

Dédicaces

Je dédie ce mémoire A mes très chers parents, qui sont la graine de mon existence et la source de ma réussite, pour leurs encouragements et leurs sacrifices. A mes cher frères Tarek, Aymen, and Sofiane. A mes amis. A mes enseignants.

Remerciements

Je remercie en première lieu **Allah** tout puissant pour la volonté, la santé, et la patience qu'il ma donné durant ces longue années d'étude et le courage pour terminer ce mémoire. Je tiens à remercier *Dr. MEKKAOUI Imen* pour le soutien et l'encadrement qu'elle ma donné. Je voudrais aussi remercier l'ensemble des membres du jury *Dr. NASRI Rafik* et *Dr. HALIMI Abderazzak* et *Dr. DJERFI Kowider* de m'avoir honoré par leurs évaluation de ce travail. Enfin, je remercie ma famille de m'avoir donné tout ce dont j'avais besoin pour réussir dans mes études.

Résumé

Les équations différentielles aux dérivées partielles peuvent modéliser un nombre important de phénomènes naturels, physiques, chimiques, etc,... Cependant, dans la pratique les paramètres de ces équations ne sont en général pas bien connus. Nous sommes donc devant un problème de contrôle ou d'estimation de paramètres. C'est le problème que nous avons essayé d'aborder dans ce mémoire. Le contrôle ou l'estimation de paramètres, fait intervenir l'optimisation de certaine fonction coût qui s'écrit en fonction de données observées ou déjà disponibles à l'avance et d'une solution d'un modèle s'écrivant sous forme d'EDP. Le problème d'optimisation est résolu en calculant le gradient de cette fonction coût qui n'est en général pas trivial. Pour le faire, on a recours à la méthode adjointe pour le calculer efficacement. Dans les deux premiers chapitres, nous faisons un rappel de toutes les notions nécessaires sur les EDPs, des résultats d'existence et d'unicité de solution, leurs formulations variationnelles qui sera la base de leur approximation numérique par la méthode des éléments finis. Dans le troisième chapitre, nous nous intéressons à l'application du calcul adjoint au contrôle de paramètres d'EDPs avec des exemples illustratifs présentés. Dans le dernier chapitre, nous introduisons une application numérique effectuée sous MatLab illustrant les résultats obtenus en s'intéressant à l'identification de paramètres d'EDPs paraboliques.

Mots-clés : Equations aux dérivées partielles (EDPs), Optimisation, Simulation numérique, Méthode des éléments finis, Calcul adjoint, Contrôle optimal ou identification de paramètres.

Abstract

Partial differential equations (PDEs) can model a large number of natural, physical, chemical phenomena, etc. However, in practice the parameters of these equations are generally not well known. We are therefore faced with a problem of control or estimation of parameters. This is the problem we have tried to address in this work. The control or estimation of parameters involves the optimization of a certain cost function which is written as a function of data observed or already available in advance and of a solution of a model which is written in the form of PDE. The optimization problem is solved by calculating the gradient of this cost function which is generally not trivial. To do this, we use the adjoint method to calculate it efficiently. In the first two chapters, we recall all the necessary notions on PDEs, results of existence and uniqueness of solution, their variational formulations which will be the basis of their numerical approximation by the finite element method. In the third chapter, we are interested in the application of the adjoint calculus to the control of parameters of PDEs with illustrative examples presented in the last sections. In the last chapter, we introduce a numerical application carried out under MatLab illustrating the results obtained by focusing on the identification of parameters of parabolic PDEs.

Keywords : Partial Differential Equations (PDEs), Optimization, Numerical Simulation, Finite Element Method, Adjoint Calculus, Optimal control or identification of parameters.

Table des matières

Introduction générale	1
1 Résultats préliminaires	3
1.1 Introduction	3
1.2 Equations aux Dérivées Partielles	3
1.2.1 Dérivées Partielles	4
1.2.2 Les Equations aux Dérivées Partielles	4
1.3 Quelque types des Equations aux Dérivées Partielles	5
1.3.1 Équations aux dérivées partielles elliptiques [1]	5
1.3.2 Équations aux dérivées partielles paraboliques [1]	6
1.3.3 Équation aux dérivées partielles hyperboliques [1]	7
1.4 Rappel sur les espaces de Banach et de Hilbert	9
1.5 Distributions	10
1.6 Espaces de Sobolev	11
1.7 Formulation variationnelle	13
1.7.1 Problème de Dirichlet homogène	13
1.7.2 Problème de Dirichlet non homogène	15
1.7.3 Problème de Neumann	16
1.8 Théorème de Lax-Milgram	17
1.8.1 Application aux problèmes elliptiques	18
1.8.2 Application aux problèmes paraboliques	21
2 Méthode des éléments finis	25
2.1 Introduction	25
2.2 Bref historique	25
2.3 Approximation interne générale	27

2.4	Méthode de Galerkin	29
2.4.1	Principe de base d'une méthode de Galerkin	30
2.5	Méthode des éléments finis	31
2.5.1	Méthode des éléments finis en dimension un	31
2.5.2	Élément fini \mathbb{P}_1	32
2.5.3	Résolution du problème de Dirichlet	34
2.5.4	Résolution de problèmes paraboliques	37
3	Méthode de calcul adjoint et contrôle de paramètres	40
3.1	Introduction	40
3.2	Motivation	41
3.3	Rappel d'optimisation	43
3.3.1	Fonctions convexes et résultat d'existence et d'unicité	43
3.3.2	Dérivée directionnelle et dérivée au sens de Fréchet	46
3.3.3	Minimisation avec contrainte d'égalité	48
3.3.4	Méthode de calcul adjointe	49
3.4	Contrôle de paramètres d'équation différentielle ordinaires	51
3.5	Ecriture générale dans le cas du contrôle de la condition initiale	53
3.5.1	Dérivée directionnelle de J	53
3.5.2	Modèle linéaire tangent	54
3.5.3	Modèle adjoint	55
3.5.4	Calcul du gradient	56
3.6	Equation de Burgers	56
3.6.1	Contrôle de condition initiale sous contrainte de moyenne nulle	56
3.6.2	Contrôle de conditions initiale, et aux limites	58
4	Application numérique à l'identification de paramètres d'EDPs	61
4.1	Introduction	61
4.2	Dérivation du modèle	62
4.3	Résultats numériques	64
4.4	Code MatLab	70
	Conclusion générale	74

[Bibliographie](#)

75

Introduction générale

Les équations aux dérivées partielles (EDP) sont des outils mathématiques puissants pour modéliser des phénomènes physiques variés, tels que la mécanique des fluides, l'électromagnétisme, la thermique ou la biologie. Leur résolution numérique nécessite des méthodes efficaces et robustes, capables de traiter des problèmes complexes et de grande taille. Parmi ces méthodes, le calcul adjoint est une technique qui permet de calculer la sensibilité d'une quantité d'intérêt par rapport à des paramètres du modèle ou du problème. Le calcul adjoint est particulièrement utile pour des applications telles que le contrôle optimal, l'optimisation de forme, l'assimilation de données ou l'identification de paramètres.

Dans ce mémoire, nous nous intéressons au calcul adjoint et à la simulation numérique des EDPs. Nous présentons les principes généraux du calcul adjoint et les différentes approches pour le mettre en oeuvre.

Dans le premier chapitre, nous allons introduire les notions préliminaires nécessaires à notre travail. Nous commençons par rappeler quelques notions sur les équations différentielles ordinaires et les équations aux dérivées partielles. Nous allons présenter quelques types d'EDPs linéaires comme les équations elliptiques, paraboliques, hyperboliques. Nous allons également faire un rappel sur les résultats d'existence et d'unicité de solution au sens faible en passant par la formulation variationnelle d'EDPs.

Dans le deuxième chapitre nous présentons la méthode des éléments finis qui est la méthode numérique de référence pour le calcul des solutions de problèmes aux limites. Le principe de cette méthode est directement issu de l'approche variationnelle que nous avons vu dans le chapitre précédent. L'idée est de remplacer l'espace de Hilbert sur lequel est posée la formulation variationnelle, par un sous-espace de dimension finie. Le problème approché

se ramène une la simple résolution d'un système linéaire. Le domaine spatial est discrétisé en utilisant un maillage d'éléments finis. Une discrétisation temporelle est aussi présentée pour les EDPs évolutives dans le temps.

Ensuite, dans le troisième chapitre nous allons introduire quelques notions fondamentales d'optimisation, puis présenter des calculs formels sur le gradient de problèmes d'optimisation. Cela sert à résoudre le problème d'identification ou d'estimation de paramètres d'EDPs. En calculant la dérivée du Lagrangien associé au problème d'optimisation, nous obtenons une expression pour le gradient de la fonction coût à minimiser. Ce gradient est calculée en fonction de la solution d'un problème de l'état adjoint associé. Nous présentons également quelques algorithmes d'optimisations utilisés dans le calcul numérique du gradient. Nous donnons enfin quelques exemples d'optimisation et de calcul adjoints liés aux solutions d'EDPs dans le but d'identifier certains paramètres d'EDPs.

Dans le dernier chapitre une application numérique est faite à l'aide du langage MATLAB. Nous allons nous intéresser à l'identifications de paramètres de problèmes paraboliques. Nous allons tout d'abord introduire la simulation numérique d'une solution d'une EDP de référence qui va jouer le rôle de données observées, puis nous introduisons le modèle à qui nous voulons identifier les paramètres. Nous présentons le problème de minimisation de la fonctionnelle coût. Nous donnons ensuite les calculs formels liés au gradient de cette fonctionnelle. Nous présentons les solutions numériques de l'EDP de référence et de l'EDP modélisant le problème adjoint par la méthode des éléments finis pour la discrétisation en espace et une méthode d'Euler implicite pour la discrétisation en temps. Ensuite nous calculons la solution du problème de minimisation par une méthode de descente de gradient. Nous donnons finalement différentes parties du code MATLAB utilisé.

Chapitre 1

Résultats préliminaires

1.1 Introduction

Dans ce chapitre, nous allons introduire les notions préliminaires nécessaires à notre étude. Nous allons commencer par rappeler quelques notions sur les équations différentielles ordinaires et les équations aux dérivées partielles. Nous allons ensuite présenter les différents types d'EDPs. Nous allons également faire un rappel sur les espaces de Banach et de Hilbert. Enfin, nous allons voir quelques notions sur les distributions, les espaces de Sobolev et les formulations variationnelles.

1.2 Equations aux Dérivées Partielles

Le caractère particulier d'une équation aux dérivées partielles (EDP) est de mettre en jeu des fonctions de plusieurs variables :

$$(x, y, \dots) \mapsto u(x, y, \dots)$$

Une EDP est alors une relation entre les variables et les dérivées partielles de u .

1.2.1 Dérivées Partielles

On introduira au fur et à mesure quelques notions sur les fonctions de plusieurs variables réelles. On se limite pour les énoncés au cas de fonctions de deux variables, mais les notions qui suivent se généralisent facilement aux fonctions de n variables réelles, où n est un entier quelconque (supérieur à 2). Pour le moment, nous n'examinons que les propriétés des applications partielles associées à une telle fonction f .

Définition 1.2.1. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ et $(x_0, y_0) \in \mathbb{R}^2$. On appelle applications partielles associées à f en (x_0, y_0) , les deux applications de \mathbb{R} dans \mathbb{R} obtenues en fixant l'une des variables :

$$f_1 : x \mapsto f_1(x) := f(x, y_0) \text{ et } f_2 : x \mapsto f_2(y) := f(x_0, y)$$

La notion de dérivée partielle de f en un point (x_0, y_0) est alors particulièrement simple : il s'agit des dérivées des applications partielles associées à f en (x_0, y_0) .

Définition 1.2.2. Soit $\Omega =]a, b[\times]c, d[$ dans \mathbb{R}^2 , et $f : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ une application. Soit $(x_0, y_0) \in \Omega$ et $f_1 :]a, b[\rightarrow \mathbb{R}$ l'application définie par :

$$f_1(x) = f(x, y_0)$$

On dit que f admet une dérivée partielle par rapport à la première variable en (x_0, y_0) lorsque f_1 est dérivable en x_0 . On note $\partial_1 f(x_0, y_0)$ ou encore $\partial_x f(x_0, y_0)$ le nombre $f_1'(x_0)$

De la même manière, si elle existe, on note $\partial_2 f(x_0, y_0)$ la dérivée partielle de f par rapport à la deuxième variable en (x_0, y_0) .

1.2.2 Les Equations aux Dérivées Partielles

Dans le cas de deux variables, une EDP d'ordre 1 s'écrit :

$$F(x, y, u(x, y), \partial_x u(x, y), \partial_y u(x, y)) = 0 \quad (1.1)$$

et une équation du second ordre s'écrit :

$$F(x, y, u(x, y), \partial_x u(x, y), \partial_y u(x, y), \partial_x^2 u(x, y), \partial_x \partial_y u(x, y), \partial_y^2 u(x, y)) = 0 \quad (1.2)$$

Plus généralement, on peut considérer des équations mettant en jeu des dérivées $\partial_x^{m_j} \partial_y^{n_j} u$. L'ordre d'une EDP est alors le plus grand ordre de dérivation $m_j + n_j$ qui apparaît dans l'équation.

Résoudre une EDP dans un domaine Ω de \mathbb{R}^d (d est le nombre de variables), c'est trouver une fonction suffisamment différentiable dans Ω , telle que la relation (1.1) soit satisfaite pour toutes les valeurs des variables dans Ω .

1.3 Quelques types des Equations aux Dérivées Partielles

1.3.1 Équations aux dérivées partielles elliptiques [1]

Nous allons dans cette partie étudier les problèmes aux limites de la forme suivante :

$$\begin{cases} Lu = f & \text{dans } \Omega, \\ u = 0 & \text{sur } \partial\Omega. \end{cases} \quad (1.3)$$

où Ω un ouvert de \mathbb{R}^n et $u : \bar{\Omega} \rightarrow \mathbb{R}$ une fonction inconnue, $u = u(x)$. La fonction $f : \Omega \rightarrow \mathbb{R}$ est donnée. L est un opérateur aux dérivées partielles de second ordre s'écrivant sous la forme :

$$Lu = - \sum_{i,j=1}^n a_{i,j}(x) \partial_{i,j}^2 u(x) + \sum_{i=1}^n b_i(x) \partial_i u(x) + c(x)u(x) \quad (1.4)$$

avec $A(x) = (a_{i,j}(x))$, $b_i(x)$, $c(x)$ sont des coefficients donnés. La condition $u = 0$ sur $\partial\Omega$ est souvent appelée condition aux limites de Dirichlet.

On dit qu'un opérateur différentiel linéaire du second ordre L est elliptique s'il existe une constante $\lambda > 0$ telle que les coefficients de la matrice $A(x) = (a_{i,j}(x))$ satisfaisant la propriété d'ellipticité :

$$\sum_{i,j=1}^n a_{i,j}(x) \xi_i \xi_j \geq \lambda |\xi|^2 \quad \text{pour tout } x \in \Omega \text{ et tout } \xi \in \mathbb{R}^n \quad (1.5)$$

Comme la matrice hessienne $D^2u = (\partial_{i,j}^2 u)$ est symétrique, on peut se restreindre au cas de matrices $A(x)$ qui sont symétriques, auquel cas la condition

d'ellipticité est équivalente au fait que la matrice $A(x)$ est définie positive et que sa plus petite valeur propre est plus grande que $\lambda > 0$.

Equation de Poisson

Une telle équation s'écrit sous la forme :

$$\begin{cases} -\Delta u = f & \text{dans } \Omega, \\ u = 0 & \text{sur } \partial\Omega. \end{cases} \quad (1.6)$$

Cette équation intervient dans de nombreux domaines des mathématiques et de ses applications. Par exemple, en dimension $n = 3$, si f représente une densité de charge électrique présente dans Ω , alors $-u$ est le potentiel électrique dans Ω quand le bord de Ω est parfaitement conducteur. Le gradient de $-u$ est le champ électrique. Plus généralement, ce problème intervient dans les questions relatives au potentiel newtonien.

En dimension $n = 2$, il s'agit de l'équation de la membrane élastique : f représente une densité volumique de forces et u représente le déplacement vertical d'une membrane qui occupe la position $\bar{\Omega}$ au repos. La condition limite $u = 0$ sur $\partial\Omega$ signifie que la membrane est fixée sur le bord.

1.3.2 Équations aux dérivées partielles paraboliques [1]

Cette section traite les équations aux dérivées partielles paraboliques. L'idée est que la solution d'une telle équation est qu'elle évolue dans le temps à partir d'une configuration initiale donnée.

Pour un ouvert $\Omega \in \mathbb{R}^n$ et un temps fixé $T > 0$, on définit le domaine spatio-temporel $\Omega_T = \Omega \times [0, T]$, et le problème suivant :

$$\begin{cases} \partial_t u + lu = f & \text{dans } \Omega_T, \\ u = 0 & \text{sur } \partial\Omega \times [0, T], \\ u = u_0 & \text{sur } \Omega \times \{t = 0\}, \end{cases} \quad (1.7)$$

où $u : \bar{\Omega}_T \rightarrow \mathbb{R}$ une fonction inconnue, $u = u(x, t)$. Les fonctions $f : \Omega_T \rightarrow \mathbb{R}$, $u_0 : \Omega \rightarrow \mathbb{R}$ sont données. L'opérateur l s'écrit, pour chaque instant

t , sous la forme :

$$lu = - \sum_{i,j=1}^n a_{i,j}(x,t) \partial_{i,j}^2 u(x,t) + \sum_{i=1}^n b_i(x,t) \partial_i u(x,t) + c(x,t)u(x,t) \quad (1.8)$$

avec $A(x,t) = (a_{i,j}(x,t))$, $b_i(x,t)$, $c(x,t)$ sont des coefficients donnés.

L'opérateur $\partial_t + l$ est dit parabolique si pour tout instant $t \in [0, T]$ l'opérateur l est elliptique en x .

Les équations paraboliques décrivent dans les applications physiques l'évolution en temps d'une densité de certaine quantité u , disant par exemple une concentration chimique dans une région Ω . Le terme d'ordre deux

$\sum_{i,j=1}^n a_{i,j}(x,t) \partial_{i,j}^2 u(x,t)$ décrit la diffusion, le terme d'ordre un $\sum_{i=1}^n b_i(x,t) \partial_i u(x,t)$ représente le transport, et le terme d'ordre zéro $c(x,t)u(x,t)$ représente la création ou l'épuisement de la quantité u .

Un cas plus simple des équations paraboliques est l'équation de diffusion ou de la chaleur.

Equation de la chaleur

Etant donné un réel positif k , on appelle équation de la chaleur, ou équation de diffusion pour le coefficient de diffusion k , la donnée suivante :

$$\partial_t u = k \Delta u.$$

Cette équation est qualifiée de parabolique (par analogie avec l'équation générale d'une parabole $y = \xi^2/a^2$).

1.3.3 Équation aux dérivées partielles hyperboliques [1]

Dans cette partie nous allons voir un autre type d'équations aux dérivées partielles, nommé hyperbolique. Cette classe d'équation modélise les phénomènes de propagation, émergeant par exemple naturellement en mécanique.

Le modèle est donné par le problème suivant :

$$\begin{cases} \partial_{tt} u + Lu = f & \text{dans } \Omega_T, \\ u = 0 & \text{sur } \partial\Omega \times [0, T], \\ u = u_0, \partial_t u = g & \text{sur } \Omega \times \{t = 0\}, \end{cases} \quad (1.9)$$

Les fonctions f , u_0 , g sont données. La fonction u est l'inconnue. L'opérateur L est défini comme dans le cas d'équations paraboliques par l'expression (1.21).

