Approximation variationnelle des fonctions
Notes du Cours de M2

Albert Cohen

Une fonction f arbitraire, appartenant à un espace de Banach X de fonctions définies sur un domaine $\Omega \subset \mathbb{R}^d$, peut être représentée par son graphe, ou de manière équivalente par la donnée de l'ensemble de ses valeurs $(x, f(x))$ pour $x \in \Omega$. Ces valeurs étant en nombre infini et il n'est pas possible en pratique de les mettre en mémoire sur un ordinateur. On peut alors chercher à remplacer f par une fonction g plus simple qui est proche de f dans X et dépend d'un nombre fini n de paramètres que l'on peut ainsi mettre en mémoire. On choisit ainsi g dans un espace $X_n \subset X$ de dimension finie n en cherchant à résoudre le problème de minimisation

$$\min_{g \in X_n} \|f - g\|_X. \tag{0.1}$$

Le qualificatif d'approximation *variationnelle* fait référence à ce problème d'optimisation. Il existe une immense variété d'espaces X_n utilisés en pratique : polynômes algébriques, trigonométriques, fonctions polynomiales par morceaux. Intuitivement, on approche de mieux en mieux f lorsque la dimension n augmente.

La *théorie de l'approximation* étudie de façon rigoureuse le compromis entre la *complexité* donnée par le nombre de paramètres n et la *précision* que l'on peut obtenir entre f et g, c'est à dire l'erreur de meilleure approximation définie comme la valeur du minimum

$$\sigma_n(f)_X := \min_{g \in X_n} \|f - g\|_X. \tag{0.2}$$

Elle s'intéresse aussi à la manière dont on construit en pratique l'approximation g à partir de f, car le problème de minimisation ci-dessus n'est pas toujours simple à résoudre numériquement. Les applications sont multiples : conception des courbes et surfaces assistée par ordinateur, compression du signal et de l'image, lissage des données en statistiques, simulation numérique des EDP. Suivant l'application, fonction f que l'on approche peut donc être (i) connue explicitement à travers une formule, (ii) connue à travers ses valeurs en certains points ou par certains de ses coefficients dans une base, (iii) connue à une certaine précision près lorsque ces valeurs sont bruitées ou (iv) inconnue mais solution d'une équation connue.

Le cours se penche en premier lieu sur deux exemples fondamentaux d'approximation, très différents, permettant de se familiariser dans un cadre simple avec les méthodes et résultats de théorie de l'approximation. Il s'agit d'une part de l'approximation par des fonctions *polynomiales par morceaux* sur une partition de Ω, et d'autre part de l'approximation par les *séries de Fourier* dans le cas de fonctions périodiques. Les espaces de Sobolev jouent dans les deux cas un rôle central dans l'analyse de l'erreur d'approximation. On s'intéresse à la *méthode des éléments finis* dans laquelle les fonctions de X_n sont polynomiales par morceaux sur une partition de Ω avec des contraintes de raccord réguliers entre les morceaux. On applique en particulier cette méthode à l'approximation numérique de solutions d'équations aux dérivées partielles linéaires qui admettent une *formulation variationnelle*, c'est à dire dont la solution u est aussi solution du problème : *Trouver* $u \in X$ *tel que* $a(u, v) = L(v)$ *pour tout* $v \in X$, où X est un espace de Hilbert, a une forme bilinéaire sur $X \times X$, et L une forme linéaire sur X. La *méthode de Galerkin* permet de calculer une approximation u_n de u dans un sous espace $X_n \subset X$ dont la précision se compare à l'erreur de meilleure approximation $\sigma_n(u)_X$.

Ces notes contiennent la totalité des résultats du cours sous une forme relativement condensée. En particulier, les démonstrations les plus simples sont esquissées ou laissées en exercice. *Ces notes sont mises à jour et corrigées en temps réel. Toutes les remarques permettant d'en améliorer la rédaction peuvent être envoyées à l'adresse* cohen@ljll.math.upmc.fr.

1 Approximation des fonctions

De façon générale, nous considérons dans ce cours l'approximation de fonctions définies sur un domaine, c'est à dire un ouvert connexe, $\Omega \subset \mathbb{R}^d$. On considérera par exemple des fonctions appartenant à l'espace de Lebesgue $L^p(\Omega)$ pour un $p \geq 1$, ou à l'espace des fonctions continues, ou à des espaces de fonctions plus régulières.

Les procédés généraux consistant à approcher des fonctions arbitraires de d variables par des fonctions particulières sont utiles au delà du domaine du calcul numérique. En effet, un grand nombre de résultats fondamentaux sur les fonctions sont prouvés en se ramenant par *densité* à des fonctions plus simples. Rappelons trois résultats de densité fondamentaux :

1. Les fonctions constantes par morceaux, sur un nombre fini de sous-domaines, sont denses dans $L^p(\mathbb{R}^d)$ pour $1 \leq p \leq \infty$: pour tout $f \in L^p(\mathbb{R}^d)$ et $\varepsilon > 0$, il existe une fonction de la forme $g = \sum_{i=1}^m c_i \chi_{A_i}$ où les A_i sont des ensembles mesurables et χ_{A_i} leurs fonctions indicatrices, telle que $\|f - g\|_{L^p} \leq \varepsilon$. Une manière simple de construire g est de prendre pour n suffisament grand une suite croissante de valeurs de la forme $c_i = -n + i/n$, pour $i = 1, \dots, m = 2n^2$, et de définir A_i par
$$A_i = \{x \in \Omega \ : \ c_i \leq f(x) < c_{i+1}\}. \tag{1.1}$$
L'approximation par les fonctions constantes par morceaux est à la base de la construction de l'intégrale de Lebesgue et des espaces L^p.
2. Les fonctions continues à support compact sont denses dans $L^p(\mathbb{R}^d)$ pour $1 \leq p < \infty$: pour tout $f \in L^p(\mathbb{R}^d)$ et $\varepsilon > 0$, il existe une fonction $g \in C_c(\mathbb{R}^d)$ telle que $\|f - g\|_{L^p} \leq \varepsilon$. C'est un résultat fondamental de la théorie des espaces L^p.
3. Les fonctions infiniment dérivables sont denses dans $L^p(\mathbb{R}^d)$ pour $1 \leq p < \infty$: pour tout $f \in L^p(\mathbb{R}^d)$ et $\varepsilon > 0$, il existe une fonction $g \in C^\infty(\mathbb{R}^d)$ telle que $\|f - g\|_{L^p} \leq \varepsilon$. Une manière de construire g consiste à régulariser f par convolution : pour φ une fonction C^∞ à support compact telle que $\int \varphi(x)dx = 1$, et n suffisament grand, on définit
$$g(x) = f * \varphi_n(x) = \int_{\mathbb{R}^d} f(x-y)\varphi_n(y)dy, \qquad \varphi_n(x) := n^d \varphi(nx), \tag{1.2}$$
En utilisant la densité des fonctions continues à support compact rappelée ci-dessus, on montre aussi que l'espace $\mathcal{D}(\mathbb{R}^d)$ des fonctions C^∞ à support compact est dense dans $L^p(\mathbb{R}^d)$ pour $1 \leq p < \infty$. Ce résultat est bien sûr faux pour l'espace l'espace $L^\infty(\mathbb{R}^d)$, mais vrai pour l'espace des fonctions continues qui tendent vers 0 à l'infini muni de la norme L^∞.

Les mêmes résultats sont vérifiés lorsqu'on remplace $\mathbb{R}^d$ par un domaine Ω.

Dans la pratique du calcul numérique, on cherche à approcher des fonctions arbitraires par des fonctions plus simples pouvant être représentée par un nombre fini n de paramètres, c'est à dire typiquement appartenant à un espace de dimension n. Ce n'est pas le cas pour les résultats d'approximation qu'on vient de citer. Un exemple élémentaire d'approximation par des espaces de dimension finie est fourni par le théorème de Weierstrass : si I est un intervalle fermé borné, pour tout $f \in C(I)$ et $\varepsilon > 0$, il existe $n > 0$ et $g \in \mathbb{P}_n$ telle que $\|f - g\|_{L^\infty(I)} \leq \varepsilon$.

Ce type de résultat est insuffisant dans la pratique du calcul numérique car il ne fournit pas une estimation quantitative de la précision ε en fonction n. L'un des objectifs de la *théorie de l'approximation* est de fournir de telles estimations. De façon générale, étant donné un espace de Banach X (typiquement de fonctions définies sur un domaine $\Omega \subset \mathbb{R}^d$), et une famille $(X_n)_{n\geq 1}$ de sous-espaces tels que $\dim(X_n) = n$, on s'intéresse à l'approximation de $f \in X$ par $g \in X_n$. L'erreur de meilleure approximation de f est alors définie par la valeur
$$\sigma_n(f)_X := \min_{g \in X_n} \|f - g\|_X. \tag{1.3}$$
Une observation importante est que le minimum est toujours atteint, c'est à dire qu'il existe un élément $g_n \in X_n$, appelé meilleure approximation de f dans X_n, tel que
$$\|f - g_n\|_X = \sigma_n(f)_X. \tag{1.4}$$

En effet, l'application $g \mapsto \|f-g\|_X$ est continue sur X_n et infinie à l'infinie, puisque $\|f-g\|_X \geq \|g\|_X - \|f\|_X$. Par compacité on en déduit l'existence d'un minimum sur X_n.

Ce résultat reste vrai lorsque X_n est un sous espace de dimension infinie fermé et que l'espace X est réflexif, en utilisant le fait que toute application continue convexe sur un Banach réflexif est faiblement semi-continue inférieurement (exercice). On remarque que le caractère fermé de l'espace X_n entraîne l'équivalence $\sigma_n(f)_X = 0 \Leftrightarrow f \in X_n$.

On montre aisément que g_n est unique lorsque l'espace X est strictement convexe, c'est à dire lorsque tout segment contenu dans la boule unité ne peut rencontrer la sphère unité qu'à ses extrémités. On rappelle que $L^p(\Omega)$ est strictement convexe et réflexif lorsque $1 < p < \infty$. Il est par ailleurs facile d'exhiber des cas de non-unicité de la meilleure approximation lorsque $X = L^\infty(\Omega)$ ou $L^1(\Omega)$ (exercice).

Pour une famille $(X_n)_{n\geq 1}$ donnée, on cherche à identifier les fonctions $f \in X$ telles que $\sigma_n(f)$ décroit à une certaine vitesse lorsque n augmente. Pour un $r > 0$, on peut ainsi introduire l'ensemble

$$\mathcal{A}^r = \mathcal{A}^r((X_n)_{n\geq 1}, X) := \{f \in X \ : \ \sup_{n\geq 1} n^r \sigma_n(f)_X < \infty\}, \tag{1.5}$$

qui contient toutes les fonctions $f \in X$ telles que $\sigma_n(f)$ décroit (au moins) à la vitesse n^{-r}. On peut montrer (exercice) que cet ensemble est un sous-espace vectoriel de X. Il est parfois possible de caractériser exactement l'espace $\mathcal{A}^r$ par des propriétés de *régularité*. Un exemple élémentaire illustrant une telle caractérisation est celui de l'approximation des fonctions continues sur un intervalle par des fonctions constantes par morceaux sur une partition uniforme.

Théorème 1.1 *Soit $X = C([0,1])$ muni de la norme L^∞ et X_n l'espace des fonctions constantes par morceaux sur les intervalles $[\frac{k}{n}, \frac{k+1}{n}[$ pour $k = 0, \dots, n-1$. L'espace $\mathcal{A}^1$ coincide alors avec celui des fonctions lipschitzienne sur $[0,1]$.*

Preuve : Supposons f lipschitzienne de constante de Lipschitz L. On choisit $g_n \in X_n$ telle que $g_n = f(\frac{k}{n})$ sur chaque intervalle $[\frac{k}{n}, \frac{k+1}{n}[$. On a alors pour $x \in [\frac{k}{n}, \frac{k+1}{n}[$,

$$|f(x) - g_n(x)| = \left| f(x) - f\Big(\frac{k}{n}\Big) \right| \leq Ln^{-1}, \tag{1.6}$$

et donc $\|f - g_n\|_{L^\infty} \leq \frac{L}{n}$, ce qui montre que $f \in \mathcal{A}^1$.

Réciproquement, soit $f \in C([0,1])$ telle que $f \in \mathcal{A}^1$. Il existe donc une constante $C \geq 0$ et une suite de fonctions $g_n \in X_n$ telle que $\|f - g_n\|_{L^\infty} \leq Cn^{-1}$, pour tout $n \geq 1$. Pour tout $x \in [0,1[$ et pour tout $0 < h < 1 - x$, on peut écrire

$$\begin{aligned} |f(x+h) - f(x)| &\leq |f(x+h) - g_n(x+h)| + |g_n(x+h) - g_n(x)| + |g_n(x) - f(x)| \\ &\leq |g_n(x+h) - g_n(x)| + 2Cn^{-1}. \end{aligned}$$

En choisissant n tel que $\frac{1}{2n} \leq h \leq \frac{1}{n}$, on voit que x et $x+h$ sont soit dans le même intervalle $[\frac{k}{n}, \frac{k+1}{n}[$, auquel cas $g_n(x+h) - g_n(x) = 0$, soit dans deux intervalles adjacents $[\frac{k-1}{n}, \frac{k}{n}[$ et $[\frac{k}{n}, \frac{k+1}{n}[$, auquel cas on peut écrire en utilisant la continuité de f

$$|g_n(x+h) - g_n(x)| \leq \left| g_n(x+h) - f\Big(\frac{k}{n}\Big) \right| + \left| g_n(x) - f\Big(\frac{k}{n}\Big) \right| \leq 2Cn^{-1}. \tag{1.7}$$

On a donc établi

$$|f(x+h) - f(x)| \leq 4Cn^{-1} \leq 8Ch, \tag{1.8}$$

ce qui montre que f est lipschitzienne. ◇

Il n'est pas toujours aussi aisé de caractériser exactement les fonctions de $\mathcal{A}^s$ pour des familles $(X_n)_{n\geq 1}$ d'intérêt, et on se contentera plus souvent d'identifier des conditions de régularité suffisantes permettant de garantir une certaine vitesse de décroissance en n^{-r} de $\sigma_n(x)_X$.

Un autre objectif de théorie de l'approximation est de proposer des méthodes pratiques pour calculer une approximation de $f \in X$ dans X_n car le problème de minimisation qui définit la meilleure approximation g_n n'est pas toujours simple à résoudre numériquement.

Dans le cas particulier où X est un espace de Hilbert (par exemple dans le cas $X = L^2(\Omega)$) la meilleure approximation est fournie par la *projection orthogonale*

$$g_n = P_n f, \tag{1.9}$$

qui est caractérisée par le fait que $g_n \in X_n$ et

$$\langle f - g_n, g\rangle_X = 0, \quad g \in X_n, \tag{1.10}$$

où $\langle \cdot, \cdot\rangle_X$ est le produit scalaire de X.

L'intérêt de la projection orthogonale est qu'elle peut se calculer explicitement quand on dispose d'une base $\{\varphi_1, \dots, \varphi_n\}$ de X_n. Si c'est une base orthonormale, on a directement

$$P_n f = \sum_{i=1}^{n} \langle f, \varphi_i\rangle \varphi_i. \tag{1.11}$$

Pour une base non orthonormale, en posant $P_n f = \sum_{i=1}^n a_i \varphi_i$, et en utilisant la caractérisation de la projection orthogonale rappelée ci-dessus avec $g = \varphi_j$, on voit que le vecteur $A := (a_1, \dots, a_n)^t$ est solution du système

$$GA = F \tag{1.12}$$

où $G := (\langle \varphi_i, \varphi_j\rangle_X)_{i,j=1,\dots,n}$ est la matrice de Gramm (qui est inversible puisque les φ_i contituent une base) et $F := (\langle f, \varphi_1\rangle_X, \dots, \langle f, \varphi_n\rangle_X)^t$.

Dans le cas d'espaces de Banach non-hilbertiens, l'application qui à $f \in X$ associe sa meilleure approximation $g_n \in X_n$ est typiquement non-linéaire et son calcul explicite est rarement possible. Il est parfois possible de remplacer la meilleure approximation par un $\tilde{g}_n \in X_n$ plus simple à calculer à partir de f et qui se comporte "presque" aussi bien, au sens où il existe une constante $C > 1$ telle que

$$\|f - \tilde{g}_n\|_X \le C\sigma_n(f)_X, \tag{1.13}$$

pour tout $f \in X$ et $n \ge 1$. De telles approximation sont dites "quasi-optimales". On peut en particulier les construire si l'on dispose d'opérateurs de projection P_n, c'est a dire tels que $P_n(X) = X_n$ et $P_n g = g$ si $g \in X_n$, qui sont *uniformément stables* au sens où

$$\sup_{n\ge 1} \|P_n\|_{\mathcal{L}(X)} < \infty. \tag{1.14}$$

En effet, si on définit

$$\tilde{g}_n = P_n f, \tag{1.15}$$

on peut écrire pour tout $g \in X_n$,

$$\|f - \tilde{g}_n\|_X \le \|f - g\|_X + \|\tilde{g}_n - g\|_X = \|f - g\|_X + \|P_n f - P_n g\|_X \le (1 + \|P_n\|_{\mathcal{L}(X)})\|f - g\|_X. \tag{1.16}$$

En prenant l'infimum sur $g \in X_n$, il vient

$$\|f - \tilde{g}_n\|_X \le C\sigma_n(f)_X, \tag{1.17}$$

pour tout $f \in X$ et $n \ge 1$, avec $C := 1 + \sup_{n\ge 1} \|P_n\|_{\mathcal{L}(X)}$.

En particulier, on voit que si $\cup_{n\ge 0} X_n$ est dense dans X, c'est à dire que pour tout $f \in X$ on a

$$\lim_{n\to+\infty} \sigma_n(f)_X = 0, \tag{1.18}$$

alors si $(P_n)_{n\ge 0}$ est une suite de projecteurs uniformément stables, on a pour tout $f \in X$

$$\lim_{n\to+\infty} P_n f = f. \tag{1.19}$$

Réciproquement, si pour une suite d'opérateurs $(P_n)_{n\ge 0}$ on a

$$\lim_{n\to+\infty} P_n f = f, \tag{1.20}$$

pour tout $f \in X$ alors le théorème de Banach-Steinhaus entraîne que ces opérateurs sont uniformément stables.

2 Espaces de Sobolev

Pour m entier positif, et Ω un domaine de $\mathbb{R}^d$, on rappelle que $C^m(\Omega)$ désigne l'espace des fonctions qui admettent des dérivées partielles d'ordre $1, \dots, m$ continues et uniformément bornées sur Ω. On peut munir cet espace de la norme

$$\|f\|_{C^m} = \max\{\|D^\alpha f\|_{L^\infty} \ : \ |\alpha| \le m\}, \tag{2.1}$$

où

$$D^\alpha f := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}, \quad \alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d, \quad |\alpha| := \alpha_1 + \cdots + \alpha_d, \tag{2.2}$$

pour laquelle il constitue un espace de Banach. Dans le cas $m = 0$ on retrouve l'espace $C(\Omega) = C^0(\Omega)$ des fonctions continues et bornées sur Ω. On peut aussi définir les espaces C^m sur un domaine fermé, en particulier les espaces $C^m(\overline{\Omega})$.

Les espaces de Sobolev permettent de décrire la régularité des fonctions à l'ordre m au sens plus faible de l'intégrabilité L^p des dérivées partielles : pour $1 \le p \le \infty$ on définit l'espace $W^{m,p}(\Omega)$ comme l'ensemble des fonctions $f \in L^p(\Omega)$ dont les dérivées partielles $D^\alpha f$ *au sens des distributions* appartiennent à $L^p(\Omega)$ pour tout $|\alpha| \le m$. Ceci correspond plus rigoureusement à la définition suivante.

Définition 2.1 *Une fonction $f \in L^p(\Omega)$ appartient à $W^{m,p}(\Omega)$ si et seulement si pour tout $|\alpha| \le m$, il existe $f_\alpha \in L^p(\Omega)$ telle que*

$$\int_\Omega f_\alpha \varphi = (-1)^{|\alpha|} \int_\Omega f D^\alpha \varphi, \tag{2.3}$$

pour tout $\varphi \in \mathcal{D}(\Omega)$ où $\mathcal{D}(\Omega)$ est l'ensemble des fonctions C^∞ à support compact dans Ω.

Il est aisé de vérifier que f_α est alors unique et on note alors $f_\alpha = D^\alpha f$. Lorsque $1 < p < \infty$, une autre façon équivalente d'exprimer cette définition est d'affirmer que pour tout $|\alpha| \le m$, il existe une constante C_α telle que

$$\Big| \int_\Omega f D^\alpha \varphi \Big| \le C_\alpha \|\varphi\|_{L^{p'}}, \tag{2.4}$$

avec $1/p' + 1/p = 1$ (exercice). Pour $m = 0$ on pose $W^{0,p} = L^p$

On munit $W^{m,p}(\Omega)$ de la norme

$$\|f\|_{W^{m,p}} := \Big(\sum_{|\alpha| \le m} \|D^\alpha f\|_{L^p}^p \Big)^{1/p}, \tag{2.5}$$

et on définit aussi la semi-norme par

$$|f|_{W^{m,p}} := \Big(\sum_{|\alpha| = m} \|D^\alpha f\|_{L^p}^p \Big)^{1/p}, \tag{2.6}$$

de sorte que $\|f\|_{W^{m,p}} = (\|f\|_{W^{m-1,p}}^p + |f|_{W^{m,p}}^p)^{1/p}$.

Remarque 2.1 *Tout au long de ce cours, on ne mentionnera pas le domaine Ω dans les normes, en écrivant en particulier $\|f\|_{L^p} = \|f\|_{L^p(\Omega)}$, lorsque celui-ci est celui sur lequel on a défini, ou sur lequel on étudie, la fonction f.*

Dans le cas $p = \infty$ ces définitions sont modifiées en remplacant les norme ℓ^p par des max. On voit en particulier que les espaces C^m et $W^{m,\infty}$ ont la même norme, mais on note qu'ils différent au sens où le premier est strictement inclus dans le second. Par exemple, l'espace $W^{1,\infty}$ est l'espace des fonctions Lipschitiziennes, qui ne sont dérivables que presque partout.

En partant de ces définitions et en utilisant le fait que $L^p(\Omega)$ est complet, on démontre facilement le résultat suivant.

Théorème 2.1 *$W^{m,p}(\Omega)$ muni de sa norme est un espace de Banach.*

Dans le cas $p = 2$ qui nous intéresse plus particulièrement on pose $H^m(\Omega) := W^{m,2}(\Omega)$. La norme H^m dérive du produit scalaire

$$\langle f, g\rangle_{H^m} := \sum_{|\alpha|\leq m} \langle D^\alpha f, D^\alpha g\rangle_{L^2}, \tag{2.7}$$

et par conséquent H^m est un espace de Hilbert.

La définition de $W^{m,p}(\Omega)$ montre facilement que si $f \in W^{m,p}(\mathbb{R}^d)$ alors sa restriction à Ω est dans $W^{m,p}(\Omega)$. Une question plus délicate est de savoir si toute fonction de $W^{m,p}(\Omega)$ est la restriction d'une fonction de $W^{m,p}(\mathbb{R}^d)$, ce qui revient à demander si toute fonction de $W^{m,p}(\Omega)$ peut être étendue en une fonction de $W^{m,p}(\mathbb{R}^d)$. Lorsque $m = 0$, c'est à dire dans le cas des espaces L^p pour $1 \leq p \leq \infty$, l'extension triviale par valeur nulle convient quelquesoit le type de domaine considéré. Lorsque $m \geq 1$, il est nécessaire de faire des hypothèses supplémentaires sur la *régularité géométrique* du domaine Ω. Une classe importante de domaines bornés est donnée par la définition suivante.

