Ainsi la nature cherchait à minimiser son effort²

$$\mathcal{J}(v) = \int_0^L \left[W(v(x), v'(x), v''(x), \dots) - \rho g v(x) \sqrt{1 + v'(x)^2} \right] dx,$$

où le premier terme décrivant l'énergie de déformation est écrit en toute généralité comme une fonction de la forme de la chaîne et de ses dérivées caractérisant l'élongation, la courbure, etc., autour de la configuration au repos.

Dans le cas le plus élémentaire où la déformation de la chaînette est infinitésimale, c'est-à-dire une configuration où la longueur totale $L_{\text{tot}} = L + o(a)$ où a est le déplacement maximal, le critère à minimiser devient

$$\mathcal{J}(v) = \int_0^L \left[\frac{\kappa}{2} v'(x)^2 - \rho g v(x) \right] dx.$$

où le premier terme s'interprète comme si la corde était constituée d'un continuum de ressorts élastiques infinitésimaux³, de raideur linéique

1: il y a aussi le même genre d'idée en optique avec Fermat en 1662 (voir plus loin dans le chapitre). Et Leibniz a introduit antérieurement des idées semblables à Maupertuis.

2: on remarquera que la définition de l'abscisse curviligne donne $ds = \sqrt{1 + u'(x)^2} dx$

3: l'énergie des ressorts est asymptotiquement $\sum_i \frac{\kappa \delta x}{2} [v(x_i + \frac{\delta x}{2}) - v(x_i - \frac{\delta x}{2})]^2 \xrightarrow{\delta x \rightarrow 0} \int_0^L \frac{\kappa}{2} v'(x)^2 dx$

$\kappa \geq 0$.

Cette fonctionnelle est convexe car on vérifie aisément que pour tout v_1 et v_2 et pour tout réel λ

$$\mathcal{J}(\lambda v_1 + (1 - \lambda)v_2) \leq \lambda \mathcal{J}(v_1) + (1 - \lambda)\mathcal{J}(v_2)$$

par convexité de la fonction $\mathbb{R} \ni u \mapsto u^2$. Nous allons chercher à la minimiser, et cherchons pour l'instant une condition nécessaire.

Nous devons nous fixer un espace de fonctions sur lequel minimiser $\mathcal{J}$. Cette question est en réalité très délicate et au cœur de la théorie variationnelle des équations aux dérivées partielles. Pour commencer, nous pouvons par exemple définir l'espace vectoriel

$$\mathcal{V} = \{v \in C^1([0, 1]) \text{ tel que } v(0) = v(1) = 0\},$$

muni de la norme

$$\forall v \in C^1([0, L]), \quad \|v\|_{C^1} = \sup_{x \in [0, L]} |v(x)| + \sup_{x \in [0, L]} |v'(x)|.$$

4: C'est donc un espace de Banach

Nous rappelons notamment que cet espace est complet⁴ en utilisant le résultat suivant.

Proposition 1.1.1 Pour tout $(a, b) \in \mathbb{R}^2$ et $k \in \mathbb{N}$, l'espace $C^k([a, b])$ est complet pour la norme

$$\|u\|_{C^k([0, L])} = \sum_{i=0}^k \sup_{x \in [0, L]} |u^{(i)}(x)|.$$

Démonstration. Montrons le résultat pour $k = 0$. Le cas général se déduit sans difficulté ensuite par récurrence. Soit $(u^n)_{n \in \mathbb{N}}$ une suite de Cauchy dans $C^0([a, b])$. Alors pour tout $\varepsilon > 0$, et pour tout $p, q \in \mathbb{N}$ assez grands, on a

$$\forall x \in [a, b], \quad |u^p(x) - u^q(x)| \leq \varepsilon.$$

Donc pour tout $x \in [a, b]$, $(u^n(x))_{n \in \mathbb{N}}$ est de Cauchy donc converge dans $\mathbb{R}$ complet vers une limite notée $u(x)$. Soit alors $\delta > 0$, et $(x, y) \in [a, b]^2$. On a par inégalité triangulaire

$$|u(x) - u(y)| \leq |u(x) - u^n(x)| + |u^n(x) - u^n(y)| + |u^n(y) - u(y)|$$

Donc en choisissant n assez grand et $\alpha = |x - y|$ assez petit on obtient bien $|u(x) - u(y)| \leq \delta$ qui assure la continuité de u . $\square$

1.1.2 Quelques rappels de calcul différentiel

On rappelle la définition de la différentielle d'une application $f : \mathcal{V} \rightarrow \mathcal{U}$, où $\mathcal{V}$ et $\mathcal{U}$ sont des espaces vectoriels normés.

Définition 1.1.1 (Différentielle) Soient $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ et $(\mathcal{U}, \|\cdot\|_{\mathcal{U}})$ deux espaces vectoriels normés, $\mathcal{B}$ un ouvert de $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$, $v \in \mathcal{B}$ et $f : \mathcal{B} \rightarrow \mathcal{U}$. On dit que l'application f est différentiable en $v \in \mathcal{B}$ s'il existe $L \in$

$\mathcal{L}(\mathcal{V}, \mathcal{U})^5$ telle que

$$\forall w \in \mathcal{V}, \quad \|f(v+w) - f(v) - L(w)\|_{\mathcal{U}} = o(\|w\|_{\mathcal{V}}). \quad (1.1)$$

5: l'ensemble des applications linéaires continues de l'espace normé $\mathcal{V}$ dans l'espace normé $\mathcal{U}$

Proposition 1.1.2 *Supposons que f est différentiable en $v \in \mathcal{B}$. Alors $L \in \mathcal{L}(\mathcal{V}, \mathcal{U})$ donnée par (1.1) est unique. Elle est appelée différentielle de f en v et est notée $Df(v)$. Pour $w \in \mathcal{V}$, on notera $L(w) = Df(v)(w)$.*

Démonstration. Soient $L_1, L_2 \in \mathcal{L}(\mathcal{V}, \mathcal{U})$ vérifiant la relation (1.1). Soit $\varepsilon > 0$. Pour $i \in \{1, 2\}$, il existe $\delta_i > 0$ tel que

$$\forall w \in \mathcal{V}, \quad \|w\|_{\mathcal{V}} < \delta_i \implies \|f(v+w) - f(v) - L_i(w)\|_{\mathcal{U}} \leq \varepsilon \|w\|_{\mathcal{V}}.$$

Pour tout $w \in \mathcal{V}$ tel que $\|w\|_{\mathcal{V}} < \min\{\delta_1, \delta_2\}$, on a

$$\begin{aligned} \|(L_1 - L_2)(w)\|_{\mathcal{U}} &\leq \|f(v+w) - f(v) - L_1(w)\|_{\mathcal{U}} \\ &\quad + \|f(v+w) - f(v) - L_2(w)\|_{\mathcal{U}} \leq 2\varepsilon \|w\|_{\mathcal{V}}. \end{aligned}$$

Par la linéarité de $L_1 - L_2$, on obtient $\|L_1 - L_2\|_{\mathcal{L}(\mathcal{V}, \mathcal{U})} \leq 2\varepsilon$. Ainsi $\varepsilon \rightarrow 0$, implique $L_1 = L_2$. $\square$

À noter que lorsque $f : \mathcal{V} \rightarrow \mathbb{R}$, on dit que f est une fonctionnelle sur $\mathcal{V}$. Dans ce cas, $Df(v)$ est une forme linéaire, donc un élément du dual de $\mathcal{V}$, noté $\mathcal{V}' = \mathcal{L}(\mathcal{V}, \mathbb{R})$. On peut alors noter, en utilisant un crochet de dualité ⁶,

$$\langle Df(v), w \rangle_{\mathcal{V}', \mathcal{V}} = Df(v)(w),$$

cette notation ayant l'avantage de souligner la dépendance linéaire de $Df(v)$ par rapport à w .

On rappelle maintenant quelques résultats classiques, dont les preuves élémentaires sont laissées au lecteur.

Proposition 1.1.3 *Si f est différentiable en $v \in \mathcal{V}$, alors elle admet une dérivée dans la direction $w \in \mathcal{V}$*

$$\lim_{\varepsilon \rightarrow 0} \frac{f(v + \varepsilon w) - f(v)}{\varepsilon} = Df(v)(w),$$

pour tout $w \in \mathcal{V}$. La réciproque est fautive⁷.

6: parfois appelé aussi un produit de dualité par opposition au produit scalaire

Théorème 1.1.4 (Théorème de composition des différentielles) *Soient $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$, $(\mathcal{U}, \|\cdot\|_{\mathcal{U}})$ et $(\mathcal{W}, \|\cdot\|_{\mathcal{W}})$ des espaces vectoriels normés. Soient $f : \mathcal{V} \rightarrow \mathcal{U}$ et $g : \mathcal{U} \rightarrow \mathcal{W}$. Si f est différentiable en v et g est différentiable en $f(v)$, alors $g \circ f$ est différentiable en v et*

$$D(g \circ f)(v) = Dg(f(v)) \circ Df(v).$$

7: Pensez à $f(x, y) = \frac{x^2 y}{x^2 + y^2}$ si $(x, y) \neq (0, 0)$ et 0 si $(x, y) = (0, 0)$ qui admet une dérivée dans toutes les directions sans être continue en $(0, 0)$

Enfin, puisque nous cherchons à caractériser des extrema, nous en profitons pour rappeler un résultat classique sur les conditions d'optimalité lorsqu'on cherche un extrema dans un espace non contraint.

Proposition 1.1.5 (Condition nécessaire d'optimalité du premier ordre) Soient $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$, un espace vectoriel normé réel et $f : \mathcal{V} \rightarrow \mathbb{R}$ une application continue différentiable. Soit u un minimum local de f , alors

$$\forall w \in \mathcal{V}, \quad Df(u)(w) = 0.$$

1.1.3 Condition nécessaire d'optimalité

8: Dans les calculs de différentielle, on pourra préférer utiliser la dérivée au sens de Gâteaux

Il est facile de montrer que $\mathcal{J}$ est différentiable⁸. Ici, on obtient pour tout $v \in \mathcal{V}$,

$$\langle D\mathcal{J}(u), w \rangle_{\mathcal{V}', \mathcal{V}} = \int_0^L [\kappa u'(x)w'(x) - \varrho g w(x)] dx,$$

et on vérifie bien par simple majoration

$$|\mathcal{J}(u + w) - \mathcal{J}(u) - \langle D\mathcal{J}(u), w \rangle_{\mathcal{V}', \mathcal{V}}| = O(\|w\|_{C^1([0, L])}^2).$$

Si u est le minimum on a $\langle D\mathcal{J}(u), w \rangle_{\mathcal{V}', \mathcal{V}} = 0$, pour tout $w \in \mathcal{V}$, qui se réécrit

$$\forall w \in \mathcal{V}, \quad \int_0^L \kappa u'(x)w'(x) dx = \int_0^L \varrho g w(x) dx. \quad (1.2)$$

Nous appellerons ce type de relation la caractérisation du minimum sous forme variationnelle, car nous avons calculé des variations de la fonctionnelle autour du minimum u . Ces variations incrémentales w sont appelées *fonctions tests*.

Supposons désormais que $u \in C^2(]0, L[)$, alors nous pouvons aller un cran plus loin dans la caractérisation en intégrant par parties sachant que $w(0) = w(1) = 0$, on obtient

$$\int_0^L [-\kappa u''(x)w(x) - \varrho g w(x)] dx = 0. \quad (1.3)$$

Pour conclure ici, nous utilisons ici le lemme⁹ suivant.

9: énoncé par Friedrich Ludwig Stegmann en 1854 et justifié succinctement; Joseph-Louis Lagrange l'avait tenu pour allant de soi. Démontré par Eduard Heine en 1870 et Paul David Gustave du Bois-Reymond en 1879 pour des fonctions régulières. Ce résultat sera étendu dans les chapitres suivants à des fonctions beaucoup moins régulières, et la preuve sera plus délicate. Ce seront ces versions étendues qui nous seront finalement les plus utiles dans ce cours et ce lemme aura surtout une valeur historique.

Lemme 1.1.6 (Lemme de Du Bois-Reymond) Soit Ω un ouvert de $\mathbb{R}^d$. Soit f une fonction continue dans Ω . Si pour toute fonction ϕ de $C^\infty(\Omega)$ à support compact dans Ω , on a

$$\int_{\Omega} f(\underline{x})\phi(\underline{x}) d\Omega = 0,$$

alors la fonction f est nulle dans Ω .

Démonstration. Supposons qu'il existe un point $\underline{x}_0 \in \Omega$ tel que $f(\underline{x}_0) \neq 0$. Sans perte de généralité, on peut supposer que $f(\underline{x}_0) > 0$ (sinon on prend $-f$). Par continuité, il existe un petit voisinage ouvert $\omega \subset \Omega$ de $\underline{x}_0$ tel que $f(\underline{x}) > 0$ pour tout $\underline{x} \in \omega$. Soit alors une fonction test positive, non nulle, ϕ à support inclus dans ω . On a

$$0 = \int_{\Omega} f(\underline{x})\phi(\underline{x}) \, d\Omega = \int_{\omega} f(\underline{x})\phi(\underline{x}) \, d\Omega > 0,$$

ce qui est absurde. Donc $f(\underline{x}) = 0$ pour tout $\underline{x} \in \Omega$. $\square$

On déduit du lemme 1.1.6 les équations dites d'*Euler-Lagrange*¹⁰ de la forme

$$-\kappa u''(x) - \varrho g = 0,$$

ce qui donne en prenant en compte les conditions aux limites

$$\begin{cases} -\kappa u''(x) = \varrho g, & x \in]0, L[, \\ u(0) = u(L) = 0. \end{cases} \quad (1.4)$$

On appelle ce système *un problème aux limites*, car à la différence d'une équation différentielle, nous avons une condition aux deux bouts du segment $[0, L]$. En fait, dans ce premier exemple élémentaire, on remarquera que cette équation est celle d'une parabole si ϱ et κ sont constants. Dans le cas général, il va nous falloir développer une théorie adaptée pour l'existence de solutions de ce type de problème aux limites.

Ici, quelques remarques s'imposent. Tout d'abord, nous avons travaillé par conditions nécessaires pour obtenir la forme variationnelle des équations d'Euler-Lagrange. et il a fallu choisir l'espace $\mathcal{V}$. Puis ensuite, nous avons dû supposer une régularité supplémentaire pour obtenir la forme forte. La question fondamentale sera dans ce cours de voir comment définir un cadre parfaitement rigoureux nous permettant de justifier ce passage entre la formulation variationnelle et la forme forte de l'équation d'Euler-Lagrange.

De plus, nous pouvons déjà murir l'intuition que choisir l'espace $\mathcal{V}$ répond par ailleurs à deux injonctions éventuellement contradictoires. Tout d'abord, nous avons contraint $\mathcal{V}$ à ne contenir que des fonctions suffisamment régulières. Ainsi, $\mathcal{V}$ doit être relativement petit. Mais par ailleurs, il faut tout de même que le minimum de la fonctionnelle tombe dans cet espace et, pour obtenir l'équation d'Euler-Lagrange, il nous faut vérifier (1.3) pour toute variation w dans un ensemble assez grand. Donc, $\mathcal{V}$ devra être relativement grand. Par ailleurs, nous pouvons aussi voir que l'espace C^1 , n'est peut-être pas le plus grand espace possible. En effet, on verra qu'on peut définir la fonctionnelle sur un espace tel que la fonction et sa dérivée sont définies presque partout pourvu qu'elles soient L^2 , puis en obtenir la formulation variationnelle ...

Remarque 1.1.1 () Historiquement, l'exemple emblématique du calcul des variations est le cas d'une chaîne inélastique toujours accrochée entre deux points.¹¹ Cependant il est plus complexe que le problème précédent car il n'y a plus d'énergie élastique et la minimisation du potentiel gravitationnel

$$\mathcal{J}(v) = - \int_0^L \varrho g v(x) \sqrt{1 + v'(x)^2} \, dx,$$

s'effectue sous la contrainte que la longueur de la chaînette reste

10: voir plus loin le complément historique au paragraphe 1.5

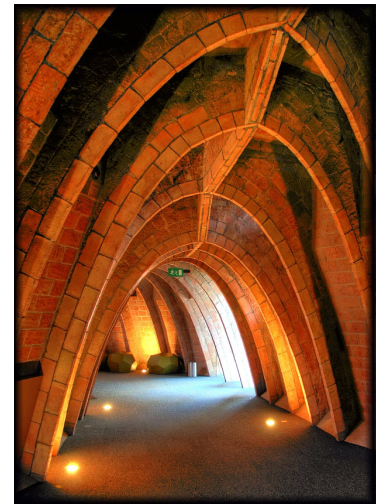


FIGURE 1.1 – Casa Milà, Barcelone, Espagne. Arcs en chaînette d'Antonio Gaudí, que celui-ci imagine en observant dans un miroir des chaînettes suspendues, Author : Matthias Ott, Source : <https://commons.wikimedia.org/>

11: il est considéré dès 1638 par Galilée et donnant lieu à de nombreux exemples en architecture, voir la figure 1.1

12: on notera que le lagrangien est ici une fonctionnelle dépendant de la fonction $v : x \mapsto v(x)$ et du multiplicateur de Lagrange $q \in \mathbb{R}$ ici car nous avons une contrainte scalaire. Précédemment la fonction de Lagrange était une fonction dépendant de quantité scalaire $v(x)$ et $q(x)$

constante :

$$\mathcal{E}(v) = \int_0^L \sqrt{1 + v'(x)^2} dx - L_{\text{tot}} = 0.$$

Dans le cas d'une minimisation sous contraintes d'égalité sous les conditions rappelées en annexe, la condition d'optimalité implique que les deux différentielles sont liées. Il existe donc $p \in \mathbb{R}$, appelé *multiplicateur de Lagrange*, telle qu'à l'optimum

$$D\mathcal{J}(u) = pD\mathcal{E}(u).$$

Cette condition est équivalente à chercher la condition d'optimalité d'une fonctionnelle modifiée

$$\mathcal{L}(v, q) = - \int_0^L \varrho g v(x) \sqrt{1 + v'(x)^2} dx + q \left(\int_0^L \sqrt{1 + v'(x)^2} dx - L_{\text{tot}} \right).$$

La fonctionnelle $\mathcal{L}$ est aussi appelée le lagrangien¹² du problème. Annuler la différentielle de $\mathcal{L}$ par rapport à q redonne la contrainte. Puisqu'ici $q \in \mathbb{R}$, on obtient en effet en calculant simplement une dérivée partielle

$$\partial_q \mathcal{L}(u, p) = 0 \Rightarrow \int_0^L \sqrt{1 + u'(x)^2} dx = L_{\text{tot}}.$$

Annuler la différentielle de $\mathcal{L}$ par rapport à v fournit les équations d'Euler-Lagrange associées à la recherche du minimum sous contrainte

$$\sqrt{1 + u'(x)^2} = \frac{u'(x)^2}{\sqrt{1 + u'(x)^2}} + \frac{u''(x)(u(x) - \frac{p}{\varrho g})}{(1 + u'(x)^2)^{\frac{3}{2}}},$$

ce qui donne finalement

$$u''(x) \left(u(x) - \frac{p}{\varrho g} \right) = 1 + u'(x)^2.$$

On peut reconnaître ici les dérivées d'un cosinus hyperbolique, soit

$$u(x) = a \cosh\left(\frac{x-b}{a}\right) + \frac{p}{\varrho g}.$$

Les conditions aux limites ainsi que la contrainte de longueur fixent les constantes (a, b) et p .

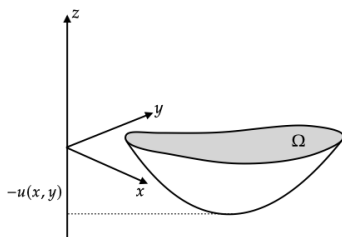


FIGURE 1.2 – Membrane élastique.

1.2 Problème modèle en dimension supérieure

1.2.1 Un exemple en dimension 2 d'espace

Supposons désormais non pas une chaînette mais une membrane élastique tendue soumise à la gravité. Dans le cas de petites déformations, on cherchera $u : \Omega \rightarrow \mathbb{R}$ qui minimise¹³, pour $\kappa \geq 0$,

$$\mathcal{J}(v) = \int_{\Omega} \left[\frac{\kappa}{2} |\nabla v(\underline{x})|^2 - \varrho g v(\underline{x}) \right] d\Omega.$$

13: On appelle parfois ce type d'énergie minimisée, énergie de Dirichlet, du nom

Cette minimisation doit s'effectuer sur un espace $\mathcal{V}$ tel que la fonctionnelle est définie. Un premier choix naturel est donc $\mathcal{V} \subset C^1(\Omega)$.

De plus si on veut prendre en compte le fait que l'on connaît u sur le bord $\Gamma = \partial\Omega$, on doit pouvoir considérer

$$u \in \mathcal{V} \cap \left\{ v \mid v|_{\partial\Omega} = u_\Gamma \right\}.$$

dont la régularité devra être compatible avec celle de la donnée u_Γ . Un premier choix peut-être $\mathcal{V} \subset C^0(\bar{\Omega})$ à supposer que $u_\Gamma \in C^0(\Gamma)$, une fonction continue en tant que fonction du bord¹⁴.

1.2.2 Notion de domaine en dimension supérieure

Nous venons donc de considérer des problèmes dans un domaine Ω ouvert de $\mathbb{R}^d$, en général borné, de frontière notée $\partial\Omega := \bar{\Omega} \setminus \Omega$ qui est supposée suffisamment régulière. Profitons en pour préciser mathématiquement cette notion.

Définition 1.2.1 Un ouvert Ω de $\mathbb{R}^d$ a une frontière $\partial\Omega$ localement lipchitzienne (on dira dans la suite suffisamment régulière) si en tout point $\underline{x}^*$ de $\partial\Omega$, il existe

- un repère orthonormé local et un système de coordonnées locales $(\underline{x}_\parallel, x_d) \in \mathbb{R}^{d-1} \times \mathbb{R}$ avec $\underline{x}^* = (\underline{x}_\parallel^*, x_d^*)$;
- un ouvert $\mathcal{O}$ de $\mathbb{R}^{d-1}$ contenant $\underline{x}_\parallel^*$ et I un ouvert de $\mathbb{R}$ contenant x_d^* ;
- une application lipchitzienne¹⁵ f de $\mathcal{O}$ dans I ;

tels qu'il existe un voisinage $\mathcal{V}(\underline{x}^*)$ de $\underline{x}^*$,

$$\Omega \cap \mathcal{V}(\underline{x}^*) = \{(\underline{x}_\parallel^*, x_d) \in \mathcal{O} \times I, \quad x_d > f(\underline{x}_\parallel^*)\}.$$

Dans ce cas, la normale unitaire extérieure est définie presque partout¹⁶ dans le système de coordonnées locales

$$p.p. \underline{x}^* = (\underline{x}_\parallel^*, f(\underline{x}_\parallel^*)) \in \partial\Omega, \quad \underline{n}(\underline{x}^*) = \frac{1}{\sqrt{1 + |\nabla f(\underline{x}_\parallel^*)|^2}} (\nabla f(\underline{x}_\parallel^*), -1).$$

On dira que la frontière est de classe C^k pour $k \geq 0$ si en tout point $\underline{x}^*$ la fonction f associée est de classe C^k .

En d'autres termes, on peut cartographier la frontière localement par une application lipchitzienne. On appelle système de cartes locales de la frontière, toutes les données associées aux points $\underline{x}^*$ de la frontière (le système de coordonnées, les ouverts et la fonction lipchitzienne). Si la frontière est bornée, seul un nombre fini de cartes locales est nécessaire pour définir un atlas de la frontière. On représente Figure 1.4 un certain nombre de domaines vérifiant cette propriété. Un exemple important est celui des polygones convexes ou non. Notons que leurs frontières sont lipchitziennes (donc de classe C^0) mais pas de classe C^1 .

Cette hypothèse de régularité de la frontière exclut de fait les domaines qui ne sont pas d'un seul côté de la frontière ou dont une des cartes locales est associée à une fonction non lipchitzienne (voir Figure 1.5).

14: nous voyons combien modélisation et régularité des fonctions considérées sont intimement liées. Relaxer les choix introduit ici à titre introductif sera l'objectif du cours.

15: Soient X et Y deux espaces normés, $f : X \rightarrow Y$ est dite lipchitzienne ssi

$$\exists \eta > 0, \forall x, x' \in X, \quad \|f(x) - f(x')\|_Y \leq \eta \|x - x'\|_X.$$

16: Comme f est lipchitzienne alors elle est dérivable presque partout : c'est le théorème de Rademacher.

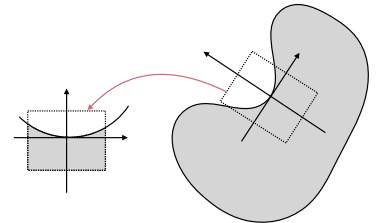


FIGURE 1.3 – Cartographie de la frontière.



FIGURE 1.4 – Exemples de domaines à frontière localement lipchitzienne.

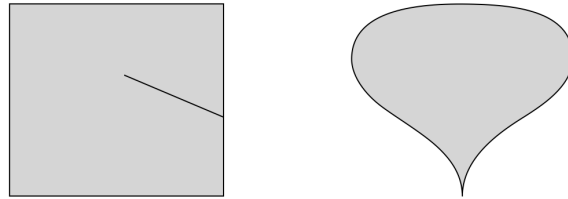


FIGURE 1.5 – Exemples de domaines dont la frontière n'est pas localement lipchitzienne.

L'analyse des équations aux dérivées partielles (EDPs) posées dans ce type de domaines dépasse le cadre de ce cours.

1.2.3 Formules de Green pour les fonctions régulières

Afin d'obtenir les conditions d'optimalité en calcul des variations en dimension supérieure à 1, nous devons introduire la formule de Stokes. Cette formule jouent un rôle essentiel dans ce cours d'analyse variationnelle et plus généralement en analyse des EDPs.

C'est une généralisation en dimension $d > 1$ de la formule d'intégration pour des fonctions réelles : soit $a, b \in \mathbb{R}$

$$\forall f \in C^1([a, b]) \cap C^0([a, b]), \quad \int_a^b f'(x) \, dx = f(b) - f(a). \quad (1.5)$$

Nous le présentons tout d'abord dans le cadre de fonctions régulières mais un des enjeux de ce cours sera ensuite de l'étendre à des espaces de fonctions moins régulières.

17: Cette formule peut être étendue au cas où Ω n'est pas borné à condition que la fonction w soit à support compact. Elle peut également être étendue à des domaines Ω plus généraux [1] mais cela dépasse le cadre de ce cours.

Théorème 1.2.1 (Théorème de Stokes) Soit Ω un ouvert borné de $\mathbb{R}^d$ de frontière suffisamment régulière¹⁷ et $w \in C^1(\Omega) \cap C^0(\overline{\Omega})$ alors pour tout $i \in \llbracket 1, d \rrbracket$

$$\int_{\Omega} \frac{\partial w}{\partial x_i} \, d\Omega = \int_{\partial\Omega} w n_i \, d\Gamma,$$

où n_i est la i -ème composante de la normale unitaire $\underline{n}$ à Ω dirigée vers l'extérieur.

Notons que pour $d = 1$, cette formule n'est rien d'autre que (1.5) puisque pour $\Omega =]a, b[$, $\partial\Omega = \{a\} \cup \{b\}$, $n = -1$ en a et $n = 1$ en b .

Démonstration. Pour simplifier la preuve, nous proposons de considérer le cas $d = 2$, $i = 2$ et un domaine Ω défini à partir de $a < b$ deux réels et

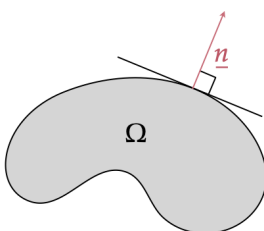


FIGURE 1.6 – Un domaine Ω et la normale unitaire extérieure associée à un point du bord.

deux fonctions g_- , g_+ de classe C^1 :

$$\Omega = \{(x_1, x_2) \in \mathbb{R}^2, a < x_1 < b, g_-(x_1) < x_2 < g_+(x_1)\}.$$

Dans le cas général, il faut utiliser une partition de l'unité de Ω , des cartes locales afin de se ramener à un calcul similaire (mais dans $\mathbb{R}^d$) à celui présenté ici. Pour une démonstration complète, nous renvoyons le lecteur à [4] (une référence parmi d'autres).

Dans ce cas, on a

$$\int_{\Omega} \frac{\partial w}{\partial x_2} d\Omega = \int_{x_1=a}^b \int_{x_2=g_-(x_1)}^{g_+(x_1)} \frac{\partial w}{\partial x_2} dx_1 dx_2 \quad (1.6)$$

$$= \int_a^b (w(x_1, g_+(x_1)) - w(x_1, g_-(x_1))) dx_1. \quad (1.7)$$

Nous cherchons maintenant à expliciter l'intégrale surfacique de $w \underline{n}_2$ qui apparaît à droite de l'égalité recherchée. L'intégrale sur le bord de $w \underline{n}_2$ est la somme de ses intégrales sur les bouts du bord définis par $\Gamma_a := \{(x_1, x_2) \in \partial\Omega, x_1 = a\}$, $\Gamma_b := \{(x_1, x_2) \in \partial\Omega, x_1 = b\}$, $\Gamma_{\pm} := \{(x_1, x_2) \in \partial\Omega, x_2 = g_{\pm}(x_1)\}$:

$$\int_{\partial\Omega} w n_2 d\Gamma = \int_{\Gamma_a} w n_2 d\Gamma + \int_{\Gamma_-} w n_2 d\Gamma + \int_{\Gamma_b} w n_2 d\Gamma + \int_{\Gamma_+} w n_2 d\Gamma.$$

Sur les bords Γ_a et Γ_b , la normale $\underline{n}$ est dans la direction e_1 donc $n_2 = 0$. Les intégrales de $w n_2$ sur Γ_a et Γ_b sont donc nulles. En utilisant la paramétrisation de Γ_- , on a par définition pour toute fonction F définie sur Γ_-

$$\int_{\Gamma_-} F d\Gamma = \int_a^b F(x_1, g_-(x_1)) \sqrt{1 + (g'_-(x_1))^2} dx_1.$$

Toujours en utilisant cette paramétrisation, on montre que la normale $\underline{n}$ au point $(x_1, g_-(x_1))$ est donnée par $(g'_-(x_1), -1)/\sqrt{1 + (g'_-(x_1))^2}$. Ceci nous donne que

$$\begin{aligned} \int_{\Gamma_-} w n_2 d\Gamma &= \int_a^b (w \underline{n}_2)(x_1, g_-(x_1)) \sqrt{1 + (g'_-(x_1))^2} dx_1 \\ &= - \int_a^b w(x_1, g_-(x_1)) dx_1. \end{aligned}$$

Pour l'intégrale sur Γ_+ , les calculs sont similaires à ceci près que la normale $\underline{n}$ au point $(x_1, g_+(x_1))$ est donnée par $(g'_+(x_1), 1)/\sqrt{1 + (g'_+(x_1))^2}$. Ainsi,

$$\int_{\Gamma_+} w n_2 d\Gamma = \int_a^b w(x_1, g_+(x_1)) dx_1.$$

On a donc

$$\int_{\partial\Omega} w n_2 d\Gamma = \int_a^b w(x_1, g_+(x_1)) dx_1 - \int_a^b w(x_1, g_-(x_1)) dx_1.$$

En utilisant (1.7), nous avons bien prouvé le théorème. $\square$

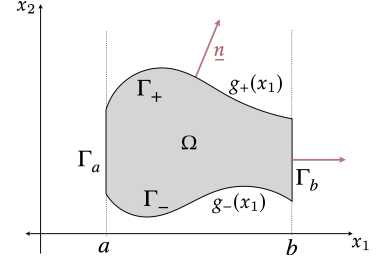


FIGURE 1.7 – Le domaine Ω pour lequel la preuve du théorème 1.2.1 est réalisée.

La formule de Stokes permet de déduire assez facilement un certain

nombre d'autres formules dites *formules de Green*. Tout d'abord la formule de Stokes vectorielle qui se démontre grâce au théorème 1.2.1 en raisonnant composante par composante.

Corollaire 1.2.2 [Théorème de la divergence] Soit Ω un ouvert borné de $\mathbb{R}^d$ de frontière suffisamment régulière et $\underline{W} \in [C^1(\Omega)]^d \cap [C^0(\overline{\Omega})]^d$ alors

$$\int_{\Omega} \underline{\nabla} \cdot \underline{W} \, d\Omega = \int_{\partial\Omega} \underline{W} \cdot \underline{n} \, d\Gamma,$$

où $\underline{n}$ est la normale unitaire à Ω dirigée vers l'extérieur.¹⁸

18: On rappelle que la divergence du champ de vecteur $\underline{W}$ est donnée par $\underline{\nabla} \cdot$

$$\underline{W} = \sum_{i=1}^d \frac{\partial W_i}{\partial x_i}.$$

De plus, en appliquant le théorème 1.2.1 à $w = uv$ avec $u, v \in C^1(\Omega) \cap C^0(\overline{\Omega})$, on obtient pour tout $i \in \llbracket 1, d \rrbracket$

$$\int_{\Omega} \left[\frac{\partial u}{\partial x_i} v + u \frac{\partial v}{\partial x_i} \right] d\Omega = \int_{\partial\Omega} u v \underline{n}_i \, d\Gamma. \quad (1.8)$$

De la même façon, en appliquant le corollaire 1.2.2 à $\underline{W} = \underline{U}v$ avec $\underline{U} \in [C^1(\Omega)]^d \cap [C^0(\overline{\Omega})]^d$, $v \in C^1(\Omega) \cap C^0(\overline{\Omega})$, on obtient

$$\int_{\Omega} [v \underline{\nabla} \cdot \underline{U} + \underline{U} \cdot \underline{\nabla} v] \, d\Omega = \int_{\partial\Omega} \underline{U} \cdot \underline{n} v \, d\Gamma. \quad (1.9)$$

Il est facile de déduire en particulier les formules de Green faisant apparaître des dérivées d'ordre 2. En effet, en appliquant le théorème 1.2.1 à $w = \partial_{x_i} u v$ avec $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$, $v \in C^1(\Omega) \cap C^0(\overline{\Omega})$, on obtient pour tout $i \in \llbracket 1, d \rrbracket$

$$\int_{\Omega} \left[\frac{\partial^2 u}{\partial x_i^2} v + \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} \right] d\Omega = \int_{\partial\Omega} \frac{\partial u}{\partial x_i} v \underline{n}_i \, d\Gamma.$$

On montre de la même façon, à partir du corollaire 1.2.2 par exemple cette formule de Green qui sera beaucoup utilisée pendant ce cours.

Corollaire 1.2.3 Soit Ω un ouvert borné de $\mathbb{R}^d$ de frontière suffisamment régulière et $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$, $v \in C^1(\Omega) \cap C^0(\overline{\Omega})$ alors

$$\int_{\Omega} [\Delta u v + \underline{\nabla} u \cdot \underline{\nabla} v] \, d\Omega = \int_{\partial\Omega} \underline{\nabla} u \cdot \underline{n} v \, d\Gamma,$$

où $\underline{n}$ est la normale unitaire à Ω dirigée vers l'extérieur.¹⁹

19: On rappelle que le laplacien d'une fonction u est donné par

$$\Delta u := \underline{\nabla} \cdot (\underline{\nabla} u) = \sum_{i=1}^d \frac{\partial^2 u_i}{\partial x_i^2}.$$

1.2.4 Des formules de Green aux équations d'Euler Lagrange

Nous disposons désormais de tous les outils pour les obtenir une condition nécessaire à notre problème de minimisation. On commence par calculer

$$\langle D\mathcal{J}(u), w \rangle_{\mathcal{V}', \mathcal{V}} = \int_{\Omega} [\kappa \underline{\nabla} u(\underline{x}) \cdot \underline{\nabla} w(\underline{x}) - \varrho g w(\underline{x})] \, d\Omega,$$

pour toute variation $w \in \mathcal{V}$ de l'élévation u . Comme $v = u + w$ reste une élévation possible nous avons $v = u_\Gamma$ sur Γ donc $w = 0$ sur Γ . Ainsi pour $u = \operatorname{argmin}_{v \in \mathcal{V}} \mathcal{J}(v)$, la condition d'optimalité donne la *formulation variationnelle*

$$\forall w \in \mathcal{V}_0, \quad \int_{\Omega} \left[\kappa \underline{\nabla} u(\underline{x}) \cdot \underline{\nabla} w(\underline{x}) - \varrho g w(\underline{x}) \right] d\Omega = 0. \quad (1.10)$$

où

$$\mathcal{V}_0 = \mathcal{V} \cap \{v \mid v|_{\partial\Omega} = 0\}.$$

La formulation précédente conduit à la définition d'un problème mis sous une forme variationnelle.

Définition 1.2.2 (Problème sous forme variationnelle) *Soit $\mathcal{V}$ un espace vectoriel, $a : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ une forme bilinéaire et $\ell : \mathcal{V} \rightarrow \mathbb{R}$ une forme linéaire. On appelle problème variationnel, le fait de chercher $u \in \mathcal{V}$ tel que*

$$\forall v \in \mathcal{V}, \quad a(u, v) = \ell(v) \quad (1.11)$$

Ainsi dans notre cas, puisque $w = 0$ sur $\partial\Omega$, en appliquant la formule (1.9) à $\underline{U} = \kappa \underline{\nabla} u$ et $v = w$, on trouve

$$\langle D\mathcal{J}(u), w \rangle_{\mathcal{V}', \mathcal{V}} = \int_{\Omega} \left[-\underline{\nabla} \cdot (\kappa \underline{\nabla} u(\underline{x})) - \varrho g \right] w(\underline{x}) d\Omega.$$

Supposons κ constant pour simplifier et rappelons que

$$\underline{\nabla} \cdot \underline{\nabla} u = \sum_{1 \leq i \leq d} \frac{\partial^2 u}{\partial x_i^2} = \Delta u,$$

alors la *formulation forte* des équations d'Euler-Lagrange est ici

$$\begin{cases} -\kappa \Delta u(\underline{x}) = \varrho g, & \underline{x} \in \Omega, \\ u(\underline{x}) = u_\Gamma(\underline{x}), & \underline{x} \in \Gamma. \end{cases}$$

C'est une *équation aux dérivées partielles (EDP) elliptique* au sens de la définition suivante.

Définition 1.2.3 (EDP elliptique) *Une équation aux dérivées partielles linéaire du second ordre, dont la forme générale est donnée par :*

$$\sum_{i,j=1}^n a_{ij}(\underline{x}) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(\underline{x}) \frac{\partial u}{\partial x_i} + c(\underline{x})u = f(\underline{x}), \quad \underline{x} \in \Omega \subset \mathbb{R}^d$$

est dite elliptique en un point donné $\underline{x}$ de l'ouvert Ω si le tenseur symétrique $\underline{a}(\underline{x}) = (a_{ij})_{1 \leq i,j \leq n}$ des coefficients du second ordre admet des valeurs propres non nulles et de même signe.

1.3 Autres exemples de problèmes variationnels

1.3.1 Un problème électrostatique

Il existe de nombreux autres phénomènes physiques dont les équations s'obtiennent par le calcul des variations. C'est notamment le cas pour l'équation du potentiel électrostatique $\phi : \mathbb{R}^3 \ni \underline{x} \mapsto \phi(\underline{x}) \in \mathbb{R}$ créé par une distribution de charge $\varrho : \mathbb{R}^3 \ni \underline{x} \mapsto \varrho(\underline{x}) \in \mathbb{R}$. Le principe variationnel implique ici de minimiser la différence entre l'énergie du champ

$$\mathcal{E}_E = \frac{\varepsilon_0}{2} \int_{\mathbb{R}^3} |\underline{E}(\underline{x})|^2 d\Omega,$$

et l'énergie potentielle électrostatique de la distribution de charge $\varrho(\underline{x})$ donnée par

$$\mathcal{E}_\rho = \int_{\mathbb{R}^3} \varrho(\underline{x}) \phi(\underline{x}) d\Omega.$$

De plus, le champ électrostatique s'exprime à partir du potentiel par $\underline{E} = -\nabla \phi$ donc la fonctionnelle à minimiser est

$$\phi = \operatorname{argmin}_{\psi \in \mathcal{V}} \left\{ \mathcal{J}(\psi) := \int_{\mathbb{R}^3} \left[\frac{\varepsilon_0}{2} |\nabla \psi(\underline{x})|^2 - \varrho(\underline{x}) \psi(\underline{x}) \right] d\Omega \right\}.$$

C'est un exemple typique de problème de minimisation posé dans un espace non-borné, ici $\Omega = \mathbb{R}^3$. Nous supposons que les charges sont distribuées dans un sous-espace borné et il faudra que l'espace $\mathcal{V}$ contienne des fonctions à décroissance suffisamment rapide pour que la fonctionnelle soit bien définie. En considérant des accroissement de la fonctionnelle pour des fonctions régulières à support-compact $C_c^\infty(\mathbb{R}^3)$, le théorème de la divergence nous permet alors de retrouver l'équation de Poisson²⁰ dans tout l'espace

$$-\Delta \phi(\underline{x}) = \frac{\varrho(\underline{x})}{\varepsilon_0}, \quad \underline{x} \in \mathbb{R}^3.$$

À noter que dans ce cours, nous nous limiterons sauf exception à des configurations en domaine borné.

1.3.2 Calcul des variations en traitement de signal

Enfin le calcul des variations n'apparaît pas que dans les systèmes physiques que nous observons, la nature n'étant pas la seule à vouloir définir une solution par un principe d'optimalité.

Supposons par exemple que nous ayons une image bruitée représentée par ses niveaux de gris $\Omega \ni \underline{x} \mapsto u_0(\underline{x})$ et que nous souhaitons la débruiter. Une stratégie possible est de minimiser la fonctionnelle

$$\mathcal{J}(v) = \frac{1}{2} \int_{\Omega} |v - u_0|^2 d\Omega + \frac{1}{2} \int_{\Omega} \alpha(\underline{x}) |\nabla v|^2 d\Omega,$$

où le premier terme dit *d'attache aux données* indique que nous cherchons u aussi proche de u_0 que possible et le deuxième terme dit de *régularisation* indique que nous voulons pénaliser les grandes variations de v que nous

20: Siméon Denis Poisson (1871-1840) est un mathématicien, géomètre et physicien français. Ancien élève de l'École polytechnique (1798) il succède à Joseph Fourier en 1806 en tant que professeur

associons au bruit présent dans u_0 . Le champ scalaire $\Omega \ni \underline{x} \mapsto \alpha(\underline{x}) \in \mathbb{R}$ permet éventuellement d'amplifier la régularisation dans certaines zones de l'image.

De nouveau, il faut définir un espace $\mathcal{V}$ dans lequel minimiser cette fonctionnelle si on veut *a minima* donner un sens à chacun des termes. Formellement pour l'instant, les équations d'Euler-Lagrange associées à la minimisation de $\mathcal{J}$, donnent la formulation forte suivante

$$\begin{cases} u - u_0 - \nabla \cdot (\alpha(\underline{x}) \nabla u) = 0, \\ \frac{\partial u}{\partial n} = 0. \end{cases}$$

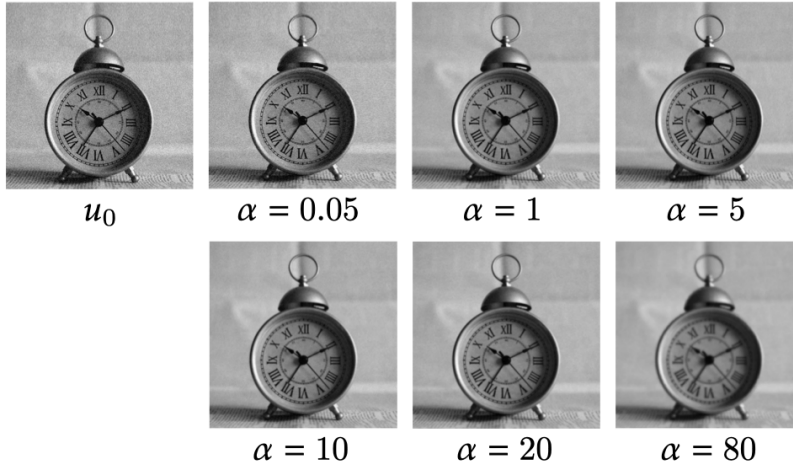


FIGURE 1.8 – Image reconstruite pour différents paramètres de régularisation α . Reproduction autorisée par A. Chambolle. Source : interactions19.sciencesconf.org

Nous pouvons même en toute généralité remplacer le champ scalaire α par un champ de tenseurs $\underline{\kappa} : \Omega \ni \underline{x} \in \underline{\kappa}(\underline{x}) \in \mathcal{T}_2(\mathbb{R})$ qu'on identifiera aux applications linéaires $\mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ ou $\cdot$ est le produit de contraction qu'on assimilera au produit scalaire. Un tel champ de tenseurs permet introduire une régularisation anisotrope. La fonctionnelle devient alors

$$\mathcal{J}(u) = \frac{1}{2} \int_{\Omega} |u - u_0|^2 \, d\Omega + \frac{1}{2} \int_{\Omega} \nabla u \cdot \underline{\kappa}(x) \cdot \nabla u \, d\Omega.$$

Une difficulté dans la régularisation d'image est donc que les contours de l'image sont aussi des zones de fort gradient et donc une régularisation en norme L^2 du gradient a tendance à filtrer le contour en même temps que le bruit. Une idée est alors d'ajuster le tenseur de régularisation spatialement en fonction du gradient de u , on parle alors de diffusion de Perona-Malik avec par exemple

$$\alpha(\underline{x}) = e^{-\lambda \|\nabla u\|_{L^2(\Omega)}^2}, \text{ où } \alpha(\underline{x}) = \frac{1}{1 + \|\nabla u\|(\underline{x})_{L^2(\Omega)}^2}.$$

L'équation d'Euler-Lagrange devient alors non linéaire. On peut dans ce cas, essayer de remplacer la non-linéarité par une stratégie itérative, par exemple²¹ en résolvant sur quelques itérations la suite $(u_k)_k$ de solutions

21: de nombreux autres schémas sont possibles et la question de leur convergence doit être étudiée dans chaque cas

des problèmes

$$\begin{cases} u_k - u_{k-1} - \nabla \cdot \left(e^{-\lambda \|\nabla u_{k-1}\|_{L^2(\Omega)}^2} \nabla u_k \right) = 0, & \text{dans } \Omega, k \geq 1, \\ \frac{\partial u_k}{\partial n} = 0, & \text{sur } \partial\Omega. \end{cases}$$

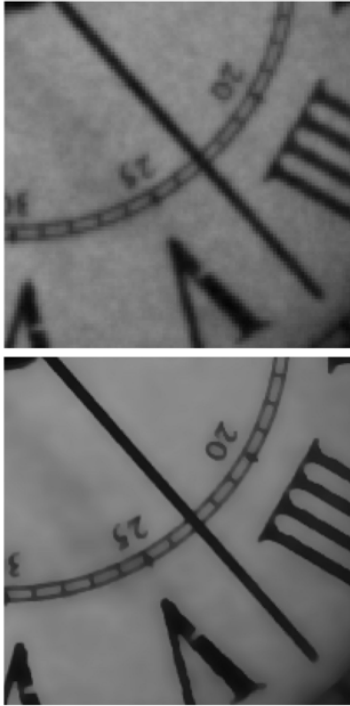


FIGURE 1.9 – Régularisation de type variation totale qui préserve mieux les contours. Reproduction autorisée par A. Chambolle. Source : interactions19.sciencesconf.org

Une autre régularisation très efficace et qui préserve mieux les contours de l'image : c'est la régularisation dite par variation totale, voir par exemple [7] pour des discrétisations par la méthode des éléments finis avec

$$\mathcal{J}(u) = \frac{1}{2} \int_{\Omega} |u - u_0|^2 \, d\Omega + \frac{1}{2} \int_{\Omega} \alpha(\underline{x}) |\nabla u| \, d\Omega.$$

Dans ce cas, on montre que l'équation d'Euler-Lagrange associée est

$$\begin{cases} u - u_0 - \nabla \cdot \left(\alpha(\underline{x}) \frac{\nabla u}{|\nabla u|} \right) = 0, & \text{sur } \Omega \\ \frac{\partial u}{\partial n} = 0, & \text{sur } \partial\Omega \end{cases}$$

qui est, elle aussi, non linéaire.

1.3.3 Équations des poutres et plaques en flexion

Présentons encore deux modèles issus de la mécanique avec les énergies de poutre et de plaque. Dans le premier cas, l'énergie de déformation d'une poutre en flexion est proportionnelle, dans la limite des petits déplacements, à sa courbure. On a donc une énergie mécanique du type

$$\mathcal{E}(v) = \int_0^L v''(x)^2 \, dx,$$

qui peut être combinée à une énergie potentielle pour donner la fonctionnelle

$$\mathcal{J}(v) = \int_0^L \left[\kappa v''(x)^2 - f v(\underline{x}) \right] \, dx.$$

Les formes fortes des équations d'Euler-Lagrange donnent dans ce cas (voir aussi le paragraphe ??)

$$u^{(4)}(x) = f, \quad x \in]0, L[.$$

pour typiquement des conditions aux limites

$$u(0) = u(L) = 0 \quad \text{et} \quad u'(0) = u'(L) = 0,$$

caractérisant l'encastrement aux bords.

En dimension 2, cette fois c'est le laplacien qui intervient dans l'énergie de courbure pour donner

$$\mathcal{J}(u) = \int_{\Omega} \left[\frac{\kappa}{2} (\Delta u(\underline{x}))^2 - f u(\underline{x}) \right] \, d\Omega,$$

qui donnera cette fois-ci une équation sur le bilaplacien

$$\kappa \Delta^2 u = f,$$

complétée par des conditions aux limites sur la fonction et sa dérivée normale, par exemple

$$u(x) = 0 \quad \text{et} \quad \frac{\partial u}{\partial n} = 0.$$

1.4 Des lois de conservations aux solutions à l'équilibre

1.4.1 Lois des conservations et lois de comportements

Nous venons de voir un certain nombre de problèmes physiques formulés par minimisation d'une fonctionnelle. Un autre point de vue est possible à partir de la notion de loi de conservation. On considère une grandeur extensive²² $\phi : \Omega \times \mathbb{R} \ni (x, t) \mapsto \phi(x, t) \in \mathbb{R}$ entraînée par un flux $\underline{j}$ et comportant un terme de production volumique s . La loi de conservation de ϕ exprime alors que la variation interne de ϕ sur tout sous-domaine suffisamment régulier ω est la somme des contributions induites par le terme source s sur ce sous-domaine et des quantités entrantes par le bord $\partial\omega$ du sous-domaine à la vitesse $\underline{v}$, soit

$$\int_{\omega} \frac{\partial \phi}{\partial t} d\Omega = \int_{\omega} s d\Omega - \int_{\partial\omega} \underline{j} \cdot \underline{n} d\Gamma,$$

où on rappelle que le vecteur normal unitaire $\underline{n}$ est orienté par convention vers l'extérieur de ω . En appliquant le théorème de la divergence, on obtient alors la formulation

$$\int_{\omega} \left[\frac{\partial \phi}{\partial t} + \underline{\nabla} \cdot \underline{j} \right] d\Omega = \int_{\omega} s d\Omega.$$

Comme cette relation est vraie pour tout $\omega \in \Omega$ alors la loi de conservation s'écrit sous la *forme forte* suivante

$$\frac{\partial \phi}{\partial t} + \underline{\nabla} \cdot \underline{j} = s, \quad \forall x \in \Omega. \quad (1.12)$$

De nombreux problèmes physiques se présentent sous cette forme. Reste ensuite à définir la quantité extensive ϕ . Par exemple, pour une densité volumique de charge, on a $\phi = \varrho$. Pour un bilan de masse, on définira $\phi = \varrho c$ où c est une concentration et ϱ une densité. Pour la chaleur, on aura par exemple à pression constante $\phi = \varrho c_p T$.

Cette loi de conservation doit ensuite être étendue en caractérisant le flux par une loi dite de comportement²³. Voici trois exemples emblématiques de couples $(u, \underline{j}_u)$:

- loi d'Ohm $\underline{j} = -\gamma \underline{E}$ où $\underline{E}$ est le champ électrique,
- loi de Fick $\underline{j} = -\underline{\kappa} \underline{\nabla} c$,
- loi de Fourier $\underline{j} = -\underline{\kappa} \underline{\nabla} T$,

où $\Omega \ni x \mapsto \underline{\kappa}(x) \in \mathbb{S}_d^{+*}(\mathbb{R})$ est un champ de tenseurs d'ordre 2 symétriques définis positifs (voir aussi la remarque 1.4.1). Ainsi le signe – traduit le fait que le flux s'oppose au gradient de la quantité pour équilibrer l'occupation spatiale.

22: une grandeur est extensive lorsque la somme des valeurs de cette grandeur pour deux systèmes disjoints est égale à la valeur de la grandeur pour la réunion des systèmes

23: Si la loi de conservation traduit un concept qu'on pourrait considérer comme universel en physique, la loi de comportement, elle, sera heuristique et traduira notre modélisation du système sous-jacent. Cette modélisation pourra cependant être contrainte par un principe physique tel que le second principe

Remarque 1.4.1 (Second Principe) Pour la loi de Fourier, le second principe indique typiquement que si

$$\delta S = \int_{\omega} \frac{\rho c_p}{T} \frac{\partial T}{\partial t} d\Omega = - \int_{\omega} \frac{\nabla \cdot \underline{j}}{T} d\Omega,$$

est la variation d'entropie lors d'une transformation (à pression constante) et

$$\delta S_e = - \int_{\partial\omega} \frac{\underline{j} \cdot \underline{n}}{T} d\Gamma,$$

est l'entropie échangée par conduction avec l'extérieur, alors on doit nécessairement avoir

$$0 < \delta S - \delta S_e = \int_{\omega} \left[\nabla \cdot \left(\frac{\underline{j}}{T} \right) - \frac{\nabla \cdot \underline{j}}{T} \right] d\Omega = - \int_{\omega} \frac{\underline{j} \cdot \nabla T}{T^2} d\Omega.$$

La loi de Fourier est la loi phénoménologique la plus simple qui réalise cette condition.

À une renormalisation près, les lois de conservation couplées aux lois de comportement vont donc se mettre sous la forme un peu générique

$$\rho \frac{\partial u}{\partial t} - \nabla \cdot (\underline{\kappa} \cdot \nabla u) = f, \quad \text{dans } \Omega. \quad (1.13)$$

De plus, au bord du domaine, les interactions avec l'extérieur impliquent de nouveau le flux normal $\underline{j} \cdot \underline{n}$ sous la forme de l'existence d'un terme *source frontière* g tel que

$$(\underline{\kappa} \cdot \nabla u) \cdot \underline{n} = g, \quad \underline{x} \in \partial\Omega. \quad (1.14)$$

On appelle cette condition *une condition de Neumann* par opposition à *une condition de Dirichlet* où on imposerait la valeur de la solution au bord.

Les termes sources f et g peuvent être imposés, mais ils peuvent tout autant modéliser un comportement du système. Par exemple $f(u) = -\alpha u + f_0$ modélisera typiquement le cas d'un système réactif où une proportion du réactif devient actif ou inerte. De nouveau pour respecter le second principe, si le système ne peut que retourner à l'équilibre on aura $\alpha \geq 0$. La dynamique de u se présente alors sous la forme éventuellement plus générale

$$\rho \frac{\partial u}{\partial t} - \nabla \cdot (\underline{\kappa} \cdot \nabla u) + \alpha u = f, \quad \text{dans } \Omega. \quad (1.15)$$

Les mêmes enjeux de modélisation prévalent au bord du domaine. On peut imaginer que g imposé est typiquement ce qu'on attend d'un radiateur. Pour un climatiseur, on définira $g(u) = -\beta(u - u_0)$ car le flux de chaleur est alors proportionnel à l'écart de température. La condition au bord devient alors

$$(\underline{\kappa} \cdot \nabla u) \cdot \underline{n} + \beta(u - u_0) = g, \quad \text{sur } \partial\Omega, \quad (1.16)$$

Cette condition de proportionnalité entre le flux et la valeur de la fonction est appelée *une condition de Robin* ou *une condition de Fourier*. C'est un cas intermédiaire entre une condition de Neumann et une condition de

Dirichlet.

1.4.2 Problèmes stationnaires et principe variationnel

Si u ne dépend pas du temps t , on dit qu'on a atteint un état stationnaire qui peut être interprété comme la limite asymptotique²⁴ quand $t \rightarrow \infty$ du problème en temps. On obtient

$$\begin{cases} -\underline{\nabla} \cdot (\underline{\kappa} \underline{\nabla} u) + \alpha u = f, & \text{dans } \Omega, \\ (\underline{\kappa} \cdot \underline{\nabla} u) \cdot \underline{n} + \beta(u - u_0) = 0, & \text{sur } \partial\Omega. \end{cases} \quad (1.17)$$

24: asymptotique est ici fondamental car il représente un cas qui n'est jamais exactement vérifié mais pertinent à étudier en tant que limite.

Et cet équilibre correspond à un principe variationnel au sens où, la solution à l'équilibre minimise

$$\begin{aligned} \mathcal{J}(v) = & \frac{1}{2} \int_{\Omega} \underline{\nabla} v \cdot \underline{\kappa} \cdot \underline{\nabla} v \, d\Omega \\ & + \frac{1}{2} \int_{\Omega} \alpha v^2 \, d\Omega + \frac{1}{2} \int_{\partial\Omega} \beta v^2 \, d\Gamma - \int_{\Omega} f v \, d\Omega + \int_{\partial\Omega} \beta u_0 v \, d\Omega. \end{aligned} \quad (1.18)$$

1.4.3 Les problèmes modèles de ce cours

Dans ce cours nous allons donc étudier principalement des équations aux dérivées partielles de la forme de

$$-\underline{\nabla} \cdot (\underline{\kappa} \underline{\nabla} u) + \alpha u = f, \quad \text{dans } \Omega, \quad (1.19)$$

où le terme source f , les fonctions $\underline{\kappa}$ et $\alpha \geq 0$ (qui modélisent des propriétés physiques du milieu) sont données. On étudiera souvent le cas où $\kappa = \bar{\kappa} = 1$ et on verra que le cas $\alpha = 0$ est un peu plus délicat à traiter que la situation $\alpha > 0$.

Les conditions aux limites modèles que nous considérons sur tout ou partie du bord sont de la forme

$$\begin{aligned} (\text{Dirichlet}) \quad & u = g_D, \quad \text{sur } \Gamma_D, \\ (\text{Neumann}) \quad & \underline{\nabla} u \cdot \underline{n} = g_N, \quad \text{sur } \Gamma_N, \\ (\text{Robin}) \quad & \underline{\nabla} u \cdot \underline{n} + \lambda u = g_R, \quad \text{sur } \Gamma_R, \end{aligned}$$

où les termes sources g_D , g_N , g_R et la fonction λ sont donnés. Nous verrons que pour les équations qui s'écrivent comme (1.19), au moins une condition aux limites est nécessaire sur tout le bord : $\overline{\Gamma_D} \cup \overline{\Gamma_N} \cup \overline{\Gamma_R} = \partial\Omega$ mais qu'il en faut une seule pour espérer trouver une solution $\Gamma_D \cap \Gamma_N = \Gamma_D \cap \Gamma_R = \Gamma_R \cap \Gamma_N = \emptyset$.

Ces équations sont souvent, mais pas toujours, associées à un principe d'optimalité (de moindre action). Étudier ces EDP sera donc intimement lié à l'étude du problème de minimisation par exemple (1.18) and le cas du problème (1.18), ou aux formulations variationnelles associées aux conditions d'optimalité du premier d'ordre d'Euler-Lagrange.

Une question centrale de ce cours est de savoir comment on souhaite que (1.19) et que les conditions aux limites soient vérifiées. En effet pour l'instant nous avons toujours supposé que les fonctions considérées étaient suffisamment régulières, notamment C^2 dans le cas de (1.19) ou

C^1 dans le cas de (1.17). Hélas, rien n'est aussi simple et ceci pour des raisons que nous pouvons anticiper dès maintenant.

Tout d'abord sur la formulation forte, nous pouvons imaginer que la modélisation du terme source (1.19) nécessite que f ne soit pas $C^0(\Omega)$. En effet, si on veut prendre en compte le fait que la source est bien localisée sur un sous domaine, on aurait plus naturellement envie de choisir $f \in L^2(\Omega)$ mais dans ce cas (1.19) ne peut pas avoir de sens pour $u \in C^2(\Omega)$. De deux choses l'une, soit le modèle n'est pas valide, soit (1.19) doit être vérifiée en un sens plus faible. Par exemple la solution, si elle existe, ne vérifie peut-être pas toujours ces équations presque partout. On verra alors que les dérivées ne sont pas toujours à comprendre au sens classique...

Réciproquement, en prenant le problème via la minimisation de fonctionnelle, nous avons supposé que la minimum était atteint dans l'espace $\mathcal{V}$. Or rien ne l'assure a priori et nous pouvons même nourrir l'intuition que ce n'est pas vrai.

Considérons en effet pour simplifier le cas pour $\kappa = \bar{\kappa} = 1$, $\alpha = \beta = 1$, $u_0 = 0$ et tentons de minimiser $\mathcal{J}$ donné par (1.18) sur les fonctions $u \in C^1(\bar{\Omega})$. Or on a vu que

— $C^1(\bar{\Omega})$ muni de la norme

$$\forall u \in C^1(\bar{\Omega}), \quad \|u\| := \sup_{x \in \bar{\Omega}} |u(x)| + \sup_{x \in \bar{\Omega}} \|\nabla u(x)\|,$$

est complet, mais cette norme n'est pas nécessairement contrôlée par la fonctionnelle

— L'espace $C^1(\bar{\Omega})$ muni du produit scalaire naturel de ce problème

$$(u, v) \rightarrow \int_{\Omega} [u v + \underline{\nabla} u \cdot \underline{\nabla} v] \, d\Omega, \quad (1.20)$$

n'est pas complet (voir notamment [4] et le chapitre suivant).

Une condition nécessaire d'optimalité est

$$\forall v \in C^1(\bar{\Omega}), \quad \int_{\Omega} [\underline{\nabla} u \cdot \underline{\nabla} v + u v] \, d\Omega = \int_{\Omega} f v \, d\Omega.$$

mais la solution u n'est pourtant pas nécessairement C^1 . L'objectif de ce cours sera de trouver l'espace de solution adapté. Il sera en fait le complété de l'espace C^1 pour la norme (1.20). Un espace dans lesquelles les fonctions ont une régularité suffisante pour continuer de donner un sens à (1.19).

1.5 Le calcul des variations d'Euler et Lagrange

Cette partie peut être sautée lors d'une première lecture. Elle apporte des compléments historiques et tente de mettre en perspective les développements du cours, tant sur les cas non linéaires que sur les impasses rencontrées lorsqu'on utilise des fonctions trop régulières.

1.5.1 Le défi de Bernoulli

Ce paragraphe est une mise en perspective historique inspirée notamment de [15] qui rappelle que nous sommes *des nains sur des épaules de géants*²⁵, ces illustres scientifiques qui ont développé le calcul des variations dès le XVII^e siècle en passant du calcul à l'abstraction mathématique jusqu'aux considérations philosophiques.

En 1696, Jean Bernoulli²⁶ lance un défi à son frère Jacques Bernoulli et au monde scientifique avec le problème suivant : *Deux points A et B étant donnés dans un plan vertical, déterminer la courbe AMB le long de laquelle un mobile M, abandonné en A, descend sous l'action de sa propre pesanteur et parvient à l'autre point B dans le moins de temps possible*. Il ajoute que si la droite est bien la courbe minimisant la distance, ce n'est pas celle qui minimise le temps de parcours. Cette question est le point de départ²⁷ du développement de la théorie du calcul des variations qui va conduire à réécrire une partie de la physique sous forme de problème de minimisation. Mais qui plus est, cette vision va aussi offrir un ensemble de stratégies permettant la résolution pratique des équations physiques résultantes.

Mais avant de répondre au défi lancé par Bernoulli, commençons tout d'abord par un problème plus simple qui est celui de la minimisation de la distance entre les points A et B dont nous connaissons déjà le résultat : la ligne droite. Nous paramétrons la courbe $s \mapsto (x(s), y(s))$ recherchée sous la forme $x \mapsto y = -u(x)$ avec par convention $A = (0, 0)$ et $B = (L, u(L) = -y_B)$, voir figure 1.10. On cherche donc à minimiser²⁸ la fonctionnelle

$$u = \operatorname{argmin}_{v \in \mathcal{V}} \left\{ \mathcal{J}(v) = \int_0^L \sqrt{1 + v'(x)^2} dx \right\}, \quad (1.21)$$

dans un ensemble $\mathcal{V}$ de fonctions suffisamment régulières pour que la fonctionnelle soit bien définie, et telles que $v(0) = 0$ et $v(L) = -y_B$. Puisque nous devons trouver la ligne droite pour résultat, le minimum est caractérisé par l'équation

$$-u''(x) = 0.$$

La question est donc de savoir pourquoi?

Pour résoudre le problème posé par Jean Bernoulli, il nous faut minimiser non pas la distance, mais le temps de parcours. La vitesse du mobile sur la courbe est $c = \sqrt{2g|y|}$ d'après la loi de Galilée avec g la constante de gravitation et y l'ordonnée de la position du mobile. En effet, en l'absence de frottement, l'énergie totale du mobile se conserve²⁹. On a donc

$$\frac{1}{2}mc^2 + mgy = 0.$$

Le temps de parcours T est donc

$$T = \int_0^S \frac{1}{c(s)} ds,$$

pour un mobile se déplaçant sur la courbe $s \mapsto (x(s), y(s))$. Si comme précédemment nous cherchons la courbe sous la forme $x \mapsto y = -u(x)$

25: L'expression (en latin : *Nanos gigantum umeris insidentes*) est attribuée à Bernard de Chartres, XII^e siècle, pour souligner l'importance de s'appuyer sur les travaux des grands penseurs du passé (les « géants »). Elle fût également utilisée par Blaise Pascal (1623-1662) et Isaac Newton (1643-1727).

26: Jean (Johann) Bernoulli (1667-1748), est un mathématicien et physicien suisse, frère de Jacques (Jacob) Bernoulli (celui de la loi de Bernoulli en probabilité) et le père de Daniel (principe de Bernoulli en mécanique des fluides), Nicolas et Jean.

27: 5 personnes vont présenter des solutions parmi lesquelles Jean lui-même et son frère Jacques, mais aussi Gottfried Wilhelm Leibniz (1646-1716), père du calcul infinitésimal, et une personne anonyme qui se révélera être Isaac Newton.

28: toujours en utilisant le fait qu'un élément de longueur infinitésimal est $\delta s = \sqrt{1 + u'(x)^2} \delta x$

29: la conservation de l'énergie est d'ailleurs aussi un principe variationnel

la fonctionnelle s'écrit cette fois

$$u = \operatorname{argmin}_{v \in \mathcal{V}} \left\{ \mathcal{J}(v) = \int_0^L \frac{\sqrt{1 + v'(x)^2}}{\sqrt{|v(x)|}} dx \right\}, \quad (1.22)$$

toujours dans un ensemble $\mathcal{V}$ de fonctions suffisamment régulières pour que la fonctionnelle soit bien définie et telles que $v(0) = 0$ et $v(L) = -y_B$.

1.5.2 Les équations d'Euler-Lagrange

Un demi-siècle après le défi lancé par Bernoulli, Euler propose de reformuler dans [13] les deux problèmes précédents dans un cadre plus général où on souhaite minimiser des fonctionnelles de la forme

$$\mathcal{J}(v) = \int_0^L \mathcal{L}(x, v(x), v'(x)) dx, \quad (1.23)$$

où $\mathcal{L} : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \ni (x, v, q) \mapsto \mathcal{L}(x, v, q)$ est une fonction régulière appelée fonction de Lagrange ou *lagrangien*. Dans le cas de la minimisation de la distance, nous avons

$$\mathcal{L}(x, v, q) = \sqrt{1 + q^2},$$

alors que dans le cas du temps minimal nous avons

$$\mathcal{L}(x, v, q) = \frac{\sqrt{1 + q^2}}{\sqrt{|v|}}.$$

On peut démontrer que ces fonctionnelles admettent des minimiseurs, ce que nous supposons pour l'instant dans le cadre de cette introduction. Par ailleurs, ces deux lagrangiens sont strictement convexes par rapport à leur variable q , ce qui assure que les minimiseurs locaux sont globaux.

Comme nous l'avons vu précédemment, la minimisation d'une telle fonctionnelle sur un espace $\mathcal{V}$ fait appel au calcul des variations historiquement imaginé par Joseph Louis de Lagrange³⁰. Supposons que $\mathcal{J}$ admette une courbe minimum u et considérons une variation par rapport à cette courbe de la forme $v : x \mapsto u(x) + \varepsilon w(x)$ pour w fixée telle que $w(0) = 0$ et $w(L) = 0$ de telle sorte que $v(0) = 0$ et $v(L) = u(L) = -y_B$.

Nous calculons alors

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathcal{J}(u + \varepsilon w) - \mathcal{J}(u)}{\varepsilon} = D\mathcal{J}(u)(w).$$

Supposons pour simplifier que $\mathcal{L}$ est une fonction C^1 . Et faisons le choix de minimiser la fonctionnelle sur l'espace

$$\mathcal{V} = \left\{ v \in C^1([0, L]) \mid v(0) = 0, v(L) = -y_B \right\}$$

muni de la norme

$$\forall u \in C^1([0, L]), \quad \|u\| = \sup_{x \in [0, L]} |u(x)| + \sup_{x \in [0, L]} |u'(x)|.$$

30: Joseph Louis de Lagrange, né Giuseppe Luigi Lagrangia (1736-1813) fut professeur à l'École polytechnique à partir de 1794

Nous introduisons l'espace

$$\mathcal{V}_0 = \left\{ v \in C^1([0, L]) \mid v(0) = 0, v(L) = 0 \right\},$$

auquel appartient les variations de la courbe recherchée. Par le théorème de composition des dérivées, nous avons

$$\begin{aligned} \langle D\mathcal{J}(u), w \rangle_{\mathcal{V}_0', \mathcal{V}_0} &= \int_0^L \left[\frac{\partial \mathcal{L}}{\partial v}(x, u(x), u'(x)) w(x) \right. \\ &\quad \left. + \frac{\partial \mathcal{L}}{\partial q}(x, u(x), u'(x)) w'(x) \right] dx. \end{aligned}$$

Une condition nécessaire d'optimalité est donc

$$\begin{aligned} \forall w \in \mathcal{V}_0, \int_0^L \left[\frac{\partial \mathcal{L}}{\partial v}(x, u(x), u'(x)) w(x) \right. \\ \left. + \frac{\partial \mathcal{L}}{\partial q}(x, u(x), u'(x)) w'(x) \right] dx = 0. \end{aligned} \quad (1.24)$$

C'est la *formulation variationnelle* du problème de minimisation.

Puis par intégration par parties³¹, en supposant la régularité supplémentaire³² $u \in \mathcal{V} \cap C^2(]0, L[)$,

$$\begin{aligned} \int_0^L \left[\frac{\partial \mathcal{L}}{\partial v}(x, u(x), u'(x)) - \frac{d}{dx} \left(\frac{\partial \mathcal{L}}{\partial q}(x, u(x), u'(x)) \right) \right] w(x) dx \\ + \left[w(x) \frac{\partial \mathcal{L}}{\partial q}(x, u(x), u'(x)) \right]_0^L = 0. \end{aligned} \quad (1.25)$$

En utilisant $w(0) = w(L) = 0$, on obtient³³

$$\begin{aligned} \forall w \in \mathcal{V}_0, \int_0^L \left[\frac{\partial \mathcal{L}}{\partial v}(x, u(x), u'(x)) \right. \\ \left. - \left(\frac{d}{dx} \frac{\partial \mathcal{L}}{\partial q}(x, u(x), u'(x)) \right) \right] w(x) dx = 0, \end{aligned} \quad (1.26)$$

Comme ceci doit être vrai pour tout $w \in \mathcal{V}_0$, on en déduit formellement l'équation dite *équation d'Euler-Lagrange*

$$\frac{\partial \mathcal{L}}{\partial v}(x, u(x), u'(x)) - \left(\frac{d}{dx} \frac{\partial \mathcal{L}}{\partial q}(x, u(x), u'(x)) \right) = 0, \quad x \in]0, L[, \quad (1.27)$$

qui fournit l'équation différentielle vérifiée par u et caractérise donc un minimiseur de $\mathcal{J}$. Cette équation est d'ordre 2 en x et nécessite deux conditions pour être résolue qui sont ici données aux deux extrémités 0 et L puisque $u(0) = 0$ et $u(L) = -y_B$. L'équation d'Euler-Lagrange couplée aux conditions aux limites définit la *formulation forte* du problème de minimisation.

Pour nos deux exemples précédents, l'équation d'Euler-Lagrange

31: En fait il est possible de démontrer a priori que $u \in C^2(]0, L[)$.

32: En fait il est possible de démontrer a priori que $u \in C^2(]0, L[)$.

33: À noter qu'en toute généralité il suffit d'avoir $w(0) = 0$ ou $\partial_q \mathcal{L}(0) = 0$ et de même en L pour simplifier les termes de bords

justifie donc pour la distance minimale, le problème aux limites

$$\begin{cases} u''(x) = 0, & x \in]0, L[, \\ u(0) = 0, \\ u(L) = -y_B. \end{cases}$$

Pour la minimisation du temps de parcours, la courbe minimisante appelée *brachistochrone*³⁴ vérifie le problème

$$\begin{cases} 2u(x)u''(x) + u'(x)^2 + 1 = 0, & x \in]0, L[, \\ u(0) = 0, \\ u(L) = -y_B. \end{cases}$$

La solution de ce problème aux limites n'est pas une droite, ni un cercle comme le pensait Galilée, mais une cycloïde.

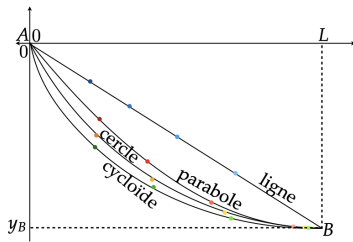


FIGURE 1.10 – Comparaison des vitesses du mobile pour différentes courbes.

Remarque 1.5.1 (Identité de Beltrami) La forme développée de l'équation d'Euler Lagrange peut être simplifiée dans le cas où $\mathcal{L}(v, p)$ ne dépend pas de x par une astuce déjà révélée par Euler. Si nous développons en effet

$$\begin{aligned} & \frac{d}{dx} \left(\mathcal{L}(u(x), u'(x)) - u'(x) \frac{\partial \mathcal{L}}{\partial q}(u(x), u'(x)) \right) \\ &= u'(x) \frac{\partial \mathcal{L}}{\partial v}(u(x), u'(x)) + u''(x) \frac{\partial \mathcal{L}}{\partial q}(u(x), u'(x)) \\ & \quad - u''(x) \frac{\partial \mathcal{L}}{\partial q}(u(x), u'(x)) - u'(x) \frac{d}{dx} \left(\frac{\partial \mathcal{L}}{\partial q}(u(x), u'(x)) \right), \end{aligned}$$

donc en utilisant l'équation d'Euler-Lagrange (1.27), on obtient

$$\begin{aligned} & \frac{d}{dx} \left(\mathcal{L}(u(x), u'(x)) - u'(x) \frac{\partial \mathcal{L}}{\partial q}(u(x), u'(x)) \right) \\ &= u'(x) \frac{\partial \mathcal{L}}{\partial v}(u(x), u'(x)) + u''(x) \frac{\partial \mathcal{L}}{\partial q}(u(x), u'(x)) \\ & \quad - u''(x) \frac{\partial \mathcal{L}}{\partial q}(u(x), u'(x)) - u'(x) \frac{\partial \mathcal{L}}{\partial v}(u(x), u'(x)) = 0. \end{aligned}$$

Donc il existe une constante c telle que

$$\mathcal{L}(u(x), u'(x)) - u'(x) \frac{\partial \mathcal{L}}{\partial q}(u(x), u'(x)) = c,$$

ce qui conduit à

$$\frac{\sqrt{1 + u'(x)^2}}{\sqrt{|u(x)|}} - \frac{u'(x)^2}{\sqrt{u(x)}\sqrt{1 + u'(x)^2}} = c,$$

autrement dit, il existe R tel que

$$u'(x) = \sqrt{\frac{R - u(x)}{u(x)}},$$

dont on démontre par le changement de variable $u' = \cot t$ que les solutions sont données par un bout de cycloïde d'équation

$$\begin{cases} x(t) = R^2(t - \sin(2t)), \\ y(t) = R^2(1 + \cos(2t)). \end{cases}$$

1.5.3 Des relation de Snell-Descartes à la nécessité d'affaiblir la régularité

Considérons ici un exemple similaire au problème de Bernoulli, pour déterminer le trajet de la lumière dans un milieu.³⁵ Le principe de moindre action consiste à minimiser le temps de parcours lors propagation de la lumière. Si nous voulons modéliser le fait que la vitesse de propagation dépendent de l'indice du milieu, on montre que

$$\mathcal{L}(x, v, q) = \frac{\sqrt{1 + q^2}}{c(x)}.$$

Supposons maintenant que l'indice du milieu $n(x) = c^{-1}(x)$ est une fonction de x donnée par

$$n(x) = n_-, \text{ si } 0 \leq x < L/2 \text{ et } n_+, \text{ si } L/2 < x \leq L,$$

ce qui implique que $\mathcal{L}$ n'est plus C^1 mais seulement C^1 par morceaux. Nous allons illustrer ici que, dans ce cas, la solution optimale n'est pas C^1 .

Toujours en fonctionnant par condition nécessaire, admettons que la formulation variationnelle (1.24) est toujours valide, alors sur chaque sous domaine $(0, L/2)$ et $(L/2, L)$, on montre directement que la trajectoire optimale vérifie $u'' = 0$ donc est une droite.

Cependant on ne peut pas le dire sur tout le segment $[0, L]$ car $\mathcal{L}$ n'est plus continue. Soit alors une fonction test $w \in \mathcal{V}_0$ non nulle en $x = L/2$, alors on a

$$\begin{aligned} & \int_0^{L/2} \left\{ \frac{\partial \mathcal{L}}{\partial v}(x, u(x), u'(x)) - \frac{d}{dx} \left[\frac{\partial \mathcal{L}}{\partial q}(x, u(x), u'(x)) \right] \right\} w(x) dx \\ & + \int_{L/2}^L \left\{ \frac{\partial \mathcal{L}}{\partial v}(x, u(x), u'(x)) - \frac{d}{dx} \left[\frac{\partial \mathcal{L}}{\partial q}(x, u(x), u'(x)) \right] \right\} w(x) dx \\ & + \left[\frac{\partial \mathcal{L}}{\partial q}(L/2^-, u(L/2^-), u'(L/2^-)) - \frac{\partial \mathcal{L}}{\partial q}(L/2^+, u(L/2^+), u'(L/2^+)) \right] w(0) = 0. \end{aligned} \quad (1.28)$$

Sur $[0, L/2]$ et $[L/2, L]$ l'équation d'Euler est vérifiée, donc on obtient

$$\frac{\partial \mathcal{L}}{\partial q}(L/2^-, u(L/2^-), u'(L/2^-)) = \frac{\partial \mathcal{L}}{\partial q}(L/2^+, u(L/2^+), u'(L/2^+)),$$

ce qui se traduit en

$$\frac{u'(L/2^-)}{c_- \sqrt{1 + u'(L/2^-)^2}} = \frac{u'(L/2^+)}{c_+ \sqrt{1 + u'(L/2^+)^2}}.$$

35: Lors de son parcours entre deux points A et B, la lumière suit un trajet qui rend la durée de celui-ci stationnaire par rapport à tout trajet infiniment voisin – Pierre de Fermat

Les angles formés par la courbe et l'axe des abscisses sont donnés par

$$\sin(\theta_-) = \frac{u'(L/2^-)}{\sqrt{1 + u'(L/2^-)^2}} \text{ et } \sin(\theta_+) = \frac{u'(L/2^+)}{\sqrt{1 + u'(L/2^+)^2}}.$$

On retrouve alors les fameuses lois de Snell-Descartes, voir Figure 1.11,

$$n_- \sin(\theta_-) = n_+ \sin(\theta_+).$$

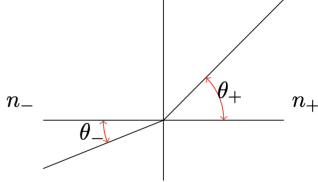


FIGURE 1.11 – Lois de Snell-Descartes

La fonction construite n'est pas C^1 si $n_- \neq n_+$ malgré la recherche d'un infimum sur l'ensemble $\mathcal{V} = C^1$. Elle vérifie pourtant

$$J(u) \leq \inf_{v \in \mathcal{V}} J(v).$$

Le lecteur intéressé pourra essayer de démontrer qu'elle réalise bien l'infimum, en considérant une limite de fonctions C^1 qui converge vers la solution C^1 par morceaux ainsi obtenue. C'est le manque de complétude de C^1 pour les normes contrôlées par la fonctionnelle qui conduit à un infimum dans un espace moins régulier. On comprend une fois encore que l'espace de minimisation était sans doute trop petit, puisque le minimum de la fonctionnelle est moins régulier. Et pourtant, la relation de Snell-Descartes a bien un sens physique correspondant à notre choix de modélisation avec deux milieux séparés par une interface.

Cet exemple illustre une nouvelle fois la nécessité d'introduire des espaces de solutions mieux adaptés aux problèmes variationnels, suffisamment grands pour que les infimas recherchés y appartiennent, mais offrant une régularité suffisante pour permettre de poursuivre les calculs de variations. C'est l'objectif du chapitre suivant.

Dans ce chapitre nous définissons les espaces de Sobolev. Ce sont les espaces « naturels » de fonctions permettant de résoudre les formulations variationnelles d'équations aux dérivées partielles. En effet, ils s'interprètent comme les plus grands espaces tels que les fonctionnelles d'énergie soient bien définies. On dit donc qu'ils sont des espaces de *fonctions d'énergie finie*. Nous verrons que ce sont des espaces de Hilbert et donc en particulier des espaces complets, ce qui nous manquait au chapitre précédent pour permettre l'analyse des problèmes variationnels. En ce sens, plus qu'une construction mathématique abstraite, ils ont un caractère essentiel, car ils traduisent ce que nous attendons en termes de régularité des solutions des problèmes variationnels. Si ces problèmes variationnels modélisent une réalité physique, alors la régularité des solutions modélisera la réalité physique de la réponse attendue.

Il est préférable pour comprendre tous les résultats de ce chapitre d'avoir suivi un cours d'intégration. Nous rappelons néanmoins tous les résultats importants pour le propos de ce cours et nous conseillons le lecteur de consulter [4, 5] pour plus de détails.

