

وزارة التعليم العالي و البحث العلمي

Badji Mokhtar-Annaba University
Université Badji Mokhtar
Annaba



جامعة باجي مختار
عناية

Faculté des Sciences

Département de Mathématiques

THESE

Présentée en vue de l'obtention du diplôme de
DOCTORAT EN MATHÉMATIQUES

Méthodes d'approximation spectrale de Galerkin-Legendre pour un problème de commande optimale elliptique

Analyse numérique et estimation d'erreur a posteriori

Option : Analyse numérique

par

GHANEM RADOUEN

Sous la direction de
Professeur : **H.SISSAOUI**

Devant le jury

Président :	F. Zohra. NOURI	Prof.	Université Badji Mokhtar, Annaba.
Examineur :	A. BOUKHMIS	Prof.	Université Badji Mokhtar, Annaba.
Examineur :	M. HAIOUR	M.C.	Université Badji Mokhtar, Annaba.
Examineur :	M. Z. AISSAOUI	M.C.	Université 8 mai 1945, Guelma.

Année : 2007

Remerciements

Je voudrais commencer par remercier mon directeur, le professeur *H. SISSAOUI*, qui m'a proposé ce sujet et qui m'a encadré pendant cette thèse et durant cette période. Il m'a beaucoup apporté par son exigence de clarté et de rigueur ainsi que par son expérience.

Je suis également redevable à *G. TRONEL* Maître de conférence à l'université Paris 6 de m'avoir invité au laboratoire d'analyse numérique Jaque Louis Lions et de ses conseils très précieux.

Je tiens à remercier *F. Z. NOURI* Professeur à l'université Annaba d'avoir accepté de présider le jury de thèse, *M. HAIOUR* Maître de conférence à l'université de Annaba, *M. Z. AISSAOUI* Maître de conférence à l'université de Guelma et *A. BOUKHMIS* Professeur à l'université de Annaba, d'avoir accepté de faire partie de ce jury.

Une mention particulière va à ma petite famille (mes chers et tendres parents et ma soeur) et à mon adorable fiancée Wafa DEBASSI. Je veux leur témoigner toute gratitude et mon immense reconnaissance pour leur soutien constant et encouragements et d'avoir toujours crus en moi tout le long de cette période difficile. Je ne cesserais jamais de remercier les deux familles *BORSZTEIN* et *CHARBONNEL* pour leur chaleur familiale et gentillesse, et particulièrement à *Nicole* et *Pascal* qui ont été pour moi comme des parents pendant la totalité de mon séjour en France.

Enfin, et pour être sûr de n'oublier personne, je remercie également tous ceux qui m'ont apportés leur aide. Je tiens aussi à remercier toutes les autres personnes qui auraient pu m'aider et qui ne l'ont pas fait, c'est un peu grâce à eux que j'ai pu aller de l'avant.

Table des matières

Remerciements	i
Introduction	v
1 Quelques notions d'analyse fonctionnelle	1
1.1 Espaces de Sobolev	1
1.2 Théorème de représentation de Riesz	4
1.2.1 Notions sur les opérateurs adjoints	4
1.3 Minimisation des fonctions convexe	5
1.4 Commande optimale d'une équation elliptique linéaire	10
1.4.1 Introduction	10
1.4.2 Formulation du problème de commande optimale	11
2 Méthodes d'approximations spectrale	19
2.1 Introduction	19
2.1.1 Polynômes orthogonaux	20
2.1.2 Polynômes de Legendre	21
2.2 Techniques de détermination des coefficients d'approximation	24
2.3 Méthode d'approximation de Galerkin	25
2.4 Erreur d'approximation polynomiale en dimension une	28
2.5 Erreur D'approximation polynomiale en dimension deux	28
2.6 Formules de quadrature	30
3 Estimations de l'erreur dans un cadre abstrait	31
3.1 Introduction	31
3.2 Approximation de Galerkin dans un cadre abstrait	32
3.2.1 Analyse a priori de l'erreur	33
3.2.2 Analyse a posteriori de l'erreur	36
3.2.3 Indicateur par résidu	38
3.3 Représentation graphique de l'erreurs	39

4	Analyse numérique du problème de commande optimale (\mathcal{P})	41
4.1	Introduction	41
4.2	Approximation du problème de commande optimale (\mathcal{P})	42
4.2.1	Approximation numérique de l'équation d'état	43
4.2.2	Approximation numérique de la fonction objectif	46
	Résultats numériques	47
4.3	Estimation d'Erreur a Posteriori	49
4.3.1	Estimation d'Erreur sur l'état et la commande	49
4.3.2	Estimation d'Erreur sur l'état, le contrôle, la fonction objectif	53
4.3.3	Evaluations des estimateurs d'erreurs	64
	Résultats numériques	66
4.4	Conclusion et perspectives	68
A	Autres techniques pour le calcul des conditions d'optimalité	71
A.1	Lagrangien et multiplicateur de Lagrange	71
A.2	Principe du maximum de Pontryagin	71
B	Calcul matriciel des nœuds ξ_j	73
B.1	Calcul des nœuds ξ_j	73
C	Estimation de l'erreur a posteriori par dualité	75
C.1	Introduction	75
C.1.1	Estimation d'erreur a posteriori dans un context général abstrait	76

Introduction

La théorie de la commande optimale est l'un des domaines de recherche les plus importants dans la théorie du contrôle. Son origine remontée après 1696, lorsque Johann Bernoulli a posé et a résolu le problème du brachystochrone [76]. A propos, ce problème est à présent vue par les théoriciens du contrôle comme un problème de commande optimale en temps minimum. Les racines de la théorie de la commande optimale se trouvent dans le domaine mathématique du calcul de variations. Les plus importantes contributions, ont été faites par Bernoulli, Newton, Euler et Lagrange et un des résultats majeur de la théorie du contrôle optimal est le célèbre Principe du Maximum du Pontryagin [70].

Une grande impulsion au développement de la théorie de la commande optimale est donnée vers 1950, en parallèle avec l'introduction des calculateurs numériques et le commencement d'explorations spatiales et l'une des grandes applications de la commande optimale a été l'application au lanceur Appolo dans les années 1960 [76]. La commande optimale reste donc un sujet de recherche d'actualité.

On commence par une introduction concernant les problèmes de commande optimale gouvernés par des équations différentielles ordinaires et en particulier leur résolution numérique. On considère l'évolution d'un système dynamique sur un intervalle de temps $I = [t_0, t_1]$. L'état du système à tout moment t dans I est décrit par une fonction x de I dans \mathbb{R}^n . On appellera la fonction $x(t)$ la trajectoire où l'état du système. La commande envoyée au système à tout moment de temps t sera représentée par une fonction u de I dans \mathbb{R}^m . Nous supposons que l'évolution du système sous l'effet de la commande est décrite par une équation différentielle ordinaire (nous notons dans la suite $\dot{x} = \frac{dx}{dt}$)

$$\dot{x} = f(t, x(t), u(t)), \quad t \in [t_0, t_1], \quad (1)$$

où, f est définie sur I à valeurs dans $\mathbb{R}^n \times \mathbb{R}^m$, qui est une fonction continûment dérivable. On associe à cette équation différentielle la connaissance des conditions initiales et/ou finales. On suppose que les fonctions de commande appartiennent à un espace de fonctions définies sur I donnée, qu'on note U . On définit également un ensemble G dans \mathbb{R}^n comme étant l'ensemble des états finaux ou des états cibles. On définit U_{ad} l'ensemble des commandes admissibles, l'ensemble des commandes qui peuvent permettre d'atteindre un état final (défini à l'avance) de l'ensemble G . On doit définir aussi, une fonction objectif, appelée également fonction coût, qui est une fonctionnelle définie sur la classe des commandes admissibles

$$J : U_{ad} \rightarrow \mathbb{R}. \quad (2)$$

Dans ce cas, le problème de commande optimale, peut être formulé comme suit :

Etant donné le système dynamique (1) et la fonction objectif (2), trouver la commande optimale u qui minimise la fonction objectif (2), et où la trajectoire optimale x (l'état du système) est solution de l'équation (1).

C'est sous cette forme que la commande optimale a été étudiée dès le *XIX^{ème}* siècle avec le calcul des variations.

Les systèmes distribués sont les systèmes, où l'équation d'état est une équation aux dérivées partielles (avec des conditions aux limites et, dans le cas d'évolution, des conditions initiales sont nécessaires pour la détermination de la solution).

Commençons par rappeler le schéma général des problèmes de contrôle optimale des systèmes distribués qu'on va étudier dans ce travail. Le problème générique est le suivant :

$$\min \{ J(\mathbf{y}, \mathbf{u}), \mathcal{L}(\mathbf{y}, \mathbf{u}) = 0, \mathbf{u} \in U_{ad}, \mathbf{y} \in K \subset \mathcal{X} \},$$

où \mathcal{U} est un espace de Hilbert, \mathcal{X} est un espace de Banach, U_{ad} et K sont des sous-ensembles convexes, fermés non vides de \mathcal{U} et \mathcal{X} respectivement, J est une fonction coût de $\mathcal{X} \times \mathcal{U}$ dans $\mathbb{R} \cup +\infty$, et \mathcal{L} est un opérateur linéaire ou non stationnaire ou d'évolution.

Pour la suite, on suppose que $\mathcal{L}(\mathbf{y}, \mathbf{u}) = \mathcal{A}\mathbf{y}(\mathbf{u}) - \mathcal{B}\mathbf{u}$. On définit alors l'équation qu'on appellera équation d'état, écrite formellement sous la forme :

$$\mathcal{A}\mathbf{y}(\mathbf{u}) = \mathcal{B}\mathbf{u}, \quad (3)$$

où dans l'équation (3) \mathcal{A} désigne dans le cas général un opérateur aux dérivées partielles, linéaire ou non, stationnaire ou d'évolution et \mathcal{B} est un opérateur linéaire ou non, où à l'équation (3) il faut ajouter des conditions aux limites et, lorsque \mathcal{A} est un opérateur d'évolution, il faut ajouter des conditions initiales. Ces conditions (limites et / ou initiales), qu'il faut préciser pour chaque classe d'opérateur \mathcal{A} , ne sont pas explicitées dans cette partie car nous présentons les choses de façon formelle.

Dans l'équation (3) la fonction \mathbf{u} est le contrôle ou commande, variable qui peut apparaître de manière distribuée dans le domaine où l'on étudie le phénomène physique modélisé par l'équation d'état (3); la variable \mathbf{u} peut également apparaître de manière frontière, par l'intermédiaire des conditions aux limites. La variable \mathbf{u} peut également intervenir de manière géométrique; qui est le cas des problèmes d'optimisation de forme [19].

Dans la théorie du contrôle des systèmes distribués, il est généralement supposé que :

(H.1) Pour \mathbf{u} donné dans un ensemble convenable, l'équation (3) admet une solution unique $\mathbf{y}(\mathbf{u})$, elle aussi dans un ensemble convenable.

Nous précisons plus loin, sous certaines mesures, quels sont ces ensembles convenables. Pour le moment, nous désignons par $\mathbf{y}(\mathbf{u})$ la solution, ainsi entendue; de l'équation (3);

qui est l'état du système ; par suite on peut alors définir l'application \mathcal{T} (du contrôle vers l'état)¹

$$\mathcal{T} : \underset{\text{commande}}{\mathbf{u}} \rightarrow \underset{\text{état}}{\mathbf{y}(\mathbf{u})},$$

qui est l'application, qui associe à chaque commande \mathbf{u} l'état \mathbf{y} . Une fois qu'on a défini l'état \mathbf{y} , on introduit la fonction coût (fonction objectif) $J(\cdot)$, qui à tout \mathbf{u} de l'ensemble des commandes lui associe un nombre donné par

$$J(\mathbf{u}) = \phi(\mathbf{y}(\mathbf{u})) + \psi(\mathbf{u}), \quad (4)$$

où dans la fonction objectif (4), les fonctionnelles ϕ et ψ sont respectivement définies sur l'ensemble des états et sur l'ensemble des commandes et sont à valeurs réelles.

Dans la plupart des applications, la fonctionnelle ψ est une fonction d'une norme de \mathbf{u} ; elle définit en fait un espace de Banach U . Le choix de l'espace U dans lequel on va chercher les contrôles \mathbf{u} est donc (au moins pour l'instant) imposé par la fonction objectif (4). Une fois U est fixé, on sait où $\mathcal{B}\mathbf{u}$ va être pour \mathbf{u} dans U , ce qui fixe (à peu près) le cadre fonctionnel pour la résolution de l'équation d'état(3). On introduit ainsi un espace de Banach Y , où l'on cherche $\mathbf{y}(\mathbf{u})$, et l'hypothèse **(H.1)** s'écrit alors :

– Pour $\mathbf{u} \in U$, l'équation (3) admet une solution unique $\mathbf{y}(\mathbf{u}) \in Y$.

Une hypothèse souvent faite (mais avec de notables exceptions, comme on va voir ci-après est que, dans un sens à préciser :

(H.2) L'application $\mathcal{T} : \mathbf{u} \mapsto \mathbf{y}(\mathbf{u})$ de $U \rightarrow Y$ est différentiable

Cette hypothèse est complétée par :

– Les fonctionnelles $\mathbf{y} \mapsto \phi(\mathbf{y})$ et $\mathbf{u} \mapsto \psi(\mathbf{u})$ sont différentiables de $Y \rightarrow \mathbb{R}$ et de $U \rightarrow \mathbb{R}$

Le problème de contrôle optimal est alors de trouver

$$\inf J(\mathbf{u}),$$

lorsque \mathbf{u} parcourt U ou un sous ensemble U_{ad} de U ;(où l'ensemble U_{ad} exprime les contraintes sur \mathbf{u} (le cas où \mathbf{u} parcourt U tout entier est dit sans contraintes).

Les objectifs de la théorie de la commande optimale sont les suivants :

1. Etudier l'existence de $\mathbf{u} \in U_{ad}$ réalisant le minimum dans (4) ; on dit alors que \mathbf{u} est le contrôle optimal, $\mathbf{y}(\mathbf{u})$ est l'état optimal, dans ce cas $(\mathbf{y}(\mathbf{u}), \mathbf{u})$ est la solution optimale.
2. Donner des conditions nécessaires et si possible suffisantes, exprimant que \mathbf{u} est la commande optimale.
3. Obtenir des propriétés du (ou des) commande(s) optimale(s) à partir de la condition 2.
4. Etablir des algorithmes de résolution numériques.