Définition 2.2 *Un domaine borné Ω est dit lipschitzien si il existe une famille finie de boules ouvertes $(B_i)_{i=1,\cdots,n}$ telle que $\partial\Omega \subset \cup_{i=1}^n B_i$ et sur chaque B_i il existe un système de coordonnées $(x_1, \cdots, x_d)$ et une fonction ψ_i lipschitzienne telle que*

$$\Omega \cap B_i = \{(x_1, \cdots, x_d) \in B_i \ ; \ x_d < \psi_i(x_1, \cdots, x_{d-1})\}. \tag{2.8}$$

Intuitivement cela signifie que la frontière de Ω peut-être vue localement comme le graphe d'une fonction lipschitzienne. La plupart des domaines classiques - en particulier les polygones en dimension 2 et presque tous les polyhèdres en dimension 3 - sont lipschitziens. Des exemples de domaines non-lipschitziens sont ceux dont la frontière présente des points de rebroussement, ou une fissure rentrant dans le domaine. On peut aussi définir des domaines plus réguliers : on dira que Ω est de classe $C^{m,1}$ lorsque les fonctions ψ_i sont C^m et leurs dérivées partielles d'ordre m sont lipschitziennes. Tous les domaines considérés dans ce cours seront au minimum de type lipschitzien.

Le résultat suivant, dont la preuve est difficile, est dû à Stein.

Théorème 2.2 *Si Ω est un domaine borné lipschitzien, il existe un opérateur d'extension E qui est borné de $W^{m,p}(\Omega)$ dans $W^{m,p}(\mathbb{R}^d)$: pour tout $f \in W^{m,p}(\Omega)$, la restriction de Ef à Ω est égale à f et $\|Ef\|_{W^{m,p}(\mathbb{R}^d)} \leq C\|f\|_{W^{m,p}(\Omega)}$, pour une constante C indépendante de f. Il en est de même pour les espaces C^m.*

A titre d'exercice on pourra montrer que ce théorème est faux dans un domaine non-lipschitzien présentant une fissure (par exemple le carré ouvert $]-1,1[^2$ privé du segment allant de $(0,0)$ à $(1,0)$).

Rappelons quelques résultats importants sur la densité des fonctions régulières dans les espaces de Sobolev : si Ω est un domaine borné lipschitzien, $\mathcal{D}(\overline{\Omega})$ est dense dans $W^{m,p}(\Omega)$ pour $p < +\infty$, où $\mathcal{D}(\overline{\Omega})$ désigne l'ensemble des restrictions à Ω des fonctions C^∞. Ce théorème peut se démontrer d'abord sur $\mathbb{R}^d$ par des opérations de troncature et de régularisation (exercice), puis sur Ω en utilisant le théorème de prolongement 2.2. Il faut bien noter que $\mathcal{D}(\overline{\Omega})$ diffère de $\mathcal{D}(\Omega)$ qui n'est pas dense dans $W^{m,p}(\Omega)$, sauf si $\Omega = \mathbb{R}^d$ ou si $m = 0$ ce qui correspond à la densité de $\mathcal{D}(\Omega)$ dans $L^p(\Omega)$. On peut aussi démontrer plus directement (par régularisation, exercice) des résultat de densité pour un domaine borné quelconque lorqu'on se place à une distance fixée de son bord : pour tout $f \in W^{m,p}(\Omega)$ et pour tout $\delta > 0$ il existe une suite $(f_n)_{n\geq 0}$ de fonctions de $\mathcal{D}(\Omega)$ qui converge vers f dans $W^{m,p}(\Omega_\delta)$ où $\Omega_\delta := \{x \in \Omega \ : \ B(x,\delta) \subset \Omega\}$. D'autre part, tous ces résultats de densité sont faux pour les espaces $W^{m,\infty}(\Omega)$ mais vrais pour les espaces $C^m(\overline{\Omega})$.

Les espaces de Sobolev peuvent aussi être définis pour des indices de régularité non-entier. Dans cette définition, tout comme dans la suite de ces notes, $\|x\|$ désigne la norme euclidienne pour tout vecteur x.

Définition 2.3 *Pour Ω un domaine de $\mathbb{R}^d$, $m < s < m+1$ et $1 \leq p < \infty$, $W^{s,p}(\Omega)$ est l'ensemble des fonctions f de $W^{m,p}(\Omega)$ telles que*

$$|f|_{W^{s,p}}^p := \sum_{|\alpha|=m} \int\int_{\Omega\times\Omega} \frac{|D^\alpha f(x) - D^\alpha f(y)|^p}{\|x-y\|^{(s-m)p+d}} dx\, dy < +\infty, \tag{2.9}$$

Cet espace est muni de la norme $\|f\|_{W^{s,p}} := (\|f\|_{W^{m,p}}^p + |f|_{W^{s,p}}^p)^{1/p}$

Dans le cas $p = 2$ et $\Omega = \mathbb{R}^d$, on peut aussi définir $H^s(\mathbb{R}^d)$ de manière équivalente en utilisant la transformée de Fourier : $f \in H^s(\mathbb{R}^d)$ si et seulement si $(1 + |\omega|^s)\hat{f} \in L^2(\mathbb{R}^d)$. Dans le cas $p = \infty$ on remplace (2.9) par la condition de Hölder

$$|f|_{W^{s,\infty}} := \sup_{|\alpha|=m} \sup_{x \neq y} \frac{|D^\alpha f(x) - D^\alpha f(y)|}{\|x-y\|^{s-m}} < \infty, \tag{2.10}$$

et on note parfois $W^{s,\infty} = C^s$.

Ces espaces ont des propriétés similaires à celles énoncées pour les espaces d'indices entiers : complétude par rapport à leur norme, opérateur d'extension, densité des fonctions régulières.

Terminons par quelques remarques sur les relations d'inclusions entre espaces de Sobolev. On a de manière évidente $W^{m,p}(\Omega) \subset W^{n,p}(\Omega)$ si $m \geq n$, ainsi $W^{m,p}(\Omega) \subset W^{m,q}(\Omega)$ si $p > q$ et Ω est borné. On rappelle aussi le théorème de Rellich.

Théorème 2.3 *Si Ω est domaine borné lipschitzien et $m > n$, l'injection de $W^{m,p}(\Omega)$ dans $W^{n,p}(\Omega)$ est compacte : toute suite bornée dans $W^{m,p}(\Omega)$ admet une sous-suite convergente dans $W^{n,p}(\Omega)$.*

En particulier, l'injection de $W^{m,p}(\Omega)$ dans $L^p(\Omega)$ est compacte pour $m > 0$. Des inclusions plus générales s'expriment par les *injections de Sobolev* : pour un domaine $\Omega \subset \mathbb{R}^d$ borné lipschitizien, $m \geq n$ et $m - n > d/p - d/q$ on a l'injection $W^{m,p}(\Omega)$ dans $W^{n,q}(\Omega)$ et cette injection est compacte si $m > n$. Le cas critique $m - n = d/p - d/q$ est à manipuler avec précaution, l'injection étant vraie si $q < \infty$ mais pouvant être fausse si $q = \infty$ et n est entier. Par exemple, en dimension $d = 2$ l'espace $H^1(\Omega)$ s'injecte dans $L^q(\Omega)$ pour tout $q < \infty$ mais pas dans L^∞, en revanche l'espace $W^{2,1}(\Omega)$ s'injecte dans L^∞ (exercice). On peut vérifier que les injections de Sobolev ne sont jamais compactes dans le cas critique (exercice).

3 Approximation polynomiale par morceaux

Dans cette section on considére l'approximation par des fonctions qui coincident avec des polynômes sur chaque élément d'une partition d'un domaine $\Omega \subset \mathbb{R}^d$ borné et lipschitzien. Par définition, une partition $\mathcal{T}$ de Ω est une famille de domaines lipschiztien contenus dans $\overline{\Omega}$, deux à deux disjoints, et tels que

$$|\Omega \setminus (\cup_{T \in \mathcal{T}} T)| = 0, \tag{3.1}$$

où $|A|$ désigne la mesure de Lebesgue d'un ensemble mesurable $A \subset \mathbb{R}^d$. On a en particulier

$$\overline{\Omega} = \overline{\cup_{T \in \mathcal{T}} T}. \tag{3.2}$$

Pour chaque élément T on notera

$$h_T := \sup_{x,y \in T} \|x - y\|, \tag{3.3}$$

son diamètre. On appelle *finesse* de la partition T la quantité

$$h_{\mathcal{T}} = \max_{T \in \mathcal{T}} h_T. \tag{3.4}$$

Pour un entier $k \geq 0$ donné, on note

$$\mathbb{P}_k = \mathrm{Vect}\{x_1^{\alpha_1} \cdots x_d^{\alpha_d} \ ; \ 0 \leq |\alpha| \leq k\}, \tag{3.5}$$

l'espace des polynômes de degré total inférieur à k en dimension d, dont la dimension est donnée (exercice) par la formule

$$\dim(\mathbb{P}_k) = \binom{k+d}{d} = \frac{1}{d!}(k+1)(k+2)\cdots(k+d). \tag{3.6}$$

Etant donné une partition $\mathcal{T}$ de Ω, et un entier $k \geq 0$, on note

$$X_{\mathcal{T},k} := \{g \ : \ g_{|T} \in \mathbb{P}_k, \ T \in \mathcal{T}\}, \tag{3.7}$$

l'espace des fonctions polynomiales de degré k par morceaux sur chaque élément de $\mathcal{T}$. La dimension de cet espace est donc

$$\dim(X_{\mathcal{T},k}) = \#(\mathcal{T})\dim(\mathbb{P}_k), \tag{3.8}$$

où par définition $\#(E)$ désigne le cardinal d'un ensemble fini E. En pratique on fixe le degré k et on considère des partitions dont la finesse tend vers 0. On peut ainsi considérer une famille de partition

$$(\mathcal{T}_h)_{h\in S}, \tag{3.9}$$

avec $h := h_{\mathcal{T}_h} = \max_{T\in\mathcal{T}_h} h_T$, où S est une suite de réels strictement positifs qui tend vers 0, et on y associe la famille des espaces

$$(X_h)_{h\in S}, \quad X_h := X_{\mathcal{T}_h,k}. \tag{3.10}$$

On désigne parfois plus simplement ces familles par $(\mathcal{T}_h)_{h>0}$ et $(X_h)_{h>0}$, tout en sachant que toutes les valeurs de h ne sont pas nécessairement représentées. Par consistance avec les notations introduites dans le chapitre, on pourra aussi considérer la famille

$$(X_n)_{n\geq 1}, \quad X_n := X_h, \quad n = \dim(X_h) = \#(\mathcal{T}_h)\dim(\mathbb{P}_k), \tag{3.11}$$

avec $n \to +\infty$ lorsque $h \to 0$, tout en sachant que toutes les valeurs de n ne sont pas nécessairement représentées.

Dans le cas $k = 0$ qui correspond à l'approximation par les fonctions constantes par morceaux, on peut établir de façon élémentaire un premier résultat d'approximation en norme L^∞. Rappelons qu'un domaine T est dit *étoilé* si et seulement si il existe $x_T \in T$ tel pour tout $y \in T$ le segment $[x_T, y]$ est contenu dans T. En particulier tout convexe est étoilé.

Théorème 3.1 *Soit $\Omega \subset \mathbb{R}^d$ et $\mathcal{T}$ une partition de Ω dont tous les éléments $T \in \mathcal{T}$ sont des étoilés. Alors pour tout $f \in W^{1,\infty}(\Omega)$, on a*

$$\inf_{g\in X_{\mathcal{T},0}} \|f-g\|_{L^\infty} \leq Ch_{\mathcal{T}}|f|_{W^{1,\infty}} \tag{3.12}$$

avec $C = \sqrt{d}$.

Preuve : On considére le cas où $f \in C^1(\Omega)$. Pour tout $T \in \mathcal{T}$, on considère $x_T \in T$ tel pour tout $y \in T$ le segment $[x_T, y]$ est contenu dans T. On peut écrire pour tout $y \in T$, avec $F(t) := f(x_T + t(y - x_T))$

$$\begin{aligned}
|f(y) - f(x_T)| &= |F(1) - F(0)| = |\int_0^1 F'(t)dt| \\
&= |\int_0^1 (y - x_T)\cdot\nabla f(x_T + t(y - x_T))dt| \\
&\leq \|x_T - y\|\|\nabla f\|_{L^\infty(T)} \leq \sqrt{d}h_{\mathcal{T}}|f|_{W^{1,\infty}}.
\end{aligned}$$

En posant $g := \sum_{T\in\mathcal{T}} f(x_T)\chi_T \in X_{\mathcal{T},0}$, on a donc $\|f - g\|_{L^\infty} \leq h_{\mathcal{T}}|f|_{W^{1,\infty}}$. La généralisation à une fonction $f \in W^{1,\infty}(\Omega)$ fait appel à l'argument de densité suivant que l'on peut démontrer par régularisation (exercice) : pour tout $f \in W^{1,\infty}(\Omega)$ et $\delta > 0$ il existe une suite $(f_n)_{n\geq 0}$ de fonctions de $\mathcal{D}(\Omega)$ qui converge uniformément vers f sur $\Omega_\delta := \{x \in \Omega \ : \ B(x,\delta) \subset \Omega\}$ et telle que $\|\nabla f_n\|_{L^\infty(\Omega_\delta)} \leq \|\nabla f\|_{L^\infty(\Omega)}$ pour tout $x \in \Omega_\delta$ (on rappelle que $\mathcal{D}(\overline{\Omega})$ n'est pas dense dans $W^{1,\infty}(\Omega)$). $\diamond$

Il est possible de montrer (exercice assez technique) que tout domaine lipschitzien admet des partitions arbitrairement fines en sous-domaines étoilés. Le théorème 3.1 montre que si Ω est un domaine lipschitzien, $(\mathcal{T}_h)_{h>0}$ une famille de partitions de Ω en domaines étoilés et $(X_h)_{h>0}$ la famille des espaces de fonctions constantes par morceaux associée à ces partitions, on a, pour tout $f \in W^{1,\infty}(\Omega)$,

$$\inf_{g\in X_h} \|f-g\|_{L^\infty} \leq h|f|_{W^{1,\infty}} \tag{3.13}$$

Notons que si $f \in W^{s,\infty}(\Omega)$ pour $0 < s < 1$, c'est à dire Hölderienne d'exposant s, un argument similaire à celui de la preuve du théorème 3.1 conduit à l'estimation

$$\inf_{g\in X_h} \|f-g\|_{L^\infty} \leq h^s|f|_{W^{s,\infty}}. \tag{3.14}$$

Afin d'exprimer ces estimations en fonction de la dimension n des espaces d'approximation, on se restreint à des partitions $\mathcal{T}_h$ où la taille de *tous* les éléments est de l'ordre de h. Plus précisément, on définit la *rondeur* d'un élément T par

$$\rho_T := \sup\{2r \,:\, \exists x \in T,\, B(x,r) \subset T\}, \tag{3.15}$$

c'est à dire le diamètre maximal d'une boule contenue dans $\overline{T}$.

Définition 3.1 *Une famille de partition $(X_h)_{h>0}$ d'un domaine $\Omega \subset \mathbb{R}^d$ est dite quasi-uniforme si et seulement si il existe une constante $c > 0$ telle que, pour tout $h > 0$ et $T \in \mathcal{T}_h$,*

$$ch \leq \rho_T \leq h_T \leq h. \tag{3.16}$$

De cette propriété, on déduit facilement qu'il existe une constante a qui dépend de c et de la dimension d telle que, pour tout $h > 0$ et $T \in \mathcal{T}_h$,

$$ah^d \leq |T| \leq h^d. \tag{3.17}$$

Puisque $\sum_{T\in\mathcal{T}_h} |T| = |\Omega|$, ceci permet d'estimer le cardinal de $\mathcal{T}_h$ par

$$|\Omega| h^{-d} \leq \#(\mathcal{T}_h) \leq B h^{-d}, \tag{3.18}$$

avec $B = \frac{|\Omega|}{a}$. En posant $X_n = X_h$ avec

$$n = \dim(X_h) = \#(\mathcal{T}_h) \leq B h^{-d}, \tag{3.19}$$

on déduit de l'estimation (3.14) que, pour $0 < s \leq 1$,

$$\sigma_n(f)_{L^\infty} \leq C |f|_{W^{s,\infty}} n^{-s/d}, \tag{3.20}$$

où $C = B^{s/d}$. On voit ainsi que pour l'approximation en norme L^∞ par les fonctions constantes par morceaux sur des partitions quasi-uniformes, on a $W^{s,\infty}(\Omega) \subset \mathcal{A}^{s/d}$ avec la terminologie de (1.5).

Afin d'augmenter l'ordre de précision, ici limité par $s \leq 1$, il faut utiliser des fonctions polynomiales par morceaux de degré plus élevé. On peut d'autre part choisir d'étudier l'erreur d'approximation en norme L^p avec $1 \leq p \leq \infty$. Le point de départ de cette étude plus générale est le résultat fondamental suivant.

Théorème 3.1 *Si R est un domaine borné lipschitzien connexe de $\mathbb{R}^d$, m un entier strictement positif, alors pour tout $f \in W^{m,p}(R)$ telle que $\int_R f\pi = 0$ pour tout $\pi \in \mathbb{P}_{m-1}$, on a*

$$\|f\|_{W^{m,p}} \leq C |f|_{W^{m,p}}. \tag{3.21}$$

où C dépend que de R, m et p.

Preuve : On raisonne par contradiction. Si le résultat était faux, il existerait une suite $(f_n)_{n>0}$ de fonctions de $W^{m,p}(R)$ telle que :

(i) $\int_R f_n \pi = 0$ pour tout $\pi \in \mathbb{P}_{m-1}$,

(ii) $\|f_n\|_{W^{m,p}} = 1$ pour tout $n > 0$,

(iii) $|f_n|_{W^{m,p}}$ tends vers 0 quand $n \to +\infty$.

Par le théorème de Rellich, quitte à extraire une sous-suite, on peut supposer que f_n converge vers une limite f dans $W^{m-1,p}(R)$. Puisque $|f_n|_{W^{m,p}}$ tends vers 0 on en déduit que $(f_n)_{n>0}$ est une suite de Cauchy dans $W^{m,p}(R)$ et qui converge donc vers f dans $W^{m,p}(R)$. Par passage à la limite on voit que $|f|_{W^{m,p}} = 0$, $\int_R f\pi = 0$ pour tout $\pi \in \mathbb{P}_{m-1}$, et $\|f\|_{W^{m,p}} = 1$. La première propriété montre que $f \in \mathbb{P}_{m-1}$ puisque R est connexe. D'après la deuxième propriété on a donc $f = 0$ ce qui est en contradiction avec la troisième propriété. $\diamond$

Remarque 3.1 *On remarque que ce résultat est faux pour un ouvert non-connexe, en prenant une fonction constante par morceaux sur chaque composante et de moyenne nulle.*

Pour k fixé, désignons par P_R la projection orthogonale sur $\mathbb{P}_k$ pour la norme $L^2(R)$. Cette projection s'écrit

$$P_R f = \sum_{j=1}^{K} \Big(\int_R f \xi_j \Big) \xi_j, \tag{3.22}$$

où $\{\xi_1, \dots, \xi_K\}$ est une base orthonormée de $\mathbb{P}_k$ pour la norme $L^2(R)$, avec $K = \dim(\mathbb{P}_k)$. Elle peut donc s'appliquer à toute fonction $f \in L^1(R)$. Pour tout $f \in W^{m,p}(R)$, on a

$$\int_R (f - P_R f)\pi = 0, \quad \pi \in \mathbb{P}_k. \tag{3.23}$$

Par conséquent, le théorème précédent s'applique à la fonction $f - P_R f$ lorsque $k = m - 1$. On en déduit un résultat d'approximation par les polynômes : le *théorème de Deny-Lions.*

Théorème 3.2 *Si R est un domaine borné lipschitzien connexe de $\mathbb{R}^d$, m un entier positif, alors pour tout $f \in W^{m,p}(R)$ on a*

$$\min_{\pi \in \mathbb{P}_{m-1}} \|f - \pi\|_{W^{m,p}} \le C |f|_{W^{m,p}}. \tag{3.24}$$

où la constante C est la même que dans le théorème précèdent.

On remarque que puisque $\mathbb{P}_{m-1} \subset \mathbb{P}_k$ pour $k \ge m - 1$, le théorème de Deny-Lions donne aussi

$$\min_{\pi \in \mathbb{P}_k} \|f - \pi\|_{W^{m,p}} \le C |f|_{W^{m,p}}, \tag{3.25}$$

sous l'hypothèse $m \le k + 1$. Ceci entraîne que pour $n \le m \le k + 1$ on a

$$\min_{\pi \in \mathbb{P}_k} |f - \pi|_{W^{n,p}} \le C |f|_{W^{m,p}}, \tag{3.26}$$

et en particulier, pour $n = 0$, on obtient

$$\min_{\pi \in \mathbb{P}_k} \|f - \pi\|_{L^p} \le C |f|_{W^{m,p}}. \tag{3.27}$$

Nous allons utiliser ce résultat d'approximation sur chaque élément T d'une partition $\mathcal{T}_h$ pour en déduire un résultat d'approximation en norme L^p par les fonctions polynomiales par morceaux sur des familles de partitions. Pour cela il est important de comprendre comment la constante C dans l'estimation ci-dessus se comporte lorsque la taille de T varie.

Dans cet objectif, on se restreint à des familles de partitions où tous les éléments se déduisent par transformation affine d'un nombre fini *d'éléments de référence.* Une telle restriction est naturelle dans la pratique du calcul numérique où l'on utilise des partitions de Ω en éléments de formes élémentaires, par exemple des triangles et des parallélogrammes dans le cas de la dimension $d = 2$.

Définition 3.2 *Une famille $(\mathcal{T}_h)_{h>0}$ de partition d'un domaine $\Omega \subset \mathbb{R}^d$ est dite affine-invariante si et seulement si il existe une famille finie $\mathcal{R} := \{\widehat{T}_1, \dots, \widehat{T}_l\}$ de domaines de référence, telle que pour tout $h > 0$ et $T \in \mathcal{T}_h$, il existe $\widehat{T} \in \mathcal{R}$ et $A_T : \mathbb{R}^d \to \mathbb{R}^d$ une application affine et inversible telle que $T = A_T(\hat{T})$.*

Pour une telle famille et pour tout élément T, on peut écrire la transformation affine A_T sous la forme

$$A_T(x) = a_T + B_T x, \tag{3.28}$$

où B_T est une matrice $d \times d$ et $a_T \in \mathbb{R}^d$. Cette matrice est inversible et on a

$$|T| = |\det(B_T)|\, |\widehat{T}|. \tag{3.29}$$

Nous allons utiliser la transformation A_T pour transporter les objets d'intérêt depuis T vers $\widehat{T}$. Pour simplifier on notera systématiquement $\hat{q}$ la quantité obtenue par transport de q. On note ainsi

$$\hat{x} = A_T^{-1}(x) = B_T^{-1}(x - a_T) \Leftrightarrow x = A_T(\hat{x}). \tag{3.30}$$

Si f est une fonction définie sur T, on définit $\hat{f}$ sur $\widehat{T}$ par

$$\hat{f}(\hat{x}) = f(x) \Leftrightarrow \hat{f} = f \circ A_T. \tag{3.31}$$

On remarque que l'espace $\mathbb{P}_k$ des polynômes de degré k est invariant par cette transformation puisque A_T est affine.

Afin d'établir des résultats d'approximation, nous allons comparer les quantités $|f|_{W^{m,p}(T)}$ et $|\hat{f}|_{W^{m,p}(\widehat{T})}$. Lorsque $m = 0$, on a simplement

$$\begin{aligned} \|f\|^p_{L^p(T)} &= \int_T |f(x)|^p dx \\ &= \int_{\widehat{T}} |f(A_T(\hat{x}))|^p |\det(B_T)| d\hat{x} \\ &= \frac{|T|}{|\widehat{T}|} \int_{\widehat{T}} |\hat{f}(\hat{x})|^p d\hat{x}, \end{aligned}$$

ce qui donne

$$\|f\|_{L^p(T)} = C|T|^{1/p} \|\hat{f}\|_{L^p(\widehat{T})}, \tag{3.32}$$

avec $C = |\widehat{T}|^{-1/p}$. Pour les semi-normes d'ordre $m > 1$, nous allons d'abord estimer les normes de B_T et B_T^{-1} en fonction des paramètres géométriques de T.