2.1 Fonctions de carré intégrable

2.1.1 L'espace $L^2(\Omega)$

Soit Ω un ouvert de $\mathbb{R}^d$ muni de la mesure de Lebesgue. On définit l'espace $L^2(\Omega)$ des fonctions mesurables de carré sommable dans Ω :

$$L^2(\Omega) := \left\{ f : \Omega \rightarrow \mathbb{R} \text{ mesurable, } \int_{\Omega} f^2 \, d\Omega < +\infty \right\}. \quad (2.1)$$

Rappelons que les fonctions mesurables dans Ω sont définies *presque partout* dans Ω . C'est en fait une classe d'équivalence puisque si on change les valeurs d'une fonction mesurable f sur un sous-ensemble de Ω de mesure nulle, on ne change pas la fonction mesurable f . Ainsi, deux fonctions mesurables f et g sont dites égales si $f(x) = g(x)$ presque partout dans Ω , c'est-à-dire s'il existe $E \subset \Omega$ tel que la mesure de Lebesgue de E est nulle et $f(x) = g(x)$ pour tout $x \in \Omega \setminus E$.

Un exemple intéressant concerne le cas où $\Omega \equiv]0, 1[\subset \mathbb{R}$. Il est facile de montrer que

$$x \mapsto x^{\beta} \in L^2(]0, 1[) \quad \text{ssi} \quad \beta > -\frac{1}{2}. \quad (2.2)$$

De plus lorsque Ω est borné, toute fonction continue sur $\overline{\Omega}$ appartient à $L^2(\Omega)$. Ces fonctions sont en effet continues sur $\overline{\Omega}$ qui est un fermé borné de $\mathbb{R}^d$, donc un compact. Elles sont donc bornées et

$$f \in C^0(\overline{\Omega}) \rightarrow \int_{\Omega} f^2 \, d\Omega \leq \sup_{x \in \overline{\Omega}} |f(x)|^2 \int_{\Omega} d\Omega < +\infty.$$

1: on vérifiera que cette forme bilinéaire est bien un produit scalaire, d'après la définition 3.2.1.

On introduit maintenant le produit scalaire¹

$$\forall f, g \in L^2(\Omega), \quad (f, g)_{L^2(\Omega)} = \int_{\Omega} f g \, d\Omega, \quad (2.3)$$

qui est bien défini, car si f et g sont dans $L^2(\Omega)$, $f g \in L^1(\Omega)$ d'après l'inégalité de Cauchy-Schwarz rappelée ci-dessous. On note

$$\|f\|_{L^2(\Omega)} = \left(\int_{\Omega} |f|^2 \, d\Omega \right)^{\frac{1}{2}}, \quad (2.4)$$

la norme associée à ce produit scalaire. On rappelle l'identité du parallélogramme

$$\|f + g\|_{L^2(\Omega)}^2 + \|f - g\|_{L^2(\Omega)}^2 = 2 \left[\|f\|_{L^2(\Omega)}^2 + \|g\|_{L^2(\Omega)}^2 \right], \quad (2.5)$$

l'inégalité de Cauchy-Schwarz,

$$|(f, g)_{L^2(\Omega)}| \leq \|f\|_{L^2(\Omega)} \|g\|_{L^2(\Omega)}, \quad (2.6)$$

et enfin l'inégalité de Minkowski

$$\|f + g\|_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)} + \|g\|_{L^2(\Omega)},$$

pour toutes fonctions $f, g \in L^2(\Omega)$. Cette dernière inégalité permet de déduire que l'espace $L^2(\Omega)$ est un espace vectoriel. On a de plus le résultat suivant très important.

Théorème 2.1.1 *Muni du produit scalaire défini en (2.3), $L^2(\Omega)$ est un espace de Hilbert.*

2: Si une suite de Cauchy a une sous suite qui converge alors toute la suite converge. La preuve se fait par l'absurde : on ne peut pas construire de sous suite qui reste à une certaine distance de la limite sans contredire le fait que la suite est de Cauchy.

Démonstration. Concentrons-nous sur la complétude. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de Cauchy de l'espace $L^2(\Omega)$. Il suffit² de montrer qu'une sous-suite converge. Pour simplifier les notations, on note cette sous-suite encore $(f_n)_n$ et on la choisit telle que

$$\|f_n - f_{n-1}\|_{L^2(\Omega)} \leq 2^{1-n}.$$

Introduisons maintenant pour presque tout $x \in \Omega$

$$g_n(x) = \sum_{k=1}^n |f_k(x) - f_{k-1}(x)|.$$

Pour tout $x \in \Omega$, $(g_n(x))_{n \in \mathbb{N}}$ est une suite croissante de $\mathbb{R}$ et pour tout $n \in \mathbb{N}$

$$\|g_n\|_{L^2(\Omega)} \leq \sum_{k=1}^n \|f_k - f_{k-1}\|_{L^2(\Omega)} \leq \sum_{k=1}^n 2^{1-k} \leq 1.$$

3: Théorème de convergence monotone de Beppo/Levi : pour toute suite croissante (u_n) de fonctions mesurables positives, sa limite simple est mesurable et

$$\int \lim u_n = \lim \int u_n.$$

D'après le théorème de convergence monotone³ appliqué à $\|g_n\|^2$, on déduit que sa limite g est dans $L^2(\Omega)$ puisque d'après l'inégalité ci-dessus $\|g\|_{L^2(\Omega)} \leq 1$. De plus, g_n converge vers g pour la norme L^2 et donc pour presque tout $x \in \Omega$, $g_n(x)$ converge vers $g(x)$.

Comme pour presque tout x , on a pour $q > p$,

$$|f_q(x) - f_p(x)| \leq \sum_{p+1}^q |f_k(x) - f_{k-1}(x)| = g_q(x) - g_p(x),$$

la suite $(f_n(x))_{n \in \mathbb{N}}$ est une suite de Cauchy dans $\mathbb{R}$ donc elle converge vers un réel que nous notons $f(x)$. De plus, en passant à la limite sur q dans l'inégalité ci-dessus, on trouve

$$|f(x) - f_p(x)| \leq 2g_p(x),$$

ce qui nous permet de déduire que f est bien une fonction de $L^2(\Omega)$. En appliquant le théorème de convergence dominée⁴ appliquée à la fonction $|f - f_n|^2$, on conclut que f_n converge vers f dans $L^2(\Omega)$. $\square$

On note $C_c^\infty(\Omega)$, ou encore $D(\Omega)$, l'espace des fonctions de classe C^∞ à support compact dans Ω ⁵. Notons aussi que les fonctions de $C_c^\infty(\Omega)$ s'annulent, ainsi que toutes leurs dérivées, sur le bord de Ω .

Nous rappelons le résultat de densité suivant, dont la preuve peut être trouvée dans [5, théorème 3.4.3].

Théorème 2.1.2 *L'espace $D(\Omega)$ est dense⁶ dans $L^2(\Omega)$, c'est-à-dire que pour tout $f \in L^2(\Omega)$ il existe une suite $f_n \in D(\Omega)$ telle que*

$$\lim_{n \rightarrow +\infty} \|f - f_n\|_{L^2(\Omega)} = 0.$$

Rappelons le résultat suivant. Supposons que $f \in L^2(\Omega)$ vérifie

$$\begin{aligned} \forall g \in L^2(\Omega), \quad (f, g)_{L^2(\Omega)} = 0 &\Leftrightarrow f = 0 \text{ dans } L^2(\Omega) \\ &\Leftrightarrow f(x) = 0 \text{ p.p. } x \in \Omega. \end{aligned} \quad (2.7)$$

Le théorème de densité permet d'étendre ce résultat en ne considérant que des fonctions g dans le sous-espace $D(\Omega)$ de $L^2(\Omega)$.

Corollaire 2.1.3 *Soit $f \in L^2(\Omega)$. Si pour toute fonction $\phi \in D(\Omega)$ on a*

$$\int_{\Omega} f \phi \, d\Omega = 0,$$

alors $f(x) = 0$ pour presque tout $x \in \Omega$.

4: Théorème de convergence dominée : soit $(f_n)_n$ une suite de fonctions mesurables telles que f_n converge simplement vers f et telle que pour tout n et pour tout x $f_n(x) \leq g(x)$ avec g intégrable alors f est intégrable et

$$\int f = \lim \int f_n.$$

5: L'espace $C_c^\infty(\Omega)$ n'est pas réduit à la seule fonction nulle partout, voir le corollaire 3.2.6 dans [5].

6: *Stricto-sensu*, il faut toujours donner la norme pour laquelle on considère la densité donc ici on veut dire $L^2(\Omega)$ muni de la norme $\|\cdot\|_{L^2(\Omega)}$

2.1.2 Théorème de prolongement par densité

Dans la suite, le théorème de prolongement par densité jouera un rôle fondamental. En effet, si on sait que l'ensemble des fonctions régulières est dense dans un espace de Sobolev, il suffira de démontrer les propriétés pour les fonctions régulières et d'utiliser le théorème de prolongement par densité pour les étendre aux fonctions de l'espace de Sobolev. Il faudra vérifier que ces propriétés sont bien continues par rapport à la norme de l'espace.

Nous énonçons le théorème de prolongement dans sa version la plus abstraite, car il sera utile dans la suite dans différents contextes.

Théorème 2.1.4 Soient $\mathcal{H}$, $\mathcal{W}$ deux espaces vectoriels normés, avec $\mathcal{W}$ complet, $\mathcal{V}$ un sous-espace vectoriel de $\mathcal{H}$, tel que $\mathcal{V}$ est dense dans $\mathcal{H}$. Soit ℓ une application linéaire de $\mathcal{V}$ dans $\mathcal{W}$, pour laquelle il existe une constante $C_\ell > 0$ telle que,

$$\forall v \in \mathcal{V}, \quad \|\ell(v)\|_{\mathcal{W}} \leq C_\ell \|v\|_{\mathcal{H}}.$$

Alors ℓ se prolonge de manière unique en une application linéaire et continue de $\mathcal{H}$ dans $\mathcal{W}$ et

$$\forall v \in \mathcal{H}, \quad \|\ell(v)\|_{\mathcal{W}} \leq C_\ell \|v\|_{\mathcal{H}}.$$

Démonstration. On raisonne en plusieurs étapes :

1. Définition du prolongement par continuité $\tilde{\ell}$.

Soit $h \in \mathcal{H}$. Par densité, il existe une suite d'éléments de $\mathcal{V}$, notée $(v_k)_{k \in \mathbb{N}}$, telle que $\lim_{k \rightarrow \infty} \|v_k - h\|_{\mathcal{H}} = 0$. Soit $(z_k)_k$ la suite d'éléments de $\mathcal{W}$ définie par $z_k = \ell(v_k)$ pour tout k . Étudions $(z_k)_k$:

$$\|z_k - z_m\|_{\mathcal{W}} = \|\ell(v_k) - \ell(v_m)\|_{\mathcal{W}} \stackrel{\ell \text{ lin.}}{=} \|\ell(v_k - v_m)\|_{\mathcal{W}} \leq C_\ell \|v_k - v_m\|_{\mathcal{H}}.$$

Comme $(v_k)_{k \in \mathbb{N}}$ converge (vers h) dans $\mathcal{H}$, c'est en particulier une suite de Cauchy, c'est-à-dire que $\lim_{k, m \rightarrow \infty} \|v_k - v_m\|_{\mathcal{H}} = 0$. Ainsi, $(z_k)_{k \in \mathbb{N}}$ est une suite de Cauchy dans $\mathcal{W}$. Or, $\mathcal{W}$ est complet : $(z_k)_{k \in \mathbb{N}}$ est donc convergente vers $z \in \mathcal{W}$. À partir de là, on définit le prolongement par continuité de ℓ en h comme étant égal à la limite de $(z_k)_{k \in \mathbb{N}}$:

$$\tilde{\ell}(h) = z.$$

Bien sûr, on peut effectuer le même raisonnement pour tout élément h de $\mathcal{H}$, ce qui permet de construire $\tilde{\ell} : \mathcal{H} \rightarrow \mathcal{W}$.

2. Unicité du prolongement $\tilde{\ell}$.

Reprenons le processus précédent de construction... Pour $h \in \mathcal{H}$, soient deux suites d'éléments de $\mathcal{V}$, $(v_k)_{k \in \mathbb{N}}$ et $(v'_k)_{k \in \mathbb{N}}$, convergeant vers h . Vérifions maintenant que la définition de $\tilde{\ell}(h)$ est indépendante de la suite choisie, ce qui prouvera l'unicité :

$$\begin{aligned} \left\| \lim_k \ell(v_k) - \lim_k \ell(v'_k) \right\|_{\mathcal{W}} &= \left\| \lim_k \{ \ell(v_k) - \ell(v'_k) \} \right\|_{\mathcal{W}} \\ &\stackrel{\ell \text{ lin.}}{=} \left\| \lim_k \ell(v_k - v'_k) \right\|_{\mathcal{W}}. \end{aligned}$$

Or,

$$\|\ell(v_k - v'_k)\|_{\mathcal{W}} \leq C_\ell \|v_k - v'_k\|_{\mathcal{H}},$$

et $\lim_k \|v_k - v'_k\|_{\mathcal{H}} = 0$ par inégalité triangulaire

$$\|v_k - v'_k\|_{\mathcal{H}} \leq \|v_k - h\|_{\mathcal{H}} + \|h - v'_k\|_{\mathcal{H}}.$$

On en conclut que $\lim_k \ell(v_k) = \lim_k \ell(v'_k)$.

3. Linéarité du prolongement $\tilde{\ell}$.

On obtient la linéarité en passant à la limite. Soient $h^1, h^2 \in \mathcal{H}$,

$\alpha^1, \alpha^2 \in \mathbb{R}$: par densité, il existe $(v_k^1)_{k \in \mathbb{N}}$ et $(v_k^2)_{k \in \mathbb{N}}$ deux suites d'éléments de $\mathcal{V}$, qui convergent respectivement vers h^1 et h^2 dans $\mathcal{H}$, et

$$\begin{aligned} \tilde{\ell}(\alpha^1 h^1 + \alpha^2 h^2) &= \lim_k \ell(\alpha^1 v_k^1 + \alpha^2 v_k^2) \\ &\stackrel{\ell \text{ lin.}}{=} \lim_k \{\alpha^1 \ell(v_k^1) + \alpha^2 \ell(v_k^2)\} = \alpha^1 \tilde{\ell}(h^1) + \alpha^2 \tilde{\ell}(h^2). \end{aligned}$$

4. Continuité du prolongement $\tilde{\ell}$ de $\mathcal{H}$ dans $\mathcal{W}$.

On veut prouver :

$$\exists C > 0, \forall h \in \mathcal{H}, \|\tilde{\ell}(h)\|_{\mathcal{W}} \leq C \|h\|_{\mathcal{H}}.$$

Pour cela, $h \in \mathcal{H}$ étant donné, soit $(v_k)_{k \in \mathbb{N}}$ une suite d'éléments de $\mathcal{V}$ qui converge vers h dans $\mathcal{H}$. On écrit

$$\|\tilde{\ell}(h)\|_{\mathcal{W}} = \|\lim_k \ell(v_k)\|_{\mathcal{W}} = \lim_k \|\ell(v_k)\|_{\mathcal{W}};$$

or, pour tout k , $\|\ell(v_k)\|_{\mathcal{W}} \leq C_\ell \|v_k\|_{\mathcal{H}}$. Comme $\lim_k \|v_k\|_{\mathcal{H}} = \|h\|_{\mathcal{H}}$, on en conclut que

$$\|\tilde{\ell}(h)\|_{\mathcal{W}} \leq C_\ell \|h\|_{\mathcal{H}}.$$

Notons que le module de continuité est identique pour ℓ et pour son prolongement $\tilde{\ell}$.

□

De ce résultat, on déduit la propriété de prolongement par densité pour les formes bilinéaires continues.

Corollaire 2.1.5 Soient $\mathcal{H}_1, \mathcal{H}_2, \mathcal{W}$ des espaces vectoriels normés, avec $\mathcal{W}$ complet, $\mathcal{V}_1$ (respectivement $\mathcal{V}_2$) un sous-espace vectoriel de $\mathcal{H}_1$ (respectivement $\mathcal{H}_2$), tel que $\mathcal{V}_1$ (respectivement $\mathcal{V}_2$) est dense dans $\mathcal{H}_1$ (respectivement $\mathcal{H}_2$). Soit a une application bilinéaire de $\mathcal{V}_1 \times \mathcal{V}_2$ dans $\mathcal{W}$, pour laquelle il existe une constante $C_a > 0$ telle que,

$$\forall (v_1, v_2) \in \mathcal{V}_1 \times \mathcal{V}_2, \quad \|a(v_1, v_2)\|_{\mathcal{W}} \leq C_a \|v_1\|_{\mathcal{H}_1} \|v_2\|_{\mathcal{H}_2}.$$

Alors a se prolonge de manière unique en une application bilinéaire et continue de $\mathcal{H}_1 \times \mathcal{H}_2$ dans $\mathcal{W}$ et

$$\forall (v_1, v_2) \in \mathcal{H}_1 \times \mathcal{H}_2, \quad \|a(v_1, v_2)\|_{\mathcal{W}} \leq C_a \|v_1\|_{\mathcal{H}_1} \|v_2\|_{\mathcal{H}_2}.$$

2.1.3 Dérivation faible des fonctions de $L^2(\Omega)$

Ce concept de dérivée faible des fonctions de $L^2(\Omega)$ généralise la notion de dérivée au sens classique pour les fonctions $C^1(\Omega)$ dans le sens où la dérivée faible d'une fonction $C^1(\Omega)$ coïncide avec sa dérivée au sens classique. Pour ceux qui connaissent la dérivée au sens des distributions, la dérivée faible n'est rien d'autre que la dérivée au sens des distributions puisque les fonctions de $L^2(\Omega)$ sont des distributions. Sur ce dernier aspect, nous renvoyons le lecteur au paragraphe 2.3, où le lien entre les 2 notions de dérivée est établi.

7: Ceci lui vaudra la médaille Fields en 1950.

Cependant, d'un point de vue historique, la notion de dérivée faible est antérieure à celle des distributions. En effet, elle date de 1934 et apparait dans un article de J. Leray sur la construction de solutions "faibles" aux équations de Navier Stokes [20]. Ce concept de dérivée faible a été ensuite généralisé dans la théorie des distributions de Laurent Schwartz [24] en 1937⁷.

Définition 2.1.1 Soit $u \in L^2(\Omega)$, on dit que u admet une dérivée faible dans $L^2(\Omega)$ dans la direction $\underline{e}_i$ si et seulement s'il existe $u_i \in L^2(\Omega)$ telle que

$$\forall \phi \in D(\Omega), \quad \left(u, \frac{\partial \phi}{\partial x_i} \right)_{L^2(\Omega)} = -(u_i, \phi)_{L^2(\Omega)}.$$

Dans ce cas, u_i est appelée la dérivée faible de u dans la direction $\underline{e}_i$ et est notée $\frac{\partial u}{\partial x_i}$ ou $\partial_{x_i} u$.

Quand u admet une dérivée faible dans toutes les directions, elle admet un gradient faible.

La dérivée faible est ainsi définie de manière unique d'après le corollaire 2.1.3.

La dérivée faible est une extension de la dérivation au sens classique puisque d'après la formule de Green (1.8)

$$\forall u \in C^1(\overline{\Omega}), \forall v \in D(\Omega), \quad \int_{\Omega} \frac{\partial u}{\partial x_i} v \, d\Omega = - \int_{\Omega} \frac{\partial v}{\partial x_i} u \, d\Omega.$$

Ainsi, quand il est possible de définir la dérivée au sens classique, celle-ci correspond également à la dérivée faible.

De plus, c'est un cas particulier de la dérivée au sens des distributions d'après la définition 2.3.2.

Donnons maintenant un critère très simple qui permet de montrer qu'une fonction de $L^2(\Omega)$ a bien une dérivée faible dans $L^2(\Omega)$.

Proposition 2.1.6 Soit u une fonction de $L^2(\Omega)$. Si pour $i \in \llbracket 1, d \rrbracket$, il existe une constante $C > 0$ telle que

$$\forall \phi \in D(\Omega), \quad \left| \int_{\Omega} u \frac{\partial \phi}{\partial x_i} \, d\Omega \right| \leq C \|\phi\|_{L^2(\Omega)},$$

alors u est dérivable au sens faible dans la direction $\underline{e}_i$.

Démonstration. Soit ℓ la forme linéaire définie par

$$\forall \phi \in D(\Omega), \quad \ell(\phi) = \int_{\Omega} u \frac{\partial \phi}{\partial x_i} \, d\Omega.$$

Elle est continue pour la norme L^2 par hypothèse. D'après le théorème 2.1.4 de prolongement par densité, puisque $D(\Omega)$ est dense dans $L^2(\Omega)$ (voir théorème 2.1.2), ℓ s'étend en une forme linéaire continue de $L^2(\Omega)$. D'après le théorème de représentation de Riesz, il existe $u_i \in L^2(\Omega)$ telle que

$$\forall v \in L^2(\Omega), \quad \ell(v) = -(u_i, v)_{L^2(\Omega)}.$$

et en particulier

$$\forall \phi \in D(\Omega), \quad \int_{\Omega} u \frac{\partial \phi}{\partial x_i} d\Omega = -(u_i, \phi)_{L^2(\Omega)},$$

ce qui prouve que u est dérivable au sens faible dans $L^2(\Omega)$. $\square$

On retrouve un résultat bien connu pour la dérivée classique.

Proposition 2.1.7 Soit $u \in L^2(\Omega)$ telle que toutes ses dérivées au sens faible sont nulles alors u est constante sur chaque composante connexe de Ω .

Démonstration. Nous détaillons la preuve dans le cas de la dimension 1. On remplace Ω par l'intervalle $I =]a, b[$ et on suppose que $u \in L^2(\Omega)$ est telle que sa dérivée au sens faible est $u' = 0$. Si $\varphi \in D(I)$ est telle que

$$\int_I \varphi(x) dx = 0,$$

la fonction ψ définie par

$$\psi(x) = \int_a^x \varphi(t) dt$$

est bien dans $D(I)$ et $\psi' = \varphi$. Ainsi

$$(u, \varphi)_{L^2(\Omega)} = (u, \psi')_{L^2(\Omega)} = -(u', \psi)_{L^2(\Omega)} = 0.$$

Considérons maintenant une fonction $\theta \in D(I)$ telle que $\int_I \theta = 1$. Pour tout $\phi \in D(I)$, on a d'après ce qui précède

$$(u, \phi - \theta \int_I \phi)_{L^2(\Omega)} = 0 \quad \Rightarrow \quad (u, \phi)_{L^2(\Omega)} = (u, \theta)_{L^2(\Omega)} \int_I \phi,$$

ce qui signifie que u est la fonction constante égale à $(u, \theta)_{L^2(\Omega)}$. $\square$

Enfin, il est possible d'étendre cette notion de dérivée faible à des opérateurs différentiels qui ne font apparaître que certaines combinaisons linéaires des dérivées. Nous définissons ainsi la divergence au sens faible d'un champ de vecteurs⁸.

Définition 2.1.2 Soit $\underline{U} \in [L^2(\Omega)]^d$, on dit que $\underline{U}$ admet une divergence faible dans $L^2(\Omega)$ si et seulement s'il existe $w \in L^2(\Omega)$ telle que

$$\forall \phi \in D(\Omega), \quad \int_{\Omega} \underline{U} \cdot \nabla \phi d\Omega = - \int_{\Omega} w \phi d\Omega.$$

Dans ce cas w est appelée la divergence faible de $\underline{U}$.

Enfin, il est possible d'obtenir un critère simple pour déterminer si un champ de vecteurs admet une divergence au sens faible. La preuve de ce résultat est très similaire à celle de la proposition 2.1.6.

8: Un champ de vecteurs dont chaque composante admet des dérivées au sens faible dans toutes les directions admet une divergence au sens faible de manière évidente. Mais il existe des champs de vecteurs qui admettent une divergence au sens faible sans que toutes ses composantes soient dérivables au sens faible.

Proposition 2.1.8 Soit $\underline{U} \in [L^2(\Omega)]^d$. S'il existe $C > 0$ telle que

$$\forall \phi \in D(\Omega), \quad \left| \int_{\Omega} \underline{U} \cdot \nabla \phi \, d\Omega \right| \leq C \|\phi\|_{L^2(\Omega)},$$

alors $\underline{U}$ admet une divergence au sens faible.

2.2 Les espaces de Sobolev

2.2.1 L'espace $H^1(\Omega)$ et l'espace $H_0^1(\Omega)$

Définition 2.2.1 L'espace $H^1(\Omega)$ est l'ensemble des fonctions de $L^2(\Omega)$ qui admettent une dérivée faible dans toutes les directions :

$$H^1(\Omega) := \left\{ v \in L^2(\Omega), \frac{\partial v}{\partial x_i} \in L^2(\Omega) \quad \forall i \in \llbracket 1, d \rrbracket \right\}.$$

Un exemple intéressant concerne le cas où $\Omega \equiv]0, 1[\subset \mathbb{R}$. Nous renvoyons à (2.2) pour comparaison. On peut montrer que

$$x \mapsto x^\beta \in H^1(]0, 1[) \quad \text{ssi} \quad \beta > \frac{1}{2}. \quad (2.8)$$

Par ailleurs pour Ω un ouvert borné, $C^1(\overline{\Omega})$ est inclus dans $H^1(\Omega)$. En effet, si $f \in C^0(\overline{\Omega})$ alors $f \in L^2(\Omega)$. Et comme f est $C^1(\overline{\Omega})$ alors ses dérivées au sens classiques sont ses dérivées faibles. Comme elles sont dans $C^0(\overline{\Omega})$, elles sont dans $L^2(\Omega)$.

Un résultat plus intéressant concerne les fonctions qui sont seulement C^1 par morceaux. Nous donnons la condition nécessaire et suffisante pour que ces fonctions soient dans $H^1(\Omega)$.

Théorème 2.2.1 Soit Ω un ouvert borné de $\mathbb{R}^d$: on le partitionne en $\overline{\Omega} = \overline{\Omega_1} \cup \overline{\Omega_2}$, où Ω_1 et Ω_2 sont deux ouverts disjoints. On note Γ l'interface entre Ω_1 et Ω_2 : $\Gamma := \overline{\Omega_1} \cap \overline{\Omega_2}$.

Soit v telle que $v_i := v|_{\Omega_i} \in C^1(\overline{\Omega_i})$, pour $i = 1, 2$, alors

$$v \in C^0(\Omega) \quad \Leftrightarrow \quad v \in H^1(\Omega),$$

et dans ce cas

$$\nabla v = \chi_{\Omega_1} \nabla v_1 + \chi_{\Omega_2} \nabla v_2,$$

où χ_{Ω_i} est la fonction indicatrice de Ω_i .

Démonstration. Comme pour $i = 1, 2$, $v_i \in C^0(\overline{\Omega_i})$, on a $v_i \in L^2(\Omega_i)$ ce qui implique que $v \in L^2(\Omega)$.

On va chercher maintenant à vérifier le critère de la proposition 2.1.6. Pour toute fonction $\varphi \in D(\Omega)$, et pour tout $i \in \llbracket 1, d \rrbracket$, on a

$$\int_{\Omega} v \frac{\partial \varphi}{\partial x_i} \, d\Omega = \int_{\Omega_1} v_1 \frac{\partial \varphi}{\partial x_i} \, d\Omega + \int_{\Omega_2} v_2 \frac{\partial \varphi}{\partial x_i} \, d\Omega.$$

En utilisant la formule de Green (1.8) et en isolant les intégrales sur le bord, on obtient

$$\begin{aligned} \int_{\Omega} v \frac{\partial \varphi}{\partial x_i} d\Omega = & - \int_{\Omega_1} \frac{\partial v_1}{\partial x_i} \varphi d\Omega - \int_{\Omega_2} \frac{\partial v_2}{\partial x_i} \varphi d\Omega \\ & + \int_{\partial\Omega_1} v_1 \varphi \underline{n}_1 \cdot \underline{e}_i d\Gamma + \int_{\partial\Omega_2} v_2 \varphi \underline{n}_2 \cdot \underline{e}_i d\Gamma, \end{aligned} \quad (2.9)$$

où $\underline{n}_1$ (respectivement $\underline{n}_2$) est la normale unitaire extérieure à $\partial\Omega_1$ (respectivement à $\partial\Omega_2$). Comme φ est à support compact, elle s'annule sur $\partial\Omega$. Or $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$, donc on obtient

$$\int_{\partial\Omega_1} v_1 \varphi \underline{n}_1 d\Gamma + \int_{\partial\Omega_2} v_2 \varphi \underline{n}_2 d\Gamma = \int_{\Gamma} (v_1 - v_2) \varphi \underline{n}_1 d\Gamma,$$

car $\underline{n}_1 = -\underline{n}_2$ en tout point de Γ et φ est continue à la traversée de Γ . On a donc

$$\begin{aligned} \int_{\Omega} v \frac{\partial \varphi}{\partial x_i} d\Omega = & - \int_{\Omega_1} \frac{\partial v_1}{\partial x_i} \varphi d\Omega - \int_{\Omega_2} \frac{\partial v_2}{\partial x_i} \varphi d\Omega \\ & + \int_{\Gamma} (v_1 - v_2) \varphi \underline{n}_1 \cdot \underline{e}_i d\Gamma. \end{aligned}$$

On observe que les deux premières intégrales du terme à droite de l'égalité vérifient le critère de la proposition 2.1.6

$$\forall j \in \{1, 2\}, \quad \left| \int_{\Omega_j} \frac{\partial v_j}{\partial x_i} \varphi d\Omega \right| \leq \|\nabla v_j\|_{L^2(\Omega_j)} \|\varphi\|_{L^2(\Omega)}.$$

Montrons maintenant que si v n'est pas continue, c'est à dire si

$$v_1|_{\Gamma} \neq v_2|_{\Gamma},$$

on ne peut pas trouver $f \in L^2(\Omega)$ telle que

$$\int_{\Gamma} (v_1 - v_2) \varphi \underline{n}_1 \cdot \underline{e}_i d\Gamma = \int_{\Omega} f \varphi d\Omega.$$

Nous allons le prouver par l'absurde. Supposons pour simplifier que $\Gamma := \{x = 0\}$, $v_1 - v_2 = 1$ sur Γ et qu'il existe une fonction $f \in L^2(\Omega)$ telle que

$$\forall \varphi \in D(\Omega), \quad \int_{\Gamma} \varphi d\Gamma = \int_{\Omega} f \varphi d\Omega, \quad (2.10)$$

Soit $\varphi \in D(\Omega)$ à support dans $\square(A, r) := \{\underline{x}, (\underline{x} - \underline{x}_A) \cdot \underline{e}_i \leq r\}$ où A est un point de Γ et r est choisi assez petit pour que $\square(A, r)$ soit inclus dans Ω . On choisit φ telle que $0 \leq \varphi \leq 1$, $\varphi = 1$ sur $\square(A, r/2)$. On pose $\varphi_n(\underline{x}) = \varphi(n\underline{x}, \underline{x}_{d-1})$ où $\underline{x} = (x, \underline{x}_{d-1})$. Alors φ_n est à support compact dans $\square(A, r/n, r) := \{\underline{x}, (\underline{x} - \underline{x}_A) \cdot \underline{e}_1 \leq r/n, (\underline{x} - \underline{x}_A) \cdot \underline{e}_i \leq r \text{ pour } 2 \leq i \leq d\}$, $0 \leq \varphi_n \leq 1$, $\varphi_n = 1$ dans $\square(A, r/2n, r/2)$. On a

$$\int_{\Gamma} \varphi_n d\Gamma \geq (r/2)^{d-1} > 0.$$

Or

$$\left| \int_{\Omega} f \varphi_n \, d\Omega \right| \leq \left[\int_{\Omega} |f|^2 \, d\Omega \right]^{\frac{1}{2}} \left[\int_{\square(A,r/n,r)} d\Omega \right]^{\frac{1}{2}} = \frac{r^{d/2}}{\sqrt{n}} \xrightarrow{n \rightarrow +\infty} 0.$$

On en déduit par contradiction que (2.10) ne peut pas être vérifiée.

Enfin si v est continue alors

$$\int_{\Omega} v \frac{\partial \varphi}{\partial x_i} \, d\Omega = - \int_{\Omega} \left[\frac{\partial v_1}{\partial x_i} \chi_{\Omega_1} + \frac{\partial v_2}{\partial x_i} \chi_{\Omega_2} \right] \varphi \, d\Omega.$$

On retrouve bien le résultat annoncé. $\square$

En utilisant exactement la même preuve et le fait qu'on étendra un peu plus tard la formule de Green (1.8) pour des fonctions qui sont seulement H^1 , il est possible de généraliser le résultat précédent à des fonctions qui sont H^1 dans chaque sous-domaine. On utilise également dans ce théorème la notion de trace $v|_{\Gamma}$ pour une fonction de H^1 , notion qui sera introduite au paragraphe 2.2.3.

Theorème 2.2.2 Soit Ω un ouvert borné de $\mathbb{R}^d$ partitionné en deux sous-domaines disjoint Ω_1 et Ω_2 tels que $\overline{\Omega} = \overline{\Omega_1} \cup \overline{\Omega_2}$. On note Γ l'interface entre Ω_1 et Ω_2 : $\Gamma := \overline{\Omega_1} \cap \overline{\Omega_2}$.

Soit v telle que $v_i := v|_{\Omega_i} \in H^1(\Omega_i)$, pour $i = 1, 2$ alors

$$v_1|_{\Gamma} = v_2|_{\Gamma} \Leftrightarrow v \in H^1(\Omega),$$

et dans ce cas

$$\nabla v = \chi_{\Omega_1} \nabla v_1 + \chi_{\Omega_2} \nabla v_2$$

où χ_{Ω_i} est la fonction caractéristique de Ω_i .

9: On vérifiera que cette forme bilinéaire est bien un produit scalaire, d'après la définition 3.2.1.

Introduisons maintenant le produit scalaire⁹

$$\forall f, g \in H^1(\Omega), \quad (f, g)_{H^1(\Omega)} = \int_{\Omega} [f g + \underline{\nabla} f \cdot \underline{\nabla} g] \, d\Omega, \quad (2.11)$$

qui est bien défini, car si f et g sont dans $H^1(\Omega)$, $f, g \in L^2(\Omega)$ et $\underline{\nabla} f, \underline{\nabla} g \in [L^2(\Omega)]^d$; et la norme induite par ce produit scalaire

$$\|f\|_{H^1(\Omega)} = \left(\int_{\Omega} [|f|^2 + |\underline{\nabla} f|^2] \, d\Omega \right)^{\frac{1}{2}}. \quad (2.12)$$

Nous avons de manière évidente

$$\forall f \in H^1(\Omega), \quad \|f\|_{L^2(\Omega)} \leq \|f\|_{H^1(\Omega)} \quad \text{et} \quad \|\underline{\nabla} f\|_{[L^2(\Omega)]^d} \leq \|f\|_{H^1(\Omega)}. \quad (2.13)$$

On peut vérifier que l'espace $H^1(\Omega)$ est un espace vectoriel. Le résultat suivant est plus précis et très important.

Theorème 2.2.3 Muni du produit scalaire défini en (2.11), $H^1(\Omega)$ est un espace de Hilbert.

Démonstration. Montrons simplement que $H^1(\Omega)$ est complet pour la norme induite par le produit scalaire. Soit $(f_n)_{n \in \mathbb{N}}$ une suite de Cauchy de $H^1(\Omega)$. D'après (2.13), c'est une suite de Cauchy de $L^2(\Omega)$ donc $(f_n)_{n \in \mathbb{N}}$ converge vers f dans $L^2(\Omega)$ d'après le théorème 2.1.1. De plus, d'après (2.13), $(\nabla f_n)_{n \in \mathbb{N}}$ est une suite de Cauchy de $[L^2(\Omega)]^d$ donc $(\nabla f_n)_{n \in \mathbb{N}}$ converge vers F dans $[L^2(\Omega)]^d$ d'après le théorème 2.1.1. Il ne reste plus qu'à montrer que f admet un gradient faible et $\nabla f = F$. En effet, pour tout $\varphi \in [D(\Omega)]^d$, on a

$$\begin{aligned} (f, \nabla \cdot \varphi)_{L^2(\Omega)} &= \lim_{n \rightarrow +\infty} (f_n, \nabla \cdot \varphi)_{L^2(\Omega)} \\ &= - \lim_{n \rightarrow +\infty} (\nabla f_n, \varphi)_{[L^2(\Omega)]^d} = (-F, \varphi)_{[L^2(\Omega)]^d}. \end{aligned}$$

On en déduit

$$\|f - f_n\|_{H^1(\Omega)}^2 = \|f - f_n\|_{L^2(\Omega)}^2 + \|F - \nabla f_n\|_{[L^2(\Omega)]^d}^2 \xrightarrow{n \rightarrow +\infty} 0.$$

□

On a enfin le résultat de densité suivant.

Théorème 2.2.4 *L'espace $D(\overline{\Omega})$ est dense dans $H^1(\Omega)$, c'est-à-dire que pour tout $f \in H^1(\Omega)$ il existe une suite $(f_n)_{n \in \mathbb{N}} \in [D(\overline{\Omega})]^\mathbb{N}$ telle que*

$$\lim_{n \rightarrow +\infty} \|f - f_n\|_{H^1(\Omega)} = 0.$$

Les fonctions de $D(\overline{\Omega})$ sont C^∞ à support compact dans $\overline{\Omega}$. Elles ne s'annulent donc pas nécessairement sur le bord de Ω quand Ω a un bord. En revanche, les fonctions de $D(\Omega)$ s'annulent sur le bord de Ω (le support, fermé, de ces fonctions est inclus dans l'ouvert Ω). Par conséquent si $\Omega \subsetneq \mathbb{R}^d$, $D(\Omega) \subsetneq D(\overline{\Omega})$.

Attention c'est $D(\overline{\Omega})$ qui est dense dans $H^1(\Omega)$ et non $D(\Omega)$. En effet, les fonctions de $D(\Omega)$ ne peuvent approcher qu'un sous-espace de fonctions de $H^1(\Omega)$. Introduisons cet espace qui est fondamental dans la suite également.

Définition 2.2.2 *On appelle $H_0^1(\Omega)$ l'adhérence¹⁰ des fonctions de $D(\Omega)$ dans $H^1(\Omega)$, i.e.*

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

10: C'est l'ensemble des limites des suites d'éléments de $D(\Omega)$ convergentes dans $H^1(\Omega)$.

Ainsi, par définition, l'espace $D(\Omega)$ est dense dans $H_0^1(\Omega)$. Si $\Omega = \mathbb{R}^d$, on a $H_0^1(\mathbb{R}^d) = H^1(\mathbb{R}^d)$ et si $\Omega \subsetneq \mathbb{R}^d$, $H_0^1(\Omega) \subsetneq H^1(\Omega)$.

Proposition 2.2.5 *Muni du produit scalaire défini en (2.11), $H_0^1(\Omega)$ est un espace de Hilbert.*

Démonstration. $H_0^1(\Omega)$ est par définition un sous-espace fermé de $H^1(\Omega)$. Donc, d'après la proposition 3.2.1, c'est un espace de Hilbert quand on le munit de la norme de $H^1(\Omega)$. □

2.2.2 Les espaces $H(\Omega, \text{div})$, $H^m(\Omega)$ et $H_0^m(\Omega)$

Commençons par introduire l'espace des champs de vecteurs dont les composantes sont dans $L^2(\Omega)$ et qui admettent une divergence faible (voir la définition 2.1.2).

Définition 2.2.3 L'espace $H(\Omega, \text{div})$ est l'ensemble des champs de vecteurs de $[L^2(\Omega)]^d$ qui admettent une divergence faible

$$H(\Omega, \text{div}) := \{\underline{U} \in [L^2(\Omega)]^d, \quad \nabla \cdot \underline{U} \in L^2(\Omega)\}.$$

11: On peut montrer que l'inclusion est stricte en trouvant des champs de vecteurs $H(\Omega, \text{div})$ qui ne sont pas dans $[H^1(\Omega)]^d$.

Remarquons que¹¹

$$[H^1(\Omega)]^d \subset H(\Omega, \text{div}). \quad (2.14)$$

Nous avons de plus les résultats suivants.

Theorème 2.2.6 Muni du produit scalaire défini par

$$\begin{aligned} \forall \underline{U}, \underline{V} \in H(\Omega, \text{div}), \\ (\underline{U}, \underline{V})_{H(\text{div})} := (\underline{U}, \underline{V})_{[L^2(\Omega)]^d} + (\nabla \cdot \underline{U}, \nabla \cdot \underline{V})_{L^2(\Omega)}. \end{aligned} \quad (2.15)$$

$H(\Omega, \text{div})$ est un espace de Hilbert.

Theorème 2.2.7 L'espace $[D(\overline{\Omega})]^d$ est dense dans $H(\Omega, \text{div})$.

On définit maintenant par récurrence les espaces de Sobolev $H^m(\Omega)$ pour $m \geq 2$.

Définition 2.2.4 L'espace $H^m(\Omega)$ est l'ensemble des fonctions de $H^{m-1}(\Omega)$ dont toutes les dérivées partielles faibles d'ordre $m-1$ admettent une dérivée faible dans toutes les directions¹² :

$$H^m(\Omega) := \{v \in H^{m-1}(\Omega), \partial^\alpha v \in L^2(\Omega) \quad \forall \alpha \in \mathbb{N}^d, |\alpha| = m\}.$$

12: On rappelle la notation :

$$\forall \alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d, |\alpha| := \sum_{k=1}^d |\alpha_k|,$$

et

$$\partial^\alpha := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}.$$

Un exemple intéressant concerne le cas où $\Omega \equiv]0, 1[\subset \mathbb{R}$. Il est facile de montrer que

$$x \mapsto x^\beta \in H^m(]0, 1[) \quad \text{ssi} \quad \beta > -\frac{1}{2} + m \quad (2.16)$$

De plus, pour Ω ouvert borné, toutes les fonctions de $C^m(\overline{\Omega})$ sont dans $H^m(\Omega)$. On a également la suite d'inclusions suivante ($m \geq 2$) :

$$D(\Omega) \subset H^m(\Omega) \subset H^1(\Omega) \subset L^2(\Omega).$$

Remarque 2.2.1 Il est possible de définir les espaces de Sobolev fractionnaires $H^s(\Omega)$, $s \in \mathbb{R}$ en utilisant la théorie dite d'interpolation (voir [21]). Cela dépasse le cadre du cours, mais on retiendra que

$$s < s' \quad \Rightarrow \quad H^{s'}(\Omega) \subset H^s(\Omega).$$

De plus, les fonctions de $H^s(\Omega)$ pour $m < s < m + 1$ avec $m \in \mathbb{N}$ sont plus régulières que les fonctions de $H^m(\Omega)$ sans être dans $H^{m+1}(\Omega)$:

$$H^{m+1}(\Omega) \subset H^s(\Omega) \subset H^m(\Omega) \quad \text{pour } m < s < m + 1.$$

Theorème 2.2.8 Muni du produit scalaire défini pour $f, g \in H^m(\Omega)$ par

$$(f, g)_{H^m(\Omega)} := (f, g)_{H^{m-1}(\Omega)} + \sum_{\substack{\alpha \in \mathbb{N}^m \\ |\alpha|=m}} (\partial^\alpha f, \partial^\alpha g)_{L^2(\Omega)}. \quad (2.17)$$

$H^m(\Omega)$ est un espace de Hilbert.

Dans la suite, nous travaillerons en particulier avec l'espace $H^2(\Omega)$ qui est donc défini par

$$H^2(\Omega) := \left\{ v \in H^1(\Omega), \frac{\partial^2 v}{\partial x_i \partial x_j} \in L^2(\Omega) \quad \forall i, j \in \llbracket 1, d \rrbracket \right\}. \quad (2.18)$$

Définissons également le sous-espace des fonctions de $H^1(\Omega)$ dont le laplacien est dans $L^2(\Omega)$

$$\begin{aligned} H(\Omega, \Delta) &:= \{v \in H^1(\Omega), \Delta v \in L^2(\Omega)\} \\ &= \{v \in H^1(\Omega), \nabla v \in H(\Omega, \text{div})\}, \end{aligned} \quad (2.19)$$

que l'on munit du produit scalaire

$$\forall f, g \in H(\Omega, \Delta), \quad (f, g)_{H(\Omega, \Delta)} := (f, g)_{H^1(\Omega)} + (\Delta f, \Delta g)_{L^2(\Omega)}. \quad (2.20)$$

Notons que

$$H^2(\Omega) \subset H(\Omega, \Delta),$$

puisque si toutes les dérivées secondes d'une fonction $u \in H^1(\Omega)$ sont dans $L^2(\Omega)$ alors son laplacien appartient à $L^2(\Omega)$. On montre que si $\Omega = \mathbb{R}^d$, il y a égalité, c'est à dire $H^2(\mathbb{R}^d) \subset H(\mathbb{R}^d, \Delta)$ alors que dès que Ω a un bord, cette inclusion est stricte.

Nous avons enfin les résultats de densités suivants.

Theorème 2.2.9 Pour tout m , l'espace $D(\overline{\Omega})$ est dense dans $H^m(\Omega)$. De plus, l'espace $D(\overline{\Omega})$ est dense dans $H(\Omega, \Delta)$.

Définition 2.2.5 On appelle $H_0^m(\Omega)$ l'adhérence des fonctions de $D(\Omega)$ dans $H^m(\Omega)$. Ceci s'écrit aussi

$$H_0^m(\Omega) = \overline{D(\Omega)}^{H^m(\Omega)}.$$

Finissons cette section par discuter la régularité des fonctions de $H^m(\Omega)$. Les fonctions de $H^m(\Omega)$ sont définies par des conditions d'intégrabilité. On peut se poser la question de la régularité au sens usuel (c'est à dire dans $C^k(\overline{\Omega})$) de telles fonctions. Les fonctions de $H^m(\Omega)$ pour $m \geq 0$ ne sont pas toujours continues (penser à $m = 0$ ou alors $H^0(\Omega) = L^2(\Omega)$). Cela dépend en fait de m et de la dimension d . Nous admettrons le

résultat suivant (voir le paragraphe suivant pour la justification pour $d = 1$, et par exemple [6, 17]).

Théorème 2.2.10 *Si Ω est un ouvert de frontière suffisamment régulière et si $m > \frac{d}{2} + k$ alors*

$$H^m(\Omega) \subset C^k(\overline{\Omega}),$$

et $H^m(\Omega)$ s'injecte de manière continue dans $C^k(\overline{\Omega})$, c'est à dire :

$$\exists C > 0, \forall u \in H^m(\Omega), \quad \|u\|_{C^k(\overline{\Omega})} := \sup_{x \in \overline{\Omega}} \sum_{|\alpha| \leq k} |\partial^\alpha u(x)| \leq C \|u\|_{H^m(\Omega)}.$$

Ainsi plus m est grand, plus les fonctions sont régulières, c'est à dire dérivables au sens classique jusqu'à un ordre strictement plus petit que m . On déduit de ce résultat que

$$\text{si } \Omega \subset \mathbb{R}, \quad H^1(\Omega) \subset C^0(\overline{\Omega}). \quad (2.21)$$

Nous démontrerons ce résultat au paragraphe suivant. Indiquons que ce résultat n'est pas vrai quand $\Omega \subset \mathbb{R}^2$ ou $\Omega \subset \mathbb{R}^3$. Ainsi, il existe des fonctions de $H^1(\Omega)$ qui ne sont pas continues et qui peuvent même exploser en certains points (voir la fonction donnée en (2.25)). Notons également que le théorème précédent donne

$$\text{si } \Omega \subset \mathbb{R}^2 \text{ ou } \Omega \subset \mathbb{R}^3, \quad H^2(\Omega) \subset C^0(\overline{\Omega}), \quad (2.22)$$

mais pour de tels domaines, les fonctions de $H^2(\Omega)$ ne sont en général pas dans $C^1(\overline{\Omega})$.

2.2.3 Théorèmes de trace

Comme toute fonction mesurable, les fonctions de $L^2(\Omega)$, $H^1(\Omega)$, $H^2(\Omega)$ ne sont définies que presque partout. On ne peut a priori pas parler de leur valeur en un point ou de leur valeur au bord (ce qui est pourtant essentiel quand il s'agira de proposer des conditions aux limites pour des fonctions de tels espaces). Heureusement pour les fonctions de $H^1(\Omega)$, il est possible de définir leur "restriction au bord", ou plus rigoureusement leur trace au bord, en tant que fonction L^2 du bord. C'est le résultat essentiel suivant appelé théorème de trace.

Énonçons le tout d'abord en dimension un c'est à dire quand $\Omega =]a, b[\subset \mathbb{R}$.

Théorème 2.2.11 *Soit $a < b$. L'application trace*

$$\gamma_0 : \begin{cases} C^\infty([a, b]) \rightarrow \mathbb{R} \\ v \mapsto v(a) \end{cases}$$

vérifie

$$\exists C_0 > 0, \quad \forall v \in C^\infty([a, b]), \quad |v(a)| \leq C_0 \|v\|_{H^1([a, b])}. \quad (2.23)$$

Par densité de $C^\infty([a, b])$ dans $H^1(]a, b[)$, l'application se prolonge¹³ en

$$\gamma_0 : \begin{cases} H^1(]a, b[) \rightarrow \mathbb{R} \\ v \mapsto v(a) \end{cases}$$

et

$$\exists C_0 > 0, \quad \forall v \in H^1(]a, b[), \quad |v(a)| \leq C_0 \|v\|_{H^1(]a, b[)}.$$

13: Voir les théorèmes 2.2.4 et 2.1.4

Démonstration. Posons $\Omega =]a, b[$ et considérons $v \in C^\infty(\overline{\Omega})$, on écrit

$$v(a) = v(x) - \int_a^x v'(t) dt.$$

On sait que si deux fonctions sont égales, leurs normes L^2 sont égales. En utilisant une inégalité triangulaire, ceci implique

$$\|v(a)\|_{L^2(\Omega)} \leq \|v\|_{L^2(\Omega)} + \left\| \int_a^x v'(t) dt \right\|_{L^2(\Omega)},$$

où le premier terme correspondant à la norme L^2 de la fonction constante égale à $v(a)$. On a

$$\|v(a)\|_{L^2(\Omega)} = \sqrt{b-a} |v(a)|.$$

On majore ensuite le dernier terme à l'aide de l'inégalité de Cauchy-Schwarz,

$$\begin{aligned} \left| \int_a^x v'(t) dt \right| &\leq \int_a^x |v'(t)| dt \leq \left(\int_a^x 1 dt \right)^{\frac{1}{2}} \left(\int_a^x |v'(t)|^2 dt \right)^{\frac{1}{2}} \\ &\leq \sqrt{x-a} \left(\int_a^b |v'(t)|^2 dt \right)^{\frac{1}{2}} \leq \sqrt{b-a} \left(\int_a^b |v'(t)|^2 dt \right)^{\frac{1}{2}}. \end{aligned}$$

On en déduit

$$\left\| \int_a^x v'(t) dt \right\|_{L^2(\Omega)} \leq (b-a) \|v'\|_{L^2(\Omega)}.$$

On en conclut que, pour tout élément v de $C^\infty([a, b])$, on a la majoration

$$\begin{aligned} |v(a)| &\leq \frac{1}{\sqrt{b-a}} \|v\|_{L^2(\Omega)} + \sqrt{b-a} \|v'\|_{L^2(\Omega)} \\ &\leq \sqrt{2} \max\left(\frac{1}{\sqrt{b-a}}, \sqrt{b-a}\right) \|v\|_{H^1(\Omega)}, \end{aligned}$$

où pour la dernière inégalité, on a utilisé $(\alpha + \beta) \leq \sqrt{2} \sqrt{\alpha^2 + \beta^2}$, valable pour tous $\alpha, \beta \in \mathbb{R}^+$.

D'après le théorème 2.1.4 avec $\mathcal{H} = H^1(]a, b[)$, $\mathcal{W} = \mathbb{R}$ et $\mathcal{V} = C^\infty([a, b])$ on peut prolonger l'application trace $\gamma_0 : v \mapsto v(a)$ de façon unique en une application linéaire et continue de $H^1(]a, b[)$ dans $\mathbb{R}$.

□

Notons que la constante C_0 ne dépend que de la longueur $(b - a)$.

Remarque 2.2.2 Nous pouvons donc parler de la valeur d'une fonction de $H^1([a, b])$ en un point. C'est assez remarquable et ceci témoigne d'une "régularité" supplémentaire des fonctions de H^1 par rapport aux fonctions qui ne sont que L^2 . Prenez les exemples (2.2) et (2.8) des fonctions de la forme $x \mapsto x^\beta$. Pour $\beta < 0$ ces fonctions (qui sont seulement dans $L^2([0, 1])$) explosent en 0 (elles n'ont pas de trace) alors que pour $\beta > \frac{1}{2}$ (elles sont dans H^1), elles ont bien une trace! De plus, il est possible de montrer que l'inégalité (2.23) quand on remplace la norme H^1 par la norme L^2 est fausse. En effet, prenons $a = 0$ et $b = 1$. Soit $(v_m)_{m \geq 1}$ la suite d'éléments de $L^2([0, 1])$ définie par $v_m(x) = e^{-mx}$. Par un calcul direct, on montre que

$$\int_0^1 v_m^2 dx = \int_0^1 e^{-2mx} dx = \left[\frac{1}{-2m} e^{-2mx} \right]_0^1 = \frac{1}{2m} (1 - e^{-2m}).$$

Ainsi $\lim_{m \rightarrow \infty} \|v_m\|_{L^2([0, 1])} = 0$, alors que $v_m(0) = 1$ pour tout $m \geq 1$. On ne peut donc pas trouver de constante $C > 0$ telle que, pour tout $m \geq 1$, on ait $|v_m(0)| \leq C \|v_m\|_{L^2([0, 1])}$.

En adaptant la preuve ci-dessus, on montre pareillement que les fonctions de $H^1([a, b])$ sont continues. C'est un cas particulier du théorème 2.2.10 pour $d = 1$, $m = 1$ et $k = 0$.

Corollaire 2.2.12 Soit $a < b$. Toute fonction v de $H^1([a, b])$ est continue dans $[a, b]$ et

$$\forall x, y \in [a, b], \quad v(x) = v(y) + \int_y^x v'(t) dt.$$

Démonstration. Soit $v \in H^1([a, b])$, d'après le théorème 2.2.3, il existe une suite $(v_n)_{n \in \mathbb{N}}$ de fonctions de $C^\infty([a, b])$ telles que

$$\lim_{n \rightarrow +\infty} \|v_n - v\|_{H^1(\Omega)} = 0.$$

Les fonctions v_n sont C^∞ donc elles vérifient pour tout $x, y \in [a, b]$

$$v_n(x) = v_n(y) + \int_y^x v'_n(t) dt.$$

On a d'une part

$$\left| \int_y^x (v'(t) - v'_n(t)) dt \right| \leq \|v' - v'_n\|_{L^2(\Omega)} \sqrt{b - a}$$

et d'autre part d'après la continuité de l'application trace en x puis en y

$$|v(x) - v_n(x)| \leq C_0 \|v - v_n\|_{H^1(\Omega)} \quad \text{et} \quad |v(y) - v_n(y)| \leq C_0 \|v - v_n\|_{H^1(\Omega)}.$$

On déduit par passage à la limite que

$$v(x) = v(y) + \int_y^x v'(t) dt.$$

Ceci permet d'ailleurs de montrer en utilisant une inégalité de Cauchy-Schwarz que

$$\forall x, y \in [a, b], \quad |v(x) - v(y)| \leq \sqrt{|x - y|} \|v'\|_{L^2(\Omega)}.$$

La fonction v est donc bien continue sur $[a, b]$. $\square$

Enonçons maintenant le théorème de trace dans $\Omega \subset \mathbb{R}^d$ pour $d \geq 1$.

Théorème 2.2.13 *L'application trace*

$$\gamma_0 : \begin{cases} C^\infty(\overline{\Omega}) \rightarrow L^2(\partial\Omega) \\ v \mapsto v|_{\partial\Omega} \end{cases}$$

est bien définie et vérifie

$$\exists C_0 > 0, \quad \forall v \in C^\infty(\overline{\Omega}), \quad \|v|_{\partial\Omega}\|_{L^2(\partial\Omega)} \leq C_0 \|v\|_{H^1(\Omega)}. \quad (2.24)$$

Par densité de $C^\infty(\overline{\Omega})$ dans $H^1(\Omega)$, l'application se prolonge¹⁴ en

$$\gamma_0 : \begin{cases} H^1(\Omega) \rightarrow L^2(\partial\Omega) \\ v \mapsto v|_{\partial\Omega} \end{cases}$$

et

$$\exists C_0 > 0, \quad \forall v \in H^1(\Omega), \quad \|v|_{\partial\Omega}\|_{L^2(\partial\Omega)} \leq C_0 \|v\|_{H^1(\Omega)}.$$

14: voir les théorèmes 2.2.4 et 2.1.4

Démonstration. Etudions le cas où Ω est un demi-espace, par exemple $\mathbb{R}_*^+ \times \mathbb{R}^{d-1}$. Soit $v \in C^\infty(\mathbb{R}^+ \times \mathbb{R}^{d-1})$. Si on note $x = (x_1, \tilde{x})$, d'après le théorème 2.2.11 appliqué à $x_1 \mapsto v(x_1, \tilde{x})$ pour $\tilde{x} \in \mathbb{R}^{d-1}$, il existe $C_0 > 0$ ¹⁵ telle que

$$|v(0, \tilde{x})|^2 \leq C_0^2 \int_{\mathbb{R}^+} \left[|v(x_1, \tilde{x})|^2 + \left| \frac{\partial v}{\partial x_1}(x_1, \tilde{x}) \right|^2 \right] dx_1.$$

En intégrant en $\tilde{x}$, on obtient

$$\int_{\mathbb{R}^{d-1}} |v(0, \tilde{x})|^2 d\tilde{x} \leq C_0^2 \int_{\mathbb{R}^+ \times \mathbb{R}^{d-1}} \left[|v(x)|^2 + \left| \frac{\partial v}{\partial x_1}(x) \right|^2 \right] dx \leq C_0^2 \|v\|_{H^1(\Omega)}^2.$$

On en déduit que $\|v|_{\partial\Omega}\|_{L^2(\partial\Omega)} \leq C_0 \|v\|_{H^1(\Omega)}$ pour tout $v \in C^\infty(\overline{\Omega})$ et donc par densité pour tout $v \in H^1(\Omega)$. Notons que la même preuve fonctionne évidemment si Ω est une bande $]a, b[\times \mathbb{R}^{d-1}$ ou un parallélépipède.

Pour démontrer le résultat pour un domaine Ω ouvert de $\mathbb{R}^d$ à frontière assez régulière, il faut utiliser des cartes locales du bord, redresser le bord puis utiliser un calcul similaire à celui présenté ci-dessus. Nous ne détaillons pas cet argument et renvoyons à [6]. $\square$

15: Il est possible de montrer que la meilleure constante est 1 dans le cas où l'intervalle est $\mathbb{R}^+$ mais ce n'est pas important pour la preuve.

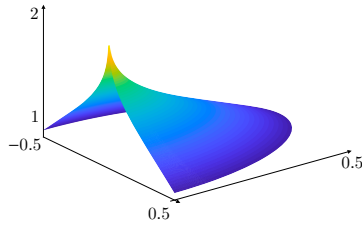


FIGURE 2.1 – Représentation de la fonction v dans Ω .

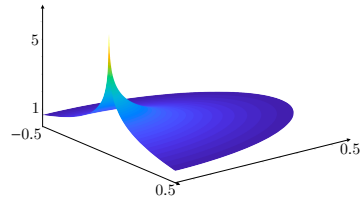


FIGURE 2.2 – Représentation de la fonction w dans Ω .

Le théorème de trace permet de donner un sens à la trace d'une fonction au bord en tant que fonction L^2 du bord, et ce même pour des fonctions qui ne sont pas continues !

Prenons par exemple la fonction

$$v : (x, y) \mapsto \ln \left(\ln \left(\frac{1}{\sqrt{x^2 + y^2}} \right) \right), \quad (2.25)$$

définie dans le demi-disque

$$\Omega := \{(x, y) \in \mathbb{R}^2, x^2 + y^2 < 1, x > 0\}.$$

Cette fonction n'est pas dans $C^0(\overline{\Omega})$ car elle explose en $(0, 0)$. On ne peut donc pas définir $v(0, 0)$. Cependant, on peut montrer que v est dans $H^1(\Omega)$. D'après le théorème de trace, on peut définir sa trace sur le bord $\partial\Omega$ et sa trace est dans $L^2(\partial\Omega)$.

Prenons maintenant la fonction

$$w : (x, y) \mapsto \ln \left(\frac{1}{\sqrt{x^2 + y^2}} \right).$$

Cette fonction explose aussi en $(0, 0)$, elle est dans $L^2(\Omega)$, mais pas dans $H^1(\Omega)$ (l'explosion est trop forte en $(0, 0)$). On ne peut donc pas définir sa trace comme une fonction de $L^2(\partial\Omega)$.

Énonçons maintenant quelques propriétés de l'image et du noyau de l'application trace.

Proposition 2.2.14 *L'application trace n'est pas surjective dans $L^2(\partial\Omega)$. En effet, l'image de l'application trace γ_0 est un sous-espace strict de $L^2(\partial\Omega)$ qui est dense dans $L^2(\partial\Omega)$. Quand on munit l'espace $\text{Im } \gamma_0$ de la norme*

$$\forall \varphi \in \text{Im } \gamma_0, \quad \|\varphi\|_{\text{Im } \gamma_0} := \inf_{\substack{u \in H^1(\Omega) \\ u|_{\partial\Omega} = \varphi}} \|u\|_{H^1(\Omega)},$$

c'est un espace de Banach. De plus, on a

$$\forall \varphi \in \text{Im } \gamma_0, \quad \exists v \in H^1(\Omega) \text{ telle que } v|_{\partial\Omega} = \varphi, \quad \|v\|_{H^1(\Omega)} \leq C \|\varphi\|_{\text{Im } \gamma_0}.$$

La fonction v est appelée relèvement de φ (et n'est pas unique).

Remarque 2.2.3 L'espace $\text{Im } \gamma_0$ est un sous-espace de $L^2(\partial\Omega)$ qui est composé de fonctions plus régulières que L^2 . On peut montrer (mais cela dépasse le cadre de ce cours) que c'est un espace de Sobolev fractionnaire (voir la remarque 2.2.1), noté $H^{\frac{1}{2}}(\partial\Omega)$ qu'il est possible de caractériser par

$$H^{\frac{1}{2}}(\partial\Omega) = \left\{ \varphi \in L^2(\partial\Omega), \frac{|\varphi(x) - \varphi(y)|}{|x - y|^{d/2}} \in L^2(\partial\Omega \times \partial\Omega) \right\}.$$

Le fait qu'il existe toujours un relèvement est une évidence. Ce qui n'est absolument pas évident c'est qu'il existe un relèvement dont

la norme H^1 est contrôlée par la norme $H^{\frac{1}{2}}$ de la trace. C'est une conséquence (théorème de l'application ouverte ou du graphe fermé) d'un théorème fondamental de l'analyse qui est le théorème de Hahn-Banach analytique.

Proposition 2.2.15 *L'application trace n'est pas injective : son noyau est l'espace $H_0^1(\Omega)$ (voir la définition 2.2.2). On a donc la caractérisation*

$$H_0^1(\Omega) = \{u \in H^1(\Omega), \quad u|_{\partial\Omega} = 0\}.$$

Démonstration. Nous ne montrerons que le fait que l'application n'est pas injective et que $H_0^1(\Omega)$ est bien inclus dans le noyau. Remarquons tout d'abord que toute fonction de $D(\Omega)$ est à trace nulle sur $\partial\Omega$. Soit $v \in H_0^1(\Omega)$. Par définition, il existe une suite $(v_n)_{n \in \mathbb{N}}$ d'éléments de $D(\Omega)$ telle que

$$\lim_{n \rightarrow +\infty} \|v - v_n\|_{H^1(\Omega)} = 0.$$

Par continuité de l'application trace et comme $v_n|_{\partial\Omega} = 0$, on a

$$\|v|_{\partial\Omega}\|_{L^2(\partial\Omega)} \leq C_0 \|v - v_n\|_{H^1(\Omega)} \xrightarrow{n \rightarrow +\infty} 0,$$

soit $v|_{\partial\Omega} = 0$. La réciproque est très technique, nous renvoyons le lecteur à [6]. $\square$

Il est possible d'établir le deuxième théorème de trace qui permet de donner un sens à la dérivée normale de fonctions de $H^2(\Omega)$. En effet, en appliquant le théorème 2.2.13 de trace à chacune des composantes du gradient de fonctions $H^2(\Omega)$ qui sont bien toutes dans $H^1(\Omega)$, on obtient facilement le deuxième théorème de trace.

Théorème 2.2.16 *L'application trace*

$$\gamma_1 : \begin{cases} H^2(\Omega) \rightarrow L^2(\partial\Omega) \\ v \mapsto \underline{\nabla} v \cdot \underline{n}|_{\partial\Omega} \end{cases}$$

est bien définie et vérifie

$$\exists C_1 > 0, \quad \forall v \in H^2(\Omega), \quad \|\underline{\nabla} v \cdot \underline{n}|_{\partial\Omega}\|_{L^2(\partial\Omega)} \leq C_1 \|v\|_{H^2(\Omega)}. \quad (2.26)$$

Il est donc possible de définir la trace de la dérivée normale pour des fonctions de $H^2(\Omega)$ en tant que fonctions de $L^2(\partial\Omega)$. Cependant, ceci n'est pas possible pour des fonctions qui sont seulement $H^1(\Omega)$.

Le lecteur pourra aisément étendre ces théorèmes de trace à la trace de dérivées d'ordre plus élevé à condition que la fonction ait une régularité suffisante. Dans la suite de cours, nous n'aurons besoin que de ces deux théorèmes de trace.

2.2.4 Formules de Green

Nous avons maintenant tous les éléments pour étendre les formules de Green initialement écrites pour des fonctions dans C^k à des fonctions des espaces de Sobolev. Commençons par la formule de Stokes.

Théorème 2.2.17 Soit Ω un ouvert borné de $\mathbb{R}^d$ de frontière suffisamment régulière et $w \in H^1(\Omega)$ alors pour tout $i \in \llbracket 1, d \rrbracket$

$$\int_{\Omega} \frac{\partial w}{\partial x_i} d\Omega = \int_{\partial\Omega} w|_{\partial\Omega} n_i d\Gamma.$$

Démonstration. Nous allons utiliser le théorème 2.1.4 de prolongement par densité. Introduisons tout d'abord $\ell_1 : C^\infty(\overline{\Omega}) \rightarrow \mathbb{R}$ défini par

$$\forall w \in C^\infty(\overline{\Omega}), \quad \ell_1(w) := \int_{\Omega} \frac{\partial w}{\partial x_i} d\Omega,$$

qui est une forme linéaire continue pour la norme H^1 :

$$\forall w \in C^\infty(\overline{\Omega}), \quad |\ell_1(w)| \leq |\Omega|^{\frac{1}{2}} \left\| \frac{\partial w}{\partial x_i} \right\|_{L^2(\Omega)} \leq |\Omega|^{\frac{1}{2}} \|w\|_{H^1(\Omega)},$$

où on a utilisé une inégalité de Cauchy-Schwarz dans $L^2(\Omega)$ et la définition de la norme de $H^1(\Omega)$. Comme $C^\infty(\overline{\Omega})$ est dense dans $H^1(\Omega)$ (voir le théorème 2.2.4), d'après le théorème 2.1.4 de prolongement par densité, ℓ_1 se prolonge en une application linéaire continue de $H^1(\Omega)$ dans $\mathbb{R}$.

Introduisons maintenant $\ell_2 : C^\infty(\overline{\Omega}) \rightarrow \mathbb{R}$ défini par

$$\forall w \in C^\infty(\overline{\Omega}), \quad \ell_2(w) := \int_{\partial\Omega} w|_{\partial\Omega} n_i d\Gamma.$$

On peut montrer que ℓ_2 est également une forme linéaire continue pour la norme H^1 . En effet, on a

$$\forall w \in C^\infty(\overline{\Omega}), \quad |\ell_2(w)| \leq |\partial\Omega|^{\frac{1}{2}} \|w|_{\partial\Omega}\|_{L^2(\partial\Omega)} \leq C_0 |\partial\Omega|^{\frac{1}{2}} \|w\|_{H^1(\Omega)},$$

où on a utilisé une inégalité de Cauchy-Schwarz dans $L^2(\partial\Omega)$ et le théorème 2.2.13 de trace. D'après le théorème 2.1.4 de prolongement par densité, ℓ_2 se prolonge en une application linéaire continue de $H^1(\Omega)$ dans $\mathbb{R}$.

Enfin, comme d'après la formule de Green donnée dans le théorème 1.2.1, nous avons

$$\forall w \in C^\infty(\overline{\Omega}), \quad \ell_1(w) = \ell_2(w).$$

Comme $C^\infty(\overline{\Omega})$ est dense dans $H^1(\Omega)$ et que $\ell_1 - \ell_2$ est continue dans $H^1(\Omega)$, on en déduit que

$$\forall w \in H^1(\Omega), \quad \ell_1(w) = \ell_2(w).$$

□

À partir de cette formule, on établit les formules de Green suivantes

$$\forall \underline{W} \in [H^1(\Omega)]^d, \int_{\Omega} \underline{\nabla} \cdot \underline{W} \, d\Omega = \int_{\partial\Omega} \underline{W}|_{\partial\Omega} \cdot \underline{n} \, d\Gamma, \quad (2.27)$$

pour tout $i \in \llbracket 1, d \rrbracket$

$$\forall u, v \in H^1(\Omega), \int_{\Omega} \left[\frac{\partial u}{\partial x_i} v + u \frac{\partial v}{\partial x_i} \right] d\Omega = \int_{\partial\Omega} u|_{\partial\Omega} v|_{\partial\Omega} \underline{n}_i \, d\Gamma, \quad (2.28)$$

et enfin

$$\forall \underline{U} \in [H^1(\Omega)]^d, \forall v \in H^1(\Omega), \int_{\Omega} [v \underline{\nabla} \cdot \underline{U} + \underline{U} \cdot \underline{\nabla} v] \, d\Omega = \int_{\partial\Omega} \underline{U}|_{\partial\Omega} \cdot \underline{n} v|_{\partial\Omega} \, d\Gamma. \quad (2.29)$$

Arrêtons-nous un instant sur cette dernière formule de Green. Remarquons qu'en utilisant le théorème 2.1.4 de prolongement par densité et le théorème 2.2.7 de densité, il est possible de montrer que la forme bilinéaire

$$(\underline{U}, v) \mapsto \int_{\Omega} [v \underline{\nabla} \cdot \underline{U} + \underline{U} \cdot \underline{\nabla} v] \, d\Omega,$$

s'étend en une forme bilinéaire continue pour des fonctions

$$\underline{U} \in H(\Omega, \text{div}), \quad v \in H^1(\Omega).$$

Il n'est en effet pas nécessaire de contrôler toutes les dérivées de toutes les composantes de $\underline{U}$ pour étendre cette intégrale. Contrôler seulement la divergence est suffisant. Cependant, il n'est pas possible de définir la trace de tous les champs de vecteurs de $H(\Omega, \text{div})$ en tant que champ de vecteurs dans $[L^2(\partial\Omega)]^d$. Il est en effet possible de construire des champs de vecteurs de $H(\Omega, \text{div}) \setminus [H^1(\Omega)]^d$ dont la trace n'est pas dans $[L^2(\partial\Omega)]^d$. Est-ce qu'il est tout de même possible d'étendre la formule de Green (1.9) à des champs de vecteurs $\underline{U} \in H(\Omega, \text{div})$? La réponse est oui, mais à condition de définir de manière encore plus faible la trace normale $\underline{U} \cdot \underline{n}|_{\partial\Omega}$. La trace normale est en fait définie par la formule de Green.

Théorème 2.2.18 *L'application $\gamma_{\underline{n}}$ définie par*

$$\forall \underline{U} \in [H^1(\Omega)]^d, \quad \gamma_{\underline{n}} \underline{U} := \underline{U}|_{\partial\Omega} \cdot \underline{n} \in L^2(\Omega),$$

s'étend en une application linéaire continue de $H(\Omega, \text{div})$ dans le dual de $\text{Im } \gamma_0$ noté¹⁶ $[\text{Im } \gamma_0]'$

$$\forall \underline{U} \in H(\Omega, \text{div}), \quad \gamma_{\underline{n}} \underline{U} := \underline{U} \cdot \underline{n}|_{\partial\Omega} \in [\text{Im } \gamma_0]'$$

où $\forall \varphi \in \text{Im } \gamma_0$,

$$\langle \underline{U} \cdot \underline{n}|_{\partial\Omega}, \varphi \rangle_{[\text{Im } \gamma_0]', \text{Im } \gamma_0} := \int_{\Omega} [v(\varphi) \underline{\nabla} \cdot \underline{U} + \underline{U} \cdot \underline{\nabla} v(\varphi)] \, d\Omega,$$

où $v(\varphi) \in H^1(\Omega)$ est telle que $v(\varphi)|_{\partial\Omega} = \varphi$.

16: Comme mentionné dans la proposition 2.2.14, $\text{Im } \gamma_0$ est l'espace de Sobolev fractionnaire $H^{\frac{1}{2}}(\partial\Omega)$. Il s'avère que son dual $[\text{Im } \gamma_0]'$ est également un espace de Sobolev fractionnaire mais d'ordre négatif : $[\text{Im } \gamma_0]' \equiv H^{-\frac{1}{2}}(\partial\Omega)$

D'après la formule de Green 2.29, cette définition de la trace normale

étend bien la trace normale pour des fonctions plus régulières. La formule de Green faisant intervenir le laplacien sera également très importante.

Remarque 2.2.4 Dans ce cours, nous proposons de ne pas utiliser cette définition délicate de la trace normale pour les fonctions de $H(\Omega, \text{div})$. Dès qu'un champ de vecteur est $H(\Omega, \text{div})$, nous allons supposer qu'il est dans $[H^1(\Omega)]^d$. Ainsi, nous ne manipulerons que des intégrales faisant intervenir des fonctions qui sont bien L^2 . Le résultat précédent permet de se convaincre que les démonstrations sont très similaires à condition de remplacer les intégrales sur le bord par des crochets de dualité.

Théorème 2.2.19 Soit Ω un ouvert borné de $\mathbb{R}^d$ de frontière suffisamment régulière et $u \in H^2(\Omega)$, $v \in H^1(\Omega)$ alors

$$\int_{\Omega} [\Delta u v + \nabla u \cdot \nabla v] \, d\Omega = \int_{\partial\Omega} \nabla u \cdot \underline{n}|_{\partial\Omega} v|_{\partial\Omega} \, d\Gamma,$$

où $\underline{n}$ est la normale unitaire à Ω dirigée vers l'extérieur.

Démonstration. La preuve est très similaire à celle du théorème 2.2.17 mais il faut utiliser le théorème de prolongement par densité donné dans le corollaire 2.1.5, la densité de $C^\infty(\bar{\Omega})$ dans $H^1(\Omega)$ d'une part (voir théorème 2.2.4) et dans $H^2(\Omega)$ d'autre part (voir le théorème 2.2.9), le premier théorème de trace (voir le théorème 2.2.13) et le deuxième théorème de trace (voir le théorème 2.2.16). $\square$

Remarque 2.2.5 Remarquons que la forme bilinéaire

$$(u, v) \mapsto \int_{\Omega} [\Delta u v + \nabla u \cdot \nabla v] \, d\Omega,$$

s'étend en une forme bilinéaire continue pour des fonctions

$$u \in H^1(\Omega, \Delta), \quad v \in H^1(\Omega).$$

En fait, par définition de $H^1(\Omega, \Delta)$, on a

$$H^1(\Omega, \Delta) = \left\{ u \in H^1(\Omega), \quad \nabla u \in H(\Omega, \text{div}) \right\}.$$

En utilisant le théorème 2.2.18, il est donc possible de donner un sens à $\nabla u \cdot \underline{n}|_{\partial\Omega}$ dans $[\text{Im } \gamma_0]'$ et d'étendre la formule de Green aux fonctions $u \in H^1(\Omega, \Delta)$, $v \in H^1(\Omega)$

$$\int_{\Omega} [\Delta u v + \nabla u \cdot \nabla v] \, d\Omega = \langle \nabla u \cdot \underline{n}|_{\partial\Omega}, v|_{\partial\Omega} \rangle_{[\text{Im } \gamma_0]', [\text{Im } \gamma_0]}. \quad (2.30)$$

Comme indiqué dans la remarque 2.2.4, nous allons éviter de manipuler cette formule de Green faisant intervenir le crochet de dualité. Pour toute fonction u dans $H^1(\Omega, \Delta)$, nous allons supposer qu'elle est dans $H^2(\Omega)$ pour nous ramener à la formule de Green du théorème 2.2.19 et ne manipuler que des intégrales faisant intervenir des fonctions qui sont L^2 . Les démonstrations sont, sans cette hypothèse, en fait très

similaires, mais un peu plus techniques.

2.2.5 Inégalités de type Poincaré

Un résultat essentiel pour la résolution des problèmes elliptiques du chapitre suivant est l'inégalité de Poincaré¹⁷. C'est une propriété importante des fonctions de $H_0^1(\Omega)$ quand Ω est un ouvert borné.¹⁸

Proposition 2.2.20 Soit $\Omega \subset \mathbb{R}^d$ un ouvert borné. Alors

$$\exists C_p > 0, \forall v \in H_0^1(\Omega), \quad \|v\|_{L^2(\Omega)} \leq C_p \|\nabla v\|_{[L^2(\Omega)]^d}. \quad (2.31)$$

Une des conséquences de cette inégalité est que la semi-norme¹⁹ H^1 notée $|\cdot|_{H^1(\Omega)}$ ²⁰ et définie par

$$\forall v \in H^1(\Omega), \quad |v|_{H^1(\Omega)} := \|\nabla v\|_{[L^2(\Omega)]^d}$$

est une norme dans $H_0^1(\Omega)$ qui est équivalente à la norme H^1 dans $H_0^1(\Omega)$.

Corollaire 2.2.21 La semi-norme H^1 est une norme dans $H_0^1(\Omega)$ qui est équivalente à la norme H^1 :

$$\exists C_1 > 0, \quad \forall v \in H_0^1(\Omega), \quad |v|_{H^1(\Omega)} \leq \|v\|_{H^1(\Omega)} \leq C_1 |v|_{H^1(\Omega)}.$$

$H_0^1(\Omega)$ muni de la semi-norme H^1 est donc un espace de Hilbert.

Démonstration. On a de manière évidente

$$\forall v \in H^1(\Omega), \quad \|\nabla v\|_{[L^2(\Omega)]^d} \leq \|v\|_{H^1(\Omega)}.$$

D'après l'inégalité de Poincaré, on a aussi

$$\forall v \in H_0^1(\Omega), \quad \|v\|_{H^1(\Omega)} \leq \sqrt{1 + C_p^2} \|\nabla v\|_{[L^2(\Omega)]^d}.$$

Dans $H_0^1(\Omega)$, la norme H^1 et la semi-norme H^1 sont donc équivalentes.

La proposition 2.2.5 assure que $H_0^1(\Omega)$ muni de la norme H^1 est un espace de Hilbert. Comme la norme H^1 et la semi-norme H^1 sont équivalentes dans $H_0^1(\Omega)$, il est facile de montrer que toute suite d'éléments de $H_0^1(\Omega)$ est de Cauchy pour la norme H^1 si et seulement si elle l'est pour la semi-norme H^1 . Et elle est convergente pour la norme H^1 si et seulement si elle l'est pour la semi-norme H^1 . On en déduit que l'espace $H_0^1(\Omega)$ muni de la semi-norme H^1 est aussi un espace de Hilbert. $\square$

L'inégalité de Poincaré est une conséquence de l'inégalité plus générale suivante.

Proposition 2.2.22 Soit Ω un ouvert borné de $\mathbb{R}^d$ alors

$$\exists C > 0, \forall v \in H^1(\Omega), \quad \|v\|_{L^2(\Omega)} \leq C \left[\|\nabla v\|_{[L^2(\Omega)]^d} + \|v|_{\partial\Omega}\|_{L^2(\partial\Omega)} \right].$$

17: Henri Poincaré (1854-1912), ancien élève de l'École polytechnique (promotion 1873) puis professeur, a eu des contributions majeures dans des domaines aussi variés que l'analyse, la topologie algébrique, la théorie du chaos et la relativité restreinte. Il ne faut pas le confondre avec son cousin, Raymond Poincaré, président de la république de 1913 à 1920.

18: Cette propriété est vraie également pour des ouverts qui sont bornés seulement dans une seule direction.

19: Une semi-norme dans un espace vectoriel V est "presque" une norme au sens où une des propriétés d'une norme n'est pas vérifiée : certains éléments non nuls peuvent avoir une semi-norme nulle. Ainsi $N : V \rightarrow \mathbb{R}^+$ est une semi-norme si

$$\forall x \in V, \forall \lambda \in \mathbb{R}, \quad N(\lambda x) = |\lambda| N(x)$$

et

$$\forall x, y \in V, \quad N(x + y) \leq N(x) + N(y).$$

20: Il est facile de voir que $|\cdot|_{H^1(\Omega)}$ est bien une semi-norme et que toutes les constantes ont une semi-norme nulle.

Commençons par démontrer cette dernière inégalité dans le cas unidimensionnel d'un intervalle $]a, b[$.

Proposition 2.2.23 Soit $a < b$. On a

$$\forall v \in H^1(]a, b[), \|v\|_{L^2(]a, b[)} \leq (b-a) \|v'\|_{L^2(]a, b[)} + \sqrt{b-a} |v(a)|.$$

Démonstration. D'après le corollaire 2.2.12, on peut écrire

$$\forall x \in]a, b[, \quad v(x) = v(a) + \int_a^x v'(y) \, dy.$$

Nous avons donc deux fonctions qui sont égales pour tout x : leurs normes L^2 sont donc égales. Par application de l'inégalité triangulaire au membre de droite (pour la norme de $L^2(]a, b[)$), on déduit

$$\begin{aligned} \|v\|_{L^2(\Omega)} &\leq \left(\int_a^b \left(\int_a^x v'(y) \, dy \right)^2 dx \right)^{\frac{1}{2}} + \left(\int_a^b |v(a)|^2 dx \right)^{\frac{1}{2}} \\ &\leq \left(\int_a^b \left(\int_a^x |v'(y)|^2 dy \right) (x-a) dx \right)^{\frac{1}{2}} + \sqrt{b-a} |v(a)| \\ &\leq (b-a) \left(\int_a^b |v'(y)|^2 dy \right)^{\frac{1}{2}} + \sqrt{b-a} |v(a)|, \end{aligned}$$

où on a utilisé l'inégalité de Cauchy-Schwarz pour obtenir la deuxième inégalité et on a majoré x par b . $\square$

Démonstration de la Proposition 2.2.22. On considère ici le cas où $\Omega :=]a, b[\times \tilde{\Omega}$ où $\tilde{\Omega}$ est un ouvert (non nécessairement borné) de $\mathbb{R}^{d-1}$.