¹en anglais on dit : control to state

Avant de passer à autre choses décrivons toujours d'une manière formelle, la structure des conditions nécessaires d'optimalités du problème de contrôle optimal décrit plus haut.

Soit donc u un contrôle optimal et soit $y(u) = y$ l'état optimal correspondant. Alors, il existe un triplet $\{u, y, p\}$ vérifiant le système d'optimalité (*S.O.*) :

$$\mathcal{A}(y) = \mathcal{B}(u), \quad (5a)$$

$$\mathcal{A}'(y)^* p = \phi'(y), \quad (5b)$$

$$(\mathcal{B}'(u)^* p + \psi'(u), v - u) \geq 0, \quad \forall v \in U_{ad}. \quad (5c)$$

Dans le système (*S.O.*), $\mathcal{A}'(y)^*$ respectivement $\mathcal{B}'(u)^*$ désigne l'opérateur adjoint de la dérivée de $\mathcal{B}(u)$, supposée exister, de \mathcal{A} respectivement de \mathcal{B} au point y respectivement au point u .

Grâce à **H.1**, l'état adjoint p est défini de manière unique par l'équation (5b). Dans (5c), on suppose généralement que U_{ad} est convexe ; l'inégalité (5c) est une inéquation variationnelle dans la terminologie introduite par [43]. Naturellement dans l'équation (5b), il faut préciser les conditions aux limites et/ou initiales, satisfaites par p , il faut aussi préciser tous les espaces fonctionnels dans lesquels on travaille. Tout repose alors sur l'application de diverses formules de Green, qu'il faudra justifier leur usage. On utilise systématiquement la théorie des problèmes non homogènes [57].

Sans prétendre d'avoir fait le tour de la théorie des problèmes de commande optimale, on peut dire que, dans les problèmes de commande optimale, il s'agit de trouver la commande minimisant un critère donné et l'étude de ce type de problèmes demeure toujours motivée par ces applications assez variées, telles que, par exemple, le contrôle des réactions enzymatiques [58], le contrôle du transport d'énergie électrique, le contrôle de la forme de plasmas, en mécanique, en physique, en aérospatiale, en économie et en sciences sociales,... ect.

Souvent ces problèmes ne sont pas faciles à résoudre car la solution exacte (i.e. explicite) de ces problèmes est parfois difficile voire impossible à déterminer. Pour cette raison, qu'on est obligé de faire appel aux méthodes numériques pour résoudre ce type de problèmes. Pour être plus précis, on fait appel aux méthodes d'approximations, pour discrétiser le problème continue de dimension infinie et le transformer en un problème discret en dimension finie qu'on pourra résoudre par les techniques du calcul scientifique et de l'analyse numérique.

Les problèmes de commande optimale peuvent être résolus généralement par deux techniques. La première dite indirecte (après optimisation) qui consiste à discrétiser les conditions d'optimalité et par la suite obtenir un système d'équations linéaires (ou non linéaires) dans la solution est souvent coûteuse et fournissant uniquement la solution optimale (commande, état) sans fournir une valeur approchée de la fonction objectif tandis que la deuxième approche dite directe (après optimisation) consiste à discrétiser le problème tel qu'il est posé, et par la suite obtenir un problème d'optimisation qu'on peut résoudre facilement par les algorithmes d'optimisation. Contrairement à la première

approche celle-ci fournit directement la valeur de la fonction objectif ainsi que la solution optimale (commande, état).

Depuis longtemps les chercheurs se sont intéressés aux différentes techniques d'approximations des équations différentielles ordinaire et aux équations aux dérivées partielles en général et aux problèmes de commande optimale en particulier. Ils ont utilisé les méthodes des différences finies (voir par exemple [67]) qui cherche à trouver les valeurs de la solution aux noeuds d'un maillage (donc à trouver des solutions ponctuelles). Mais cependant un problème immédiat s'est posé à savoir celui du calcul de la solution en dehors des noeuds du maillage. Bien plus tard les chercheurs se sont intéressés à l'utilisaiton de la méthode des éléments finis laquelle est une méthode variationnelle qui cherche une solution appartenant à des sous domaines comme combinaison linéaire de polynômes connus à l'avance sur chaque sous domaine [66]. Parfois, cette méthode ne converge pas rapidement et donne des matrices qui sont souvent des matrices creuses mais de grandes tailles. Cette méthode s'avère parfois coûteuse.

En outre, des difficultés numériques telle que l'accumulation des erreurs d'arrondis sont souvent rencontrées dans les méthodes des éléments finis et des différences finies ; d'où l'utilisation des méthodes spectrales pour l'approximation des modèles mathématiques. Les méthodes d'approximation spectrale (Tau, Galerkin, pseudo-spectrale ou collocation) sont des méthodes qui recherchent la solution comme une combinaison linéaire finie de polynômes de degré N . En effet, il est connu que, les polynômes orthogonaux possèdent des propriétés intéressantes d'approximation. Ils sont utilisés depuis longtemps dans l'approximation des équations aux dérivées partielles. L'avantage principal des méthodes spectrales réside dans leur convergence rapide et dans leur implémentation numérique relativement plus simple que les autres méthodes [79], surtout si on se place dans un rectangle.

Malgré leur énorme succès dans l'approximation et la résolution des équations aux dérivées partielles, les méthodes spectrales n'ont pas été beaucoup utilisées dans l'approximation des problèmes de commande optimale notamment en $2 - D$.

Les méthodes spectrales ont été introduites par [32], [33] et [86], en utilisant la méthode d'approximation de Tau pour l'approximation d'un problème de commande optimale gouverné par une équation différentielle ordinaire. Bien que les résultats numériques obtenus sont encourageants et satisfaisants, l'expression **méthode spectrale** pour ce travail semble inappropriée car l'auteur n'a utilisé que les polynômes de Tchebycheff de première espèce sans faire appel à leur fonction poids qui compliquera la mise en oeuvre numérique. En 2000, dans le cadre d'un travail de Magister [41], nous avons appliqué cette technique pour l'approximation d'un problème de commande optimale gouverné par une équation parabolique [52] (toujours en $1 - D$) où l'on a considéré une semi discrétisation (par la méthode des différences finies) par rapport à la variable temporaire. Ceci a conduit à l'obtention d'un problème de commande optimale gouverné par un système d'équations différentielles ordinaires. Ensuite, on a conclu avec une discrétisation complète (par la méthode d'approximation spectrale de Tau) par rapport à la variable spatiale. Les résultats numériques obtenus ont été satisfaisants sur le plan du volume de calcul comparés avec les résultats obtenus par Abdeli [1] utilisant les éléments fini et par Huntly [52] utilisant les

différences finies. Nous avons également essayé de généraliser la méthode en l'appliquant à un problème de commande optimale elliptique en $2 - D$ toujours par l'utilisation des polynômes de Tchebycheff en se basant sur la méthode de Tau. Cependant, dans cette étude, les résultats numériques ont été beaucoup moins satisfaisants que les précédents. La question immédiate qui se pose alors est de savoir pourquoi?. La réponse est due aux conditions de régularités concernant la fonction contrôle u qui n'apparaissent pas directement dans la formulation du problème de départ et par la suite qui ne peuvent pas être imposées dans l'approximation numérique. D'où des perturbations et des instabilités numériques sur le bord qui en résultent.

Dans ce travail, on s'intéresse à l'application des méthodes spectrales d'approximation par la méthode de Galerkin en utilisant les polynômes de Legendre pour l'approximation du même problème déjà proposé dans [63]. Notre travail consiste à analyser numériquement le problème de commande optimale elliptique en se basant sur la formulation directe.

Dans l'analyse numérique du problème, il s'agit particulièrement d'évaluer l'erreur de calcul réalisée durant l'approximation qu'on note habituellement par \mathbf{e} . Cette erreur n'est autre en réalité que la différence entre la solution exacte et la solution approchée du problème dans une norme appropriée. Généralement, il y a deux types d'estimations de l'erreur : l'estimation de l'erreur à priori qui nécessite la connaissance de la solution exacte du problème à résoudre numériquement ainsi que les paramètres d'approximations ; et l'estimation de l'erreur à posteriori qui nécessite la connaissance des paramètres d'approximations ainsi que la solution approchée mais par contre ne nécessitant pas forcément la connaissance de la solution exacte du problème à résoudre.

Comme il a été signalé plus haut, on est souvent confronté à des problèmes qu'on ne peut pas résoudre explicitement. Ainsi donc on est incapable de calculer leur solutions exactes. Donc pour évaluer l'erreur commise durant l'approximation, on doit trouver une estimation à posteriori de l'erreur. Pour être plus clair : supposons qu'on s'intéresse à la résolution numérique d'un problème quelconque donné admettant une solution u (qu'on ne connaît pas forcément !). Soit u_σ , l'approximation de la solution u (la solution numérique), par suite, l'estimation à priori de l'erreur $\|\mathbf{e}\| = \|u - u_\sigma\|$ est donnée par l'inégalité générale suivante

$$\|\mathbf{e}\| \leq C\eta(\sigma, u).$$

Par opposition, l'estimation à posteriori de l'erreur, peut être donnée par l'inégalité générale suivante

$$\|\mathbf{e}\| \leq C\eta(\sigma, u_\sigma),$$

où, C est une constante indépendante du paramètre d'approximation σ .

En général, dans la détermination de l'estimation de l'erreur à posteriori, on utilise deux méthodes différentes mais équivalentes soit la méthode du résidu [5], soit la méthode de dualité [48]. La méthode du résidu, a été beaucoup utilisée, particulièrement dans les problèmes modélisés par des équations différentielles (par exemple, des équations aux

dérivée partielles de type elliptique) [5]. Elle repose essentiellement sur l'estimation de la différence entre le problème continu et le problème discret (Le résidu) qui engendre par la suite l'estimation de $\|\mathbf{e}\|$. La méthode de dualité est généralement utilisée, dans les cas où l'estimation par résidu ne s'applique pas (par exemple, des inéquations variationnelles) et consiste à définir le problème dual du problème à étudier et par la suite évaluer $\|\mathbf{e}\|$. En ce qui nous concerne, on va appliquer la méthode du résidu, pour la simple et unique raison qu'on va avoir affaire à des équations aux dérivées partielles de type elliptique. Maintenant, on peut se poser des questions sur l'intérêt de l'analyse a posteriori de l'erreur. Cette dernière nous aide à adapter la méthode d'approximation afin d'obtenir la précision de calcul voulue dans tel ou tel endroit du domaine ou tout simplement de réduire la taille et le temps de calcul.

Dans ces dernières années, les méthodes d'approximations adaptatives par éléments finis ont eu un grand succès dans l'analyse à posteriori de l'erreur concernant plusieurs problèmes (problèmes de contact, optimisation de forme,...) provenant de domaines variés (mécanique, physique,...) [39]. En particulier, pour les problèmes de commande optimale [47] où l'estimation de l'erreur est donnée de différentes manières (locale ou globale), et sous différentes formes (voir par exemple [60]), en estimant, par exemple, juste $\|\mathbf{e}\|$ (voir par exemple [59]), ou en donnant une estimation du critère du gradient² (voir par exemple [89]), ou en utilisant le lagrangien et les multiplicateurs de Lagrange [72].

Dans ce travail, on étudiera l'analyse à posteriori de l'erreur globale de la méthode d'approximation spectrale du problème de commande optimale elliptique en utilisant, bien sûr, la méthode du résidu. On obtiendra alors une estimation de l'erreur qui dépend des solutions discrètes et du seul paramètre d'approximation N le degré du polynôme de Legendre.

Le mémoire se présente comme suit. Dans le premier chapitre on présente une introduction générale sur les outils théoriques nécessaires pour la compréhension des chapitres ultérieurs. Dans le chapitre 2, on construit la méthode d'approximation avec des résultats théoriques originaux nécessaires à l'approximation du problème de commande optimale. Le chapitre 3 contient, d'une part, une définition abstraite et générale concernant les estimations d'erreurs en générale et les estimations à posteriori en particulier et, d'autre part introduit les différentes méthodes utilisées pour cette étude. Enfin, le chapitre 4 est consacré à l'analyse numérique d'un problème de commande optimale elliptique. Des résultats originaux concernant les estimations d'erreurs à posteriori sont obtenus. Le mémoire s'achève par une conclusion, des perspectives, des appendices et des références.

ملخص

في هذا البحث نعطي الإرتياب الملحق بالراسب المرتكب بالتقريب الطيفي و هذا في تقريب مسألة التحكم الفضلي الإهليجي بإستعمال طريقة غلركين لجندر.

²qui est une estimation de $\nabla u - \nabla u_\sigma$

Abstract

In this study, a residual a posteriori error estimates for the Galerkin-Legendre spectral approximation of an elliptic control problem.

Résumer

Dans cette étude, on donne les estimations de l'erreur à posteriori par résidu pour l'approximation par la méthode de Galerkin-Legendre d'un problème de commande optimale elliptique.

Liste des publications parues ou acceptées

- (P1). R. GHANEM ET H. SISSAOUI, *Residual a posteriori error estimates for an elliptic optimal control problem*, International Journal of Mathematics and Computer Science, **1**(2006), 227-242.
- (P2). R. GHANEM ET H. SISSAOUI, *A posteriori error estimate by a spectral method of an elliptic optimal control problem*, Journal of Computational Mathematics and Optimization, **2**,(2006), 111-125.
- (P3). R. GHANEM ET H. SISSAOUI, *On $h - N$ approximation of an optimal control problem governed by a linear parabolic equation*, Advances in Algebra and Analysis, **1**(3),(2006), 203-217.

Chapitre 1

Quelques notions d'analyse fonctionnelle

1.1 Espaces de Sobolev

On commence par introduire quelques notions de base et rappeler quelques définitions indispensables à la compréhension du manuscrit. On introduit par exemple les espaces dans lesquelles les solutions recherchées. Naturellement ce sont les espaces de Sobolev qui sont utilisés dans l'analyse numérique de ce problème [2].