Lemme 3.1 *Pour tout élément T, on a*

$$\|B_T\| \le C_1 h_T, \tag{3.33}$$

avec $C_1 = \rho_{\widehat{T}}^{-1}$ et

$$\|B_T^{-1}\| \le C_2 \rho_T^{-1}, \tag{3.34}$$

avec $C_2 = h_{\widehat{T}}$.

Preuve : Par définition, on a

$$\|B_T\| = \sup_{x \in \mathbb{R}^d \setminus \{0\}} \frac{\|B_T x\|}{\|x\|} = \sup_{\|x\| = \rho_{\widehat{T}}} \frac{\|B_T x\|}{\rho_{\widehat{T}}}, \tag{3.35}$$

et le sup est atteint. Soit $\hat{x}$ un vecteur qui réalise ce sup. Il existe donc deux points $\hat{a}$ et $\hat{b}$ dans la boule inscrite de $\overline{\widehat{T}}$ tels que $\hat{x} = \hat{a} - \hat{b}$. On a donc

$$B_T x = A_T(\hat{a}) - A_T(\hat{b}) = a - b, \tag{3.36}$$

où a et b sont dans $\overline{T}$ et par conséquent $\|B_T \hat{x}\| \le h_T$. L'autre inégalité se démontre de la même manière en échangeant les rôles de T et $\widehat{T}$. $\diamond$

Examinons à présent la semi-norme Sobolev $W^{m,p}$, tout d'abord dans le cas $m = 1$ et $p = 2$, c'est à dire la semi-norme H^1. On remarque que si $f(x) = \hat{f}(\hat{x})$ c'est à dire $\hat{f} = f \circ A_T$ on a

$$\nabla \hat{f}(\hat{x}) = B_T^t \nabla f(x), \tag{3.37}$$

où B_T^t est la transposée de B_T. Ceci nous montre que

$$\begin{aligned} |f|^2_{H^1(T)} &= \int_T \|\nabla f(x)\|^2 dx \\ &= \int_{\widehat{T}} \|(B_T^t)^{-1} \nabla \hat{f}(\hat{x})\|^2 |\det(B_T)| d\hat{x} \\ &\le C_2^2 \frac{|T|}{|\widehat{T}| \rho_T^2} \int_{\widehat{T}} \|\nabla \hat{f}(\hat{x})\|^2 d\hat{x}, \end{aligned}$$

et par conséquent

$$|f|_{H^1(T)} \leq C\frac{|T|^{1/2}}{\rho_T}|\hat{f}|_{H^1(\widehat{T})}, \tag{3.38}$$

où C dépend de $\widehat{T}$. En inversant le rôle de T et $\hat{T}$ dans le calcul, on obtient aussi

$$|\hat{f}|_{H^1(\widehat{T})} \leq C\frac{h_T}{|T|^{1/2}}|f|_{H^1(T)}. \tag{3.39}$$

Dans le cas plus général de la semi-norme $W^{1,p}$, on effectue un calcul de changement de variable similaire, en utilisant en plus les équivalences des norme ℓ^p dans $\mathbb{R}^d$

$$\|z\|_p \sim \|z\|_2 = \|z\|, \tag{3.40}$$

pour $1 \leq p \leq \infty$ où les constantes d'équivalence ne dépendent que de d. On aboutit ainsi à

$$|f|_{W^{1,p}(T)} \leq C\frac{|T|^{1/p}}{\rho_T}|\hat{f}|_{W^{1,p}(\widehat{T})}, \tag{3.41}$$

et

$$|\hat{f}|_{W^{1,p}(\widehat{T})} \leq C\frac{h_T}{|T|^{1/p}}|f|_{W^{1,p}(T)}, \tag{3.42}$$

où C ne dépend que de $\widehat{T}$ et d. Pour les dérivées d'ordre supérieur, on montre (exercice) que

$$(\sum_{|\alpha|=m}|D^\alpha f(x)|^p)^{1/p} \leq C\|(B_T^t)^{-1}\|^m(\sum_{|\alpha|=m}|D^\alpha \hat{f}(\hat{x})|^p)^{1/p}, \tag{3.43}$$

où C dépend de $\widehat{T}$, m, d et p. Par un changement de variable similaire au cas $m = 1$ on obtient ainsi

$$|f|_{W^{m,p}(T)} \leq C\frac{|T|^{1/p}}{\rho_T^m}|\hat{f}|_{W^{m,p}(\widehat{T})}, \tag{3.44}$$

et

$$|\hat{f}|_{W^{m,p}(\widehat{T})} \leq C\frac{h_T^m}{|T|^{1/p}}|f|_{W^{m,p}(T)}, \tag{3.45}$$

où C dépend de $\widehat{T}$, m, d et p.

Remarque 3.2 *Les constantes C qui apparaissent dans les inégalités ci-dessus dépendent de l'élément de référence $\widehat{T}$ associé à T. Pour une famille affine invariante $(\mathcal{T}_h)_{h>0}$, ces éléments appartiennent à un ensemble fini $\mathcal{R}$. On peut prendre pour C la constante maximale parmis tous les éléments de références, et les inégalités sont ainsi valables pour tout $T \in \mathcal{T}_h$ et $h > 0$.*

En combinant le théorème de Deny-Lions appliqué sur l'élément de référence $\hat{T}$ et ces résultats de changement de variable, nous pouvons obtenir des résultats d'approximation locale par les polynômes sur un élément quelconque $T \in \mathcal{T}_h$ pour une famille affine-invariante $(\mathcal{T}_h)_{h>0}$. Pour $n \leq m \leq k+1$, on peut en effet écrire

$$\begin{aligned}
\min_{\pi\in\mathbb{P}_k}|f-\pi|_{W^{n,p}(T)} &\leq C\frac{|T|^{1/p}}{\rho_T^n}\min_{\hat{\pi}\in\mathbb{P}_k}|\hat{f}-\hat{\pi}|_{W^{n,p}(\widehat{T})}\\
&\leq C\frac{|T|^{1/p}}{\rho_T^n}|\hat{f}|_{W^{m,p}(\widehat{T})}\\
&\leq C\frac{|T|^{1/p}}{\rho_T^n}\frac{h_T^m}{|T|^{1/p}}|f|_{W^{m,p}(T)},
\end{aligned}$$

où C varie d'une ligne à l'autre et dépend que de $\mathcal{R}$, m, n, d et p, soit

$$\min_{\pi\in\mathbb{P}_k}|f-\pi|_{W^{n,p}(T)} \leq C\frac{h_T^m}{\rho_T^n}|f|_{W^{m,p}(T)}. \tag{3.46}$$

Remarque 3.3 *On peut généraliser ces résultats d'approximation polynomiale en faisant appel aux injections de Sobolev : si $n < m \le k+1$ et $m - n > d/p - d/q$, on a sur tout domaine borné lipschitzien connexe R*

$$\begin{aligned}\min_{\pi\in\mathbb{P}_k} \|f-\pi\|_{W^{n,q}(R)} &\le C \min_{\pi\in\mathbb{P}_k} \|f-\pi\|_{W^{m,p}(R)} \\ &\le C|f|_{W^{m,p}(R)},\end{aligned}$$

où C dépend de R, m, n, p et q. En combinant cette estimation sur $\widehat{T}$ et les résultats de changement de variable, on obtient ainsi pour tout $T \in \mathcal{T}_h$ et $h > 0$,

$$\min_{\pi\in\mathbb{P}_k} |f-\pi|_{W^{n,q}(T)} \le C|T|^{1/q-1/p}\frac{h_T^m}{\rho_T^n}|f|_{W^{m,p}(T)}. \tag{3.47}$$

où C dépend de $\mathcal{R}$, m, n, d, p et q.

Dans le cas $n = 0$, l'estimation (3.46) s'écrit, si $0 < m \le k+1$,

$$\min_{\pi\in\mathbb{P}_k} \|f-\pi\|_{L^p(T)} \le Ch_T^m|f|_{W^{m,p}(T)}. \tag{3.48}$$

où C dépend de $\mathcal{R}$, m, d et p.

En sommant cette estimation élevée à la puissance p sur les différents éléments $T \in \mathcal{T}_h$, on obtient pour $1 \le p < \infty$, une estimation globale d'approximation en norme L^p pour l'espace X_h :

$$\begin{aligned}\min_{g\in X_h} \|f-g\|^p_{L^p(\Omega)} &= \sum_{T\in\mathcal{T}_h} \min_{\pi\in\mathbb{P}_k} \|f-\pi\|^p_{L^p(T)} \\ &\le C^p \sum_{T\in\mathcal{T}_h} h_T^{mp}|f|^p_{W^{m,p}(T)} \\ &\le C^p h^{mp}|f|^p_{W^{m,p}(\Omega)},\end{aligned}$$

soit

$$\min_{g\in X_h} \|f-g\|_{L^p(\Omega)} \le Ch^m|f|_{W^{m,p}(\Omega)}, \tag{3.49}$$

où C dépend de $\mathcal{R}$, m, d et p. Notons que le minimum g dans cette estimation globale est donné par

$$g := \sum_{T\in\mathcal{T}_h} \pi_T\chi_T, \tag{3.50}$$

où π_T est le polynôme qui donne le minimum dans l'estimation locale (3.48). Dans le cas $p = \infty$, on obtient aussi (3.49) en prenant le max dans (3.48) sur tous les $T \in \mathcal{T}_h$. L'estimation (3.49) généralise (3.13) qui correspond au cas $k = 0$, $m = 1$ et $p = \infty$.

Il est aussi possible d'étendre cette généralisation aux espaces de Sobolev d'indice de régularité non-entier : si $0 < s \le k+1$, on a

$$\min_{g\in X_h} \|f-g\|_{L^p} \le Ch^s|f|_{W^{s,p}}, \tag{3.51}$$

mais la preuve est beaucoup plus difficile, et hors du cadre de ce cours.

Dans le cas de familles de partitions affine-invariantes et quasi-uniformes, en posant $X_n = X_h$ avec

$$n = \dim(X_h) = \dim(\mathbb{P}_k)\#(\mathcal{T}_h) \le \dim(\mathbb{P}_k)Bh^{-d}, \tag{3.52}$$

l'estimation (3.51) peut s'écrire

$$\sigma_n(f)_{L^p} \le C|f|_{W^{s,p}}n^{-s/d}, \tag{3.53}$$

où C dépend de $\mathcal{R}$, s, d, p, $|\Omega|$ et de la constante c dans (3.16). On voit ainsi que pour l'approximation en norme L^p par les fonctions polynomiales par morceaux de degré $k \ge s-1$ sur des partitions affine-invariantes et quasi-uniformes, on a $W^{s,p}(\Omega) \subset \mathcal{A}^{s/d}$ avec la terminologie de (1.5).

Remarque 3.4 *Contrairement à l'estimation* (3.48)*, on ne peut pas agréger les estimations* (3.46) *pour $n > 0$. En effet la fonction polynomiale par morceau g définie par* (3.50) *n'appartient généralement pas à $W^{n,p}$, à cause des discontinuités entre les différents π_T aux interfaces entre les éléments (exercice). Dans la méthodes des éléments finis de Lagrange, étudiée plus loin, la continuité est imposée, ce qui permet d'avoir des estimations pour $n = 1$.*

Nous terminons cette étude par quelques remarques sur la projection orthogonale sur les espaces $X_h = X_{\mathcal{T}_h,k}$ associés à une famille affine-invariante $(\mathcal{T}_h)_{h>0}$. On notera P_h ces projecteurs. Par application de (3.50) dans le cas $p = 2$, on a donc

$$P_h f := \sum_{T \in \mathcal{T}_h} \pi_T \chi_T, \quad \pi_T := P_T f, \tag{3.54}$$

où P_T est l'opérateur de projection orthogonale dans $L^2(T)$ sur $\mathbb{P}_k$.

On remarque que $\pi_T = P_T f$ est caractérisé par

$$\int_T \pi_T \pi = \int_T f\pi, \quad \pi \in \mathbb{P}_k, \tag{3.55}$$

ce qui par changement de variable est équivalent à

$$\int_{\widehat{T}} \widehat{\pi_T} \hat{\pi} = \int_T \hat{f} \hat{\pi}, \quad \hat{\pi} \in \mathbb{P}_k. \tag{3.56}$$

ce qui montre que $\widehat{P_T f} = P_{\widehat{T}} \widehat{f}$. En d'autres termes, on a la relation de commutation

$$P_T f \circ A_T = P_{\widehat{T}}(f \circ A_T). \tag{3.57}$$

entre la projection orthogonale et le transport d'une fonction par changement de variable.

On peut étudier la stabilité en norme L^p du projecteur orthogonal. Pour $p = 2$, on a

$$\|P_h\|_{\mathcal{L}(L^2(\Omega))} = 1. \tag{3.58}$$

Pour $1 \le p \le \infty$ quelconque, on va montrer que les projecteurs P_h sont uniformément stables en norme L^p. Pour cela, on se place sur un élément de référence $\hat{T}$, et on peut écrire

$$P_{\widehat{T}} \widehat{f} = \sum_{j=1}^{K} \Big(\int_{\widehat{T}} \hat{f} \xi_j \Big) \xi_j. \tag{3.59}$$

où $\{\xi_1, \dots, \xi_K\}$ est une base orthonormée de $\mathbb{P}_k$ pour la norme $L^2(\widehat{T})$ avec $K = \dim(\mathbb{P}_k)$. Par conséquent

$$\begin{aligned} \|P_{\widehat{T}} \widehat{f}\|_{L^p(\widehat{T})} &\le \sum_{j=1}^{K} \Big| \int_{\widehat{T}} \hat{f} \xi_j \Big| \|\xi_j\|_{L^p(\widehat{T})} \\ &\le \sum_{j=1}^{K} \|\hat{f}\|_{L^p(\widehat{T})} \|\xi_j\|_{L^{p'}(\widehat{T})} \|\xi_j\|_{L^p(\widehat{T})}, \end{aligned}$$

où p' est tel que $\frac{1}{p} + \frac{1}{p'} = 1$. Ceci montre que

$$\|P_{\widehat{T}} \hat{f}\|_{L^p(\widehat{T})} \le C \|\hat{f}\|_{L^p(\widehat{T})}, \tag{3.60}$$

où $C := \sum_{j=1}^{K} \|\xi_j\|_{L^{p'}(\widehat{T})} \|\xi_j\|_{L^p(\widehat{T})}$ dépend de $\widehat{T}$, k, d et p. Par changement de variable, et en utilisant la commutation entre la projection orthogonale et le transport d'une fonction par changement de variable, on obtient pour tout élément $T \in \mathcal{T}_h$,

$$\begin{aligned} \|P_T f\|_{L^p(T)} &= \det(B_T)^{1/p} \|\widehat{P_T f}\|_{L^p(\widehat{T})} \\ &= \det(B_T)^{1/p} \|P_{\widehat{T}} \hat{f}\|_{L^p(\widehat{T})} \\ &\le C \det(B_T)^{1/p} \|\hat{f}\|_{L^p(\widehat{T})} \\ &= C \|f\|_{L^p(T)}, \end{aligned}$$

où C dépend de $\mathcal{R}$, k, d et p. En sommant ces estimations élevées à la puissance p sur tous les éléments $T \in \mathcal{T}_h$ (ou en prenant le max lorsque $p = \infty$) on obtient l'estimation de stabilité uniforme

$$\|P_h f\|_{L^p(\Omega)} \le C \|f\|_{L^p(\Omega)}. \tag{3.61}$$

Par le raisonnement invoqué à la fin de la section §1, on peut donc écrire

$$\|f - P_h f\|_{L^p} \leq C \inf_{g \in X_h} \|f - g\|_{L^p}, \tag{3.62}$$

où C dépend de $\mathcal{R}$, k, d et p. En combinant ceci avec l'estimation d'erreur de meilleure approximation (3.51), on obtient une estimation similaire pour l'erreur de projection c'est à dire

$$\|f - P_h f\|_{L^p} \leq C h^s |f|_{W^{s,p}}; \tag{3.63}$$

où C dépend de $\mathcal{R}$, k, s, d et p.

4 Approximation par les séries de Fourier

Les séries de Fourier sont adaptées à la représentation des fonctions périodiques. En dimension $d = 1$, pour les fonctions de période 2π, elles peuvent s'écrire de façon équivalente sous la forme réelle

$$f(x) = \sum_{k \geq 0} a_k \cos(kx) + \sum_{k \geq 1} b_k \sin(kx), \tag{4.1}$$

ou la forme complexe

$$f(x) = \sum_{k \in \mathbb{Z}} c_k e^{ikx}. \tag{4.2}$$

On adoptera ici la forme complexe, plus simple, tout en remarquant que $c_{-k} = \overline{c}_k$ dans le cas où f est réelle. On rappelle que toute fonction $f \in L^2([-\pi, \pi])$ admet la décomposition ci-dessus avec

$$c_k = c_k(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx. \tag{4.3}$$

Cette décomposition s'écrit aussi

$$f = \sum_{k \in \mathbb{Z}} \langle f, e_k \rangle e_k, \quad e_k(x) := \frac{1}{\sqrt{2\pi}} e^{ikx}, \tag{4.4}$$

où $\langle \cdot, \cdot \rangle$ est le produit scalaire de $L^2([-\pi, \pi])$ et apparaît ainsi comme la représentation de f dans la base orthonormale $(e_k)_{k \in \mathbb{Z}}$. Il faut dans ce cas entendre la convergence des séries (4.1) ou (4.2) au sens L^2. L'étude générale de la convergence des série de Fourier, en particulier au sens ponctuel, L^p ou de la norme de la convergence uniforme, est à l'origine de nombreux développements mathématiques depuis le début du XIXème siècle.

Dans le cas de fonctions multidimensionelles et 2π-périodiques en chaque variable, c'est à dire définies sur le tore $T = [-\pi, \pi[^d$, les séries de Fourier s'écrivent

$$f(x) = \sum_{k \in \mathbb{Z}^d} c_k e^{ik \cdot x}, \quad c_k = c_k(f) := (2\pi)^{-d} \int_T f(x) e^{-ik \cdot x} dx, \tag{4.5}$$

où $k \cdot x = k_1 x_1 + \cdots + k_d x_d$ si $x = (x_1, \ldots, x_d)$ et $k = (k_1, \ldots, k_d)$, ou de manière équivalente

$$\sum_{k \in \mathbb{Z}^d} \langle f, e_k \rangle e_k, \quad e_k(x) := (2\pi)^{-d/2} e^{ik \cdot x}. \tag{4.6}$$

On peut facilement adapter les séries de Fourier à des fonctions de période quelconque en chaque variable. Par exemple, en dimension 1, pour des fonctions de période F, il suffite de modifier la définition de e_k suivant $e_k := \frac{1}{\sqrt{F}} e^{i \frac{\pi k \cdot x}{F}}$. L'analyse étant tout à fait similaire, on considére ici uniquement le cas $F = 2\pi$ qui a l'avantage de simplifier les notations.

L'égalité de Parseval

$$(2\pi)^{-d} \int_T |f(x)|^2 dx = \sum_{k \in \mathbb{Z}^d} |c_k(f)|^2, \tag{4.7}$$

exprime une isométrie ente $L^2(T)$ et $\ell^2(\mathbb{Z}^d)$: pour tout $f \in L^2(T)$, la suite $(c_k(f))_{k\in\mathbb{Z}^d}$ appartient à $\ell^2(\mathbb{Z}^d)$, et réciproquement, pour toute suite $(c_k)_{k\in\mathbb{Z}^d} \in \ell^2(\mathbb{Z}^d)$, la série de Fourier $\sum_{k\in\mathbb{Z}^d} c_k e^{ik\cdot x}$ converge dans $L^2(T)$ vers une fonction $f \in L^2(T)$.

La question de la précision l'approximation des fonctions par les *sommes partielles* obtenues en tronquant leur série de Fourier est au coeur de nombreux problèmes d'analyse théorique et appliquée. En dimension $d \geq 1$ quelconque, on considérera ici les sommes partielles

$$P_K f(x) = \sum_{\|k\|_\infty \leq K} c_k(f) e^{ik\cdot x} = \sum_{\|k\|_\infty \leq K} \langle f, e_k \rangle e_k, \tag{4.8}$$

où $\|k\|_\infty = \max_{i=1,\dots,d} |k_i|$, c'est à dire la projection orthogonale sur l'espace des *polynômes trigonométriques* de degré K en chaque variable

$$\mathbb{T}_K := \mathrm{Vect}\{e_k \ : \ \|k\|_\infty \leq K\}. \tag{4.9}$$

Il existe bien entendu d'autres manières de tronquer les séries de Fourier, par exemple en remplacant $\|k\|_\infty$ par $\|k\|_1$ ou $\|k\|$.

Afin d'étudier l'erreur d'approximation et la relier à la régularité de la fonction f, on introduit la version périodique des espaces de Sobolev $W^{m,p}$ dont la définition est très similaire.

Définition 4.1 *Soit $1 \leq p \leq \infty$ et m un entier positif. Une fonction $f \in L^p(T)$ appartient à $W^{m,p}_{per}(T)$ si et seulement si pour tout $|\alpha| \leq m$, il existe $f_\alpha \in L^p(T)$ telle que*

$$\int_T f_\alpha \varphi = (-1)^{|\alpha|} \int_T f D^\alpha \varphi, \tag{4.10}$$

pour tout $\varphi \in C^\infty_{per}(T)$ où $C^\infty_{per}(T)$ est l'ensemble des fonctions C^∞ sur $\mathbb{R}^d$ et de période 2π en chaque variable.

Comme pour les espaces de Sobolev non-périodiques on note alors $f_\alpha = D^\alpha f$, on définit de la même manière les normes et semi-normes $W^{m,p}$ et on note $H^m_{per} = W^{m,2}_{per}$. Les propriétés des espaces de Sobolev périodiques (densité des fonctions $C^\infty_{per}(T)$, injections de Sobolev, théorème de Rellich) s'énoncent de manière identique à celles des espaces non-périodiques. On peut vérifier que $W^{m,p}_{per}(T) \subset W^{m,p}(T)$ et que cette inclusion est stricte lorsque $m \geq 1$ à cause de la condition de périodicité imposée par la définition (exercice). On a cependant $W^{0,p}_{per}(T) = W^{0,p}(T) = L^p(T)$.

En appliquant la définition ci-dessus à la fonction $\varphi(x) = e^{-ik\cdot x}$, on obtient la relation importante

$$c_k(D^\alpha f) = i^{|\alpha|} k^\alpha c_k(f), \quad k^\alpha := k_1^{\alpha_1} \cdots k_d^{\alpha_d}, \tag{4.11}$$

qui permet de relier la régularité de f aux propriétés de décroissance de ses coefficients de Fourier. En particulier, l'égalité de Parseval nous indique qu'une fonction f appartient à l'espace de Sobolev $H^m_{per}(T)$ si et seulement si

$$\sum_{k\in\mathbb{Z}^d} |k^\alpha|^2 |c_k|^2 < \infty, \quad |\alpha| \leq m, \quad c_k = c_k(f) \tag{4.12}$$

et la norme H^m_{per} s'écrit

$$\|f\|^2_{H^m_{per}} = (2\pi)^d \sum_{k\in\mathbb{Z}^d} \Big(\sum_{|\alpha|\leq m} |k^\alpha|^2 \Big) |c_k|^2. \tag{4.13}$$

Il est facile de vérifier que

$$1 + \|k\|_\infty^{2m} \leq \sum_{|\alpha|\leq m} |k^\alpha|^2 \leq C(1 + \|k\|_\infty^{2m}), \tag{4.14}$$

où la constante C dépend de m et d. Par conséquent, f appartient à $H^m_{per}(T)$ si et seulement si

$$\sum_{k\in\mathbb{Z}^d} \|k\|_\infty^{2m} |c_k|^2 < \infty, \tag{4.15}$$

et on a l'équivalence de norme

$$(2\pi)^d \sum_{k\in\mathbb{Z}^d} (1+\|k\|_\infty^{2m})|c_k|^2 \leq \|f\|^2_{H^m_{per}} \leq C(2\pi)^d \sum_{k\in\mathbb{Z}^d} (1+\|k\|_\infty^{2m})|c_k|^2. \tag{4.16}$$

Ceci permet définir les espaces de Sobolev $H^s_{per}(T)$ pour un réel s quelconque, à travers la norme

$$\|f\|^2_{H^s_{per}} := (2\pi)^d \sum_{k\in\mathbb{Z}^d} (1+\|k\|_\infty^{2s})|c_k|^2, \tag{4.17}$$

ainsi que la semi-norme

$$|f|^2_{H^s_{per}} := (2\pi)^d \sum_{k\in\mathbb{Z}^d} \|k\|_\infty^{2s}|c_k|^2. \tag{4.18}$$

On obtient bien entendu une norme équivalente si l'on remplace $\|\cdot\|_\infty$ par toute autre norme sur $\mathbb{R}^d$, en particulier la norme euclidienne.