Soit v une fonction $D(\overline{\Omega})$. D'après la proposition précédente, on a

$$\forall y \in \tilde{\Omega}, \quad \|v(\cdot, y)\|_{L^2(]a, b[)}^2 \leq 2C_{a,b}^2 \left[\|\partial_x v(\cdot, y)\|_{L^2(]a, b[)}^2 + |v(a, y)|^2 \right],$$

avec $C_{a,b} := \max(b-a, \sqrt{b-a})$ et où on a utilisé que $(a+b)^2 \leq 2(a^2+b^2)$. En intégrant en y , on obtient

$$\begin{aligned} \|v\|_{L^2(\Omega)}^2 &\leq 2C_{a,b}^2 \left[\|\partial_x v\|_{L^2(\Omega)}^2 + \int_{\tilde{\Omega}} |v(a, y)|^2 dy \right] \\ &\leq 2C_{a,b}^2 \left[\|\nabla v\|_{[L^2(\Omega)]^d}^2 + \|v|_{\partial\Omega}\|_{L^2(\partial\Omega)}^2 \right], \end{aligned}$$

où on a utilisé que $\{a\} \times \tilde{\Omega} \subset \partial\Omega$. Puisque $\sqrt{\alpha^2 + \beta^2} \leq (\alpha + \beta)$ pour tout $\alpha, \beta \geq 0$, on trouve

$$\|v\|_{L^2(\Omega)} \leq \sqrt{2} C_{a,b} \left[\|\nabla v\|_{[L^2(\Omega)]^d} + \|v|_{\partial\Omega}\|_{L^2(\partial\Omega)} \right].$$

Nous avons donc démontré l'inégalité pour des fonctions $v \in D(\overline{\Omega})$. En utilisant la densité de $D(\overline{\Omega})$ dans $H^1(\Omega)$ (voir le théorème 2.2.4), cette inégalité s'étend aux fonctions dans $H^1(\Omega)$. Détaillons ce point. Soit $v \in H^1(\Omega)$, il existe une suite $(v_n)_{n \in \mathbb{N}}$ d'éléments de $D(\overline{\Omega})$ telle que

$$\lim_{n \rightarrow +\infty} \|v_n - v\|_{H^1(\Omega)} = 0.$$

Chaque élément v_n vérifie

$$\|v_n\|_{L^2(\Omega)} \leq \sqrt{2} C_{a,b} \left[\|\nabla v_n\|_{[L^2(\Omega)]^d} + \|v_n|_{\partial\Omega}\|_{L^2(\partial\Omega)} \right].$$

Il suffit ensuite d'écrire que pour tout n

$$\begin{aligned} \|v\|_{L^2(\Omega)} &\leq \|v_n\|_{L^2(\Omega)} + \|v - v_n\|_{L^2(\Omega)} \\ &\leq \sqrt{2} C_{a,b} \left[\|\nabla v_n\|_{[L^2(\Omega)]^d} + \|v_n|_{\partial\Omega}\|_{L^2(\partial\Omega)} \right] + \|v - v_n\|_{L^2(\Omega)} \\ &\leq \sqrt{2} C_{a,b} \left[\|\nabla v\|_{[L^2(\Omega)]^d} + \|v|_{\partial\Omega}\|_{L^2(\partial\Omega)} \right] + \|v - v_n\|_{L^2(\Omega)} \\ &\quad + \sqrt{2} C_{a,b} \|\nabla v - \nabla v_n\|_{[L^2(\Omega)]^d} + \sqrt{2} C_{a,b} \|v|_{\partial\Omega} - v_n|_{\partial\Omega}\|_{L^2(\partial\Omega)}. \end{aligned}$$

En faisant tendre n vers l'infini et en utilisant que

$$\|v - v_n\|_{L^2(\Omega)} + \|\nabla v - \nabla v_n\|_{L^2(\Omega)} \leq 2\|v - v_n\|_{H^1(\Omega)}$$

par définition des normes L^2 et H^1 et

$$|v|_{\partial\Omega} - v_n|_{\partial\Omega}| \leq C_0 \|v - v_n\|_{H^1(\Omega)}$$

d'après la continuité de l'application trace (voir le théorème 2.2.13), on en déduit que

$$\|v\|_{L^2(\Omega)} \leq \sqrt{2} C_{a,b} \left[\|\nabla v\|_{[L^2(\Omega)]^d} + \|v|_{\partial\Omega}\|_{L^2(\partial\Omega)} \right].$$

□

Pour démontrer l'inégalité de Poincaré de la Proposition 2.2.22 dans un ouvert borné quelconque de $\mathbb{R}^d$, il faut utiliser un résultat de compacité des espaces de Sobolev qui joue un rôle essentiel dans la théorie spectrale des opérateurs intervenant dans les problèmes aux limites que nous étudions dans ce cours (dans le chapitre 8 plus précisément). Ce résultat est appelé le théorème de Rellich. Nous l'admettons et renvoyons le lecteur à [6] pour une preuve. Attention, ce résultat n'est valable que lorsque l'ouvert Ω est borné.

Théorème 2.2.24 (Rellich) *Soit Ω un ouvert borné de $\mathbb{R}^d$ de frontière suffisamment régulière. De toute suite bornée de $H^1(\Omega)$, on peut extraire une sous-suite convergente dans $L^2(\Omega)$ ²¹. Plus généralement, pour $p > q$, de toute suite bornée de $H^p(\Omega)$, on peut extraire une sous-suite convergente dans $H^q(\Omega)$.*

Ce résultat est faux si l'ouvert Ω n'est pas borné. Ainsi, pour $\Omega = \mathbb{R}^d$, considérons la suite u_n définie par $u_n(x) = u(x + n\vec{e})$ où $\vec{e}$ est un vecteur non nul et $u \in H^1(\mathbb{R}^d) \setminus \{0\}$. La fonction u_n n'est rien d'autre que la translatée de u . Il est clair qu'aucune sous-suite de u_n ne peut converger dans L^2 .

Nous pouvons maintenant proposer une nouvelle démonstration de la proposition 2.2.22.

Démonstration alternative de la Proposition 2.2.22. Raisonnons par l'absurde.

21: On dit aussi que l'application appelée injection définie par

$$\begin{array}{ccc} \mathcal{I} : H^1(\Omega) & \rightarrow & L^2(\Omega) \\ u & \mapsto & u \end{array}$$

est compacte.

On aurait alors le résultat :

$$\forall C, \exists v \in H^1(\Omega), \|v\|_{L^2(\Omega)} > C [\|\nabla v\|_{[L^2(\Omega)]^d} + \|v\|_{L^2(\partial\Omega)}].$$

On peut choisir en particulier $C = n$, pour $n \in \mathbb{N} \setminus \{0\}$. Ainsi, il existe $v_n \in H^1(\Omega)$ tel que

$$\|v_n\|_{L^2(\Omega)} = 1, \quad \|\nabla v_n\|_{[L^2(\Omega)]^d} + \|v_n\|_{L^2(\partial\Omega)} < \frac{1}{n}. \quad (2.32)$$

Ceci implique que

$$\lim_{n \rightarrow +\infty} \|\nabla v_n\|_{[L^2(\Omega)]^d} = 0 \quad \text{et} \quad \lim_{n \rightarrow +\infty} \|v_n\|_{L^2(\partial\Omega)} = 0. \quad (2.33)$$

La suite $(v_n)_{n \in \mathbb{N}}$ est bornée dans $H^1(\Omega)$, puisque

$$\|v_n\|_{H^1(\Omega)}^2 = \|v_n\|_{L^2(\Omega)}^2 + \|\nabla v_n\|_{[L^2(\Omega)]^d}^2 < 1 + \frac{1}{n^2} < 2.$$

D'après le théorème 2.2.24 de Rellich, il existe une sous-suite extraite de $(v_n)_{n \in \mathbb{N}}$, toujours notée $(v_n)_{n \in \mathbb{N}}$, qui converge fortement dans $L^2(\Omega)$, vers une limite $v \in L^2(\Omega)$. On a bien sûr $\|v\|_{L^2(\Omega)} = 1$.

Est-ce que $v \in H^1(\Omega)$? Que vaut ∇v , pris au sens faible? Soit $\varphi \in [D(\Omega)]^d$. Nous avons :

$$\begin{aligned} (v, \nabla \cdot \varphi)_{L^2(\Omega)} &= \int_{\Omega} v \nabla \cdot \varphi \, d\Omega \\ &= \lim_{n \rightarrow +\infty} \int_{\Omega} v_n \nabla \cdot \varphi \, d\Omega \quad \text{car } v = \lim_{n \rightarrow +\infty} v_n \text{ dans } L^2(\Omega) \\ &= - \lim_{n \rightarrow +\infty} \int_{\Omega} \nabla v_n \cdot \varphi \, d\Omega \quad \text{car } v_n \in H^1(\Omega) \\ &= 0 \quad \text{car } \lim_{n \rightarrow +\infty} \nabla v_n = 0 \text{ dans } [L^2(\Omega)]^d. \end{aligned}$$

Ainsi, v admet un gradient faible et $\nabla v = 0$! Comme $0 \in L^2(\Omega)^N$, v appartient donc à $H^1(\Omega)$. Et puisque son gradient est nul sur Ω , on en déduit que v est donc constante sur chaque composante connexe de Ω . Comme d'une part $\lim_{n \rightarrow +\infty} v_n = v$ dans $L^2(\Omega)$ et d'autre part

$$\lim_{n \rightarrow +\infty} \nabla v_n = 0 = \nabla v \text{ dans } [L^2(\Omega)]^d,$$

on a en fait

$$\lim_{n \rightarrow +\infty} v_n = v \text{ dans } H^1(\Omega).$$

Quel est l'avantage, par rapport à la convergence de $(v_n)_{n \in \mathbb{N}}$ dans $L^2(\Omega)$? L'application trace est continue de $H^1(\Omega)$ dans $L^2(\partial\Omega)$, mais elle n'est pas continue de $L^2(\Omega)$ dans $L^2(\partial\Omega)$.

De la convergence dans $H^1(\Omega)$, on en déduit la convergence des traces $(v_n|_{\partial\Omega})_n$ vers $v|_{\partial\Omega}$ dans $L^2(\partial\Omega)$. D'après (2.33), on en déduit que $v|_{\partial\Omega} = 0$. On sait que v est constante sur chaque composante connexe, d'où finalement $v = 0$ dans Ω . Ceci contredit le fait que $\|v\|_{L^2(\Omega)} = 1$.²² $\square$

22: Avec cette preuve par contradiction, il n'est malheureusement pas possible d'estimer la constante intervenant dans l'inégalité.

Cette démonstration élégante nous permet de démontrer facilement une inégalité de Poincaré généralisée.

Proposition 2.2.25 Soit $\Omega \subset \mathbb{R}^d$ un ouvert borné et $\ell : H^1(\Omega) \rightarrow \mathcal{V}$ une application linéaire continue, où $\mathcal{V}$ est un espace normé, telle que

$$\|\ell(v)\|_{\mathcal{V}}^2 + \|\nabla v\|_{[L^2(\Omega)]^d}^2 = 0 \quad \Rightarrow \quad v = 0.$$

Alors

$$\exists C > 0, \forall v \in H^1(\Omega), \quad \|v\|_{L^2(\Omega)} \leq C \left[\|\nabla v\|_{[L^2(\Omega)]^d} + \|\ell(v)\|_{\mathcal{V}} \right].$$

Remarque 2.2.6 En prenant pour tout $v \in H^1(\Omega)$, $\ell(v) := v|_{\Gamma}$ où Γ est une partie du bord $\partial\Omega$ de mesure non nulle avec Ω connexe, on en déduit

$$\exists C_1 > 0, \forall v \in H^1(\Omega), \quad \|v\|_{L^2(\Omega)} \leq C_1 \left[\|\nabla v\|_{[L^2(\Omega)]^d} + \|v|_{\Gamma}\|_{L^2(\Gamma)} \right].$$

En prenant pour tout $v \in H^1(\Omega)$, $\ell(v) := \int_{\Omega} v \, d\Omega$, on en déduit

$$\exists C_2 > 0, \forall v \in H^1(\Omega), \quad \|v\|_{L^2(\Omega)} \leq C_2 \left[\|\nabla v\|_{[L^2(\Omega)]^d} + \left| \int_{\Omega} v \, d\Omega \right| \right].$$

On déduit de ces inégalités en particulier que

$$\exists C_1 > 0, \forall v \in \left\{ v \in H^1(\Omega), v|_{\Gamma} = 0 \right\}, \quad \|v\|_{L^2(\Omega)} \leq C_1 \|\nabla v\|_{[L^2(\Omega)]^d}.$$

$$\begin{aligned} \exists C_2 > 0, \forall v \in \left\{ v \in H^1(\Omega), \int_{\Omega} v \, d\Omega = 0 \right\}, \\ \|v\|_{L^2(\Omega)} \leq C_2 \|\nabla v\|_{[L^2(\Omega)]^d}. \end{aligned} \quad (2.34)$$

Cette dernière inégalité est connue sous le nom d'inégalité de Poincaré-Wirtinger.

2.3 Lien entre la dérivée au sens faible et la dérivée au sens des distributions

La théorie des distributions mériterait bien entendu un cours à part entière (voir par exemple [4, 5]). Nous allons dans cette section définir les distributions, donner quelques exemples et propriétés. Nous ne détaillerons pas toute la théorie, mais seulement les outils qui nous permettront de définir la dérivée au sens des distributions et de faire le lien avec la dérivée au sens faible de la définition 2.1.1.

Définition 2.3.1 Une distribution T est une forme définie sur $D(\Omega)$ – et on notera pour tout $v \in D(\Omega)$, $T(v) := \langle T, v \rangle_{D'(\Omega), D(\Omega)}$ – qui est linéaire

$$\forall u, v \in D(\Omega), \forall \lambda \in \mathbb{R},$$

$$\langle T, \lambda u + v \rangle_{D'(\Omega), D(\Omega)} = \lambda \langle T, u \rangle_{D'(\Omega), D(\Omega)} + \langle T, v \rangle_{D'(\Omega), D(\Omega)};$$

et continue au sens suivant : pour tout compact $K \subset \Omega$

$$\exists C_K > 0, \exists m_K \in \mathbb{N}, \forall v \in D(K),$$

$$|\langle T, v \rangle_{D'(\Omega), D(\Omega)}| \leq C_K \sup_{|\alpha| \leq m_K} \sup_{x \in K} |\partial^\alpha v(x)|.$$

Une distribution est dite d'ordre fini si la constante m_K peut être choisie indépendamment de K . Dans ce cas, l'ordre est le plus petit m_K pour lequel la continuité est satisfaite.

Donnons maintenant deux exemples de distributions fondamentaux pour ce cours.

Soient $f \in L^2(\Omega)$ et $T_f : v \mapsto (f, v)_{L^2(\Omega)}$ alors $T_f \in D'(\Omega)$. En effet, T_f est une forme linéaire (par linéarité de l'intégrale) qui vérifie : pour tout compact $K \subset \Omega$ et pour tout $v \in D(K)$

$$|\langle T_f, v \rangle_{D'(\Omega), D(\Omega)}| \leq \int_{\Omega} |f v| \, d\Omega \leq \sup_K |v| \int_K |f| \, d\Omega$$

$$\leq |K|^{\frac{1}{2}} \sup_K |v| \|f\|_{L^2(\Omega)}.$$

On en déduit que T_f est une distribution d'ordre 0 et puisque pour

$$v \in D(\Omega), \quad \langle T_f, v \rangle_{D'(\Omega), D(\Omega)} = (f, v)_{L^2(\Omega)},$$

on peut identifier f et T_f en sous-entendant que c'est le produit scalaire avec f qui est la distribution. Nous venons donc de montrer que²³

$$L^2(\Omega) \subset D'(\Omega).$$

23: Les mêmes arguments permettent de montrer que $L^1_{\text{loc}}(\Omega) \subset D'(\Omega)$ où on rappelle que L^1_{loc} est l'ensemble des fonctions localement intégrables sur Ω . C'est en ce sens que les distributions généralisent la notion de fonction.

Un autre exemple fondamental de distribution est une distribution qui n'est pas une fonction. Soit $M \in \Omega$ on définit la distribution notée δ_M appelée mesure (ou masse) de Dirac au point M par

$$\forall v \in D(\Omega), \quad \langle \delta_M, v \rangle_{D'(\Omega), D(\Omega)} = v(M).$$

Il est facile de montrer que c'est bien une distribution et qu'elle est d'ordre 0. Cependant, ce n'est pas une fonction mesurable. La preuve se fait par l'absurde. Supposons qu'il existe une fonction $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ telle que $\delta_M = T_f$ c'est-à-dire

$$\forall v \in D(\mathbb{R}^d), \quad v(M) = \langle \delta_M, v \rangle_{D'(\Omega), D(\Omega)} = \int_{\Omega} f v \, d\Omega.$$

Soit $v \in D(\mathbb{R}^d)$ à support dans la boule centrée en 0 de rayon 1, $B(0, 1)$, telle que $v(0) = 1$ et $|v| \leq 1$. On pose $v_n(x) = v(n(x - M))$ alors v_n est à support dans $B(M, 1/n)$, $v_n(M) = 1$ et $|v_n| \leq 1$. On a pour tout n

$$1 = \int_{\mathbb{R}^d} f v_n \, d\Omega \leq \int_{B(M, 1/n)} |f| \, d\Omega.$$

Comme f est intégrable dans $B(M, 1)$, on a d'après le théorème de

convergence dominée

$$\lim_{n \rightarrow +\infty} \int_{B(M, 1/n)} |f| \, d\Omega = 0,$$

d'où la contradiction.

Il est possible de définir une mesure de Dirac sur toute variété $\Gamma \subset \Omega$ de dimension inférieure à d (par exemple une ligne si $d = 2$, une ligne ou une surface si $d = 3, \dots$) par

$$\forall v \in D(\Omega), \quad \langle \delta_\Gamma, v \rangle_{D'(\Omega), D(\Omega)} = \int_\Gamma v \, d\Gamma.$$

Ces mesures ne sont, pour les mêmes raisons que précédemment, pas des fonctions mesurables.

On sait que certaines fonctions ne sont pas dérivables au sens usuel. Ce n'est pas le cas des distributions. Plus précisément, toute distribution est indéfiniment dérivable au sens suivant.

Définition 2.3.2 Soit $T \in D'(\Omega)$, on note pour tout i , $\frac{\partial T}{\partial x_i}$ la distribution définie par

$$\forall v \in D(\Omega), \quad \left\langle \frac{\partial T}{\partial x_i}, v \right\rangle_{D'(\Omega), D(\Omega)} = - \left\langle T, \frac{\partial v}{\partial x_i} \right\rangle_{D'(\Omega), D(\Omega)}.$$

La dérivation au sens des distributions n'est qu'une extension de la dérivation au sens classique puisque

$$\forall u \in C^1(\Omega) \subset D'(\Omega), \quad v \in D(\Omega), \quad \int_\Omega \frac{\partial u}{\partial x_i} v \, d\Omega = - \int_\Omega u \frac{\partial v}{\partial x_i} \, d\Omega.$$

On retrouve un résultat bien connu pour la dérivée classique.

Proposition 2.3.1 Soit $T \in D'(\Omega)$ telle que toutes ses dérivées au sens des distributions sont nulles alors T est constante sur chaque composante connexe de Ω .

Démonstration. Il suffit de comprendre le cas de la dimension 1 et on remplace Ω par l'intervalle $I =]a, b[$ et on suppose que $T \in D'(I)$ est telle que $T' = 0$. Soit $\varphi \in D(I)$ alors si

$$\int_I \varphi(x) \, dx = 0,$$

la fonction ψ définie par

$$\psi(x) = \int_a^x \varphi(t) \, dt,$$

est bien dans $D(I)$ et $\psi' = \varphi$. Ainsi

$$\langle T, \varphi \rangle_{D'(I), D(I)} = \langle T, \psi' \rangle_{D'(I), D(I)} = \langle -T', \psi \rangle_{D'(I), D(I)} = 0.$$

Considérons maintenant une fonction $\theta \in D(I)$ telle que $\int_I \theta = 1$. Pour

tout $\phi \in D(I)$, on a d'après ce qui précède

$$\langle T, \phi - \theta \int_I \phi \rangle_{D'(I), D(I)} = 0 \quad \Rightarrow \quad \langle T, \phi \rangle_{D'(I), D(I)} = \langle T, \theta \rangle_{D'(I), D(I)} \int_I \phi,$$

ce qui signifie que T est la fonction constante égale à $\langle T, \theta \rangle_{D'(I), D(I)}$. $\square$

La dérivée au sens des distributions est ainsi toujours définie, ce qui n'est pas le cas de la dérivée au sens classique qui nécessite que les fonctions soient dérivables. Nous allons introduire une formule dite des sauts qui donne la dérivée au sens des distributions d'une fonction qui n'est pas dérivable globalement, mais seulement par morceaux.

Theorème 2.3.2 (Formule des sauts) *Soit Ω un ouvert borné de $\mathbb{R}^d$: on le partitionne en $\overline{\Omega} = \overline{\Omega_1} \cup \overline{\Omega_2}$, où Ω_1 et Ω_2 sont deux ouverts disjoints. On note Γ l'interface entre Ω_1 et Ω_2 : $\Gamma := \overline{\Omega_1} \cap \overline{\Omega_2}$.*

Soit v telle que $v|_{\Omega_i} \in C^1(\overline{\Omega_i})$, pour $i = 1, 2$ alors son gradient au sens des distributions est donné par

$$\underline{\nabla} v = \chi_{\Omega_1} \underline{\nabla} v_1 + \chi_{\Omega_2} \underline{\nabla} v_2 - \delta_\Gamma (v_1|_\Gamma - v_2|_\Gamma) \underline{n}_1$$

où χ_{Ω_i} est la fonction indicatrice de Ω_i et $\underline{n}_1$ est la normale extérieure à $\partial\Omega_1$.

Démonstration. On sait pour commencer que $v \in L^2(\Omega) \subset D'(\Omega)$. On va calculer $\underline{\nabla} v$ au sens des distributions. On pose $v_i = v|_{\Omega_i}$ puis on dérive au sens des distributions. Pour toute fonction $\underline{\varphi} \in [D(\Omega)]^d$, on a

$$\begin{aligned} \langle \underline{\nabla} v, \underline{\varphi} \rangle_{[D'(\Omega)]^d, [D(\Omega)]^d} &= \sum_{i=1}^N \left\langle \frac{\partial v}{\partial x_i}, \varphi_i \right\rangle_{D'(\Omega), D(\Omega)} = - \sum_{i=1}^N \left\langle v, \frac{\partial \varphi_i}{\partial x_i} \right\rangle_{D'(\Omega), D(\Omega)} \\ &= - \langle v, \underline{\nabla} \cdot \underline{\varphi} \rangle_{D'(\Omega), D(\Omega)} \\ &= - \int_{\Omega} v \underline{\nabla} \cdot \underline{\varphi} \, d\Omega \quad \text{car } v \in L^2(\Omega) \\ &= - \int_{\Omega_1} v_1 \underline{\nabla} \cdot \underline{\varphi} \, d\Omega - \int_{\Omega_2} v_2 \underline{\nabla} \cdot \underline{\varphi} \, d\Omega. \end{aligned}$$

En utilisant la formule de Green (1.9) et en isolant les intégrales sur le bord, on obtient

$$\begin{aligned} \langle \underline{\nabla} v, \underline{\varphi} \rangle_{[D'(\Omega)]^d, [D(\Omega)]^d} &= \int_{\Omega_1} \underline{\nabla} v_1 \cdot \underline{\varphi} \, d\Omega + \int_{\Omega_2} \underline{\nabla} v_2 \cdot \underline{\varphi} \, d\Omega \\ &\quad - \int_{\partial\Omega_1} v_1 \underline{\varphi} \cdot \underline{n}_1 \, d\Gamma - \int_{\partial\Omega_2} v_2 \underline{\varphi} \cdot \underline{n}_2 \, d\Gamma, \end{aligned}$$

où $\underline{n}_1$ (respectivement $\underline{n}_2$) est la normale extérieure à $\partial\Omega_1$ (respectivement $\partial\Omega_2$). Comme $\underline{\varphi}$ est à support compact elle s'annule sur la frontière et donc puisque $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$, on a

$$\int_{\partial\Omega_1} v_1 \underline{\varphi} \cdot \underline{n}_1 \, d\Gamma + \int_{\partial\Omega_2} v_2 \underline{\varphi} \cdot \underline{n}_2 \, d\Gamma = \int_{\Gamma} (v_1 - v_2) \underline{\varphi} \cdot \underline{n}_1 \, d\Gamma,$$

car $\underline{n}_1 = -\underline{n}_2$ en tout point de Γ et $\underline{\varphi}$ est continue à la traversée de Γ . On

a donc

$$\begin{aligned} \langle \underline{\nabla} v, \underline{\varphi} \rangle_{[D'(\Omega)]^d, [D(\Omega)]^d} &= \int_{\Omega_1} \underline{\nabla} v_1 \cdot \underline{\varphi} \, d\Omega + \int_{\Omega_2} \underline{\nabla} v_2 \cdot \underline{\varphi} \, d\Omega \\ &\quad - \int_{\Gamma} (v_1 - v_2) \underline{\varphi} \cdot \underline{n}_1 \, d\Gamma, \end{aligned}$$

soit

$$\underline{\nabla} v = \chi_{\Omega_1} \underline{\nabla} v_1 + \chi_{\Omega_2} \underline{\nabla} v_2 - \delta_{\Gamma} (v_1|_{\Gamma} - v_2|_{\Gamma}) \underline{n}_1.$$

□

Remarque 2.3.1 Une fonction $f \in L^2(\Omega)$ n'admet pas toujours de dérivée au sens faible alors qu'en tant que distribution, elle en admet toujours au sens des distributions. Elle admet une dérivée au sens faible si et seulement si sa dérivée au sens des distributions est dans $L^2(\Omega)$.

Analyse hilbertienne des problèmes variationnels

3

Ce chapitre couvre un ensemble de résultats fondamentaux de l'analyse hilbertienne en général, dont les applications aux problèmes d'EDP sont nombreuses. Ainsi, grâce à l'introduction des espaces de Sobolev, nous pourrions bénéficier de ces outils pour l'analyse des problèmes aux limites. Il existe également d'autres cadres d'application de ces outils en analyse, notamment autour des problèmes d'optimisation en dimension infinie.

3.1 Introduction

Reprenons donc un des problèmes modèles rencontrés au chapitre 1 en cherchant pour l'instant $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ solution de

$$\begin{cases} -\Delta u + u = f, & \text{dans } \Omega, \\ \nabla u \cdot \underline{n} = 0, & \text{sur } \partial\Omega. \end{cases} \quad (3.1)$$

Le problème est écrit pour l'instant au sens fort. Pour l'écrire au sens faible, il suffit de choisir une fonction test $v \in C^1(\overline{\Omega})$, de multiplier l'équation écrite sur Ω par cette fonction v et d'intégrer sur Ω . On obtient

$$\int_{\Omega} (-\Delta u v + u v) \, d\Omega = \int_{\Omega} f v \, d\Omega.$$

Si on veut diminuer l'ordre de dérivation sur u , il suffit alors d'utiliser la formule de Green du corollaire 1.2.3. En prenant en compte les conditions aux limites, ceci nous donne

$$\forall v \in C^1(\overline{\Omega}), \quad \int_{\Omega} (\nabla u \cdot \nabla v + u v) \, d\Omega = \int_{\Omega} f v \, d\Omega. \quad (3.2)$$

En utilisant la formule de Green dans l'autre sens, il est facile de voir que chercher $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ vérifiant (3.2) est équivalent à chercher $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ vérifiant (3.1).

On vient donc de mettre le problème sous une forme variationnelle

$$\forall v \in \mathcal{V}, \quad a(u, v) = \ell(v) \quad (3.3)$$

C'est cette formulation qui va nous permettre dans ce cours de montrer le caractère bien posé du problème. Rappelons tout d'abord la définition du caractère bien posé d'un problème au sens de Hadamard.

Définition 3.1.1 (Caractère bien posé au sens de Hadamard) *On dit qu'un problème est bien posé dans $\mathcal{V}$, un espace vectoriel normé, au sens de Hadamard si et seulement si*

- **Existence** : il existe au moins une solution dans $\mathcal{V}$;
- **Unicité** : il existe au plus une solution dans $\mathcal{V}$;
- **Stabilité** : la solution dépend de façon continue des données dans le cadre de la topologie de $\mathcal{V}$, ou dit autrement, sa norme dans $\mathcal{V}$ est

contrôlée par les données du problème.

Grâce à cette formulation variationnelle, il est très facile de montrer qu'il ne peut y avoir au plus qu'une solution (autrement dit l'*unicité* de la solution) de (3.2) et donc de (3.1). Supposons que u_1 et u_2 sont solutions de (5.1). En soustrayant les deux formulations, comme le second membre est indépendant de u_1 et u_2 , on trouve

$$\forall v \in C^1(\overline{\Omega}), \quad \int_{\Omega} [\nabla (u_1 - u_2) \cdot \nabla v + (u_1 - u_2)v] \, d\Omega = 0.$$

En particulier pour $v = u_1 - u_2 \in C^1(\overline{\Omega})$, on obtient

$$\int_{\Omega} [|\nabla (u_1 - u_2)|^2 + |u_1 - u_2|^2] \, d\Omega = 0.$$

ce qui nous permet de déduire que $u_1 = u_2$.

1: En dimension infinie (celle de $\mathcal{V}$), même pour un problème linéaire, l'unicité n'implique pas l'existence.

Il n'est pas du tout aussi simple de montrer qu'il existe au moins¹ une solution (autrement dit l'*existence* d'une solution). On verra d'ailleurs qu'il faut absolument rajouter des hypothèses de régularité sur le terme source f ou sur le bord du domaine Ω pour espérer montrer l'existence d'une solution $C^1(\overline{\Omega})$ (voir la section 4.4.2).

Enfin notons que grâce à la formulation variationnelle, il est possible de démontrer un premier résultat de stabilité s'il existe une solution. En effet, supposons qu'il existe une solution u de (5.1) alors en choisissant $v = u$, on obtient

$$\int_{\Omega} [|\nabla u|^2 + |u|^2] \, d\Omega = \int_{\Omega} f u \, d\Omega.$$

En utilisant une inégalité de Cauchy-Schwarz sur le terme de droite, nous obtenons

$$\int_{\Omega} [|\nabla u|^2 + |u|^2] \, d\Omega \leq \|f\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)},$$

ce qui implique

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

et donc

$$\left[\int_{\Omega} [|\nabla u|^2 + |u|^2] \, d\Omega \right]^{\frac{1}{2}} \leq \|f\|_{L^2(\Omega)}.$$

C'est donc cette norme de u qui est naturellement contrôlée par les données et non sa norme C^1 ou C^2 .

Ce sont ces deux considération d'existence et de stabilité qui sous-tendent l'introduction des espaces de Sobolev que nous avons au chapitre 2. En généralisant les formules de Green aux espaces de Sobolev, nous sommes désormais capable de rechercher des solutions $u \in H^2(\Omega) \cap H^1(\Omega)$ du problème (3.1) qui conduisent à la recherche de solution de la formulation variationnelle

$$\forall v \in H^1(\Omega), \quad \int_{\Omega} (\nabla u \cdot \nabla v + u v) \, d\Omega = \int_{\Omega} f v \, d\Omega. \quad (3.4)$$

Désormais l'espace des fonctions tests $\mathcal{V}$ dans (3.3) est plus gros et c'est un espace de Hilbert.

Nous allons maintenant montrer que cette structure hilbertienne (produit scalaire et complétude) est d'une grande utilité pour l'analyse des formulations variationnelles. En effet, les théorèmes les plus importants pour prouver qu'un problème écrit sous sa forme variationnelle est bien posé sont le théorème de Riesz et le théorème de Lax-Milgram, ainsi que les espaces de Hilbert. Ce cadre est parfaitement adapté aux EDP linéaires. Les cas non linéaires sont beaucoup plus délicats, mais certains peuvent s'étudier à travers la minimisation de fonctionnelles, pourvu que l'on dispose de propriétés de continuité, de convexité, voire de compacité amené via la notion de convergence faible.

3.2 Espaces de Hilbert et propriétés élémentaires

Commençons par rappeler ici la définition et les propriétés de ces espaces utiles pour ce cours. Nous nous concentrons ici sur les espaces vectoriels définis sur $\mathbb{R}$.

Commençons par définir un produit scalaire sur un espace vectoriel $\mathcal{H}$.

Définition 3.2.1 Une forme bilinéaire b définie sur $\mathcal{H} \times \mathcal{H}$ est un produit scalaire si et seulement si

- b est symétrique : $\forall (u, v) \in \mathcal{H} \times \mathcal{H}, \quad b(u, v) = b(v, u)$;
- b est positive : $\forall u \in \mathcal{H}, \quad b(u, u) \geq 0$;
- b est définie : $\forall u \in \mathcal{H}, \quad b(u, u) = 0 \Rightarrow u = 0$.

Si b est un produit scalaire alors

$$\begin{aligned} b(u+v, u+v) + b(u-v, u-v) &= 2[b(u, u) + b(v, v)] && \text{(Identité du parallélogramme)} \\ \forall (u, v) \in \mathcal{H} \times \mathcal{H}, \quad |b(u, v)| &\leq [b(u, u)]^{\frac{1}{2}} [b(v, v)]^{\frac{1}{2}} && \text{(Inégalité de Cauchy-Schwarz)} \\ [b(u+v, u+v)]^{\frac{1}{2}} &\leq [b(u, u)]^{\frac{1}{2}} + [b(v, v)]^{\frac{1}{2}} && \text{(Inégalité de Minkowski)} \end{aligned}$$

Il est facile de démontrer ainsi que si b est un produit scalaire alors

$$\mathcal{H} \ni u \mapsto [b(u, u)]^{\frac{1}{2}} \quad \text{est une norme sur } \mathcal{H}.$$

Définition 3.2.2 Un espace de Hilbert est un espace vectoriel muni d'un produit scalaire et qui est complet pour la norme associée.²

Dans la suite pour un espace de Hilbert $\mathcal{H}$, on notera $\|\cdot\|_{\mathcal{H}}$ la norme pour lequel $\mathcal{H}$ est complet et $(\cdot, \cdot)_{\mathcal{H}}$ le produit scalaire dont elle dérive.

Proposition 3.2.1 Soit $\mathcal{F}$ un sous-espace fermé d'un espace de Hilbert $\mathcal{H}$ alors muni de la norme induite par le produit scalaire de $\mathcal{H}$, $\mathcal{F}$ est un espace de Hilbert.

2: Un espace complet est un espace vectoriel normé dans lequel toutes les suites de Cauchy convergent.

Démonstration. Soit une suite de Cauchy d'éléments de $\mathcal{F}$. Comme $\mathcal{F}$ est inclus dans $\mathcal{H}$ et que $\mathcal{H}$ est complet, la suite de Cauchy converge dans $\mathcal{H}$. Comme $\mathcal{F}$ est fermé, la suite converge dans $\mathcal{F}$. $\mathcal{F}$ est donc complet. $\square$

Une des propriétés les plus importantes des espaces de Hilbert est le théorème de projection qui établit l'existence d'un opérateur de projection sur tout sous-ensemble convexe fermé. Rappelons tout d'abord la définition d'un sous-ensemble convexe.

Définition 3.2.3 On dit qu'un sous-ensemble $\mathcal{K}$ de $\mathcal{H}$ est convexe si et seulement si pour tout couple (u, v) d'éléments de $\mathcal{K}$ et pour tout réel t dans $[0, 1]$, $(1 - t)u + tv$ est dans $\mathcal{K}$.

À noter qu'un espace vectoriel est toujours convexe. Nous pouvons maintenant énoncer et démontrer le résultat.

Théorème 3.2.2 (Théorème de projection) Soient $\mathcal{H}$ un espace de Hilbert et $\mathcal{K} \subset \mathcal{H}$ un convexe fermé non vide. Pour tout u dans $\mathcal{H}$, il existe un unique élément de $\mathcal{K}$ que l'on notera Pu tel que

$$\|u - P(u)\|_{\mathcal{H}} = \inf_{v \in \mathcal{K}} \|u - v\|_{\mathcal{H}}.$$

Ce théorème établit donc que l'infimum est atteint par un élément de $\mathcal{K}$ et que cet élément est unique. On appelle P l'opérateur qui à u dans $\mathcal{H}$ associe Pu dans $\mathcal{K}$. Cet opérateur est un opérateur de projection sur $\mathcal{K}$ car il est facile de voir que pour tout $u \in \mathcal{H}$, $P^2 u = Pu$.

Démonstration. Montrons tout d'abord que l'infimum est atteint. Soit $(v_n)_{n \in \mathbb{N}}$ une suite d'éléments de $\mathcal{K}$ telle que³

3: On dit que $(v_n)_{n \in \mathbb{N}}$ est une suite minimisante.

$$\forall n \in \mathbb{N}, \quad \|u - v_n\|_{\mathcal{H}}^2 \leq \inf_{w \in \mathcal{K}} \|u - w\|_{\mathcal{H}}^2 + \frac{1}{n^2}. \quad (3.5)$$

Autrement dit,

$$\forall n \in \mathbb{N}, \forall w \in \mathcal{K}, \quad \|u - v_n\|_{\mathcal{H}}^2 - \|u - w\|_{\mathcal{H}}^2 \leq \frac{1}{n^2}. \quad (3.6)$$

Montrons que $(v_n)_{n \in \mathbb{N}}$ est une suite de Cauchy. L'identité du parallélogramme appliquée à $u - v_n$ et $u - v_m$ donne

$$4\|u - \frac{v_n + v_m}{2}\|_{\mathcal{H}}^2 + \|v_n - v_m\|_{\mathcal{H}}^2 = 2\|u - v_n\|_{\mathcal{H}}^2 + 2\|u - v_m\|_{\mathcal{H}}^2.$$

En réorganisant, on trouve

$$\begin{aligned} \|v_n - v_m\|_{\mathcal{H}}^2 &= 2 \left(\|u - v_n\|_{\mathcal{H}}^2 - \|u - \frac{v_n + v_m}{2}\|_{\mathcal{H}}^2 \right) \\ &\quad + 2 \left(\|u - v_m\|_{\mathcal{H}}^2 - \|u - \frac{v_n + v_m}{2}\|_{\mathcal{H}}^2 \right) \leq \frac{2}{n^2} + \frac{2}{m^2}, \end{aligned}$$

où on a utilisé (3.6) pour établir la dernière inégalité puisque $(v_n + v_m)/2$ est bien un élément de $\mathcal{K}$, $\mathcal{K}$ étant convexe. La suite $(v_n)_{n \in \mathbb{N}}$ est donc une suite de Cauchy dans $\mathcal{K}$. Comme $\mathcal{K}$ est complet en tant que sous-espace fermé d'un espace complet (voir la proposition 3.2.1), $(v_n)_{n \in \mathbb{N}}$ converge dans $\mathcal{K}$ vers $v \in \mathcal{K}$. En passant à la limite dans (3.5), on déduit

$$\|u - v\|_{\mathcal{H}}^2 \leq \inf_{w \in \mathcal{K}} \|u - w\|_{\mathcal{H}}^2.$$

Montrons maintenant que l'infimum est atteint en un seul élément de $\mathcal{K}$.

Supposons que v et v' vérifient

$$\|u - v\|_{\mathcal{H}} = \|u - v'\|_{\mathcal{H}} = d := \inf_{w \in \mathcal{K}} \|u - w\|_{\mathcal{H}}.$$

Par l'identité du parallélogramme appliquée à $u - v$ et $u - v'$ et en réorganisant on obtient

$$\|v - v'\|_{\mathcal{H}}^2 = 2\|u - v\|_{\mathcal{H}}^2 + 2\|u - v'\|_{\mathcal{H}}^2 - 4\|u - \frac{v + v'}{2}\|_{\mathcal{H}}^2 \leq 2d^2 + 2d^2 - 4d^2 = 0,$$

où l'on a utilisé que $\|u - \frac{v + v'}{2}\|_{\mathcal{H}} \geq d$. On a donc $v = v'$. $\square$

Proposition 3.2.3 Soient $\mathcal{H}$ un espace de Hilbert et $\mathcal{K} \subset \mathcal{H}$ un convexe fermé non vide et $P : \mathcal{H} \rightarrow \mathcal{K}$ l'opérateur de projection. Alors on a la caractérisation

$$\forall u \in \mathcal{H}, \forall v \in \mathcal{K}, \quad (u - P(u), v - P(u)) \leq 0. \quad (3.7)$$

Démonstration. Soient $u \in \mathcal{H}$ et $v \in \mathcal{K}$. Comme $P(u) \in \mathcal{K}$, alors pour tout $\lambda \in [0, 1]$, $(1 - \lambda)P(u) + \lambda v \in \mathcal{K}$. Par définition du projecteur

$$\|u - [(1 - \lambda)P(u) + \lambda v]\|_{\mathcal{H}} \geq \|u - P(u)\|_{\mathcal{H}},$$

qui reste vrai en élevant au carré des deux côtés de l'inégalité. On développe

$$\begin{aligned} & \|u - [(1 - \lambda)P(u) + \lambda v]\|_{\mathcal{H}}^2 \\ &= \|u - P(u)\|_{\mathcal{H}}^2 + \lambda^2 \|v - P(u)\|_{\mathcal{H}}^2 - 2\lambda(u - P(u), v - P(u)). \end{aligned}$$

On en déduit

$$\lambda^2 \|v - P(u)\|_{\mathcal{H}}^2 - 2\lambda(u - P(u), v - P(u)) \geq 0,$$

soit

$$2\lambda(u - P(u), v - P(u)) \leq \lambda^2 \|v - P(u)\|_{\mathcal{H}}^2.$$

En divisant par $\lambda \in [0, 1]$ et en passant à la limite en $\lambda \rightarrow 0$, on obtient bien (3.7). $\square$

Le théorème de projection s'applique bien entendu dans le cas particulier où $\mathcal{K}$ est un espace vectoriel fermé de $\mathcal{H}$ car un espace vectoriel est convexe. Dans ce cas, l'opérateur de projection a d'autres propriétés importantes et une caractérisation plus précise.

Proposition 3.2.4 Soient $\mathcal{H}$ un espace de Hilbert, $\mathcal{F}$ un sous-espace vectoriel fermé. L'opérateur P de projection sur $\mathcal{F}$ est un opérateur linéaire de $\mathcal{H}$ dans $\mathcal{H}$ qui est continu c'est-à-dire

$$\forall u \in \mathcal{H}, \quad \|Pu\|_{\mathcal{H}} \leq \|u\|_{\mathcal{H}},$$

et qui est caractérisé par

$$\forall u \in \mathcal{H}, \forall v \in \mathcal{F}, \quad (u - Pu, v) = 0. \quad (3.8)$$

On dit que P est l'opérateur de projection orthogonale.

Démonstration. Commençons par montrer la caractérisation. Soient $v \in \mathcal{F}$ et $\lambda \in \mathbb{R}^+$. Comme $\lambda v \in \mathcal{F}$, on a par définition de Pu ,

$$\|u - Pu \pm \lambda v\|_{\mathcal{H}} \geq \|u - Pu\|_{\mathcal{H}},$$

avec

$$\|u - Pu \pm \lambda v\|_{\mathcal{H}}^2 = \|u - Pu\|_{\mathcal{H}}^2 + \lambda^2 \|v\|_{\mathcal{H}}^2 \pm 2\lambda(u - Pu, v)_{\mathcal{H}}.$$

Ces deux dernières relations impliquent pour $\lambda \neq 0$ que

$$\lambda \|v\|_{\mathcal{H}}^2 \pm 2(u - Pu, v)_{\mathcal{H}} \geq 0.$$

En faisant tendre λ vers 0, on déduit $(u - Pu, v)_{\mathcal{H}} = 0$. Supposons maintenant que $w \in \mathcal{F}$ vérifie pour tout $v \in \mathcal{F}$, $(u - w, v)_{\mathcal{H}} = 0$. On a donc pour tout $v \in \mathcal{F}$, $(u - w, v - w)_{\mathcal{H}} = 0$ puisque $v - w \in \mathcal{F}$. Ceci implique que pour tout $v \in \mathcal{F}$,

$$\|u - v\|^2 = \|u - w\|^2 + \|v - w\|^2,$$

et donc

$$\|u - w\|_{\mathcal{H}} = \inf_{v \in \mathcal{F}} \|u - v\|_{\mathcal{H}}.$$

On a bien $w = Pu$. Ceci prouve que (3.8) caractérise bien l'opérateur de projection.

Pour établir que l'opérateur P est linéaire, on montre grâce à (3.8) et la linéarité du produit scalaire que pour tout $u, w \in \mathcal{H}$ et pour tout $\lambda \in \mathbb{R}$,

$$(u + \lambda w - [Pu + \lambda Pw], v)_{\mathcal{H}} = 0 \quad \forall v \in \mathcal{F}.$$

Comme (3.8) caractérise $P(u + \lambda w)$, ceci prouve que $P(u + \lambda w) = Pu + \lambda Pw$.

Enfin comme $Pu \in \mathcal{F}$, on a $(u - Pu, Pu)_{\mathcal{H}} = 0$ et donc

$$\|u\|_{\mathcal{H}}^2 = \|u - Pu\|_{\mathcal{H}}^2 + \|Pu\|_{\mathcal{H}}^2 \geq \|Pu\|_{\mathcal{H}}^2.$$

On en déduit que P est bien un opérateur linéaire continu sur $\mathcal{H}$. □

Finissons cette section par énoncer quelques résultats liés à la notion d'espace orthogonal.

Définition 3.2.4 Soit $\mathcal{E}$ un sous-espace d'un espace de Hilbert $\mathcal{H}$. On note l'espace orthogonal de $\mathcal{E}$, $\mathcal{E}^\perp := \{u \in \mathcal{H} \mid \forall v \in \mathcal{E}, (u, v)_{\mathcal{H}} = 0\}$.

Il est facile de voir que $\mathcal{E}^\perp$ est un sous-espace fermé de $\mathcal{H}$. En effet, si une suite $(u_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{E}^\perp$ converge vers $u \in \mathcal{H}$ alors $(u - u_n, v)_{\mathcal{H}} = 0$ pour tout $v \in \mathcal{H}$ d'après l'inégalité de Cauchy-Schwarz. Par conséquent, $(u, v)_{\mathcal{H}} = 0$ pour tout v dans $\mathcal{E}$.

Proposition 3.2.5 Soit $\mathcal{F}$ un sous-espace vectoriel fermé de $\mathcal{H}$ alors

$$\mathcal{H} = \mathcal{F} \oplus \mathcal{F}^\perp. \quad (3.9)$$

Soit $\mathcal{E}$ un sous-espace d'un espace de Hilbert $\mathcal{H}$ alors

$$\mathcal{H} = \overline{\mathcal{E}} \oplus \mathcal{E}^\perp. \quad (3.10)$$

Démonstration. Montrons tout d'abord (3.9). Soit $\mathcal{F}$ un sous-espace vectoriel fermé de $\mathcal{H}$. Soit P le projecteur orthogonal sur $\mathcal{F}$. Pour tout u dans $\mathcal{H}$, on a

$$u = \underbrace{Pu}_{\in \mathcal{F}} + \underbrace{u - Pu}_{\in \mathcal{F}^\perp \text{ d'après la proposition 3.2.4}}.$$

De plus, il y a unicité de la décomposition, car si $u_1 + u_2 = 0$ avec $u_1 \in \mathcal{F}$ et $u_2 \in \mathcal{F}^\perp$ alors $0 = \|u_1 + u_2\|_{\mathcal{H}}^2 = \|u_1\|_{\mathcal{H}}^2 + \|u_2\|_{\mathcal{H}}^2$ donc $u_1 = u_2 = 0$.

Établissons maintenant (3.10). Soit $\mathcal{E}$ un sous-espace d'un espace de Hilbert $\mathcal{H}$. Nous venons de montrer que $\mathcal{H} = \overline{\mathcal{E}} \oplus \overline{\mathcal{E}}^\perp$. Prouvons que $\overline{\mathcal{E}}^\perp = \mathcal{E}^\perp$. Tout d'abord il est facile de voir que comme $\mathcal{E} \subset \overline{\mathcal{E}}$ alors $\overline{\mathcal{E}}^\perp \subset \mathcal{E}^\perp$. Montrons que $\mathcal{E}^\perp \subset \overline{\mathcal{E}}^\perp$. Soit $v \in \mathcal{E}^\perp$ et $w \in \overline{\mathcal{E}}$. Comme $w \in \overline{\mathcal{E}}$, il existe une suite $(w_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{E}$ qui tend vers w pour la norme de $\mathcal{H}$. Comme

$$(v, w - w_n)_{\mathcal{H}} \leq \|v\|_{\mathcal{H}} \|w - w_n\|_{\mathcal{H}},$$

on a que $(v, w - w_n)_{\mathcal{H}}$ tend vers 0. De plus $v \in \mathcal{E}^\perp$ et pour tout n , $w_n \in \mathcal{E}$, donc $(v, w_n)_{\mathcal{H}} = 0$. On en déduit que $(v, w)_{\mathcal{H}} = 0$, soit $v \in \overline{\mathcal{E}}^\perp$. En conclusion, on a $\overline{\mathcal{E}}^\perp = \mathcal{E}^\perp$ et donc (3.10). $\square$

Corollaire 3.2.6 Soit $\mathcal{E}$ un sous-espace vectoriel de $\mathcal{H}$. Alors $\mathcal{E}$ est dense⁴ dans $\mathcal{H}$ si et seulement si $\mathcal{E}^\perp = \{0\}$.

4: $\mathcal{E}$ est dense dans $\mathcal{H}$ si et seulement si $\overline{\mathcal{E}} = \mathcal{H}$.

3.3 Le théorème de Représentation de Riesz et ses conséquences

Le théorème de représentation de Riesz est un résultat très important, car il montre qu'un espace de Hilbert est isomorphe à son dual topologique, ce qui permet d'identifier l'espace et son dual. On rappelle que le dual topologique $\mathcal{E}'$ d'un espace normé $\mathcal{E}$ est l'ensemble des formes linéaires $\ell : \mathcal{E} \rightarrow \mathbb{R}$ continues sur $\mathcal{E}$, c'est-à-dire telles que,

$$\exists C > 0, \forall u \in \mathcal{E}, \quad |\ell(u)| \leq C \|u\|_{\mathcal{E}}.$$

On munit cet espace de la norme duale

$$\forall \ell \in \mathcal{E}', \quad \|\ell\|_{\mathcal{E}'} := \sup_{v \in \mathcal{E}} \frac{|\ell(v)|}{\|v\|_{\mathcal{E}}}.$$

Théorème 3.3.1 Soit $\mathcal{V}$ un espace de Hilbert et $\ell \in \mathcal{V}'$, alors il existe un unique $u \in \mathcal{V}$ tel que

$$\forall v \in \mathcal{V}, \quad \ell(v) = (u, v)_{\mathcal{V}}.$$

De plus,

$$\|\ell\|_{\mathcal{V}'} = \|u\|_{\mathcal{V}}.$$

On appelle u le représentant de Riesz de ℓ .

Démonstration. Il est facile de montrer que si un tel u existe, il est unique. En effet, si u et u' sont des représentants de Riesz alors

$$\forall v \in \mathcal{V}, \quad (u, v)_{\mathcal{V}} = (u', v)_{\mathcal{V}} \Rightarrow \forall v \in \mathcal{V}, \quad (u - u', v)_{\mathcal{V}} = 0,$$

ce qui implique que $u = u'$.

Supposons qu'un tel u existe alors, d'une part,

$$\ell(u) = \|u\|_{\mathcal{V}}^2 \leq \|\ell\|_{\mathcal{V}'} \|u\|_{\mathcal{V}} \Rightarrow \|u\|_{\mathcal{V}} \leq \|\ell\|_{\mathcal{V}'},$$

et d'autre part d'après l'inégalité de Cauchy-Schwarz

$$\forall v \in \mathcal{V}, \quad \ell(v) = (u, v)_{\mathcal{V}} \leq \|u\|_{\mathcal{V}} \|v\|_{\mathcal{V}} \Rightarrow \|\ell\|_{\mathcal{V}'} \leq \|u\|_{\mathcal{V}}.$$

Ainsi $\|\ell\|_{\mathcal{V}'} = \|u\|_{\mathcal{V}}$.

Montrons maintenant l'existence d'un représentant de Riesz. Notons que si $\ell = 0$ le résultat est immédiat avec $u = 0$. Supposons maintenant que $\ell \neq 0$. On introduit⁵ le sous-espace de $\mathcal{V}$:

$$\mathcal{E} := \ker \ell = \{v \in \mathcal{V}, \ell(v) = 0\}.$$

Comme ℓ est continu, $\mathcal{E}$ est un sous-espace fermé de $\mathcal{V}$. Puisque $\ell \neq 0$, $\mathcal{E}^\perp$ n'est pas réduit à 0 et on peut choisir un élément non nul $w \in \mathcal{E}^\perp$, qui vérifie donc $\ell(w) \neq 0$. En renormalisant, il est possible de choisir w tel que $\ell(w) = 1$. Il est facile de voir que

$$\forall v \in \mathcal{V}, \quad v = \underbrace{\ell(v)w}_{\in \mathcal{E}^\perp} + \underbrace{v - \ell(v)w}_{\in \mathcal{E}},$$

ce qui implique que

$$\forall v \in \mathcal{V}, \quad (w, v)_{\mathcal{V}} = \ell(v) \|w\|_{\mathcal{V}}^2 \Rightarrow \forall v \in \mathcal{V}, \quad \ell(v) = \left(\frac{w}{\|w\|_{\mathcal{V}}^2}, v \right)_{\mathcal{V}}.$$

Donc $u = \frac{1}{\|w\|_{\mathcal{V}}^2} w$ convient. $\square$

Un corollaire du théorème de représentation de Riesz est qu'il est possible d'associer à toute forme bilinéaire continue un opérateur continu.

Corollaire 3.3.2 Soit $\mathcal{V}$ un espace de Hilbert et a une forme bilinéaire $a : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ continue c'est-à-dire qu'elle vérifie

$$\exists C_a > 0, \quad \forall (u, v) \in \mathcal{V} \times \mathcal{V}, \quad |a(u, v)| \leq C_a \|u\|_{\mathcal{V}} \|v\|_{\mathcal{V}}.$$

5: en effet, si un tel représentant de Riesz u existe alors pour tout v tel que $\ell(v) = 0$ on a $(u, v)_{\mathcal{V}} = 0$. Ceci implique qu'un tel u est nécessairement dans l'orthogonal du noyau de ℓ .

Alors il existe $A \in \mathcal{L}(\mathcal{V})$ tel que

$$\forall (u, v) \in \mathcal{V} \times \mathcal{V}, \quad a(u, v) = (Au, v)_{\mathcal{V}}.$$

Démonstration. Fixons $u \in \mathcal{V}$. Pour tout v dans $\mathcal{V}$, l'application $v \mapsto a(u, v)$ est linéaire continue. D'après le théorème de représentation de Riesz, il existe un élément de $\mathcal{V}$ qui dépend de u que nous notons donc $A(u)$ tel que

$$\forall v \in \mathcal{V}, \quad a(u, v) = (A(u), v)_{\mathcal{V}}, \quad (3.11)$$

ceci étant vrai pour tout u dans $\mathcal{V}$.

Montrons que l'application $A : \mathcal{V} \rightarrow \mathcal{V}$ ainsi construite est linéaire et continue. Soit $u, w \in \mathcal{V}$, $\lambda \in \mathbb{R}$, on a

$$\begin{aligned} \forall v \in \mathcal{V}, \quad (A(\lambda u + w), v)_{\mathcal{V}} &= a(\lambda u + w, v) \\ &= \lambda a(u, v) + a(w, v) = (\lambda A(u) + A(w), v)_{\mathcal{V}}, \end{aligned}$$

où on a utilisé la linéarité de a par rapport à la première composante. Ceci nous permet de déduire que $A(\lambda u + w) = \lambda A(u) + A(w)$ donc A est linéaire⁶. Pour la continuité de A , il suffit de prendre dans (3.11) $v = Au$ et d'utiliser la continuité de a :

$$(Au, Au)_{\mathcal{V}} = a(u, Au) \leq C_a \|u\|_{\mathcal{V}} \|Au\|_{\mathcal{V}}.$$

On a donc pour tout $u \in \mathcal{V}$, $\|Au\|_{\mathcal{V}} \leq C_a \|u\|_{\mathcal{V}}$. □

6: Nous avons de nouveau utilisé ici la propriété suivante : soit $u \in \mathcal{V}$,

$$\forall v \in \mathcal{V}, (u, v)_{\mathcal{V}} = 0 \Rightarrow (u, u)_{\mathcal{V}} = 0 \Rightarrow u = 0.$$

Ceci implique en particulier que si $u_1, u_2 \in \mathcal{V}$ vérifie $(u_1, v)_{\mathcal{V}} = (u_2, v)_{\mathcal{V}}$ pour tout $v \in \mathcal{V}$ alors $u_1 = u_2$.

3.4 Convergence faible

La convergence faible dans les espaces de Hilbert joue un rôle fondamental. Plus souple que la convergence forte (en norme) elle permet également de généraliser les principaux théorèmes de la dimension finie.

Définition 3.4.1 (Convergence faible) *On dit qu'une suite $(v_n)_{n \in \mathbb{N}}$ d'éléments d'un espace de Hilbert $\mathcal{V}$ converge faiblement vers $v_{\infty} \in \mathcal{V}$ si*

$$\forall v \in \mathcal{V}, \quad \lim_{n \rightarrow +\infty} (v_n, v)_{\mathcal{V}} = (v_{\infty}, v)_{\mathcal{V}}.$$

On note dans ce cas $v_n \xrightarrow[n \rightarrow \infty]{} v_{\infty}$.

La convergence faible est plus générale que la convergence forte.

Proposition 3.4.1 *Soit $(v_n)_{n \in \mathbb{N}}$ une suite d'éléments d'un espace de Hilbert $\mathcal{V}$ qui converge fortement vers $v_{\infty} \in \mathcal{V}$. Alors $(v_n)_{n \in \mathbb{N}}$ converge faiblement vers v_{∞} .*

Démonstration. C'est une application de l'inégalité de Cauchy-Schwarz. En effet pour tout $v \in \mathcal{V}$, on a

$$|(v_n - v_{\infty}, v)_{\mathcal{V}}| \leq \|v_n - v_{\infty}\|_{\mathcal{V}} \|v\|_{\mathcal{V}},$$

qui tend vers 0 lorsque n tend vers $+\infty$. $\square$

7: aussi appelé le *Uniform Boundedness Principle* en anglais ce qui est plus explicite

Commençons par rappeler le théorème fondamental de Banach-Steinhaus⁷.

Théorème 3.4.2 (Banach-Steinhaus) Soient $(\mathcal{E}, \|\cdot\|_{\mathcal{E}})$ et $(\mathcal{F}, \|\cdot\|_{\mathcal{F}})$ des espaces vectoriels normés complets (des Banach) et $(T_n)_{n \in \mathbb{N}}$ une suite de $\mathcal{L}(\mathcal{E}, \mathcal{F})$. Si,

$$\forall x \in \mathcal{E}, \quad \sup_n \{\|T_n x\|_{\mathcal{F}}\} < \infty,$$

alors, on a

$$\sup_n \{\|T_n\|_{\mathcal{L}(\mathcal{E}, \mathcal{F})}\} < \infty.$$

Le théorème de Banach-Steinhaus nous permet de démontrer un résultat très utile.

Proposition 3.4.3 Toute suite faiblement convergente est bornée.

Démonstration. On considère une suite $(v_n)_{n \in \mathbb{N}}$ faiblement convergente dans $\mathcal{V}$. On considère aussi la forme linéaire ℓ_n définie sur $\mathcal{V}$ par

$$\forall v \in \mathcal{V}, \quad \ell_n(v) = (v_n, v)_{\mathcal{V}}.$$

Comme $(v_n)_{n \in \mathbb{N}}$ est faiblement convergente, pour tout v dans $\mathcal{V}$ la suite $(\ell_n(v))_{n \in \mathbb{N}}$ est convergente, donc elle est bornée. Par application du théorème de Banach-Steinhaus, on obtient que la suite $(\ell_n)_{n \in \mathbb{N}}$ est bornée dans $\mathcal{V}' = \mathcal{L}(\mathcal{V}, \mathbb{R})$. Enfin, comme on a

$$\|\ell_n\|_{\mathcal{V}'} = \sup_{v \in \mathcal{V} \setminus \{0\}} \frac{|\ell_n(v)|}{\|v\|_{\mathcal{V}}} = \|v_n\|_{\mathcal{V}},$$

on en déduit que $(v_n)_{n \in \mathbb{N}}$ est bornée dans $\mathcal{V}$. $\square$

Plus généralement, on a le théorème dit de convergence forte-faible suivant.

Théorème 3.4.4 Soient $(u_n)_{n \in \mathbb{N}}$ et $(v_n)_{n \in \mathbb{N}}$ deux suites telles que

$$u_n \xrightarrow[n \rightarrow +\infty]{} u_{\infty} \quad (\text{fortement quand } n \rightarrow +\infty),$$

$$v_n \xrightarrow[n \rightarrow +\infty]{} v_{\infty} \quad (\text{faiblement quand } n \rightarrow +\infty).$$

Alors

$$\lim_{n \rightarrow +\infty} (u_n, v_n)_{\mathcal{V}} = (u_{\infty}, v_{\infty})_{\mathcal{V}}.$$

Démonstration. Le théorème serait évident si l'on avait une convergence forte plutôt que faible pour la suite $(v_n)_{n \in \mathbb{N}}$. On écrit

$$(u_n, v_n)_{\mathcal{V}} - (u_{\infty}, v_{\infty})_{\mathcal{V}} = (u_n - u_{\infty}, v_n)_{\mathcal{V}} + (u_{\infty}, v_n - v_{\infty})_{\mathcal{V}}.$$

Le deuxième terme tend vers 0 par convergence faible. Pour le premier, on utilise le fait que $(v_n)_{n \in \mathbb{N}}$ convergeant faiblement est bornée grâce, à

la proposition 3.4.3, par une constante C . L'inégalité de Cauchy-Schwarz fournit alors

$$|(u_n - u_\infty, v_n)_V| \leq C \|u_n - u_\infty\|_V \rightarrow 0 \text{ quand } n \rightarrow +\infty.$$

□

Par passage à la limite faible, on peut éventuellement “perdre de la norme” comme le montre la proposition suivante.

Proposition 3.4.5 Soit $(u_n)_{n \in \mathbb{N}}$ une suite qui converge faiblement vers u_∞ dans un espace de Hilbert V . Alors⁸

- $\|u_\infty\|_V \leq \liminf_{n \rightarrow +\infty} \|u_n\|_V$.
- Si de plus $\|u_\infty\|_V = \lim_{n \rightarrow +\infty} \|u_n\|_V$, alors $(u_n)_{n \in \mathbb{N}}$ converge fortement vers u_∞ .

$$8: \liminf_{n \rightarrow \infty} \|u_n\| := \lim_{n \rightarrow \infty} \left(\inf_{m \geq n} \|u_m\| \right)$$

Démonstration. On développe

$$\|u_n - u_\infty\|_V^2 = \|u_n\|_V^2 - 2(u_n, u_\infty)_V + \|u_\infty\|_V^2.$$

Or comme $u_n \rightharpoonup u_\infty$ faiblement, on a

$$\lim_{n \rightarrow +\infty} (u_n, u_\infty)_V = \|u_\infty\|_V^2,$$

dont on déduit

$$0 \leq \liminf_{n \rightarrow +\infty} \|u_n - u_\infty\|_V^2 = \liminf_{n \rightarrow +\infty} \|u_n\|_V^2 - \|u_\infty\|_V^2,$$

et le premier résultat en découle. Pour le deuxième, l'hypothèse et le calcul précédent permettent de déduire

$$\lim_{n \rightarrow +\infty} \|u_n - u_\infty\|_V^2 = 0.$$

□

Enfin, nous terminons par un théorème fondamental. Il précise que les bornés sont compacts pour la convergence faible⁹.

Théorème 3.4.6 Soit $(v_n)_{n \in \mathbb{N}}$ une suite bornée de V . On peut extraire de $(v_n)_{n \in \mathbb{N}}$ une sous-suite qui converge faiblement.

9: généralisant ainsi ce que l'on avait en dimension finie où les compacts sont les fermés bornés

Ainsi, les suites qui convergent faiblement sont bornées et des suites bornées on peut extraire une sous-suite qui converge faiblement. Cette propriété fondamentale permettra, à partir d'une borne, de trouver une limite potentielle. Nous l'utiliserons à la fin de ce chapitre et au chapitre suivant quand nous étudierons directement l'existence de minimiseurs des fonctionnelles, avec des généralisations possibles à des fonctionnelles non quadratiques. Mais avant, les résultats du paragraphe suivant vont nous permettre de traiter tous les cas de fonctionnelles quadratiques et plus généralement de nombreuses formulations variationnelles. Nous renvoyons à [6] pour plus de détails sur la convergence faible.

3.5 Lax-Milgram

3.5.1 Le théorème de Lax-Milgram

Nous pouvons maintenant énoncer et démontrer le théorème de Lax-Milgram qui est la base des résultats d'existence et d'unicité de formulations variationnelles.

Théorème 3.5.1 (Lax-Milgram) *Soient*

- $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ un espace de Hilbert,
- $a : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ une forme bilinéaire continue

$$\exists C_a > 0, \forall (u, v) \in \mathcal{V} \times \mathcal{V}, \quad |a(u, v)| \leq C_a \|u\|_{\mathcal{V}} \|v\|_{\mathcal{V}}, \quad (3.12)$$

et coercive :

$$\exists \alpha > 0, \forall v \in \mathcal{V}, \quad a(v, v) \geq \alpha \|v\|_{\mathcal{V}}^2. \quad (3.13)$$

- $\ell : \mathcal{V} \rightarrow \mathbb{R}$ une forme linéaire continue

$$\exists C_{\ell} > 0, \forall v \in \mathcal{V}, \quad |\ell(v)| \leq C_{\ell} \|v\|_{\mathcal{V}}. \quad (3.14)$$

Alors il existe une unique solution dans $\mathcal{V}$ au problème variationnel suivant : Trouver $u \in \mathcal{V}$ tel que

$$\forall v \in \mathcal{V}, \quad a(u, v) = \ell(v). \quad (3.15)$$

De plus cette solution vérifie

$$\|u\|_{\mathcal{V}} \leq \frac{C_{\ell}}{\alpha}. \quad (3.16)$$

Remarque 3.5.1 La condition de coercivité sur a est essentielle dans ce théorème. La constante α est appelée constante de coercivité. Notons que si a est coercive alors elle est définie positive, c'est-à-dire

$$\forall v \in \mathcal{V} \setminus \{0\}, \quad a(v, v) > 0.$$

On peut montrer qu'il y a équivalence si $\mathcal{V}$ est un espace de dimension finie. Mais si $\mathcal{V}$ n'est pas de dimension finie, ce n'est plus vrai et l'hypothèse de coercivité sur a dans le théorème de Lax-Milgram ne peut pas être remplacée par une hypothèse sur son caractère défini positif.

Si la forme bilinéaire est négative alors il faut tenter de montrer que $-a$ est coercive.

S'il existe un $v \in \mathcal{V}$ tel que $a(v, v) = 0$ alors il est clair que a n'est pas coercive. De plus s'il existe une suite $(v_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{V}$ telle que

$$\lim_{n \rightarrow +\infty} \frac{a(v_n, v_n)}{\|v_n\|^2} = 0,$$

alors a n'est pas coercive non plus.

Démonstration. D'après le théorème de représentation de Riesz (voir le

théorème 3.3.1), comme ℓ est une forme linéaire continue sur $\mathcal{V}$, il existe $F \in \mathcal{V}$ telle que

$$\forall v \in \mathcal{V}, \quad \ell(v) = (F, v)_{\mathcal{V}}.$$

D'après le corollaire 3.3.2, comme a est une forme bilinéaire continue, il existe un opérateur $A : \mathcal{V} \rightarrow \mathcal{V}$ tel que

$$\forall (u, v) \in \mathcal{V} \times \mathcal{V}, \quad a(u, v) = (Au, v)_{\mathcal{V}}.$$

Le problème variationnel : trouver $u \in \mathcal{V}$ tel que (3.15) est vérifié est donc équivalent à :

$$\text{Trouver } u \in \mathcal{V} \text{ tel que } Au = F.$$

L'unicité d'une telle solution de (3.15) est donc équivalente à l'injectivité de l'opérateur A et l'existence est donc équivalente à la surjectivité de A .

Montrons tout d'abord que la coercivité de a implique l'injectivité de A et donc l'unicité d'une solution :

$$\forall v \in \mathcal{V}, \quad \alpha \|v\|_{\mathcal{V}}^2 \underset{(1)}{\leq} a(v, v) = (Av, v)_{\mathcal{V}} \underset{(2)}{\leq} \|Av\|_{\mathcal{V}} \|v\|_{\mathcal{V}},$$

où (1) est dû à la coercivité de a et (2) est une inégalité de Cauchy-Schwarz dans $\mathcal{V}$. Ceci implique que

$$\forall v \in \mathcal{V}, \quad \|Av\|_{\mathcal{V}} \geq \alpha \|v\|_{\mathcal{V}}, \quad (3.17)$$

L'opérateur A est donc injectif.

Pour montrer que A est surjectif, commençons par montrer que $\text{Im } A$ est fermé. Soit $(v_n)_{n \in \mathbb{N}}$ une suite d'éléments de $\text{Im } A$ qui converge dans $\mathcal{V}$. On appelle $v \in \mathcal{V}$ sa limite. Comme c'est une suite d'éléments de $\text{Im } A$, il existe une suite $(u_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{V}$ telle que pour tout n , $v_n = Au_n$. D'après (3.17), on a

$$\forall n < p, \quad \|v_n - v_p\|_{\mathcal{V}} \geq \alpha \|u_n - u_p\|_{\mathcal{V}}.$$

Comme $(v_n)_{n \in \mathbb{N}}$ converge c'est une suite de Cauchy et l'inégalité ci-dessus implique que $(u_n)_{n \in \mathbb{N}}$ est également une suite de Cauchy dans $\mathcal{V}$. Comme $\mathcal{V}$ est un espace de Hilbert, $(u_n)_{n \in \mathbb{N}}$ converge et on appelle $u \in \mathcal{V}$ sa limite. A étant un opérateur linéaire continu, on a :

$$\|Au_n - Au\|_{\mathcal{V}} = \|A(u_n - u)\|_{\mathcal{V}} \leq C_a \|u_n - u\|_{\mathcal{V}}.$$

$(v_n = Au_n)_{n \in \mathbb{N}}$ converge donc vers Au et par unicité de la limite $Au = v$. La limite v est donc dans $\text{Im } A$. $\text{Im } A$ est donc fermé.

Comme $\text{Im } A$ est un sous-espace vectoriel fermé de $\mathcal{V}$, d'après la proposition 3.2.5, on a $\mathcal{V} = \text{Im } A \oplus [\text{Im } A]^{\perp}$. Soit $v \in [\text{Im } A]^{\perp}$, on a d'après la coercivité de a

$$\alpha \|v\|_{\mathcal{V}}^2 \leq a(v, v) = (Av, v)_{\mathcal{V}} = 0,$$

puisque v est dans l'orthogonal de $\text{Im } A$. Ceci implique donc que $v = 0$. L'espace $[\text{Im } A]^{\perp}$ est donc réduit à $\{0\}$. On en déduit $\text{Im } A = \mathcal{V}$. A est donc surjectif.

Il ne reste plus qu'à montrer le résultat de stabilité (3.16). Pour cela, il suffit de prendre $v = u$ dans (3.15) ce qui donne

$$\alpha \|u\|_{\mathcal{V}}^2 \underset{a \text{ est coercive}}{\leq} a(u, u) = \ell(u) \underset{\ell \text{ est continue}}{\leq} C_{\ell} \|u\|_{\mathcal{V}}.$$

On retrouve bien le résultat de stabilité énoncé. $\square$

Remarque 3.5.2 (Une autre preuve du théorème de Lax-Milgram) Cette preuve commence comme la précédente avec l'introduction de l'élément $F \in \mathcal{V}$ représentant de Riesz de ℓ et l'opérateur A . On montre que l'application

$$T : w \in \mathcal{V} \mapsto w - \rho(Aw - F),$$

avec $0 < \rho < 2\alpha/C_a^2$ est une application contractante. En effet, on a pour tout v, w dans $\mathcal{V}$

$$\begin{aligned} \|T(v) - T(w)\|_{\mathcal{V}}^2 &= \|v - w - \rho A(v - w)\|_{\mathcal{V}}^2 \\ &= \|v - w\|_{\mathcal{V}}^2 - 2\rho a(v - w, v - w) \\ &\quad + \rho^2 \|A(v - w)\|_{\mathcal{V}}^2 \\ &\leq (1 - 2\rho\alpha + \rho^2 C_a^2) \|v - w\|_{\mathcal{V}}^2 \\ &< \beta \|v - w\|_{\mathcal{V}}^2 \quad \text{avec } 0 < \beta < 1. \end{aligned}$$

D'après le théorème de point fixe de Picard¹⁰, il existe une unique solution u dans $\mathcal{V}$ de $T(u) = u$, c'est à dire de $Au = F$.

10: Théorème du point fixe de Picard : Soit T une application linéaire sur un espace de Hilbert $\mathcal{V}$ qui est contractante alors T a un unique point fixe, c'est à dire un unique $u \in \mathcal{V}$ tel que $T(u) = u$.

Preuve : Soit la suite $(u_n)_{n \in \mathbb{N}}$ définie pour tout n par $u_{n+1} = T(u_n)$. comme T est contractante, cette suite est de Cauchy dans $\mathcal{V}$ et comme $\mathcal{V}$ est complet, cette suite converge vers une fonction $u \in \mathcal{V}$. En passant à la limite dans $u_{n+1} = T(u_n)$, comme T est continue, on déduit que $T(u) = u$. Il n'existe qu'un seul point fixe car T est contractante.

Enfin, établissons maintenant l'équivalence entre la résolution d'une formulation variationnelle et la minimisation d'une énergie quand l'espace de travail est un espace de Hilbert.

Proposition 3.5.2 Supposons que les hypothèses du théorème de Lax-Milgram (voir théorème 3.5.1) sont vérifiées et que la forme bilinéaire est symétrique, c'est-à-dire

$$\forall (u, v) \in \mathcal{V}^2, \quad a(u, v) = a(v, u). \quad (3.18)$$

Alors la fonctionnelle $\mathcal{J}$ définie par

$$\mathcal{J} : \mathcal{V} \ni u \mapsto \frac{1}{2}a(u, u) - \ell(u) \in \mathbb{R},$$

admet un unique minimiseur $u = \operatorname{argmin}_{\mathcal{V}} \mathcal{J}$.

Démonstration. Notons tout d'abord que pour tout u, v dans $\mathcal{V}$, on a

$$\begin{aligned} \mathcal{J}(u + v) - \mathcal{J}(u) &= \frac{1}{2}a(u + v, u + v) - \ell(u + v) - \frac{1}{2}a(u, u) - \ell(u) \\ &= \frac{1}{2}(a(u, v) + a(v, u)) - \ell(v) + \frac{1}{2}a(v, v) \\ &= a(u, v) - \ell(v) + \frac{1}{2}a(v, v) \quad \text{car } a \text{ vérifie (3.18)}. \end{aligned}$$

Comme a vérifie (3.12), on a

$$\forall v \in \mathcal{V}, \quad a(v, v) = o(\|v\|_{\mathcal{V}}).$$

On en déduit que $\mathcal{J}$ est différentiable et que la différentielle de $\mathcal{J}$ en u , notée $D\mathcal{J}(u)$, est donnée par

$$\forall v \in \mathcal{V}, \quad \langle D\mathcal{J}(u), v \rangle_{\mathcal{V}', \mathcal{V}} = a(u, v) - \ell(v).$$

Ainsi si u est l'unique solution de (3.15) alors $D\mathcal{J}(u) = 0$ et le calcul précédent donne

$$\mathcal{J}(u + v) - \mathcal{J}(u) = \frac{1}{2}a(v, v) > 0 \quad \text{car } a \text{ vérifie (3.13).}$$

On en déduit que u minimise $\mathcal{J}$. Si u minimise $\mathcal{J}$ alors comme on vient de montrer que $\mathcal{J}$ est différentiable, on a

$$\forall v \in \mathcal{V}, \quad \langle D\mathcal{J}(u), v \rangle_{\mathcal{V}', \mathcal{V}} = 0,$$

et donc u vérifie (3.15). Comme il n'existe qu'une solution de (3.15), il n'existe qu'un minimiseur. $\square$

Remarque 3.5.3 Nous devons insister sur le fait que la plupart des exemples que nous avons rencontrés correspondent à des formes bilinéaires symétriques. Dans ce cas, le théorème de Riesz suffit à montrer l'existence, l'unicité et la continuité par rapports aux données. Le théorème de Lax-Milgram permet de traiter les cas non symétriques

3.6 Approximation interne générale

Revenons au cas des formulations variationnelles du type (3.3) (où la fonctionnelle éventuellement associée est quadratique) et anticipons sur les objectifs d'ingénierie qui consiste à trouver des stratégies pour résoudre numériquement le problème variationnel (3.3). *L'approximation interne* de (3.3) consiste à remplacer l'espace de Hilbert $\mathcal{V}$ par un sous-espace de dimension finie $\mathcal{V}_N \subset \mathcal{V}$ muni de la norme de $\mathcal{V}$ où on cherche la solution de

$$\text{trouver } u_N \in \mathcal{V}_N \text{ tel que } a(u_N, v_N) = \ell(v_N) \quad \forall v_N \in \mathcal{V}_N. \quad (3.19)$$

Tout la difficulté est de trouver un tel sous-espace de dimension finie. Cependant, le caractère bien posé de cette formulation approchée (3.19) est ensuite immédiate comme le montre le paragraphe suivant.

3.6.1 Caractère bien posé et propriétés du système linéaire

Proposition 3.6.1 Soit $\mathcal{V}$ un espace de Hilbert réel, et $\mathcal{V}_N \subset \mathcal{V}$ un sous-espace de dimension finie. Soit $a : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ une forme bilinéaire continue et coercive sur $\mathcal{V} \times \mathcal{V}$, et $\ell : \mathcal{V} \rightarrow \mathbb{R}$ une forme linéaire continue sur $\mathcal{V}$. Alors l'approximation interne (3.19) admet une unique solution.

Démonstration. L'existence et l'unicité de $u_N \in \mathcal{V}_N$, solution de (3.19), découlent directement du théorème 3.5.1 de Lax-Milgram appliqué à $\mathcal{V}_N$. En effet $\mathcal{V}_N$ muni de la norme de $\mathcal{V}$ est un espace de Hilbert comme

sous-espace fermé de $\mathcal{V}$ (car de dimension finie). De plus les propriétés de coercivité de a et de continuité de a et ℓ sont directement vraies dans l'espace $\mathcal{V}_N$ puisqu'elles le sont dans $\mathcal{V}$ et que $\mathcal{V}_N \subset \mathcal{V}$. $\square$

Introduisons maintenant une base $(\phi_j)_{1 \leq j \leq N}$ de $\mathcal{V}_N$. Le problème (3.19) est équivalent à :

$$\text{Trouver } u_N \in \mathcal{V}_N \text{ tel que } a(u_N, \phi_i) = \ell(\phi_i) \quad \forall 1 \leq i \leq N. \quad (3.20)$$

En effet, il est évident que si u_N vérifie (3.19) alors en prenant $v_N = \phi_i$ pour i entre 1 et N , on obtient que u_N vérifie (3.20). Réciproquement supposons que u_N vérifie (3.20). Soit $v_N \in \mathcal{V}_N$, alors v_N se décompose dans la base $(\phi_j)_{1 \leq j \leq N}$: $v_N = \sum_{i=1}^N v^i \phi_i$. En multipliant la i -ième équation de (3.20) par v^i , en sommant toutes les équations et en utilisant la linéarité de $a(u_N, \cdot)$ et de ℓ , on retrouve bien que u_N vérifie (3.19).

Comme $u_N \in \mathcal{V}_N$, on a $u_N = \sum_{j=1}^N u^j \phi_j$. Posons $\vec{u}_N := (u^1, \dots, u^N)$ le vecteur dans $\mathbb{R}^N$ des degrés de liberté de u_N . Le problème (3.20) se réécrit

$$\text{Trouver } \vec{u}_N \in \mathbb{R}^N \text{ tel que } a\left(\sum_{j=1}^N u^j \phi_j, \phi_i\right) = \ell(\phi_i), \quad \forall 1 \leq i \leq N,$$

qui est strictement équivalent à la résolution du système linéaire

$$\mathbb{A}_N \vec{u}_N = \vec{\ell}_N, \quad (3.21)$$

avec, pour $1 \leq i, j \leq N$,

$$(\mathbb{A}_N)_{ij} = a(\phi_j, \phi_i), \quad (\vec{\ell}_N)_i = \ell(\phi_i).$$

La coercivité de la forme bilinéaire $a(u, v)$ entraîne le caractère défini positif de la matrice $\mathbb{A}_N$. En effet, pour tout vecteur $\vec{v}_N \in \mathbb{R}^N$, en posant

$$v_N = \sum_{j=1}^N v^j \phi_j \quad \text{où} \quad \vec{v}_N := (v^1, \dots, v^N),$$

on a

$$\vec{v}_N^\top \mathbb{A}_N \vec{v}_N = a(v_N, v_N) \geq \nu \left\| \sum_{j=1}^N v^j \phi_j \right\|_{\mathcal{V}}^2 \geq C |\vec{v}_N|^2 \quad \text{avec } C > 0,$$

car toutes les normes sont équivalentes en dimension finie ($|\cdot|$ désignant la norme euclidienne dans $\mathbb{R}^N$).

Remarque 3.6.1 On retrouve que le problème matriciel (3.21) qui est équivalent au problème (3.19) possède une unique solution puisque la matrice $\mathbb{A}_N$ est définie positive donc inversible.

