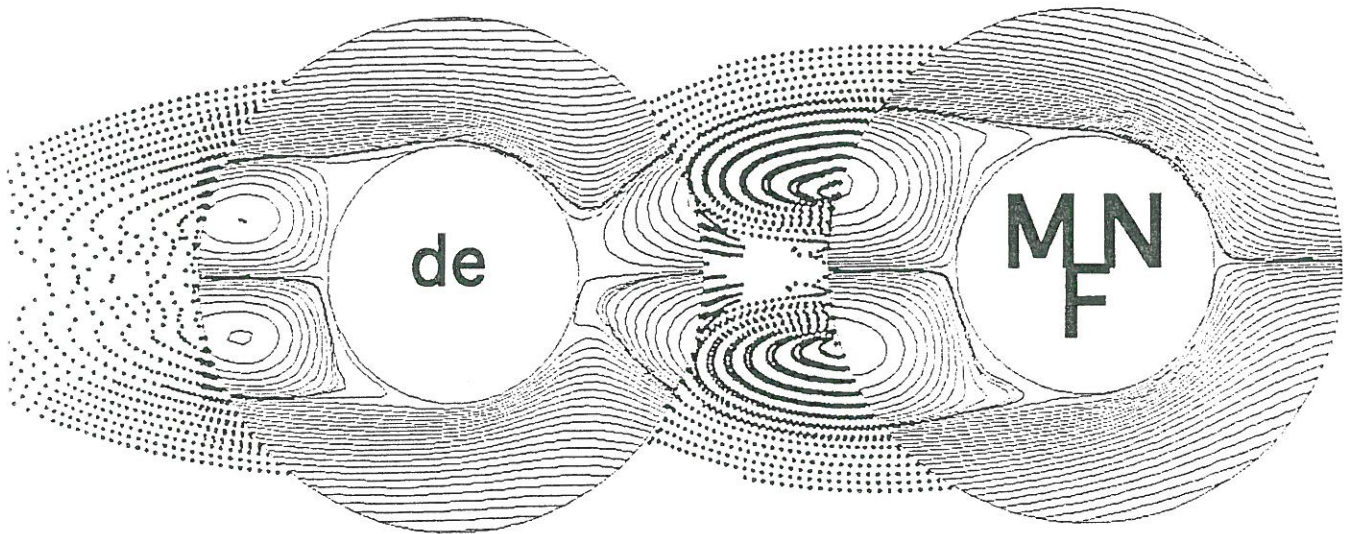


C0494

Ecole de Printemps



Mécanique des Fluides Numérique

Carcans-Maubuisson
9-15 Mai 1993

C. N. R. S. LIMSI	N° INVENTAIRE - 8/06/1993
SERVICE DOCUMENTATION	N° entrée - C0494

INTRODUCTION A L'APPROXIMATION
NUMERIQUE DES EQUATIONS
DE NAVIER-STOKES

Jean-Luc GUERMOND

Ecole organisée avec le soutien du CNRS et de la DRET
Comité d'organisation: O. Daube, P. Le Quéré, S. Candel

Table des matières

1	Introduction-Généralités	3
1.1	Introduction	3
1.1.1	Régularité du domaine fluide	3
1.1.2	Les équations	3
1.1.3	Compatibilité, existence, unicité	4
1.2	Quelques problèmes types	5
1.2.1	La diffusion	5
1.2.2	La convection	6
1.2.3	L'incompressibilité	7
1.2.4	L'instationnarité	8
1.2.5	La non-linéarité de la convection	9
2	Diffusion – Formulations Fortes – Différences Finies	11
2.1	Etude de la formulation forte	11
2.1.1	Définitions et généralités sur les problèmes elliptiques	11
2.1.2	Les conditions aux limites	12
2.1.3	Formules de Green-Ostrogradski	13
2.2	L'équation de la neutronique (ou de "Helmholtz")	14
2.2.1	Principe du maximum	14
2.2.2	Unicité et stabilité pour le problème de Dirichlet	15
2.2.3	Problèmes de Neumann - Robin - Mêlés	16
2.3	Approximation par différences finies	17
2.3.1	Introduction	17
2.3.2	Mise en œuvre	17
2.3.3	Existence et unicité de la solution approchée	18
2.4	Consistance - Stabilité - Convergence	19
2.4.1	Définitions des espaces fonctionnels	19
2.4.2	Définition d'un cadre abstrait	20
2.4.3	Consistance	21
2.4.4	Stabilité	21
2.4.5	Convergence	23
2.5	Commentaires	23
3	Diffusion – Formulations Variationnelles	25
3.1	Introduction	25
3.2	Cadre fonctionnel des formulations variationnelles	26
3.2.1	Espaces de Sobolev	26
3.2.2	Les problèmes de trace	27

3.2.3	Rappels d'analyse Hilbertienne	29
3.3	Exemples de problèmes variationnels	32
3.3.1	Problème de Dirichlet	32
3.3.2	Problèmes de transmission	34
3.3.3	Problème mêlé	35
3.3.4	Régularité des solutions variationnelles	36
3.4	Approximation interne	36
3.4.1	Le problème discret	37
3.4.2	Convergence de l'approximation interne	38
3.4.3	Approximation de Galerkin – Approximation spectrale	40
3.4.4	Approximation par éléments finis	42
4	Incompressibilité – Approximation – Stabilité	43
4.1	Le problème de Stokes	43
4.1.1	Formulation du problème continu	43
4.1.2	Formulation discrète du problème de Stokes	44
4.1.3	Existence-Unicité (Condition “inf-sup”)	46
4.2	Exemples, contre-exemples	48
4.2.1	Instabilité en damier	48
4.2.2	L'élément P_1/P_1 -bulle	50
4.2.3	L'élément de Taylor–Hood: P_1/P_2	50
4.3	Stabilité – Convergence	51
4.3.1	Stabilité	51
4.3.2	Convergence	52
4.4	Résolution du système linéaire	53
4.4.1	Introduction	53
4.4.2	Compressibilité artificielle	53
4.4.3	Opérateur d'Uzawa	53
4.5	Approximation par différences finies	54

Chapitre 1

Généralités sur les équations de Navier-Stokes

1.1 Introduction

Ce chapitre est consacré à la présentation des équations de la mécanique des fluides newtoniens incompressibles (autrement dit les équations de Navier-Stokes incompressibles) et aux problèmes mathématiques et numériques qui leurs sont spécifiques.

1.1.1 Régularité du domaine fluide

Considérons Ω le domaine physique occupé par le fluide. Ce domaine est supposé ouvert dans \mathbb{R}^d (en pratique $d = 2$ ou 3) et connexe, c'est à dire d'un seul morceau. Pour simplifier l'exposé on s'intéresse au cas où Ω est borné. Le cas des domaines non bornés présente des difficultés supplémentaires (surmontables) liées au comportement à l'infini des solutions.

Encore dans un souci de simplification on supposera toujours que Ω possède la régularité qu'on lui demandera. On ne se posera pas trop de questions sur la forme des bords de Ω . Pour fixer les idées, disons qu'un ouvert de classe C^1 par morceaux et localement d'un seul côté de sa frontière satisfait à la plupart des exigences de régularité. Les ouverts pathologiques sont ceux avec des arrêtes effilées (en lames de rasoir) ou avec des fissures ou des pointes effilées (on ne s'intéresse pas aux écoulements autour des hérissons). En fait les domaines fluides sympathiques sont de classe C^m avec $m \geq 2$.

1.1.2 Les équations

Pour les fluides newtoniens (c'est-à-dire régis par des lois de comportement linéaires) incompressibles ou faiblement compressibles les inconnues sont le champ de vitesse et le champ de pression. Le bilan de masse et de quantité de mouvement local s'expriment par:

$$(1.1) \quad \begin{cases} \frac{\partial u}{\partial t} + u \cdot \nabla u - \nu \nabla^2 u + \nabla p = f \\ \nabla \cdot u = g \end{cases}$$

où f est une force volumique externe et g un terme de flux de masse. Le cas purement incompressible correspond à $g = 0$. Lorsque les équations sont dimensionnées le paramètre ν est la viscosité cinématique. Si les équations sont adimensionnées l'inverse du paramètre ν est appelé nombre de Reynolds et on note $Re = 1/\nu$.

Dans le cadre des écoulements incompressibles ou faiblement compressibles l'équation de bilan d'énergie est généralement découplée des bilans de masse et de quantité de mouvement. En fait le présent modèle, connu sous le nom d'équations de Navier-Stokes, peut admettre un couplage faible avec le bilan d'énergie par l'intermédiaire des termes source f et g . Par la suite f et g sont supposées être des données du problème.

Pour avoir quelques chances d'avoir un problème bien posé (existence et unicité d'une solution dans un cadre à définir) il faut imposer des conditions initiales et des conditions aux limites. Un premier type de conditions aux limites consiste à imposer la valeur de la vitesse sur la paroi $\partial\Omega$ ou seulement une portion $\partial\Omega_1$ de celle-ci. Ce type de condition est appelé condition de Dirichlet:

$$(1.2) \quad u(t, x) = a(t, x), \quad \forall x \in \partial\Omega_1$$

La vitesse a est généralement nulle dans un repère lié au bord $\partial\Omega$, c'est la condition d'adhérence du fluide à la paroi. On peut aussi imposer une condition de non friction sur une portion $\partial\Omega_2$ de $\partial\Omega$. Cette condition se met sous la forme:

$$(1.3) \quad \nu(\nabla u + \nabla^t u) \cdot n + pn = 0, \quad \text{sur } \partial\Omega_2$$

Cette condition peut être utilisée pour simuler une frontière artificielle comme une sortie de canal. Le problème des frontières artificielles est en soit un problème très difficile qui suscite actuellement de nombreuses recherches. Certains modèles de turbulence remplacent la condition d'adhérence à la paroi par une condition de glissement: $u \cdot n = 0$ sur une portion $\partial\Omega_3$ de $\partial\Omega$ et une condition imposée sur la composante tangentielle de la vitesse:

$$(1.4) \quad (au + b\partial u/\partial n) \times n = c, \quad \text{sur } \partial\Omega_3$$

où a , b et c sont des fonctions données. Pour simplifier l'analyse des modèles numériques qu'on introduira par la suite, on se placera généralement dans le cadre des conditions de Dirichlet homogènes.

Pour assurer l'unicité d'une solution au problème (1.1) il faut aussi imposer une condition initiale. Pour ce type de problème d'évolution cette condition est généralement du type

$$(1.5) \quad u(t = 0, x) = u_0(x), \quad \forall x \in \Omega$$

C'est-à-dire qu'on impose le champ de vitesse à l'instant initial.

1.1.3 Compatibilité, existence, unicité

Bien sûr les termes sources, les données aux frontières et la condition initiale ne peuvent pas être totalement indépendantes. Toutes ces données doivent être compatibles entre elles.

Par exemple, si on veut espérer une certaine continuité de la vitesse lorsque $t \rightarrow 0$ (ce qui paraît raisonnable), le champ de vitesse initial doit alors satisfaire la condition de compatibilité:

$$(1.6) \quad \nabla \cdot u_0 = g(t \rightarrow 0, x), \quad \forall x \in \Omega$$

De même, dans le cadre d'une donnée à la frontière du type Dirichlet, la donnée de frontière et la condition initiale doivent satisfaire la condition de compatibilité

$$(1.7) \quad u_0 \cdot n = a(t \rightarrow 0, x) \cdot n, \quad \forall x \in \partial\Omega$$

Remarquer que seules les composantes normales de la vitesse doivent être compatibles. On peut admettre une discontinuité de la vitesse tangentielle à l'instant initial sans mettre en

où $\Sigma^- = \{x \in \partial\Omega, u \cdot n < 0\}$ est la portion de frontière sur laquelle le champ de vitesse est entrant, la normale étant dirigée vers l'extérieur du domaine.

Cette équation est le prototype des problèmes hyperboliques du premier ordre. Pour cette classe de problème l'information se propage dans une seule direction. En fait, si $f = 0$ (pas de terme source) l'équation aux dérivées partielles se réduit à $u \cdot \nabla \phi = 0$, c'est-à-dire que le gradient de ϕ est nul dans la direction de la vitesse u . En d'autres termes cela signifie que ϕ est constant sur les lignes de champ de u et donc que la valeur de ϕ se propage dans le domaine Ω à partir de Σ^- en suivant les lignes de champ de u . D'après ce raisonnement on conçoit bien que le problème aura une solution unique si toutes les lignes de champ passent par Σ^- . Si ce n'est pas le cas, c'est-à-dire s'il y a des lignes de champ qui s'enroulent sur elles-mêmes (présence de tourbillons) les valeurs constantes de ϕ sur ces lignes de champ sont arbitraires, on n'a donc pas unicité de la solution.

Une autre façon de regarder le problème consiste à intégrer (1.12) exactement. Pour un point x donné de Ω introduisons $\{X_x(s)\}$ la ligne de champ passant par x . Cette ligne est solution de l'équation différentielle ordinaire

$$(1.13) \quad \begin{cases} \frac{dX_x(s)}{ds} = u(X_x(s)) \\ X_x(0) = x \end{cases}$$

Soit $X_{|\partial\Omega}(x)$ l'intersection de $\{X_x(s), s < 0\}$ avec $\partial\Omega$ et $s^-(x)$ l'abscisse curviligne de ce point: $X_{|\partial\Omega}(x) = X_x(s^-(x))$. Le point $X_{|\partial\Omega}(x)$ est l'intersection de Σ^- avec l'unique ligne de champ passant par x . Cette intersection existe bien si toutes les lignes de champ entrent dans le domaine par Σ^- . Muni de ces nouvelles définitions on montre que $u \cdot \nabla \phi = d[\phi(X_x(s))]/ds$; en intégrant le long de la ligne de champ de $X_{|\partial\Omega}(x)$ à x on obtient la solution explicite:

$$(1.14) \quad \phi(x) = a(X_{|\partial\Omega}(x)) + \int_{s^-(x)}^0 f(X_x(s)) ds$$

qui montre bien que la valeur de ϕ en un point x ne dépend que de l'information qui s'est propagée et accumulée en amont sur la ligne de champ passant par x . Cette propriété montre que contrairement aux équations elliptiques qui sont régularisantes les équations hyperboliques produisent des solutions qui ont, au mieux, la même régularité que les données.

Les approximations numériques de telles équations doivent tenir compte de ce phénomène sous peine de graves problèmes de stabilité. Remarquons enfin que, modulo les problèmes de l'incompressibilité, les équations de Navier-Stokes tendent vers le régime convectif pur lorsque le nombre de Reynolds devient grand.

1.2.3 L'incompressibilité

Un modèle simple de l'incompressibilité et de son interaction avec la diffusion visqueuse est donné par les équations de Stokes.

$$(1.15) \quad \begin{cases} -\nu \nabla^2 u + \nabla p = f \\ \nabla \cdot u = g \\ u_{|\partial\Omega} = 0 \end{cases}$$

où la donnée g doit satisfaire la condition de compatibilité $\int_{\Omega} g = 0$.

Ce problème fait partie de la classe des problèmes elliptiques. Il s'interprète comme un problème de point selle où la pression est la variable de Lagrange qui permet d'imposer une contrainte sur la divergence de la vitesse (voir Brezzi [6] pour d'autres détails sur cette interprétation). On peut aussi l'interpréter comme un problème elliptique classique en quotientant l'équation de conservation de la quantité de mouvement par tous les gradients, quitte à relever le terme source g , en restreignant la recherche de la vitesse dans l'espace des fonctions à divergence nulle (cf. Ladyzhenskaya [12] ou Temam [20] pour une analyse détaillée de cette approche).

Dans la perspective de l'approximation numérique du problème de Stokes, il convient de remarquer que la pression n'est définie qu'à une constante près, c'est-à-dire que le noyau de l'opérateur gradient est réduit aux constantes. Cette caractéristique est essentielle et doit être reproduite par les approximations numériques de ce problème. Dans le même ordre d'idée remarquons aussi que la solution doit nécessairement satisfaire la condition $\int_{\Omega} \nabla \cdot u = 0$, l'opérateur divergence n'est donc pas de rang maximum, son image est orthogonale à la fonction caractéristique de Ω . Cette propriété doit être respectée par les approximations numériques. On remarque enfin que $\int_{\Omega} u \nabla p = - \int_{\Omega} p \nabla \cdot u$ dès que u est nulle sur le bord de Ω . Cette relation signifie que les opérateurs gradient et divergence sont transposés l'un de l'autre au signe près, cette propriété se révèle par la suite très importante et doit être respectée par les approximations discrètes sous peine de graves difficultés.

Dans le cadre de l'approximation temporelle des équations de Navier–Stokes instationnaires on a à résoudre à chaque pas de temps un problème de Stokes généralisé de la forme:

$$(1.16) \quad \begin{cases} u - \sigma \nabla^2 u + \nabla p = f \\ \nabla \cdot u = g \\ u|_{\partial\Omega} = 0 \end{cases}$$

où le réel σ est strictement positif et proportionnel à $\nu \delta t$, où δt est le pas de temps de l'approximation temporelle.

1.2.4 L'instationnarité

Les modèles instationnaires se déduisent de ceux des sections précédentes en ajoutant un terme $\partial u / \partial t$. Ainsi l'équation prototype de la diffusion instationnaire est l'équation de la chaleur:

$$(1.17) \quad \begin{cases} \frac{\partial u}{\partial t} - \nu \nabla^2 u = f \\ u|_{\partial\Omega}(t) = a(t) \\ u(t=0) = u_0 \end{cases}$$

Ces équations sont répertoriées dans la classe des équations paraboliques. Pour ce type d'équation l'information se propage selon les temps positifs. Il est possible de montrer une certaine stabilité au cours du temps (ie. continuité) de la solution en fonction des données initiales, aux limites et sources. Si on change le sens du temps l'équation change complètement de nature, la solution de cette autre équation devient explosive dans le temps.

Les équations de convection instationnaires sont de la forme

$$(1.18) \quad \begin{cases} \frac{\partial \phi}{\partial t} + u \cdot \nabla \phi = f \\ u|_{\Sigma_t^-} = a \\ \phi(t=0) = \phi_0 \end{cases}$$

où Σ_t^- est la portion de frontière à l'instant t par laquelle entrent les lignes de champ de la vitesse u . Ces équations sont classées parmi les équations hyperboliques du premier ordre. Comme dans le cas de la convection stationnaire l'information se propage le long des lignes de champ de $(u, 1)$ compris comme un champ de vecteur de $\mathbb{R}^d \times [0, \infty[$ où la $d+1$ ème dimension représente le temps.

On peut exhiber la solution explicite en introduisant les trajectoires $\{X_{x,t}(\tau)\}$ telles que:

$$(1.19) \quad \begin{cases} \frac{dX_{x,t}(\tau)}{d\tau} = \begin{cases} u(X_{x,t}(\tau), \tau) & \text{si } X_{x,t}(\tau) \in \Omega \\ 0 & \text{sinon} \end{cases} \\ X_{x,t}(t) = x \end{cases}$$

Le problème d'évolution sur ϕ se réécrit sous la forme condensée $d[\phi(X_{x,t}(\tau), \tau)]/d\tau = f(X_{x,t}(\tau), \tau)$. En intégrant le long de la trajectoire on obtient

$$(1.20) \quad \phi(x, t) = \lambda(x, t) + \int_{t_{x,t}^-}^t f(X_{x,t}(\tau), \tau) d\tau$$

où $\lambda(x, t) = \phi_0(x)$ et $t_{x,t}^- = 0$ si $X_{x,t}(0) \in \Omega$, c'est-à-dire si la trajectoire n'est pas sortie du domaine en remontant le temps jusqu'à $t = 0$. Dans le cas contraire, c'est-à-dire si $X_{x,t}(0) \in \Sigma_t^-$, on a $\lambda(x, t) = a(X_{x,t}(t_{x,t}^-))$ et $t_{x,t}^-$ est le temps au bout duquel $X_{x,t}(\tau)$ atteint la frontière du domaine en remontant la trajectoire.

1.2.5 La non-linéarité de la convection

La non-linéarité du terme convectif des équations de Navier-Stokes est à l'origine des phénomènes de bifurcation des solutions en fonction du nombre de Reynolds, plus généralement elle est à l'origine de la turbulence. En quelques mots on peut dire que la non-linéarité a tendance à pomper l'énergie des grandes structures de l'écoulement induites par la géométrie du domaine et les termes sources vers les petites échelles. Le frein à ce phénomène est la diffusion visqueuse qui opère efficacement aux très petites échelles. Autrement dit, le terme $u \cdot \nabla u$ a tendance à faire apparaître et accentuer des fronts de vitesse alors que la diffusion visqueuse a tendance à régulariser ces fronts.

Ce phénomène est assez bien représenté par l'équation de Burgers en dimension 1:

$$(1.21) \quad \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} - \nu \frac{\partial^2 u}{\partial x^2} = 0$$

avec les conditions limites $u(-1, t) = -1$, $u(1, t) = 1$ et la condition initiale $u(x, 0) = x$. La solution limite stationnaire est donnée par $u(x, t \rightarrow \infty) = \tanh(x/2\nu)$. Le graphe de la solution limite est d'autant plus raide que ν (qui joue le rôle de l'inverse du nombre de

Reynolds) est petit. Cet exemple montre bien que l'analyse linéaire des schémas numériques ne pourra pas être suffisante pour s'assurer des comportements des solutions discrètes conformes aux solutions des problèmes continus. Ceci sera d'autant plus vrai que la non-linéarité de la convection sera un facteur dominant du problème.