La caractérisation par les séries de Fourier des espaces de Sobolev $H^s_{per}(T)$ nous permet d'établir un résultat d'approximation élémentaire par les sommes partielles de Fourier en norme L^2.

Théorème 4.1 *Soit $s > 0$. Pour tout $f \in H^s_{per}(T)$, on a*

$$\|f - P_K f\|_{L^2} \leq K^{-s}|f|_{H^s_{per}}, \quad K > 0. \tag{4.19}$$

Preuve : En utilisant Parseval, on écrit avec $c_k = c_k(f)$,

$$\begin{aligned}\|f - P_K f\|^2_{L^2} &= (2\pi)^d \sum_{\|k\|_\infty > K} |c_k|^2 \\ &\leq (K+1)^{-2s}(2\pi)^d \sum_{k\in\mathbb{Z}^d} \|k\|_\infty^{2s}|c_k|^2 \\ &\leq K^{-2s}|f|^2_{H^s_{per}},\end{aligned}$$

qui est le résultat annoncé. ◇

En posant $X_n = \mathbb{T}_K$ avec

$$n = \dim(\mathbb{T}_K) = (2K+1)^d. \tag{4.20}$$

on déduit de l'estimation (4.19) que, pour $s > 0$,

$$\sigma_n(f)_{L^2} \leq C|f|_{H^s_{per}} n^{-s/d}, \tag{4.21}$$

où la constante C dépend de d et s. On voit ainsi que pour l'approximation en norme L^2 par les polynômes trigonométriques on a $H^s_{per}(T) \subset \mathcal{A}^{s/d}$ avec la terminologie de (1.5). On remarque l'analogie avec les estimations d'approximation obtenues pour les fonctions polynomiales de degré k par morceaux sur des partitions quasi-uniformes. Cependant, pour ces dernières, la vitesse de convergence est limitée par la condition $s-1 \leq k$, qui traduit le fait qu'on utilise une méthode *d'ordre* $k+1$. Dans le cas des polynômes trigonométriques, il n'y a pas de limitation de ce type et on parle alors de convergence *spectrale*.

Remarque 4.1 *L'approximation par les polynômes trigonométriques permet aussi d'obtenir des estimations d'erreur dans des normes plus régulières que L^2, ce qui n'était pas possible avec les fonctions polynomiales par morceaux qui ne sont pas suffisament régulières, voir Remarque 3.4. Plus précisément, si $0 \leq t \leq s$ et $f \in H^s_{per}(T)$, on peut écrire*

$$\begin{aligned}|f - P_K f|^2_{H^t_{per}} &= (2\pi)^d \sum_{\|k\|_\infty > K} \|k\|_\infty^{2t}|c_k|^2 \\ &\leq (K+1)^{2(t-s)}(2\pi)^d \sum_{k\in\mathbb{Z}^d} \|k\|_\infty^{2s}|c_k|^2 \\ &\leq CK^{-2(s-t)}|f|^2_{H^s_{per}},\end{aligned}$$

où C dépend de s et t. En modifiant la constante C, on obtient ainsi les estimations

$$\|f - P_K f\|_{H^t_{per}} \leq C K^{-(s-t)} |f|_{H^s_{per}}, \quad K > 0, \tag{4.22}$$

ou encore

$$\sigma_n(f)_{H^t_{per}} \leq C |f|_{H^s_{per}} n^{-(s-t)/d}. \tag{4.23}$$

Pour l'approximation en norme H^t_{per} par les polynômes trigonométriques on a donc $H^s_{per}(T) \subset \mathcal{A}^{(s-t)/d}$.

Nous cherchons à présent à établir des estimations d'approximation par les polynômes trigonométriques en norme L^p pour $1 \leq p \leq \infty$. Une difficulté est que pour $p \neq 2$ on ne peut pas caractériser de façon simple l'espace $L^p(T)$ par une propriété portant sur les coefficients de Fourier. En particulier, il n'existe pas d'analogue à la formule de Parseval pour la norme L^p lorsque $p \neq 2$.

Afin d'analyse l'action du projecteur orthogonal P_K du point de vue de la norme L^p, il est utile de l'identifier à un opérateur de *convolution*. En dimension $d = 1$, on peut en effet écrire

$$P_K f(x) = \sum_{k=-K}^{K} \Big(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(y) e^{-iky} dy\Big) e^{ikx} = \int_{-\pi}^{\pi} f(y) \frac{1}{2\pi} (\sum_{k=-K}^{K} e^{ik(x-y)}) dy, \tag{4.24}$$

soit

$$P_K f(x) = f * d_K(x) = \int_{-\pi}^{\pi} f(y) d_K(x-y) dy, \tag{4.25}$$

où

$$d_K(z) := \frac{1}{2\pi} \sum_{k=-K}^{K} e^{ikz} = \frac{\sin((2K+1)z/2)}{2\pi \sin(z/2)}, \tag{4.26}$$

est appelé *noyau de Dirichlet*. En dimension supérieure, un calcul analogue donne

$$P_K f(x) = \int_T f(y) D_K(x-y) dy, \tag{4.27}$$

avec

$$D_K(z) = \prod_{j=1}^{d} d_K(z_j), \quad z = (z_1, \dots, z_d). \tag{4.28}$$

D'après les remarques générales énoncées à la fin de la section §1, la convergence en norme L^p de $P_K f$ pour tout $f \in L^p(T)$ nécessite d'avoir la stabilité uniforme des opérateurs P_K en norme L^p. Afin de vérifier une telle propriété, on peut utiliser l'inégalité de Young pour les convolutions qui affirme que

$$\|f * g\|_{L^r} \leq \|f\|_{L^p} \|g\|_{L^q}, \quad 1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}, \tag{4.29}$$

aussi bien pour les espaces de Lebesgue sur $\mathbb{R}^d$ que sur le tore T. On a donc en particulier

$$\|P_K f\|_{L^p(T)} \leq \|f\|_{L^p(T)} \|D_K\|_{L^1(T)}, \tag{4.30}$$

c'est à dire

$$\|P_K\|_{\mathcal{L}(L^p)} \leq \|D_K\|_{L^1}. \tag{4.31}$$

Dans le cas de la norme L^∞, cette borne est optimale ce qui signifie que l'on a (exercice)

$$\|P_K\|_{\mathcal{L}(L^\infty)} = \|D_K\|_{L^1}. \tag{4.32}$$

On note que $\|D_K\|_{L^1(T)} = \|d_K\|^d_{L^1([-\pi,\pi[)}$. Il est facile de montrer que $\|d_K\|_{L^1}$ n'est pas uniformément borné et croît logarithmiquement avec K, suivant

$$c \log(K) \leq \|d_K\|_{L^1([-\pi,\pi[)} \leq C \log(K), \quad K \geq 2, \tag{4.33}$$

pour deux constantes $0 < c \leq C$, ce qui entraîne

$$c^d \log(K)^d \leq \|D_K\|_{L^1(T)} \leq C^d \log(K)^d, \quad K \geq 2. \tag{4.34}$$

On voit ainsi que les projecteurs P_K ne sont pas uniformément stable pour la norme L^∞. Afin de contourner cette difficulté, on peut modifier le procédé d'approximation dans $\mathbb{T}_K$ en remplacant les projecteurs P_K par des opérateur plus stables. On remarque que P_K peut s'écrire

$$P_K f(x) = \sum_{k\in\mathbb{Z}^d} \chi_{[-K,K]^d}(k) c_k(f) e^{ik\cdot x} = \sum_{k\in\mathbb{Z}^d} a\Big(\frac{k}{K}\Big) c_k(f) e^{ik\cdot x}, \tag{4.35}$$

où $a = \chi_{[-1,1]^d}$. L'idée est alors de définir un nouveau procédé de sommabilité en remplacant a par une fonction plus régulière afin d'obtenir une troncature "plus douce" des coefficients de Fourier.

Pour cela, on considère une fonction b dans $\mathcal{D}(]-1,1[^d)$, c'est à dire de classe C^∞ et nulle en dehors de $]-1,1[^d$, et qui vaut 1 au voisinage de 0. On peut par exemple supposer

$$b(x) = 1, \quad \text{pour } \|x\|_\infty \leq \frac{1}{2}. \tag{4.36}$$

On définit alors la somme partielle modifiée

$$S_K f(x) = \sum_{k\in\mathbb{Z}^d} b\Big(\frac{k}{K}\Big) c_k(f) e^{ik\cdot x}. \tag{4.37}$$

On remarque que $S_K f \in \mathbb{T}_K$ et que S_K est lui aussi un opérateur de convolution puisque l'on a

$$S_K f(x) = \int_T f(y)(2\pi)^{-d} \sum_{k\in\mathbb{Z}^d} b\Big(\frac{k}{K}\Big) e^{ik\cdot(x-y)} dy = f * E_K(x), \tag{4.38}$$

avec $E_K \in \mathbb{T}_K$ définie par

$$E_K(z) := (2\pi)^{-d} \sum_{k\in\mathbb{Z}^d} b\Big(\frac{k}{K}\Big) e^{ik\cdot z}. \tag{4.39}$$

Afin de dégager les propriétés de E_K, on utilise à présent un résultat important qui relie les séries de Fourier et les transformées de Fourier : la *formule de Poisson*.

Théorème 4.2 *Soit g une fonction définie sur $\mathbb{R}^d$ telle que*

$$|g(x)| \leq C(1+\|x\|)^{-a}, \tag{4.40}$$

avec $a > d$, et soit $\hat{g}$ sa transformée de Fourier

$$\hat{g}(\omega) = \int_{\mathbb{R}^d} g(x) e^{-i\omega\cdot x} dx. \tag{4.41}$$

On a alors, pour presque tout x,

$$\sum_{n\in\mathbb{Z}^d} g(x+2n\pi) = (2\pi)^{-d} \sum_{k\in\mathbb{Z}^d} \hat{g}(k) e^{ikx}. \tag{4.42}$$

Preuve : La propriété de décroissance (4.40) assure que $g \in L^1(\mathbb{R}^d)$ ce qui implique que $\hat{g}$ est continue. D'autre même elle assure que la série du membre de gauche dans (4.42) est convergente pour tout $x \in \mathbb{R}^d$. Sa limite est une fonction 2π périodique en chaque variable et qui appartient à $L^\infty(T)$ et donc à $L^2(T)$. Ses coefficients de Fourier sont donnés par

$$\begin{aligned}
c_k &= (2\pi)^{-d} \int_T \Big(\sum_{n\in\mathbb{Z}^d} g(x+2n\pi)\Big) e^{-ik\cdot x} dx \\
&= (2\pi)^{-d} \sum_{n\in\mathbb{Z}^d} \int_{T+2n\pi} g(x) e^{-ik\cdot x} dx \\
&= (2\pi)^{-d} \int_{\mathbb{R}^d} g(x) e^{-ik\cdot x} dx \\
&= (2\pi)^{-d} \hat{g}(k),
\end{aligned}$$

ce qui entraîne (4.42). $\diamond$

Pour appliquer la formule de Poisson dans notre cadre, on considére la fonction φ dont la transformée de Fourier vaut la fonction b, c'est à dire telle que

$$\hat{\varphi}(\omega) = \int_{\mathbb{R}^d} \varphi(x) e^{-i\omega\cdot x} dx = b(\omega), \tag{4.43}$$

ou de manière équivalente.

$$\varphi(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} b(\omega) e^{i\omega\cdot x} d\omega. \tag{4.44}$$

On peut écrire

$$E_K(z) := (2\pi)^{-d} \sum_{k\in\mathbb{Z}^d} \hat{\varphi}\Big(\frac{k}{K}\Big) e^{ik\cdot z}. \tag{4.45}$$

On remarque que $\omega \mapsto \hat{\varphi}(\frac{\omega}{K})$ est la transformée de Fourier de la fonction φ_K définie par

$$\varphi_K(x) = K^d \varphi(Kx), \tag{4.46}$$

et on a donc

$$E_K(z) := (2\pi)^{-d} \sum_{k\in\mathbb{Z}^d} \hat{\varphi}_K(k) e^{ik\cdot z}. \tag{4.47}$$

On note que, puisque b est C^∞ à support compact, φ appartient à la classe de Schwartz $\mathcal{S}(\mathbb{R}^d)$: elle est de classe C^∞ et vérifie

$$\sup_{x\in\mathbb{R}^d} \|x\|^a |D^\alpha \varphi(x)| < \infty, \tag{4.48}$$

pour tout $a > 0$ et $\alpha \in \mathbb{N}^d$. Il en est évidemment de même pour φ_K, ce qui nous permet d'appliquer la formule de Poisson qui nous donne

$$E_K(z) = \sum_{n\in\mathbb{Z}^d} \varphi_K(z + 2n\pi). \tag{4.49}$$

Une première conséquence est la stabilité uniforme en norme $L^p(T)$ des opérateurs S_K.

Théorème 4.3 *Pour $1 \le p \le \infty$, on a*

$$\|S_K\|_{\mathcal{L}(L^p(T))} \le \|\varphi\|_{L^1(\mathbb{R}^d)} < \infty, \tag{4.50}$$

pour tout $K > 0$.

Preuve : Comme pour l'opérateur P_K, la formule de Young implique

$$\|S_K\|_{\mathcal{L}(L^p(T))} \le \|E_K\|_{L^1(T)}. \tag{4.51}$$

On écrit alors

$$\begin{aligned} \|E_K\|_{L^1} &= \int_T |E_K(z)| dz \le \int_T \sum_{n\in\mathbb{Z}^d} |\varphi_K(z + 2n\pi)| \\ &= \int_{\mathbb{R}^d} |\varphi_K(z)| dz = \int_{\mathbb{R}^d} |\varphi(z)| dz = \|\varphi\|_{L^1(\mathbb{R}^d)}, \end{aligned}$$

ce qui donne l'estimation annoncée. $\diamond$.

On remarque ensuite que φ_K est une approximation de l'identité quand $K \to +\infty$: on a

$$\int_{\mathbb{R}^d} \varphi(x) dx = \hat{\varphi}(0) = b(0) = 1, \tag{4.52}$$

et pour tout $\delta, \eta > 0$, il existe $K(\delta, \eta)$ tel que

$$K \geq K(\delta, \eta) \Rightarrow \int_{\delta \leq \|x\|_\infty} |\varphi_K(x)| dx = \int_{K\delta \leq \|x\|_\infty} |\varphi(x)| dx \leq \eta. \tag{4.53}$$

La fonction E_K est donc une version périodisée de l'approximation de l'identité : on a

$$\int_T E_K(x) dx = \int_{\mathbb{R}^d} \varphi_K(x) dx = 1, \tag{4.54}$$

et on montre facilement que, pour tout $\delta, \eta > 0$, il existe $K(\delta, \eta)$ tel que

$$K \geq K(\delta, \eta) \Rightarrow \int_{\delta \leq \|x\|_\infty \leq \pi} |E_K(x)| dx \leq \eta. \tag{4.55}$$

En utilisant ces propriétés, on peut établir la convergence dans L^p de $S_K f$ vers f pour tout $f \in L^p(T)$, en remplacant dans le cas $p = \infty$ l'espace $L^\infty(T)$ par l'espace $C_{per}(T)$ des fonctions continues et 2π périodique en chaque variable.

Théorème 4.4 *Soit* $1 \leq p < \infty$. *Pour tout* $f \in L^p(T)$, *on a*

$$\lim_{K \to +\infty} \|f - S_K f\|_{L^p} = 0. \tag{4.56}$$

Pour tout $f \in C_{per}(T)$, *on a*

$$\lim_{K \to +\infty} \|f - S_K f\|_{L^\infty} = 0. \tag{4.57}$$

Preuve : On traite d'abord le cas $f \in C_{per}(T)$ par un raisonnement très classique d'approximation de l'identité que nous rappelons ici. Puisque $\int_T E_K(x) dx = 1$, on peut écrire, pour tout $x \in T$,

$$|f(x) - S_K f(x)| = |\int_T (f(x) - f(y)) E_K(x-y) dy| = |\int_T (f(x) - f(x-z)) E_K(z) dz|. \tag{4.58}$$

Soit $\varepsilon > 0$ arbitraire. On remarque que f est uniformément continue sur $\mathbb{R}^d$, et il existe par conséquent δ tel que

$$\|x - y\|_\infty \leq \delta \Rightarrow |f(x) - f(y)| \leq \frac{\varepsilon}{2\|\varphi\|_{L^1}} \tag{4.59}$$

Pour tout $x \in T$, on écrit alors

$$\begin{aligned} |f(x) - S_K f(x)| &\leq \Big| \int_{\|z\|_\infty \leq \delta} (f(x) - f(x-z)) E_K(z) dz \Big| + \Big| \int_{\delta \leq \|z\|_\infty \leq \pi} (f(x) - f(x-z)) E_K(z) dz \Big| \\ &\leq \frac{\varepsilon}{2\|\varphi\|_{L^1}} \int_{\|z\|_\infty \leq \delta} |E_K(z)| dz + 2\|f\|_{L^\infty} \int_{\delta \leq \|z\|_\infty \leq \pi} |E_K(z)| dz \\ &\leq \frac{\varepsilon}{2} + 2\|f\|_{L^\infty} \int_{\delta \leq \|z\|_\infty \leq \pi} |E_K(z)| dz. \end{aligned}$$

La propriété de concentration (4.55) montre que pour K suffisament grand le second terme est plus petit que $\frac{\varepsilon}{2}$. On a ainsi établi la convergence uniforme de $S_K f$ vers f.

Dans le cas $f \in L^p(T)$, on remarque que pour tout $\varepsilon > 0$ il existe une fonction $g \in C_{per}(T)$ telle que

$$\|f - g\|_{L^p} \leq \frac{\varepsilon}{2 + 2\|\varphi\|_{L^1}}, \tag{4.60}$$

On peut alors écrire

$$\begin{aligned} \|f - S_K f\|_{L^p} &\leq \|f - g\|_{L^p} + \|S_K f - S_K g\|_{L^p} + \|g - S_K g\|_{L^p} \\ &\leq (1 + \|\varphi\|_{L^1}) \|f - g\|_{L^p} + \|g - S_K g\|_{L^p} \\ &\leq \frac{\varepsilon}{2} + |T|^{1/p} \|g - S_K g\|_{L^\infty}, \end{aligned}$$

où l'on a utilisé (4.50). D'après la première partie de la preuve, le second membre est plus petit que $\frac{\varepsilon}{2}$ pour K suffisament grand. On a ainsi établi la convergence dans $L^p(T)$ de $S_K f$ vers f. $\diamond$

Nous pouvons à présent établir un résultat concernant la vitesse d'approximation par les polynômes trigonométriques en norme L^p, qui s'apparente au Theorème 4.1 dans le cas $p = 2$.

Théorème 4.5 *Soit m un entier positif et $1 \leq p \leq \infty$. Pour tout $f \in W^{m,p}_{per}(T)$, on a*

$$\|f - S_K f\|_{L^p} \leq C K^{-m} |f|_{W^{m,p}_{per}}, \quad K > 0, \tag{4.61}$$

où la constante C dépend de m, p, d et de la fonction b utilisée dans la définition de S_K.

Preuve : Nous donnons la preuve uniquement en dimension $d = 1$, sa généralisation à la dimension supérieure étant technique et laissée en exercice aux étudiants courageux. On sait d'après le Théorème 4.4 que $S_K f$ converge vers f dans L^p. Ceci permet d'écrire l'erreur sous la forme d'une série convergente

$$f - S_K f = \sum_{j \geq 0} (S_{2^{j+1}K} f - S_{2^j K} f). \tag{4.62}$$

Pour chaque terme, on peut écrire

$$S_{2^{j+1}K} f - S_{2^j K} f = \sum_{k \in \mathbb{Z}} \Big(\hat{\varphi}\Big(\frac{k}{2^{j+1}K}\Big) - \hat{\varphi}\Big(\frac{k}{2^j K}\Big)\Big) c_k(f) e^{ik\cdot x} = \sum_{k \in \mathbb{Z}} \hat{\psi}\Big(\frac{k}{2^j K}\Big) c_k(f) e^{ik\cdot x}, \tag{4.63}$$

où ψ est définie par

$$\hat{\psi}(\omega) = \hat{\varphi}(\omega/2) - \hat{\varphi}(\omega). \tag{4.64}$$

Le même calcul que celui effectué pour S_K montre ainsi que

$$S_{2^{j+1}K} f - S_{2^j K} f = f * F_{2^j K}, \tag{4.65}$$

avec

$$F_K(z) = \sum_{n \in \mathbb{Z}} \psi_K(z + 2n\pi), \tag{4.66}$$

où $\psi_K(x) := K\psi(Kx)$. Comme φ, la fonction ψ est dans la classe de Schwartz. Par construction, on a

$$\hat{\psi}(\omega) = 0, \quad |\omega| \leq \frac{1}{2}. \tag{4.67}$$

Ceci nous permet de définir une nouvelle fonction $\tilde{\psi}$ dans la classe de Schwartz telle que

$$\hat{\psi}(\omega) = (i\omega)^m \hat{(\tilde{\psi}(\omega))} \tag{4.68}$$

ce qui équivaut à

$$\tilde{\psi}^{(m)} = \psi. \tag{4.69}$$

Autrement dit, $\tilde{\psi}$ est une primitive d'ordre m de ψ. Avec la notation $\tilde{\psi}_K(x) := K\tilde{\psi}(Kx)$, on a donc

$$\tilde{\psi}_K^{(m)} = K^m \psi_K. \tag{4.70}$$

On peut ainsi écrire

$$F_K(z) = K^{-m} \sum_{n \in \mathbb{Z}^d} \tilde{\psi}_K^{(m)}(z + 2n\pi) = K^{-m} \tilde{F}_K^{(m)}(z), \quad \tilde{F}_K(z) := \sum_{n \in \mathbb{Z}} \tilde{\psi}_K(z + 2n\pi) \tag{4.71}$$

et par conséquent

$$f * F_K = K^{-m} f * \tilde{F}_K^{(m)} = K^{-m} f^{(m)} * \tilde{F}_K. \tag{4.72}$$

En remarquant que comme pour E_K on a $\|\tilde{F}_K\|_{L^1(]-\pi,\pi[)} \leq \|\tilde{\psi}\|_{L^1(\mathbb{R})}$, on obtient ainsi l'estimation

$$\|f * F_K\|_{L^p} \leq CK^{-m}\|f^{(m)}\|_{L^p}, \tag{4.73}$$

avec $C := \|\tilde{\psi}\|_{L^1(\mathbb{R})}$. En remplacant K par 2^jK, il vient

$$\|S_{2^{j+1}K}f - S_{2^jK}f\|_{L^p} = \|f * F_{2^jK}\|_{L^p} \leq C2^{-mj}K^{-m}\|f^{(m)}\|_{L^p}. \tag{4.74}$$

On obtient ainsi

$$\|f - S_Kf\|_{L^p} \leq \sum_{j\geq 0}\|S_{2^{j+1}K}f - S_{2^jK}f\|_{L^p} \leq C(\sum_{j\geq 0} 2^{-jm})K^{-m}\|f^{(m)}\|_{L^p}, \tag{4.75}$$

soit finalement quitte à diviser la constante C par $1 - 2^{-m}$,

$$\|f - S_Kf\|_{L^p} \leq CK^{-m}|f|_{W^{m,p}}, \tag{4.76}$$

qui est le résultat annoncé. $\diamond$

Remarque 4.2 *Comme dans le cas de l'approximation en norme L^2, il est possible (mais plus difficile) d'étendre ce théorème à une régularité $s > 0$ quelconque : pour tout $f \in W^{s,p}_{per}(T)$, on a*

$$\|f - S_Kf\|_{L^p} \leq CK^{-s}|f|_{W^{s,p}_{per}}, \quad K > 0. \tag{4.77}$$

En posant $X_n = \mathbb{T}_K$ avec

$$n = \dim(\mathbb{T}_K) = (2K+1)^d. \tag{4.78}$$

on déduit de l'estimation (4.77) que, pour $0 < s$,

$$\sigma_n(f)_{L^p} \leq C|f|_{W^{s,p}_{per}}n^{-s/d}. \tag{4.79}$$

Pour l'approximation en norme L^p par les polynômes trigonométriques on a $W^{s,p}_{per}(T) \subset \mathcal{A}^{s/d}$, avec la terminologie de (1.5). De même, on peut obtenir des estimations d'erreur dans des normes plus régulières que L^p : si $0 \leq t \leq s$ et $f \in W^{s,p}_{per}(T)$, on a l'estimation

$$\|f - S_Kf\|_{W^{t,p}_{per}} \leq CK^{-(s-t)}|f|_{W^{s,p}_{per}}, \quad K > 0, \tag{4.80}$$

ou encore

$$\sigma_n(f)_{W^{t,p}_{per}} \leq C|f|_{W^{s,p}_{per}}n^{-(s-t)/d}. \tag{4.81}$$

Pour l'approximation en norme $W^{t,p}_{per}$ par les polynômes trigonométriques on a donc $W^{s,p}_{per}(T) \subset \mathcal{A}^{(s-t)/d}$.