De plus, si $a(u, v)$ est symétrique alors $\mathbb{A}_N$ l'est aussi. Dans tous les cas, résoudre le problème variationnel de type (3.3) par approximation interne consiste *in fine* à résoudre un problème d'algèbre linéaire. La définition d'un bon espace d'approximation interne sera le sujet des chapitres 5 et 6. Ce seront souvent des espaces de fonctions continues

mais seulement C^1 par morceaux donc si l'approximation est interne cela veut dire que l'espace $\mathcal{V}$ devra au moins les contenir...

3.6.2 Lemme de Céa

Nous allons maintenant comparer l'erreur commise quand on cherche une solution $u_N \in \mathcal{V}_N$ de (3.19) plutôt que $u \in \mathcal{V}$ dans (3.3). Le lemme suivant, dû à Jean Céa¹¹, montre que la distance entre la solution exacte u et la solution approchée u_N est majorée *uniformément*¹² par rapport au sous-espace $\mathcal{V}_N$ par la distance de u à $\mathcal{V}_N$.

Theorème 3.6.2 (Lemme de Céa) *Soit $\mathcal{V}$ un espace de Hilbert réel, et $\mathcal{V}_N$ un sous-espace de dimension finie. Soit $a(u, v)$ une forme bilinéaire continue et coercive sur $\mathcal{V}$, et $\ell(v)$ une forme linéaire continue sur $\mathcal{V}$. Soit u la solution de (3.3) et u_N celle de (3.19). On a*

$$\|u - u_N\|_{\mathcal{V}} \leq \frac{M}{\nu} \inf_{v_N \in \mathcal{V}_N} \|u - v_N\|_{\mathcal{V}}, \quad (3.22)$$

où M est la constante de continuité de a , et ν sa constante de coercivité.

Démonstration. Puisque $\mathcal{V}_N \subset \mathcal{V}$, on déduit, par soustraction des formulations variationnelles (3.3) et (3.19), que¹³

$$a(u - u_N, w_N) = 0 \quad \forall w_N \in \mathcal{V}_N.$$

En considérant $w_N = v_N - u_N$, on obtient alors

$$\begin{aligned} \nu \|u - u_N\|_{\mathcal{V}}^2 &\leq a(u - u_N, u - u_N) = a(u - u_N, u - v_N) + a(u - u_N, v_N - u_N) \\ &= a(u - u_N, u - v_N) \\ &\leq M \|u - u_N\|_{\mathcal{V}} \|u - v_N\|_{\mathcal{V}}, \end{aligned}$$

qui prouve (3.22). $\square$

Remarque 3.6.2 Il est facile de voir que $M/\nu \geq 1$ puisque pour tout $v \in \mathcal{V}$

$$\nu \|v\|_{\mathcal{V}}^2 \leq a(v, v) \leq M \|v\|_{\mathcal{V}}^2.$$

L'estimation (3.22) est ainsi cohérente avec le fait que nous avons par définition

$$\inf_{v_N \in \mathcal{V}_N} \|u - v_N\|_{\mathcal{V}} \leq \|u - u_N\|_{\mathcal{V}}.$$

Le lemme de Céa est l'outil fondamental pour étudier la convergence de l'approximation interne. Ce résultat est remarquable car il nous indique que l'erreur de toute approximation interne est contrôlée par la meilleur approximation que l'on peut trouver de u dans $\mathcal{V}_N$.

Il peut être amélioré¹⁴ dans le cas où a est symétrique.

Proposition 3.6.3 *On se place sous les hypothèses du lemme 3.6.2 de Céa*

11: Jean Céa 1932-2023 dont le titre de la thèse était *Approximation variationnelle des problèmes aux limites*

12: uniformément signifie ici que la distance ne dépend pas de u_N mais bien de $\mathcal{V}_N$ et de sa capacité à approcher l'espace

13: L'égalité suivante est fondamentale et elle signifie que l'erreur $u - u_N$ est "orthogonale" à $\mathcal{V}_N$ pour la forme bilinéaire a .

14: Puisque $M \geq \nu$ donc $\sqrt{\frac{M}{\nu}} < \frac{M}{\nu}$

auxquelles on ajoute le fait que a est symétrique, alors

$$\|u - u_N\|_{\mathcal{V}} \leq \sqrt{\frac{M}{\nu}} \inf_{v_N \in \mathcal{V}_N} \|u - v_N\|_{\mathcal{V}}. \quad (3.23)$$

Démonstration. Nous savons que dans le cas où a est symétrique alors elle définit un produit scalaire sur $\mathcal{V}$ équivalent au produit scalaire usuel. Par le théorème 3.2.2 de projection orthogonale, on a

$$\nu \|u - u_N\|_{\mathcal{V}}^2 \leq a(u - u_N, u - u_N) \leq a(u - v_N, u - v_N) \leq M \|u - v_N\|_{\mathcal{V}}^2,$$

d'où le résultat. $\square$

Théorème 3.6.4 (Convergence de l'approximation interne) *On se place sous les hypothèses du lemme 3.6.2. On suppose qu'il existe un sous-espace $\mathcal{W}$ dense de $\mathcal{V}$ et, pour chaque N , une application $r_N : \mathcal{W} \rightarrow \mathcal{V}_N$ tels que*

$$\forall v \in \mathcal{W}, \quad \lim_{N \rightarrow +\infty} \|v - r_N v\|_{\mathcal{V}} = 0, \quad (3.24)$$

alors

$$\lim_{N \rightarrow +\infty} \|u - u_N\|_{\mathcal{V}} = 0, \quad (3.25)$$

où u est la solution de (3.3) et u_N celle de (3.19).

Démonstration. Soit $\varepsilon > 0$, $\mathcal{W}$ est dense dans $\mathcal{V}$ donc, il existe $w \in \mathcal{W}$ tel que

$$\|u - w\|_{\mathcal{V}} \leq \frac{\nu \varepsilon}{2M}.$$

D'après l'hypothèse (3.24), il existe $N(\varepsilon) > 0$ tel que

$$\forall N \geq N(\varepsilon), \quad \|w - r_N w\|_{\mathcal{V}} \leq \frac{\nu \varepsilon}{2M}.$$

D'après le lemme (3.6.2) de Céa, nous avons donc

$$\|u - u_N\|_{\mathcal{V}} \leq \frac{M}{\nu} \|u - r_N w\|_{\mathcal{V}} \leq \frac{M}{\nu} \left[\|u - w\|_{\mathcal{V}} + \|w - r_N w\|_{\mathcal{V}} \right] \leq \varepsilon,$$

ce qu'il fallait démontrer. $\square$

Ce résultat est important car il indique que la méthode d'approximation interne converge tout le temps, sans hypothèse supplémentaire sur les données ou sur la solution à approcher. Cependant, si une information qualitative de la convergence est nécessaire, on verra que des hypothèses supplémentaires sont nécessaires¹⁵. Ce sera le sujet du chapitre 7.

15: On dit souvent en français "On n'a rien sans rien" ou encore en anglais "there is no free lunch"!

Étude mathématique des problèmes aux limites

4

Les outils d'analyse hilbertienne et les espaces de Sobolev adaptés vont nous fournir un cadre puissant et relativement général pour l'analyse des problèmes aux limites introduits au premier chapitre. C'est l'objectif de ce chapitre, avant d'aborder les questions de discrétisation des chapitres suivants.

4.1 Introduction

Nous allons maintenant traiter plusieurs problèmes aux limites élémentaires où les EDPs considérées sont de la forme

$$-\Delta u + \alpha u = f, \quad \text{dans } \Omega$$

avec $\alpha \geq 0$ complétées d'une condition aux limites de type Dirichlet homogène

$$u = 0, \quad \text{sur } \partial\Omega$$

ou de type Dirichlet non homogène

$$u = g, \quad \text{sur } \partial\Omega$$

pour une certaine fonction g du bord ou de type Neumann non homogène

$$\underline{\nabla} u \cdot \underline{n} = g, \quad \text{sur } \partial\Omega$$

pour une certaine fonction g du bord. La condition de Robin non homogène

$$\underline{\nabla} u \cdot \underline{n} + \lambda u = g, \quad \text{sur } \partial\Omega$$

pour $\lambda \geq 0$ et une certaine fonction g du bord sera étudiée en exercice.

Nous allons utiliser l'approche variationnelle présentée au chapitre 3 et le cadre fonctionnel présenté au chapitre 2. Il suffit alors de supposer peu de régularité sur les données, par exemple pour le terme source f , il suffit qu'il soit dans $L^2(\Omega)$. Nous verrons que pour les données au bord, cela dépendra de si c'est une condition aux bords de Dirichlet ou Neumann. Comme nous allons manipuler des fonctions qui sont typiquement seulement dans L^2 , l'équation aux dérivées partielles et la condition aux limites sont donc supposées être vérifiées seulement presque partout, respectivement dans Ω et dans $\partial\Omega$. On verra à la section 4.4.2 si on peut montrer a posteriori que la solution est régulière si les données le sont et par exemple à quelles conditions sur les données et le domaine il existe une solution classique de ces problèmes aux limites.

Nous allons également mettre en évidence comment les conditions aux limites interviennent dans ces différentes formulations. Ainsi certaines conditions aux limites apparaissent de manière naturelle dans la formulation variationnelle et la minimisation de l'énergie, on parlera de **conditions aux limites naturelles**; alors que d'autres apparaissent de manière explicite dans l'espace variationnel, on parlera de **conditions**

aux limites essentielles. Un soin particulier sera consacré à l'équivalence entre chaque problème aux limites et sa formulation variationnelle associée.

Pour étudier le caractère bien posé (existence, unicité et stabilité par rapport aux données) de ces formulations variationnelles, nous allons essayer d'appliquer le théorème de Lax-Milgram. Ainsi nous verrons que la situation est plus simple pour $\alpha > 0$. Pour $\alpha = 0$, c'est plus subtil, mais le théorème de Lax-Milgram dans le cas des conditions de Dirichlet s'applique toujours¹ et une des hypothèses du théorème n'est plus vérifiée dans le cas des conditions de Neumann.

1: le cas des conditions de Robin sera traité en exercice

4.2 Étude d'un problème elliptique scalaire

On suppose Ω un ouvert borné de $\mathbb{R}^d$ à frontière suffisamment régulière (voir la définition 1.2.1).

4.2.1 Conditions de Dirichlet homogènes

Nous considérons le problème aux limites suivant :

Trouver $u \in H^1(\Omega)$ telle que

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

où f est un terme source connu supposé dans $L^2(\Omega)$ et $\alpha \geq 0$.

Construction de la formulation variationnelle On suppose que $u \in H^1(\Omega)$ vérifie $(\mathcal{P}_D)$. Inspiré par ce que nous avons fait à la section ??, pour obtenir une formulation variationnelle, il faut multiplier l'équation volumique écrite sur Ω par une fonction test v et intégrer sur Ω . Il faut ensuite utiliser la formule de Green écrite pour des fonctions vivant dans les espaces de Sobolev : celle du théorème 2.2.19 a priori. Pour l'utiliser, il faut que u soit dans $H^2(\Omega)$ et v dans $H^1(\Omega)$.

Comme nous pouvons choisir les fonctions tests, prenons v dans $H^1(\Omega)$. Que savons-nous de u ? Nous savons que u est dans $H^1(\Omega)$ et d'après l'équation volumique de $(\mathcal{P}_D)$, que $-\Delta u = f - \alpha u \in L^2(\Omega)$. On en déduit que $u \in H^1(\Omega, \Delta)$. Cependant, rien ne nous assure que u est dans $H^2(\Omega)$. Il serait donc plus pertinent d'utiliser la formule (2.30). Cependant, comme expliqué à la remarque 2.2.5, nous allons supposer pour simplifier la formule² que $u \in H^2(\Omega)$. Nous pouvons appliquer la formule de Green du théorème 2.2.19 et nous trouvons

$$\int_{\Omega} [\nabla u \cdot \nabla v + \alpha u v] \, d\Omega - \int_{\partial\Omega} \nabla u \cdot \underline{n}|_{\partial\Omega} v|_{\partial\Omega} \, d\Gamma = \int_{\Omega} f v \, d\Omega. \quad (4.1)$$

Que faire de l'intégrale surfacique ? Nous n'avons aucune information sur la dérivée normale de u sur le bord $\nabla u \cdot \underline{n}|_{\partial\Omega}$ ³ Il faut donc éliminer ce terme. Pour cela, il suffit de choisir des fonctions tests v qui s'annulent au

2: et manipuler seulement des intégrales

3: Notons qu'il n'est pas possible de garder ce terme dans la formulation variationnelle car l'hypothèse de continuité de la forme bilinéaire dans le théorème de Lax-Milgram nécessiterait de contrôler ce terme en fonction de la norme H^1 . Ceci est impossible (voir le théorème de trace 2.2.18).

bord, c'est-à-dire $v \in H_0^1(\Omega)$ (voir la proposition 2.2.15). Nous obtenons que u vérifie

$$\forall v \in H_0^1(\Omega), \quad \int_{\Omega} [\nabla u \cdot \nabla v + \alpha u v] \, d\Omega = \int_{\Omega} f v \, d\Omega.$$

D'après la deuxième relation de $(\mathcal{P}_D)$ et d'après la proposition 2.2.15, on a que $u \in H_0^1(\Omega)$.

Ainsi si u vérifie $(\mathcal{P}_D)$ alors u vérifie la formulation variationnelle

Trouver $u \in H_0^1(\Omega)$ telle que

$$\forall v \in H_0^1(\Omega), \quad \int_{\Omega} [\nabla u \cdot \nabla v + \alpha u v] \, d\Omega = \int_{\Omega} f v \, d\Omega. \quad (FV_D)$$

La condition aux limites de Dirichlet intervient donc directement dans l'espace variationnel dans lequel on recherche la solution et dans lequel vivent les fonctions tests. On parle dans ce cas de **conditions aux limites essentielles**.

Remarque 4.2.1 Si on ne veut pas supposer que $u \in H^2(\Omega)$ pour obtenir (4.1), il faut utiliser la formule avec les crochets de dualité (2.30). Montrons que c'est un peu plus technique, mais que la conclusion est la même. Dans ce cas, il faut remplacer dans (4.1), l'intégrale surfacique par le crochet de dualité

$$\langle \nabla u \cdot \underline{n} |_{\partial\Omega}, v |_{\partial\Omega} \rangle_{[\text{Im } \gamma_0]', \text{Im } \gamma_0}$$

Comme nous n'avons aucune information sur $\nabla u \cdot \underline{n} |_{\partial\Omega}$, nous devons choisir $v \in H_0^1(\Omega)$ et le crochet de dualité disparaît. Nous aboutissons donc bien à la même formulation variationnelle.

Remarque 4.2.2 Une autre méthode pour obtenir cette formulation variationnelle est d'écrire que comme d'après l'EDP, u a un laplacien faible ($\Delta u \in L^2(\Omega)$) $\underline{U} := \nabla u$ a une divergence faible (puisque $\Delta = \nabla \cdot \nabla$) et par la définition 2.1.2

$$\forall v \in C_c^\infty(\Omega), \quad \int_{\Omega} \Delta u v \, d\Omega = - \int_{\Omega} \nabla u \cdot \nabla v \, d\Omega.$$

On en déduit que

$$\forall v \in C_c^\infty(\Omega), \quad \int_{\Omega} [\nabla u \cdot \nabla v + \alpha u v] \, d\Omega = \int_{\Omega} f v \, d\Omega.$$

Puisque la définition 2.2.2 implique que $C_c^\infty(\Omega)$ est dense dans $H_0^1(\Omega)$, ceci est vrai pour tout $v \in H_0^1(\Omega)$. Mais attention, cela ne fonctionne qu'avec des conditions de Dirichlet où les fonctions tests sont dans $H_0^1(\Omega)$.

Équivalence Montrons maintenant qu'en résolvant la formulation variationnelle (FV_D) , on aura bien résolu le problème aux limites $(\mathcal{P}_D)$. C'est une étape importante, car elle permet de vérifier qu'on a considéré "assez" de fonctions tests et que nous n'avons perdu aucune information

sur le problème aux limites. Cela permet de justifier en particulier dans ce cas qu'en choisissant des fonctions tests dans $H_0^1(\Omega)$ pour éliminer le terme de bord, nous n'avons perdu aucune information pour autant. Supposons donc que u vérifie (FV_D) et montrons que u vérifie $(\mathcal{P}_D)$.

Tout d'abord comme u est dans $H_0^1(\Omega)$, d'après la proposition 2.2.15, on a

$$u = 0, \quad \text{sur } \partial\Omega.$$

Montrons maintenant que (FV_D) permet de déduire que u a un laplacien faible dans $L^2(\Omega)$ ou en d'autres termes que $\underline{U} := \underline{\nabla} u$ a une divergence faible (nous rappelons que $\underline{\nabla} \cdot \underline{\nabla} = \Delta$). Pour cela, il suffit de montrer qu'il existe une fonction $L \in L^2(\Omega)$ telle que

$$\forall v \in C_c^\infty(\Omega), \quad \int_{\Omega} \nabla u \cdot \nabla v \, d\Omega = - \int_{\Omega} L v \, d\Omega,$$

dans ce cas $L = \Delta u$. Dans (FV_D) , si on choisit $v \in C_c^\infty(\Omega) \subset H^1(\Omega)$, on a alors :

$$\int_{\Omega} (\nabla u \cdot \nabla v + \alpha u v) \, d\Omega = \int_{\Omega} f v \, d\Omega,$$

soit

$$\int_{\Omega} \nabla u \cdot \nabla v \, d\Omega = \int_{\Omega} (f - \alpha u) v \, d\Omega.$$

où $f - \alpha u \in L^2(\Omega)$. La fonction u a bien un laplacien dans $L^2(\Omega)$ et $-\Delta u = f - \alpha u$ dans $L^2(\Omega)$ et donc presque partout dans Ω :

$$-\Delta u + \alpha u - f = 0, \quad \text{presque partout dans } \Omega.$$

Remarque 4.2.3 Pour les personnes qui sont familières avec les distributions et la dérivation au sens des distributions (voir Section 2.3), il n'est pas indispensable d'utiliser la formulation variationnelle comme ci-dessus pour définir le laplacien au sens faible. En effet, le laplacien est toujours défini au sens des distributions. En utilisant la définition (2.3.2), on a en effet pour tout $u \in \mathcal{D}'(\Omega)$ et pour tout $v \in C_c^\infty(\Omega)$

$$\langle \Delta u, v \rangle_{\mathcal{D}'(\Omega), \mathcal{D}(\Omega)} = - \sum_{i=1}^d \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial v}{\partial x_i} \right\rangle_{\mathcal{D}'(\Omega), \mathcal{D}(\Omega)} = - \langle \underline{\nabla} u, \underline{\nabla} v \rangle_{\mathcal{D}'(\Omega), \mathcal{D}(\Omega)}.$$

Ainsi, pour $u \in H^1(\Omega)$, on a pour tout $v \in C_c^\infty(\Omega)$,

$$\int_{\Omega} \underline{\nabla} u \cdot \underline{\nabla} v \, d\Omega \underset{\text{car } L^2(\Omega) \subset \mathcal{D}'(\Omega)}{=} \langle \underline{\nabla} u, \underline{\nabla} v \rangle_{\mathcal{D}'(\Omega), \mathcal{D}(\Omega)} = - \langle \Delta u, v \rangle_{\mathcal{D}'(\Omega), \mathcal{D}(\Omega)}.$$

De plus, pour tout $v \in C_c^\infty(\Omega)$ comme $L^2(\Omega) \subset \mathcal{D}'(\Omega)$ on a également

$$\int_{\Omega} u v \, d\Omega = \langle u, v \rangle_{\mathcal{D}'(\Omega), \mathcal{D}(\Omega)} \quad \text{et} \quad \int_{\Omega} f v \, d\Omega = \langle f, v \rangle_{\mathcal{D}'(\Omega), \mathcal{D}(\Omega)}.$$

Ainsi, si u vérifie (FV_D) , alors

$$\forall v \in C_c^\infty(\Omega), \quad \langle -\Delta u + \alpha u - f, v \rangle_{\mathcal{D}'(\Omega), \mathcal{D}(\Omega)} = 0,$$

soit

$$-\Delta u + \alpha u - f = 0, \quad \text{dans } D'(\Omega).$$

Comme f et u sont dans $L^2(\Omega)$, cette égalité est vraie dans $L^2(\Omega)$ et donc presque partout dans Ω .

Caractère bien posé de (FV_D) dans le cas $\alpha > 0$. Nous allons montrer maintenant que la formulation variationnelle (FV_D) admet une solution unique en utilisant le théorème de Lax-Milgram (voir le théorème 3.5.1). On introduit donc la forme bilinéaire $a : H_0^1(\Omega) \times H_0^1(\Omega) \rightarrow \mathbb{R}$ et la forme linéaire $\ell : H_0^1(\Omega) \rightarrow \mathbb{R}$ définies par

$$\forall v, w \in H_0^1(\Omega), \quad a(w, v) := \int_{\Omega} [\nabla w \cdot \nabla v + \alpha w v] \, d\Omega,$$

$$\forall v \in H_0^1(\Omega), \quad \ell(v) = \int_{\Omega} f v \, d\Omega.$$

- $(H_0^1(\Omega), \|\cdot\|_{H^1(\Omega)})$ est, d'après la proposition 2.2.5, un espace de Hilbert (en tant que sous-espace fermé de $H^1(\Omega)$).
- la forme bilinéaire a est continue. En effet, en utilisant des inégalités de Cauchy-Schwarz, on trouve

$$\begin{aligned} \forall v, w \in H_0^1(\Omega), \quad |a(w, v)| &\leq \|\nabla w\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} + \alpha \|w\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\ &\leq (1 + \alpha) \|w\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}; \end{aligned}$$

- la forme bilinéaire a est coercive puisque $\alpha > 0$ et

$$\forall v \in H_0^1(\Omega), \quad |a(v, v)| \geq \min(1, \alpha) \|v\|_{H^1(\Omega)}^2;$$

- la forme linéaire ℓ est continue. En effet, en utilisant une inégalité de Cauchy-Schwarz, on trouve

$$\forall v \in H_0^1(\Omega), \quad |\ell(v)| \leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)} \|v\|_{H^1(\Omega)}.$$

Toutes les hypothèses du théorème de Lax-Milgram sont donc satisfaites et on peut conclure qu'il existe une unique solution $u \in H_0^1(\Omega)$ de (FV_D) .

Il est facile de (re-)démontrer la stabilité de la solution par rapport aux données. En effet, l'unique solution $u \in H_0^1(\Omega)$ de (FV_D) vérifie la formulation variationnelle pour tout $v \in H_0^1(\Omega)$ et donc en particulier pour $v = u$. On obtient dans ce cas :

$$a(u, u) = \ell(u),$$

en utilisant d'une part la coercivité de a et d'autre part la continuité de ℓ , on obtient

$$\min(1, \alpha) \|u\|_{H^1(\Omega)}^2 \leq a(u, u) = \ell(u) \leq \|f\|_{L^2(\Omega)} \|u\|_{H^1(\Omega)}.$$

Cela signifie que si $f = 0$ alors $u = 0$ (ce qui était évident) et si $f \neq 0$ alors $u \neq 0$ ⁴ et

$$\|u\|_{H^1(\Omega)} \leq \frac{\|f\|_{L^2(\Omega)}}{\min(1, \alpha)}.$$

4: $u = 0$ ne peut pas être solution de (FV_D) si $f \neq 0$

5: On peut écrire également que l'application

$$f \in L^2(\Omega) \mapsto u \in H_0^1(\Omega)$$

où u est l'unique solution dans $H_0^1(\Omega)$ de (FV_D) ou de $(\mathcal{P}_D)$, est une application linéaire continue.

La formulation variationnelle (FV_D) et donc le problème aux limites $(\mathcal{P}_D)$ (puisqu'il y a équivalence) est donc bien posé dans $H_0^1(\Omega)$ au sens de Hadamard (voir la définition 3.1.1)⁵.

Notons que dans cette preuve, on utilise de manière cruciale (notamment pour établir la coercivité de a) que $\alpha > 0$. Le cas $\alpha = 0$ doit donc être traité à part.

Caractère bien posé de (FV_D) dans le cas $\alpha = 0$. Dans le cas, où $\alpha = 0$, le problème aux limites s'écrit

Trouver $u \in H^1(\Omega)$

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

et d'après ce qui précède, ce problème est équivalent à la formulation variationnelle

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

La différence principale avec le cas $\alpha > 0$ concerne la preuve du caractère bien posé du problème.

On montre comme précédemment que la forme bilinéaire a est continue dans $H_0^1(\Omega) \times H_0^1(\Omega)$ et que la forme linéaire ℓ est continue dans $H_0^1(\Omega)$ quand $H_0^1(\Omega)$ est muni de la norme $\|\cdot\|_{H^1(\Omega)}$. La difficulté se concentre donc sur la preuve de la coercivité de a . Rappelons que

$$\forall v \in H_0^1(\Omega), \quad a(v, v) = \int_{\Omega} |\nabla v|^2 \, d\Omega.$$

Nous reconnaissons donc la semi-norme $H^1, |\cdot|_{H^1(\Omega)}$ dont nous savons grâce à l'inégalité de Poincaré (voir la proposition 2.2.20) qu'elle est équivalente à la norme $\|\cdot\|_{H^1(\Omega)}$ (voir le corollaire 2.2.21). Plus précisément, nous avons

$$\forall v \in H_0^1(\Omega), \quad |v|_{H^1(\Omega)}^2 \geq \frac{1}{C_p^2 + 1} \|v\|_{H^1(\Omega)}^2,$$

où C_p est la constante de Poincaré. La forme bilinéaire est donc coercive, et la constante de coercivité est égale à $(C_p^2 + 1)^{-1}$.

Toutes les hypothèses du théorème de Lax-Milgram sont donc satisfaites. On peut donc conclure qu'il existe une unique solution $u \in H_0^1(\Omega)$ de (FV_D) même dans le cas $\alpha = 0$. Pour démontrer la stabilité, il suffit d'utiliser la même démarche que précédemment et on trouve :

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

Remarque 4.2.4 D'après le corollaire 2.2.21, $H_0^1(\Omega)$ muni de la semi-norme H^1 est un espace de Hilbert. Nous aurions pu donc démontrer que toutes les hypothèses du théorème de Lax-Milgram sont vérifiées pour $|\cdot|_{H^1(\Omega)}$.

4.2.2 Conditions de Dirichlet non homogènes

Nous considérons le problème aux limites suivant :

Trouver $u \in H^1(\Omega)$

$$\begin{cases} -\Delta u + u = f, & \text{dans } \Omega, \\ u = g, & \text{sur } \partial\Omega, \end{cases} \quad (\mathcal{P}_g)$$

où f est un terme source connu supposé dans $L^2(\Omega)$. Notons que par rapport au cas précédent, nous avons choisi $\alpha = 1$, la difficulté de ce cas se situant surtout sur la condition aux limites qui est maintenant non homogène.

La première question est l'espace auquel doit appartenir la donnée g . Si une solution u de $(\mathcal{P}_g)$ existe, g est par définition sa trace sur $\partial\Omega$. Donc g est dans $L^2(\partial\Omega)$. Mais il est important de se souvenir (voir la proposition 2.2.14) que toutes les fonctions de $L^2(\partial\Omega)$ ne sont pas des traces de fonctions de $H^1(\Omega)$. Pour espérer trouver une solution dans $H^1(\Omega)$, il faut nécessairement que

$$g \in \text{Im } \gamma_0.$$

Construction de la formulation variationnelle. On suppose que u vérifie $(\mathcal{P}_g)$. Pour construire la formulation variationnelle, les étapes sont strictement identiques à celles de la section précédente. Nous obtenons alors

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

Dans le cas présent, u n'est pas dans $H_0^1(\Omega)$ puisque sa trace n'est pas nulle sur le bord. Comme nous n'avons pas utilisé la condition aux limites, il faut donc garder cette information dans la formulation variationnelle qui est donc

Trouver $u \in H^1(\Omega)$ telle que $u|_{\partial\Omega} = g$ et

$$\forall v \in H_0^1(\Omega), \quad \int_{\Omega} [\nabla u \cdot \nabla v + u v] \, d\Omega = \int_{\Omega} f v \, d\Omega. \quad (FV_g)$$

Équivalence. Pour démontrer l'équivalence, là encore les étapes sont strictement équivalentes à celle de la section précédente pour les conditions de Dirichlet homogènes. Nous ne redétaillerons pas cette étape.

Caractère bien posé de (FV_g) . Pour appliquer le théorème de Lax-Milgram, il faut que l'espace où la solution est recherchée soit le même que l'espace où vivent les fonctions tests. Ici, $v \in H_0^1(\Omega)$ alors que, sauf si $g = 0$, ce n'est pas le cas de u . Pour y remédier, nous allons utiliser un relèvement de $g \in \text{Im } \gamma_0$. D'après la proposition 2.2.14,

$$\exists u_g \in H^1(\Omega) \text{ telle que } u_g|_{\partial\Omega} = g \text{ et } \|u_g\|_{H^1(\Omega)} \leq C \|g\|_{\text{Im}(\gamma_0)}. \quad (4.2)$$

Il suffit alors d'écrire la formulation variationnelle vérifiée par $u_0 := u - u_g$ qui par linéarité est

Trouver $u_0 \in H_0^1(\Omega)$ telle que

$$\begin{aligned} \forall v \in H_0^1(\Omega), \quad \int_{\Omega} [\nabla u_0 \cdot \nabla v + u_0 v] \, d\Omega = \int_{\Omega} f v \, d\Omega \\ + \int_{\Omega} [\nabla u_g \cdot \nabla v + u_g v] \, d\Omega. \quad (FV'_g) \end{aligned}$$

Nous pouvons maintenant tenter de vérifier que toutes les hypothèses du théorème de Lax-Milgram sont vérifiées. On introduit donc la forme bilinéaire $a : H_0^1(\Omega) \times H_0^1(\Omega) \rightarrow \mathbb{R}$ et la forme linéaire $\ell : H_0^1(\Omega) \rightarrow \mathbb{R}$ définis par

$$\forall v, w \in H_0^1(\Omega), \quad a(w, v) := \int_{\Omega} [\nabla w \cdot \nabla v + w v] \, d\Omega,$$

$$\forall v \in H_0^1(\Omega), \quad \ell(v) = \int_{\Omega} f v \, d\Omega + \int_{\Omega} [\nabla u_g \cdot \nabla v + u_g v] \, d\Omega.$$

- $(H_0^1(\Omega), \|\cdot\|_{H^1(\Omega)})$ est, d'après la proposition 2.2.5, un espace de Hilbert (en tant que sous-espace fermé de $H^1(\Omega)$).
- la forme bilinéaire a est continue. En effet, en utilisant l'inégalité de Cauchy-Schwarz dans $H^1(\Omega)$, on trouve

$$\forall v, w \in H_0^1(\Omega), \quad |a(w, v)| = |(v, w)_{H^1(\Omega)}| \leq \|w\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)};$$

- la forme bilinéaire a est coercive puisque

$$\forall v \in H_0^1(\Omega), \quad |a(v, v)| = \|v\|_{H^1(\Omega)}^2;$$

- la forme linéaire ℓ est continue. En effet, en utilisant une inégalité de Cauchy-Schwarz dans L^2 pour la première intégrale et H^1 pour la deuxième, on trouve

$$\begin{aligned} \forall v \in H_0^1(\Omega), \quad |\ell(v)| &\leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} + \|u_g\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} \\ &\leq [\|f\|_{L^2(\Omega)} + C \|g\|_{\text{Im}(\gamma_0)}] \|v\|_{H^1(\Omega)}, \end{aligned}$$

où on a utilisé (4.2).

Toutes les hypothèses du théorème de Lax-Milgram sont donc satisfaites et on peut conclure qu'il existe une unique solution $u_0 \in H_0^1(\Omega)$ de (FV'_g) .

Remarque 4.2.5 En remarquant que la forme bilinéaire a n'est rien d'autre que le produit scalaire H^1 , nous aurions pu utiliser le théorème de représentation de Riesz (voir théorème 3.3.1). En effet, comme d'après ce qui précède ℓ est une forme linéaire continue de $H_0^1(\Omega)$ qui est un espace de Hilbert, il existe un unique $u_0 \in H_0^1(\Omega)$ telle que

$$\forall v \in H_0^1(\Omega), \quad (u_0, v)_{H^1(\Omega)} = \ell(v).$$

Ceci nous donne donc que (FV'_g) a une unique solution qui est le représentant de Riesz dans $H_0^1(\Omega)$ de ℓ .

Pour (re-)démontrer la stabilité de u_0 par rapport aux données, il suffit de prendre $v = u_0$ dans (FV'_g) et on obtient

$$\|u_0\|_{H^1(\Omega)} \leq \|f\|_{L^2(\Omega)} + C \|g\|_{\text{Im}(\gamma_0)}.$$

Il existe donc au moins une solution $u \in H^1(\Omega)$ de (FV_g) qui est donnée par $u = u_0 + u_g$ où u_g vérifie (FV'_g) et u_0 est l'unique solution de (FV'_g) . Comme il n'y a pas unicité du relèvement, il est naturel de se poser la question de l'unicité. Supposons u et u' deux solutions dans $H^1(\Omega)$ de (FV_g) . Par linéarité $u - u' \in H_0^1(\Omega)$ vérifie

$$\forall v \in H_0^1(\Omega), \quad \int_{\Omega} [\nabla(u - u') \cdot \nabla v + (u - u') v] \, d\Omega = 0.$$

En prenant $v = u - u'$, on aboutit à $\|u - u'\|_{H^1(\Omega)} = 0$ et donc $u = u'$ presque partout dans Ω .

Nous venons de prouver qu'il existe une unique solution $u \in H^1(\Omega)$ de (FV_g) et donc de $(\mathcal{P}_g)$.

4.2.3 Conditions de Neumann non homogènes

Nous considérons maintenant le problème aux limites suivant :

Trouver $u \in H^1(\Omega)$ telle que

$$\begin{cases} -\Delta u + \alpha u = f, & \text{dans } \Omega, \\ \nabla u \cdot \underline{n} = g, & \text{sur } \partial\Omega, \end{cases} \quad (\mathcal{P}_N)$$

où f est un terme source connu supposé dans $L^2(\Omega)$, $\alpha \geq 0$ et $g \in L^2(\partial\Omega)$. L'approche variationnelle est un peu différente de celle vue précédemment c'est pour cela que nous redétaillons les différentes étapes.

Construction de la formulation variationnelle. On suppose que $u \in H^1(\Omega)$ vérifie $(\mathcal{P}_N)$. Comme précédemment, pour obtenir une formulation variationnelle, il faut multiplier l'équation volumique écrite sur Ω par une fonction test v et intégrer sur Ω . Il faut ensuite utiliser la formule de Green écrite pour des fonctions vivant dans les espaces de Sobolev : celle du théorème 2.2.19 a priori. Pour l'utiliser, il faut que u soit dans $H^2(\Omega)$ et v dans $H^1(\Omega)$.

Comme nous pouvons choisir les fonctions tests, prenons v dans $H^1(\Omega)$. Que savons-nous de u ? Nous savons que u est dans $H^1(\Omega)$ et d'après l'équation volumique de $(\mathcal{P}_D)$, que $-\Delta u = f - \alpha u \in L^2(\Omega)$. On en déduit que $u \in H^1(\Omega, \Delta)$. Cependant, rien ne nous assure que u est dans $H^2(\Omega)$. Il est donc plus pertinent d'utiliser la formule (2.30). Comme expliqué à la remarque 2.2.5, nous allons supposer pour simplifier la formule ⁶ que $u \in H^2(\Omega)$. Nous pouvons appliquer la formule de Green du théorème 2.2.19 et nous trouvons

$$\int_{\Omega} [\nabla u \cdot \nabla v + \alpha u v] \, d\Omega - \int_{\partial\Omega} \nabla u \cdot \underline{n} \big|_{\partial\Omega} v \big|_{\partial\Omega} \, d\Gamma = \int_{\Omega} f v \, d\Omega. \quad (4.3)$$

⁶: et manipuler seulement des intégrales

7: Notons que si nous avions pris une fonction test $v \in H_0^1(\Omega)$, ce terme aurait disparu. Nous aurions abouti à une autre formulation variationnelle mais qui n'est pas équivalente au problème aux limites. En effet, il n'aurait pas été possible de retrouver la condition aux limites.

Dans le cas du problème avec des conditions de Neumann, nous avons une condition sur la dérivée normale de u sur le bord $\nabla u \cdot \underline{n}|_{\partial\Omega} = g$. Il suffit donc d'utiliser cette information⁷. Nous obtenons que u vérifie

$$\forall v \in H^1(\Omega), \quad \int_{\Omega} [\nabla u \cdot \nabla v + \alpha u v] \, d\Omega = \int_{\Omega} f v \, d\Omega + \int_{\partial\Omega} g v|_{\partial\Omega} \, d\Gamma.$$

Ainsi si u vérifie $(\mathcal{P}_N)$ alors u vérifie la formulation variationnelle :
Trouver $u \in H^1(\Omega)$ telle que

$$\begin{aligned} \forall v \in H^1(\Omega), \quad \int_{\Omega} [\nabla u \cdot \nabla v + \alpha u v] \, d\Omega \\ = \int_{\Omega} f v \, d\Omega + \int_{\partial\Omega} g v|_{\partial\Omega} \, d\Gamma. \quad (FV_N) \end{aligned}$$

La condition aux limites de Neumann a été prise en compte dans la formulation variationnelle (et non dans l'espace variationnel). On parle dans ce cas de **conditions aux limites naturelles**.

Remarque 4.2.6 Si on ne veut pas supposer que $u \in H^2(\Omega)$ pour obtenir (4.3), il faut utiliser la formule avec les crochets de dualité (2.30). C'est un peu plus technique, mais la conclusion est la même. Dans ce cas, il faut remplacer dans (4.3), l'intégrale surfacique par le crochet de dualité

$$\langle \nabla u \cdot \underline{n}|_{\partial\Omega}, v|_{\partial\Omega} \rangle_{[Im\gamma_0]', Im\gamma_0}.$$

Nous pouvons utiliser la condition aux limites $\nabla u \cdot \underline{n}|_{\partial\Omega} = g$ dans le crochet de dualité. Comme g et $v|_{\partial\Omega}$ sont dans $L^2(\partial\Omega)$, ce crochet de dualité n'est rien d'autre qu'une intégrale sur $\partial\Omega$. Nous aboutissons donc bien à la même formulation variationnelle.

Équivalence. Montrons maintenant qu'en résolvant la formulation variationnelle (FV_N) , on aura bien résolu le problème aux limites $(\mathcal{P}_N)$. Supposons donc que u vérifie (FV_N) et montrons que u vérifie $(\mathcal{P}_N)$.

Pour retrouver l'EDP volumique, il suffit de procéder comme précédemment. En d'autres termes, il suffit de considérer des fonctions tests $v \in C_c^\infty(\Omega) \subset H^1(\Omega)$. Dans ce cas, (FV_N) se simplifie puisque $v|_{\partial\Omega} = 0$

$$\forall v \in C_c^\infty(\Omega), \quad \int_{\Omega} [\nabla u \cdot \nabla v + \alpha u v] \, d\Omega = \int_{\Omega} f v \, d\Omega. \quad (4.4)$$

Nous retrouvons une formulation qui permet encore de montrer que u a un laplacien faible dans $L^2(\Omega)$ puisque

$$\forall v \in C_c^\infty(\Omega), \quad \int_{\Omega} \nabla u \cdot \nabla v \, d\Omega = - \int_{\Omega} (\alpha u - f) v \, d\Omega.$$

où $\alpha u - f \in L^2(\Omega)$. Cela signifie par définition (voir la définition 2.1.2) que ∇u a bien une divergence faible dans $L^2(\Omega)$ qui n'est rien d'autre que $\alpha u - f$. On a donc que $\Delta u = \alpha u - f$ dans $L^2(\Omega)$ et donc

$$-\Delta u + \alpha u - f = 0 \quad \text{presque partout dans } \Omega.$$

Nous avons donc retrouvé l'EDP volumique de $(\mathcal{P}_N)$.

Pour retrouver la condition aux limites, il faut continuer à exploiter (FV_N) pour des fonctions tests qui ne vont maintenant plus s'annuler au bord. Nous savons maintenant que u vérifie l'EDP de $(\mathcal{P}_N)$ donc $u \in H^1(\Omega, \Delta)$. Comme expliqué à la remarque 2.2.5, nous allons supposer pour simplifier la formule (et manipuler seulement des intégrales) que $u \in H^2(\Omega)$. Nous pouvons appliquer la formule de Green du théorème 2.2.19 et nous trouvons

$$\begin{aligned} \forall v \in H^1(\Omega), \quad \int_{\Omega} [-\Delta u + \alpha u] v \, d\Omega + \int_{\partial\Omega} \underline{\nabla} u \cdot \underline{n}|_{\partial\Omega} v|_{\partial\Omega} \, d\Gamma \\ = \int_{\Omega} f v \, d\Omega + \int_{\partial\Omega} g v|_{\partial\Omega} \, d\Gamma. \end{aligned} \quad (4.5)$$

Comme $-\Delta u + \alpha u - f = 0$ presque partout dans Ω , les intégrales volumiques se simplifient et on obtient

$$\forall v \in H^1(\Omega), \quad \int_{\partial\Omega} [\underline{\nabla} u \cdot \underline{n}|_{\partial\Omega} - g] v|_{\partial\Omega} \, d\Gamma = 0.$$

D'après la proposition 2.2.14, pour tout $v \in H^1(\Omega)$, $v|_{\partial\Omega} \in \text{Im } \gamma_0$ et

$$\forall \psi \in \text{Im } \gamma_0, \exists v \in H^1(\Omega) \quad v|_{\partial\Omega} = \psi.$$

On en déduit que

$$\forall \varphi \in \text{Im } \gamma_0, \quad \int_{\partial\Omega} [\underline{\nabla} u \cdot \underline{n}|_{\partial\Omega} - g] \varphi \, d\Gamma = 0,$$

autrement dit

$$\underline{\nabla} u \cdot \underline{n}|_{\partial\Omega} - g \in L^2(\partial\Omega) \text{ est telle que } \underline{\nabla} u \cdot \underline{n}|_{\partial\Omega} - g \in [\text{Im } \gamma_0]^\perp,$$

où l'orthogonal est défini pour le produit scalaire de $L^2(\partial\Omega)$. Comme $\text{Im } \gamma_0$ est dense dans $L^2(\partial\Omega)$, d'après le corollaire 3.2.6, on en déduit que

$$\underline{\nabla} u \cdot \underline{n}|_{\partial\Omega} - g = 0 \text{ dans } L^2(\partial\Omega) \text{ et donc presque partout sur } \partial\Omega.$$

On retrouve donc la condition aux limites de $(\mathcal{P}_N)$.

Remarque 4.2.7 Si on ne veut pas supposer que $u \in H^2(\Omega)$ pour obtenir (4.5), il faut utiliser la formule avec les crochets de dualité (2.30). Montrons que c'est un peu plus technique, mais que la conclusion est la même. Il faut tout d'abord remplacer dans (4.5) l'intégrale surfacique par le crochet de dualité. Comme u vérifie l'EDP, les intégrales volumiques se simplifient et on obtient

$$\forall v \in H^1(\Omega), \quad \langle \underline{\nabla} u \cdot \underline{n}|_{\partial\Omega} - g, v|_{\partial\Omega} \rangle_{[\text{Im } \gamma_0]', \text{Im } \gamma_0} = 0.$$

Pour les mêmes raisons que précédemment, on en déduit que

$$\forall \varphi \in \text{Im } \gamma_0, \quad \langle \underline{\nabla} u \cdot \underline{n}|_{\partial\Omega} - g, \varphi \rangle_{[\text{Im } \gamma_0]', \text{Im } \gamma_0} = 0$$

soit

$$\underline{\nabla} u \cdot \underline{n}|_{\partial\Omega} - g = 0 \text{ dans } [\text{Im } \gamma_0]'.$$

Comme $g \in L^2(\partial\Omega)$, l'égalité est vraie dans $L^2(\partial\Omega)$ et donc presque partout sur $\partial\Omega$.

Caractère bien posé de (FV_N) dans le cas $\alpha > 0$. Nous allons montrer maintenant que la formulation variationnelle (FV_N) admet une solution unique en utilisant le théorème de Lax-Milgram (voir le théorème 3.5.1). On introduit donc la forme bilinéaire $a : H^1(\Omega) \times H^1(\Omega) \rightarrow \mathbb{R}$ et la forme linéaire $\ell : H^1(\Omega) \rightarrow \mathbb{R}$ définies par

$$\forall v, w \in H^1(\Omega), \quad a(w, v) := \int_{\Omega} [\nabla w \cdot \nabla v + \alpha w v] \, d\Omega,$$

$$\forall v \in H^1(\Omega), \quad \ell(v) = \int_{\Omega} f v \, d\Omega + \int_{\partial\Omega} g v|_{\partial\Omega} \, d\Gamma.$$

- $(H^1(\Omega), \|\cdot\|_{H^1(\Omega)})$ est un espace de Hilbert.
- la forme bilinéaire a est continue. En effet, en utilisant des inégalités de Cauchy-Schwarz, on trouve

$$\begin{aligned} \forall v, w \in H^1(\Omega), \quad |a(w, v)| &\leq \|\nabla w\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} + \alpha \|w\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \\ &\leq (1 + \alpha) \|w\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}; \end{aligned}$$

- la forme bilinéaire a est coercive puisque $\alpha > 0$ et

$$\forall v \in H^1(\Omega), \quad |a(v, v)| \geq \min(1, \alpha) \|v\|_{H^1(\Omega)}^2;$$

- la forme linéaire ℓ est continue. En effet, en utilisant des inégalités de Cauchy-Schwarz dans $L^2(\Omega)$ et $L^2(\partial\Omega)$, on trouve

$$\begin{aligned} \forall v \in H_0^1(\Omega), \quad |\ell(v)| &\leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} + \|g\|_{L^2(\partial\Omega)} \|v|_{\partial\Omega}\|_{L^2(\partial\Omega)} \\ &\leq [\|f\|_{L^2(\Omega)} + C_0 \|g\|_{L^2(\partial\Omega)}] \|v\|_{H^1(\Omega)}, \end{aligned}$$

où on a utilisé la continuité de l'application trace (voir le théorème 2.2.13)

Toutes les hypothèses du théorème de Lax-Milgram sont donc satisfaites et on peut conclure qu'il existe une unique solution $u \in H^1(\Omega)$ de (FV_N) .

Pour (re-)démontrer la stabilité de u par rapport aux données, il suffit de prendre $v = u$ dans (FV_N) et on obtient en utilisant la coercivité de a d'une part et la continuité de ℓ d'autre part que

$$\|u\|_{H^1(\Omega)} \leq \frac{\|f\|_{L^2(\Omega)} + C_0 \|g\|_{L^2(\partial\Omega)}}{\min(1, \alpha)}.$$

Caractère bien posé de (FV_N) dans le cas $\alpha = 0$? Dans le cas, où $\alpha = 0$, le problème aux limites s'écrit :

Trouver $u \in H^1(\Omega)$

$$\begin{cases} -\Delta u = f, & \text{dans } \Omega, \\ \nabla u \cdot \underline{n} = g, & \text{sur } \partial\Omega, \end{cases} \quad (4.6)$$

et d'après ce qui précède, ce problème est équivalent à la formulation variationnelle

$$\forall v \in H^1(\Omega), \quad \int_{\Omega} \nabla u \cdot \nabla v \, d\Omega = \int_{\Omega} f v \, d\Omega + \int_{\partial\Omega} g v|_{\partial\Omega} \, d\Gamma. \quad (4.7)$$

Est-ce que ce problème est bien posé dans $H^1(\Omega)$? Il est facile de voir en utilisant le problème aux limites ou la formulation variationnelle que s'il existe une solution u alors pour toute fonction constante notée C , $u + C$ est également solution. Il n'y a donc pas **unicité**. De plus, en prenant $v = 1$ dans (4.7) qui est bien dans $H^1(\Omega)$, on trouve

$$\int_{\Omega} f \, d\Omega + \int_{\partial\Omega} g|_{\partial\Omega} \, d\Gamma = 0. \quad (4.8)$$

On appelle cette condition, la *condition de compatibilité*. Si (4.8) n'est pas satisfaite, il n'y a pas **existence** d'une solution.

Puisque le problème n'est pas bien posé, cela signifie en particulier qu'une des hypothèses du théorème de Lax-Milgram n'est pas vérifiée. En effet la coercivité de la forme bilinéaire associée

$$\forall v, w \in H^1(\Omega), \quad a(w, v) := \int_{\Omega} \nabla w \cdot \nabla v \, d\Omega$$

fait défaut puisque l'on a

$$\forall C \in \mathbb{R}, \quad v = C \text{ dans } \Omega, \quad a(v, v) = 0 \quad \text{et} \quad \|v\|_{H^1(\Omega)}^2 = C^2 |\Omega|.$$

Cette difficulté est intimement liée au fait qu'il ne peut exister qu'une solution définie à une constante additive près.

Nous allons montrer qu'il est possible de rajouter des conditions au problème pour qu'il devienne bien posé. Tout d'abord, il faut nécessairement supposer que f et g vérifient la condition de compatibilité (4.8) pour espérer construire une solution. De plus, il ne peut exister qu'une solution à une constante additive près. Pour lever cette indétermination, il faut "fixer" la constante en fixant par exemple la moyenne. Par linéarité, il suffit de supposer que la solution est recherchée dans

$$\mathcal{V} := \{v \in H^1(\Omega), \quad m(v) = 0\} \quad \text{où} \quad m(v) = \int_{\Omega} v \, d\Omega.$$

Nous introduisons donc la formulation variationnelle

Trouver $u \in \mathcal{V}$ telle que

$$\forall v \in \mathcal{V}, \quad \int_{\Omega} \nabla u \cdot \nabla v \, d\Omega = \int_{\Omega} f v \, d\Omega + \int_{\partial\Omega} g v|_{\partial\Omega} \, d\Gamma. \quad (FV'_N)$$

Dans cette nouvelle formulation variationnelle, l'espace des fonctions tests est donc strictement inclus dans $H^1(\Omega)$. Il n'est donc absolument pas évident que cette nouvelle formulation variationnelle est bien équivalente à (4.7) (qui a été démontrée, précédemment, être équivalente à (4.6)). Il est facile de voir que si u vérifie (4.7) alors u vérifie en particulier (FV'_N) . Supposons maintenant que u vérifie (FV'_N) , alors pour tout $v \in H^1(\Omega)$,

$v - m(v)|\Omega|^{-1}$ est dans $\mathcal{V}$. On a donc

$$\begin{aligned} & \int_{\Omega} \underline{\nabla} u \cdot \underline{\nabla} (v - m(v)|\Omega|^{-1}) \, d\Omega \\ &= \int_{\Omega} f (v - m(v)|\Omega|^{-1}) \, d\Omega + \int_{\partial\Omega} g (v - m(v)|\Omega|^{-1})|_{\partial\Omega} \, d\Gamma \end{aligned}$$

soit

$$\begin{aligned} & \int_{\Omega} \underline{\nabla} u \cdot \underline{\nabla} v \, d\Omega \\ &= \int_{\Omega} f v \, d\Omega + \int_{\partial\Omega} g v|_{\partial\Omega} \, d\Gamma - m(v)|\Omega|^{-1} \left[\int_{\Omega} f \, d\Omega + \int_{\partial\Omega} g \, d\Gamma \right]. \end{aligned}$$

Comme la condition de compatibilité (4.8) est satisfaite, on retrouve bien que u vérifie (4.7).

Montrons maintenant que pour (FV'_N) les hypothèses du théorème de Lax-Milgram sont bien satisfaites. (voir le théorème 3.5.1). On introduit donc la forme bilinéaire $a : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ et la forme linéaire $\ell : \mathcal{V} \rightarrow \mathbb{R}$ définies par

$$\forall v, w \in \mathcal{V}, \quad a(w, v) := \int_{\Omega} \underline{\nabla} w \cdot \underline{\nabla} v \, d\Omega,$$

$$\forall v \in \mathcal{V}, \quad \ell(v) = \int_{\Omega} f v \, d\Omega + \int_{\partial\Omega} g v|_{\partial\Omega} \, d\Gamma.$$

- $(\mathcal{V}, \|\cdot\|_{H^1(\Omega)})$ est un espace de Hilbert en tant que sous-espace fermé de $H^1(\Omega)$. En effet, $\mathcal{V}$ est le noyau de la forme linéaire

$$m : \begin{cases} L^2(\Omega) \rightarrow \mathbb{R} \\ v \mapsto \int_{\Omega} v \, d\Omega \end{cases}$$

qui est bien continue dans $L^2(\Omega)$ (d'après l'inégalité de Cauchy-Schwarz) et donc dans $H^1(\Omega)$.

- la forme bilinéaire a est continue. En effet, en utilisant l'inégalité de Cauchy-Schwarz, on trouve

$$\forall v, w \in \mathcal{V}, \quad |a(w, v)| \leq \|\underline{\nabla} w\|_{L^2(\Omega)} \|\underline{\nabla} v\|_{L^2(\Omega)} \leq \|w\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)};$$

- la forme bilinéaire a est coercive puisque dans V , nous pouvons utiliser l'inégalité de Poincaré-Wirtinger (2.34)

$$\forall v \in \mathcal{V}, \quad |a(v, v)| \geq \frac{1}{C_2^2 + 1} \|v\|_{H^1(\Omega)}^2;$$

où C_2 est la constante qui intervient dans (2.34).

- la forme linéaire ℓ est continue. En effet, en utilisant des inégalités de Cauchy-Schwarz dans $L^2(\Omega)$ et $L^2(\partial\Omega)$, on trouve

$$\begin{aligned} \forall v \in \mathcal{V}, \quad |\ell(v)| &\leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} + \|g\|_{L^2(\partial\Omega)} \|v|_{\partial\Omega}\|_{L^2(\partial\Omega)} \\ &\leq (\|f\|_{L^2(\Omega)} + C_0 \|g\|_{L^2(\partial\Omega)}) \|v\|_{H^1(\Omega)}, \end{aligned}$$

où on a utilisé la continuité de l'application trace (voir le théorème 2.2.13)

Toutes les hypothèses du théorème de Lax-Milgram sont donc satisfaites. Il existe donc une unique solution u dans $\mathcal{V}$ de (FV'_N) et donc de (4.7) et de (4.6) à condition que la condition de compatibilité (4.8) soit satisfaite.

4.3 La méthode directe du calcul des variations

La méthode directe du calcul des variations est une méthode permettant de trouver des solutions à des problèmes de minimisation par minimisation directe et sans passer par la résolution d'une équation de type Euler-Lagrange. À des fins d'illustration de la méthode, on reprend un exemple déjà étudié précédemment avec la fonctionnelle

$$\forall v \in H_0^1(\Omega), \quad \mathcal{J}(v) = \frac{1}{2} \int_{\Omega} |\nabla v(\underline{x})|^2 \, d\Omega - \int_{\Omega} f(\underline{x})v(\underline{x}) \, d\Omega,$$

et l'on souhaite montrer que $\mathcal{J}$ admet un minimum sur $H_0^1(\Omega)$. La méthode comporte 3 étapes :

- (1) On montre que $\mathcal{J}$ est minorée sur $H_0^1(\Omega)$. Ceci entraîne que l'infimum $m = \inf_{v \in H_0^1(\Omega)} \mathcal{J}(v)$ existe.
- (2) On considère une suite minimisante de $\mathcal{J}$ sur $H_0^1(\Omega)$, c'est-à-dire une suite $(v_n)_{n \in \mathbb{N}}$ telle que

$$\lim_{n \rightarrow +\infty} \mathcal{J}(v_n) = m. \quad (4.9)$$

On extrait une sous-suite éventuelle de $(v_n)_{n \in \mathbb{N}}$ qui converge (éventuellement faiblement) dans $H_0^1(\Omega)$. Pour cela, il suffit de montrer que puisque $(\mathcal{J}(v_n))_{n \in \mathbb{N}}$ est borné alors $(v_n)_{n \in \mathbb{N}}$ est bornée dans $H_0^1(\Omega)$. On aura alors $v_{\phi(n)} \rightharpoonup v_{\infty}$ faiblement dans $H_0^1(\Omega)$.

- (3) On montre enfin que

$$\mathcal{J}(v_{\infty}) \leq \liminf_{n \rightarrow +\infty} \mathcal{J}(v_{\phi(n)}) = m.$$

Ceci conclut que v_{∞} réalise le minimum de $\mathcal{J}$ sur $H_0^1(\Omega)$ et donc que le minimum est atteint.

Les trois étapes précédentes sont résolues successivement très simplement avec les outils vus dans les chapitres précédents.

- (1) Par l'inégalité de Poincaré (2.31), on a pour tout $v \in H_0^1(\Omega)$

$$\left| \int_{\Omega} f(\underline{x})v(\underline{x}) \, d\Omega \right| \leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \leq C \|f\|_{L^2(\Omega)} \|v\|_{H_0^1(\Omega)},$$

de sorte que⁸

$$\mathcal{J}(v) \geq \frac{1}{2} \|v\|_{H_0^1(\Omega)}^2 - C \|f\|_{L^2(\Omega)} \|v\|_{H_0^1(\Omega)} \geq -\frac{C^2 \|f\|_{L^2(\Omega)}^2}{2},$$

et $\mathcal{J}$ est minorée sur $H_0^1(\Omega)$.

- (2) Du calcul précédent, on déduit que si $\mathcal{J}(v_n) < M$, alors

$$\frac{1}{2} \|v_n\|_{H_0^1(\Omega)}^2 - C \|f\|_{L^2(\Omega)} \|v_n\|_{H_0^1(\Omega)} < M.$$

8: en utilisant que pour tout $a, b \in \mathbb{R}$,
 $2ab \leq a^2 + b^2$

Mais comme $C\|f\|_{L^2}\|v_n\|_{H_0^1(\Omega)} \leq C^2\|f\|_{L^2(\Omega)}^2 + \frac{1}{4}\|v_n\|_{H_0^1(\Omega)}^2$, on a

$$\frac{1}{4}\|v_n\|_{H_0^1(\Omega)}^2 < M + C^2\|f\|_{L^2(\Omega)}^2,$$

et donc $(v_n)_n$ est bornée dans $H_0^1(\Omega)$. On peut, quitte à extraire une sous-suite, considérer que

$$v_n \rightharpoonup v_\infty \text{ faiblement dans } H_0^1(\Omega),$$

puis, grâce au théorème 2.2.24 de Rellich, que

$$v_n \rightarrow v_\infty \text{ fortement dans } L^2(\Omega).$$

(3) On sait, par propriété de la convergence faible que

$$\|v_\infty\|_{H_0^1(\Omega)} \leq \liminf_{n \rightarrow +\infty} \|v_n\|_{H_0^1(\Omega)},$$

et, grâce à la convergence forte dans L^2 que

$$\lim_{n \rightarrow +\infty} \int_{\Omega} f(\underline{x})v_n(\underline{x}) \, d\Omega = \int_{\Omega} f(\underline{x})v_\infty(\underline{x}) \, d\Omega.$$

Ceci permet de conclure que

$$\mathcal{J}(v_\infty) \leq \liminf_{n \rightarrow +\infty} \mathcal{J}(v_n) = m,$$

et le résultat.

En utilisant le fait que v_∞ réalise le minimum de $\mathcal{J}$ sur $H_0^1(\Omega)$ on a que v_∞ est solution de l'équation d'Euler-Lagrange du problème qui n'est autre que la formulation variationnelle. On a ainsi résolu le problème directement par minimisation sans passer par l'application du théorème de Lax-Milgram.

9: voir les opérateurs compacts au chapitre 8

Remarque 4.3.1 (Energie non-linéaire) Lorsque le domaine Ω est borné, nous avons vu que le théorème 2.2.24 de Rellich permet d'extraire de toute suite bornée de $H^1(\Omega)$ une sous-suite convergeant fortement dans $L^2(\Omega)$. On dit que l'injection de $H^1(\Omega)$ dans $L^2(\Omega)$ est compacte⁹. Il se trouve que plus généralement pour Ω borné, l'injection de $H^1(\Omega)$ dans $L^p(\Omega)$ est compacte pour tout $p < \bar{p}$ où $\frac{1}{\bar{p}} = \frac{1}{2} - \frac{1}{d}$. Dans le cas où $d=2$ (donc $\bar{p} = +\infty$), l'injection de $H^1(\Omega)$ dans $L^p(\Omega)$ reste compacte pour tout p fini, mais il n'y a pas d'injection continue dans $L^\infty(\Omega)$ contrairement au cas $d = 1$. Pour tout $p < \bar{p}$, l'inégalité suivante appelée *inégalité de Sobolev* stipule alors qu'il existe une constante $C(p)$ telle que l'on ait

$$\forall v \in H^1(\Omega), \quad \|v\|_{L^p} \leq C(p)\|v\|_{H^1}.$$

Ce résultat permet facilement d'étendre la méthode directe du calcul des variations à des énergies éventuellement non linéaires à l'image de celles vues au chapitre 1. Considérons en effet

$$\mathcal{J}(v) = \frac{1}{2} \int_{\Omega} |\nabla v(\underline{x})|^2 \, d\Omega + \int_{\Omega} F(v(\underline{x})) \, d\Omega$$

où le dernier terme est continu dans L^p pour un certain $p < \bar{p}$. C'est typiquement le cas pour le modèle de transition de phase suivant où il s'agit de minimiser sur $H^1(\Omega)$ l'énergie

$$\mathcal{J}(v) = \frac{1}{2} \int_{\Omega} |\nabla v(\underline{x})|^2 d\Omega + \frac{1}{4\varepsilon} \int_{\Omega} (v(\underline{x})^2 - 1)^2 d\Omega$$

En effet on peut montrer qu'il existe pour tout $\varepsilon > 0$ un minimiseur u_ε , qui vérifie l'équation d'Euler

$$\begin{cases} -\Delta u_\varepsilon + \frac{1}{\varepsilon}(u_\varepsilon^2 - 1)u_\varepsilon = 0 & \text{sur } \Omega, \\ \frac{\partial u_\varepsilon}{\partial n} = 0 & \text{sur } \partial\Omega. \end{cases} \quad (4.10)$$

Il s'agit là d'une équation non-linéaire dont on aurait du mal à démontrer l'existence d'une solution sans passer par la procédure de minimisation précédente. Notez que cette solution n'est en général pas unique, du moins si ε est assez petit, puisque l'on peut montrer que dans ce cas u_ε n'est pas identiquement nulle et qu'ainsi $-u_\varepsilon$ est aussi solution du problème.

4.4 Propriétés des solutions faibles

Nous finissons ce chapitre en donnant des propriétés qualitatives des solutions faibles des problèmes étudiés précédemment : principe du maximum et régularité ou non des solutions faibles.

4.4.1 Principe du maximum

Nous considérons dans cette section le problème de Dirichlet non homogène étudié à la section (4.2.2)

$$\begin{cases} -\Delta u + \alpha u = f, & \text{dans } \Omega, \\ u = g, & \text{sur } \partial\Omega. \end{cases} \quad (\mathcal{P})$$

où $f \in L^2(\Omega)$, $\alpha \geq 0$ et $g \in \text{Im } \gamma_0$.

Énonçons tout d'abord un résultat de positivité de la solution.

Proposition 4.4.1 *Supposons que $f \geq 0$ presque partout dans Ω et $g \geq 0$ presque partout sur $\partial\Omega$. Alors la solution u de $(\mathcal{P})$ vérifie $u \geq 0$ presque partout dans Ω .*

Puisque u est solution de $(\mathcal{P})$ si et seulement si $-u$ est solution de $(\mathcal{P})$ en remplaçant f et g respectivement par $-f$ et $-g$, il est facile de déduire de la proposition 4.4.1, que si $f \leq 0$ presque partout dans Ω et $g \leq 0$ presque partout sur $\partial\Omega$ alors la solution u de $(\mathcal{P})$ vérifie $u \leq 0$ presque partout dans Ω .

Démonstration. On introduit¹⁰

$$\Omega_- := \text{Int}(\text{Support } u \leq 0) \quad \text{et} \quad \Omega_+ := \text{Int}(\text{Support } u \geq 0).$$

10: Pour des fonctions continues, on a

$$\text{Support } u := \overline{\{\underline{x} \in \Omega, u(\underline{x}) \neq 0\}}.$$

Pour des fonctions de L^2 , cette définition est plus délicate car les fonctions ne sont définies que presque partout. Il faut définir le support de u de telle façon qu'il ne dépende que de la classe d'équivalence des fonctions égales à u . On note ω l'union de tous les ouverts où u s'annule. Le support (dit essentiel dans ce cas) est

$\Omega_{\pm}$ sont donc des ouverts disjoints de Ω tels que $\overline{\Omega} = \overline{\Omega_-} \cup \overline{\Omega_+}$.

Introduisons maintenant $u_- := \min(u, 0)$. Si $u_- = 0$ dans Ω , cela signifie que $\Omega_- = \emptyset$ et donc $u \geq 0$ presque partout dans Ω . Supposons maintenant que ce ne soit pas le cas et donc $\Omega_- \neq \emptyset$. Dans ce cas, il est facile de montrer par l'absurde que $u|_{\partial\Omega_-} = 0$. D'après le théorème 2.2.2, u_- est donc dans $H^1(\Omega)$ et

$$\nabla u_- = \chi_{\Omega_-} \nabla u,$$

où χ_{Ω_-} est la fonction caractéristique de Ω_- . De plus $u_- \in H_0^1(\Omega)$ puisque

$$u_-|_{\partial\Omega} = \min(g, 0) = 0 \quad \text{car } g \geq 0.$$

Dans la formulation variationnelle équivalente à $(\mathcal{P})$ (voir la section 4.2.2), il est possible de choisir u_- comme fonction test. On obtient

$$\int_{\Omega} [\nabla u \cdot \nabla u_- + \alpha u u_-] \, d\Omega = \int_{\Omega} f u_- \, d\Omega.$$

Or, comme $u_- = 0$ dans Ω_+ et $u_- = u$ dans Ω_- , on a, d'une part,

$$\int_{\Omega} [\nabla u \cdot \nabla u_- + \alpha u u_-] \, d\Omega = \int_{\Omega_-} [|\nabla u_-|^2 + \alpha |u_-|^2] \, d\Omega \geq 0,$$

et comme $f \geq 0$ et $u_- \leq 0$ presque partout dans Ω , on a, d'autre part,

$$\int_{\Omega} f u_- \, d\Omega \leq 0,$$

ce qui aboutit à une contradiction. $\square$

On déduit de cette proposition le principe du maximum faible.

Théorème 4.4.2 Supposons $f \in L^2(\Omega)$ et $g \in \text{Im } \gamma_0$.

Si $\alpha > 0$ alors la solution u de $(\mathcal{P})$ vérifie presque partout dans Ω

$$\min(\alpha^{-1} \inf_{\Omega} f, \inf_{\partial\Omega} g) \leq u \leq \max(\alpha^{-1} \sup_{\Omega} f, \sup_{\partial\Omega} g).$$

Si $\alpha = 0$ alors la solution u de $(\mathcal{P})$ vérifie presque partout dans Ω

$$\begin{aligned} f \geq 0 &\Rightarrow \inf_{\partial\Omega} g \leq u, \\ f \leq 0 &\Rightarrow u \leq \sup_{\partial\Omega} g, \\ f = 0 &\Rightarrow \inf_{\partial\Omega} g \leq u \leq \sup_{\partial\Omega} g. \end{aligned}$$

Démonstration. Supposons $\alpha > 0$. Soit $K = \min(\alpha^{-1} \inf_{\Omega} f, \inf_{\partial\Omega} g)$. Si $K = -\infty$, il est clair que $u \geq -\infty$. Sinon K vérifie

$$-\Delta K + \alpha K = \min(\inf_{\Omega} f, \alpha \inf_{\partial\Omega} g) \quad \text{dans } \Omega$$

. La fonction $v = u - K$ vérifie donc

$$\begin{cases} -\Delta v + \alpha v = f - \min \left(\inf_{\Omega} f, \alpha \inf_{\partial\Omega} g \right) & (\geq 0) \quad \text{dans } \Omega, \\ v = g - \min \left(\alpha^{-1} \inf_{\Omega} f, \inf_{\partial\Omega} g \right) & (\geq 0) \quad \text{sur } \partial\Omega. \end{cases}$$

D'après la proposition précédente $v \geq 0$ et donc $u \geq K$. Toutes les autres inégalités se démontrent de la même façon. $\square$

4.4.2 Résultats de régularité

L'approche variationnelle vue dans ce chapitre ne fournit que des solutions faibles qui sont pour tous les problèmes étudiés dans ce chapitre dans $H^1(\Omega)$. Une question naturelle est maintenant de se demander si ces solutions, définies de manière unique, ne sont pas plus régulières et dans quelles conditions elles pourraient l'être.

Cette question est très technique, car elle dépend de la régularité de la frontière et de la régularité des données. Dans les premiers résultats que nous énonçons, nous supposons que la frontière est assez régulière. Dans un second temps, nous considérerons le cas de domaines seulement polygonaux dans $\mathbb{R}^2$ ou polyédrique dans $\mathbb{R}^3$. Dans ce dernier cas, nous verrons que la régularité des solutions dépend notamment du caractère convexe ou non du domaine! Nous verrons en particulier dans un cas particulier de domaine polygonal non convexe, le type de singularité que peut présenter la solution faible. Nous finirons par énoncer un résultat de régularité locale des solutions.

Tous les résultats de cette section ne seront pas démontrés. Nous renvoyons le lecteur à [17, 18] qui sont, il nous semble, des livres de référence dans ce domaine.

Régularité globale des solutions. Nous n'énonçons que les résultats de régularité pour les problèmes de Dirichlet.

Remarque 4.4.1 Des résultats similaires peuvent être obtenus pour les problèmes de Neumann. Dans le cas où les conditions aux limites sont mixtes (Dirichlet sur une partie du bord et Neumann sur l'autre partie du bord), les résultats de régularité sont beaucoup plus délicats et dépendent de la nature du "raccord" des conditions aux limites (voir encore [17, 18]).

Théorème 4.4.3 Soit Ω un ouvert borné de $\mathbb{R}^d$ dont la frontière est de classe C^2 . Soient $f \in L^2(\Omega)$ et $u \in H_0^1(\Omega)$ l'unique solution de $(\mathcal{P}_D)$. Alors $u \in H^2(\Omega)$ et

$$\exists C > 0, \quad \|u\|_{H^2} \leq C \|f\|_{L^2}.$$

Corollaire 4.4.4 Si Ω est un ouvert borné de $\mathbb{R}^d$ dont la frontière est de classe

C^{m+2} et si $f \in H^m(\Omega)$ alors l'unique solution de $(\mathcal{P}_D)$ est dans $H^{m+2}(\Omega)$ et

$$\exists C > 0, \quad \|u\|_{H^{m+2}} \leq C \|f\|_{H^m}.$$

Corollaire 4.4.5 Si Ω est un ouvert borné de $\mathbb{R}^d$ dont la frontière est de classe C^p , si $f \in H^m(\Omega)$ et si g est tel qu'il existe un de ces relèvements qui est $H^l(\Omega)$ alors l'unique solution de $(\mathcal{P}_g)$ est dans $H^q(\Omega)$ avec $q = \min(p, m + 2, l)$.

D'après le théorème 2.2.10, on déduit enfin quand les solutions sont des solutions classiques.

Corollaire 4.4.6 Si Ω est un ouvert borné de $\mathbb{R}^d$ dont la frontière est de classe C^{m+2} et si $f \in H^m(\Omega)$ avec $m > d/2$ alors l'unique solution de $(\mathcal{P}_D)$ est dans $C^2(\Omega)$ et est une solution classique de $(\mathcal{P}_D)$.

Tous ces résultats s'étendent aux problèmes aux limites avec conditions de Neumann ou de Robin, mais pas aux problèmes avec conditions mixtes (Dirichlet sur un bout de bord et Neumann sur le reste). Dans ce dernier cas, la réponse va dépendre de la nature des raccords et des données.

Nous allons maintenant énoncer une extension du théorème 4.4.3 dans le cas où Ω n'est pas de classe C^2 .

Théorème 4.4.7 Soit Ω un ouvert de $\mathbb{R}^d$ de frontière lipchitzienne. Soit $f \in L^2(\Omega)$, et $u \in H_0^1(\Omega)$ l'unique solution de $(\mathcal{P}_D)$. Si Ω est convexe alors $u \in H^2(\Omega)$ et

$$\exists C > 0, \quad \|u\|_{H^2(\Omega)} \leq C \|f\|_{L^2(\Omega)}.$$

Sinon, il n'est pas garanti que u soit dans $H^2(\Omega)$.

Ainsi, même si l'existence d'une solution $u \in H_0^1(\Omega)$ est garantie, cette fonction n'est pas toujours dans H^2 même pour un terme source très régulier.

Remarque 4.4.2 Pour un ouvert polygonal non convexe de $\mathbb{R}^2$ ou $\mathbb{R}^3$, on montre que pour tout $f \in L^2(\Omega)$, la solution de $(\mathcal{P}_D)$ est dans $H^{1+s}(\Omega)$ (voir la remarque 2.2.1 sur les espaces de Sobolev fractionnaires) avec $s < \sigma$ et $\sigma \in]\frac{1}{2}, 1[$ ne dépend que du domaine Ω . Cette perte de régularité est localisée au niveau des coins rentrants en 2D et des coins rentrants et des arêtes rentrantes en 3D. On montre en effet que $u \in H^2(\Omega \setminus \omega)$ où ω est un voisinage des coins rentrants en 2D et des coins rentrants et des arêtes rentrantes en 3D (voir aussi le théorème 4.4.8).

Exemple de solutions singulières. Pour s'en convaincre, nous allons construire pour un ouvert non convexe de $\mathbb{R}^2$ particulier, une solution de $(\mathcal{P}_D)$ qui n'est pas dans $H^2(\Omega)$ car elle présente un comportement singulier au voisinage des "coins" rentrants du domaine.

Soit le domaine angulaire suivant

$$\Omega := \{(r, \theta) \in \mathbb{R}^+ \times [0, 2\pi], 0 < r < R, 0 < \theta < \Theta\},$$

avec $0 < R < +\infty, 0 \leq \Theta \leq 2\pi$. Pour $\Theta \in (0, \pi)$ le domaine est convexe et pour $\Theta \in (\pi, 2\pi)$, le domaine ne l'est pas.

On note $\Gamma_0 := \{(r, 0), 0 < r < R\}$, $\Gamma_\Theta := \{(r, \Theta), 0 < r < R\}$ et $\Gamma_R := \{(R, \theta), 0 < \theta < \Theta\}$ les trois bouts du bord $\partial\Omega$. Cet ouvert Ω est régulier par morceaux, il comporte trois coins, mais un seul dit rentrant. C'est au voisinage de ce coin que la solution que nous allons construire n'est pas régulière¹¹. Introduisons la fonction pour tout $k \geq 1$ un entier

$$u_k(r, \theta) = \left(\frac{r}{R}\right)^{\frac{k\pi}{\Theta}} \sin\left(\frac{k\pi\theta}{\Theta}\right).$$

Montrons tout d'abord que $u_k \in H^1(\Omega)$ pour tout k . En effet, $u_k \in C^0(\overline{\Omega}) \subset L^2(\Omega)$. De plus,¹²

$$\nabla u_k(r, \theta) = \frac{k\pi}{R\Theta} \left(\frac{r}{R}\right)^{\frac{k\pi}{\Theta}-1} \left(\sin\left(\frac{k\pi\theta}{\Theta}\right) \underline{e}_r + \cos\left(\frac{k\pi\theta}{\Theta}\right) \underline{e}_\theta \right).$$

Ainsi, si $k \geq 2$ ou $\Theta \leq \pi$, ∇u_k est dans $[C^0(\overline{\Omega})]^2 \subset [L^2(\Omega)]^2$. Si $k = 1$ et $\Theta > \pi$, ∇u_1 n'est pas bornée au voisinage de l'origine, mais

$$\|\nabla u_1\|_{L^2}^2 = \int_0^\Theta \int_0^R |\nabla u(r, \theta)|^2 r dr d\theta \leq C \int_0^R r^{\frac{2\pi}{\Theta}-1} dr < +\infty.$$

Nous souhaitons calculer maintenant¹³ Δu . Un calcul simple montre que

$$\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r}\right) \left(\frac{r}{R}\right)^{\frac{k\pi}{\Theta}} = \left(\frac{k\pi}{\Theta}\right)^2 \frac{1}{r^2} \left(\frac{r}{R}\right)^{\frac{k\pi}{\Theta}},$$

ce qui permet de déduire que

$$\Delta u_k = 0 \quad \text{dans } \Omega.$$

Pour tout $k \geq 1$, nous avons donc construit une solution dans $H^1(\Omega)$ de

$$\begin{cases} \Delta u_k = 0 & \text{dans } \Omega, \\ u_k = 0 & \text{sur } \Gamma_0 \cup \Gamma_\Theta, \\ u_k = \sin\left(\frac{k\pi\theta}{\Theta}\right) & \text{sur } \Gamma_R. \end{cases}$$

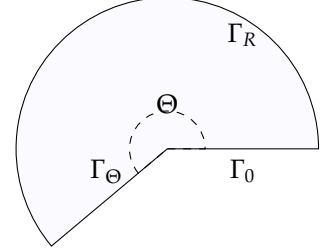
D'après la section 4.2.2, c'est l'unique solution¹⁴.

Est-ce que u_k est dans $H^2(\Omega)$? Calculons par exemple

$$\frac{1}{r^2} \frac{\partial^2 u_k}{\partial \theta^2} = -\left(\frac{k\pi}{R\Theta}\right)^2 \left(\frac{r}{R}\right)^{k\pi/\Theta-2} \sin\left(\frac{k\pi\theta}{\Theta}\right).$$

Cette fonction n'est pas dans $L^2(\Omega)$ pour $k = 1$ et $\Theta > \pi$. En conclusion, u_1 n'est pas dans $H^2(\Omega)$ dès que $\Theta > \pi$. On pourrait montrer néanmoins que $u_1 \in H^{1+s}(\Omega)$ avec $s < \pi/\Theta$.

Régularité locale des solutions Nous finissons ce chapitre par un résultat de régularité locale précisant que si le terme source est assez



11: C'est d'ailleurs un résultat général : dans le cas où Ω est un polygone, la solution de $(\mathcal{P}_D)$ est régulière, en l'occurrence H^2 au voisinage des coins sortants et ne l'est pas au voisinage des coins entrants, son comportement dépendant, comme nous allons le voir dans cet exemple, de l'angle du coin.

12: Rappelons que

$$\nabla u(r, \theta) := \frac{\partial u}{\partial r} \underline{e}_r + \frac{1}{r} \frac{\partial u}{\partial \theta} \underline{e}_\theta$$

13: Rappelons que

$$\Delta u(r, \theta) = \frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2}.$$

14: Notons qu'ici la donnée de Dirichlet est bien dans $\text{Im } \gamma_0$ puisque nous en avons construit un relèvement.

régulier, ce type de singularité ne peut apparaître que proche des bords.

Theorème 4.4.8 Soit Ω un ouvert borné de $\mathbb{R}^d$ de frontière lipchitzienne. Soit $f \in L^2(\Omega)$, et $u \in H^1(\Omega)$ l'unique solution de $(\mathcal{P}_D)$, $(\mathcal{P}_g)$ ou $(\mathcal{P}_N)$. Soit $\mathcal{B}(C, R)$ une boule de centre C et de rayon R strictement incluse dans Ω . Alors si $f \in H^m(\mathcal{B}(C, R))$ alors $u \in H^{m+2}(\mathcal{B}(C, R'))$ avec $R' < R$.

La méthode des éléments finis constitue l'une des grandes classes de méthodes pour l'approximation de problèmes aux limites. Elle est sans doute la plus célèbre stratégie d'approximation interne d'une formulation variationnelle, comme introduite au paragraphe 3.6. L'objectif de ce chapitre est de présenter la méthode dans le cadre unidimensionnel, pour lequel les calculs sont simplifiés compte tenu de la géométrie linéique et ordonnée.

5.1 Introduction

Afin de présenter la méthode des éléments finis, nous allons considérer un exemple caractéristique. Nous cherchons dans ce chapitre à approcher le problème en dimension 1 d'espace posé sur le domaine $\Omega =]0, L[$, qui s'écrit à l'aide de la formulation variationnelle :

$$\text{Trouver } u \in \mathcal{V} \text{ tel que } a(u, v) = \ell(v) \quad \forall v \in \mathcal{V}, \quad (5.1)$$

où $\mathcal{V} = H^1(\Omega)$,

$$\forall (u, v) \in \mathcal{V}^2, \quad a(u, v) = \int_0^L \kappa(x) u'(x) v'(x) + u(x) v(x) \, dx, \quad (5.2)$$

et

$$\forall v \in \mathcal{V}, \quad \ell(v) = \int_0^L f(x) v(x) \, dx, \quad (5.3)$$

avec $\kappa \in C^0(\Omega)$ et $f \in C^0(\Omega)$.

Nous allons introduire des sous-espaces de dimension finie de l'espace $\mathcal{V}$ construits à partir de la discrétisation de la grille linéique indexée par un petit paramètre h caractérisant la taille du pas de discrétisation de la grille et inversement proportionnel au nombre de points constituant la grille, et donc à la dimension N de l'espace d'approximation de $\mathcal{V}$. Le problème variationnel discret sera typiquement une approximation interne au sens de (3.19) de la forme

$$\text{Trouver } u_h \in \mathcal{V}_h \text{ tel que } a(u_h, v_h) = \ell(v_h) \quad \forall v_h \in \mathcal{V}_h, \quad (5.4)$$

où $\mathcal{V}_h \subset \mathcal{V}$ est l'espace d'approximation.

5.2 Élément fini de Lagrange $\mathbb{P}_1$ en dimension 1

5.2.1 Maillage et fonctions de forme

Soit $N_i \in \mathbb{N}$, nous considérons tout d'abord une discrétisation du domaine $]0, L[$ en un ensemble de points $(x_i)_{0 \leq i \leq N_i+1}$, appelé *sommets* tels que

$$0 = x_0 < x_1 < \dots < x_{N_i} < x_{N_i+1} = L.$$

Cette grille comporte donc N_i sommets intérieurs et $N_s = N_i + 2$ sommets au total. On introduit le *pas de discrétisation*

$$h = \max_{0 \leq i \leq N_i} (|x_{i+1} - x_i|).$$

La discrétisation sera dite *uniforme* si les x_i sont équidistants, c'est-à-dire

$$\forall 0 \leq i \leq N_i + 1, \quad x_i = ih \text{ où } h = \frac{1}{N_i + 1}.$$

Le pas de discrétisation tend vers 0 quand le nombre de sommets augmente. Cet ensemble de sommets permet de construire un *maillage* $\mathcal{I}_h$ du domaine constitué des $N_e = N_i + 1$ segments jointifs $\{[x_l, x_{l+1}]\}_{l=0}^{N_i}$. L'indice h permet d'identifier la discrétisation au travers de son pas de discrétisation.