Chapitre 2

Formulations fortes des problèmes de diffusion et approximation par différences finies

Ce chapitre est consacré à l'étude de la formulation forte des équations elliptiques d'ordre deux à coefficients constants. Dans la première partie on donne quelques résultats généraux de régularité et de stabilité. Dans la seconde partie on montre comment on peut approcher numériquement les solutions des problèmes de diffusion à l'aide des différences finies. Dans la dernière partie on montre un résultat général de convergence.

2.1 Etude de la formulation forte

2.1.1 Définitions et généralités sur les problèmes elliptiques

Etant donné Ω un ouvert connexe borné régulier de \mathbb{R}^d , considérons l'équation aux dérivées partielles du second ordre

$$(2.1) \quad - \sum_{i,j=1}^d a_{ij} \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_i b_i \frac{\partial u}{\partial x_i} + cu = f$$

où les coefficients a_{ij} , b_i , c et f sont des fonctions à valeur réelle sur Ω . On appelle solution classique ou solution forte de (2.1) une fonction $u \in C^2(\Omega)$ vérifiant (2.1) pour tout x dans Ω . Etant donné que pour tout $u \in C^2(\Omega)$ on a $\partial^2 u / \partial x_i \partial x_j = \partial^2 u / \partial x_j \partial x_i$, quitte à remplacer a_{ij} par $(a_{ij} + a_{ji})/2$ on peut toujours supposer $a_{ij} = a_{ji}$. Définissons pour tout $x \in \Omega$, $A(x) = (a_{ij}(x))_{i,j=1,\dots,d}$ la matrice symétrique des coefficients du second ordre de (2.1). Cette matrice est diagonalisable. Soient $\lambda(x)$ et $\Lambda(x)$ la plus petite et la plus grande des valeurs propres de $A(x)$. On dit que l'équation (2.1) est elliptique sur Ω si pour tout $x \in \Omega$

$$(2.2) \quad \lambda(x) > 0$$

L'équation est strictement elliptique si

$$(2.3) \quad \inf_{x \in \Omega} \lambda(x) > 0$$

L'équation est dite uniformément elliptique si

$$(2.4) \quad \inf_{x \in \Omega} \frac{\lambda(x)}{\Lambda(x)} > 0$$

Lorsque tous les coefficients a_{ij} , b_i et c sont constants et si l'équation (2.1) est elliptique, on peut toujours se ramener par des changements d'axes de coordonnées et de fonction à l'équation:

$$(2.5) \quad -\nabla^2 u + \alpha u = f$$

où α est un réel.

Exemple: A titre d'exemple en dimension 1, en posant $v = ue^{bx/2a}$, l'équation $a\partial^2 u/\partial x^2 + b\partial u/\partial x + cu = f$ devient:

$$(2.6) \quad -\frac{\partial^2 v}{\partial x^2} + \frac{b^2 - 4ac}{4a^2}v = -\frac{e^{bx/2a}}{a}f.$$

On obtient l'équation type en posant $\alpha = (b^2 - 4ac)/4a^2$.

- Le cas $\alpha < 0$ est appelé équation de Helmholtz. Cette équation intervient dans l'étude de l'équation des ondes (rayonnement électromagnétique, acoustique, vibrations de plaques, etc...). L'étude de ce type d'équation sort du cadre du présent exposé. Retenons simplement que lorsque $\alpha < 0$ le noyau de l'opérateur $-\nabla^2 + \alpha Id$ avec des conditions de Dirichlet homogènes, par exemple, n'est pas forcément réduit à zéro. C'est-à-dire que dans ce cas peuvent se poser des problèmes d'existence et d'unicité non triviaux.
- Le cas $\alpha = 0$ est connu sous le nom d'équation de Poisson.
- Le cas $\alpha > 0$ présente pratiquement toutes les propriétés de l'équation de Poisson. On lui donne parfois le nom d'équation de la neutronique. On l'écrit généralement sous la forme

$$(2.7) \quad -\nabla^2 u + k^2 u = f$$

En fait un modèle plus réaliste de la diffusion neutronique s'écrit plutôt

$$(2.8) \quad -\nabla \cdot (a(x) \cdot \nabla u) + k^2 u = f$$

où $a(x)$ est une fonction positive et définie sur Ω . Par abus de langage dans la communauté de la mécanique de fluides numérique (MFN), l'équation (2.7) est aussi appelée équation de Helmholtz.

2.1.2 Les conditions aux limites

Afin d'assurer l'unicité des solutions des problèmes du type (2.7) il faut imposer des conditions supplémentaires sur le bord du domaine Ω . Ces conditions sont généralement de trois types:

- Condition de Dirichlet: si u représente un champ de température dans Ω , cette condition revient à imposer la valeur de la température au bord du domaine $\partial\Omega$:

$$(2.9) \quad u|_{\partial\Omega} = a$$

- Condition de Neumann: en conservant l'image du champ de température pour u , la condition de Neumann revient à imposer le flux de chaleur sur $\partial\Omega$. Par exemple pour des parois adiabatiques (ie. parfaitement isolantes) on aurait

$$(2.10) \quad \frac{\partial u}{\partial n}|_{\partial\Omega} = 0$$

- Condition de Robin ou de Fourier: on impose une combinaison linéaire des deux conditions précédentes. Dans le cas où u serait un champ de température, cette condition consiste à imposer un flux de chaleur proportionnel à l'écart de température entre la paroi et le fluide:

$$(2.11) \quad \frac{\partial T}{\partial n}|_{\partial\Omega} = a(T - T_e)$$

Bien sûr on peut faire un panachage de ces conditions, par exemple en imposant une condition de Dirichlet sur une portion $\partial\Omega_1$ de $\partial\Omega$, puis une condition de Neumann sur $\partial\Omega_2$ et une condition de Fourier sur le complémentaire de $\partial\Omega_1 \cup \partial\Omega_2$.

Remarquons enfin que le nombre de conditions ainsi que leur support dépendent de la nature de l'équation aux dérivées partielles (EDP) à laquelle on s'intéresse. Pour un problème elliptique d'ordre 2, on applique une condition sur tout le bord du domaine. Ce n'est plus vrai si l'EDP est d'un autre type, par exemple pour une équation hyperbolique du premier ordre on applique une condition de Dirichlet sur la partie du bord de Ω où le champ convecteur est entrant (cf. chapitre 1). Pour une équation elliptique du quatrième ordre, par exemple $\nabla^2 \nabla^2 u = f$, on impose deux conditions sur tout le bord: une condition de Dirichlet et une condition de Neumann.

2.1.3 Formules de Green-Ostrogradski

L'un des outils les plus performants de l'analyse est certainement la technique d'intégration par parties. Dans le cadre des problèmes elliptiques du second ordre on a systématiquement recours à cette technique par le biais de la formule d'Ostrogradski et ses corollaires.

Proposition 1 *Soit u un champ de vecteur de $(C^1(\Omega) \cap C^0(\bar{\Omega}))^d$ tel que $\nabla \cdot u$ soit intégrable sur Ω , alors on a:*

$$(2.12) \quad \int_{\Omega} \nabla \cdot u = \int_{\partial\Omega} u \cdot n$$

Dans un souci de concision on omet les éléments différentiels $d\Omega$ et $d\partial\Omega$ dans les intégrales sachant que par la suite on utilisera implicitement la mesure de Lebesgue systématiquement (sauf avis explicite contraire). La formule d'Ostrogradski permet de démontrer les formules de Green qui sont véritablement des formules d'intégration par parties:

Proposition 2 (i) *Soient u un champ de vecteur de $(C^1(\Omega) \cap C^0(\bar{\Omega}))^d$ et v une fonction scalaire de $C^1(\Omega) \cap C^0(\bar{\Omega})$ telles que $v \nabla \cdot u$ et $u \cdot \nabla v$ soient intégrables sur Ω , alors:*

$$(2.13) \quad \int_{\Omega} v \nabla \cdot u = - \int_{\Omega} u \cdot \nabla v + \int_{\partial\Omega} (u \cdot n) v$$

(ii) Soient u une fonction de $(C^2(\Omega) \cap C^1(\bar{\Omega}))^d$ et v une fonction scalaire de $C^1(\Omega) \cap C^0(\bar{\Omega})$ telles que $v \nabla^2 u$ et $\nabla u \cdot \nabla v$ soient intégrables sur Ω , alors:

$$(2.14) \quad \int_{\Omega} v \nabla^2 u = - \int_{\Omega} \nabla u \cdot \nabla v + \int_{\partial\Omega} (\nabla u \cdot n) v$$

(iii) Soient u et v deux fonctions scalaires de $C^2(\Omega) \cap C^1(\bar{\Omega})$ telles que $v \nabla^2 u$, $u \nabla^2 v$ et $\nabla u \cdot \nabla v$ soient intégrables sur Ω , alors:

$$(2.15) \quad \int_{\Omega} v \nabla^2 u = \int_{\Omega} u \nabla^2 v + \int_{\partial\Omega} [(\nabla u \cdot n) v - u (\nabla v \cdot n)]$$

Démonstration: (i) Utiliser la formule d’Ostrogradski en remarquant que $\nabla \cdot (uv) = u \cdot \nabla v + v \nabla \cdot u$.

(ii) Appliquer (i) à ∇u et v .

(iii) Appliquer (ii) au couple (u, v) puis au couple (v, u) puis retrancher une équation à l’autre.

□

Les formules ci-dessus donnent des informations importantes sur les opérateurs divergence, gradient et laplacien. La formule (2.13) montre que l’opérateur divergence agissant sur les champs de vecteurs à trace normale nulle sur le bord du domaine est au signe près le transposé de l’opérateur gradient. En termes d’approximation numérique cela signifie que les matrices associées aux approximations de la divergence et du gradient doivent être symétriques l’une de l’autre au signe près (ou quasi-symétriques), et ce quel que soit le type de l’approximation. La formule (2.15) signifie que ∇^2 opérant sur les fonctions nulles sur $\partial\Omega$ ou de dérivée normale nulle est autoadjoint. Cela veut dire que les matrices d’approximation de ∇^2 seront vraisemblablement symétriques. D’autre part en identifiant les opérateurs transposés de la divergence, du gradient et du laplacien on ramène l’étude du noyau de l’un quelconque de ces opérateurs à l’étude de l’orthogonal (au sens du produit scalaire de $L^2(\Omega)$) de son transposé grâce à la formule classique:

$$(2.16) \quad \ker T = (\text{Im } T^t)^\perp$$

Ces formules mettent d’ores et déjà en lumière le fait que le cadre classique n’est pas bien adapté pour bénéficier pleinement de la théorie de la dualité. Pour travailler dans ce cadre il faut plutôt utiliser les espaces de fonctions munis d’une structure hilbertienne induite par le produit scalaire de $L^2(\Omega)$.

2.2 L’équation de la neutronique (ou de “Helmholtz”)

Dans cette section on s’intéresse à l’équation de la neutronique (on dira aussi équation de Helmholtz par abus de langage):

$$(2.17) \quad -\nu \nabla^2 u + u = f$$

On étudie plus particulièrement l’unicité d’une solution éventuelle et sa stabilité par rapport aux termes source et données de frontière. On montre l’importance du principe du maximum pour l’étude des problèmes avec données de Dirichlet. On termine enfin en montrant l’importance des techniques de dualité (ou énergétiques) pour étudier l’existence des solutions dans le cadre de données à la frontière arbitraires.

2.2.1 Principe du maximum

Le principe du maximum est un des outils les plus performants pour étudier les problèmes elliptiques d’ordre deux. Ce principe se comprend assez bien intuitivement lorsqu’on pense à un champ de température induit par chauffage au bords d’un domaine Ω borné. En effet, dans ce cadre il paraît assez naturel que la température en tout point du domaine ne peut pas dépasser la température maximum qu’on lui impose à la frontière. De même la température minimum ne peut pas être inférieure à la température minimum imposée sur $\partial\Omega$. En termes plus formels on a le résultat suivant:

Proposition 3 *Si u est une solution classique de (2.17) alors*

$$(2.18) \quad \forall x \in \Omega, \quad \min \left\{ \inf_{\partial\Omega} u, \inf_{\Omega} f \right\} \leq u(x) \leq \max \left\{ \sup_{\partial\Omega} u, \sup_{\Omega} f \right\}$$

Démonstration: Soit $x_0 \in \bar{\Omega}$ tel que $u(x_0) = \sup_{x \in \Omega} u(x)$. Deux possibilités se présentent, ou bien x_0 est sur la frontière, ou bien x_0 appartient à l'ouvert Ω . Dans le premier cas on a

$$\sup_{\Omega} u(x) = u(x_0) \leq \sup_{\partial\Omega} u(x) \leq \max \left\{ \sup_{\partial\Omega} u(x), \sup_{\Omega} f(x) \right\},$$

le résultat est acquis. Dans le second cas x_0 réalise un maximum de u ; aussi en ce point le gradient de u est nul et les dérivées secondes $\partial^2 u / \partial x_i^2$ sont toutes négatives. On en déduit $\nabla^2 u \leq 0$, d'où le résultat suivant:

$$\sup_{\Omega} u(x) = u(x_0) = f(x_0) + \nu \nabla^2 u(x_0) \leq f(x_0) \leq \sup_{\Omega} f(x) \leq \max \left\{ \sup_{\partial\Omega} u(x), \sup_{\Omega} f(x) \right\}.$$

On a donc démontré l'inégalité pour la borne supérieure de u . L'inégalité pour la borne inférieure se démontre de la même façon. \square

Par la suite, en supposant les régularités voulues, on adopte les notations suivantes:

$$(2.19) \quad |f|_{\infty, \Omega} = \sup_{x \in \Omega} f(x), \quad |g|_{\infty, \partial\Omega} = \sup_{x \in \partial\Omega} g(x)$$

2.2.2 Unicité et stabilité pour le problème de Dirichlet

Considérons l'équation de Helmholtz à laquelle on associe une donnée de Dirichlet (par abus de langage on parle de problème de Dirichlet):

$$(2.20) \quad \begin{cases} -\nu \nabla^2 u + u = f \\ u|_{\partial\Omega} = g \end{cases}$$

On suppose par la suite que ce problème admet une solution classique $u \in C^2(\Omega) \cap C^0(\bar{\Omega})$. En fait la démonstration d'un tel résultat n'est pas du tout évidente et est très technique. Par contre la démonstration de l'unicité est à notre portée puisqu'elle résulte directement du principe du maximum.

Théorème 1 (i) Pour toutes solutions de (2.20) on a

$$(2.21) \quad |u|_{\infty, \Omega} \leq \max \{ |g|_{\infty, \partial\Omega}, |f|_{\infty, \Omega} \}$$

(ii) La solution du problème est unique.

Démonstration: (i) Appliquer le principe du maximum (2.18).

(ii) Soient u_1 et u_2 deux solutions du problème (2.20), la différence $u_2 - u_1$ satisfait le même problème avec une condition de frontière homogène et un terme source nul. D'après le principe du maximum on a donc $|u_2 - u_1|_{\infty, \Omega} \leq 0$, c'est-à-dire $u_2 = u_1$. \square

L'inégalité (2.21) montre que la solution dépend continuellement des données au sens de la topologie de la convergence uniforme. On dit aussi qu'on a stabilité de la solution par rapport aux données. C'est-à-dire que de faibles variations sur les données f ou g entraînent des variations sur la solution qui sont au plus du même ordre de grandeur.

Le principe du maximum est bien adapté à l'étude des problèmes avec condition de Dirichlet. Il l'est beaucoup moins pour les autres types de conditions limites. En effet pour les problèmes de Neumann et de Fourier on est incapable de contrôler la valeur de la fonction sur le bord du domaine. On est donc obligé, même pour les problèmes d'unicité et de stabilité,

de se tourner vers d'autres techniques. On montre par la suite comment les techniques de dualité associées aux formules de Green permettent aussi de répondre assez simplement aux questions de stabilité et d'unicité. En laissant tomber un peu du lest de nos exigences sur la régularité de la solution on verra que ces techniques permettent aussi de montrer l'existence d'une solution faible.

2.2.3 Problèmes de Neumann - Robin - Mêlés

Considérons le problème suivant avec des conditions aux limites mêlées:

$$(2.22) \quad \begin{cases} -\nu \nabla^2 u + u = f \\ u|_{\partial\Omega_1} = g_1 \\ \frac{\partial u}{\partial n}|_{\partial\Omega_2} = g_2 \\ \lambda u|_{\partial\Omega_3} + \frac{\partial u}{\partial n}|_{\partial\Omega_3} = g_3 \end{cases}$$

En supposant pour l'instant l'existence d'une solution classique à ce problème, on peut montrer:

Théorème 2 *La solution de (2.22) est unique si $\lambda > 0$.*

Démonstration La preuve de l'unicité repose sur les formules de Green (intégration par parties). En multipliant l'équation aux dérivées partielles par u et en intégrant sur Ω on obtient

$$\int_{\Omega} [\nu |\nabla u|^2 + |u|^2] = \int_{\partial\Omega} \nu \frac{\partial u}{\partial n} u + \int_{\Omega} f u$$

d'où

$$\int_{\Omega} [\nu |\nabla u|^2 + |u|^2] + \int_{\partial\Omega_3} \lambda |u|^2 = \int_{\partial\Omega_1} \nu g_1 \frac{\partial u}{\partial n} + \int_{\partial\Omega_2} \nu u g_2 + \int_{\partial\Omega_3} \nu u g_3 + \int_{\Omega} f u$$

Soient u_2 et u_1 deux solutions du problèmes, alors la différence $u_2 - u_1$ est solution du problème avec des condition aux limites homogènes et un terme source nul. On tire donc de l'égalité ci-dessus:

$$\int_{\Omega} [\nu |\nabla(u_2 - u_1)|^2 + |u_2 - u_1|^2] + \int_{\partial\Omega_3} \lambda |u_2 - u_1|^2 = 0$$

Ce qui donne $u_2 = u_1$ si λ est une fonction positive $\lambda \geq 0$. Lorsque $\lambda < 0$ on ne peut rien dire sur l'unicité.

La technique utilisée est très générale. Elle permet de traiter d'un seul coup tous les types de conditions aux limites qui nous intéressent. Cette façon de procéder est souvent appelée technique énergétique en référence au terme $\int_{\Omega} [\nu |\nabla u|^2 + |u|^2]$ qui représente en quelque sorte l'énergie du système. Cette technique entre dans le cadre plus large des techniques variationnelles.

Les questions d'existence de solutions classiques aux équations aux dérivées partielles considérées ici sont très techniques. Elles font intervenir de façon déterminante la régularité du bord $\partial\Omega$ (cf. l'ouvrage collectif de Dautray et Lions [8, t.2 p.387-401] pour une revue introductive à ces problèmes). En fait, si on est moins exigeant sur la régularité de la solution,

c'est-à-dire si on se satisfait de l'existence d'une solution dans un sous-espace de $L^2(\Omega)$, les techniques variationnelles permettent de répondre simplement au problème d'existence. Avec un peu de chance et un peu de travail on peut alors montrer, sous réserve de régularité du bord et des données, que la solution du problème variationnel est aussi solution au sens classique. Le cadre variationnel est probablement le cadre fonctionnel le plus utilisé par les consommateurs d'analyse numérique.

2.3 Approximation par différences finies

2.3.1 Introduction

Bien qu'on ait vu que la voie classique n'est pas la plus simple pour montrer l'existence des solutions, on persiste dans cette direction et on va chercher à approcher numériquement les opérateurs de dérivation par des différences finies.

Pour illustrer la technique dans le cadre des problèmes elliptiques de diffusion on se restreint en dimension 2 au problème de Dirichlet suivant:

$$(2.23) \quad \begin{cases} -\nu \nabla^2 u + u = f \\ u|_{\partial\Omega} = g \end{cases}$$

où Ω est le rectangle $]0, a[\times]0, b[$ et où les données sont régulières: $f \in C^2(\bar{\Omega})$, $g \in C^0(\partial\Omega)$.

2.3.2 Mise en œuvre

La mise en œuvre de la technique des différences finies consiste à remplacer les opérateurs différentiels par des quotients aux différences. Dans cette perspective on se donne deux entiers I, J et on définit un maillage du rectangle Ω par l'ensemble de points $(x_i = i\delta x, y_j = j\delta y)$ pour $0 \leq i \leq I + 1$ et $0 \leq j \leq J + 1$ et par définition $\delta x = a/(I + 1)$ et $\delta y = b/(J + 1)$. Les points (x_i, y_j) sont appelés nœuds du maillage. Afin de conserver l'analogie avec le problème continu on note Ω_h l'ensemble des nœuds intérieurs du maillage, $\partial\Omega_h$ l'ensemble des nœuds sur la frontière et $\bar{\Omega}_h = \Omega_h \cup \partial\Omega_h$.

La méthode des différences finies consiste à obtenir une approximation de la solution u aux nœuds du maillage. Pour ce faire on approche les opérateurs aux nœuds du maillage à l'aide des formules de Taylor. Pour le cas qui nous intéresse, en notant $u_{ij} = u(x_i, x_j)$ on a pour tout $u \in C^4(\bar{\Omega})$

$$(2.24) \quad \begin{aligned} \nabla^2 u_{ij} &= \frac{1}{\delta x^2} [2u_{ij} - u_{i+1,j} - u_{i-1,j}] + \frac{\delta x^2}{12} \frac{\partial^4 u}{\partial x^4}(\xi_i, y_j) \\ &+ \frac{1}{\delta y^2} [2u_{ij} - u_{i,j+1} - u_{i,j-1}] + \frac{\delta y^2}{12} \frac{\partial^4 u}{\partial y^4}(x_i, \eta_j). \end{aligned}$$

où ξ et η vérifient $x_i < \xi_i < x_{i+1}$ et $y_j < \eta_j < y_{j+1}$. Les deux principes de base sont alors:

- (i) Négliger le reste du développement en δx^2 et δy^2 .
- (ii) Ecrire que le système (2.23) est satisfait en tous les nœuds intérieurs du maillage.

