Approximations variationnelles des EDP
Notes du Cours de M2

Albert Cohen

Dans ce cours, on s'intéresse à l'approximation numérique d'équations aux dérivées partielles linéaires qui admettent une *formulation variationnelle*, c'est à dire dont la solution u est aussi solution du problème

(V) *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. 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 reflettent 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 : *trouver* $u_h \in X_h$ *tel que* $a(u_h, v_h) = L(v_h)$ *pour tout* $v_h \in X_h$, où X_h est un sous-espace de dimension finie de X.

Le cours se concentre autour de ce dernier aspect qui pose la question du contrôle de l'erreur $u - u_h$ entre la solution exacte et la solution approchée. On s'interessera tout particulièrement à la *méthode des éléments finis* dans laquelle les fonctions de X_h sont polynomiales par morceaux sur une partition du domaine de la solution de l'équation.

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@ann.jussieu.fr.

1 Formulations et approximations variationnelles

1.1 Un exemple fondamental

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

$$-\Delta u = f \text{ dans } \Omega \text{ et } u_{|\partial\Omega} = 0, \tag{1.1}$$

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

Il est important de faire des hypothèses supplémentaires sur la *régularité géométrique* du domaine Ω.

Définition 1.1.1 *Le domaine Ω est dit lipschitzien si il existe une 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*

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

Intuitivement cela signifie que la frontière de Ω peut-être vue localement comme le graphe d'une fonction lipschitzienne. La pluspart 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 $\mathcal{C}^{m,1}$ lorsque les fonctions ψ_i sont $\mathcal{C}^m$ et leurs dérivées partielles d'ordre m sont lipschitziennes. Tous les domaines considérés dans ce cours seront au minimum de type lipschitzien.

Afin de donner un sens à l'équation (1.1) il faut préciser *l'espace* dans lequel on cherche la solution. Un premier choix intuitivement possible est de chercher u dans $\mathcal{C}^2$ en supposant alors f continue. En multipliant l'équation par une fonction v arbitraire de classe $\mathcal{C}^1$, et en intégrant sur le domaine, on obtient

$$-\int_\Omega \Delta u \, v = \int_\Omega f v,$$

et en appliquant la formule de Green

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

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

$$\int_\Omega \nabla u \cdot \nabla v = \int_\Omega f v.$$

L'équation ci-dessus garde un sens lorsque u est seulement de classe $\mathcal{C}^1$. En introduisant l'espace $X = \{v \in \mathcal{C}^1 \; ; \; v_{|\partial\Omega} = 0\}$, on obtient ainsi que toute solution de (1.1) est aussi solution de la formulation variationnelle

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

avec

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

(*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 1.1.1 *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 (1.1), on obtient $\nabla u = 0$, donc u est constante et forcément nulle puisque $u_{\partial\Omega} = 0$.*

Remarque 1.1.2 *Il est important de vérifier que réciproquement, toute solution suffisament régulière de (1.1) est aussi solution de la formulation initiale (1.1). En supposant u de classe $\mathcal{C}^2$ et solution de (1.1), on obtient en appliquant la formule de Green dans le sens inverse que*

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

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

Remarque 1.1.3 *Il est assez naturel que l'espace X où l'on cherche la solution soit le même que celui que parcours les fonctions v. Remarquons que si X est un espace de dimension finie, toute équation de type (1.1) 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 (1.1) va découler de la théorie de Lax-Milgram.*

1.2 Théorie de Lax-Milgram

Dans cette section, on considère un problème général pouvant se mettre sous la forme variationnelle

(V) *Trouver $u \in X$ tel que $a(u, v) = L(v)$ pour tout $v \in X$,*

Le théorème de Lax-Milgram apporte une réponse à l'existence, l'unicité et la stabilité de la solution dans un cadre précis.

Théorème 1.2.1 *On suppose que X est un espace de Hilbert et que les formes a et L vérifient les hypothèses suivantes :*

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

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

$$\|u\|_X \le \frac{\|L\|_{X'}}{\alpha},$$

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

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

$$\alpha\|u\|_X^2 \le C_L\|u\|_X.$$

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

$$\alpha\|v\|_X \le \|Av\|_X,$$

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

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

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

Remarque 1.2.1 *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 1.2.2 *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 fonctionelle*

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

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

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

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 (1.1) et à sa formulation variationnelle (1.1) dans l'espace $X = \{v \in \mathcal{C}^1 \; ; \; v_{|\partial\Omega} = 0\}$, on voit que

$$\|v\|_X := \|\nabla v\|_{L^2},$$

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

$$\|v\|_{L^2} \leq C_P \|\nabla v\|_{L^2}. \tag{1.2}$$

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

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

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 sur les espaces de Sobolev.

1.3 Rappels sur les espaces de Sobolev

Pour m entier positif, $1 \leq p \leq \infty$ et Ω un domaine ouvert de $\mathbb{R}^d$ on définit

$$W^{m,p}(\Omega) := \{v \in L^p(\Omega) \; ; \; D^\alpha v \in L^p(\Omega), \; |\alpha| \leq m\},$$

les dérivées $D^\alpha v = \frac{\partial^{|\alpha|} v}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}$ étant prises au sens des distributions avec la notation usuelle $|\alpha| = \alpha_1 + \cdots + \alpha_d$. Autrement dit $v \in W^{m,p}(\Omega)$ si et seulement si pour tout $|\alpha| \leq m$, il existe $v_\alpha \in L^p(\Omega)$ telle que

$$\int_\Omega v_\alpha \varphi = (-1)^{|\alpha|} \int v D^\alpha \varphi,$$

pour tout $\varphi \in \mathcal{D}(\Omega)$ où $\mathcal{D}(\Omega)$ est l'ensemble des fonctions $\mathcal{C}^\infty$ à support compact dans Ω, et on note alors $v_\alpha = D^\alpha v$. Lorsque $1 < p < \infty$, une autre façon équivalente d'exprimer ceci est d'affirmer que pour tout $|\alpha| \leq m$, il existe une constante C_α telle que

$$|\int v D^\alpha \varphi| \leq C_\alpha \|\varphi\|_{L^{p'}},$$

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

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

et on définit aussi la semi-norme par

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

de sorte que $\|v\|_{W^{m,p}} = (\|v\|_{W^{m-1,p}}^p + \|v\|_{W^{m,p}}^p)^{1/p}$. Dans le cas $p = \infty$ ces définitions sont modifiées en remplacant les norme ℓ^p par des max. En partant de ces définitions et en utilisant le fait que $L^p(\Omega)$ est complet, on démontre facilement le résultat suivant.

Théorème 1.3.1 *$W^{m,p}(\Omega)$ 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

$$\langle v, w \rangle_{H^m} := \sum_{|\alpha| \leq m} \langle D^\alpha u, D^\alpha v \rangle_{L^2},$$

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

La définition de $W^{m,p}(\Omega)$ montre facilement que si $v \in W^{m,p}(\mathbb{R}^d)$ 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 $W^{m,p}(\mathbb{R}^d)$. Nous admettrons le résultat difficile suivant du à Stein.

Théorème 1.3.2 *Si Ω est un domaine lipschitzien, il existe un opérateur linéaire d'extension E borné de $W^{m,p}(\Omega)$ dans $W^{m,p}(\mathbb{R}^d)$, c'est à dire tel que pour tout $v \in W^{m,p}(\Omega)$, la restriction de Ev à Ω est égale à v et $\|Ev\|_{W^{m,p}(\mathbb{R}^d)} \leq C\|v\|_{W^{m,p}(\Omega)}$.*

A titre d'exercice on pourra essayer de montrer que ce théorème est faux dans un domaine non-lipschitzien présentant une fissure.

Rappelons quelques résultats importants sur la densité des fonctions régulières dans les espaces de Sobolev : si Ω est un ouvert 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 $\mathcal{C}^\infty$. Ce théorème peut se démontrer d'abord sur $\mathbb{R}^d$ par des opérations de troncature et de régularisation, puis sur Ω en utilisant le théorème de prolongement 1.3.2. 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 $L^p(\Omega)$. D'autre part, tous ces résultats de densité sont faux dans le cas $p = \infty$.

Rappelons à présent les 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

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

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 1.1.1. 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

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

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 (1.2) pour les fonctions de $H^1_0(\Omega)$. Cette inégalité nous indique que la semi-norme $\|\nabla v\|_{L^2}$ est une norme équivalente à la norme H^1 sur $H^1_0(\Omega)$. On la désigne parfois comme "norme H^1_0" et on note donc

$$\|v\|_{H^1_0} = |v|_{H^1} = \|\nabla v\|_{L^2}.$$

On désigne par $H^{-1}(\Omega)$ l'espace dual de $H^1_0(\Omega)$ muni de la norme

$$\|v\|_{H^{-1}} := \sup_{\|w\|_{H^1_0}=1} \langle v, w\rangle_{H^{-1},H^1_0}.$$

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

$$\|v\|_{H^{-1}} := \sup_{\|w\|_{H^1_0}=1} \int_\Omega vw.$$

On a ainsi la chaine d'inclusion

$$H^1_0(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega).$$

Remarque 1.3.1 *Quand $\Omega = \mathbb{R}^{\mathrm{d}}$ la densité de $\mathcal{D}(\Omega)$ montre que $H^1_0(\mathbb{R}^{\mathrm{d}}) = \mathrm{H}^1(\mathbb{R}^{\mathrm{d}})$ et $(H^1(\mathbb{R}^{\mathrm{d}}))' = \mathrm{H}^{-1}(\mathbb{R}^{\mathrm{d}})$. 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'annulle sur toutes les fonctions de $\mathcal{D}(\Omega)$).*

Les espaces de Sobolev peuvent aussi être définis pour des indices de régularité non-entier.

Définition 1.3.1 *Pour Ω un domaine de $\mathbb{R}^{\mathrm{d}}$, $m < s < m+1$ et $1 \le p < \infty$, $W^{s,p}(\Omega)$ est l'ensemble des fonctions v de $W^{m,p}(\Omega)$ telles que*

$$|v|^p_{W^{s,p}} := \sum_{|\alpha|=m} \int\int_{\Omega\times\Omega} \frac{|D^\alpha v(x) - D^\alpha v(y)|^p}{|x-y|^{(s-m)p+d}} dx\, dy < +\infty, \tag{1.3}$$

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

Dans le cas $p = 2$ et $\Omega = \mathbb{R}^{\mathrm{d}}$, on peut aussi définir $H^s(\mathbb{R}^{\mathrm{d}})$ de manière équivalente en utilisant la transformée de Fourier : $v \in H^s(\mathbb{R}^{\mathrm{d}})$ si et seulement si $(1 + |\omega|^s)\hat{v} \in L^2(\mathbb{R}^{\mathrm{d}})$. Dans le cas $p = \infty$ on remplace (1.3) par la condition de Hölder $|D^\alpha v(x) - D^\alpha v(y)| \le C|x-y|^{s-m}$ pour $|\alpha| = m$ et on note $W^{s,\infty} = \mathcal{C}^s$.

Ces espaces ont propriétés similaires à celles énoncées pour les espaces d'indices entiers : complétude par rapport à leur norme, opérateur d'extension, densité des fonctions régulières. L'espace $H^{1/2}$ joue un rôle important dans l'étude de l'opérateur de trace, comme l'indique le résultat admis suivant.

Théorème 1.3.3 *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*

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

qui est équivalente à la norme $H^{1/2}$ obtenue par la définition 1.3.1 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 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 1.3.4 *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*

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

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

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

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

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

En utilisant l'inégalité de Schwarz et le théorème 1.3.5, on obtient donc

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

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 $w \in H_0^1(\Omega)$. $\diamond$

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

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).

1.4 Exemples de formulations variationnelles

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

Exemple 1. Equation du laplacien
On a déjà vu en section 1.1 la mise sous forme variationnelle de l'équation du laplacien (1.1). Dans ce cas on travaille avec l'espace de Sobolev $X = H_0^1(\Omega)$, et toutes les hypothèses du théorème de Lax-Milgram sont vérifiées. Le laplacien définit ainsi un isomorphisme de $H_0^1(\Omega)$ dans $H^{-1}(\Omega)$ et on a l'estimation a-priori

$$\|u\|_{H_0^1} \leq \|f\|_{H^{-1}}.$$

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

$$\|u\|_{H_0^1} \leq C_P \|f\|_{L^2},$$

où C_P est la constante de l'inégalité de Poincaré qui s'étend aux fonctions de $H_0^1(\Omega)$. 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 1.4.1 *Soit Ω un domaine $\mathcal{C}^{1,1}$ ou un domaine convexe. Alors si $f \in L^2(\Omega)$, la solution u de la formulation variationnelle de (1.1) est dans H^2 et il existe une constante C qui ne dépend que de Ω telle que*

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

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

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

On verra par la suite que ce résultat n'est pas valide pour des domaines moins réguliers.

Exemple 2. Laplacien avec conditions non-homogènes
Considérons à présent le problème de Laplace avec une donnée de Dirichlet non-homogène au bord.

$$-\Delta u = f \text{ dans } \Omega \text{ et } \gamma_0 u = g, \tag{1.4}$$

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

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

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

$$\int_\Omega \nabla\tilde{u}\cdot\nabla v = \int_\Omega fv - \int_\Omega \nabla R(g)\cdot\nabla v \text{ pour tout } v \in H_0^1(\Omega). \tag{1.5}$$

On peut vérifier que le choix du relèvement ne change rien à la solution de (1.5). On voit que dans cette formulation variationnelle on peut supposer $g \in H^{1/2}(\partial\Omega)$ et utiliser le théorème 1.3.5, ce qui montre que la forme lineaire du membre de droite est bornée suivant

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

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

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

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

$$\sigma u - \mathrm{div}(A.\nabla u) = f \text{ dans } \Omega \text{ et } \gamma_0 u = 0, \tag{1.6}$$

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

$$0 < \alpha \leq \lambda_{\min}(A(x)) \leq \lambda_{\max}(A(x)) \leq C, \tag{1.7}$$

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

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

et les hypothèses de Lax-Milgram dans $H_0^1(\Omega)$ découlent aisément de la propriété (1.7).

Exemple 5. Equation de convexion diffusion
On considère l'équation

$$-\Delta u + b\cdot\nabla u = f \text{ dans } \Omega \text{ et } \gamma_0 u = 0, \tag{1.8}$$

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

$$\operatorname{div}(b(x)) = 0, \quad x \in \Omega.$$

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

$$\int_\Omega \nabla u \cdot \nabla v + \int_\Omega (b \cdot \nabla u)v = \int_\Omega fv \text{ pour tout } v \in H_0^1(\Omega). \tag{1.9}$$

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 $\operatorname{div}(buv) = (b \cdot \nabla u)v + (b \cdot \nabla v)u$, ce qui entraine 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 6. Problème de Neumann
Le problème

$$-\Delta u + u = f \text{ dans } \Omega \text{ et } \gamma_1 u = g, \tag{1.10}$$

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

$$\int_\Omega \nabla u \cdot \nabla v + \int_\Omega uv = \int_\Omega fv + \int_{\partial\Omega} g\gamma_0 v \text{ pour tout } v \in H^1(\Omega). \tag{1.11}$$

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 (1.11), on élimine le terme $\int_{\partial\Omega} gv$ et on retrouve l'équation (1.10), puis on obtient que pour tout $v \in H^1(\Omega)$ on a

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

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. Au vu du théorème 1.3.4, on prend $f \in L^2$ et $f \in H^{-1/2}(\Omega)$, ce qui entraine la continuité sur $H^1(\Omega)$ de la forme linéaire dans (1.11). La continuité et la coercivité de la forme bilinéaire est immédiate, et les hypothèses de Lax-Milgram sont donc vérifiées.

Exemple 7. Laplacien avec conditions de Neumann
Toujours avec $f \in L^2(\Omega)$ et $g \in H^{-1/2}(\Omega)$ on considère le problème

$$-\Delta u = f \text{ dans } \Omega \text{ et } \gamma_1 u = g, \tag{1.12}$$

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

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

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

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

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

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

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

$$\int_\Omega \nabla u \cdot \nabla v = \int_\Omega fv + \int_{\partial\Omega} g\gamma_0 v \text{ pour tout } v \in H^1(\Omega)/\mathbb{R}. \tag{1.13}$$

En utilisant la relation entre f et g, on voit que l'égalité dans (1.13) 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 (1.13) est aussi solution de (1.12).

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.

Théorème 1.4.2 *Soit Ω borné, lipschitzien et connexe. Il existe une constante C_Q qui ne dépend que du domaine Ω telle que pour tout $v \in H^1(\Omega)/\mathbb{R}$*

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

Preuve : On procède par contradiction. Si le résultat est faux, il existe une suite $(v_n)_{n>0}$ telle que $v_n \in H^1(\Omega)/\mathbb{R}$, $\|v_n\|_{H^1} = 1$ et $\|\nabla v_n\|_{L^2}$ tends vers 0. Par le théorème de Rellich on peut extraire une sous suite w_n qui converge vers w dans $L^2(\Omega)$. Puisque $\|\nabla w_n\|_{L^2}$ tends vers 0 on en déduit que c'est une suite de Cauchy dans $H^1(\Omega)$ et qui converge donc vers w dans $H^1(\Omega)$. Par passage à la limite on voit que $\|\nabla w\|_{L^2} = 0$, $w \in H^1(\Omega)/\mathbb{R}$ et $\|w\|_{H^1} = 1$. La première propriété montre que w est constante puisque Ω est connexe. D'après la deuxième propriété on a donc $w = 0$ ce qui est en contradiction avec la troisième propriété. ◇

Remarque 1.4.1 *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.*

Remarque 1.4.2 *On peut aussi utiliser ce mode de raisonnement pour démontrer l'inégalité de Poincaré sur $H_0^1(\Omega)$ (exercice).*

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.

Exemple 8. Bi-Laplacien
La formulation variationnelle du problème

$$\Delta^2 u = f, \quad \gamma_0 u = \gamma_1 u = 0, \tag{1.14}$$

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

$$H_0^2(\Omega) := \{v \in H^2(\Omega)\ ;\ \gamma_0 v = \gamma_1 v = 0\},$$

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 $u \in H_0^2(\Omega)$ telle que

$$\int_\Omega \Delta u \Delta v = \int_\Omega fv \ \text{ pour tout } \ v \in H_0^2(\Omega). \tag{1.15}$$

On vérifie aisément que toute solution de (1.15) suffisament régulière est aussi solution de (1.14).

Sur $H_0^2(\Omega)$ on peut travailler avec la norme

$$\|u\|_{H_0^2(\Omega)} := \|\Delta u\|_{L^2},$$

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 $v \in H_0^2(\Omega)$,

$$\|v\|_{L^2} \le C_B \|\Delta v\|_{L^2}.$$

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 $\mathcal{C}^{1,1}$. 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 $H_0^2(\Omega)$ que l'on note $H^{-2}(\Omega)$, et pour $f \in L^2$ on a donc

$$\|f\|_{H^{-2}} = \sup_{\|v\|_{H_0^2}=1} \int_\Omega fv \le C_B \|f\|_{L^2}.$$

1.5 Approximation variationnelle

Considérons un problème 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* $X_h \subset X$. Pour cela on définit un problème approché en substituant X_h à X dans (V) :

(G) *Trouver* $u_h \in X_h$ *tel que* $a(u_h, v_h) = L(v_h)$ *pour tout* $v_h \in X_h$.

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

$$u_h = \sum_{k=1}^{N_h} u_{h,k} \varphi_k,$$

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

$$a(u_h, \varphi_l) = \sum_{k=1}^{N_h} u_{k,h} a(\varphi_k, \varphi_l) = L(\varphi_l), \quad l = 1, \cdots, N_h.$$

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

$$A_h U_h = F_h \tag{1.16}$$

où $U_h := (u_{h,k})_{k=1,\cdots,N_h}$, $F_h := (L(\varphi_k))_{k=1,\cdots,N_h}$ et $(A_h)_{k,l} = a(\varphi_k, \varphi_l)$ pour $k, l = 1, \cdots, N_h$. La matrice A_h 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_h, au sens où l'erreur $\|u - u_h\|_X$ est comparable au minimum de $\|u - v_h\|_X$ quand v_h parcourt X_h. Ceci est exprimé par le lemme de Cea.

Lemme 1.5.1 *On suppose que les hypothèses du théorème de Lax-Milgram sont vérifiées. Alors, il existe une solution unique u_h au problème approché (G) et la matrice A_h est inversible. On a l'estimation d'erreur*

$$\|u - u_h\|_X \le \frac{C_a}{\alpha} \min_{v_h \in X_h} \|u - v_h\|_X.$$

Dans le cas où a est symétrique on a

$$\|u - u_h\|_X \le \left(\frac{C_a}{\alpha}\right)^{1/2} \min_{v_h \in X_h} \|u - v_h\|_X.$$

Preuve : Comme X_h 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_h (ainsi que l'estimation a-priori $\|u_h\|_X \leq \frac{\|L\|_{X'}}{\alpha}$). Comme (φ_k) est une base, l'unicité montre que $U_h = 0$ si $F_h = 0$ ce qui signifie que A_h est inversible.