Dans ce qui suit, on désigne par $\mathbf{x} = (x, y)$ le point générique de Ω qui est un ouvert borné de \mathbb{R}^d , où $d \geq 2$. Deux définitions sont nécessaires pour caractériser la géométrie des ouverts que l'on considère.

Définition 1.1.1. *On dit que la frontière $\partial\Omega$ de Ω est Lipschitzienne si, pour tout point \mathbf{x} de $\partial\Omega$, il existe un système de coordonnées orthogonales (y_1, \dots, y_d) centré en \mathbf{x} (c'est à dire tel que les coordonnées y_i de \mathbf{x} soient toutes égales à 0), un pavé $U^{\mathbf{x}} = \prod_{i=1}^d]-a_i, a_i[$ et une application lipschitzienne $\mathcal{F}^{\mathbf{x}}$ du sous-produit $\prod_{i=1}^{d-1}]-a_i, a_i[$ dans $] -\frac{a_d}{2}, \frac{a_d}{2} [$ tels que*

$$\Omega \cap U^{\mathbf{x}} = \{(y_1, \dots, y_d) \in U^{\mathbf{x}}; y_d > \mathcal{F}^{\mathbf{x}}(y_1, \dots, y_{d-1})\},$$

$$\partial\Omega \cap U^{\mathbf{x}} = \{(y_1, \dots, y_d) \in U^{\mathbf{x}}; y_d = \mathcal{F}^{\mathbf{x}}(y_1, \dots, y_{d-1})\}.$$

Cette propriété signifie que la frontière coïncide localement avec le graphe d'une fonction lipschitzienne.

Définition 1.1.2. *Soit m un entier positif, et soit Ω un ouvert borné à frontière $\partial\Omega$ lipschitzienne. Une partie ouverte Γ de cette frontière est dite de classe $\mathcal{C}^{m,1}$ si, pour tout point \mathbf{x} de Γ , l'application $\mathcal{F}^{\mathbf{x}}$ de la Définition 1.1.1 peut être choisie différentiable jusqu'à l'ordre m avec la différentielle d'ordre m lipschitzienne. Elle est dite de classe \mathcal{C}^∞ si elle est de classe $\mathcal{C}^{m,1}$ pour tout entier m positif. Pour la suite on désigne par Ω un ouvert borné convexe à frontière lipschitzienne de \mathbb{R}^d . On note par \mathbf{x} le point générique de Ω , et*

(x_1, \dots, x_d) ses coordonnées et on utilise la mesure de Lebesgue dans \mathbb{R}^d , que l'on écrit $d\mathbf{x}$.

On rappelle que $\mathcal{D}(\Omega)$ désigne l'espace des fonctions indéfiniment différentiables à support compact dans Ω , et que $\mathcal{D}(\overline{\Omega})$ désigne l'espace des restrictions sur $\overline{\Omega}$ des fonctions indéfiniment différentiables à support compact dans \mathbb{R}^d . Le dual $\mathcal{D}'(\Omega)$ de $\mathcal{D}(\Omega)$ est l'espace des distributions sur Ω . On introduit également l'espace $\mathcal{C}^0(\overline{\Omega})$ des fonctions continues sur $\overline{\Omega}$ [77].

Soit p tel que $1 \leq p \leq +\infty$. On note par $L^p(\Omega)$ l'espace des fonctions v de Ω dans \mathbb{R} mesurables telles que

$$\|v\|_{L^p(\Omega)} < +\infty,$$

où $\|v\|_{L^p(\Omega)}$ désigne la norme

$$\|v\|_{L^p(\Omega)} = \left(\int_{\Omega} |v(\mathbf{x})|^p d\mathbf{x} \right)^{\frac{1}{p}} \text{ si } p < +\infty \quad \text{et} \quad \|v\|_{L^\infty(\Omega)} = \sup_{\mathbf{x} \in \Omega} \text{ess } |v(\mathbf{x})|. \quad (1.1.1)$$

L'espace $L^p(\Omega)$ est un espace de Banach pour cette norme, et il est réflexif si et seulement si $1 < p < +\infty$. On sait également que

- Pour $1 \leq p < +\infty$, l'espace $L^p(\Omega)$ contient les deux espaces $\mathcal{D}(\Omega)$ et $\mathcal{D}(\overline{\Omega})$ comme sous-espaces denses,
- Pour $1 \leq p \leq +\infty$, l'espace $L^p(\Omega)$ est contenu dans l'espace $\mathcal{D}'(\Omega)$.

Dans le cas particulier $p = 2$, $L^p(\Omega)$ est un espace de Hilbert pour le produit scalaire

$$(u, v) \mapsto \int_{\Omega} u(\mathbf{x}) v(\mathbf{x}) d\mathbf{x}.$$

Le produit de dualité entre les espaces $\mathcal{D}(\Omega)$ et $\mathcal{D}'(\Omega)$ étant alors une extension du produit scalaire dans $L^2(\Omega)$ [77].

La théorie des distributions [77] permet de définir, pour les fonctions de $L^p(\Omega)$, des dérivées d'ordre quelconque à valeurs dans $\mathcal{D}'(\Omega)$: pour tout d -uplet $\alpha = (\alpha_1, \dots, \alpha_d)$ de \mathbb{N}^d , $|\alpha|$ représente la longueur $\alpha_1 + \dots + \alpha_d$ et on note ∂^α la dérivée partielle d'ordre total $|\alpha|$ et d'ordre α_j par rapport à la j -ème variable, $1 \leq j \leq d$.

Définition 1.1.3. Soit p tel que $1 \leq p \leq +\infty$, et m un entier positif. On définit l'espace de Sobolev $W^{m,p}(\Omega)$ par

$$W^{m,p}(\Omega) = \{v \in L^p(\Omega); \forall \alpha \in \mathbb{N}^n, |\alpha| \leq m, \partial^\alpha v \in L^p(\Omega)\}.$$

Dans le cas particulier $p = 2$, l'espace $W^{m,2}(\Omega)$ est noté par $H^2(\Omega)$, où Ω est un ouvert borné convexe à frontière lipschitzienne de \mathbb{R}^n .

On le munit de la norme

$$\|v\|_{W^{m,p}(\Omega)} = \left(\int_{\Omega} \sum_{|\alpha| \leq m} |\partial^{\alpha} v(\mathbf{x})|^p d\mathbf{x} \right)^{\frac{1}{p}} \quad \text{si } p < +\infty. \quad (1.1.2)$$

On peut vérifier que l'espace $W^{m,p}(\Omega)$ est un espace de Banach, réflexif lorsque $1 < p < +\infty$. Dans le cas particulier où $p = 2$, l'espace $H^m(\Omega)$ est un espace de Hilbert pour le produit scalaire associé à la norme (1.1.2) :

$$(u, v) \mapsto \int_{\Omega} \sum_{|\alpha| \leq m} (\partial^{\alpha} u)(\mathbf{x}) (\partial^{\alpha} v)(\mathbf{x}) d\mathbf{x}.$$

On note par $W_0^{m,p}(\Omega)$ l'adhérence de l'espace $\mathcal{D}(\Omega)$ dans l'espace $W^{m,p}(\Omega)$.

Corollaire 1.1.1. *Pour tout nombre réel p , tel que $1 \leq p < +\infty$, et tout entier positif m , la semi-norme*

$$|v|_{W^{m,p}(\Omega)} = \left(\int_{\Omega} \sum_{|\alpha|=m} |\partial^{\alpha} v(\mathbf{x})|^p d\mathbf{x} \right)^{\frac{1}{p}},$$

est une norme sur l'espace $W_0^{m,p}(\Omega)$, équivalente à la norme $\|\cdot\|_{W^{m,p}(\Omega)}$.

Démonstration. [2]. □

Définition 1.1.4. *Soit p un nombre réel, $1 \leq p < +\infty$, et m un entier positif. On définit le nombre réel p' tel que $\frac{1}{p} + \frac{1}{p'} = 1$. On note $W^{-m,p'}(\Omega)$ le dual de l'espace $W_0^{m,p}(\Omega)$, et on le munit de la norme dual*

$$\|f\|_{W^{-m,p'}(\Omega)} = \sup_{v \in W_0^{m,p}(\Omega), v \neq 0} \frac{\langle f, v \rangle}{|v|_{W^{m,p}(\Omega)}},$$

où $\langle \cdot, \cdot \rangle$ désigne le produit de dualité entre $W^{-m,p'}(\Omega)$ et $W_0^{m,p}(\Omega)$. Dans le cas particulier où $p = 2$, il est évident que $p' = 2$. On note respectivement $H_0^m(\Omega)$ et $H^{-m}(\Omega)$ les espaces $W_0^{m,2}(\Omega)$ et $W^{-m,2}(\Omega)$, et on utilise la même notation pour les normes associées; on pose

$$H^{-m}(\Omega) = (H_0^m(\Omega))',$$

alors

$$H_0^m(\Omega) \subset L^2(\Omega) \subset H^{-m}(\Omega),$$

où les inclusions précédentes sont considérées algébriquement et topologiquement.

1.2 Théorème de représentation de Riesz

Soit H un espace de Hilbert, considérons pour chaque x fixé dans H la fonction \mathbf{l}_x définie par le produit scalaire

$$\mathbf{l}_x(y) = \mathbf{l}(y) = (x, y), \quad \forall y \in H. \quad (1.2.1)$$

Il est facile de vérifier que \mathbf{l} est dans H' le dual de H et que

$$\|\mathbf{l}\|_{H'} = \|x\|_H.$$

En plus, l'application qui à tout x de H associe \mathbf{l} dans H' est linéaire, et on va voir que nous pouvons associer d'une manière naturelle une correspondance entre l'espace de Hilbert H et son dual H' .

Théorème 1.2.1 (Théorème de représentation de Riesz). *Soit H un espace de Hilbert sur le corps K et \mathbf{l} dans H' , alors il existe un et un seul élément x dans H , tel que*

$$\mathbf{l}(y) = (x, y), \quad \forall x, y \in H, \quad (1.2.2)$$

de plus

$$\|\mathbf{l}\|_{H'} = \|x\|_H. \quad (1.2.3)$$

Démonstration. [21]. □

La fonction qui à tout x de H associe \mathbf{l} de H' est souvent rencontrée dans la suite du travail. cette fonction sera appelée *représentation de Riesz* et sera notée par \mathcal{R}_H . En particulier, \mathcal{R}_H est une isométrie de H sur H' qui est définie par

$$(\mathcal{R}_H x, y) = (x, y)_H, \quad \forall x, y \in H. \quad (1.2.4)$$

1.2.1 Notions sur les opérateurs adjoints

Soient V et W deux espaces normés sur le même corps K et soit $\mathcal{L}(V, W)$ l'espace normé des opérateurs linéaires et continus de V dans W . Pour tout T dans $\mathcal{L}(V, W)$ on associe l'opérateur T^* dans $\mathcal{L}(W', V')$.

Théorème 1.2.2. *Pour tout T dans $\mathcal{L}(V, W)$ et pour toute fonctionnelle linéaire continue y^* dans W' , l'application x^* de V dans K , donnée par*

$$x^*(x) = y^*(Tx), \quad \forall x \in V, \quad (1.2.1)$$

est une fonctionnelle linéaire continue, i.e. x^ est dans V'*

Démonstration. [90]. □

Théorème 1.2.3. *L'application T^* de W' dans V' donnée par*

$$T^*y^* = x^*, \quad (1.2.2)$$

(au sens du Théorème 1.2.2) est un opérateur linéaire continu. De plus, nous avons pour tout x dans V

$$(T^*y^*)(x) = x^*(x) = y^*(Tx). \quad (1.2.3)$$

Démonstration. [90]. □

Définition 1.2.1. *L'opérateur T^* est appelé l'opérateur conjugué ou adjoint de l'opérateur T .*

Théorème 1.2.4. *Soit T dans $\mathcal{L}(V, W)$. Alors*

$$\|T^*\|_{V'} = \|T\|_W. \quad (1.2.4)$$

Démonstration. [90]. □

1.3 Minimisation des fonctions convexe

L'objectif de cette partie est d'amener peu à peu les lecteurs dans le "vif du sujet" de la théorie de la commande optimale. On traitera essentiellement les outils de base qui nous permettront d'étudier les conditions nécessaires et suffisantes d'optimalité dans un cadre abstrait.

Définition 1.3.1. *Une suite $\{u_n\}_n$ dans un espace de Hilbert H est dite faiblement convergente vers u dans H si $\lim_{n \rightarrow \infty} (u_n, v)_H = (u, v)_H$ pour tout v dans H . Il est clair que la limite faible u de la suite $\{u_n\}_n$ est unique. De la même manière, on dit que la suite $\{u_n\}_n$ est faiblement bornée si $|(u_n, v)_H|$ est bornée pour tout v dans H .*

Théorème 1.3.1. *Soit H un espace de Hilbert et soit D un sous ensemble de H dense et dénombrable. Si $\{u_n\}_n$ est une suite bornée dans H , alors elle est faiblement convergente.*

Tout espace muni d'une semi norme et qui admet un sous ensemble dénombrable et dense est dit séparable. Le théorème 1.3.1 montre que tout ensemble borné dans un espace de Hilbert séparable est séquentiellement relativement faiblement compact. Ce résultat existe dans tout espace de Banach réflexive, mais tous les espaces de fonctions qu'on va considérer ici sont des espaces de Hilbert séparables, donc le théorème 1.3.1 est suffisant pour la suite.

A partir de maintenant, on désigne par H un espace de Hilbert et par U un sous ensemble non vide de H et par $J(\cdot)$ une fonction définie de U dans \mathbb{R} . On rappelle que l'espace H est faiblement (séquentiellement) compact. Il vaut mieux considérer des sous

ensembles de U qui possède cette propriété. Donc, U est dit faiblement (séquentiellement) fermé si la limite de toute suite dans U converge faiblement dans U . Il est bien de savoir que la convergence forte (en norme) implique la convergence faible et qu'un ensemble faiblement fermé est nécessairement fermé.