Un cas particulier d'équations hyperboliques est l'équation d'onde suivante :

Equation d'ondes

Etant donné un réel positif c , on appelle équation d'ondes pour la vitesse c :

$$\partial_{tt}u - c^2\Delta u = 0.$$

Cette équation est qualifiée d'hyperbolique (par analogie avec l'équation générale d'une hyperbole $\xi_1^2/a^2 - \xi_2^2/b^2 = 1$)

Remarque 1.3.1. *Les équations d'ondes et de la chaleur sont dites d'évolution car elles modélisent en général un phénomène instationnaire, évoluant avec le temps t . L'équation de Poisson quant à elle est stationnaire. Elle modélise en général un phénomène à l'équilibre. Les trois équations sont linéaires, c'est-à-dire qu'elles dépendent linéairement de l'inconnue u . Nous n'étudierons pas ici les équations non-linéaires. Les équations de Poisson et de la chaleur modélisent des phénomènes de diffusion, comme celle de la chaleur de la matière (par exemple un polluant dans une rivière, ou des bactéries dans un organe, etc.), ou encore d'une charge électrique. L'équation de Poisson pour $f = 0$, aussi appelée équation de Laplace, peut être vue comme un cas particulier d'équation de la chaleur lorsque l'équilibre est atteint, c'est-à-dire lorsque l'inconnue u ne dépend plus de t . L'équation d'ondes modélise des phénomènes de propagation, comme celle du son, de la lumière.*

1.4 Rappel sur les espaces de Banach et de Hilbert

Définition 1.4.1. Soit $\{u_n\}_{n \in \mathbb{N}}$ une suite dans un espace vectoriel normé $(V, \|\cdot\|)$. On dit que cette suite est de Cauchy si :

$$\lim_{n, m \rightarrow +\infty} \|u_n - u_m\| = 0.$$

L'espace normé $(V, \|\cdot\|)$ est dit de Banach si toute suite de Cauchy dans V converge vers un élément de V (pour la norme $\|\cdot\|$). En d'autres termes, un espace de Banach est un espace vectoriel normé complet.

Définition 1.4.2. Un espace de Hilbert est un espace de Banach (donc complet) dont la norme $\|\cdot\|$ découle d'un produit scalaire ou hermitien $\langle \cdot, \cdot \rangle$ par la formule $\|x\| = \sqrt{\langle x, x \rangle}$. C'est la généralisation en dimension quelconque d'un espace euclidien ou hermitien.

Théorème 1.4.1. [2] Un espace de Banach (respectivement espace vectoriel normé) est un espace de Hilbert (respectivement espace préhilbertien) ssi sa norme vérifie l'égalité :

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

qui signifie que la somme des carrés de côtés d'un parallélogramme est égale à la somme des carrés des diagonales.

Théorème 1.4.2 (Inégalité de Cauchy-Schwarz[2]). Soit $(V, \langle \cdot, \cdot \rangle)$ un espace préhilbertien réel muni d'un produit scalaire. Soit $\|\cdot\|$ la norme euclidienne associée à $\langle \cdot, \cdot \rangle$. Alors on a :

$$|\langle u, v \rangle| \leq \|u\| \|v\|, \quad \forall u, v \in V$$

Définition 1.4.3 (Espaces $L^p(\Omega)$). Considérons l'espace $L^p(\Omega)$ des fonctions, pour tout $(1 \leq p \leq \infty)$ est défini par :

$$L^p(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{R} \mid f \text{ mesurable tel que } \int_{\Omega} |f(x)|^p dx < \infty \right\}$$

Remarque 1.4.1. *l'espace $L^2(\Omega)$ est un espace de Hilbert.*

$$\langle f, g \rangle_{L^2(\Omega)} = \int_{\Omega} f(x)g(x) dx.$$

La norme de $L^2(\Omega)$ associée à ce produit scalaire est donc donnée par :

$$\|f\|_{L^2(\Omega)} = (\langle f, f \rangle_{L^2(\Omega)})^{\frac{1}{2}} = \left(\int_{\Omega} |f(x)|^2 dx \right)^{\frac{1}{2}}$$

Théorème 1.4.3 (Inégalité de Hölder[3]). *Pour toutes $f \in L^p(\Omega)$ et $g \in L^{p'}(\Omega)$ avec $1 = \frac{1}{p} + \frac{1}{p'}$, on a : $fg \in L^1(\Omega)$ et ainsi l'inégalité :*

$$\int_{\Omega} |f(x)g(x)| dx \leq \|f\|_{L^p} \|g\|_{L^{p'}}.$$

1.5 Distributions

Définition 1.5.1. *Soient Ω un ouvert de \mathbb{R}^n ($n \geq 1$) et f une fonction $f : \Omega \rightarrow \mathbb{R}$. On appelle support de f l'ensemble :*

$$\text{supp } f = \overline{\{x \in \Omega \text{ tel que } :f(x) \neq 0\}}$$

Définition 1.5.2. *On dit que la fonction f est à support compact s'il existe un sous-ensemble compact $K \in \Omega$ tel que $\text{supp } f \in K$. On définit ainsi l'espace $D(\Omega)$ par $D(\Omega) = \{f \in C^\infty(\Omega) \text{ tel que } : f \text{ est à support compact}\}$.*

Définition 1.5.3. *On peut maintenant définir l'espace des distributions sur Ω . Soit T une application linéaire de $D(\Omega)$ dans \mathbb{R} . On note par $\langle T, \varphi \rangle$ la valeur de T correspondante à l'élément $\varphi \in D(\Omega)$, c'est-à-dire $\langle T, \varphi \rangle = T(\varphi)$. On dit que T est continue si :*

$$\lim_{n \rightarrow +\infty} \langle T, \varphi_n \rangle = \langle T, \varphi \rangle$$

où $\{\varphi_k\}_{k \in \mathbb{N}}$ est une suite arbitraire de $D(\Omega)$ qui converge vers φ dans $D(\Omega)$.

On appelle distribution sur $D(\Omega)$ toute application $T : D(\Omega) \rightarrow \mathbb{R}$ linéaire et continue.

L'espace des distributions sur Ω est l'espace $D'(\Omega)$, dual de $D(\Omega)$, c'est-à-dire, l'espace des applications $T : D(\Omega) \rightarrow \mathbb{R}$ linéaires et continues.

Si à une fonction $f \in L^p(\Omega)$ on associe une distribution $T_f \in D'(\Omega)$, alors elle va être définie par :

$$\langle T_f, \varphi \rangle = \int_{\Omega} f(x)\varphi(x)dx, \quad \forall \varphi \in D(\Omega).$$

Définition 1.5.4. Soit $T \in D(\Omega)$ avec Ω un ouvert de \mathbb{R}^n . La dérivée partielle de T par rapport à x_i avec $i = (1, \dots, n)$ au sens des distributions est une nouvelle distribution définie de la manière suivante :

$$\left\langle \frac{\partial T}{\partial x_i}, \varphi \right\rangle = - \left\langle T, \frac{\partial \varphi}{\partial x_i} \right\rangle \quad \forall \varphi \in D(\Omega)$$

Remarque 1.5.1. De manière analogue on peut définir les dérivées successives de T . Pour chaque multi-indice $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ d'ordre $|\alpha| = \alpha_1 + \dots + \alpha_n$, on définit :

$$\langle D^\alpha T, \varphi \rangle = (-1)^{|\alpha|} \langle T, D^\alpha \varphi \rangle \quad \forall \varphi \in D(\Omega)$$

1.6 Espaces de Sobolev

Définition 1.6.1 (Espace $H^1(\Omega)$). Soit Ω un ouvert de \mathbb{R}^n . On pose

$$H^1(\Omega) := \{u \in L^2(\Omega) : \partial u / \partial x_i \in L^2(\Omega), \forall i = 1, \dots, n\},$$

Bien entendu, la dérivation est à comprendre au sens des distributions. En d'autres termes, une fonction $u \in L^2(\Omega)$ est dans $H^1(\Omega)$ s'il existe des fonctions v_1, \dots, v_n dans $L^2(\Omega)$ telles que :

$$\int_{\Omega} u \frac{\partial \varphi}{\partial x_i} dx = - \int_{\Omega} v_i \varphi dx, \quad \forall \varphi \in D(\Omega), \quad \forall i = 1, \dots, n.$$

On sait, d'après la section de distributions, que les fonctions v_i sont alors notées $\partial u / \partial x_i$.

On associe à $H^1(\Omega)$ le produit scalaire :

$$(u, v)_{H^1(\Omega)} = \int_{\Omega} \left(uv + \sum_{i=1}^n \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} \right) dx$$

et la norme :

$$\|v\|_{H^1(\Omega)} = (v, v)_{H^1}^{\frac{1}{2}} = \left(\int_{\Omega} \left(|v|^2 + \sum_{i=1}^n \left(\frac{\partial v}{\partial x_i} \right)^2 \right) dx \right)^{\frac{1}{2}}$$

Proposition 1.6.1. [4] *L'espace $H^1(\Omega)$ muni de la norme $\|\cdot\|_{H^1}$ est un espace de Hilbert.*

Définition 1.6.2 (Espace $H^m(\Omega)$). *On définit les espaces de Sobolev $H^m(\Omega)$ où m est un entier strictement positif par :*

$$H^m(\Omega) := \{u \in L^2(\Omega) : D^\alpha u \in L^2(\Omega), \quad \alpha \in \mathbb{N}^n, \quad |\alpha| \leq m\}.$$

On le munit de la norme :

$$\|u\|_{H^m} := \left(\sum_{\alpha: |\alpha| \leq m} \int_{\Omega} |D^\alpha u|^2 dx \right)^{\frac{1}{2}}.$$

Remarque 1.6.1. *De façon plus générale, pour tout $1 \leq p \leq \infty$ et pour tout $m \in \mathbb{N}$, $1 \leq m$, on peut définir les espaces de Sobolev :*

$$W^{m,p}(\Omega) := \{u \in L^p(\Omega) : D^\alpha u \in L^p(\Omega), \quad \forall \alpha \in \mathbb{N}^n, \quad |\alpha| \leq m\}.$$

que l'on munit de la norme :

$$\|u\|_{W^{1,p}} := \left(\sum_{\alpha: |\alpha| \leq m} \int_{\Omega} |D^\alpha u|^p dx \right)^{1/p}$$

Définition 1.6.3 (Espace $H_0^1(\Omega)$). *L'espace $H_0^1(\Omega)$ est très utile pour les problèmes avec conditions aux limites. L'espace de Sobolev $H_0^1(\Omega)$ est défini comme étant l'adhérence de $D(\Omega)$ dans $H^1(\Omega)$.*

En général, on note $H_0^m(\Omega)$ l'adhérence de $D(\Omega)$ dans $H^m(\Omega)$ pour la norme de $H^m(\Omega)$.

Lorsque $\Omega \neq \mathbb{R}^n$, les fonctions de $H_0^m(\Omega)$ sont, dans un certain sens, les fonctions qui s'annulent sur le bord de Ω .

Définition 1.6.4. *Nous dirons qu'un ouvert Ω est borné dans une direction s'il existe un vecteur unitaire $d \in \mathbb{R}^n$, et une constante $A > 0$ tels que :*

$$\Omega \subset \{x \in \mathbb{R}^n : -A \leq x \cdot d \leq A\}.$$

Théorème 1.6.1 (Inégalité de Poincaré [4]). *Soit Ω un ouvert borné dans une direction. Alors il existe une constante $C > 0$ ne dépendant que de Ω telle que :*

$$\forall u \in H_0^1(\Omega), \quad \|u\|_{L^2(\Omega)} \leq C \|\nabla u\|_{L^2(\Omega)}.$$

1.7 Formulation variationnelle

On considère le problème elliptique avec conditions aux limites de Dirichlet (1.3) avec l'opérateur elliptique L donné par (1.4). Notre but dans cette section est de définir et construire ce que nous allons appeler "une solution faible" de ce problème. On suppose que les coefficients $a_{i,j}, b_i, c \in L^\infty(\Omega)$, $i, j = 1, \dots, n$, et $f \in L^2(\Omega)$. Supposons pour le moment que u une solution suffisamment régulière. Multiplions l'EDP (1.3) par une fonction test $v \in D(\Omega)$ et intégrons sur Ω , pour trouver :

$$\int_{\Omega} \left(\sum_{i,j=1}^n a_{i,j} \partial_{x_i} u \partial_{x_j} v + \sum_{i=1}^n b_i \partial_{x_i} uv + cuv \right) dx = \int_{\Omega} f v dx.$$

Nous avons intégré par parties le premier terme du membre gauche de l'équation. Il n'y a pas de termes d'intégration sur le bord puisque $v = 0$ sur $\partial\Omega$. Le même résultat est trouvé si on prend une fonction $v \in H_0^1(\Omega)$, et cela aura un sens si la fonction u est dans $H_0^1(\Omega)$. On choisit l'espace $H_0^1(\Omega)$ pour incorporer la condition aux limites de l'équation (1.3) : $u = 0$ sur $\partial\Omega$.

Définition 1.7.1. *La forme bilinéaire $a(.,.)$ associée à l'opérateur elliptique L est définie par la formule :*

$$a(u, v) = \int_{\Omega} \left(\sum_{i,j=1}^n a_{i,j} \partial_{x_i} u \partial_{x_j} v + \sum_{i=1}^n b_i \partial_{x_i} uv + cuv \right) dx, \quad \forall u, v \in H_0^1(\Omega).$$

On dit que $u \in H_0^1(\Omega)$ est une solution faible du problème (1.3) si :

$$a(u, v) = (f, v) \tag{1.10}$$

pour tout $v \in H_0^1(\Omega)$, où $(.,.)$ définit le produit scalaire dans $L^2(\Omega)$.

L'identité (1.10) est nommée parfois la formulation variationnelle du problème original (1.3). Elle forme la base de l'approximation numérique par la méthode des éléments finis que nous allons voir dans le deuxième chapitre.

1.7.1 Problème de Dirichlet homogène

On appelle problème de Dirichlet une équation de Laplace avec des conditions aux limites de type Dirichlet c'est à dire on spécifie les valeurs de la

solution sur la frontière. Pour tout Ω ouvert de \mathbb{R}^n et de frontière $\Gamma := \partial\Omega$ le problème de Dirichlet s'énonce de la façon suivante : étant donné une fonction $f \in L^2(\Omega)$, trouver une fonction u définie dans Ω et solution de :

$$\begin{cases} -\Delta u = f & \text{dans } \Omega, \\ u = 0 & \text{sur } \Gamma. \end{cases} \quad (1.11)$$

Supposons qu'il existe une solution $u \in H^1(\Omega)$ de (1.11) (une solution assez régulière). Soit $v \in H_0^1(\Omega)$; en multipliant la première équation de (1.11) par la fonction test v et en intégrant sur (Ω) ; on obtient :

$$\int_{\Omega} -(\Delta u)v dx = \int_{\Omega} f v dx,$$

Et en utilisant la formule de Green, on a :

$$\int_{\Omega} \nabla u \nabla v dx - \int_{\Gamma} \frac{\partial u}{\partial n} v d\sigma = \int_{\Omega} f v dx,$$

La notation $\frac{\partial u}{\partial n} = \nabla u \cdot n$ est la dérivée normale de u et n est la normale extérieure du domaine Ω . Comme $v = 0$ sur le bord, on obtient ce que l'on appelle la formulation variationnelle du problème (1.11) :

$$\begin{cases} \text{Trouvez } u \in H_0^1(\Omega), \text{ tel que :} \\ \int_{\Omega} \nabla u \nabla v dx = \int_{\Omega} f v dx, \quad \forall v \in H_0^1(\Omega). \end{cases} \quad (1.12)$$

Réciproquement, si $u \in H_0^1(\Omega)$ est solution de (1.12), alors puisque $D(\Omega) \subset H_0^1(\Omega)$; on a pour tout $\varphi \in D(\Omega)$:

$$\int_{\Omega} \nabla u \cdot \nabla \varphi dx = \sum_{i=1}^n \int_{\Omega} \frac{\partial u}{\partial x_i} \frac{\partial \varphi}{\partial x_i} dx = \int_{\Omega} f \varphi dx,$$

cette égalité s'écrit encore :

$$\sum_{i=1}^n \left\langle \frac{\partial u}{\partial x_i}, \frac{\partial \varphi}{\partial x_i} \right\rangle = \langle f, \varphi \rangle,$$

et par définition de la dérivation au sens des distributions il vient :

$$-\sum_{i=1}^n \left\langle \frac{\partial^2 u}{\partial x_i^2}, \varphi \right\rangle = \langle f, \varphi \rangle,$$

c'est à dire :

$$-\Delta u = f$$

La formulation variationnelle est donc équivalente pour les solutions régulières, à la formulation classique du problème (1.11) de Dirichlet.

1.7.2 Problème de Dirichlet non homogène

Soit Ω ouvert de \mathbb{R}^n et de frontière $\Gamma := \partial\Omega$ on considère le problème suivant : étant donné une fonction $f \in L^2(\Omega)$, trouver une fonction u définie dans Ω et solution de :

$$\begin{cases} -\Delta u = f & \text{dans } \Omega, \\ u = d & \text{sur } \Gamma. \end{cases} \quad (1.13)$$

Pour se ramener au problème de Dirichlet homogène, on va construire un relèvement, c'est-à-dire une fonction $u_0 \in H^1(\Omega)$ telle que $\gamma_0(u_0) = d$ où γ_0 est l'application trace.

Si $d \in \text{Im}(\gamma_0) = \gamma_0(H^1(\Omega))$ qui est un sous-espace de $L^2(\Gamma)$, on sait qu'il existe $u_0 \in H^1(\Omega)$ telle que $d = \gamma_0(u_0)$, On cherche donc u sous la forme $u = \tilde{u} + u_0$ avec $\tilde{u} \in H_0^1(\Omega)$ et $u_0 \in H^1(\Omega)$ telle que $\gamma_0(u_0) = d$, soit $v \in H_0^1(\Omega)$, on multiplie la première équation de (1.13) par v et on intègre sur Ω on obtient :

$$\int_{\Omega} -(\Delta u)v dx = \int_{\Omega} f v dx,$$

Comme $u = \tilde{u} + u_0$; on trouve :

$$-\int_{\Omega} \nabla \tilde{u} \cdot \nabla v dx - \int_{\Omega} \nabla u_0 \cdot \nabla v dx = \int_{\Omega} f v dx,$$

Et en utilisant la formule de Green, on a :

$$\int_{\Omega} \nabla \tilde{u} \nabla v dx - \int_{\Gamma} \frac{\partial \tilde{u}}{\partial n} v d\sigma + \int_{\Omega} \nabla u_0 \nabla v dx - \int_{\Gamma} \frac{\partial u_0}{\partial n} v d\sigma = \int_{\Omega} f v dx,$$

Et puisque $\tilde{u} \in H_0^1(\Omega)$, alors $\tilde{u} = 0$ sur Γ , et on se ramène donc au problème suivant :

$$\begin{cases} \text{Trouvez } \tilde{u} \in H_0^1(\Omega), \text{ tel que :} \\ \int_{\Omega} \nabla \tilde{u} \nabla v dx = \int_{\Omega} f v dx - \int_{\Omega} \nabla u_0 \nabla v dx, \quad \forall v \in H_0^1(\Omega). \end{cases} \quad (1.14)$$

C'est la formulation variationnelle du problème (1.13) de Dirichlet non homogène.

1.7.3 Problème de Neumann

Toujours avec les mêmes hypothèses sur Ω ; on cherche maintenant une fonction u telle que :

$$\begin{cases} -\Delta u + u = f & \text{dans } \Omega, \\ \frac{\partial u}{\partial n} = 0 & \text{sur } \Gamma. \end{cases} \quad (1.15)$$

Supposons qu'il existe une solution $u \in H^2(\Omega)$ de (1.15) (une solution assez régulière). Soit $v \in H^1(\Omega)$; en multipliant la première équation de (1.15) par la fonction test v et en intégrant sur Ω , on obtient :

$$\int_{\Omega} -(\Delta u)v dx + \int_{\Omega} uv dx = \int_{\Omega} f v dx,$$

Et en utilisant la formule de Green, on a :

$$\int_{\Omega} \nabla u \nabla v dx - \int_{\Gamma} \frac{\partial u}{\partial n} v d\sigma + \int_{\Omega} uv dx = \int_{\Omega} f v dx,$$

comme $\frac{\partial u}{\partial n} = 0$ sur le bord, on obtient ce que l'on appelle la formulation variationnelle du problème (1.15) :

$$\begin{cases} \text{Trouvez } u \in H_0^1(\Omega), \text{ tel que :} \\ \int_{\Omega} \nabla u \nabla v dx + \int_{\Omega} uv dx = \int_{\Omega} f v dx, \quad \forall v \in H^1(\Omega). \end{cases} \quad (1.16)$$

Réciproquement, si $u \in H^2(\Omega)$ est solution de (1.16), alors en particulier on a en prenant $v \in D(\Omega)$ dans (1.16) ; et on utilisant la formule de Green :

$$\int_{\Omega} -(\nabla u)v \, dx + \int_{\Omega} uv \, dx - \int_{\Gamma} \frac{\partial u}{\partial n} v d\sigma = \int_{\Omega} fv \, dx$$

D'après l'égalité obtenue précédemment, $\int_{\Gamma} \frac{\partial u}{\partial n} v = 0$ pour tout $v \in H^1$; on retrouve alors la formulation variationnelle.