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

$$a(u - u_h, v_h) = 0. \tag{1.17}$$

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

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

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

Dans le cas où a est symétrique, on remarque que (1.17) signifie que u_h est la projection orthogonale de u sur X_h au sens du produit scalaire $a(\cdot, \cdot)$, et donc

$$\|u - u_h\|_a = \min_{v_h \in X_h} \|u - v_h\|_a.$$

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

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

qui découle de la continuité et de la coercivité de a. $\diamond$

Notre but est à présent de définir des espaces X_h pertinents pour appliquer la méthodes de Galerkin. En premier lieu, il s'agit d'obtenir un compromis raisonable entre la complexité du système décrit par la dimension N_h de X_h et la précision dans l'approximation de la solution. D'autre part, il est important que les espaces X_h 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 (1.16). Enfin, il est important que ce système puisse être résolu aisément, en particulier lorsque N_h devient grand. Il peut ainsi être intéressant que X_h possède une base qui rend la matrice de rigidité *creuse*, c'est à dire contenant peu d'éléments non-nuls, ou encore *bien conditionnée*. La *méthode des éléments finis* répond à l'ensemble de ces exigences.

2 La méthode des éléments finis

2.1 Présentation de la méthode

Dans la méthode des éléments finis, le domaine Ω est décomposé en une partition ou *maillage* $\mathcal{T}_h$ et l'espace d'approximation X_h est constitués de fonctions polynomiales par morceaux sur chaque élément de la partition $T \in \mathcal{T}_h$. Le paramètre h représente la finesse de la discrétisation au sens où

$$h = \max_{K \in \mathcal{T}_h} \mathrm{diam}(K).$$

Avant d'aborder cette méthode plus en détail, il convient de la situer par rapport à d'autres approches. Il existe deux autres classes importantes de méthodes numériques pour les EDP qui ne rentrent pas dans le cadre de l'approximation variationnelle : la méthode des *différences finies* et celle des *volumes finis*.

Dans la méthode des différences finies, on cherche à calculer une solution sur une grille de points, au moyen d'une formule de discrétisation pour l'opérateur au dérivée partielles. Prenons ici l'exemple du laplacien (1.1) sur un domaine bidimensionel (d=2) et avec une grille de discrétisation régulière du type $x_{k,l} = (hk, hl)$ où (k,l) sont les couples d'entiers tels que $x_{k,l} \in \Omega$. En notant $u_{k,l}$ l'inconnue qui doit approcher $u(x_{k,l})$, on peut par exemple discrétiser le laplacien par la formule à 5 points :

$$\frac{4u_{k,l} - u_{k,l+1} - u_{k,l-1} - u_{k+1,l} - u_{k-1,l}}{h^2} = f(x_{k,l}).$$

Dans la méthode des volumes finis, on se donne un maillage $\mathcal{T}_h$ et des inconnues sont associées à chaque maille $T \in \mathcal{T}_h$. Cette méthode se prête bien à la discrétisation des équations sous forme divergence qui peuvent être re-interprêtées par une formule de flux. Dans le cas du laplacien (1.1), on peut en effet écrire pour tout $T \in \mathcal{T}_h$,

$$-\int_{\partial T} \frac{\partial u}{\partial n} = \int_T f$$

L'idée est alors d'approcher le flux du membre de gauche au moyen d'une formule faisant intervenir les inconnues sur T et les mailles voisines. Dans le cas d'un domaine bidimensionel et d'un maillage carré

$$T_{k,l} = [(k - 1/2)h, (k + 1/2)h] \times [(l - 1/2)h, (l + 1/2)h],$$

si l'inconnue $u_{k,l}$ approche u sur $T_{k,l}$ on peut discrétiser le flux intégré sur l'interface $T_{k,l} \cap T_{k+1,l}$ par $u_{k+1,kl} - u_{k,l}$. En procédant de même pour les trois autres flux et en approchant $\int_{T_{k,l}} f$ par $h^2 f(x_{k,l})$, on obtient ainsi l'équation

$$4u_{k,l} - u_{k,l+1} - u_{k,l-1} - u_{k+1,l} - u_{k-1,l} = h^2 f(x_{k,l}),$$

et on voit ainsi dans ce cas que le système discrèt obtenu est similaire à celui des différences finies.

Il est d'autre part assez fréquent que les méthodes de différences finies et de volumes finis aboutissent aux mêmes équations discrètes que les méthodes d'éléments finis. Ces dernières présentent néammoins plusieurs avantages qui justifient leur succès :

- La prise en compte des conditions aux limites est particulièrement simple puisqu'elle est intégrée dans l'espace X que l'on choisit dans la formulation variationnelle ainsi que par son approximation interne X_h.
- L'utilisation de maillages non-rectangulaires et non-uniformes rend souvent difficile la discrétisation des opérateurs différentiels par la méthode des différences finies.
- L'analyse d'erreur de la méthodes des éléments finis bénéficie du cadre général des approximations variationnelles présenté en §1.5.

Il existe par ailleurs des méthodes variationnelles utilisant d'autres espaces d'approximation que les espaces d'éléments finis, en particulier des espaces de polynômes ou de polynômes trigonométriques (séries de Fourier partielles) définis globalement sur le domaine Ω. Il s'agit des *méthodes spectrales* qui sont particulièrement performantes lorsque la solution que l'on souhaite approcher est très régulière, mais n'offrent pas les possibilités d'adaptation locale du maillage qui sont utiles pour des solutions présentant des singularités locales et que nous étudierons dans le chapitre 3. Leur mise en oeuvre pose en outre des difficultés pour la prise en compte dans le cas où le domaine Ω ne possède pas une géométrie simple. Un dernier type de méthodes variationnelles, plus proche de la méthode des éléments finis, utilise les *bases d'ondelettes* et sera évoqué dans le chapitre 3.

Décrivons rapidement la méthode des éléments finis de Lagrange dans le cas unidimensionel $d = 1$. Si $\Omega =]a, b[$ est le domaine de calcul, on le divise en intervalle $T_i = [x_i, x_{i+1}]$, avec $a = x_0 < x_1 < \cdots < x_{N-1} < x_N = b$ et on note $h = \max |x_{i+1} - x_i|$. Pour un entier $m > 0$ fixé, on définit

$$X_h = \{v \in \mathcal{C}^0([a, b]) \; ; \; v_{|T_i} \in \Pi_m, \; i = 0, \cdots, N-1\}$$

l'espace des fonctions globalement continues sur $[a, b]$ dont la restriction sur chaque T_i est un polynôme de degre m. Cet espace est contenu dans $H^1([a, b])$. Sur chaque intervalle T_i, un polynôme de degré m peut -être caractérisé par ses valeurs en m points distincts. Pour chaque T_i on introduit ainsi les noeuds

$$\gamma_{i,k} = x_i + \frac{k}{m}(x_{i+1} - x_i), \;\; k = 0, \cdots, m.$$

et l'on voit que toute fonctions de X_h est uniquement déterminée par ses valeurs aux points du maillage

$$\Gamma_h = \{\gamma_{i,k}, \; i = 0, \cdots N-1, \; k = 0, \cdots, m\}.$$

Le fait d'avoir inclu les extrémités x_i et x_{i+1} dans le choix des noeuds pour chaque intervalle est important : il assure que réciproquement n'importe quel ensemble de valeurs prises sur Γ_h peut être associé à une fonction de X_h puisque la continuité est assurée entre chaque intervalle. L'ensemble des valeurs prises sur Γ_h est pour cette raison appelé ensemble des *degrés de liberté* de X_h. En d'autre termes l'application qui à v associe ses valeurs sur Γ_h est un isomorphisme entre X_h et $\mathbb{R}^{\Gamma_h}$. Ceci nous permet d'introduire la *base nodale* $(\varphi_\gamma)_{\gamma \in \Gamma_h}$ de X_h définie par

$$\varphi_\gamma(\mu) = 1 \;\text{ si } \mu = \gamma, \; 0 \text{ si } \mu \in \Gamma_h - \{\gamma\}.$$

Dans le cas $m = 1$, on retrouve les classiques fonctions "chapeaux" affines par morceaux associés aux noeuds x_i. Dans le cas général on vérifie aisément que ces fonctions sont supportés dans T_i si γ est un noeud intérieur à T_i et dans $T_{i-1} \cup T_i$ si $\gamma = x_i$. On voit aussi que les coordonées de $v_h \in X_h$ dans la base nodale sont données par ses valeurs aux noeuds correspondants

$$v_h = \sum_{\gamma \in \Gamma_h} v_h(\gamma)\varphi_\gamma.$$

Pour une fonction $v \in \mathcal{C}^0([a, b])$, on peut définir son interpolation sur X_h par

$$I_h v := \sum_{\gamma \in \Gamma_h} v(\gamma)\varphi_\gamma.$$

On voit ainsi que I_h est un opérateur de $\mathcal{C}^0$ sur X_h qui laisse X_h invariant.

Afin de généraliser ces notions au cadre multidimensionel, nous suivrons d'abord une approche locale en nous concentrant sur un seul élément, avant d'en déduire la construction des espaces d'éléments finis et leurs propriétés d'approximation.

2.2 L'élément fini triangulaire de Lagrange

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 2.2.1 *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 2.2.1 *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 $(y_1, \cdots, y_M) \in \mathbb{R}^{\mathrm{M}}$, il existe $v \in X_T$ tel que $\psi_i(v) = v_i$.

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

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

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 tétraedre 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'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-planaire). Cette condition entraine que le volume $|T|$ de T est strictement positif.

Remarque 2.2.2 *Nous nous placerons 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.*

Définissons trois paramètres géométriques importants :
- Le diamètre h_T : longueur du plus grand côté de T.
- La rondeur ρ_T : diamètre de la boule inscrite dans T.
- L'excentricité $\sigma_T = \frac{h_T}{\rho_T}$ qui mesure la non-dégénérescence de T.

L'espace choisi X_T est celui des polynôme de degré total inférieur à k :

$$X_T = \Pi_k = \mathrm{Vect}\{x_1^{\alpha_1} \cdots x_d^{\alpha_d} \; ; \; 0 \leq |\alpha| \leq k\},$$

dont la dimension est donnée (exercice) par la formule

$$\dim(\Pi_k) = \frac{1}{d!}(k+1)(k+2)\cdots(k+d).$$

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

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

L'existence et l'unicité se vérifie en développant la première équation sur chaque coordonées 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,

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

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

$$\pi(x) = \sum_{i=0}^{d} \pi(a_i)\lambda_i(x). \tag{2.18}$$

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

$$\Sigma_k := \{x \in \mathrm{I\!R}^{\mathrm{d}} \; ; \; \lambda_{\mathrm{i}}(\mathrm{x}) \in \{0, \frac{1}{\mathrm{k}}, \frac{2}{\mathrm{k}}, \cdots, 1\}\}.$$

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

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

Théorème 2.2.1 *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

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

où les α_i sont des entiers positifs. Il suffit donc de prouver que si $\pi \in \Pi_k$ 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 (2.18). 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

$$\Sigma'_k = \{x \in \Sigma_k \; ; \; \lambda_0(x) = 0\},$$

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

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

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

$$\Sigma''_k = \Sigma_k - \Sigma'_k.$$

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 treilli 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 coordonées 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.

2.3 L'élément de référence

On définit le *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 coordonées 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

$$A_T(x) = a_0 + B_T x,$$

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

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

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

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

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

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

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

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

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

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

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 $\widehat{\Sigma}$.

On voit finalement que l'espace des polynômes de degré k est invariant par la transformation A_T, et on peut ainsi écrire

$$X_{\hat{T}} = \Pi_k = \widehat{X}_T.$$

Ceci entraine que les interpolants $I_{\widehat{T}}$ et I_T vérifient l'identité

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

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

$$I_T v \circ A_T = I_{\hat{T}}(v \circ A_T).$$

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 2.3.1 *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, nous allons à présent comparer les quantités $|v|_{W^{m,p}(T)}$ et $|\hat{v}|_{W^{m,p}(\widehat{T})}$. Lorsque $m = 0$, on a simplement

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

ce qui donne
$$\|v\|_{L^p(T)} = C|T|^{1/p}\|\hat{v}\|_{L^p(\widehat{T})},$$
avec $C = (d!)^{1/p}$. 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 2.3.1 *Pour tout simplexe T, on a*
$$\|B_T\| \leq C_1 h_T,$$
avec $C_1 = \rho_{\widehat{T}}^{-1}$ qui dépend seulement de d et
$$\|B_T^{-1}\| \leq C_2 \rho_T^{-1},$$
avec $C_2 = h_{\widehat{T}} = \sqrt{2}$.

Preuve : Par définition, on a
$$\|B_T\| = \sup_{x \in \mathbb{R}^d} \frac{\|B_T x\|}{\|x\|} = \sup_{\|x\|=\rho_{\widehat{T}}} \frac{\|B_T x\|}{\rho_{\widehat{T}}},$$
et le sup est atteint. Soit x un vecteur qui réalise ce sup. Il existe donc deux points $\hat{a}$ et $\hat{b}$ dans la boule inscrite de $\widehat{T}$ tels que $x = \hat{a} - \hat{b}$. On a donc
$$B_T x = A_T(\hat{a}) - A_T(\hat{b}) = a - b,$$
où a et b sont dans T et par conséquent $\|B_T x\| \leq h_T$. L'autre inégalité se démontre de la même manière en échangeant les rôles de T et $\widehat{T}$. $\diamond$

Examinons à présent les semi-normes Sobolev dans le cas $m = 1$. On remarque que si $v(x) = \hat{v}(\hat{x})$ c'est à dire $\hat{v} = v \circ A_T$ on a
$$\nabla \hat{v}(\hat{x}) = B_T^t \nabla v(x),$$
où B_T^t est la transposée de B_T. Ceci nous montre que
$$\begin{aligned} |v|_{H^1(T)}^2 &= \int_T \|\nabla v(x)\|^2 dx \\ &= \int_{\widehat{T}} \|(B_T^t)^{-1} \nabla \hat{v}(\hat{x})\|^2 |\det(B_T)| d\hat{x} \\ &\leq C_2 d! \tfrac{|T|}{\rho_T^2} \int_{\widehat{T}} \|\nabla \hat{v}(\hat{x})\|^2 d\hat{x}, \end{aligned}$$
et par conséquent
$$|v|_{H^1(T)} \leq C \frac{|T|^{1/2}}{\rho_T} |\hat{v}|_{H^1(\widehat{T})},$$
où C ne dépend que de d. En inversant le rôle de T et $\hat{T}$ dans le calcul, on obtient aussi
$$|\hat{v}|_{H^1(\widehat{T})} \leq C \frac{h_T}{|T|^{1/2}} |v|_{H^1(T)}.$$
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 de norme dans $\mathbb{R}^d$
$$\|z\|_{\ell^p} \leq d^{1/p-1/2}\|z\| \leq \|z\|_{\ell^p},$$
pour $1 \leq p \leq 2$ et
$$\|z\|_{\ell^p} \leq \|z\| \leq d^{1/2-1/p}\|z\|_{\ell^p},$$
pour $p \geq 2$. On aboutit ainsi à
$$|v|_{W^{1,p}(T)} \leq C \frac{|T|^{1/p}}{\rho_T} |\hat{v}|_{W^{1,p}(\widehat{T})},$$

et

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

où C ne dépend que de d. Pour les dérivées d'ordre supérieur, on montre (exercice) que

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

où C ne dépend que de d et de m. Par un calcul de changement de variable similaire au cas $m=1$ on obtient ainsi

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

et

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

où C ne dépend que de d et de m.

2.4 Approximation polynomiale locale

On part du résultat fondamental suivant.

Théorème 2.4.1 *Si Ω est un domaine borné lipschitzien connexe de $\mathbb{R}^{\rm d}$, m un entier positif, alors pour tout $v \in W^{m,p}(\Omega)$ telle que $\int_\Omega v\pi = 0$ pour tout $\pi \in \Pi_{m-1}$, on a*

$$\|v\|_{W^{m,p}} \leq C|v|_{W^{m,p}}.$$

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

Preuve : c'est une généralisation immédiate de la preuve du théorème 1.4.2, en utilisant le même argument de contradiction. $\diamond$

Une conséquence importante de ce résultat est le *théorème de Deny-Lions*.

Théorème 2.4.2 *Si Ω est un domaine borné lipschitzien connexe de $\mathbb{R}^{\rm d}$, m un entier positif, alors pour tout $v \in W^{m,p}(\Omega)$ on a*

$$\min_{\pi\in\Pi_{m-1}} \|v-\pi\|_{W^{m,p}} \leq C|v|_{W^{m,p}}.$$

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

Preuve : Il suffit d'appliquer le théorème précèdent à la fonction $v - P_{m-1}v$ où P_{m-1} est la projection orthogonale pour $L^2(\Omega)$ sur Π_{m-1}. $\diamond$

On remarque que puisque $\Pi_{m-1} \subset \Pi_k$ pour $k \geq m-1$, le théorème de Deny-Lions donne aussi

$$\min_{\pi\in\Pi_k} \|v-\pi\|_{W^{m,p}} \leq C|v|_{W^{m,p}},$$

sous l'hypothèse $m \leq k+1$. En combinant le théorème de Deny-Lions appliqué sur l'élément de référence $\hat{T}$ et les résultats de la section précèdente, nous pouvons obtenir des résultats d'approximation locale par les polynômes sur un élément quelconque T. Pour $n \leq m \leq k+1$, on peut en effet écrire

$$\begin{aligned}
\min_{\pi\in\Pi_k} |v-\pi|_{W^{n,p}(T)} &\leq C\tfrac{|T|^{1/p}}{\rho_T^n}\min_{\hat{q}\in\Pi_k}|\hat{v}-\hat{\pi}|_{W^{n,p}(\widehat{T})}\\
&\leq C\tfrac{|T|^{1/p}}{\rho_T^n}|\hat{v}|_{W^{m,p}(\widehat{T})}\\
&\leq C\tfrac{|T|^{1/p}}{\rho_T^n}\tfrac{h_T^m}{|T|^{1/p}}|v|_{W^{m,p}(T)},
\end{aligned}$$

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

$$\min_{\pi \in \Pi_k} |v - \pi|_{W^{n,p}(T)} \leq C \frac{h_T^m}{\rho_T^n} |v|_{W^{m,p}(T)}. \tag{2.19}$$

La preuve du théorème de Deny-Lions nous indique que l'estimation ci-dessus peut être obtenue en remplacant $\min_{\pi \in \Pi_k} |v - \pi|_{W^{n,p}(T)}$ par $|v - P_T v|_{W^{n,p}(T)}$ où P_T est la projection orthogonale pour $L^2(T)$ sur Π_{m-1} (il faut pour cela remarquer que l'on a $P_{\widehat{T}} \hat{v}(\hat{x}) = P_T v(x)$).

Pour des raisons de raccord continu entre les différents triangles qui constitueront le maillage dans la méthode des éléments finis, on est amené à étudier l'erreur d'interpolation plutôt que l'erreur de projection orthogonale. Pour cela, nous aurons besoin du *lemme de Bramble-Hilbert* qui est une conséquence du théorème de Deny-Lions.

Lemme 2.4.1 *Soit Ω un domaine borné lipschitzien connexe de $\mathbb{R}^d$, m un entier positif, et R un opérateur borné de $W^{m,p}$ dans Π_k avec $m \leq k+1$, tel que $R\pi = \pi$ pour tout $\pi \in \Pi_k$. Alors pour tout $v \in W^{m,p}(\Omega)$ on a*

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

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

Preuve : Pour tout $\pi \in \Pi_k$, 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 Π_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 Lagrance (T, X_T, Σ_T) avec $X_T = \Pi_k$, il faut que l'espace $W^{m,p}$ s'injecte continuement dans l'espace des fonctions continues, 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 $n \leq m \leq k+1$,

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

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

$$|v - I_T v|_{W^{n,p}(T)} \leq C \frac{h_T^m}{\rho_T^n} |v|_{W^{m,p}(T)}. \tag{2.20}$$

Remarquons finalement qu'on peut généraliser ces résultats d'approximation polynomiale en faisant appel aux injections de Sobolev : si $n < m \leq k+1$ et $m - n > d/p - d/q$, on a sur tout domaine borné lipschitzien connexe Ω

$$\begin{aligned} \min_{\pi \in \Pi_k} \|v - \pi\|_{W^{n,q}} &\leq C \min_{\pi \in \Pi_k} \|v - \pi\|_{W^{m,p}} \\ &\leq C |v|_{W^{m,p}}, \end{aligned}$$

où la constante C varie d'une ligne à l'autre mais ne dépend que de Ω, m, n, p et q. En combinant cette estimation sur $\widehat{T}$ et les résultats de la section précèdente, on obtient ainsi

$$\min_{\pi \in \Pi_k} |v - \pi|_{W^{n,q}(T)} \leq C |T|^{1/q - 1/p} \frac{h_T^m}{\rho_T^n} |v|_{W^{m,p}(T)}. \tag{2.21}$$

On a de même pour l'interpolant

$$|v - I_T v|_{W^{n,q}(T)} \leq C |T|^{1/q - 1/p} \frac{h_T^m}{\rho_T^n} |v|_{W^{m,p}(T)}, \tag{2.22}$$

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

2.5 Approximation globale dans les espaces d'éléments finis

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 c'est à dire un ensemble fini de simplexes disjoints (aux frontières près) et dont l'union est égale à Ω. Le paramètre h désigne la finesse du maillage, c'est à dire

$$h = \max_{T\in\mathcal{T}_h} h_T.$$

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 T" entre deux arêtes.