La quantité $u - \nu \nabla^2 u$ est alors approchée par

$$(2.25) \quad u_{ij} + \frac{\nu}{\delta x^2} [2u_{ij} - u_{i+1,j} - u_{i-1,j}] + \frac{\nu}{\delta y^2} [2u_{ij} - u_{i,j+1} - u_{i,j-1}]$$

En posant $f_{ij} = f(x_i, y_j)$ et $g_{kl} = g(x_k, y_l)$ on obtient donc le système de $I \times J$ équations aux $I \times J$ inconnues u_{ij} , $1 \leq i \leq I$, $1 \leq j \leq J$:

$$(2.26) \quad u_{ij} + \frac{\nu}{\delta x^2} [2u_{ij} - u_{i+1,j} - u_{i-1,j}] + \frac{\nu}{\delta y^2} [2u_{ij} - u_{i,j+1} - u_{i,j-1}] = f_{ij}$$

avec les données au bord:

$$(2.27) \quad u_{kl} = g_{kl}, \quad \text{pour } (x_k, y_l) \in \partial\Omega_h$$

qui sont prises en compte dans le système (2.26) en remplaçant u_{kl} par g_{kl} dès que (x_k, y_l) est sur le bord du domaine et en rejetant ce terme dans le second membre. On obtient ainsi un système linéaire

$$(2.28) \quad AU = F$$

où le vecteur colonne U contient les inconnues (u_{ij}) . Le rangement des inconnues dans le vecteur U définit une numérotation des nœuds de 1 à $I \times J$. La structure de A dépend de cette numérotation. On adopte généralement une numérotation lignes par lignes ou colonnes par colonnes. Quelque soit la numérotation choisie la matrice A est creuse. Sur chaque ligne (resp. colonne) elle a seulement 3 coefficients non nuls si l'indice de ligne (resp. colonne) correspond à un nœud de coin, 4 coefficients non nuls si l'indice correspond à un nœud de bord et 5 non nuls dans le cas général.

La question qui se pose à ce niveau est de savoir si le système linéaire discret (2.28) admet une solution unique et si c'est le cas est-ce que cette solution discrète converge en un certain sens vers la solution classique du problème continu (2.23)?

2.3.3 Existence et unicité de la solution approchée

En s'inspirant de la démonstration de l'unicité de la solution du problème continu on est conduit à démontrer une version discrète du principe du maximum:

Lemme 1 Pour tout nœud intérieur $x_{ij} \in \Omega_h$ on a:

$$(2.29) \quad \min \left\{ \min_{\partial\Omega_h} g, \min_{\Omega_h} f \right\} \leq u_{ij} \leq \max \left\{ \max_{\partial\Omega_h} g, \max_{\Omega_h} f \right\}$$

et en particulier en introduisant la norme du sup et en notant u_h la famille des $\{u_{ij}\}$, on a

$$(2.30) \quad |u_h|_{\infty, \Omega_h} \leq \max \{|g|_{\infty, \partial\Omega_h}, |f|_{\infty, \Omega_h}\}$$

Démonstration: On procède comme dans le cas continu. Soit $x_0 \in \bar{\Omega}_h$ tel que $u_h(x_0) = \max_{x \in \bar{\Omega}_h} u_h(x)$. Deux possibilités se présentent: x_0 est sur la frontière $\partial\Omega_h$ ou bien x_0 appartient à Ω_h . Dans le premier cas on a

$$\max_{\Omega_h} u_h(x) \leq \max_{\bar{\Omega}_h} u_h(x) = u_h(x_0) = \max_{\partial\Omega_h} g(x) \leq \max \left\{ \max_{\partial\Omega_h} g(x), \max_{\Omega_h} f(x) \right\}$$

et le résultat est acquis. Dans le second cas en posant $x_0 = x_{ij}$ on a

$$u_{ij} + \frac{\nu}{\delta x^2} [u_{ij} - u_{i+1,j}] + \frac{\nu}{\delta x^2} [u_{ij} - u_{i-1,j}] + \frac{\nu}{\delta y^2} [u_{ij} - u_{i,j+1}] + \frac{\nu}{\delta y^2} [u_{ij} - u_{i,j-1}] = f_{ij}$$

et comme u_{ij} est la plus grande valeur de u_h sur le domaine $\bar{\Omega}_h$, on déduit de l'égalité ci-dessus:

$$u_{ij} \leq f_{ij}.$$

D'où

$$\sup_{\Omega} u(x) = u(x_0) \leq f(x_0) \leq \max_{\Omega} f(x) \leq \max \left\{ \max_{\partial\Omega} g(x), \max_{\Omega} f(x) \right\}.$$

On a donc démontré l'inégalité pour la borne supérieure de u . L'autre inégalité se démontre de la même façon. \square

Remarquer que la démonstration ci-dessus ne fait pas intervenir la forme du domaine. La technique des différences finies est applicable à tout domaine Ω_h du moment que celui-ci est une union de mailles élémentaires cartésiennes. A partir du principe du maximum on déduit immédiatement le résultat d'existence et d'unicité espéré:

Théorème 3 *Le problème discret (2.28) admet une solution unique. En d'autres termes, la matrice A est inversible.*

Démonstration: La matrice A est carrée ($I.J \times I.J$) et d'après le théorème du rang, l'inversibilité de la matrice est équivalente à l'injectivité (unicité) ou à la surjectivité (existence) de l'opérateur linéaire associé (ie. unicité \Leftrightarrow existence). Il suffit donc de montrer l'unicité de la solution pour aussi conclure à son existence. Pour montrer l'unicité considérons le problème homogène associé: $f_h = 0, g_h = 0$. D'après le principe du maximum (2.30) on a $|u_h|_{\infty, \Omega_h} = 0$, c'est-à-dire $u_h = 0$, d'où l'unicité, et l'existence en résulte. \square

Une autre façon de voir les choses est de remarquer que A est à diagonale dominante et de ce fait est inversible. Remarquons toutefois que ce raisonnement ne serait plus valable si au lieu d'approcher l'opérateur de Helmholtz on approchait le problème $-\nabla^2 u = f$. Par contre la technique du principe du maximum est encore applicable dans ce cas. En fait le principe du maximum est un des outils les plus puissants pour étudier les problèmes elliptiques d'ordre deux avec conditions de Dirichlet.

2.4 Consistance - Stabilité - Convergence

Dans cette section on définit une notion de convergence (ie. on définit une topologie) et on dégage des conditions suffisantes pour assurer la convergence des solutions discrètes vers les solutions des problèmes continus. En particulier on démontre une version d'un des théorèmes (principes) les plus importants de l'analyse numérique des EDP:

$$\text{Consistance} + \text{Stabilité} \implies \text{Convergence}$$

2.4.1 Définitions des espaces fonctionnels

Avant d'introduire une notion de convergence il faut définir correctement les espaces dans lesquels les solutions vivent. Pour ce faire on introduit M (resp. N) le nombre de nœuds dans Ω_h (resp. $\partial\Omega_h$). Définissons alors D_h (resp. F_h) l'espace vectoriel réel des fonctions définies sur Ω_h (resp. $\partial\Omega_h$) et à valeur dans \mathbb{R}^M (resp. \mathbb{R}^N). On munit ces espaces d'une structure d'espace vectoriel normé en les équipant de normes $|\cdot|_{D_h}, |\cdot|_{F_h}$. Par commodité d'écriture on introduit aussi l'espace produit $\overline{D}_h = D_h \times F_h$.

Exemple 1 : Dans le cas d'une grille cartésienne régulière ($\delta x = \delta y = h$) on peut choisir les normes l^2 associées à cette grille:

$$|v_h|_{D_h}^2 = h^2 \sum_{\Omega_h} v_{ij}^2$$

$$|v_h|_{F_h}^2 = h \sum_{\partial\Omega_h} v_{ki}^2$$

$$|v_h|_{\overline{D}_h}^2 = |v_h|_{D_h}^2 + |v_h|_{F_h}^2$$

Ces normes sont intrinsèques au sens où elles fournissent des normes équivalentes à celles de $L^2(\Omega)$ et $L^2(\partial\Omega)$ lorsque $h \rightarrow 0$.

Exemple 2 : Dans le même cadre discret que précédemment on peut aussi choisir la norme de l^∞ (ie. la norme du sup):

$$|v_h|_{D_h} = \max_{\Omega_h} |v_{ij}|$$

$$|v_h|_{F_h} = \max_{\partial\Omega_h} |v_{kl}|$$

$$|v_h|_{\overline{D}_h} = \max\{|v_h|_{D_h}, |v_h|_{F_h}\}$$

Ces normes sont particulièrement adaptées à l'approximation des EDP par des différences finies.

2.4.2 Définition d'un cadre abstrait

Pour mieux saisir la portée du résultat de convergence qui est donné par la suite il convient de se placer d'emblée dans un cadre abstrait. Après avoir introduit les espaces D_h , F_h et \overline{D}_h il semble naturel de les relier aux espaces des fonctions continues $C^0(\Omega)$, $C^0(\partial\Omega)$ et $C^0(\overline{\Omega})$. Pour ce faire on définit les opérateurs ρ_{1h} et ρ_{2h} tels que

$$(2.31) \quad \begin{aligned} \rho_{1h} : C^0(\Omega) &\longrightarrow D_h \\ v &\longmapsto \rho_{1h}(v) : \{x_{ij}\} \longmapsto \{v(x_{ij})\} \end{aligned}$$

$$(2.32) \quad \begin{aligned} \rho_{2h} : C^0(\partial\Omega) &\longrightarrow F_h \\ v &\longmapsto \rho_{2h}(v) : \{x_{kl}\} \longmapsto \{v(x_{kl})\} \end{aligned}$$

On définit de même l'opérateur $\rho = (\rho_{1h}, \rho_{2h})$.

Considérons maintenant un problème d'équations aux dérivées partielles écrit sous la forme abstraite générale

$$(2.33) \quad \begin{cases} L_1 u = f \\ L_2 u|_{\partial\Omega} = g \end{cases}$$

L'approximation du problème (2.33) par des différences finies conduit toujours à un problème linéaire du type:

$$(2.34) \quad \begin{cases} L_{1h} u_h = f_h \\ L_{2h} u_h|_{\partial\Omega_h} = g_h \end{cases}$$

où par définition on a posé $f_h = \rho_{1h}(f)$ et $g_h = \rho_{2h}(g)$ et L_{1h} et L_{2h} sont des opérateurs aux différences finies approchant L_1 et L_2 . Dans le cas de l'approximation du problème (2.23) on a $L_{1h} = Id_h - \nu \nabla_h^2$ et $L_{2h} = Id_h$. Dans un souci de simplification des notations introduisons les opérateurs $L = (L_1, L_2)$ et $L_h = (L_{1h}, L_{2h})$, alors le problème discret (2.34) prend la forme générale

$$(2.35) \quad L_h u_h = \rho_h(f, g)$$

Par la suite on suppose que le système linéaire discret (2.35) est carré (ie. il y a autant d'inconnues que d'équations) pour qu'il ait quelques chances d'être inversible.

2.4.3 Consistance

Pour légitimement espérer une certaine convergence de la solution discrète u_h vers u la solution du problème continu, il faut que les opérateurs L_{1h} et L_{2h} approchent L_1 et L_2 en un certain sens. On formalise cette idée en introduisant la notion de consistance.

Définition 1 On dit que L_h est une approximation consistante de L pour la norme de \overline{D}_h si

$$(2.36) \quad |L_h(\rho_h u) - \rho_h L(u)|_{\overline{D}_h} \xrightarrow{h \rightarrow 0} 0$$

On dit que L_h est une approximation consistante de L d'ordre $p > 0$ si

$$(2.37) \quad |L_h(\rho_h u) - \rho_h L(u)|_{\overline{D}_h} \leq c(u)h^p$$

où la constante $c(u) \geq 0$ est indépendante de h .

Exemple 1 : Dans le cas du problème de Helmholtz, en notant $h = \max\{\delta x, \delta y\}$ et en équipant D_h et F_h de la norme du sup on a :

$$|L_h(\rho_h u) - \rho_h L(u)|_{\overline{D}_h} \leq \frac{h^2}{12} \left| \frac{\partial^4 u}{\partial x^4}(\xi_i, y_j) + \frac{\partial^4 u}{\partial y^4}(x_i, \eta_j) \right|_{D_h} \leq c(u)h^2$$

L'approximation (2.26) est donc consistante à l'ordre 2.

2.4.4 Stabilité

On dit qu'un problème d'EDP est bien posé si son unique solution est continue (au sens d'une topologie à définir) par rapport aux données. C'est-à-dire que si on perturbe un peu les données, l'écart de la solution perturbée à la solution initiale doit rester commensurable aux perturbations; il n'y a pas amplification dramatique des perturbations. On généralise cette notion aux problèmes discrets

Définition 2 On dit qu'une famille d'approximations discrètes $\{L_h\}$ est stable si la continuité des solutions par rapport aux données est indépendante de h . En d'autres termes on a stabilité de la famille $\{L_h\}$ si :

$$(2.38) \quad |u_h|_{\overline{D}_h} \leq c|\rho_h(f, g)|_{\overline{D}_h} = c|L_h(u_h)|_{\overline{D}_h}.$$

Remarque: La stabilité implique l'unicité de la solution du problème discret, en effet pour le problème homogène (ie. $f = 0, g = 0$) l'inégalité (2.38) donne $|u_h|_{\overline{D}_h} = 0$. Lorsque la matrice associée à L_h est carrée, la stabilité implique donc l'inversibilité du système linéaire. En notant L_h^{-1} l'inverse de L_h , la condition de stabilité implique que la famille d'opérateurs $\{L_h^{-1}\}$ est uniformément bornée par rapport à h . En effet l'inégalité (2.38) s'écrit de façon équivalente

$$|L_h^{-1}(f_h, g_h)|_{\overline{D}_h} \leq c|(f_h, g_h)|_{\overline{D}_h}.$$

Ce qui signifie que la norme de L_h^{-1} est majorée par une constante c indépendamment de h .

Exemple 1 : Si L_h est un opérateur symétrique, et si la norme de \overline{D}_h est induite par un produit scalaire (ie. on a choisi une norme l^2) alors L_h est diagonalisable dans une base orthonormée de \overline{D}_h . Soit λ_{1h} la plus petite valeur propre de L_h (qu'on suppose être positive), on a :

$$\lambda_{1h}|u_h|_{\overline{D}_h} \leq |L_h(u_h)|_{\overline{D}_h}.$$

L'inversibilité de L_h implique $\lambda_{1h} \neq 0$. Si λ_{1h} est minorée indépendamment de h par une constante λ_1 , on a stabilité de la famille d'approximation avec une constante de stabilité égale à $1/\lambda_1$.

Exemple 2 : Pour le problème de Helmholtz, en équipant D_h et F_h de la norme du sup, le principe du maximum (2.30) fournit une constante de stabilité égale à 1.

Exemple 3 : Pour l'équation de Poisson $\nabla^2 u = f$ avec une donnée de Dirichlet $u|_{\partial\Omega} = g$ on peut montrer la stabilité (en norme du sup) de l'approximation par différences finies en vérifiant le principe du maximum. Démontrons ce résultat en dimension 2 sur $\Omega =]0, a[\times]0, b[$ en reprenant la même discrétisation que pour le problème de Helmholtz.

Lemme 2 *Toute fonction discrète u_h telle que*

$$(2.39) \quad \nabla_h^2 u_{ij} \geq 0, \quad \text{aux nœuds intérieurs}$$

atteint sa plus grande valeur en certains points de $\partial\Omega_h$. On a aussi la proposition symétrique, toute fonction discrète u_h telle que

$$(2.40) \quad \nabla_h^2 u_{ij} \leq 0, \quad \text{aux nœuds intérieurs}$$

atteint sa plus petite valeur en certains points de $\partial\Omega_h$.

Démonstration: Montrons la propriété pour la plus grande valeur de u_h , la démonstration dans l'autre sens est identique. Raisonnons par l'absurde, supposons que u_h atteint sa plus grande valeur strictement en un nœud intérieur, il est donc possible par hypothèse de choisir ce point x_{ij} de telle sorte que la valeur de u_h en ce point soit strictement plus grande que l'une des valeurs au moins prises aux quatre points voisins. En ce point x_{ij} on a:

$$0 \geq \frac{1}{\delta x^2} [u_{ij} - u_{i+1,j}] + \frac{1}{\delta x^2} [u_{ij} - u_{i-1,j}] + \frac{1}{\delta y^2} [u_{ij} - u_{i,j+1}] + \frac{1}{\delta y^2} [u_{ij} - u_{i,j-1}] > 0$$

La dernière inégalité est stricte puisque tous les termes sont positifs ou nuls et un des termes au moins est strictement positif, d'où la contradiction. \square

On peut établir maintenant le résultat de stabilité.

Théorème 4 *Les solutions du problème de Poisson discret $\nabla^2 u_h = f_h$ avec une donnée de Dirichlet $u|_{\partial\Omega_h} = g_h$ satisfont l'inégalité de stabilité:*

$$(2.41) \quad |u_h|_{\infty, \bar{\Omega}_h} \leq c \max \{ |g_h|_{\infty, \partial\Omega_h}, |f_h|_{\infty, \Omega_h} \}$$

Démonstration : (cf. Godounov-Riabenki [11, p. 269]) Introduisons $R = (a^2 + b^2)^{1/2}$ le diamètre du domaine et construisons la fonction polynômiale auxiliaire

$$P_h(x, y) = \frac{1}{4}(R^2 - x^2 - y^2)|f_h|_{\infty, \Omega_h} + |g_h|_{\infty, \partial\Omega_h}.$$

Dans un premier temps on vérifie aisément que

$$\nabla^2(u_h - P_h(x, y)) = f_h - |f_h|_{\infty, \Omega_h} \geq 0$$

d'où on déduit à l'aide du lemme ci-dessus que la fonction $u_h - p_h$ prend sa plus grande valeur sur $\partial\Omega_h$. Or sur la frontière on a

$$(u_h - P_h)|_{\partial\Omega_h} = g_h - |g_h|_{\infty, \partial\Omega_h} - \frac{1}{4}(R^2 - x^2 - y^2)|f_h|_{\infty, \Omega_h} \leq 0$$

d'où la fonction $u_h - P_h$ est partout négative: ie. $u_h \leq P_h$. Dans un second temps, en considérant la fonction $u_h + P_h$ et en utilisant le lemme ci-dessus, on montre comme précédemment que $u_h + P_h$ est partout positive: ie. $-P_h \leq u_h$. En conclusion on tire $|u_h|_{\infty, \Omega_h} \leq |P_h|_{\infty, \Omega_h} = R^2|f_h|_{\infty, \Omega_h}/4 + |g_h|_{\infty, \partial\Omega_h}$, d'où l'inégalité de stabilité désirée avec une constante égale à $\max\{1, R^2/4\}$. \square

2.4.5 Convergence

On conserve les notations introduites dans les sections précédentes. Une notion de convergence est introduite par:

Définition 3 Une approximation numérique (on dit aussi schéma) est dite convergente pour la norme $|\cdot|_{\overline{D}_h}$ si

$$(2.42) \quad |\rho_h u - u_h|_{\overline{D}_h} \xrightarrow{h \rightarrow 0} 0$$

On dit que l'approximation est convergente d'ordre $s > 0$ si il existe une constante $c(u)$ indépendante de h telle que

$$(2.43) \quad |\rho_h u - u_h|_{\overline{D}_h} \leq c(u)h^s$$

Muni de cette définition on peut alors montrer

Théorème 5 (i) Si L_h est une approximation consistante de L et si elle est stable alors on a convergence de u_h vers u pour la norme $|\cdot|_{\overline{D}_h}$.

(ii) Si on a la consistance à l'ordre p et la stabilité, alors on a la convergence à l'ordre p .

Démonstration: On a par définition de u_h

$$L_h(u_h) = \rho_h(f, g) = \rho_h L(u)$$

d'où

$$L_h(u_h) - L_h(\rho_h u) = \rho_h L(u) - L_h(\rho_h u)$$

La stabilité implique

$$|u_h - \rho_h u|_{\overline{D}_h} \leq c |L_h(u_h - \rho_h u)|_{\overline{D}_h} = |\rho_h L(u) - L_h(\rho_h u)|_{\overline{D}_h}$$

et la consistance (resp. la consistance à l'ordre p) implique la convergence (resp. convergence à l'ordre p). \square

2.5 Commentaires

(i) Les opérateurs de différences finies s'appuient sur des grilles structurées cartésiennes, ce qui limite leur utilisation aux domaines de géométrie simple. Dans la cas où Ω est à frontière courbe on peut plaquer une grille cartésienne sur Ω_h et approcher la frontière $\partial\Omega$ par la ligne brisée inscrite sur la grille et s'approchant au "mieux" de $\partial\Omega$. On obtient ainsi un domaine avec une frontière en escalier. Dans le cas général cette approximation de $\partial\Omega$ est d'ordre 1 si le maillage est uniforme.