1.8 Théorème de Lax-Milgram

Le théorème de Lax-Milgram est l'outil essentiel permettant de démontrer des résultats d'existence et d'unicité de solutions de la formulation variationnelle dans un espace de Hilbert.

On note V un espace de Hilbert réel de produit scalaire $\langle \cdot, \cdot \rangle$ et de norme $\|\cdot\|$. Nous considérons une formulation variationnelle du type :

$$\left\{ \begin{array}{l} \text{Trouvez } u \in V, \text{ tel que :} \\ a(u, v) = \langle f, v \rangle, \quad \forall v \in V. \end{array} \right. \quad (1.17)$$

avec $f : V \rightarrow \mathbb{R}$, tel que :

1. $a(\cdot, \cdot)$ est une forme bilinéaire sur V , c'est-à-dire que $u \rightarrow a(u, v)$ est une forme linéaire de V dans \mathbb{R} pour tout $v \in V$, et $v \rightarrow a(u, v)$ est une forme linéaire de V dans \mathbb{R} pour tout $u \in V$.
2. $a(\cdot, \cdot)$ est continue sur $V \times V$, c'est-à-dire, qu'il existe une constante $M > 0$ telle que :

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

3. $a(\cdot, \cdot)$ est coercive, c'est-à-dire, qu'il existe une constante strictement positive α telle que :

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

4. La forme linéaire $l(v) = \langle f, v \rangle$ est continue, c'est-à-dire, il existe une constante $M > 0$ telle que :

$$|l(v)| \leq M \|v\|_V, \quad \forall v \in V.$$

Théorème 1.8.1 (Théorème de Lax-Milgram [1]). *Soit V un espace de Hilbert, et soit $a(.,.)$ une forme bilinéaire sur $V \times V$. Si $a(.,.)$ est continue et coercive, alors il existe une unique solution $u \in V$ du problème variationnel (1.17).*

Théorème 1.8.2. [1] *Sous les mêmes hypothèses du théorème (1.8.1) et si de plus la forme bilinéaire $a(.,.)$ est symétrique, c'est-à-dire :*

$$a(u, v) = a(v, u), \quad \forall u, v \in V$$

nous avons :

Le problème variationnel (1.17) est équivalent au Problème de minimisation suivant :

$$\left\{ \begin{array}{l} \text{Trouvez une fonction } u \in V, \text{ tel que :} \\ J(u) = \min_{v \in V} J(v) = \min_{v \in V} \left(\frac{1}{2} a(v, v) - \langle f, v \rangle \right), \quad \forall v \in V. \end{array} \right. \quad (1.18)$$

Dans les deux prochaines sections, on va se restreindre à l'application du théorème de Lax-Milgram aux cas de problèmes elliptiques et paraboliques puisque ce sont les modèles qui nous intéresseront le plus dans les chapitres qui suivent.

1.8.1 Application aux problèmes elliptiques

- **Problème de Dirichlet**

On rappelle qu'il s'agit de trouver $u \in H_0^1(\Omega)$ tel que :

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

pour :

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v \, dx \quad \text{et} \quad l(v) = \int_{\Omega} f v \, dx,$$

Ceci sur l'espace $V = H_0^1(\Omega)$ muni de la norme :

$$\|v\|_V = |v|_{H_0^1(\Omega)} = \|\nabla v\|_{L^2(\Omega)}$$

La forme bilinéaire $a(., .)$ est continue sur $H_0^1(\Omega) \times H_0^1(\Omega)$, en effet, à l'aide de l'inégalité de Cauchy-Schwarz, on a :

$$\begin{aligned} |a(u, v)| &= \left| \int_{\Omega} \nabla u \cdot \nabla v \, dx \right| \\ &\leq \int_{\Omega} |\nabla u \cdot \nabla v| \, dx \\ &\leq \left(\int_{\Omega} |\nabla u|^2 \, dx \right)^{1/2} \left(\int_{\Omega} |\nabla v|^2 \, dx \right)^{1/2} \\ &= \|\nabla u\|_{L^2(\Omega)} \|\nabla v\|_{L^2(\Omega)} \\ &= \|u\|_V \|v\|_V \end{aligned}$$

D'où :

$$|a(u, v)| \leq \|u\|_V \|v\|_V,$$

- D'autre part, $a(., .)$ est coercive, en effet :

$$\begin{aligned} a(v, v) &= \int_{\Omega} (\nabla v)^2 \, dx \\ &= \int_{\Omega} |\nabla v|^2 \, dx \\ &= \|\nabla v\|_{L^2(\Omega)}^2 \\ &= \|v\|_V^2 \end{aligned}$$

D'où :

$$|a(v, v)| = \|v\|_V^2,$$

Et enfin, $l(.)$ est continue sur $H_0^1(\Omega)$, en effet à l'aide des inégalités de Cauchy-Schwarz et de Poincaré, on a :

$$\begin{aligned} |L(v)| &= \left| \int_{\Omega} f v \, dx \right| \\ |L(v)| &\leq \left(\int_{\Omega} |f|^2 \, dx \right)^{1/2} \cdot \left(\int_{\Omega} |v|^2 \, dx \right)^{1/2} \\ &= \|f\|_{L^2(\Omega)} \cdot \|v\|_{L^2(\Omega)} \\ &\leq C(\Omega) \|f\|_{L^2(\Omega)} \cdot \|\nabla v\|_{L^2(\Omega)} \end{aligned}$$

On pose : $M = C(\Omega) \|f\|_{L^2(\Omega)}$. D'où :

$$|l(v)| \leq M \|v\|_V.$$

Le théorème de Lax-Milgram nous assure donc l'existence et l'unicité de la solution $u \in H_0^1(\Omega)$ du problème de Dirichlet.

De plus, puisque la forme bilinéaire $a(.,.)$ est symétrique, le théorème nous assure que cette fonction u minimise la fonctionnelle quadratique suivante :

$$J(v) = \frac{1}{2} \int_{\Omega} (\nabla v)^2 dx - \int_{\Omega} f v dx, \quad \text{sur l'espace } H_0^1(\Omega).$$

- **Problème de Neumann**

On rappelle qu'il s'agit de trouver $u \in H^1(\Omega)$ tel que :

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

On applique donc le théorème de Lax-Milgram, avec :

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v dx + \int_{\Omega} uv dx \quad \text{et} \quad l(v) = \int_{\Omega} f v dx,$$

Ceci sur l'espace $V = H^1(\Omega)$ muni de la norme :

$$\|v\|_V = \|v\|_{H^1(\Omega)}$$

La forme bilinéaire $a(.,.)$ est continue sur $H^1(\Omega) \times H^1(\Omega)$, en effet, à l'aide de l'inégalité de Cauchy-Schwarz, on a :

$$a(u, v) = \int_{\Omega} \nabla u \cdot \nabla v dx + \int_{\Omega} uv dx \leq \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}$$

D'où :

$$|a(u, v)| \leq \|u\|_V \|v\|_V,$$

- D'autre part, $a(.,.)$ est coercive, en effet :

$$\begin{aligned} a(v, v) &= \int_{\Omega} |\nabla v|^2 dx + \int_{\Omega} |v|^2 dx \\ &\geq \|v\|_V^2 \end{aligned}$$

D'où :

$$a(v, v) \geq \|v\|_{H^1(\Omega)}^2,$$

Et enfin, $l(\cdot)$ est continue sur $H^1(\Omega)$, en effet à l'aide des inégalités de Cauchy-Schwarz et de Poincaré, on a :

$$\begin{aligned} |l(v)| &= \left| \int_{\Omega} f v \, dx \right| \\ |l(v)| &\leq \left(\int_{\Omega} |f|^2 \, dx \right)^{1/2} \cdot \left(\int_{\Omega} |v|^2 \, dx \right)^{1/2} \\ &= \|f\|_{L^2(\Omega)} \cdot \|v\|_{L^2(\Omega)} \\ &\leq C(\Omega) \|f\|_{L^2(\Omega)} \cdot \|v\|_{L^2(\Omega)} \end{aligned}$$

On pose :

$$M = C(\Omega) \|f\|_{L^2(\Omega)}$$

D'où :

$$|l(v)| \leq M \|v\|_V$$

Le théorème de Lax-Milgram nous assure donc l'existence et l'unicité de la solution $u \in H^1(\Omega)$ du problème de Neumann.

De plus, puisque la forme bilinéaire $a(\cdot, \cdot)$ est symétrique, le théorème nous assure que cette fonction u minimise la fonctionnelle quadratique suivante :

$$J(v) = \frac{1}{2} \int_{\Omega} |\nabla v|^2 \, dx - \int_{\Omega} f v \, dx, \quad \text{sur l'espace } H^1(\Omega).$$

1.8.2 Application aux problèmes paraboliques

Dans cette section un résultat d'existence et d'unicité de solutions des EDP paraboliques est introduit. Soit Ω un ouvert borné de \mathbb{R}^n , de frontière notée $\partial\Omega$, $T > 0$ un réel fixé, et Q un domaine espace-temps défini par $Q = \Omega \times [0, T]$. On note par $\partial Q = \partial\Omega \times [0, T]$ sa frontière. Soient $V \subset H$ deux espaces de Hilbert séparables tels que l'injection de V dans H soit continue et que V soit dense dans H . On identifie H à son dual H' , et on note V' le dual de V , alors H peut être identifié à un sous-espace de V' :

$V \subset H \subset V'$. Si $f \in V'$ et $v \in V$, $\langle f, v \rangle$ désigne leur produit de dualité. Pour $t \in [0, T]$, soit $a(t; u, v) : V \times V \rightarrow \mathbb{R}$ une forme bilinéaire définie par :

$$a(t; u, v) = \langle L(t)u, v \rangle; \quad L(t)u \in V'; \quad L(t) \in \mathcal{L}(V, V')$$

et vérifie les propriétés suivantes :

- (p1) $\forall u, v \in V$, la fonction $t \mapsto a(t; u, v)$ est mesurable,
- (p2) (coercivité) il existe des constantes $\alpha > 0$ et $\sigma \in \mathbb{R}$ telles que :

$$\alpha \|v\|_V^2 \leq a(t; v, v) + \sigma \|v\|_H^2, \quad \forall v \in V, \quad \forall t \in [0, T].$$

- (p3) (continuité) il existe une constante $\beta > 0$ telle que :

$$|a(t; u, v)| \leq \beta \|u\|_V \|v\|_V, \quad \forall t \in [0, T], \quad \forall u, v \in V,$$

Nous considérons le problème faible suivant :

$$\left\{ \begin{array}{l} \text{Trouver } u \in W(V) \text{ tel que} \\ \langle \partial_t u, v \rangle + a(t; u, v) = \langle f(t), v \rangle \quad p.p. \ t \in [0, T], \quad \forall v \in V, \\ u(0) = u_0, \end{array} \right. \quad (1.19)$$

où $u_0 \in H$, et $W(V) = \{v : [0, T] \rightarrow V; v \in L^2([0, T]; V); \partial_t v \in L^2([0, T]; V')\}$.

Nous avons le résultat suivant qui représente l'équivalent du Théorème de Lax-Milgram pour EDP paraboliques :

Théorème 1.8.3. (J.L. Lions). *Si les conditions (p1)–(p3) sont vérifiées, alors le problème (1.19) admet une solution unique.*

Maintenant, le théorème 1.8.3 est utilisé pour monter l'existence et l'unicité d'une solution faible du problème parabolique (2.10) :

$$\left\{ \begin{array}{ll} \partial_t u + Lu = f & \text{dans } \Omega_T, \\ u = 0 & \text{sur } \partial\Omega \times [0, T], \\ u = u_0 & \text{sur } \Omega \times \{t = 0\}, \end{array} \right. \quad (1.20)$$

avec $u_0 \in L^2(\Omega, \mathbb{R})$, $f \in L^2(Q, \mathbb{R})$ et $a_{i,j}, b_i, c \in L^\infty(Q, \mathbb{R})$.

Prenons $V = H_0^1(\Omega, \mathbb{R})$, et $H = L^2(\Omega, \mathbb{R})$. Pour chaque $t \in [0, T]$, L est l'opérateur différentiel donné dans (1.21) :

$$Lu = - \sum_{i,j=1}^n a_{i,j}(x,t) \partial_{i,j}^2 u(x,t) + \sum_{i=1}^n b_i(x,t) \partial_i u(x,t) + c(x,t)u(x,t) \quad (1.21)$$

Afin de formuler le problème (2.10) dans un sens faible, nous nous donnons une fonction test v dans l'espace V muni du produit scalaire :

$$(u, v)_V = \int_{\Omega} (\nabla u \cdot \nabla \bar{v} + u \bar{v}) d\mathbf{x}$$

et de la norme induite :

$$\|u\|_V^2 = \int_{\Omega} (|\nabla u|^2 + |u|^2) d\mathbf{x}.$$

En multipliant (2.10) par une fonction test $v \in H_0^1(\Omega, \mathbb{R})$, en intégrant par parties sur Ω , et en utilisant la formule de Green, on obtient :

$$\langle \partial_t u(t), v \rangle + a(t; u, v) = 0, \quad \forall t \in [0, T],$$

où la forme bilinéaire a est donnée par :

$$a(t; u, v) = \int_{\Omega} (\mathbf{A}(x,t) \nabla u \cdot \nabla v + b(x,t) \cdot (\nabla u)v + c(x,t)uv) dx, \quad \forall t \in [0, T].$$

Une formulation faible du problème (2.10) est obtenue :

$$\begin{cases} \text{Trouver } u \in W(V) \text{ tel que} \\ \langle \partial_t u(t), v \rangle + a(t; u, v) = 0, \quad \forall v \in V, t \in [0, T], \\ u(x, 0) = u_0. \end{cases} \quad (1.22)$$

Théorème 1.8.4. [5] *Si l'opérateur L donné par l'expression (1.4) est elliptique et que les coefficients $b_i, i = 1, \dots, n$, et c sont dans $L^\infty(Q, \mathbb{R})$, et a_{ij} est dans $L^\infty(Q, \mathbb{R})$, alors le problème (1.22) admet une solution unique.*

Preuve. Il est clair que $\forall u, v \in V$, la fonction $t \mapsto a(t; u, v)$ est mesurable. En utilisant l'ellipticité de L , nous avons :

$$\begin{aligned} \delta \int_{\Omega} |\nabla v|^2 dx &\leq a(t; v, v) - \int_{\Omega} (b \cdot \nabla v v + c|v|^2) dx \\ &\leq a(t; v, v) + \|b\|_{L^\infty(Q)} \int_{\Omega} |\nabla v| |v| dx + \|c\|_{L^\infty(Q)} \int_{\Omega} |v|^2 dx \end{aligned} \quad (1.23)$$

En utilisant l'inégalité de Young avec $\epsilon > 0$ ¹, on observe :

$$\int_{\Omega} |\nabla v| |v| dx \leq \epsilon \int_{\Omega} |\nabla v|^2 dx + \frac{1}{4\epsilon} \int_{\Omega} |v|^2 dx.$$

On insère cette estimation dans (1.23) :

$$\delta \int_{\Omega} |\nabla v|^2 dx \leq a(t; v, v) + \|b\|_{L^\infty} \int_{\Omega} (\epsilon |\nabla v|^2 + \frac{1}{4\epsilon} |v|^2) dx + \|c\|_{L^\infty} \int_{\Omega} |v|^2 dx.$$

Choisissons un $\epsilon > 0$ vérifiant $\delta - \epsilon \|b\|_{L^\infty} > 0$. Alors :

$$(\delta - \epsilon \|b\|_{L^\infty}) \int_{\Omega} |\nabla v|^2 dx \leq a(t; v, v) + (\frac{1}{4\epsilon} \|b\|_{L^\infty} + \|c\|_{L^\infty}) \int_{\Omega} |v|^2 dx.$$

Il vient :

$$(\delta - \epsilon \|b\|_{L^\infty}) \|v\|_V^2 \leq a(t; v, v) + (\frac{1}{4\epsilon} \|b\|_{L^\infty} + \|c\|_{L^\infty} + \delta - \epsilon \|b\|_{L^\infty}) \|v\|_H^2.$$

Donc, il existe des constantes $\sigma = \frac{1}{4\epsilon} \|b\|_{L^\infty} + \|c\|_{L^\infty} + \delta - \epsilon \|b\|_{L^\infty}$,
et $\alpha = \delta - \epsilon \|b\|_{L^\infty}$ telles que la forme a soit coercive.

Montrons maintenant que a est continue :

$$|a(t; u, v)| \leq \int_{\Omega} (|A \nabla u| |\nabla v| + |b| |\nabla u| |v| + |c| |u| |v|) dx,$$

on pose $C_1 = \max\{\|A\|_2, \|c\|_{L^\infty(Q)}\} < \infty$, et $C_2 = \|b\|_{L^\infty(Q)} < \infty$.

En vertu de $|A \nabla u| \leq \|A\|_2 |\nabla u|$, nous avons :

$$|a(t; M, v)| \leq C_1 \int_{\Omega} (|\nabla u| |\nabla v| + |u| |v|) dx + C_2 \int_{\Omega} |\nabla u| |v| dx.$$

Par l'inégalité de Cauchy-Schwarz, on obtient :

$$\begin{aligned} |a(t; u, v)| &\leq C_1 \int_{\Omega} (|\nabla u|^2 + |u|^2)^{1/2} (|\nabla v|^2 + |v|^2)^{1/2} dx \\ &\quad + C_2 \left(\int_{\Omega} |\nabla u|^2 dx \right)^{1/2} \left(\int_{\Omega} |v|^2 dx \right)^{1/2} \\ &\leq C_1 \left(\int_{\Omega} |u|^2 + |\nabla u|^2 dx \right)^{1/2} \left(\int_{\Omega} |v|^2 + |\nabla v|^2 dx \right)^{1/2} + C_2 \|u\|_V \|v\|_H \\ &\leq (C_1 + C_2) \|u\|_V \|v\|_V, \quad t \in [0, T]. \end{aligned}$$

Donc il existe une constante $\beta = C_1 + C_2$ telle que la forme a soit continue.

Alors par le théorème 1.8.3, nous déduisons que le problème (1.22) admet une solution unique. \square

1. Inégalité de Young est donnée par : $\forall a, b \geq 0, \forall \epsilon > 0$, on a : $ab \leq \frac{a^2}{2\epsilon} + \frac{\epsilon b^2}{2}$.

Chapitre 2

Méthode des éléments finis

2.1 Introduction

Dans ce chapitre nous présentons la méthode des éléments finis ([10], [11]) qui est la méthode numérique de référence pour le calcul des solutions de problèmes aux limites. Le principe de cette méthode est directement issu de l'approche variationnelle que nous avons vu dans le chapitre précédent. L'idée de base est de remplacer l'espace de Hilbert V , sur lequel est posée la formulation variationnelle, par un sous-espace V_h de dimension finie. Le problème approché posé sur V_h se ramène à la simple résolution d'un système linéaire, qui fait intervenir des matrices appelées "matrices de rigidité et de masse". Par ailleurs, on peut choisir le mode de construction de V_h de manière à ce que le sous-espace V_h soit une bonne approximation de V ; et que la solution u_h dans V_h de la formulation variationnelle, soit proche de la solution exacte u dans V .

2.2 Bref historique

L'idée fondamentale derrière la méthode des éléments finis remonte loin en arrière. Les grecs par exemple avaient reconnu que l'on peut approcher la solution d'un problème complexe en le divisant en problèmes plus simples. On peut par exemple approcher le périmètre d'un cercle en calculant le périmètre d'un polygone à n côtés, pourvu que n soit suffisamment grand. Il

suffit alors de connaître la longueur d'un segment de droite, problème beaucoup plus simple que celui de la longueur d'un arc de cercle.

L'application à la solution des équations aux dérivées partielles est évidemment plus récente et est intimement liée au développement de l'informatique. Courant [6] a introduit le concept de formulation variationnelle, qui est à la base de toute méthode d'éléments finis.

Pour la méthode de Ritz [7], on part d'un problème posé dans un espace de dimension infinie. On approche ensuite la solution du problème initial en cherchant une solution dans une suite croissante de sous-espaces de dimension finie. Ces problèmes approchés sont en général beaucoup plus facile à résoudre. On peut de plus espérer que la solution du problème en dimension infinie peut être obtenue par un passage à la limite. Le choix des fonctions de base constituant ces espaces de dimension finie est délicat et initialement on les construisait globalement sur le domaine. C'est Courant [6] qui eut l'idée d'introduire des fonctions à support local qui simplifient grandement leur construction.