Revenons à présent sur l'approximation de f par ses sommes partielles de Fourier, c'est à dire par P_Kf. Nous avons déja remarqué que les projecteurs P_K ne sont pas uniformément borné en norme L^∞, puisque l'on a

$$c\log(K)^d \leq \|P_K\|_{\mathcal{L}(L^\infty)} = \|D_K\|_{L^1} \leq C\log(K)^d, \quad K \geq 2. \tag{4.82}$$

Par le raisonnement invoqué à la fin de la section §1, on peut donc écrire, quitte à modifier la constante C

$$\|f - P_Kf\|_{L^\infty} \leq C\log(K)^d \inf_{g\in\mathbb{T}_K}\|f - g\|_{L^\infty}, \quad K \geq 2. \tag{4.83}$$

D'après (4.77), ceci nous montre que pour tout $f \in W^{s,\infty}_{per}(T)$, on a

$$\|f - P_Kf\|_{L^\infty} \leq C\log(K)^dK^{-s}|f|_{W^{s,\infty}_{per}}, \quad K \geq 2. \tag{4.84}$$

On note que le facteur logarithmique dans l'estimation est asymptotiquement négligeable devant K^ε pour $\varepsilon > 0$ arbitrairement petit. Par conséquent, on peut aussi écrire

$$\|f - P_Kf\|_{L^\infty} \leq CK^{-(s-\varepsilon)}|f|_{W^{s,\infty}_{per}}, \quad K > 0, \tag{4.85}$$

pour tout ε, avec une constante C qui dépend de ε.

En ce qui concerne la norme L^p, il est possible de montrer que pour $1 < p < \infty$ les projecteurs P_K sont uniformément stable : il existe une constante C qui dépend de p et d, telle que

$$\|P_K\|_{\mathcal{L}(L^p)} \leq C, \quad K \geq 0. \tag{4.86}$$

La preuve de ce résultat utilise des méthode d'analyse beaucoup plus fines que l'inégalité de Young pour étudier la stabilité de la convolution par D_K : la théorie des intégales singulières de Calderon-Zygmund, qui sort du cadre de ce cours.

Un résultat encore plus fascinant concerne les sommes partielles obtenues si l'on remplace la troncature $\|k\|_\infty \leq K$ par $\|k\| \leq K$ dans la définition de P_K, lorsque $d > 1$. Charles Fefferman a montré que dans ce cas les projecteurs obtenus ne sont pas borné dans L^p lorsque $p \neq 2$, résultat parmi ceux pour lesquels il recu la médaille Fields en 1978. Ainsi la stabilité dans L^p des sommes partielles de Fourier est fortement influencée par la *forme* (en particulier la courbure) du domaine de troncature.

Nous terminons cette section en présentant un résultat important concernant les polynômes trigonométriques : *l'inégalité de Bernstein* ou *inégalité inverse*.

Théorème 4.6 *Soit m un entier strictement positif et $1 \leq p \leq \infty$. Il existe une constante C qui dépend de m et d telle que*

$$|g|_{W^{m,p}_{per}} \leq CK^m \|g\|_{L^p} \tag{4.87}$$

pour tout $g \in \mathbb{T}_K$.

Preuve : En utilisant (4.36) dans la définition de l'opérateur S_K, on note que, pour tout $g \in \mathbb{T}_K$, on a

$$g = S_{2K} g = g * E_{2K}. \tag{4.88}$$

Pour tout multi-indice de dérivation α tel que $|\alpha| = m$, on peut donc écrire

$$D^\alpha g = D^\alpha (g * E_{2K}) = g * \partial^\alpha E_{2K}. \tag{4.89}$$

D'après (4.49), en notant $\tilde{\varphi} = \partial^\alpha \varphi$, on peut écrire

$$D^\alpha E_{2K}(z) = (2K)^m \tilde{E}_{2K}(z), \tag{4.90}$$

avec la notation $\tilde{E}_K(z) := \sum_{n \in \mathbb{Z}^d} \tilde{\varphi}_K(z + 2n\pi)$ et $\tilde{\varphi}_K(x) = K^d \tilde{\varphi}(Kx)$. L'inégalité de Young donne alors

$$\|D^\alpha g\|_{L^p} \leq \|g\|_{L^p} \|D^\alpha E_{2K}\|_{L^1} = (2K)^m \|g\|_{L^p} \|\tilde{E}_K\|_{L^1} \leq 2^m \|\tilde{\varphi}\|_{L^1} K^m \|g\|_{L^p}. \tag{4.91}$$

En prenant la puissance p et en sommant sur tous les $|\alpha| = m$ (ou en prenant le max dans le cas $p = \infty$), on obtient l'inégalité annoncée. $\diamond$

L'inégalité de Bernstein montre que si g est un polynôme trigonométrique de degré K, on peut controler les tailles de ses dérivées par celle de g, mesurées en norme L^p, au prix d'une constante qui augmente avec K. Cette inégalité explique en particulier qu'on ne peut pas bien approcher une fonction discontinue par un polynôme trigonométrique car elle limite la taille de la dérivée. Une autre illustration de cette difficulté est le *phénomène de Gibbs* - oscillations qui apparaissent dans la reconstruction par la somme partielle de Fourier d'une fonction constante par morceaux - bien connu en traitement du signal.

Dans l'inégalité de Bernstein, les normes sont dans le sens opposé par rapport à l'inégalité d'approximation (4.61), ce qui justifie l'appélation d'inégalité inverse, l'inégalité d'approximation étant parfois appelée *inégalité directe*.

Remarque 4.3 *Il est possible (mais plus difficile) d'étendre l'inégalité inverse à des régularité non entières, c'est à dire*

$$|g|_{W^{s,p}_{per}} \leq CK^s \|g\|_{L^p}, \quad g \in \mathbb{T}_K, \tag{4.92}$$

ainsi qu'à un controle par des normes de dérivées d'ordre inférieur, c'est-à-dire

$$|g|_{W^{s,p}_{per}} \leq CK^{s-t} |g|_{W^{t,p}_{per}}, \quad g \in \mathbb{T}_K, \tag{4.93}$$

pour $0 \leq t < s$.

5 Approximation par éléments finis

Les approximations polynomiales par morceaux étudiées dans la section §3 permettent d'approcher en norme L^p des fonctions définies sur des domaines Ω de géométrie arbitraire, puisqu'il suffit de disposer de partitions de ces domaines en éléments de géométries simples. L'inconvénient principal de ces approximations est qu'elles ne sont pas continues aux interfaces entre les éléments et a-fortiori non-différentiables même au sens des espaces de Sobolev $W^{1,p}(\Omega)$. Cela rend leur usage difficile pour certaines applications telles que l'approximation numérique des solutions d'équations aux dérivées partielles.

A l'inverse, les approximations par les polynômes trigonométriques étudiées dans la section §4 sont de classes C^∞ sur le tore T et par conséquent peuvent être dérivées à tous les ordres. Leur inconvénient est qu'elles s'appliquent aux fonctions périodiques sur le tore, tous les résultats d'approximations étant formulés dans les espaces $W^{1,p}_{per}(T)$. Or la plupart des problèmes pratiques font intervenir des fonctions qui ne sont pas périodiques et qui sont définies sur domaines plus généraux que T.

L'approximation par les éléments finis consiste à conserver la souplesse géométrique des fonctions polynomiales par morceaux, tout en imposant à ces fonctions des conditions de raccords réguliers permettant de les dériver. Cette approche a été introduite par les ingénieurs pour le calcul des structures, puis analysée et systématisée par les mathématiciens.

De manière générale, un élément fini se définit localement par la donnée d'un triplet (T, X_T, Σ_T) où T est un domaine simple, X_T un espace de fonctions de dimension finie M défini sur T, et Σ_T un espace de M forme linéaires indépendantes $(\psi_i)_{i=1,\cdots,M}$ dont le domaine de définition contient X_T. On exige de plus la propriété *d'unisolvence*.

Définition 5.1 *Le triplet (T, X_T, Σ_T) est dit "unisolvent" si et seulement si l'application de X_T dans $\mathbb{R}^M$ qui à v associe le vecteur $(\psi_1(v), \cdots, \psi_M(v))$ est un isomorphisme.*

Remarque 5.1 *Afin de vérifier l'unisolvence, il suffit de vérifier l'une des deux propriétés équivalentes suivantes :*
(i) Si $v \in X_T$ est telle que $\psi_i(v) = 0$ pour tout $i = 1, \cdots, M$, alors $v = 0$.
(ii) Pour tout vecteur $(y_1, \cdots, y_M) \in \mathbb{R}^M$, il existe $v \in X_T$ tel que $\psi_i(v) = y_i$.

Les formes linéaires de Σ_T sont appelés *degrés de liberté* de l'élément fini. La propriété d'unisolvence nous permet de définir une base $(\varphi_i)_{i=1,\cdots,M}$ associée aux degrés de libertés par les relations

$$\psi_i(\varphi_j) = \delta_{i,j}, \quad i, j = 1, \cdots, M. \tag{5.1}$$

On peut aussi définir aussi un *opérateur d'interpolation* qui agit sur les fonctions définis sur T qui n'appartiennent pas forcément à X_T mais au domaine de définition Y_T des formes linéaires Σ_T : pour tout $v \in Y_T$ on définit

$$I_T v = \sum_{i=1}^{M} \psi_i(v)\varphi_i.$$

On note que $I_T v$ est l'unique élément de X_T tel que $\psi_i(I_T v) = \psi_i(v)$ pour $i = 1, \cdots, M$.

L'élément fini triangulaire de Lagrange est le plus utilisé en pratique. Dans ce cas, le domaine simple T est un triangle de sommets (a_0, a_1, a_2) en dimension 2, un tétraèdre de sommets (a_0, a_1, a_2, a_3) en dimension 3, et plus généralement un d-simplexe de sommets $(a_0, \cdots, a_d)$ en dimension d. Le domaine T est donc défini comme l'intérieur de l'enveloppe convexe de ces $d+1$ points. On suppose le simplexe *non-dégénéré*, i.e. les points a_i ne sont pas tous contenus dans un même hyperplan (dans le cas d'un triangle cela signifie que les trois points ne sont pas alignés, et dans le cas d'un tétraèdre que les quatres points ne sont pas co-planaires). Cette condition entraine que le volume $|T|$ de T est strictement positif. Si h_T et ρ_T désignent le diamètre et la rondeur de T, on définit *l'excentricité*

$$\sigma_T = \frac{h_T}{\rho_T}, \tag{5.2}$$

qui quantifie la non-dégénérescence de T.

Remarque 5.2 *Nous nous plaçons dans toute la suite dans le cas général de la dimension d quelconque, mais il peut être utile pour l'intuition de traduire la construction dans les cas $d = 2$ ou 3.*

On choisi ici l'espace $X_T := \mathbb{P}_k$ des polynôme de degré total inférieur à k, dont on a déja vu que la dimension est $M = \binom{k+d}{d}$.

Il est utile d'introduire les *coordonnées barycentriques* afin de décrire le simplexe T : pour tout point $x \in \mathbb{R}^d$, il existe un unique vecteur $(\lambda_i(x))_{i=0,\cdots,d}$ qui vérifie les équations

$$x = \sum_{i=0}^{d} a_i \lambda_i(x) \;\text{ et }\; \sum_i \lambda_i(x) = 1. \tag{5.3}$$

L'existence et l'unicité se vérifient en développant la première équation sur chaque coordonnées cartésienne, on obtient ainsi un système $d+1 \times d+1$ dont on vérifie facilement que la matrice est inversible si et seulement si le simplexe est non-dégénéré. On peut alors écrire,

$$T = \{x \in \mathbb{R}^d \; ; \; \lambda_i(x) > 0, \; i = 0, \cdots, d\}. \tag{5.4}$$

On remarque que le passage des coordonnées cartésienne aux coordonnées barycentriques est une transformation affine. De ce fait, un polynôme de degré total k en coordonnées cartésiennes s'exprime comme un polynôme de degré total k en coordonnées barycentriques et vice versa. On voit en particulier que pour tout polynôme π de degré 1, on a

$$\pi(x) = \sum_{i=0}^{d} \pi(a_i)\lambda_i(x). \tag{5.5}$$

Il reste à définir l'ensemble des formes linéaire Σ_T. Pour cela on définit, pour $k \geq 1$ le *treillis principal d'ordre* k de T comme l'ensemble

$$\Sigma_k := \Big\{x \in \mathbb{R}^d \; ; \; \lambda_i(x) \in \Big\{0, \frac{1}{k}, \frac{2}{k}, \cdots, 1\Big\}\Big\}. \tag{5.6}$$

Notons que Σ_1 correspond aux sommets du simplexe, et qu'on obtient Σ_2 en ajoutant les milieux des arêtes. On prend pour degrés de libertés les valeurs aux points de Σ_k, c'est à dire

$$\Sigma_T := \{v \mapsto v(\gamma), \; \gamma \in \Sigma_k\}. \tag{5.7}$$

Théorème 5.1 *Le triplet (T, X_T, Σ_T) définissant l'élément fini triangulaire de Lagrange est unisolvent.*

Preuve : On remarque tout d'abord que le cardinal de Σ_T est bien égal à la dimension de X_T. En effet, on peut aussi écrire les éléments de Σ_k sous la forme

$$\Sigma_k = \Big\{ \sum_{i=1}^{d} \frac{\alpha_i}{k} a_i + (1 - \sum_{i=1}^{d} \frac{\alpha_i}{k}) a_0, \;\; 0 \leq \alpha_1 + \cdots + \alpha_d \leq k \Big\}, \tag{5.8}$$

où les α_i sont des entiers positifs. Il suffit donc de prouver que si $\pi \in \mathbb{P}_k$ est tel que $\pi(x) = 0$ pour tout $x \in \Sigma_k$ alors $\pi = 0$. On raisonne pour cela par récurrence sur la dimension d. Pour $d = 1$, la propriété est immédiate puisqu'un polynôme de degré k qui s'annule en $k+1$ points distincts est identiquement nul. On suppose la propriété prouvée en dimension $d-1$ et on cherche à la prouver en dimension d.

On raisonne alors par récurrence sur le degré k. Lorsque $k = 1$, le fait qu'une fonction affine qui s'anulle au sommet d'un simplexe non-dégénéré est identiquement nulle est une conséquence de (5.5). Supposons la propriété prouvée pour les polynômes de degrés $k-1$, et considérons un polynôme π de degré k qui s'anulle sur Σ_k. On remarque que Σ_k contient en particulier l'ensemble

$$\Sigma_k' = \{x \in \Sigma_k \; ; \; \lambda_0(x) = 0\}, \tag{5.9}$$

qui n'est rien d'autre que le treillis principal d'ordre k du $d-1$-simplexe de sommets $(a_1, \cdots, a_d)$. Comme la restriction de π à l'hyperplan engendré par $(a_1, \cdots, a_d)$ est un polynôme de degré k de $d-1$ variables, on peut utiliser l'hypothèse de récurrence sur d qui nous montre de π est identiquement nul sur cet hyperplan. En prenant un système de coordonnées $(x_1, \cdots, x_d)$ tel que l'hyperplan ait pour équation $x_d = 0$, ceci implique que

$$\pi(x_1, \cdots, x_d) = x_d \tilde{\pi}(x_1, \cdots, x_{d-1}), \tag{5.10}$$

où $\tilde{\pi}$ est de degré total $k-1$. Considérons à présent le reste des points de Σ_k c'est à dire

$$\Sigma_k'' = \Sigma_k - \Sigma_k'. \tag{5.11}$$

Le polynôme $\tilde{\pi}$ s'anulle nécessairement sur cet ensemble puisque x_d ne s'y anulle pas. Or on remarque que Σ_k'' est un treillis principal d'ordre $k-1$. L'hypothèse de récurrence sur k entraine que $\tilde{\pi} = 0$. On a ainsi prouvé que $\pi = 0$. $\diamond$

L'unisolvence permet d'introduire les fonctions de bases φ_i et l'interpolant I_T. On peut exprimer ces fonctions à l'aide des coordonnées barycentriques :

- Dans le cas $k=1$, on trouve simplement les fonctions λ_i associées à chaque point a_i qui coincident avec les degrés de liberté.
- Dans le cas $k=2$, on trouve les fonctions quadratiques $2\lambda_i(\lambda_i - 1/2)$ pour les degrés de libertés associés aux sommets a_i, et les fonctions $4\lambda_i\lambda_j$ pour les degrés de libertés associés aux milieux $(a_i+a_j)/2$.
- Dans le cas $k=3$, on trouve des fonctions cubiques du même type pour les degrés de libertés sur les sommets et sur les arêtes, et on a des fonctions supplémentaires associée à des degrés de libertés qui ne sont pas sur les arêtes. Dans le cas $d=2$, on trouve la *fonction bulle* $27\lambda_0\lambda_1\lambda_2$ qui vaut 1 au barycentre du triangle et 0 sur le bord.

On introduit un *simplexe de référence* $\widehat{T}$ dont les sommets sont donnés par l'origine $\hat{a}_0 = (0,\cdots,0)$ et les points $\hat{a}_i = (0,\cdots,1,\cdots,0)$ dont toutes les coordonnées sont nulles sauf la i-ème qui vaut 1. On note A_T l'unique transformation affine qui envoie $\hat{a}_i$ sur a_i pour tout $i = 0,\cdots,d$. On peut donc écrire

$$A_T(x) = a_T + B_T x, \quad a_T = a_0, \tag{5.12}$$

où B_T est une matrice $d \times d$ dont la i-ème colonne est donnée par les coordonnées de $a_i - a_0$. Comme le simplexe T est non-dégénéré, A_T est une bijection qui envoie $\widehat{T}$ sur T, la matrice B_T est inversible et on a

$$|T| = |\det(B_T)|\,|\widehat{T}| = \frac{|\det(B_T)|}{d!}. \tag{5.13}$$

Nous allons utiliser la transformation A_T pour transporter les objets d'intérêt depuis T vers $\widehat{T}$. On reprend les notations de la section §3 en notant systématiquement $\hat{q}$ la quantité obtenue par transport de q :

$$\hat{x} = A_T^{-1}(x) = B_T^{-1}(x - a_0) \Leftrightarrow x = A_T(\hat{x}), \tag{5.14}$$

pour les points, et

$$\hat{v}(\hat{x}) = v(x) \Leftrightarrow \hat{v} = v \circ A_T. \tag{5.15}$$

pour les fonctions. De même si ψ est une forme linéaire qui agit sur les fonctions définies sur T, on définit la forme linéaire transportée $\hat{\psi}$ qui agit sur les fonctions définies sur $\widehat{T}$ suivant

$$\hat{\psi}(\hat{v}) = \psi(v). \tag{5.16}$$

On remarque que les coordonnées barycentriques sont laissées invariantes par transformation affine :

$$\lambda_i(\hat{x}) = \lambda_i(x) = \hat{\lambda}_i(\hat{x}). \tag{5.17}$$

Ceci entraine en particulier que le treillis principal d'ordre k du simplexe T est l'image par A_T du treillis principal d'ordre k de $\widehat{T}$ que l'on note $\widehat{\Sigma}_k$. Ainsi l'ensemble des forme linéaire $\Sigma_{\widehat{T}}$ pour l'élément de référence $\widehat{T}$ coincide avec le transport $\widetilde{\Sigma}_T$ de l'ensemble Σ_T des formes linéaires pour l'élément T. On le note parfois $\widehat{\Sigma}$.

L'espace $\mathbb{P}_k$ étant invariant par la transformation A_T, ceci entraine que les interpolants $I_{\widehat{T}}$ et I_T vérifient l'identité

$$I_{\widehat{T}}\hat{v}(\hat{x}) = I_T v(x), \tag{5.18}$$

qui peut être vue comme une relation de *commutation* entre le transport d'une fonction par A_T et son interpolation :

$$I_T v \circ A_T = I_{\widehat{T}}(v \circ A_T), \tag{5.19}$$

similaire à l'identité (3.57) que nous avons observée pour le projecteur orthogonal P_T. On voit aussi que les fonctions de base φ et $\tilde{\varphi}$ associées aux degrés de liberté γ et $\hat{\gamma}$ sur T et $\widehat{T}$ vérifient $\tilde{\varphi} = \hat{\varphi}$.

Remarque 5.3 *Nous avons montré que les objets d'intérêt de l'élément fini se transportent entre T et $\hat{T}$ au moyen de la transformation A_T. On peut inversement construire des éléments finis en partant d'un triplet unisolvent $(\widehat{T}, \widehat{X}, \widehat{\Sigma})$ et en définissant directement (T, X_T, Σ_T) comme le transport de ces objets par A_T. On vérifie aisément que le nouveau triplet obtenu est alors unisolvent.*

Afin d'établir des résultat d'approximation entre f et son interpolation $I_T f$, nous aurons besoin du *lemme de Bramble-Hilbert* qui est une conséquence du théorème de Deny-Lions.

Lemme 5.1 *Soit Ω un domaine borné lipschitzien connexe de $\mathbb{R}^d$, m un entier positif, et R un opérateur linéaire borné de $W^{m,p}(\Omega)$ dans $\mathbb{P}_k$ avec $m \leq k+1$, tel que $R\pi = \pi$ pour tout $\pi \in \mathbb{P}_k$. Alors pour tout $v \in W^{m,p}(\Omega)$ on a*

$$\|v - Rv\|_{W^{m,p}} \leq C|v|_{W^{m,p}}. \tag{5.20}$$

où la constante C ne dépend que de Ω, m, p et de la norme de R.