Une technique plus répandue consiste à transformer le domaine physique Ω en un cube de \mathbb{R}^d (pas toujours possible pour des raisons topologiques). Dans la mesure du possible on utilise une transformation conforme.

En résumé les techniques de différences finies sont techniquement limitées par la complexité géométrique du domaine de calcul.

(ii) Une autre limite beaucoup plus structurelle vient du fait que la solution (et donc les données) doit être régulière. Une régularité $C^2(\Omega) \cap C^0(\overline{\Omega})$ est au minimum souhaitable. Cette contrainte exclue a priori les problèmes à coefficients fortement variables voire discontinus du type:

$$(2.44) \quad u - \nabla \cdot (\sigma(x) \cdot \nabla u) = f$$

où $\sigma(x)$ est une matrice symétrique définie positive (éventuellement proportionnelle à l'identité) dont les coefficients peuvent être fortement variables voire discontinus. Dans le cadre d'une simulation de la turbulence $\sigma(x)$ pourrait être une (matrice de) viscosité turbulente. Dans le cadre des problèmes de transferts thermiques à travers une interface entre deux milieux différents, $\sigma(x)$ est la diffusivité thermique de chaque milieu (différente pour chaque milieu). En fait les techniques variationnelles sont particulièrement adaptées pour traiter ce genre de problèmes, comme elles sont adaptées aussi pour traiter les problèmes à géométrie complexe.

(iii) Pour une étude détaillée des problèmes d'approximation par différences finies voir par exemple Godounov-Riabenki [11], Strikwerda [19] ou Forsythe-Wasow [9] et la littérature citée.

Chapitre 3

Formulations variationnelles des problèmes de diffusion et approximation

Dans ce chapitre on introduit la notion de formulation variationnelle des problèmes de diffusion et on montre comment approcher les solutions des problèmes variationnels par des techniques d'approximations internes: méthodes spectrales, éléments finis conformes.

3.1 Introduction

Considérons le problème de Helmholtz avec une condition de Dirichlet homogène:

$$(3.1) \quad \begin{cases} u - \nu \nabla^2 u = f \\ u|_{\partial\Omega} = 0 \end{cases}$$

En supposant que le terme source est régulier on peut dualiser l'équation par une fonction test v régulière et faire une intégration par parties (appliquer une formule de Green). On obtient alors:

$$\int_{\Omega} [uv + \nu \nabla u \cdot \nabla v] = \int_{\partial\Omega} \nu \frac{\partial u}{\partial n} v + \int_{\Omega} f v$$

Afin de respecter une certaine symétrie entre u et les fonctions tests on peut se restreindre aux fonctions tests vérifiant $v|_{\partial\Omega} = 0$, comme pour la solution cherchée. On peut alors se poser le problème suivant: existe-t-il une fonction u avec une régularité à définir et $u|_{\partial\Omega} = 0$ telle que pour toute fonction test v de régularité équivalente et vérifiant $v|_{\partial\Omega} = 0$, on ait:

$$(3.2) \quad \int_{\Omega} [uv + \nu \nabla u \cdot \nabla v] = \int_{\Omega} f v$$

Pour mieux cerner les problèmes de régularité qu'on est en droit d'exiger on peut toujours prendre $v = u$ dans l'équation ci-dessus et demander que ce qui est écrit ait un sens, c'est-à-dire que toutes les intégrales soient finies. On est ainsi conduit à chercher u dans l'espace des fonctions de carré sommable et de dérivées partielles de carré sommables. Si on retient ce cadre de régularité on doit alors vérifier s'il est suffisant pour donner un sens à la condition de Dirichlet $u|_{\partial\Omega} = 0$. En effet par analogie avec le cadre des fonctions continues, on peut

effectivement parler de $u|_{\partial\Omega}$ lorsque $u \in C^0(\overline{\Omega})$, mais $u|_{\partial\Omega}$ n'a plus de sens si on a seulement $u \in C^0(\Omega)$.

L'objectif de ce chapitre est de montrer que le problème ci-dessus avec le cadre fonctionnel qu'il implique a un sens et l'unique solution du problème en question est en fait aussi la solution du problème d'Helmholtz du début sous certaines hypothèses raisonnables. Pour atteindre cet objectif on introduit d'abord (succintement) la notion d'espace de Sobolev et on montre que dans certains de ces espaces les problèmes de trace sur la frontière $\partial\Omega$ ont un sens. Ensuite on rappelle quelques notions d'analyse hilbertienne qui permettent de démontrer le théorème de Lax-Milgram. Ce théorème est l'outil de base qui sert systématiquement à démontrer l'existence et l'unicité des solutions aux problèmes qui nous intéressent (problèmes elliptiques). Ces rappels théoriques sont suivis d'exemples illustrant à la fois la puissance et la simplicité de la méthode. On termine ce chapitre en présentant quelques aspects de la théorie de l'approximation conforme.

3.2 Cadre fonctionnel des formulations variationnelles

3.2.1 Espaces de Sobolev

L'introduction a mis en évidence l'importance des fonctions de carré sommable. On introduit donc $L^2(\Omega)$ l'espace vectoriel réel (de dimension infinie) des fonctions réelles de carré sommable sur Ω . Lorsqu'on muni $L^2(\Omega)$ du produit scalaire

$$(3.3) \quad (u, v)_0 = \int_{\Omega} uv$$

et de la norme associée $|\cdot|_0$ on confère à $L^2(\Omega)$ une structure d'espace de Hilbert. C'est-à-dire que, modulo le problème de la dimension infinie, on peut faire dans $L^2(\Omega)$ de la géométrie comme dans un espace euclidien (ie. $L^2(\Omega)$ est préhilbertien) et les suites de Cauchy de fonctions de carré sommable convergent vers des fonctions de carré sommable ($L^2(\Omega)$ est complet). Voir l'ouvrage de Rudin [18] pour une revue détaillée des problèmes connexes à la structure hilbertienne de $L^2(\Omega)$. Lorsqu'aucune ambiguïté est possible on note (\cdot, \cdot) le produit scalaire dans $L^2(\Omega)$.

A partir de $L^2(\Omega)$ on construit une multitude de sous-espaces. Par exemple on définit $H^1(\Omega)$ tel que

$$(3.4) \quad H^1(\Omega) = \{u \in L^2(\Omega), \frac{\partial u}{\partial x_i} \in L^2(\Omega), i = 1, \dots, d\}$$

Les dérivées étant présent au sens des distributions. En pratique on fait comme si les dérivées existaient et on vérifie après coup si elles sont de carré sommable (en réalité on vérifie si les dérivées en question sont des formes linéaires continues de $L^2(\Omega)$, et si c'est le cas on les identifie à des fonctions de $L^2(\Omega)$). $H^1(\Omega)$ muni du produit scalaire

$$(3.5) \quad (u, v)_1 = \int_{\Omega} uv + \int_{\Omega} \nabla u \cdot \nabla v$$

est un espace de Hilbert.

Plus généralement on introduit les espaces $H^m(\Omega)$ constitués des fonctions de carré sommable et dont toutes les dérivées partielles jusqu'à l'ordre m sont aussi de carré sommable. Muni du produit scalaire correspondant ce sont des espaces de Hilbert. Cette famille d'espace appartient à la famille plus large des espaces dits de Sobolev dont les espaces de base sont les $L^p(\Omega)$, $1 \leq p \leq \infty$. Les espaces de Sobolev construits à partir de $L^2(\Omega)$ (ie. les $H^m(\Omega)$)

sont particulièrement intéressants car ce sont les seuls à posséder une structure hilbertienne. Pour travailler avec des notations homogènes on note parfois $H^0(\Omega) = L^2(\Omega)$. Pour d'autres détails sur les espaces de Sobolev le lecteur est renvoyé à Adams [1], Aubin [2], Brézis [5], Lions-Magenes [14] ou encore Raviart-Thomas [17].

3.2.2 Les problèmes de trace

Pour pouvoir parler de condition de Dirichlet ou de Neumann sur une fonction il faut que la fonction en question ait un minimum de régularité. Par exemple la condition de Dirichlet (resp. Neumann) a un sens si $u \in C^0(\bar{\Omega})$ (resp. $C^1(\bar{\Omega})$), par contre elle n'en a plus si u est seulement dans $C^0(\Omega)$ (resp. $C^1(\Omega)$). Plus généralement quel sens peut-on donner aux conditions de Dirichlet/Neumann dans les espaces $H^m(\Omega)$?

Avant d'entamer l'étude théorique du problèmes de traces, analysons l'exemple suivant. Considérons la fonction $u(x_1, x_2) = x_1^k$ avec $-1/2 < k < 0$ et $(x_1, x_2) \in \Omega =]0, 1[\times]0, 1[$. La fonction u est manifestement dans $L^2(\Omega)$ par contre elle tend vers l'infini en approchant le segment $x_1 = 0$. C'est-à-dire qu'on ne peut pas définir la trace de u sur le segment $x_1 = 0$. Il est donc clair que l'appartenance à $L^2(\Omega)$ n'est pas suffisante pour parler de trace sur $\partial\Omega$.

En fait le cadre des fonctions de $C^0(\bar{\Omega})$ est un très agréable pour traiter le problème des traces, il serait donc judicieux d'y revenir. Un premier pas dans cette direction peut être fait à l'aide de:

Théorème 6 *Pour tout $j \geq 0$, $C^j(\bar{\Omega})$ est dense dans $H^m(\Omega)$ pour $m \geq 0$.*

Ce théorème signifie que pour toute fonction u de $H^m(\Omega)$ on peut trouver une suite de Cauchy (dans $H^m(\Omega)$) de fonctions de $C^j(\bar{\Omega})$ qui approchent aussi près qu'on le souhaite la fonction u . Pour chaque élément de la suite de Cauchy en question on peut définir sa trace sur $\partial\Omega$ et la trace de toutes ses dérivées jusqu'à l'ordre j . La question qui se pose alors est de savoir si les suites des traces en question convergent en un certain sens lorsque les éléments de la suite approchent u de mieux en mieux. En d'autres termes, peut-on trouver un espace raisonnable (par exemple un sous-espace de $L^2(\partial\Omega)$) dans lequel les suites des traces soient des suites de Cauchy? Dans l'affirmative on pourrait raisonnablement définir la trace de u et la trace de ses dérivées comme étant les limites des traces correspondantes des fonctions de la suite de $C^j(\bar{\Omega})$ qui approchent u . Dans le cas où $m = 0$ la réponse à l'existence d'un tel sous-espace est manifestement négative d'après le contre-exemple fourni précédemment. Par contre pour $m \geq 1$ on a le résultat positif suivant (voir Lions-Magenes [14, p. 38]):

Théorème 7 (i) *Soit Ω un ouvert borné régulier, toute fonction u de $H^1(\Omega)$ a une trace sur $\partial\Omega$, cette trace est dans $L^2(\partial\Omega)$ et c'est la limite dans $L^2(\partial\Omega)$ des traces des suites de Cauchy de fonctions de $C^0(\bar{\Omega})$ qui convergent vers u au sens de la topologie de $H^1(\Omega)$. En d'autres termes pour les initiés, l'application linéaire*

$$(3.6) \quad \begin{aligned} \gamma_0 : C^0(\bar{\Omega}) \subset H^1(\Omega) &\longrightarrow L^2(\partial\Omega) \\ u &\longmapsto u|_{\partial\Omega} \end{aligned}$$

est continue (uniformément en fait) et se prolonge donc par densité à $H^1(\Omega)$.

(ii) *De même pour $m \geq 2$, et pour $0 \leq j \leq m - 1$ l'application linéaire*

$$(3.7) \quad \begin{aligned} \gamma_j : C^j(\bar{\Omega}) \subset H^m(\Omega) &\longrightarrow L^2(\partial\Omega) \\ u &\longmapsto \partial^j u / \partial n^j|_{\partial\Omega} \end{aligned}$$

est continue (uniformément en fait) et se prolonge donc par densité à $H^m(\Omega)$.

En fait les traces sur $\partial\Omega$ des fonctions de $H^1(\Omega)$ ont un peu plus de régularité que $L^2(\partial\Omega)$ et un peu moins que $H^1(\partial\Omega)$ (espace défini comme $H^1(\Omega)$ mais avec les dérivées tangentielles). Avec un peu d'effort on peut construire par interpolation un espace de Hilbert coincé entre $L^2(\partial\Omega)$ et $H^1(\partial\Omega)$ et qui contient toutes les traces des fonctions de $H^1(\Omega)$, cet espace est conventionnellement noté $H^{1/2}(\partial\Omega)$ (cf. Lions-Magenes [14]) et est équipé de la norme

$$(3.8) \quad |f|_{1/2,\partial\Omega} = \inf_{\substack{v \in H^1(\Omega) \\ v|_{\partial\Omega} = f}} |v|_{1,\Omega}$$

Ainsi toutes les fonctions de $H^1(\Omega)$ ont une trace sur $\partial\Omega$ et cette trace est dans $H^{1/2}(\partial\Omega)$. Pour les fonctions de $H^2(\Omega)$ on peut définir la trace de la dérivée normale et on peut montrer qu'elle appartient à $H^{1/2}(\partial\Omega)$ alors que la trace de la fonction quant à elle appartient à un espace qu'on note conventionnellement $H^{3/2}(\partial\Omega)$ (parce qu'on peut montrer que dans un certain sens il est coincé entre $H^1(\partial\Omega)$ et $H^2(\partial\Omega)$).

Un des intérêts des résultats de trace ci-dessus est de permettre la généralisation des formules d'intégration par parties, valables pour les fonctions continues, aux fonctions de $H^m(\Omega)$. Par exemple, si $u \in H^2(\Omega)$ et $v \in H^1(\Omega)$, le théorème 7 permet de donner un sens à la formule de Green suivante

$$(3.9) \quad \int_{\Omega} v \nabla^2 u = - \int_{\Omega} \nabla u \cdot \nabla v + \int_{\partial\Omega} \frac{\partial u}{\partial n} v$$

Cette formule n'est toutefois pas encore très pratique car elle exige que la fonction u soit dans $H^2(\Omega)$ (ie. une condition sur la dérivée seconde) alors que la formulation variationnelle du problème de Helmholtz met uniquement à notre disposition des fonctions de $H^1(\Omega)$. En regardant la formule d'intégration par partie ci-dessus d'un peu plus près on remarque que si u est dans $H^1(\Omega)$ et si $\nabla^2 u$ est dans $L^2(\Omega)$ alors les termes de l'équation ont tous un sens sauf $\partial u / \partial n$, mais comme deux des trois intégrales sont finies la troisième doit l'être aussi. En d'autres termes, il est certainement possible de donner un sens à la dérivée normale de u . Dans cette perspective introduisons $H^{-1/2}(\partial\Omega)$ le dual de $H^{1/2}(\partial\Omega)$, c'est-à-dire l'espace des formes linéaires continues sur $H^{1/2}(\partial\Omega)$. Cet espace est muni de la norme

$$(3.10) \quad |f|_{-1/2,\partial\Omega} = \sup_{g \in H^{1/2}(\partial\Omega) - \{0\}} \frac{(f, g)}{|g|_{1/2,\partial\Omega}}$$

où (f, g) désigne le résultat dans \mathbb{R} de l'action de f sur g . Modulo quelques subtilités qu'on explicitera par la suite, on identifie l'action de f sur g au produit scalaire dans $L^2(\partial\Omega)$. Montrons maintenant

Proposition 4 *Pour $u \in H^1(\Omega)$ tel que $\nabla^2 u \in L^2(\Omega)$, la dérivée normale $\partial u / \partial n$ a un sens dans $H^{-1/2}(\partial\Omega)$.*

Démonstration: D'après la définition de la norme dans $H^{-1/2}(\partial\Omega)$ on a

$$\left| \frac{\partial u}{\partial n} \right|_{-1/2,\partial\Omega} = \sup_{f \in H^{1/2}(\partial\Omega) - \{0\}} \frac{(\partial u / \partial n, f)}{|f|_{1/2,\partial\Omega}} = \sup_{f \in H^{1/2}(\partial\Omega) - \{0\}} \inf_{\substack{v \in H^1(\Omega) - \{0\} \\ v|_{\partial\Omega} = f}} \frac{(\partial u / \partial n, v)}{|f|_{1/2,\partial\Omega}}$$

et à l'aide de la formule d'intégration par parties (3.9) et en utilisant les inégalités de Cauchy-Schwarz on obtient

$$\left| \frac{\partial u}{\partial n} \right|_{-1/2, \partial\Omega} \leq \sup_{f \in H^{1/2}(\partial\Omega) - \{0\}} \frac{1}{|f|_{1/2, \partial\Omega}} \inf_{\substack{v \in H^1(\Omega) - \{0\} \\ v|_{\partial\Omega} = f}} [|\nabla^2 u|_{0, \Omega} + |u|_{1, \Omega}] |v|_{1, \Omega}$$

La définition (3.8) de la norme 1/2 de f fournie finalement

$$(3.11) \quad \left| \frac{\partial u}{\partial n} \right|_{-1/2, \partial\Omega} \leq |\nabla^2 u|_{0, \Omega} + |u|_{1, \Omega} < \infty$$

Ce qui prouve la proposition. \square

Afin de fixer les idées et montrer l'utilité de la dualité, poursuivons cet exercice de style un peu plus loin. Considérons une fonction à valeur vectorielle dans \mathbb{R}^d de carré sommable et de divergence de carré sommable (ie. $u \in L^2(\Omega)^d$ et $\nabla \cdot u \in L^2(\Omega)$). On se pose la question de savoir si la trace normale $u \cdot n$ a un sens. Pour ce faire on utilise la formule d'intégration par parties qui fait intervenir la divergence et la trace normale:

$$(3.12) \quad \int_{\Omega} p \nabla \cdot u = - \int_{\Omega} u \cdot \nabla p + \int_{\partial\Omega} (u \cdot n) p$$

Si on prend $p \in H^1(\Omega)$, les deux intégrales de volume ont un sens et dans l'intégrale de bord seul $p|_{\partial\Omega}$ a un sens pour l'instant dans $H^{1/2}(\partial\Omega)$. Le lecteur pourra vérifier en procédant comme pour la proposition précédente que $u \cdot n$ a un sens dans $H^{-1/2}(\partial\Omega)$. Remarquer qu'on ne peut absolument rien dire sur la trace tangentielle de u . Le contre-exemple suivant montre en fait que la trace tangentielle n'a pas de sens a priori si u ne possède pas plus de régularité. Soit $\Omega =]0, 1[\times]0, 1[$ et $u(x_1, x_2) = (0, x_1^k)$ avec $-1/2 < k < 0$. On vérifie que u est de carré sommable et à divergence nulle, donc de carré sommable. La trace normale a partout un sens sauf sur l'ensemble négligeable constitué des deux coins $\{(0, 0), (0, 1)\}$. Par contre la trace tangentielle $u_2(x_1, x_2) = x_1^k$ est infinie sur tout le segment $x_1 = 0$.

Pour finir cette sensibilisation aux problèmes des traces, le lecteur pourra vérifier en se servant de la formule de Stokes

$$(3.13) \quad \int u \cdot \text{rot}(v) = \int \text{rot}(u) \cdot v + \int_{\partial\Omega} (u \times n) \cdot v$$

qu'on peut donner un sens dans $H^{-1/2}(\partial\Omega)^d$ à la trace tangentielle $u \times n$, si u est de carré sommable et de rotationnel de carré sommable.

Pour une étude détaillée de tous ces problèmes de traces le lecteur est renvoyé à Lions-Magenes [14] et à Girault-Raviart [10].

3.2.3 Rappels d'analyse Hilbertienne

On désigne par H un espace de Hilbert muni du produit scalaire $(\cdot, \cdot)_H$, c'est-à-dire que H est préhilbertien et complet pour la topologie induite par la norme $|\cdot|_H$ associée au produit scalaire. On introduit et on note H' le dual de H , (espace des formes linéaires continues sur H). H' est un espace de Hilbert lorsqu'on le muni de la norme

$$(3.14) \quad |f|_{H'} = \sup_{v \in H - \{0\}} \frac{f(v)}{|v|_H}$$

où $f(v)$ désigne le résultat dans \mathbb{R} de l'action de f sur v . L'ensemble des propriétés essentielles des espaces de Hilbert résultent de l'existence d'une projection sur tout convexe fermé non vide de H .

Théorème 8 (cf. Brézis [5, p. 79]) – Soit $K \subset H$ un convexe fermé non vide. Alors pour tout $f \in H$ il existe un unique $u \in K$ tel que

$$(3.15) \quad |f - u|_H = \inf_{v \in K} |f - v|_K$$

Cette projection qu'on note $u = P_K(f)$ est continue et satisfait

$$(3.16) \quad |P_K(f_1) - P_K(f_2)|_H \leq |f_1 - f_2|_H$$

Elle est de plus caractérisée par

$$(3.17) \quad \forall v \in K, \quad (f - u, v - u)_H \leq 0$$

Lorsque K est un sous-espace fermé de H , la caractérisation de la projection devient

$$(3.18) \quad \forall v \in K, \quad (f - u, v)_H = 0$$

Une application importante de ce résultat est le théorème de représentation de Riesz-Frechet qui permet d'identifier tout élément du dual H' à un élément de H et ce de façon unique:

Théorème 9 (Riesz-Frechet) (cf. Brézis [5, p.]) – Etant donné $f \in H'$, il existe un unique $u \in H$ tel que

$$(3.19) \quad \forall v \in H \quad f(v) = (u, v)_H$$

De plus, la correspondance est isométrique, c'est à dire $|f|_{H'} = |u|_H$.