La théorie derrière la méthode des éléments finis a pris une forme plus rigoureuse avec les travaux de Strang et Fix [8]. On retrouve les premières applications véritables de la méthode des éléments finis en 1956 en mécanique des structures. Un groupe de chercheurs (Turner, Clough, Martin et Topp [9]) de Boeing utilisent cette méthode pour calculer la voilure d'un avion. La méthode des éléments finis est maintenant reconnue comme l'une des principales méthodes de résolution des équations aux dérivées partielles (EDP) dans des géométries quelconques, que ce soit en dimension un, deux ou trois. On trouve même des méthodes d'éléments finis en dimension quatre, soit en espace-temps.

Les applications sont tout aussi nombreuses et variées. Les ingénieurs de diverses disciplines utilisent les éléments finis, que ce soit en mécanique des fluides ou des solides, mais aussi pour les problèmes thermiques, électromagnétiques, chimiques, etc. On retrouve aussi des applications en physique,

astrophysique, etc.

2.3 Approximation interne générale

Nous considérons à nouveau le cadre général de la formulation variationnelle des problèmes elliptiques introduite dans le Chapitre 1. Étant donné un espace de Hilbert V , une forme bilinéaire continue et coercive $a(u, v)$; et une forme linéaire continue $l(v)$; on considère la formulation variationnelle suivante :

$$\left\{ \begin{array}{l} \text{Trouvez } u \in V, \text{ tel que :} \\ a(u, v) = l(v), \quad \forall v \in V. \end{array} \right. \quad (2.1)$$

dont on sait qu'elle admet une unique solution par le Théorème de Lax-Milgram. L'approximation interne de (2.1) consiste à remplacer l'espace de Hilbert V par un sous-espace de dimension finie V_h ; c'est-à-dire à chercher la solution de :

$$\left\{ \begin{array}{l} \text{Trouvez } u_h \in V_h, \text{ tel que :} \\ a(u_h, v_h) = l(v_h), \quad \forall v_h \in V_h. \end{array} \right. \quad (2.2)$$

La résolution de l'approximation interne (2.2) est facile comme le montre le théorème suivant.

Théorème 2.3.1. *Soit V un espace de Hilbert réel, et V_h un sous-espace de dimension finie. Soit $a(u, v)$ une forme bilinéaire continue et coercive sur V ; et $l(v)$ une forme linéaire continue sur V . Alors l'approximation interne (2.2) admet une unique solution. Par ailleurs cette solution peut s'obtenir en résolvant un système linéaire de matrice définie positive et symétrique si $a(u, v)$ est symétrique.*

Preuve. L'existence et l'unicité de $u_h \in V_h$, solution de (2.2), découle du théorème de Lax-Milgram appliqué à V_h . Pour mettre le problème sous une forme plus simple, on introduit une base $(\varphi_j)_{1 \leq j \leq N_h}$ de V_h . Si $u_h = \sum_{j=1}^{N_h} u_j \varphi_j$ on pose $U = (u_1, \dots, u_{N_h})$ le vecteur dans \mathbb{R}^{N_h} des coordonnées de u_h . Le problème (2.2) est équivalent à :

$$\text{Trouver } U \in \mathbb{R}^{N_h} \quad \text{tel que} \quad a\left(\sum_{j=1}^{N_h} u_j \varphi_j, \varphi_i\right) = l(\varphi_i) \quad \forall i = 1, \dots, N_h$$

Ce qui s'écrit sous la forme d'un système linéaire :

$$AU = b, \quad (2.3)$$

pour $1 \leq i, j \leq N_h$, $(A)_{ij} = a(\varphi_j, \varphi_i)$, et $(b)_i = l(\varphi_i)$. Dans les applications mécaniques la matrice A est appelée *matrice de rigidité*. La proposition suivante permet de montrer que la matrice A est inversible.

Proposition 2.3.1. *La matrice A est définie positive c'est-à-dire qu'elle vérifie la propriété suivante :*

$$\forall U \in \mathbb{R}^{N_h} \setminus \{0\}, \quad \langle AU, U \rangle > 0.$$

Avant de donner la démonstration de cette proposition, expliquons pourquoi elle implique l'inversibilité de la matrice A . En effet, si A n'est pas inversible, alors il existe un vecteur $U \neq 0$ dans \mathbb{R}^{N_h} tel que $AU = 0$: Ceci implique en particulier que $\langle AU, U \rangle = 0$ avec $U \neq 0$; ce qui contredit le fait que A est définie positive.

Preuve. soit $U = (u_1, \dots, u_{N_h}) \neq 0$ dans \mathbb{R}^{N_h} . Le vecteur AU a pour coordonnées

$$\left(\sum_{j=1}^{N_h} A_{1j} u_j, \dots, \sum_{j=1}^{N_h} A_{N_h j} u_j \right) = \left(\sum_{j=1}^{N_h} u_j a(\varphi_j, \varphi_1), \dots, \sum_{j=1}^{N_h} u_j a(\varphi_j, \varphi_{N_h}) \right).$$

On a donc

$$\langle AU, U \rangle = \sum_{i=1}^{N_h} \sum_{j=1}^{N_h} u_i u_j a(\varphi_j, \varphi_i) = a \left(\sum_{j=1}^{N_h} u_j \varphi_j, \sum_{i=1}^{N_h} u_i \varphi_i \right) = a(u_h, u_h) = \int_0^1 |u'_h|^2 dx,$$

où $u_h(x)$ est une fonction de V_h définie par $u_h(x) = \sum_{j=1}^{N_h} u_j \varphi_j(x)$. Comme les coefficients (u_1, \dots, u_{N_h}) sont non nuls, on vérifie aisément que u_h est non nulle, ce qui implique que $\int_0^1 |u'_h|^2 dx > 0$.

Nous allons maintenant comparer l'erreur commise en remplaçant l'espace V par son sous-espace V_h : Plus précisément, nous allons majorer la différence $\|u - u_h\|$ où u est la solution dans V de (2.1) et u_h dans V_h de (2.2). Le lemme suivant, dû à Jean Céa, montre que la distance entre la solution exacte u et la solution approchée u_h est majorée uniformément par rapport au sous-espace V_h par la distance entre u et V_h .

Lemme 2.3.1. (*lemme de Céa*)

Soit u une solution du problème (2.1) et u_h une solution du problème (2.2). Alors il existe une constante $C > 0$ indépendante de h tel que :

$$\|u - u_h\|_V \leq \inf_{v_h \in V_h} \|u - v_h\|_V$$

Preuve. Soit $v_h \in V_h$ quelconque, et $w_h = v_h - u_h$, $V_h \subset V$. Comme w_h est dans V et v_h c'est une fonction test valide dans (2.1) et (2.3) On a donc :

$$a(u_h, w_h) = l(w_h),$$

Et

$$a(u, w_h) = l(w_h),$$

D'où :

$$a(u - u_h, w_h) = 0,$$

Par la coercivité et la continuité de $a(., .)$; on a :

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

Donc

$$\|u - u_h\| \leq \frac{M}{\alpha} \|u - v_h\|$$

et ceci pour tout v_h ; d'où le résultat avec $C = \frac{M}{\alpha}$.

2.4 Méthode de Galerkin

La méthode de Galerkin est une méthode très générale et très robuste. L'idée de la méthode est la suivante. Partant d'un problème posé dans un espace de dimension infinie, on procède d'abord à une approximation dans une suite croissante de sous-espaces de dimension finie. On résout ensuite

le problème approché, ce qui est en général plus facile que de le résoudre directement en dimension infinie. Enfin, on passe d'une façon ou d'une autre à la limite quand on fait tendre la dimension des espaces d'approximation vers l'infini pour construire une solution du problème de départ. Il convient de noter que, outre son intérêt théorique, la méthode de Galerkin fournit également un procédé constructif d'approximation.

2.4.1 Principe de base d'une méthode de Galerkin

Le principe d'une méthode de Galerkin est d'approcher les éléments de l'espace X par ceux d'un sous-espace X^h de dimension finie. C'est donc, à la base, la construction d'un processus d'approximation des éléments d'un espace de Hilbert X . L'exposant h est un paramètre réel > 0 caractérisant la discrétisation et tendant vers 0 au fur et à mesure que la discrétisation devient de plus en plus fine. Cette convention est utile pour les méthodes d'éléments finis où ce paramètre a un sens géométrique précis. La propriété d'approximation est décrite par la condition suivante dit d'approximation interne :

$$\forall v \in X, \quad \lim_{h \rightarrow 0} \inf_{v^h \in X^h} \|v - v^h\|_X = 0,$$

La quantité $\inf_{v^h \in X^h} \|v - v^h\|_X$ est la distance de v à X^h : Comme X^h est un espace de dimension finie, dès que X est un espace normé (même s'il n'est pas complet), cette distance est caractérisée de la façon suivante. Il existe $z^h \in X^h$ tel que

$$\inf_{v^h \in X^h} \|v - v^h\|_X = \|v - z^h\|_X,$$

Si X est un espace préhilbertien, i.e. si sa norme est associée à un produit scalaire, la meilleure approximation est unique : c'est la projection de v sur X^h : Elle est caractérisée par les équations dites normales

$$\begin{cases} z^h \in X^h, & \forall v^h \in X^h \\ (z^h, v^h)_X = (v, v^h)_X \end{cases}$$

Le problème variationnel fait intervenir aussi le sous-espace vectoriel V de X : Une façon simple, mais qui ne fonctionne pas toujours, est d'introduire le sous-espace de X^h

$$V^h := X^h \cap V,$$

Pour $v \in V$; on a cependant seulement :

$$\inf_{v^h \in X^h} \|v - v^h\|_X \leq \inf_{v^h \in V^h} \|v - v^h\|_X,$$

puisqu'on prend la borne inférieure sur un espace plus restreint. On est amené à faire l'hypothèse supplémentaire que la famille $\{V^h\}_{h>0}$ constitue une approximation interne de V

$$\lim_{h \rightarrow 0} \inf_{v^h \in V^h} \|v - v^h\|_V = 0.$$

2.5 Méthode des éléments finis

2.5.1 Méthode des éléments finis en dimension un

Nous choisissons le domaine $\Omega =]0, 1[$. En dimension 1 un maillage est simplement constitué d'une collection de points $(x_j)_{0 \leq j \leq N+1}$ tels que :

$$0 = x_0 \leq x_1 \leq x_2 \leq \dots \leq x_N \leq x_{N+1} = 1.$$

Le maillage sera dit uniforme si les points x_j sont équidistants, c'est-à-dire que :

$$x_j = jh, \quad \text{avec} \quad h = \frac{1}{N+1}, \quad 0 \leq j \leq N+1.$$

Les points x_j sont aussi appelés les sommets (ou nœuds) du maillage. Nous considérons le problème modèle suivant :

$$\begin{cases} -u''(x) = f(x), & \text{dans }]0, 1[\\ u(0) = u(1) = 0, \end{cases} \quad (2.4)$$

dont nous savons qu'il admet une solution unique dans $H_0^1(\Omega)$ si $f \in L^2(\Omega)$. Dans ce qui suit on notera \mathbb{P}_k l'ensemble des polynômes à coefficients réels d'une variable réelle de degré inférieur ou égal à k :

$$\mathbb{P}_k = \left\{ p(x) = \sum_{j=0}^k a_j x^j, a_j \in \mathbb{R} \right\}.$$

Remarque 2.5.1. Généralement, le sous-espace V_h de V est noté :

$$V_h^k = \left\{ v_h \in C^0(\Omega) \quad \text{tel que} \quad v_{h|_k} \in \mathbb{P}_k(k) \right\}.$$

Sa dimension est égale à :

$$\dim V_h^k = (k+1)(N+1) - N.$$

2.5.2 Élément fini \mathbb{P}_1

Pour $k=1$ et en notant $V_h^1 = V_h$ on a :

$$V_h = \{v \in C^0([0; 1]), \quad v_{[x_j, x_{j+1}]} \in \mathbb{P}_1, \quad 0 \leq j \leq N \quad v(0) = v(1) = 0\}. \quad (2.5)$$

La méthode des éléments finis \mathbb{P}_1 est alors simplement une méthode d'approximation variationnelle interne appliquée à l'espace V_h défini par (2.5).

On peut représenter les fonctions de V_h , affines par morceaux, à l'aide d'une fonctions de base très simple. Introduisons la "fonction chapeau" φ_j définie par :

$$\varphi_j(x) = \begin{cases} \frac{x-x_{j-1}}{x_j-x_{j-1}}, & \text{si } x \in [x_{j-1}, x_j] \\ \frac{x_{j+1}-x}{x_{j+1}-x_j}, & \text{si } x \in [x_j, x_{j+1}] \\ 0, & \text{sinon.} \end{cases}$$

Lemme 2.5.1. L'espace V_h défini par (2.5), est un sous-espace de $H^1(]0, 1[)$ de dimension $N+2$, et toute fonction $v_h \in V_h$ est définie de manière unique

par ses valeurs aux sommets $(x_j)_{0 \leq j \leq N+1}$:

$$v_h(x) = \sum_{j=0}^{N+1} v_h(x_j) \varphi_j(x) \quad \forall x \in [0, 1].$$

Preuve. Les fonctions continues et de classe C^1 par morceaux appartiennent $H^1(\Omega)$. Donc V_h est bien un sous-espace de $H^1(]0; , 1[)$. Le reste de la preuve est immédiat en remarquant que $\varphi(x_i) = \delta_{ij}$; où δ_{ij} est le symbole de Kronecker qui vaut 1 si $i = j$ et 0 sinon.

Corollaire 2.1. *Les fonctions $\{\varphi_j\}_{0 \leq i \leq N+1}$ constituent une base de V_h .*

Preuve. Puisque la famille $\{\varphi_j\}_{0 \leq i \leq N+1}$ est de dimension finie, et sa dimension est égale à la dimension de V_h ; on montre seulement que cette famille est libre. Soient donc α_j des scalaires dans \mathbb{R} ; avec $j \in \{0, \dots, N+1\}$ On suppose que :

$$\sum_{j=0}^{N+1} \alpha_j \varphi_j(x_i) = 0 \quad \text{avec } i \in \{0, \dots, N+1\}.$$

Pour $i = 0$; on a :

$$\alpha_0 \varphi_0(x_0) + \alpha_1 \varphi_1(x_0) + \dots + \alpha_{N+1} \varphi_{N+1}(x_0) = 0 \implies \alpha_0 \varphi_0(x_0) = 0$$

d'où :

$$\alpha_0 = 0$$

Pour $i = 1$; on a :

$$\alpha_0 \varphi_0(x_1) + \alpha_1 \varphi_1(x_1) + \dots + \alpha_{N+1} \varphi_{N+1}(x_1) = 0 \implies \alpha_1 \varphi_1(x_1) = 0$$

d'où :

$$\alpha_1 = 0$$

Pour $i = N+1$; on a :

$$\alpha_0 \varphi_0(x_{N+1}) + \alpha_1 \varphi_1(x_{N+1}) + \dots + \alpha_{N+1} \varphi_{N+1}(x_{N+1}) = 0 \implies \alpha_{N+1} \varphi_{N+1}(x_{N+1}) = 0$$

d'où :

$$\alpha_{N+1} = 0$$

Donc : $\forall j \in \{0, \dots, N+1\}$, $\alpha_j = 0$ Ainsi $\{\varphi_j\}_{0 \leq i \leq N+1}$ est une famille libre.

2.5.3 Résolution du problème de Dirichlet

Décrivons la résolution pratique du problème de Dirichlet (2.4) par la méthode des éléments finis \mathbb{P}_1 . La formulation variationnelle (2.2) de l'approximation interne devient ici :

$$\left\{ \begin{array}{l} \text{Trouvez } u_h \in V_h, \text{ tel que :} \\ \int_0^1 u_h'(x)v_h'(x)dx = \int_0^1 f_h(x)v_h(x)dx, \quad \forall v_h \in V_h. \end{array} \right. \quad (2.6)$$

On décompose u_h sur la base des $(\varphi_j)_{1 \leq j \leq N}$ et on prend $v_h = \varphi_i$; ce qui donne :

$$\sum_{j=1}^N u_h(x_j) \int_0^1 \varphi_j'(x)\varphi_i'(x)dx = \int_0^1 f(x)\varphi_i(x)dx,$$

En notant $U = (u_h(x_j))_{1 \leq j \leq N}$, $(b = \int_0^1 f(x)\varphi_i(x)dx)_{1 \leq i, j \leq N}$, et en introduisant la matrice de rigidité :

$$A = \left((x_j) \int_0^1 \varphi_j'(x)\varphi_i'(x)dx \right)_{1 \leq i, j \leq N}$$

la formulation variationnelle dans V_h revient à résoudre dans \mathbb{R}^N le système linéaire suivant :

$$AU = b$$

Comme les fonctions de base φ_j ont un "petit" support, l'intersection des supports de φ_j et φ_i est souvent vide et la plupart des coefficients de A sont nuls. Les coefficients non nuls se calculent facilement :

$$\begin{aligned} A_{i,i-1} &= a(\varphi_{i-1}, \varphi_i) = \int_0^1 \varphi_i'(x)\varphi_{i-1}'(x)dx = \int_{x_{i-1}}^{x_i} \frac{(-1)}{h} \frac{1}{h} dx = \frac{(-1)}{h}, \\ A_{i,i} &= a(\varphi_i, \varphi_i) = \int_0^1 (\varphi_i'(x))^2 dx = \int_{x_{i-1}}^{x_i} \frac{1}{h^2} dx + \int_{x_i}^{x_{i+1}} \frac{1}{(-h)^2} dx = \frac{2}{h}, \\ A_{i,i+1} &= a(\varphi_{i+1}, \varphi_i) = \int_0^1 \varphi_i'(x)\varphi_{i+1}'(x)dx = \int_{x_i}^{x_{i+1}} \frac{(-1)}{h} \frac{1}{h} dx = \frac{(-1)}{h}, \end{aligned}$$

D'où :

$$A_{i,j} = \begin{cases} \frac{-1}{h}, & \text{si } j = i - 1 \text{ ou } j = i + 1 \\ \frac{2}{h}, & \text{si } j = i \\ 0, & \text{si } j \neq \{i - 1, i, i + 1\} \end{cases} \quad (2.7)$$

Finalement, la matrice de rigidité est une matrice tridiagonale :

$$A = \frac{1}{h} \begin{pmatrix} 2 & -1 & 0 & & \ddots & 0 \\ -1 & 2 & -1 & 0 & & \\ 0 & -1 & 2 & -1 & 0 & \vdots \\ & & \ddots & \ddots & \ddots & 0 \\ \vdots & & 0 & -1 & 2 & -1 \\ 0 & & \ddots & 0 & -1 & 2 \end{pmatrix} \quad (2.8)$$

Remarque 2.5.2. Si le problème modèle est sous la forme : $-u'' + u = f$; on a au plus de la matrice de rigidité (2.8) ; la matrice de masse qui s'écrit de la forme suivante :

$$M = \frac{h}{6} \begin{pmatrix} 4 & 1 & 0 & & \ddots & 0 \\ 1 & 4 & 1 & 0 & & \\ 0 & 1 & 4 & 1 & 0 & \vdots \\ & & \ddots & \ddots & \ddots & 0 \\ \vdots & & 0 & 1 & 4 & 1 \\ 0 & & \ddots & 0 & 1 & 4 \end{pmatrix} \quad (2.9)$$

Pour obtenir le second membre b , il faut calculer l'intégrale :

$$b = \int_{x_{i-1}}^{x_{i+1}} f(x)\varphi_i(x)dx \quad \text{pour tout } 1 \leq i \leq N,$$

L'évaluation exacte du second membre b peut être difficile ou impossible si la fonction f est compliquée. En pratique on a recours à des formules de quadrature (ou formules d'intégration numérique) qui donnent une approximation des intégrales définissant b . Par exemple, on peut utiliser la formule du point milieu :

$$\frac{1}{x_{i+1} - x_i} \int_{x_i}^{x_{i+1}} \psi(x)dx \approx \psi\left(\frac{x_{i+1} + x_i}{2}\right),$$

ou la formule des trapèzes :

$$\frac{1}{x_{i+1} - x_i} \int_{x_i}^{x_{i+1}} \psi(x) dx \approx \frac{1}{2}(\psi(x_i) + \psi(x_{i+1})),$$

ou bien encore la formule de Simpson :

$$\frac{1}{x_{i+1} - x_i} \int_{x_i}^{x_{i+1}} \psi(x) dx \approx \frac{1}{6}\psi(x_i) + \frac{1}{6}\psi(x_{i+1}) + \frac{2}{3}\psi\left(\frac{x_i + x_{i+1}}{2}\right),$$

On utilise la formule du trapèze, et on trouve :

$$\begin{aligned} \int_{x_{i-1}}^{x_i} f(x)\varphi_i(x) dx &= \frac{x_i - x_{i-1}}{2}(f(x_i)\varphi_i(x_i) + f(x_{i-1})\varphi_i(x_{i-1})). \\ &= \frac{h}{2}f(x_i) \end{aligned}$$

et de la même manière on trouve :

$$\begin{aligned} \int_{x_i}^{x_{i+1}} f(x)\varphi_i(x) dx &= \frac{x_{i+1} - x_i}{2}(f(x_{i+1})\varphi_i(x_{i+1}) + f(x_i)\varphi_j(x_j)), \\ &= \frac{h}{2}f(x_i) \end{aligned}$$

D'où :

$$b = h \times f(x_i).$$

Ainsi, le système à résoudre s'écrit :

$$\frac{1}{h^2} \begin{pmatrix} 2 & -1 & 0 & & \cdots & 0 \\ -1 & 2 & -1 & 0 & & \\ 0 & -1 & 2 & -1 & 0 & \vdots \\ & & \ddots & \ddots & \ddots & 0 \\ \vdots & & 0 & -1 & 2 & -1 \\ 0 & & \ddots & 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_N \end{pmatrix} = \begin{pmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ \vdots \\ f(x_N) \end{pmatrix}.$$

On remarque que la matrice A est tridiagonale, symétrique et définie positive, alors le système ci-dessus admet une unique solution.