Définition 1.3.2. *Un ensemble U dans un espace vectoriel H est dit convexe, si pour tout u, v dans U et pour tout τ compris entre 0 et 1, on a $\tau u + (1 - \tau)v$ est dans U , i.e si toute paire de vecteurs est dans U , alors le segment qui les relie l'est aussi.*

Lemme 1.3.1. *Si U est fermé et convexe, alors il est faiblement fermé.*

Le Lemme 1.3.1 et le Théorème 1.3.1 montrent que chaque sous ensemble fermé est faiblement séquentiellement compact. Il faut savoir que parfois on est ramené à considérer des cas où U n'est pas borné; par suite les notions suivantes sont parfois nécessaires.

Définition 1.3.3. *Une fonction $J(\cdot)$ définie sur U à valeurs réelles est faiblement semi-continue- inférieurement en u dans U , si pour toute suite $\{u_n\}_n$ dans U qui converge faiblement vers u dans U , on a*

$$J(u) \leq \liminf_{n \rightarrow \infty} J(u_n).$$

Définition 1.3.4. *On dit que la fonction $J(\cdot)$ a la propriété de croissance en u dans U , si pour r strictement positif et pour tout v dans U tels que $\|v - u\| \geq r$, on a*

$$J(v) > J(u). \quad (1.3.1)$$

Théorème 1.3.2. *Soit U un convexe fermé et soit $J(\cdot)$ une fonction définie sur U à valeurs réelles et faiblement semi-continue-inférieurement pour tout u dans U , soit si*

(a.) U est borné

soit si

(b.) $J(\cdot)$ vérifie la propriété de croissance dans U ,

alors il existe un u_0 dans U tel que $J(u_0) \leq J(u)$ pour tout u dans U ; donc $J(\cdot)$ atteint son minimum sur U .

Démonstration. Soient $d = \inf \{J(u) : u \in U\}$ et $\{u_n\}_n$ une suite dans U , tel que $d = \lim_{n \rightarrow \infty} J(u_n)$ si (a.) est vérifiée, alors par la faible compacité séquentielle il existe une sous suite de la suite $\{u_n\}$ noté par $\{u_{n_k}\}$ qui converge faiblement vers u_0 dans H ; le lemme 1.3.1 montre que u_0 est dans U . La faible continuité inférieure de $J(\cdot)$ montre que $J(u) \leq \liminf_{n \rightarrow \infty} J(u_{n_k}) = d$, d'où, $J(u_0) = d$ d'où le résultat. Pour le cas (b.), supposons que $J(\cdot)$ a la propriété de croissance en v qui appartient à U et soit r un réel strictement positif tel que $J(u) > J(v)$ dès que on a $\|v - u\| \geq r$ et u dans U . On pose $B = \{u \in V : \|u - v\| \leq r\}$ et on applique (a.) pour l'ensemble $B \cap U$ qui est un ensemble fermé, convexe et borné, d'où le résultat d'après l'observation que $\inf \{J(u) : u \in U\} = \inf \{J(u) : u \in B \cap U\}$. \square

On note que si U est borné alors $J(\cdot)$ à la propriété de croissance en tout les points de U ; donc le cas **(b.)** du théorème 1.3.2 inclus le cas **(b.)** comme un cas particulier.

Lemme 1.3.2. *Si U est un convexe fermé, alors il est faiblement fermé.*

Démonstration. [21]. □

La condition qu'une fonction soit faiblement semi-continue-inférieurement est difficile à vérifier. Cependant pour certaines fonctions qui sont convexes la semi-continuité-inférieure est la même pour la notion faible et forte, et cela peut être démontré directement d'après le Lemme 1.3.1.

Définition 1.3.5. *Une fonction $J(\cdot)$ définie sur U à valeurs réelles est dite convexe si U est convexe et si pour tout u et v dans U et τ compris entre 0 et 1 on a*

$$J(\tau u + (1 - \tau)v) \leq \tau J(u) + (1 - \tau)J(v), \quad (1.3.2)$$

et si on remplace l'inégalité large par une inégalité stricte, on dit que la fonction $J(\cdot)$ est strictement convexe.

Définition 1.3.6. *Une fonction $J(\cdot)$ définie sur U à valeurs réelles est dite différentiable au sens de Gâteaux en u qui appartenant à U , si U est convexe et si on a $J'(\cdot)$ appartient à H' tels que*

$$\lim_{\tau \rightarrow 0^+} \frac{1}{\tau} [J(u + \tau(v - u)) - J(u)] = J'(u)(v - u),$$

pour tout v appartenant à U . Dans ce cas, $J'(\cdot)$ est appelée la différentielle au sens de Gâteaux de $J(\cdot)$ au point u . Si en plus $J(\cdot)$ Gâteaux différentiable on tous les points de U , alors J' défini de U dans H' est le gradient de $J(\cdot)$ dans U . La différentielle $J'(\cdot)$ au sens de Gâteaux de $J(\cdot)$ est plus précisément la dérivée directionnelle de $J(\cdot)$ au point u suivant la direction v .

Ce qui va suivre nous indiquera comment caractériser la convexité de $J(\cdot)$.

Théorème 1.3.3. *Soit Une $J(\cdot)$ une fonction définie sur un convexe U et à valeurs réelles, alors les propriétés suivantes sont équivalentes :*

(a.) $J(\cdot)$ est convexe.

(b.) Pour tout u et v dans U on a :

$$J'(u)(v - u) \leq J(v) - J(u) \quad (1.3.3)$$

(c.) Pour tout u et v dans U on a :

$$(J'(u) - J'(v))(u - v) \geq 0. \quad (1.3.4)$$

Démonstration. Si $J(\cdot)$ est convexe, pour tout u et v dans U et tout τ dans $[0, 1]$ on a

$$J(u + \tau(v - u)) \leq J(u) + \tau(J(v) - J(u)),$$

dans ce cas on déduit 1.3.3. Par suite **(a.)** implique **(b.)**. Si **(b.)** est vérifiée, on a $J'(v)(u - v) \leq J(u) - J(v)$ et $J(u) - J(v) \leq J'(u)(u - v)$, par suite on déduit **(c.)**.

Finalement, on montre que **(c.)** implique **(a.)**. Soient (x, y) dans K , on définit φ de $[0, 1]$ dans \mathbb{R} par

$$\varphi(\tau) = J(\tau u + (1 - \tau)v) = J(v + \tau(u - v)), \quad \forall \tau \in [0, 1].$$

Alors $\varphi'(\tau) = J'(v + \tau(u - v))(u - v)$ et on a pour tout κ et τ dans $[0, 1]$, la relation

$$(\varphi'(\tau) - \varphi'(\kappa)) = (J'(v + \tau(u - v)) - J'(v + \kappa(u - v)))((\tau - \kappa)(u - v)) \geq 0,$$

on a d'après **(c.)** que φ' est non décroissante. Le Théorème des valeurs intermédiaires implique que

$$\frac{\varphi(1) - \varphi(\tau)}{1 - \tau} \geq \frac{\varphi(\tau) - \varphi(0)}{\tau - 0}, \quad \forall \tau \in]0, 1[.$$

Par suite,

$$\varphi(\tau) \leq \tau\varphi(1) + (1 - \tau)\varphi(0),$$

ce qui est exactement (1.3.5). □

Corollaire 1.3.1. *Soit $J(\cdot)$ une fonction Gâteaux différentiable et convexe. Alors $J(\cdot)$ est faiblement semi-continue-inférieurement dans U , et pour tout u et v dans U la fonction*

$$\tau \mapsto J'(u + \tau(v - u))(v - u), \quad \tau \in [0, 1],$$

est continue.

Démonstration. Soit la suite $\{u_n\}_n$ qui converge faiblement dans U . Comme $J'(u)$ est dans H' , on a $\lim_{u \rightarrow +\infty} J'(u)(u_n) = J'(u)(u)$, donc d'après (1.3.3), on obtient

$$\liminf_{n \rightarrow \infty} (J(u_n) - J(u)) \geq \liminf_{n \rightarrow \infty} J'(u)(u_n - u) = 0.$$

Ce qui montre que $J(\cdot)$ est faiblement semi-continue-inférieurement pour tout u dans U . Pour la continuité il faut voir la démonstration du théorème 1.3.3. □

Lemme 1.3.3. *Soit $J(\cdot)$ une fonction Gâteaux différentiable et convexe. On suppose qu'on a*

$$\lim_{\|u\| \rightarrow +\infty} \frac{J'(u)(u)}{\|u\|} = +\infty, \quad \forall u \in K.$$

par suite $\lim_{\|x\| \rightarrow +\infty} J(x) = +\infty$, (on dit que J est infinie à l'infini), dans ce cas la fonction $J(\cdot)$ a la propriété de croissance pour tous les points de U ; on dit aussi que $J(\cdot)$ est coercive.

Définition 1.3.7. Soit D un sous ensemble non vide de U , et soit $J(\cdot)$ une fonction définie sur D à valeur dans V' (le dual de V), on dit que $J(\cdot)$ est monotone si

$$(J(u) - J(v))(u - v) \geq 0, \quad \forall u, v \in D,$$

et strictement monotone si l'égalité est stricte.

Théorème 1.3.4. Soit U un sous ensemble non vide convexe et fermé d'un espace de Hilbert séparable, et soit une fonction $J(\cdot)$ définie sur U à valeurs réelles et Gâteaux différentiable dans U , on suppose que son gradient $J'(\cdot)$ est monotone et que l'une des propriétés suivantes est satisfaite

(a.) U est borné,

ou

(b.) $J'(\cdot)$ est coercive.

Alors l'ensemble $M = \{u \in U : J(u) \leq J(v) \text{ pour tout } v \text{ dans } U\}$ est non vide, fermé et convexe, et u est dans M si et seulement si

$$u \in U : J'(u)(v - u) \geq 0, \quad \forall v \in U. \quad (1.3.5)$$

Démonstration. D'après les Théorèmes 1.3.2 et 1.3.3, le Corollaire 1.3.1 et le Lemme 1.3.3, on déduit que l'ensemble M est non vide. Chaque ensemble $M_v = \{u \in U : J(u) \leq J(v)\}$ est fermé et convexe ainsi que leur intersection M . Si u est dans M alors (1.3.5) se déduit d'après la définition de $J'(u)$, réciproquement (1.3.3) montre que (1.3.5) implique que u est dans M . \square

On termine cette partie avec une condition suffisante sur l'unicité du minimum.

Théorème 1.3.5. Une fonction $J(\cdot)$ définie sur U à valeurs réelles qui est strictement convexe admet au plus un minimum.

Démonstration. Soient u_1 et u_2 dans U tels que u_1 est distinct de u_2 avec

$$J(u_1) = J(u_2) = \inf \{J(u) : u \in U\}.$$

Comme $\frac{1}{2}(u_1 + u_2)$ est dans U , la stricte convexité de $J(\cdot)$ donne

$$J\left(\frac{1}{2}(u_1 + u_2)\right) < \frac{1}{2}(J(u_1) + J(u_2)) = \inf \{J(u) : u \in U\},$$

qui est une contradiction. \square

La troisième partie de la démonstration du théorème 1.3.3 nous permet de donner le théorème suivant

Théorème 1.3.6. Soit $J(\cdot)$ Gâteaux différentiable dans U . Si le gradient $J'(\cdot)$ est strictement monotone, alors $J(\cdot)$ est strictement convexe.

Démonstration. [6]. \square

1.4 Commande optimale d'une équation elliptique linéaire

1.4.1 Introduction

Dans cette partie, on va modéliser un problème de commande optimale en introduisant une équation aux dérivées partielles de type elliptique qui est l'équation d'état du système qui admet une solution y qui est la fonction état du système, ensuite on introduit également une fonction de contrôle u , grâce à laquelle on va pouvoir agir sur le système (le contrôler ou le commander). On va justement montrer comment contrôler ce système via l'équation d'état. L'action ne doit pas se faire n'importe comment mais de manière optimale par rapport à un certain critère $J(\cdot)$ (ou coût) donné, et ce qui justifie la terminologie de **contrôle optimal** ou **commande optimale**. Nous allons, à présent modéliser ceci et en donnant une traduction mathématiques précise.

On commence par introduire, le problème variationnel abstrait suivant

$$\begin{aligned} \text{Trouver } u \in V \text{ tel que :} \\ a(u, v) = \langle f, v \rangle_{V', V}, \quad \forall v \in V, \end{aligned} \tag{\mathcal{E}_a}$$

où, V est un espace vectoriel de fonctions définies sur Ω associé à son dual V' et $\langle \cdot, \cdot \rangle$ désigne le crochet de dualité entre V et son dual V' . Ces fonctions peuvent être à valeurs scalaires (comme pour le cas du Laplacien) ou vectorielles (comme pour le cas du problème de Stokes [17]). Nous nous placerons toujours dans le cas où V est un espace de Hilbert, muni de la norme $\|\cdot\|_V$, où V est contenu dans H avec injection continue et dense; on identifie l'espace pivot H avec son dual de telle façon à avoir

$$V \hookrightarrow H \hookrightarrow V',$$

$a(u, v)$ est une forme bilinéaire sur $V \times V$, où la linéarité de $a(\cdot, \cdot)$ en v résulte de l'approche par des fonctions tests alors que la linéarité de $a(\cdot, \cdot)$ en u provient du fait que le problème de départ est linéaire et nous supposons que f est continue sur V .

On dit que la forme bilinéaire $a(\cdot, \cdot)$ est continue sur V si elle satisfait la condition de continuité suivante

$$\exists C > 0, \forall u, v \in V, |a(u, v)| \leq C \|u\|_V \|v\|_V. \tag{1.4.1}$$

où C sera appelée constante de continuité de la forme $a(\cdot, \cdot)$.

On dit que la forme bilinéaire $a(\cdot, \cdot)$ est coercive sur V si elle satisfait la condition de coercivité suivante

$$\exists c > 0, \forall u \in V, a(u, u) \geq c \|u\|_V^2. \tag{1.4.2}$$

où c sera appelée la constante de coercivité de la forme $a(\cdot, \cdot)$ et on associe à la forme bilinéaire $a(\cdot, \cdot)$ l'opérateur \mathcal{A} (par exemple l'opérateur elliptique), tel que \mathcal{A} est dans $\mathcal{L}(V, V')$ et défini comme suit

$$a(u, v) = \langle \mathcal{A}u, v \rangle_{V', V}, \quad \forall u, v \in V, \quad (1.4.3)$$

où \mathcal{A} est une bijection continue qui admet un inverse \mathcal{A}^{-1} continue.