Remarque 1: En dimension finie ce théorème est utilisé naturellement lorsqu'on identifie la différentielle d'une fonction avec son vecteur gradient.

Remarque 2: Par la suite on fait systématiquement l'identification de $L^2(\Omega)$ avec son dual.

Remarque 3: Identifier ou ne pas identifier H à H' ? Si on travaille avec un seul espace H il n'y a pas lieu de distinguer H et H' . Par contre si on travaille avec plusieurs espaces imbriqués les uns dans les autres, disons deux pour fixer les idées, il peut être intéressant de faire la distinction. Par exemple considérons $V \subset H$ deux espaces de Hilbert le premier étant dense dans le second et se plongeant continuellement dans le second. Il peut être intéressant d'identifier les formes linéaires sur H aux fonctions de H et l'action de ces formes linéaires au produit scalaire de H (c'est possible d'après le théorème de Riesz). Remarquons à ce niveau qu'il est alors possible de plonger H , et donc H' puisqu'on a fait l'identification, dans V' en identifiant l'action des formes linéaires de V' sur V à un prolongement du produit scalaire de H . Pour ce faire, on considère toute fonction f de H comme un élément de H' en posant $f(u) = (f, u)_H$ pour tout $u \in V$. Comme V est dense dans H on peut montrer que H est aussi dense dans V' et en conséquence pour tout f dans V' on peut trouver une suite de Cauchy (f_n) de fonctions de H qui convergent (au sens de H') vers f . On peut alors représenter l'action de f sur tout $u \in V$ comme la limite des produits scalaires $(f_n, u)_H$. En pratique il faut donc imaginer l'action des éléments de V' sur V comme le prolongement du produit scalaire de H . On a donc construit la suite $V \subset H \equiv H' \subset V'$ où les injections sont denses et continues.

Remarque 4: Le lecteur intéressé par ces notions aura remarquer que le théorème de Riesz-Fréchet met en perspective une propriété spécifique des espaces de dimension infinie. Supposons un instant que $V \neq H$. D'après ce qui précède on a $V \subset V'$ mais $V \neq V'$. D'autre part le théorème de Riesz nous dit que V' et V sont isomorphes. Nous sommes donc en présence d'un espace V' qui est isomorphe à un sous-espace V qu'il contient strictement. Cette propriété un peu surprenante au premier abord est spécifique des ensembles de dimension infinie. C'est-à-dire qu'elle ne peut être vraie que si V' (et donc V et H) est de dimension infinie. On ne pourra donc avoir injection continue de V dans H avec densité et $V \neq H$ que si V est de dimension infinie. Penser à \mathbb{Q} qui contient \mathbb{N} strictement et qui est aussi en bijection avec \mathbb{N} .

Pour nos besoins ultérieurs concernant l'analyse des problèmes avec des conditions de Dirichlet homogènes il est utile de définir le sous-espace de $H^1(\Omega)$ des fonctions dont la trace sur $\partial\Omega$ est nulle (c'est possible d'après les théorèmes de trace). On note cet espace $H_0^1(\Omega)$, et on peut montrer

Proposition 5 $C_c^\infty(\Omega)$, l'espace des fonction C^∞ à support compact sur Ω , est dense dans $H_0^1(\Omega)$

Remarque 1: Dans les présentations mathématiques des problèmes de trace, cette proposition est prise comme définition de $H_0^1(\Omega)$ et on interprète a posteriori cet espace comme étant l'espace des fonctions de trace nulle sur $\partial\Omega$ si Ω est régulier. La présentation inverse a été choisie car elle semble plus intuitive (elle est toutefois un peu moins puissante).

Remarque 2: Sachant que $C_c^\infty(\Omega)$ est dense dans $L^2(\Omega)$ on tire de la proposition ci-dessus que $H_0^1(\Omega)$ est lui aussi dense dans $L^2(\Omega)$.

Exemple 1: Pour des raisons techniques et un certain souci d'optimalité des résultats à venir il est utile d'introduire le dual de $H_0^1(\Omega)$, c'est-à-dire l'espace des formes linéaires continues sur $H_0^1(\Omega)$. On note $H^{-1}(\Omega)$ cet espace. Par la suite on identifie le dual de $L^2(\Omega)$ avec $L^2(\Omega)$, c'est-à-dire qu'on identifie les formes linéaires sur $L^2(\Omega)$ aux fonctions de $L^2(\Omega)$ et l'action de ces formes linéaires au produit scalaire de $L^2(\Omega)$ (c'est tout à fait possible d'après le théorème de Riesz). Appliquons maintenant la théorie de l'identification qui a été développée plus haut pour deux espaces emboîtés. Pour pouvoir plonger $L^2(\Omega)$ dans $H^{-1}(\Omega)$ on identifie l'action des formes linéaires de $H^{-1}(\Omega)$ à un prolongement du produit scalaire de $L^2(\Omega)$. En d'autres termes on considère toute fonction f de $L^2(\Omega)$ comme un élément de $H^{-1}(\Omega)$ en posant $f(u) = \int_\Omega fu$ pour tout $u \in H_0^1(\Omega)$. Comme $H_0^1(\Omega)$ est dense dans $L^2(\Omega)$ (cf. remarque ci-dessus) on peut montrer que $L^2(\Omega)$ est aussi dense dans $H^{-1}(\Omega)$ et en conséquence pour tout f dans $H^{-1}(\Omega)$ on peut trouver une suite de Cauchy (f_n) de fonctions de $L^2(\Omega)$ qui convergent vers f (au sens de de la topologie de $H^{-1}(\Omega)$). On peut alors représenter l'action de f sur tout $u \in H_0^1(\Omega)$ comme la limite des produits scalaires (f_n, u) . En pratique il faut donc imaginer l'action des éléments de $H^{-1}(\Omega)$ sur $H_0^1(\Omega)$ comme le prolongement du produit scalaire de $L^2(\Omega)$. L'espace $H^{-1}(\Omega)$ est muni de la norme

$$(3.20) \quad |f|_{-1,\Omega} = \sup_{v \in H_0^1(\Omega) - \{0\}} \frac{(f, v)}{|v|_{1,\Omega}}$$

où (f, v) est, modulo les subtilités ci-dessus, le produit scalaire de $L^2(\Omega)$.

Définition 4 On dit qu'une forme bilinéaire $a(u, v) : H \times H \rightarrow \mathbb{R}$ est continue si il existe une constante c telle que $a(u, v) \leq c|u|_H|v|_H$. On dit qu'elle est coercive s'il existe une constante $\alpha > 0$ telle que

$$(3.21) \quad a(u, u) \geq \alpha|u|_H^2$$

Une conséquence importante du théorème de Riesz est le théorème de Lax-Milgram sur lequel repose une bonne partie de la théorie variationnelle des problèmes elliptiques.

Théorème 10 (Lax-Milgram) (cf. Brézis [5, p. 84]) – Soit $a(u, v)$ une forme bilinéaire, continue et coercive. Alors pour tout $f \in H'$ il existe $u \in H$ unique tel que

$$(3.22) \quad \forall v \in H \quad a(u, v) = f(v).$$

De plus il y a continuité de la solution par rapport à la donnée f au sens où u satisfait l'inégalité de stabilité

$$(3.23) \quad |u|_H \leq \frac{|f|_{H'}}{\alpha}$$

Démonstration : Voir par exemple Dautray-Lions [8, t.4 p. 1217]. \square

Donnons quelques exemples de formes bilinéaires continues coercives (on dit aussi H -elliptique).

Exemple 1: Soit $H = H^1(\Omega)$ et posons $a(u, v) = (u, v)_0 + \nu(\nabla u, \nabla v)_0$. On a trivialement la continuité de a et la coercivité est obtenue par

$$(3.24) \quad a(u, u) = |u|_0^2 + \nu|\nabla u|_0^2 \geq \min(1, \nu)|u|_1^2$$

Exemple 2: Inégalité de Poincaré dans $H^1(\Omega)$ (cf. Dautray-Lions [8, t.3 p. 922]). Soit Ω un ouvert borné régulier, il existe $c > 0$ tel que pour tout $u \in H^1(\Omega)$:

$$(3.25) \quad |u|_0^2 - \frac{1}{\text{mes}(\Omega)} \left| \int_{\Omega} u \right|^2 \leq c|\nabla u|_0^2$$

d'où la forme $a(u, v) = (\nabla u, \nabla v)$ est coercive sur le sous espace de $H^1(\Omega)$ des fonctions de moyenne nulle. Soit $H^1(\Omega)/\mathbb{R} = \{u \in H^1(\Omega), \int_{\Omega} u = 0\}$ cet espace. Pour les fonctions de $H^1(\Omega)/\mathbb{R}$ l'inégalité de Poincaré donne $a(u, u) \geq |u|_1^2/(1+c)$.

Exemple 3: Inégalité de Poincaré dans $H_0^1(\Omega)$ (cf. Dautray-Lions [8, t.3 p. 920]). Soit Ω un ouvert contenu dans une bande de \mathbb{R}^d de largeur l , on a alors

$$(3.26) \quad |u|_0^2 \leq \frac{l^2}{2} |\nabla u|_0^2$$

d'où la coercivité de $a(u, v) = (\nabla u, \nabla v)$ sur $H_0^1(\Omega)$. En effet l'inégalité de Poincaré implique $a(u, u) \geq |u|_1^2/(1+l^2/2)$.

3.3 Exemples de problèmes variationnels

3.3.1 Problème de Dirichlet

Revenons maintenant au problème de Helmholtz avec une condition de Dirichlet homogène considéré en introduction de ce chapitre.

$$(3.27) \quad \begin{cases} u - \nu \nabla^2 u = f \\ u|_{\partial\Omega} = 0 \end{cases}$$

En dualisant l'EDP par des fonctions tests régulières v nulles sur le bord et en intégrant par parties on est conduit à l'égalité:

$$(u, v) + \nu(\nabla u, \nabla v) = (f, v)$$

Pour que tout ceci ait un sens, d'après les sections qui précèdent il suffit de prendre v dans $H_0^1(\Omega)$ et chercher u dans $H_0^1(\Omega)$. Le second membre (f, v) a un sens si f est dans $L^2(\Omega)$, il a même un sens si f est identifiable à une forme linéaire continue de $H_0^1(\Omega)$ c'est-à-dire si $f \in H^{-1}(\Omega)$.

On pose donc le problème suivant: pour $f \in L^2(\Omega)$ (ou même dans $f \in H^{-1}(\Omega)$) chercher $u \in H_0^1(\Omega)$ tel que

$$(3.28) \quad \forall v \in H_0^1(\Omega) \quad (u, v) + \nu(\nabla u, \nabla v) = (f, v)$$

Théorème 11 *Le problème ci-dessus a une solution unique. De plus cette solution satisfait les inégalités de stabilité*

$$(3.29) \quad |u|_0 \leq |f|_0, \text{ et } |u|_1 \leq \min(1, \nu)|f|_{-1}$$

Démonstration : (i) La continuité des formes linéaires et bilinéaires impliquées est évidente, la coercivité a été montrée dans la section précédente. L'unicité et l'existence résultent du théorème de Lax-Milgram.

(ii) Les inégalités de stabilité sont obtenues en prenant $v = u$ dans (3.28) et en utilisant l'inégalité de Cauchy-schwarz.

La connexion entre le problème aux dérivées partielles (3.27) et la formulation variationnelle (3.28) est donnée par

Proposition 6 *Si $u \in H^2(\Omega)$ on a l'équivalence entre les propositions suivantes:*

- (i) *u est solution du problème $u - \nu \nabla^2 u = f$ dans $L^2(\Omega)$ avec $u|_{\partial\Omega} = 0$.*
- (ii) *u est solution du problème variationnel (3.28).*

Démonstration : (i) \Rightarrow (ii) a déjà été montré.

(ii) Soit $u \in H^2(\Omega)$ solution du problème variationnel. En intégrant (3.28) par parties on obtient

$$\forall v \in H_0^1(\Omega), \quad \int_{\Omega} [u - \nu \nabla^2 u - f]v = 0$$

Or $H_0^1(\Omega)$ étant dense dans $L^2(\Omega)$, l'égalité ci-dessus est valable pour tout v dans $L^2(\Omega)$, c'est-à-dire $u - \nu \nabla^2 u = f$ dans $L^2(\Omega)$. La condition limite $u|_{\partial\Omega}$ résulte de l'appartenance de u à $H_0^1(\Omega)$. \square

Si au lieu de supposer $u \in H^2(\Omega)$ on suppose a posteriori $u \in C^2(\Omega) \cap C^1(\bar{\Omega})$, en procédant de la même façon que précédemment on montre l'égalité de la solution variationnelle et de la solution classique.

Remarque: Le théorème 11 montre que $Id - \nu \nabla^2$ est un isomorphisme de $H_0^1(\Omega)$ sur $H^{-1}(\Omega)$. En fait, en utilisant les mêmes arguments que précédemment on peut aussi montrer que $-\nu \nabla^2$ est un isomorphisme de $H_0^1(\Omega)$ sur $H^{-1}(\Omega)$. Ceci nous permet a posteriori de caractériser les fonctions (on dit aussi distributions) de $H^{-1}(\Omega)$: elles sont toutes de la forme $f = -\nabla^2 u$ avec $u \in H_0^1(\Omega)$. Par exemple le lecteur peut vérifier que $f(x) = x^{-1-1/k}$ avec $2 < k$ est dans $H^{-1}(]0, 1[)$ puisqu'à une constante près c'est la dérivée seconde de $u(x) = x^{1-1/k}$ qui est dans $H_0^1(]0, 1[)$. Cet exemple permet de démystifier un peu l'espace $H^{-1}(\Omega)$.

En conclusion, la méthode variationnelle donne bien la solution du problème qu'on s'était posé. Elle a de plus l'avantage de fournir une solution même s'il n'existe pas de solution classique (ie. très régulière $u \in C^2(\Omega) \cap C^0(\bar{\Omega})$). Alors que la démonstration de l'existence d'une solution classique est très difficile, l'obtention d'une solution variationnelle (moins régulière a priori) repose quant à elle sur des arguments (énergétiques) simples.

3.3.2 Problèmes de transmission

Un autre avantage significatif de la méthode variationnelle se révèle lorsqu'on étudie des problèmes du type

$$(3.30) \quad \begin{cases} -\nabla \cdot (\sigma(x) \nabla u) = f \\ u|_{\partial\Omega} = 0 \end{cases}$$

où $\sigma(x)$ est une fonction discontinue. On suppose par la suite que Ω est partitionné en deux sous-domaines $\Omega = \Omega_1 \cup \Omega_2$ et que $\sigma(x) > 0$ est une fonction régulière sur chaque sous-domaine Ω_1 et Ω_2 mais discontinue à travers l'interface $\partial\Omega_{12}$ les séparant. Ce problème correspond par exemple à l'équilibre de la température dans un domaine composé de deux milieux de diffusivité thermique différente. Dans ce cas l'approche classique consiste à séparer les milieux et à résoudre le problème de couplage suivant:

$$(3.31) \quad \begin{cases} -\nabla \cdot (\sigma_1(x) \nabla u_1) = f_1 \\ -\nabla \cdot (\sigma_2(x) \nabla u_2) = f_2 \\ u_1|_{\partial\Omega} = 0 \\ u_2|_{\partial\Omega} = 0 \\ u_1|_{\partial\Omega_{12}} = u_2|_{\partial\Omega_{12}} \\ \sigma_1(x) \partial u_1 / \partial n_1|_{\partial\Omega_{12}} = -\sigma_2(x) \partial u_2 / \partial n_2|_{\partial\Omega_{12}} \end{cases}$$

Ce problème peut devenir très compliqué si l'interface $\partial\Omega_{12}$ n'est pas connue explicitement.

En supposant $f \in L^2(\Omega)$ (voire dans $H^{-1}(\Omega)$), l'approche variationnelle consiste à chercher u dans $H_0^1(\Omega)$ tel que

$$(3.32) \quad \forall v \in H_0^1(\Omega) \quad \int_{\Omega} \sigma(x) \nabla u \cdot \nabla v = \int_{\Omega} f v$$

Théorème 12 *Le problème (3.32) a une solution unique. De plus la solution satisfait l'inégalité de stabilité*

$$(3.33) \quad |u|_1 \leq \frac{1}{\inf_{x \in \Omega} \sigma(x)} |f|_{-1}$$

Démonstration : Tout ceci résulte du théorème de Lax-Milgram. \square

Proposition 7 *Sous réserve de régularité suffisante, la solution du problème variationnel est solution du problème de couplage (3.31).*

Démonstration : En prenant $v \in C_c^\infty(\Omega_1)$, en intégrant par parties l'égalité variationnelle, et par densité de $C_c^\infty(\Omega_1)$ dans $L^2(\Omega_1)$ on obtient l'égalité $-\nabla \cdot (\sigma_1(x) \nabla u_1) = f_1$ dans $L^2(\Omega_1)$. On procède de la même façon dans Ω_2 .

Montrons maintenant l'égalité $u_1|_{\partial\Omega_{12}} = u_2|_{\partial\Omega_{12}}$. Soit $v \in C_c^\infty(\Omega)$, on a:

$$\begin{aligned} - \int_{\Omega} u \nabla^2 v &= \int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega_1} \nabla u_1 \cdot \nabla v + \int_{\Omega_2} \nabla u_2 \cdot \nabla v \\ &= - \int_{\Omega_1} u_1 \nabla^2 v - \int_{\Omega_2} u_2 \nabla^2 v + \int_{\partial\Omega_{12}} u_1 \partial v / \partial n_1 + \int_{\partial\Omega_{12}} u_2 \partial v / \partial n_2 \\ &= - \int_{\Omega} u \nabla^2 v + \int_{\partial\Omega_{12}} [u_1 - u_2] \partial v / \partial n_1 \end{aligned}$$

d'où on tire $u_1|_{\partial\Omega_{12}} = u_2|_{\partial\Omega_{12}}$ (en un sens qu'il faudrait préciser) en admettant que $\partial v/\partial n_1|_{\partial\Omega_{12}}$ parcourt un sous-espace dense de $L^2(\partial\Omega_{12})$ lorsque v parcourt $C_c^\infty(\Omega_1)$.

On montre maintenant l'égalité des flux. En prenant encore une fois $v \in C_c^\infty(\Omega)$, on a :

$$\begin{aligned} \int_{\Omega} f v &= - \int_{\Omega_1} \nabla \cdot (\sigma_1(x) \nabla u_1) v - \int_{\Omega_2} \nabla \cdot (\sigma_2(x) \nabla u_2) v \\ &= \int_{\Omega_1} \sigma_1(x) \nabla u_1 \cdot \nabla v + \int_{\Omega_2} \sigma_2(x) \nabla u_2 \cdot \nabla v - \int_{\partial\Omega_{12}} \sigma_1(x) v \partial u_1 / \partial n_1 - \int_{\partial\Omega_{12}} \sigma_2(x) v \partial u_2 / \partial n_2 \\ (f, v)_0 &= a(u, v) - \int_{\partial\Omega_{12}} [\sigma_1(x) \partial u_1 / \partial n_1 + \sigma_2(x) \partial u_2 / \partial n_2] v \end{aligned}$$

D'où l'égalité des flux $\sigma_1(x) \partial u_1 / \partial n_1 = -\sigma_2(x) \partial u_2 / \partial n_2$ en vertu du fait que u est solution du problème variationnel $a(u, v) = (f, v)_0$.

3.3.3 Problème mêlé

Pour finir la série des exemples considérons le problème suivant :

$$(3.34) \quad \begin{cases} u - \nu \nabla^2 u = f \\ u|_{\partial\Omega_1} = 0 \\ \frac{\partial u}{\partial n}|_{\partial\Omega_2} = g \\ \lambda(x) u|_{\partial\Omega_3} + \frac{\partial u}{\partial n}|_{\partial\Omega_3} = 0 \end{cases}$$

où $(\partial\Omega_1, \partial\Omega_2, \partial\Omega_3)$ est une partition de $\partial\Omega$. Les conditions de Dirichlet et de Fourier sont prises homogènes pour simplifier la présentation technique des arguments. Tout ce qui suit peut être repris avec des conditions non-homogènes. La formulation variationnelle du problème est obtenue en dualisant l'EDP par des fonctions v de $H^1(\Omega)$ nulles sur $\partial\Omega_1$ et en intégrant par parties. Introduisons l'espace de Hilbert

$$(3.35) \quad V = \{v \in H^1(\Omega), v|_{\partial\Omega_1} = 0\}$$

et la forme bilinéaire continue

$$(3.36) \quad a(u, v) = \int_{\Omega} [uv + \nu \nabla u \cdot \nabla v] + \int_{\partial\Omega_3} \lambda(x) uv$$

Pour f dans $L^2(\Omega)$ et g dans $L^2(\partial\Omega_2)$, on cherche $u \in V$ tel que :

$$(3.37) \quad \forall v \in V, \quad a(u, v) = \int_{\Omega} f v + \int_{\partial\Omega_2} g v$$

Théorème 13 *Si $\inf_{x \in \partial\Omega_3} \lambda(x) > 0$, le problème ci-dessus a une solution unique, et cette solution vérifie l'inégalité de stabilité :*

$$(3.38) \quad |u|_{1, \Omega} \leq \min(1, \nu) (|f|_{0, \Omega} + |g|_{0, \partial\Omega_2})$$

Démonstration : Vérifier que a est continue et coercive et appliquer le théorème de Lax-Milgram. \square

L'équivalence de la formulation variationnelle avec la formulation classique se fait comme pour les autres exemples.