La méthode des éléments finis repose sur l'utilisation de fonctions définies sur chaque segment du maillage. Pour les éléments finis de Lagrange de degré 1, l'approximation sur chaque segment $[x_l, x_{l+1}]$ est affine c'est-à-dire est un polynôme de degré 1. On notera $\mathbb{P}_1$ l'espace des polynômes de degré 1. L'espace éléments finis $\mathbb{P}_1$ associé est alors¹

1: l'usage est d'utiliser aussi h pour identifier l'espace éléments finis et non pas sa dimension N comme précédemment

$$\mathcal{V}_h = \mathcal{V}_h^{\mathbb{P}_1}(\mathcal{I}_h) := \left\{ v_h \in C^0([0, L]) \mid \forall 0 \leq l \leq N_i, \right. \\ \left. v_h|_{[x_l, x_{l+1}]} \in \mathbb{P}_1([x_l, x_{l+1}]) \right\}, \quad (5.5)$$

et si on souhaite fixer les valeurs aux extrémités

$$\mathcal{V}_{0h} = \left\{ v_h \in \mathcal{V}_h^{\mathbb{P}_1}(\mathcal{I}_h) \mid v_h(0) = v_h(L) = 0 \right\}. \quad (5.6)$$

En utilisant la théorème 2.2.2, on montre aisément la proposition suivante.

Proposition 5.2.1 Soient $\mathcal{V}_h$ et $\mathcal{V}_{0h}$ définis respectivement en (5.5) et (5.6). On a $\mathcal{V}_h \subset H^1([0, L])$ et $\mathcal{V}_{0h} \subset H_0^1([0, L])$.

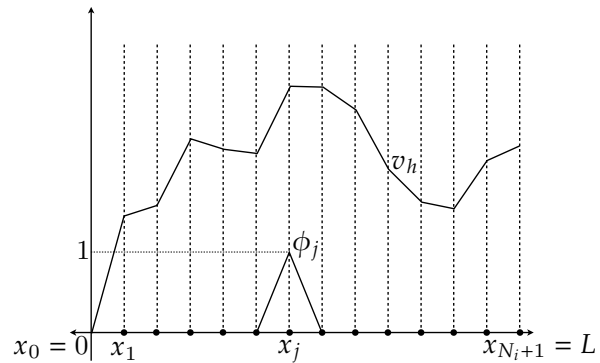


FIGURE 5.1 – Approximation par des fonctions de base $\mathbb{P}_1$

Cas des conditions de Dirichlet. Commençons par étudier le cas où les extrémités sont justement fixées. Une base naturelle de $\mathcal{V}_{0h}$ est alors

construite à partir des « fonctions chapeau » suivantes

$$\forall 1 \leq i \leq N_i, \quad \phi_j(x) = \begin{cases} \frac{x - x_{j-1}}{x_j - x_{j-1}}, & \text{si } x \in [x_{j-1}, x_j], \\ \frac{x_{j+1} - x}{x_{j+1} - x_j}, & \text{si } x \in [x_j, x_{j+1}], \\ 0, & \text{sinon.} \end{cases}$$

Theorème 5.2.2 La famille $(\phi_j)_{1 \leq j \leq N_i}$ est une base de $\mathcal{V}_{0h}$ de dimension $N = N_i$.

Démonstration. Supposons que $(v^j)_{1 \leq j \leq N_i}$ vérifiant

$$\forall x \in [0, L], \quad \sum_{1 \leq j \leq N_i} v^j \phi_j(x) = 0.$$

On a donc pour $x = x_i$

$$0 = \sum_{1 \leq j \leq N_i} v^j \phi_j(x_i) = v^i,$$

donc la famille est libre.

Considérons alors $v_h \in \mathcal{V}_{0h}$ et un point $x \in (x_i, x_{i+1})$. Il existe $t \in [0, 1]$ tel que $x = tx_i + (1 - t)x_{i+1}$. Comme v_h est affine, alors

$$v_h(x) = tv_h(x_i) + (1 - t)v_h(x_{i+1}).$$

Si on note alors $w_h = \sum_{1 \leq j \leq N} v_h(x_j) \phi_j$, on a par linéarité

$$w_h(x) = tw_h(x_i) + (1 - t)w_h(x_{i+1}) = tv_h(x_i) + (1 - t)v_h(x_{i+1}).$$

Donc v_h et w_h coïncident sur tout segment (x_i, x_{i+1}) . Elles coïncident donc partout, et la famille $(\phi_j)_{1 \leq j \leq N}$ est donc génératrice. $\square$

Ainsi toute fonction $u_h \in \mathcal{V}_{0h}$ se décompose dans la base des $(\phi_j)_{1 \leq j \leq N_i}$ sous la forme

$$\forall x \in \Omega, \quad u_h(x) = \sum_{1 \leq j \leq N_i} u_h(x_j) \phi_j(x). \quad (5.7)$$

5.2.2 Matrices de raideur et de masse – cas constant et conditions de Dirichlet

On déduit directement, avec a et ℓ données respectivement par (5.2) et (5.3), les expressions spécifiques de (3.21) avec $\vec{u}_h = (u_h(x_j))_{1 \leq j \leq N}$ et

$$\vec{\ell}_h = \left(\int_0^L f(x) \phi_i(x) dx \right)_{1 \leq i \leq N},$$

ainsi que $\mathbb{A} = \mathbb{K}_h + \mathbb{M}_h$, où

$$\mathbb{K}_h = \left(\int_0^L \kappa(x) \phi_j'(x) \phi_i'(x) dx \right)_{1 \leq i, j \leq N},$$

et

$$\mathbb{M}_h = \left(\int_0^L \phi_j(x) \phi_i(x) \, dx \right)_{1 \leq i, j \leq N}.$$

On appelle couramment $\mathbb{M}_h$ la *matrice de masse*, car elle correspond à la représentation matricielle du terme associé à la masse de la corde déformée au chapitre 1, paragraphe 1.1.1. Et respectivement, on appelle $\mathbb{K}_h$ la *matrice de raideur ou de rigidité* du problème, car elle correspond à la représentation matricielle du terme associé à la raideur de la corde.

En résolvant le système linéaire $(\mathbb{K}_h + \mathbb{M}_h) \vec{u}_h = \vec{\ell}_h$, on obtient le vecteur $\vec{u}_h$ dont les composantes sont les valeurs de u_h aux sommets $(x_i)_{1 \leq i \leq N_i}$ du maillage.

Pour obtenir des expressions plus explicites des matrices de raideur et de masse, plaçons-nous tout d'abord dans le cas où $\kappa \equiv \bar{\kappa} \in \mathbb{R}_+^*$ est une fonction constante. Nous pouvons alors exploiter la structure particulière des fonctions de base $\{\phi_j\}_{j=1}^N$ dont le support est fini et réduit à deux intervalles consécutifs. L'intersection des supports de ϕ_j et ϕ_i est vide dès que $|i - j| > 2$, et la plupart des coefficients de $\mathbb{K}_h$ et $\mathbb{M}_h$ sont nuls. Un calcul simple montre que pour tout $(i, j) \in \llbracket 1, N \rrbracket^2$,

$$\int_0^L \phi_j'(x) \phi_i'(x) \, dx = \begin{cases} -\frac{1}{x_{j+1} - x_j} & \text{si } j = i - 1, \\ \frac{1}{x_j - x_{j-1}} + \frac{1}{x_{j+1} - x_j}, & \text{si } j = i, \\ -\frac{1}{x_j - x_{j-1}}, & \text{si } j = i + 1, \\ 0, & \text{sinon.} \end{cases}$$

et, dans le cas du maillage uniforme, on a

$$\forall (i, j) \in \llbracket 1, N \rrbracket^2, \quad \int_0^L \phi_j'(x) \phi_i'(x) \, dx = \begin{cases} -h^{-1}, & \text{si } j = i - 1, \\ 2h^{-1}, & \text{si } j = i, \\ -h^{-1}, & \text{si } j = i + 1, \\ 0 & \text{sinon,} \end{cases}$$

La matrice $\mathbb{K}_h$ est donc tridiagonale et on a en particulier pour le maillage uniforme

$$\mathbb{K}_h = \mathbb{K}_{0h} := \frac{\bar{\kappa}}{h} \begin{pmatrix} 2 & -1 & & 0 \\ -1 & 2 & -1 & \\ & \ddots & \ddots & \ddots \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix}.$$

De même dans le cas général, en intégrant des polynômes du second degré, on a pour tout $(i, j) \in \llbracket 1, N \rrbracket^2$

$$\int_0^L \phi_j(x) \phi_i(x) \, dx = \begin{cases} \frac{x_{j+1} - x_j}{6} & \text{si } j = i - 1, \\ \frac{x_j - x_{j-1}}{3} + \frac{x_{j+1} - x_j}{3}, & \text{si } j = i, \\ \frac{x_j - x_{j-1}}{6}, & \text{si } j = i + 1, \\ 0, & \text{sinon,} \end{cases}$$

et, dans le cas du maillage uniforme, on a

$$\forall (i, j) \in \llbracket 1, N \rrbracket^2, \quad \int_0^L \phi_j(x) \phi_i(x) dx = \begin{cases} \frac{h}{6}, & \text{si } j = i - 1, \\ \frac{2h}{3}, & \text{si } j = i, \\ \frac{h}{6}, & \text{si } j = i + 1, \\ 0 & \text{sinon.} \end{cases}$$

La matrice $\mathbb{M}_h$ est donc elle aussi tridiagonale et on a en particulier pour le maillage uniforme

$$\mathbb{M}_h = \mathbb{M}_{0h} := \frac{h}{6} \begin{pmatrix} 4 & 1 & & 0 \\ 1 & 4 & 1 & \\ & \ddots & \ddots & \ddots \\ & & 1 & 4 & 1 \\ 0 & & & 1 & 4 \end{pmatrix}.$$

Proposition 5.2.3 Les matrices $\mathbb{K}_{0h}$ et $\mathbb{M}_{0h}$ sont inversibles.

Nous donnons ici une preuve directement fondée sur la connexion entre les matrices $\mathbb{K}_{0h}$ et $\mathbb{M}_{0h}$ et les formulations bilinéaires associées.²

Démonstration. Commençons par $\mathbb{M}_{0h}$. Soit $u_h \in \mathcal{V}_{0h}$ et $\vec{u}_h$ le vecteur des degrés de liberté associé. D'après la définition de la matrice de masse, on a

$$\vec{u}_h^\top \mathbb{M}_{0h} \vec{u}_h = \int_0^L u_h(x)^2 dx \geq 0,$$

donc $\vec{u}_h \in \ker \mathbb{M}_{0h}$ implique $u_h \equiv 0$ et donc $\vec{u}_h = 0$.

De même pour $\mathbb{K}_{0h}$, on a

$$\vec{u}_h^\top \mathbb{K}_{0h} \vec{u}_h = \int_0^L u_h'(x)^2 dx \geq 0.$$

Ainsi $\vec{u}_h \in \ker \mathbb{K}_{0h}$ implique qu'il existe une constante $c \in \mathbb{R}$ tel que $u_h \equiv c$. Or $u_h(0) = u_h(L) = 0$ implique $c = 0$ et donc $\vec{u}_h = 0$. $\square$

2: Il existe des preuves qui exploitent plus directement la structure des matrices pour démontrer leur inversibilité. Pour $\mathbb{M}_{0h}$, elle est à diagonale strictement dominante et pour $\mathbb{K}_{0h}$ voir les derniers paragraphes du chapitre 7.

5.2.3 Calcul du second membre.

Il nous reste à calculer le second membre $\vec{\ell}_h$ correspondant à la forme linéaire $\ell \in \mathcal{V}_h'$. Il faut alors calculer les intégrales

$$\forall 1 \leq i \leq N, \quad \vec{\ell}_h^i = \int_{x_{i-1}}^{x_{i+1}} f(x) \phi_i(x) dx. \quad (5.8)$$

Ces intégrales n'ont pas nécessairement de forme analytique. Cependant si f appartient à $\mathcal{V}_h$ alors $f = f_h := \sum_{1 \leq j \leq N} f^j \phi_j$ et

$$\vec{\ell}_h = \mathbb{M}_h \vec{f}_h. \quad (5.9)$$

Cette formule peut être utilisée dans le cas général et correspond à approcher f par son *interpolé* sur l'espace élément finis $\mathcal{V}_h$, c'est à dire un élément de $\mathcal{V}_h$ reconstruit à partir des valeurs de f aux nœuds de la grille,

$$f \simeq f_h = \sum_{1 \leq j \leq N} f(x_j) \phi_j,$$

Ceci revient alors à remplacer dans la formulation variationnelle la forme linéaire ℓ par

$$\forall v \in \mathcal{V}_h, \quad \ell_h(v) = \int_0^L f_h(x) v(x) \, dx,$$

5.2.4 Procédure d'assemblage.

Ce que nous venons de voir dans l'exemple précédent est un premier pas dans la méthode des éléments finis avec une première construction de système linéaire de la forme (3.21). Cependant, la méthode de construction des opérateurs peut exploiter beaucoup plus intensément la forme de chaque fonction de base par élément afin de proposer un algorithme remarquablement général et efficace de construction des opérateurs matriciels associés aux formulations variationnelles.

Pour bien comprendre la procédure, reprenons notre exemple de problème variationnel à partir des formes bilinéaire (5.2) et linéaire (5.3) mais cette fois avec des extrémités libres. Dans ce cas le problème variationnel est formulé sur $\mathcal{V}_h$ où désormais

$$\mathcal{V}_h = \text{Vect} \left\{ \phi_j, j \in \llbracket 0, N_i + 1 \rrbracket \right\} \text{ et } N = \dim \mathcal{V}_h = N_i + 2 = N_s,$$

avec, en sus des fonctions $(\phi_j)_{1 \leq j \leq N}$ déjà introduites,

$$\phi_0(x) = \frac{x_1 - x}{x_1}, \quad \text{sur } [0, x_1],$$

et

$$\phi_{N_i+1}(x) = \frac{x - x_{N_i}}{x_{N_i+1} - x_{N_i}}, \quad \text{sur } [x_{N_i}, x_{N_i+1}].$$

Le principe est de décomposer les calculs d'intégrales comme une somme par élément, ici un intervalle $[x_i, x_{i+1}]$. Nous dénombrons $N_e = N_i + 1$ segments pour notre cas. Pour la matrice de raideur, on obtient

$$\int_0^L \phi'_j(x) \phi'_i(x) \, dx = \sum_{l=0}^{N_i} \int_{x_l}^{x_{l+1}} \phi'_j(x) \phi'_i(x) \, dx.$$

Or la fonction ϕ_i , respectivement ϕ_j , a un support sur chaque segment seulement si $i \in \{l, l+1\}$, respectivement $j \in \{l, l+1\}$. Et dans ce cas, on a tout de suite

$$\phi_l|_{[x_l, x_{l+1}]}(x) = \frac{x_{l+1} - x}{x_{l+1} - x_l} \Rightarrow \phi'_l(x) = \frac{-1}{x_{l+1} - x_l},$$

et

$$\phi_{l+1}|_{[x_l, x_{l+1}]}(x) = \frac{x - x_l}{x_{l+1} - x_l} \Rightarrow \phi'_{l+1}(x) = \frac{1}{x_{l+1} - x_l}.$$

Un calcul simple donne alors

$$\int_{x_l}^{x_{l+1}} \phi'_j(x) \phi'_i(x) dx = \begin{cases} 0, & (i, j) \notin \{l, l+1\}^2, \\ \frac{1}{x_{l+1} - x_l}, & (i, j) \in \{l, l+1\}^2 \text{ et } i = j, \\ \frac{-1}{x_{l+1} - x_l}, & (i, j) \in \{l, l+1\}^2 \text{ et } i \neq j. \end{cases}$$

Dans le cas du maillage uniforme, ce calcul se simplifie et donne

$$\mathbb{K}_l^{\text{el}} := \left(\int_{x_l}^{x_{l+1}} \phi'_j(x) \phi'_i(x) dx \right)_{(i,j) \in \{l, l+1\}^2} = \frac{1}{h} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Reste alors à sommer toutes les contributions dans la matrice correspondante. On le fait successivement dans l'algorithme 1 en ajoutant les contributions une à une lors d'une boucle sur les éléments. Cette procédure est appelée *procédure d'assemblage*. On note dans cet algorithme pour toute matrice $\mathbb{A}$, la sous-matrice

$$\mathbb{A}_{\llbracket p, q \rrbracket, \llbracket p, q \rrbracket} := (A_{ij})_{p \leq i, j \leq q}.$$

Algorithme 1: Procédure d'assemblage 1D

```

1  $\mathbb{K} \leftarrow 0$ ;
2  $\mathbb{M} \leftarrow 0$ ;
3 for  $l \in \{1, \dots, N_e\}$  do
4    $\mathbb{K}_{\llbracket l, l+1 \rrbracket, \llbracket l, l+1 \rrbracket} \leftarrow \mathbb{K}_{\llbracket l, l+1 \rrbracket, \llbracket l, l+1 \rrbracket} + \mathbb{K}_l^{\text{el}}$ ;
5    $\mathbb{M}_{\llbracket l, l+1 \rrbracket, \llbracket l, l+1 \rrbracket} \leftarrow \mathbb{M}_{\llbracket l, l+1 \rrbracket, \llbracket l, l+1 \rrbracket} + \mathbb{M}_l^{\text{el}}$ ;

```

La matrice de raideur résultante est alors

$$\mathbb{K}_h = \frac{\bar{\kappa}}{h} \begin{pmatrix} 1 & -1 & & & & & 0 \\ -1 & \ddots & \ddots & & & & \\ & \ddots & 1+1 & -1 & & & \\ & & -1 & +1+1 & -1 & & \\ & & & -1 & +1+1 & \ddots & \\ & & & & \ddots & \ddots & -1 \\ 0 & & & & & -1 & 1 \end{pmatrix},$$

c'est-à-dire

$$\mathbb{K}_h = \frac{\bar{\kappa}}{h} \begin{pmatrix} 1 & -1 & & & & & 0 \\ -1 & \ddots & \ddots & & & & \\ & \ddots & 2 & -1 & & & \\ & & -1 & 2 & -1 & & \\ & & & -1 & 2 & \ddots & \\ & & & & \ddots & \ddots & -1 \\ 0 & & & & & -1 & 1 \end{pmatrix}.$$

On remarque que le premier coefficient diagonal et le dernier sont égaux

à 1 alors que tous les autres coefficients sur la diagonale sont égaux à 2. Ceci est dû au fait que la procédure d'assemblage n'a donné qu'une contribution élémentaire pour les fonctions de bases ϕ_0 et ϕ_{N+1} qui correspondent désormais aux premières et dernières colonnes de la matrice.

De même on obtient pour la matrice de masse,

$$\int_{x_l}^{x_{l+1}} \phi_j(x) \phi_i(x) dx = \begin{cases} 0, & (i, j) \notin \{l, l+1\}^2, \\ \int_{x_l}^{x_{l+1}} \frac{(x - x_i)^2}{(x_{l+1} - x_l)^2} dx, & (i, j) \in \{l, l+1\}^2 \text{ et } i = j, \\ \int_{x_l}^{x_{l+1}} \frac{(x - x_l)(x_{l+1} - x)}{(x_{l+1} - x_l)^2} dx, & (i, j) \in \{l, l+1\}^2 \text{ et } i \neq j. \end{cases}$$

qui donne

$$\int_{x_l}^{x_{l+1}} \phi_j(x) \phi_i(x) dx = \begin{cases} 0, & (i, j) \notin \{l, l+1\}^2, \\ \frac{x_{l+1} - x_l}{3}, & (i, j) \in \{l, l+1\}^2 \text{ et } i = j, \\ \frac{x_{l+1} - x_l}{6}, & (i, j) \in \{l, l+1\}^2 \text{ et } i \neq j. \end{cases}$$

Dans le cas du maillage uniforme, ce calcul se simplifie et donne

$$\mathbb{M}_l^{\text{el}} := \left(\int_{x_l}^{x_{l+1}} \phi_j(x) \phi_i(x) dx \right)_{(i,j) \in \{l, l+1\}^2} = \frac{h}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Reste alors à sommer toutes les contributions dans la matrice correspondante. On le fait successivement en ajoutant les contributions une à une comme pour la matrice de raideur comme indiqué dans l'algorithme 1 d'assemblage. On obtient finalement

$$\mathbb{M}_h = \frac{h}{6} \begin{pmatrix} 2 & 1 & & & & & 0 \\ 1 & \ddots & \ddots & & & & \\ & \ddots & \textcolor{teal}{+2} \textcolor{teal}{+2} & \textcolor{teal}{+1} & & & \\ & & \textcolor{teal}{+1} & \textcolor{teal}{+2} \textcolor{teal}{+2} & \textcolor{red}{+1} & & \\ & & & \textcolor{red}{+1} & \textcolor{teal}{+2} \textcolor{teal}{+2} & \ddots & \\ & & & & \ddots & \ddots & 1 \\ 0 & & & & & 1 & 2 \end{pmatrix},$$

c'est-à-dire

$$\mathbb{M}_h = \frac{h}{6} \begin{pmatrix} 2 & 1 & & & & & 0 \\ 1 & \ddots & \ddots & & & & \\ & \ddots & 4 & 1 & & & \\ & & 1 & 4 & 1 & & \\ & & & 1 & 4 & \ddots & \\ & & & & \ddots & \ddots & 1 \\ 0 & & & & & 1 & 2 \end{pmatrix}.$$

De nouveau le premier et le dernier coefficient sur la diagonale diffèrent.

5.2.5 Prise en compte des conditions aux limites après la procédure d'assemblage.

Dans la construction précédente, nous avons pris les éléments successivement et obtenu les matrices associées à la forme bilinéaire (5.2) sur $\mathcal{V}_h$ et donc avec des extrémités libres. Pour retrouver la formulation posée sur $\mathcal{V}_{0h}$, nous remarquons alors que

$$\mathbb{K}_h = \begin{pmatrix} \frac{\bar{\kappa}}{h} & -\frac{\bar{\kappa}}{h} & 0 & \cdots & 0 \\ -\frac{\bar{\kappa}}{h} & & & & \vdots \\ 0 & & \mathbb{K}_{0h} & & 0 \\ \vdots & & & & -\frac{\bar{\kappa}}{h} \\ 0 & \cdots & 0 & -\frac{\bar{\kappa}}{h} & \frac{\bar{\kappa}}{h} \end{pmatrix},$$

et de même

$$\mathbb{M}_h = \begin{pmatrix} \frac{h}{3} & \frac{h}{6} & 0 & \cdots & 0 \\ \frac{h}{6} & & & & \vdots \\ 0 & & \mathbb{M}_{0h} & & 0 \\ \vdots & & & & \frac{h}{6} \\ 0 & \cdots & 0 & \frac{h}{6} & \frac{h}{3} \end{pmatrix},$$

ainsi que le second membre

$$\vec{\ell}_h = \mathbb{M}_h \vec{f}_h = \begin{pmatrix} \frac{h}{3}f(x_0) + \frac{h}{6}f(x_1) \\ \vec{\ell}_{0h} \\ \frac{h}{6}f(x_{N_i-1}) + \frac{h}{3}f(x_{N_i}) \end{pmatrix}.$$

On en déduit qu'un algorithme pour calculer les matrices de masse et de raideur sur $\mathcal{V}_{0h}$ est d'assembler d'abord sur $\mathcal{V}_h$ puis d'extraire les sous matrices $\mathbb{K}_{0h}$ et $\mathbb{M}_{0h}$ correspondant aux degrés de liberté qui ne sont pas associés à des nœuds du bord soit ici les ϕ_i pour $i \notin \{0, N_i + 1\}$. Cette stratégie est appelée la prise en compte des conditions aux limites *par élimination* de telle sorte qu'on cherche finalement $\vec{u} \in \mathbb{R}^{N_i}$ solution de

$$(\mathbb{K}_{0h} + \mathbb{M}_{0h})\vec{u} = \vec{\ell}_h.$$

Un algorithme alternatif appelé *pseudo-élimination* consiste à modifier $\mathbb{K}_h$ et $\mathbb{M}_h$ en

$$\tilde{\mathbb{K}}_{0h} = \begin{pmatrix} \varrho & 0 & 0 & \cdots & 0 \\ 0 & & & & \vdots \\ 0 & & \mathbb{K}_{0h} & & 0 \\ \vdots & & & & 0 \\ 0 & \cdots & 0 & 0 & \varrho \end{pmatrix},$$

et

$$\tilde{\mathbb{M}}_{0h} = \begin{pmatrix} \varrho & 0 & 0 & \cdots & 0 \\ 0 & & & & \vdots \\ 0 & & \mathbb{M}_{0h} & & 0 \\ \vdots & & & & 0 \\ 0 & \cdots & & 0 & 0 & \varrho \end{pmatrix},$$

alors que pour le second membre devient

$$\vec{\ell}_{0h} = \begin{pmatrix} 0 \\ \vec{\ell}_{0h} \\ 0 \end{pmatrix}.$$

Dans ce cas on cherche $\vec{u} \in \mathbb{R}^{N_i+2}$ solution de

$$(\tilde{\mathbb{K}}_{0h} + \tilde{\mathbb{M}}_{0h})\vec{u} = \vec{\ell}_{0h}.$$

et on vérifie immédiatement que $\vec{u} = (0 \quad \vec{u} \quad 0)^\top$ donne à la fois la solution du problème de Dirichlet pour les degrés de liberté intérieurs ainsi que la condition $\vec{u}_0 = \vec{u}_{N_i+1} = 0$ aux extrémités.

Remarque 5.2.1 Le paramètre ϱ est libre et peut être fixé comme on le souhaite. Il peut cependant polluer le spectre de la matrice $\mathbb{K}_{0h} + \mathbb{M}_{0h}$ avec deux valeurs propres égales à ϱ supplémentaires pour la matrice $\tilde{\mathbb{K}}_{0h} + \tilde{\mathbb{M}}_{0h}$, voir aussi le chapitre 8.

5.2.6 Passage à l'élément de référence.

Les différents calculs analytiques d'intégrales que nous venons de voir peuvent être simplifiés par un changement de variable. C'est le principe du passage à l'*élément de référence* qui ici sera un segment générique $[0, 1]$.

3: Parfois, le segment de référence est choisi comme étant $[-1, 1]$ sans que cela ne modifie les calculs autrement que par un facteur d'échelle.

On définit la transformation passant du segment de référence³ $[0, 1]$ au segment $[x_l, x_{l+1}] \subset [0, L]$

$$\varphi_l : [0, 1] \ni \hat{x} \mapsto x = x_l + (x_{l+1} - x_l)\hat{x} \in [x_l, x_{l+1}].$$

On a notamment $\varphi'_l(\hat{x}) = x_{l+1} - x_l > 0$. Par changement de variable, on obtient donc pour la masse

$$\int_{x_l}^{x_{l+1}} \phi_i(x)\phi_j(x) dx = \int_0^1 \phi_i(\varphi_l(\hat{x}))\phi_j(\varphi_l(\hat{x}))\varphi'_l(\hat{x}) d\hat{x},$$

où on est ramené à manipuler

$$\hat{\lambda}_0(\hat{x}) := \phi_l|_{[x_l, x_{l+1}]} \circ \varphi_l(\hat{x}) = 1 - \hat{x},$$

et

$$\hat{\lambda}_1(\hat{x}) := \phi_{l+1}|_{[x_l, x_{l+1}]} \circ \varphi_l(\hat{x}) = \hat{x}.$$

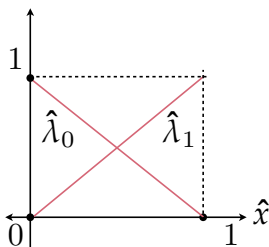


FIGURE 5.2 – Fonctions de formes $\mathbb{P}_1$ en 1D

Les $(\hat{\lambda}_m)_{0 \leq m \leq 1}$ sont appelées *fonctions de base locales* ou *fonctions de forme*.
On obtient ainsi immédiatement

$$\mathbb{M}_l^{\text{el}} := \left(\int_{x_l}^{x_{l+1}} \phi_i(x) \phi_j(x) \, dx \right)_{(i,j) \in \{l, l+1\}^2} = (x_{l+1} - x_l) \hat{\mathbb{M}}^{\text{el}},$$

où

$$\hat{\mathbb{M}}^{\text{el}} = \begin{pmatrix} \int_0^1 \hat{\lambda}_0(\hat{x})^2 \, d\hat{x} & \int_0^1 \hat{\lambda}_0(\hat{x}) \hat{\lambda}_1(\hat{x}) \, d\hat{x} \\ \int_0^1 \hat{\lambda}_0(\hat{x}) \hat{\lambda}_1(\hat{x}) \, d\hat{x} & \int_0^1 \hat{\lambda}_1(\hat{x})^2 \, d\hat{x} \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

est la *matrice de masse élémentaire* sur l'élément de référence.

De même pour la matrice de raideur, le changement de variable donne cette fois

$$\int_{x_l}^{x_{l+1}} \phi'_i(x) \phi'_j(x) \, dx = \int_0^1 \phi'_i(\varphi_l(\hat{x})) \phi'_j(\varphi_l(\hat{x})) \varphi'_l(\hat{x}) \, d\hat{x}.$$

Or

$$(\phi_i \circ \varphi_l)'(\hat{x}) = \varphi'_l(\hat{x}) \phi'_i(\varphi_l(\hat{x})),$$

permet d'écrire

$$\int_{x_l}^{x_{l+1}} \phi'_i(x) \phi'_j(x) \, dx = \int_0^1 (\phi_i \circ \varphi_l)'(\hat{x}) (\phi_j \circ \varphi_l)'(\hat{x}) (\varphi'_l(\hat{x}))^{-1} \, d\hat{x}.$$

On obtient cette fois

$$\mathbb{K}_l^{\text{el}} := \bar{\kappa} \left(\int_{x_l}^{x_{l+1}} \phi'_i(x) \phi'_j(x) \, dx \right)_{(i,j) \in \{l, l+1\}^2} = \frac{\bar{\kappa}}{x_{l+1} - x_l} \hat{\mathbb{K}}^{\text{el}},$$

où

$$\hat{\mathbb{K}}^{\text{el}} = \begin{pmatrix} \int_0^1 \hat{\lambda}'_0(\hat{x})^2 \, d\hat{x} & \int_0^1 \hat{\lambda}'_0(\hat{x}) \hat{\lambda}'_1(\hat{x}) \, d\hat{x} \\ \int_0^1 \hat{\lambda}'_0(\hat{x}) \hat{\lambda}'_1(\hat{x}) \, d\hat{x} & \int_0^1 \hat{\lambda}'_1(\hat{x})^2 \, d\hat{x} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},$$

et ces intégrales étaient effectivement très simples à calculer puisque $\lambda'_0(\hat{x}) = -1$ et $\lambda'_1(\hat{x}) = 1$.

5.2.7 Cas des coefficients variables et formules de quadrature

Recourir à l'élément de référence est d'autant plus intéressant lorsqu'on considère des paramètres eux-mêmes fonctions de l'espace. Reprenons donc notre exemple avec cette fois pour la matrice de raideur

$$\mathbb{K}_h = \left(\int_0^1 \kappa(x) \phi'_j(x) \phi'_i(x) \, dx \right)_{1 \leq i, j \leq N}.$$

Après décomposition sur chaque segment et passage à l'élément de référence, on est ramené à calculer

$$\mathbb{K}_l^{\text{el}} = \begin{pmatrix} \int_0^1 (\kappa \circ \varphi_l)(\hat{x}) \hat{\lambda}'_0(\hat{x})^2 d\hat{x} & \int_0^1 (\kappa \circ \varphi_l)(\hat{x}) \hat{\lambda}'_0(\hat{x}) \hat{\lambda}'_1(\hat{x}) d\hat{x} \\ \int_0^1 (\kappa \circ \varphi_l)(\hat{x}) \hat{\lambda}'_0(\hat{x}) \hat{\lambda}'_1(\hat{x}) d\hat{x} & \int_0^1 (\kappa \circ \varphi_l)(\hat{x}) \hat{\lambda}'_1(\hat{x})^2 d\hat{x} \end{pmatrix},$$

qui dépend désormais de $l \in \llbracket 0, N_i \rrbracket$. Comme il n'existe pas de solution analytique pour ces intégrales, on a recours à des formules d'intégration numérique appelées *formules de quadratures*.

Définition 5.2.1 (Formule de quadrature en dimension 1) Soit $[a, b] \subset \mathbb{R}$ un intervalle borné et $\psi \in C^{m+1}([a, b])$. Considérons l'intégrale définie par $\mathcal{J} = \int_a^b \psi(x) dx$ dont on cherche à estimer la valeur numérique. Une formule de quadrature se présente sous la forme

$$\mathcal{J}_{N_q} = \sum_{1 \leq k \leq N_q} \alpha_k \psi(q_k)$$

où $(q_k)_{1 \leq k \leq N_q} \in (a, b)^{N_q}$ est un ensemble de points dits de quadratures et $(\alpha_k)_{1 \leq k \leq N_q}$ est un ensemble de poids associés. On dira que la formule est d'ordre m si elle est exacte pour tout polynôme ψ de degré m .

Proposition 5.2.4 Si $(q_k)_{1 \leq k \leq N_q} \in (a, b)^{N_q}$ sont les points et $(\alpha_k)_{1 \leq k \leq N_q}$ les poids associés à une quadrature d'ordre m sur $[a, b]$ alors il existe une constante $C \in \mathbb{R}_+$ telle que l'erreur de quadrature

$$\psi \in C^{m+1}([a, b]), \quad \mathcal{R}_{N_q} := |\mathcal{J} - \mathcal{J}_{N_q}| \leq C(b-a)^{m+2} \|\psi^{(m+1)}\|_\infty.$$

Démonstration. Comme $\psi \in C^{m+1}([a, b])$, alors pour tout $x \in [a, b]$, il existe $\xi(x)$ tel que

$$\psi(x) = \sum_{k=0}^m \frac{\psi^{(k)}(a)}{k!} (x-a)^k + \frac{\psi^{(m+1)}(\xi(x))}{(m+1)!} (x-a)^{m+1}.$$

On note $p_m(x) = \sum_{k=0}^m \frac{\psi^{(k)}(a)}{k!} (x-a)^k$. Comme $p_m \in \mathbb{P}_m([a, b])$, alors

$$\mathcal{J}(\psi) - \mathcal{J}_{N_q} = \mathcal{J}(\psi - p_m) - \mathcal{J}_{N_q}(\psi - p_m).$$

Or on a

$$\mathcal{J}(\psi - p_m) \leq (b-a) \|\psi - p_m\|_\infty,$$

et de même en utilisant le fait que $\sum_{k=1}^{N_q} \alpha_k = b-a$ (car la formule de quadrature est exacte sur les constantes)

$$\mathcal{J}_{N_q}(\psi - p_m) \leq (b-a) \|\psi - p_m\|_\infty.$$

On en déduit que

$$\mathcal{R}_{N_q} \leq 2(b-a) \|\psi - p_m\|_\infty \leq 2 \frac{(b-a)^{m+2}}{(m+1)!} \|\psi^{(m+1)}\|_\infty.$$

□

Nous proposons d'introduire les *formules dites de Gauss*⁴ particulièrement bien adaptées aux calculs qui nous intéressent. Le principe des formules de Gauss est de choisir les points de quadrature tel que la précision soit la plus élevée possible, c'est-à-dire d'ordre $2N_q + 1$.

Historiquement dans les formules de Gauss, les intégrales sont calculées sur $[-1, 1]$ mais il n'est pas difficile de replacer ensuite les points sur $[0, 1]$ via une simple dilatation de l'intégrale. Du changement de variable

$$\int_{-1}^1 \psi(x) dx = 2 \int_0^1 \psi(\hat{x}) d\hat{x}$$

on déduit en effet que si $(q_i)_{1 \leq i \leq N_q}$ sont des points de quadrature et $(\alpha_i)_{1 \leq i \leq N_q}$ sont des poids adaptés sur $[-1, 1]$, alors $(\frac{1-q_i}{2})_{1 \leq i \leq N_q}$ et $(\frac{\alpha_i}{2})_{1 \leq i \leq N_q}$ sont des points de quadrature et des poids adaptés sur $[0, 1]$.

Proposition 5.2.5 (Formule de Gauss à 2 points) Soit $\psi \in C^2([-1, 1])$, alors les points de quadrature de Gauss $q_0 = \frac{-1}{\sqrt{3}}$ et $q_1 = \frac{1}{\sqrt{3}}$ et les poids $\alpha_0 = \alpha_1 = 1$ assurent que la formule est d'ordre 3 :

$$\forall p \in \mathbb{P}_3([-1, 1]), \quad \int_{-1}^1 p(x) dx = \sum_{k=0}^1 \alpha_k p(q_k).$$

Démonstration. Il suffit de vérifier que l'ordre est préservé pour les monômes

$$\forall n \in \{0, 1, 2, 3\}, \quad \int_{-1}^1 x^n dx = \sum_{k=0}^1 \alpha_k q_k^n$$

ce qui donne le système suivant

$$\left| \begin{array}{lcl} \int_{-1}^1 1 dx & = & 2 = \alpha_0 + \alpha_1 \\ \int_{-1}^1 x dx & = & 0 = \alpha_0 q_0 + \alpha_1 q_1 \\ \int_{-1}^1 x^2 dx & = & \frac{2}{3} = \alpha_0 q_0^2 + \alpha_1 q_1^2 \\ \int_{-1}^1 x^3 dx & = & 0 = \alpha_0 q_0^3 + \alpha_1 q_1^3 \end{array} \right.$$

pour lequel $\alpha_0 = \alpha_1 = 1$ et $q_0 = \frac{-1}{\sqrt{3}}$ et $q_1 = \frac{1}{\sqrt{3}}$ est solution. $\square$

Plus généralement, les points de Gauss sont donnés par les racines de polynômes de Gauss-Legendre définis par

$$\begin{aligned} \forall m \in \mathbb{N}, \quad P_m(x) &= \frac{1}{2^m m!} \frac{d^m}{dx^m} \left[(x^2 - 1)^m \right] \\ &= \frac{1}{2^m} \sum_{k=0}^m \left(\frac{m!}{k!(m-k)!} \right)^2 (x-1)^k (x+1)^{m-k}. \end{aligned}$$

Ces polynômes sont orthogonaux pour le produit scalaire canonique de $L^2(-1, 1)$:

$$\forall (m, n) \in \mathbb{N}^2, \quad \int_{-1}^1 P_m(x) P_n(x) dx = \delta_m^n$$

et se calcule, partant de $P_0(x) = 1$ et $P_1(x) = x$, par la formule de récurrence

$$(m+1) P_{m+1}(x) = (2m+1) x P_m(x) - m P_{m-1}(x).$$

4: Johann Carl Friedrich Gauss, 1777-1855 surnommé le prince des mathématiciens

TABEAU 5.1 – Points de quadrature de Gauss-Legendre sur $(-1, 1)$

Ordre	α_k			q_k		
1	2			0		
3	1	1		$-\frac{1}{\sqrt{3}}$	$\frac{1}{\sqrt{3}}$	
5	$\frac{5}{9}$	$\frac{8}{9}$	$\frac{5}{9}$	$-\sqrt{\frac{3}{5}}$	0	$\sqrt{\frac{3}{5}}$

On note $(q_k^m)_{1 \leq k \leq m}$ les racines du polynôme P_m .

Theorème 5.2.6 (Formules de quadrature de Gauss-Legendre) Soit $\psi \in C^2([-1, 1])$, alors les points de quadrature de Gauss $(q_k := q_k^m)_{1 \leq k \leq m}$ où $(q_k^m)_{1 \leq k \leq m}$ sont les racines du polynôme P_m et les poids

$$\alpha_k = \frac{2(1 - (q_k^m)^2)}{(m+1)^2 P_{m+1}(q_k^m)^2} = \frac{2}{(1 - (q_k^m)^2) P'_m(q_k^m)^2},$$

assurent une formule de quadrature d'ordre $2m - 1$:

$$\forall p \in \mathbb{P}_{2m-1}([-1, 1]), \quad \int_{-1}^1 p(x) dx = \sum_{k=1}^m \alpha_k p(q_k).$$

On renvoie au tableau 5.1 pour une liste des points de quadratures aux premiers ordres.

5.3 Autres exemples d'éléments finis en dimension 1

5.3.1 Éléments finis de Lagrange $\mathbb{P}_k$

Éléments finis de Lagrange $\mathbb{P}_2$. Nous supposons désormais que

$$\mathcal{V}_h = \mathcal{V}_h^{\mathbb{P}_2} := \left\{ v_h \in C^0([0, L]) \mid \forall 0 \leq l \leq N_i, v_h|_{[x_l, x_{l+1}]} \in \mathbb{P}_2([x_l, x_{l+1}]) \right\}.$$

et toujours

$$\mathcal{V}_{0h} = \mathcal{V}_{0h}^{\mathbb{P}_2} := \left\{ v_h \in \mathcal{V}_h^{\mathbb{P}_2} \mid v_h(0) = v_h(L) = 0 \right\}.$$

Le maillage est toujours constitué des $N_e := N_i + 1$ segments $([x_l, x_{l+1}])_{0 \leq l \leq N_i}$. Sur chacun de ces segments, les degrés de liberté sont les points x_l et x_{l+1} auquel nous allons ajouter un point de contrôle au milieu

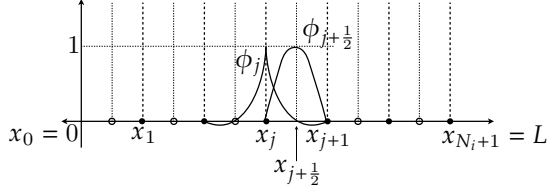
$$x_{l+\frac{1}{2}} = \frac{x_l + x_{l+1}}{2}.$$

En effet, 3 valeurs sont nécessaires pour définir de manière unique un polynôme de degré 2 sur $[x_l, x_{l+1}]$. Nous introduisons

$$\mathcal{I}_h = \left\{ 0, \frac{1}{2}, 1, \dots, N_i, N_i + \frac{1}{2}, N_i + 1 \right\},$$

l'ensemble ordonné des points correspondants de $\text{card}(\mathcal{I}_h) = 2N_i + 3$. Pour tout $i \in \llbracket 1, 2N_i + 3 \rrbracket$, on appellera dans la suite $\mathcal{I}_h(i)$ le i -ème

élément de $\mathcal{I}_h$. La correspondance $i \mapsto x_{\mathcal{I}_h(i)}$ donne la *numérotation globale*⁵ des points d'interpolation $(x_{\mathcal{I}_h(i)})_i$.



5: La numérotation que nous avons choisie ici respecte l'ordre et nous verrons les avantages que cela procure. Un autre numérotation globale aurait été possible en ayant d'abord les indices $\llbracket 0, N+1 \rrbracket$ puis les nœuds intérieurs à chaque segment tel que par exemple $x_{N+2} = \frac{x_0+x_1}{2}$, etc.

FIGURE 5.3 – fonctions de forme $\mathbb{P}_2$

Comme pour les éléments finis de Lagrange $\mathbb{P}_1$, nous introduisons des fonctions de formes $(\phi_j)_{j \in \llbracket 1, 2N_i+3 \rrbracket}$ telles que

$$\phi_j(x_{\mathcal{I}_h(i)}) = \delta_i^j, \quad (i, j) = \llbracket 1, 2N_i+3 \rrbracket^2.$$

Proposition 5.3.1 L'espace $\mathcal{V}_h^{\mathbb{P}_2}$ est de dimension $N = 2N_i + 3$ et les $(\phi_j)_{1 \leq j \leq N}$ forment une base de l'espace $\mathcal{V}_h^{\mathbb{P}_2}$.

Démonstration. Soit $\vec{v} = (v^j)_{1 \leq j \leq N}$ tel que

$$\sum_{1 \leq j \leq N} v^j \phi_j \equiv 0$$

c'est-à-dire

$$\forall x \in [0, L], \quad \sum_{1 \leq j \leq N} v^j \phi_j(x) = 0.$$

Alors, on a

$$v^j \phi_j(x_{\mathcal{I}_h(j)}) = 0$$

donc la famille $(\phi_j)_{1 \leq j \leq N}$ est libre. De plus, sur chaque segment, une fonction $u_h \in \mathcal{V}_h$ est un polynôme de degré 2 donc défini par 3 coefficients, soit ici par ses valeurs aux extrémités du segment et au milieu du segment. Alors $u_h(x)$ et $\sum_{1 \leq j \leq N} u^j \phi_j(x)$ sont égales sur chaque segment et donc sur $[0, L]$. Ainsi la famille libre des $(\phi_j)_{1 \leq j \leq N}$ est aussi génératrice. $\square$

Ainsi à toute fonction $u_h \in \mathcal{V}_h^{\mathbb{P}_2}$, on associe un vecteur de degrés de liberté $\vec{u} = (u^j)_{j \in \mathcal{I}_h}$ et on vient de voir que

$$\forall x \in [0, L], \quad u_h(x) = \sum_{1 \leq j \leq N} u^j \phi_j(x), \quad \text{avec } u^j = u_h(x_{\mathcal{I}_h(j)}).$$

Nous pouvons maintenant reprendre le cours de la méthode des éléments finis en calculant les intégrales nécessaires sur chaque élément via le passage à l'élément de référence. Les désormais 3 fonctions de forme locales associées à l'élément de référence sont cette fois-ci données par

$$\begin{cases} \hat{\lambda}_0(\hat{x}) &= (1 - 2\hat{x})(1 - \hat{x}), \\ \hat{\lambda}_1(\hat{x}) &= 4\hat{x}(1 - \hat{x}), \\ \hat{\lambda}_2(\hat{x}) &= \hat{x}(2\hat{x} - 1). \end{cases} \quad (5.10)$$

On remarquera que les fonctions $\hat{\lambda}_0$ et $\hat{\lambda}_2$ valent 1 à un des sommets de l'élément, et 0 à l'autre sommet et au milieu du segment. Inversement $\hat{\lambda}_1$ est la fonction de base associée au milieu⁶.

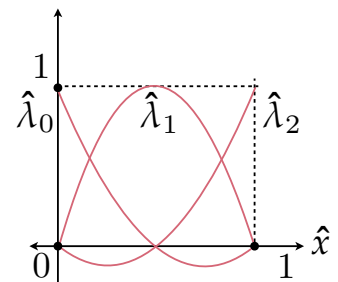


FIGURE 5.4 – Fonctions de formes $\mathbb{P}_2$ en 1D

6: Ici nous avons choisi une numérotation locale similaire à la numérotation globale choisie mais nous aurions pu tout autant prendre le point milieu en dernier.

Une fois ces fonctions de forme définies, il ne reste plus qu'à construire les matrices élémentaires qui ici sont des matrices de $\mathbb{M}_3(\mathbb{R})$. On obtient finalement

$$\mathbb{K}^{\text{el}} = \left(\int_0^1 \hat{\lambda}'_i(\hat{x}) \hat{\lambda}'_j(\hat{x}) d\hat{x} \right)_{1 \leq i, j \leq 3} = \frac{1}{3} \begin{pmatrix} 7 & -8 & 1 \\ -8 & 16 & -8 \\ 1 & -8 & 7 \end{pmatrix},$$

et

$$\mathbb{M}^{\text{el}} = \left(\int_0^1 \hat{\lambda}_i(\hat{x}) \hat{\lambda}_j(\hat{x}) d\hat{x} \right)_{1 \leq i, j \leq 3} = \frac{1}{30} \begin{pmatrix} 4 & 2 & -1 \\ 2 & 16 & 2 \\ -1 & 2 & 4 \end{pmatrix}.$$

Lors de la procédure d'assemblage, on suit de nouveau l'algorithme 1 qui donne finalement

$$\mathbb{M}_h = \mathbb{M}_h^{\mathbb{P}_2} := \frac{h}{30} \begin{pmatrix} \ddots & & & & & \\ & \ddots & & & & \\ & & 4+4 & 2 & -1 & \\ & & 2 & 16 & 2 & \\ & & -1 & 2 & 4+4 & 2 & -1 \\ & & & 2 & 16 & 2 & \ddots \\ & & & -1 & 2 & 4+4 & \ddots \\ & & & & \ddots & \ddots & \ddots \end{pmatrix},$$

et

$$\mathbb{K}_h = \mathbb{K}_h^{\mathbb{P}_2} := \frac{h}{30} \begin{pmatrix} \ddots & & & & & \\ & \ddots & & & & \\ & & 7+7 & -8 & 1 & \\ & & -8 & 16 & 2 & \\ & & 1 & 2 & 7+7 & -8 & 1 \\ & & & -8 & 16 & 2 & \ddots \\ & & & 1 & 2 & 7+7 & \ddots \\ & & & & \ddots & \ddots & \ddots \end{pmatrix},$$

et on remarquera que la numérotation globale choisie permet de limiter la largeur de bande de la matrice $\mathbb{K}_h + \mathbb{M}_h$.

Définition 5.3.1 (Largeur de bande) On appelle largeur de bande d'une matrice $\mathbb{A} = (\mathbb{A}_{i,j})_{1 \leq (i,j) \leq N} \in \mathbb{M}_N(\mathbb{R})$, l'entier

$$\delta(\mathbb{A}) = \inf_{k \in \llbracket 1, N \rrbracket} \left\{ k \in \llbracket 1, N \rrbracket \mid |i - j| > k \Rightarrow \mathbb{A}_{i,j} = 0 \right\}. \quad (5.11)$$

7: par opposition aux méthodes itératives

Or la largeur de bande a un impact sur les méthodes directes⁷ d'inversion matricielle [2, 8] dont la complexité est souvent en $\delta(\mathbb{A})N^2$.

Généralisation aux ordres supérieurs. Ce que nous venons de voir se généralise à tout ordre avec

$$\mathcal{V}_h^{\mathbb{P}_k} := \left\{ v_h \in C^0([0, L]) \mid \forall 0 \leq l \leq N_i, v_h|_{[x_l, x_{l+1}]} \in \mathbb{P}_k([x_l, x_{l+1}]) \right\}.$$

et

$$\mathcal{V}_{0h}^{\mathbb{P}_k} := \left\{ v_h \in \mathcal{V}_h^{\mathbb{P}_k} \mid v_h(0) = v_h(L) = 0 \right\},$$

sur une grille définie avec des points intérieurs à chaque élément de la forme

$$x_{l+\frac{i}{k}} = x_l + i \frac{x_{l+1} - x_l}{k}, \quad l \in \llbracket 0, N_i \rrbracket, i \in \llbracket 1, k-1 \rrbracket.$$

Les fonctions de formes locales pour l'élément de référence sont alors données par les polynômes de Lagrange de formule

$$\forall i \in \llbracket 0, k \rrbracket, \quad \hat{\lambda}_i^{\mathbb{P}_k}(\hat{x}) = \prod_{\substack{j \in \llbracket 0, k \rrbracket \\ j \neq i}} \frac{k\hat{x} - j}{i - j}. \quad (5.12)$$

5.3.2 Élément fini de Hermite pour des problèmes de type : "poutre"

On considère, dans cette section, un problème déjà rencontré au chapitre précédent correspondant à une poutre en flexion encastree à ses deux extrémités. La fonctionnelle à minimiser est dans ce cas

$$\mathcal{J}(u) = \frac{1}{2} \int_{\Omega} \kappa |\nabla^2 u|^2 \, d\Omega - \int_{\Omega} f u \, d\Omega,$$

où $\Omega =]0, L[$. L'espace de minimisation est naturellement ici $\mathcal{V} = H^2(]0, L[)$ et donne lieu à la formulation variationnelle

$$\forall v \in \mathcal{V}, \quad \int_0^L \kappa u''(x) v''(x) \, dx = \int_0^L f(x) v(x) \, dx.$$

Supposons désormais que nous souhaitons résoudre une telle formulation variationnelle sur un sous-espace de dimension finie $\mathcal{V}_h$ construit à partir d'une partition fixée par les sommets $(x_i)_{0 \leq i \leq N_i+1}$, telle que

$$0 = x_0 < x_1 < \dots < x_N < x_{N_i+1} = L.$$

de pas uniforme $h = \frac{1}{N_i+1}$. Soit donc l'espace

$$\mathcal{V}_{0h} = \left\{ u_h \in C^1([0, L]) \text{ tel que } \forall l \in \llbracket 0, N_i \rrbracket, u_h|_{[x_l, x_{l+1}]} \in \mathbb{P}_3([x_l, x_{l+1}]) \right. \\ \left. \text{et } u_h(0) = u_h(L) = u_h'(0) = u_h'(L) = 0 \right\},$$

dans lequel les fonctions sont désormais $C^1([0, L])$ et C^2 par morceaux⁸. On pourra montrer la proposition suivante.

Proposition 5.3.2 L'espace $\mathcal{V}_{0h}$ défini ci-dessus est inclus dans $H^2(]0, L[)$.

8: $u|_{[x_l, x_{l+1}]}$ doit être un polynôme de degré au moins 3 pour espérer construire une fonction $C^1([0, L])$ où il faudra que les valeurs et les dérivées coïncident de part et d'autre des nœuds

Sur cet espace, la forme bilinéaire

$$a(u_h, v_h) = \int_0^L u_h''(x) v_h''(x) dx,$$

est une forme bilinéaire continue positive. Elle est même définie positive, car si $u'' \equiv 0$ et $u \in C^1([0, L])$ alors $u = ax + b$ et par les conditions aux extrémités on obtient finalement $u \equiv 0$. Donc par application de la proposition 3.6.1, la formulation variationnelle admettra une solution et une seule dans $\mathcal{V}_{0h}$.

Pour construire le système linéaire associé à la résolution de la formulation variationnelle, nous introduisons alors la famille de fonctions de $\mathcal{V}_{0h}$ constituée des $\{\phi_1, \dots, \phi_{N_i}\}$ et $\{\psi_1, \dots, \psi_{N_i}\}$ avec,

$$\begin{aligned} \forall (i, j) \in \llbracket 1, N_i \rrbracket^2, \quad \phi_i(x_j) &= \delta_i^j, \quad \phi_i'(x_j) = 0 \\ \text{et } \psi_i(x_j) &= 0, \quad \psi_i'(x_j) = \delta_i^j, \end{aligned}$$

où δ_i^j désigne toujours le symbole de Kronecker.

Proposition 5.3.3 La famille $\{\phi_1, \dots, \phi_{N_i}, \psi_1, \dots, \psi_{N_i}\}$ forme une base de $\mathcal{V}_{0h}$.

Démonstration. Rappelons succinctement quelques propriétés élémentaires. L'espace $\mathbb{P}_3$ est de dimension 4. Il suffit donc de montrer que si $p \in \mathbb{P}_3(a, b)$ est tel que $p(a) = p'(a) = p(b) = p'(b) = 0$ avec $a \neq b$, alors p est identiquement nul. Le fait que $p(a) = p'(a) = 0$ implique que $(x - a)^2$ divise p ; de même, $(x - b)^2$ divise p . Comme $a \neq b$, cela implique que $(x - a)^2(x - b)^2$ divise p et comme p est de degré ≤ 3 , cela implique que p est identiquement nul.

Pour tout $1 \leq i \leq n$, le support de ϕ_i et ψ_i est réduit à $[x_{i-1}, x_{i+1}]$. Comme dans les paragraphes précédents, il suffit de connaître les fonctions de forme sur l'élément de référence. Celles-ci sont données par

$$\begin{aligned} \hat{\lambda}_1(\hat{x}) &= (1 + 2\hat{x})(1 - \hat{x})^2, \\ \hat{\lambda}_2(\hat{x}) &= \hat{x}(1 - \hat{x})^2, \\ \hat{\lambda}_3(\hat{x}) &= \hat{x}^2(3 - 2\hat{x}), \\ \hat{\lambda}_4(\hat{x}) &= \hat{x}^2(\hat{x} - 1), \end{aligned}$$

définies pour $\hat{x} \in [0, 1]$.

Les fonctions de forme sont telles que, pour tout $1 \leq m, n \leq 4$,

$$\sigma_m(\hat{\lambda}_n) = \delta_m^n$$

où $\sigma_1(p) = p(0)$, $\sigma_2(p) = p'(0)$, $\sigma_3(p) = p(1)$ et $\sigma_4(p) = p'(1)$ sont des formes linéaires sur $\mathbb{P}_3$. Par conséquent, on obtient⁹

$$\phi_i(x) = \begin{cases} \hat{\lambda}_3\left(\frac{x-x_{i-1}}{h}\right) & x \in [x_{i-1}, x_i], \\ \hat{\lambda}_1\left(\frac{x-x_i}{h}\right) & x \in [x_i, x_{i+1}], \\ 0 & \text{sinon.} \end{cases}$$

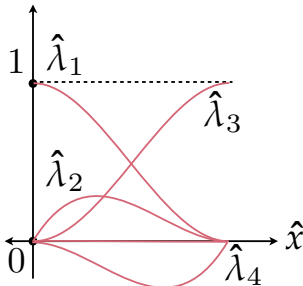


FIGURE 5.5 – Fonctions de formes de Hermite en 1D

9: comme précédemment il y a un facteur d'échelle lorsqu'on passe de l'élément de référence $[0, 1]$ où sont définies les fonctions de forme de référence au domaine $]0, L[$

et

$$\psi_i(x) = \begin{cases} h\hat{\lambda}_4\left(\frac{x-x_{i-1}}{h}\right) & x \in [x_{i-1}, x_i], \\ h\hat{\lambda}_2\left(\frac{x-x_i}{h}\right) & x \in [x_i, x_{i+1}], \\ 0 & \text{sinon.} \end{cases}$$

Donnons-nous désormais $(\alpha_i)_{1 \leq i \leq n}$, et $(\beta_i)_{1 \leq i \leq n}$ tels que

$$\forall x \in]0, L[, \quad \sum_{i=1}^n \alpha_i \phi_i(x) + \sum_{i=1}^n \beta_i \psi_i(x) = 0.$$

En évaluant cette combinaison en x_i , $\forall 1 \leq i \leq n$, il vient $\alpha_i = 0$, puis en évaluant la dérivée de cette combinaison en x_i , il vient $\beta_i = 0$, donc la famille $\{\phi_1, \dots, \phi_n, \psi_1, \dots, \psi_n\}$ est libre.

Pour montrer que la famille est génératrice, on se donne $v_h \in \mathcal{V}_{0h}$ et on pose

$$w_h(x) = \sum_{i=1}^n v_h(x_i) \phi_i(x) + \sum_{i=1}^n v'_h(x_i) \psi_i(x)$$

Sur chaque élément, les valeurs de v_h et de w_h , ainsi que celles de leur dérivée, coïncident aux deux extrémités de la maille. Les restrictions de v_h et w_h à cette maille étant dans $\mathbb{P}_3$, on en déduit d'après les premières remarques de la preuve que $v_h|_{[x_l, x_{l+1}]} \equiv w_h|_{[x_l, x_{l+1}]}$ pour tout l , et donc sur $v_h = w_h$ sur $]0, L[$. $\square$

On considère la matrice de rigidité $\mathbb{K}_h \in \mathbb{M}_{2N_i}(\mathbb{R})$ que l'on peut décomposer sous la forme

$$\mathbb{K}_h = \begin{pmatrix} \mathbb{K}_h^{\phi\phi} & \mathbb{K}_h^{\phi\psi} \\ \mathbb{K}_h^{\psi\phi} & \mathbb{K}_h^{\psi\psi} \end{pmatrix}$$

où les quatre sous-matrices sont de taille N_i et tridiagonales. Comme la forme bilinéaire a est symétrique, la matrice $\mathbb{K}_h$ est symétrique si bien que $\mathbb{K}_h^{\psi\phi} = \mathbb{K}_h^{\phi\psi\top}$. On vérifie aisément

$$\begin{aligned} \hat{\lambda}_1''(\hat{x}) &= 12\hat{x} - 6, \\ \hat{\lambda}_2''(\hat{x}) &= 6\hat{x} - 4, \\ \hat{\lambda}_3''(\hat{x}) &= -12\hat{x} + 6, \\ \hat{\lambda}_4''(\hat{x}) &= 6\hat{x} - 2, \end{aligned}$$

qui donne

$$\left(\int_0^1 \hat{\lambda}_m''(\hat{x}) \hat{\lambda}_n''(\hat{x}) d\hat{x} \right)_{1 \leq m, n \leq 4} = \begin{pmatrix} 12 & 6 & -12 & 6 \\ 6 & 4 & -6 & 2 \\ -12 & -6 & 12 & -6 \\ 6 & 2 & -6 & 4 \end{pmatrix}.$$

En revenant aux contributions sur chaque segment, on obtient

$$\mathbb{K}_h^{\phi\phi} = \frac{12}{h^3} \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix},$$

et par ailleurs

$$\mathbb{K}_h^{\phi\psi} = \frac{6}{h^2} \begin{pmatrix} 0 & 1 & & & 0 \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 0 & 1 \\ 0 & & & -1 & 0 \end{pmatrix}, \quad \mathbb{K}_h^{\psi\psi} = \frac{2}{h} \begin{pmatrix} 4 & 1 & & & 0 \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ 0 & & & 1 & 4 \end{pmatrix}$$

Comme pour les éléments finis de Lagrange $\mathbb{P}_k$ la numérotation a des conséquences directes sur la structure générale de la matrice. En effet, si nous choisissons au contraire d'ordonner la famille sous la forme

$$\{\phi_1, \psi_1, \dots, \phi_i, \psi_i, \dots, \phi_N, \psi_N\},$$

alors la matrice de rigidité est cette fois tridiagonale avec des blocs de taille 2, soit

$$\mathbb{K}_h = \begin{pmatrix} \mathbb{A} & \mathbb{B} & & & 0 \\ \mathbb{B} & & \ddots & \ddots & \\ & \ddots & \ddots & \ddots & \mathbb{B} \\ 0 & & & \mathbb{B} & \mathbb{A} \end{pmatrix},$$

où

$$\mathbb{A} = \begin{pmatrix} 24h^{-3} & 0 \\ 0 & 8h^{-1} \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} -12h^{-3} & 6h^{-2} \\ -6h^{-2} & 2h^{-1} \end{pmatrix}.$$

La largeur de bande de cette nouvelle matrice est bien plus réduite.

5.4 Méthode spectrale

Comme mentionné au chapitre précédent, une alternative à la méthode des éléments finis est une méthode de Galerkin parfois appelée méthode spectrale car fondée sur la connaissance analytique de certains modes propres de l'opérateur étudié (cf. chapitre 8).

Revenons par exemple au problème variationnel défini par les formes bilinéaire (5.2) et linéaire (5.3) et cette fois nous considérons les fonctions définies par

$$\forall i \in \llbracket 1, N \rrbracket, \quad \phi_i(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{i\pi x}{L}\right) = \sqrt{\frac{2}{L}} \sin(\omega_i x).$$

Un résultat classique montre que les $(\phi_i)_{1 \leq i \leq N}$ forment une famille

orthonormée pour le produit scalaire sur $L^2(]0, L[)$,

$$\int_0^L \phi_i(x) \phi_j(x) dx = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

On peut alors considérer $\mathcal{V}_{0h} = \text{Vect}(\phi_i)_{1 \leq i \leq N}$. Ainsi on obtient directement

$$\mathbb{M}_h^{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases} \text{ et } \mathbb{K}_h^{ij} = \begin{cases} 0 & \text{si } i \neq j \\ \omega_i^2 & \text{si } i = j \end{cases}$$

qui sont désormais des matrices diagonales. La solution du problème variationnel est donnée par

$$u_h(x) = \sum_{1 \leq i \leq N} \frac{1}{1 + \omega_i^2} \left(\int_0^1 f(x) \phi_i(x) dx \right) \phi_i(x).$$

La projection L^2 de f sur $\mathcal{V}_{0h}$ consiste à chercher $f_h \in \mathcal{V}_{0h}$

$$\forall v_h \in \mathcal{V}_{0h}, \quad \int_0^L f_h(x) v_h(x) dx = \int_0^L f(x) v_h(x) dx.$$

On vérifie aisément que cette formule définit bien un opérateur idempotent sur $\mathcal{V}_{0h}$. Étant donné la structure de la matrice de masse, nous obtenons

$$f_h(x) = \sum_{1 \leq i \leq N} \left(\int_0^L f(x) \phi_i(x) dx \right) \phi_i(x).$$

Nous voyons ainsi que la formulation variationnelle consiste à faire décroître chaque coefficient de f_h par un facteur $1 + \omega_i^2$. Or chaque coefficient correspond à un mode de Fourier et donc plus la fréquence est élevée, plus le coefficient correspondant est atténué. On retrouve ainsi la propriété de régularisation du problème variationnel vue au paragraphe 1.3.2.

Cette méthode de Galerkin est très attirante en 1D car nous avons pu introduire des fonctions parfaitement adaptées au problème que nous considérons. Si la méthode des éléments finis s'adapte naturellement à la dimension supérieure, cette approche spectrale nécessitera des domaines et problèmes spécifiques. Nous reverrons plus en détail ces questions au chapitre 8 avec l'analyse spectrale des problèmes variationnels.

La méthode des éléments finis en 2D

6

Nous allons maintenant généraliser ce que nous venons de voir en dimension 1 à la dimension 2, en commençant par traiter le cas de l'élément fini de Lagrange $\mathbb{P}_1$ défini sur une discrétisation du domaine en triangles élémentaires. Nous esquisserons ensuite les éléments de généralisation pour les éléments finis $\mathbb{P}_k$, avant de voir succinctement (et en complément) que la discrétisation aurait pu s'appuyer sur des quadrilatères, conduisant à des éléments finis un peu différents, appelés éléments finis $\mathbb{Q}_k$. Enfin, le cadre général des éléments finis est présenté à titre de perspective.

6.1 Introduction

Pour décrire les étapes fondamentales de la méthode des éléments finis en dimension 2, nous étudions maintenant le problème variationnel associé à la forme bilinéaire définie sur un espace $\mathcal{V}$

$$a(u, v) := a_{\mathcal{R}}(u, v) = \int_{\Omega} [\kappa \nabla u \cdot \nabla v + uv] \, d\Omega + \int_{\partial\Omega} uv \, d\Gamma,$$

et la forme linéaire

$$\ell(v) := \ell_{\mathcal{R}}(v) = \int_{\Omega} f v \, d\Omega + \int_{\partial\Omega} g v \, d\Gamma,$$

sur un domaine Ω que nous supposons pour l'instant être un polygone connexe de $\mathbb{R}^2$.

Comme pour le problème 1D, l'objectif est de proposer typiquement une approximation interne de la forme

$$\text{Trouver } u_h \in \mathcal{V}_h \text{ tel que } a(u_h, v_h) = \ell(v_h) \quad \forall v_h \in \mathcal{V}_h, \quad (6.1)$$

où $\mathcal{V}_h \subset \mathcal{V}$ est l'espace d'approximation. Le paramètre h caractérisera de nouveau la finesse de la décomposition du domaine en éléments géométriques élémentaires. La dimension N de l'espace d'approximation sera d'autant plus grande que ce paramètre h tendra vers 0.

6.2 Élément fini de Lagrange $\mathbb{P}_1$ en dimension 2

6.2.1 Notion de maillage triangulaire en dimension 2

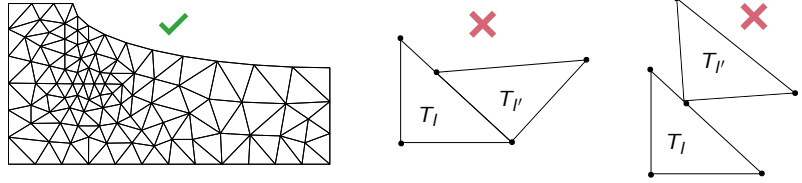
Nous allons d'abord créer une partition du domaine Ω qui sera appelée *maillage*, voir une illustration à la figure 6.1.

Définition 6.2.1 (Maillage triangulaire) Soit Ω un ouvert connexe polyédrique de $\mathbb{R}^2$. Un maillage¹ triangulaire de $\overline{\Omega}$ est un ensemble de N_e triangles $\{T_l\}_{1 \leq l \leq N_e}$ tels que

1: la Définition 6.2.1 contient un certain nombre de restrictions sur le maillage : dans ce cas on parle souvent de *maillage conforme*.

- Tout triangle T_k est d'intérieur non vide (c'est-à-dire $\overset{\circ}{T}_k \neq \emptyset$). On définit alors ρ_k le rayon du plus grand cercle contenu dans T_k et h_k le rayon du plus petit cercle contenant T_k .
- $\overset{\circ}{T}_k \cap \overset{\circ}{T}_{k'} = \emptyset$ si $k \neq k'$,
- $\bigcup_{\ell} T_k = \overline{\Omega}$,
- toute arête d'un triangle est soit une arête d'un autre triangle soit une arête portée par la frontière $\partial\Omega$.

FIGURE 6.1 – Exemple de maillage conforme (gauche), et configurations non conformes (droite)



Les sommets $(M_i)_{1 \leq i \leq N_s}$ aussi appelés nœuds du maillage triangulaire sont les sommets des triangles qui le composent. De plus, on définit

$$h = \max_{1 \leq k \leq N_e} h_k, \quad (6.2)$$

qui caractérise la finesse du maillage. Le maillage triangulaire, ou *triangulation* sera alors noté $\mathcal{T}_h = \{T_l\}_{1 \leq l \leq N_e}$ en utilisant son degré de finesse comme un paramètre. Pour continuer avec le vocabulaire, les triangles T_l seront aussi appelés *mailles* du maillage triangulaire en dimension 2, et les arêtes sont les *faces*². À noter qu'il est aussi possible de définir la finesse du maillage h à partir du maximum des longueurs des arêtes.

2: Ces correspondances maille-triangle et arêtes-faces sont spécifiques à la dimension 2 et changeront en dimension 3 où les mailles correspondront aux tétraèdres et les faces aux triangles.

Soit N_s le nombre de sommets dont on distingue les points intérieurs au nombre de N_i des N_b sommets du bord tels que $N_s = N_i + N_b$. En dimension 2, il est possible d'estimer le nombre d'arêtes N_a et de triangles N_e en fonction de N_s grâce aux relations d'Euler.

Proposition 6.2.1 (Relations d'Euler) Si Ω est simplement connexe (c'est-à-dire sans trou), et $\mathcal{T}_h$ une triangulation, alors on a les relations d'Euler suivantes entre N_e le nombre de triangles, N_a le nombre d'arêtes, N_s le nombre de sommets et N_b le nombre de sommets du bord :

$$N_e = 2N_s - N_b - 2, \quad (6.3)$$

$$N_a = 3N_s - N_b - 3. \quad (6.4)$$

Démonstration. On commence par raisonner sur les angles. La somme des angles internes à chaque triangle est πN_e . Pour chaque point intérieur la somme des angles est 2π et la somme des angles internes aux sommets d'un polygone de N_b côtés ou sommets est égale à $(N_b - 2)\pi$. Donc on obtient

$$\pi N_e = 2\pi N_i + \pi(N_b - 2),$$

ce qui justifie (6.3).

Raisonnons désormais sur les arêtes. Chaque triangle contient 3 arêtes, mais en dénombrant $3N_e$ arêtes, on compte en fait deux fois les arêtes

intérieures contre une seule fois les N_b arêtes du bord, donc

$$3N_e + N_b = 2N_a.$$

En combinant ce résultat à (6.3), on obtient

$$2N_a = 3(2N_s - N_b - 2) + N_b = 6N_s - 2N_b - 2,$$

qui justifie (6.4) et conclut la preuve. $\square$

Quand on crée un maillage ou quand on utilise un logiciel pour en créer, on peut récupérer en général les données suivantes :

- Un tableau que nous noterons *tabneu* avec N_s (le nombre de sommets) lignes et 2 ou 3 colonnes suivant que l'on est dans $\mathbb{R}^2$ ou $\mathbb{R}^3$: il indique les coordonnées des sommets composants le maillage. Ce tableau donne également naturellement une *numérotation* que nous qualifierons de *globale* des sommets. Ainsi à la N -ième ligne de *tabneu*, on trouvera les coordonnées du N -ième sommet.
- Un tableau que nous noterons *tabtri* avec N_e (le nombre de triangles/éléments) lignes et 3 colonnes : il indique la composition de chaque triangle à partir des numéros globaux des sommets le composant. Ainsi à la ℓ -ième ligne de *tabtri*, on trouvera trois entiers correspondants aux numéros des sommets de ce triangle. Les sommets sont donnés dans un certain ordre : nous qualifierons cette *numérotation* de *locale*.
- Un tableau que nous noterons *tabarete* avec N_a (le nombre d'arêtes) lignes et 2 colonnes : il indique la composition de chaque arête du bord à partir des numéros globaux des sommets la composant.
- Un tableau que nous noterons *refneu* avec N_s lignes et 1 colonne : ce tableau composé d'entiers appelés références indique si un sommet est intérieur ou sur le bord ou sur une certaine partie du bord. Cela permet de faire la différence entre les différentes parties du bord si des conditions aux limites différentes y sont imposées.
- Un tableau que nous noterons *reftri* avec N_e lignes et 1 colonne : ce tableau composé d'entiers appelés encore références, mais des triangles, indique où se situe un triangle si plusieurs sous-domaines composent le domaine. Cela permet de faire la différence entre les différentes parties du domaine si les coefficients ou les équations changent en fonction des sous-domaines.

6.2.2 Définition de l'espace d'approximation $\mathbb{P}_1$ et fonctions de forme

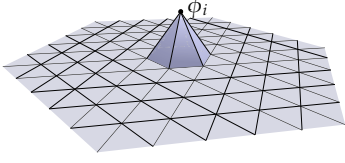
Étant donné un maillage $\mathcal{T}_h = \{T_l\}_{1 \leq l \leq N_e}$ de Ω , nous allons introduire³ l'espace d'approximation

3: comme en dimension 1

$$\mathcal{V}_h = \left\{ v_h \in C^0(\overline{\Omega}) \mid \forall 0 \leq l \leq N_e, v_h|_{T_l} \in \mathbb{P}_1(T_l) \right\}, \quad (6.5)$$

où $\mathbb{P}_1(T_l)$ désigne l'ensemble des fonctions affines sur T_l ,

$$\mathbb{P}_1(T_l) = \left\{ u \in C^0(T_l) \mid u(x, y) = ax + by + c, \quad (a, b, c) \in \mathbb{R}^3 \right\}.$$

FIGURE 6.2 – Fonctions de base $\mathbb{P}_1$ en 2D

La proposition suivante rappelle que sur chaque élément T_l , la restriction affine d'une fonction de $\mathbb{P}_1(T_l)$ est uniquement déterminée par les valeurs aux sommets.

Proposition 6.2.2 Un polynôme $p \in \mathbb{P}_1(T_l)$ est déterminé de manière unique par la valeur qu'il prend en trois points non-alignés, par exemple aux sommets de T_l . De plus, sa restriction à un segment non réduit à un point est déterminée de manière unique par les valeurs qu'elle prend aux deux extrémités de ce segment.

Démonstration. Le polynôme $p \in \mathbb{P}_1(T_l)$ est une fonction affine donc il existe $(a, b, c) \in \mathbb{R}^3$ tel que

$$\forall \underline{x} = (x, y) \in \mathbb{P}_1(T_l), p(\underline{x}) = ax + by + c.$$

Soit $\underline{M}_1 = (x_1, y_1)$, $\underline{M}_2 = (x_2, y_2)$, $\underline{M}_3 = (x_3, y_3)$ les trois points non alignés pour lesquels nous connaissons les valeurs p_1, p_2, p_3 de p . On a

$$\begin{cases} ax_1 + by_1 + c = p_1, \\ ax_2 + by_2 + c = p_2, \\ ax_3 + by_3 + c = p_3. \end{cases}$$

4: le produit vectoriel en dimension 2 est défini par $\underline{M} \wedge \underline{\tilde{M}} = x_M y_{\tilde{M}} - y_M x_{\tilde{M}}$.

Ce système linéaire a pour déterminant⁴

$$\begin{aligned} \det \begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{pmatrix} &= \det \begin{pmatrix} 0 & 0 & 1 \\ x_2 - x_1 & y_2 - y_1 & 1 \\ x_3 - x_1 & y_3 - y_1 & 1 \end{pmatrix} \\ &= (\underline{M}_2 - \underline{M}_1) \wedge (\underline{M}_3 - \underline{M}_1) = \pm 2 \text{aire}(\widehat{\underline{M}_2 \underline{M}_1 \underline{M}_3}). \end{aligned}$$

Ainsi si les trois points $\underline{M}_1, \underline{M}_2, \underline{M}_3$ ne sont pas alignés, le système est inversible et (a, b, c) sont définis de manière unique.

Soit $A = [M_1, M_2]$ un segment et $p|_A$ la restriction de p à ce segment. Soit $M \in A$, il existe $t \in [0, 1]$ tel que $M = tM_1 + (1-t)M_2$ et

$$\begin{aligned} p|_A(\underline{M}) &= a(tx_1 + (1-t)x_2) + b(tx_1 + (1-t)x_2) + c \\ &= tp(\underline{M}_1) + (1-t)p(\underline{M}_2) \end{aligned}$$

et donc $p|_A(\underline{M})$ est bien uniquement déterminé par les valeurs $p(\underline{M}_1)$ et $p(\underline{M}_2)$. $\square$

Proposition 6.2.3 Soient T_l et $T_{l'}$ deux triangles ayant une arête commune $\Gamma = \partial T_l \cap \partial T_{l'}$. Étant donnés $p_l \in \mathbb{P}_1(T_l)$ et $p_{l'} \in \mathbb{P}_1(T_{l'})$ deux polynômes, la fonction v définie par

$$v(x) = \begin{cases} p_l|_{T_l}(\underline{x}) & \text{si } \underline{x} \in T_l \\ p_{l'}|_{T_{l'}}(\underline{x}) & \text{si } \underline{x} \in T_{l'} \end{cases}$$

est continue sur $T_l \cup T_{l'}$, si et seulement si p_l et $p_{l'}$ ont des valeurs qui coïncident aux sommets de leur arête commune Γ .

Démonstration. D'après la proposition précédente, nous avons vu que pour

$$\forall \underline{x} \in \Gamma = [S_1, S_2], \exists t \in [0, 1], \underline{x} = t\underline{S}_1 + (1-t)\underline{S}_2,$$

on a

$$p_l|_{\Gamma}(\underline{x}) = tp|_{T_l}(\underline{S}_1) + (1-t)p|_{T_l}(\underline{S}_2),$$

ainsi que

$$p_{l'}|_{\Gamma}(\underline{x}) = tp|_{T_{l'}}(\underline{S}_1) + (1-t)p|_{T_{l'}}(\underline{S}_2).$$

Donc si les valeurs aux sommets coïncident alors nécessairement, ces deux fonctions sont égales. Réciproquement v ne peut être continue que si p_l et $p_{l'}$ ont des valeurs qui coïncident aux sommets de leur arête commune Γ . $\square$

Proposition 6.2.4 L'espace $\mathcal{V}_h$ est un espace vectoriel de dimension $N = N_s$ dont une base est donnée par les fonctions affines par morceaux sur chaque triangle $(\phi_i)_{1 \leq i \leq N_s} \in \mathcal{V}_h$ définies⁵ par

$$\phi_i(\underline{S}_j) = \delta_{ij} = \begin{cases} 1, & \text{si } i = j, \\ 0, & \text{sinon.} \end{cases}$$

5: Notons que la numérotation naturelle des fonctions de base est celle des sommets!

Démonstration. Soit $\vec{v} = (v^1, \dots, v^{N_s})^\top$, tel que

$$\forall \underline{x} \in \Omega, \quad \sum_{1 \leq i \leq N_s} v^i \phi_i(\underline{x}) = 0$$

alors en particulier pour tout $1 \leq j \leq N_s$

$$\sum_{1 \leq i \leq N_s} v^i \phi_i(\underline{S}_j) = v^j = 0$$

et la famille $(\phi_i)_{1 \leq i \leq N_s}$ est libre. Soit désormais $v_h \in \mathcal{V}_h$ et $\sum_{1 \leq i \leq N_s} v_h(\underline{S}_i) \phi_i(\underline{x})$ qui coïncident donc sur chaque sommet. D'après les deux propositions précédentes, elles coïncident sur chaque triangle et sur chaque arête. Elles sont donc égales sur $\bar{\Omega}$. Ainsi

$$\forall \underline{x} \in \Omega, \quad v_h(\underline{x}) = \sum_{1 \leq i \leq N_s} v_h(\underline{S}_i) \phi_i(\underline{x}),$$

et la famille $(\phi_i)_{1 \leq i \leq N_s}$ est donc génératrice. On en déduit immédiatement que $\dim \mathcal{V}_h = N_s$. $\square$

En utilisant la théorème 2.2.2 et les trois propositions précédentes, on montre finalement aisément la proposition suivante.

Proposition 6.2.5 L'espace $\mathcal{V}_h$ défini en (6.5) est un sous-espace vectoriel de $H^1(\Omega)$.

Nous allons chercher maintenant à caractériser les fonctions de forme sur chaque triangle T_l qui sont ici les *coordonnées barycentriques* $(\hat{\lambda}_i^l)_{1 \leq i \leq 3}$. En effet, les *coordonnées barycentriques* sont les uniques fonctions affines vérifiant pour les 3 sommets $(S_i^l)_{1 \leq i \leq 3}$ du triangle T_l

$$\forall (\hat{i}, \hat{j}) \in \llbracket 1, 3 \rrbracket^2, \quad \lambda_i^l(\underline{S}_j^l) = \delta_{\hat{i}}^{\hat{j}}.$$

Par définition des coordonnées barycentriques on a

$$\forall \underline{M} \in T_l, \quad \underline{M} = \lambda_1^l(\underline{M})\underline{S}_1^l + \lambda_2^l(\underline{M})\underline{S}_2^l + \lambda_3^l(\underline{M})\underline{S}_3^l,$$

avec

$$\begin{aligned} &— \forall \underline{M} \in T_l, \quad \lambda_1^l(\underline{M}) + \lambda_2^l(\underline{M}) + \lambda_3^l(\underline{M}) = 1. \\ &— \forall \underline{M} \in T_l, \quad \forall \hat{i} \in \{1, 2, 3\}, \quad 0 \leq \lambda_{\hat{i}}^l(\underline{M}) \leq 1. \end{aligned}$$

De plus on vérifie aisément que

$$\begin{aligned} &— \forall \hat{i} \in \{1, 2, 3\}, \quad \lambda_{\hat{i}|A_{\hat{i}}}^l \equiv 0 \text{ où } A_{\hat{i}}^l \text{ est l'arête opposée au sommet } S_{\hat{i}}^l. \\ &— \text{Soit } G_l \text{ le barycentre de } T_l, \text{ alors } \lambda_{\hat{i}}^l(G_l) = \frac{1}{3} \leq 1, \quad \hat{i} \in \{1, 2, 3\}. \end{aligned}$$

Ceci nous permet de déduire les formules analytiques suivantes

$$\lambda_1^l(x, y) = \frac{1}{D_l} \left[(y_2^l - y_3^l)(x - x_3^l) - (x_2^l - x_3^l)(y - y_3^l) \right], \quad (6.6)$$

$$\lambda_2^l(x, y) = \frac{1}{D_l} \left[(y_3^l - y_1^l)(x - x_1^l) - (x_3^l - x_1^l)(y - y_1^l) \right], \quad (6.7)$$

$$\lambda_3^l(x, y) = \frac{1}{D_l} \left[(y_1^l - y_2^l)(x - x_2^l) - (x_1^l - x_2^l)(y - y_2^l) \right], \quad (6.8)$$

où

$$D_l = (x_2^l - x_3^l)(y_3^l - y_1^l) - (y_2^l - y_3^l)(x_3^l - x_1^l), \quad (6.9)$$

est, au signe près, deux fois l'aire du triangle T_l . Autrement dit, on a

$$\lambda_1^l(\underline{M}) = \frac{1}{D_l} (\underline{S}_2^l - \underline{S}_3^l) \wedge (\underline{M} - \underline{S}_3^l), \quad (6.10)$$

$$\lambda_2^l(\underline{M}) = \frac{1}{D_l} (\underline{S}_3^l - \underline{S}_1^l) \wedge (\underline{M} - \underline{S}_1^l), \quad (6.11)$$

$$\lambda_3^l(\underline{M}) = \frac{1}{D_l} (\underline{S}_1^l - \underline{S}_2^l) \wedge (\underline{M} - \underline{S}_2^l), \quad (6.12)$$

6: A noter qu'on utilisera indifféremment la notation avec $\underline{M}$ ou encore $\underline{x} = (x, y)^T = (x_1, x_2)^T$

où $D_l = (\underline{S}_2^l - \underline{S}_3^l) \wedge (\underline{S}_1^l - \underline{S}_3^l)^6$.