Lemme 1.4.1 (Théorème de Lax-Milgram). *Soit V un espace de Hilbert séparable muni de la norme $\|\cdot\|_V$. Soit $a(\cdot, \cdot)$ une forme bilinéaire qui vérifie la condition de continuité (1.4.1) et la condition de coercivité (1.4.2), par suite pour tout f dans V' (l'espace dual de V), le problème (\mathcal{E}_a) est bien posé et il existe un unique y dans V telle qu'on a l'estimation suivante*

$$a(u, v) = \langle f, v \rangle_{V', V}, \quad \forall v \in V,$$

et

$$\forall f \in V', \quad \|u\|_V \leq \frac{1}{c} \|f\|_{V'}.$$

Démonstration. [28]. □

1.4.2 Formulation du problème de commande optimale

Pour définir le problème de commande optimale qu'on veut étudier, on doit spécifier en plus de l'espace de l'état V et de l'espace des données V' , un espace de Hilbert U qui est l'espace de toutes les commandes et un opérateur \mathcal{B} (dans notre cas linéaire) dans $\mathcal{L}(U, V')$, tels que pour chaque contrôle u dans U , on lui fait correspondre l'état $y = y(u)$ solution de l'équation d'état

$$\mathcal{A}y - \mathcal{B}u = f. \quad (1.4.4)$$

L'équation (1.4.4), peut s'écrire aussi sous la forme variationnelle suivante :

$$\langle \mathcal{A}y, v \rangle_{V', V'} - (\mathcal{B}u, v) = \langle f, v \rangle_{V', V'}, \quad \forall v \in V. \quad (\mathcal{E}_e)$$

Lemme 1.4.2. *Pour un certain u dans U , le problème variationnel (\mathcal{E}_e) est bien posé, i.e pour un certain u dans U , il existe une unique solution y qui vérifie l'équation variationnelle (\mathcal{E}_e) et qui dépend de u et de f .*

Démonstration. Puisque la forme bilinéaire $a(\cdot, \cdot)$ est continue et coercive, d'après le Théorème de Lax-Milgram, pour un certain u dans U , il existe une unique solution de (\mathcal{E}_e) , telque

$$\|y\|_V \leq C (\|f\|_{V'} + \|u\|_V).$$

□

On définit également l'espace de Hilber W , comme étant l'espace des observations et l'opérateur \mathcal{C} dans $\mathcal{L}(V, W)$, tels que pour chaque état y dans V , on lui fait correspondre l'état observable $\mathcal{C}y$ dans W , qu'on cherche à ramener le plus proche possible d'une observation donnée et connue à l'avance y_d dans W , le coût d'appliquer la commande u dans U est donné par $(Nu, u)_U$, où N est dans $\mathcal{L}(U, V')$ qui est un opérateur Hermitien (coercif) défini positif,

$$(Nu, u)_U \geq c \|u\|_U^2, \quad (1.4.5)$$

où, c est une constante positive.

Par suite, on définit la fonction objectif $J(\cdot)$, par l'addition d'une erreur au coût $(Nu, u)_U$, comme suit

$$J(u) = \|\mathcal{C}y(u) - y_d\|_V^2 + \alpha (Nu, u)_U, \quad (\mathcal{J})$$

où α est une constante positive.

On introduit maintenant U_{ad} le domaine des commandes admissibles qui est un sous ensemble fermé convexe et non vide de U . Dans ce cas, on définit le problème de commande optimale (\mathcal{P}) : il s'agit de minimiser la fonctionnelle (\mathcal{J}) sur l'ensemble U_{ad} , soumis à l'équation d'état (\mathcal{E}_e) . de manière équivalente le problème de commande optimale (\mathcal{P}) est défini par

$$u \in U_{ad} : J(u) \leq J(v), \quad \forall v \in U_{ad}. \quad (1.4.6)$$

Notre objectif principal est de trouver les conditions suffisantes d'existence et si possible d'unicité de la commande optimale u solution du problème (\mathcal{P}) . Il est utile de caractériser ces conditions sous une forme simple, qui peut fournir par la suite plus d'informations. Pour atteindre notre but, on fait appel au théorème 1.3.4, afin de pouvoir calculer la différentielle de Gâteaux de $J(\cdot)$, et grâce à la relation (1.4.20a), on obtient

$$\mathcal{C}y(u) - y_d = \mathcal{C}\mathcal{A}^{-1}\mathcal{B}u + \mathcal{C}\mathcal{A}^{-1}f - y_d. \quad (1.4.7)$$

Par suite, la relation (1.3.3) de la fonction objectif $J(\cdot)$ s'écrit sous la forme

$$J(u) = \|\mathcal{C}\mathcal{A}^{-1}\mathcal{B}u\|_W^2 + 2(\mathcal{C}\mathcal{A}^{-1}\mathcal{B}u, \mathcal{C}\mathcal{A}^{-1}f - y_d)_W + \|\mathcal{C}\mathcal{A}^{-1}f - y_d\|_V^2 + \alpha (Nu, u)_U. \quad (1.4.8)$$

Il est clair d'après la relation (1.4.8) que la fonction objectif $J(\cdot)$ est forme quadratique. Ainsi donc la différentielle de Gâteaux de $J(\cdot)$, s'écrit facilement sous la forme

$$\begin{aligned} (\nabla J(u), \chi) &= 2 [(\mathcal{C}\mathcal{A}^{-1}\mathcal{B}u, \mathcal{C}\mathcal{A}^{-1}\mathcal{B}\chi)_W + (\mathcal{C}\mathcal{A}^{-1}\mathcal{B}\chi, \mathcal{C}\mathcal{A}^{-1}f - y_d)_W + \alpha (Nu, \chi)_U] \\ &= 2 [(\mathcal{C}y(u) - y_d, \mathcal{C}\mathcal{A}^{-1}\mathcal{B}\chi)_W + \alpha (Nu, \chi)_U], \end{aligned}$$

puisque

$$\left(\frac{1}{2}\nabla J(u), \chi\right) = (\mathcal{C}y(u) - y_d, \mathcal{C}\mathcal{A}^{-1}\mathcal{B}\chi)_W + \alpha(Nu, \chi)_U \quad (1.4.9)$$

il vient que le gradient $\nabla J(\cdot)$ de $J(\cdot)$ est monotone et

$$\frac{1}{2}(\nabla J(u), u) \geq \alpha(Nu, u)_U - c\|u\|_U. \quad (1.4.10)$$

D'après la coercivité de $N(\cdot, \cdot)$ définie par la relation (1.4.5), on déduit facilement la coercivité de la fonctionnelle $J(\cdot)$.

D'après le théorème 1.3.4, on obtient le théorème suivant :

Théorème 1.4.1. *Soit le problème de commande optimale (\mathcal{P}), qui se ramène ramené à minimiser (\mathcal{J}) soumise à l'équation d'état (\mathcal{E}_e) sur l'ensemble fermé convexe et non vide U_{ad} . Si*

(a) U_{ad} est borné,
ou si

(b) N est coercive sur U_{ad} ,

alors l'ensemble des commandes optimales U est non vide, fermé et convexe.

Corollaire 1.4.1. *Dans le cas où la condition (b) du Théorème 1.4.1 est satisfaite, la commande optimale u est unique.*

Démonstration. Se déduit facilement du Théorème 1.3.6 et (1.4.5) impliquent que $\nabla J(\cdot)$ est strictement monotone. \square

D'après la relation (1.4.10), la commande optimale u peut être caractérisée par une inéquation variationnelle, i.e, u est une commande optimale si et seulement si

$$u \in U_{ad} : \nabla J(u, v - u) \geq 0, \quad \forall v \in U_{ad}, \quad (1.4.11)$$

qui est exactement (1.3.5). Il est clair que cette inéquation variationnelle est donnée par (1.4.9), mais la fomulation du résultat est difficile à interpréter et la difficulté réside dans le fait qu'on est entrain de comparer des éléments de l'espace d'observation W avec ceux de l'espace des commandes U . On doit alors chercher une caractérisation équivalente sous la forme d'une inéquation variationnelle mais cette fois uniquement sur l'espace des commandes U . Afin de pouvoir transformer le membre de droite dans (1.4.9) sous une forme plus simple, on fait appel au Théorème de représentation de Riesz.

Pour la suite, on introduit l'opérateur adjoint \mathcal{C}^* de \mathcal{C} dans $\mathcal{L}(W', V')$, dans ce cas, d'après (1.4.9), il vient

$$\begin{aligned} \left(\frac{1}{2}\right) (J'(u), \chi) &= (\mathcal{C}y(u) - y_d, \mathcal{C}\mathcal{A}^{-1}\mathcal{B}\chi)_W + \alpha(Nu, \chi)_U \\ &= (\mathcal{R}_W\mathcal{C}y(u) - y_d, \mathcal{C}\mathcal{A}^{-1}\mathcal{B}\chi)_W + \alpha(Nu, \chi)_U \\ &= (\mathcal{C}^*\mathcal{R}_W\mathcal{C}y(u) - y_d, \mathcal{A}^{-1}\mathcal{B}\chi)_W + \alpha(Nu, \chi)_U, \quad \forall u, \chi \in U. \end{aligned}$$

On introduit également l'opérateur adjoint \mathcal{A}^* de \mathcal{A} dans $\mathcal{L}(V, V')$, où il faut noter que l'opérateur \mathcal{A}^* se déduit de la forme adjointe $a^*(\cdot, \cdot)$ de $a(\cdot, \cdot)$. Donc pour chaque commande u dans U , on définit l'état adjoint correspondant $p = p(u)$ comme étant la solution de l'équation suivante

$$\mathcal{A}^*p = \mathcal{C}^*\mathcal{R}_W(\mathcal{C}y(u) - y_d), \quad p = p(u). \quad (1.4.12)$$

Dans la suite cette équation sera appelée équation adjointe. D'après (1.4.9), on obtient

$$\begin{aligned} \left(\frac{1}{2}\right) (\nabla J(u), \chi) &= (\mathcal{A}^*p, \mathcal{A}^{-1}\mathcal{B}\chi)_W + \alpha(Nu, \chi)_U \\ &= (p, \mathcal{B}\chi)_W + \alpha(Nu, \chi)_U \\ &= (\mathcal{B}^*p, \chi)_W + \alpha(Nu, \chi)_U. \end{aligned}$$

Par suite, on a

$$\left(\frac{1}{2}\right) (\nabla J(u), \chi) = (\mathcal{B}^*p + \alpha Nu, \chi)_U. \quad (1.4.13)$$

D'après (1.4.11), on déduit l'inéquation suivante

$$(\mathcal{B}^*p(u) + \alpha Nu, v - u)_U \geq 0. \quad (1.4.14)$$

La relation (1.4.14), qu'on appelle équation de projection, peut être écrite sous la forme

$$\inf_{v \in U_{ad}} (\mathcal{B}^*p(u) + \alpha Nu, v - u)_U. \quad (1.4.15)$$

D'une manière équivalente, on écrit

$$u = \frac{N^{-1}}{\alpha} \pi_{U_{ad}}(\mathcal{B}^*p(u)),$$

où $\pi_{U_{ad}}$ est l'opérateur de projection de U sur U_{ad} et \mathcal{B}^* est dans $\mathcal{L}(V, U')$, qui est l'opérateur adjoint de l'opérateur \mathcal{B} dans $\mathcal{L}(U, V')$. Ce calcul nous amène à la formulation de (4.8) qu'on résume sous la forme du théorème suivant :

Théorème 1.4.2. *Soit le problème de commande optimale (\mathcal{P}) donné ci-dessus. Alors une condition nécessaire et suffisante pour que la commande u soit solution du problème (\mathcal{P}) (i.e une commande optimale) est qu'elle satisfait le système (\mathcal{S}) suivant :*

$$u \in U_{ad} : \mathcal{A}y(u) - \mathcal{B}u = f, \quad (1.4.16a)$$

$$\mathcal{A}^*p(u) = \mathcal{C}^*\mathcal{R}_W(\mathcal{C}y(u) - y_d), \quad (1.4.16b)$$

$$(\mathcal{B}^*p(u) + \alpha Nu, v - u) \geq 0, \quad \forall v \in U_{ad}. \quad (1.4.16c)$$

Le système (1.4.16), s'appelle : système d'optmalité du problème de commande optimale (\mathcal{P}) .

Corollaire 1.4.2. *Si on suppose que les conditions du Théorème 1.4.2 restent satisfaites et si l'ensemble des commandes admissibles U_{ad} est égal à tout l'espace U , alors le système des conditions d'optimalités (\mathcal{S}) s'écrit sous la forme*

$$u \in U_{ad} : \mathcal{A}y(u) + \mathcal{B}u = f, \quad (1.4.17a)$$

$$\mathcal{A}^*p(u) = \mathcal{C}^*\mathcal{R}_W(\mathcal{C}y(u) - y_d), \quad (1.4.17b)$$

$$\mathcal{B}^*p(u) + \alpha Nu = 0. \quad (1.4.17c)$$

Dans ce qui précède le choix typique des espaces est de prendre $V = H^1(\Omega)$, $H = L^2(\Omega)$ et $V' = H^{-1}(\Omega)$. Soit α_i , $i = 1, 2$ deux constantes positives tel que $\alpha_1 + \alpha_2 > 0$. On définit pour tout u et v dans $H^1(\Omega)$ la forme bilinéaire $a(\cdot, \cdot)$ par

$$a(u, v) = \sum_{i,j=1}^N \int_{\Omega} a_{ij}(\mathbf{x}) \frac{\partial u(\mathbf{x})}{\partial \mathbf{x}_i} \frac{\partial v(\mathbf{x})}{\partial \mathbf{x}_i} + \int_{\Omega} a_0(\mathbf{x}) u(\mathbf{x}) v(\mathbf{x}) d\mathbf{x} + \frac{\alpha_1}{\alpha_2} \int_{\partial\Omega} u(\boldsymbol{\sigma}) v(\boldsymbol{\sigma}) d\boldsymbol{\sigma},$$

si $\alpha_1 \geq 0$ et $\alpha_2 > 0$, où a_0 et a_{ij} sont dans $L^\infty(\Omega)$ pour tout $i, j = 1, \dots, n$ et

$$a_0(\mathbf{x}) \geq c > 0; \sum_{i,j=1}^N a_{ij}(\mathbf{x}) \xi_i \xi_j \geq c \|\xi\|_N^2, \forall \xi \in \mathbb{R}^N, \text{ p.p. } \mathbf{x} \in \Omega. \quad (1.4.18)$$

Dans le cas où, $\alpha_1 > 0$ et $\alpha_2 = 0$, la forme bilinéaire $a(\cdot, \cdot)$ sera définie par

$$a(u, v) = \sum_{i,j=1}^N \int_{\Omega} a_{ij}(\mathbf{x}) \frac{\partial u(\mathbf{x})}{\partial \mathbf{x}_i} \frac{\partial v(\mathbf{x})}{\partial \mathbf{x}_i} + \int_{\Omega} a_0(\mathbf{x}) u(\mathbf{x}) v(\mathbf{x}) d\mathbf{x}, \forall u, v \in H_0^1(\Omega),$$

tel que $a(\cdot, \cdot)$ est une forme bilinéaire définie par

$$a(u, v) = \langle \mathcal{A}u, v \rangle,$$

où

$$\mathcal{A}v = - \sum_{i,j=1}^n \frac{\partial}{\partial \mathbf{x}_i} \left(a_{ij}(\mathbf{x}) \frac{\partial v}{\partial \mathbf{x}_j} \right) + a_0(\mathbf{x}) v.$$

On appelle $\mathcal{A} \in \mathcal{L}(H_0^1(\Omega), H_0^{-1}(\Omega))$ l'opérateur linéaire (elliptique) associé à la forme bilinéaire $a(\cdot, \cdot)$.