3.3.4 Régularité des solutions variationnelles

Après avoir trouver assez facilement une solution variationnelle à notre problème, il peut être satisfaisant intellectuellement de se poser la question de savoir si la solution trouvée ne possède pas en fin de compte plus de régularité qu'on en a exigé. On peut même essayer de voir si ce n'est pas en fait la solution classique au problème.

Une première réponse allant dans ce sens est donnée par un théorème indiquant la connexion qui existe entre les espaces $H^m(\Omega)$ et $C^k(\bar{\Omega})$:

Théorème 14 *Si Ω est un ouvert borné régulier de \mathbb{R}^d , et $m > n/2 + k$ alors $H^m(\Omega)$ s'injecte continuellement dans $C^k(\bar{\Omega})$.*

Les problèmes de régularité globale des solution variationnelles sont très techniques et font intervenir de façon déterminante la régularité du bord. A titre d'exemple donnons un résultat de régularité pour le problème de Dirichlet:

Théorème 15 *Si $f \in H^m(\Omega)$, et $g \in L^2(\partial\Omega)$ est tel que il existe un relèvement $\tilde{g} \in H^n(\Omega)$ avec $n \geq 1$ et $\tilde{g}|_{\partial\Omega} = g$, alors la solution variationnelle $u \in H^1(\Omega)$ du problème de Dirichlet*

$$(3.39) \quad \begin{cases} -\nabla^2 u = f \\ u|_{\partial\Omega} = g \end{cases}$$

appartient à $H^p(\Omega)$ avec $p = \min(m + 2, n)$.

Remarquons qu'en vertu de ce théorème et du précédent il suffit que $m \geq d/2$ et $n \geq d/2 + 2$ pour que $H^p(\Omega) \subset C^2(\bar{\Omega})$; ce qui fournit des conditions suffisantes de régularité pour que u soit solution classique du problème.

Pour le problème de Neumann on a quasiment les mêmes conclusions:

Théorème 16 *Si $f \in H^m(\Omega)$, et $g \in L^2(\partial\Omega)$ est tel que il existe un relèvement $\tilde{g} \in H^n(\Omega)$ avec $n \geq 2$ et $\tilde{g}|_{\partial\Omega} = g$, alors la solution variationnelle $u \in H^1(\Omega)$ du problème de Neumann*

$$(3.40) \quad \begin{cases} u - \nabla^2 u = f \\ \frac{\partial u}{\partial n}|_{\partial\Omega} = g \end{cases}$$

appartient à $H^p(\Omega)$ avec $p = \min(m + 2, n)$.

3.4 Approximation interne

Dans cette partie on met tous nos problèmes de diffusion sous une forme variationnelle abstraite. On cherche ensuite une approximation de la solution dans un sous-espace de dimension finie contenu dans l'espace de Hilbert qui contient la solution. Cette technique est connue sous les noms d'approximation interne ou conforme. On montre la convergence de l'approximation en se basant sur le principe (théorème) général en analyse numérique:

$$\text{Stabilité} + \text{Consistance} \implies \text{Convergence}$$

La stabilité est assurée par l'ellipticité des opérateurs, la consistance est assurée par le fait que les espaces de dimensions finies qu'on utilise sont des espaces d'approximation interne.

3.4.1 Le problème discret

On a vu dans les sections précédentes que les problèmes de diffusion avec des conditions aux limites du type Dirichlet, Neumann ou Fourier admettent des formulations variationnelles qui peuvent toutes se mettre sous la forme abstraite suivante: trouver $u \in V$ tel que

$$(3.41) \quad \forall v \in V, \quad a(u, v) = \langle f, v \rangle$$

où V est un espace de Hilbert réel (généralement un espace de Sobolev), $a(.,.)$ est une forme bilinéaire continue V -elliptique et $\langle f, . \rangle$ est une forme linéaire continue sur V . Par la suite on note $(.,.)_V$ le produit scalaire sur V et $|\cdot|_V$ la norme associée. La norme de la forme bilinéaire continue $a(.,.)$ est notée $|a|$. L'hypothèse de V -ellipticité sur $a(.,.)$ implique qu'il existe $\alpha > 0$ tel que

$$(3.42) \quad \forall v \in V, \quad a(v, v) \geq \alpha |v|_V^2.$$

Le théorème de Lax-Milgram fournit l'existence, l'unicité d'une solution u à notre problème, il fournit aussi la stabilité de la solution par rapport aux données. On se fixe maintenant comme objectif de construire une approximation numérique de u (ie. en dimension finie: rappelons que les ordinateurs actuels sont bornés dans leurs capacités). L'approche la plus couramment employée actuellement (c'est aussi la plus simple conceptuellement) pour construire une telle approximation consiste à se donner une suite $\{V_h\}$ de sous-espaces de V de dimensions finies, qui approchent V lorsque $h \rightarrow 0$ au sens où il existe un sous-espace dense \mathcal{N} de V et une application $r_h : \mathcal{N} \rightarrow V_h$ tels que pour tout $v \in \mathcal{N}$, $\lim_{h \rightarrow 0} |v - r_h(v)|_1 = 0$.

Remarque: En général l'espace \mathcal{N} est du type $C^k(\bar{\Omega})$ ou $H^m(\Omega)$ avec $m > d/2$ de telle sorte que d'après le théorème 14, les fonctions de $H^m(\Omega)$ soient continues sur $\bar{\Omega}$, on peut ainsi définir leurs valeurs en certains points du domaine et reconstruire une approximation polynômiale (ie. de dimension finie) convergente à partir de ces points (ce sera l'application r_h recherchée).

Une fois qu'on s'est donné un V_h , on cherche $u_h \in V_h \subset V$ tel que

$$(3.43) \quad \forall v_h \in V_h, \quad a(u_h, v_h) = \langle f, v_h \rangle.$$

Remarque: Tout ceci a un sens puisque V_h est inclus dans V ; l'inclusion est ici fondamentale, on montre par la suite que c'est elle qui assure la consistance de l'approximation (ie. c'est la moitié de la convergence assurée). On dit que V_h réalise une approximation interne de V (ie. on cherche à approcher u par une suite d'approximations u_h internes à V).

Théorème 17 *le problème discret (3.43) admet une solution unique $u_h \in V_h$.*

Démonstration: La restriction $a(.,.)$ à V_h est continue et V_h -elliptique, idem pour la restriction de f à V_h . Le théorème de Lax-Milgram permet de conclure. \square

L'autre façon de voir les choses consiste à se donner une base de V_h , $(v_h^i)_{i=1, \dots, I}$ et à projeter le problème discret dans cette base. On cherche donc u_h sous la forme $u_h = \sum_{i=1}^I u_i v_h^i$ tel que

$$(3.44) \quad \forall i = 1, \dots, I, \quad \sum_{j=1}^I u_j a(v_h^j, v_h^i) = \langle f, v_h^i \rangle.$$

En introduisant les notations matricielles $U = (u_1, \dots, u_I)$ et $F = (\langle f, u_1 \rangle, \dots, \langle f, u_I \rangle)$ le système ci-dessus est équivalent au problème matriciel

$$(3.45) \quad AU = F$$

où les coefficients de A sont définis par $A_{ij} = a(v_h^j, v_h^i)$.

Remarque 1: L'inversibilité de la matrice A résulte directement de la coercivité de $a(.,.)$. En effet, la dimension d'espace étant finie, l'existence-unicité de la solution du système linéaire (ie. l'inversibilité de A) est équivalente à l'unicité ou à l'existence. Autrement dit A est inversible si on montre que l'unique solution du problème $AX = 0$ est la solution triviale $X = 0$. En posant $x_h = \sum_{j=1}^I x_j v_h^j$, on vérifie aisément que $(X, AX) = a(x_h, x_h)$. Il s'ensuit que $AX = 0$ entraîne $a(x_h, x_h) = 0$ et l'hypothèse de V_h -ellipticité implique $x_h = 0$, d'où les x_i sont tous nuls puisque $(v_h^i)_{i=1, \dots, I}$ est une base de V_h .

Remarque 2: En pratique c'est le système matriciel (3.45) qu'on exploite numériquement. La mise en œuvre de cette technique se fait en trois étapes

- La première consiste à se donner V_h et une base de V_h , cette étape consiste en pratique à construire un maillage du domaine Ω et l'opérateur d'approximation associé r_h .
- Ensuite il faut calculer les coefficients de A et F ; généralement les formes $a(.,.)$ et $\langle f, . \rangle$ font intervenir des intégrales de volume ou de bord, il faut alors se donner des formules d'approximation de ces intégrales.
- La dernière étape consiste à résoudre le système linéaire.

Tout cela bien sûr est plus facile à dire qu'à faire.

3.4.2 Convergence de l'approximation interne

Une formulation du problème variationnel (3.41) plus proche de la formulation matricielle consiste à introduire l'opérateur A tel que:

$$(3.46) \quad \begin{aligned} A : V &\longrightarrow V' \\ u &\longmapsto Au : v \longmapsto a(u, v) \end{aligned}$$

De même dans le cadre discret on introduit A_h :

$$(3.47) \quad \begin{aligned} A_h : V_h &\longrightarrow V_h' \\ u_h &\longmapsto Au_h : v_h \longmapsto a(u_h, v_h) \end{aligned}$$

où V_h' désigne le dual de V_h .

L'espace V_h étant de dimension finie, les normes (et donc les topologies) de V_h et V_h' sont bien sûr équivalentes, mais les constantes d'équivalence tendent généralement vers l'infini ou zéro lorsque la dimension de V_h devient très grande. En pratique la dimension de V_h est comprise entre 10^4 et 10^6 ; la limite supérieure ne dépend que la capacité des ordinateurs. Il y a donc lieu, en pratique, de distinguer V_h et V_h' .

On projete V' sur V_h' en définissant l'opérateur linéaire ρ_h :

$$(3.48) \quad \begin{aligned} \rho_h : V' &\longrightarrow V_h' \\ f &\longmapsto \rho_h f : v_h \longmapsto \langle f, v_h \rangle \end{aligned}$$

La notation $\langle f, v_h \rangle$ désigne le résultat dans \mathbb{R} de l'action de f sur v_h . La définition de ρ_h a un sens puisque V_h est inclus dans V . On vérifie aisément que ρ_h est une contraction, ie. $|\rho_h| \leq 1$.

On rappelle finalement qu'on dispose aussi de l'opérateur d'approximation $r_h : \mathcal{N} \rightarrow V_h$. La consistance de la méthode est confirmée par:

Lemme 3 (i) *L'approximation interne est consistante; plus précisément:*

$$(3.49) \quad \forall u \in \mathcal{N}, \quad |\rho_h Au - A_h r_h u|_{V_h} \xrightarrow{h \rightarrow 0} 0$$

(ii) *Si r_h est un opérateur d'approximation d'ordre p , ie. si pour tout $u \in \mathcal{N}$ il existe $c(u) > 0$ tel que $|r_h u - u|_V < c(u)h^p$, alors l'approximation interne est consistante à l'ordre p :*

$$(3.50) \quad \forall u \in \mathcal{N}, \quad |\rho_h Au - A_h r_h u|_{V_h} \leq c(u)h^p$$

Démonstration: Soit $u \in \mathcal{N}$, on a

$$\rho_h Au - A_h r_h u = \rho_h A(u - r_h u) + [\rho_h A - A_h](r_h u)$$

vérifions que $[\rho_h A - A_h](w_h) = 0$ dans V'_h pour tout $w_h \in V_h$:

$$\forall w_h \in V_h, \quad \langle \rho_h A w_h, v_h \rangle = \langle A w_h, v_h \rangle = a(w_h, v_h) = \langle A_h w_h, v_h \rangle$$

d'où le résultat. Remarquer que la première et la dernière égalité sont dues au fait que l'approximation est interne. Finalement on a:

$$\begin{aligned} |\rho_h Au - A_h r_h u|_{V_h} &= |\rho_h A(u - r_h u)|_{V_h} \\ &\leq |A(u - r_h u)|_{V'} \quad \text{car } |\rho_h| \leq 1, \\ &\leq |a||u - r_h u|_V \quad \text{car } a \text{ est continu,} \end{aligned}$$

ce qui prouve la consistance. \square

Lemme 4 *L'approximation interne est stable:*

$$(3.51) \quad \forall v_h \in V_h, \quad |v_h|_{V_h} \leq \frac{1}{\alpha} |A_h v_h|_{V'_h}$$

Démonstration: L'inégalité de stabilité résulte de la coercivité de a :

$$\forall v_h \in V_h, \quad |a_h v_h|_{V'_h} = \sup_{|w_h|_{V_h} \neq 0} \frac{\langle A_h v_h, w_h \rangle}{|w_h|_{V_h}} \geq \frac{\langle A_h v_h, v_h \rangle}{|v_h|_{V_h}} \geq \alpha |v_h|_{V_h}. \square$$

Théorème 18 (i) *L'approximation interne est convergente.*

(ii) *L'approximation interne est convergente à l'ordre p si elle est consistante à l'ordre p .*

Démonstration: (ii) Soit u la solution du problème continu et u_h la solution du problème discret. Rappelons que u_h est solution du problème $A_h u_h = \rho_h Au$. Supposons dans un premier temps que $u \in \mathcal{N}$, on a alors:

$$|u - u_h|_V \leq |u - r_h u|_V + |r_h u - u_h|_{V_h}$$

La stabilité permet d'écrire

$$|r_h u - u_h|_{V_h} \leq \frac{1}{\alpha} |A_h r_h u - A_h u_h|_{V'_h} = \frac{1}{\alpha} |A_h r_h u - \rho_h Au|_{V'_h}$$

et la consistance à l'ordre p permet d'écrire:

$$|A_h r_h u - \rho_h A u|_{V'_h} \leq |a| |u - r_h u|_V \leq |a| c(u) h^p$$

D'où finalement:

$$|u - u_h|_V \leq \left(1 + \frac{|a|}{\alpha}\right) c(u) h^p$$

(i) Si $u \notin \mathcal{N}$, la densité de \mathcal{N} dans V permet pour tout $\epsilon > 0$ de trouver $w \in \mathcal{N}$ tel que $|u - w|_V \leq \epsilon$. Introduisons w_h la solution du problème $A_h w_h = \rho_h A w$. En utilisant l'inégalité triangulaire on a:

$$|u - u_h|_V \leq |u - w|_V + |w - w_h|_V + |w_h - u_h|_{V_h}$$

Puisque w est dans \mathcal{N} on peut appliquer le résultat de (ii) autrement dit, il existe $h(\epsilon)$ tel que pour tout $h \leq h(\epsilon)$, on a

$$|w - w_h|_V \leq \epsilon$$

Le terme $|w_h - u_h|_{V_h}$ est contrôlé grâce à la stabilité de A :

$$|w_h - u_h|_{V_h} \leq \frac{1}{\alpha} |A_h w_h - A_h u_h|_{V'_h} = \frac{1}{\alpha} |\rho_h A w - \rho_h A u|_{V'_h} \leq \frac{|a|}{\alpha} \epsilon.$$

D'où la majoration finale:

$$|u - u_h|_V \leq \left(2 + \frac{|a|}{\alpha}\right) \epsilon$$

ce qui prouve la convergence. \square

La qualité de la convergence de l'approximation interne dépend de façon déterminante de la qualité d'approximation de l'opérateur r_h ; la méthode est d'autant plus puissante que l'opérateur d'approximation est excellent. En pratique il faut faire un compromis entre la qualité d'approximation et les difficultés de mise en œuvre qui lui sont liées (principe de conservation des difficultés...). On distingue en gros deux type de méthodes en fonction des qualités approximantes:

- Les méthodes spectrales. Ces méthodes ont un degré de précision infini en ce sens que l'erreur d'approximation ne dépend que de la régularité de la fonction à approcher. En contrepartie elles ne sont (actuellement) pratiquement utilisables que sur des domaines cylindriques ou parallélépipédiques (cf. section suivante et Bernardi-Maday [4]).
- Les méthodes d'éléments finis. Ce sont des méthodes de bas degré (du premier ou second ordre en pratique) mais elles sont très flexibles au sens où on peut les utiliser sur des domaines dont la géométrie est assez complexe.

3.4.3 Approximation de Galerkin – Approximation spectrale

Dans cette section on donne un exemple d'espace d'approximation interne V_h basé sur la notion de base hilbertienne. Plus précisément, on appelle base hilbertienne de V toute suite $(w_j)_{j \geq 1}$ d'éléments de V telle que

(i) $\forall m \geq 1$, la famille (w_1, \dots, w_m) est libre.

(ii) L'espace engendré par les combinaisons linéaires finies des (w_j) est dense dans V .

Supposons que l'espace de Hilbert V soit séparable, c'est à dire qu'il contient un sous-ensemble dense dénombrable, on peut montrer alors que V admet une base hilbertienne (il suffit d'extraire des éléments linéairement indépendants du sous-ensemble dense). La notion de séparabilité n'est pas restrictive car on peut montrer que tous les espaces de Sobolev usuels sont séparables.

Exemple 1: Les polynômes de Legendre forment une base hilbertienne de $L^2(]-1, 1[, dx)$, où dx est la mesure de Lebesgue. Les polynômes de Tchebycheff forment une base hilbertienne de $L^2(]-1, 1[, d\mu)$, où $d\mu = (1 - x^2)^{-1/2} dx$.

Exemple 2: Dans le carré $\Omega =]0, 1[\times]0, 1[$ on peut montrer que la suite des fonctions

$$(3.52) \quad i, j \geq 0, \quad w_{ij}(x_1, x_2) = \sin(i\pi x_1) \sin(j\pi x_2)$$

est une base hilbertienne de $H_0^1(\Omega)$. C'est en fait l'ensemble des fonctions propres du laplacien avec condition de Dirichlet homogène sur $\partial\Omega$.

Exemple 3: La théorie des opérateurs compacts auto-adjoints (pour ceux qui connaissent) fournit une grande quantité (en pratique la quasi-totalité) de bases hilbertiennes. On peut montrer que les opérateurs compacts auto-adjoints sont diagonalisables et l'ensemble de leurs vecteurs propres forme une base hilbertienne. On appelle parfois l'ensemble des valeurs propres le spectre de l'opérateur. Voir par exemple Dautray-Lions [8, t. 5, p. 22-135] pour une introduction au sujet.

S'étant donné $(w_j)_{j \geq 1}$ une base hilbertienne de V , une idée simple (l'idée est de Galerkin semble-t-il) consiste à poser $h = 1/m$ et $V_m = \langle w_1, \dots, w_m \rangle$. Le problème discret s'écrit: trouver $u_m \in V_m$ tel que

$$(3.53) \quad \forall j = 1, \dots, m \quad a(u_m, v_j) = l(v_j)$$

On a le résultat général suivant:

Théorème 19 *Pour tout $m \geq 1$ le problème (3.53) a une solution unique, et on a la convergence:*

$$(3.54) \quad \lim_{m \rightarrow +\infty} |u - u_m|_V = 0$$

Démonstration: L'existence, l'unicité et la stabilité de la solution de ce problème sont bien sûr garanties par le théorème de Lax-Milgram (ici en dimension finie la matrice est définie positive). La convergence de cette approximation est donnée par le théorème général de convergence 18 en prenant $h = 1/m$, $\mathcal{N} = V$ et $r_h = \Pi_m$ est la projection orthogonale de V sur V_m . La propriété (ii) de définition de la base hilbertienne implique

$$\forall v \in V, \quad |v - \Pi_m v|_V \xrightarrow{h \rightarrow 0} 0.$$

ce qui prouve la qualité approximante de Π_m . \square

Cette technique est aussi connue sous le nom d'approximation spectrale, car en pratique les bases hilbertiennes qu'on utilise sont toujours les vecteurs propres (d'où l'origine du mot spectrale) d'un opérateur compact auto-adjoint. L'ouvrage de Bernardi-Maday [4] fait une analyse (théorique) détaillée de toutes ces techniques.

Remarque: En pratique les formes $a(\cdot, \cdot)$ et $\langle f, \cdot \rangle$ font intervenir des intégrales de volume ou de bord, elles sont généralement remplacées par des quadratures numériques quasi-exactes pour les éléments de V_m .

En pratique on sait construire des bases hilbertiennes sur des domaines de géométrie simple: carré, cube, disque, cylindre, etc... Mais les choses se corsent lorsqu'on s'intéresse aux

problèmes à géométrie complexe. L'alternative consiste à abandonner les propriétés d'approximation globale des bases hilbertienne et à se contenter de bases qui ont des propriétés d'approximation locale. L'idée est donc d'approcher la solution par morceaux: on coupe le domaine en petits morceaux et sur chaque petit morceau on cherche une approximation, polynômiale par exemple, de la solution. Cette idée est à la base de l'approximation par éléments finis.

3.4.4 Approximation par éléments finis

Pour une introduction détaillée à cette technique le lecteur est renvoyé aux ouvrages de Ciarlet [7], Dautray-Lions [8, t.6 ch. XII], Pironneau [16], Raviart-Thomas [17].

Chapitre 4

Formulation variationnelle du problème de Stokes et approximation.

4.1 Le problème de Stokes

4.1.1 Formulation du problème continu

On s'intéresse dans ce chapitre au problème de Stokes généralisé dans Ω un ouvert borné connexe régulier de \mathbb{R}^d (en pratique $d = 2$ ou 3). La formulation classique (on dit aussi formulation forte) de ce problème consiste à chercher une vitesse u et une pression p suffisamment régulières telles que:

$$(4.1) \quad \begin{cases} u - \sigma \nabla^2 u + \nabla p = f \\ \nabla \cdot u = g, \quad u|_{\partial\Omega} = 0 \end{cases}$$

où f et g sont des termes sources donnés suffisamment réguliers et σ un réel strictement positif. La donnée g doit en plus vérifier la condition de compatibilité $\int_{\Omega} g = 0$. Pour les écoulements incompressibles on prend $g = 0$.