2.5.4 Résolution de problèmes paraboliques

On cherche une solution numérique de l'équation parabolique suivant avec conditions aux limites de Dirichlet :

$$\begin{cases} \partial_t u + Lu = f & \text{dans } \Omega_T, \\ u = 0 & \text{sur } \partial\Omega \times [0, T], \\ u = u_0 & \text{sur } \Omega \times \{t = 0\}, \end{cases} \quad (2.10)$$

Pour cela on utilise sa formulation variationnelle qui peut s'écrire sous la forme :

$$\frac{d}{dt} \int_{\Omega} uv dx + \int_{\Omega} A \nabla u \cdot \nabla v dx + \int_{\Omega} b \cdot (\nabla u) v dx + \int_{\Omega} cuv dx = 0. \quad (2.11)$$

Nous allons donner une approximation numérique du système (2.10) par la méthode des éléments finis. Pour cela, on définit un maillage du domaine Ω grâce auquel on va introduire un espace d'approximation V_h qui est un sous-espace de V de dimension finie. Pour discrétiser le domaine Ω , nous choisissons un maillage triangulaire \mathcal{T}_h , où l'indice h désigne la taille maximale des mailles. L'espace d'approximation V_h est l'ensemble des fonctions continues sur $\bar{\Omega}$ et qui sont polynômiales sur chaque triangle K du maillage. En choisissant une approximation par polynômes de degré inférieur ou égal à 1, on a :

$$V_h = \{v \text{ continue sur } \bar{\Omega}, v|_K \in \mathbb{P}_1 \text{ pour tout } K \in \mathcal{T}_h\}.$$

Cet espace admet des fonctions de base $(\varphi_j)_{1 \leq j \leq N}$ qui prennent la valeur 1 au noeud x_j et 0 aux autres noeuds x_k du maillage. Cherchons maintenant la solution du problème approché :

$$\begin{cases} \text{Trouver } u_h \in \mathcal{C}^1([0, T], V_h) \text{ tel que :} \\ \frac{d}{dt} \int_{\Omega} u_h v_h dx + \int_{\Omega} A \nabla u_h \cdot \nabla v_h dx + \int_{\Omega} b \cdot (\nabla u_h) v_h dx + \\ \int_{\Omega} cu_h v_h dx = 0 \quad \text{pour } t \in [0, T], \forall v_h \in V_h \\ u_h(0) = u_{0h} \end{cases} \quad (2.12)$$

$u_{0h} \in V_h$ est une approximation de u_0 . La fonction inconnue u_h est discrétisée sous la forme :

$$u_h(x, t) = \sum_{j=1}^N u_j(t) \varphi_j(x), \quad t > 0,$$

u_j est la valeur de la solution u_h au noeud x_j , et N est le nombre total de noeuds du domaine discrétisé Ω . On pose $v_h = \varphi_i$, la formulation variationnelle dans (2.12) donne :

$$\sum_{j=1}^N \left[d_t u_j(t) \int_{\Omega} \varphi_j \varphi_i dx + u_j(t) \left(\int_{\Omega} A \nabla \varphi_j \cdot \nabla \varphi_i dx + \int_{\Omega} b \cdot (\nabla \varphi_j) \varphi_i dx + \int_{\Omega} c \varphi_j \varphi_i dx \right) \right] = 0, \quad (2.13)$$

pour $i = 1, \dots, N$. Considérons $\mathbf{u}(t) = (u_i(t))_{1 \leq i \leq N}$. Pour $t \in [0, T]$, on introduit les matrices :

$$\begin{aligned} \mathcal{M}_{ij} &= \int_{\Omega} \varphi_j \varphi_i dx, \\ \mathcal{S}_{ij}(t) &= \int_{\Omega} A(x, t) \nabla \varphi_j \cdot \nabla \varphi_i dx, \\ \mathcal{K}_{ij}(t) &= \int_{\Omega} b(x, t) \cdot (\nabla \varphi_j) \varphi_i dx, \\ \mathcal{H}_{ij}(t) &= \int_{\Omega} c(x, t) \varphi_j \varphi_i dx, \end{aligned}$$

pour $i, j = 1, \dots, N$. La formulation (2.13) peut être écrite sous forme matricielle :

$$\mathcal{M} \frac{d}{dt} \mathbf{u}(t) + \mathcal{A}(t) \mathbf{u}(t) = 0, \quad t > 0, \quad (2.14)$$

avec

$$\mathcal{A}(t) = \mathcal{S}(t) + \mathcal{K}(t) + \mathcal{H}(t),$$

et la condition initiale $\mathbf{u}(0)$ est donnée. Pour résoudre l'EDO (2.14), on utilise un schéma de discrétisation en temps. On peut choisir par exemple le θ -schéma avec $\theta \in [0, 1]$ un paramètre donné. L'intervalle $[0, T]$ est subdivisé en $n_T \in \mathbb{N}$ sous-intervalles de même longueur $\Delta t := T/n_T$. On pose $t^n = n\Delta t$ pour $0 \leq n \leq n_T$. Le terme $\frac{d}{dt} \mathbf{u}(t)$ est approché par la différence finie $\frac{1}{\Delta t} (\mathbf{u}^{n+1} - \mathbf{u}^n)$. Nous obtenons donc la suite $(\mathbf{u}^n)_{0 \leq n \leq n_T}$ qui vérifie :

$$\frac{1}{\Delta t} \mathcal{M} (\mathbf{u}^{n+1} - \mathbf{u}^n) + \theta \mathcal{A}^{n+1} \mathbf{u}^{n+1} + (1 - \theta) \mathcal{A}^n \mathbf{u}^n = 0, \quad (2.15)$$

avec $\mathcal{A}^n = \mathcal{A}(n\Delta t)$ et $\mathbf{u}^0 = \mathbf{u}(0)$ est donné. Dans les simulations numériques présentées au Chapitre 4, un schéma d'Euler implicite ($\theta = 1$) est choisi pour

la discrétisation en temps. Dans ce cas, l'équation (2.15) donne :

$$\frac{1}{\Delta t} \mathcal{M}(\mathbf{u}^{n+1} - \mathbf{u}^n) + \mathcal{A}^{n+1} \mathbf{u}^{n+1} = 0$$

qui peut être résolue à chaque pas de temps n pour trouver la solution \mathbf{u} .

Chapitre 3

Méthode de calcul adjoint et contrôle de paramètres

3.1 Introduction

Dans ce chapitre nous présentons la méthode adjointe qui est une méthode numérique pour calculer efficacement le gradient d'une fonction ou d'un opérateur dans un problème d'optimisation numérique. Elle intervient aussi dans l'optimisation avec contraintes de solutions des EDPs. Elle a des applications en géophysique, en imagerie sismique, en photonique et plus récemment dans les réseaux de neurones. La méthode adjointe formule le gradient d'une fonction dépendant de plusieurs paramètres sous une forme d'optimisation sous contraintes [13]. La forme adjointe du problème d'optimisation sous contraintes, peut être utilisé pour calculer le gradient très rapidement. Une belle propriété est que le nombre de calculs est indépendant du nombre de paramètres pour lesquels on veut calculer gradient.

L'avantage de la méthode adjointe est qu'elle peut calculer le gradient par rapport à de nombreux paramètres avec une seule solution de l'équation d'état et une solution de l'équation d'état adjoint, quel que soit le nombre de paramètres. Cela la rend plus efficace que les méthodes aux différences finies, par exemple, qui nécessitent de résoudre l'équation d'état pour chaque paramètre de perturbation [14].

La méthode adjointe peut être utilisée dans les problèmes d'identification de paramètres à partir d'observations expérimentales. Ceci est traité comme un problème de minimisation d'une fonction coût. Dans ce chapitre nous allons introduire quelques notions fondamentales d'optimisation, puis présenter des calculs formels sur le gradient de problèmes d'identification de paramètres. En annulant la dérivée du Lagrangien associé au problème d'optimisation, on obtient une formule pour le gradient de la fonction coût à minimiser. Ce gradient dépendra de la solution du problème de l'état adjoint associé. Nous présentons quelques algorithmes d'optimisations de fonctions utilisés dans les simulations numériques. Nous donnons enfin quelques exemples d'optimisation et de calcul adjoints liés aux solutions d'EDPs.

3.2 Motivation

L'idée des problèmes d'estimation ou d'identification de paramètres réside dans le fait qu'il s'agit d'optimiser une fonctionnelle dépendant de la solution d'une équation aux dérivées partielles. En effet ces problèmes sont formulés comme l'optimisation de l'écart quadratique entre les observations expérimentales et les quantités correspondantes calculées comme la solution d'EDPs. Les variables à contrôler (variable de contrôle) sont dans ce cas les paramètres ou les fonctions à identifier. L'optimisation de telles fonctionnelles relève de "la théorie du contrôle optimal" des EDPs développée par J.L.Lions [15].

Les méthodes d'optimisation requièrent le calcul du gradient de la "fonction coût" J :

$$J(x_0) = \int_0^T \|H(x(t)) - y^o(t)\|^2 dt + \|x_0 - x_b\|^2, \quad (3.1)$$

où J est définie d'un certain espace vectoriel V dans \mathbb{R} . On cherche à résoudre le problème :

$$\text{Trouver } x_0^* \in U \text{ tel que } J(x_0^*) = \min_{x_0 \in U} J(x_0),$$

U est un convexe fermé de V , appelée espace "des solutions admissibles", ou "espace de contrôle". Si $U = V$ le problème est sans contraintes, si $U \neq V$, le

problème est avec contraintes. La difficulté d'optimisation vient généralement du fait que J ne dépend de x_0 que de façon assez indirecte, via par exemple un modèle et un opérateur d'observation plus ou moins complexes. Nous allons étudier le problème d'optimisation pour des observations distribuées dans le temps. On tient donc compte d'un modèle d'évolution dans le temps qui permettra de comparer l'état du système avec les observations à l'instant approprié. La fenêtre d'optimisation est un intervalle de temps donné $[0, T]$, l'analyse est réalisée à l'instant initial. On notera $y^o(t)$ les observations, $x(t)$ l'état du système. La minimisation de cette fonction coût est réalisée sous contrainte. Il s'agit d'une contrainte forte de modèle, puisque l'écriture de $J(x)$ dépend des valeurs $x(t)$, qui elles-mêmes dépendent de la condition initiale x_0 .

Le vecteur des observations y^o est souvent de taille inférieure à celle du vecteur d'état x . L'espace des observations et l'espace d'état sont en général différents, et donc de dimensions différentes. Le passage de l'espace d'état où vit le vecteur x à l'espace des observations où vit y^o se fait grâce à un opérateur d'observation H .

Le problème est souvent sous-déterminé, il faut donc rajouter des contraintes pour le résoudre. Si le modèle et l'opérateur d'observation sont linéaires, alors J est quadratique et son minimum est unique. Sinon, il peut exister des minima locaux, ce qui crée des difficultés pour les algorithmes de minimisation, d'où l'importance de l'ébauche x_b qui présente une estimation à priori de l'état à analyser, donc à x_0 . Elle est issue d'un modèle ou d'une prévision. Le terme en x_b agit comme une régularisation de J et pour ne pas partir trop loin du minimum.

Si la dépendance de J vis à vis de la variable de contrôle est complexe et/ou indirecte, ce calcul peut être difficile. Numériquement, on peut toujours s'en sortir par un calcul de taux d'accroissement. Mais pour avoir le gradient, il faut faire le calcul pour chaque direction de perturbation possible. Si x_0 est la condition initiale du modèle, donc l'état à analyser. Le gradient $\nabla_{x_0} J$ va demander plusieurs simulations du modèle pour chaque composante du gradient, ce qui donne un nombre de simulations supérieure à la taille de l'état initial, et ce à chaque itération de l'algorithme de descente. En pratique c'est le plus souvent hors de portée comme en météorologie ou en océanographie

où l'état initial vit dans un espace de dimension 10^6 à 10^7 . Cependant, si on contrôle un nombre très restreint de paramètres, c'est alors faisable, sinon il est préférable de mettre en œuvre des méthodes plus lourdes comme le calcul adjoint du gradient.

Lorsque la fonctionnelle est choisie (ébauche, normes), le problème est entièrement défini, et sa solution aussi. Toute la partie "physique" est donc dans la définition de la fonctionnelle. Le reste, c'est-à-dire, trouver le minimum, n'est que de la technique.

La fonction coût peut être rendue quadratique en prenant, un opérateur d'observation H linéaire. On pourra alors remplacer le modèle par son approximation linéaire tangente (ou sa dérivée), en introduisant le modèle linéaire tangent. Cela permet de se ramener à un problème d'optimisation linéaire sans contraintes, ce qui est plus simple à résoudre. Le premier terme de J dans l'équation (4.4) inclut la résolution du modèle d'évolution, et le deuxième terme c'est un terme de rappel à l'ébauche, qui agit également comme une régularisation de J .

Une fois la fonction coût et son gradient sont calculés, la minimisation est réalisée à l'aide d'une méthode de gradient. Le minimum de la fonction coût est donc trouvé après un nombre fini d'itérations dans l'algorithme de minimisation. La prévision de $x(t)$ est déduite en résolvant simplement le modèle direct, initialisé par l'état d'analyse x_0 que l'on a trouvé.

3.3 Rappel d'optimisation

3.3.1 Fonctions convexes et résultat d'existence et d'unicité

Soit H un espace de Hilbert muni de la norme $\|\cdot\|$ et du produit scalaire (\cdot, \cdot) .

Définition 3.3.1. Une fonction $J : H \rightarrow \mathbb{R}$ est dite convexe si :

$$\forall u, v \in H, \forall \theta \in [0, 1], J(\theta u + (1 - \theta)v) \leq \theta J(u) + (1 - \theta)J(v),$$

et strictement convexe si :

$$\forall u \neq v \in H, \forall \theta \in]0, 1[, J(\theta u + (1 - \theta)v) < \theta J(u) + (1 - \theta)J(v).$$

Proposition 3.3.1. Soit $J : H \rightarrow \mathbb{R}$ une fonction différentiable. Alors :

– J est convexe si et seulement si :

$$J(v) \geq J(u) + \langle \nabla J(u), v - u \rangle, \forall u, v \in H.$$

– J est strictement convexe si et seulement si :

$$J(v) > J(u) + \langle \nabla J(u), v - u \rangle, \forall u \neq v \in H.$$

Définition 3.3.2. Une fonction $J : H \rightarrow \mathbb{R}$ est dite fortement convexe ou α -convexe s'il existe $\alpha > 0$ tel que :

$$J(\theta u + (1 - \theta)v) \leq \theta J(u) + (1 - \theta)J(v) - \frac{\alpha}{2}\theta(1 - \theta)\|v - u\|^2, \forall u, v \in H.$$

Proposition 3.3.2. Soit $J : H \rightarrow \mathbb{R}$ une fonction différentiable. Alors les propositions suivantes sont équivalentes

- J est fortement convexe.
- La fonction $J - \frac{\alpha}{2}\|\cdot\|^2$ est convexe.
- J est α -elliptique, autrement dit :

$$\langle \nabla J(v) - \nabla J(u), v - u \rangle \geq \alpha\|v - u\|^2, \forall u, v \in H.$$

Preuve. Posons $g(x) = J(x) - \frac{\alpha}{2}\|x\|^2$. En développant $\|tx + (1 - t)y\|^2$ et en regroupant les termes correctement, on trouve :

$$tg(x) + (1 - t)g(y) - g(tx + (1 - t)y) = tJ(x) + (1 - t)J(y) - J(tx + (1 - t)y) - \frac{\alpha}{2}t(1 - t)\|x - y\|^2,$$

ce qui prouve la première équivalence annoncée. La deuxième équivalence résulte de la proposition : si $g : H \rightarrow \mathbb{R}$ est différentiable, alors g est convexe si et seulement si :

$$g(y) \geq g(x) + \langle \nabla g(x), y - x \rangle, \forall (x, y) \in H^2$$

ou encore si, et seulement si $\langle \nabla g(y) - \nabla g(x), y - x \rangle \geq 0, \forall (x, y) \in H^2$.

Théorème 3.3.1. Soient $J : H \rightarrow \mathbb{R}$ une fonction continue, α -convexe et \mathcal{U}_{ad} une partie non-vide convexe fermée de H . Alors le problème $\inf\{J(v), v \in \mathcal{U}_{ad}\}$ admet une unique solution.

Preuve. Soit (u_k) une suite minimisante, i.e. $\lim_{k \rightarrow \infty} J(u_k) = \inf_{v \in \mathcal{U}_{ad}} J(v)$. La α -convexité de J implique que :

$$\frac{\alpha}{4} \|u_k - u_l\|^2 \leq J(u_k) + J(u_l) - 2J\left(\frac{u_k + u_l}{2}\right) \leq J(u_k) + J(u_l) - 2 \inf_{v \in \mathcal{U}_{ad}} J(v),$$

puisque $\frac{u_k + u_l}{2} \in \mathcal{U}_{ad}$ qui est convexe. On en déduit que la suite (u_k) est de Cauchy dans H , i.e. $\|u_k - u_l\|^2 \rightarrow 0$ si $k, l \rightarrow +\infty$. L'ensemble \mathcal{U}_{ad} étant fermé, la suite de Cauchy (u_k) converge vers un élément $u \in \mathcal{U}_{ad}$. Il s'ensuit que

$$J(u) = \lim_{k \rightarrow \infty} J(u_k) = \inf_{v \in \mathcal{U}_{ad}} J(v).$$

L'unicité découle de la stricte convexité de la fonctionnelle J .

Théorème 3.3.2. *Supposons que $J : H \rightarrow \mathbb{R}$ est une fonction différentiable et convexe. Alors $u \in \mathcal{U}_{ad}$ est solution du problème $\inf\{J(v), v \in \mathcal{U}_{ad}\}$ si et seulement s'il satisfait l'inéquation d'Euler :*

$$(\nabla J(u), v - u) \geq 0, \forall v \in \mathcal{U}_{ad}.$$

Preuve. Supposons que $u \in \mathcal{U}_{ad}$ satisfait l'inéquation d'Euler. Alors la convexité de J implique :

$$J(v) \geq J(u) + (\nabla J(u), v - u) \geq J(u), \forall v \in \mathcal{U}_{ad}.$$

Réciproquement, supposons que u est solution de $\inf\{J(v), v \in \mathcal{U}_{ad}\}$. Alors pour tout $0 \leq \theta \leq 1$, la convexité de \mathcal{U}_{ad} implique :

$$u + \theta(v - u) = (1 - \theta)u + \theta v \in \mathcal{U}_{ad}.$$

Par minimalité de $J(u)$, on a $J(u) \leq J(u + \theta(v - u))$. Puisque J est dérivable, la formule de Taylor fournit :

$$J(u) \leq J(u) + \theta(\nabla J(u), v - u) + o(\theta).$$

La conclusion s'obtient en simplifiant par $\theta > 0$, puis en faisant tendre $\theta \rightarrow 0^+$ dans la relation ci-dessus.