Preuve : Pour tout $\pi \in \mathbb{P}_k$, on écrit

$$\begin{aligned} \|v - Rv\|_{W^{m,p}} &\leq \|v - \pi\|_{W^{m,p}} + \|Rv - R\pi\|_{W^{m,p}} \\ &\leq (1 + \|R\|)\|v - \pi\|_{W^{m,p}}, \end{aligned}$$

avec $\|R\| := \sup_{\|v\|_{W^{m,p}}=1} \|Rv\|_{W^{m,p}}$. Comme π est arbitraire dans $\mathbb{P}_k$, on obtient l'estimation voulue en appliquant le théorème de Deny-Lions. $\diamond$

Pour appliquer le théorème de Bramble-Hilbert à l'interpolant I_T associé à l'élément fini de Lagrange (T, X_T, Σ_T) avec $X_T = \mathbb{P}_k$, il faut que l'espace $W^{m,p}(T)$ s'injecte continuement dans l'espace des fonctions uniformément continues sur T et pouvant ainsi être prolongées par continuité sur $\overline{T}$, ce qui est vrai uniquement lorsque $m > d/p$ si $p > 1$, et $m \geq d$ si $p = 1$. On a dans ce cas pour $n \leq m \leq k+1$,

$$\begin{aligned} |v - I_T v|_{W^{n,p}(T)} &\leq C\frac{|T|^{1/p}}{\rho_T^n}|\hat{v} - I_{\widehat{T}}\hat{v}|_{W^{n,p}(\widehat{T})} \\ &\leq C\frac{|T|^{1/p}}{\rho_T^n}|\hat{v}|_{W^{m,p}(\widehat{T})} \\ &\leq C\frac{|T|^{1/p}}{\rho_T^n}\frac{h_T^m}{|T|^{1/p}}|v|_{W^{m,p}(T)}, \end{aligned}$$

où la constante C varie d'une ligne à l'autre mais ne dépend que de d, m, n, k et p, soit l'estimation

$$|v - I_T v|_{W^{n,p}(T)} \leq C\frac{h_T^m}{\rho_T^n}|v|_{W^{m,p}(T)}, \tag{5.21}$$

qui est du même type que (3.46) obtenue pour l'erreur de meilleure approximation, mais avec la contrainte $m > d/p$ liée à l'utilisation de l'interpolant.

Remarque 5.4 *Comme pour la meilleure approximation (voir Remarque 3.3), on peut généraliser ces résultats en faisant appel aux injections de Sobolev : si $n < m \leq k+1$ et $m - n > d/p - d/q$, on obtient ainsi*

$$|v - I_T v|_{W^{n,q}(T)} \leq C|T|^{1/q-1/p}\frac{h_T^m}{\rho_T^n}|v|_{W^{m,p}(T)}, \tag{5.22}$$

sous l'hypothèse supplémentaire $m > d/p$ si $p > 1$, et $m \geq d$ si $p = 1$.

Afin de construire les *espaces d'éléments finis*, nous supposerons pour simplifier que le domaine Ω dans lequel on travaille peut-être partitionné en un nombre fini de simplexes. C'est le cas pour un domaine polygonal en dimension $d = 2$ et pour un polyhèdre en dimension $d = 3$. Soit $\mathcal{T}_h$ une telle partition de finesse $h = \max_{T\in\mathcal{T}_h} h_T$. On fera en plus l'hypothèse essentielle suivante :

Toute face d'un élément de T est soit contenue dans la frontière $\partial\Omega$, soit égale à une face d'un autre élément T'.

Une telle partition est dite *conforme*. Dans le cas d'une triangulation lorsque $d = 2$ cela signifie que l'on interdit les jonctions en forme $\perp$ dans le maillage de l'intérieur du domaine.

Etant donné un entier $k > 0$ fixé, *l'espace des éléments finis de degré k de Lagrange* (où éléments $\mathbb{P}_k$) associé à la partition $\mathcal{T}_h$ est défini en considérant l'ensemble des noeuds constitué par l'union de tous les treillis principaux d'ordre k pour chaque triangle

$$\Gamma_h = \cup_{T \in \mathcal{T}_h} \Sigma_k(T), \tag{5.23}$$

où de façon équivalente, l'ensemble des formes linéaires

$$\Sigma_h = \cup_{T \in \mathcal{T}_h} \Sigma_T, \tag{5.24}$$

où l'on ne compte qu'une seule fois les degrés de liberté communs à plusieurs éléments adjacents (par exemple les sommets des simplexes). Une fonction $v \in X_h$ est alors définie comme une fonction polynomiale de degré k sur chaque T et caractérisée par une suite de valeurs $v(\gamma)$ pour $\gamma \in \Gamma_h$, i.e. $\psi(v)$ pour $\psi \in \Sigma_h$.

Théorème 5.2 *Une définition équivalente de X_h est donnée par*

$$X_h := \{v \in C(\Omega) \ ; \ v_{|T} \in \mathbb{P}_k, \ T \in \mathcal{T}_h\}. \tag{5.25}$$

Preuve : Si v est continue et polynomiale sur chaque T, elle admet une valeur bien définie en chaque point de Γ_h et est donc uniquement caractérisée par sa suite de valeurs sur $v(\gamma)$ pour $\gamma \in \Gamma_h$. Réciproquement supposons v polynomiale de degré k sur chaque T et avec une valeur uniquement définie en chaque point de Γ_h. Si on considère deux simplexes adjacents T et T', alors sur l'interface $\partial T \cap \partial T'$ les polynômes $v_{|T}$ et $v_{|T'}$ prennent les mêmes valeurs aux points de $\Sigma_k(T) \cap \Sigma_k(T')$. Comme cet ensemble est aussi le treillis principal d'ordre k du $d-1$-simplexe $\partial T \cap \partial T'$, et que $v_{|T} - v_{|T'}$ est un polynôme de degré $k-1$ qui s'y annule, on en déduit la continuité de v à l'interface entre T et T'. $\diamond$

Ce résultat nous montre que

$$X_h \subset W^{1,p}(\Omega),$$

pour tout $p \geq 1$, et en particulier $X_h \subset H^1(\Omega)$. Par définition l'application qui à v associe ses valeurs $(v(\gamma))_{\gamma \in \Gamma_h}$ est un isomorphisme de X_h dans $\mathbb{R}^{\#(\Gamma_h)}$ et on a ainsi $\dim(X_h) = \#(\Gamma_h)$. On peut définir une base nodale pour l'espace X_h : pour chaque degré de liberté $\gamma \in \Gamma_h$ la fonction de base associée est définie par

$$\varphi_\gamma(\mu) = \delta_{\gamma,\mu}, \quad \mu \in \Gamma_h.$$

La fonction de base φ_γ a donc pour support l'union des simplexes qui contiennent γ, et sur chacun de ces simplexes elle est égale à la fonction de base pour l'élément fini (T, X_T, Σ_T) associée au degré de liberté γ. Dans le cas des éléments $\mathbb{P}_1$ sur une triangulation en dimension $d = 2$, on trouve les "fonctions chapeau" associées à chaque sommet du maillage. On remarque que la coordonée de $v \in X_h$ pour la fonction de base φ_γ est donnée par $v(\gamma)$.

On définit finalement l'interpolant I_h comme l'opérateur qui envoie $C(\Omega)$ dans X_h suivant

$$I_h v = \sum_{\gamma \in \Gamma_h} v(\gamma) \varphi_\gamma, \tag{5.26}$$

c'est à dire l'unique élément de X_h tel que $I_h v(\gamma) = v(\gamma)$. Notons que la restriction de $I_h v$ sur T est égale à l'interpolant local $I_T v$:

$$I_h v(v) = I_T v(x), \ \ x \in T. \tag{5.27}$$

Afin d'établir des résultats d'approximation, on va considérer une famille de partitions $(\mathcal{T}_h)_{h>0}$ et s'intéresser à l'erreur d'approximation lorsque h tend vers 0. On peut écrire pour $n = 0$ ou 1,

$$|v - I_h v|_{W^{n,p}} = \Big(\sum_{T \in \mathcal{T}_h} |v - I_T v|^p_{W^{n,p}(T)} \Big)^{1/p}. \tag{5.28}$$

Avant d'aller plus loin, on introduit une propriété importante.

Définition 5.2 *La famille* $(\mathcal{T}_h)_{h>0}$ *est dite "régulière" si et seulement si il existe une constante* C_r *telle que*

$$\sigma_T \leq C_r, \tag{5.29}$$

pour tout $T \in \mathcal{T}_h$ *et pour tout* $h > 0$.

Une propriété équivalente est l'existence d'une constante $c > 0$ telle que

$$ch_T^d \leq |T|, \tag{5.30}$$

où $|T|$ est le volume de T (exercice). Dans le cas d'une triangulation en dimension $d = 2$, cette propriété équivaut aussi au fait qu'il existe un $\theta_0 > 0$ tel que tous les angles des triangles vérifient $\theta \geq \theta_0$ (exercice).

Notons que toute famille quasi-uniforme est régulière. En dimension $d = 2$, il est facile de construire une famille quasi-uniforme de triangulations en partant d'une triangulation grossière $\mathcal{T}_0 = \mathcal{T}_{h_0}$ de Ω et en divisant chaque triangle en 4 sous-triangles semblables à partir des milieux des arêtes. En itérant ce découpage, on obtient des triangulations $\mathcal{T}_j = \mathcal{T}_{h_j}$ avec $h_j = 2^{-j}h_0$, pour tout $j \geq 0$.

Si $(\mathcal{T}_h)_{h>0}$ est une famille régulière, on peut reformuler l'estimation locale (5.21), sous la forme

$$|v - I_T v|_{W^{n,p}(T)} \leq Ch^{m-n}|v|_{W^{m,p}(T)}, \tag{5.31}$$

pour $n = 0$ ou 1, où C dépend de C_r, k, m, d et p. En combinant cette derniere estimation avec (5.28), on obtient un résultat d'approximation global sur X_h.

Théorème 5.3 *Sous les conditions* $n \leq m \leq k+1$, *et* $m > d/p$ *si* $p > 1$ *ou* $m \geq d$ *si* $p = 1$, *on a pour tout* $v \in W^{m,p}(\Omega)$, *et pour* $n = 0$ *ou* 1,

$$|v - I_h v|_{W^{n,p}} \leq Ch^{m-n}|v|_{W^{m,p}}, \tag{5.32}$$

où C *dépend de* C_r, k, m, d *et* p.

Remarque 5.5 *Il est possible (mais plus difficile) d'étendre ce résultat à des régularités non-entières : sous les conditions* $n \leq s \leq k+1$, *et* $s > d/p$ *si* $p > 1$ *ou* $s \geq d$ *si* $p = 1$, *on a pour tout* $v \in W^{s,p}(\Omega)$, *et pour* $n = 0$ *ou* 1,

$$|v - I_h v|_{W^{n,p}} \leq Ch^{s-n}|v|_{W^{s,p}}, \tag{5.33}$$

où C *dépend de* C_r, k, s, d *et* p.

Dans le cas de familles de partitions affine-invariantes et quasi-uniformes, en posant $X_n = X_h$ avec

$$n = \dim(X_h) \leq \dim(\mathbb{P}_k)\#(\mathcal{T}_h) \leq \dim(\mathbb{P}_k)Bh^{-d}, \tag{5.34}$$

l'estimation ci-dessus entraîne

$$\sigma_n(f)_{L^p} \leq C|f|_{W^{s,p}}n^{-s/d}, \tag{5.35}$$

et

$$\sigma_n(f)_{W^{1,p}} \leq C|f|_{W^{s,p}}n^{-(s-1)/d}. \tag{5.36}$$

On voit ainsi que pour l'approximation en norme L^p (resp. $W^{1,p}$) par les éléments finis de Lagrange de degré $k \geq s-1$ sur des partitions affine-invariantes et quasi-uniformes, on a $W^{s,p}(\Omega) \subset \mathcal{A}^{s/d}$ (resp. $\mathcal{A}^{(s-1)/d}$) avec la terminologie de (1.5).

Comme pour l'approximation par les polynômes trigonométriques, il est possible détablir une *inégalité inverse* qui compare les normes L^p et $W^{1,p}$ d'une fonction de l'espace X_h. On part de la remarque que toutes les normes sur l'espace $\mathbb{P}_k$ sont équivalentes puisque cet espace est de dimension finie. On obtient ainsi sur l'élément de référence $\hat{T}$ que pour tout $\hat{\pi} \in \mathbb{P}_k$ et pour tout entier $m \geq 0$ et $1 \leq p \leq \infty$,

$$|\hat{\pi}|_{W^{m,p}(\widehat{T})} \leq C\|\hat{\pi}\|_{L^p(\widehat{T})},$$

où la constante C ne dépend que de m et p. En combinant ceci avec les formules de changement de variable on obtient ainsi sur tout élément $T \in \mathcal{T}_h$

$$|\pi|_{W^{m,p}(T)} \leq \frac{C}{\rho_T^m}\|\pi\|_{L^p(T)}$$

Si la famille $(\mathcal{T}_h)_{h>0}$ est régulière et quasi-uniforme, on a donc

$$|\pi|_{W^{m,p}(T)} \leq Ch^{-m}\|\pi\|_{L^p(T)}.$$

Pour obtenir une estimation globale pour les fonctions de X_h sur Ω, il faut supposer $m = 1$ (ou $m = 0$ mais cela n'a aucun intérêt) et on obtient ainsi le résultat suivant.

Théorème 5.4 *Si $(\mathcal{T}_h)_{h>0}$ est une famille quasi-uniforme, on a pour tout $h > 0$ et $v_h \in X_h$,*

$$|v_h|_{W^{1,p}} \leq Ch^{-1}\|v_h\|_{L^p},$$

où C ne dépend que de p et de la constante c dans l'hypothèse de quasi-uniformité. On a en particulier $|v_h|_{H^1} \leq Ch^{-1}\|v_h\|_{L^2}$ dans le cas $p = 2$.

Nous terminons cette section en donnant trois exemples d'éléments finis d'usage moins courant que les éléments triangulaires de Lagrange, et nous indiquons brièvement leurs propriétés.

Exemple 1 : l'élément quadrilatéral
Dans cette construction on va partir de l'élément de référence pour aller vers un élément quelconque. On se place à nouveau en dimension $d = 2$, et on prend comme élément de référence $\widehat{T}$ le carré de sommet $\hat{a}_0 = (0,0)$, $\hat{a}_1 = (1,0)$, $\hat{a}_2 = (1,1)$ et $\hat{a}_3 = (0,1)$. L'espace $X_{\widehat{T}}$ est celui des polynôme de degré global inférieur à k que l'on note

$$\mathbb{Q}_k = \text{Vect}\{x_1^{k_1}x_2^{k_2} \; ; \; 0 \leq k_1, k_2 \leq k\}$$

et qui est de dimension $(k+1)^2$. On prend

$$\Sigma_{\widehat{T}} = \{v \mapsto v(\gamma) \; ; \; \gamma \in \hat{\Sigma}_k\}$$

où le treillis principal $\hat{\Sigma}_k$ est ici donné par

$$\Sigma_k = \{(\hat{x}_1, \hat{x}_2) \; ; \; \hat{x}_i \in \{0, \frac{1}{k}, \frac{2}{k}, \cdots, 1\}\}.$$

Il est très facile de prouver l'unisolvence. Afin de transporter cet élément sur un quadrilatère T de sommets (a_0, a_1, a_2, a_3), on suppose que celui-ci est convexe et n'est pas réduit à un triangle. On suppose aussi que les sommets sont décrit dans le sens direct (inverse des aiguilles d'une montre) comme pour $\widehat{T}$. Pour aller de $\widehat{T}$ vers T on utilise la transformation *bilinéaire*

$$x = A_T(\hat{x}) := a_0 + \hat{x}_1(a_1 - a_0) + \hat{x}_2(a_3 - a_0) + \hat{x}_1\hat{x}_2(a_2 - a_1 - a_3 + a_0).$$

Il faut faire attention au fait que $\mathbb{Q}_k$ n'est pas laissé invariant par transport de $\hat{T}$ vers T et on a donc en général $X_T \neq \mathbb{Q}_k$. On désigne par Σ_T l'image par A_T de $\Sigma_{\widehat{T}}$ et on obtient ainsi un triplet unisolvent. On peut d'autre part vérifier que pour une partition $\mathcal{T}_h$ *conforme* en quadrilatères, les $k+1$ degrés de libertés sur une arête commune à deux éléments adacents sont situés aux mêmes points ce qui entraine la continuité globale des fonctions de l'espace d'élément fini X_h obtenu.

Cette construction se généralise de façon naturelle en dimension $d > 2$. Les éléments finis quadrilatéraux possèdent des propriétés d'approximation similaires à celles des éléments triangulaires de Lagrange, mais celles-ci sont plus difficiles à établir, en particulier du fait que le jacobien de A_T n'est pas constant et que l'inverse de A_T n'est pas du même type que A_T.

Exemple 2 : l'élément de Hermite
On peut chercher à construire des espaces d'éléments finis possédant plus de régularité que la continuité globale et l'appartenance à $H^1(\Omega)$. En dimension 1, l'élément de Hermite se définit sur chaque intervalle $T = [a,b]$ en prenant $X_T = \mathbb{P}_{2k+1}$ et des degrés de libertés donnés par les formes linéaires $v \mapsto v^{(n)}(a)$ et $v \mapsto v^{(n)}(b)$ pour $n = 0, \cdots, k$. On assure ainsi la régularité C^k et H^{k+1} des fonctions de X_h.

Cette construction ne se généralise pas de manière naturelle en dimension $d > 1$ sur des partitions quelconques. En dimension $d = 2$, on peut utiliser une partition $\mathcal{T}_h$ conforme en rectangles T, ce qui impose que les côtés de tous les rectangles sont situés sur la même paire d'axe orthogonaux. En appelant

x_1 et x_2 les coordonnées sur ces axes, on choisi $X_T = \mathbb{Q}_{2k+1}$ où $\mathbb{Q}_k$ a été défini dans l'exemple précèdent. Si T a pour sommets (a_0, a_1, a_2, a_3), les degrés de libertés sont donnés par

$$\Sigma_T := \{v \mapsto D^\alpha v(a_i) \ ; \ i = 0, \cdots, 3, \ 0 \leq \alpha_1, \alpha_2 \leq k\}.$$

On vérifie aisément l'unisolvence de ce triplet, ainsi que le fait que les fonctions de l'espace X_h obtenu possèdent la régularité C^k et H^{k+1}. Cette construction se généralise naturellement à la dimension $d > 2$, mais elle présente l'inconvénient que l'on ne peut pas mailler ainsi un polygone ou polyèdre quelconque.

Exemple 3 : l'élément triangulaire d'Argyris
Il est possible de revenir à des triangulations conforme et d'obtenir la régularité C^1 et H^2 par la construction suivante : sur un triangle T de sommet (a_0, a_1, a_2), on prend $X_T = \mathbb{P}_5$, et pour les formes linéaires

$$\Sigma_T := \{v \mapsto D^\alpha v(a_i) \ ; \ i = 0, 1, 2, \ 0 \leq |\alpha| \leq 2\} \cup \{v \mapsto \frac{\partial v}{\partial n}(b_i) \ ; \ i = 0, 1, 2\},$$

où b_i est le milieu de l'arête e_i opposée à a_i.

Afin de démontrer l'unisolvence, on remarque d'abord que le cardinal de Σ_T est bien la dimension de $\mathbb{P}_5$. Puis on suppose que $v \in \mathbb{P}_5$ annule toutes les formes linéaires, et on montre d'abord que la restriction de v à toutes les arêtes e_i est nulle, puis qu'il en est de même pour $\frac{\partial v}{\partial n}$. On doit donc pouvoir factoriser $(\lambda_0(x)\lambda_1(x)\lambda_2(x))^2$ dans v ce qui entraine $v = 0$.

Cette preuve montre aussi que les fonctions de l'espace X_h associé sont C^1 aux interfaces entre les triangles et ont par conséquent la régularité $W^{2,p}$ pour tout $1 \leq p \leq \infty$. Le principal défaut de cet élément est sa complexité : 21 degrés de liberté par triangle.

6 Formulation variationnelle des EDP

Un problème important est l'approximation de fonctions qui sont sont pas données explicitement mais solutions de problèmes telles que des EDP. On s'intéresse ici à des EDP admettant une *formulation variationnelle* :

$$\textit{Trouver } u \in X \textit{ tel que } a(u, v) = L(v) \textit{ pour tout } v \in X,$$

où X est un espace de Hilbert, a une forme bilinéaire sur $X \times X$, et L une forme linéaire sur X. Ces formulations sont importantes pour les raisons suivantes :

1. De nombreux problèmes issus de la physique et de la mécanique admettent de telles formulations, et celles-ci reflètent souvent une propriété fondamentale du modèle, typiquement la minimisation d'une énergie sous-jacente.
2. Ces formulations donnent accès à des résultats fondamentaux sur le caractère bien posé de l'équation, c'est à dire l'existence et l'unicité de la solution, et la stabilité de cette solution par rapport à des perturbations des données.
3. Elles sont à la base de méthodes performantes pour l'approximation numérique des solutions, par la résolution d'un problème approché :

$$\textit{Trouver } u_n \in X_n \textit{ tel que } a(u_n, v_n) = L(v_n) \textit{ pour tout } v_n \in X_n,$$

   où X_n est un sous-espace de dimension finie de X.

Nous allons illustrer le passage à une formulation variationnelle sur l'exemple simple mais important du problème du laplacien (ou équation de Poisson) : on cherche une fonction u telle que

$$-\Delta u = f \ \text{ dans } \ \Omega \ \text{ et } \ u_{|\partial\Omega} = 0, \tag{6.1}$$

où Ω désigne un ouvert borné lipschitzien de $\mathbb{R}^d$, $\partial\Omega$ désigne la frontière de Ω et f est une fonction définie sur Ω.

Afin de donner un sens à l'équation (6.1) il faut préciser *l'espace* dans lequel on cherche la solution, ce qui entraîne une restriction sur le second membre f. Un premier choix intuitivement possible est de chercher u dans $C^2(\Omega)$ en supposant alors $f \in C(\Omega)$. En multipliant l'équation par une fonction v arbitraire de classe C^1, et en intégrant sur le domaine, on obtient

$$-\int_\Omega \Delta u \, v = \int_\Omega fv, \tag{6.2}$$

et en appliquant la formule de Green

$$\int_\Omega \nabla u \cdot \nabla v - \int_{\partial\Omega} \frac{\partial u}{\partial n} v = \int_\Omega fv, \tag{6.3}$$

où $\frac{\partial u}{\partial n} = \nabla u \cdot \mathbf{n}$ est la dérivée normale de u, avec $\mathbf{n}$ le vecteur unitaire normal extérieur au bord $\partial\Omega$. Dans le cas d'un domaine lipschitzien, cette normale est définie en presque tout point de $\partial\Omega$ et la formule de Green s'applique. En supposant de plus que v s'annulle au bord on obtient ainsi

$$\int_\Omega \nabla u \cdot \nabla v = \int_\Omega fv. \tag{6.4}$$

L'équation ci-dessus garde un sens lorsque u est seulement de classe C^1. En introduisant l'espace $X = \{v \in C^1 \ ; \ v_{|\partial\Omega} = 0\}$, on obtient ainsi que toute solution de (6.1) est aussi solution de la formulation variationnelle

$$\text{Trouver } u \in X \text{ tel que } a(u,v) = L(v) \text{ pour tout } v \in X, \tag{6.5}$$

avec

$$a(u,v) := \int_\Omega \nabla u \cdot \nabla v = \langle \nabla u, \nabla v \rangle_{L^2} \text{ et } L(v) := \int_\Omega fv = \langle f, v \rangle_{L^2}, \tag{6.6}$$

(*dans l'ensemble du cours, on ne considère que des fonctions à valeurs réelles, ce qui explique l'absence de quantités conjuguées dans la définition du produit scalaire*). On voit aisément que a est bilinéaire sur $X \times X$ et L est linéaire sur X.

Remarque 6.1 *La formulation variationnelle nous permet aisément d'obtenir un résultat d'unicité pour l'équation, en prouvant, ce qui est équivalent, que $u = 0$ si $f = 0$. En effet, en prenant dans ce cas $v = u$ dans* (6.5)*, on obtient $\nabla u = 0$, donc u est constante et forcément nulle puisque $u_{\partial\Omega} = 0$.*

Remarque 6.2 *Il est important de vérifier que réciproquement, toute solution suffisament régulière de* (6.5) *est aussi solution de la formulation initiale* (6.1)*. En supposant u de classe C^2 et solution de* (6.5)*, on obtient en appliquant la formule de Green dans le sens inverse que*

$$\int_\Omega (\Delta u + f)v = 0, \ \textit{pour tout } v \in X, \tag{6.7}$$

qui implique immédiatement que $-\Delta u = f$ (exercice : on peut au choix évoquer l'égalité $\Delta u + f = 0$ au sens des distributions, ou la densité de X dans $L^2(\Omega)$, ou prouver directement que $\Delta u + f$ est nul en tout point de Ω).