Pour comprendre les difficultés spécifiques à ce problème il est préférable de considérer sa formulation variationnelle (on dit aussi parfois formulation faible). Compte tenu des conditions aux limites et des opérateurs en présence, la formulation variationnelle du problème consiste à chercher u dans $H_0^1(\Omega)^d$ et p dans $L^2(\Omega)$ satisfaisant les équations précédentes au sens de la dualité dans $L^2(\Omega)^d$ et $L^2(\Omega)$ respectivement. Cependant, la pression intervenant par son gradient n'est définie qu'à une constante près. Afin d'assurer l'unicité de la pression (ie. au bout du compte avoir un système linéaire inversible) il faut restreindre la recherche de la pression dans un espace ne contenant pas les constantes. On peut imposer par exemple la pression en un point particulier du domaine. Cette solution est valable pour une solution forte du problème, elle ne l'est plus pour une solution faible puisque dans ce cas on impose uniquement à p d'être de carré sommable et de ce fait p n'est a priori défini que presque partout (et non pas partout). Une autre solution consiste à se restreindre à des pressions de moyenne nulle. On introduit donc l'espace

$$(4.2) \quad L_0^2(\Omega) = \{p \in L^2(\Omega), \int_{\Omega} p = 0\}.$$

Dans ce contexte la formulation faible du problème ci-dessus s'écrit: pour $f \in L^2(\Omega)^d$ (on pourrait aussi prendre $f \in H^{-1}(\Omega)^d$) et $g \in L_0^2(\Omega)$, chercher $u \in H_0^1(\Omega)^d$ et $p \in L_0^2(\Omega)$ tels que

$$(4.3) \quad \begin{cases} \forall v \in H_0^1(\Omega)^d, & (u, v) + \sigma(\nabla u, \nabla v) - (p, \nabla \cdot v) = (f, v) \\ \forall q \in L^2(\Omega), & (q, \nabla \cdot u) = (q, g) \end{cases}$$

où (\cdot, \cdot) désigne indifféremment le produit scalaire dans $L^2(\Omega)$ ou $L^2(\Omega)^d$.

Une certaine connexion entre la formulation classique et la formulation variationnelle est assurée par

Proposition 8 *Si f et g sont dans $C^0(\overline{\Omega})^d$ et $C^0(\overline{\Omega})$ respectivement, si Ω est suffisamment régulier (ie. C^2 par exemple) et si la solution variationnelle (u, p) vérifie $u \in C^2(\overline{\Omega})^d$ et $p \in C^1(\overline{\Omega})$, alors le couple (u, p) est solution classique du problème de Stokes.*

Démonstration: Pour la première équation on se restreint aux fonctions tests dans $C_c^\infty(\Omega)^d$, espace des fonctions C^∞ à support compact sur Ω et à valeur dans \mathbb{R}^d . En intégrant par parties on obtient

$$\forall v \in C_c^\infty(\Omega)^d, \quad \int_{\Omega} (u - \sigma \nabla^2 u + \nabla p) v = \int_{\Omega} f v$$

D'où par densité de $C_c^\infty(\Omega)^d$ dans $L^2(\Omega)^d$ on déduit l'égalité $u - \sigma \nabla^2 u + \nabla p = f$ dans $L^2(\Omega)$, d'où l'égalité presque partout; en fait partout puisque les fonctions sont continues. On procède de même pour la seconde égalité. La condition limite sur u résulte de son appartenance à $H_0^1(\Omega)^d$. \square

Il est possible de démontrer que le problème (4.3) est bien posé. La démonstration est assez technique. La difficulté essentielle vient du fait que le problème est posé en dimension infinie (cf. eg. Girault-Raviart [10]). Par la suite on étudie le problème en dimension finie pour contourner les subtilités qui sont spécifiques de la dimension infinie.

4.1.2 Formulation discrète du problème de Stokes

On se donne une suite $\{V_h\}$ d'espaces de dimensions finies qui approchent $H_0^1(\Omega)^d$ au sens où il existe un sous-espace dense \mathcal{N} de $H_0^1(\Omega)^d$ et une application $r_h : \mathcal{N} \rightarrow V_h$ tels que pour tout $v \in \mathcal{N}$, $\lim_{h \rightarrow 0} \|v - r_h(v)\|_1 = 0$. V_h est muni du produit scalaire et de la norme de $H_0^1(\Omega)^d$. On définit aussi V_h' le dual de V_h qu'on muni de la norme:

$$(4.4) \quad \|f_h\|_{V_h'} = \sup_{|v_h|_{V_h} \neq 0} \frac{(f_h, v_h)}{|v_h|_{V_h}}$$

Les normes de V_h et V_h' sont bien sûr équivalentes puisque V_h est de dimension finie. Cependant les constantes d'équivalence tendent vers l'infini ou zéro lorsque I tend vers l'infini. Il y a donc lieu de les distinguer lorsque I est très grand. Dans la pratique I est compris entre 10000 et quelques millions. En pratique V_h' est constitué des éléments de V_h mais il est équipé de la structure d'espace de Hilbert induite par la dualité; l'action des éléments de V_h' sur V_h est identifiée au produit scalaire de $L^2(\Omega)^d$.

On se donne de même une suite d'espaces $\{P_h\}$ approchant $L_0^2(\Omega)$. Les espaces P_h sont équipés de la structure de Hilbert induite par $L_0^2(\Omega)$. Par la suite on identifie P_h et P_h' .

Pour h fixé on se donne une base $(v_h^i)_{i=1,\dots,I}$ de V_h et une base $(q_h^k)_{k=1,\dots,K}$ de P_h . En posant $u_h = \sum_1^I u_i v_h^i$ et $p_h = \sum_1^K p_k q_h^k$, on peut alors considérer la projection du problème (4.3) sur $V_h \times P_h$ en prenant pour fonctions tests les éléments des bases. On obtient:

$$(4.5) \quad \begin{cases} \forall i = 1, \dots, I, & \sum_{j=1}^I u_j [(v_h^j, v_h^i) + \sigma(\nabla v_h^j, \nabla v_h^i)] - \sum_{l=1}^K p_l (q_h^l, \nabla \cdot v_h^i) = (f, v_h^i) \\ \forall k = 1, \dots, K, & \sum_{j=1}^I u_j (q_h^k, \nabla \cdot v_h^j) = (g, q_h^k) \end{cases}$$

C'est à dire en introduisant les notations vectorielles $U = (u_1, \dots, u_I)$ et $P = (p_1, \dots, p_K)$ on obtient le système matriciel suivant:

$$(4.6) \quad \begin{bmatrix} A & B \\ C & 0 \end{bmatrix} \begin{bmatrix} U \\ P \end{bmatrix} = \begin{bmatrix} F \\ G \end{bmatrix}$$

où $F = ((f, v_h^1), \dots, (f, v_h^I))$, $P = ((g, q_h^1), \dots, (g, q_h^K))$ et les coefficients de A , B et C sont définis par

$$(4.7) \quad \begin{aligned} A_{ij} &= (v_h^j, v_h^i) + \sigma(\nabla v_h^j, \nabla v_h^i) \\ B_{il} &= -(q_h^l, \nabla \cdot v_h^i) \\ C_{ki} &= (q_h^k, \nabla \cdot v_h^i) \end{aligned}$$

Remarquons que les matrices B et C sont transposées l'une de l'autre au signe près, c'est à dire $B = -C^t$. La raison fondamentale de cette relation est que les opérateurs divergence et gradient sont eux-mêmes transposés (au signe près) au sens de la dualité $H_0^1(\Omega)^d / H^{-1}(\Omega)^d$. Ce fait aura des conséquences très importantes par la suite. Notons au passage que la matrice A est symétrique définie positive (classique).

Une autre façon (équivalente) de regarder le problème consiste à définir les opérateurs A_h et B_h tels que

$$(4.8) \quad \begin{aligned} A_h : V_h &\longrightarrow V_h' \\ u_h &\longmapsto A_h(u_h) : v_h \longmapsto (A_h u_h, v_h) = (u_h, v_h) + \sigma(\nabla u_h, \nabla v_h) \end{aligned}$$

$$(4.9) \quad \begin{aligned} B_h : P_h &\longrightarrow V_h' \\ p_h &\longmapsto B_h(p_h) : v_h \longmapsto (B_h p_h, v_h) = -(p_h, \nabla \cdot v_h) \end{aligned}$$

On introduit aussi $B_h^t : V_h \longrightarrow P_h$ l'opérateur transposé de B_h défini par

$$(4.10) \quad \forall u_h \in V_h, \forall p_h \in P_h, \quad (B_h p_h, u_h) = (p_h, B_h^t u_h)$$

Bien sûr, les opérateurs A_h et B_h étant linéaires et les dimensions finies, A_h et B_h sont continus. Cependant en différenciant les espaces V_h et V_h' on a

$$(4.11) \quad |A_h| = \max_{u_h \in V_h - \{0\}} \frac{|A_h u_h|_{V_h'}}{|u_h|_1} \leq \max(1, \sigma)$$

$$(4.12) \quad |B_h| = \max_{p_h \in P_h - \{0\}} \frac{|B_h p_h|_{V'_h}}{|p_h|_0} \leq 1$$

C'est à dire que les constantes de continuité de A_h et B_h sont majorées indépendamment de la dimension des espaces (ie. indépendantes de h), ce qui ne pourrait être le cas si on ne différencie pas les topologies de V_h et V'_h . En effet, si V_h et P_h sont des espaces d'éléments finis et si on ne différencie pas V_h et V'_h , c'est à dire si on équipe V_h du produit scalaire de $L^2(\Omega)^d$ on peut montrer que la norme de A_h se comporte comme $1/h^2$ et la norme de B_h comme $1/h$. Lorsque $h \rightarrow 0$ ces normes explosent et ne sont d'aucune utilité.

On définit enfin $f_h \in V'_h$ et $g_h \in P_h$ (cf. théorème de projection) tels que

$$(4.13) \quad \forall v_h \in V_h, \quad (f, v_h) = (f_h, v_h)$$

$$(4.14) \quad \forall q_h \in P_h, \quad (g, q_h) = (g_h, q_h)$$

Ces nouvelles définitions permettent de réécrire le problème discret sous la forme abstraite suivante: trouver $u_h \in V_h$ et $p_h \in P_h$ tels que

$$(4.15) \quad \begin{cases} A_h u_h + B_h p_h = f_h & \text{dans } V'_h \\ -B_h^t u_h = g_h & \text{dans } P_h \end{cases}$$

Cette forme du problème discret est très proche du problème continue. Elle se révélera par la suite très efficace pour obtenir les résultats d'existence, d'unicité et de stabilité de la solution discrète.

Maintenant qu'on a un problème discret (posé sous forme matricielle et sous forme abstraite) il faut se poser trois questions:

- Le problème discret est-il bien inversible?
- A-t-on la stabilité de la solution par rapport aux données?
- A-t-on convergence?

4.1.3 Existence-Unicité (Condition "inf-sup")

Il convient de remarquer dans un premier temps que le problème matriciel et le problème abstrait sont équivalents. En effet, à tout vecteur vitesse $U = (u_1, \dots, u_I)$ de \mathbb{R}^I et à tout vecteur pression $P = (p_1, \dots, p_K)$ de \mathbb{R}^K on associe, et ce de façon unique, le champ de vitesse u_h de V_h et le champ p_h de P_h tels que $u_h = \sum_1^I u_i v_h^i$ et $p_h = \sum_1^K p_k q_h^k$. Tous les résultats d'existence, d'unicité et de stabilité pour le problème abstrait se transposent immédiatement au problème matriciel.

Lemme 5 *L'opérateur A_h est symétrique définie positif. Plus précisément on a:*

$$(4.16) \quad \forall u_h \in V_h, \quad (A_h u_h, u_h) \geq \min(\sigma, 1) |u_h|_1, \text{ et } (A_h u_h, u_h) \geq |u_h|_0$$

Démonstration: La symétrie (ie. $A_h = A_h^t$) résulte de la définition de A_h . Pour montrer les inégalités il suffit de revenir aux définitions des normes impliquées en remarquant que $(A_h u_h, u_h) = |u_h|_0^2 + \sigma |\nabla u_h|_0^2$. \square

Remarque: Le lemme se transpose en termes matriciels: la matrice A est symétrique définie positive et

$$(4.17) \quad \forall U \in \mathbb{R}^I, \quad (U, AU) \geq \min(\sigma, 1)|U|_1, \text{ et } (U, AU) \geq |U|_0$$

où (\cdot, \cdot) désigne le produit scalaire canonique de \mathbb{R}^I et on a posé par définition $|U|_0^2 = |u_h|_0^2$ et $|U|_1^2 = |u_h|_1^2$

Le problème (4.6) est posé en dimension finie et le système linéaire est carré d'ordre $I.K \times I.K$. Pour montrer l'existence et l'unicité il suffit de montrer ou bien l'existence (surjectivité) ou bien l'unicité (injectivité). Ce résultat bien connu résulte du théorème du rang qu'on écrit souvent sous la forme: $\dim \text{Ker} + \dim \text{Im} = \dim E$.

Une autre version du théorème du rang, qui sera très utile par la suite, se met sous la forme abstraite suivante:

Lemme 6 Soit $B_h : P_h \rightarrow V_h'$ défini comme précédemment, on a:

$$(4.18) \quad \ker B_h^t = (\text{Im} B_h)^{\perp_0}$$

où \perp_0 désigne l'orthogonal dans V_h' au sens de la dualité, c'est à dire au sens du produit scalaire $L^2(\Omega)^d$.

Démonstration: (i) Soit $v \in (\text{Im} B_h)^{\perp_0}$, pour tout $q_h \in P_h$ on a $0 = (B_h q_h, v_h) = (q_h, B_h^t v_h)$, d'où $v_h \in \ker B_h^t$.

(ii) Réciproquement, soit $v_h \in \ker B_h^t$, pour tout $q_h \in P_h$ on a $0 = (q_h, B_h^t v_h) = (B_h q_h, v_h)$, d'où $v_h \in (\text{Im} B_h)^{\perp_0}$. \square

Ce résultat se généralise en dimension infinie modulo quelques subtilités de régularité sur B (cf. Brezis [5, p. 29]). Ce résultat compte parmi les plus importants de l'algèbre linéaire, il est valable quelque soit l'opérateur linéaire; il va se révéler d'une grande importance pour l'étude de notre problème discret (4.6).

Théorème 20 Le problème (4.15) (resp. (4.6)) admet une solution unique si l'opérateur B_h est injectif (resp. la matrice B est de rang K).

Démonstration: Soit (u_{0h}, p_{0h}) une solution du problème homogène:

$$\begin{cases} A_h u_{0h} + B_h p_{0h} = 0 \\ -B_h^t u_{0h} = 0 \end{cases}$$

La seconde équation (équation de conservation de la masse) implique $u_{0h} \in \ker B^t$, or le lemme 6 implique: $\forall q_h \in P_h, (u_{0h}, B_h q_h) = 0$. D'autre part en dualisant la première équation (conservation de la quantité de mouvement) par u_{0h} on obtient: $(u_{0h}, A_h u_{0h}) + (u_{0h}, B_h p_{0h}) = 0$, où le second terme est nul d'après ce qui précède. On en déduit $(u_{0h}, A_h u_{0h}) = 0$, qui conduit à $u_{0h} = 0$ (ie. unicité de la vitesse) puisque A_h est défini positif. On a donc $B_h p_{0h} = 0$, ce qui donne $p_{0h} = 0$ si l'opérateur B_h est injectif. \square

Remarquer ici que la transposition naturelle entre les opérateurs divergence et gradient se transporte naturellement aux opérateurs discrets grâce à la formulation variationnelle. Noter aussi que la transposition en question est intervenue de façon déterminante pour conclure à l'unicité de la solution discrète. On verra par la suite que la formulation forte et son approximation par des différences finies ne permettent pas en général de conserver ces propriétés importantes. Les approximations par différences finies sont de ce fait très difficilement exploitables et réservent souvent des surprises désagréables.

Regardons d'un peu plus près ce que signifie l'injectivité de l'opérateur B_h .

Proposition 9 Soit $B_h : P_h \longrightarrow V'_h$, B_h est injectif si et seulement si il existe une constante $\beta > 0$ telle que:

$$(4.19) \quad \inf_{|p_h|_0 \neq 0} \sup_{|u_h|_1 \neq 0} \frac{(B_h p_h, u_h)}{|p_h|_0 |u_h|_1} \geq \beta$$

Démonstration: Puisqu'on est en dimension finie, B_h est injectif si et seulement si $\min_{|q_h|_0} |B_h q_h|_{V'_h} > 0$, autrement dit si et seulement si il existe une constante $\beta > 0$ telle que $\min_{|q_h|_0 \neq 0} |B_h q_h|_{V'_h} / |q_h|_0 > \beta$. D'autre part, en revenant à la définition de la norme $|\cdot|_{V'_h}$ on a $|B_h q_h|_{V'_h} = \max_{|v_h|_1 \neq 0} (B_h q_h, v_h) / |v_h|_1$. On a donc injectivité de B_h si et seulement si $\min_{|q_h|_0 \neq 0} \max_{|v_h|_1 \neq 0} (B_h q_h, v_h) / |q_h|_0 |v_h|_1 \geq \beta$. \square

Remarquer que les dimensions étant finies les “inf” et les “sup” sont des “min” et des “max”, c'est à dire les bornes sont nécessairement atteintes (continuité et compacité). La condition (4.20) se généralise en dimension infinie sous quelques réserves de régularité sur B . Cette condition est connue dans le microcosme de la Mécanique des Fluides Numérique sous les noms de condition “inf-sup” ou condition de “Babuška-Brezzi” (cf. par exemple Babuška [3] et Brezzi [6]). C'est généralement en vérifiant explicitement cette condition qu'on montre que l'opérateur B_h est injectif.

En revenant aux définitions de la matrice B et des vecteurs P et U on vérifie que $(BP, U) = (B_h p_h, u_h)$. Ainsi en introduisant $|P|_0 = |p_h|_0$, la condition “inf-sup” s'écrit de façon équivalente:

$$(4.20) \quad \min_{|P|_0 \neq 0} \max_{|U|_1 \neq 0} \frac{(BP, U)}{|P|_0 |U|_1} \geq \beta$$

La caractéristique la plus étonnante de la condition “inf-sup” est qu'elle est satisfaite dans le cas continu (d'où existence et unicité la solution pour le problème continu, modulo les constantes pour la pression) alors qu'elle peut ne pas l'être en dimension finie. Dans le cadre continu l'opérateur B_h est remplacé par le gradient ∇ et on peut montrer (cf. Girault-Raviart [10, p. 20]):

Lemme 7 Soit Ω un ouvert connexe borné régulier de \mathbb{R}^d , il existe $\beta > 0$ tel que

$$(4.21) \quad \forall p \in L_0^2(\Omega), \quad |\nabla p|_{-1} \geq \beta |p|_0$$

ce qui s'écrit encore

$$(4.22) \quad \inf_{|p|_0 \neq 0} \sup_{|v|_1 \neq 0} \frac{(p, \nabla \cdot u)}{|p|_0 |v|_1} \geq \beta$$

En fait la condition “inf-sup” est une condition de compatibilité des espaces d'approximation V_h et P_h , en d'autres termes les espaces d'approximation de la vitesse et de la pression ne peuvent pas être choisis indépendamment l'un de l'autre sous peine de graves déconvenues (ie. perte d'unicité de la pression). Afin d'illustrer la notion de condition “inf-sup” on donne quelques exemples et contre-exemples.

4.2 Exemples, contre-exemples

4.2.1 Instabilité en damier

L'exemple le plus connu d'espaces incompatibles est celui pour lequel la vitesse est approchée par des éléments polynomiaux Q_1 et la pression est approchée par des éléments P_0 (ie. des constantes par morceaux). Cette paire d'espaces donne lieu a une incompatibilité connue sous le nom d'instabilité en damier (cf. Girault-Raviart [10, p. 160]).

Considérons en dimension 2 le carré $\Omega =]0, 1[\times]0, 1[$ et définissons un maillage cartésien uniforme $\{x_{ij}\}$ pour $0 \leq i, j \leq I$. Posons $h = 1/I$ et notons M_{ij} la maille dont le nœud du bas à gauche a pour indices (i, j) . On note \mathcal{M}_h le maillage ainsi constitué. On définit maintenant les espaces d'approximation comme suit:

$$(4.23) \quad V_h = \{u_h \in C^0(\overline{\Omega})^2; \forall M \in \mathcal{M}_h \quad u_h|_M \in Q_1^2, u_h|_{\partial\Omega} = 0\}$$

$$(4.24) \quad P_h = \{p_h \in L_0^2(\Omega); \forall M \in \mathcal{M}_h \quad p_h|_M \in P_0\}$$

Pour simplifier on note $p_{i+1/2, j+1/2}$ la valeur de p_h sur M_{ij} et (u_{ij}, v_{ij}) les valeurs des composantes de u_h au nœud x_{ij} .