6.2.3 Matrices de masse et raideur élémentaires.

Matrice de raideur. Considérons tout d'abord la matrice de raideur où

$$\forall (i, j) \in \llbracket 1, N_s \rrbracket^2, \quad \mathbb{K}_h^{ij} = \int_{\Omega} \kappa(\underline{x}) \underline{\nabla} \phi_j(\underline{x}) \cdot \underline{\nabla} \phi_i(\underline{x}) \, d\Omega,$$

et commençons par détailler le cas $\kappa = \bar{\kappa}$. Puisque $\cup_{1 \leq l \leq N_e} T_l = \Omega$, chaque coefficient de la matrice se réécrit

$$\forall (i, j) \in \llbracket 1, N_s \rrbracket^2, \quad \mathbb{K}_h^{ij} = \sum_{T_l \in \mathcal{T}_h} \bar{\kappa} \int_{T_l} \underline{\nabla} \phi_i(\underline{x}) \cdot \underline{\nabla} \phi_j(\underline{x}) \, d\Omega,$$

où nous avons vu que $\phi_i|_{T_l}$ et $\phi_j|_{T_l}$ sont soit nulles, soit égales à un des 3 λ_i^l , $i \in \{1, 2, 3\}$. Nous avons donc besoin de calculer $\mathbb{K}_l^{\text{el}} \in \mathbb{M}_3(\mathbb{R})$ où⁷

$$\forall (\hat{i}, \hat{j}) \in \llbracket 1, 3 \rrbracket^2, \quad (\mathbb{K}_l^{\text{el}})_{\hat{i}, \hat{j}} = \bar{\kappa} \int_{T_l} \underline{\nabla} \lambda_{\hat{j}}^l(\underline{x}) \cdot \underline{\nabla} \lambda_{\hat{i}}^l(\underline{x}) \, d\Omega. \quad (6.13)$$

7: A noter que dans le cas des éléments finis $\mathbb{P}^1$, le gradient de chaque coordonnée barycentrique est constant par élément.

À partir des formules 6.6, 6.7 et 6.8, on obtient

$$\nabla \lambda_1^l(x, y) = \frac{1}{D_l} \begin{pmatrix} y_2^l - y_3^l \\ x_3^l - x_2^l \end{pmatrix}, \quad (6.14)$$

$$\nabla \lambda_2^l(x, y) = \frac{1}{D_l} \begin{pmatrix} y_3^l - y_1^l \\ x_1^l - x_3^l \end{pmatrix}, \quad (6.15)$$

$$\nabla \lambda_3^l(x, y) = \frac{1}{D_l} \begin{pmatrix} y_1^l - y_2^l \\ x_2^l - x_1^l \end{pmatrix}. \quad (6.16)$$

Ces formules indiquent que, pour tout $\hat{i} \in \llbracket 1, 3 \rrbracket$, le gradient de chaque fonction de forme $\lambda_{\hat{i}}^l$ est proportionnel au vecteur $\underline{v}_{\hat{i}}^l$ orthogonal au segment opposé au sommet $\underline{S}_{\hat{i}}$. On en déduit, avec les notations des figures 6.3 et 6.4, sachant que $D_l = 2|T_l| = b_1^l h_1^l$ et $|\underline{v}_1^l| = b_1^l$,

$$\begin{aligned} \int_{T_l} |\nabla \lambda_1^l(\underline{x})|^2 d\Omega &= \frac{1}{D_l^2} \underline{v}_1^l \cdot \underline{v}_1^l |T_l| = \frac{b_1^l}{2h_1^l} \\ &= \frac{b_{1,2}^l}{2h_1^l} + \frac{b_{1,3}^l}{2h_1^l} = \frac{1}{2} (\cot \theta_2^l + \cot \theta_3^l), \end{aligned}$$

et de même pour λ_2^l et λ_3^l . De plus sachant que $|T_l| = \frac{1}{2} b_2^l b_3^l \sin \theta_1^l$, on a

$$\begin{aligned} \int_{T_l} \nabla \lambda_2^l(\underline{x}) \cdot \nabla \lambda_3^l(\underline{x}) d\Omega &= \frac{1}{D_l^2} \underline{v}_2^l \cdot \underline{v}_3^l |T_l| = \frac{1}{4|T_l|} |\underline{v}_2^l| |\underline{v}_3^l| \cos(\pi - \theta_1^l) \\ &= -\frac{1}{2} \frac{|\underline{v}_2^l| |\underline{v}_3^l| \cos(\theta_1^l)}{|\underline{v}_2^l| |\underline{v}_3^l| \sin \theta_1^l} = -\frac{1}{2} \cot \theta_1^l, \end{aligned}$$

et de même pour tous les autres couples $(\lambda_1^l, \lambda_2^l)$ et $(\lambda_1^l, \lambda_3^l)$. On obtient

$$\mathbb{K}_l^{\text{el}} = \frac{\bar{\kappa}}{2} \begin{pmatrix} \cot \theta_2^l + \cot \theta_3^l & -\cot \theta_3^l & -\cot \theta_2^l \\ -\cot \theta_3^l & \cot \theta_1^l + \cot \theta_3^l & -\cot \theta_1^l \\ -\cot \theta_2^l & -\cot \theta_1^l & \cot \theta_1^l + \cot \theta_2^l \end{pmatrix}. \quad (6.17)$$

On en déduit la formule dite *formule des cotangentes* pour la matrice de raideur $\mathbb{P}_1$

$$\mathbb{K}_h^{ij} = \begin{cases} \frac{\bar{\kappa}}{2} \sum_{k \in \mathcal{N}_i} (\cot \theta_{ik} + \cot \theta_{ki}), & \text{si } i = j \\ -\frac{\bar{\kappa}}{2} (\cot \theta_{ij} + \cot \theta_{ji}), & \text{sinon} \end{cases} \quad (6.18)$$

où $\mathcal{N}_i$ est l'ensemble des sommets $\underline{S}_j$ adjacents à $\underline{S}_i$, c'est à dire tel que $[\underline{S}_i, \underline{S}_j]$ est une arête du maillage. Dans cette formule, on fera attention que typiquement il n'y a parfois qu'un seul θ_{ij} pour les segments touchant le bord.

Remarque 6.2.1 Dans le cas de conditions de Dirichlet, avec $\underline{u}$ le vecteur des degrés de liberté intérieurs (ne touchant pas le bord du

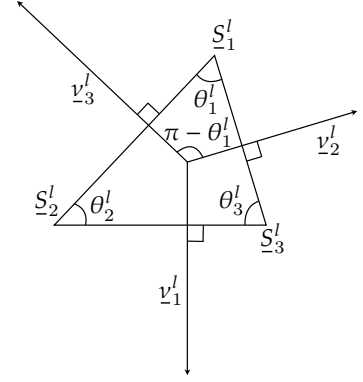


FIGURE 6.3 – Triangle T_l , vecteurs normaux et angles

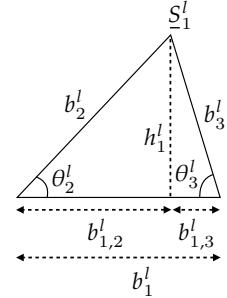
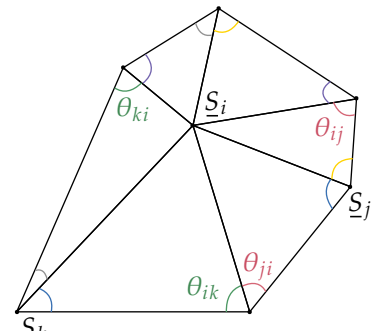


FIGURE 6.4 – hauteur et base associées à $\underline{S}_1$ sur T_l



domaine), on a alors la formule

$$\mathbb{K}\vec{u} = \frac{\bar{\kappa}}{2} \sum_{j \in \mathcal{N}_i} (\cot \theta_{ij} + \cot \theta_{ji})(u_i - u_j). \quad (6.19)$$

Matrice de masse. Concernant la matrice de masse, elle est définie par

$$\forall (i, j) \in \llbracket 1, N_s \rrbracket^2, \quad \mathbb{M}_h^{ij} = \int_{\Omega} \phi_j(\underline{x}) \phi_i(\underline{x}) \, d\Omega.$$

En décomposant les contributions suivant les triangles, on peut écrire

$$\forall (i, j) \in \llbracket 1, N_s \rrbracket^2, \quad \mathbb{M}_h^{ij} = \sum_{T_l \in \mathcal{T}_h} \int_{T_l} \phi_j(\underline{x}) \phi_i(\underline{x}) \, d\Omega.$$

Nous devons donc calculer $\mathbb{M}_l^{\text{el}} \in \mathbb{M}_3(\mathbb{R})$ où

$$\forall (\hat{i}, \hat{j}) \in \llbracket 1, 3 \rrbracket^2, \quad (\mathbb{M}_l^{\text{el}})_{\hat{i}, \hat{j}} = \int_{T_l} \lambda_{\hat{j}}^l(\underline{x}) \lambda_{\hat{i}}^l(\underline{x}) \, d\Omega, \quad (6.20)$$

en nous appuyant cette fois sur la proposition suivante.

Proposition 6.2.6 Soit $(k_1, k_2, k_3) \in \mathbb{N}^3$, alors

$$\int_{T_l} (\lambda_1^l)^{k_1} (\lambda_2^l)^{k_2} (\lambda_3^l)^{k_3} \, d\Omega = 2 \frac{k_1! k_2! k_3!}{(k_1 + k_2 + k_3 + 2)!} |T_l| \quad (6.21)$$

Démonstration. À un changement de variable près, on peut se ramener sur le triangle $\hat{T} = \widehat{\hat{S}_1 \hat{S}_2 \hat{S}_3}$ où $\hat{S}_1 = (0, 0)$, $\hat{S}_2 = (0, 1)$ et $\hat{S}_3 = (1, 0)$,

$$\int_{T_l} (\lambda_1^l)^{k_1} (\lambda_2^l)^{k_2} (\lambda_3^l)^{k_3} \, d\Omega = 2|T_l| \int_{\hat{T}} \hat{\lambda}_1^{k_1} \hat{\lambda}_2^{k_2} \hat{\lambda}_3^{k_3} \, d\Omega,$$

8: voir le paragraphe suivant sur le passage à l'élément de référence où nous détaillons le même calcul

où $2|T_l|$ est le déterminant du changement de variable⁸ et pour $(\hat{x}, \hat{y}) \in \hat{T}$

$$\begin{aligned} \hat{\lambda}_1(\hat{x}, \hat{y}) &= 1 - \hat{x} - \hat{y}, \\ \hat{\lambda}_2(\hat{x}, \hat{y}) &= \hat{x}, \\ \hat{\lambda}_3(\hat{x}, \hat{y}) &= \hat{y}. \end{aligned}$$

Or, on a

$$\int_{\hat{T}} \hat{\lambda}_1^{k_1} \hat{\lambda}_2^{k_2} \hat{\lambda}_3^{k_3} \, d\Omega = \int_0^1 \hat{x}^{k_2} \int_0^{1-\hat{x}} \hat{y}^{k_3} (1 - \hat{x} - \hat{y})^{k_1} \, d\hat{y} \, d\hat{x}.$$

avec

$$\begin{aligned} \int_0^{1-\hat{x}} \hat{y}^{k_3} (1-\hat{x}-\hat{y})^{k_1} d\hat{y} &= \int_0^{1-\hat{x}} k_1 \frac{\hat{y}^{k_3+1}}{k_3+1} (1-\hat{x}-\hat{y})^{k_1-1} d\hat{y} \\ &= \int_0^{1-\hat{x}} k_1! k_3! \frac{\hat{y}^{k_3+k_1}}{(k_3+k_1)!} d\hat{y} \\ &= \frac{k_1! k_3!}{(k_3+k_1+1)!} (1-\hat{x})^{k_3+k_1+1}, \end{aligned}$$

et ainsi

$$\begin{aligned} \int_0^1 \hat{x}^{k_2} \int_0^{1-\hat{x}} \hat{y}^{k_3} (1-\hat{x}-\hat{y})^{k_1} d\hat{y} d\hat{x} &= k_1! k_3! \int_0^1 \frac{1}{(k_3+k_1+1)!} (1-\hat{x})^{k_3+k_1+1} \hat{x}^{k_2} d\hat{x} \\ &= k_1! k_2! k_3! \int_0^1 \frac{(1-\hat{x})^{k_3+k_2+k_1+1}}{(k_3+k_2+k_1+1)!} d\hat{x} \\ &= \frac{k_1! k_2! k_3!}{(k_3+k_2+k_1+2)!}, \end{aligned}$$

ce qui conclut la preuve. $\square$

De la proposition 6.2.6, on déduit donc

$$\mathbb{M}_l^{\text{el}} = \frac{|T_l|}{12} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \quad (6.22)$$

ce qui donne finalement sur les degrés de liberté

$$\mathbb{M}_h^{ij} = \begin{cases} \frac{1}{6} \sum_{S_i \in T_l} |T_l|, & \text{si } i = j, \\ \frac{1}{12} \sum_{[S_i S_j] \in T_l} |T_l|, & \text{sinon.} \end{cases} \quad (6.23)$$

Matrice de masse linéique. Enfin, la matrice de masse linéique est définie par

$$\forall (i, j) \in \llbracket 1, N_s \rrbracket^2, \quad \mathbb{S}_h^{ij} = \int_{\partial\Omega} \phi_j(\underline{x}) \phi_i(\underline{x}) d\Gamma.$$

Cette fois il faut décomposer le bord du domaine sur ses arêtes $(A_m)_{1 \leq m \leq N_a}$, et $\partial\Omega = \cup_{1 \leq m \leq N_a} A_m$ implique

$$\forall (i, j) \in \llbracket 1, N_s \rrbracket^2, \quad \mathbb{S}_h^{ij} = \sum_{1 \leq m \leq N_a} \int_{A_m} \phi_j(\underline{x}) \phi_i(\underline{x}) d\Gamma,$$

qui conduit à calculer $\mathbb{S}_m^{\text{el}} \in \mathbb{M}_2(\mathbb{R})$ définie par

$$\forall (\hat{i}, \hat{j}) \in \llbracket 1, 2 \rrbracket^2, \quad (\mathbb{S}_m^{\text{el}})_{\hat{i}, \hat{j}} = \int_{A_m} \mu_{\hat{j}}^m(\underline{x}) \mu_{\hat{i}}^m(\underline{x}) d\Gamma, \quad (6.24)$$

où cette fois les fonctions de forme sur l'arête $A_m = [\underline{S}_1^m, \underline{S}_2^m]$ sont

$$\mu_1^m(\underline{x}) = \frac{|\underline{x} - \underline{S}_1^m|}{|\underline{S}_2^m - \underline{S}_1^m|}, \text{ et } \mu_2^m(\underline{x}) = \frac{|\underline{x} - \underline{S}_2^m|}{|\underline{S}_1^m - \underline{S}_2^m|}.$$

via des calculs en tout point similaires à ce que nous avons fait en 1D, nous obtenons⁹

$$\forall(\hat{t}, \hat{j}) \in \llbracket 1, 2 \rrbracket^2, \quad (\mathbb{S}_m^{\text{el}}) = \frac{|A_l|}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}. \quad (6.25)$$

9: en fait le calcul est élémentaire par changement de variable pour se ramener sur le segment de référence $[0, 1]$, voir le paragraphe suivant, d'où la similarité avec les calculs en 1D

6.2.4 Passage à l'élément de référence.

Encore une fois, une stratégie plus élégante et surtout plus systématique¹⁰ pour calculer chacune des contributions par élément est de se ramener à l'élément de référence. Commençons par les contributions surfaciques où on se ramène au triangle de référence $\hat{T}$ défini par les trois sommets $\hat{S}_1 = (0, 0)$, $\hat{S}_2 = (0, 1)$ et $\hat{S}_3 = (1, 0)$. Soit $\underline{F}_l$, le tenseur dont la représentation matricielle a pour colonnes $\underline{S}_2^l - \underline{S}_1^l$ et $\underline{S}_3^l - \underline{S}_1^l$,

$$\underline{F}_l = \begin{pmatrix} x_2^l - x_1^l & x_3^l - x_1^l \\ y_2^l - y_1^l & y_3^l - y_1^l \end{pmatrix}.$$

La transformation pour passer de l'élément de référence $\hat{T}$ à T_l est donnée par

$$\varphi_l : \hat{T} \ni \hat{\underline{x}} \mapsto \underline{x} = \underline{S}_1^l + \underline{F}_l \cdot \hat{\underline{x}} \in T_l.$$

Nous noterons $\varphi_l^{-1} : T_l \rightarrow \hat{T}$, la transformation réciproque. Le calcul différentiel donne

$$\nabla \underline{\varphi}_l(\underline{x}) := \left(\frac{\partial \varphi_{l_i}}{\partial x_j} \right)_{1 \leq i, j \leq d} = \underline{F}_l.$$

Sur ce triangle de référence, les fonctions de base élémentaires sont

$$\begin{aligned} \hat{\lambda}_1(\hat{\underline{x}}) &= \hat{\lambda}_1(\hat{x}, \hat{y}) = 1 - \hat{x} - \hat{y}, \\ \hat{\lambda}_2(\hat{\underline{x}}) &= \hat{\lambda}_2(\hat{x}, \hat{y}) = \hat{x}, \\ \hat{\lambda}_3(\hat{\underline{x}}) &= \hat{\lambda}_3(\hat{x}, \hat{y}) = \hat{y}, \end{aligned} \quad (6.26)$$

10: et déjà utilisé en filigrane dans la preuve de la proposition 6.2.6 et pour calculer la matrice de masse linéique

d'où on déduit immédiatement¹¹

$$\hat{\mathbb{K}}^{\text{el}} = \frac{1}{2} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad \hat{\mathbb{M}}^{\text{el}} = \frac{1}{24} \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

Un changement de variable dans les intégrales donne

$$\begin{aligned} \forall(\hat{t}, \hat{j}) \in \llbracket 1, 3 \rrbracket^2, \quad (\mathbb{M}_l^{\text{el}})^{\hat{t}, \hat{j}} &= \int_{T_l} \lambda_{\hat{j}}(\underline{x}) \lambda_{\hat{t}}(\underline{x}) \, d\Omega \\ &= \int_{\hat{T}} |\det(\underline{F}_l)| \hat{\lambda}_{\hat{j}}(\hat{\underline{x}}) \hat{\lambda}_{\hat{t}}(\hat{\underline{x}}) \, d\hat{\Omega} \\ &= 2|T_l| \int_{\hat{T}} \hat{\lambda}_{\hat{j}}(\hat{\underline{x}}) \hat{\lambda}_{\hat{t}}(\hat{\underline{x}}) \, d\hat{\Omega} = 2|T_l| (\hat{\mathbb{M}}^{\text{el}})^{\hat{t}, \hat{j}}. \end{aligned}$$

11: pour la matrice de raideur, on peut vérifier que $\cot \hat{\theta}_2 = \cot \hat{\theta}_3 = 1$ et $\cot \hat{\theta}_1 = 0$

Pour la raideur, on rappelle que par composition des différentielles, on a pour tout $\hat{i} \in \llbracket 1, 3 \rrbracket$,

$$\forall 1 \leq n \leq d, \quad \frac{\partial}{\partial \hat{x}_n} (\lambda_{\hat{i}} \circ \varphi_I)(\hat{x}) = \sum_{1 \leq k \leq d} \frac{\partial \lambda_{\hat{i}}}{\partial x_k} (\varphi_I(\hat{x})) \frac{\partial \varphi_{I_k}}{\partial \hat{x}_n} (\hat{x}),$$

soit

$$\underline{\nabla} (\lambda_{\hat{i}} \circ \varphi_I)(\hat{x}) = \underline{\nabla} \varphi_I(\hat{x})^\top \cdot \underline{\nabla} \lambda_{\hat{i}}(\varphi_I(\hat{x})) = \underline{E}_I^\top \cdot \underline{\nabla} \lambda_{\hat{i}}(\varphi_I(\hat{x})).$$

Donc pour tout $(\hat{i}, \hat{j}) \in \llbracket 1, 3 \rrbracket^2$,

$$\begin{aligned} (\mathbb{K}_I^{\text{el}})^{\hat{i}, \hat{j}} &= \bar{\kappa} \int_{T_I} \underline{\nabla} \lambda_{\hat{j}}(\underline{x}) \cdot \underline{\nabla} \lambda_{\hat{i}}(\underline{x}) \, d\Omega \\ &= \bar{\kappa} \int_{\hat{T}} |\det \underline{E}_I| (\underline{\nabla} \lambda_{\hat{j}})(\varphi_I(\hat{x})) \cdot (\underline{\nabla} \lambda_{\hat{i}})(\varphi_I(\hat{x})) \, d\Omega \\ &= 2\bar{\kappa} |T_I| \int_{\hat{T}} (\underline{E}_I^{-\top} \underline{\nabla} \hat{\lambda}_{\hat{j}}(\hat{x})) \cdot (\underline{E}_I^{-\top} \underline{\nabla} \hat{\lambda}_{\hat{i}}(\hat{x})) \, d\Omega, \end{aligned}$$

où $\underline{E}_I^{-\top} := (\underline{E}_I^\top)^{-1}$. Et, avec des éléments finis de Lagrange $\mathbb{P}_1$ où le gradient est constant par élément, on obtient

$$(\mathbb{K}_I^{\text{el}})^{\hat{i}, \hat{j}} = \bar{\kappa} |T_I| \underline{\nabla} \hat{\lambda}_{\hat{j}} \cdot (\underline{E}_I^{-1} \cdot \underline{E}_I^{-\top}) \cdot \underline{\nabla} \hat{\lambda}_{\hat{i}},$$

dont on remarquera qu'elle ne s'exprime jamais directement en fonction de $\hat{\mathbb{K}}^{\text{el}}$.

Regardons maintenant le passage à l'élément de référence pour la matrice de masse linéique. On repart de (6.24) pour trouver, par le changement de variable,

$$[0, 1] \ni t \mapsto t \underline{S}_1 + (1 - t) \underline{S}_2,$$

que

$$\mathbb{S}_I^{\text{el}} = |A_I| \hat{\mathbb{S}}_I^{\text{el}},$$

où

$$\hat{\mathbb{S}}^{\text{el}} = \begin{pmatrix} \int_0^1 t^2 \, dt & \int_0^1 t(1-t) \, dt \\ \int_0^1 t(1-t) \, dt & \int_0^1 (1-t)^2 \, dt \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

On retrouve donc (6.25).

6.2.5 Formules de quadrature en dimension 2

Comme en dimension 1, il nous reste à comprendre comment étendre les résultats précédents à des cas où les coefficients ne sont plus homogènes, par exemple quand il s'agit de calculer

$$\begin{aligned} (\mathbb{K}_I^{\text{el}})^{\hat{i}, \hat{j}} &= \int_{T_I} \kappa(\underline{x}) \underline{\nabla} \lambda_{\hat{j}}(\underline{x}) \cdot \underline{\nabla} \lambda_{\hat{i}}(\underline{x}) \, d\Omega \\ &= 2|T_I| \int_{\hat{T}} \kappa(\varphi_I(\hat{x})) \underline{\nabla} \hat{\lambda}_{\hat{j}}(\hat{x}) \cdot (\underline{E}_I^{-1} \cdot \underline{E}_I^{-\top}) \cdot \underline{\nabla} \hat{\lambda}_{\hat{i}}(\hat{x}) \, d\Omega. \end{aligned}$$

Et comme en dimension 1, il suffit donc de formuler les règles de quadrature sur le triangle de référence pour obtenir

$$(\mathbb{K}_l^{\text{el}})^{i,j} = \sum_{1 \leq n \leq N_q} \alpha_n^l \kappa(\varphi_l(\hat{q}_n)) \nabla \hat{\lambda}_{\hat{l}}(\hat{q}_n) \cdot (\underline{F}_l^{-1} \cdot \underline{F}_l^{-\top}) \cdot \nabla \hat{\lambda}_{\hat{l}}(\hat{q}_n).$$

Le tableau 6.1 récapitule différents choix.

TABLEAU 6.1 – Formule de quadrature sur un triangle T . Les coordonnées des points de quadratures sont données à partir des coordonnées barycentriques

Ordre	N_q	Poids α_n^l	Nœuds $\hat{q}_n$
1	1	$ T_l $	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$
1	3	$\frac{1}{3} T_l $	$(1, 0, 0) (0, 1, 0) (0, 0, 1)$
2	3	$\frac{1}{3} T_l $	$(\frac{1}{2}, \frac{1}{2}, 0) (0, \frac{1}{2}, \frac{1}{2}) (\frac{1}{2}, 0, \frac{1}{2})$
4	4	$-\frac{9}{16} T_l $ $\frac{25}{48} T_l $	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ $(\frac{3}{5}, \frac{1}{5}, \frac{1}{5}) (\frac{1}{5}, \frac{3}{5}, \frac{1}{5}) (\frac{1}{5}, \frac{1}{5}, \frac{3}{5})$

6.2.6 Assemblage.

Une fois calculées les matrices élémentaires, nous pouvons construire la matrice globale en sommant chaque contribution sur chaque élément. On obtient l'algorithme 2. A noter que dans un code éléments finis 2D, l'assemblage sur les triangles est complété par un assemblage sur les arêtes pour les termes linéiques tels que celui étudié permettant de prendre en compte la condition de Robin.

6.2.7 Calcul du second membre

Il nous reste à calculer le second membre $\vec{\ell}$ correspondant à la forme linéaire $\ell \in \mathcal{V}_h'$. Il faut alors calculer les intégrales

$$\forall 1 \leq i \leq N, \quad \vec{\ell}_h^i = \int_{\Omega} f \phi_i \, d\Omega + \int_{\partial\Omega} g \phi_i \, d\Gamma. \quad (6.27)$$

Comme en 1D, on assimile f à son interpolée

$$f \simeq f_h = \sum_{1 \leq j \leq N} f(\underline{S}_j) \phi_j,$$

et g à la sienne

$$g \simeq g_h = \sum_{1 \leq j \leq N, \underline{S}_j \in \partial\Omega} g(\underline{S}_j) \phi_j.$$

On construit ainsi le second membre :

$$\vec{\ell}_h = \mathbb{M}_h \vec{f}_h + \mathbb{S}_h \vec{g}_h,$$

où on a approché les termes sources par leurs interpolées.

Algorithme 2: Procédure d'assemblage des matrices de raideur et de masse en 2D

```

1  $\mathbb{M} \leftarrow 0$ ;
2  $\mathbb{K} \leftarrow 0$ ;
  /* Boucle sur les triangles */
3 for  $l \in \{1, \dots, N_e\}$  do
  /* Mettre à jour la transformation entre l'élément
    de référence et l'élément courant */
  /* Récupération des coordonnées des sommets
    composant le triangle à partir de tabneu et
    tabtri */
4   $\underline{\varphi} \leftarrow \underline{F}\underline{x} + S_1$ ;
5   $|T| \leftarrow \det(\underline{F})$ ;
6   $BB \leftarrow \underline{F}^\top \cdot \underline{F}^{-\top}$ ;
  /* Boucle sur les indices */
7  for  $\hat{i} \in \{1, 2, 3\}$  do
8     $I = \text{tabtri}(l, \hat{i})$ ;
9    for  $\hat{j} \in \{1, 2, 3\}$  do
10      $J = \text{tabtri}(l, \hat{j})$ ;
11     /* Calcul de la masse en fonction de  $|T|$  */
12      $\mathbb{M}_{I,J} \leftarrow \mathbb{M}_{I,J} + |T|(\hat{\mathbb{M}}_k^{\text{el}})_{\hat{i},\hat{j}}$ ;
13      $(\mathbb{K}_l^{\text{el}})_{\hat{i},\hat{j}} \leftarrow 0$ ;
14     /* Boucle sur les points de quadrature */
15     for  $n \in \{1, \dots, N_q\}$  do
      /* Calcul de la raideur en fonction de
         $|T|, BB, \kappa \circ \varphi^{-1}$  au point de quadrature */
       $(\mathbb{K}_l^{\text{el}})_{\hat{i},\hat{j}} \leftarrow (\mathbb{K}_l^{\text{el}})_{\hat{i},\hat{j}} + (\mathbb{K}_l^{\text{el}})_{\hat{i},\hat{j}}(\hat{q}_n)$ 
16      $\mathbb{K}_{I,J} \leftarrow \mathbb{K}_{I,J} + (\mathbb{K}_l^{\text{el}})_{\hat{i},\hat{j}}$ 

```

6.2.8 Prise en compte des conditions aux limites de Dirichlet

Comme en 1D, la prise en compte des conditions aux limites de Dirichlet se fait a posteriori en identifiant les degrés de liberté sur le bord Γ concernés par une condition aux limites de Dirichlet. On note I l'ensemble des indices intérieurs et B l'ensemble des indices du bord Γ concernés par les conditions de Dirichlet. Pour récupérer tous les indices du bord, il suffit de récupérer dans `refneu` tous les indices qui ont une référence correspondant au bord.

Conditions de Dirichlet homogènes Considérons dans un premier temps la formulation variationnelle suivante : Chercher $u \in \mathcal{V}_{0h}$ tel que

$$\forall v \in \mathcal{V}_{0h}, \quad \int_{\Omega} [\kappa \underline{\nabla} u \cdot \underline{\nabla} v + uv] \, d\Omega = \int_{\Omega} f v \, d\Omega,$$

où

$$\mathcal{V}_{0h} = \left\{ v \in \mathcal{V}_h \text{ tel que } v|_{\Gamma} = 0 \right\}. \quad (6.28)$$

Il est facile de voir que toutes les fonctions de base définies à la

proposition 6.2.4 ne sont pas dans V_{0h} . En effet,

$$\underline{S}_j \in \partial\Omega, \Rightarrow \phi_j \notin V_{0h}.$$

Cependant, il est facile de montrer que V_{0h} est engendré par les fonctions de base associées à des sommets intérieurs :

$$\mathcal{V}_{0h} = \text{vect}\{\phi_j, \underline{S}_j \notin \partial\Omega\}.$$

On note dans la suite N_i le nombre de sommets intérieurs, ou encore la dimension de $\mathcal{V}_{0h}$. Pour définir les matrices de masse et de rigidité du problème, notées respectivement $\mathbb{M}_0$ et $\mathbb{K}_0$, (qui sont de taille $N_i \times N_i$) et le second membre, noté $\vec{\ell}_0$ (de taille $N_i \times 1$) il faudrait donc introduire une nouvelle numérotation des sommets (et donc des fonctions de base)¹². Pour calculer et assembler ces matrices, il faudrait également traiter de manière différente les triangles qui touchent le bord des autres.

12: Il n'y a aucune raison que les sommets intérieurs soient numérotés en premier !

La méthode que nous vous présentons ici est l'extension de la méthode vue en 1D qui s'appelle la méthode de *pseudo-élimination*. La démarche est la suivante.

On construit les matrices de masse et de raideur sur $\mathcal{V}_h$ comme indiqué plus haut sans nous soucier des conditions aux limites. La matrice de la formulation variationnelle sans prendre en compte les conditions aux limites est de taille $N \times N$ et est ici simplement

$$\mathbb{A} = \mathbb{K}_h + \mathbb{M}_h.$$

13: on omettra les indices h pour simplifier l'écriture

Le second membre¹³ $\vec{\ell}_h$ est calculé à partir de toutes les fonctions de base également, il est de taille $N \times 1$.

Il faut ensuite modifier cette matrice pour que la solution calculée soit bien celle qui vérifie les conditions aux limites du problème. Pour comprendre cette étape, commençons par supposer que tous les nœuds intérieurs sont numérotés en premier :

$$\mathcal{V}_{0h} = \text{Vect}\{\phi_j, 1 \leq j \leq N_i\}.$$

Dans ce cas, la matrice $\mathbb{A}$ peut s'écrire bloc par bloc (en utilisant (3.21))

$$\mathbb{A} = \begin{pmatrix} \mathbb{A}_{II} & \mathbb{A}_{IB} \\ \mathbb{A}_{BI} & \mathbb{A}_{BB} \end{pmatrix},$$

avec

$$\begin{cases} \mathbb{A}_{II}^{ij} := a(\phi_j, \phi_i), & i, j \in \llbracket 1, N_i \rrbracket, \\ \mathbb{A}_{IB}^{ij} := a(\phi_j, \phi_i), & i \in \llbracket 1, N_i \rrbracket, j \in \llbracket N_i + 1, N \rrbracket, \\ \mathbb{A}_{BI}^{ij} := a(\phi_j, \phi_i), & i \in \llbracket N_i + 1, N \rrbracket, j \in \llbracket 1, N_i \rrbracket, \\ \mathbb{A}_{BB}^{ij} := a(\phi_j, \phi_i), & i \in \llbracket N_i + 1, N \rrbracket, j \in \llbracket N_i + 1, N \rrbracket, \end{cases} \quad (6.29)$$

et $\vec{\ell}$ s'écrit

$$\vec{\ell} = \begin{pmatrix} \vec{\ell}_I \\ \vec{\ell}_B \end{pmatrix} \quad \text{avec} \quad \begin{cases} \vec{\ell}_I^i = \ell(\phi_i), & i \in \llbracket 1, N_i \rrbracket, \\ \vec{\ell}_B^i = \ell(\phi_i), & i \in \llbracket N_i + 1, N \rrbracket. \end{cases} \quad (6.30)$$

Résoudre le problème de Dirichlet est quant à lui équivalent à résoudre

le système linéaire suivant :

$$\mathbb{A}_{II} \vec{u}_0 = \vec{\ell}_I. \quad (6.31)$$

On observe donc qu'il suffit de modifier la matrice $\mathbb{A}$ et le second membre $\vec{\ell}$ de la façon suivante

$$\mathbb{A} \rightarrow \tilde{\mathbb{A}} = \begin{pmatrix} \mathbb{A}_{II} & 0 \\ 0 & \mathbb{I} \end{pmatrix}, \quad \text{et } \vec{\ell} \rightarrow \tilde{\vec{\ell}} = \begin{pmatrix} \vec{\ell}_I \\ 0 \end{pmatrix},$$

avec $\mathbb{I}$ la matrice identité. Ainsi, résoudre le système linéaire $\tilde{\mathbb{A}} \tilde{\vec{u}} = \tilde{\vec{\ell}}$ est équivalent à résoudre le problème de Dirichlet (6.31) avec

$$\tilde{\vec{u}} = \begin{pmatrix} \vec{u}_0 \\ 0 \end{pmatrix}.$$

Comme dit précédemment, les nœuds du bord ne sont pas toujours numérotés à la fin ou au début. Pour les distinguer des nœuds intérieurs, il suffit d'utiliser le tableau `Refneu`. L'algorithme de pseudo-élimination (voir l'algorithme 3) consiste donc à éliminer les interactions entre nœuds du bord et tous les autres nœuds dans la matrice `EF` et de remplacer la valeur du second membre pour les nœuds du bord par 0 pour finalement se ramener au problème de Dirichlet.

Algorithme 3: Pseudo-élimination de la matrice $\mathbb{A}$ et du second membre $\vec{\ell}_h$ - Dirichlet homogène

```

/* Boucle sur les lignes                                     */
1 for i ∈ {1, ..., N} do
    /* Utilisation du tableau refneu pour distinguer
       les nœuds du bord                                     */
2   if  $S_i$  est un nœud du bord then
3        $\mathbb{A}_{ii} \leftarrow 1$ ;
4        $\vec{\ell}_{hi} \leftarrow 0$ ;
       /* Boucle sur les colonnes et les lignes             */
5       for j ∈ {1, ..., N}, j ≠ i do
6            $\mathbb{A}_{ij} \leftarrow 0$ ;
7            $\mathbb{A}_{ji} \leftarrow 0$ ;

```

Notons qu'il est possible également de réaliser la même action en introduisant une matrice de projection $\mathbb{P}_0 : \mathcal{V}_h \rightarrow \mathcal{V}_{0h}$ telle $\mathbb{P}_0 \mathbb{P}_0 = \mathbb{P}_0$, de taille $N_i \times N$ et de résoudre

$$\mathbb{P}_0 \mathbb{A} \mathbb{P}_0^\top \vec{u}_0 = \mathbb{P}_0 \vec{\ell}.$$

Le nombre d'opérations pour construire la matrice de projection (creuse) et pour réaliser les produits matrice-vecteur est plus élevé que pour l'algorithme proposé (sauf éventuellement pour des langages de programmation où des fonctions optimisées peuvent être utilisées pour réaliser ces opérations).

Conditions de Dirichlet non homogènes Considérons enfin le problème avec des conditions non homogènes : Chercher $u \in \mathcal{V}_h$ tel que $u = g_h$ sur $\partial\Omega$ et

$$\forall v \in \mathcal{V}_{0h}, \quad \int_{\Omega} \left[\kappa \underline{\nabla} u \cdot \underline{\nabla} v + uv \right] d\Omega = \int_{\Omega} f v d\Omega,$$

où on suppose pour simplifier que $g_h \in V_h$ (sinon il faut interpoler la donnée).

Ecrire $u = g_h$ sur $\partial\Omega$ est équivalent à écrire

$$u(\underline{S}_j) = g_h(\underline{S}_j), \quad \underline{S}_j \in \partial\Omega.$$

Nous présentons tout de suite la méthode de pseudo-élimination dans ce cas. Comme précédemment, il faut assembler les matrices et calculer le second membre sans se soucier des conditions aux limites dans un premier temps.

Supposons tout d'abord que les nœuds intérieurs sont numérotés en premier. Résoudre le problème précédent est équivalent à résoudre le système linéaire suivant

$$\mathbb{A}_{II} \vec{u}_g + \mathbb{A}_{BI} \vec{g} = \vec{\ell}_I \text{ où } \vec{g} = (g(\underline{S}_i))_{i \in \llbracket N_i+1, N \rrbracket}, \quad (6.32)$$

et où les matrices $\mathbb{A}_{II}$, $\mathbb{A}_{BI}$ sont données dans (6.29) et $\vec{\ell}_I$ est donnée par (6.30).

Il suffit cette fois de modifier la matrice $\mathbb{A}$ et le second membre $\vec{\ell}$ de la façon suivante

$$\mathbb{A} \rightarrow \widehat{\mathbb{A}} = \begin{pmatrix} \mathbb{A}_{II} & \mathbb{A}_{BI} \\ 0 & \mathbb{I} \end{pmatrix}, \quad \text{et } \vec{\ell} \rightarrow \vec{\widehat{\ell}}_g = \begin{pmatrix} \vec{\ell}_I \\ \vec{g} \end{pmatrix},$$

et résoudre le système linéaire $\widehat{\mathbb{A}} \vec{\widehat{u}}_g = \vec{\widehat{\ell}}_g$ est équivalent à résoudre le problème de Dirichlet (6.32) avec

$$\vec{\widehat{u}}_g = \begin{pmatrix} \vec{u}_g \\ \vec{g} \end{pmatrix}.$$

Algorithme 4: Pseudo-élimination de la matrice $\mathbb{A}$ et du second membre $\vec{\ell}_h$ - Dirichlet non homogène

```

/* Boucle sur les lignes */
1 for  $i \in \{1, \dots, N\}$  do
    /* Utilisation du tableau refneu pour distinguer
       les nœuds du bord */
2   if  $\underline{S}_i$  est un nœud du bord then
3      $\mathbb{A}_{ii} \leftarrow 1$ ;
4      $\vec{\ell}_{hi} \leftarrow g(\underline{S}_i)$ ;
    /* Boucle sur les colonnes et les lignes */
5     for  $j \in \{1, \dots, N\}, j \neq i$  do
6        $\mathbb{A}_{ij} \leftarrow 0$ ;

```

À noter que dans l'algorithme 4, la matrice résultante n'est pas symétrique même si le problème l'était initialement. Ce n'est pas forcément très satisfaisant pour des algorithmes d'inversion qui peuvent bénéficier du caractère symétrique de la matrice. Dans ce cas, on pseudo-élimine aussi sur les colonnes en prenant soin de retrancher d'abord la contribution des nœuds du bord :

$$\mathbb{A} \rightarrow \tilde{\mathbb{A}} = \begin{pmatrix} \mathbb{A}_{II} & 0 \\ 0 & \mathbb{I} \end{pmatrix}, \quad \text{et } \vec{\ell} \rightarrow \vec{\tilde{\ell}}_g = \begin{pmatrix} \vec{\ell}_I - \mathbb{A}_{BI}\vec{g} \\ \vec{g} \end{pmatrix},$$

On calcule dans ce cas non pas u mais $u - \sum_{i, \underline{S}_i \in \partial\Omega} g(\underline{S}_i)\phi_i$.

6.3 Éléments finis de Lagrange généraux

En fait, il existe une formulation abstraite de la notion d'élément fini de Lagrange que nous proposons maintenant de présenter succinctement. Cette définition est générale sur $\mathbb{R}^d$ mais nécessite d'étendre la notion de maillage dans $\mathbb{R}^d$.

6.3.1 Éléments finis de Lagrange $\mathbb{P}_k$

Commençons par généraliser ce que nous avons vu à $\mathbb{R}^d$ pour $d = 3$, sachant que le cas d quelconque se déduira sans peine. Le cadre général nécessite de pouvoir définir la notion de maillage symplectique dans $\mathbb{R}^d$. En dimension 3, la généralisation du maillage de triangles est donc un maillage de tétraèdres.

Définition 6.3.1 (Maillage tétraédrique) Soit Ω un ouvert connexe polyédrique de $\mathbb{R}^3$. Un maillage tétraédrique de $\overline{\Omega}$ est un ensemble noté $\mathcal{T}_h$ de N_e tétraèdres $\{T_l\}_{1 \leq l \leq N_e}$ tels que

- Tout tétraèdre T_k est d'intérieur non vide (c'est-à-dire $\overset{\circ}{T}_k \neq \emptyset$). On définit alors ρ_k le rayon de la plus grande boule contenue dans T_k et h_k le rayon de la plus petite boule contenant T_k .
- $\overset{\circ}{T}_k \cap \overset{\circ}{T}_{k'} = \emptyset$ si $k \neq k'$,
- $\bigcup_l \overset{\circ}{T}_l = \overset{\circ}{\Omega}$,
- toute face triangulaire d'un tétraèdre est soit une face d'un autre tétraèdre soit un triangle porté par la frontière $\partial\Omega$,

On notera la similitude avec la définition 6.2.1 et on pourra donc généraliser cette définition à n'importe quelle dimension en remplaçant tétraèdre par d -simplexe et face par $d - 1$ simplexe.

On peut ainsi introduire l'espace d'approximation éléments finis $\mathbb{P}_k$ pour $k \geq 1$ par

$$\mathcal{V}_h = \mathcal{V}_h^{\mathbb{P}_k} = \left\{ v_h \in C^0(\overline{\Omega}) \text{ tel que } v_h|_{T_i} \in \mathbb{P}_k \forall T_i \in \mathcal{T}_h \right\}, \quad (6.33)$$

où $\mathbb{P}_k(T_l)$ désigne l'ensemble des polynômes multivariés sur T_l de degré au maximum k .

On peut maintenant donner une définition générale et abstraite d'un élément fini de Lagrange, qui s'appliquera notamment aux éléments finis de Lagrange $\mathbb{P}_k$.

Définition 6.3.2 On appelle élément fini de Lagrange d'ordre m un triplet $(K, \Sigma, \mathcal{P})$ où K est un fermé borné non vide de $\mathbb{R}^d$, Σ un ensemble de N_K nœuds $(M_i^K)_{i=1, N_K}$ appartenant à K et $\mathcal{P}$ un espace vectoriel de polynômes contenant $\mathcal{P}^m(K)$ (espace des polynômes de degré au plus m sur K) tels que :

$$\forall (\alpha^1, \alpha^2, \dots, \alpha^{N_K})^\top \in \mathbb{R}^{N_K}, \exists ! p \in \mathcal{P} \text{ tel que} \\ \forall i \in \llbracket 1, N_K \rrbracket, \quad p(M_i^K) = \alpha^i. \quad (6.34)$$

Un point de Σ s'appelle un degré de liberté de Lagrange et lorsque la propriété (6.34) est vérifiée on dit que $\mathcal{P}$ est Σ -unisolvant. Cette propriété exprime le fait que l'application :

$$\Phi_\Sigma : \begin{cases} \mathcal{P} \rightarrow \mathbb{R}^{N_K} \\ p \mapsto (p(M_1^K), \dots, p(M_{N_K}^K)) \end{cases}$$

est bijective, c'est-à-dire que toute fonction de l'espace d'approximation est déterminée de façon unique par les valeurs qu'elle prend aux degrés de liberté.

Revoyons quelques exemples en dimension 2 afin de mettre en pratique cette définition générale.

Élément fini $\mathbb{P}_0$: Soit $\hat{T}$ le triangle de référence constitué par les points $\hat{S}_1 = (0, 0)$, $\hat{S}_2 = (1, 0)$ et $\hat{S}_3 = (0, 1)$

$$\left| \begin{array}{l} K = \hat{T} \\ \Sigma = \left\{ \left(\frac{1}{3}, \frac{1}{3} \right) \right\} \\ \mathcal{P}(K) = \mathbb{P}_0(\hat{T}) = \{c, c \in \mathbb{R}\} \end{array} \right.$$

14: Il est important de remarquer que cet élément fini permet de définir des fonctions qui seront constantes par élément mais pas continues. Elles ne rentreront donc pas dans le cadre de l'approximation interne que nous verrons au chapitre 7, mais pourront s'avérer utiles pour des approximations de fonctionnelles définies uniquement sur $L^2(\Omega)$.

La fonction de forme associée¹⁴ est

$$\hat{\lambda}_{123} = 1.$$

Élément fini $\mathbb{P}_1$:

$$\left| \begin{array}{l} K = \hat{T} \\ \Sigma = \{\hat{S}_1, \hat{S}_2, \hat{S}_3\} \\ \mathcal{P}(K) = \mathbb{P}_1(\hat{T}) = \{ax + by + c, (a, b, c) \in \mathbb{R}^3\} \end{array} \right.$$

Les fonctions de forme associées sont

$$\left| \begin{array}{l} \hat{\lambda}_1(x, y) = 1 - \hat{x} - \hat{y}, \\ \hat{\lambda}_2(x, y) = \hat{x}, \\ \hat{\lambda}_3(x, y) = \hat{y}. \end{array} \right.$$

Élément fini $\mathbb{P}_2$: On définit de plus les milieux $\hat{S}_{12}, \hat{S}_{23}, \hat{S}_{13}$ des segments $(\hat{S}_1, \hat{S}_2), (\hat{S}_2, \hat{S}_3)$ et $(\hat{S}_1, \hat{S}_3)$. On a alors

$$\left| \begin{array}{l} K = \hat{T} \\ \Sigma = \{\hat{S}_1, \hat{S}_2, \hat{S}_3, \hat{S}_{12}, \hat{S}_{23}, \hat{S}_{13}\} \\ \mathcal{P}(K) = \mathbb{P}_2(\hat{T}) = \{ax^2 + by^2 + cxy + dx + ey + f, (a, b, c, d, e, f) \in \mathbb{R}^6\} \end{array} \right.$$

Les fonctions de forme associées sont

$$\left| \begin{array}{l} \hat{\lambda}_1(x, y) = (1 - \hat{x} - \hat{y})(1 - 2\hat{x} - 2\hat{y}), \\ \hat{\lambda}_2(x, y) = \hat{x}(2\hat{x} - 1), \\ \hat{\lambda}_3(x, y) = \hat{y}(2\hat{y} - 1), \\ \hat{\lambda}_{12}(x, y) = 4\hat{x}(1 - \hat{x} - \hat{y})(1 - 2\hat{x} - 2\hat{y}), \\ \hat{\lambda}_{23}(x, y) = 4\hat{x}\hat{y}, \\ \hat{\lambda}_{13}(x, y) = 4\hat{y}(1 - \hat{x} - \hat{y}). \end{array} \right.$$

Une stratégie similaire à la dimension 1 permet de retrouver rapidement ces fonctions. Il suffit d'écrire les équations des droites sur lesquelles les fonctions doivent s'annuler. Par exemple $\hat{\lambda}_1$ doit s'annuler sur $(\hat{S}_2\hat{S}_3)$ d'équation $(1 - \hat{x} - \hat{y}) = 0$ et sur $(\hat{S}_{12}\hat{S}_{23})$ d'équation $(1 - 2\hat{x} - 2\hat{y}) = 0$.

Remarque 6.3.1 On notera que les éléments finis $\mathbb{P}^k$ vus en dimension 1 rentrent aussi dans le cadre de la définition 6.3.2.

— Ordre 0 :

$$\left| \begin{array}{l} K = [0, 1] \\ \Sigma = \left\{\frac{1}{2}\right\} \\ \mathcal{P}(K) = \mathbb{P}_0([0, 1]) = \{c, c \in \mathbb{R}\} \end{array} \right.$$

— Ordre 1 :

$$\left| \begin{array}{l} K = [0, 1] \\ \Sigma = \{0, 1\} \\ \mathcal{P}(K) = \mathbb{P}_1([0, 1]) = \{bx + c, (b, c) \in \mathbb{R}^2\} \end{array} \right.$$

— Ordre 2 :

$$\left| \begin{array}{l} K = [0, 1] \\ \Sigma = \left\{0, 1, \frac{1}{2}\right\} \\ \mathcal{P}(K) = \mathbb{P}_2([0, 1]) = \{ax^2 + bx + c, (a, b, c) \in \mathbb{R}^3\} \end{array} \right.$$

Revenons à la définition de $\mathcal{V}_h^{\mathbb{P}^k}$. On peut montrer que dès que le triplet $(K, \Sigma, \mathbb{P}_k)$ est bien défini et est unisolvant, il est possible de caractériser les degrés de liberté du maillage construit à partir de Σ , introduire une base de fonctions dite de Lagrange qui permet de caractériser l'espace.

Proposition 6.3.1 L'ensemble des degrés de liberté de l'espace $\mathcal{V}_h^{\mathbb{P}^k}$ est donné

par

$$\Sigma_h := \{F_l(\Sigma), \quad F_l(\hat{T}) = T_l, \quad l \in \llbracket 1, N_e \rrbracket\} = \{\underline{M}_1, \dots, \underline{M}_N\}$$

Une base de $\mathcal{V}_h^{\mathbb{P}_k}$ est donnée par les fonctions $(\phi_i)_{1 \leq i \leq N}$ telles que

$$\phi_i(\underline{M}_j) = \delta_i^j \quad i, j \in \llbracket 1, N \rrbracket,$$

et

$$\forall v \in \mathcal{V}_h^{\mathbb{P}_k}, \quad v = \sum_{i=1}^N v(\underline{M}_i) \phi_i(x).$$

La preuve est très similaire à celle de la proposition 6.2.4 dès qu'on s'est assuré que $\mathbb{P}_k$ est Σ -unisolvant.

Ainsi en dimension 2, d'après ce qui précède, on déduit que

$$\dim \mathcal{V}_h^{\mathbb{P}_1} = N_s, \quad \dim \mathcal{V}_h^{\mathbb{P}_2} = N_s + N_a, \quad \dim \mathcal{V}_h^{\mathbb{P}_3} = N_s + 2N_a + N_e, \dots$$

6.3.2 Éléments finis de Lagrange $\mathbb{Q}_k$

En dimension 2, il est possible de définir des éléments finis, notamment de Lagrange, sur des fermés bornés K qui ne sont pas des triangles, mais des quadrangles. On présente ici comme exemples les deux éléments $\mathbb{Q}_1$ et $\mathbb{Q}_2$ en dimension 2, conformément à la définition 6.3.2.

Élément fini $\mathbb{Q}_1$:

$$\left| \begin{array}{l} K = \hat{Q} \\ \Sigma = \{\hat{S}_1, \hat{S}_2, \hat{S}_3, \hat{S}_4\} \\ \mathcal{P}(K) = \mathbb{Q}_1(\hat{T}) = \{ax + by + cxy + d, (a, b, c, d) \in \mathbb{R}^4\} \end{array} \right.$$

Les fonctions de forme associées sont

$$\left| \begin{array}{l} \hat{\lambda}_1(x, y) = (1 - \hat{x})(1 - \hat{y}), \\ \hat{\lambda}_2(x, y) = \hat{x}(1 - \hat{y}), \\ \hat{\lambda}_3(x, y) = \hat{x}\hat{y}, \\ \hat{\lambda}_4(x, y) = \hat{y}(1 - \hat{x}). \end{array} \right.$$

dont on voit qu'elles correspondent à la tensorisation des fonctions de bases $\mathbb{P}_1([0, 1])$ en dimension 1. On remarquera que les fonctions de $\mathbb{Q}_1(\hat{Q})$ ne sont pas affines mais contiennent des termes quadratiques de type $\hat{x}\hat{y}$.

Élément fini $\mathbb{Q}_2$: On définit de plus les milieux $\hat{S}_{12}, \hat{S}_{23}, \hat{S}_{34}, \hat{S}_{41}$, des segments $(\hat{S}_1, \hat{S}_2), (\hat{S}_2, \hat{S}_3), (\hat{S}_3, \hat{S}_4)$ et $(\hat{S}_1, \hat{S}_4)$. Et on ajoute enfin le

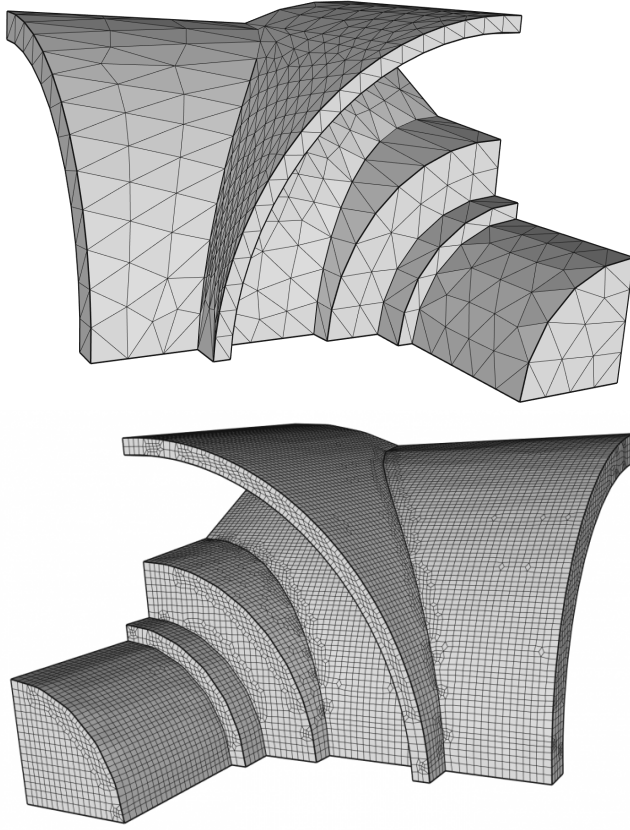


FIGURE 6.6 – Maillage tétrahédrique et conversion en maillage hexaédrique réalisée par le code Hexotic. Reproduction autorisée par Loïc Maréchal, source team.inria.fr/gamma/gamma-software/hexotic

barycentre de quadrilatères $\hat{S}_{1234}$ On a alors

$$\left| \begin{array}{l} K = \hat{Q} \\ \Sigma = \{\hat{S}_1, \hat{S}_2, \hat{S}_3, \hat{S}_4, \hat{S}_{12}, \hat{S}_{23}, \hat{S}_{34}, \hat{S}_{14}, \hat{S}_{1234}\} \\ \mathcal{P}(K) = \mathbb{Q}_2(\hat{T}) = \left\{ \sum_{0 \leq i, j \leq 2} \alpha_{ij} \hat{x}^i \hat{y}^j (\alpha_{ij}) \in \mathbb{R}^9 \right\} \end{array} \right.$$

Les fonctions de forme associées sont de nouveau construites par tensorisation des fonctions de base $\mathbb{P}_2([0, 1])$. On obtient donc 9 fonctions de bases chacune s'annulant sur 8 nœuds de l'élément et valant 1 sur le nœud restant.

En dimension 3, on peut définir le même type d'éléments finis de Lagrange, mais cette fois K devient un hexaèdre.

L'utilisation de tels éléments finis nécessite cependant de construire des maillages « cubiques ». Si cela peut être élémentaire sur des domaines eux-mêmes cubiques, c'est beaucoup plus délicat pour des géométries plus générales [22], voir un exemple en dimension 3 à la figure 6.6.

Définition 6.3.3 Soit Ω un ouvert connexe polyédrique de $\mathbb{R}^d$. Un maillage cubique de $\overline{\Omega}$ est un ensemble $\mathcal{Q}_h$ de d -cubes (non dégénérés) $(Q_i)_{1 \leq i \leq n}$ qui vérifient

1. $Q_i \subset \overline{\Omega}$ et $\overline{\Omega} = \cup_{i=1}^n Q_i$,
2. l'intersection $Q_i \cap Q_j$ de deux d -cubes distincts est une face (ou un segment), dont tous les sommets sont aussi des sommets de Q_i et Q_j .

Les sommets du maillage $\mathcal{Q}_h$ sont les sommets des d -cubes Q_i qui le composent. Par convention, le paramètre h désigne le maximum des diamètres des d -cubes Q_i .

En toute généralité, nous définissons l'ensemble $\mathbb{Q}_k$ des polynômes à coefficients réels de $\mathbb{R}^d$ dans $\mathbb{R}$ de degré inférieur ou égal à k par rapport à chaque variable, c'est-à-dire que tout $p \in \mathbb{Q}_k$ s'écrit sous la forme

$$p(\underline{x}) = \sum_{0 \leq i_1 \leq k, \dots, 0 \leq i_N \leq k} \alpha_{i_1, \dots, i_N} x_1^{i_1} \cdots x_N^{i_N} \text{ avec } \underline{x} = (x_1, \dots, x_N).$$

Et nous soulignons que le degré de p peut être supérieur à k , ce qui différencie l'espace $\mathbb{Q}_k$ de $\mathbb{P}_k$.

Soit $Q = \Pi_{1 \leq i \leq d} [l_i, L_i]$ un d -cube généralisé de $\mathbb{R}^d$. Pour tout entier $k \geq 1$, on définit le treillis d'ordre k du d -cube Q comme l'ensemble

$$\Sigma_k = \left\{ \underline{x} \in Q \text{ tel que } \frac{x_j - l_j}{L_j - l_j} \in \left\{ 0, \frac{1}{k}, \dots, \frac{k-1}{k}, 1 \right\}, 1 \leq j \leq N \right\}. \quad (6.35)$$

Le treillis est unisolvant pour les polynômes $\mathbb{Q}_k$

Lemme 6.3.2 Soit Q un d -cube. Soit un entier $k \geq 1$. Alors, tout polynôme de $\mathbb{Q}_k$ est déterminé de manière unique par ses valeurs aux points du treillis d'ordre k , Σ_k , défini par (6.35).

Lemme 6.3.3 Soit Q et Q' deux d -cubes ayant une face commune $\Gamma = \partial Q \cap \partial Q'$. Soit un entier $k \geq 1$. Alors, leur treillis d'ordre k Σ_k et Σ'_k coïncident sur cette face Γ . De plus, étant donné p_Q et $p_{Q'}$ deux polynômes de $\mathbb{Q}_k$, la fonction v définie par

$$v(\underline{x}) = \begin{cases} p_Q(\underline{x}) & \text{si } \underline{x} \in Q \\ p_{Q'}(\underline{x}) & \text{si } \underline{x} \in Q' \end{cases}$$

est continue sur $Q \cup Q'$, si et seulement si p_Q et $p_{Q'}$ ont des valeurs qui coïncident aux points du treillis sur la face commune Γ .

On peut ainsi introduire l'espace d'approximation éléments finis $\mathbb{Q}_k$ par

$$\mathcal{V}_h = \mathcal{V}_h^{\mathbb{Q}_k} = \left\{ v \in C^0(\bar{\Omega}) \text{ tel que } v|_{Q_i} \in \mathbb{Q}_k \text{ pour tout } Q_i \in \mathcal{Q}_h \right\}. \quad (6.36)$$

Pour caractériser cet espace, il nous faudra aussi définir la transformation de l'élément de référence à l'élément courant. Pour passer du cube unité $\hat{Q} = [0, 1]^d$ à un d -cube convexe Q_ℓ quelconque, on introduit une transformation $\underline{\varphi}_\ell$, comme pour les triangles. Cette fois-ci, $\underline{\varphi}_\ell$ n'est pas systématiquement affine¹⁵. Par exemple, en 2D, elle est de la forme

$$\underline{\varphi}_\ell(\hat{x}, \hat{y}) = \begin{pmatrix} a_1^l \hat{x} \hat{y} + b_1^l \hat{x} + c_1^l \hat{y} + d_1^l \\ a_2^l \hat{x} \hat{y} + b_2^l \hat{x} + c_2^l \hat{y} + d_2^l \end{pmatrix},$$

où les 8 coefficients sont définis tels que

$$\underline{\varphi}_\ell(\hat{S}_i) = \underline{S}_i^l \quad 1 \leq i \leq 4.$$

15: On peut vérifier que la transformation est affine si, et seulement si, Q_ℓ est un parallélogramme

où $\{\hat{S}_i, 1 \leq i \leq 4\}$ sont les sommets du carré $\hat{Q} = [0, 1]^2$ et $\{S_i^l, 1 \leq i \leq 4\}$ les sommets du quadrilatère Q_l . On vérifiera que comme pour les triangles, si le d -cube n'est pas dégénéré alors cette transformation est bien définie et inversible. Nous renvoyons à [16] pour plus de détails.

Enfin, nous pouvons caractériser l'espace $\mathcal{V}_h^{\mathbb{Q}_k}$ à partir de la base d'éléments finis associée.

Proposition 6.3.4 *L'ensemble des degrés de liberté de l'espace $\mathcal{V}_h^{\mathbb{Q}_k}$ est donné par*

$$\Sigma_k := \left\{ \varphi_\ell(\Sigma), \quad \varphi_\ell(\hat{Q}) = Q_\ell, \quad Q_\ell \in \mathcal{Q}_h \right\} = \{\underline{M}_1, \dots, \underline{M}_N\}.$$

Une base de $\mathcal{V}_h^{\mathbb{Q}_k}$ est donnée par les fonctions $(\phi_i)_{1 \leq i \leq N}$ telles que

$$\phi_i \in \mathcal{V}_h^{\mathbb{Q}_k}, \quad \phi_i(\underline{M}_j) = \delta_i^j \quad i, j \in \llbracket 1, N \rrbracket,$$

et

$$\forall v \in \mathcal{V}_h^{\mathbb{Q}_k}, \quad v = \sum_{i=1}^{N_k} v(\underline{M}_i) \phi_i(x).$$

6.4 Éléments finis : cadre général

Nous pouvons conclure ce chapitre sur les éléments finis en passant au cadre permettant de définir un espace éléments finis en toute généralité.

En effet, les éléments finis de Lagrange ne permettent pas de traiter toutes les situations. Nous avons commencé à le voir avec le problème du 4e ordre traité avec les éléments finis de Hermite. Certains éléments finis pourront aussi être construits en contrôlant les flux aux interfaces et ainsi être plus adaptés à certaines lois de conservations, etc.

Définition 6.4.1 (Élément fini) *Soit K un fermé borné connexe d'intérieur non vide de $\mathbb{R}^d$. On appelle élément fini sur K la donnée d'un triplet $(K, \Sigma, \mathcal{P})$, où*

- $\mathcal{P}$ est un espace vectoriel de dimension N_K de fonctions $C^\infty(K)$
- Σ est un ensemble de N_K formes linéaires¹⁶ $(\sigma^i)_{1 \leq i \leq N_K}$ sur $\mathcal{P}$ tels que

$$\forall (\alpha^1, \dots, \alpha^{N_K}) \in \mathbb{R}^{N_K}, \quad \exists ! p \mid \sigma^i(p) = \alpha^i. \quad (6.37)$$

Les formes linéaires $(\sigma^i)_{1 \leq i \leq N_K}$ sont appelées les degrés de liberté.

16: c'est à dire une distribution sur $C^\infty(K)$ avec K un fermé

Une conséquence directe de la définition est que

$$\Phi_\Sigma : \mathcal{P} \ni p \mapsto (\sigma^i(p))_{1 \leq i \leq N_K}, \quad (6.38)$$

est un isomorphisme. Notamment, on a

$$\forall p \in \mathcal{P}, \quad [\forall i \in \llbracket 1, N_K \rrbracket, \sigma^i(p) = 0] \Rightarrow [p = 0], \quad (6.39)$$

qu'on appelle la *propriété d'unisolvance*.

La proposition suivante est immédiate.

Proposition 6.4.1 (Fonctions de forme) *Il existe une base $(\lambda_i)_{1 \leq i \leq N_K}$ de $\mathcal{P}$ telle que*

$$\forall i, j \in \llbracket 1, N_K \rrbracket^2, \quad \sigma^i(\lambda_j) = \delta_i^j.$$

Les (λ_i) sont appelées fonctions de formes.

Évidemment la définition 6.3.2 des éléments finis de Lagrange est compatible avec la définition générale pour peu qu'on définisse non pas le support des points $(M_i)_{1 \leq i \leq N_K}$ mais les formes linéaires associées à l'évaluation en ces points $\Sigma = (\delta_{M_i})_{1 \leq i \leq N_K}$ définies¹⁷ par

$$\forall p \in C^0(K), \quad \delta_{M_i}(p) = p(\underline{M}_i).$$

On vérifiera aussi que l'élément fini de Hermite en 1D rentre aussi dans cette définition, où nous avons déjà introduit les formes linéaires $(\sigma^i)_{1 \leq i \leq 4}$. Comme dernier exemple d'éléments finis généraux, nous présentons une classe d'éléments finis très importante constituée des éléments finis de type moments.

Proposition 6.4.2 (Éléments de type moments) *Soit K un fermé borné connexe d'intérieur non vide de $\mathbb{R}^d$. Soit $\mathcal{P}$ un sous-espace de dimension finie de $L^2(K)^d$ et $(\underline{\lambda}_i)_{1 \leq i \leq N_K}$ une base de $\mathcal{P}$. On définit $\Sigma = (\sigma^i)_{1 \leq i \leq N_K}$ où pour tout $i \in \llbracket 1, N_K \rrbracket$*

$$\sigma^i : \mathcal{P} \ni \underline{p} \mapsto \frac{1}{|K|} \int_K \underline{\lambda}_i \cdot \underline{p} \, d\Omega.$$

Alors $(K, \Sigma, \mathcal{P})$ est un élément fini appelé élément fini modal

Démonstration. La preuve est classique. On a d'abord $\dim \mathcal{P} = \text{card } \Sigma$. De plus soit $\underline{p} = \sum_{1 \leq j \leq N_K} \alpha^j \underline{\lambda}_j \in \mathcal{P}$, tel que $\sigma^i(\underline{p}) = 0$ pour tout $i \in \llbracket 1, N_K \rrbracket$, alors

$$0 = \sum_{1 \leq i \leq N_K} \alpha^i \sigma^i(\underline{p}) = \frac{1}{|K|} \int_K |\underline{p}|^2 \, dK,$$

donc $\underline{p} \equiv 0$ ce qui démontre la propriété d'unisolvance. $\square$

Il existe de nombreux autres éléments finis comme les éléments finis de Nédélec¹⁸, de Raviart-Thomas, de Crouzeix-Raviart, etc. Nous renvoyons notamment à [12] pour une présentation beaucoup plus exhaustive.

17: ce sont des distributions de Dirac

18: Jean-Claude Nédélec et Pierre-Arnault Raviart ont été enseignants-chercheurs à l'École polytechnique

Analyse numérique de la méthode des éléments finis

7

Dans ce chapitre nous procédons à l'*analyse numérique* de la méthode des *éléments finis* qui a été présentée dans les chapitres 5 et 6. L'analyse numérique consiste à contrôler quantitativement l'erreur d'approximation en fonction du raffinement du paramètre h . Nous avons déjà vu au chapitre¹ 3 avec le lemme de Céa 3.6.2 puis le théorème 3.6.4 que la convergence ne nécessite pas d'hypothèse sur la solution du problème dans $\mathcal{V}$. Nous allons voir dans ce chapitre que pour estimer la vitesse de convergence, des hypothèses supplémentaires sur la solution sont nécessaires. De plus, cette vitesse de convergence est éventuellement limitée par la régularité de la solution. Ce type d'analyse est fondamental dès qu'on souhaite offrir des garanties sur l'approximation.

1: Attention dans le chapitre 3, les espaces de dimension finie sont indexés par la dimension de l'espace alors que dans la méthode des éléments et donc dans ce chapitre, les espaces sont indexés par le pas du maillage h .

7.1 Introduction

Nous considérons à nouveau le cadre général du formalisme variationnel introduit au Chapitre 3. Étant donné un espace de Hilbert $\mathcal{V}$, une forme bilinéaire continue et coercive $\mathcal{V}^2 \ni (u, v) \mapsto a(u, v) \in \mathbb{R}$, et une forme linéaire continue $\mathcal{V} \ni v \mapsto \ell(v) \in \mathbb{R}$. On notera dans toute la suite $\nu > 0$ la constante de coercivité et $M > 0$ la constante de continuité de la forme bilinéaire $a(u, v)$ qui vérifie donc

$$\begin{aligned} a(u, u) &\geq \nu \|u\|_{\mathcal{V}}^2 \quad \forall u \in \mathcal{V}, \\ |a(u, v)| &\leq M \|u\|_{\mathcal{V}} \|v\|_{\mathcal{V}} \quad \forall u, v \in \mathcal{V}. \end{aligned}$$

On considère la formulation variationnelle :

$$\text{Trouver } u \in \mathcal{V} \text{ tel que } a(u, v) = \ell(v), \quad \forall v \in \mathcal{V}, \quad (7.1)$$

dont on sait qu'elle admet une unique solution par le théorème 3.5.1 de Lax-Milgram.

Nous avons vu que l'*approximation interne* de (7.1) consiste à remplacer l'espace de Hilbert $\mathcal{V}$ par un sous-espace de dimension finie $\mathcal{V}_h \subset \mathcal{V}$ muni de la norme de $\mathcal{V}$ où on cherche la solution de

$$\text{trouver } u_h \in \mathcal{V}_h \text{ tel que } a(u_h, v_h) = \ell(v_h) \quad \forall v_h \in \mathcal{V}_h. \quad (7.2)$$

Nous avons au chapitre 3 et le théorème 3.6.4 que

$$\lim_{h \rightarrow 0} \|u - u_h\|_{\mathcal{V}} = 0, \quad (7.3)$$

où u est la solution de (7.1) et u_h celle de (7.2).

Dans la suite de ce chapitre, nous allons justement effectuer l'analyse numérique qualitative, en étudiant la vitesse de convergence de l'approximation.

Définition 7.1.1 (Vitesse de convergence) *On dit que l'approximation interne est convergente à l'ordre k s'il existe une constante $C > 0$, indépendante de h telle que*

$$\|u - u_h\|_{\mathcal{V}} \leq Ch^k.$$

7.2 Vitesse de convergence de la méthode des éléments finis

7.2.1 Une estimation d'erreur en dimension 1

Problème modèle Pour fixer les idées, nous nous appuyons sur un problème modèle, mais les résultats que nous allons présenter sont en fait très généraux. Soient $L > 0$ et $f \in L^2(]0, L[)$, on considère le problème modèle suivant : On cherche la solution $u \in \mathcal{V} = H^1(]0, L[)$ de

$$\begin{aligned} \forall v \in H^1(]0, L[), \quad \int_0^L [u(x)v(x) + u'(x)v'(x)] dx \\ = \int_0^L f(x)v(x) dx, \end{aligned} \quad (7.4)$$

correspondant à la formulation forte

$$\begin{cases} -u'' + u = f & \text{dans }]0, L[, \\ u'(0) = u'(L) = 0. \end{cases} \quad (7.5)$$

La formulation variationnelle (7.4) est approchée par une méthode de Galerkin de type éléments finis de Lagrange d'ordre 1. Nous considérons donc une discrétisation $\mathcal{I}_h$ du domaine $[0, L]$ de la forme,

$$0 = x_0 < x_1 < \dots < x_{N_i} < x_{N_i+1} = L.$$

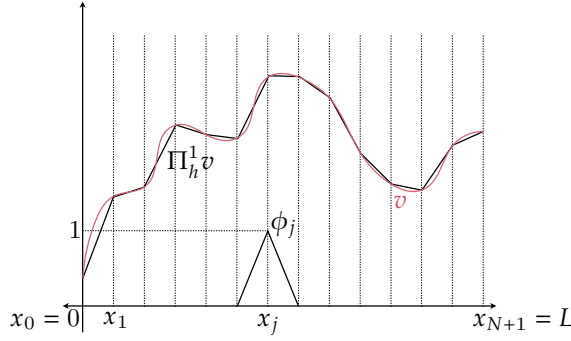
On note $h = \max_{0 \leq i \leq N_i} (x_{i+1} - x_i)$, et $\mathcal{V}_h$ l'espace d'approximation défini en (5.5), voir le chapitre 5 pour plus de détails. On rappelle que l'espace $\mathcal{V}_h$ des éléments finis de Lagrange d'ordre 1 est inclus dans les fonctions continues par ailleurs C^1 par morceaux sur une partition finie du domaine. Donc par le théorème 2.2.2, l'espace $\mathcal{V}_h$ est un sous-espace de $H^1(\Omega)$.

Opérateur d'interpolation Dans le cadre d'une approximation $\mathbb{P}_1$, les fonctions de base $(\phi_j)_j$ sont alors les fonctions de forme vues au paragraphe 5.2.1. Pour démontrer la convergence de la méthode des éléments finis, nous définissons tout d'abord un opérateur d'interpolation.

Définition 7.2.1 (Opérateur d'interpolation $\mathbb{P}_1$) *On appelle opérateur d'interpolation $\mathbb{P}_1$ l'application linéaire Π_h^1 de $H^1(]0, L[)$ dans $\mathcal{V}_h = \mathbb{P}_1(\mathcal{I}_h)$ définie, pour tout $v \in H^1(]0, L[)$, par*

$$(\Pi_h^1 v)(x) = \sum_{j=0}^{N_i+1} v(x_j) \phi_j(x).$$

Cette définition a bien un sens, car, en vertu du théorème 2.2.10, les fonctions de $H^1([0, L])$ sont continues et leurs valeurs ponctuelles sont donc bien définies². L'interpolée $\Pi_h^1 v$ d'une fonction v est simplement la fonction affine par morceaux qui coïncide avec v sur les sommets du maillage x_j , comme illustré sur la figure 7.1.



2: En dimension 1, l'interpolée est définie sur $H^1([0, L])$, et non pas seulement pour des fonctions régulières (continues) de $H^1([0, L])$. Ce sera différent en dimension supérieure

FIGURE 7.1 – Opérateur d'interpolation Π_h^1

L'estimation de la vitesse de convergence de la méthode des éléments finis $\mathbb{P}_1$ va reposer sur la quantification de l'erreur d'interpolation du type $\|v - \Pi_h^1 v\|_{H^1}$. Pour ce faire nous allons suivre une méthode systématique qui s'appuie sur le passage à l'élément de référence. Cette approche se généralisera ainsi aux dimensions supérieures. Commençons par introduire des semi-normes locales sur chaque élément.

Définition 7.2.2 (Semi-normes locales) *Sur chaque intervalle ouvert $I_l = (x_l, x_{l+1})$ avec $l \in \llbracket 0, N_i \rrbracket$, on notera pour tout $m \in \mathbb{N}$ les semi-normes³*

$$|v|_{m, I_l} = \left(\int_{x_l}^{x_{l+1}} |v^{(m)}|^2 dx \right)^{\frac{1}{2}}.$$

3: et notamment pour $m = 0$, on retrouve la norme $L^2(x_l, x_{l+1})$

Lemme 7.2.1 (Lemme de Bramble-Hilbert en dimension 1) *Il existe une constante⁴ C telle que pour tout $l \in \llbracket 0, N_i \rrbracket$*

$$\forall v \in H^2(I_l), \quad |v - \Pi_h^1 v|_{0, I_l} \leq C h_l^2 |v|_{2, I_l}, \quad (7.6)$$

et

$$\forall v \in H^2(I_l), \quad |v - \Pi_h^1 v|_{1, I_l} \leq C h_l |v|_{2, I_l}. \quad (7.7)$$

4: indépendante donc du diamètre $h_l = |x_{l+1} - x_l|$ de l'intervalle ouvert $I_l = (x_l, x_{l+1})$

Démonstration. Supposons que $v \in D(\bar{I}_l) := C_c^\infty([x_l, x_{l+1}])$ et on conclura par densité. On définit

$$w = v - \Pi_h^1 v \in D(\bar{I}_l).$$

La transformation entre l'élément de référence $\hat{I} = [0, 1]$ et I_l est donnée par

$$\varphi : [0, 1] \ni \hat{x} \mapsto h_l \hat{x} + x_l \in [x_l, x_{l+1}],$$

telle que

$$|w|_{0, I_l}^2 = \int_{x_l}^{x_{l+1}} |w(x)|^2 dx = h_l \int_0^1 |\hat{w}(\hat{x})|^2 d\hat{x} = h_l |\hat{w}|_{0, \hat{I}}^2, \quad (7.8)$$

où $\hat{w} = w \circ \varphi$. Puisque $\hat{w}' = h_l w' \circ \varphi$, on obtient aussi

$$|w|_{1,I_l}^2 = \int_{x_l}^{x_{l+1}} |w'(x)|^2 dx = \frac{1}{h_l} \int_0^1 |\hat{w}'(\hat{x})|^2 d\hat{x} = \frac{1}{h_l} |\hat{w}|_{1,\hat{I}}^2. \quad (7.9)$$

Et de même, on obtiendrait

$$|v|_{2,I_l}^2 = \frac{1}{h_l^3} |\hat{v}|_{2,\hat{I}}^2. \quad (7.10)$$

Pour démontrer le lemme, il nous suffit donc de montrer qu'il existe une constante $C \in \mathbb{R}_+^*$ telle que pour tout $\hat{v} \in D([0, 1])$ et pour $m \in \{0, 1\}$

$$\left| \hat{v} - \hat{\Pi}^1 \hat{v} \right|_{m,\hat{I}} \leq C |\hat{v}|_{2,\hat{I}},$$

où sur l'élément de référence $\hat{I}$, on a

$$\hat{\Pi}^1 \hat{v} = (\hat{v}(1) - \hat{v}(0))\hat{x} + \hat{v}(0).$$

Remarquons tout d'abord que $\hat{v} - \hat{\Pi}^1(\hat{v}) \in H_0^1(]0, 1[)$ et par l'inégalité de Poincaré (2.31), on en déduit

$$\left| \hat{v} - \hat{\Pi}^1 \hat{v} \right|_{0,\hat{I}} \leq \left| \hat{v} - \hat{\Pi}^1 \hat{v} \right|_{1,\hat{I}}, \quad (7.11)$$

car la constante de Poincaré est égale à la longueur de l'intervalle d'après la proposition 2.2.23, ici 1. Puis on remarque que

$$(\hat{\Pi}^1 \hat{v})'(x) = \hat{v}(1) - \hat{v}(0) = \int_0^1 \hat{v}'(\hat{x}) d\hat{x},$$

donc $\hat{v}' - (\hat{\Pi}^1 \hat{v})' \in H^1(]0, 1[)$ est à moyenne nulle. Par l'inégalité de Poincaré-Wirtinger (2.34), il existe une constante $C \in \mathbb{R}_+^*$ telle que

$$\left| \hat{v}' - (\hat{\Pi}^1 \hat{v})' \right|_{0,\hat{I}} \leq C \left| \hat{v}'' - (\hat{\Pi}^1 \hat{v})'' \right|_{0,\hat{I}} = C |\hat{v}''|_{0,\hat{I}} \quad (7.12)$$

puisque $(\hat{\Pi}^1 \hat{v})'' = 0$. On obtient donc finalement, d'après (7.8), (7.12) puis (7.10),

$$|v - \Pi_h^1 v|_{0,I_l}^2 \leq h_l \left| \hat{v} - \hat{\Pi}^1 \hat{v} \right|_{0,\hat{I}}^2 \leq h_l |\hat{v}''|_{0,\hat{I}}^2 \leq h_l^4 |v|_{2,I_l}^2,$$

qui justifie (7.6). De même d'après (7.9), (7.11) puis (7.10), on a

$$|v - \Pi_h^1 v|_{1,I_l}^2 \leq \frac{1}{h_l} \left| \hat{v} - \hat{\Pi}^1 \hat{v} \right|_{1,\hat{I}}^2 \leq \frac{C^2}{h_l} |\hat{v}''|_{0,\hat{I}}^2 \leq C^2 h_l^2 |v|_{2,I_l}^2,$$

qui justifie (7.7). □

On déduit du lemme de Bramble-Hilbert le résultat suivant sur tout l'intervalle $]0, L[$.