Remarque 3.3.1. *Si $\mathcal{U}_{ad} = H$, l'inéquation d'Euler :*

$$(\nabla J(u), v - u) \geq 0, \forall v \in \mathcal{U}_{ad}.$$

se réécrit $\nabla J(u) = 0$.

En effet, choisissons $v = u - \nabla J(u) \in H$. Alors l'inéquation d'Euler se réécrit $-\|\nabla J(u)\|^2 \geq 0$. D'où le résultat.

3.3.2 Dérivée directionnelle et dérivée au sens de Fréchet

Définition 3.3.3. Soit f une fonction définie d'un espace vectoriel E dans \mathbb{R} . On appelle dérivée de Gâteaux, ou dérivée directionnelle, de f au point x dans la direction $d \in E$ la limite, si elle existe :

$$\lim_{\alpha \rightarrow 0} \frac{f(x + \alpha d) - f(x)}{\alpha}.$$

Cette dérivée directionnelle sera notée $\frac{\partial f}{\partial d}(x)$, ou encore $f'_d(x)$, ou encore $\hat{f}(x)[d]$.

Exemple. 1. Les dérivées partielles dans \mathbb{R}^n sont les dérivées directionnelles dans la direction des vecteurs de la base canonique :

$$\frac{\partial f}{\partial x_i} = \lim_{\alpha \rightarrow 0} \frac{f(x + \alpha e_i) - f(x)}{\alpha}.$$

2. Soit $f(x, y) = 4x^2 - 2y$ et $d = (1, 2)$:

$$\frac{\partial f}{\partial d} = \lim_{\alpha \rightarrow 0} \frac{f(x + \alpha, y + 2\alpha) - f(x, y)}{\alpha} = \lim_{\alpha \rightarrow 0} 4(2x + \alpha) - 4 = 8x - 4.$$

On peut remarquer que :

$$\frac{\partial f}{\partial d}(x) = 8x - 4 = \begin{pmatrix} 8x \\ -2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \nabla f \cdot d$$

3. Considérons maintenant la fonctionnelle :

$$J(K) = \int_{\Omega} (u - u^{obs})^2$$

où u dépend indirectement des paramètres K .

$$\begin{aligned} \hat{J}(K)[k] &= \lim_{\alpha \rightarrow 0} \frac{J(K + \alpha k) - J(K)}{\alpha} \\ &= \lim_{\alpha \rightarrow 0} \int_{\Omega} \frac{u_{K+\alpha k} - u_K}{\alpha} (u_{K+\alpha k} + u_K - 2u^{obs}) \\ &= 2 \int_{\Omega} (u - u^{obs}) \hat{u} \end{aligned}$$

où

$$\hat{u} = \lim_{\alpha \rightarrow 0} \frac{u_{K+\alpha k} - u_K}{\alpha}.$$

Remarque 3.3.2. En pratique, le calcul d'une dérivée directionnelle peut être fait par taux d'accroissement. On calcule ainsi :

$$T(\alpha) = \frac{f(x + \alpha d) - f(x)}{\alpha}$$

pour différentes valeurs de α de plus en plus petites. En général, on a convergence quand α devient suffisamment petit, jusqu'à ce qu'il devienne trop petit et qu'on observe une divergence numérique due à la précision machine. On reviendra là-dessus pour le "test du gradient", qui permet de tester le modèle adjoint.

Définition 3.3.4. Soit E un espace de Hilbert. On dit que f est Fréchet différentiable en x s'il existe $p \in E$ tel que :

$$f(x + h) = f(x) + (p, h) + o(h), \quad \forall h \in E$$

où $o(h)$ signifie que $\lim_{h \rightarrow 0} \frac{\|o(h)\|}{\|h\|} = 0$.

On dit alors que p est la dérivée, la différentielle, ou encore le gradient de f en x , noté $f'(x)$ ou $\nabla f(x)$.

Remarque 3.3.3. Si le gradient existe, alors la dérivée directionnelle dans la direction d est donnée par :

$$f'_d(x) = f'(x).d$$

(c'est immédiat d'après la définition).

La Fréchet-différentiabilité implique la Gâteaux différentiabilité. La réciproque est fautive. Par exemple :

$$f(x, y) = \frac{x^6}{(y - x^2)^2 + x^8} \text{ si } (x, y) \neq (0, 0), \text{ et } f(0, 0) = 0.$$

Remarque 3.3.4. Si on revient à l'exemple de

$$J(K) = \int_{\Omega} (u - u^{obs})^2$$

la dérivée directionnelle :

$$\hat{J}(K)[k] = 2 \int_{\Omega} (u - u^{obs}) \hat{u}$$

ne s'écrit pas sous la forme $(\nabla J, k)$, le gradient n'est pas trivial. On utilisera la méthode adjointe pour l'obtenir.

3.3.3 Minimisation avec contrainte d'égalité

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ différentiable. Soit K l'ensemble des contraintes :

$$K = \{x \in \mathbb{R}^n, h_1(x) = 0, h_2(x) = 0, \dots, h_p(x) = 0\}$$

où les fonctions $h_i : \mathbb{R}^n \rightarrow \mathbb{R}$ sont au moins de classe \mathcal{C}^1 . On cherche alors $\hat{x} \in K$ tel que

$$f(\hat{x}) = \min_{x \in K} f(x)$$

Théorème 3.3.3. *Si $\hat{x} \in K$ est un minimum local de f sur K , si les vecteurs $\nabla h_1(\hat{x}), \nabla h_2(\hat{x}), \dots, \nabla h_p(\hat{x})$ sont linéairement indépendants, alors il existe $\hat{\lambda} = (\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_p) \in \mathbb{R}^p$ tel que :*

$$\nabla f(\hat{x}) + \sum_{i=1}^p \hat{\lambda}_i \nabla h_i(\hat{x}) = 0.$$

On introduit "le Lagrangien" \mathcal{L} :

$$\mathcal{L}(x, \lambda) = f(x) + \sum_{i=1}^p \lambda_i h_i(x).$$

Minimiser f dans K , donc résoudre $\nabla f(x) = 0$ dans K revient à résoudre $\nabla \mathcal{L} = 0$, donc à chercher un point selle de \mathcal{L} . En effet, $\nabla \mathcal{L} = 0$ signifie :

$$\begin{aligned} \nabla_x \mathcal{L} &= \nabla f + \sum_i \nabla h_i = 0 \\ \nabla_\lambda \mathcal{L} &= (h_1, h_2, \dots, h_p)^T = 0 \end{aligned}$$

La deuxième ligne $\nabla_\lambda \mathcal{L} = 0$ donne les contraintes, et donc $x \in K$, la première donne la condition du théorème.

Exemple. Supposons inconnues x et y , avec une observation $x + y = 2$, et cherchons la solution de norme minimale. Le problème peut se voir comme celui de minimiser :

$$f(x, y) = x^2 + y^2, \text{ sous la contrainte } x + y = 2.$$

Ecrivons le Lagrangien :

$$\mathcal{L}(x, y, \lambda) = x^2 + y^2 + \lambda(x + y - 2)$$

La condition de point selle donne :

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x} = 0 = 2x + \lambda \\ \frac{\partial \mathcal{L}}{\partial y} = 0 = 2y + \lambda \\ \frac{\partial \mathcal{L}}{\partial \lambda} = 0 = x + y - 2 \end{cases}$$

On obtient

$$\lambda = -2, x = 1, y = 1.$$

3.3.4 Méthode de calcul adjointe

On considère le problème suivant :

$$\text{Trouver le minimum } \hat{x} \text{ tel que } J(\hat{x}) = \min_{x \in \mathbb{R}^n} J(x)$$

On appelle méthode de descente tout algorithme du type :

$$x_{k+1} = x_k + \alpha_k d_k, \text{ tel que } J(x_{k+1}) < J(x_k)$$

où :

- $d_k \in \mathbb{R}^n$ est la direction de descente à l'itération k ,
- $\alpha_k \in \mathbb{R}$ est le pas de descente à l'itération k ,

Les méthodes de descente diffèrent dans le choix des α_k et des d_k .

Méthodes à pas optimal, à pas constant

On suppose que d_k est choisi. Les méthodes à pas optimal choisissent α_k tel que :

$$J(x_k + \alpha_k d_k) = \min_{\alpha \in \mathbb{R}} J(x_k + \alpha d_k)$$

Autrement dit on minimise J dans la direction d_k .

Cela peut être coûteux, et cela n'est pas forcément intéressant, dans la mesure où ce minimum n'est pas celui de J mais juste une valeur intermédiaire.

Il existe également les méthodes à pas constant : $\alpha_k = \alpha, \forall k$.

Méthodes de relaxation

L'idée est de choisir pour les d_k successivement les vecteurs de la base canonique :

$$x_{k+1} = x_k + \alpha_k e_k$$

On modifie donc x_i simplement composante par composante.

La méthode a l'avantage d'être simple, mais elle peut être lente à converger.

Méthodes de gradient

Par définition du gradient, on a :

$$J(x_k + h) = J(x_k) + (\nabla J(x_k), h) + o(h)$$

Donc, si $\nabla J(x_k) \neq 0$, la partie principale de l'accroissement de J est $(\nabla J(x_k), h)$, qui sera la plus négative possible lorsque $h = -\alpha \nabla J(x_k)$, ce qui revient à poser :

$$d_k = -\nabla J(x_k)$$

Méthodes de Newton

On se pose en dimension 1. On cherche à résoudre $f(x) = 0$. On suppose connu x_k , on définit x_{k+1} comme le point d'intersection de l'axe des abscisses avec la tangente à la courbe de f en x_k :

$$\frac{f(x_k) - 0}{x_k - x_{k+1}} = f'(x_k) \Rightarrow x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

La généralisation de la méthode précédente en dimension n est :

$$x^{(k+1)} = x^{(k)} - [f'(x^{(k)})]^{-1} f(x^{(k)})$$

où $f'(x^{(k)})$ est la matrice jacobienne $[\partial_j f_i(x^{(k)})]$. A chaque itération, il faut donc calculer cette matrice jacobienne, et résoudre un système linéaire $[f'(x^{(k)})] \delta x^{(k)} = -f(x^{(k)})$.

On applique la méthode de Newton à l'équation d'Euler $\nabla J(x) = 0$. Cela qui donne :

$$x_{k+1} = x_k - [\nabla^2 J(x_k)]^{-1} \nabla J(x_k),$$

3.4 Contrôle de paramètres d'équation différentielle ordinaires 51

où $\nabla^2 J(x_k)$ est la Hessienne de J . La différence majeure avec les problèmes de gradient est donc que la direction de descente n'est plus $\nabla J(x_k)$, mais $[\nabla^2 J(x_k)]^{-1} \nabla J(x_k)$.

Il faut calculer la hessienne à chaque itération, et résoudre un système linéaire de matrice $\nabla^2 J(x_k)$. Pour des gros problèmes, c'est hors de portée. Il existe donc des algorithmes comme ceux de Quasi-Newton qui fournissent des approximations de la hessienne, qui s'améliorent au fil des itérations, à un coût raisonnable.

3.4 Contrôle de paramètres d'équation différentielle ordinaires

On considère l'équation différentielle ordinaire suivante :

$$\begin{cases} -bu''(x) + cu'(x) = f(x), & x \in]0, 1[\\ u(0) = 0, & u(1) = 0 \end{cases}$$

avec f une fonction donnée de $L^2(]0, 1[)$, b et c des paramètres inconnus, que l'on cherche à identifier grâce à une mesure de $u(x)$ sur $]0, 1[$. La fonction coût s'écrit

$$J(b, c) = \int_0^1 (u(x) - u^{obs}(x))^2 dx$$

Calculons son gradient :

$$\begin{aligned} J(b + \alpha\delta b, c + \alpha\delta c) - J(b, c) &= \int_0^1 (u_{b+\alpha\delta b, c+\alpha\delta c} - u^{obs})^2 - (u_{b,c} - u^{obs})^2 \\ &= \int_0^1 (u_{b+\alpha\delta b, c+\alpha\delta c} + u_{b,c} - 2u^{obs}) (u_{b+\alpha\delta b, c+\alpha\delta c} - u_{b,c}) \end{aligned}$$

En notant $\tilde{u} = u_{b+\alpha\delta b, c+\alpha\delta c}$, $u = u_{b,c}$, on a :

$$J(b + \alpha\delta b, c + \alpha\delta c) - J(b, c) = \int_0^1 (\tilde{u} + u - 2u^{obs}) (\tilde{u} - u)$$

En divisant par α et en passant à la limite quand α tend vers 0, on obtient :

3.4 Contrôle de paramètres d'équation différentielle ordinaires 52

$$\hat{J}[b, c](\delta b, \delta c) = 2 \int_0^1 (u - u^{obs}) \hat{u}, \quad \text{avec } \hat{u} = \lim_{\alpha \rightarrow 0} \frac{\tilde{u} - u}{\alpha}$$

Cherchons maintenant l'équation vérifiée par \hat{u} . On a

$$\begin{cases} -(b + \alpha \delta b) \tilde{u}'' + (c + \alpha \delta c) \tilde{u}' = f \\ \tilde{u}(0) = 0, \quad \tilde{u}(1) = 0 \end{cases}$$

$$\begin{cases} -bu'' + cu' = f \\ u(0) = 0, \quad u(1) = 0 \end{cases}$$

D'où

$$\begin{cases} -b\hat{u}'' - \delta b u'' + c\hat{u}' + \delta c u' = 0 \\ \hat{u}(0) = 0, \quad \hat{u}(1) = 0 \end{cases}$$

On obtient ainsi le modèle linéaire tangent :

$$\begin{cases} -b\hat{u}'' + c\hat{u}' = \delta b u'' - \delta c u' \\ \hat{u}(0) = 0, \quad \hat{u}(1) = 0 \end{cases}$$

On veut pouvoir reformuler $\int_0^1 (u - u^{obs}) \hat{u}$, donc on multiplie le modèle linéaire tangent par une variable p et on intègre :

$$-b \int_0^1 \hat{u}'' p + c \int_0^1 \hat{u}' p = \int_0^1 (\delta b u'' - \delta c u') p$$

Calculons séparément :

$$\begin{aligned} \int_0^1 \hat{u}'' p &= [\hat{u}' p]_0^1 - \int_0^1 \hat{u}' p' \\ &= [\hat{u}' p - \hat{u} p']_0^1 + \int_0^1 \hat{u} p'' \\ &= \hat{u}'(1)p(1) - \hat{u}'(0)p(0) + \int_0^1 \hat{u} p'' \\ \int_0^1 \hat{u}' p &= [\hat{u} p]_0^1 - \int_0^1 \hat{u} p' \\ &= - \int_0^1 \hat{u} p' \end{aligned}$$

Ce qui donne :

$$\begin{aligned}
 & -b \left(\hat{u}'(1)p(1) - \hat{u}'(0)p(0) + \int_0^1 \hat{u}p'' \right) + c \left(- \int_0^1 \hat{u}p' \right) = \int_0^1 (\delta b u'' - \delta c u') p \\
 \Leftrightarrow & \int_0^1 (-bp'' - cp') \hat{u} = b\hat{u}'(1)p(1) - b\hat{u}'(0)p(0) + \int_0^1 (\delta b u'' - \delta c u') p
 \end{aligned}$$

Posons maintenant

$$\begin{cases} -bp'' - cp' = 2(u - u^{obs}) \\ p(0) = 0, \quad p(1) = 0 \end{cases}$$

On appelle ces équations le modèle adjoint. On a alors

$$2 \int_0^1 (u - u^{obs}) \hat{u} = \int_0^1 (-bp'' - cp') \hat{u} = \delta b \left(\int_0^1 pu'' \right) + \delta c \left(- \int_0^1 pu' \right)$$

D'où

$$\nabla J(b, c) = \left(\int_0^1 pu'', - \int_0^1 pu' \right)$$

On vient donc de calculer le gradient grâce au modèle adjoint.

3.5 Ecriture générale dans le cas du contrôle de la condition initiale

On considère le modèle suivant :

$$\begin{cases} \frac{dX}{dt} = M(X), & \text{dans } \Omega \times [0, T] \\ X(t=0) = U \end{cases}$$

avec la fonction coût

$$J(U) = \frac{1}{2} \int_0^T \|HX - Y^o\|^2$$

3.5.1 Dérivée directionnelle de J

On perturbe U dans la direction u . On note \tilde{X} la trajectoire correspondante :

$$\begin{cases} \frac{d\tilde{X}}{dt} = M(\tilde{X}) \\ \tilde{X}(t=0) = U + \alpha u \end{cases}$$

On a alors

$$\begin{aligned}
J(U + \alpha u) - J(u) &= \frac{1}{2} \int_0^T \|H\tilde{X} - Y\|^2 - \|HX - Y\|^2 \\
&= \frac{1}{2} \int_0^T (H\tilde{X} - Y, H\tilde{X} - HX + HX - Y) - (HX - Y, HX - Y) \\
&= \frac{1}{2} \int_0^T (H\tilde{X} - Y, H(\tilde{X} - X)) + (H\tilde{X} - Y - (HX - Y), HX - Y) \\
&= \frac{1}{2} \int_0^T (H\tilde{X} - Y, H(\tilde{X} - X)) + (H(\tilde{X} - X), HX - Y)
\end{aligned}$$

On pose alors

$$\hat{X} = \lim_{\alpha \rightarrow 0} \frac{\tilde{X} - X}{\alpha}$$

et on calcule

$$\begin{aligned}
\hat{J}[U](u) &= \lim_{\alpha \rightarrow 0} \frac{J(U + \alpha u) - J(u)}{\alpha} \\
&= \frac{1}{2} \int_0^T (HX - Y, H\hat{X}) + (H\hat{X}, HX - Y) \\
&= \int_0^T (H\hat{X}, HX - Y) \\
&= \int_0^T (\hat{X}, H^T(HX - Y))
\end{aligned}$$

3.5.2 Modèle linéaire tangent

En soustrayant les équations vérifiées par \tilde{X} et X on obtient :

$$\begin{cases} \frac{d(\tilde{X} - X)}{dt} = M(\tilde{X}) - MX = \left[\frac{\partial M}{\partial X} \right] (\tilde{X} - X) + \frac{1}{2} (\tilde{X} - X)^T \left[\frac{\partial^2 M}{\partial X^2} \right] (\tilde{X} - X) + \dots \\ (\tilde{X} - X)(t=0) = \alpha u \end{cases}$$

En divisant par α et en faisant tendre α vers 0 , on obtient :

$$\begin{cases} \frac{d\hat{X}}{dt} = \left[\frac{\partial M}{\partial X} \right] \hat{X} \\ \hat{X}(t=0) = u \end{cases}$$

Ces équations sont appelés modèle linéaire tangent.

3.5.3 Modèle adjoint

Comme dans l'exemple précédent, on multiplie l'équation du modèle tangent par P et on intègre par parties sur $[0, T]$:

$$\begin{aligned} \int_0^T \left(\frac{d\hat{X}}{dt}, P \right) &= - \int_0^T \left(\hat{X}, \frac{dP}{dt} \right) + [(\hat{X}, P)]_0^T \\ &= - \int_0^T \left(\hat{X}, \frac{dP}{dt} \right) + (\hat{X}(T), P(T)) - (\hat{X}(0), P(0)) \\ &= - \int_0^T \left(\hat{X}, \frac{dP}{dt} \right) + (\hat{X}(T), P(T)) - (u, P(0)) \\ \int_0^T \left(\left[\frac{\partial M}{\partial X} \right] \hat{X}, P \right) &= \int_0^T \left(\hat{X}, \left[\frac{\partial M}{\partial X} \right]^T P \right) \end{aligned}$$

D'où

$$\int_0^T \left(\frac{d\hat{X}}{dt} - \left[\frac{\partial M}{\partial X} \right] \hat{X}, P \right) = 0 = \int_0^T \left(\hat{X}, -\frac{dP}{dt} - \left[\frac{\partial M}{\partial X} \right]^T P \right) + (\hat{X}(T), P(T)) - (u, P(0))$$

En identifiant avec

$$\hat{J}[U](u) = \int_0^T \left(\hat{X}, H^T(HX - Y) \right)$$

On obtient les équations du modèle adjoint :

$$\begin{cases} \frac{dP}{dt} + \left[\frac{\partial M}{\partial X} \right]^T P = H^T(HX - Y) \\ P(t = T) = 0 \end{cases}$$

On remarque que le modèle adjoint est rétrograde en temps : l'équation se résout de T jusqu'à 0 .