Il est assez naturel que l'espace X où l'on cherche la solution soit le même que celui que parcourent les fonctions v. Remarquons que si X est un espace de dimension finie, toute équation de type (6.5) peut se reformuler sous la forme d'une équation linéaire $Au = f$ où A est l'unique endomorphisme de X tel que $a(u,v) = \langle Au, v \rangle$ et f l'unique élément de X tel que $L(v) = \langle f, v \rangle$, avec $\langle \cdot, \cdot \rangle$ un produit scalaire fixé dans X. Dans un tel cas, l'unicité de la solution signifie que A est injective et donc un isomorphisme ce qui assure l'existence d'une solution. Cependant nous travaillons ici en dimension infinie et ce raisonement n'est plus valable.

L'existence de la solution de (6.5) va découler de la théorie de Lax-Milgram, qui s'applique à des problèmes généraux pouvant se mettre sous la forme variationnelle

$$\textit{Trouver } u \in X \textit{ tel que } a(u,v) = L(v) \textit{ pour tout } v \in X, \tag{V}$$

où X est un *espace de Hilbert*. Le théorème de Lax-Milgram apporte une réponse à l'existence, l'unicité et la stabilité de la solution dans ce cadre précis.

Théorème 6.1 *On suppose que X est un espace de Hilbert et que les formes a et L vérifient les hypothèses suivantes :*

1. *Continuité de L : $|L(v)| \leq C_L \|v\|_X$ pour tout $v \in X$.*
2. *Continuité de a : $|a(u,v)| \leq C_a \|u\|_X \|v\|_X$ pour tout $u, v \in X$.*
3. *Coercivité de a : $a(u,u) \geq \alpha \|u\|_X^2$ pour tout $u \in X$, avec $\alpha > 0$.*

Alors il existe une solution unique u au problème (V) qui vérifie l'estimation a-priori

$$\|u\|_X \leq \frac{\|L\|_{X'}}{\alpha}, \tag{6.8}$$

avec $\|L\|_{X'} = \sup_{\|v\|_X=1} |L(v)|$ la norme de L dans le dual X' de X.

Preuve : L'estimation a-priori s'établit en prenant $v = u$ dans (V) puis en appliquant la continuité de L et la coercivité de a ce qui donne

$$\alpha \|u\|_X^2 \leq C_L \|u\|_X. \tag{6.9}$$

Il suffit alors de prendre $C_L = \|L\|_{X'}$. Cette estimation nous donne aussi l'unicité de la solution.

Pour l'existence, considérons d'abord le cas simple où a est une forme symétrique. Dans ce cas, la continuité et la coercivité de a montrent qu'il s'agit d'un produit scalaire sur $X \times X$ et que la norme $\|v\|_a := \sqrt{a(v,v)}$ est équivalente à la norme $\|\cdot\|_X$. Puisque L est continue, elle l'est aussi par rapport à $\|\cdot\|_a$ et le théorème de représentation de Riesz nous assure donc l'existence d'un unique $u \in X$ tel que $L(v) = a(u,v)$ pour tout $v \in X$.

Dans le cas non-symétrique, on remarque que puisque $v \mapsto a(u,v)$ et $v \mapsto L(v)$ sont continues, on peut écrire $a(u,v) = \langle Au, v \rangle$ et $L(v) = \langle f, v \rangle$ où A est un opérateur continu sur X, $f \in X$ et $\langle \cdot, \cdot \rangle$ un produit scalaire dans X. L'équation (V) s'écrit donc $Au = f$ dans X. L'hypothèse de coercivité nous permet d'affirmer que

$$\alpha \|v\|_X \leq \|Av\|_X, \tag{6.10}$$

pour tout $v \in X$, ce qui entraine que $\mathrm{Im}(A)$ est un sous-espace fermé de X qui se décompose donc suivant $X = \mathrm{Im}(A) \oplus (\mathrm{Im}(A))^\perp$. Considérons à présent $w \in (\mathrm{Im}(A))^\perp$. La coercivité nous montre que

$$\alpha \|w\|_X^2 \leq a(w,w) = \langle Aw, w \rangle = 0. \tag{6.11}$$

Par conséquent $\mathrm{Im}(A) = X$ ce qui prouve l'existence de la solution u. $\diamond$

Remarque 6.3 *Bien qu'un espace de Hilbert s'identifie avec son dual, il sera souvent pertinent de distinguer X et X'. On écrit donc plutôt $a(u,v) = \langle Au, v \rangle_{X',X}$ et $L(v) = \langle f, v \rangle_{X',X}$ où A est un opérateur continu de X dans X', $f \in X'$ et $\langle \cdot, \cdot \rangle_{X',X}$ le produit de dualité entre X' et X. L'équation (V) s'écrit alors $Au = f$ dans X' et le théorème de Lax-Milgram montre que A est isomorphisme de X dans X'.*

Remarque 6.4 *Dans le cas où la forme a est symétrique, on vérifie que toute solution u de (V) est aussi un minimiseur sur X de la fonctionnelle*

$$J(v) := \frac{1}{2} a(v,v) - L(v). \tag{6.12}$$

On pourra vérifier (exercice) que la propriété de coercivité est alors équivalente à la propriété dite de α-convexité

$$J\Big(\frac{v+w}{2}\Big) \leq \frac{J(v)+J(w)}{2} - \frac{\alpha}{8} \|v-w\|_X^2, \tag{6.13}$$

et que toute fonction α-convexe continue sur un espace de Hilbert atteint un minimum qui est unique (indication : prouver que les suites minimisantes sont de Cauchy).

Si l'on revient à présent au problème du laplacien (6.1) et à sa formulation variationnelle (6.5) dans l'espace $X = \{v \in C^1 \; ; \; v_{|\partial\Omega} = 0\}$, on voit que

$$\|v\|_X := \|\nabla v\|_{L^2} = |v|_{H^1}, \tag{6.14}$$

définit une norme sur X telle que les propriétés de continuité et de coercivité sont trivialement satisfaites par a. La continuité de $L(v) = \langle f, v\rangle_{L^2}$ est une conséquence de *l'inégalité de Poincaré* : si Ω est un domaine lipschitzien borné, il existe une constante C_P telle que pour toute fonction f de X on a

$$\|v\|_{L^2} \leq C_P \|\nabla v\|_{L^2}. \tag{6.15}$$

En appliquant successivement l'inégalité de Cauchy-Schwarz et l'inégalité de Poincaré on a donc

$$|L(v)| \leq \|f\|_{L^2}\|v\|_{L^2} \leq C_L \|v\|_X, \tag{6.16}$$

avec $C_L = C_P\|f\|_{L^2}$. Toutes les conditions du théorème de Lax-Milgram sont vérifiées à l'exception du fait que X muni de la norme définie ci-dessus n'est pas un espace complet.

Il est donc nécessaire d'introduire un cadre fonctionnel mieux approprié. Celui-ci se fonde naturellement sur l'espace de Sobolev $H^1(\Omega)$ dont on voit apparaître la semi-norme dans (6.14). Afin de prendre en compte la condition aux limites $v = 0$ sur $\partial\Omega$, nous aurons besoin des résultats concernant la restriction au bord ou *trace* des fonctions des espaces de Sobolev : si Ω est un ouvert lipschitzien alors l'opérateur γ_0 qui à une fonction régulière u définie sur Ω associe sa restriction $\gamma_0 u$ au bord $\partial\Omega$ vérifie l'estimation

$$\|\gamma_0 u\|_{L^2(\partial\Omega)} \leq C\|u\|_{H^1(\Omega)}. \tag{6.17}$$

Cette estimation se prouve facilement dans le cas où Ω est le demi-espace $\mathbb{R}^{d-1} \times \mathbb{R}_+^*$ (exercice), le passage à un ouvert Ω plus général est technique et utilise les ouverts recouvrant la frontière de Ω dans la définition 2.2. Par un argument de densité ceci montre que la trace γ_0 définit un opérateur continu de $H^1(\Omega)$ dans $L^2(\partial\Omega)$. Remarquons que la trace d'une fonction de $L^2(\Omega)$ n'a en général pas de sens puisque le $\partial\Omega$ est un ensemble de mesure nulle.

La continuité de l'opérateur de trace permet de définir l'espace

$$H_0^1(\Omega) := \{v \in H^1(\Omega)\ ;\ \gamma_0 v = 0\}, \tag{6.18}$$

qui est un sous-espace hilbertien de $H^1(\Omega)$. On peut vérifier que $\mathcal{D}(\Omega)$ est dense dans $H_0^1(\Omega)$ qui peut ainsi être défini de manière équivalente comme la fermeture de $\mathcal{D}(\Omega)$ au sens de la norme H^1. Ceci entraine en particulier la validité de l'inégalité de Poincaré (6.15) pour les fonctions de $H_0^1(\Omega)$. Cette inégalité nous indique que la semi-norme $\|\nabla v\|_{L^2}$ est une norme équivalente à la norme H^1 sur $H_0^1(\Omega)$. On la désigne parfois comme "norme H_0^1" et on note donc

$$\|v\|_{H_0^1} = |v|_{H^1} = \|\nabla v\|_{L^2}. \tag{6.19}$$

On désigne par $H^{-1}(\Omega)$ l'espace dual de $H_0^1(\Omega)$ muni de la norme

$$\|v\|_{H^{-1}} := \sup_{\|w\|_{H_0^1}=1} \langle v, w\rangle_{H^{-1},H_0^1}. \tag{6.20}$$

Lorsque v est une fonction de $L^2(\Omega)$, on l'identifie à un élément de $H^{-1}(\Omega)$ en utilisant la forme linéaire définie par le produit scalaire L^2, c'est à dire

$$\|v\|_{H^{-1}} := \sup_{\|w\|_{H_0^1}=1} \int_\Omega vw. \tag{6.21}$$

On a ainsi la chaine d'inclusion

$$H_0^1(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega). \tag{6.22}$$

Remarque 6.5 *Quand $\Omega = \mathbb{R}^d$ la densité de $\mathcal{D}(\Omega)$ montre que $H_0^1(\mathbb{R}^d) = H^1(\mathbb{R}^d)$ et $(H^1(\mathbb{R}^d))' = H^{-1}(\mathbb{R}^d)$. En revanche si Ω est borné lipschitzien ces espaces diffèrent, et l'espace $(H^1(\Omega))'$ n'est pas un espace de distribution (par exemple $v \mapsto \int_{\partial\Omega} \gamma_0 v$ est dans cet espace mais s'annule sur toutes les fonctions de $\mathcal{D}(\Omega)$).*

Si l'on revient à présente à la formulation variationnelle (6.5) de l'équation du laplacien (6.1), on voit qu'en travaillant avec l'espace de Sobolev

$$X := H^1_0(\Omega), \tag{6.23}$$

toutes les hypothèses du théorème de Lax-Milgram sont vérifiées. Le laplacien définit ainsi un isomorphisme de $H^1_0(\Omega)$ dans $H^{-1}(\Omega)$ et on a l'estimation a-priori

$$\|u\|_{H^1_0} \leq \|f\|_{H^{-1}}. \tag{6.24}$$

Si f est dans L^2, on peut aussi écrire

$$\|u\|_{H^1_0} \leq C_P \|f\|_{L^2}, \tag{6.25}$$

où C_P est la constante de l'inégalité de Poincaré qui s'étend aux fonctions de $H^1_0(\Omega)$. Intuitivement la solution u du laplacien gagne deux ordres de dérivabilité par rapport à la donnée f. Ceci est confirmé par le résultat admis suivant.

Théorème 6.2 *Soit Ω un domaine $C^{1,1}$ ou un domaine convexe. Alors si $f \in L^2(\Omega)$, la solution u de la formulation variationnelle de* (6.1) *est dans H^2 et il existe une constante C qui ne dépend que de Ω telle que*

$$\|u\|_{H^2} \leq C\|f\|_{L^2}. \tag{6.26}$$

Plus généralement, si Ω est de classe $C^{m+1,1}$ alors $f \in H^m(\Omega)$ implique $u \in H^{m+2}(\Omega)$, avec

$$\|u\|_{H^{m+2}} \leq C\|f\|_{H^m}. \tag{6.27}$$

Remarque 6.6 *Il est possible de montrer que ce résultat n'est pas valide pour des domaines moins réguliers, en particulier des polygones non convexes.*

Voici à présent quelques exemples d'application de la théorie de Lax-Milgram dans les espaces de Sobolev pour d'autres équations aux dérivées partielles mises sous forme variationnelle.

Exemple 1. Laplacien avec conditions non-homogènes
On considère le problème de Laplace avec une donnée de Dirichlet non-homogène au bord.

$$-\Delta u = f \ \text{ dans } \ \Omega \ \text{ et } \ \gamma_0 u = g, \tag{6.28}$$

Pour étudier ce problème, il est utile de comprendre quelle fonctions de $L^2(\partial\Omega)$ sont les traces de fonctions de $H^1(\Omega)$. La réponse est donnée par le résultat admis suivant.

Théorème 6.3 *Soit Ω un domaine lipschitzien. L'image de $H^1(\Omega)$ par l'opérateur de trace γ_0 est l'espace $H^{1/2}(\partial\Omega)$. On peut définir une norme sur $H^{1/2}(\partial\Omega)$ par*

$$\|v\|_{H^{1/2}(\partial\Omega)} := \inf_{\gamma_0 w = v} \|w\|_{H^1(\Omega)}, \tag{6.29}$$

qui est équivalente à la norme $H^{1/2}$ obtenue par la définition 2.3 adaptée à $\partial\Omega$. Il existe un opérateur de relevement R linéaire et continu de $H^{1/2}(\partial\Omega)$ dans $H^1(\Omega)$ tel que $\gamma_0 \circ R(v) = v$ pour tout $v \in H^{1/2}(\partial\Omega)$ et tel que $\|R(v)\|_{H^1} = \|v\|_{H^{1/2}}$.

On peut vérifier (exercice) que l'infimum dans la définition ci-dessus de la norme $H^{1/2}$ est atteint par la fonction $w \in H^1$ solution de $-\Delta w + w = 0$ dans Ω avec condition aux limites $\gamma_0 w = v$. L'opérateur R de relèvement minimal est donc donné par la résolution de cette équation.

On peut reformuler l'équation (6.28) en considérant un relèvement $R(g)$ de g, et en écrivant $u = \tilde{u} + R(g)$. On voit ainsi que $\tilde{u}$ s'annule au bord et est solution de l'équation

$$-\Delta\tilde{u} = f + \Delta R(g), \tag{6.30}$$

que l'on peut mettre sous la forme variationnelle suivante : trouver $\tilde{u} \in H_0^1(\Omega)$ tel que

$$\int_\Omega \nabla\tilde{u} \cdot \nabla v = \int_\Omega fv - \int_\Omega \nabla R(g) \cdot \nabla v, \quad v \in H_0^1(\Omega). \tag{6.31}$$

On peut vérifier que le choix du relèvement ne change rien à la solution de (6.31).

On voit que dans cette formulation variationnelle on peut supposer $g \in H^{1/2}(\partial\Omega)$ et utiliser le théorème 6.3, ce qui montre que la forme lineaire du membre de droite est bornée suivant

$$|L(v)| \leq \|v\|_{H_0^1}(\|f\|_{H^{-1}} + \|g\|_{H^{1/2}}). \tag{6.32}$$

On obtient ainsi $\|\tilde{u}\|_{H_0^1} \leq \|f\|_{H^{-1}} + \|g\|_{H^{1/2}}$ et finalement en ajoutant $R(g)$,

$$\|u\|_{H^1} \leq \|f\|_{H^{-1}} + 2\|g\|_{H^{1/2}}. \tag{6.33}$$

On voit ainsi que $u \mapsto (\Delta u, \gamma_0 u)$ définit un isomorphisme de $H^1(\Omega)$ sur $H^{-1}(\Omega) \times H^{1/2}(\partial\Omega)$.

Exemple 2. Coefficients variables
On peut généraliser l'équation du laplacien en introduisant des coefficients variables :

$$\sigma u - \mathrm{div}(A.\nabla u) = f \;\text{ dans }\; \Omega \;\text{ et }\; \gamma_0 u = 0, \tag{6.34}$$

où $\sigma \in L^\infty(\Omega)$ est une fonction positive, et A est une fonction de Ω à valeur dans l'ensemble des matrices symétriques définies positives $d \times d$, telle que les valeurs propres maximales et minimales de $A(x)$ vérifient

$$0 < \alpha \leq \lambda_{\min}(A(x)) \leq \lambda_{\max}(A(x)) \leq C, \tag{6.35}$$

où α et C sont indépendantes de x. La forme bilinéaire est alors donnée par

$$a(u,v) = \int_\Omega \sigma uv + \int_\Omega A\nabla u \cdot \nabla v, \tag{6.36}$$

et les hypothèses de Lax-Milgram dans $H_0^1(\Omega)$ découlent aisément de la propriété (6.35).

Exemple 3. Equation de convection diffusion
On considère l'équation

$$-\Delta u + b \cdot \nabla u = f \;\text{ dans }\; \Omega \;\text{ et }\; \gamma_0 u = 0, \tag{6.37}$$

où $b(x)$ est un champ de vecteur, i.e. une application de Ω à valeur dans $\mathbb{R}^d$ que l'on suppose de classe C^1 et uniformément bornée sur Ω et telle que

$$\mathrm{div}(b(x)) = 0, \;\; x \in \Omega. \tag{6.38}$$

La formulation variationnelle est alors : trouver $u \in H_0^1(\Omega)$ tel que

$$\int_\Omega \nabla u \cdot \nabla v + \int_\Omega (b \cdot \nabla u)v = \int_\Omega fv, \quad v \in H_0^1(\Omega). \tag{6.39}$$

Dans ce cas, la forme bilinéaire n'est pas symétrique. Il est facile de vérifier sa continuité sur H_0^1. Pour la coercivité, on utilise la condition de divergence nulle de b qui implique $\mathrm{div}(buv) = (b \cdot \nabla u)v + (b \cdot \nabla v)u$, ce qui entraîne l'antisymétrie du terme $\beta(u,v) = \int (b \cdot \nabla u)v$. Il s'en suit que $a(u,u) = \|\nabla u\|_{L^2}^2$ et la coercivité est donc vérifiée.

Exemple 4. Problème de Neumann
Le problème

$$-\Delta u + u = f \;\text{ dans }\; \Omega \;\text{ et }\; \gamma_1 u = g, \tag{6.40}$$

peut se mettre sous forme variationnelle en multipliant par une fonction arbitraire de $H^1(\Omega)$ et en appliquant la formule de Green. En remplacant $\frac{\partial u}{\partial n}$ par sa valeur au bord, on obtient ainsi : trouver $u \in H^1(\Omega)$ tel que

$$\int_\Omega \nabla u \cdot \nabla v + \int_\Omega uv = \int_\Omega fv + \int_{\partial\Omega} g\gamma_0 v, \quad v \in H^1(\Omega). \tag{6.41}$$

Notons que la condition de Neumann aux limites n'apparait plus dans l'espace dans lequel on travaille, mais elle est implicitement contenue dans la formulation variationnelle. En effet si u est suffisament régulière, en considérant d'abord $v \in H_0^1(\Omega)$ dans (6.41), on élimine le terme $\int_{\partial\Omega} gv$ et on retrouve l'équation (6.40), puis on obtient que pour tout $v \in H^1(\Omega)$ on a

$$\int_{\partial\Omega} (\frac{\partial u}{\partial n} - g)\gamma_0 v = 0, \tag{6.42}$$

ce qui entraine la validité de la condition aux limites.

Il convient de préciser dans quel espace on choisit la donnée f et la condition au bord g. On définit $H^{-1/2}(\partial\Omega)$ l'espace dual de $H^{1/2}(\partial\Omega)$. Cet espace joue un rôle naturel dans la définition de la trace de la dérivée normale. On note γ_1, l'opérateur qui à u associe la fonction $\frac{\partial u}{\partial n}$ définie sur $\partial\Omega$.

Théorème 6.4 *Soit $u \in H^1(\Omega)$ telle que $\Delta u \in L^2(\Omega)$, alors $\gamma_1 u$ définit un élément de $H^{-1/2}(\partial\Omega)$ par la formule*

$$\langle \gamma_1 u, v\rangle_{H^{-1/2},H^{1/2}} := \int_\Omega \nabla u \cdot \nabla w + \int_\Omega \Delta u w, \tag{6.43}$$

pour tout $w \in H^1(\Omega)$ tel que $v = \gamma_0 w$.

Preuve : La définition est justifiée car si u et w sont suffisament régulières, la formule de Green donne

$$\int_{\partial\Omega} (\gamma_1 u)v = \int_\Omega \nabla u w + \int_\Omega \Delta u w \tag{6.44}$$

En considérant le relèvement $R(v)$, posont d'abord

$$\langle \gamma_1 u, v\rangle_{H^{-1/2},H^{1/2}} := \int_\Omega \nabla u \cdot \nabla R(v) + \int_\Omega \Delta u R(v), \tag{6.45}$$

En utilisant l'inégalité de Schwarz et le théorème 6.3, on obtient donc

$$|\int_{\partial\Omega} (\gamma_1 u)v| \leq (\|\nabla u\|_{L^2}^2 + \|\Delta u\|_{L^2}^2)^{1/2}\|v\|_{H^{1/2}}, \tag{6.46}$$

ce qui montre que $\|\gamma_1 u\|_{H^{-1/2}} \leq (\|\nabla u\|_{L^2}^2 + \|\Delta u\|_{L^2}^2)^{1/2}$. Il reste à vérifier que la définition de $\gamma_1 u$ est indépendante du choix de w tel que $\gamma_0 w = v$. Ceci provient du fait que par définition du Laplacien au sens des distribution $\int_\Omega \Delta u w = -\int_\Omega \nabla u \cdot \nabla w$ pour tout $w \in \mathcal{D}(\Omega)$ et donc pour tout $w \in H_0^1(\Omega)$. $\diamond$

Au vu du théorème 6.4, on prend $f \in L^2$ et $g \in H^{-1/2}(\Omega)$, ce qui entraine la continuité sur $H^1(\Omega)$ de la forme linéaire dans (6.41). La continuité et la coercivité de la forme bilinéaire est immédiate, et les hypothèses de Lax-Milgram sont donc vérifiées.

Exemple 5. Laplacien avec conditions de Neumann

Toujours avec $f \in L^2(\Omega)$ et $g \in H^{-1/2}(\Omega)$ on considère le problème

$$-\Delta u = f \text{ dans } \Omega \text{ et } \gamma_1 u = g, \tag{6.47}$$

On fait ici l'hypothèse supplémentaire que Ω est *connexe*. La solution de ce problème est définie à une constante près. Afin de lever cette indétermination, on va chercher u dans l'espace quotient

$$H^1(\Omega)/\mathbb{R} := \{u \in H^1(\Omega)\ ;\ \int_\Omega u = 0\}. \tag{6.48}$$

On vérifie aisément que cet espace est un sous espace hilbertien de $H^1(\Omega)$. Comme $H_0^1(\Omega)$, on peut le munir de la norme

$$\|v\|_{H^1/\mathbb{R}} := \|\nabla v\|_{L^2}. \tag{6.49}$$

On remarque que les données de l'équation ne sont pas indépendantes : on a nécessairement

$$\int_\Omega f + \int_{\partial\Omega} g = -\int_\Omega \Delta u + \int_{\partial\Omega} \gamma_1 u = 0. \tag{6.50}$$

Sous cette hypothèse, examinons la formulation variationnelle : trouver u dans $H^1(\Omega)/\mathbb{R}$ telle que

$$\int_\Omega \nabla u \cdot \nabla v = \int_\Omega fv + \int_{\partial\Omega} g\gamma_0 v, \quad v \in H^1(\Omega)/\mathbb{R}. \tag{6.51}$$

En utilisant la relation entre f et g, on voit que l'égalité dans (6.51) est automatiquement vérifiée pour tout $v \in H^1$ puisqu'on peut lui ajouter n'importe quelle constante. Cela permet de montrer en raisonnant comme dans l'exemple précédent qu'une solution u suffisament régulière de (6.51) est aussi solution de (6.47).