Proposition 7.2.2 *Il existe une constante C indépendante de h telle que,*

pour tout $v \in H^2([0, L])$,

$$\|v - \Pi_h^1 v\|_{L^2([0, L])} \leq Ch^2 \|v''\|_{L^2([0, L])}, \quad (7.13)$$

et

$$\|v' - (\Pi_h^1 v)'\|_{L^2([0, L])} \leq Ch \|v''\|_{L^2([0, L])}. \quad (7.14)$$

Démonstration. À partir du lemme précédent de Bramble-Hilbert, on écrit simplement

$$\begin{aligned} \|v - \Pi_h^1 v\|_{L^2([0, L])}^2 &= \sum_{0 \leq l \leq N_i} |v - \Pi_h^1 v|_{0, I_l}^2 \\ &\leq \sum_{0 \leq l \leq N_i} h_l^4 |v|_{2, I_l}^2 \leq h^4 \|v''\|_{L^2([0, L])}^2 \end{aligned}$$

et de même

$$\begin{aligned} \|v' - (\Pi_h^1 v)'\|_{L^2([0, L])}^2 &= \sum_{0 \leq l \leq N_i} |v' - \Pi_h^1 v'|_{1, I_l}^2 \\ &\leq \sum_{0 \leq l \leq N_i} h_l^2 |v'|_{2, I_l}^2 \leq h^2 \|v''\|_{L^2([0, L])}^2. \end{aligned}$$

□

Remarque 7.2.1 Il existe aussi des preuves directes du résultat précédent, c'est-à-dire sans passer par l'élément de référence. Cependant, les ingrédients de la preuve sont en réalité similaires. Soit $v \in C^\infty([0, 1])$. Par construction, l'interpolée $\Pi_h^1 v$ est une fonction affine et, pour tout $x \in (x_l, x_{l+1})$, on a

$$\begin{aligned} v(x) - \Pi_h^1 v(x) &= v(x) - \left(v(x_l) + \frac{v(x_{l+1}) - v(x_l)}{x_{l+1} - x_l} (x - x_l) \right) \\ &= \int_{x_l}^x v'(t) dt - \frac{x - x_l}{x_{l+1} - x_l} \int_{x_l}^{x_{l+1}} v'(t) dt \\ &= (x - x_l) v'(x_l + \theta_x) - (x - x_l) v'(x_l + \theta_l) \\ &= (x - x_l) \int_{x_l + \theta_l}^{x_l + \theta_x} v''(t) dt, \end{aligned} \quad (7.15)$$

par la formule des accroissements finis avec $0 \leq \theta_x \leq x - x_l$ et $0 \leq \theta_l \leq h$. On en déduit, par l'inégalité de Cauchy-Schwarz, que

$$\begin{aligned} |v(x) - \Pi_h^1 v(x)|^2 &\leq h^2 \left(\int_{x_l}^{x_{l+1}} |v''(t)| dt \right)^2 \\ &\leq h^3 \int_{x_l}^{x_{l+1}} |v''(t)|^2 dt. \end{aligned} \quad (7.16)$$

En intégrant par rapport à x sur l'intervalle $[x_l, x_{l+1}]$, on obtient

$$\int_{x_l}^{x_{l+1}} |v(x) - \Pi_h^1 v(x)|^2 dx \leq h^4 \int_{x_l}^{x_{l+1}} |v''(t)|^2 dt,$$

ce qui, en sommant sur l , prouve (7.13). Par densité ce résultat est encore vrai pour tout $v \in H^2(]0, 1[)$. La démonstration de (7.14) est tout à fait similaire : pour $v \in C_c^\infty([0, 1])$ et $x \in (x_l, x_{l+1})$, on a

$$\begin{aligned} v'(x) - (\Pi_h^1 v)'(x) &= v'(x) - \frac{v(x_{l+1}) - v(x_l)}{h} \\ &= \frac{1}{h} \int_{x_l}^{x_{l+1}} [v'(x) - v'(t)] dt \\ &= \frac{1}{h} \int_{x_l}^{x_{l+1}} \int_t^x v''(y) dy dt. \end{aligned}$$

Élevant au carré cette inégalité, appliquant Cauchy-Schwarz deux fois et sommant en l on obtient (7.14), qui est aussi valide pour tout $v \in H^2(]0, 1[)$ par densité.

Convergence On peut maintenant étudier la convergence de la méthode des éléments finis $\mathbb{P}_1$.

Théorème 7.2.3 Soient $u \in H^1(]0, L[)$ et $u_h \in \mathcal{V}_h$ les solutions respectives des formulations variationnelles (7.1) et (7.2), définies pour le problème modèle (7.4). Alors, la méthode des éléments finis $\mathbb{P}_1$ converge, c'est-à-dire que

$$\lim_{h \rightarrow 0} \|u - u_h\|_{H^1(]0, L[)} = 0. \quad (7.17)$$

On suppose de plus que $f \in L^2(]0, L[)$, alors $u \in H^2(]0, L[)$ et il existe une constante C indépendante de h telle que

$$\|u - u_h\|_{H^1(]0, L[)} \leq Ch \|u''\|_{L^2(]0, L[)} \leq Ch \|f\|_{L^2(]0, L[)}. \quad (7.18)$$

Démonstration. Soient $\mathcal{W} = H^2(]0, L[)$ et

$$r_h : \mathcal{W} \ni v \mapsto \Pi_h^1 v \in \mathcal{V}_h.$$

D'après la proposition 7.2.2, si $v \in H^2(]0, L[)$ il existe une constante $C \in \mathbb{R}_+^*$ indépendante de h telle que

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

donc en particulier pour tout $v \in H^2(]0, L[)$ dense dans $H^1(]0, L[)$

$$\lim_{h \rightarrow 0} \|v - r_h v\|_{H^1(]0, L[)} = 0.$$

Ainsi d'après le théorème 3.6.4, on a bien la convergence

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

Si $f \in L^2(]0, L[)$, alors on déduit immédiatement de $-u'' = f - u$ que $u \in H^2(]0, L[)$ et

$$\|u''\|_{L^2(]0, L[)} \leq \|f\|_{L^2(]0, L[)} + \|u\|_{L^2(]0, L[)}.$$

De plus, par la formulation variationnelle (7.4), on a

$$\|u\|_{L^2(]0,L])}^2 \leq \|u\|_{H^1(]0,L])}^2 \leq \|f\|_{L^2(]0,L])} \|u\|_{L^2(]0,L])}.$$

Ces deux dernières inégalités nous permettent de déduire que

$$\|u''\|_{L^2(]0,L])} \leq 2\|f\|_{L^2(]0,L])}.$$

Comme $u \in H^2(]0, L])$, on a d'après le lemme 3.6.2 de Cea

$$\begin{aligned} \|u - u_h\|_{H^1(]0,L])} &\leq C \inf_{v_h \in \mathcal{V}_h} \|u - v_h\|_{H^1(]0,L])} \\ &\leq C \|u - \Pi_h^1 u\|_{H^1(]0,L])} \leq Ch \|u''\|_{L^2(]0,L])} \leq 2Ch \|f\|_{L^2(]0,L])}. \end{aligned}$$

□

7.2.2 Convergence et estimation d'erreur en dimension $d \geq 2$

Problème modèle. De nouveau, nous considérons un problème modèle, mais les résultats que nous allons présenter sont en fait très généraux. Soient $\Omega \subset \mathbb{R}^d$ un ouvert de frontière polyédrique et $f \in L^2(\Omega)$, on considère le problème modèle suivant : On cherche la solution $u \in \mathcal{V} = H^1(\Omega)$ de

$$\begin{aligned} \forall v \in H^1(\Omega), \quad \int_{\Omega} \left[u(\underline{x})v(\underline{x}) + \underline{\nabla} u(\underline{x}) \cdot \underline{\nabla} v(\underline{x}) \right] d\Omega \\ = \int_{\Omega} f(\underline{x})v(\underline{x}) d\Omega. \end{aligned} \quad (7.19)$$

de formulation forte correspondante

$$\begin{cases} -\Delta u + u = f & \text{dans } \Omega, \\ \frac{\partial u}{\partial n} = 0 & \text{sur } \partial\Omega. \end{cases} \quad (7.20)$$

La formulation variationnelle (7.4) est approchée par une méthode de Galerkin de type éléments finis de Lagrange sur des maillages triangulaires. On considère l'espace de discrétisation

$$\mathcal{V}_h = \mathcal{V}_h^{\mathbb{P}_k} = \left\{ v_h \in C^0(\bar{\Omega}) \text{ tel que } v_h|_{T_i} \in \mathbb{P}_k \text{ pour tout } T_i \in \mathcal{T}_h \right\},$$

tel qu'introduit au paragraphe 6.3.1. On rappelle que l'espace $\mathcal{V}_h$ est inclus dans les fonctions continues par ailleurs C^1 par morceaux sur une partition finie du domaine. Donc par le théorème 2.2.2, on a bien que l'espace $\mathcal{V}_h$ est un sous-espace de $H^1(\Omega)$.

Maillage et raffinement Il nous faut commencer par des hypothèses sur les maillages triangulaires considérés. On introduit deux paramètres géométriques caractérisant les éléments du maillage : pour tout d -simplexe T , on note $h(T)$ le rayon de la plus petite boule contenant T et $\rho(T)$, le

rayon de la plus grande boule contenue dans T ,

$$h(T) = \inf_{T \subset B(\underline{x}, r)} r, \quad \rho(T) = \sup_{B(\underline{x}, r) \subset T} r.$$

Bien sûr, on a toujours $h(T)/\rho(T) > 1$. Ce rapport est d'autant plus grand que T est « aplati » : il mesure en quelque sorte la tendance à la dégénérescence de T .⁵

5: Que ce soit en pratique, comme en théorie, il faudra donc éviter d'utiliser des d -simplexes T trop aplatis.

Dans toute la suite, on considèrera une séquence dénombrable de maillages $(\mathcal{T}_{h_n})_{h_n \in \mathbb{N}}$, paramétrée par

$$h_n = \max_{T_l \in \mathcal{T}_{h_n}} h(T_l), \quad (7.21)$$

et qui, sauf mention contraire, seront tels que $\overline{\Omega} = \cup \{\overline{T_l}, T_l \in \mathcal{T}_{h_n}\}$. Le paramètre h_n a vocation à tendre vers 0, ce qui implique des maillages d'autant plus raffinés que h_n est proche de 0. Il est d'usage pour simplifier l'écriture de noter génériquement cette suite $(\mathcal{T}_h)_{h>0}$.

Définition 7.2.3 Soit $(\mathcal{T}_h)_{h>0}$ une suite de maillages de Ω . On dit qu'il s'agit d'une suite de maillages réguliers si

1. la suite $h = \max_{T_l \in \mathcal{T}_h} h(T_l)$ tend vers 0,
2. il existe une constante C telle que, pour tout $h > 0$ et tout $T \in \mathcal{T}_h$,

$$\frac{h(T)}{\rho(T)} \leq C. \quad (7.22)$$

Remarque 7.2.2 Parfois, on définit h comme étant $\max_{T_l \in \mathcal{T}_h} \text{diam}(T_l)$ où

$$\text{diam}(T) = \max_{(\underline{x}, \underline{y}) \in T^2} |\underline{x} - \underline{y}|.$$

On se convaincra sans peine que si ces définitions diffèrent, elles sont équivalentes au sens où il existe C_1 et C_2 telle que

$$C_1 \max_{T_l \in \mathcal{T}_h} \text{diam}(T_l) \leq h \leq C_2 \max_{T_l \in \mathcal{T}_h} \text{diam}(T_l).$$

Remarque 7.2.3 En dimension $d = 2$ la condition (7.22) est équivalente à la condition suivante sur les angles du triangle T : il existe un angle minimum $\theta_0 > 0$ qui minore (uniformément en h) tous les angles de tout $T \in \mathcal{T}_h$.

Opérateur d'interpolation Commençons par définir les opérateurs d'interpolation sur les espaces $\mathcal{V}_h^{\mathbb{P}_k}$ des éléments finis $\mathbb{P}_k$ introduit en (6.33).

Définition 7.2.4 On appelle opérateur d'interpolation $\mathbb{P}_k$ l'application linéaire $\Pi_h^k \in \mathcal{L}(C^0(\overline{\Omega}), \mathcal{V}_h^{\mathbb{P}_k})$ définie par

$$\forall v \in C^0(\overline{\Omega}), \quad (\Pi_h^k v)(\underline{x}) = \sum_{j=1}^N v(\underline{S}_j) \phi_j(\underline{x}),$$

où la famille $\{\phi_j, j \in \llbracket 1, N \rrbracket\}$ est définie à la proposition 6.3.1, et $\underline{S}_j$ sont les nœuds du maillage.

Proposition 7.2.4 L'opérateur Π_h^k est un opérateur de projection.

Démonstration. D'après le lemme 6.3.3, $\Pi_h^k v$ est continu, notamment au passage d'un élément à l'autre. Par ailleurs, par construction de la base d'éléments finis de Lagrange, on a

$$(\Pi_h^k v)(\underline{S}_i) = \sum_{j=1}^N v(\underline{S}_j) \phi_j(\underline{S}_i) = \sum_{j=1}^N \delta_i^j v(\underline{S}_j) = v(\underline{S}_i),$$

d'où le résultat. $\square$

La propriété précédente était déjà vraie en 1d. Cependant, quelque chose a changé en dimension $d \geq 2$. L'opérateur d'interpolation ne s'applique pas aux fonctions de $H^1(\Omega)$ qui ne sont en général pas continues⁶.

6: voir le théorème 2.2.10

Comme en dimension 1, on introduit différentes semi-normes. Rappelons tout d'abord que pour $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$ on note $|\alpha| = \sum_i \alpha_i$, et pour tout $v \in H^m$ avec $|\alpha| = m$ on a

$$\partial^\alpha v := \partial_{x_1^{\alpha_1} \dots x_d^{\alpha_d}} v := \frac{\partial^{|\alpha|} v}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}.$$

Définition 7.2.5 (Semi-normes locales) Sur chaque élément T_l de $\mathcal{T}_h$, on définit les semi-normes

$$|v|_{m, T_l} = \sum_{\substack{\alpha \in \mathbb{N}^d \\ |\alpha|=m}} \left(\int_{T_l} \left| \partial_{x_1^{\alpha_1} \dots x_d^{\alpha_d}} v \right|^2 d\Omega \right)^{\frac{1}{2}}.$$

Comme en dimension 1, nous allons nous ramener à l'élément de référence. On introduit la transformation affine entre l'élément de référence et l'élément T_l , soit

$$\hat{\phi}_l : \hat{T} \ni \underline{x} \mapsto \hat{\phi}_l(\underline{x}) = \underline{F}_l \cdot \underline{x} + \underline{S}_{l,1} \in T_l.$$

C'est une application affine où $\nabla \hat{\phi}_l(\underline{x}) = \underline{F}_l$ ne dépend pas de $\underline{x}$ et $\underline{S}_{l,1}$ est le premier sommet du triangle T_l . De plus, si le triangle n'est pas dégénéré, on contrôle la norme de $\underline{F}_l$ comme on va le montrer dans le lemme qui suit.⁷

Lemme 7.2.5 Pour tout $T_l \in \mathcal{T}_h$ supposé non dégénéré, on note $h_l = h(T_l) \neq 0$ et $\rho_l = \rho(T_l) \neq 0$. On a

$$|\underline{F}_l|_2 \leq \frac{h_l}{\hat{\rho}} \quad \text{et} \quad |\underline{F}_l^{-1}|_2 \leq \frac{\hat{h}}{\rho_l}, \quad (7.24)$$

7: En calcul tensoriel, on rappelle qu'il est d'usage de noter pour tout tenseur $\underline{F}$

$$|\underline{F}|^2 = \underline{F} : \underline{F} = \sum_{1 \leq i, j \leq d} F_{ij}^2.$$

qui vérifie notamment $|\underline{F} \cdot \underline{G}| \leq |\underline{F}| |\underline{G}|$ pour tous les tenseurs $\underline{F}$ et $\underline{G}$. Mais on peut définir une autre norme subordonnée au produit scalaire dans $\mathbb{R}^d$, soit

$$|\underline{F}|_2 = \sup_{\underline{x} \neq 0} \frac{|\underline{F} \cdot \underline{x}|}{|\underline{x}|}$$

qui vérifie par construction $|\underline{F} \cdot \underline{x}| \leq |\underline{F}|_2 |\underline{x}|$. Dans $\mathbb{R}^d$, ces deux normes sont équivalentes et on montre facilement dans [8, Theorem 1.4-4] que

où $\hat{h}$ (respectivement $\hat{\rho}$) est le rayon de la plus grande boule contenant (de la plus petite boule incluse dans) le triangle de référence.

Démonstration. Pour tout $\underline{x} \in \mathbb{R}^d$ tel que $|\underline{x}| = 2\hat{\rho}$, il existe deux points $\underline{M}_1$ et $\underline{M}_2$ situés sur le cercle inscrit au triangle de référence tels que $\underline{x} = \underline{M}_2 - \underline{M}_1$. On a alors $\underline{F}_I \cdot \underline{x} = \underline{F}_I \cdot \underline{M}_2 - \underline{F}_I \cdot \underline{M}_1$. Comme $\underline{F}_I \cdot \underline{M}_1 \in T_I$ et $\underline{F}_I \cdot \underline{M}_2 \in T_I$, on a alors $|\underline{F}_I \cdot \underline{M}_2 - \underline{F}_I \cdot \underline{M}_1| \leq 2h_I$ d'où le premier résultat.

Réciproquement, on raisonne de même à partir de l'application réciproque $\hat{\varphi}_I^{-1} : T_I \ni \underline{x} \mapsto \underline{F}_I^{-1}(\underline{x} - \underline{S}_{I,1})^* \in \hat{T}$. $\square$

Le lemme précédent va nous être utile pour ramener les calculs de normes locales sur l'élément de référence.

Lemme 7.2.6 Pour tout triangle $T_I \in \mathcal{T}_h$ supposé non dégénéré, on note $h_I = h(T_I) \neq 0$ et $\rho_I = \rho(T_I) \neq 0$. Soient $m \in \mathbb{N}$, $v \in H^m(T_I)$ et $\hat{v} = v \circ \hat{\varphi}_I \in H^m(\hat{T})$, il existe des constantes C_m et C'_m telles que

$$|v|_{m,T_I} \leq C_m \rho_I^{-m} |\det \underline{F}_I|^{\frac{1}{2}} |\hat{v}|_{m,\hat{T}}, \quad (7.25)$$

$$|\hat{v}|_{m,\hat{T}} \leq C'_m h_I^m |\det \underline{F}_I|^{-\frac{1}{2}} |v|_{m,T_I}. \quad (7.26)$$

Démonstration. On rappelle que

$$\hat{\varphi}_I : \hat{T} \ni \underline{x} \mapsto \underline{\varphi}(\underline{x}) = \underline{F}_I \underline{x} + \underline{S}_{I,1} \in T_I.$$

est la transformation affine entre l'élément de référence et l'élément T_I . Nous nous limiterons à la preuve de (7.25) et au cas $m = 0, 1$ et 2 . Les cas $m > 2$ sont similaires au cas $m = 2$. De plus (7.26) s'obtient comme (7.25) en utilisant la transformation réciproque à la place de la transformation.

Cas $m = 0$. On a, pour tout $v \in L^2(T_I)$,

$$|v|_{0,T_I}^2 \leq \int_{\hat{T}} |\hat{v}|^2 |\det(\underline{F}_I)| \, d\Omega$$

où $\hat{v} = v \circ \varphi$. Comme $\underline{F}_I$ est constant par élément, on a

$$|v|_{0,T_I} \leq |\det(\underline{F}_I)|^{\frac{1}{2}} |\hat{v}|_{0,\hat{T}}.$$

Cas $m = 1$. Puisque $\underline{\nabla} \hat{v} = \underline{\nabla} \hat{\varphi}_I^\top \cdot \underline{\nabla} (v \circ \hat{\varphi}_I) = \underline{F}_I^\top \cdot \underline{\nabla} (v \circ \hat{\varphi}_I)$, alors on obtient

$$\begin{aligned} |v|_{1,T_I}^2 &\leq \int_{\hat{T}} |\underline{F}_I^{-\top} \cdot \underline{\nabla} \hat{v}|^2 |\det(\underline{F}_I)| \, d\Omega, \\ &\leq |\underline{F}_I^{-1}|_2^2 |\det(\underline{F}_I)| \int_{\hat{T}} |\underline{\nabla} \hat{v}|^2 \, d\Omega, \end{aligned}$$

dont on déduit d'après le lemme 7.2.5 que

$$|v|_{1,T_I}^2 \leq \frac{\hat{h}^2}{\hat{\rho}_I^2} |\det(\underline{F}_I)| \int_{\hat{T}} |\underline{\nabla} \hat{v}|^2 \, d\Omega.$$

Cas $m = 2$. On raisonne dérivée partielle par dérivée partielle. On a par équivalence des normes, l'existence d'une constante C telle que

$$\forall 1 \leq i, j \leq d, \quad |\partial_{x_i x_j}^2 v(\underline{x})| \leq |D^2 v(\underline{x})|_2 \leq \max_{\substack{\alpha \in \mathbb{N}^d \\ |\alpha|=m}} |\partial^\alpha v(\underline{x})|$$

et par définition de la dérivée seconde comme succession de deux différentielles, on a

$$D^2 v(\underline{x})(y, z) = D^2 \hat{v}(\hat{\underline{x}})(\underline{F}_l^{-1} \cdot y, \underline{F}_l^{-1} \cdot z),$$

donc

$$\forall 1 \leq i, j \leq d, \quad |\partial_{x_i x_j}^2 v(\underline{x})| \leq |D^2 \hat{v}(\hat{\underline{x}})|_2 |\underline{F}_l^{-1}|_2^2.$$

On obtient finalement

$$\begin{aligned} |v|_{2, T_l}^2 &= \sum_{\substack{\alpha \in \mathbb{N}^d \\ |\alpha|=m}} \left(\int_{T_l} \left| \partial_{x_1^{\alpha_1} \dots x_d^{\alpha_d}}^{|\alpha|} v \right|^2 d\Omega \right) \\ &\leq d |\underline{F}_l^{-1}|_2^4 \int_{T_l} |D^2 \hat{v} \circ \hat{\phi}_l^{-1}(\underline{x})|^2 d\Omega \\ &\leq d |\det(\underline{F}_l)| \frac{\hat{h}^4}{\rho_l^4} \int_{\hat{T}} |D^2 \hat{v}|^2 d\Omega \\ &\leq C d |\det(\underline{F}_l)| \frac{\hat{h}^4}{\rho_l^4} \max_{\substack{\alpha \in \mathbb{N}^d \\ |\alpha|=m}} \left(\int_{T_l} \left| \partial_{x_1^{\alpha_1} \dots x_d^{\alpha_d}}^{|\alpha|} \hat{v} \right|^2 d\Omega \right). \end{aligned}$$

□

Enfin il nous faut un dernier résultat nous permettant de comparer des semi-normes à des normes à l'image des inégalités de Poincaré utilisées dans la preuve du lemme 7.2.1 de Bramble-Hilbert⁸ en dimension 1.

8: et dont la démonstration va reposer sur le même type de raisonnement que la preuve des inégalités de Poincaré vue au paragraphe 2.2.5

Théorème 7.2.7 Soit $\hat{\Pi}$ un projecteur de $H^{k+1}(\hat{T})$ sur l'espace des $\mathbb{P}_k(\hat{T})$. Il existe une constante C telle que pour tout $\hat{v} \in H^{k+1}(\hat{T})$ on a

$$\|\hat{v} - \hat{\Pi}\hat{v}\|_{H^k(\hat{T})} \leq C |\hat{v}|_{k+1, \hat{T}}. \quad (7.27)$$

Démonstration. Nous procédons par l'absurde. Supposons en effet qu'il existe une suite $(\hat{v}_n)_{n \in \mathbb{N}} \in [H^{k+1}(\hat{T})]^\mathbb{N}$ telle que $\hat{w}_n := \hat{v}_n - \hat{\Pi}\hat{v}_n$ vérifie

$$\|\hat{w}_n\|_{H^k(\hat{T})} = 1 \quad \text{et} \quad \lim_{n \rightarrow \infty} |\hat{v}_n|_{k+1, \hat{T}} = 0.$$

Commençons par remarquer que $\hat{\Pi}\hat{v}_n$ est un polynôme de degré k donc

$$\lim_{n \rightarrow \infty} |\hat{v}_n|_{k+1, \hat{T}} = 0 \quad \Rightarrow \quad \lim_{n \rightarrow \infty} |\hat{w}_n|_{k+1, \hat{T}} = 0. \quad (7.28)$$

On en déduit qu'il existe une constante C telle que pour tout n

$$\|\hat{w}_n\|_{H^{k+1}(\hat{T})}^2 = \sum_{0 \leq q \leq k} \|\hat{w}_n\|_{H^q(\hat{T})}^2 + |\hat{w}_n|_{H^{k+1}(\hat{T})}^2 \leq C$$

et ainsi, par le théorème 2.2.24 de Rellich, et à une extraction près, $(\hat{w}_n)_{n \in \mathbb{N}}$ converge dans $H^k(\hat{T})$. Soit $\hat{w}$ la limite. Soit désormais $\phi \in D(\hat{T})$, on a

donc pour tout $\alpha \in \mathbb{N}^d$ avec $|\alpha| = k$,

$$\left| \left(\partial_{x_1^{\alpha_1} \dots x_d^{\alpha_d}} \hat{w}, \partial_{x_i} \phi \right)_{L^2(\hat{T})} \right| = \lim_{n \rightarrow \infty} \left| \left(\partial_{x_1^{\alpha_1} \dots x_d^{\alpha_d}} \hat{w}_n, \partial_{x_i} \phi \right)_{L^2(\hat{T})} \right|.$$

Or si on note $\tilde{\alpha} \in \mathbb{N}^d$ tel que $\tilde{\alpha}_j = \alpha_j + \delta_{ij}$, on a $|\tilde{\alpha}| = k + 1$ et

$$\begin{aligned} \left| \left(\partial_{x_1^{\alpha_1} \dots x_d^{\alpha_d}} \hat{w}_n, \partial_{x_i} \phi \right)_{L^2(\hat{T})} \right| &= \left| \left(\partial_{x_1^{\tilde{\alpha}_1} \dots x_d^{\tilde{\alpha}_d}} \hat{w}_n, \phi \right)_{L^2(\hat{T})} \right| \\ &\leq \|\hat{w}_n\|_{k+1, \hat{T}} \|\phi\|_{L^2(\hat{T})}. \end{aligned}$$

Donc d'après (7.28)

$$\left(\partial_{x_1^{\alpha_1} \dots x_d^{\alpha_d}} \hat{w}, \partial_{x_i} \phi \right)_{L^2(\hat{T})} = 0,$$

et ainsi $\hat{w}$ admet toutes ses dérivées faibles d'ordre $k + 1$ et pour tout $\tilde{\alpha} \in \mathbb{N}^{k+1}$ avec $|\tilde{\alpha}| = k + 1$

$$\partial_{x_1^{\tilde{\alpha}_1} \dots x_d^{\tilde{\alpha}_d}} \hat{w} = 0.$$

Ainsi $\hat{w} \in H^{k+1}(\hat{T})$ et comme toutes ses dérivées d'ordre k sont nulles, d'après la proposition 2.1.7, c'est un polynôme multivarié de degré k . On en déduit donc que $\hat{\Pi} \hat{w} = \hat{w}$.

Or pour tout $n \in \mathbb{N}$, $\hat{w}_n \in \text{Im}(\text{Id} - \hat{\Pi}) = \ker(\hat{\Pi})$ fermé, car $\hat{\Pi}$ est une application continue dans $H^{k+1}(\hat{T})$. Donc la suite $(\hat{w}_n)_{n \in \mathbb{N}}$ converge dans $\ker(\hat{\Pi})$, ce qui implique $\hat{\Pi} \hat{w} = 0$.

D'après ce qui précède, $\hat{w} = 0$, ce qui contredit le fait que par ailleurs

$$\|\hat{w}\|_{H^k(\hat{T})} = \lim_{n \rightarrow \infty} \|\hat{w}_n\|_{H^k(\hat{T})} = 1.$$

□

Nous pouvons désormais démontrer le théorème de Bramble-Hilbert pour contrôler localement l'erreur d'interpolation.

Théorème 7.2.8 (Bramble-Hilbert) *Pour tout triangle $T_l \in \mathcal{T}_h$ supposé non dégénéré, on note $h_l = h(T_l) \neq 0$ et $\rho_l = \rho(T_l) \neq 0$. Il existe une constante C indépendante de T_l telle que pour tout $v \in C^0(\overline{T_l}) \cap H^{k+1}(T_l)$ on a⁹*

$$\forall m \in \llbracket 0, k \rrbracket, \quad |v - \Pi_h^k v|_{m, T_l} \leq C \frac{h_l^{k+1}}{\rho_l^m} |v|_{k+1, T_l}. \quad (7.29)$$

9: Notons que d'après le théorème 2.2.10, dès que $v \in H^{k+1}(T_l)$ avec $k + 1 > d/2$ alors $v \in C^0(\overline{T_l})$.

Démonstration. Soit $m \in \llbracket 0, k \rrbracket$, commençons par appliquer le lemme 7.2.6 à $w = v - \Pi_h^k v$. Puisque $\hat{\phi}_l$ est affine, on a alors $\hat{w} = w \circ \hat{\phi}_l = \hat{v} - \hat{\Pi}_h^k \hat{v}$ et ainsi (7.25) donne l'existence d'une constante C_1 telle que

$$|v - \Pi_h^k v|_{m, T_l} \leq C_1 \rho_l^{-m} |\det \underline{\underline{F}}_l|^{\frac{1}{2}} |\hat{v} - \hat{\Pi}_h^k \hat{v}|_{m, \hat{T}}.$$

Par application du théorème 7.2.7 sur l'élément de référence $\hat{T}$ pour le terme de droite, il existe alors une constante C_2 telle que

$$|v - \Pi_h^k v|_{m, T_l} \leq C_2 \rho_l^{-m} |\det \underline{F}_l|^{\frac{1}{2}} |\hat{v} - \hat{\Pi}_h^k \hat{v}|_{k+1, \hat{T}}.$$

Puis on revient sur l'élément courant T_l à droite en utilisant (7.26). Il existe donc une constante C_3 telle que

$$|v - \Pi_h^k v|_{m, T_l} \leq C_3 \frac{h_l^{k+1}}{\rho_l^m} |v - \Pi_h^k v|_{k+1, T_l},$$

ce qu'il fallait démontrer. $\square$

Nous pouvons maintenant déduire de ce théorème le résultat pour $m = 1$ suivant.

Théorème 7.2.9 Soit $(\mathcal{T}_h)_{h>0}$ une suite de maillages réguliers de Ω . Soit

$$v \in C^0(\overline{\Omega}) \quad \text{avec} \quad \forall T_l \in \mathcal{T}_h, \quad v \in H^{k+1}(T_l)$$

alors

$$\|v - \Pi_h^k v\|_{H^1(\Omega)} \leq C h^k \left(\sum_{1 \leq l \leq N_e} |v|_{H^{k+1}(T_l)}^2 \right)^{\frac{1}{2}}, \quad (7.30)$$

où C est une constante indépendante de h et de v .

Démonstration. On note $h_l = h(T_l) \neq 0$ et $\rho_l = \rho(T_l) \neq 0$. D'après le théorème 7.2.8 de Bramble-Hilbert, il existe $C \in \mathbb{R}_+^*$ telle que

$$|v - \Pi_h^k v|_{0, T_l}^2 \leq C h_l^{2k+2} |v|_{k+1, T_l}^2,$$

donc

$$\int_{\Omega} (v - \Pi_h^k v)^2 \, d\Omega = \sum_{l=1}^{N_e} |v - \Pi_h^k v|_{0, T_l}^2 \leq C \sum_{l=1}^{N_e} h_l^{2k+2} |v|_{k+1, T_l}^2.$$

De même,

$$|v - \Pi_h^k v|_{1, T_l}^2 \leq C \frac{h_l^{2k+2}}{\rho_l^2} |v|_{k+1, T_l}^2$$

donc

$$\int_{\Omega} |\nabla (v - \Pi_h^k v)|^2 \, d\Omega = \sum_{l=1}^{N_e} |v - \Pi_h^k v|_{1, T_l}^2 \leq C \sum_{l=1}^{N_e} \frac{h_l^{2k+2}}{\rho_l^2} |v|_{k+1, T_l}^2.$$

Comme $(\mathcal{T}_h)_{h>0}$ est une suite de maillages réguliers de Ω , il existe, d'après la définition 7.2.3, $\delta \in \mathbb{R}_+^*$ tel que

$$\forall \mathcal{T}_h, \quad \forall T_l \in \mathcal{T}_h, \quad \frac{h(T_l)}{\rho(T_l)} \leq \delta,$$

on en déduit donc

$$\|v - \Pi_h^k v\|_{H^1(\Omega)}^2 \leq C(\delta^2 + h^2) h^{2k} \sum_{l=1}^{N_e} |v|_{k+1, T_l}^2,$$

qui conclut la preuve. $\square$

On déduit facilement de ce résultat et du théorème 2.2.10 le corollaire suivant.

Corollaire 7.2.10 Soit $(\mathcal{T}_h)_{h>0}$ une suite de maillages réguliers de Ω . Si $v \in H^{k+1}(\Omega)$ avec $k+1 > d/2$ alors

$$\|v - \Pi_h^k v\|_{H^1(\Omega)} \leq Ch^k |v|_{H^{k+1}(\Omega)}, \quad (7.31)$$

où C est une constante indépendante de h et de v .

Étude de convergence. Nous pouvons maintenant énoncer le résultat principal de ce paragraphe qui affirme la convergence de la méthode des éléments finis $\mathbb{P}_k$. Nous donnerons ensuite celui qui donne une estimation de la vitesse de convergence lorsque la solution est régulière.

Théorème 7.2.11 Soit $(\mathcal{T}_h)_{h>0}$ une suite de maillages réguliers de Ω . Soit $u \in H^1(\Omega)$, la solution du problème variationnel (7.1) associée au problème modèle (7.19), et $u_h \in \mathcal{V}_h$, celle de son approximation interne (7.2) par la méthode des éléments finis $\mathbb{P}_k$. Alors la méthode converge, c'est-à-dire que

$$\lim_{h \rightarrow 0} \|u - u_h\|_{H^1(\Omega)} = 0. \quad (7.32)$$

Démonstration. La preuve est en tout point similaire au cas 1d.

Soit $\mathcal{W} = D(\overline{\Omega})$ et

$$r_h : \mathcal{W} \ni v \mapsto \Pi_h^k v \in \mathcal{V}_h. \quad (7.33)$$

D'après le théorème 7.2.8 de Bramble-Hilbert,

$$\lim_{h \rightarrow 0} \|v - r_h v\|_{H^1(\Omega)} = 0.$$

Donc d'après le théorème 3.6.4, on a bien la convergence

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

$\square$

Passons à la vitesse de convergence.

Théorème 7.2.12 Soit $(\mathcal{T}_h)_{h>0}$ une suite de maillages réguliers de Ω . Soit $u \in H^1(\Omega)$, la solution du problème variationnel (7.1) associée au problème modèle (7.19), et $u_h \in \mathcal{V}_h^{\mathbb{P}_k}$, celle de son approximation interne (7.2) par la méthode des éléments finis $\mathbb{P}_k$. De plus, si

$$u \in C^0(\overline{\Omega}) \quad \text{et} \quad \forall T_l \in \mathcal{T}_h, \quad u \in H^{k+1}(T_l)$$

alors on a la vitesse de convergence

$$\|u - u_h\|_{H^1(\Omega)} \leq C h^k \left(\sum_{l=1}^{N_e} |u|_{H^{k+1}(T_l)}^2 \right)^{\frac{1}{2}}, \quad (7.34)$$

où C est une constante indépendante de h et de u .

Démonstration. On sait, d'après le lemme 3.6.2 de Céa, qu'il existe $C \in \mathbb{R}_+^*$ telle que

$$\|u - u_h\|_{H^1(\Omega)} \leq C \inf_{v_h \in \mathcal{V}_h} \|u - v_h\|_{H^1(\Omega)} \leq C \|u - \Pi_h^k u\|_{H^1(\Omega)},$$

où $\Pi_h^k u$ est bien défini puisque $u \in C^0(\overline{\Omega})$. Il suffit ensuite d'utiliser le théorème 7.2.9. $\square$

Il est essentiel de garder à l'esprit que le théorème demande une certaine régularité sur la solution, qu'il faut pouvoir démontrer par ailleurs, en fonction des données du problème, c'est-à-dire de la régularité du domaine, du terme source et des conditions aux limites. Évidemment, si la solution $u \in H^{k+1}(\Omega)$ avec $k+1 > d/2$ alors d'après le corollaire 7.2.10

$$\|u - u_h\|_{H^1(\Omega)} \leq C h^k \|u\|_{H^{k+1}(\Omega)}, \quad (7.35)$$

qui est l'estimation la plus souvent rencontrée.

Remarque 7.2.4 On notera tout de même qu'en toute généralité, l'hypothèse sur la régularité est locale. Cela permet de considérer des problèmes inhomogènes avec des coefficients réguliers par morceaux sur N_D domaines :

$$\sum_i \int_{\Omega_i} \left[\kappa_{1 \leq i \leq N_D} \nabla u \cdot \nabla v + uv \right] d\Omega = \int_{\Omega} f v d\Omega$$

où $\Omega = \bigcup_{1 \leq i \leq N_D} \Omega_i$, pour tout i , Ω_i est polygonal convexe, $\kappa_i \in C^1(\Omega_i)$. Dans ce cas, si $f \in L^2(\Omega)$, alors on peut avoir dans certains cas¹⁰ $u \in H^2(\Omega_i)$ mais pour autant $u \notin H^2(\Omega)$. Si le maillage est construit de telle façon que pour chaque triangle T_l , il existe $i \in \llbracket 1, N_D \rrbracket$ tel que $T_l \subset \Omega_i$, alors le théorème 7.36, s'applique pour $k = 1$.

10: cela dépendra de l'agencement des sous-domaines

Remarque 7.2.5 Le théorème 7.2.11 s'applique en fait à toute méthode des éléments finis de type Lagrange, notamment, les éléments finis quadrangulaires du paragraphe 6.3.2.

Ce résultat peut être affiné dans le cas où la solution n'est pas aussi régulière, ce qui va être le cas pour des domaines polygonaux non convexes.

Théorème 7.2.13 Soit $(\mathcal{T}_h)_{h>0}$ une suite de maillages réguliers de Ω . Soit $u \in H^1(\Omega)$, la solution du problème variationnel (7.1) associée au problème modèle (7.19), et $u_h \in \mathcal{V}_h$, celle de son approximation interne (7.2) par la

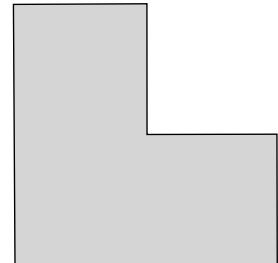


FIGURE 7.2 – Une configuration où typi-

méthode des éléments finis $\mathbb{P}_k$. Si

$$u \in C^0(\overline{\Omega}) \quad \text{et} \quad \forall T_l \in \mathcal{T}_h, \quad u \in H^{p+1}(T_l) \quad \text{pour } p \in \mathbb{N}$$

alors on a pour $\beta = \min(p, k)$ la vitesse de convergence

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch^\beta \left(\sum_{l=1}^{N_e} |u|_{H^{\beta+1}(T_l)}^2 \right)^{\frac{1}{2}}, \quad (7.36)$$

où C est une constante indépendante de h et de u .

Démonstration. Le cas $p = k$ découle du théorème 7.2.11 donc on considère $p \leq k$. La preuve reste tout à fait similaire à celle du théorème 7.2.11, une fois remarqué que

$$\inf_{v_h \in \mathcal{V}_h^k} \|u - v_h\|_{H^1(\Omega)} \leq \|u - \Pi_h^p u\|_{H^1(\Omega)}.$$

□

Remarque 7.2.6 Dans le cas d'un coin, pour une erreur donnée, on peut utiliser des maillages adaptés pour contrer le manque de régularité imposé par le coin. En effet, on diminuera la taille des éléments en s'approchant du coin.

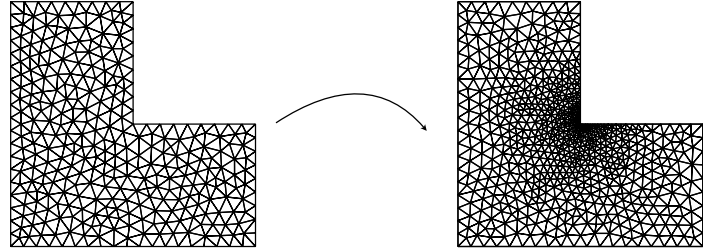


FIGURE 7.3 – Maillage adaptatif près d'un coin

7.2.3 Prise en compte de la régularité des données

Nous pouvons passer dans ce paragraphe à l'étude de la vitesse de convergence en fonction du domaine. On considère ici toujours le problème modèle (7.19) mais le résultat se généralise à de nombreuses configurations.

Théorème 7.2.14 Soit Ω un ouvert polygonal et $(\mathcal{T}_h)_{h>0}$ une suite de maillages réguliers de Ω . Soit $u \in H^1(\Omega)$ la solution du problème variationnel (7.19), et $u_h \in \mathcal{V}_h$, celle de son approximation interne (7.2) par la méthode des éléments finis $\mathbb{P}_1$ pour $d \in \{2, 3\}$. On suppose par ailleurs que $f \in L^2(\Omega)$. Si Ω est convexe alors

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch \|f\|_{L^2(\Omega)}, \quad (7.37)$$

où C est une constante indépendante de h et de u .

Démonstration. On combine ici les résultats sur l'existence de solutions du problème continu et les résultats d'approximation. D'après le théorème 4.4.7, nous savons que si Ω est convexe alors la solution u du problème (7.19) est dans $H^2(\Omega)$ et il existe une constante $C \in \mathbb{R}_+^*$ telle que

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

Dans ce cas, par application du théorème 7.2.12 sur les éléments finis $\mathbb{P}_1$, on sait qu'il existe une constante $C' \in \mathbb{R}_+^*$ telle que

$$\|u - u_h\|_{H^1(\Omega)} \leq C'h\|u\|_{H^2(\Omega)} \leq CC'h\|f\|_{L^2(\Omega)}$$

ce qu'il fallait démontrer. $\square$

Remarque 7.2.7 Si Ω n'est pas convexe, comme on peut démontrer (voir la remarque 4.4.2) en toute généralité qu'il existe $0 < \bar{s} < 1$ tel que tout $0 \leq s \leq \bar{s}$, $u \in H^{s+1}(\Omega)$ alors on obtient la vitesse de convergence

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch^s\|f\|_{L^2}. \quad (7.38)$$

Pour aller un cran plus loin dans l'analyse numérique de la méthode, nous avons vu que dans un certain nombre de situations nous devons approcher la forme linéaire ℓ par une forme linéaire ℓ_h lorsque le terme source f n'est pas polynomiale. Par exemple, soit

$$\ell(v) = \int_{\Omega} f v \, d\Omega,$$

où f est localement H^2 . Dans ce cas, d'après le théorème 7.2.8, il existe une constante C' telle que

$$\|f - \Pi_h^1 f\|_{L^2(\Omega)} \leq C'h^2 \left(\sum_{l \in \llbracket 1, N_e \rrbracket} |f|_{2, T_l}^2 \right)^{\frac{1}{2}}.$$

Supposons alors que le problème variationnel approché consiste à chercher la solution $\tilde{u}_h \in \mathcal{V}_h$ telle que

$$\forall v_h \in \mathcal{V}_h, \quad a(\tilde{u}_h, v_h) = \ell_h(v_h),$$

où

$$\ell_h(v_h) = \int_{\Omega} (\Pi_h^1 f) v_h \, d\Omega,$$

alors que u_h vérifie $a(u_h, v_h) = \ell(v_h)$ pour tout $v_h \in \mathcal{V}_h$.

On a, par l'inégalité triangulaire,

$$\|u - \tilde{u}_h\|_{H^1(\Omega)} \leq \|u - u_h\|_{H^1(\Omega)} + \|u_h - \tilde{u}_h\|_{H^1(\Omega)}$$

avec d'après le théorème 7.2.11, l'existence d'une constante C telle que

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch \left(\sum_{l=1}^{N_e} |u|_{2, T_l}^2 \right)^{\frac{1}{2}}.$$

Par ailleurs

$$\begin{aligned} \nu \|u_h - \tilde{u}_h\|_{H^1(\Omega)}^2 &\leq a(u_h - \tilde{u}_h, u_h - \tilde{u}_h) \\ &= \int_{\Omega} (f - (\Pi_h^1 f))(u_h - \tilde{u}_h) \, d\Omega \leq \|f - (\Pi_h^1 f)\|_{L^2(\Omega)} \|u_h - \tilde{u}_h\|_{L^2(\Omega)}, \end{aligned}$$

dont on déduit

$$\|u_h - \tilde{u}_h\|_{H^1(\Omega)} \leq \frac{1}{\nu} \|f - (\Pi_h^1 f)\|_{L^2(\Omega)}.$$

On obtient finalement

$$\|u - \tilde{u}_h\|_{H^1(\Omega)} \leq Ch \left(\sum_{l=1}^{N_e} |u|_{2,T_l}^2 \right)^{\frac{1}{2}} + \frac{C'}{\nu} h^2 \left(\sum_{l=1}^{N_e} |f|_{2,T_l}^2 \right)^{\frac{1}{2}}. \quad (7.39)$$

Dans ce cas, on voit que la vitesse de convergence n'est pas affectée par l'erreur d'interpolation sur les données. En particulier, dans les hypothèses du théorème 7.2.14,

$$\|u - \tilde{u}_h\|_{H^1(\Omega)} \leq Ch \|u\|_{H_2(\Omega)} + \frac{C'}{\nu} h^2 \|f\|_{L^2(\Omega)} \leq C'' h (1 + h) \|f\|_{L^2(\Omega)},$$

pour des constantes indépendantes de u et h .

7.2.4 Lemme de Strang

Plus généralement, les formes bilinéaire et linéaire sont sujettes à approximation. Nous venons de le voir pour l'approximation des données, mais cela peut aussi être dû à la différence entre le domaine triangulé et le domaine initialement modélisé, ou encore au calcul des coefficients de la matrice par des méthodes de quadrature numérique.

Dans ce cas, le problème approché n'est plus l'approximation interne du problème initial. Nous le mettons alors sous la forme

$$\text{Trouver } u_h \in \mathcal{V}_h \text{ tel que, } \forall v_h \in \mathcal{V}_h, a_h(u_h, v_h) = \ell_h(v_h), \quad (7.40)$$

où $\mathcal{V}_h$ est encore un sous-espace de dimension finie de $\mathcal{V}$.

Nous allons toujours chercher à estimer $\|u - u_h\|$. Cependant, il faudra d'abord vérifier que les différents problèmes (7.40) sont bien posés; en effet, nous ne disposons plus de la propriété de coercivité qui nous assurait que (7.2) était bien posé dès que a est coercive.

Pour ce faire nous introduisons la notion de famille stable de formes bilinéaires approchées.

Définition 7.2.6 Une famille de forme bilinéaire $(a_h)_h$ est dite stable s'il existe une constante $\beta > 0$, indépendante de h , telle que :

$$\sup_{w_h \in \mathcal{V}_h} \frac{a_h(v_h, w_h)}{\|w_h\|_{\mathcal{V}}} \geq \beta \|v_h\|_{\mathcal{V}}, \quad \forall v_h \in \mathcal{V}_h. \quad (7.41)$$

Lemme 7.2.15 Si les $(a_h)_h$ forment une famille stable de formes bilinéaires, alors les problèmes (7.40) sont bien posés.

Démonstration. Soit u_h , solution de (7.40), homogène ; on aura $a_h(u_h, w_h) = 0 \forall w_h \in \mathcal{V}_h$. On en déduit que $\|u_h\|_{\mathcal{V}} = 0$, d'après (7.41). On a donc démontré l'unicité de u_h , solution de (7.40) ; l'existence s'en déduit, puisque $\mathcal{V}_h$ est de dimension finie. $\square$

Théorème 7.2.16 (Lemme de Strang) Si les $(a_h)_h$ forment une famille stable de formes bilinéaires, alors il existe une constante C , indépendante de h , telle que

$$\|u - u_h\|_{\mathcal{V}} \leq C \left\{ \inf_{v_h \in \mathcal{V}_h} \left(\|u - v_h\|_{\mathcal{V}} + \sup_{w_h \in \mathcal{V}_h} \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_{\mathcal{V}}} \right) + \sup_{w_h \in \mathcal{V}_h} \frac{|\ell(w_h) - \ell_h(w_h)|}{\|w_h\|_{\mathcal{V}}} \right\},$$

où u_h est la solution de (7.40).

Démonstration. Notons tout d'abord que

$$\|u - u_h\|_{\mathcal{V}} \leq \|u - v_h\|_{\mathcal{V}} + \|v_h - u_h\|_{\mathcal{V}}, \quad \forall v_h \in \mathcal{V}_h,$$

et par conséquent

$$\|u - u_h\|_{\mathcal{V}} \leq \inf_{v_h \in \mathcal{V}_h} (\|u - v_h\|_{\mathcal{V}} + \|v_h - u_h\|_{\mathcal{V}}). \quad (7.42)$$

D'après la stabilité de la famille des $(a_h)_h$, on a

$$\begin{aligned} \beta \|u_h - v_h\|_{\mathcal{V}} &\leq \sup_{w_h \in \mathcal{V}_h} \frac{|a_h(u_h - v_h, w_h)|}{\|w_h\|_{\mathcal{V}}} \\ &\leq \sup_{w_h \in \mathcal{V}_h} \frac{|a(u, w_h) - a_h(v_h, w_h) + \ell_h(w_h) - \ell(w_h)|}{\|w_h\|_{\mathcal{V}}}, \end{aligned}$$

puisque $a(u, w_h) = \ell(w_h)$ et $a_h(u_h, w_h) = \ell_h(w_h)$.

On en déduit que

$$\begin{aligned} \beta \|u_h - v_h\|_{\mathcal{V}} &\leq \sup_{w_h \in \mathcal{V}_h} \frac{|a(u - v_h, w_h)|}{\|w_h\|_{\mathcal{V}}} \\ &\quad + \sup_{w_h \in \mathcal{V}_h} \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_{\mathcal{V}}} + \sup_{w_h \in \mathcal{V}_h} \frac{|\ell(w_h) - \ell_h(w_h)|}{\|w_h\|_{\mathcal{V}}}, \end{aligned}$$

et la conclusion en résulte, en raison de la continuité de a , et de l'inégalité (7.42). $\square$

Toute la difficulté se réduit donc à trouver des conditions simples telles que la famille des $(a_h)_h$ soit stable.

Lemme 7.2.17 Si a est une forme bilinéaire coercive, et s'il existe une fonction $C(h)$ vérifiant

$$C(h) \rightarrow 0 \quad \text{quand } h \rightarrow 0,$$

et

$$|a(v_h, w_h) - a_h(v_h, w_h)| \leq C(h) \|v_h\|_{\mathcal{V}} \|w_h\|_{\mathcal{V}}, \quad \forall v_h, w_h \in \mathcal{V}_h,$$

alors, pour h suffisamment petit, la famille des a_h est stable.

Démonstration. On a

$$|a(v_h, w_h)| \leq |a_h(v_h, w_h)| + |a(v_h, w_h) - a_h(v_h, w_h)|;$$

on en déduit que

$$\sup_{w_h \in \mathcal{V}_h} \frac{|a(v_h, w_h)|}{\|w_h\|_{\mathcal{V}}} \leq \sup_{w_h \in \mathcal{V}_h} \frac{|a_h(v_h, w_h)|}{\|w_h\|_{\mathcal{V}}} + \sup_{w_h \in \mathcal{V}_h} \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_{\mathcal{V}}},$$

soit encore

$$\sup_{w_h \in \mathcal{V}_h} \frac{|a_h(v_h, w_h)|}{\|w_h\|_{\mathcal{V}}} \geq \sup_{w_h \in \mathcal{V}_h} \frac{|a(v_h, w_h)|}{\|w_h\|_{\mathcal{V}}} - \sup_{w_h \in \mathcal{V}_h} \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_{\mathcal{V}}}.$$

Mais la forme bilinéaire a est coercive, donc

$$\sup_{w_h \in \mathcal{V}_h} \frac{|a(v_h, w_h)|}{\|w_h\|_{\mathcal{V}}} \geq \frac{a(v_h, v_h)}{\|v_h\|_{\mathcal{V}}} \geq \alpha \|v_h\|_{\mathcal{V}},$$

et en tenant compte de l'hypothèse,

$$\sup_{w_h \in \mathcal{V}_h} \frac{|a_h(v_h, w_h)|}{\|w_h\|_{\mathcal{V}}} \geq (\alpha - C(h)) \|v_h\|_{\mathcal{V}},$$

ce qu'il fallait démontrer. $\square$

7.2.5 Estimations de l'erreur en norme L^2

Nous poursuivons ici avec notre formulation variationnelle (7.19) dans $\mathcal{V} = H^1(\Omega)$, pour laquelle nous venons de démontrer, sous certaines conditions, des estimations de l'erreur dans la norme de l'espace $\mathcal{V}$, ici « en norme H^1 ». Ainsi, l'erreur en norme H^1 se comporte comme h pour les éléments finis $\mathbb{P}_1$, comme l'indique le théorème 7.2.14. Cependant, le théorème de Bramble–Hilbert 7.2.8 montre que si les estimations sont d'ordre h , c'est parce que, dans la norme de $\mathcal{V}$, nous cherchons à contrôler les dérivées premières, et donc la semi-norme $|\cdot|_1$. En revanche, $\|u - \Pi_h^1 u\|_{L^2}$ se comporte comme h^2 . On pourrait donc gagner un ordre de convergence en ne cherchant à contrôler que la convergence dans une norme plus faible. C'est l'objectif du lemme d'Aubin–Nitsche¹¹ suivant¹².

Théorème 7.2.18 (Aubin–Nitsche) Soit Ω un ouvert polygonal et $(\mathcal{T}_h)_{h>0}$ une suite de maillages réguliers de Ω . Soit $u \in H^1(\Omega)$ la solution du problème variationnel (7.19), et $u_h \in \mathcal{V}_h$, celle de son approximation interne (7.2) par

11: Joachim A. Nitsche (1926–1996) est un mathématicien allemand qui a apporté d'importantes contributions à l'analyse et à l'analyse numérique. Il en va de même pour Jean-Pierre Aubin (né en 1939), à ne pas confondre avec son homonyme Thierry Aubin (1942–2009), également mathématicien et académicien, ayant travaillé sur les EDP non linéaires. En anglais, on parle de Aubin–Nitsche trick.

12: Notons que ce résultat ne peut pas être déduit directement des résultats précédents, car le lemme de Céa ne permet de contrôler que $\|u - u_h\|_{H^1}$ à partir de l'erreur de meilleure approximation en

la méthode des éléments finis $\mathbb{P}_1$ pour $d \in \{2, 3\}$. On suppose par ailleurs que $f \in L^2(\Omega)$. Si Ω est convexe alors

$$\|u - u_h\|_{L^2(\Omega)} \leq Ch^2 \|f\|_{L^2(\Omega)}, \quad (7.43)$$

où C est une constante indépendante de h et de u .

Le théorème est donné ici dans le cadre spécifique du problème variationnel (7.19) mais la preuve est très générale et s'adapte à de nombreux cas.

Démonstration. On commence d'abord par une estimation d'ordre 1 en utilisant le théorème 7.2.14. Il existe une constante C telle que

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

Pour le résultat d'estimation à l'ordre 2, on introduit un problème auxiliaire sous la forme :

Trouver $w \in H^1(\Omega)$ tel que

$$a(v, w) = \int_{\Omega} (u - u_h) v \, d\Omega, \quad v \in H^1(\Omega). \quad (7.44)$$

Par le théorème 3.5.1 de Lax-Milgram, ce problème admet une unique solution $w \in H^1(\Omega)$. Comme $u - u_h \in L^2(\Omega)$, on déduit d'après le théorème 4.4.7, que si Ω est convexe alors la solution $w \in H^2(\Omega)$. En choisissant $v = u - u_h$, on déduit

$$\begin{aligned} \|u - u_h\|_{L^2(\Omega)}^2 &= a(u - u_h, w) \\ &= a(u - u_h, w - \Pi_h^1 w) \\ &\leq M \|w - \Pi_h^1 w\|_{H_0^1(\Omega)} \|u - u_h\|_{H_0^1(\Omega)}, \end{aligned}$$

où M est la constante de continuité de a et Π_h^1 est l'opérateur d'interpolation $\mathbb{P}_1$ qui est bien défini sur $H^2(\Omega) \subset C^0(\bar{\Omega})$. En utilisant l'estimation d'erreur à l'ordre 1 ainsi que l'erreur d'interpolation de w par les éléments finis $\mathbb{P}_1$, il vient

$$\|u - u_h\|_{L^2(\Omega)}^2 \leq Ch \|w\|_{H^2(\Omega)} C_2 h \|f\|_{L^2(\Omega)}.$$

Comme w est solution de (7.44) pour $f = u - u_h$, on a par hypothèse que $\|w\|_{H^2(\Omega)} \leq C \|u - u_h\|_{L^2(\Omega)}$, et on obtient l'estimation annoncée. $\square$

Remarque 7.2.8 (Problème non symétrique) Dans la preuve, nous remarquons que le problème auxiliaire (7.44) est posé sous la forme

$$a(v, w) = \ell(v) \text{ et non pas } a(w, v) = \ell(v).$$

Ainsi dans le cas général où a n'est pas symétrique, le problème auxiliaire est formulé à partir de la forme bilinéaire « adjointe » a^* que nous pouvons définir par

$$\forall (u, v) \in \mathcal{V}^2, \quad a^*(w, v) := a(v, w).$$

La fin du théorème nécessite alors d'appliquer un théorème de régularité sur w comme solution d'un problème défini à partir de a^* .

7.3 Principe du maximum discret

On se place dans le cadre du paragraphe 4.4.1 où nous avons pu montrer le principe du maximum pour le laplacien. Pour $\alpha \geq 0$, la forme bilinéaire était dans ce cas

$$a_\alpha(u, v) = \int_{\Omega} \left[\underline{\nabla} u \cdot \underline{\nabla} v + \alpha uv \right] d\Omega \quad (7.45)$$

et la forme bilinéaire

$$\ell(v) = \int_{\Omega} f v d\Omega. \quad (7.46)$$

Nous cherchons à vérifier le même type de principe du maximum, mais au niveau discret. Nous nous limitons au cas de conditions de Dirichlet homogènes. Le problème variationnel est : on cherche $u^\alpha \in \mathcal{V} = H_0^1(\Omega)$ tel que

$$\forall v \in H_0^1(\Omega), \quad a_\alpha(u^\alpha, v) = \ell(v), \quad (\mathcal{F}_\alpha)$$

qui est bien posé d'après le chapitre 4.

Définition 7.3.1 (M-Matrice) On dit que la matrice réelle $\mathbb{A} = (A^{ij})_{1 \leq i, j \leq N} \in M_N(\mathbb{R})$ est une M-matrice si et seulement si pour tout $i \in \llbracket 1, N \rrbracket$,

$$A^{ii} > 0, \quad \forall j \in \llbracket 1, N \rrbracket \setminus \{i\}, \quad A^{ij} \leq 0, \quad \text{et} \quad \sum_{1 \leq j \leq N} A^{ij} > 0.$$

Remarque 7.3.1 Il existe d'autres définitions de M-matrices dont on démontre qu'elles sont équivalentes à la définition 7.3.1. Notamment celle pour laquelle une M-matrice est de la forme $\mathbb{A} = s\mathbb{I} - \mathbb{B}$ où $\mathbb{B} = (B^{ij})_{1 \leq i, j \leq N} \in M_N(\mathbb{R})$ est telle que $B^{ij} \geq 0$ et $s > \rho(\mathbb{B})$, le rayon spectral de $\mathbb{B}$.

Définition 7.3.2 (Matrice monotone) Pour tout $\vec{u} = (u^i)_{1 \leq i \leq N} \in \mathbb{R}^N$, on note

$$\vec{u} \geq 0 \Rightarrow \left(\forall i \in \llbracket 1, N \rrbracket, \quad u_i \geq 0 \right).$$

On dit alors que la matrice réelle $\mathbb{A} = (A^{ij})_{1 \leq i, j \leq N} \in M_N(\mathbb{R})$ est monotone si et seulement si

$$\forall \vec{u} \in \mathbb{R}^N, \quad \mathbb{A}\vec{u} \geq 0 \Rightarrow \vec{u} \geq 0.$$

Proposition 7.3.1 Soit $\mathbb{A} = (A^{ij})_{1 \leq i, j \leq N} \in M_N(\mathbb{R})$ une matrice monotone alors $\mathbb{A}$ est inversible et son inverse $\mathbb{D} = (D^{ij})_{1 \leq i, j \leq N} \in M_N(\mathbb{R})$ a tous ses coefficients positifs, c'est à dire pour tout $(i, j) \in \llbracket 1, N \rrbracket^2$, $D^{ij} \geq 0$.

Démonstration. Commençons par l'inversibilité. Si $\mathbb{A}\underline{u} = 0$ alors $\mathbb{A}\underline{u} \geq 0$

et $-\mathbb{A}u \geq 0$ ce qui implique que $\underline{u} \geq 0$ et $-\underline{u} \geq 0$. Donc $u = 0$ et le noyau de $\mathbb{A}$ est réduit à 0.

Soit $i \in \llbracket 1, N \rrbracket$ soit $\vec{b} = (b^j)_{1 \leq j \leq N}$ tel que $b^j = \delta_{i,j}^j$. On considère $\vec{u} = (u^k)_{1 \leq k \leq N}$ tel que $\vec{u} = \mathbb{D}\vec{b}$. Par monotonie de $\mathbb{A}$, on a $\vec{b} = \mathbb{A}\vec{u} \geq 0 \Rightarrow \vec{u} \geq 0$. Et donc pour tout $j \in \llbracket 1, N \rrbracket$, $0 \leq u^j = D^{ij}b^i = D^{ij}$. Ceci étant vrai pour tout $i \in \llbracket 1, N \rrbracket$ alors on a le résultat. $\square$

Proposition 7.3.2 Soit $\mathbb{A} = (A^{ij})_{1 \leq i, j \leq N} \in M_N(\mathbb{R})$ une M-matrice, alors $\mathbb{A}$ est monotone.

Démonstration. Soit un vecteur $\vec{b} = (b^i)_{1 \leq i \leq N} \geq 0$ et $\vec{u} = (u^i)_{1 \leq i \leq N}$ tel que $\mathbb{A}\vec{u} = \vec{b}$. On veut montrer que $\vec{u} \geq 0$. Si on minore, on a

$$\begin{aligned} b^i &= \sum_{1 \leq j \leq N} A^{ij} u^j \leq A^{ii} u^i + \sum_{j \neq i} A^{ij} \min_k(u^k) \\ &\leq A^{ii}(u^i - \min_k(u^k)) + \sum_{1 \leq j \leq N} A^{ij} \min_k(u^k). \end{aligned}$$

Notamment pour i_m tel que $u^{i_m} = \min_k u^k$, on a

$$0 \leq b^{i_m} \leq \min_k(u^k) \sum_{1 \leq j \leq N} A^{ij} \Rightarrow \min_k(u^k) \geq 0$$

car $\sum_{1 \leq j \leq N} A^{ij} \geq 0$. Ainsi on a bien $\vec{u} \geq 0$. $\square$

Theorème 7.3.3 (Principe du maximum discret en dimension 1) Soient $\alpha \geq 0$, $f \in C^0([0, L])$ et u_h^α la solution approchée par éléments finis $\mathbb{P}_1$ du problème $(\mathcal{F}_\alpha)$ sur le domaine $\Omega =]0, L[$. Alors, il existe $h_0 \in \mathbb{R}_+^*$ tel que pour tout $h \leq h_0$

$$\left(\forall x \in]0, L[, \quad f(x) \geq 0 \right) \Rightarrow \left(\forall x \in]0, L[, \quad u_h^\alpha(x) \geq 0 \right).$$

Démonstration. Reprenons la matrice associée à la formulation variationnelle du problème $\mathcal{F}_\alpha$ sur le domaine $[0, L]$,

$$\mathbb{A}_\alpha = \frac{1}{h} \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix} + \frac{\alpha h}{6} \begin{pmatrix} 4 & 1 & & & 0 \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ 0 & & & 1 & 4 \end{pmatrix}.$$

On a donc, pour h suffisamment petit, $\mathbb{A}_\alpha$ une M-matrice pour tout $\alpha > 0$. Ainsi $\mathbb{A}$ est monotone. Si maintenant pour tout $x \geq 0$, $f(x) \geq 0$ alors pour tout sommet du maillage $f_h(x_i) = (\Pi_h^1 f)(x_i) = f(x_i) \geq 0$. Ainsi $\vec{f}_h \geq 0$ donc $\vec{u} \geq 0$. Enfin pour tout $x \in [0, 1]$, on note $T_k(x) \in \mathcal{T}_h$ l'élément contenant x . On a

$$u_h^\alpha(x) = \sum_{S_i \in T_k(x)} \lambda_i(x) u_h^\alpha(S_i),$$

où $\sum_{1 \leq i \leq 3} \lambda_i(x) = 1$ et $\forall 1 \leq i \leq 3, \quad 0 \leq \lambda_i(x) \leq 1$ donc $u_h^\alpha(x) \geq 0$. En passant à la limite en $\alpha \rightarrow 0$ dans cette dernière inégalité, le résultat reste vrai pour $\alpha = 0$. $\square$

Ce résultat peut s'étendre à la dimension supérieure avec des hypothèses supplémentaires. Nous limiterons notre présentation à la dimension $d = 2$ par simplicité. Il faut ajouter une condition sur le maillage qui doit s'appuyer sur une triangulation spécifique dite triangulation de Delaunay¹³

13: Boris Delaunay ou Delone (1890-1980) est un mathématicien russe dont l'ancêtre, l'officier de Launay, fut capturé en Russie pendant la campagne de Napoléon

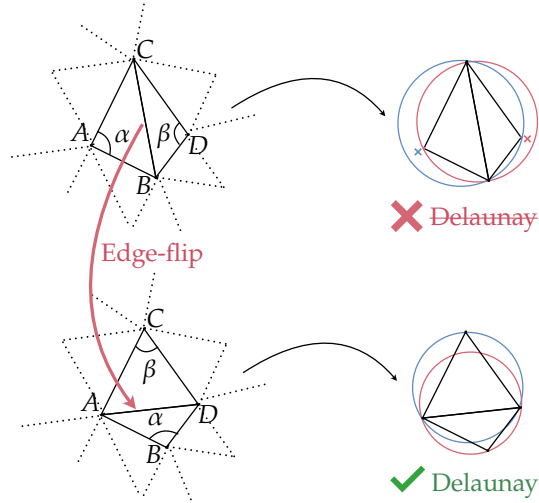


FIGURE 7.4 – Triangulation non Delaunay avant le *edge-flip*, et qui devient Delaunay ensuite.

Définition 7.3.3 (Triangulation de Delaunay) Une triangulation de Delaunay d'un ensemble $\mathcal{P}$ de points du plan est une triangulation $\mathcal{T}$ telle qu'aucun point de $\mathcal{P}$ n'est à l'intérieur du cercle circonscrit d'un des triangles de $\mathcal{T}$.

Proposition 7.3.4 Une triangulation est de Delaunay si pour tout segment, la somme des angles faisant face à ce segment pour les triangles qui contiennent ce segment sont plus petits que π .

Autrement dit avec les notations de la figure 7.4, on doit avoir

$$\alpha + \beta \leq \pi.$$

Un fait remarquable, illustré figure 7.4, est que lorsqu'un maillage n'est pas Delaunay parce qu'un point d'un triangle se retrouve à l'intérieur d'un triangle voisin alors une procédure assez simple dit de *edge-flip* permet de rendre le maillage Delaunay. En effet pour un triangle (A, B, C) , où A, B , et C sont placés dans le sens trigonométrique, le point D est dans le cercle circonscrit si le déterminant

$$\begin{vmatrix} A_x & A_y & A_x^2 + A_y^2 & 1 \\ B_x & B_y & B_x^2 + B_y^2 & 1 \\ C_x & C_y & C_x^2 + C_y^2 & 1 \\ D_x & D_y & D_x^2 + D_y^2 & 1 \end{vmatrix} \geq 0.$$

Ainsi, si ce déterminant est négatif et qu'on change le triangle formé

$\widehat{ABC}$ en un triangle $\widehat{ABD}$ alors on échange les deux dernières lignes du déterminant ce qui rend le nouveau déterminant positif. Cette propriété a permis d'imaginer des algorithmes de construction de triangulation de Delaunay, potentiellement très efficaces par des algorithmes diviser pour régner de complexité en $O(N_e \ln N_e)$, [19].

Theorème 7.3.5 (Principe du maximum discret en dimension 2) Soient $\mathcal{T}_h$ une triangulation de Delaunay sans angle droit intérieur, $\alpha \geq 0$, $f \in C^0(\overline{\Omega})$ et u_h^α la solution approchée par éléments finis $\mathbb{P}_1$ du problème $(\mathcal{F}_\alpha)$, alors, il existe $h_0 \in \mathbb{R}_+^*$ tel que pour tout $h \leq h_0$

$$\left(\forall x \in \Omega, \quad f(x) \geq 0 \right) \Rightarrow \left(\forall x \in \Omega, \quad u_h^\alpha(x) \geq 0 \right).$$

Démonstration. Reprenons la matrice associée à la formulation variationnelle du problème (6.18) : $\mathbb{A}_\alpha = \mathbb{K} + \alpha \mathbb{M}$. Comme nous avons considéré des conditions de Dirichlet homogènes, alors toutes les fonctions de bases de $\mathcal{V}_h$ sont associées à des points intérieurs du maillage tel que représenté sur la figure 7.5. En particulier, chaque sommet est entouré par un ensemble de triangles et chaque arête a exactement deux triangles voisins.

La matrice de masse $\mathbb{M}$ est donnée par (6.23), c'est-à-dire

$$\forall (i, j) \in \llbracket 1, N \rrbracket^2, \quad \mathbb{M}_h^{ij} = \begin{cases} \frac{1}{6} \sum_{S_i \in T_l} |T_l|, & \text{si } i = j, \\ \frac{1}{12} \sum_{[S_i S_j] \in T_l} |T_l|, & \text{sinon.} \end{cases}$$

On rappelle que pour un triangle $T = \widehat{ABC}$ de côtés de longueur a, b et c , on a $|T| = \frac{abc}{4h}$. Or on a $|T|^3 \leq a^2 b^2 c^2 \leq 16|T|^2 h^2$, soit $|T| \leq 16h^2$. Ainsi on obtient $\mathbb{M}_h^{ii} > 0$ et $\forall (i, j) \in \llbracket 1, N \rrbracket^2$ $i \neq j$, $0 \leq \mathbb{M}_h^{ij} \leq \frac{4}{3}h^2$. Enfin on a immédiatement que

$$\forall i \in \llbracket 1, N \rrbracket, \quad \sum_{1 \leq j \leq N} \mathbb{M}_h^{ij} > 0.$$

Pour la raideur, on a

$$\forall (i, j) \in \llbracket 1, N \rrbracket^2, \quad \mathbb{K}_h^{ij} = \begin{cases} \frac{1}{2} \sum_{k \in \mathcal{N}_i} (\cot \theta_{ik} + \cot \theta_{ki}), & \text{si } i = j, \\ -\frac{1}{2} (\cot \theta_{ij} + \cot \theta_{ji}), & \text{sinon,} \end{cases}$$

où $\mathcal{N}_i$ est l'ensemble des sommets S_j adjacents à S_i , c'est à dire tels que $[S_i, S_j]$ est une arête du maillage, et, pour tout $(i, j) \in \llbracket 1, N \rrbracket^2$, alors θ_{ij} et $\cot \theta_{ji}$ sont les angles opposés à $[S_i, S_j]$ des deux triangles ayant $[S_i, S_j]$ pour arêtes. Si on réordonne les termes, on a que

$$\mathbb{K}_h^{ii} = \frac{1}{2} \sum_{\substack{S_i \in T_l \\ T_l = \widehat{S_i S_j S_k}}} \cot \theta_j + \cot \theta_k.$$

FIGURE 7.5 – Angles θ_{ij} autour d'une arête $[S_i S_j]$ et angles $\theta_i, \theta_j, \theta_k$ dans le triangle $\overline{S_i S_j S_k}$

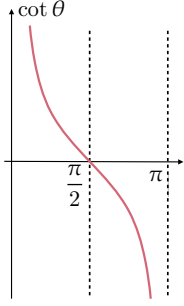
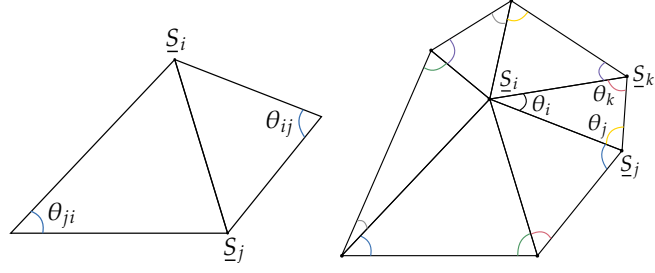


FIGURE 7.6 – fonction $\theta \mapsto \cot \theta$

Or dans le triangle $T_l = S_i S_j S_k$, $\theta_i + \theta_j + \theta_k = \pi$ donc $\theta_j + \theta_k < \pi$. Supposons alors que le plus petit des angles est θ_j . On a $\theta_j \leq \frac{\pi}{2}$ et donc $\theta_k \leq \frac{\pi}{2} + (\frac{\pi}{2} - \theta_j)$. Or la fonction $\cot$ est symétrique par rapport à $\frac{\pi}{2}$ donc

$$\cot \theta_j + \cot \theta_k > 0.$$

Le même résultat reste valable si le plus petit des angles est θ_k , donc $\mathbb{K}_h^{ii} > 0$. De même pour $i \neq j$, la triangulation étant de Delaunay, on a $\theta_{ij} + \theta_{ji} \leq \pi$. Et par hypothèse sur les angles droits intérieurs, on a $\theta_{ij} + \theta_{ji} < \pi$. Donc $\cot \theta_{ij} + \cot \theta_{ji} > 0$ qui implique $\mathbb{K}_h^{ij} < 0$. Enfin on voit directement que $\mathbb{K}_h(1 \cdots 1)^T = 0$ qui est d'ailleurs aussi justifié par le fait que les constantes sont dans le noyau de la forme bilinéaire associée à $\mathbb{K}_h$.

On conclut que, pour h suffisamment petit et pour tout $\alpha > 0$, $\mathbb{A}_\alpha$ est monotone. Exactement comme dans la preuve du théorème 7.3.3 en 1D, on a bien pour tout $\alpha > 0$

$$\left(\forall x \in \Omega, \quad f(x) \geq 0 \right) \Rightarrow \left(\forall x \in \Omega, \quad u_h^\alpha(x) \geq 0 \right).$$

Cette inégalité reste alors vraie en passant à la limite $\alpha \rightarrow 0$.

□

Analyse spectrale des problèmes aux limites

8

8.1 Introduction

Dans ce chapitre, on s'intéresse aux valeurs λ telles qu'un problème du type

$$-\Delta u = \lambda u, \quad \text{sur } \Omega,$$

assortis de conditions aux limites homogènes de type Dirichlet, Neumann ou Robin sur la frontière $\partial\Omega$ de Ω supposé borné, admette des solutions u non nulles. On dit alors que λ est une valeur propre et u une fonction propre associée. L'analogie avec les valeurs propres d'une matrice est assez naturelle dès qu'on a associé au problème aux limites¹ un opérateur linéaire. L'ensemble des valeurs propres de cet opérateur est appelé spectre de l'opérateur². On verra qu'alors que le spectre d'une matrice est fini, il ne l'est pas pour les opérateurs étudiés dans ce chapitre.

1: qui définit une équation aux dérivées partielles dite elliptique puisque tous les coefficients devant les dérivées d'ordre 2 sont positives

2: d'où le titre du chapitre

Nous avons vu dans le chapitre 1 quelques motivations de ce chapitre : étude des solutions dites harmoniques de l'équation des ondes, ou encore la recherche des modes de vibrations d'une structure élastique.

8.2 Un problème aux valeurs propres modèle

Soient Ω un ouvert régulier dans $\mathbb{R}^d$ et $f \in L^2(\Omega)$. On cherche $\lambda \in \mathbb{R}$, telle qu'il existe $u \neq 0$ solution dans $H^1(\Omega)$ de

$$\begin{cases} -\Delta u - \lambda u = 0, & \text{dans } \Omega, \\ u = 0, & \text{sur } \partial\Omega. \end{cases} \quad (8.1)$$

Il est possible d'étendre tous les résultats de ce chapitre au cas des conditions de Neumann, des conditions de Robin, des conditions mixtes (Dirichlet, Neumann, Robin dans un bout du bord).

En utilisant les mêmes arguments que ceux développés au chapitre 4, on montre que ce problème est équivalent à la formulation variationnelle suivante :

On cherche $\lambda \in \mathbb{R}$, telle qu'il existe $u \neq 0$ solution dans $H_0^1(\Omega)$

$$\forall v \in H_0^1(\Omega), \quad \int_{\Omega} [\nabla u \cdot \nabla v - \lambda uv] \, d\Omega = 0. \quad (8.2)$$

Introduisons la forme bilinéaire $a : H_0^1(\Omega) \times H_0^1(\Omega) \rightarrow \mathbb{R}$ définie par

$$\forall v, w \in H_0^1(\Omega), \quad a(w, v) := \int_{\Omega} \nabla w \cdot \nabla v \, d\Omega.$$

Nous avons démontré à la section 4.2.1 que cette forme bilinéaire est continue et coercive. Notons enfin $a_{\lambda} : H_0^1(\Omega) \times H_0^1(\Omega) \rightarrow \mathbb{R}$ la forme

bilinéaire définie par

$$\forall v, w \in H_0^1(\Omega), \quad a_\lambda(w, v) := a(w, v) - \lambda \int_\Omega wv \, d\Omega.$$

- Si $\lambda < 0$, a_λ est évidemment coercive. On déduit immédiatement que la solution possible de (8.1) est la solution nulle. Il n'existe donc pas de valeurs propres négatives.
- Si $\lambda \geq 0$ mais suffisamment petit a_λ reste coercive. En effet, si $\lambda < C_\Omega^{-1}$, avec C_Ω la constante de Poincaré de Ω , c'est-à-dire la plus petite constante C non nulle telle que l'inégalité de Poincaré

$$\forall u \in \mathcal{V}, \quad \int_\Omega u^2 \, d\Omega \leq C \int_\Omega |\nabla u|^2 \, d\Omega$$

est vérifiée alors a_λ est coercive. On déduit comme précédemment qu'il ne peut exister de valeurs propres $\lambda < C_\Omega^{-1}$.

- Dès que $\lambda \geq C_\Omega^{-1}$, il est possible de montrer que la forme bilinéaire a_λ n'est pas coercive. Nous allons montrer que pour certaines valeurs de λ , le problème (8.1) peut admettre des solutions non nulles.

8.3 Problème aux valeurs propres sous forme opérateur

3: et des problèmes variationnels similaires

4: On suppose ici que les espaces de Hilbert sont séparables. Un espace de Hilbert est dit séparable s'il contient un sous-ensemble dénombrable et dense. On peut alors montrer que l'on peut construire une famille dénombrable $(e_k)_{k \in \mathbb{N}}$ qui en constitue une base orthonormale. Ainsi, on peut écrire tout élément v de $\mathcal{H}$ sous la forme

$$v = \sum_{k=0}^{\infty} v_k e_k,$$

avec $\forall k \in \mathbb{N}, v_k = (v|e_k)_V$, et

$$\|v\|_{\mathcal{H}} = \left(\sum_{k=0}^{\infty} v_k^2 \right)^{\frac{1}{2}}.$$

Une telle base est appelée base hilbertienne. Nous avons même la caractérisation

$$\mathcal{H} \equiv \left\{ \sum_{k=0}^{\infty} v_k e_k, \quad \sum_{k=0}^{\infty} v_k^2 < +\infty \right\}.$$

5: On rappelle que *injection continue* signifie que l'application identité $\iota_{\mathcal{V}} : \mathcal{V} \rightarrow \mathcal{H}$ est continue, c'est à dire (voir section 8.4.1) qu'il existe une constante $C \in \mathbb{R}_+$ telle que

$$\forall u \in \mathcal{V}, \quad \|u\|_{\mathcal{H}} \leq C \|u\|_{\mathcal{V}}.$$

6: par exemple en considérant $\mathcal{H} = L^2(\Omega)$ et $\mathcal{V} = H_0^1(\Omega)$

Il est possible de reformuler le problème variationnel précédent³ sous la forme opérateur. Dans la suite nous le faisons dans un cas général.

Nous considérons deux espaces de Hilbert séparables⁴ $\mathcal{H}$ et $\mathcal{V}$ tels que⁵

$$\begin{cases} \mathcal{V} \subset \mathcal{H} \text{ avec injection continue}^5, \\ \mathcal{V} \text{ est dense dans } \mathcal{H}. \end{cases} \quad (8.3)$$

À noter que les espaces $\mathcal{H}$ et $\mathcal{V}$ ne partagent pas le même produit scalaire, nous notons leurs produits scalaires respectifs $(\cdot, \cdot)_{\mathcal{H}}$ et $(\cdot, \cdot)_{\mathcal{V}}$. À noter aussi que puisque nous définissons deux espaces, nous ne pouvons identifier qu'un seul espace avec son dual via le théorème de Riesz. Ici typiquement nous identifierons $\mathcal{H}$ avec son dual $\mathcal{H}'$ ce qui donne alors le jeu d'inclusions

$$\mathcal{V} \subset \mathcal{H} \equiv \mathcal{H}' \subset \mathcal{V}'.$$

Cette structure d'espace se retrouve dans de nombreux problèmes variationnels et s'appelle un *Triplet de Gelfand*. En particulier dans ce cadre, on a

$$(u, v) \in \mathcal{H} \times \mathcal{V}, \quad \langle u, v \rangle_{\mathcal{V}', \mathcal{V}} = (u, v)_{\mathcal{H}}$$

qui est au coeur de nombreux raisonnements sur les distributions ou dans les espaces de Sobolev⁶.