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 principal d'ordre k pour chaque triangle

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

où de façon équivalente, l'ensemble des formes linéaires

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

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 élement fini polynomiale de degré k sur chaque T et caractérisée par par une suite de valeurs $v(\gamma)$ pour $\gamma \in \Gamma_h$, i.e. $\psi(v)$ pour $\psi \in \Sigma_h$.

Théorème 2.5.1 *Une définition équivalente de X_h est donnée par*

$$X_h := \{v \in \mathcal{C}(\Omega)\ ;\ v_{|T} \in \Pi_k,\ T \in \mathcal{T}_h\}.$$

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 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 $\mathcal{C}(\Omega)$ dans X_h suivant

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

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 $I_T v$:

$$I_h v(v) = I_T v(x), \quad x \in T.$$

Il est facile d'intégrer des conditions aux limites homogènes de Dirichlet dans la construction de l'espace X_h : on définit

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

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

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

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 $X_{0,h}$.

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,

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

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

Définition 2.5.1 *La famille $(\mathcal{T}_h)_{h>0}$ est dite "régulière" si et seulement si il existe une constante C_r telle que*

$$\sigma_T \leq C_r,$$

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

Remarque 2.5.1 *Dans le cas d'une triangulation en dimension $d = 2$, cette propriété équivaut au fait qu'il existe un $\theta_0 > 0$ tel que tous les angles des triangles vérifient $\theta \geq \theta_0$ (exercice).*

Remarque 2.5.2 *Une propriété équivalente est l'existence d'une constante $c > 0$ telle que*

$$ch_T^d \leq |T|,$$

où $|T|$ est le volume de T (exercice).

Définition 2.5.2 *La famille $(\mathcal{T}_h)_{h>0}$ est dite "quasi-uniforme" si et seulement il existe une constante $c > 0$ telle que*

$$ch \leq \rho_T \leq h_T \leq h$$

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

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 (2.20), sous la forme

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

où C dépend de (C_r, k, m, n, d, p). En combinant cette derniere estimation avec (2.5), on obtient un résultat d'approximation global sur X_h.

Théorème 2.5.2 *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 $v \in W^{m,p}(\Omega)$,*

$$|v - I_h v|_{W^{n,p}} \leq C h^{m-n} |v|_{W^{m,p}},$$

où C dépend de (C_r, k, m, n, d, p).

On a en particulier

$$\|v - I_h v\|_{L^2} \leq C h^m |v|_{H^m},$$

et

$$|v - I_h v|_{H^1} \leq C h^{m-1} |v|_{H^m}.$$

Un exemple immédiat d'application de ces résultats concerne l'approximation variationnelle du problème de Laplace (1.1). D'après le lemme de Cea, la solution discrète $u_h \in X_{h,0}$ vérifie

$$\|u - u_h\|_{H_0^1} \leq C \|u - I_h u\|_{H_0^1},$$

et donc si la solution est dans $H^m(\Omega)$ et $m \leq k+1$ et $m > d/2$

$$\|u - u_h\|_{H_0^1} \leq C h^{m-1} |v|_{H^m}. \tag{2.23}$$

Remarque 2.5.3 *La restriction $m > d/2$ 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^m pour tout $m \geq 0$. Un exemple d'un tel opérateur sera donné dans la section 3.2.*

Remarque 2.5.4 *Si Ω est convexe et $f \in L^2(\Omega)$, on peut utiliser le théorème de régularité 1.4.1 qui nous indique que la norme H^2 de u est controlée par la norme L^2 de f. On a ainsi*

$$\|u - u_h\|_{H_0^1} \leq C h \|f\|_{L^2}. \tag{2.24}$$

2.6 Estimation L^2 et inégalité inverse

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

$$\min_{v_h \in X_{h,0}} \|u - v_h\|_{L^2} \leq \|u - I_h u\|_{L^2} \leq C h^2 |u|_{H^2},$$

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 2.6.1 *Si Ω est un polygone ou polyèdre convexe, l'approximation u_h de la solution de* (1.1) *par la méthode de Galerkin dans l'espace $X_{h,0}$ des éléments finis $\mathbb{P}_{\mathrm{k}}$ vérifie l'estimation*

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

Preuve : On introduit la fonction w solution du problème auxiliaire

$$-\Delta w = u - u_h \text{ dans } \Omega, \ \ w_{|\partial\Omega} = 0,$$

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 $w_h \in X_h$. En prenant par exemple $w_h = I_h w$ on obtient par le théorème 2.5.2 $\|w - w_h\|_{H_0^1} \le Ch|w|_{h^2}$ et on en déduit en utilisant le théorème 1.4.1 $\|w - w_h\|_{H_0^1} \le Ch\|u - u_h\|_{L^2}$. Le résultat s'en déduit. $\diamond$

Un autre résultat important est *l'inégalité inverse* qui établit des liens entre les différentes normes d'une fonction dans l'espace des éléments finis X_h. Le point de départ en est la remarque que toutes les normes sur l'espace Π_k sont équivalentes puisque cet espace est de dimension finie. On obtient ainsi sur l'élément de référence $\hat{T}$ que pour tout $\hat{\pi} \in \Pi_k$ et pour tout $m \ge 0$, $p \ge 1$,

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

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 $T \in \mathcal{T}_h$

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

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

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

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

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

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

2.7 Résolution et conditionnement

La résolution numérique du système

$$A_h U_h = F_h$$

qui calcule la solution de Galerkin d'une formulation variationnelle (V) dans la base des éléments finis 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 $U_h^0 = 0$, on pose

$$U_h^n = U_h^{n-1} + \tau(F_h - A_h U_h^{n-1}) \tag{2.25}$$

avec $\tau > 0$. Dans le cas où A_h est définie positive, ce n'est rien d'autre qu'un algorithme de gradient à pas fixe pour la minimisation de

$$J(V_h) = \frac{1}{2}\langle A_h V_h, V_h\rangle - \langle F_h, V_h\rangle.$$

Comme U_h est un point fixe de l'itération (2.25), en définissant l'erreur d'itération $E_h = U_h - U_h^n$, on a

$$E_h^n = (I - \tau A_h)E_h^{n-1} = \cdots = (1 - \tau A_h)^n E_h^0 = (1 - \tau A_h)^n U_h.$$

On peut choisir diverse normes pour mesurer l'erreur E_h, mais un choix naturel est la norme d'énergie

$$\|V_h\|_A := \langle A_h V_h, V_h\rangle = a(v_h, v_h) = \|v_h\|_a,$$

où v_h est la fonction de X_h de vecteur de coordonées V_h dans la base nodale. Si u_h^n est la fonction de X_h de coordonées U_h^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_h est symmétrique définie positive. On a alors $\|\cdot\|$

$$\|E_h^n\|_A \leq \rho^n \|U_h\|_A$$

où le facteur de réduction ρ est donné par

$$\rho := \max\{1 - \tau\lambda_{\min}(A_h), \tau\lambda_{\max}(A_h) - 1\}.$$

Le choix de τ minimisant ρ est

$$\tau = \frac{2}{\lambda_{\min}(A_h) + \lambda_{\max}(A_h)},$$

qui donne ainsi

$$\rho = \frac{\kappa(A_h) - 1}{\kappa(A_h) + 1},$$

où

$$\kappa(A_h) = \frac{\lambda_{\max}(A_h)}{\lambda_{\min}(A_h)},$$

est le *nombre de conditionnement* de la matrice A_h. Si on utilise l'algorithme du *gradient conjugué* qui est plus performant, le facteur de réduction est alors

$$\rho = \frac{\sqrt{\kappa(A_h)} - 1}{\sqrt{\kappa(A_h)} + 1}.$$

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 partiulier du laplacien que ce nombre augmente vers $+\infty$ lorsque l'on augmente la résolution du maillage.

Théorème 2.7.1 *Dans le cas de la résolution de l'équation du laplacien par la méthode des éléments finis $\mathbb{P}_{\mathrm{k}}$ dans une famille régulière de triangulations, on a*

$$\kappa(A_h) \geq ch^{-2}. \tag{2.26}$$

Si la famille est quasi-uniforme, on a aussi

$$\kappa(A_h) \leq Ch^{-2}. \tag{2.27}$$

Preuve : pour simplifier, on se place dans le cas des éléments $\mathbb{P}_1$, l'adaptation de la preuve pour les éléments $\mathbb{P}_{\mathrm{k}}$ est laissée en exercice. On va estimer les valeurs propres $\lambda_{\min}(A_h)$ et $\lambda_{\max}(A_h)$. On a

$$\lambda_{\max}(A_h) = \max_{V_h} \frac{\langle A_h V_h, V_h\rangle}{\|V_h\|^2} \ \text{ et } \ \lambda_{\min}(A_h) = \min_{V_h} \frac{\langle A_h V_h, V_h\rangle}{\|V_h\|^2}.$$

Pour $\lambda_{\max}$ on prend un sommet $\gamma \in \Gamma_h$ tel que l'un des simplexes T contenant γ vérifie $h_T = h$, et on considère la fonction de base associée $v_h = \varphi_\gamma$ et son vecteur de coordonnée V_h, qui vérifie clairement $\|V_h\| = 1$. Pour ce vecteur on a

$$\langle A_h V_h, V_h\rangle = a(v_h, v_h) = \int \|\nabla\varphi_\gamma\|^2.$$

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

$$\|\nabla\varphi_\gamma\| \geq h_T^{-1}.$$

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

$$\langle A_h V_h, V_h\rangle \geq C \min_{\gamma\in T} \rho_T^d h_T^{-2}.$$

En prenant le simplexe T tel que $h_T = h$ et en utilisant la régularité des partitions, on trouve donc

$$\lambda_{\max}(A_h) \geq C h^{d-2}. \tag{2.28}$$

Pour λmin 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 coordonées V_h. Les coordonées $v_h(\gamma)$ de v_h sont positives, et valent 1 si $\gamma \in B$. Par conséquent

$$\|V_h\|^2 \geq M_h,$$

où M_h est le nombre de points de Γ_h contenu dans B. Comme le maillage est de finesse h, on a $M_h \geq C h^{-d}$ où la constante C ne dépend que de d, et donc

$$\|V_h\|^2 \geq C h^{-d}.$$

On a d'autre part

$$\begin{aligned} \langle A_h V_h, V_h \rangle &= \|v_h\|_{H_0^1}^2 \\ &\leq (\|v\|_{H_0^1} + \|v - v_h\|_{H_0^1})^2 \\ &\leq (\|v\|_{H_0^1} + h\|v\|_{H^2})^2 \\ &\leq C, \end{aligned}$$

où $C = (\|v\|_{H_0^1} + \mathrm{diam}(\Omega)\|v\|_{H^2})^2$ est une constante. On en déduit

$$\lambda_{\min}(A_h) \leq C h^d. \tag{2.29}$$

En combinant (2.28) et (2.29), on obtient (2.26).

Dans le cas de partitions régulières et quasi-uniforme, on obtient (2.27) en prouvant d'abord l'existence de deux constantes $C_1, C_2 \geq 0$ telle que pour tout $h > 0$ et $v_h \in X_h$ de vecteur de coordonées V_h, on a

$$C_1 h^d \|V_h\|^2 \leq \|v_h\|_{L^2} \leq C_2 h^d \|V_h\|^2. \tag{2.30}$$

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 $\hat{\pi} \in \Pi_1$

$$c_1 \sum_{i=0}^{d} |\hat{\pi}(\hat{a}_i)|^2 \leq \|\hat{\pi}\|_{L^2(\hat{T})}^2 \leq c_2 \sum_{i=0}^{d} |\hat{\pi}(\hat{a}_i)|^2,$$

ce qui par changement de variable donne

$$c_1 \frac{|T|}{|\hat{T}|} \sum_{i=0}^{d} |\pi(a_i)|^2 \leq \|\pi\|_{L^2(T)}^2 \leq c_2 \frac{|T|}{|\hat{T}|} \sum_{i=0}^{d} |\pi(a_i)|^2.$$

En sommant sur tous les simplexes, on obtient

$$c_1 \sum_{\gamma \in \Gamma_h} w_\gamma |v(\gamma)|^2| \leq \|v_h\|_{L^2} \leq c_2 \sum_{\gamma \in \Gamma_h} w_\gamma |v(\gamma)|^2$$

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

$$\tilde{c}_1 h^d \leq w_\gamma \leq \tilde{c}_2 h^d,$$

avec deux constantes $\tilde{c}_2, \tilde{c}_2 > 0$ indépendantes de γ et h, ce qui nous permet d'aboutir à (2.30).

On écrit ensuite pour tout $v_h \in X_h$ de vecteur de coordonées V_h

$$\langle A_h V_h, V_h \rangle = \|v_h\|_{H_0^1}^2 \geq \frac{1}{C_P^2} \|v_h\|_{L^2}^2,$$

où C_P est la constante de Poincaré sur Ω. En combinant ceci avec l'inégalité de gauche dans (2.30), on obtient

$$\lambda_{\min}(A_h) \geq \frac{C_1}{C_P^2} h^d.$$

D'autre part, l'inégalité inverse établie dans le théorème 2.6.1 nous montre que

$$\langle A_h V_h, V_h \rangle = \|v_h\|_{H_0^1}^2 \leq C h^{-2} \|v_h\|_{L^2}^2,$$

ce qui combiné avec l'inégalité de droite dans (2.30), nous donne

$$\lambda_{\max}(A_h) \leq C h^{d-2}.$$

On obtient ainsi l'estimation (2.27). $\diamond$

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).

2.8 Autres exemples d'éléments finis

Dans cette section nous passons en revue quelques exemples d'éléments finis d'usage moins courant que les éléments triangulaires de Lagrange, et nous indiquons brievement leurs propriétés.

Exemple 1 : l'élément triangulaire de Thomas
On se place en dimension $d = 2$ et on considère le triplet (T, X_T, Σ_T) où T est le triangle de sommets (a_0, a_1, a_2), $X_T := \Pi_k$ avec $k \geq 3$ et Σ_T est défini par l'union des formes linéaires

$$v \mapsto \int_T v \pi_\lambda,$$

où (π_λ) est une base de Π_{k-3},

$$v \mapsto \int_{e_i} v \pi_\mu, \; i = 0, 1, 2,$$

où e_i est l'arête de T opposée à a_i et (π_μ) une base de Π_{k-2} sur chaque e_i, ainsi que

$$v \mapsto v(a_i), \;\; i = 0, 1, 2.$$

Il est facile de vérifier l'unisolvence pour ces choix en remarquant tout d'abord que le cardinal de Σ_T est égal à la dimension de Π_k. On montre ensuite que si $v \in \Pi_k$ est annulé par toutes les formes de Σ_T, on a sur chaque arête $\int_{e_i} v v'' = 0$, d'où $\int_{e_i} v'^2 = 0$ soit $v' = 0$ ce qui entraine la nullité de v sur le bord ∂T. On en déduit que $v(x) = \lambda_0(x)\lambda_1(x)\lambda_2(x)w(x)$ òu $w \in \Pi_{k-3}$. Puisque $\int_T vw = 0$, on en déduit que v est identiquement nul.

La démonstration de l'unisolvence montre aussi que l'espace d'élément finis obtenu X_h est celui des fonctions Π_k par morceaux et globalement continues, c'est à dire le même que celui des éléments de Lagrange. Ce sont les formes linéaires choisies, et donc la base et interpolant, qui diffèrent.

Exemple 2 : 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

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

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 Q_k n'est pas laissé invariant par transport de $\hat{T}$ vers T et on a donc en général $X_T \neq Q_k$. 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 3 : 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 $X_T = \Pi_{2k+1}$ et des degres 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 $\mathcal{C}^k$ 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 coordonées sur ces axes, on choisi $X_T = Q_{2k+1}$ où 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

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

On vérifie aisément l'unisolvence de ce triplet, ainsi que le fait que les fonctions de l'espace X_h obtenu possèdent la régularité $\mathcal{C}^k$ 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 4 : l'élément triangulaire d'Argyris

Il est possible de revenir à des triangulations conforme et d'obtenir la régularité $\mathcal{C}^1$ et H^2 par la construction suivante : sur un triangle T de sommet (a_0, a_1, a_2), on prend $X_T = \Pi_5$, et pour les formes linéaires

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

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 Π_5. Puis on suppose que $v \in \Pi_5$ 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 $\mathcal{C}^1$ aux interfaces entre les triangles et ont par conséquent la régularité H^2. Le principal défaut de cet élément est sa complexité : 21 degrés de liberté par triangle.

Exemple 5 : l'élément triangulaire de Crouzeix-Raviart

Il existe des espaces d'éléments finis intéressant qui au contraire des précedents ont moins de régularité. Pour le triangle de Crouzeix-Raviart en dimension $d = 2$, on prend $X_T = \Pi_1$, et

$$\Sigma_T := \{v \mapsto v(b_i)\ ;\ i = 0, 1, 2\},$$

où b_i est le milieu de l'arête e_i opposée à a_i. L'unisolvence est évidente. Un autre choix qui est équivalent à celui-ci puisqu'on travaille avec des polynômes de degré 1 est de prendre les formes $v \mapsto \int_{e_i} v$. Les fonctions de l'espace d'élément fini X_h associé à ces choix ne sont pas globalement continues. Cependant entre deux triangles adjacents, on a continuité de la valeur au milieu de l'arête où de l'intégrale sur l'arête. Cette propriété s'appelle le *patch test* et permet d'utiliser ces éléments pour résoudre des problèmes tels que le laplacien en sortant du cadre de l'approximation interne, comme nous le montrons dans la section suivante. Le patch test se généralise pour les éléments de degré k par la propriété

$$\int_{e_i} [v_h]\pi = 0, \quad \pi \in \Pi_{k-1},$$

où $[v_h]$ est le saut à l'interface.

2.9 Approximation non-conforme

Afin d'utiliser les éléments de Crouzeix-Raviart pour discrétiser le problème du laplacien (1.1) sur un domaine $\Omega \subset \mathbb{R}^2$ muni d'une famille de triangulations régulière $(\mathcal{T}_h)_{h\geq 0}$, on introduit la forme bilinéaire brisée

$$a_h(u, v) := \sum_{T\in\mathcal{T}_h} \int_T \nabla u \cdot \nabla v,$$

qui coincide avec la forme bilinéaire $a(u, v)$ de la formulation variationnelle lorsque u et v sont dans $H^1(\Omega)$. On introduit aussi l'espace $X_{h,0}$ des éléments de Crouzeix-Raviart dont les degrés de libertés situés sur $\partial\Omega$ sont nuls. Le problème discret est alors : chercher $u_h \in X_{h,0}$ tel que pour tout $v_h \in X_{h,0}$,

$$a_h(u_h, v_h) = L(v_h), \tag{2.31}$$

où L est la forme linéaire de la formulation variationnelle. Une première remarque est que ce problème a toujours une solution unique car les hypothèses de Lax-Milgram sont satisfaites sur X_h. La seule propriété non-évidente est la coercivité de a_h. Comme on est en dimension finie, il suffit de prouver que cette forme est définie positive. Pour cela on remarque que si $a_h(u_h, u_h)$ est nulle, alors u_h est constante sur chaque triangle. Comme il y a continuité au milieu de chaque arête u_h est globalement constante. Comme elle s'annule sur les milieux des arêtes contenues dans la frontière $\partial\Omega$, elle est identiquement nulle. Cette remarque nous montre par le même biais que

$$\|u\|_h := \sqrt{a_h(u, u)},$$

définit une norme sur $X_{h,0}$. C'est dans cette norme que nous allons établir l'estimation d'erreur $u - u_h$ puisqu'il est impossible de l'avoir dans la norme H_0^1. Il sera au préalable utile d'établir une inégalité de Poincaré pour les fonctions de $X_{0,h}$.

Théorème 2.9.1 *Si $(\mathcal{T}_h)_{h>0}$ est une famille régulière, il existe une constante C indépendante de h telle que*

$$\|v_h\|_{L^2} \leq C\|v_h\|_h$$

pour tout $v_h \in X_{h,0}$.