Pour l'analyse numérique, on considère un modèle simple [58], mais loin d'être trivial pour illustrer un problème de commande optimale gouverné par une équation variationnelle.

Le problème que l'on étudiera ci-dessous est un problème de commande optimale pour un état stable d'un conducteur de transfert de chaleur gouverné par l'équation de Poisson

par rapport à une plaque carrée avec des conditions au bord homogènes de type Dirichlet. L'objectif dans les problèmes de transfert de chaleur est d'obtenir le contrôle optimal u , i.e., la source de chaleur dans la plaque qui minimise la fonction coût J , il se pose sous la forme

$$\begin{cases} \min_{u \in U_{ad}} J(y, u) = \frac{1}{2} \int_{\Omega} (y - y_d)^2 d\Omega + \frac{\alpha}{2} \int_{\Omega} u^2 d\mathbf{x}, \\ a(y, w) = (f + Bu, w)_{L^2(\Omega)}, \quad \forall w \in H_0^1(\Omega), \end{cases} \quad (\mathcal{P})$$

où f et y_d sont dans $L^2(\Omega)$ et tel que $a(y, w) = \langle -\Delta y, w \rangle_{H_0^1(\Omega), H_0^{-1}(\Omega)}$, ci-dessus, on a supposé que $V = H_0^1(\Omega)$ et $U = L^2(\Omega)$, et que l'opérateur \mathcal{A} est égal à l'inverse de l'opérateur de Laplace $\Delta \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$. Donc les fonctions d'état $y(u)$ et de commande u sont couplées par l'équation d'état suivante :

$$-\Delta y = Bu + f.$$

Par suite, l'opérateur \mathcal{T} est ici $-(\Delta)^{-1}$, qui est continu, linéaire et bijectif de $H^{-1}(\Omega)$ dans $H_0^1(\Omega)$ et de $L^2(\Omega)$ dans $H^2(\Omega) \cap H_0^1(\Omega)$. Ainsi le problème s'écrit sous la forme :

D'après le développement théorique introduit plus haut, on déduit facilement les conditions nécessaires et suffisantes d'optimalité. La paire $(y(u), u)$ est solution optimale du problème (\mathcal{P}) s'il existe $p(u)$ qui est l'état adjoint, tel que le triplet $(y(u), p(u), u_N)$ satisfait le système (\mathcal{S}) suivant

$$\begin{aligned} a(y(u), w) &= (f + Bu, w)_{L^2(\Omega)}, \quad \forall w \in H_0^1(\Omega), \\ a(p(u), q) &= (y(u) - y_d, q)_{L^2(\Omega)}, \quad \forall q \in H_0^1(\Omega), \\ (\alpha u + B^*p(u), v - u)_{L^2(\Omega)} &\geq 0, \quad \forall v \in U_{ad}. \end{aligned} \quad (1.4.19a)$$

Dans le cas particulier où $U_{ad} = L^2(\Omega)$, l'équation de projection(1.4.19a) s'écrit sous la forme

$$\alpha u + B^*p(u) = 0.$$

Alors le système des conditions d'optimalité (1.4.19) s'écrit sous la forme

$$a(y(u), w) = (f + Bu, w)_{L^2(\Omega)}, \quad \forall w \in H_0^1(\Omega), \quad (1.4.20a)$$

$$a(p(u), q) = (y(u) - y_d, q)_{L^2(\Omega)}, \quad \forall q \in H_0^1(\Omega), \quad (1.4.20b)$$

$$\alpha u + B^*p(u) = 0, \quad (1.4.20c)$$

en substituant l'équation de projection (1.4.20c), dans l'équation d'état (1.4.20a), on obtient le système équivalent suivant

$$a(y(u), w) = (f - \frac{1}{\alpha} B^* B p(u), w), \quad \forall w \in H_0^1(\Omega), \quad (1.4.21a)$$

$$a(p(u), q) = (y(u) - y_d, q), \quad \forall q \in H_0^1(\Omega), \quad (1.4.21b)$$

où

$$a(v, w) = \int_{\Omega} \nabla v \nabla w \, d\mathbf{x} \quad \forall v, w \in H_0^1(\Omega).$$

Ici, on a utilisé le principe du calcul des variations pour construire les conditions d'optimalité, mais on peut aussi utiliser d'autres conduisant à des conditions d'optimalité équivalentes (voir Annexe A).

Chapitre 2

Méthodes d'approximations spectrale

2.1 Introduction

Les méthodes spectrales introduites par D.Gottlieb et S.Orszag [69] ont été utilisées pour l'approximation des solutions d'équations aux dérivées partielles (stationnaire ou d'évolution). Elles furent initialement introduites par l'utilisation des séries de *Fourier* tronquées qui ont été essentiellement utilisées pour l'approximation des problèmes avec des conditions aux limites périodiques, ensuite par l'utilisation des polynômes de haut degré formant des bases tensorisées des espaces d'approximations qui est une propriété fondamentale dans la construction des méthodes spectrales. L'approximation par les méthodes spectrales généralement fait appel à des polynômes de la famille des polynômes de **Jacobi**,

$$P_k^{(\alpha, \beta)}(x) = \frac{1}{2^k} \sum_{l=0}^k \binom{k+\alpha}{l} \binom{k+\beta}{k-l} (x-1)^l (x+1)^{k-l}, \quad \forall x \in [-1, 1], \quad (2.1.1)$$

par rapport à la fonction poids

$$\omega(x) = (1-x)^\alpha (1+x)^\beta.$$

Bien sûr, la nature et le choix des polynômes d'approximation dépendent particulièrement du domaine sur lequel on cherche la solution. On fait appel par exemple aux polynômes de Legendre (pour $\alpha = 0$ et $\beta = 0$), ou aux polynômes de Tchebycheff de première espèce (pour $\alpha = \frac{-1}{2}$ et $\beta = \frac{-1}{2}$ à un facteur multiplicatif près qui dépend essentiellement du degré k) si on se place sur le segment $\Lambda = [-1, 1]$ dans le cas $1 - D$ et si on se place sur le carré $\Omega = [-1, 1]^2$ dans le cas $2 - D$.

En ce qui nous concerne, on ne va considérer que des problèmes en $2 - D$. Cependant, le choix des polynômes, et la méthode d'approximation sont également soumis à certaines conditions qui dépendent dans la majorité des cas de la nature du problème à résoudre

numériquement. Par exemple, les méthodes spectrales font appel à l'utilisation des polynômes de Tchebycheff sont parfois difficiles à mettre en œuvre et analyser numériquement en particulier lorsqu'il s'agit des problèmes de commande optimale. Il faut savoir aussi que la définition et l'analyse numérique des méthodes spectrales reposent en grande partie sur les propriétés des polynômes orthogonaux.

Pour une brève introduction aux méthodes spectrales, on considère le problème formel suivant

$$\begin{aligned} \mathcal{L}u + f &= 0 \\ + \text{Conditions au bord et / ou initiales.} \end{aligned} \quad (2.1.2)$$

Il faut se rendre compte qu'on ne considère ici que les choses d'une façon formelle, où \mathcal{L} est un opérateur différentiel (stationnaire ou d'évolution) ou intégrale. Notre première préoccupation est de pouvoir approximer la fonction inconnue u (la solution du problème (2.1.2), en l'écrivant sous la forme d'une combinaison linéaire de certaines fonctions polynomiales φ_N connues à l'avance de degré inférieur ou égale à N , appartenant à la famille des polynômes de Jacobi définis par (2.1.1) sous la forme suivante

$$u \simeq u_N = \sum_{i=0}^{N+1} u_i \varphi_N, \quad (2.1.3)$$

où le choix des fonctions polynomiales φ_N ne se fait pas d'une façon arbitraire, mais dépend de plusieurs paramètres à prendre en considération dès la formulation du problème (par exemple les conditions aux bords), en remplaçant la relation (2.1.3) dans l'équation (2.1.2), on définit de façon naturelle la quantité

$$R_{N+1}(u_0, u_1, \dots, u_{N+1}) = \mathcal{L}u_N + f, \quad (2.1.4)$$

qui n'est autre que le résidu (l'erreur) de l'approximation numérique. Notre souci principal justement est d'étudier la famille des méthodes d'approximation, afin de pouvoir calculer les valeurs des coefficients u_i (inconnus) où $i = 0, \dots, N+1$ et d'évaluer la quantité R_{N+1} et pourquoi pas la rendre la plus petite possible.

2.1.1 Polynômes orthogonaux

Soit $\omega(\mathbf{x})$, une fonction mesurable au sens de Lebesgue sur le domaine Ω , tel que $\int_{\Omega} \omega(\mathbf{x}) d\mathbf{x} > 0$ et on note par $L_{\omega}^2(\Omega)$ l'espace des fonctions mesurables telle que leur norme induite par le produit scalaire défini sur le même espace doit être borné. On définit la norme sur $L_{\omega}^2(\Omega)$ par

$$\|u\|_{\omega} = \left(\int_{\Omega} |u(\mathbf{x})| \omega(\mathbf{x}) d\mathbf{x} \right)^{\frac{1}{2}}, \quad (2.1.5)$$

qui est la norme induite par le produit scalaire défini sur le même espace par l'intégrale de Lebesgue-Steiltjes :

$$(u, v)_\omega = \int_\Omega u(\mathbf{x}) v(\mathbf{x}) \omega(\mathbf{x}) d\mathbf{x}. \quad (2.1.6)$$

Un ensemble de polynômes $\{p_k\}_k$ est dit orthonormal si la relation

$$\begin{aligned} (p_i, p_j)_\omega &= \int_\Omega p_i(\mathbf{x}) p_j(\mathbf{x}) \omega(\mathbf{x}) d\mathbf{x} \\ &= \delta_{ij}, \end{aligned} \quad (2.1.7)$$

est vérifiée pour tout polynôme $p_i(\mathbf{x})$ dans $\{p_k\}_k$, où δ_{ij} est le symbole de Kronecker. Par suite, on dit que les polynômes $p_i \in \{p_k\}_k$, sont des polynômes orthogonaux (dans ce cas on dit orthonormaux) par rapport à la fonction poids $\omega(\mathbf{x})$, où, la famille des polynômes $\{p_k\}_k$ peut être construite par le procédé *d'orthogonalisation de Schmidt* à partir de la suite des polynômes $1, x, x^2, \dots, x^n$ [81]. On note par la suite $\mathbb{P}(\Lambda)$ l'espace des polynômes à une variable de degré inférieur ou égal à N et où Λ est l'intervalle $[-1, 1]$

2.1.2 Polynômes de Legendre

Les polyômes de Legendre sont notés habituellement par $L_k(x)$, où k désigne le degré du polynôme $L_k(x)$ par rapport à la variable x , telle que la parité de $L_k(x)$ est la même que celle du degré k . Ils forment une famille de polynômes $\{L_k\}_k$ deux à deux orthogonaux dans l'espace $L^2(\Lambda)$ par rapport à la fonction poids $\omega(x) = 1$, par suite on a

$$\int_\Lambda L_i(x) L_j(x) dx = \frac{2}{2n+1} \delta_{ij},$$

où δ_{ij} est le symbole de Kronecker.

Les polynômes de Legendre peuvent être définis de plusieurs façons (par exemple qu'ils appartiennent à la famille des polynômes de Jacobi) et peuvent être construit par plusieurs procédés (par exemple par le procédé d'orthogonalisation de Schmidt) qui sont tous équivalents. Ils peuvent être aussi vu tout simplement comme étant fonctions propres de l'opérateur discret de Sturm-Liouville [17],

$$\mathcal{A}\varphi = -\frac{d}{d\zeta} ((1 - \zeta^2)\varphi'). \quad (2.1.8)$$

L'opérateur \mathcal{A} est un opérateur auto-adjoint positif¹ dans $L^2(\Lambda)$ [17], ce qui explique l'origine du qualificatif **spectral**. Pour des raisons numériques, ils peuvent aussi être construits grâce à la relation de récurrence à trois termes suivante :

$$L_{k+1}(x) = \frac{2k+1}{k+1} x L_k(x) - \frac{k}{k+1} L_{k-1}(x), \quad (2.1.9)$$

¹il ne faut pas le confondre avec l'opérateur elliptique

où $L_0(x) = 1$ et $L_1(x) = x$. Telle que pour l'implémentation et la mise on œuvre numérique des polynômes de Legendre $L_k(x)$, on peut faire appel à la relation de récurrence (2.1.9). Si on veut, par exemple, utiliser un logiciel de calcul scientifique comme **Fortran**, **Matlab** ou **Scilab**,..., etc, qui nécessite un code permettant la construction des polynômes de Legendre. Cependant on peut faire plus simple en utilisant un logiciel de calcul symbolique, par exemple **Maple** ou **Mupad**,...,etc, où il faut juste appeler la bibliothèque interne des polynômes orthogonaux du logiciel utilisé (par exemple Maple) par la commande

with(orthopoly)

pour utiliser pleinement les propriétés des polynômes de Legendre, on exécute simplement la commande

P(k,x)

où k , désigne le degré du polynôme de Legendre noté dans Maple par P(k,x) par rapport à la variable x .