Pour montrer que la condition "inf-sup" n'est pas satisfaite il suffit de montrer qu'il existe un champ de pression $p_h \in P_h$ tel que $(p_h, \nabla \cdot u_h)$ est nul pour tout $u_h \in V_h$. Par définition de l'espace d'approximation de la pression, p_h est constant par maille et on a:

$$\begin{aligned} \int_{M_{ij}} p_h \nabla \cdot u_h &= \int_{\partial M_{ij}} p_h (u_h \cdot n) \\ &= hp_{i+1/2, j+1/2} \left(\frac{u_{i+1, j} + u_{i+1, j+1}}{2} + \frac{v_{i+1, j+1} + v_{i, j+1}}{2} - \right. \\ &\quad \left. \frac{u_{i, j} + u_{i, j+1}}{2} - \frac{v_{i, j} + v_{i+1, j}}{2} \right) \end{aligned}$$

En faisant la sommation sur toutes les mailles et en sommant par parties on obtient:

$$\int_{\Omega} p_h \nabla \cdot u_h = -h^2 \sum_{ij} (u_{i, j} (\partial_1 p)_{ij} + v_{i, j} (\partial_2 p)_{ij})$$

où la sommation est réalisée sur tous les nœuds intérieurs et par définition on a posé:

$$(4.25) \quad \begin{aligned} (\partial_1 p)_{ij} &= (1/2h)[p_{i+1/2, j+1/2} + p_{i+1/2, j-1/2} - p_{i-1/2, j+1/2} - p_{i-1/2, j-1/2}] \\ (\partial_2 p)_{ij} &= (1/2h)[p_{i+1/2, j+1/2} + p_{i-1/2, j+1/2} - p_{i+1/2, j-1/2} - p_{i-1/2, j-1/2}] \end{aligned}$$

Ainsi, le champ p_h appartient au noyau de B_h si $p_{i+1/2, j+1/2} = p_{i-1/2, j-1/2}$ et $p_{i-1/2, j+1/2} = p_{i+1/2, j-1/2}$. Ce système engendre un espace de dimension 2 dont l'une des directions de base est le champ constant, cette solution n'est pas gênante puisqu'on l'a exclue a priori de l'espace d'approximation P_h (ie. on veut un champ de pression a moyenne nulle). Une autre direction de base est engendrée par la fonction qui vaut +1 ou -1 alternativement sur deux mailles adjacentes comme sur un damier noir/blanc. Le noyau de B_h n'est donc pas réduit aux constantes. On conclue donc que la condition "inf-sup" n'est pas satisfaite, en d'autre termes les espaces V_h et P_h sont incompatibles pour résoudre le problème de Stokes.

On peut essayer de récupérer la situation en réduisant la taille de P_h , c'est à dire en imposant à la pression d'être orthogonale au mode gênant (on dit aussi mode parasite). Ce faisant on remplace un mal par un autre, on obtient ainsi effectivement deux espaces d'approximations compatibles mais la constante "inf-sup" tend vers zero avec h . Plus précisément on montre (cf. Girault-Raviart [10, p. 164]) que la constante β vérifie: $ch \leq \beta \leq c'h$ où les constantes c et c' sont strictement positives indépendantes de h . On verra par la suite que cela a des incidences néfastes sur la stabilité de l'approximation.

4.2.2 L'élément P_1/P_1 -bulle

Considérons Ω un ouvert polygonal de \mathbb{R}^2 (pour simplifier) et construisons sur $\bar{\Omega}$ une triangulation \mathcal{T}_h régulière. Pour chaque triangle T de la triangulation on introduit $(\lambda_1, \lambda_2, \lambda_3)$ les coordonnées barycentriques par rapport aux sommets et on définit $\mathcal{P}_1(T)$ l'espace des polynômes sur T engendrés par $(1, \lambda_1, \lambda_2, \lambda_1\lambda_2\lambda_3)$. Le degré de liberté $\lambda_1\lambda_2\lambda_3$ est parfois appelé degré de liberté "bulle" par référence à la forme de son graphe.

On définit maintenant les espaces d'approximations de la vitesse et de la pression comme suit:

$$(4.26) \quad V_h = \{u_h \in C^0(\bar{\Omega})^2; \forall T \in \mathcal{T}_h \quad u_{h|T} \in \mathcal{P}_1(T), u_{h|\partial\Omega} = 0\}$$

$$(4.27) \quad P_h = \{p_h \in L_0^2(\Omega); \forall T \in \mathcal{T}_h \quad p_{h|T} \in P_1\}$$

Les degrés de liberté sont les sommets et les barycentres des triangles pour la vitesse et les sommets des triangles pour la pression. En notant $u(1)$, $u(2)$, $u(3)$ et $u(4)$ la valeur de la vitesse respectivement aux sommets d'un triangle et au barycentre on a l'expression suivante de la vitesse sur le triangle en question:

$$u|_T = \lambda_1(1 - 9\lambda_2\lambda_3)u(1) + \lambda_2(1 - 9\lambda_3\lambda_1)u(2) + \lambda_3(1 - 9\lambda_1\lambda_2)u(3) + 27\lambda_1\lambda_2\lambda_3u(4)$$

de même la pression s'exprime localement par:

$$p|_T = \lambda_1p(1) + \lambda_2p(2) + \lambda_3p(3)$$

Pour les espaces d'approximation V_h et P_h ainsi définis on a le résultat important suivant (cf. Girault-Raviart [10, p. 175]):

Lemme 8 *Si la triangulation \mathcal{T}_h est régulière, les espaces V_h et P_h sont compatibles uniformément par rapport à h , c'est à dire il existe une constante β vérifiant (4.19) indépendamment de h .*

4.2.3 L'élément de Taylor–Hood: P_1/P_2

Le résultat de la section précédente peut être amélioré en prenant une meilleure approximation de la vitesse. En gardant pour la pression une représentation locale P_1 mais en prenant une représentation P_2 pour la vitesse on obtient encore deux espaces compatibles uniformément par rapport à h (cf. Girault-Raviart [10, p. 176]). Cet élément fini est connu sous le nom d'élément de Taylor–Hood.

On pourrait allonger encore la liste des éléments finis compatibles. Le lecteur intéressé trouvera de bonnes introductions à tous ces problèmes dans les ouvrages de Girault-Raviart [10] et Pironneau [16].

En conclusion, sous réserve de vérifier que l'opérateur B_h est injectif, c'est à dire en choisissant convenablement le couple (V_h, P_h) on a un problème discret (4.6) ou (4.15) bien posé. On peut maintenant se poser la question de la stabilité de la solution par rapport aux données et répondre ensuite au problème de la convergence de l'approximation discrète vers la solution du problème continu.

4.3 Stabilité – Convergence

4.3.1 Stabilité

Rappelons que le problème à résoudre s'écrit sous forme abstraite:

$$(4.28) \quad \begin{cases} A_h u_h + B_h p_h = f_h & \text{dans } V'_h \\ -B_h^t u_h = g_h & \text{dans } P_h \end{cases}$$

où l'opérateur A_h est symétrique défini positif (cf. lemme 5) et B_h satisfait la condition "inf-sup" (4.19). Afin de rester dans un cadre général on note α la constante de coercivité de A_h (ie. $\alpha = \min(1, \sigma)$ pour le problème de stokes).

Lemme 9 Soit $B : V_h \rightarrow V'_h$, les propriétés suivantes sont équivalentes:

(i) Il existe $\beta > 0$ tel que

$$(4.29) \quad \inf_{|q_h|_0 \neq 0} \sup_{|v_h|_1 \neq 0} \frac{(B_h q_h, v_h)}{|q_h|_0 |v_h|_1} \geq \beta$$

(ii) L'opérateur B_h est un isomorphisme de P_h sur $(\ker B_h^t)^{\perp_0}$ et pour tout $q_h \in P_h$, $|B_h q_h|_{V'_h} \geq \beta |q_h|_0$.

(iii) L'opérateur B_h^t est un isomorphisme de $(\ker B_h^t)^{\perp_1}$ sur P_h où la notation \perp_1 signifie l'orthogonal dans V_h au sens du produit scalaire $(\cdot, \cdot)_1$. De plus, pour tout $v_h \in (\ker B_h^t)^{\perp_1}$, $|B_h^t v_h|_0 \geq \beta |v_h|_1$.

Démonstration: (i) \Leftrightarrow (ii) L'inégalité (i) est équivalente à

$$\forall q_h \in P_h, \quad |B_h q_h|_{V'_h} = \sup_{|v_h|_1 \neq 0} \frac{(B_h q_h, v_h)}{|v_h|_1} \geq \beta |q_h|_0$$

Ce qui est aussi équivalent à l'injectivité de B_h (ie. en dimension finie) d'où B_h est un isomorphisme de P_h sur $\text{Im } B_h = (\ker B_h^t)^{\perp_0}$ d'après (4.18).

(ii) \Leftrightarrow (iii) Définissons l'opérateur Λ tel que

$$(4.30) \quad \begin{aligned} \Lambda : V_h &\longrightarrow V'_h \\ u_h &\longmapsto \Lambda(u_h) : v_h \longmapsto (\Lambda u_h, v_h) = (u_h, v_h)_1 \end{aligned}$$

Λ est manifestement un isomorphisme d'après le théorème de Lax-Milgram, c'est même une isométrie: $|\Lambda u_h|_{V'_h} = |u_h|_1$. En termes naïfs Λ est quasiment l'opérateur $Id_h - \nabla_h^2$. Λ réalise un isomorphisme entre $(\ker B_h^t)^{\perp_1}$ et $(\ker B_h^t)^{\perp_0}$. En effet, soit $f_h \in (\ker B_h^t)^{\perp_0} \subset V'_h$, il existe un unique $u_h \in V_h$ tel que $\Lambda u_h = f_h$; pour tout $w_h \in \ker B_h^t$ on a $(u_h, w_h)_1 = (\Lambda u_h, w_h) = (f_h, w_h) = 0$, d'où $u_h \in (\ker B_h^t)^{\perp_1}$. En conséquence dire que B_h est un isomorphisme de P_h sur $(\ker B_h^t)^{\perp_0}$ est équivalent à dire que B_h^t est un isomorphisme de $(\ker B_h^t)^{\perp_1}$ sur P_h .

Supposons (ii) vérifié. Soit $v_h \in (\ker B_h^t)^{\perp_1}$, d'après ce qui précède on a alors $\Lambda v_h \in (\ker B_h^t)^{\perp_0}$ et d'après (ii) il existe $q_h \in P_h$ tel $B_h q_h = \Lambda v_h$, d'où

$$|v_h|_1^2 = (\Lambda v_h, v_h) = (B_h q_h, v_h) = (q_h, B_h^t v_h) \leq |q_h|_0 |B_h^t v_h|_0$$

Mais d'après (ii) on a aussi

$$|q_h|_0 \leq \frac{1}{\beta} |B_h q_h|_{V'_h} = \frac{1}{\beta} |\Lambda v_h|_{V'_h} = \frac{1}{\beta} |v_h|_{V'_h}$$

Des deux inégalités ci-dessus on déduit le résultat cherché $\beta |v_h|_1 \leq |B_h^t v_h|_0$.

Réciproquement supposons (iii) vérifié. Soit $q_h \in P_h$, d'après (iii) il existe $v_h \in (\ker B_h^t)^{\perp 1}$ tel que $B_h^t v_h = q_h$, d'où

$$|B_h q_h|_{V'_h} = \sup_{w_h \in V_h - \{0\}} \frac{(B_h q_h, w_h)}{|w_h|_1} \geq \frac{(q_h, B_h^t v_h)}{|v_h|_1} \geq \beta \frac{|q_h|_0^2}{|B_h^t v_h|_0} = \beta |q_h|_0$$

et le résultat est établi. \square

On est maintenant en mesure d'établir le résultat de stabilité de la solution du problème discret par rapport aux données:

Théorème 21 *La solution de (4.28) satisfait les inégalités de stabilité suivantes:*

$$(4.31) \quad |u_h|_1 \leq \frac{|f_h|_{V'_h}}{\alpha} + \frac{|g_h|_0}{\beta} \left(1 + \frac{|A_h|}{\alpha}\right)$$

$$(4.32) \quad |p_h|_1 \leq \frac{|f_h|_{V'_h}}{\alpha} \left(1 + \frac{|A_h|}{\alpha}\right) + \frac{|g_h|_0}{\beta^2} |A_h| \left(1 + \frac{|A_h|}{\alpha}\right)$$

Démonstration: (i) B_h étant surjectif, il existe $v_{gh} \in (\ker B_h^t)^{\perp 1}$ tel que $B_h^t v_{gh} = g_h$ et $|g_h|_0 \geq \beta |v_{gh}|_1$ d'après le lemme 9.

(ii) On peut réécrire le problème sous la forme

$$\begin{cases} A_h(u_h - v_{gh}) + B_h p_h = f_h - A_h v_{gh} & \text{dans } V'_h \\ -B_h^t(u_h - v_{gh}) = 0 & \text{dans } P_h \end{cases}$$

En dualisant la première équation par $(u_h - v_{gh})$ et en utilisant la coercivité de A_h on obtient:

$$\alpha |u_h - v_{gh}|_1^2 \leq (|f_h|_{V'_h} + |A_h| |v_{gh}|_1) |u_h - v_{gh}|_1$$

ce qui fournit l'inégalité de stabilité sur u_h compte tenu de l'inégalité $|v_{gh}|_1 \leq |g_h|_0 / \beta$.

(iii) L'inégalité sur la pression est obtenue grâce à la condition "inf-sup"

$$\beta |p_h|_0 \leq |B_h p_h|_{V'_h} = |f - A_h u_h|_{V'_h} \leq |f|_{V'_h} + |A_h| |u_h|_1$$

Ce qui donne le résultat cherché en combinant cette inégalité avec la précédente. \square

Remarque: Ces inégalités montrent bien que la stabilité de l'approximation est contrôlée par les constantes $1/\alpha$ et $1/\beta$. Si $\beta = 0$ on n'a pas stabilité (continuité) de la pression par rapport aux données. En fait dans ce cas on perd l'unicité de la pression, il y a seulement unicité modulo n'importe quelle pression dans le noyau de B_h .

4.3.2 Convergence

La démonstration de la convergence se fait exactement comme dans le cas de l'approximation des solutions variationnelles des problèmes de diffusion. La consistance est assurée par le caractère interne de l'approximation par le fait que V_h et P_h approchent V et P lorsque $h \rightarrow 0$. La stabilité est assurée par la coercivité de A_h et la condition "inf-sup" sur B_h . La convergence résulte de la consistance et de la stabilité de l'approximation. Le lecteur peut le vérifier à titre d'exercice.

4.4 Résolution du système linéaire

4.4.1 Introduction

On a vu précédemment que l'approximation par éléments finis du problème de Stokes se ramène à résoudre un système linéaire du type:

$$(4.33) \quad \begin{cases} AU + BP = F \\ -B^t U = G \end{cases}$$

La résolution de ce système par des méthodes directes est généralement très coûteuse vu la taille du système. On est donc conduit à rechercher des méthodes itératives.

Une des difficultés du système linéaire ci-dessus est que la matrice qui lui est associée est positive mais non-définie. On ne peut donc pas utiliser les techniques du type Gauss-Seidel ou Jacobi par blocs. Les méthodes de gradient non plus ne sont pas très efficaces sur ce genre de matrice. Il faut donc trouver autre chose! En fait la structure du système linéaire (4.33) est, avec les problèmes de compatibilité d'espaces d'approximation évoqués plus haut, l'une des spécificités de l'approximation des problèmes d'écoulements incompressibles.

On donne brièvement par la suite quelques exemples de techniques très utilisées pour résoudre (4.33).

4.4.2 Compressibilité artificielle

La première technique consiste à pénaliser la contrainte d'incompressibilité. C'est à dire, au lieu de résoudre (4.33) on considère plutôt

$$(4.34) \quad \begin{cases} AU + BP = F \\ -B^t U + \epsilon P = G \end{cases}$$

où $\epsilon > 0$ est un coefficient de pénalisation assez petit. On peut alors éliminer la pression dans l'équation de la quantité de mouvement, il reste:

$$(4.35) \quad \left(A + \frac{1}{\epsilon} BB^t \right) U = F - \frac{1}{\epsilon} BG.$$

On résout ce système par une technique standard (gradient conjugué, etc...) puisque la matrice du nouveau système est symétrique définie positive. On peut montrer la convergence de la solution du système pénalisé vers celle du système non pénalisé lorsque ϵ tend vers zéro.

Cette technique est connue sous le nom de méthode de compressibilité artificielle.

4.4.3 Opérateur d'Uzawa

Une autre technique assez populaire consiste à éliminer la vitesse de (4.33), on obtient alors:

$$(4.36) \quad B^t A^{-1} B P = B^t A^{-1} F + G.$$

La matrice $B^t A^{-1} B$ est souvent appelée opérateur d'Uzawa par référence à l'algorithme d'Uzawa qui fait intervenir cet opérateur implicitement pour résoudre (4.33) itérativement. L'opérateur d'Uzawa est symétrique défini positif. Il est quasiment exclu de résoudre ce

nouveau système par une méthode directe car cela nécessiterait le calcul explicite de la matrice A^{-1} , un tel calcul serait très coûteux. En conséquence on est contraint d'utiliser des méthodes itératives standard pour cette approche.

4.5 Approximation par différences finies

L'approximation du problème de Stokes par différences finies conduit toujours à un système linéaire du type

$$(4.37) \quad \begin{cases} AU + B_1 P = F \\ -B_2^t U = G \end{cases}$$

où la matrice A est le plus souvent symétrique définie positive. Par contre il y a assez peu de chance que les matrices B_1 et B_2 soient identiques si on ne fait pas vraiment attention à la discrétisation de ces opérateurs.

Si les nœuds de pression et de vitesse sont collocalifs, il est alors quasiment impossible d'avoir la transposition des opérateurs gradient et divergence sans perdre de consistance. Pour retrouver la symétrie salvatrice on est généralement conduit à désolidariser les nœuds de pression et de vitesse. On dit qu'on des maillages décalés ou entrelacés ("staggered grids" en anglais).

Pour illustrer la difficulté du problème dans le cas général donnons un résultat général d'existence et d'unicité pour le problème (4.37). Soient $K_1 = \ker B_1^t$ et $K_2 = \ker B_2^t$. Soit π_1 la projection des vecteurs vitesse sur K_1 .

Proposition 10 *Le problème (4.37) est bien posé ssi*

- (i) $\pi_1 A : K_2 \longrightarrow K_1$ est un isomorphisme.
- (ii) B_1 et B_2 satisfont chacun une condition "inf-sup".

La condition (i) est en général très difficile à vérifier car la projection π_1 fait perdre la symétrie. De plus, au lieu d'avoir à satisfaire une condition "inf-sup" on doit en satisfaire deux.

En résumé l'approximation du problème de Stokes par des différences finies est possible mais requière beaucoup de prudence dans le choix des discrétisations (cf. Peyret-Taylor pour d'autres détails [15] sur ces problèmes et une large revue des techniques numériques utilisées en MFN).

Bibliographie

- [1] Adams, R. A., *Sobolev spaces*, Academic Press, New-York, San Francisco, London, 1975.
- [2] Aubin J.-P., *Analyse fonctionnelle appliquée*, Vol. 1 et 2, Presses Universitaires de France, Paris, 1987.
- [3] Babuška I., The finite element method with Lagrangian multipliers, *Numer. Math.*, 20, p. 179-192, 1973.
- [4] Bernardi C., Maday Y., Approximations spectrales des problèmes aux limites elliptiques, *Mathématiques&Applications* 10, Springer-Verlag, 1992.
- [5] Brezis H., *Analyse fonctionnelle. Théorie et applications*, Masson, Paris, 1983.
- [6] Brezzi F., On the existence, uniqueness and approximation of saddle-point problems arising from Lagrange multipliers, *R.A.I.R.O., Anal. Numer. R2*, p. 129-151, 1974.
- [7] Ciarlet, P. G., *The finite element method for elliptic problems*, North-Holland, Amsterdam, 1978.
- [8] Dautray R., Lions J.-L., *Analyse mathématique et calcul numérique*, Vol. 1 à 9, Masson, Paris, 1984.
- [9] Forsythe G. E., Wasow W. R., 1960, *Finite-Difference Methods for Partial Differential Equations*, Wiley, New-York, 1960.
- [10] Girault V., Raviart P.-A., 1986, *Finite Element Methods for Navier–Stokes Equations*, Springer Series in Computational Mathematics, Vol. 5, Springer–Verlag, Berlin, 1986.
- [11] Godounov S., Riabenki V., 1973, *Schémas aux différences*, Traduction française Editions MIR, 1977.
- [12] Ladyzhenskaya O. A., *The Mathematical Theory of Viscous Incompressible Flow*, Gordon and Breach, New York, 1969.
- [13] Lions J.-L., *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod–Gauthier–Villar, 1969.
- [14] Lions J.-L., Magenes E., *Problèmes aux limites non homogènes et applications*, Vol. 1, Dunod, 1968.
- [15] Peyret R., Taylor C., *Methods in Fluid Flow*, Springer-Verlag, New-York, 1985.
- [16] Pironneau O., *Méthodes des éléments finis pour les fluides*, Masson, 1983.

- [17] Raviart P.-A., Thomas J.-M., *Introduction à l'analyse numérique des équations aux dérivées partielles*, Masson, 1983.
- [18] Rudin W., *Real and Complex Analysis*, Edition française, Masson, 1975
- [19] Strikwerda J. C., *Finite Difference Schemes and Partial Differential Equations*, Wadsworth & Brooks/Cole, 1989.
- [20] Temam R., *Navier-Stokes Equations*, Studies in Mathematics and its Applications, 2 (revised edition), North-Holland, Amsterdam, 1979.