Preuve : on va utiliser l'estimation par dualité

$$\|v_h\|_{L^2} = \sup_{g\in L^2(\Omega)} \frac{\int_\Omega v_h g}{\|g\|_{L^2}}.$$

On remarque d'abord que pour tout $g \in L^2(\Omega)$, il existe un vecteur $w \in (H^1(\Omega))^2$ tel que

$$\operatorname{div}(w) = g \text{ et } \|w\|_{H^1} \leq C\|g\|_{L^2},$$

où la constante C ne dépend pas de g. Une manière de construire w est la suivante : on étend g à un domaine convexe ou régulier $\widetilde{\Omega}$ contenant Ω (par exemple un carré) en posant

$$\tilde{g}(x) = g(x) \text{ si } x \in \Omega,\ 0 \text{ sinon.}$$

puis on résoud l'équation du laplacien $-\Delta v = \tilde{g}$ sur $\widetilde{\Omega}$ avec condition homogènes de Dirichlet au bord, et on pose $w = \nabla v$. On a ainsi en utilisant le théorème 1.4.1

$$\|w\|_{H^1} \leq \|v\|_{H^2} \leq C\|\tilde{g}\|_{L^2} = C\|g\|_{L^2}.$$

En utilisant cette construction, on en déduit que l'on a

$$\|v_h\|_{L^2} \leq C \sup_{w\in L^2(\Omega)} \frac{\int_\Omega v_h \mathrm{div}(w)}{\|w\|_{H^1}}.$$

On écrit maintenant

$$\int_\Omega v_h \mathrm{div}(w) = \sum_{T\in\mathcal{T}_h} \int_{\partial T} (w\cdot \mathbf{n})v_h - \sum_{T\in\mathcal{T}_h} \int_T \nabla v_h \cdot w.$$

La deuxième somme peut-être majorée par

$$|\sum_{T\in\mathcal{T}_h} \int_T \nabla v_h \cdot w| \leq \|v_h\|_h \|w\|_{L^2}.$$

La première somme peut s'écrire sous la forme

$$\sum_{e\in\mathcal{E}_h} \int_e (w\cdot\mathbf{n})[v_h],$$

où $\mathcal{E}_h$ est l'ensemble des arêtes. Dans cette expression, si e est une arête entre T et T' et si $\mathbf{n}$ est la normale à e vers T' alors $[v_h] := v_{h|T} - v_{h|T'}$. Dans le cas où e est sur le bord du domaine alors $\mathbf{n}$ est la normale extérieure au domaine et $[v_h] := v_{h|T}$.

Afin d'estimer $\int_e (w\cdot\mathbf{n})[v_h]$, on utilise le fait que $[v_h]$ est d'intégrale nulle sur e, ce qui permet d'écrire pour tout $c \in \mathbb{R}$

$$|\int_e (w\cdot\mathbf{n})[v_h]| = |\int_e ((w\cdot\mathbf{n}) - c)[v_h - v_h(b)]| \leq \|(w\cdot\mathbf{n}) - c\|_{L^2(e)} \|[v_h - v_h(b)]\|_{L^2(e)},$$

où b est le milieu de e. Comme $v_h(b)$ est aussi la moyenne de v_h sur e, on a

$$|\int_e (w\cdot\mathbf{n})[v_h]| \leq \min_{c\in\mathbb{R}} \|(w\cdot\mathbf{n}) - c\|_{L^2(e)} \min_{c\in\mathbb{R}} \|[v_h - c]\|_{L^2(e)}$$

Afin d'évaluer les normes L^2 qu'on a fait apparaitre, on utilise l'élément de référence $\widehat{T}$ et on écrit que pour toute fonction $\varphi \in H^1(T)$, on a pour toute arête e de T

$$\begin{aligned}
\min_{c\in\mathbb{R}} \|\varphi - c\|_{L^2(e)} &= |e|^{1/2} \min_{c\in\mathbb{R}} \|\hat{\varphi} - c\|_{L^2(\hat{e})} \\
&\leq Ch_T^{1/2} \min_{c\in\mathbb{R}} \|\hat{\varphi} - c\|_{H^1(\widehat{T})} \\
&\leq Ch_T^{1/2} |\hat{\varphi}|_{H^1(\widehat{T})} \\
&\leq C\frac{h_T^{3/2}}{|T|^{1/2}} |\varphi|_{H^1(T)} \\
&\leq Ch_T^{1/2} |\varphi|_{H^1(T)},
\end{aligned}$$

où on a utilisé le théorème de trace, le théorème de Deny-Lions, les formules de changement de variable et la régularité des triangulations, avec une constante C qui varie d'une ligne à l'autre. Ceci nous permet d'obtenir

$$\min_{c\in\mathbb{R}} \|(w\cdot\mathbf{n}) - c\|_{L^2(e)} \leq Ch_T^{1/2} |w|_{H^1(T)},$$

et

$$\min_{c\in\mathbb{R}} \|[v_h - c]\|_{L^2(e)} \leq C(h_T^{1/2}|v_h|_{H^1(T)}+, h_{T'}^{1/2}|v_h|_{H^1(T')}),$$

où T et T' sont les triangles commun à e (dans le cas où e est contenu dans $\partial\Omega$ il n'y a qu'un seul terme). En sommant sur les arêtes et en appliquant l'inégalité de Cauchy-Schwarz, on obtient finalement l'estimation

$$|\sum_{e\in\mathcal{E}_h} \int_e (w\cdot\mathbf{n})[v_h]| \leq Ch|w|_{H^1}\|v_h\|_h.$$

On a donc obtenu

$$\int_\Omega v_h \mathrm{div}(w) \leq C\|v_h\|_h(\|w\|_{L^2} + h|w|_{H^1}) \leq C\|v_h\|_h\|w\|_{H^1},$$

ce qui prouve le résultat souhaité. ◇

Une première conséquence de l'inégalité de Poincaré sur $X_{0,h}$ est une estimation a-priori pour la solution discrète : on a en effet

$$\|u_h\|_h^2 = a_h(u_h, u_h) = \int_\Omega f u_h \leq \|f\|_{L^2}\|u_h\|_{L^2} \leq C\|f\|_{L^2}\|u_h\|_h,$$

soit

$$\|u_h\|_h \leq C\|f\|_{L^2}.$$

On a finalement une estimation d'erreur donnée par le théorème suivant.

Théorème 2.9.2 *Si la triangulation est régulière, et si la solution u est dans $H^2(\Omega)$ on a l'estimation d'erreur*

$$\|u - u_h\|_h \leq Ch|u|_{H^2}.$$

Preuve : on écrit tout d'abord que pour tout $v_h \in X_{h,0}$ on a

$$a_h(u_h, v_h) = L(v_h) = -\int_\Omega \Delta u\, v_h = a_h(u, v_h) - \sum_{e\in\mathcal{E}_h} \int_e \frac{\partial u}{\partial n}[v_h],$$

soit

$$a_h(u - u_h, v_h) = \sum_{e\in\mathcal{E}_h} \int_e \frac{\partial u}{\partial n}[v_h].$$

En utilisant la même estimation que dans la preuve de l'inégalité de poincaré avec ici $w = \nabla u$, on trouve ainsi

$$a_h(u - u_h, v_h) \leq Ch\|v_h\|_h|u|_{H^2}.$$

On introduit alors l'interpolant I_h sur $X_{h,0}$ pour lequel on peut prouver par les mêmes techniques que pour les éléments de Lagrange l'estimation

$$\|u - I_h u\|_h \leq Ch|u|_{H^2}.$$

En prenant $v_h = I_h u - u_h$, on obtient ainsi

$$a_h(v_h, v_h) = a_h(I_h u - u, v_h) + a_h(u - u_h, v_h) \leq Ch\|v_h\|_h|u|_{H^2},$$

Soit

$$\|I_h u - u_h\|_h \leq Ch|u|_{H^2},$$

ce qui, combiné avec l'estimation pour l'erreur d'interpolation donne l'estimation d'erreur annoncée. ◇

3 Estimation a-postériori et adaptativité

3.1 Pourquoi et comment adapter le maillage ?

Les estimations d'erreur que nous avons obtenues telles que

$$\|u-u_h\|_{H_0^1} \leq Ch^{m-1}|u|_{H^m}, \tag{3.32}$$

font intervenir la finesse du maillage et régularité H^m de la solution u que l'on cherche à approcher. Dans le cas d'une triangulation quasi-uniforme, la dimension $N = N_h$ de X_h est du même ordre que nombre de simplexe de $\mathcal{T}_h$ c'est à dire

$$N \sim h^{-d},$$

pour un domaine d-dimensionel. On peut ainsi reformuler l'estimation (3.32) comme un compromis entre la précision et la taille des calcul :

$$\|u-u_h\|_{H_0^1} \leq C|u|_{H^m} N^{-\frac{m-1}{d}}.$$

Dans le cas particulier de l'approximation par des éléments $\mathbb{P}_1$ en dimension $d = 2$, on a

$$\|u-u_h\|_{H_0^1} \leq C|u|_{H^2} N^{-1/2}. \tag{3.33}$$

On voit que ce compromis se dégrade lorsque la solution u présente des singularités qui rendent très grande ou infinie la quantité $|u|_{H^m}$. Il peut alors être intéressant d'utiliser des maillages qui ne sont pas uniformes, mais localement raffinés au voisinage de ces singularités.

Un exemple simple d'apparition de singularité peut-être illustré dans le cas de l'équation du laplacien (1.1) en dimension $d = 2$ lorsque le domaine Ω n'est pas régulier. Pour cela on part de la remarque suivante : si K est un cône infini d'ouverture $\omega \in]0, 2\pi[$, c'est à dire

$$K = \{(r\sin\theta, r\cos\theta)\ ;\ r > 0,\ 0 < \theta < \omega\},$$

il existe des solutions non-triviale à l'équation

$$\Delta u = 0 \ \text{ dans } \ K, \ \ u_{|\partial K} = 0.$$

Il suffit en effet de considérer l'expression du laplacien en coordonées pôlaires

$$\Delta u = \frac{1}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2},$$

et de chercher les solutions sous la forme

$$u(r,\theta) = r^\alpha s(\theta).$$

on aboutit ainsi à l'équation

$$\alpha^2 s(\theta) + s''(\theta) = 0, \ \ s(0) = s(\omega) = 0,$$

dont les solution existent pour

$$\alpha = \alpha_n := \frac{n\pi}{\omega}, \ \ n \in \mathbb{Z}$$

et sont données par

$$u_n(r,\theta) = r^{\alpha_n}\sin(\alpha_n\theta).$$

Considérons à présent le domaine borné

$$\Omega := \{(r\sin\theta, r\cos\theta)\ ;\ 0 < r < R,\ 0 < \theta < \omega\},$$

pour $R > 0$, et soit φ une fonction $\mathcal{C}^\infty([0,R])$ telle que $\varphi(x) = 1$ sur $[0, R/3]$ et $\varphi(x) = 0$ sur $[2R/3, R]$. On définit alors les fonctions

$$v_n = v_n(r,\theta) = \varphi(r)u_n(r,\theta).$$

La fonction v_n est solution de l'équation du laplacien (1.1) avec second membre $f = f_n = -\Delta v_n \in \mathcal{C}^\infty(\overline{\Omega})$. Considérons en particulier la fonction v_1. On voit que pour $0 < r < R/3$,

$$\frac{\partial^2 v_1}{\partial r^2} = r^{\alpha_1 - 2} \sin(\alpha_1 \theta).$$

Comme $\alpha_1 = \pi/\omega$, on en déduit que pour $\omega > \pi$ la fonction $\frac{\partial^2 v_1}{\partial r^2}$ n'appartient pas à $L^2(\Omega)$. Comme

$$\frac{\partial^2 v_1}{\partial r^2} = \langle D^2 v_1 e_r, e_r \rangle$$

où $D^2 v_1$ est la hessienne de v_1 et $e_r = (\cos\theta, \sin\theta)$, ceci entraine que v_1 n'appartient pas à $H^2(\Omega)$ même si le second membre f_1 est de classe $\mathcal{C}^\infty$. Par conséquent l'estimation (3.33) ne s'applique pas.

Remarque 3.1.1 *Pour $n > 1$, on peut vérifier que les solutions v_n appartiennent à H^2, mais pas à tous les espaces H^m. On ne considère pas les fonctions v_n pour $n < 0$ qui sont moins régulières que v_1 et n'appartiennent pas à $H^1(\Omega)$. Il s'agit donc de solutions "non-variationelles", et on voit ainsi que l'unicité de telles solution n'est plus garantie puisqu'il existe toujours par ailleurs une solution dans $H^1(\Omega)$.*

Remarque 3.1.2 *On peut citer de multiples autres sources de singularités dans les EDP elliptiques : lorsque le second membre f possède des singularités, lorsque les opérateurs intervenant dans l'équation sont à coefficients discontinus, lorsque la condition aux limites est de type Dirichlet sur une partie du bord et de type Neumann sur une autre partie.*

Afin de comprendre le gain potentiel d'une discrétisation adaptative, nous allons partir d'un résultat d'approximation polynomiale locale sur le triangle de référence.

Lemme 3.1.1 *Pour tout $\hat{u} \in W^{2,1}(\hat{T})$, si $I_{\widehat{T}}$ est l'interpolant sur Π_k avec $k \geq 1$, on a*

$$|\hat{u} - I_{\widehat{T}} \hat{u}|_{H^1} \leq C |\hat{u}|_{W^{2,1}}.$$

où C est une constante fixée qui ne dépend que de k.

Preuve : on part de l'injection de Sobolev $W^{2,1} \subset H^1$ en dimension $d = 2$ pour écrire

$$|\hat{u} - I_{\widehat{T}} \hat{u}|_{H^1} \leq C \|\hat{u} - I_{\widehat{T}} \hat{u}\|_{W^{2,1}}.$$

On utilise ensuite le fait que $W^{2,1}$ s'injecte continuement dans $\mathcal{C}$ en dimension $d = 2$ pour appliquer le lemme de Bramble-Hilbert qui nous donne le résultat. $\diamond$

Par changement de variable, on obtient sur tout triangle T,

$$|u - I_T u|_{H^1} \leq C |T|^{-1/2} \frac{h_T^2}{\rho_T} |u|_{W^{2,1}(T)}, \tag{3.34}$$

qui est un cas particulier de (2.22). Lorsque la triangulation $\mathcal{T}_h$ auquel appartient T est régulière, cette inégalité nous donne simplement

$$|u - I_T u|_{H^1} \leq C |u|_{W^{2,1}(T)},$$

où la constante C ne dépend pas de la taille du triangle T

Supposons à présent que la solution u appartienne à $W^{2,1}(\Omega)$ et que l'on puisse construire une triangulation régulière $\mathcal{T}_h$ qui *équidistribue* la semi-norme $W^{2,1}$ sur Ω c'est à dire

$$|u|_{W^{2,1}(T)} \leq \frac{|u|_{W^{2,1}(\Omega)}}{N},$$

où $N = \#(\mathcal{T}_h)$ (notons qu'une telle triangulation sera en général non-uniforme). En utilisant (3.34), on a alors

$$\begin{aligned}|u - I_h u|_{H^1} &= \Big(\textstyle\sum_{T\in\mathcal{T}_h} |u - I_T u|^2_{H^1(T)}\Big)^{1/2}\\ &\leq C\Big(\textstyle\sum_{T\in\mathcal{T}_h} |u|^2_{W^{2,1}(T)}\Big)^{1/2}\\ &\leq C|u|_{W^{2,1}} N^{-1/2},\end{aligned}$$

et par conséquent, d'après le lemme de Cea,

$$\|u - u_h\|_{H^1_0} \leq C|u|_{W^{2,1}} N^{-1/2}. \tag{3.35}$$

Si on compare (3.35) et (3.33), on observe qu'une triangulation adaptative bien construite permet de retrouver la vitesse d'approximation $N^{-1/2}$ lorsque u n'est pas dans $H^2(\Omega)$ mais seulement dans $W^{2,1}(\Omega)$. On peut vérifier que c'est le cas pour la fonction v_1 considérée précèdemment, dont les dérivées du second ordre ne sont pas dans L^2 pour $\omega > \pi$ mais sont cependant dans L^1 (exercice).

En pratique, lorsque u est la solution d'une équation telle que celle du laplacien, il n'est pas possible de contruire un triangulation $\mathcal{T}_h$ qui équidistribue la semi-norme $W^{2,1}$ de u, puisque que u est précisément inconnue. Afin de s'affranchir de l'évaluation de cette semi-norme sur chaque triangle, on remarque tout d'abord qu'une autre manière de construire une bonne triangulation adaptative consiste à équidistribuer *l'erreur d'approximation locale* sur chaque triangle. Supposons en effet que l'on ait construit une triangulation $\mathcal{T}_h$ telle que pour un certain $\varepsilon > 0$ on ait pour tout $T \in \mathcal{T}_h$

$$c\varepsilon \leq |u - I_T u|_{H^1(T)} \leq \varepsilon,$$

avec $c \in]0, 1[$ une constante indépendante de u et ε (par exemple $c = 1/2$). On voit alors qu'avec $N = \#(\mathcal{T}_h)$, on a

$$\|u - I_h u\|_{H^1_0} = \Big(\sum_{T\in\mathcal{T}_h} |u - I_T u|^2_{H^1(T)}\Big)^{1/2} \leq N^{1/2}\varepsilon. \tag{3.36}$$

Par ailleurs, si $u \in W^{2,1}(T)$, en utilisant (3.34) on peut écrire

$$c\varepsilon \leq |u - I_T u|_{H^1(T)} \leq C|u|_{W^{2,1}(T)},$$

ce qui en sommant sur tous les triangle nous montre que

$$\varepsilon \leq \frac{C}{cN}|u|_{W^{2,1}}.$$

En combinant avec (3.36) en conclut que l'estimation (3.35) est aussi vérifiée par la triangulation $\mathcal{T}_h$.

De façon plus générale et en dimension d quelconque, une partition adaptative $\mathcal{T}_h$ optimale au sens du compromis entre la complexité et la précision doit tendre à équidistribuer l'erreur locale entre u et son approximation $I_h u$ dans X_h. A nouveau, en pratique on ne connait ni u ni $I_h u$. On peut essayer d'estimer l'erreur locale entre u et la solution discrète u_h, à partir des données du problème et de u_h. C'est l'objectif des techniques *d'estimation a-postériori* qui fournissent des *indicateurs d'erreur* $\eta_T(u_h)$ visant à approcher $|u - u_h|_{H^1(T)}$. On peut alors choisir de découper les éléments T dont les indicateurs sont les plus grand afin d'obtenir une nouvelle triangulation mieux adaptée à la solution.

Remarque 3.1.3 *Lors d'un raffinement de maillage local, il est souvent important de préserver les propriétés de conformité et de régularité géométrique. Cela ne va pas de soi : en dimension $d = 2$, si on raffine un seul triangle en le découpant en quatre à partir des milieux (b_0, b_1, b_2) des arêtes, on introduit de la non-conformité en ces trois points. On peut corriger ce défaut en bisectant les trois triangles adjacents depuis les sommets opposés à ces points mais on remarque alors que l'on tend à diminuer l'angle minimum des triangles. Il est possible de définir des procédures plus sophistiquées qui garantissent la régularité de la famille de triangulations obtenue. Dans ces procédures, le raffinement d'un triangle peut entrainer celui d'autres triangles que ceux qui lui sont adjacent. On observe ainsi un phénomène de "pollution" du raffinement, qui a pour effet d'augmenter la complexité, et que l'on peut parfois quantifier plus précisément.*

Remarque 3.1.4 *Notons que dans une partition non-uniforme mais régulière, la taille des éléments adjacents ne peut pas varier trop brutalement : si T et T' ont une partie de leur bord en commun on a nécessairement*

$$h_T \leq C h'_T$$

où la constante C ne dépend de la constante de régularité (exercice).

3.2 Estimation a-postériori par indicateurs d'erreur

Nous allons utiliser l'équation du laplacien (1.1) et sa discrétisation dans les espaces d'éléments finis $\mathbb{P}_k$ afin de décrire la construction des indicateurs d'erreur locaux. On se placera ici en dimension $d = 2$ et avec une donnée $f \in L^2(\Omega)$. Cette construction se généralise en dimension supérieure et à d'autres problèmes.

Soit u la solution exacte et u_h la solution approchée. Afin d'évaluer $\|u - u_h\|_{H_0^1}$ on dispose déjà des estimations a-priori (2.23) et (2.24) mais celles-ci font intervenir la connaissance de u. Nous allons établir une estimation a-postériori qui ne dépend que de u_h et des données du problème, c'est à dire ici le second membre f.