3.5.4 Calcul du gradient

Le modèle adjoint permet de réécrire le gradient :

$$\begin{aligned}\hat{J}[U](u) &= \int_0^T \left(\hat{X}, H^T(HX - Y) \right) \\ &= \int_0^T \left(\hat{X}, \frac{dP}{dt} + \left[\frac{\partial M}{\partial X} \right]^T P \right) \\ &= -(u, P(0))\end{aligned}$$

Or

$$\hat{J}[U](u) = (\nabla J_U, u)$$

D'où

$$\nabla J_U = -P(0)$$

Remarque : Le gradient se calcule dont en une intégration rétrograde du modèle adjoint. Ceci est à comparer avec le nombre gigantesque d'intégrations du modèle direct requis pour calculer le gradient par la méthode des taux d'accroissement !

3.6 Equation de Burgers

Nous allons dans cette section introduire le calcul adjoint pour le gradient d'une fonctionnelle coût afin de minimiser l'écart entre une solution de l'équation de Burgers et des observations. Nous allons nous intéresser au début à la minimisation par rapport à la condition initiale sous contrainte de moyenne nulle, puis à la minimisation par rapport à la condition initiale et aux conditions aux limites.

3.6.1 Contrôle de condition initiale sous contrainte de moyenne nulle

On considère l'équation de Burgers en dimension 1 :

$$\begin{aligned}\partial_t u + \frac{1}{2} \partial_x (u^2) - \nu \partial_{xx} u &= f, \quad x \in]0, L[, t \in [0, T], \\ u(x, 0) &= u_0(x), x \in]0, L[\\ u(0, t) = u(L, t) &= 0, t \in [0, T]\end{aligned}$$

On souhaite contrôler la condition initiale u_0 en minimisant la fonction coût :

$$J(u_0) = \frac{1}{2} \int_0^T \int_0^L (u - u^{obs})^2 dx dt$$

sous la contrainte de valeur moyenne nulle :

$$\int_0^L u_0(x) dx = 0$$

Pour cela on écrit le Lagrangien :

$$\mathcal{L}(u_0, \mu) = \frac{1}{2} \int_0^T \int_0^L (u - u^{obs})^2 dx dt + \mu \int_0^L u_0(x) dx$$

Regardons maintenant les conditions de point selle :

$$\nabla_{u_0} \mathcal{L} = 0, \quad \nabla_{\mu} \mathcal{L} = 0$$

Dérivons d'abord par rapport à μ :

$$\frac{\mathcal{L}(u_0, \mu + \alpha \delta \mu) - \mathcal{L}(u_0, \mu)}{\alpha} = \delta \mu \int_0^L u_0(x) dx$$

ce qui donne pour la deuxième condition de point selle :

$$\nabla_{\mu} \mathcal{L} = \int_0^L u_0(x) dx$$

On retrouve bien la contrainte. Dérivons maintenant par rapport à u_0 :

$$\lim_{\alpha \rightarrow 0} \frac{\mathcal{L}(u_0 + \alpha \delta u_0, \mu) - \mathcal{L}(u_0, \mu)}{\alpha} = \int_0^T \int_0^L \hat{u} (u - u^{obs}) dx dt + \mu \int_0^L u_0(x) dx$$

où \hat{u} vérifie l'équation suivante :

$$\partial_t \hat{u} + \partial_x (u \hat{u}) - \nu \partial_{xx} \hat{u} = 0, \quad x \in]0, L[, \quad t \in [0, T]$$

$$\hat{u}(x, 0) = \delta u_0 \quad x \in]0, L[$$

$$\hat{u}(0, t) = \hat{u}(L, t) = 0, \quad t \in [0, T]$$

En intégrant par parties on trouve l'équation du modèle adjoint :

$$\begin{aligned} \int_0^T (\partial_t \hat{u} + \partial_x (u \hat{u}) - \nu \partial_{xx} \hat{u}; p) &= \int_0^T \int_0^L (\partial_t \hat{u} + \partial_x (u \hat{u}) - \nu \partial_{xx} \hat{u}) p \\ &= \int_0^T \int_0^L -\partial_t p \hat{u} + \int_0^L \hat{u}(T, x) p(T, x) - \delta u_0 p(0, x) \\ &+ \int_0^T \int_0^L -u \hat{u} \partial_x p - \nu \left(\int_0^T \int_0^L \hat{u} \partial_{xx} p \right. \\ &\left. + \int_0^T \partial_x u(L, t) p(L, t) - \partial_x u(0, t) p(0, t) \right) \end{aligned}$$

ce qui nous amène à poser comme équation du modèle adjoint :

$$\begin{aligned}\partial_t p + u \partial_x p + \nu \partial_{xx} p &= u - u^{obs}, \quad x \in]0, L[, t \in [0, T] \\ p(x, T) &= 0, \quad x \in]0, L[\\ p(0, t) = p(L, t) &= 0, t \in [0, T]\end{aligned}$$

Grâce à cet adjoint, on peut réécrire le gradient du Lagrangien par rapport à u_0 :

$$\begin{aligned}\lim_{\alpha \rightarrow 0} \frac{\mathcal{L}(u_0 + \alpha \delta u_0, \mu) - \mathcal{L}(u_0, \mu)}{\alpha} &= \int_0^T \int_0^L \hat{u}(u - u^{obs}) dx dt + \mu \int_0^L u_0(x) dx \\ &= \int_0^T \int_0^L (\partial_t p + u \partial_x p + \nu \partial_{xx} p) \hat{u} dx dt + \mu \int_0^L \delta u_0 dx \\ &= - \int_0^L \delta u_0(x) p(x, 0) dx + \mu \int_0^L \delta u_0(x) dx \\ &= (\nabla \mathcal{L}_{u_0}; \delta u_0)\end{aligned}$$

D'où

$$\nabla \mathcal{L}_{u_0} = -p(\cdot, 0) + \mu$$

3.6.2 Contrôle de conditions initiale, et aux limites

Dans ce partie, on va traiter l'exemple de l'équation de Burgers avec contrôle de la condition initiale et des conditions aux limites. Soient L et T des réels strictement positifs. On considère l'équation de Burgers visqueuse suivante, pour $x \in]0, L[$ et $t \in]0, T[$:

$$\begin{cases} \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} - \nu \frac{\partial^2 u}{\partial x^2} = f \\ u(x = 0, t) = \psi_1(t) \\ u(x = L, t) = \psi_2(t) \\ u(x, t = 0) = u_0(x) \end{cases}$$

Le vecteur de contrôle est (u_0, ψ_1, ψ_2) , et la fonction coût est donnée par

$$J(u_0, \psi_1, \psi_2) = \frac{1}{2} \int_0^T \int_0^L (u - u_{\text{obs}})^2$$

où u est la solution de l'équation de Burgers ci-dessus associées aux conditions (u_0, ψ_1, ψ_2) . On sait que la dérivée de J dans la direction de (h_u, h_1, h_2) est

donnée par

$$\hat{J}(u_0, \psi_1, \psi_2)[h_u, h_1, h_2] = \int_0^T \int_0^L \hat{u}(u - u_{\text{obs}})$$

où \hat{u} est défini par

$$\hat{u} = \lim_{\alpha \rightarrow 0} \frac{u(u_0 + \alpha h_u, \psi_1 + \alpha h_1, \psi_2 + \alpha h_2) - u(u_0, \psi_1, \psi_2)}{\alpha}$$

et \hat{u} est solution de l'équation aux dérivées partielles du modèle linéaire tangent :

$$\begin{cases} \frac{\partial \hat{u}}{\partial t} + \frac{\partial(u\hat{u})}{\partial x} - \nu \frac{\partial^2 \hat{u}}{\partial x^2} = 0 \\ \hat{u}(x=0, t) = h_1(t) \\ \hat{u}(x=L, t) = h_2(t) \\ \hat{u}(x, t=0) = h_u(x) \end{cases}$$

Cherchons maintenant l'équation du modèle adjoint. Pour cela, on multiplie l'équation du modèle tangent par p et on intègre par parties sur $[0, T]$:

$$\begin{aligned} \int_0^T \left(\frac{\partial \hat{u}}{\partial t}, p \right) &= \int_0^T \int_0^L \frac{\partial \hat{u}}{\partial t} p \\ &= \int_0^L [\hat{u}p]_0^T - \int_0^L \int_0^T \frac{\partial p}{\partial t} \hat{u} \\ &= \int_0^L (\hat{u}(T)p(T, x) - h_u p(0, x)) - \int_0^L \int_0^T \frac{\partial p}{\partial t} \hat{u} \\ \int_0^T \left(\frac{\partial(u\hat{u})}{\partial x}, p \right) &= \int_0^T \int_0^L \frac{\partial(u\hat{u})}{\partial x} p \\ &= \int_0^T [u\hat{u}p]_0^L - \int_0^T \int_0^L u\hat{u} \frac{\partial p}{\partial x} \\ &= \int_0^T (\psi_2 h_2 p(L, t) - \psi_1 h_1 p(0, t)) - \int_0^T \int_0^L u\hat{u} \frac{\partial p}{\partial x} \end{aligned}$$

$$\begin{aligned}
\int_0^T \left(\frac{\partial^2 \hat{u}}{\partial x^2}, p \right) &= \int_0^T \int_0^L \frac{\partial^2 \hat{u}}{\partial x^2} p \\
&= \int_0^T \left[p \frac{\partial \hat{u}}{\partial x} \right]_0^L - \int_0^T \int_0^L \frac{\partial \hat{u}}{\partial x} \frac{\partial p}{\partial x} \\
&= \int_0^T \left[p \frac{\partial \hat{u}}{\partial x} - \hat{u} \frac{\partial p}{\partial x} \right]_0^L + \int_0^T \int_0^L \hat{u} \frac{\partial^2 p}{\partial x^2} \\
&= \int_0^T \left(p(L, t) \frac{\partial \hat{u}}{\partial x}(L, t) - h_2 \frac{\partial p}{\partial x}(L, t) - p(0, t) \frac{\partial \hat{u}}{\partial x}(0, t) + h_1 \frac{\partial p}{\partial x}(0, t) \right) \\
&\quad + \int_0^T \int_0^L \hat{u} \frac{\partial^2 p}{\partial x^2}
\end{aligned}$$

Les conditions aux limites naturelles pour p sont alors :

$$p(x, t = T) = 0, \quad p(x = 0, t) = p(x = L, t) = 0$$

Et on obtient ainsi

$$\begin{aligned}
0 &= \int_0^T \int_0^L \left(\frac{\partial \hat{u}}{\partial t} + \frac{\partial(u\hat{u})}{\partial x} - \nu \frac{\partial^2 \hat{u}}{\partial x^2} \right) p \\
&= \int_0^L \int_0^T \hat{u} \left(-\frac{\partial p}{\partial t} - u \frac{\partial p}{\partial x} - \nu \frac{\partial^2 p}{\partial x^2} \right) \\
&\quad + \int_0^L -h_u p(0, x) + \int_0^T \nu h_2 \frac{\partial p}{\partial x}(L, t) - \nu h_1 \frac{\partial p}{\partial x}(0, t)
\end{aligned}$$

autrement dit

$$\int_0^T \int_0^L \hat{u} \left(\frac{\partial p}{\partial t} + u \frac{\partial p}{\partial x} + \nu \frac{\partial^2 p}{\partial x^2} \right) = - \int_0^L h_u p(t = 0) + \int_0^T \nu h_2 \frac{\partial p}{\partial x}(x = L) - \nu h_1 \frac{\partial p}{\partial x}(x = 0)$$

Posons donc pour le modèle adjoint :

$$\begin{cases} \frac{\partial p}{\partial t} + u \frac{\partial p}{\partial x} + \nu \frac{\partial^2 p}{\partial x^2} = u - u_{\text{obs}} \\ p(x = 0, t) = 0 \\ p(x = L, t) = 0 \\ p(x, t = T) = 0 \end{cases}$$

On peut aussi réécrire le gradient de J sous la forme :

$$\hat{J}(u_0, \varphi_1, \varphi_2)[h_u, h_1, h_2] = - \int_0^L h_u p(0) - \mu \int_0^T h_2 \frac{\partial p}{\partial x}(L) - h_1 \frac{\partial p}{\partial x}(0).$$

Ce qui donne :

$$\nabla_{u_0} J = -p(x, t = 0), \quad \nabla_{\varphi_1} J = -\mu \frac{\partial p}{\partial x}(x = 0, t), \quad \nabla_{\varphi_2} J = -\mu \frac{\partial p}{\partial x}(x = L, t).$$

Chapitre 4

Application numérique à l'identification de paramètres d'EDPs

4.1 Introduction

Dans ce chapitre, nous allons nous intéresser à l'identification du coefficient de diffusion dans les EDPs paraboliques. Nous allons utiliser le langage MATLAB pour les simulations numériques. Tout d'abord nous allons introduire le calcul formel pour le modèle d'EDP que nous avons choisi, pour la fonctionnelle coût, ainsi que pour le modèle adjoint qui sera utilisé pour obtenir l'expression du gradient de la fonctionnelle coût. Ensuite, nous allons présenter la simulation numérique de la solution u^{obs} d'une EDP de référence qui va être utilisée comme données observées. Nous présentons les solutions numériques de l'EDP modélisant le problème adjoint et l'EDP à qui nous voulons identifier le paramètre de diffusion par une méthode des éléments finis pour une discrétisation en espace et une méthode d'Euler implicite pour une discrétisation en temps. Ensuite nous calculons la solution du problème de minimisation par une méthode de descente de gradient. Nous donnons à la fin le code MATLAB utilisé dans ce chapitre.

4.2 Dérivation du modèle

On considère le modèle d'EDP parabolique suivant qui a pour solution la fonction u^{obs} :

$$\begin{cases} \partial_t u^{obs} - \operatorname{div}(D_1 \nabla u^{obs}) + c_1 u^{obs} = f_1 & \text{dans } \Omega \times [0, T] \\ u^{obs} = 0 & \text{sur } \partial\Omega \times [0, T], \\ u^{obs} = u_0^{obs} & \text{sur } \Omega \times \{t = 0\}, \end{cases} \quad (4.1)$$

Nous avons choisi des conditions homogènes de Dirichlet sur le bord du domaine de calcul Ω . Les coefficients D_1 , c_1 et la fonctions f_1 et la condition initiale u_0^{obs} sont donnés. La solution u^{obs} sera utilisée comme données observées.

On introduit aussi le modèle d'EDP calculant la solution u :

$$\begin{cases} \partial_t u - \operatorname{div}(D \nabla u) + cu = f & \text{dans } \Omega \times [0, T] \\ u = 0 & \text{sur } \partial\Omega \times [0, T], \\ u = u_0 & \text{sur } \Omega \times \{t = 0\}, \end{cases} \quad (4.2)$$

Les coefficients D , c et la fonctions f et la condition initiale u_0 sont donnés. On se donne maintenant la fonctionnelle coût J suivante :

$$J(D) = \frac{1}{2} \int_0^T \int_{\Omega} (u(D) - u^{obs})^2 dx dt$$

où le paramètre à contrôler ici est le coefficient de diffusion D . La fonctionnelle J dépend indirectement du paramètre D via la solution u .

Le but est de chercher le D qui minimise J . Calculons la dérivée de J par rapport à D dans la direction d :

$$J(D+\alpha d) - J(D) = \frac{1}{2} \int_0^T \int_{\Omega} (u(D+\alpha d) - u^{obs})^2 dx dt - \frac{1}{2} \int_0^T \int_{\Omega} (u(D) - u^{obs})^2 dx dt.$$

En divisant par α et en passant à la limite quand $\alpha \rightarrow 0$, on obtient la dérivée directionnelle de J :

$$J[D](d) = \int_0^T \int_{\Omega} (u - u^{obs}) \hat{u} dx dt. \quad (4.3)$$

avec

$$\hat{u} = \lim_{\alpha \rightarrow 0} \frac{u(D + \alpha d) - u(D)}{\alpha}.$$

La solution \hat{u} vérifie le modèle linéaire tangent suivant :

$$\begin{cases} \partial_t \hat{u} - \operatorname{div}(D\nabla \hat{u}) + c\hat{u} = \operatorname{div}(d\nabla u) \text{ dans } \Omega \times [0, T], \\ \hat{u} = 0 \text{ sur } \partial\Omega \times [0, T] \\ \hat{u} = 0 \text{ sur } \Omega \times \{0\}. \end{cases} \quad (4.4)$$

On note par L l'opérateur

$$L(\hat{u}) = \partial_t \hat{u} - \operatorname{div}(D\nabla \hat{u}) + c\hat{u}.$$

On multiplie (4.4) par une fonction p puis on intègre, pour obtenir :

$$\int_0^T \int_{\Omega} (\partial_t \hat{u})p - \operatorname{div}(D\nabla \hat{u})p + c\hat{u}p = \int_0^T \int_{\Omega} \operatorname{div}(d\nabla u)p$$

Ce qui donne :

$$\int_{\Omega} [p\hat{u}]_0^T dx - \int_0^T \int_{\Omega} (\partial_t p)\hat{u} dx dt - \int_0^T \int_{\Omega} \operatorname{div}(D\nabla \hat{u})p + \int_0^T \int_{\Omega} c\hat{u}p = \int_0^T \int_{\Omega} \operatorname{div}(d\nabla u)p$$

En posant $p(x, T) = 0$ sur Ω et $p(x, t) = 0$ sur $\partial\Omega \times [0, T]$, on obtient :

$$- \int_0^T \int_{\Omega} (\partial_t p)\hat{u} dx dt - \int_0^T \int_{\Omega} \operatorname{div}(D\nabla \hat{u})p + \int_0^T \int_{\Omega} c\hat{u}p = \int_0^T \int_{\Omega} \operatorname{div}(d\nabla u)p$$

Ainsi :

$$\begin{aligned} \int_0^T \int_{\Omega} L(\hat{u})p dx dt &= \int_0^T \int_{\Omega} [-\partial_t p - \operatorname{div}(D\nabla p) + cp]\hat{u} dx dt \\ &= \int_0^T \int_{\Omega} \operatorname{div}(d\nabla u)p dx dt \\ &= \int_0^T \int_{\Omega} L^*(p)\hat{u} dx dt, \end{aligned}$$

avec L^* est l'opérateur adjoint de L donné par :

$$L^*(p) = -\partial_t p - \operatorname{div}(D\nabla p) + cp$$

Par identification avec l'expression du gradient (4.3), on déduit que

$$L^*(p) = (u - u^{obs}).$$

Alors le modèle adjoint s'écrit :

$$\begin{cases} -\partial_t p - \operatorname{div}(D\nabla p) + cp = u - u^{obs} \text{ dans } \Omega \times [0, T], \\ p = 0 \text{ sur } \partial\Omega \times [0, T] \\ p = 0 \text{ sur } \Omega \times \{T\}. \end{cases} \quad (4.5)$$

On obtient pour la dérivée de J :

$$J[D](d) = \int_0^T \int_{\Omega} \operatorname{div}(d\nabla u) p dx dt.$$

Ce qui donne pour le gradient de J :

$$\nabla J(D) = \int_0^T \operatorname{div}(\nabla u) p dt.$$

Remarquons que le modèle adjoint (4.5) est rétrogradée, on introduit le changement de variable $\tilde{p}(x, t) = p(x, T - t)$, il résulte que la fonction \tilde{p} vérifie le problème suivant avec conditions initiales :

$$\begin{cases} \partial_t \tilde{p} - \operatorname{div}(D\nabla \tilde{p}) + c(x, T - t)\tilde{p} = u(x, T - t) - u^{obs}(x, T - t) \text{ dans } \Omega \times [0, T], \\ \tilde{p} = 0 \text{ sur } \partial\Omega \times [0, T] \\ \tilde{p} = 0 \text{ sur } \Omega \times \{0\}. \end{cases} \quad (4.6)$$

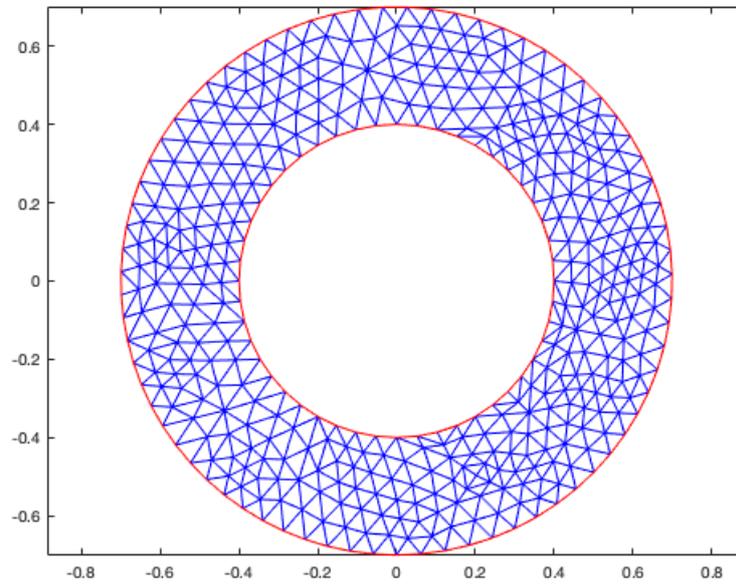
Finalement le coefficient de diffusion estimé est calculé numériquement à l'aide d'une méthode de descente de gradient :

$$D_{k+1} = D_k - \alpha \nabla J(D_k)$$

où k représente les itérations successifs et α est le pas de la méthode du gradient. Un critère d'arrêt sous forme $\|\nabla J(D_k)\|_2 < tol$ est utilisé dans les calculs numériques de la section suivante.