Dans ce qui suit, on donne quelques relations sans démonstration qui sont importantes dans l'approximation numérique et dans la démonstration de certaines propriétés qui vont suivre. Pour une lecture plus approfondie on peut se référer par exemple à [22].

- Les polynômes de Legendre sont des fonctions bornées sur Λ ;

$$|L_k(x)| \leq 1, \quad \text{pour } x \text{ dans } \Lambda.$$

- Pour la dérivée première L'_k de L_k , on a ;

$$|L'_k(x)| \leq (\pm 1)^k \frac{1}{2} k(k+1),$$

$$L'_k(\pm 1) = (\pm 1)^{k-1} \frac{k(k+1)}{2}.$$

Les polynômes $L_k(x)$ de Legendre forment un sous-espace dense dans l'espace des fonctions sur $\bar{\Lambda}$. Par conséquent, la famille $\{L_k\}_k$ des polynômes de Legendre est une famille totale de l'espèce $L^2(\Lambda)$. Par suite, si on suppose que u est une fonction dans $L^2(\Lambda)$, alors $u(x)$ peut être écrite sous la forme d'une combinaison linéaire de polynômes de Legendre, comme suit

$$u(x) = \sum_{k=0}^{\infty} u_k L_k(x),$$

où

$$u_k = \left(k + \frac{1}{2}\right) \int_{\Lambda} u(x) L_k(x) dx,$$

et dans le cas où on suppose que u' est dans $L^2(\Lambda)$, on peut écrire $u'(x)$ sous la forme

$$u'(x) = \sum_{k=0}^{\infty} u_k^{(1)} L_k(x),$$

où

$$u_k^{(1)} = (2m+1) \sum_{\substack{p=m+1 \\ p+m \text{ impaire}}} u_p,$$

Dans ce cas, on déduit la relation suivante

$$u_k = \frac{u_{k-1}^{(1)}}{2k-1} - \frac{u_{k+1}^{(1)}}{2k+3}, \quad \text{pour } k \geq 1,$$

qui est généralisée par la relation suivante

$$u_k^{(q-1)} = \frac{u_{k-1}^{(q)}}{2k-1} - \frac{u_{k+1}^{(q)}}{2k+3},$$

$$u_m^{(2)} = \left(m + \frac{1}{2}\right) \sum_{\substack{p=m+1 \\ p+m \text{ paire}}} [p(p+1) - m(m+1)] u_p.$$

D'après ce qui précède, on donne la relation des dérivées premières $L'(x)$ et la dérivée seconde $L''(x)$ sous la forme

$$L'(x) = \sum_{\substack{k=0 \\ k+n \text{ impair}}}^{n-1} (2k+1) L_k(x),$$

et

$$L''_n(x) = \sum_{\substack{k=0 \\ k+n \text{ paire}}}^{n-2} \left(k + \frac{1}{2}\right) [n(n+1) - k(k+1)] L_k(x).$$

Les polynômes de Legendre en $2-D$ ainsi que leurs propriétés se déduisent de celles des polynômes de Legendre en $1-D$ par un paramètre de tensorisation [17]. Pour tout entier $N \geq 0$, on note par $\mathbb{P}_N(\Omega)$ l'espace des polynômes sur Ω de degré inférieur ou égal à N par rapport à chaque variable x et y , qui n'est autre que l'espace des polynômes en x sur Λ de degré inférieur ou égal à N , à coefficients dans l'espace $\mathbb{P}_N(\Lambda)$ de degré inférieur ou égale à N par rapport à la variable y .

Dans ce cas, on définit les polynômes de Legendre en $2-D$, sous la forme

$$L_{kl}(\mathbf{x}) = \{L_k(x) L_l(y), 0 \leq k, l \leq N\},$$

où, les éléments de la famille $\{L_{kl}(\mathbf{x})\}_{kl}$ forment une base sur l'espace $\mathbb{P}_N(\Omega)$, où k et l sont respectivement, les degrés des polynômes L_k et L_l .

Par suite on déduit la relation d'orthogonalité en dimension deux, sous la forme

$$(L_{kl}, L_{mn}) = \left(\frac{2}{2k+1}\right) \left(\frac{2}{2l+1}\right) \delta_{km} \delta_{nl}, \quad (2.1.10)$$

où $\delta_{..}$, est le symbole de Kronecker.

2.2 Techniques de détermination des coefficients d'approximation

Pour la détermination des coefficients d'approximations par les méthodes spectrale (par exemple les coefficients dans (2.1.3)), il y a en général trois techniques basées sur les méthodes d'approximation qui sont : la méthode d'approximation de Tau, la méthode d'approximation de Galerkin et la méthode d'approximation pseudo-spectrale (méthode de collocation). Dans tous les cas, on considère le résidu (le reste) R_{N+1} donné par la relation (2.1.4) qu'on veut rendre le plus petit possible. D'une manière formelle, on doit avoir :

$$\lim_{N \rightarrow \infty} R_{N+1} \rightarrow 0.$$

Dans ce qui suit, on donne les définitions et les différentes techniques d'approximations qu'on peut rencontrer dans les méthodes d'approximation spectrale. Pour fixer les idées, on considère le problème formel (2.1.2) dans le cas $1 - D$.

- **Méthode d'approximation de Tau** : Cette méthode a été introduite par *C.Lanczos* en 1938, où les coefficients de l'approximation sont déterminés de façon à ce que les conditions au bord soient vérifiées et le résidu R_{N+1} soit orthogonal aux fonctions polynomiales de Legendre $L_k(x)$. Dans ce cas, la solution approchée u_N de u s'écrit sous la forme

$$u \simeq u_N = \sum_{i=0}^{N+1} u_i L_i. \quad (2.2.1)$$

En substituant la relation (2.2.1) directement dans l'équation (2.1.2), on obtient

$$\begin{aligned} \mathcal{L}u_N + f &= 0 \\ &+ \text{Conditions au bord,} \end{aligned} \quad (2.2.2)$$

- **Méthode d'approximation de Galerkin** : Cette méthode a été introduite par *Boris Grigoryevich Galerkin* en 1871, il s'agit de définir de nouvelles fonctions de base $\{\Phi_i\}_i$, où $i = 0, \dots, N + 1$ qui sont une combinaison linéaire des fonctions

de base originales (dans notre cas, ce sont les polynômes de Legendre $L_k(x)$) tel que ces nouvelles fonctions de base $\{\Phi_i\}_i$, où $i = 0, \dots, N + 1$ doivent satisfaire les conditions au bord imposées dans le problème (2.1.2). Dans ce cas, la solution approchée u_N de u s'écrit sous la forme

$$u \simeq u_N = \sum_{i=0}^{N+1} u_i \Phi_i. \quad (2.2.3)$$

En substituant la relation (2.2.3), dans l'équation (2.1.2), écrite sous sa formulation variationnelle, on obtient

$$(\mathcal{L}u_N, \Phi_i) + (f, \Phi_i) = 0, \quad i = 0, \dots, N + 1. \quad (2.2.4)$$

Le résidu (2.1.4) doit être orthogonal aux nouvelles fonctions de base $\{\Phi_i\}_i$.

- **Méthode d'approximation pseudo-spectral** : Elle ressemble à la méthode d'approximation de Tau ; où les coefficients d'approximation sont déterminés tels que les conditions aux bords soient satisfaites en des points ζ_i où $i = 0, \dots, N + 1$ qu'on appelle souvent : **points de collocation**. Le résidu (2.1.4) doit être nul en ces points. Il reste à choisir les fonctions de base, qui sont généralement les polynômes de Lagrange $\{\psi_i\}_i$, $i = 0, \dots, N + 1$, où chaque polynôme ψ_i prend la valeur 1 au point ζ_i et la valeur 0 ailleurs. Dans ce cas, la solution approchée u_N de u du problème (2.1.2), s'écrit sous la forme

$$u \simeq u_N = \sum_{i=0}^{N+1} u_i \psi_i, \quad (2.2.5)$$

et en substituant la relation (2.2.3) dans l'équation (2.1.2) évaluée aux points de collocation $\zeta_0, \zeta_1, \zeta_2, \dots, \zeta_N$, on obtient

$$\begin{aligned} \mathcal{L}u_N(\zeta_i) + f(\zeta_i) &= 0, \quad i = 0, \dots, N + 1 \\ &+ \text{Conditions au bord aux points } \zeta_i. \end{aligned} \quad (2.2.6)$$

2.3 Méthode d'approximation de Galerkin

Soit le domaine $\Omega = [-1, 1]^2$. On désire trouver la solution u du problème (2.1.2), dans un certain espace V de dimension infinie définissant les conditions au bord et des conditions de régularités. D'après ces conditions, pour l'approximation numérique du problème (2.1.2), i.e. pour trouver la solution approchée u_N de u , on définit l'espace de dimension fini V_N par

$$V_N = \mathbb{P}_N(\Omega) \cap V,$$

qui est l'espace des polynômes satisfaisant les conditions au bord imposées dans le problème (2.1.2). Dans ce cas l'espace V_N est défini sous la forme

$$V_N = \text{span} \{ \Phi_i(x) \otimes \Phi_j(y); i, j = 1, 2, \dots, N-2 \},$$

où les éléments de V_N sont linéairement indépendants. Ensuite on propose de chercher une base sous la forme [79]

$$\Phi_{ij}(\mathbf{x}) = \Phi_i(x) \otimes \Phi_j(y)$$

où

$$\Phi_i(x) = L_i(x) + a_i L_{i+1}(x) + b_i L_{i+2}(x), \quad (2.3.1)$$

et

$$\Phi_j(y) = L_j(y) + a_j L_{j+1}(y) + b_j L_{j+2}(y). \quad (2.3.2)$$

Afin de ne pas encombrer la présentation, on ne considère que les polynômes $\Phi_i(x)$ ainsi que leur propriétés, et pour le passage en $2-D$, on utilise simplement un paramètre de tensorisation [17].

Lemme 2.3.1. *Pour tout $i \geq 0$, il existe une solution unique $\{a_i, b_i\}$, tel que*

$$\Phi_i(x) = L_i(x) + a_i L_{i+1}(x) + b_i L_{i+2}(x),$$

satisfaisant les conditions au bord homogènes de type (Dirichlet et/ou Neumann).

Démonstration. Pour fixer les idées, on donne la démonstration pour les conditions mixtes de type *Dirichlet-Neumann*

$$V_N = \text{span} \{ \Phi'_i(-1) = \Phi_i(1) = 0; i, j = 1, 2, \dots, N-2 \}.$$

Pour les conditions de type *Dirichlet*,

$$V_N = \text{span} \{ \Phi_i(-1) = \Phi_i(1) = 0; i, j = 1, 2, \dots, N-2 \},$$

$$\begin{cases} \Phi'_i(-1) = L'_i(-1) + a_i L'_{i+1}(-1) + b_i L'_{i+2}(-1) = 0, \\ \Phi_i(1) = L_i(1) + a_i L_{i+1}(1) + b_i L_{i+2}(1) = 0. \end{cases}$$

D'après les propriétés des polynômes de Legendre, on obtient le système suivant :

$$\begin{cases} (-1)^i \frac{(i+2)(i+1)}{2} a_i + (\pm 1)^{i+1} \frac{(i+2)(i+3)}{2} b_i = (-1)^{i-1} \frac{i(i+1)}{2} \\ a_i + b_i = -1. \end{cases}$$

Il est facile ensuite de trouver la solution

$$\{a_i, b_i\} = \left\{ \frac{i^2 + (i+1)^2}{(i+1)^2 + (i+2)^2} - 1, -\frac{i^2 + (i+1)^2}{(i+1)^2 + (i+2)^2} \right\},$$

satisfaisant les conditions mixtes de type *Dirichlet-Neumann*, et pour le cas des conditions de *Dirichlet*, on a

$$\{a_i, b_i\} = \{0, -1\}.$$

□

Pour l'approximation numérique suivante, on a besoin du Lemme suivant qui donne des résultats de dérivée seconde et d'intégration qui sont nécessaires pour le calcul et l'implémentation numérique.

Lemme 2.3.2. Soit $\Phi_i(x) = L_i(x) + a_i L_{i+1}(x) + b_i L_{i+2}(x)$, alors $a_{kl} = -(\Phi_i''(x), \Phi_k(x))$, $b_{kl} = (\Phi_i(x), \Phi_k(x))$
où

$$a_{kl} = \begin{cases} b_k (L''_{k+2}, L_l) & \text{si } k = l, \\ a_k (L''_{k+1}, L_l) + b_k a_l (L''_{k+2}, L_{l+1}) & \text{si } l + 1 = k, \\ (L''_k, L_l) + a_k a_l (L''_{k+1}, L_{l+1}) + b_k (L''_{k+2}, L_l) + b_k b_l (L''_{k+2}, L_{l+2}) & \text{si } l + 2 = k, \\ a_l (L''_{l+1}, L_{l+1}) + a_k (L''_{k+1}, L_l) + a_k b_l (L''_{k+1}, L_{l+2}) + b_k b_l (L''_{k+2}, L_{l+1}) & \text{si } l + 1 < k \text{ et } k + l \text{ est pair,} \\ (L''_{l+1}, L_l) + b_k (L''_k, L_{l+2}) + a_k b_l (L''_{k+1}, L_{l+1}) + b_k (L''_{k+2}, L_l) + b_k b_l (L''_{k+2}, L_{l+2}) & \text{si } l + 2 < k \text{ et } k + l \text{ est pair,} \\ 0 & \text{sinon,} \end{cases} \quad (2.3.3)$$

et

$$b_{kl} = b_{lk} = \begin{cases} L_{i,j} + a_i a_j L_{i+1,i+1} + b_i b_j L_{i+2,i+2} & \text{si } k = l, \\ a_j L_{i,i} + a_i b_j L_{i+1,i+1} & \text{si } k = j + 1, \\ b_i L_{j,j} & \text{si } j = k + 2, \\ 0 & \text{sinon,} \end{cases} \quad (2.3.4)$$

où

$$(L_k'', L_l) = \begin{cases} (k + \frac{1}{2})(l(l+1) - k(k+1)) L_{k,k} & \text{si } k+l \text{ est pair,} \\ 0 & \text{sinon.} \end{cases} \quad (2.3.5)$$

Démonstration. on utilise la récurrence. □

2.4 Erreur d'approximation polynomiale en dimension une

Pour la suite, on note par π_N l'opérateur de la projection orthogonale de $L^2(\Lambda)$ sur $\mathbb{P}_N(\Lambda)$.