Pour cela on écrit

$$\|u - u_h\|_{H_0^1} = \sup_{w \in H_0^1(\Omega)} \frac{a(u - u_h, w)}{\|w\|_{H_0^1}}.$$

On évalue le numérateur en écrivant d'abord que pour tout $w_h \in X_{h,0}$ on a

$$\begin{aligned} a(u - u_h, w) &= a(u - u_h, w - w_h) \\ &= \int_\Omega \nabla(u - u_h) \cdot \nabla(w - w_h) \\ &= \int_\Omega (f(w - w_h) - \nabla u_h \cdot \nabla(w - w_h)) \\ &= \sum_{T \in \mathcal{T}_h} \int_T (f(w - w_h) - \nabla u_h \cdot \nabla(w - w_h)) \\ &= \sum_{T \in \mathcal{T}_h} \int_T (f + \Delta u_h)(w - w_h) - \sum_{T \in \mathcal{T}_h} \int_{\partial T} \frac{\partial u_h}{\partial n}(w - w_h) \\ &= \Sigma_1 + \Sigma_2 \end{aligned}$$

En introduisant l'ensemble $\mathcal{E}_h$ des arêtes de $\mathcal{T}_h$ et en convenant d'une orientation pour la normale sur chaque élément de $\mathcal{E}_h$ comme dans la preuve du théorème 2.9.1, on peut écrire

$$\Sigma_2 = \sum_{e \in \mathcal{E}_h} \int_e [\frac{\partial u_h}{\partial n}](w - w_h).$$

où $[\frac{\partial u_h}{\partial n}]$ est le saut de la dérivée le long de la normale choisie pour e et vaut $\frac{\partial u_h}{\partial n}$ lorsque $e \subset \partial\Omega$. On peut ensuite utiliser l'inégalité de Cauchy-Schwarz pour écrire

$$\Sigma_1 \leq \sum_{T \in \mathcal{T}_h} \|f + \Delta u_h\|_{L^2(T)} \|w - w_h\|_{L^2(T)},$$

et

$$\Sigma_2 \leq \sum_{e \in \mathcal{E}_h} \|[\frac{\partial u_h}{\partial n}]\|_{L^2(e)} \|w - w_h\|_{L^2(e)}.$$

Rappelons que w_h est arbitraire. On ne peut pas prendre $w_h = I_h w$ car w est simplement dans $H_0^1(\Omega)$ et il n'est donc pas possible en général de l'interpoler. Nous allons cependant utiliser un opérateur d'approximation de type *quasi-interpolant* dont la définition a été proposée par Clément.

Définition 3.2.1 *Soit $\mathcal{T}_h$ une partition en triangles conformes, $\Gamma_{0,h}$ l'ensemble de ses sommets internes et $(\varphi_\gamma)_{\gamma \in \Gamma_{0,h}}$ la base des éléments finis $\mathbb{P}_1$ avec conditions aux limites nulles. On définit l'opérateur A_h par*

$$A_h v = \sum_{\gamma \in \Gamma_{0,h}} m_\gamma(v) \varphi_\gamma,$$

avec

$$m_\gamma(v) = |K_\gamma|^{-1} \int_{K_\gamma} v,$$

la moyenne de v sur l'ensemble K_γ défini comme l'union des triangles de $\mathcal{T}_h$ dont γ est l'un des sommets.

Autrement dit on remplace la valeur ponctuelle $v(\gamma)$ dans la définition de l'interpolant I_h par une moyenne locale autour de γ. L'intérêt de l'opérateur de Clément est qu'il s'applique à toute fonction de $L^2(\Omega)$ et qu'il vérifie des propriétés d'approximation locale comme l'exprime le théorème suivant.

Théorème 3.2.1 *Si $(\mathcal{T}_h)$ est une famille régulière, l'opérateur A_h vérifie les estimations locales*

$$\|v - A_h v\|_{L^2(T)} \leq C h_T |v|_{H^1(K_T)}, \tag{3.37}$$

où K_T est l'union de T et des triangles dont l'un des sommets est commun à T, et

$$\|v - A_h v\|_{L^2(e)} \leq C h_e^{1/2} |v|_{H^1(K_e)}, \tag{3.38}$$

où K_e est l'union des triangles dont l'un des sommet est contenu dans e et h_e est la longueur de e. La constante C ne dépend pas de v, T, e et h.

Avant d'aborder la preuve de ces inégalités appliquons les à l'estimation de $\|u - u_h\|_{H_0^1}$ en prenant $w_h = A_h w$. On obtient ainsi

$$\Sigma_1 \leq C \sum_{T \in \mathcal{T}_h} h_T \|f + \Delta u_h\|_{L^2(T)} |w|_{H^1(K_T)},$$

et

$$\Sigma_2 \leq C \sum_{e \in \mathcal{E}_h} h_e^{1/2} \|[\frac{\partial u_h}{\partial n}]\|_{L^2(e)} |w|_{H^1(K_e)}.$$

On applique ensuite l'inégalité de Cauchy-Schwarz qui nous donne,

$$\Sigma_1 + \Sigma_2 \leq C(A_1 A_2)^{1/2},$$

avec

$$A_1 := \sum_{T \in \mathcal{T}_h} h_T^2 \|f + \Delta u_h\|_{L^2(T)}^2 + \sum_{e \in \mathcal{E}_h} h_e \|[\frac{\partial u_h}{\partial n}]\|_{L^2(e)}^2,$$

et

$$A_2 := \sum_{T \in \mathcal{T}_h} |w|_{H^1(K_T)}^2 + \sum_{e \in \mathcal{E}_h} |w|_{H^1(K_e)}^2.$$

En introduisant les indicateurs d'erreur locaux

$$\eta_T^2 := h_T^2 \|f + \Delta u_h\|_{L^2(T)}^2 + \frac{1}{2} \sum_{e \in \partial T} h_e \|[\frac{\partial u_h}{\partial n}]\|_{L^2(e)}^2,$$

avec la convention que le facteur $\frac{1}{2}$ est remplacé par 1 pour une arête contenue sur le bord, on peut exprimer le premier facteur A_1 suivant

$$A_1 = \sum_{T \in \mathcal{T}_h} \eta_T^2.$$

D'autre part, le facteur A_2 s'ecrit aussi

$$A_2 := \sum_{T \in \mathcal{T}_h} n(T) |w|_{H^1(K_T)}^2$$

où

$$\begin{aligned} n(T) &= \#\{T' \,;\, T \subset K_{T'}\} + \#\{e \,;\, T \in K_e\} \\ &= \#\{T' \,;\, T' \subset K_T\} + \#\{e \,;\, e \subset \mathrm{int}(K_T)\}. \end{aligned}$$

Comme la triangulation est régulière, on a un angle minimal $\theta_{\min} > 0$ pour tous les triangles, et on vérifie aisément que

$$\#\{T' \,;\, T' \subset K_T\} \leq \frac{5\pi}{\theta_{\min}} - 2,$$

et

$$\#\{e\ ;\ e \subset \mathrm{int}(K_T)\} \leq \frac{5\pi}{\theta_{\min}}.$$

On a donc

$$A_2 \leq C\|w\|^2_{H^1_0(\Omega)},$$

avec $C = \frac{10\pi}{\theta_{\min}} - 2$. Nous avons ainsi obtenu

$$a(u-u_h, w) = \Sigma_1 + \Sigma_2 \leq C\Big(\sum_{T\in\mathcal{T}_h} \eta_T^2\Big)^{1/2}\|w\|_{H^1_0(\Omega)},$$

ce qui nous donne finalement l'estimation a-postériori

$$\|u-u_h\|_{H^1_0} \leq C\Big(\sum_{T\in\mathcal{T}_h} \eta_T^2\Big)^{1/2}. \tag{3.39}$$

Remarque 3.2.1 *Les quantités η_T associées à chaque triangle T sont appelés indicateurs d'erreurs locaux. Cependant, on ne dispose pas d'estimation locale du type $|u-u_h|_{H^1(T)} \leq C\eta_T$.*

Remarque 3.2.2 *La quantité $f + \Delta u_h$ est le résidu de l'équation du laplacien quand $u = u_h$. Lorsque qu'on utilise les éléments $\mathbb{P}_1$ ce résidu se réduit à f. Notons que le terme de saut $[\frac{\partial u_h}{\partial n}]$ est lui aussi nul lorsque $u = u_h$.*

Donnons finalement la preuve du théorème 3.2.1. Pour l'estimation (3.37), supposons d'abord que le triangle $T = (a_0, a_1, a_2)$ n'ait aucun sommet sur le bord du domaine. On peut alors écrire en tout point $x \in T$

$$v(x) - A_h v(x) = v(x) - \sum_{\gamma=a_0,a_1,a_2} m_\gamma \varphi_\gamma(x) = \sum_{\gamma=a_0,a_1,a_2} (v(x) - m_\gamma)\varphi_\gamma(x),$$

où nous avons utilisé le fait que $\sum_{\gamma=a_0,a_1,a_2} \varphi_\gamma(x) = 1$ sur T. On a par conséquent

$$\|v - A_h v\|_{L^2(T)} \leq \sum_{\gamma=a_0,a_1,a_2} \|v - m_\gamma\|_{L^2(T)}.$$

Comme $K_\gamma \subset K_T$ pour $\gamma = a_0, a_1, a_2$, il nous suffit donc de prouver que pour tout $\gamma \in \Gamma_{h,0}$ on a

$$\|v - m_\gamma\|_{L^2(K_\gamma)} \leq C h_T |v|_{H^1(K_\gamma)}, \tag{3.40}$$

où h_T est le diamètre d'un triangle dont γ est l'un des sommets. La difficulté est ici que K_γ n'est pas un simple triangle mais une union de triangles :

$$K_\gamma = T_1 \cup \cdots \cup T_n.$$

Comme la famille de triangulations est régulière, il existe $N \leq \frac{2\pi}{\theta_{\min}}$ tel que

$$3 \leq n \leq N,$$

pour tout $\gamma \in \Gamma_{h,0}$ et $h > 0$. On peut ainsi introduire $N - 2$ configurations de références

$$\widehat{K}_n = \widehat{T}_1 \cup \cdots \cup \widehat{T}_n,$$

avec $\widehat{T}_i$ le triangle dont l'un des sommets est $\hat{a}_0 = 0$ et les deux autres sont $\hat{a}_1 = (1, \frac{i-1}{n})$ et $\hat{a}_2 = (1, \frac{i}{n})$ en coordonées polaires (r, θ).

On définit une transformation A de $\widehat{K}_n$ vers K_γ qui est globalement continue et affine par morceaux et envoie $\hat{T}_i$ vers T_i. Sur $\widehat{K}$, on peut appliquer le théorème de Deny-Lions qui donne

$$\min_{c\in\mathbb{R}} \|\hat{v} - c\|_{L^2(\widehat{K}_n)} \leq C|\hat{v}|_{H^1(\widehat{K}_n)}.$$

Comme il n'y a qu'un nombre fini de configurations on peut supposer que la constante C est indépendante de n. Les changements de variables nous donnent ici pour toute fonction v définie sur K_γ

$$\|v\|^2_{L^2(K_\gamma)} = \sum_{i=1}^{n} \frac{|T_i|}{|\hat{T}_i|} \int_{\hat{T}_i} |v|^2 \leq C h_T^2 \|v\|^2_{L^2(\hat{K}_n)},$$

où on a utilisé la régularité de la triangulation soit

$$\|v\|_{L^2(K_\gamma)} \leq C h_T \|v\|_{L^2(\hat{K}_n)},$$

et de même

$$\|v\|^2_{L^2(\hat{K}_n)} \leq C h_T^{-1} \|v\|_{L^2(K_\gamma)}.$$

On a ainsi

$$\begin{aligned}
\min_{c\in\mathbb{R}} \|v-c\|_{L^2(K_\gamma)} &\leq C h_T \min_{c\in\mathbb{R}} \|\hat{v}-c\|_{L^2(\widehat{K}_n)} \\
&\leq C h_T |\hat{v}|_{H^1(\widehat{K}_n)} \\
&\leq C \max_{\hat{x}\in\widehat{K}_n} \|DA(\hat{x})\| \, |v|_{H^1(K_\gamma)} \\
&\leq C \max_{i=1,\cdots,n} \frac{h_{T_i}}{\rho_{\widehat{T}_i}} |v|_{H^1(K_\gamma)}.
\end{aligned}$$

En utilisant à nouveau la régularité de la triangulation, on a donc

$$\min_{c\in\mathbb{R}} \|v-c\|_{L^2(K_\gamma)} \leq C h_T |\hat{v}|_{H^1(K_\gamma)}.$$

Comme le minimum est atteint pour $c = m_\gamma$, on en déduit l'estimation (3.40) et donc (3.37).

Le cas où le triangle T a certains sommets sur le bord de Ω doit être traité spécifiquement puisque les valeurs de $A_h v$ aux sommets contenus dans $\partial\Omega$ sont nulles et ne sont donc plus données par les moyennes locales de v. Notons $S_T = \{a_0, a_1, a_2\} \cap \partial\Omega$ l'ensemble des sommets de T situés sur le bord. On écrit alors

$$\|v - A_h v\|_{L^2(T)} \leq \sum_{\gamma=a_0,a_1,a_2} \|v - m_\gamma\|_{L^2(T)} + \sum_{\gamma\in S_T} \|m_\gamma \varphi_\gamma\|_{L^2(T)}.$$

Le premier terme se majore comme dans le cas d'un triangle qui ne touche pas le bord. Pour le deuxième terme, on écrit

$$\begin{aligned}
\|m_\gamma \varphi_\gamma\|_{L^2(T)} &\leq |m_\gamma| \, \|\varphi_\gamma\|_{L^2(T)} \\
&\leq |T|^{1/2} |m_\gamma| \\
&= |T|^{1/2} |K_\gamma|^{-1} |\textstyle\int_{K_\gamma} v| \\
&\leq |T|^{1/2} |K_\gamma|^{-1/2} \|v\|_{L^2(K_\gamma)} \\
&\leq \|v\|_{L^2(K_\gamma)}.
\end{aligned}$$

On passe finalement à l'élément de référence $\widehat{K}_n$ sur lequel on utilise l'inégalité de Poincaré (puisque v s'annulle sur l'un des cotés de K_γ, donc $\hat{v}$ s'annulle sur l'un des cotés de $\widehat{K}_n$) ce qui donne

$$\begin{aligned}
\|v\|_{L^2(K_\gamma)} &\leq C h_T \|\hat{v}\|_{L^2(\widehat{K}_n)} \\
&\leq C h_T |v|_{H^1(\widehat{K}_n)} \\
&\leq C h_T |v|_{H^1(K_\gamma)}.
\end{aligned}$$

En sommant sur les $\gamma \in S_T$ on a ainsi prouvé (3.37).

La preuve de (3.38) s'effectue de façon très similaire. En notant (a_0, a_1) les extremités de e, on traite d'abord le cas où aucun de ces points est situé sur le bord de Ω. Par le même raisonnement on arrive d'abord à

$$\|v - A_h v\|_{L^2(e)} \leq \sum_{\gamma=a_0,a_1} \|v - m_\gamma\|_{L^2(e)}.$$

On passe à la configuration de référence et on écrit

$$\begin{aligned}
\|v - m_\gamma\|_{L^2(e)} &= h_e^{1/2}\|\hat{v} - m_\gamma\|_{L^2(\hat{e})} \\
&\leq Ch_e^{1/2}\|\hat{v} - m_\gamma\|_{H^1(\widehat{K}_n)} \\
&= Ch_e^{1/2}(\|\hat{v} - m_\gamma\|^2_{L^2(\widehat{K}_n)} + |\hat{v}|^2_{H^1(\widehat{K}_n)})^{1/2} \\
&\leq Ch_e^{1/2}(h_e^2\|v - m_\gamma\|^2_{L^2(K_\gamma)} + |v|^2_{H^1(K_\gamma)})^{1/2} \\
&\leq Ch_e^{1/2}(h_e^2 \min_{c\in\mathbb{R}} \|v - c\|^2_{L^2(K_\gamma)} + |v|^2_{H^1(K_\gamma)})^{1/2} \\
&\leq Ch_e^{1/2}(\min_{c\in\mathbb{R}} \|\hat{v} - c\|^2_{L^2(\widehat{K}_n)} + |v|^2_{H^1(K_\gamma)})^{1/2} \\
&\leq Ch_e^{1/2}(|\hat{v}|^2_{H^1(\widehat{K}_n)} + |v|^2_{H^1(K_\gamma)})^{1/2} \\
&\leq Ch_e^{1/2}|v|_{H^1(K_\gamma)} \\
&\leq Ch_e^{1/2}|v|_{H^1(K_e)}
\end{aligned}$$

où on a utilisé le théorème de trace, le théorème de Deny-Lions, les changements de variables et la régularité des triangulations. On obtient ainsi (3.38). Dans le cas où l'une des extremité de e est sur le bord de Ω, on adapte la preuve comme pour celle de (3.37) en utilisant l'inégalité de Poincaré et on aboutit à la même conclusion.

3.3 Fiabilité des indicateurs d'erreur et saturation

Les indicateurs d'erreurs η_T à la fois à évaluer la taille de l'erreur $u - u_h$ et à choisir les éléments qui seront raffinés dans l'étape suivante du calcul. Il est important que ces indicateurs soient fiable au sens où ils ne surévaluent pas l'erreur de manière exagérée ce qui aurait pour conséquence la mise en oeuvre de raffinement de maillages inutiles.

Dans cette optique, nous allons montrer que η_T peut être lui même controlé par l'erreur $u - u_h$ mesurée localement autour du triangle T. On introduit pour cela une approximation polynomiale par morceaux f_h de la donnée f qui est définie sur chaque triangle T comme la projection $L^2(T)$-orthogonale de f sur Π_{k-1}.

Théorème 3.3.1 *On a sur chaque triangle*

$$\eta_T \leq C(|u - u_h|_{H^1(\Omega_T)} + h_T\|f - f_h\|_{L^2(\Omega_T)}), \tag{3.41}$$

où la constante C est indépendente de T, h et f et où $\Omega_T = T \cup T_1 \cup T_2 \cup T_3$ où (T_1, T_2, T_3) sont les triangles adjacents à T.

Preuve : Elle est technique du fait de l'approximation de f par f_h, et il peut être utile à la compréhension de l'effectuer dans le cas plus simple où f est polynomiale sur chaque triangle et où l'on prend directement $f = f_h$.

Nous allons majorer séparément les contributions $h_T\|f + \Delta u_h\|_{L^2(T)}$ et $h_e^{1/2}\|[\frac{\partial u_h}{\partial n}]\|_{L^2(e)}$ à l'indicateur d'erreur η_T. Pour la première contribution, on écrit tout d'abord

$$h_T\|f + \Delta u_h\|_{L^2(T)} \leq h_T\|f_h + \Delta u_h\|_{L^2(T)} + h_T\|f - f_h\|_{L^2(T)}).$$

On introduit ensuite la fonction bulle

$$\psi_T = 27\lambda_0\lambda_1\lambda_2.$$

Par équivalence de norme, pour le triangle de référence il existe une constante $C = C(n)$ telle que pour tout $\hat{\pi} \in \Pi_n$,

$$\|\hat{\pi}\|_{L^2(\widehat{T})} \leq C\|\hat{\pi}\sqrt{\hat{\psi}_T}\|_{L^2(\widehat{T})}$$

soit par changement de variable

$$\|\pi\|_{L^2(T)} \leq C\|\pi\sqrt{\psi_T}\|_{L^2(T)}$$

où C est indépendante du triangle T. On peut appliquer ceci à $\pi = f_h + \Delta u_h$ ce qui donne

$$\begin{aligned}\|f_h + \Delta u_h\|^2_{L^2(T)} &\leq C^2\|(f_h + \Delta u_h)\sqrt{\psi_T}\|^2_{L^2(T)}\\ &= C^2 \int_T (f_h + \Delta u_h)(f_h + \Delta u_h)\psi_T\\ &= C^2(\int_T (f + \Delta u_h)(f_h + \Delta u_h)\psi_T\\ &\quad + \int_T (f_h - f)(f_h + \Delta u_h)\psi_T).\end{aligned}$$

Pour le premier terme, on écrit

$$\begin{aligned}\int_T (f + \Delta u_h)(f_h + \Delta u_h)\psi_T &= \int_T \Delta(u_h - u)(f_h + \Delta u_h)\psi_T\\ &= \int_T \nabla(u - u_h)\nabla((f_h + \Delta u_h)\psi_T)\\ &\leq |u - u_h|_{H^1(T)}\|(f_h + \Delta u_h)\psi_T\|_{H^1(T)}\\ &\leq Ch_T^{-1}|u - u_h|_{H^1(T)}\|(f_h + \Delta u_h)\psi_T\|_{L^2(T)}\\ &\leq Ch_T^{-1}|u - u_h|_{H^1(T)}\|f_h + \Delta u_h\|_{L^2(T)},\end{aligned}$$

où on a utilisé l'inégalité inverse pour passer de la semi-norme H^1 à la norme L^2. Pour le deuxième terme, on écrit simplement

$$\int_T (f_h - f)(f_h + \Delta u_h)\psi_T) \leq \|f - f_h\|_{L^2(T)}\|f_h + \Delta u_h\|_{L^2(T)}.$$

Nous avons donc

$$\|f_h + \Delta u_h\|^2_{L^2(T)} \leq C(h_T^{-1}|u - u_h|_{H^1(T)}\|f_h + \Delta u_h\|_{L^2(T)}, + \|f - f_h\|^2_{L^2(T)}),$$

ce qui entraine

$$h_T\|f_h + \Delta u_h\|_{L^2(T)} \leq C(|u - u_h|_{H^1(T)} + h_T\|f - f_h\|_{L^2(T)}). \tag{3.42}$$