La continuité et la coercivité de la forme bilinéaire est immédiate pour la norme de $H^1(\Omega)/\mathbb{R}$. La continuité de la forme linéaire utilise le résultat suivant qui est l'analogue de l'inégalité de Poincaré pour les fonctions de moyenne nulle, et un cas particulier du Théorème 3.1.

Théorème 6.5 *Soit Ω borné, lipschitzien et connexe. Il existe une constante C_Q qui ne dépend que du domaine Ω telle que pour tout $v \in H^1(\Omega)/\mathbb{R}$*

$$\|v\|_{H^1} \leq C_Q \|\nabla v\|_{L^2}. \tag{6.52}$$

On peut à présent écrire

$$\begin{aligned} |\textstyle\int_\Omega fv + \int_{\partial\Omega} g\gamma_0 v| &\leq \|f\|_{L^2}\|v\|_{L^2} + \|g\|_{H^{-1/2}}\|\gamma_0 v\|_{H^{1/2}} \\ &\leq C_Q\|f\|_{L^2}\|\nabla v\|_{L^2} + \|g\|_{H^{-1/2}}\|v\|_{H^1} \\ &\leq C_Q(\|f\|_{L^2} + \|g\|_{H^{-1/2}})\|\nabla v\|_{L^2}. \end{aligned}$$

La continuité de la forme linéaire est ainsi démontrée et le théorème de Lax-Milgram s'applique.

Remarque 6.7 *Une analyse très proche permet de résoudre le problème du Laplacien sur le tore T avec conditions aux limites périodiques en prenant $X = H^1_{per}/\mathbb{R}$ (exercice).*

Exemple 6. Bi-Laplacien

La formulation variationnelle du problème

$$\Delta^2 u = f, \;\; \gamma_0 u = \gamma_1 u = 0, \tag{6.53}$$

utilise le sous-espace espace hilbertien de $H^2(\Omega)$

$$H^2_0(\Omega) := \{v \in H^2(\Omega) \; ; \; \gamma_0 v = \gamma_1 v = 0\}, \tag{6.54}$$

qui est aussi la fermeture de $\mathcal{D}(\Omega)$ pour la norme H^2.

En appliquant deux fois la formule de Green, on obtient la formulation variationnelle : trouver $u \in H^2_0(\Omega)$ telle que

$$\int_\Omega \Delta u \Delta v = \int_\Omega fv, \quad v \in H^2_0(\Omega). \tag{6.55}$$

On vérifie aisément que toute solution de (6.55) suffisament régulière est aussi solution de (6.53).

Sur $H^2_0(\Omega)$ on peut travailler avec la norme

$$\|u\|_{H^2_0(\Omega)} := \|\Delta u\|_{L^2}, \tag{6.56}$$

qui rend immédiate la continuité et la coercivité de la forme bi-linéaire (il s'agit bien d'une norme puisque $\Delta u = 0$ et $\gamma_0 u = 0$ implique $u = 0$). Pour la continuité de la forme linéaire, on a besoin d'une équation analogue à Poincaré pour le laplacien : pour tout $v \in H^2_0(\Omega)$,

$$\|v\|_{L^2} \leq C_B \|\Delta v\|_{L^2}. \tag{6.57}$$

Cette estimation est une conséquence directe de la régularité H^2 des solutions du laplacien dans le cas où Ω est convexe ou de classe $C^{1,1}$. Elle est aussi vérifiée pour un domaine lipschitzien.

La formulation variationnelle permet de donner un sens aux solutions du bi-laplacien quand f est dans le dual de $H^2_0(\Omega)$ que l'on note $H^{-2}(\Omega)$, et pour $f \in L^2$ on a donc

$$\|f\|_{H^{-2}} = \sup_{\|v\|_{H^2_0}=1} \int_\Omega fv \leq C_B \|f\|_{L^2}. \tag{6.58}$$

7 Approximation variationnelle des EDP

On s'intéresse à présent à l'approximation numérique des solutions de problèmes admettant une formulation variationnelle de type (V) dans un espace de Hilbert X. Une méthode *d'approximation interne* ou méthode de Galerkin, consiste à rechercher une solution approchée dans un *sous-espace hilbertien* $X_n \subset X$. Pour cela on définit un problème approché en substituant X_n à X dans (V) :

$$\textit{Trouver } u_n \in X_n \textit{ tel que } a(u_n, v_n) = L(v_n) \textit{ pour tout } v_n \in X_n. \tag{G}$$

L'un des objectif est d'aboutir à une solution calculable en un nombre fini d'opérations. Pour cela, on va supposer X_n de dimension finie n et on introduit une base $(\varphi_k)_{k=1,\cdots,n}$ de X_n. La solution u_n peut donc se décomposer suivant :

$$u_n = \sum_{k=1}^{n} u_{n,k}\varphi_k, \tag{7.1}$$

et on voit que u_n est solution de (G) si et seulement si

$$a(u_n, \varphi_n) = \sum_{k=1}^{n} u_{n,h} a(\varphi_k, \varphi_l) = L(\varphi_l), \quad l = 1, \cdots, n. \tag{7.2}$$

Ces équations peuvent se reformuler sous la forme d'un système

$$AU = F \tag{7.3}$$

où $U := (u_{n,k})_{k=1,\cdots,n}$, $F := (L(\varphi_k))_{k=1,\cdots,n}$ et $A_{k,l} = a(\varphi_k, \varphi_l)$ pour $k, l = 1, \cdots, n$. La matrice A est appelée *matrice de rigidité* du problème. Cette matrice est symétrique lorsque a est symétrique.

L'un des intérêts de la méthode de Galerkin est qu'elle permet d'aboutir à une estimation d'erreur optimale entre la solution exacte u du problème (V) et la solution approchée u_n, au sens où l'erreur $\|u - u_n\|_X$ est comparable à l'erreur de meilleure approximation $\sigma_n(u)_X := \min_{v_n \in X_n} \|u - v_n\|_X$. Ceci est exprimé par le lemme de Cea.

Lemme 7.1 *On suppose que les hypothèses du théorème de Lax-Milgram sont vérifiées. Alors, il existe une solution unique u_n au problème approché (G) et la matrice A est inversible. On a l'estimation d'erreur*

$$\|u - u_n\|_X \leq \frac{C_a}{\alpha} \min_{v_n \in X_n} \|u - v_n\|_X. \tag{7.4}$$

Dans le cas où a est symétrique on a

$$\|u - u_n\|_X \leq \left(\frac{C_a}{\alpha}\right)^{1/2} \min_{v_n \in X_n} \|u - v_n\|_X. \tag{7.5}$$

Preuve : Comme X_n est un sous-espace Hilbertien de X, les hypothèses de Lax-Milgram pour (V) entrainent la validité des mêmes hypothèses pour (G). On en déduit l'existence et l'unicité de u_n (ainsi que l'estimation a-priori $\|u_n\|_X \leq \frac{\|L\|_{X'}}{\alpha}$). Comme (φ_k) est une base, l'unicité montre que $U = 0$ si $F = 0$ ce qui signifie que A est inversible.

Pour l'estimation d'erreur, on remarque en combinant (V) et (G) que pour tout $v_n \in X_n$ on a

$$a(u - u_n, v_n) = 0. \tag{7.6}$$

En utilisant cette remarque ainsi que la coercivité et la continuité de a, on obtient

$$\begin{aligned} \alpha\|u - u_n\|_X^2 &\leq a(u - u_n, u - u_n) \\ &= a(u - u_n, u_n - v_n) \\ &\leq C_a\|u - u_n\|_X\|u - v_n\|_X, \end{aligned}$$

ce qui donne l'estimation recherchée dans le cas général.

Dans le cas où a est symétrique, on remarque que (7.6) signifie que u_n est la projection orthogonale de u sur X_n au sens du produit scalaire $a(\cdot,\cdot)$, et donc

$$\|u - u_n\|_a = \min_{v_n \in X_n} \|u - v_n\|_a. \tag{7.7}$$

On en déduit l'estimation recherchée en utilisant l'équivalence de norme

$$\alpha^{1/2}\|v\|_X \leq \|v\|_a \leq C_a^{1/2}\|v\|_X, \tag{7.8}$$

qui découle de la continuité et de la coercivité de a. ◇

Bien entendu, il convient d'utiliser des espaces X_n pertinents pour appliquer la méthode de Galerkin. En premier lieu, il s'agit d'obtenir un compromis raisonable entre la complexité du système décrit par la dimension n de X_n et la précision dans l'approximation de la solution. D'autre part, il est important que les espaces X_n aient des propriétés permettant de calculer de façon simple et rapide les quantités $a(\varphi_k, \varphi_l)$ et $L(\varphi_k)$ qui interviennent dans le système (7.3). Enfin, il est important que ce système puisse être résolu aisément, en particulier lorsque n devient grand. Il peut ainsi être intéressant que X_n possède une base qui rend la matrice de rigidité *creuse*.

Dans le cas du problème du laplacien, comme pour de nombreux autres exemples d'EDP, les espaces X_h d'éléments finis de Lagrange étudiés dans la section §5 répondent à l'ensemble de ces exigences. Dans ce cas, on notera $u_h \in X_h$ la solution approchée définie par

$$a(u_h, v_h) = L(v_h), \quad v_h \in X_h. \tag{7.9}$$

Dans le cas du problème de Laplace (6.1), il faut ajouter les conditions aux limites homogènes de Dirichlet dans la construction de l'espace X_h : on définit

$$X_{h,0} = X_h \cap H_0^1(\Omega). \tag{7.10}$$

C'est aussi le sous-espace des fonctions de X_h dont les degrés de libertés aux bords sont nuls. Une base de cet espace est donné par la sous famille $(\varphi_\gamma)_{\gamma \in \Gamma_{0,h}}$ avec

$$\Gamma_{0,h} := \{\gamma \in \Gamma_h \setminus \partial\Omega\}, \tag{7.11}$$

l'ensemble des degrés de liberté intérieur au domaine. On remarque que l'interpolant I_h envoie les fonctions continues et nulles aux bord dans $X_{h,0}$. La solution approchée $u_h \in X_{h,0}$ est maintenant définie par

$$a(u_h, v_h) = L(v_h), \quad v_h \in X_{h,0}. \tag{7.12}$$

D'après le lemme de Cea, elle vérifie

$$\|u - u_h\|_{H_0^1} \leq C \min_{v_h \in X_{h,0}} C\|u - I_h u\|_{H_0^1}, \tag{7.13}$$

et donc si la solution exacte u est dans $H^s(\Omega)$ avec $d/2 < s \leq k+1$, on a

$$\|u - u_h\|_{H_0^1} \leq C h^{s-1} |v|_{H^s}. \tag{7.14}$$

Remarque 7.1 *La restriction $s > d/2$ nécessaire pour controler l'erreur d'interpolation est en fait artificielle pour controler $\|u - u_h\|_{H_0^1}$ car on peut utiliser d'autres opérateurs que l'interpolant qui s'appliquent aux fonctions de H^s pour tout $s \geq 0$. Il suffit donc d'avoir $s > 1$.*

Remarque 7.2 *Si Ω est convexe et $f \in L^2(\Omega)$, on peut utiliser le théorème de régularité 6.2 qui nous indique que la norme H^2 de u est controlée par la norme L^2 de f. On a ainsi*

$$\|u - u_h\|_{H_0^1} \leq C h \|f\|_{L^2}. \tag{7.15}$$

Le lemme de Cea fournit une estimation d'erreur dans la norme de l'espace de Hilbert X pour lequel la formulation variationnelle vérifie les hypothèse de Lax-Milgram, mais on peut-être intéressé par des estimations d'erreur dans d'autres normes. Dans le cas du laplacien évoqué à la fin de la section précèdente, on sait en particulier que si $u \in H^2(\Omega)$, on a

$$\min_{v_h \in X_{h,0}} \|u - v_h\|_{L^2} \leq \|u - I_h u\|_{L^2} \leq Ch^2 |u|_{H^2},$$

mais cela ne nous permet pas de conclure à une estimation similaire pour $\|u - u_h\|_{L^2}$. On peut l'obtenir par un argument de dualité : le lemme d'Aubin-Nitsche.

Lemme 7.2 *Si Ω est un polygone ou polyèdre convexe, l'approximation u_h de la solution de* (6.1) *par la méthode de Galerkin dans l'espace $X_{h,0}$ des éléments finis $\mathbb{P}_k$ vérifie l'estimation*

$$\|u - u_h\|_{L^2} \leq Ch \|u - u_h\|_{H_0^1}. \tag{7.16}$$

Preuve : On introduit la fonction w solution du problème auxiliaire

$$-\Delta w = u - u_h \text{ dans } \Omega, \ \ w_{|\partial\Omega} = 0, \tag{7.17}$$

et on a alors d'après la formule de Green

$$\begin{aligned} \|u - u_h\|_{L^2}^2 &= -\textstyle\int_\Omega (u - u_h) \Delta w \\ &= a(u - u_h, w) \\ &= a(u - u_h, w - w_h) \\ &\leq \|u - u_h\|_{H_0^1} \|w - w_h\|_{H_0^1}, \end{aligned}$$

pour tout $w_h \in X_{h,0}$. En prenant par exemple $w_h = I_h w$ on obtient par le théorème 5.3

$$\|w - w_h\|_{H_0^1} \leq Ch |w|_{h^2}, \tag{7.18}$$

et on en déduit en utilisant le théorème 6.2

$$\|w - w_h\|_{H_0^1} \leq Ch \|u - u_h\|_{L^2}. \tag{7.19}$$

Le résultat s'en déduit. $\diamond$

Nous terminons cette section en discutant la résolution numérique du système

$$AU = F \tag{7.20}$$

qui calcule la solution de Galerkin d'une formulation variationnelle (V) dans la base des éléments finis. Celle ci peut parfois se faire par des méthodes directes d'inversion (telles que la méthode de Gauss, ou la factorisation en matrices triangulaires inférieures et supérieures) mais dans le cas de systèmes de très grande taille, on est souvent amené à utiliser une *méthode itérative.* L'exemple le plus élémentaire est l'itération de Richardson ; en partant de $U^0 = 0$, on pose

$$U^n = U^{n-1} + \tau(F - AU^{n-1}) \tag{7.21}$$

avec $\tau > 0$. Dans le cas où A est définie positive, ce n'est rien d'autre qu'un algorithme de gradient à pas fixe pour la minimisation de

$$J(V) = \frac{1}{2}\langle AV, V \rangle - \langle F, V \rangle. \tag{7.22}$$

Comme U est un point fixe de l'itération (7.21), en définissant l'erreur d'itération $E = U - U^n$, on a

$$E^n = (I - \tau A)E^{n-1} = \cdots = (1 - \tau A)^n E^0 = (1 - \tau A)^n U. \tag{7.23}$$

On peut choisir diverse normes pour mesurer l'erreur E, mais un choix naturel est la norme d'énergie

$$\|V\|_A := \langle AV, V \rangle = a(v, v) = \|v\|_a, \tag{7.24}$$

où v est la fonction de X_h de vecteur de coordonnées V dans la base nodale. Si u_h^n est la fonction de X_h de coordonnées U^n, on souhaite typiquement que la précision $\|u_h - u_h^n\|_a$ soit du même ordre que l'erreur d'approximation $\|u - u_h\|_a$.

Afin de comprendre la convergence de l'erreur d'itération, placons nous dans le cas où a est symmétrique, i.e. A est symmétrique définie positive. On a alors $\| \cdot \|$

$$\|E^n\|_A \leq \rho^n \|U\|_A \tag{7.25}$$

où le facteur de réduction ρ est donné par

$$\rho := \max\{|1 - \tau\lambda_{\min}(A)|, |1 - \tau\lambda_{\max}(A)|\} = \max\{1 - \tau\lambda_{\min}(A), \tau\lambda_{\max}(A) - 1\}. \tag{7.26}$$

Le choix de τ minimisant ρ est

$$\tau = \frac{2}{\lambda_{\min}(A) + \lambda_{\max}(A)}, \tag{7.27}$$

qui donne un facteur de réduction

$$\rho = \frac{\kappa(A) - 1}{\kappa(A) + 1}, \tag{7.28}$$

où

$$\kappa(A) = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}, \tag{7.29}$$

est le *nombre de conditionnement* de la matrice A. Si on utilise l'algorithme du *gradient conjugué* qui est plus performant, le facteur de réduction est alors

$$\rho = \frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1}. \tag{7.30}$$

Le nombre de conditionnement joue donc un rôle important dans la complexité de la résolution numérique puisqu'il gouverne le nombre d'itérations nécessaire pour atteindre une précision donnée. Le résultat suivant montre dans le cas particulier du laplacien que le nombre de conditionnement augmente vers $+\infty$ lorsque l'on augmente la résolution du maillage.

Théorème 7.1 *Dans le cas de la résolution de l'équation du laplacien par la méthode des éléments finis $\mathbb{P}_k$ dans une famille régulière de triangulations, on a*

$$\kappa(A) \geq ch^{-2}. \tag{7.31}$$

Si la famille est quasi-uniforme, on a aussi

$$\kappa(A) \leq Ch^{-2}. \tag{7.32}$$

Preuve : pour simplifier, on se place dans le cas des éléments $\mathbb{P}_1$, l'adaptation de la preuve pour les éléments $\mathbb{P}_k$ est laissée en exercice. On va estimer les valeurs propres $\lambda_{\min}(A)$ et $\lambda_{\max}(A)$. On a

$$\lambda_{\max}(A) = \max_V \frac{\langle AV, V\rangle}{\|V\|^2} \text{ et } \lambda_{\min}(A) = \min_V \frac{\langle AV, V\rangle}{\|V\|^2}. \tag{7.33}$$

Pour $\lambda_{\max}$ on prend un sommet $\gamma \in \Gamma_{h,0}$ et on considère la fonction de base associée $v_h = \varphi_\gamma$ et son vecteur de coordonnée V, qui vérifie clairement $\|V\| = 1$. Pour ce vecteur on a

$$\langle AV, V\rangle = a(v_h, v_h) = \int \|\nabla\varphi_\gamma\|^2. \tag{7.34}$$

La fonction φ_γ est supportée sur les simplexes T dont l'un des sommets est γ. Sur chacun de ces triangles son gradient est constant et vérifie

$$\|\nabla\varphi_\gamma\| \geq h_T^{-1}. \tag{7.35}$$

Comme le volume de T est au moins égale à celui de la boule inscrite qui vaut $C\rho_T^d$ avec C une constante qui dépend de d, on a

$$\langle AV, V\rangle \geq C \sum_{T \text{ t.q. } \gamma\in T} \rho_T^d h_T^{-2}. \tag{7.36}$$

En prenant un sommet γ tel que pour un T dans la somme ci-dessus on a $h_T = h$ et en utilisant la régularité des partitions, on trouve donc

$$\lambda_{\max}(A) \geq Ch^{d-2}. \tag{7.37}$$

Pour $\lambda_{\min}$ on se fixe une fonction $v \in H^2 \cap H_0^1$ telle que v est positive ou nulle et identiquement égale à 1 à l'intérieur d'une boule B contenue dans Ω (il est facile de vérifier qu'une telle fonction existe). On considère alors $v_h = I_h v$ et son vecteur de coordonnées V. Les coordonnées $v_h(\gamma)$ de v_h sont positives, et valent 1 si $\gamma \in B$. Par conséquent

$$\|V\|^2 \geq M_h, \tag{7.38}$$

où M_h est le nombre de points de Γ_h contenu dans B. Comme le maillage est de finesse h, on a $M_h \geq Ch^{-d}$ où la constante C ne dépend que de d, et donc

$$\|V\|^2 \geq Ch^{-d}. \tag{7.39}$$

On a d'autre part

$$\begin{aligned}\langle AV, V\rangle &= \|v_h\|_{H_0^1}^2 \\ &\leq (\|v\|_{H_0^1} + \|v - v_h\|_{H_0^1})^2 \\ &\leq (\|v\|_{H_0^1} + h\|v\|_{H^2})^2 \\ &\leq C,\end{aligned}$$

où $C = (\|v\|_{H_0^1} + \mathrm{diam}(\Omega)\|v\|_{H^2})^2$ est une constante. On en déduit

$$\lambda_{\min}(A) \leq Ch^d. \tag{7.40}$$

En combinant (7.37) et (7.40), on obtient (7.31).

Dans le cas de partitions régulières et quasi-uniforme, on obtient (7.32) en prouvant d'abord l'existence de deux constantes $C_1, C_2 \geq 0$ telle que pour tout $h > 0$ et $v_h \in X_{h,0}$ de vecteur de coordonnées V, on a

$$C_1 h^d \|V\|^2 \leq \|v_h\|_{L^2} \leq C_2 h^d \|V\|^2. \tag{7.41}$$

Pour cela, on remarque d'abord que par l'équivalence des normes en dimension finie, il existe deux constantes $c_1, c_2 \geq 0$ telles que pour tout $\hat{\pi} \in \mathbb{P}_1$

$$c_1 \sum_{i=0}^d |\hat{\pi}(\hat{a}_i)|^2 \leq \|\hat{\pi}\|_{L^2(\hat{T})}^2 \leq c_2 \sum_{i=0}^d |\hat{\pi}(\hat{a}_i)|^2, \tag{7.42}$$

ce qui par changement de variable donne

$$c_1 \frac{|T|}{|\hat{T}|} \sum_{i=0}^d |\pi(a_i)|^2 \leq \|\pi\|_{L^2(T)}^2 \leq c_2 \frac{|T|}{|\hat{T}|} \sum_{i=0}^d |\pi(a_i)|^2. \tag{7.43}$$

En sommant sur tous les simplexes, on obtient

$$c_1 \sum_{\gamma\in\Gamma_{h,0}} w_\gamma |v(\gamma)|^2 \leq \|v_h\|_{L^2} \leq c_2 \sum_{\gamma\in\Gamma_{h,0}} w_\gamma |v(\gamma)|^2 \tag{7.44}$$

où le poids w_γ est égal au volume total des simplexes dont γ est un sommet multiplié par $d!$. Comme la famille est quasi-uniforme, on a

$$\tilde{c}_1 h^d \leq w_\gamma \leq \tilde{c}_2 h^d, \tag{7.45}$$

avec deux constantes $\tilde{c}_2, \tilde{c}_2 > 0$ indépendantes de γ et h, ce qui nous permet d'aboutir à (7.41).

On écrit ensuite pour tout $v_h \in X_{h,0}$ de vecteur de coordonnées V

$$\langle AV, V\rangle = \|v_h\|^2_{H^1_0} \geq \frac{1}{C_P^2}\|v_h\|^2_{L^2}, \tag{7.46}$$

où C_P est la constante de Poincaré sur Ω. En combinant ceci avec l'inégalité de gauche dans (7.41), on obtient

$$\lambda_{\min}(A) \geq \frac{C_1}{C_P^2} h^d. \tag{7.47}$$

D'autre part, l'inégalité inverse établie dans le théorème 5.4 nous montre que

$$\langle AV, V\rangle = \|v_h\|^2_{H^1_0} \leq Ch^{-2}\|v_h\|^2_{L^2}, \tag{7.48}$$

ce qui combiné avec l'inégalité de droite dans (7.41), nous donne

$$\lambda_{\max}(A) \leq Ch^{d-2}. \tag{7.49}$$

On obtient ainsi l'estimation (7.32). $\diamond$

La dégradation du nombre de conditionnement avec le raffinement du maillage est un phénomène très général dans l'approximation des EDP. Elle justifie des travaux importantes sur des méthodes permettant de *préconditionner* les systèmes linéaires (par exemple les méthodes multigrilles).

Quelques références bibliographiques

R.A. Adams, "Sobolev spaces", Academic Press (1975).

C. Bernardi, Y. Maday et F. Rapetti, "Discrétisations variationnelles de problèmes aux limites elliptiques", Springer-Verlag (2004).

H. Brézis, "Analyse Fonctionnelle - théorie et applications", Dunod (1983)

P.G. Ciarlet, "The finite element method for elliptic problems", North Holland (1978)

R. DeVore and G. Lorentz, "Constructive approximation", Springer Verlag (1993)

T. Korner, "Fourier analysis", Cambridge University Press (1988)