4.3 Résultats numériques

Nous allons présenter dans cette section les résultats numériques du modèle présenté dans la section précédente. On commence tout d'abord par la

FIGURE 4.1 – Maillage du domaine de calcul Ω .

spécification du domaine de calcul Ω qui a été choisi sous forme de couronne et discrétisé à l'aide d'un maillage triangulaire d'éléments finis. Il est présenté dans la figure 4.1.

Dans les simulations numériques, nous avons pris les valeurs suivantes pour les coefficients des équations : $u_0(x) = u_0^{obs}(x) = 1$ pour $x \in \Omega$, $c_1 = 0$, $f_1 = 0$, $D_1 = 5mm^2/s$, $D = 3.5mm^2/s$, $f = 0$, $c = 0$, $T = 5ms$.

On présente la solution u^{obs} à des instants différents dans la figure 4.2

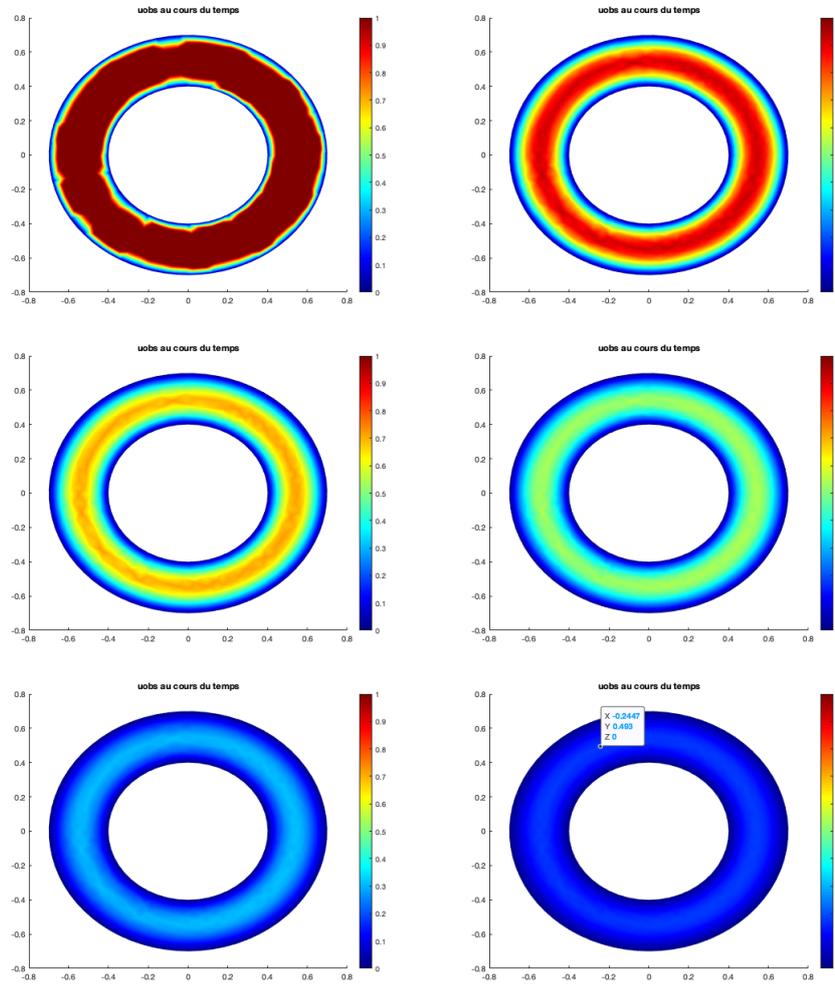


FIGURE 4.2 – La solution u^{obs} aux instants : 0ms, 0.5ms, 1ms, 1.5ms, 2.5ms, 3.5ms. Lecture de gauche à droite, de haut en bas

La solution u est présenté dans la figure 4.3 aux mêmes instants que u^{obs} .

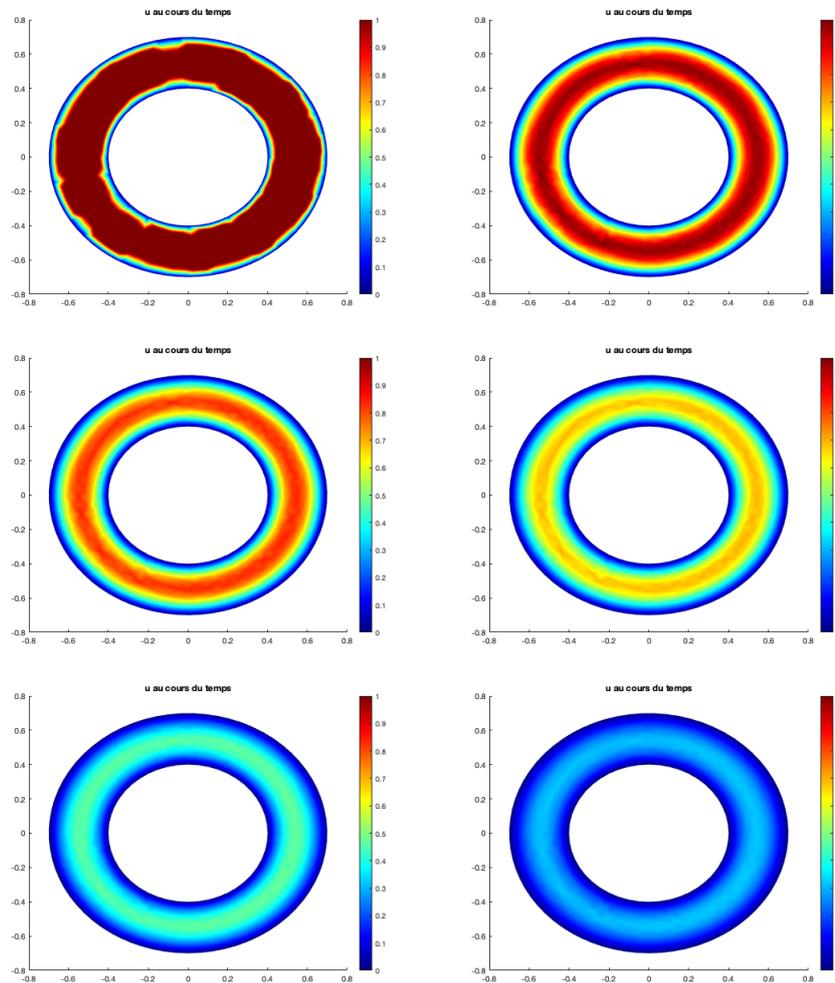


FIGURE 4.3 – La solution u aux instants : 0ms, 0.5ms, 1ms, 1.5ms, 2.5ms, 3.5ms. Lecture de gauche à droite, de haut en bas

La figure 4.4 donne les valeurs de la fonctionnelle coût J au cours des itérations.

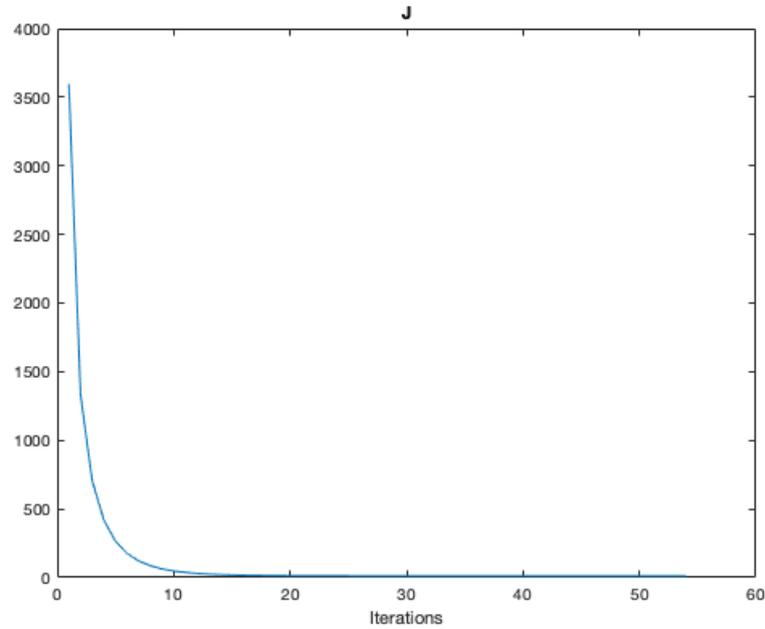


FIGURE 4.4 – La fonctionnelle J au cours des itérations.

Dans la figure 4.5 on présente les valeurs du gradient de la fonctionnelle coût en norme au cours des itérations. Le critère d'arrêt $\|\nabla J\|_2 < tol$ est utilisé dans les calculs. tol est une valeur de tolérance qui a été fixée à 3.

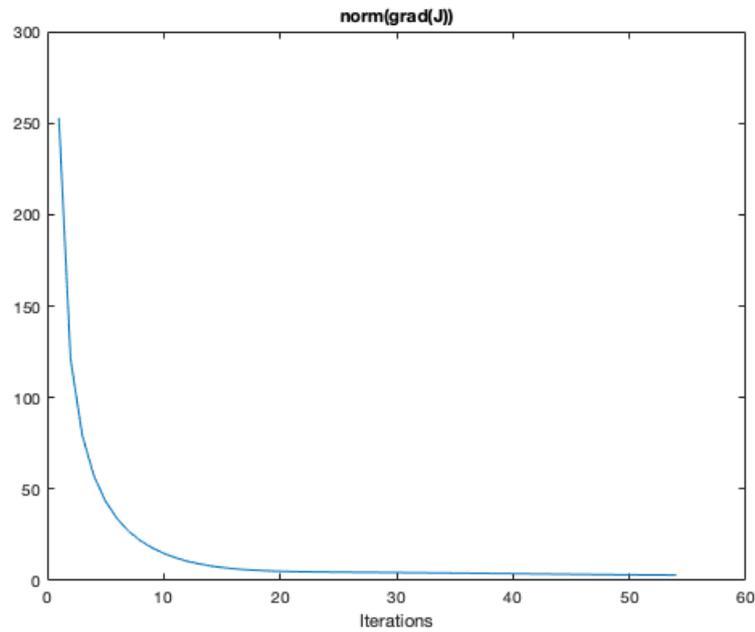


FIGURE 4.5 – Norme du gradient de la fonctionnelle J au cours des itérations.

La figure 4.6 donne les valeurs du coefficient de diffusion D estimé au cours des itérations. On voit qu'au fur et à mesure des itérations, le coefficient D qui était fixé au début à $3.5mm^2/s$, s'approche de la valeur de $5mm^2/s$ que l'on devrait obtenir.

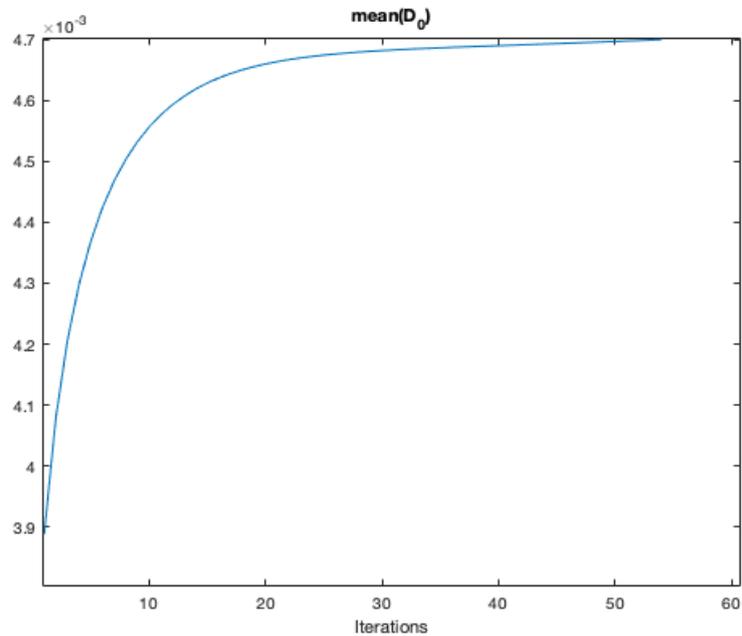


FIGURE 4.6 – Les valeurs du coefficient de diffusion estimées au cours des itérations.

4.4 Code MatLab

Le code MatLab que nous avons utilisé dans ce chapitre est donné par :

```
1 clear
2 clc
3 tic
4 % Load mesh points ans tirangles
5 load('p');
6 load('e');
7 load('b');
8 load('t');
9
10 nt=size(t,2);
```

```
11 np=size(p,2);
12 dt=0.005;
13 time=0:dt:5;
14 % Initial data
15 u0=ones(np,1);
16 D0 = 0.0035*ones(1,nt);
17 % Controled parameter
18 D = 0.005;
19 % Observed data
20 c=0;
21 uobs=parabolic(u0,time,b,p,e,t,D,c,0,1);
22
23 tol=3;
24 norme=280;
25 n=0;
26 while norme>tol
27 u=parabolic(u0,time,b,p,e,t,D0,c,0,1);
28 J=sum(sum((u-uobs).^2,1));
29 n=n+1;
30 VJ(n)=J;
31 %Adjoint model :
32 U=pdeintrp(p,t,u-uobs);
33 [K1,M1,F1,Q,G,H,R]=asempde(b,p,e,t,0,1,0);
34 for k=1:length(time)
35 [K,M,F(:,k),Q,G,H,R]=asempde(b,p,e,t,-D0,-c,U(k,:))
    ;
36 end
37 T=dt*K+M1+dt*M;
38 Nt=length(time);
39 % Retrograde solution
40 P(:,Nt)=zeros(1,np);
41 for i=2:Nt
42 P(:,Nt+1-i)=M1\((T*P(:,Nt+2-i))-dt*F(:,Nt+2-i));
43 end
```

```
44 | % Cost function J and its gradient
45 | for k=1:Nt
46 | [ux(k,:),uy(k,:)]=pdegrad(p,t,u(:,k));
47 | [Px(k,:),Py(k,:)]=pdegrad(p,t,P(:,k));
48 | end
49 | L=Px.*ux+Py.*uy;
50 | for k=1:nt
51 | gradJ(:,k)=dt*trapz(L(:,k));
52 | end
53 | norme=norm(gradJ,2);
54 | NORM(n)=norme;
55 | alpha=0.5e-4;
56 | % find the controled parameter by a descent gradient
57 | D0=D0-alpha*gradJ;
58 | D0_mean(n)=mean(D0);
59 | end
60 | % Graphical presentation of the results
61 | figure, plot(time,u(300,:),time,uobs(300,:),time,P
    | (300,:)),
62 | legend('u','u_{OBS}','P')
63 | figure, plot(1:n,VJ),xlabel('Iterations'),title('J')
64 | figure, plot(1:n,NORM),xlabel('Iterations'),
65 | title('norm(grad(J))')
66 | figure, plot(1:n,D0_mean),xlabel('Iterations'),
67 | title('mean(D_0)')
68 |
69 | for k=1:50:length(time)
70 | figure,
71 | pdeplot(p,e,t,'xydata',uobs(:,k))
72 | caxis([0 1]);colorbar, colormap jet
73 | title ('uobs au cours du temps')
74 | V(:,k)=getframe(gcf);
75 | end
76 |
```

```
77  % Solution of the observed solution of PDE uobs
78
79  for k=1:50:length(time)
80  figure ,
81  pdeplot(p,e,t,'xydata', uobs(:,k))
82  caxis([0 1]);colorbar, colormap jet
83  title ('uobs au cours du temps')
84  V(:,k)=getframe(gcf);
85  end
86
87  % Solution of the PDE u
88
89  for k=1:50:length(time)
90  figure ,
91  pdeplot(p,e,t,'xydata', u(:,k))
92  caxis([0 1]);colorbar, colormap jet
93  title ('u au cours du temps')
94  V(:,k)=getframe(gcf);
95  end
96
97  % Solution of the adjoint model p
98  for k=1:50:length(time)
99  figure ,
100 pdeplot(p,e,t,'xydata', p(:,k))
101 caxis([0 1]);colorbar, colormap jet
102 title ('p au cours du temps')
103 V(:,k)=getframe(gcf);
104 end
105 toc
```

Conclusion générale

Dans ce mémoire nous nous sommes intéressés aux problèmes d'estimation ou d'identification de paramètres d'équations aux dérivées partielles (EDPs). Cela fait intervenir la minimisation de certaine fonction coût qui s'écrit en fonction de données observées ou déjà disponibles à l'avance et d'une solution d'un modèle s'écrivant sous forme d'EDP. Ce problème d'optimisation n'est pas simple à résoudre. Il mis en oeuvre les techniques de calcul adjoint pour calculer le gradient de la fonctionnelle en question. Pour que ce travail soit clair et compréhensible, nous avons présenté dans les deux premiers chapitres toutes les notions nécessaires, à savoir l'écriture générale des EDPs, des résultats d'existence et d'unicité, leurs formulations variationnelles qui a été utilisée dans l'approximation numérique par la méthode des éléments finis. Dans le troisième chapitre nous avons présenté le but de ce mémoire qui est l'application du calcul adjoint dans le contrôle de paramètres d'EDPs. Nous avons donné quelques exemples illustratifs. Ensuite dans le dernier chapitre, nous avons introduit une application numérique effectuée sous MatLab illustrant les résultats obtenus en s'intéressant à l'identification de paramètres d'EDPs paraboliques.

Bibliographie

- [1] L.C. Evans. *Partial Differential Equations. Graduate Studies in Mathematics*. Volume 19. American Mathematical Society. 1998.
- [2] G. Bachman, L. Narici, E. Beckenstein. *Fourier and wavelet analysis*. Universitext, Berlin, New York, Springer-Verlag, 2000.
- [3] W. Rudin. *Analyse réelle et complexe*. Masson. Paris. 1977.
- [4] Mikhail S. Agranovich, *Sobolev Spaces, Their Generalizations and Elliptic Problems in Smooth and Lipschitz Domains*. Springer, 2015.
- [5] J.-L. Lions, E. Magenes. *Problèmes aux limites non homogènes et applications*. V1, Dunod, Paris, 1968.
- [6] R.L. Courant, *Variational methods for the solution of problem of equilibrium and vibration*, Bulletin of the American Mathematical Society, 49 :1-23, 1943.
- [7] Ritz, W. : *The Ueber eine neue Methode zur Losung gewisser Variations probleme der mathematischen Physik*. , Addison-Wesley Professional. J. Reine Angew. Math., 135 :1–61, 1908.
- [8] Strang, G. et G. Fix : *An analysis of the finite element method*. , Prentice-Hall, Englewood Cliffs, 1973.
- [9] Turner, M.J., R.W. Clough, H.C. Martin et L.C. Topp : *Stiffness and deflection analysis of complex structures*. Journal of the Aeronautical Sciences, 23 :805-882, 1956.
- [10] A. Ern. *Éléments finis*, Dunod, 2005.
- [11] P.A. Rariart, J.M.Thomas. *Introduction à l'analyse numérique des équations aux dérivées partielles*, Masson, Paris,1983.

-
- [12] J. Tromp, C. Tape, Q. Liu, *Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels*, Geophys. J. Int., 2005.
- [13] Y. Cao, S. Li, L. Petzold, R. Serban, *Adjoint sensitivity analysis for differential algebraic equations : The adjoint DAE system and its numerical solution*, SIAM, J. Sci. Comput. 2003.
- [14] Y. Jarny, M.N. Ozisik, J.P. Bardon, J.P, *A general optimization method using adjoint equation for solving multidimensional inverse heat conduction*. International Journal of Heat and Mass Transfer, 34(11), 2911-2919, 1991.
- [15] J.L.Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, 1971.