Examinons à présent la contribution $h_e^{1/2}\|[\frac{\partial u_h}{\partial n}]\|_{L^2(e)}$. Si e est d'extrémité (a_0, a_1), on introduit la fonction

$$\psi_e = 4\lambda_0\lambda_1,$$

dont le support est l'union des deux triangles T et T' dont e est la frontière commune. Par un raisonnement d'équivalence de norme des polynômes définis sur le segment de référence similaire à celui effectué plus haut sur $\widehat{T}$, on peut écrire

$$\begin{aligned}\|[\tfrac{\partial u_h}{\partial n}]\|^2_{L^2(e)} &\leq C\int_e |[\tfrac{\partial u_h}{\partial n}]|^2\psi_e\\ &= C\int_e [\tfrac{\partial(u_h - u)}{\partial n}]([\tfrac{\partial u_h}{\partial n}]\psi_e)\\ &= C\int_e [\tfrac{\partial(u_h - u)}{\partial n}]g_e\psi_e,\end{aligned}$$

où C est indépendante de e et où g_e est égale à $[\frac{\partial u_h}{\partial n}]$ sur e, et étendue sur $T \cup T'$ de la façon suivante : sur T elle est constante le long des droites de directions parallèlle à l'un des côtés de T autre que e, et de même sur T'. On peut alors utiliser la formule de Green qui nous donne

$$\|g_e\|^2_{L^2(e)} = C\Big(\int_{T\cup T'} (\Delta u_h + f)g_e\psi_e + \int_{T\cup T'} (\nabla u_h - \nabla u)\nabla(g_e\psi_e)\Big). \tag{3.43}$$

Le premier terme du membre de droite de (3.43) se majore suivant

$$|\int_{T\cup T'} (\Delta u_h + f)g_e\psi_e| \leq \|\Delta u_h + f\|_{L^2(T\cup T')}\|g_e\psi_e\|_{L^2(T\cup T')}.$$

En passant à l'élément de référence et en utilisant la régularité des triangulations, on vérifie aisément que

$$\|g_e\psi_e\|_{L^2(T)} \leq Ch_e^{1/2}\|g_e\|_{L^2(e)}.$$

On en déduit que

$$|\int_{T\cup T'} (\Delta u_h + f)g_e\psi_e| \leq Ch_e^{1/2}(\|\Delta u_h + f\|_{L^2(T)} + \|\Delta u_h + f\|_{L^2(T')})\|g_e\|_{L^2(e)}.$$

Le deuxième terme du membre de droite de (3.43) se majore suivant

$$\begin{aligned}|\int_{T\cup T'}(\nabla u_h-\nabla u)\nabla(g_e\psi_e)| &\leq |u-u_h|_{H^1(T\cup T')}|g_e\psi_e|_{H^1(T\cup T')}\\ &\leq |u-u_h|_{H^1(T\cup T')}(|g_e\psi_e|_{H^1(T)}+|g_e\psi_e|_{H^1(T')})\\ &\leq |u-u_h|_{H^1(T\cup T')}(h_T^{-1}\|g_e\psi_e\|_{L^2(T)}\\ &\qquad\qquad\qquad +h_{T'}^{-1}\|g_e\psi_e\|_{L^2(T')})\\ &\leq C|u-u_h|_{H^1(T\cup T')}h_e^{-1/2}\|g_e\|_{L^2(e)},\end{aligned}$$

où nous avons utilisé l'inégalité inverse sur les triangles T et T' et la régularité des triangulations. En réinjectant ces estimations dans (3.43) et en simplifiant par $\|g_e\|_{L^2(e)}$ nous obtenons donc

$$h_e^{1/2}\|g_e\|_{L^2(e)}\leq C\Big(h_T\|\Delta u_h+f\|_{L^2(T)}+h_{T'}\|\Delta u_h+f\|_{L^2(T')}+|u-u_h|_{H^1(T\cup T')}\Big).$$

Les termes $h_T\|\Delta u_h+f\|_{L^2(T)}+h_{T'}\|\Delta u_h+f\|_{L^2(T')}$ étant déjà majorés par (3.42), on trouve

$$h_e^{1/2}\|[\frac{\partial u_h}{\partial n}]\|_{L^2(e)}\leq C\Big(|u-u_h|_{H^1(T\cup T')}+h_T\|f-f_h\|_{L^2(T\cup T')}\Big). \tag{3.44}$$

En sommant les carrés de (3.42) et de (3.44) pour les trois cotés de T, on obtient finalement l'estimation (3.41) annoncée. ◇

Une conséquence immédiate de ce résultat est une estimation globale de l'erreur par valeur inférieure :

$$c(\sum_{T\in\mathcal{T}_h}\eta_T^2)^{1/2}\leq\|u-u_h\|_{H_0^1}+\varepsilon_h(f) \tag{3.45}$$

où le terme $\varepsilon_h(f):=(\sum_T h_T^2\|f-f_h\|_{L^2(T)}^2)^{1/2}$ décrit l'oscillation de la donnée f à l'intérieur des triangles de $\mathcal{T}_h$.

Une stratégie adaptative naturelle consiste à raffiner les éléments de $\mathcal{T}_h$ dont les indicateurs sont les plus grands. Plus précisément, on peut définir $\mathcal{M}_h\subset\mathcal{T}_h$ le plus petit ensemble de triangles dont les indicateurs d'erreur retiennent une proportion $0<\mu<1$ fixée de l'énergie totale des indicateurs, c'est à dire tel que

$$\sum_{T\in\mathcal{M}_h}\eta_T^2\geq\mu\sum_{T\in\mathcal{T}_h}\eta_T^2. \tag{3.46}$$

Un choix typique est $0.05\leq\mu\leq 0.2$. On raffine ensuite les triangles de $\mathcal{M}_h$ en prenant soin de conserver la conformité et la régularité, et on obtient une nouvelle triangulation $\mathcal{T}_{\tilde h}$ avec $\tilde h\leq h$ et $X_{h,0}\subset X_{\tilde h,0}$ et une nouvelle solution $u_{\tilde h}$. En partant d'une triangulation initiale $\mathcal{T}_h=\mathcal{T}_0$ et en notant $X_0=X_{h,0}$ l'espace $u_0=u_h$ la solution approchée pour cette triangulation, on obtient ainsi une suite de triangulations $(\mathcal{T}_j)_{j\geq 0}$, qui correspond à une suite d'espace d'éléments finis emboités

$$X_0\subset X_1\subset\cdots\subset X_j\subset\cdots,$$

et à une suite de solutions approchées u_j. Il est légitime de se poser la question de la convergence de u_j vers u quand $j\to+\infty$ qui n'est pas garantie par les théorèmes d'approximation du chapitre précèdent puisqu'on ne raffine pas la triangulation de façon uniforme. Une première remarque est que l'on a toujours

$$\|u-u_{\tilde h}\|_{H_0^1}\leq\|u-u_h\|_{H_0^1},$$

puisque $\|u-u_h\|_{H_0^1}=\min_{v_h\in X_{h,0}}\|u-v_h\|_{H_0^1}$ et que $X_{h,0}\subset X_{\tilde h,0}$. Plus précisément, par le théorème de Pythagore, on a

$$\|u-u_h\|_{H_0^1}^2=\|u-u_{\tilde h}\|_{H_0^1}^2+\|u_{\tilde h}-u_h\|_{H_0^1}^2. \tag{3.47}$$

On en déduit que $\|u-u_j\|_{H_0^1}$ décroit mais cela n'est pas suffisant pour assurer la convergence. Voici une propriété qui permet d'assurer la convergence de la méthode adaptative.

Définition 3.3.1 *Le procédé de raffinement adaptatif vérifie la "propriété de saturation" si et seulement si, il existe* $0 < \nu < 1$ *tel que*

$$\|u_{\tilde{h}} - u_h\|_{H_0^1} \geq \nu \|u - u_h\|_{H_0^1}.$$

pour toute triangulation $\mathcal{T}_h$ *et son raffinement* $\mathcal{T}_{\tilde{h}}$.

Si une telle propriété est satisfaite, on a d'après (3.47)

$$\|u - u_{\tilde{h}}\|^2_{H_0^1} = (1 - \nu^2)\|u - u_h\|^2_{H_0^1},$$

et par conséquent

$$\|u - u_n\|_{H_0^1} \leq \rho \|u - u_{n-1}\|_{H_0^1} \leq \cdots \leq \rho^n \|u - u_0\|_{H_0^1},$$

avec un facteur de réduction $\rho := (1-\nu^2)^{1/2}$. La propriété de saturation nous assure donc une convergence géométrique de l'algorithme adaptatif. Le problème est donc de savoir si cette propriété est satisfaite, en particulier lorsqu'on applique la stratégie décrite par (3.46) pour la sélection des éléments à raffiner. Les premières réponses positives à cette question ont été apportées par des résultats récents dus à Dorfler ainsi qu'à Morin, Nochetto et Siebert, dans des cas particuliers d'EDP et de discrétisations. Voici un résultat qui porte sur la discrétisation de l'équation du laplacien par les éléments finis $\mathbb{P}_1$.

Théorème 3.3.2 *On suppose que l'ensemble des triangles raffinés vérifie (3.46), et que si* T *fait partie de cet ensemble, la nouvelle triangulation* $\mathcal{T}_{\tilde{h}}$ *contient parmis ses sommets les milieux des arêtes de* T *ainsi qu'un point intérieur à* T*. Alors, il existe une constante* $c > 0$ *telle que*

$$c(\sum_{T \in \mathcal{M}_h} \eta_T^2)^{1/2} \leq \|u_{\tilde{h}} - u_h\|_{H_0^1} + \varepsilon_h(f),$$

Avant donner la preuve de ce résultat, examinons ses conséquences. D'après (3.46), on a alors

$$c\mu^{1/2}(\sum_{T \in \mathcal{T}_h} \eta_T^2)^{1/2} \leq \|u_{\tilde{h}} - u_h\|_{H_0^1} + \varepsilon_h(f),$$

et en appliquant l'estimation d'erreur a-posteriori (3.39), on a donc l'existence d'une constante $0 < \nu < 1$, tel que

$$\nu \|u - u_h\|_{H_0^1} \leq \|u_{\tilde{h}} - u_h\|_{H_0^1} + \varepsilon_h(f).$$

On obtient ainsi la propriété de saturation au terme d'oscillation des données près, et on en déduit facilement

$$\|u - u_{\tilde{h}}\|^2_{H_0^1} \leq (1 - \frac{\nu^2}{2})\|u - u_h\|^2_{H_0^1} + \varepsilon_h(f)^2.$$

Cette propriété nous montre que l'algorithme adaptatif réduit l'erreur tant que celle-ci est plus grande que le terme d'oscillation des données. Une stratégie possible afin d'atteindre une erreur d'ordre $\varepsilon > 0$ prescrite est d'imposer que la triangulation $\mathcal{T}_0 = \mathcal{T}_h$ de départ vérifie

$$\varepsilon_0 = \varepsilon_h(f) \leq \varepsilon.$$

On vérifie facilement que le terme d'oscillation décroit lorsque l'on raffine la triangulation, et on a par conséquent

$$\|u - u_n\|^2_{H_0^1} \leq \tilde{\rho}^2 \|u - u_{n-1}\|^2_{H_0^1} + \varepsilon^2.$$

avec $\tilde{\rho} := (1 - \frac{\nu^2}{2})^{1/2}$. Comme la suite $a_n = \tilde{\rho}^2 a_{n-1} + \varepsilon^2$ converge vers $(1 - \tilde{\rho}^2)^{-1}\varepsilon^2$, on est assuré d'avoir

$$\|u - u_n\|_{H_0^1} \leq C\varepsilon,$$

avec une constante $C > (1 - \tilde{\rho}^2)^{-1/2}$ pour n suffisament grand.

Donnons finalement la preuve du théorème 3.3.2. Il suffit pour cela de prouver que sur chaque triangle raffiné, on a

$$\eta_T \leq C(|u_{\tilde{h}} - u_h|_{H^1(\Omega_T)} + h_T \|f - f_h\|_{L^2(\Omega_T)}), \tag{3.48}$$

Comme dans la preuve du théorème 3.3.1, on va majorer séparément les contributions $h_T\|f+\Delta u_h\|_{L^2(T)}$ et $h_e^{1/2}\|[\frac{\partial u_h}{\partial n}]\|_{L^2(e)}$. Pour la première, on procède de manière similaire mais on prend comme fonction bulle ψ_T la fonction de base $\mathbb{P}_1$ de $X_{\tilde{h},0}$ associé au sommet de $\mathcal{T}_{\tilde{h}}$ interne à T. De la même manière, cela nous conduit à l'évaluation du terme $\int_T(f+\Delta u_h)(f_h+\Delta u_h)\psi_T$ pour lequel on écrit cette fois

$$\begin{aligned}
\int_T(f+\Delta u_h)(f_h+\Delta u_h)\psi_T &= \int_T \nabla(u-u_h)\nabla((f_h+\Delta u_h)\psi_T) \\
&= \int_T \nabla(u_{\tilde{h}}-u_h)\nabla((f_h+\Delta u_h)\psi_T) \\
&\le |u_{\tilde{h}}-u_h|_{H^1(T)}\|(f_h+\Delta u_h)\psi_T\|_{H^1(T)} \\
&\le Ch_T^{-1}|u_{\tilde{h}}-u_h|_{H^1(T)}\|(f_h+\Delta u_h)\psi_T\|_{L^2(T)} \\
&\le Ch_T^{-1}|u_{\tilde{h}}-u_h|_{H^1(T)}\|f_h+\Delta u_h\|_{L^2(T)}.
\end{aligned}$$

On aboutit ainsi à

$$h_T\|f_h+\Delta u_h\|_{L^2(T)} \le C\Big(|u_{\tilde{h}}-u_h|_{H^1(T)}+h_T\|f-f_h\|_{L^2(T)}\Big). \tag{3.49}$$

Pour la contribution $h_e^{1/2}\|[\frac{\partial u_h}{\partial n}]\|_{L^2(e)}$, on procède à nouveau comme dans la preuve du théorème 3.3.1 mais on prend comme fonction bulle ψ_T la fonction de base $\mathbb{P}_1$ de $X_{\tilde{h},0}$ associé au milieu de e. Cela nous permet de modifier l'évaluation du deuxième terme du membre de droite de (3.43), suivant

$$\begin{aligned}
|\int_{T\cup T'}(\nabla u_h-\nabla u)\nabla(g_e\psi_e)| &= |\int_{T\cup T'}(\nabla u_h-\nabla u_{\tilde{h}})\nabla(g_e\psi_e)| \\
&\le |u_{\tilde{h}}-u_h|_{H^1(T\cup T')}|g_e\psi_e|_{H^1(T\cup T')} \\
&\le |u_{\tilde{h}}-u_h|_{H^1(T\cup T')}(|g_e\psi_e|_{H^1(T)}+|g_e\psi_e|_{H^1(T')}) \\
&\le |u_{\tilde{h}}-u_h|_{H^1(T\cup T')}(h_T^{-1}\|g_e\psi_e\|_{L^2(T)} \\
&\qquad\qquad +h_{T'}^{-1}\|g_e\psi_e\|_{L^2(T')}) \\
&\le C|u_{\tilde{h}}-u_h|_{H^1(T\cup T')}h_e^{-1/2}\|g_e\|_{L^2(e)}.
\end{aligned}$$

On aboutit ainsi à

$$h_e^{1/2}\|[\frac{\partial u_h}{\partial n}]\|_{L^2(e)} \le C\Big(|u_{\tilde{h}}-u_h|_{H^1(T\cup T')}+h_T\|f-f_h\|_{L^2(T\cup T')}\Big). \tag{3.50}$$

En sommant les carrés de (3.49) et de (3.50) pour les trois cotés de T, on obtient finalement l'estimation (3.48).

Remarque 3.3.1 *Le résultat que nous avons établi montre que l'algorithme de raffinement local itératif fondé sur les valeurs des indicateurs d'erreur donne une suite de solution $(u_n)_{n\ge 0}$ qui converge vers u. Une question plus difficile est d'établir un taux de convergence au sens du compromis entre la précision $\|u-u_n\|_{H_0^1}$ et la complexité $N=\#(\mathcal{T}_n)$. Cette question a fait l'objet de travaux récents qui montrent que sous certaines hypothèses, on obtient le même taux de convergence que si l'on utilisait des triangulations optimalement adaptées à la solution u au sens de l'équidistribution de l'erreur locale.*

4 Approximation des problèmes mixtes

4.1 Le problème de Stokes

Le problème de Stokes est posé dans un ouvert borné Lipschitzien $\Omega \subset \mathbb{R}^d$, avec typiquement $d = 2$ ou $d = 3$. Etant donnée une fonction à valeur vectorielle $f = (f_1, \cdots, f_d)$ définie sur Ω et un réel $\nu > 0$, on cherche une fonction à valeur vectorielle $u = (u_1, \cdots, u_d)$ et une fonction à valeur scalaire p telles que

$$\begin{cases} -\nu\Delta u + \nabla p = f & \text{dans } \Omega, \\ \operatorname{div}(u) = 0 & \text{dans } \Omega, \\ u = 0 & \text{sur } \Gamma \end{cases} \tag{4.51}$$

avec $\Delta u := (\Delta u_1, \cdots, \Delta u_n)$. Les inconnues u et p de ce problème représentent le champs de vitesse et de pression d'un fluide visqueux incompressible en régime permanent soumis à un champ de force volumique f. Ce modèle peut-être vu comme une version simplifiée du modèle de Navier-Stokes dans lequel la première équation s'écrit

$$-\nu\Delta u + (u \cdot \nabla)u + \nabla p = f$$

On voit que p est définie à une constante additive près et il est donc naturel de supposer p d'intégrale nulle. Afin d'obtenir une formulation variationelle, on multiplie la première équation par une fonction vectorielle v qui s'annulle au bord et on intègre par partie ce qui nous donne

$$\nu \int_\Omega \nabla u \cdot \nabla v - \int p \operatorname{div}(v) = \int_\Omega f v,$$

avec $\nabla u \cdot \nabla v := \nabla u_1 \cdot \nabla v_1 + \cdots + \nabla u_d \cdot \nabla v_d$. Comme $\int_\Omega \operatorname{div}(u) = 0$, on peut multiplier la deuxième équation par une fonction q de moyenne nulle, ce qui nous donne

$$\int q \operatorname{div}(u) = 0.$$

Il est ainsi naturel de prendre f dans $(H^{-1}(\Omega))^d$ et de chercher u dans $(H_0^1(\Omega))^d$ et p dans l'espace de $L_0^2(\Omega)$ défini comme l'ensemble des fonctions de $L^2(\Omega)$ de moyenne nulle. La formulation variationelle s'écrit donc : trouver $(u, p) \in (H_0^1(\Omega))^d \times L_0^2(\Omega)$ tel que

$$\begin{cases} a(u, v) + b(v, p) & = L(v) \\ b(u, q) & = 0 \end{cases} \tag{4.52}$$

pout tout $(v, q) \in (H_0^1(\Omega))^d \times L_0^2(\Omega)$, avec

$$a(u, v) := \nu \int_\Omega \nabla u \cdot \nabla v,$$

et

$$b(v, p) := -\int p \operatorname{div}(v).$$

On vérifie aisément que toute solution suffisament régulière de (4.52) est aussi solution de (4.53).

On peut éliminer l'inconnue p en introduisant le sous-espace des fonctions de $(H_0^1(\Omega))^d$ à divergence nulle

$$V := \{v \in (H_0^1(\Omega))^d \; ; \; \operatorname{div}(v) = 0\}.$$

On voit ainsi que u est aussi solution de la formulation variationelle : trouver $u \in V$ telle que pour tout $v \in V$

$$a(u, v) = L(v).$$

Par le théorème de Lax-Milgram appliqué à V on a l'existence d'une unique solution u telle que

$$\|u\|_{H_0^1} \leq \frac{1}{\nu}\|L\|_{H^{-1}} = \frac{1}{\nu}\|f\|_{H^{-1}}.$$

La difficulté est à présent de retrouver la pression p. Plus précisément, la solution u que l'on vient d'exhiber vérifie que $\nu\Delta u + f$ est orthogonal à tous les éléments de V au sens de la dualité entre $(H^{-1}(\Omega))^d$ et $(H_0^1(\Omega))^d$. Autrement dit, la forme $F(v) = -a(u,v) + L(v)$ est dans le *polaire* de V qui est l'espace des formes linéaires qui s'anullent sur V, i.e.

$$V^0 := \{F \in (H^{-1}(\Omega))^d \ ; \ F(v) = 0 \ , \ v \in V\}.$$

Il est facile de voir que les formes du type $v \mapsto b(p,v) = \int_\Omega p \operatorname{div}(v)$ sont dans V^0, mais peut-on affirmer réciproquement qu'il existe toujours un $p \in L_0^2(\Omega)$ tel que F est de cette forme ? La réponse est liée à un théorème difficile du à de Rham dont on admet ici la validité.

Théorème 4.1.1 *On suppose que Ω est lipschitzien et connexe. Il existe un opérateur T continu de $L_0^2(\Omega)$ dans $(H_0^1(\Omega))^d$ tel que*

$$\operatorname{div}(Tq) = q,$$

pour tout $q \in L_0^2(\Omega)$.