Dans ce cadre, nous introduisons une forme bilinéaire $a(\cdot, \cdot)$, *symétrique* c'est-à-dire que pour tout w, v , $a(w, v) = a(v, w)$; continue, c'est-à-dire qu'il existe une constante $C_a > 0$ telle que

$$\forall w, v \in \mathcal{V}, \quad |a(w, v)| \leq C_a \|w\|_{\mathcal{V}} \|v\|_{\mathcal{V}};$$

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

$$\forall v \in \mathcal{V}, \quad a(v, v) \geq \alpha \|v\|_{\mathcal{V}}^2.$$

Nous considérons le problème aux valeurs propres suivant :

$$\text{Trouver } \lambda \in \mathbb{R} \text{ et } u \in \mathcal{V} \setminus \{0\} \text{ tel que } a(u, v) = \lambda(u, v)_{\mathcal{H}}. \quad (8.4)$$

Puisque $a(u, \cdot)$ est une forme linéaire dans $\mathcal{V}$, on peut lui associer un élément $A(u) \in \mathcal{V}'$. De plus $u \mapsto A(u)$ est linéaire en u puisque $a(u, \cdot)$ est linéaire par rapport à u . Donc il existe $A \in \mathcal{L}(\mathcal{V}, \mathcal{V}')$ tel que

$$\langle Au, v \rangle_{\mathcal{V}', \mathcal{V}} = a(u, v). \quad (8.5)$$

Ainsi le problème aux valeurs propres sous forme variationnelle s'écrit de manière strictement équivalente

$$\text{Trouver } \lambda \in \mathbb{R} \text{ et } u \in \mathcal{V} \text{ tels que } Au = \lambda u \text{ dans } \mathcal{V}'. \quad (8.6)$$

Mais comme $u \in \mathcal{V} \subset \mathcal{H}$, alors $Au \in \mathcal{H}$ et donc l'égalité précédente est vraie dans $\mathcal{H}$. Ainsi le problème aux valeurs propres s'écrit sous forme variationnelle comme suit

$$\text{Trouver } \lambda \in \mathbb{R} \text{ et } u \in \mathcal{V} \text{ tels que } Au = \lambda u \text{ dans } \mathcal{H}. \quad (8.7)$$

Si $\mathcal{H}$ (et donc $\mathcal{V}$) est un espace de dimension finie (de dimension notée N), A peut être assimilée à une matrice $\mathbb{A}$ et on retrouve le problème classique de recherche de valeurs propres d'une matrice symétrique et définie positive (car la forme bilinéaire est symétrique coercive). Dans ce cas, on sait que $\mathbb{A}$ est diagonalisable, toutes ses valeurs propres $\lambda_1 \leq \dots \leq \lambda_N$ (que nous avons répétées selon leur multiplicité) sont strictement positives. Toujours pour les matrices, le quotient de Rayleigh défini⁷ par

$$\forall \vec{u} \in \mathbb{R}^N \setminus \{0\}, \quad R_{\mathbb{A}}(\vec{u}) := \frac{\vec{u}^T \mathbb{A} \vec{u}}{\vec{u}^T \vec{u}} \in \mathbb{R}$$

donne des informations sur le spectre de la matrice. En effet, quand on l'applique à un vecteur propre, il fournit la valeur propre correspondante. Les valeurs propres peuvent même être déterminées en utilisant le principe dit *du min-max* ou *Courant-Fisher*

$$\forall k, \quad \lambda_k = \min_{W \in \mathcal{E}_k} \max_{\vec{u} \in W \setminus \{0\}} R_{\mathbb{A}}(\vec{u}) = \max_{W \in \mathcal{E}_{N-k+1}} \min_{\vec{u} \in W \setminus \{0\}} R_{\mathbb{A}}(\vec{u}),$$

où $\mathcal{E}_k$ est l'ensemble des sous-espaces vectoriels de $\mathbb{R}^N$.

7: une définition plus générale pour les matrices est parfois donnée sous la forme $R_{\mathbb{A}}(\vec{u}) := \frac{\vec{u}^T \mathbb{A} \vec{u}}{\vec{u}^T \mathbb{M} \vec{u}}$, pour une matrice $\mathbb{M}$ définie positive donnée.

Nous allons maintenant considérer le cas où $\mathcal{H}$ et $\mathcal{V}$ sont de dimension infinie. Il faut savoir qu'en dimension infinie, la nature du spectre des opérateurs peut être bien plus compliquée : il n'existe pas toujours des valeurs propres comme il peut en exister de multiplicité infinie, certaines valeurs ressemblent à des valeurs propres dans le sens où elles admettent des vecteurs propres qui ne sont pas dans l'espace d'étude,... Nous allons néanmoins introduire une classe d'opérateurs pour lesquels la nature du spectre est très similaire à celle des opérateurs en dimension finie. Comme vous le verrez, c'est une classe d'opérateurs très particulière.

8.4 Analyse spectrale des opérateurs auto-adjoints compacts

8.4.1 Opérateurs bornés

Avant de définir les opérateurs compacts, rappelons un concept essentiel en analyse fonctionnelle avec la notion d'opérateur borné.

On notera dans la suite pour tout ensemble $\mathcal{H}$, $\mathcal{H}^* := \mathcal{H} \setminus \{0\}$.

Définition 8.4.1 Soient $\mathcal{H}_1$ et $\mathcal{H}_2$ deux espaces de Hilbert. L'opérateur linéaire $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ est dit borné si l'image de la boule unité de $\mathcal{H}_1$ par T est bornée dans $\mathcal{H}_2$. On notera alors la norme d'opérateur

$$\|T\|_{\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)} = \sup_{u \in \mathcal{H}_1^*} \frac{\|Tu\|_{\mathcal{H}_2}}{\|u\|_{\mathcal{H}_1}} = \sup_{\substack{u \in \mathcal{H}_1^* \\ \|u\|_{\mathcal{H}_1} = 1}} \|Tu\|_{\mathcal{H}_2}.$$

On vérifiera aisément que

$$\|T\|_{\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)} = \sup_{(u,v) \in \mathcal{H}_1^* \times \mathcal{H}_2^*} \frac{(Tu, v)_{\mathcal{H}_2}}{\|u\|_{\mathcal{H}_1} \|v\|_{\mathcal{H}_2}} = \sup_{\substack{(u,v) \in \mathcal{H}_1^* \times \mathcal{H}_2^* \\ \|u\|_{\mathcal{H}_1} = \|v\|_{\mathcal{H}_2} = 1}} (Tu, v)_{\mathcal{H}_2},$$

mais aussi que $\|T\|$ est le rayon de la plus petite boule contenant l'image de la boule unité.

On notera aussi que s'agissant d'opérateurs linéaires, être un opérateur borné signifie être continue en 0 et donc être continue sur tout l'espace $\mathcal{H}_1$ puisque

$$\forall (u, v) \in \mathcal{H}_1 \times \mathcal{H}_1, \quad \|Tu - Tv\|_{\mathcal{H}_2} = \|T(u-v)\|_{\mathcal{H}_2} \leq \|T\|_{\mathcal{L}(\mathcal{H}_1, \mathcal{H}_2)} \|u-v\|_{\mathcal{H}_1}.$$

Réciproquement si un opérateur est continu alors pour $\varepsilon > 0$, il existe une constante α telle que

$$\|u\|_{\mathcal{H}_1} \leq \alpha \quad \Rightarrow \quad \|Tu\|_{\mathcal{H}_2} \leq \varepsilon$$

soit pour tout $u \in \mathcal{H}_1$

$$\|Tu\|_{\mathcal{H}_2} = \frac{\|u\|_{\mathcal{H}_1}}{\alpha} \left\| T \alpha \frac{u}{\|u\|_{\mathcal{H}_1}} \right\|_{\mathcal{H}_2} \leq \frac{\varepsilon}{\alpha} \|u\|_{\mathcal{H}_1},$$

qui implique bien que T est borné.

Exemple 8.4.1 Donnons quelques exemples d'opérateurs bornés. En dimension finie ($\mathcal{H} = \mathbb{R}^n$), tout opérateur linéaire est borné et

$$\|T\| = \max_{1 \leq i \leq N} \|Te_i\|,$$

où $(e_1, \dots, e_N)$ est la base hilbertienne de l'espace.

Dans $\mathcal{H} = L^2(\Omega)$, l'opérateur de multiplication par une fonction $\kappa \in L^\infty(\Omega)$, $T : \mathcal{H} \ni v \mapsto \kappa v \in \mathcal{H}$ est un opérateur borné avec

$$\|T\| = \|\kappa\|_{L^\infty(\Omega)}.$$

Dans $\mathcal{H} = H^1(\Omega)$, l'opérateur de dérivation n'est pas borné, mais en prenant $\mathcal{H} = L^2(\Omega)$ l'opérateur solution $T : \mathcal{H} \ni f \mapsto u \in H^1(\Omega) \subset \mathcal{H}$ où $u \in H^1(\Omega)$ est l'unique solution de $(\mathcal{P}_D)$ ou $(\mathcal{P}_N)$ est borné.

8.4.2 Opérateurs auto-adjoints

Commençons par définir la notion d'adjoint d'un opérateur linéaire borné.

Définition 8.4.2 (Adjoint d'un opérateur borné) *Soit T un opérateur linéaire borné sur un espace de Hilbert, noté $\mathcal{H}$. L'opérateur adjoint de T noté T^* est défini par*

$$\forall u, v \in \mathcal{H}, \quad (Tu, v)_{\mathcal{H}} = (u, T^*v)_{\mathcal{H}}$$

Il est possible de définir l'adjoint d'opérateurs non bornés, mais cela n'est pas utile pour ce cours.

Définition 8.4.3 (Opérateur auto-adjoint) *Soit T un opérateur linéaire borné sur un espace de Hilbert, noté $\mathcal{H}$. On dit que T est auto-adjoint si $T^* = T$.*

Exemple 8.4.2 Donnons quelques exemples d'opérateurs auto-adjoints.

En dimension finie ($\mathcal{H} = \mathbb{R}^n$), toute matrice $\mathbb{A}$ symétrique ($\mathbb{A}^{ij} = \mathbb{A}^{ji}$) est auto-adjointe. Le caractère auto-adjoint étend celui de la symétrie en dimension finie.

Dans $\mathcal{H} = L^2(\Omega)$, l'opérateur de multiplication par une fonction $\kappa \in L^\infty(\Omega)$, $T : \mathcal{H} \ni v \mapsto \kappa v \in \mathcal{H}$ est un opérateur auto-adjoint.

Toujours dans $\mathcal{H} = L^2(\Omega)$ l'opérateur solution $T : \mathcal{H} \ni f \mapsto u \in H^1(\Omega) \subset \mathcal{H}$ où $u \in H^1(\Omega)$ est l'unique solution de $(\mathcal{P}_D)$ ou $(\mathcal{P}_N)$ est autoadjoint. En effet,

$$\forall (f, g) \in L^2(\Omega), \quad (Tf, g)_{\mathcal{H}} = a(Tg, Tf) = a(Tf, Tg) = (f, Tg)_{\mathcal{H}}$$

avec a la forme bilinéaire associée à $(\mathcal{P}_D)$ ou $(\mathcal{P}_N)$ qui est symétrique.

8.4.3 Opérateurs compacts

Passons maintenant aux questions de compacité en commençant par un rappel essentiel.

Proposition 8.4.1 *Soit $\mathcal{H}$ un espace de Hilbert et $B_{\mathcal{H}} := \{v \in \mathcal{H}, \|v\| \leq 1\}$ sa boule unité fermée. Si $B_{\mathcal{H}}$ est compacte alors $\mathcal{H}$ est de dimension finie.*

Nous pouvons maintenant définir la notion d'opérateur compact.

Définition 8.4.4 Soient $\mathcal{H}_1$ et $\mathcal{H}_2$ deux espaces de Hilbert. L'opérateur $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ est dit compact si, de toute suite bornée, on peut extraire une sous-suite dont l'image par T converge dans $\mathcal{H}_2$.

On vérifiera aisément que la définition précédente est équivalente à la définition suivante : l'image de la boule unité par un opérateur compact T est compacte.

Proposition 8.4.2 Dans un espace de Hilbert $\mathcal{H}$, de toute suite bornée on peut extraire une sous-suite faiblement convergente.

Proposition 8.4.3 Si un opérateur $T : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ est compact alors pour toute suite $(u_n)_{n \in \mathbb{N}}$ qui converge faiblement dans $\mathcal{H}_1$, $(Tu_n)_{n \in \mathbb{N}}$ converge fortement dans $\mathcal{H}_2$:

$$u_n \xrightarrow[n \rightarrow \infty]{\mathcal{H}_1} u \quad \Rightarrow \quad Tu_n \xrightarrow[n \rightarrow \infty]{\mathcal{H}_2} Tu.$$

Théorème 8.4.4 Soient $\mathcal{H}$ un espace de Hilbert réel de dimension infinie et $T : \mathcal{H} \rightarrow \mathcal{H}$ un opérateur compact, auto-adjoint, injectif et positif⁸.

- Les valeurs propres de T sont toutes de multiplicité finie et constituent une suite de réels positifs qui tend vers 0.
- Il existe une base de hilbertienne $(u_n)_{n \in \mathbb{N}^*}$ de $\mathcal{H}$ formée des fonctions propres de T .

8: un opérateur T est positif si et seulement si

$$\forall u \in \mathcal{H}, \quad (Tu, u)_{\mathcal{H}} \geq 0.$$

Démonstration. (i) Soit μ une valeur propre et $u \neq 0$ une fonction propre associée, alors

$$(Tu, u)_{\mathcal{H}} = \mu(u, u)_{\mathcal{H}} \quad \Rightarrow \quad \mu = \frac{(Tu, u)_{\mathcal{H}}}{\|u\|_{\mathcal{H}}^2} \geq 0.$$

(ii) Montrons que l'ensemble des valeurs propres n'est pas vide. Puisque

$$\forall u \in \mathcal{H}, \quad \|u\|_{\mathcal{H}} = 1, \quad (Tu, u)_{\mathcal{H}} \leq \|T\|_{\mathcal{L}(\mathcal{H})},$$

on peut introduire

$$\bar{\mu} = \sup_{\substack{u \in \mathcal{H} \\ \|u\|_{\mathcal{H}} = 1}} (Tu, u)_{\mathcal{H}} > 0$$

dont nous allons montrer qu'elle est une valeur propre de T . Tout d'abord, nous remarquons que

$$\bar{\mu} \leq \sup_{\substack{(u,v) \in \mathcal{H} \times \mathcal{H} \\ \|u\|_{\mathcal{H}} = \|v\|_{\mathcal{H}} = 1}} (Tu, v)_{\mathcal{H}} = \|T\|_{\mathcal{L}(\mathcal{H})}.$$

De plus, par l'inégalité de Cauchy-Schwarz appliquée à $(u, v) \mapsto (Tu, v)_{\mathcal{H}}$ qui est une forme bilinéaire et positive (puisque T est auto-adjoint et positif) :

$$(Tu, v)_{\mathcal{H}} \leq (Tu, u)_{\mathcal{H}}^{\frac{1}{2}} (Tv, v)_{\mathcal{H}}^{\frac{1}{2}},$$

donc en prenant les bornes supérieures en u et v , on a $\|T\|_{\mathcal{L}(\mathcal{H})} \leq \bar{\mu}$ donc $\bar{\mu} = \|T\|_{\mathcal{L}(\mathcal{H})}$. Par définition de $\bar{\mu}$, il existe une suite $(v_n)_{n \in \mathbb{N}^*}$ telle

que $\|v_n\|_{\mathcal{H}} = 1$ et $\lim_{n \rightarrow \infty} (Tv_n, v_n)_{\mathcal{H}} = \bar{\mu}$. T étant compact, on peut considérer à une extraction près que la suite $(Tv_n)_{n \in \mathbb{N}^*}$ converge vers u . Puisque $\|Tv_n\|_{\mathcal{H}} \leq \|T\|_{\mathcal{L}(\mathcal{H})} \|v_n\|_{\mathcal{H}} = \bar{\mu}$, on a alors

$$\begin{aligned} \|Tv_n - \bar{\mu}v_n\|_{\mathcal{H}}^2 &= \|Tv_n\|_{\mathcal{H}}^2 + \bar{\mu}^2 \|v_n\|_{\mathcal{H}}^2 - 2\bar{\mu}(Tv_n, v_n)_{\mathcal{H}} \\ &\leq 2(\bar{\mu}^2 - \bar{\mu}(Tv_n, v_n)_{\mathcal{H}}) \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Donc $(\bar{\mu}v_n)_{n \in \mathbb{N}^*}$ converge également vers u et finalement puisque T est borné

$$u = \lim_{n \rightarrow \infty} Tv_n = \bar{\mu}^{-1}Tu,$$

qui prouve que $\bar{\mu}$ est bien une valeur propre de T .

(iii) Vérifions alors que les sous-espaces propres $\mathcal{V}_{\mu}$ sont de dimension finie et orthogonaux entre eux. Comme T est compacte, $T|_{\mathcal{V}_{\mu}}$ l'est et l'image de la boule unité de $\mathcal{V}_{\mu}$ est donc compacte. Comme $T|_{\mathcal{V}_{\mu}} = \mu \text{Id}$, ceci implique que la boule unité de ce sous espace est compacte et donc $\mathcal{V}_{\mu}$ est de dimension finie par la proposition 8.4.1. De plus, soient deux valeurs propres différentes μ et μ' et deux fonctions propres respectivement associées u et u' , on a alors puisque T est auto-adjoint,

$$\begin{aligned} \mu(u, u')_{\mathcal{H}} &= (Tu, u')_{\mathcal{H}} = (u, Tu')_{\mathcal{H}} = \mu'(u, u')_{\mathcal{H}} \\ &\Rightarrow (u, u')_{\mathcal{H}} = 0 \text{ si } \mu \neq \mu'. \end{aligned}$$

Donc les sous-espaces propres de T associées à deux valeurs propres distinctes sont orthogonaux.

(iv) Montrons maintenant que quand $\mathcal{H}$ est de dimension infinie, les valeurs propres ne peuvent que s'accumuler en 0. On raisonne par l'absurde en considérant $(\mu_n)_{n \in \mathbb{N}^*}$ une suite de valeurs propres distinctes qui s'accumule en $\mu \neq 0$ et $(u_n)_{n \in \mathbb{N}^*}$ une suite des fonctions propres associées de norme 1. Quitte à extraire une sous-suite, la suite des $(Tu_n)_{n \in \mathbb{N}^*}$ converge. Comme la limite de $(\mu_n)_{n \in \mathbb{N}^*}$ est non nulle, pour n assez grand, $\mu_n \neq 0$. Or $\frac{1}{\mu_n}Tu_n = u_n$ donc la suite des $(u_n)_{n \in \mathbb{N}^*}$ converge. Cependant, les u_n sont tous orthogonaux deux à deux donc à distance $\sqrt{2}$ les uns des autres. On aboutit donc à une contradiction.

(v) Pour conclure, il nous reste à montrer que les fonctions propres forment une base hilbertienne. On introduit $\mathcal{G}$, le sous-espace engendré par les fonctions propres. Par construction $\mathcal{G}$ est stable par T donc $\mathcal{G}^{\perp}$ est stable par T . Soit $\tilde{T} = T|_{\mathcal{G}^{\perp}}$ qui est toujours un opérateur compact. Donc $\|\tilde{T}\|$ est une valeur propre de $\tilde{T}$. Si $\tilde{T}$ n'est pas nul alors $\tilde{T}$ est aussi une valeur propre de T ce qui serait une contradiction. On en déduit que $\mathcal{G}^{\perp} = 0$ ce qui conclut la preuve. $\square$

Theorème 8.4.5 Soit $\mathcal{V}$ et $\mathcal{H}$ deux espaces de Hilbert où $\mathcal{V}$ s'injecte de manière compacte⁹ dans $\mathcal{H}$. Soit a une forme bilinéaire continue symétrique et coercive de $\mathcal{V} \times \mathcal{V}$ et $A \in \mathcal{L}(\mathcal{V}, \mathcal{V})$ l'opérateur associé défini dans (8.5), alors

- le problème aux valeurs propres (8.7) admet des valeurs propres qui sont toutes de multiplicité finie et constituent une suite de réels positifs $(\lambda_n)_{n \in \mathbb{N}^*}$ qui tend vers $+\infty$.
- Il existe une base hilbertienne $(u_n)_{n \in \mathbb{N}^*}$ de $\mathcal{H}$ formée des fonctions

9: Injection compacte signifie que l'application identité $\iota_{\mathcal{V}} : \mathcal{V} \rightarrow \mathcal{H}$ est compacte, c'est à dire que de toute suite d'éléments de $\mathcal{V}$ on peut extraire une sous suite qui converge dans $\mathcal{H}$.

propres de (8.7).

Démonstration. Soit $u \in \mathcal{H}$, l'application linéaire $\ell_u : \mathcal{V} \ni v \rightarrow (u, v)_{\mathcal{H}} \in \mathbb{R}$ est continue puisque par injection (compacte donc) continue de $\mathcal{V}$ dans $\mathcal{H}$, il existe une constante $C \in \mathbb{R}_+^*$ telle que

$$|(u, v)_{\mathcal{H}}| \leq \|u\|_{\mathcal{H}} \|v\|_{\mathcal{H}} \leq C \|u\|_{\mathcal{H}} \|v\|_{\mathcal{V}}.$$

D'après le théorème de Lax-Milgram, il existe donc un unique élément w solution de

$$a(w, v) = \ell_u(v), \quad v \in \mathcal{V}.$$

Par linéarité de ℓ_u par rapport à u , il existe un opérateur linéaire T tel que $w = Tu \in \mathcal{V} \subset \mathcal{H}$. La continuité par rapport aux données dans le théorème de Lax-Milgram implique de plus que

$$\|Tu\|_{\mathcal{V}} = \|w\|_{\mathcal{V}} \leq \frac{C}{\nu} \|u\|_{\mathcal{H}}.$$

L'application T en tant qu'application de $\mathcal{H}$ dans $\mathcal{V}$ est donc continue. Elle l'est donc aussi en tant qu'application de $\mathcal{H}$ dans $\mathcal{H}$. Mais mieux elle est compacte puisque dans ce cas c'est la composée d'un opérateur continu de $\mathcal{H}$ dans $\mathcal{V}$ avec l'injection compacte de $\mathcal{V}$ dans $\mathcal{H}$.

L'opérateur T est donc compact. Il est également auto-adjoint car a est symétrique

$$\begin{aligned} \forall u, v \in \mathcal{H}, \quad (Tu, v)_{\mathcal{H}} &= \ell_v(Tu) = a(Tv, Tu) \\ &= a(Tu, Tv) = \ell_u(Tv) = (u, Tv)_{\mathcal{H}}. \end{aligned}$$

Il est injectif par unicité de la solution dans le théorème de Lax-Milgram. Il est de plus clairement positif, car

$$(u, Tu)_{\mathcal{H}} = \ell_u(Tu) = a(Tu, Tu) \geq 0.$$

On peut donc appliquer le théorème 8.4.4 à T qui fournit une suite de valeurs propres $(\mu_n)_{n \in \mathbb{N}^*}$ strictement positives qui tendent vers 0 et de fonctions propres orthonormalisées associées $(u_n)_{n \in \mathbb{N}^*}$, telles que

$$Tu_n = \mu_n u_n.$$

On vérifie alors directement par définition (8.5) de A que

$$ATu_n = u_n = \frac{1}{\mu_n} Tu_n.$$

Les $(\lambda_n = \mu_n^{-1})_{n \in \mathbb{N}^*}$ sont donc des valeurs propres et Tu_n des fonctions propres associées. Puisque Tu_n est proportionnel à u_n alors u_n est aussi fonction propre. Réciproquement tout élément propre (u_n, λ_n) de A est aussi élément propre de T . Ainsi les $(u_n, \lambda_n = \mu_n^{-1})$ sont donc les seuls éléments propres de A . En particulier les λ_n sont positifs et de multiplicité finie et tendent vers $+\infty$.

Enfin puisque les $(u_n)_{n \in \mathbb{N}}$ sont déjà orthonormalisées par rapport au produit scalaire dans $\mathcal{H}$, elles forment la base hilbertienne que nous recherchons. $\square$

Nous pouvons maintenant caractériser les espaces $\mathcal{H}$ et $\mathcal{V}$ à partir de la base hilbertienne $(u_n)_{n \in \mathbb{N}^*}$ de vecteurs propres de A et des valeurs propres $(\lambda_n)_{n \in \mathbb{N}^*}$ associées. Par définition d'une base hilbertienne, on a

$$\mathcal{H} \equiv \left\{ \sum_{n \geq 1} \alpha_n u_n \text{ tel que } \sum_{n \geq 1} \alpha_n^2 < \infty \right\}. \quad (8.8)$$

Nous pouvons écrire

$$\forall u \in \mathcal{H}, \quad u = \sum_{n \geq 1} (u, u_n)_{\mathcal{H}} u_n \quad \text{avec} \quad \|u\|_{\mathcal{H}}^2 = \sum_{n \geq 1} |(u, u_n)_{\mathcal{H}}|^2 < \infty.$$

On introduit

$$P_N(u) =: \sum_{n=1}^N (u, u_n)_{\mathcal{H}} u_n \xrightarrow{N \rightarrow \infty} u,$$

le projecteur orthogonal de u sur $\text{Vect}((u_n)_{0 \leq n \leq N})$.

Soit maintenant $u \in \mathcal{V}$, on a

$$a(P_N(u), P_N(u)) = \sum_{n=1}^N \lambda_n |(u, u_n)_{\mathcal{H}}|^2,$$

qui est une suite croissante. Par orthogonalité des $(u_n)_{n \in \mathbb{N}^*}$, on a

$$a(u, u) - a(P_N(u), P_N(u)) = a(u - P_N(u), u - P_N(u)) \geq 0,$$

la suite $a(P_N(u), P_N(u))$ est donc majorée par $a(u, u)$. Par conséquent, elle converge. Comme

$$u - P_N(u) \xrightarrow{N \rightarrow \infty} 0 \quad \Rightarrow \quad a(u - P_N(u), u - P_N(u)) \xrightarrow{N \rightarrow \infty} 0$$

ce qui implique que $a(P_N(u), P_N(u))$ converge vers $a(u, u)$ et

$$a(u, u) = \sum_{n \geq 1} \lambda_n (u, u_n)_{\mathcal{H}}^2 < \infty. \quad (8.9)$$

On en déduit que

$$\mathcal{V} \subset \left\{ u = \sum_{n \geq 1} \alpha_n u_n \text{ tel que } \sum_{n \geq 1} \lambda_n \alpha_n^2 < \infty \right\}.$$

Réciproquement, soit la suite

$$(\alpha_n)_{n \in \mathbb{N}^*} \text{ telle que } \sum_{n \geq 1} \lambda_n \alpha_n^2 < \infty,$$

alors pour tout N , si on note

$$u^{(N)} := \sum_{n=1}^N \alpha_n u_n \in \mathcal{V},$$

on a

$$\forall p < q, \quad \alpha \|u^{(p)} - u^{(q)}\|_{\mathcal{V}}^2 \leq a(u^{(p)} - u^{(q)}, u^{(p)} - u^{(q)}) = \sum_{n=p+1}^q \lambda_n \alpha_n^2,$$

où α est la constante de coercivité de a . On en déduit que $(u^{(N)})_{N \in \mathbb{N}^*}$ est une suite de Cauchy dans $\mathcal{V}$ donc elle converge dans $\mathcal{V}$. Nous venons

donc de prouver la caractérisation de $\mathcal{V}$:

$$\mathcal{V} \equiv \left\{ u = \sum_{n \geq 1} \alpha_n u_n \text{ tel que } \sum_{n \geq 1} \lambda_n \alpha_n^2 < \infty \right\}. \quad (8.10)$$

Exemple 8.4.3 On considère le problème aux valeurs propres (8.1) où nous pouvons appliquer le théorème (8.4.5) avec

$$\mathcal{H} = L^2(\Omega), \quad \mathcal{V} = H_0^1(\Omega) \quad \text{et} \quad a(w, v) = \int_{\Omega} \underline{\nabla} w \cdot \underline{\nabla} v \, d\Omega.$$

On note les valeurs propres $(\lambda_n)_{n \in \mathbb{N}^*}$ et les fonctions propres qui forment une base hilbertienne $(u_n)_{n \in \mathbb{N}^*}$. On a

$$L^2(\Omega) \equiv \left\{ \sum_{n \in \mathbb{N}^*} \alpha_n u_n, \quad \sum_{n \in \mathbb{N}^*} |\alpha_n|^2 < +\infty \right\}$$

avec

$$\forall u \in L^2(\Omega), \quad \|u\|_{L^2(\Omega)}^2 = \sum_{n \geq 1} |(u, u_n)_{L^2(\Omega)}|^2,$$

et

$$H_0^1(\Omega) \equiv \left\{ \sum_{n \in \mathbb{N}^*} \alpha_n u_n, \quad \sum_{n \in \mathbb{N}^*} \lambda_n |\alpha_n|^2 < +\infty \right\},$$

avec

$$\forall u \in H_0^1(\Omega), \quad \|u\|_{H^1(\Omega)}^2 = \sum_{n \geq 1} (1 + \lambda_n) |(u, u_n)_{L^2(\Omega)}|^2,$$

Dans le cas où $\Omega =]0, L[$, les valeurs propres $(\lambda_n)_{n \in \mathbb{N}^*}$ et les fonctions propres $(u_n)_{n \in \mathbb{N}^*}$ sont données par

$$\lambda_n = \frac{n^2 \pi^2}{L^2}, \quad u_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right), \quad n \in \mathbb{N}^*,$$

Ainsi, on peut caractériser pour tout $u \in L^2(]0, L[)$

$$\|u\|_{L^2(]0, L[)}^2 = \frac{2}{L} \sum_{n \geq 1} \left| \int_0^L u(x) \sin\left(\frac{n\pi x}{L}\right) dx \right|^2,$$

et pour tout $u \in H_0^1(]0, L[)$

$$\|u\|_{H^1(]0, L[)}^2 = \frac{2}{L} \sum_{n \geq 1} \left(1 + \frac{n^2 \pi^2}{L^2} \right) \left| \int_0^L u(x) \sin\left(\frac{n\pi x}{L}\right) dx \right|^2.$$

Nous retrouvons la décomposition en séries de Fourier.

8.4.4 Principe du Min-Max

Par analogie avec la dimension finie, nous introduisons le quotient de Rayleigh associé à une forme bilinéaire sur $\mathcal{V}$:

$$\forall v \in \mathcal{V}^*, \quad R_a(v) := \frac{a(v, v)}{\|v\|_{\mathcal{H}}^2}. \quad (8.11)$$

Proposition 8.4.6 (Courant-Fisher) Soient $\mathcal{V}$ et $\mathcal{H}$ deux espaces de Hilbert réels de dimension infinie. On suppose que $\mathcal{V} \subset \mathcal{H}$ avec injection compacte. Soit $a(\cdot, \cdot)$ une forme bilinéaire symétrique continue et coercive sur $\mathcal{V}$. On note $(\lambda_n)_{n \geq 1}$ la suite croissante des valeurs propres du problème aux valeurs propres (8.4) comptées avec leur multiplicité. Alors, pour tout $n \in \mathbb{N}^*$, la n -ème valeur propre est donnée par

$$\lambda_n = \min_{\mathcal{W} \in \mathcal{E}_n} \left(\max_{v \in \mathcal{W}^*} R_a(v) \right) = \max_{\mathcal{W} \in \mathcal{E}_{n-1}} \left(\min_{v \in \mathcal{W}^{\perp*}} R_a(v) \right), \quad (8.12)$$

où $\mathcal{E}_n$ est l'ensemble des sous-espaces vectoriels de dimension n de $\mathcal{V}$ et R_a est le quotient de Rayleigh de a , défini dans (8.11). En particulier, la première valeur propre vérifie

$$\lambda_1 = \min_{v \in \mathcal{V}^*} R_a(v), \quad (8.13)$$

et tout $\bar{v}$ minimum dans (8.13) est un vecteur propre associé à λ_1 .

Démonstration. Soit $(u_n)_{n \in \mathbb{N}^*}$ la base hilbertienne de $\mathcal{H}$ formée des vecteurs propres de (8.4). D'après (8.9), on peut réécrire le quotient de Rayleigh

$$R_a(v) = \frac{\sum_{n \geq 1} \lambda_n \alpha_n^2}{\sum_{n \geq 1} \alpha_n^2},$$

ce qui démontre immédiatement le résultat pour la première valeur propre.

Introduisons le sous-espace $\mathcal{W}_n \in \mathcal{E}_n$ engendré par $(u_1, \dots, u_n)$. On a

$$\forall v \in \mathcal{W}_n^*, \quad R_a(v) = \frac{\sum_{k=1}^n \lambda_k \alpha_k^2}{\sum_{k=1}^n \alpha_k^2},$$

et

$$\forall v \in \mathcal{W}_{n-1}^{\perp*}, \quad R_a(v) = \frac{\sum_{k \geq n} \lambda_k \alpha_k^2}{\sum_{k \geq n} \alpha_k^2},$$

d'où l'on déduit

$$\lambda_n = \max_{v \in \mathcal{W}_n^*} R_a(v) = \min_{v \in \mathcal{W}_{n-1}^{\perp*}} R_a(v).$$

Soit $\mathcal{W}$ un sous-espace quelconque dans $\mathcal{E}_n$. Comme $\mathcal{W}$ est de dimension n et $\mathcal{W}_{n-1}$ de dimension $n-1$, l'intersection $\mathcal{W} \cap \mathcal{W}_{n-1}^{\perp}$ n'est pas réduite à $\{0\}$. Par conséquent,

$$\begin{aligned} \max_{v \in \mathcal{W}^*} R_a(v) &\geq \max_{v \in \mathcal{W} \cap \mathcal{W}_{n-1}^{\perp*}} R_a(v) \\ &\geq \min_{v \in \mathcal{W} \cap \mathcal{W}_{n-1}^{\perp*}} R_a(v) \geq \min_{v \in \mathcal{W}_{n-1}^{\perp*}} R_a(v) = \lambda_n, \end{aligned}$$

ce qui prouve la première égalité dans (8.12). De même, si $\mathcal{W}$ est un sous-espace dans $\mathcal{E}_{n-1}$, alors $\mathcal{W}^{\perp} \cap \mathcal{W}_n^*$ n'est pas réduit à $\{0\}$, et

$$\begin{aligned} \min_{v \in \mathcal{W}^{\perp*}} R_a(v) &\leq \min_{v \in \mathcal{W}^{\perp} \cap \mathcal{W}_n^*} R_a(v) \\ &\leq \max_{v \in \mathcal{W}^{\perp} \cap \mathcal{W}_n^*} R_a(v) \leq \max_{v \in \mathcal{W}_n^*} R_a(v) = \lambda_n, \end{aligned}$$

ce qui prouve la deuxième égalité dans (8.12).

Soit maintenant u un minimum dans (8.13). Pour $v \in \mathcal{V}$, on introduit la fonction $f : \mathbb{R} \ni t \mapsto R(u + tv) \in \mathbb{R}$ qui admet un minimum en $t = 0$. Par conséquent sa dérivée s'annule en $t = 0$. Puisque $f(0) = \lambda_1$, un simple calcul montre que

$$f'(0) = 2 \frac{a(u, v) - \lambda_1(u, v)_{\mathcal{H}}}{\|u\|_{\mathcal{H}}^2}.$$

Comme v est quelconque dans $\mathcal{V}$, la condition $f'(0) = 0$ n'est rien d'autre que la formulation variationnelle (8.4), c'est-à-dire que u est un vecteur propre associé à la valeur propre λ_1 . $\square$

Le principe du Min-Max est particulièrement utile lorsque l'on souhaite comparer des valeurs propres de différents problèmes. Par exemple, soit $a(\cdot, \cdot)$ et $\tilde{a}(\cdot, \cdot)$ deux formes bilinéaires continues symétriques et coercives agissant sur le même espace $\mathcal{V}$ d'injection compacte dans $\mathcal{H}$. On note $(\lambda_n)_{n \in \mathbb{N}^*}$ et $(\tilde{\lambda}_n)_{n \in \mathbb{N}^*}$ les valeurs propres respectives de a et $\tilde{a}$, les deux suites étant ordonnées par valeur croissantes. Si de plus on a

$$\forall u \in \mathcal{V}, \quad a(u, u) \leq \tilde{a}(u, u),$$

alors on déduit du principe du Min-Max que

$$\forall n \in \mathbb{N}^*, \quad \lambda_n \leq \tilde{\lambda}_n.$$

Si on considère maintenant une seule forme bilinéaire $a(\cdot, \cdot)$ définie sur deux espaces $\mathcal{V}$ et $\tilde{\mathcal{V}}$. On note $(\lambda_n)_{n \in \mathbb{N}^*}$ et $(\tilde{\lambda}_n)_{n \in \mathbb{N}^*}$ les valeurs propres de a définie respectivement sur $\mathcal{V}$ et $\tilde{\mathcal{V}}$. Alors si $\mathcal{V}$ et $\tilde{\mathcal{V}}$ sont tels que

$$\mathcal{V} \subset \tilde{\mathcal{V}},$$

on déduit

$$\forall n \in \mathbb{N}^*, \quad \lambda_n \geq \tilde{\lambda}_n. \quad (8.14)$$

Ce raisonnement nous sera utile notamment pour l'approximation numérique des valeurs propres.

8.5 Approximation numérique des problèmes spectraux

8.5.1 Existence des éléments propres discrets

On va considérer une approximation interne de la formulation variationnelle (8.4). Étant donné un sous-espace $\mathcal{V}_h$ de l'espace de Hilbert $\mathcal{V}$, de dimension finie, on cherche les solutions $(\lambda_h, u_h) \in \mathbb{R} \times \mathcal{V}_h$ de

$$a(u_h, v_h) = \lambda_h(u_h, v_h)_{\mathcal{H}} \quad \forall v_h \in \mathcal{V}_h. \quad (8.15)$$

Typiquement, $\mathcal{V}_h$ est un espace d'éléments finis, et $\mathcal{H}$ est toujours l'espace $L^2(\Omega)$. La résolution de l'approximation interne (8.15) est facile comme le montre le lemme suivant.

Lemme 8.5.1 On se place sous les hypothèses du corollaire 8.4.5. Alors les valeurs propres de (8.15) forment une suite croissante finie

$$0 < \lambda_{1,h} \leq \dots \leq \lambda_{N,h} \quad \text{avec } N = \dim \mathcal{V}_h,$$

et il existe une base de $\mathcal{V}_h$, orthonormale dans $\mathcal{H}$, $(u_{k,h})_{1 \leq k \leq N}$ de vecteurs propres associés, c'est-à-dire que

$$u_{k,h} \in \mathcal{V}_h, \quad \text{et } a(u_{k,h}, v_h) = \lambda_{k,h} (u_{k,h}, v_h)_{\mathcal{H}} \quad \forall v_h \in \mathcal{V}_h.$$

Démonstration. Ce lemme peut être considéré comme une variante¹⁰ évidente du corollaire 8.4.5. Néanmoins, nous en donnons une démonstration différente, purement algébrique et élémentaire. Soit $(\phi_i)_{1 \leq i \leq N}$ une base¹¹ de $\mathcal{V}_h$. On cherche u_h solution de (8.15) sous la forme

$$u_h(x) = \sum_{i=1}^N u_h^i \phi_i(x).$$

Introduisant la *matrice de masse* $\mathbb{M}_h$ définie par

$$\mathbb{M}_h^{ij} = (\phi_j, \phi_i)_{\mathcal{H}} \quad 1 \leq i, j \leq N,$$

et la matrice¹² $\mathbb{A}_h$ définie par

$$\mathbb{A}_h^{ij} = a(\phi_j, \phi_i) \quad 1 \leq i, j \leq N, \quad (8.16)$$

le problème (8.15) est équivalent à trouver $(\lambda_h, \vec{u}_h) \in \mathbb{R} \times \mathbb{R}^N$ solution de

$$\mathbb{A}_h \vec{u}_h = \lambda_h \mathbb{M}_h \vec{u}_h. \quad (8.17)$$

On vérifie immédiatement que les matrices $\mathbb{M}_h$ et $\mathbb{A}_h$ sont symétriques et définies positives. Le système (8.17) est un problème aux valeurs propres dit *généralisé* du fait de la présence de la matrice de masse $\mathbb{M}_h$ à droite de l'égalité. Le théorème de réduction simultanée (voir par exemple [2]) affirme qu'il existe une matrice inversible $\mathbb{P}_h$ telle que

$$\mathbb{M}_h = \mathbb{P}_h \mathbb{P}_h^T, \quad \text{et } \mathbb{A}_h = \mathbb{P}_h \text{diag}(\lambda_{1,h}, \dots, \lambda_{N,h}) \mathbb{P}_h^T.$$

Par conséquent, les solutions de (8.17) sont les valeurs propres $(\lambda_{k,h})$ et les vecteurs propres $(\vec{u}_{k,h})_{1 \leq k \leq N}$ qui sont les vecteurs colonnes de l'inverse de $\mathbb{P}_h^T$. Ces vecteurs colonnes forment donc une base, orthogonale pour $\mathbb{A}_h$ et orthonormale pour $\mathbb{M}_h$. Finalement, les vecteurs $\vec{u}_{k,h}$ sont simplement les vecteurs des coordonnées dans la base $(\phi_i)_{1 \leq i \leq N}$ des fonctions $u_{k,h}$ qui forment une base orthonormale de $\mathcal{V}_h$ pour le produit scalaire de $\mathcal{H}$. $\square$

Remarque 8.5.1 Pour calculer les valeurs et vecteurs propres du problème spectral (8.17) il faut, en général, commencer par calculer la factorisation dite *de Cholesky*¹³ de la matrice de masse $\mathbb{M}_h = \mathbb{L}_h \mathbb{L}_h^T$, pour se ramener au cas classique

$$\tilde{\mathbb{A}}_h \vec{\tilde{u}}_h = \lambda_h \vec{\tilde{u}}_h \quad \text{avec } \tilde{\mathbb{A}}_h = \mathbb{L}_h^{-1} \mathbb{A}_h \mathbb{L}_h^{-T} \quad \text{et } \vec{\tilde{u}}_h = \mathbb{L}_h^T \vec{u}_h,$$

pour lequel on dispose d'algorithmes de calcul de valeurs et vecteurs

10: à la différence près qu'en dimension finie il existe un nombre fini de valeurs propres

11: par exemple une base éléments-finis $\mathbb{P}_1$

12: Dans le cas des exemples rencontrés dans cet ouvrage la matrice $\mathbb{A}_h$ est typiquement une combinaison linéaire de la matrice de masse et de la matrice de rigidité

13: André-Louis Cholesky (1875-1918) polytechnicien et officier français, ingénieur topographe et géodésien. Il est célèbre pour sa méthode de résolution des systèmes d'équations linéaires. Il est décédé des suites de ses blessures sur le champ de bataille.

14: méthode dite du *mass lumping* en anglais

propres (voir la section 8.5.3).

On peut éviter de construire la matrice $\tilde{\mathbb{K}}_h$ et faire l'économie de la factorisation de Cholesky de $\mathbb{M}_h$ si on utilise une formule de quadrature pour évaluer les coefficients de la matrice $\mathbb{M}_h$ qui la rende *diagonale*. Ce procédé d'intégration numérique est appelé *condensation de masse*¹⁴ et est fréquemment utilisé. Par exemple, si on utilise une formule de quadrature qui utilise uniquement les valeurs aux nœuds d'une fonction pour calculer une intégrale, on voit facilement que la matrice de masse $\mathbb{M}_h$ ainsi obtenue est diagonale.

8.5.2 Analyse d'erreur

La question désormais est de savoir comment les valeurs propres du problème discret approchent celles du problème continu, la même question se posant aussi pour les vecteurs propres. Nous commençons par un premier résultat élémentaire à partir du principe du Min-Max.

Lemme 8.5.2 *Pour tout $n \in \llbracket 1, N \rrbracket$, on a $\lambda_{n,h} \geq \lambda_n$.*

Démonstration. Ceci résulte directement du raisonnement ayant conduit à (8.14). En effet $\mathcal{V}_h \subset \mathcal{V}$ donc par le principe du Min-Max on a bien $\lambda_{n,h} \geq \lambda_n$. $\square$

Nous allons maintenant procéder à l'analyse d'erreur pour le premier couple propre. Pour cela, nous allons nous concentrer sur le problème modèle (8.1) pour lequel tous les résultats du chapitre 7 s'appliquent. On rappelle que dans ce cas,

$$\mathcal{H} = L^2(\Omega), \quad \mathcal{V} = H_0^1(\Omega), \quad \text{et } a(w, v) = \int_{\Omega} \nabla w \cdot \nabla v \, d\Omega.$$

Le domaine Ω est polygonal et l'espace d'approximation est $\mathcal{V}_h = \mathcal{V}^{\mathbb{P}_1}$.

Proposition 8.5.3 *Supposons que la première valeur propre du problème (8.4) soit simple et que pour tout $f \in L^2(\Omega)$, le problème :*

$$\text{trouver } z \in \mathcal{V} \text{ tel que } \forall v \in \mathcal{V}, \quad a(z, v) = (f, v)_{\mathcal{H}} \quad (8.18)$$

admet une solution $z \in H^2(\Omega)$. Alors, il existe deux constantes $h_0 > 0$ et $C > 0$ indépendantes de h tels que

$$\forall h \in (0, h_0), \quad 0 \leq \lambda_{1,h} - \lambda_1 \leq Ch^2. \quad (8.19)$$

De plus pour $u_{1,h}$ choisi tel que $(u_1, u_{1,h})_{L^2(\Omega)} > 0$, on a

$$\forall h \in (0, h_0), \quad \|u_{1,h} - u_1\|_{\mathcal{V}} \leq Ch.$$

Démonstration. D'après le lemme 8.5.2, on sait que $0 \leq \lambda_{1,h} - \lambda_1$. Pour majorer l'écart, on introduit $f = \lambda_1 u_1$ (et donc la solution de (8.18) est $z = u_1$) et z_h la solution du problème discrétisé associé à (8.18) :

$$\forall v_h \in \mathcal{V}_h, \quad a(z_h, v_h) = (f, v_h)_{\mathcal{H}}.$$

On a supposé que $z \in H^2(\Omega)$ donc, d'après le théorème 7.2.18 d'Aubin-Nitsche, il existe une constante C_1 ne dépendant que de λ_1 telle que

$$\|u_1 - z_h\|_{L^2(\Omega)} \leq C_1 h^2.$$

Comme par ailleurs $\|u_1\|_{L^2(\Omega)} = 1$, on en déduit qu'il existe $h_0 > 0$ tel que pour $h < h_0$

$$\|z_h\|_{L^2(\Omega)} \geq \frac{1}{2}.$$

D'après le principe du Min-Max dans $\mathcal{V}_h$, on a

$$\lambda_{1,h} - \lambda_1 \leq \frac{a(z_h, z_h)}{\|z_h\|_{L^2(\Omega)}^2} - \lambda_1 = \frac{a(z_h, z_h) - \lambda_1 \|z_h\|_{L^2(\Omega)}^2}{\|z_h\|_{L^2(\Omega)}^2}.$$

Or par définition de z_h , on a $a(z_h, z_h) = \lambda_1(u_1, z_h)_{L^2(\Omega)}$. On en déduit

$$\lambda_{1,h} - \lambda_1 \leq \frac{\lambda_1(u_1 - z_h, z_h)_{L^2(\Omega)}}{\|z_h\|_{L^2(\Omega)}^2} \leq \frac{\lambda_1}{\|z_h\|_{L^2(\Omega)}} \|u_1 - z_h\|_{L^2(\Omega)} \leq 2C_1 \lambda_1 h^2, \quad (8.20)$$

qui justifie (8.19).

Soit maintenant $u_{1,h}$ une fonction propre discrète associée à $\lambda_{1,h}$ de norme $\|u_{1,h}\|_{L^2(\Omega)} = 1$, ce qui implique $a(u_{1,h}, u_{1,h}) = \lambda_{1,h}$. On choisit $u_{1,h}$ tel que $(u_1, u_{1,h})_{L^2(\Omega)} \geq 0$. Notre objectif est d'estimer $a(u_1 - u_{1,h}, u_1 - u_{1,h})$. Si on décompose $u_{1,h}$ sur la base hilbertienne des fonctions propres $(u_n)_{n \in \mathbb{N}^*}$, on a

$$u_{1,h} = \eta u_1 + w \text{ où } \begin{cases} \eta = (u_{1,h}, u_1)_{L^2(\Omega)}, \\ w = \sum_{n \geq 2} (u_{1,h}, u_n)_{L^2(\Omega)} u_n, \\ 1 = \eta^2 + \|w\|_{L^2(\Omega)}^2. \end{cases}$$

Par orthogonalité par rapport à la forme bilinéaire a , on a aussi

$$a(u_1 - u_{1,h}, u_1 - u_{1,h}) = (1 - \eta)^2 \lambda_1 + a(w, w) \text{ et } \lambda_{1,h} = \lambda_1 \eta^2 + a(w, w).$$

Commençons par utiliser la deuxième égalité, à laquelle on applique le principe du Min-Max sur λ_2 pour obtenir

$$\lambda_{1,h} - \lambda_1 = -(1 - \eta^2) \lambda_1 + a(w, w) \geq -(1 - \eta^2) \lambda_1 + \lambda_2 \|w\|_{L^2(\Omega)}^2 = (1 - \eta^2) (\lambda_2 - \lambda_1).$$

Puisque λ_1 est simple (on a $\lambda_1 \neq \lambda_2$) et on obtient en utilisant également (8.20)¹⁵

$$(1 - \eta^2) \leq \frac{\lambda_{1,h} - \lambda_1}{\lambda_2 - \lambda_1} \leq \frac{2C_1 \lambda_1}{\lambda_2 - \lambda_1} h^2.$$

15: À noter que si $\lambda_2 = \lambda_1$, on obtient seulement $\lambda_{1,h} - \lambda_1 \geq 0$ ce qu'on savait déjà

Donc η^2 tend vers 1 quand h tend vers 0, et ayant choisi $(u_1, u_{1,h})_{L^2(\Omega)} \geq 0$, on a donc η qui tend vers 1. On peut alors estimer

$$a(w, w) = \lambda_{1,h} - \lambda_1 \eta^2 \leq \lambda_{1,h} - \lambda_1 + \frac{2C_1 \lambda_1^2}{\lambda_2 - \lambda_1} h^2 \leq 2C_1 \frac{\lambda_2}{\lambda_2 - \lambda_1} \lambda_1 h^2.$$

où on a utilisé (8.20). Et on en déduit que

$$\begin{aligned} a(u_1 - u_{1,h}, u_1 - u_{1,h}) &= (1 - \eta)^2 \lambda_1 + a(w, w) \\ &\leq (1 - \eta^2) \lambda_1 + 2C_1 \frac{\lambda_2}{\lambda_2 - \lambda_1} \lambda_1 h^2 \\ &\leq 2C_1 \lambda_1 h^2 \frac{\lambda_1 + \lambda_2}{\lambda_2 - \lambda_1}. \end{aligned}$$

La forme bilinéaire a étant coercive de constante de coercivité ν , on en déduit finalement le résultat souhaité puisque

$$\|u_1 - u_{1,h}\|_{\mathcal{V}} \leq C' \left(\frac{\lambda_1(\lambda_1 + \lambda_2)}{(\lambda_2 - \lambda_1)} \right)^{\frac{1}{2}} h.$$

□

Le résultat précédent peut se généraliser aux valeurs propres supérieures, mais la démonstration est plus subtile, voir par exemple [3]. Si λ_n est simple et u_n est un vecteur propre tel que $\|u_n\|_{\mathcal{H}} = 1$, on montre que pour $u_{n,h}$ choisi tel que $(u_{n,h}, u_n)_{\mathcal{H}} > 0$, il existe des constantes h_0 et C_n indépendantes de h telles que

$$\forall h \in (0, h_0), \quad \|u - u_{n,h}\| \leq C_n \varepsilon_n(h),$$

où

$$\varepsilon_n(h) = \inf_{v_h \in \mathcal{V}_h} \|u_n - v_h\|_{\mathcal{V}},$$

et de même on montre qu'il existe C' indépendante de h telle que

$$0 \leq \lambda_{n,h} - \lambda_n \leq C' [\varepsilon_n(h)]^2.$$

Si maintenant la valeur propre a une certaine multiplicité, le résultat s'étend en choisissant correctement le vecteur propre et en prenant $\varepsilon(h)$ le sup sur l'espace propre associé à la valeur propre considérée.

Une remarque essentielle dans tous les cas est que même si les constantes sont indépendantes de h , elles dépendent de n . En particulier la constante

$$C_n \propto \max_{\lambda \neq \lambda_n} \frac{\lambda_n^{\frac{1}{2}}}{|\lambda - \lambda_n|^{\frac{1}{2}}},$$

qui généralise le terme $\lambda_1^{\frac{1}{2}} |\lambda_2 - \lambda_1|^{-\frac{1}{2}}$ que nous avons vu dans la preuve de la proposition 8.5.3.

Enfin, les estimations dépendent fortement de la façon dont sont approchées les fonctions propres u_n . On peut montrer (et l'exemple en 8.4.3 donné plus tôt le confirme), que plus la valeur propre est grande, plus les vecteurs propres associés oscillent vite et plus fin doit être le maillage pour bien les approcher.

8.5.3 Calcul de valeurs et vecteurs propres en pratique

Dans cette section nous expliquons comment calculer les valeurs propres et les vecteurs propres d'une matrice symétrique réelle. Pour plus de détails, nous renvoyons aux ouvrages [2] et [8].

Puisque les valeurs propres d'une matrice $\mathbb{A}$ sont les racines de son polynôme caractéristique $\det(\mathbb{A} - \lambda \mathbb{I})$, on pourrait penser naïvement que, pour les calculer, il « suffit » de factoriser son polynôme caractéristique. Cependant de telles méthodes qui donnent le résultat en un nombre fini d'opérations élémentaires n'existent pas pour des polynômes quelconques. Il n'existe en fait que des méthodes itératives pour calculer des valeurs propres. Nous nous limiterons ici à la stratégie dite de la puissance itérée.

Pour calculer la plus grande ou la plus petite valeur propre (en module) d'une matrice (et un vecteur propre associé) on peut utiliser la méthode de la puissance itérée. Une limitation de la méthode est que la valeur propre extrême que l'on calcule doit être simple¹⁶. Soit $\mathbb{A}$ une matrice symétrique réelle d'ordre n , de valeurs propres $(\lambda_1, \dots, \lambda_n)$ avec $|\lambda_n| > |\lambda_i|$ pour tout $1 \leq i \leq n-1$. La méthode de la puissance pour calculer la plus grande valeur propre λ_n et un vecteur propre associé $\vec{u}_n$ est définie par l'algorithme ci-dessous :

16: ou de multiplicité égale à 1, c'est-à-dire que la dimension du sous-espace propre correspondant est 1.

1. Initialisation : $\vec{x}_0 \in \mathbb{R}^n$ tel que $\|\vec{x}_0\| = 1$.
2. Itérations : pour $k \geq 1$
 1. $\vec{y}_k = \mathbb{A}\vec{x}_{k-1}$
 2. $\vec{x}_k = \vec{y}_k / \|\vec{y}_k\|$
 3. Si $\|\vec{x}_k - \vec{x}_{k-1}\| \leq \varepsilon$, on arrête et $|\lambda_n| \approx \|\mathbb{A}\vec{x}_k\|$ et $\vec{u}_n \approx \vec{x}_k$.

Dans le test de convergence ε est choisi assez petit. Si $\vec{\delta}_k = \vec{x}_k - \vec{x}_{k-1}$ est petit, alors $\vec{x}_k$ est un vecteur propre approché de $\mathbb{A}$ de valeur propre approchée $\|\vec{y}_k\|$ car $\mathbb{A}\vec{x}_k - \|\vec{y}_k\|\vec{x}_k = \mathbb{A}\vec{\delta}_k \approx 0$.

Proposition 8.5.4 On suppose que la matrice $\mathbb{A}$ est symétrique réelle, de valeurs propres $(\lambda_1, \dots, \lambda_n)$, associées à une base orthonormée de vecteurs propres $(\vec{e}_1, \dots, \vec{e}_n)$, et que la valeur propre de plus grand module λ_n est simple et positive, c'est-à-dire que $|\lambda_1|, \dots, |\lambda_{n-1}| < \lambda_n$. On suppose aussi que le vecteur initial $\vec{x}_0$ n'est pas orthogonal à $\vec{e}_n$. Alors la méthode de la puissance converge, c'est-à-dire que

$$\lim_{k \rightarrow +\infty} \|\vec{y}_k\| = \lambda_n, \quad \lim_{k \rightarrow +\infty} \vec{x}_k = \vec{x}_\infty \quad \text{avec} \quad \vec{x}_\infty = \pm \vec{e}_n.$$

La vitesse de convergence est proportionnelle au rapport $|\lambda_{n-1}|/|\lambda_n|$

$$\left| \|\vec{y}_k\| - \lambda_n \right| \leq C \left| \frac{\lambda_{n-1}}{\lambda_n} \right|^{2k}, \quad \|\vec{x}_k - \vec{x}_\infty\| \leq C \left| \frac{\lambda_{n-1}}{\lambda_n} \right|^k.$$

Démonstration. Soit $\vec{x}_0 = \sum_{i=1}^n \beta^i \vec{e}_i$ le vecteur initial, avec $\beta^n \neq 0$. Le vecteur $\vec{x}_k$ est proportionnel à $\mathbb{A}^k \vec{x}_0 = \sum_{i=1}^n \beta^i \lambda_i^k \vec{e}_i$ et est de norme 1, donc

$$\vec{x}_k = \frac{\beta^n \vec{e}_n + \sum_{i=1}^{n-1} \beta^i \left(\frac{\lambda_i}{\lambda_n} \right)^k \vec{e}_i}{\left((\beta^n)^2 + \sum_{i=1}^{n-1} (\beta^i)^2 \left(\frac{\lambda_i}{\lambda_n} \right)^{2k} \right)^{\frac{1}{2}}}.$$

Comme $|\lambda_i| < \lambda_n$ on en déduit que $\vec{x}_k$ converge vers $\text{sign}(\beta^n) \vec{e}_n$. De

même, on a

$$\|\vec{y}_{k+1}\| = \lambda_n \frac{\left((\beta^n)^2 + \sum_{i=1}^{n-1} (\beta^i)^2 \left(\frac{\lambda_i}{\lambda_n} \right)^{2(k+1)} \right)^{\frac{1}{2}}}{\left((\beta^n)^2 + \sum_{i=1}^{n-1} (\beta^i)^2 \left(\frac{\lambda_i}{\lambda_n} \right)^{2k} \right)^{\frac{1}{2}}},$$

qui converge vers λ_n . □

En pratique (et notamment pour le calcul des valeurs propres de la discrétisation d'un problème aux limites elliptique), on rappelle qu'on est le plus souvent intéressé par la *plus petite* valeur propre, en module, de $\mathbb{A}$. On peut adapter les idées précédentes, ce qui donne la méthode de la puissance inverse dont l'algorithme est écrit ci-dessous. On considère une matrice symétrique réelle $\mathbb{A}$ dont la plus petite valeur propre en module est simple et strictement positive $0 < \lambda_1 < \lambda_i$ pour tout $2 \leq i \leq n$.

1. Initialisation : $\vec{x}_0 \in \mathbb{R}^n$ tel que $\|\vec{x}_0\| = 1$.
2. Itérations : pour $k \geq 1$
 1. résoudre $\mathbb{A}\vec{y}_k = \vec{x}_{k-1}$
 2. $\vec{x}_k = \vec{y}_k / \|\vec{y}_k\|$
 3. test de convergence : si $\|\vec{x}_k - \vec{x}_{k-1}\| \leq \varepsilon$, on arrête et $\lambda_1 \approx \|\vec{y}_k\|^{-1}$ et $\vec{u}_1 \approx \vec{x}_k$.

Si $\vec{\delta}_k = \vec{x}_k - \vec{x}_{k-1}$ est petit, alors $\vec{x}_{k-1}$ est un vecteur propre approché de valeur propre approchée $1/\|\vec{y}_k\|$ car $\mathbb{A}\vec{x}_{k-1} - \frac{\vec{x}_{k-1}}{\|\vec{y}_k\|} = -\mathbb{A}\vec{\delta}_k$.

Proposition 8.5.5 *On suppose que la matrice $\mathbb{A}$ est symétrique réelle, de valeurs propres $(\lambda_1, \dots, \lambda_n)$, associées à une base orthonormée de vecteurs propres $(\vec{e}_1, \dots, \vec{e}_n)$, et que la valeur propre de plus petit module λ_1 est simple et strictement positive, c'est-à-dire que $0 < \lambda_1 < |\lambda_2|, \dots, |\lambda_n|$. On suppose aussi que le vecteur initial $\vec{x}_0$ n'est pas orthogonal à $\vec{e}_1$. Alors la méthode de la puissance inverse converge, c'est-à-dire que*

$$\lim_{k \rightarrow +\infty} \frac{1}{\|\vec{y}_k\|} = |\lambda_1|, \quad \lim_{k \rightarrow +\infty} \vec{x}_k = \vec{x}_\infty \text{ avec } \vec{x}_\infty = \pm \vec{e}_1.$$

La vitesse de convergence est proportionnelle au rapport $\lambda_1/|\lambda_2|$

$$\left| \|\vec{y}_k\|^{-1} - \lambda_1 \right| \leq C \left| \frac{\lambda_1}{\lambda_2} \right|^{2k}, \quad \|\vec{x}_k - \vec{x}_\infty\| \leq C \left| \frac{\lambda_1}{\lambda_2} \right|^k.$$

La démonstration est similaire à celle de la proposition 8.5.4.

Remarque 8.5.2 Pour accélérer la convergence, on peut toujours procéder à une translation de la matrice $\mathbb{A}$ qu'on remplace par $\mathbb{A} - \sigma \mathbb{I}$ avec σ une approximation de λ_1 .

8.6 Analyse modale en mécanique

L'analyse modale est au cœur de nombreux problèmes d'ingénierie mécanique pour concevoir des structures de génie civil (bâtiments parasismiques, ponts), ou des pièces mécaniques d'industrie soumises à des vibrations. C'est aussi autour de problèmes de mécanique que l'analyse modale est apparue historiquement, et fait écho au contexte historique de la fin du XVIIIe siècle évoqué en introduction.

8.6.1 Modes de vibration d'une structure élastique

Le cas des modes de vibration en élasticité est la généralisation directe de (8.1) au cas vectoriel, soit

$$\begin{cases} -\nabla \cdot (\underline{\underline{A}} : \underline{\underline{\varepsilon}}(\underline{u})) + \alpha \underline{u} = \lambda \underline{u}, & \text{dans } \Omega, \\ \underline{u} = 0, & \text{sur } \Gamma_D, \\ (\underline{\underline{A}} : \underline{\underline{\varepsilon}}(\underline{u})) \cdot \underline{n} + \beta \underline{u} = 0, & \text{sur } \Gamma_N. \end{cases} \quad (8.21)$$

L'analyse modale est en effet un outil générique en génie civil de nouveau pour éviter les phénomènes de résonances qui peuvent fragiliser les structures. Les plus fameux exemples sont certainement le pont de la Basse-Chaine suspendu sur la Maine à Angers, ou encore le pont du Millenium à Londres¹⁷. Le premier, sollicité par un régiment au pas a entraîné une résonance verticale et son effondrement. Le second est plus récent et encore plus spectaculaire. Le pont a commencé à entrer en résonance alors que la foule non ordonnée générait une sollicitation transverse avec un spectre large. Mais les individus se retrouvant eux-mêmes déséquilibrés par les mouvements du pont se sont mis à ajuster leur démarche pour compenser leur perte d'équilibre et ont donc favorisé l'entrée en résonance du pont qui aurait pu s'effondrer le jour même de son inauguration.

17: Voire aussi le pont du détroit de Tacoma près de Washington soumis à des vents forts et une résonance en torsion cette fois

8.6.2 Modes de vibration de la plaque

Historiquement, le calcul des modes sur une plaque a donné lieu à des recherches spectaculaires et au développement même de la méthode des éléments finis [15]. En 1787 à Wittenberg en Allemagne, le musicien et physicien Ernst Florence Friedrich Chladni remarqua que lorsqu'il essayait d'exciter une plaque de métal avec l'archet de son violon, il pouvait émettre des sons de différentes natures. Il a soigneusement fait des dessins de chacune d'entre elles, voire la figure 8.2, aujourd'hui appelées figures de Chladni. En 1808, Chladni les reproduit à l'académie de Paris, devant l'empereur Napoléon Ier. Ce dernier, impressionné, organise alors un concours pour établir une modélisation mathématique du phénomène. La solution sera apportée par Sophie Germain¹⁸ et correspond aux modes associés à l'équation

$$-\Delta^2 u(x, y) = \lambda u(x, y), \quad (x, y) \in (-1, 1)^2$$

18: Sophie Germain (1776-1831) se procure les cours de l'École polytechnique réservée alors aux hommes et correspond avec Lagrange sous le nom d'emprunt Antoine Auguste Le Blanc jusqu'à ce qu'il découvre sa véritable identité et devienne son ami.

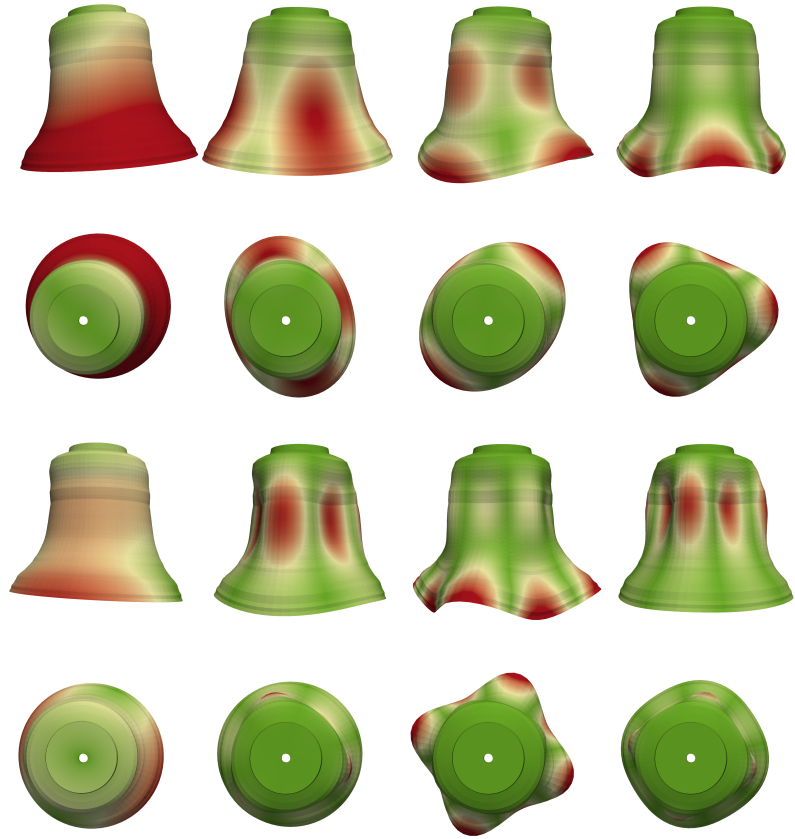


FIGURE 8.1 – Quelques modes de déformation d’une cloche d’église ($E = 10^{11}$ et $\nu = 0.34$) réalisés avec le code Cast3M développé par le CEA et disponible en version éducation pour l’enseignement et la recherche. Source : www-cast3m.cea.fr

couplée aux conditions de bord libre

$$\begin{aligned} \frac{\partial}{\partial x} \left(\frac{\partial^2 u}{\partial x^2} + (2 - \mu) \frac{\partial^2 u}{\partial y^2} \right) &= 0, & \frac{\partial^2 u}{\partial x^2} + \mu \frac{\partial^2 u}{\partial y^2} &= 0, & x \in \{-1, 1\} \\ \frac{\partial}{\partial y} \left(\frac{\partial^2 u}{\partial x^2} + (2 - \mu) \frac{\partial^2 u}{\partial y^2} \right) &= 0, & \frac{\partial^2 u}{\partial y^2} + \mu \frac{\partial^2 u}{\partial x^2} &= 0, & y \in \{-1, 1\} \end{aligned}$$

où μ est un coefficient de cisaillement.

De nouveau, la recherche de valeurs propres peut se mettre sous la forme de la minimisation d’une fonctionnelle de la forme

$$\begin{aligned} \mathcal{J}(v) = \int_{-1}^1 \int_{-1}^1 & \left[\left(\frac{\partial^2 v}{\partial x^2} \right)^2 + \left(\frac{\partial^2 v}{\partial y^2} \right)^2 \right. \\ & \left. + 2\mu \frac{\partial^2 v}{\partial x^2} \frac{\partial^2 v}{\partial y^2} + 2(1 - \mu) \left(\frac{\partial^2 v}{\partial x \partial y} \right)^2 \right] dx dy \end{aligned}$$

sous la contrainte

$$\int_{-1}^1 \int_{-1}^1 v(x, y)^2 dx dy = 1$$

Il est alors fascinant que ce problème ait ensuite donné lieu au développement de la première méthode de résolution approchée des solutions observées. En effet Walther Ritz¹⁹ en 1909 a cherché la solution sous forme de coordonnées dans une base de fonctions adaptées. Sa méthode

19: 1878-1909, physicien suisse

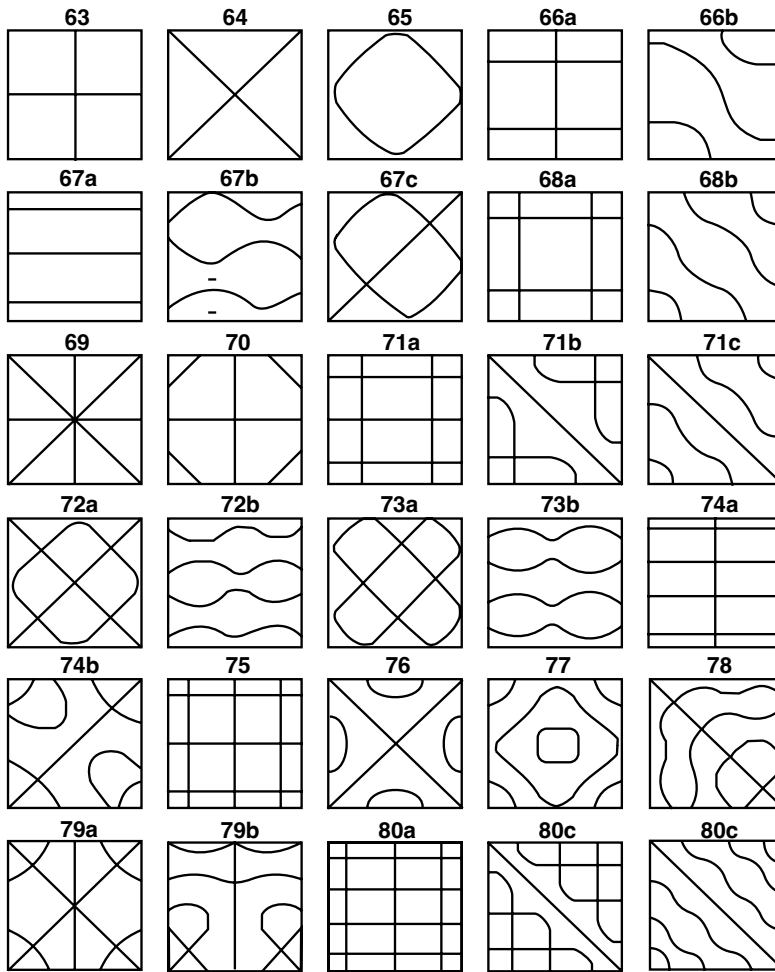


FIGURE 8.2 – Reproduction des figures de E.F.F. Chladni tirée de son livre "Die Akustik"

se veut générale et consiste à approcher toute fonction $v \in \mathcal{V}$ par

$$v(x, y) \simeq v_N(x, y) = \sum_{k=1}^{N^2} v^k \phi_k(x, y). \quad (8.22)$$

Il indique que le choix de la base est primordial. Ce peut être une base de polynômes ou de fonctions hyperboliques, prises typiquement à variables séparées

$$v_N(x, y) = \sum_{n=1}^N \sum_{m=1}^N v^{m,n} \phi_m(x) \psi_n(y).$$

Il peut alors injecter cet *ansatz*²⁰ et calcule

Le détail de ce calcul est le suivant.

$$\begin{aligned} \int_{-1}^1 \int_{-1}^1 \left(\frac{\partial^2 v}{\partial x^2} \right)^2 dx dy &= \int_{-1}^1 \int_{-1}^1 \left(\frac{\partial^2 \sum_{m,n} v^{m,n} \phi_m(x) \phi_n(y)}{\partial x^2} \right)^2 dx dy \\ &= \sum_{m,n} \sum_{i,j} v^{m,n} v^{i,j} \underbrace{\int_{-1}^1 \int_{-1}^1 \frac{\partial^2 \phi_m(x)}{\partial x^2} \phi_n(y) \frac{\partial^2 \phi_i(x)}{\partial x^2} \phi_j(y) dx dy}_{a_{mnij}}. \end{aligned}$$

En faisant de même pour les autres termes, il obtient une approximation

20: de l'allemand, signifiant « racine »

A. Lösungen, die in x und y ungerade und symmetrisch sind.

I. Grundton. $\lambda = 12,43 - 18,0 \delta \mu$.

$$w = u_1 v_1 + 0,0394 (u_1 v_3 + v_1 u_3) - 0,0040 u_3 v_3 - 0,0034 (u_1 v_5 + v_1 u_5) + 0,0011 (u_3 v_5 + u_5 v_3) - 0,0019 u_5 v_5.$$



Fig. 1.

II. $\lambda = 378 - 57 \delta \mu$.

$$w = -0,075 u_1 v_1 + (u_1 v_3 + u_3 v_1) + 0,178 u_3 v_3 + 0,045 (u_1 v_5 + u_5 v_1) - 0,015 (u_3 v_5 + u_5 v_3) - 0,029 u_5 v_5.$$



Fig. 2.

III. $\lambda = 1554$.

$$w = 0,009 u_1 v_1 - 0,075 (u_1 v_3 + u_3 v_1) + u_3 v_3 - 0,057 (u_1 v_5 + u_5 v_1) + 0,121 (u_3 v_5 + u_5 v_3) - 0,007 u_5 v_5.$$



Fig. 3.

Messungen fehlen.

IV. $\lambda = 2945$.

$$w = u_1 v_3 + u_3 v_1.$$

Diese Figur fehlt bei Chladni.



Fig. 4.

B. Lösungen, die in x und y ungerade und antisymmetrisch sind.

I. $\lambda = 316,1 - 270 \delta \mu$.



de la fonctionnelle

$$\mathcal{J}(v) \simeq \mathcal{J}_N(\vec{v}) = \vec{v}^\top \mathbb{A} \vec{v},$$

avec respectivement $\vec{v} = (v^{1,1}, v^{1,2}, \dots, v^{N,N-1}, v^{N,N})^\top \in \mathbb{R}^{N^2}$ et $\mathbb{A} \in \mathbb{M}_{N^2}(\mathbb{R})$ une matrice définie par

$$\mathbb{A} = \begin{pmatrix} a_{1111} & a_{1112} & \cdots & a_{11NN-1} & a_{11NN} \\ a_{1211} & & & & a_{12NN} \\ \vdots & & & & \vdots \\ a_{NN-111} & & & & a_{NN-1NN} \\ a_{NN11} & a_{NN12} & \cdots & a_{NNNN-1} & a_{NNNN} \end{pmatrix}$$

Le critère $\mathcal{J}_N$ est désormais posé en dimension finie et doit être minimisé sous la contrainte

$$\vec{v}^\top \mathbb{M} \vec{v} = 1$$

où $\mathbb{M} \in \mathbb{M}_{N^2}(\mathbb{R})$ a pour coefficients

$$b_{mnij} = \int_{-1}^1 \int_{-1}^1 \phi_m(x) \phi_n(y) \phi_i(x) \phi_j(y) dx dy.$$

Ainsi, en écrivant la condition de stationnarité du Lagrangien associé à cette minimisation sous contrainte, Walter Ritz se ramène au problème matriciel aux valeurs propres généralisé.

$$\mathbb{A} \vec{v} = \lambda \mathbb{M} \vec{v},$$

pour lequel il effectue les calculs qu'il compare aux expériences. C'est un des tout premiers exemples de méthode numérique effective se fondant sur les principes variationnels.

Pour conclure, disons quelques mots sur les résultats de Ritz qui ont eu un accueil contrasté, voir [15] pour tous les détails. Certains, dont David Hilbert, y voyaient avant tout une approche pratique sans rien de plus. D'autres, tels que Lord Rayleigh, ont accusé injustement Ritz de plagiat ce qui a finalement conduit à appeler l'approche de *méthode de Rayleigh-Ritz*. La méthode a en revanche rencontré un certain succès en Russie où Boris Grigor'evich Galerkin étudia les travaux de Ritz avant de publier un papier célèbre en 1915 [14] sur l'utilisation d'une décomposition de la forme (8.22) qui a conduit à appeler ce type d'approche une *méthode de Galerkin*. Enfin, presque 30 ans plus tard, Richard Courant²¹ qui au départ ne s'intéressait qu'aux aspects théoriques des solutions des problèmes variationnels comprit aussi toute la puissance de l'approche proposée par Ritz. Il compléta alors l'approche en proposant d'autres choix judicieux pour les fonctions ϕ_k . En décomposant le domaine Ω en sous-domaines élémentaires (préférentiellement des triangles) et en choisissant des fonctions affines par morceaux sur chacun de ces sous-domaines, il conçut dans [9] les premières étapes d'une méthode d'approximation qui se montrerait extrêmement efficace en pratique et plus encore avec l'avènement du calcul informatique : la méthode des éléments finis²².

21: R. Courant (1988-1972) a publié avec David Hilbert le livre *Methoden der mathematischen Physik* [10, 11] qui systématise la formulation variationnelle des équations aux dérivées partielles.

22: R. Courant a dit : *As Henri Poincaré once remarked, "solution of a mathematical problem" is a phrase of indefinite meaning. Pure mathematicians sometimes are satisfied with showing that the nonexistence of a solution implies a logical contradiction, while engineers might consider a numerical result as the only reasonable goal. Such one-sided views seem to reflect human limitations rather than objective values.*

Bibliographie

- [1] R. ABRAHAM, J. E. MARSDEN et T. RATTU. *Manifolds, tensor analysis, and applications*. T. 75. Springer Science & Business Media, 2012 (cf. p. 8).
- [2] G. ALLAIRE et S. M. KABER. *Numerical Linear Algebra*. Texts in Applied Mathematics. Springer New York, 2014 (cf. p. 112, 181, 184).
- [3] I. BABUŠKA et J. E. OSBORN. « Finite element-Galerkin approximation of the eigenvalues and eigenvectors of selfadjoint problems ». In : *Mathematics of Computation* 52.186 (1989), p. 275-297 (cf. p. 184).
- [4] A.-S. BONNET-BEN DHIA, L. BOURGEOIS et C. HAZARD. « Outils élémentaires d'analyse pour les Equations aux Dérivées Partielles ». Engineering school. Lecture. France, 2021 (cf. p. 9, 18, 25, 51).
- [5] J.-M. BONY. *Théorie des distributions et analyse de Fourier*. 2001 (cf. p. 25, 27, 51).
- [6] H. BREZIS. *Analyse fonctionnelle, Théorie et applications*. Masson, 1983 (cf. p. 38, 41, 43, 49, 67).
- [7] A. CHAMBOLLE et T. POCK. « Chapter 6 - Approximating the total variation with finite differences or finite elements ». In : *Geometric Partial Differential Equations - Part II*. Sous la dir. d'A. BONITO et R. H. NOCHETTO. T. 22. Handbook of Numerical Analysis. Elsevier, 2021, p. 383-417 (cf. p. 14).
- [8] P. G. CIARLET. *Introduction à l'analyse numérique matricielle et à l'optimisation*. Collection Mathématiques appliquées pour la maîtrise. Dunod, 1998 (cf. p. 112, 151, 184).
- [9] R. COURANT. « Variational methods for the solution of problems of equilibrium and vibrations ». In : *Bulletin of the American Mathematical Society* 49 (1943), p. 1-23 (cf. p. 190).
- [10] R. COURANT et D. HILBERT. *Methods of mathematical physics*. T. I. New York : Interscience Publishers, 1953 (cf. p. 190).
- [11] R. COURANT et D. HILBERT. *Methods of mathematical physics*. T. II. New York : Interscience Publishers, 1962 (cf. p. 190).
- [12] A. ERN et J.-L. GUERMOND. *Finite Elements I : Approximation and Interpolation*. Springer, 2021 (cf. p. 142).
- [13] L. EULER. *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes sive solutio problematis isoperimetrici latissimo sensu accepti*. T. 24. Enestr. 65, Opera Omnia, Ser. I. Lausannae & Genevae, 1744 (cf. p. 20).
- [14] B. G. GALERKIN. « Rods and Plates. Series occurring in various questions concerning the elastic equilibrium of rods and plates (in Russian), » in : *Engineers Bulletin (Vestnik Inzhenerov)* 19 (1915), p. 897-908 (cf. p. 190).
- [15] M. J. GANDER et G. WANNER. « From Euler, Ritz, and Galerkin to Modern Computing ». In : *SIAM Review* 54.4 (2012), p. 627-666 (cf. p. 19, 187, 190).
- [16] V. GIRAULT et P. RAVIART. *Finite Element Methods for Navier-Stokes Equations : Theory and Algorithms*. Springer Series in Computational Mathematics. Springer Berlin Heidelberg, 2012 (cf. p. 141).
- [17] P. GRISVARD. *Elliptic problems in nonsmooth domains*. T. 24. Monographs and Studies in Mathematics. Boston, MA : Pitman (Advanced Publishing Program), 1985, p. xiv+410 (cf. p. 38, 93).
- [18] P. GRISVARD. *Singularities in boundary value problems*. T. 22. Recherches en Mathématiques Appliquées [Research in Applied Mathematics]. Paris : Masson, 1992, p. xiv+199 (cf. p. 93).
- [19] L. GUIBAS et J. STOLFI. « Primitives for the manipulation of general subdivisions and the computation of Voronoi ». In : *ACM Trans. Graph.* 4.2 (1985), p. 74-123 (cf. p. 167).
- [20] J. LERAY. « Sur le mouvement d'un liquide visqueux emplissant l'espace ». In : *Acta mathematica* 63 (1934), p. 193-248 (cf. p. 30).
- [21] J.-L. LIONS et E. MAGENES. *Problèmes aux limites non homogènes et applications*. Dunod, 1968 (cf. p. 36).

- [22] L. MARÉCHAL. « Advances in octree-based all-hexahedral mesh generation : handling sharp features ». In : *Proc. IMR* 18 (2009), p. 65-84 (cf. p. 139).
- [23] W. RITZ. « Theorie der Transversalschwingungen einer quadratischen Platte mit freien Rändern ». In : *Annalen der Physik* 333.4 (1909), p. 737-786 (cf. p. 189).
- [24] L. SCHWARTZ. *Théorie des distributions*, Hermann, Paris (2ème édition). 1966 (cf. p. 30).