Remarque 4.1.1 *Dans la preuve du théorème 2.9.1 nous avions construit une fonction w de $(H^1(\Omega))^d$ telle que $\operatorname{div}(w) = q$ et $\|w\|_{H^1} \leq C\|q\|_{L^2}$. La difficulté supplémentaire est ici que l'on souhaite que w soit nulle au bord.*

Voyons à présent comment ce théorème nous permet de retrouver p. Une première remarque est qu'il est toujours possible de supposer que T est à valeur dans $V^\perp$, c'est à dire l'orthogonal de V dans $(H_0^1(\Omega))^d$ qu'il ne pas confondre ici avec V^0 puisqu'on fait la distinction entre H_0^1 et H^{-1}. En effet on peut toujours remplacer Tq par $\widetilde{T}q = (I - P_V)Tq$ où P_V est la projection orthogonale de sur V, et on a clairement

$$\|\widetilde{T}q\|_{H_0^1} \leq \|Tq\|_{H_0^1}.$$

Ceci nous montre que l'opérateur de divergence définit un isomorphisme de $V^\perp$ sur $L_0^2(\Omega)$ dont T est l'inverse. On voit aussi que pour tout $v \in (H_0^1(\Omega))^d$, on a $T\operatorname{div}(v) = P_{V^\perp}v$. Si $F \in V^0$, on remarque que F est completement définie par sa restriction à $V^\perp$ d'après

$$F(v) = F(P_{V^\perp}v),$$

et en utilisant l'isomorphisme T on peut écrire

$$F(v) = F(T\operatorname{div}(v)) = G(\operatorname{div}(v)),$$

où $G = F \circ T$ est une forme linéaire continue sur $L_0^2(\Omega)$. Par conséquent, il existe un unique $p \in L_0^2(\Omega)$ tel que

$$F(v) = -\int_\Omega p \operatorname{div}(v),$$

avec l'estimation $\|p\|_{L^2} \leq C_T\|F\|_{H^{-1}}$ où $C_T = \|T\|$. En appliquant ceci à $F(v) = -a(u,v) + L(v)$, on trouve l'existence d'un unique $p \in L_0^2(\Omega)$ telle que (u,p) est solution de la formulation variationelle (4.52). En prenant $v \in V^\perp$ tel que $\operatorname{div}(v) = p$ (c'est à dire $v = Tp$) on a

$$\|p\|_{L^2}^2 = \int_\Omega p \operatorname{div}(v) = \langle f, v\rangle \leq \|f\|_{H^{-1}}\|v\|_{H^1} \leq C_T\|f\|_{H^{-1}}\|p\|_{L^2},$$

soit

$$\|p\|_{L^2} \leq C_T\|f\|_{H^{-1}}.$$

Remarque 4.1.2 *La forme $b(p,v)$ permet de définir ∇p comme un élément de $(H^{-1}(\Omega))^d$ pour tout $p \in L_0^2(\Omega)$ et l'équation $a(u,v) + b(p,v) = L(v)$ peut-êre vue comme la première équation de (4.53) prise au sens de H^{-1}. La discussion précèdente nous indique que le gradient définit un isomorphisme de $L_0^2(\Omega)$ dans V^0.*

Il est intéressant d'étudier la version non-homogène du problème de Stokes, soit

$$\begin{cases} -\nu\Delta u + \nabla p = f & \text{dans } \Omega, \\ \operatorname{div}(u) = g & \text{dans } \Omega, \\ u = h & \text{sur } \Gamma \end{cases} \tag{4.53}$$

avec $g \in L^2_0(\Omega)$ et $h = (H^{1/2}(\Gamma))^d$. Il est nécessaire de supposer la condition de compatibilité

$$\int_\Gamma h \cdot \mathrm{n} = \int_\Omega g$$

Afin de se ramener à la version homogène, on commence par relever la trace de u par une fonction $u_r = R(h)$. On relève ensuite la divergence de $u - u_r$ par $u_g = T(g - \operatorname{div}(u_r))$. En écrivant

$$u = u_0 + u_r + u_g,$$

on voit que (u_0, p) est solution du système de Stokes homogène avec second membre $f_0 := f + \nu\Delta(u_r + u_g)$. On a ainsi existence d'une unique solution. En remarquant que

$$\begin{aligned} \|f_0\|_{H^{-1}} &\leq \|f\|_{H^{-1}} + \nu\|u_r\|_{H^1} + \nu\|u_g\|_{H^1_0} \\ &\leq \|f\|_{H^{-1}} + \nu\|h\|_{H^{1/2}} + \nu C_T(\|g\|_{L^2_0} + \|u_r\|_{H^1}) \\ &\leq \|f\|_{H^{-1}} + C_T\nu\|g\|_{L^2_0} + (1 + C_T)\nu\|h\|_{H^{1/2}}, \end{aligned}$$

on obtient des estimations a-priori sur u et p qui montrent qu'on a un isomorphisme reliant les données $(f, g, h) \in (H^{-1})^d \times L^2_0 \times (H^{1/2})^d$ satisfaisant la relation de compatibilité et la solution $(u, p) \in (H^1)^d \times L^2_0$. Citons finalement un résultat difficile de régularité pour le problème de Stokes homogène.

Théorème 4.1.2 *On suppose Ω connexe et de classe $C^{1,1}$ ou convexe. Alors si $f \in (L^2(\Omega))^d$, le couple (u, p) solution de (4.53) appartient à $(H^2(\Omega))^d \times H^1(\Omega)$ avec dépendance continue en fonction de f. Si Ω est de classe $C^{m+1,1}$ et $f \in H^m(\Omega)$ alors (u, p) appartient à $(H^{m+2})^d \times H^{m+1}$.*

4.2 Les problèmes mixtes abstraits

Afin d'aborder la discrétisation du problème de Stokes, il est intéressant de se placer dans un cadre abstrait plus général. Etant donné un couple (X, M) d'espaces de Hilbert, on suppose que a et b sont deux formes bilinéaires respectivement symétrique, continue et coercive sur $X \times X$, et continue sur $X \times M$, et que L et H sont deux formes linéaires sont respectivement continues sur X et M. On s'intéresse alors au problème : trouver $(u, p) \in X \times M$ tel que

$$\begin{cases} a(u, v) + b(v, p) &= L(v) \\ b(u, q) &= G(q) \end{cases} \tag{4.54}$$

pour tout $(v, q) \in X \times M$. On peut reformuler ce problème dans le langage des opérateurs en posant

$$a(u, v) = \langle Au, v\rangle_{X',X}$$

où $A \in \mathcal{L}(X, X')$, et

$$b(v, p) = \langle Bv, p\rangle M', M = \langle B'p, v\rangle_{X',X},$$

où $B \in \mathcal{L}(X, M')$ et $B' \in \mathcal{L}(M, X')$ est l'adjoint de B. La formulation s'écrit alors

$$\begin{cases} Au + B'p &= L \\ Bu &= G \end{cases} \tag{4.55}$$

Remarque 4.2.1 *Il est intéressant d'examiner l'analogue de ce système en dimension finie, lorsque A et B sont des matrices. On vérifie alors que ce système est équivalent à la minimisation de l'énergie*

$$E(u) = \frac{1}{2}\langle Au, u\rangle - \langle L, v\rangle,$$

sous la contrainte affine $Bu = H$, l'inconnue p représentant le multiplicateur de Lagrange pour cette contrainte. Dans le cas du système de Stokes, la pression peut ainsi être interpretée comme le multiplicateur de Lagrange pour la contrainte d'incompressibilité et on vérifie que (u, p) est l'unique minimiseur de $E(u) = \frac{1}{2}a(u, u) - L(v)$ sur l'espace V.

La propriété décrite par le théorème 4.1.1 dans le cas du système de Stokes se généralise sous le nom de condition LBB (Ladyzenskaia-Babushka-Brezzi) ou condition “inf-sup”.

Définition 4.2.1 *La forme b vérifie la condition LBB si et seulement si il existe $\beta > 0$ tel que*

$$\inf_{p\in M} \sup_{v\in X} \frac{b(v,p)}{\|v\|_X \|p\|_M} \geq \beta. \tag{4.56}$$

Par analogie avec le système de Stokes, on peut définir l'espace $V = \mathrm{Ker(B)}$ qui est un sous-espace de X, son orthogonal $V^\perp$ dans X et son polaire V^0 dans X'. Il est facile de vérifier que la condition LBB est équivalente à chacune des propriétés suivantes :

1. B' est un isomorphisme de M dans V^0 avec
$$\|B'p\|_{X'} \geq \beta\|p\|_M,$$
c'est à dire $\|(B')^{-1}\| \leq \frac{1}{\beta}$.
2. B est un isomorphisme de $V^\perp$ dans M' avec
$$\|Bv\|_{M'} \geq \beta\|v\|_X,$$
c'est à dire $\|B^{-1}\| \leq \frac{1}{\beta}$.

En effet si la condition (4.56) est satisfaite, on a d'une part pour tout $p \in X$,

$$\|B'p\|_{X'} = \sup_{v\in X, \|v\|_X=1} b(v,p) \geq \beta\|p\|_M.$$

Ceci nous montre que B' est injective et d'image fermée dans X'. D'après le théorème de l'image fermée de Banach on a

$$\mathrm{Im}(B') = (\mathrm{Ker}(B))^0 = V^0.$$

on a donc bien la propriété d'isomorphisme. Celle-ci est équivalente à la propriété d'isomorphisme de B par dualité en remarquant que $(V^\perp)' = V^0$. Réciproquement si B est un isomorphisme, on a pour tout $p \in M$ en faisant l'identification $M = M'$,

$$b(p, B^{-1}p) = \langle p, BB^{-1}p\rangle_M = \|p\|_M^2 \geq \beta\|p\|_M\|B^{-1}p\|$$

ce qui entraine

$$\sup_{v\in X, \|v\|_X=1} \frac{b(v,p)}{\|p\|_M} \geq \beta,$$

c'est à dire (4.56).

Théorème 4.2.1 *Sous la condition LBB, le problème (4.54) admet une unique solution $(u,p) \in X \times M$ qui satisfait les estimations*

$$\|u\|_X \leq \frac{1}{\alpha}\|L\|_{X'} + \frac{1}{\beta}(1 + \frac{C_a}{\alpha})\|G\|_{M'},$$

et

$$\|p\|_M \leq \frac{1}{\beta}(1 + \frac{C_a}{\alpha})\|L\|_{X'} + \frac{C_a}{\beta^2}(1 + \frac{C_a}{\alpha})\|G\|_{M'},$$

où α et C_a sont les constantes d'ellipticité et de continuité de a.

Preuve : c'est simplement une version abstraite du travail déjà effectué dans le cas du système de Stokes. On commence on se ramene au cas $G = 0$ en écrivant $u = u_0 + B^{-1}G$ avec $u_0 \in V$ solution de

$$a(u_0, v) = L(v) - a(B^{-1}G, v),$$

pour tout $v \in V$. En utilisant la coercivité de a sur l'espace V, on trouve une solution unique u_0 qui vérifie l'inégalité

$$\|u_0\|_X \leq \frac{1}{\alpha}(\|L\|_{X'} + \frac{C_a}{\beta}\|G\|_M),$$

d'où l'estimation pour u. Comme $L - Au$ est dans V^0 on en déduit l'existence et l'unicité de p, tel que $Bp = L - Au$ et on peut écrire

$$\|p\|_M \leq \frac{1}{\beta}\|L - Au\|_{X'} \leq \frac{1}{\beta}(\|L\|_{X'} + C_a\|u\|_X)$$

d'où l'estimation pour p. $\diamond$

Remarque 4.2.2 *Il est suffisant de supposer a est coercive sur V plutôt que sur tout X.*

Remarque 4.2.3 *La condition LBB est en réalité une condition nécessaire à l'existence, l'unicité et la stabilité d'une solution du problème (4.54). Pour s'en convaincre, regardons le cas où il existe un $q \neq 0$ dans M tel que $b(v,q) = 0$ pour tout $v \in X$. On voit dans ce cas que si (u,p) est solution alors $(u, p + tq)$ est aussi solution pour tout $t \in \mathbb{R}$ ce qui contredit l'unicité. Plus généralement si pour tout $\varepsilon > 0$, il existe $p_\varepsilon \in M$ tel que pour tout $v \in X$,*

$$b(v, p_\varepsilon) \leq \varepsilon\|v\|_X\|p_\varepsilon\|_M,$$

alors on peut montrer que la solution ne peut pas dépendre de manière stable des données.

4.3 Approximation interne des problèmes mixtes

Afin de discrétiser le problème, on introduit des sous-espaces de dimension finie $M_h \subset M$ et $X_h \subset X$, et on applique la méthode de Galerkin : trouver $(u_h, p_h) \in X_h \times M_h$ tel que

$$\begin{cases} a(u_h, v_h) + b(v_h, p_h) & = L(v_h) \\ b(u_h, q_h) & = G(q_h) \end{cases} \tag{4.57}$$

pour tout $(v_h, q_h) \in X_h \times M_h$.

Si on introduit une base (φ_γ) pour X_h et une base (ψ_λ) pour M_h, on voit que la recherche de la solution discrete équivaut à la résolution du système

$$\begin{cases} A_h U_h + B_h^t P_h & = L_h \\ B_h U_h & = G_h, \end{cases} \tag{4.58}$$

avec U_h et P_h les vecteur de coordonées de u et p, L_h et G_h les vecteurs de coordonées $L(\varphi_\gamma)$ et $G(\psi_\lambda)$, A_h la matrice de rigidité de a dans la base (φ_γ), et B_h la matrice de coefficients $B_{\lambda,\gamma} := b(\varphi_\gamma, \psi_\lambda)$.

Une première remarque importante est contrairement aux problèmes elliptiques qui entrent dans le cadre du théorème de Lax-Milgram, le caractère bien posé du problème discrèt (4.57) *ne découle pas naturellement* du caractère bien posé du problème continu (4.54). Ceci suggère la définition suivante.

Définition 4.3.1 *La forme b vérifie la condition LBB discrète pour les espaces (X_h, M_h) si et seulement si il existe $\beta_h > 0$ tel que*

$$\inf_{p_h \in M_h} \sup_{v_h \in X_h} \frac{b(v_h, p_h)}{\|v_h\|_X \|p_h\|_M} \geq \beta_h. \tag{4.59}$$

On a bien entendu un théorème est l'analogue du résultat obtenu pour le problèm continu.

Théorème 4.3.1 *Sous la condition LBB discrète, le problème (4.57) admet une unique solution $(u_h, p)_h \in X \times M$ qui satisfait les estimations*

$$\|u_h\|_X \leq \frac{1}{\alpha}\|L\|_{X'} + \frac{1}{\beta_h}(1 + \frac{C_a}{\alpha})\|G\|_{M'},$$

et

$$\|p_h\|_M \leq \frac{1}{\beta}(1 + \frac{C_a}{\alpha})\|L\|_{X'} + \frac{C_a}{\beta_h^2}(1 + \frac{C_a}{\alpha})\|G\|_{M'},$$

où α et C_a sont les constantes d'ellipticité et de continuité de a.

Notons que la condition LBB discrète signifie en particulier que l'on peut toujours trouver une solution à $B_hU_h = G_h$ pour tout G_h, c'est à dire que B_h est surjective, ce qui impose $\dim(X_h) \geq \dim(Mh)$. Cependant il peut arriver que (4.59) ne soit pas vérifiée même pour des espaces vérifiant cette contrainte de dimension. Dans ce cas, il existe des $q_h \in M_h$ tels que $b(v_h, q_h) = 0$ pour tout $v_h \in X_h$ ce qui signifie que si p_h est solution alors $p_h + tq_h$ l'est aussi pour tout $t \in \mathbb{R}$. On dit parfois que q_h est un *mode parasite*. Dans la dernière section de ce chapitre nous présentons des exemples de paires d'espaces d'éléments finis qui vérifient (4.59) dans cas du problème de Stokes.

Nous allons à présent utiliser la condition LBB discrète afin d'établir un résultat qui est l'analogue pour les problèmes mixtes du Lemme de Cea.

Théorème 4.3.2 *Sous la condition LBB discrète, la solution du problème (4.57) vérifie l'estimation*

$$\|u - u_h\|_X + \|p - p_h\|_M \leq C(\inf_{v_h \in X_h} \|u - v_h\|_X + \inf_{q_h \in M_h} \|p - q_h\|_M)$$

où la constante C ne dépend que de (α, β_h) et des constantes de continuité (C_a, C_b) de a et b.

Preuve : elle est technique du fait de la présence du second membre G et il peut être utile à la compréhension de l'effectuer dans le cas plus simple où $G = 0$. On travaille d'abord sur l'estimation de $u - u_h$. On note V_h l'ensemble des éléments de X_h qui vérifient

$$b(v_h, q_h) = 0, \;\; q_h \in M_h,$$

(c'est à dire les vitesses à divergence discrète nulle dans le cas du problème de Stokes). On note aussi V_h^G l'ensemble des éléments de X_h qui vérifient la deuxième équation de (4.57), c'est à dire

$$b(v_h, q_h) = G(q_h), \;\; q_h \in M_h.$$

Notons que $u_h \in V_h^G$. D'après la remarque précèdente, la condition LBB discrète entraine que cet ensemble est non-nul. Plus précisément, pour tout $v_h \in X_h$, on peut le corriger en lui ajoutant $z_h \in X_h$ solution de

$$b(z_h, q_h) = G(q_h) - b(v_h, q_h) = b(u - v_h, q_h), \; q_h \in M_h,$$

et on obtient ainsi $w_h = v_h + z_h \in V_h^G$. De plus, en utilisant les mêmes remarques que dans le cas continu, il existe une unique solution $z_h \in V_h^\perp$ et qui vérifie

$$\|z_h\|_X \leq \frac{1}{\beta_h} C_b \|u - v_h\|_X.$$

On en déduit que

$$\|u - w_h\|_X \leq \|u - v_h\|_X + \|z_h\|_X \leq (1 + \frac{C_b}{\beta_h})\|u - v_h\|_X. \tag{4.60}$$

On note à présent que pour tout $w_h \in V_h^G$ et pour tout $q_h \in M_h$, on a

$$\begin{aligned}
\alpha\|u_h - w_h\|_X^2 &\leq a(u_h - w_h, u_h - w_h) \\
&= a(u_h - u, u_h - w_h) + a(u - w_h, u_h - w_h) \\
&= b(u_h - w_h, p - p_h) + a(u - w_h, u_h - w_h) \\
&= b(u_h - w_h, p - q_h) + a(u - w_h, u_h - w_h) \\
&\leq C_b\|u_h - w_h\|_X\|p - q_h\|_M + C_a\|u - w_h\|_X\|u_h - w_h\|_X,
\end{aligned}$$

où on a utilisé le fait que $b(u_h - w_h, q_h) = 0$ pour tout $q_h \in M_h$. En utilisant l'inégalité triangulaire, on a donc

$$\|u - u_h\|_X \leq \frac{C_b}{\alpha} \inf_{q_h \in M_h} \|p - q_h\|_M + (1 + \frac{C_a}{\alpha}) \inf_{w_h \in V_h^G} \|u - w_h\|_X.$$

En combinant ceci avec (4.60) on obtient l'estimation

$$\|u - u_h\|_X \leq \frac{C_b}{\alpha} \inf_{q_h \in M_h} \|p - q_h\|_M + (1 + \frac{C_b}{\beta_h})(1 + \frac{C_a}{\alpha}) \inf_{v_h \in X_h} \|u - v_h\|_X. \tag{4.61}$$

Pour établir l'estimation sur $p - p_h$, on écrit que pour tout $q_h \in M_h$ et pour tout $v_h \in X_h$,

$$b(v_h, q_h - p_h) = b(v_h, q_h - p) + b(v_h, p - p_h) = b(v_h, q_h - p) - a(u - u_h, v_h).$$

En utilisant la condition LBB discrète, on en déduit

$$\begin{array}{rl} \|q_h - p_h\|_X & \le \frac{1}{\beta_h} \sup_{v_h \in X_h} \frac{b(v_h, q_h - p_h)}{\|v_h\|_X} \\ & = \frac{1}{\beta_h} \sup_{v_h \in X_h} \frac{b(v_h, q_h - p) - a(u - u_h, v_h)}{\|v_h\|_X} \\ & \le \frac{1}{\beta_h} (C_b \|q_h - p\|_X + C_a \|u - u_h\|_X). \end{array}$$

En utilisant l'inégalité triangulaire, on a donc

$$\|p - p_h\|_M \ \le (1 + \frac{C_b}{\beta_h}) \inf_{q_h \in M_h} \|p - q_h\|_M + \frac{C_a}{\alpha} \|u - u_h\|_X.$$

En combinant avec l'estimation (4.61) on obtient une estimation pour $p - p_h$ ce qui conclut la preuve du théorème. $\diamond$

Remarque 4.3.1 *La preuve du théorème nous montre que la constante C dans l'estimation se dégrade lorsque β_h tend vers 0. Il est donc important de construire des paires (X_h, M_h) qui vérifient la condition LBB discrète et telles que $\beta_h \ge \beta^* > 0$ indépendemment de h. Nous allons exhiber de telles paires dans le cas spécifique du problème de Stokes.*

4.4 Eléments finis pour le problème de Stokes

Dans le cas particulier du problème de Stokes, la condition LBB discrète prend la forme :

$$\inf_{q_h \in M_h} \sup_{v_h \in X_h} \frac{\int_\Omega q_h \mathrm{div}(v_h)}{\|v_h\|_{H_0^1} \|q_h\|_{L^2}} \ge \beta_h. \tag{4.62}$$

La mise au point de paires d'espaces vitesse-pression (X_h, M_h) satisfaisant (4.62) avec $\beta_h \ge \beta^* > 0$ pour tout $h > 0$ a constitué un programme de recherche important en analyse numérique.

Un premier exemple négatif en dimension $d = 2$ nous permet de comprendre que cette condition ne va pas de soi même lorsque $\dim(X_h) > \dim(M_h)$. On prend ici un domaine carré Ω et une partition $\mathcal{T}_h$ en carrés (ou rectangles) conforme. On prend les vitesses de type Q_1

$$X_h := \{v = (v_1, v_2) \ ; \ v_i \in H_0^1(\Omega), \ v_{i|T} \in Q_1, \ T \in \mathcal{T}_h\},$$

et les pression constante par morceaux

$$M_h := \{q \in L_0^2(\Omega) \ ; \ q_{|T} \in \Pi_0, \ T \in \mathcal{T}_h\}.$$

Il est facile de vérifier que la dimension de X_h est de l'ordre de deux fois celle de M_h ce qui autorise en théorie la condition LBB. Cependant, on peut exhiber un mode parasite de type "oscillant", en prenant une pression p_h qui vaut alternativement 1 et -1 sur chaque carré, avec un changement de signe entre deux carrés adjacents. Afin de le prouver, on remarque que pour chaque noeud interne au domaine $\gamma \in \Gamma_{h,0}$, on a deux fonctions de bases pour X_h, qui sont

$$\varphi_{\gamma,1} = (\varphi_\gamma, 0) \ \text{ et } \ \varphi_{\gamma,2} = (0, \varphi_\gamma),$$

où φ_γ est la fonction de base scalaire pour le noeud γ. Pour tout $\gamma \in \Gamma_{0,h}$ et $i = 1, 2$, on peut écrire en appliquant la formule de Green

$$\int_\Omega p_h \mathrm{div}(\varphi_{\gamma,i}) = 2 \sum_{j=1}^{4} \int_{e_j} \varphi_{\gamma,i} \cdot n_j,$$

où $(e_1, \cdots, e_4)$ sont les côtés qui touchent γ et n_j est la normale orientée dans le sens de la valeur 1 vers la valeur -1 pour p_h. On voit que $\varphi_{\gamma,i} \cdot n_j$ est nul pour les deux j tels que e_j est dans la direction i, et que d'autre part les deux autres termes s'annullent entre eux car $\varphi_{\gamma,i}$ est affine et a la même intégrale sur chaque côté e_j. Par conséquent

$$\int_\Omega p_h \mathrm{div}(u_h) = 0,$$

pour tout $v_h \in X_h$.

On peut exhiber des contres-exemples du même type dans le cas des éléments finis triangulaires si on prend des vitesses $\mathbb{P}_1$ et des pressions $\mathbb{P}_0$, (qui est pourtant le choix d'espaces qui vient le plus naturellement à l'esprit). Le *lemme de Fortin* donne un critère dans le cadre des problèmes mixtes abstraits pour la vérification de la condition LBB discrète.

Lemme 4.4.1 *Si le problème (4.54) vérifie la condition LBB continue, alors les espaces* (X_h, M_h) *vérifient la condition LBB discrète avec une constante* β^* *indépendante de* $h > 0$, *si et seulement si il existe un opérateur de "restriction"* $R_h \in \mathcal{L}(X, X_h)$ *et une constante* C *indépendante de* h *tels que*

$$\|R_h v\|_X \leq C\|v\|_X, \ \ v \in X,$$

et

$$b(R_h v - v, q_h) = 0, \ \ v \in X, \ q_h \in M_h$$

Preuve : si le critère est satisfait, on peut écrire pour tout $q_h \in M_h$

$$\begin{aligned} \sup_{v_h \in X_h} \frac{b(v_h, q_h)}{\|v_h\|_X} &\geq \sup_{v \in X} \frac{b(R_h v, q_h)}{\|R_h v\|_X} \\ &\geq \frac{1}{C} \sup_{v \in X} \frac{b(v, q_h)}{\|v\|_X} \\ &\geq \frac{\beta}{C}\|q_h\|_M, \end{aligned}$$

où on a utilisé les propriétés de R_h et la condition LBB continue. La condition LBB discrète est donc satisfaite avec $\beta_h \geq \beta^* := C/\beta$.

Réciproquement, si la condition LBB discrète est satisfaite, alors pour tout $v \in X$, il existe un unique $v_h = R_h v \in (V_h)^\perp$ tel que

$$b(R_h v, q_h) = b(v, q_h), \ \ q_h \in M_h. \tag{4.63}$$

L'application $v \mapsto R_h v$ est linéaire de X dans X_h et on a

$$\|R_h v\|_X \leq \frac{C_b}{\beta^*}\|v\|_X.$$

L'opérateur R_h possède donc les propriétés voulues. $\diamond$

Notons que l'opérateur R_h préserve en quelque sorte la propriété de divergence nulle, puisque si $u \in V$ alors $R_h u \in V_h$, i.e. la divergence discrète de $R_h u$ est nulle. La preuve du lemme de Fortin nous suggère une manière de vérifier la condition LBB discrète : construire l'opérateur R_h par la résolution du système lié à l'équation (4.63) et vérifier que celui-ci est borné dans X. Cette approche est souvent difficile à mettre en oeuvre car l'opérateur obtenu est non-local et d'expression compliquée. Il est possible de construire un opérateur R_h plus simple pour des paires spécifiques d'éléments finis et de montrer ainsi que ces paires satisfont la condition LBB discrète. Nous présentons ici trois exemples de telles paires en dimension $d = 2$ (Certains d'entre eux se généralisent sans difficulté à la dimension $d > 2$).

Exemple 1 : Elément "mini" $\mathbb{P}_1$ et $\mathbb{P}_1$ bulle.
Etant donné une triangulation $\mathcal{T}_h$, on prend pour M_h l'espace des éléments $\mathbb{P}_1$ sur $\mathcal{T}_h$, i.e.

$$M_h := \{q \in \mathcal{C}(\Omega) \ ; \ q_{|T} \in \Pi_1, \ T \in \mathcal{T}_h\}.$$

Pour l'espaces X_h, on considère l'espace des éléments finis $\mathbb{P}_1$ sur la même triangulation

$$E_h := \{v = (v_1, v_2) \in (H_0^1(\Omega))^2 \ ; \ v_{j_{|T}} \in \Pi_1, \ T \in \mathcal{T}_h, \ j = 1, 2\}.$$

et on prend

$$X_h = E_h \oplus \mathrm{Vect}\{(b_T, 0), (0, b_T), \ T \in \mathcal{T}_H\},$$

où $b_T = \lambda_0\lambda_1\lambda_2$ est la fonction bulle. Il est immédiat que $X_h \subset (H_0^1(\Omega))^2$.

On définit à présent l'opérateur R_h sur chaque triangle T par

$$R_T v = A_h v + \alpha_T b_T,$$

où $A_h v := (A_h v_1, A_h v_2)$ est l'opérateur de Clément introduit dans la section 3.2, et $\alpha_T \in \mathbb{R}^2$ est choisi de manière à avoir

$$\int_T (R_T v - v) = 0.$$

Pour cela, il suffit de prendre

$$\alpha_T = (\int_T b_T)^{-1} \int_T (v - A_h v).$$

Pour tout $v \in X$ et $q_h \in M_h$, on peut alors écrire

$$\int_\Omega q_h \mathrm{div}(v - R_h v) = \int_\Omega \nabla q_h (R_h v - v) = \sum_{T\in\mathcal{T}_h} \int_T \nabla q_h (R_h v - v) = 0,$$

puique ∇q_h est constant sur T. Il nous reste à prouver que l'opérateur R_h ainsi défini est continu de X dans X_h pour la norme X.

Théorème 4.4.1 *Si la famille $(\mathcal{T}_h)_{h>0}$ est régulière, il existe $C > 0$ indépendante de h, telle que*

$$\|R_h v\|_{H_0^1} \le C\|v\|_{H_0^1},$$

pour tout $v \in X$.

Preuve : On traite séparément les parties $A_h v$ et $\alpha_T b_T$. Pour $A_h v$, lorsque le triangle T n'a pas de sommet sur le bord, on écrit pour chaque composante $w = v_1, v_2$

$$\begin{aligned} |A_h w|_{H^1(T)} &= |\textstyle\sum_{\gamma=a_0,a_1,a_2} m_\gamma \varphi_\gamma|_{H_1(T)} \\ &= |\textstyle\sum_{\gamma=a_0,a_1,a_2} (m_\gamma - m_T)\varphi_\gamma|_{H_1(T)} \\ &\le \textstyle\sum_{\gamma=a_0,a_1,a_2} |m_\gamma - m_T|\,|\varphi_\gamma|_{H_1(T)} \\ &\le C\textstyle\sum_{\gamma=a_0,a_1,a_2} |m_\gamma - m_T|. \end{aligned}$$

avec $m_\gamma := |K_\gamma|^{-1}\int_{K_\gamma} w$ et $m_T := |T|^{-1}\int_T w$, et où on a utilisé un passage à l'élément de référence pour établir $|\varphi_\gamma|_{H_1(T)} \le C$ avec C indépendante de T et h. On écrit ensuite

$$\begin{aligned} |m_\gamma - m_T| &= ||K_\gamma|^{-1}\textstyle\int_{K_\gamma}(w - m_\gamma) - |T|^{-1}\int_T (w - m_\gamma)| \\ &\le |K_\gamma|^{-1/2}\|w - m_\gamma\|_{L^2(K_\gamma)} + |T|^{-1/2}\|w - m_\gamma\|_{L^2(T)} \\ &\le C h_T^{-1}\|w - m_\gamma\|_{L^2(K_\gamma)}. \end{aligned}$$

En utilisant l'estimation $\|w - m_\gamma\|_{L^2(K_\gamma)} \le C|w|_{H^1(K_\gamma)}$ établie dans la preuve du théorème 3.2.1. Nous avons ainsi établi

$$|A_h v|_{H^1(T)} \le C|v|_{H^1(K_T)},$$

avec C indépendante de T et h. Dans le cas où T a un sommet sur le bord, on obtient la même estimation du même type que dans la preuve du théorème 3.2.1 en faisant intervenir l'inégalité de Poincaré.

Pour la partie $\alpha_T b_T$, on obtient par changement d'échelle

$$|\alpha_T b_T|_{H^1(T)} \le C|\alpha_T|,$$

avec C indépendant de T et h. On écrit ensuite

$$\begin{aligned} |\alpha_T| &= (\textstyle\int_T b_T)^{-1}|\int_T (v - A_h v)| \\ &\le C|T|^{-1}|\textstyle\int_T (v - A_h v)| \\ &\le C|T|^{-1/2}\|v - A_h v\|_{L^2(T)}. \end{aligned}$$

En utilisant l'estimation $\|v - A_h v\|_{L^2(T)} \leq C h_T |v|_{H^1(K_T)}$ établie dans le théorème 3.2.1, on obtient ainsi

$$|\alpha_T| \leq C|v|_{H^1(K_T)},$$

On a finalement

$$|R_T v|_{H^1(T)} \leq C|v|_{H^1(K_T)}.$$

En élevant au carré et en sommant sur tous les T on conclut la preuve. $\diamond$

Exemple 2 : Elément $\mathbb{P}_0$ et $\mathbb{P}_2$ incomplet.
Etant donné une triangulation $\mathcal{T}_h$, on prend pour M_h l'espace des pressions constantes par morceaux

$$M_h := \{q \in L^2_0(\Omega) \; ; \; q_{|T} \in \Pi_0, \; T \in \mathcal{T}_h\},$$

et pour X_h un espace intermédiaire entre les éléments finis $\mathbb{P}_1$ et $\mathbb{P}_2$ sur la même triangulation. Localement cet espace est défini par :

$$X_T := (\Pi_1)^2 \oplus \mathrm{Vect}\{p_0, p_1, p_2\},$$

avec

$$p_0 := \lambda_1\lambda_2 n_0, \; p_1 := \lambda_0\lambda_2 n_1, \; p_2 := \lambda_0\lambda_1 n_2,$$

où n_i est la normale extérieure au côté e_i opposé au sommet a_i. On peut vérifier que X_T est l'espace des $v \in (\Pi_2)^2$ tels que $v \cdot \tau_i\_|e_i in \Pi_1$ pour $i = 0, 1, 2$ où τ_i est le vecteur tangent à e_i. On peut prendre comme degrés de liberté les six formes linéaires

$$v = (v_1, v2) \mapsto v_j(a_i), \quad j = 1, 2, \; i = 0, 1, 2,$$

qui caractérise la partie Π_1 de $v \in X_T$, auxquelles on adjoint trois formes lináires associées aux côtés e_i et définies par

$$\psi_i(v) = 4(v \cdot n_i)(b_i) - 2[(v \cdot n_i)(a_j) + (v \cdot n_i)(a_k)],$$

où $\{j, k\} = \{0, 1, 2\} - \{i\}$ et b_i est le milieu de côté e_i. Avec ces choix, on remarque que la base de X_T associée est donnée par les 9 fonctions $\{(\lambda_i, 0), (0, \lambda_i), p_i\}$ pour $i = 0, 1, 2$. On vérifie aussi facilement que les fonctions de l'espace d'éléments finis X_h se raccordent continuement aux interfaces entre les triangles et qu'on a donc $X_h \subset (H^1_0(\Omega))^2$ en imposant la nullité des degrés de libertés situés au bord du domaine.

On définit à présent l'opérateur R_h sur chaque triangle T par

$$R_T v = A_h v + \sum_{i=0,1,2} \alpha_i p_i,$$

où $A_h v = (A_h v_1, A_h v_2)$ est à nouveau l'opérateur de Clément, et les coefficients α_i sont choisis de manière à avoir

$$\int_T \mathrm{div}(R_h v - v) = 0.$$

On vérifie aisément qu'il suffit pour cela de prendre

$$\alpha_i := (\int_{e_i} \lambda_j \lambda_k)^{-1} \int_{e_i} (v - A_h v) \cdot n_i, \quad i = 0, 1, 2,$$

avec $\{j, k\} := \{0, 1, 2\} - \{i\}$. En effet ceci nous assure que

$$\int_{e_i} R_h v - v = 0, \; i = 0, 1, 2,$$

et par conséquent

$$\int_T \mathrm{div}(R_h v - v) = \int_{\partial T} (R_h v - v) \cdot n = 0.$$

D'autre part, comme les degrés de liberté α_i uniquement déterminés pour chaque côté, la fonction $R_h v$ obtenue en recollant les morceaux $R_T v$ est bien dans l'espace X_h, et vérifie

$$\int_\Omega q_h \mathrm{div}(R_h v - v) = 0,$$

pour tout $q_h \in M_h$. Il nous reste à prouver que l'opérateur R_h ainsi défini est continu de X dans X_h pour la norme X.

Théorème 4.4.2 *Si la famille $(\mathcal{T}_h)_{h>0}$ est régulière, il existe $C > 0$ indépendante de h, telle que*

$$\|R_h v\|_{H_0^1} \leq C\|v\|_{H_0^1},$$

pour tout $v \in X$.

Preuve : On traite séparément les parties $A_h v$ et $\alpha_i p_i$. Pour $A_h v$, nous avons déjà établi

$$|A_h v|_{H^1(T)} \leq C|v|_{H^1(K_T)},$$

dans la preuve du théorème de l'exemple précédent, avec C indépendante de T et h. Pour les termes $\alpha_i p_i$, on écrit

$$\begin{aligned} |\alpha_i p_i|_{H^1(T)} &\leq C|\alpha_i| \\ &\leq C h_{e_i}^{-1} |\textstyle\int_{e_i} (v - A_h v) \cdot n_i| \\ &\leq C h_{e_i}^{-1/2} \|v - A_h v\|_{L^2(e_i)}. \end{aligned}$$

En utilisant l'estimation $\|v - A_h v\|_{L^2(e_i)} \leq C h_{e_i}^{1/2} |v|_{H^1(K_{e_i})}$ établie dans le théorème 3.2.1, on obtient ainsi

$$|\sum_{i=0,1,2} \alpha_i p_i|_{H^1(T)} \leq C|v|_{H^1(K_T)},$$

soit finalement

$$|R_T v|_{H^1(T)} \leq C|v|_{H^1(K_T)}.$$

En élevant au carré et en sommant sur tous les T on conclut la preuve. $\diamond$

Exemple 3 : Elément $\mathbb{P}_0$ et $\mathbb{P}_1$ raffiné.

Etant donné une triangulation $\mathcal{T}_h$, on prend pour M_h l'espace des pressions constantes par morceaux

$$M_h := \{q \in L_0^2(\Omega) \; ; \; q_{|T} \in \Pi_0, \; T \in \mathcal{T}_h\},$$

et pour X_h l'espace des éléments finis $\mathbb{P}_1$ sur la triangulation $\mathcal{T}_{h/2}$ obtenu par un raffinement de chaque triangle de $\mathcal{T}_h$ en quatre sous-triangles par la règle du point milieu. On a donc

$$X_h := \{v = (v_1, v_2) \in (H_0^1(\Omega))^2 \; ; \; v_{j\,|T} \in \Pi_1, \; T \in \mathcal{T}_{h/2}, \; j = 1, 2\}.$$

On définit à présent l'opérateur R_h sur chaque triangle $T \in \mathcal{T}_h$ par

$$R_T v = A_h v + \sum_{i=0,1,2} \alpha_i \varphi_i,$$

où $A_h v = (A_h v_1, A_h v_2)$ est l'opérateur de Clément pour la triangulation $\mathcal{T}_h$, les φ_i sont les fonctions de bases de l'espace d'élément fini $\mathbb{P}_1$ sur la triangulation $\mathcal{T}_{h/2}$ associées aux milieux b_i des côtés e_i de T, et les vecteurs α_i sont définis par

$$\alpha_i := (\int_{e_i} \varphi_i)^{-1} \int_{e_i} (v - A_h v), \quad i = 0, 1, 2,$$

de sorte que l'on a

$$\int_{e_i} (R_T v - v) = 0, \quad i = 0, 1, 2.$$

On a par conséquent

$$\int_T \mathrm{div}(R_T v - v) = 0,$$

pour tout $T \in \mathcal{T}_h$ soit

$$\int_\Omega q_h \mathrm{div}(R_T v - v) = 0,$$

pour tout $q_h \in M_h$. Il nous reste à prouver que l'opérateur R_h ainsi défini est continu de X dans X_h pour la norme X.

Théorème 4.4.3 *Si la famille* $(\mathcal{T}_h)_{h>0}$ *est régulière, il existe* $C > 0$ *indépendante de* h*, telle que*

$$\|R_h v\|_{H_0^1} \leq C\|v\|_{H_0^1},$$

pour tout $v \in X$*.*

Preuve : On traite séparément les parties $A_h v$ et $\alpha_i \varphi_i$. Pour $A_h v$, nous avons déjà établi

$$|A_h v|_{H^1(T)} \leq C|v|_{H^1(K_T)},$$

dans la preuve du théorème de l'exemple précédent, avec C indépendante de T et h. Pour les termes $\alpha_i \varphi_i$, on écrit

$$\begin{aligned} |\alpha_i \varphi_i|_{H^1(T)} &\leq C|\alpha_i| \\ &\leq C h_{e_i}^{-1} \left| \int_{e_i} (v - A_h v) \right| \\ &\leq C h_{e_i}^{-1/2} \|v - A_h v\|_{L^2(e_i)}, \end{aligned}$$

et on conclut comme pour l'élément $\mathbb{P}_0$ et $\mathbb{P}_2$ incomplet. $\diamond$

Quelques références bibliographiques

R.A. Adams, “Sobolev spaces”, Academic Press (1975).

C. Bernardi, Y. Maday et F. Rapetti, “Discrétisations variationnelles des problèmes elliptiques”, Springer-SMAI (2004)

S.C. Brenner and L.R. Scott, “The mathematical theory of finite element methods”, Springer (2002)

H. Brézis, “Analyse Fonctionelle - théorie et applications”, Dunod (1983)

P.G. Ciarlet, “The finite element method for elliptic problems”, North Holland (1978)

V. Girault and P.A. Raviart, “Finite element methods for Navier-Stokes equations”, Springer (1986)

P. Grisvard, “Elliptic problems in non-smooth domains”, Pitman (1985)