Ceci signifie que pour toute fonction u de $L^2(\Lambda)$, $\pi_N u$ appartient à $\mathbb{P}_N(\Lambda)$ et vérifie

$$\forall \psi_N \in \mathbb{P}_N(\Lambda), \int_{\Lambda} (u - \pi_N u)(\zeta) \psi_N(\zeta) d\zeta = 0. \quad (2.4.1)$$

Pour des raisons de calcul, on ne considère que la série tronquée d'ordre N , et en appliquant l'opérateur π_N , on obtient

$$\pi_N u \simeq \sum_{i=0}^N u_i L_i.$$

Théorème 2.4.1. *Pour tout entier $m \geq 0$, il existe une constante c positive ne dépendant que de m telle que, pour toute fonction u de $H^m(\Lambda) \cap H_0^1(\Lambda)$, on ait*

$$\|u - \pi_N u\|_{L^2(\Lambda)} \leq c N^{-m} \|u\|_{H^m(\Lambda)}. \quad (2.4.2)$$

Démonstration. [17]. □

On ne saurait trop insister sur l'importance du Théorème 2.4.1, sur lequel reposent tout les résultats de l'analyse de l'erreur a posteriori.

2.5 Erreur D'approximation polynomiale en dimension deux

Pour la suite, on note par Π_N l'opérateur de la projection orthogonale de $L^2(\Omega)$ sur $\mathbb{P}_N(\Omega)$ Le but de ce paragraphe est d'établir des majorations analogues à celles de la section précédente. Le passage de $1 - D$ à $2 - D$ repose essentiellement sur un argument de tensorisation. Ceci signifie que l'on va faire appel à la propriété suivante :

$$\begin{aligned}
L^2(\Omega) &= \left\{ u : \Omega \rightarrow \mathbb{R}; \int_{\Omega} u^2(\mathbf{x}) d\mathbf{x} < +\infty \right\} \\
&= \left\{ u : \Lambda \times \Lambda \rightarrow \mathbb{R}; \int_{-1}^1 \left(\int_{-1}^1 u^2(x, y) dy \right) dx < +\infty \right\} \\
&= \left\{ u : \Lambda \rightarrow L^2(\Lambda); \int_{-1}^1 \|u(x, \cdot)\|_{L^2(\Lambda)}^2 dx < +\infty \right\} = L^2(\Lambda; L^2(\Lambda)).
\end{aligned} \tag{2.5.1}$$

De la même façon, on voit finalement que

$$H^1(\Omega) = L^2(\Lambda; H^1(\Lambda)) \cap H^1(\Lambda; L^2(\Lambda)).$$

Dance ce qui suit le symbole (x) ou (y) après un opérateur monodimensionnel indiquera que l'on fait agir cet opérateur par rapport à la variable x ou y respectivement. Étant donnée une fonction u dans $L^2(\Omega)$, on a par exemple pour presque tout y dans Λ :

$$\int_{\Lambda} \left(u(x, y) - \pi_N^{(x)} u(x, y) \right) L_k(x) dx = 0, \quad 0 \leq k \leq N.$$

On applique cette relation avec u remplacé par $\pi_N^{(y)} u$ et on déduit, pour tout $0 \leq k \leq N$ et $0 \leq l \leq N$ que

$$\begin{aligned}
&\int_{\Omega} \left(u(\mathbf{x}) - \pi_N^{(x)} \circ \pi_N^{(y)} u(\mathbf{x}) \right) L_k(x) L_l(y) d\mathbf{x} \\
&= \int_{\Lambda} L_l(y) \left(\int_{\Lambda} \left(u(x, y) - \pi_N^{(x)} \circ \pi_N^{(y)} u(x, y) \right) L_k(x) dx \right) dy \\
&= \int_{\Lambda} L_l(y) \left(\int_{\Lambda} \left(u(x, y) - \pi_N^{(y)} u(x, y) \right) L_k(x) dx \right) dy \\
&= \int_{\Lambda} L_k(x) \left(\int_{\Lambda} \left(u(x, y) - \pi_N^{(y)} u(x, y) \right) L_l(y) dy \right) dx = 0.
\end{aligned} \tag{2.5.2}$$

Comme $\pi_N^{(x)} \circ \pi_N^{(y)} u$ appartient à $\mathbb{P}_N(\Omega)$ et que les $L_k(x) \cdot L_l(y)$, $0 \leq k, l \leq N$, forment une base de l'espace $\mathbb{P}_N(\Omega)$, on obtient l'identité :

$$\Pi_N = \pi_N^{(x)} \circ \pi_N^{(y)}. \tag{2.5.3}$$

On peut aussi facilement vérifier que les opérateurs $\pi_N^{(x)}$ et $\pi_N^{(y)}$ commutent.

Théorème 2.5.1. *Pour tout entier $m \geq 0$, il existe une constante c positive ne dépendant que de m telle que, pour toute fonction u de $H^m(\Omega) \cap H_0^1(\Omega)$, on ait*

$$\|u - \Pi_N^{1,0} u\|_{L^2(\Omega)} \leq cN^{-m} \|u\|_{H^m(\Omega)}. \tag{2.5.4}$$

Démonstration. [17]. □

2.6 Formules de quadrature

Il est bien connu que les zéros et les extrema des polynômes de Legendre (ou ceux appartenant à une famille quelconque de polynômes orthogonaux) servent à la construction de *formules de quadrature numérique* de grande précision, c'est-à-dire qui sont exacte sur un espace de polynômes de degré élevé : il s'agit principalement des formules de **Gauss** et de **Gauss-Lobatto** [22].

Dans cette partie, on rappelle quelques propriétés des formules de Gauss et de Gauss-Lobatto pour approcher l'intégrale sur Λ qui sont utiles pour la suite. On commence d'abord par le Lemme suivant qui nous permet de connaître la valeur des poids ρ_i pour $i = 1, \dots, N + 1$, qui sont indispensables pour les formules de quadratures, qu'on va étudier dans cette section.

Lemme 2.6.1. *Les poids ρ_0 et ρ_N sont égaux à $\frac{2}{N(N+1)}$ et les poids ρ_j , $1 \leq j \leq N - 1$, sont donnés par*

$$\rho_j = \frac{2}{N(N+1)L_N^2(\xi_j)}. \quad (2.6.1)$$

Démonstration. [17]. □

Définition 2.6.1. *Soit N un entier positif fixé. On pose $\xi_0 = -1$ et $\xi_N = 1$. Il existe un unique ensemble de $N - 1$ points ξ_j de Λ , $1 \leq j \leq N - 1$, et un unique ensemble de $N + 1$ réels ρ_j , $0 \leq j \leq N$, tels que l'égalité suivante ait lieu pour tout polynôme de Φ de $\mathbb{P}_{2N-1}(\Lambda)$:*

$$\forall \Phi \in \mathbb{P}_{2N-1}(\Lambda), \quad \int_{\Lambda} \Phi(\zeta) d\zeta = \sum_{j=1}^N \Phi(\zeta_j) \rho_j. \quad (2.6.2)$$

Définition 2.6.2. *Dans ce qui suit, on notera par ξ_j , $0 \leq j \leq N$, les zéros de $(1 - \zeta^2)L'_N$ rangés dans un ordre croissant et ρ_j , $0 \leq j \leq N$, les poids qui leur sont associés de façon unique d'après le Lemme 2.6.1. La formule de quadrature :*

$$\int_{\Lambda} \Phi(\zeta) d\zeta \simeq \sum_{j=1}^N \Phi(\zeta_j) \rho_j, \quad (2.6.3)$$

est appelée formule de Gauss-Lobatto de type Legendre à $N + 1$ points.

Pour le calcul des valeurs des nœuds ξ_j , $1 \leq j \leq N - 1$, on peut utiliser un logiciel de calcul formel (par exemple Maple). On peut aussi faire appel aux techniques de calcul numérique (voir Annexe B).

Chapitre 3

Estimations de l'erreur dans un cadre abstrait

3.1 Introduction

Dans l'étude des problèmes mathématiques, il existe généralement plusieurs sources d'erreurs :

- Erreurs dues à la modélisation mathématique d'un problème physique.
- Erreurs dues à l'approximation du problème continu et à la qualité de la méthode de résolution de ce dernier.

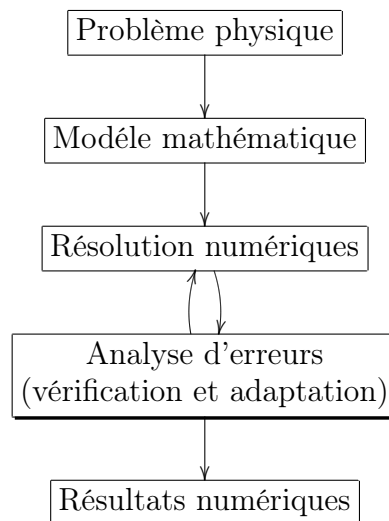


FIG. 3.1 – Analyse numérique d'un problème physique.

Un souci majeur, est de vérifier si un modèle physique est bien résolu ou non, et avec quelle précision sa solution est calculé. Une question immédiate se pose est : qu'est ce qu'on entend par modèle ? La réponse nous amène au rôle fondamental de la vérification.

Il est alors essentiel, de comprendre que le modèle dont on fait référence est un modèle mathématique qui peut être un système d'équations (différentielles ou à dérivées partielles, intégrales ...etc) avec des conditions au bord et/ou initiales de type égalités et/ou inégalités qui constitue la modélisation mathématique du phénomène physique qui nous intéresse.

Peu importe comment représenter un modèle physique par un modèle mathématique caractérisant le phénomène naturel qui nous intéresse. Il existe toujours une différence entre le modèle physique et le modèle mathématique. Ainsi donc que tous les résultats obtenus après l'utilisation de certaines techniques mathématiques, seront entachés d'erreurs dues à la modélisation.

Mais, fondamentalement, les méthodes telles que les méthodes des différences finies, éléments finis, est spectrales, les ondelettes, ou autres méthodes (par exemple le couplage des méthodes précédentes), transforment le problème mathématique sous une forme appropriée qu'on peut traiter par des outils de l'analyse numérique et les techniques du calcul scientifique. Il est cependant clair que le modèle mathématique ainsi transformé perd quelques informations, puisque les problèmes approchés numériquement (en dimension finie) sont différents des problèmes continus (en dimension infinie). Cette différence s'interprète particulièrement par l'erreur d'approximation, qu'on note habituellement par \mathbf{e} .

L'erreur d'approximation numérique \mathbf{e} peut être importante. Parfois des perturbations numériques sont dues tout simplement à la méthode même d'approximation. Ces erreurs sont souvent significatives et difficiles à détecter et à évaluer.

Nous considérons, dans ce chapitre, le problème des approximations par des modèles numériques et l'adaptation des paramètres d'approximations pour minimiser ou néanmoins contrôler l'erreur \mathbf{e} .

Nous commençons par des résultats de majoration de l'erreur a priori. Nous présentons ensuite les aspects pratiques de l'estimation d'erreur a posteriori et de l'adaptation des paramètres d'approximation. Il est important de remarquer que les estimations a priori supposent la connaissance des solutions continues (exactes) et parfois de ses dérivées. Tandis que les estimations a posteriori utilisent les solutions discrètes (approchées) et permettent en pratique l'adaptation des paramètres d'approximation. Il est bien de savoir que dans cette partie, on ne considère que l'estimation de l'erreur d'une manière formelle. Pour fixer les idées, nous considérons une méthode d'approximation conforme.

3.2 Approximation de Galerkin dans un cadre abstrait

L'objet de cette partie, est de construire à l'aide d'un opérateur d'approximation Π_σ et de certaines fonctions $e_1, e_2, \dots, e_{N(\sigma)}$ une solution approchée de la solution exacte. On considère ici des discrétisations construites par une méthode de Galerkin et par souci de généralité, le paramètre de discrétisation est ici noté par σ . On définit par V_σ l'espace approché de V .

Une méthode de Galerkin consiste à fixer $N(\sigma)$ fonctions linéairement indépendantes

$e_1, e_2, \dots, e_{N(\sigma)}$ dans V et d'écrire le sous-espace vectoriel V_σ de V comme étant l'ensemble des combinaisons linéaire des fonctions $e_1, e_2, \dots, e_{N(\sigma)}$ et par la suite chercher une fonction v_σ (par exemple la solution approchée) dans V_σ c'est à dire chercher une combinaison linéaire $v_\sigma = \sum_{i=1}^{N(\sigma)} v_i e_i$, où les v_i sont des nombres réels inconnus.

Définition 3.2.1. *Soit V un espace de Banach et V_σ est une approximation de V , dans le sens suivant : on peut trouver un espace \mathcal{V} dense dans V tel que, pour tout v dans \mathcal{V} , on puisse construire v_σ dans V_σ avec v_σ tend vers v dans V lorsque σ tend vers 0 ou vers $+\infty$. Dans ce qui suit, on considère une approximation conforme dans le sens où V_σ est inclus dans V ou interne [44].*

Pour l'instant, on ne s'intéresse pas à des résultats d'approximation locale, mais plutôt à des résultats d'approximation globale. On définit l'opérateur Π_σ de la manière suivante : pour tout v suffisamment régulier sur Ω , $\Pi_\sigma v$ appartient à V_σ et

$$\Pi_\sigma v = \sum_{i=1}^{N(\sigma)} v_i e_i.$$

On considère le problème (\mathcal{E}_a) où $a(u, v) = \langle Au, v \rangle_{V, V'}$. Pour donner une portée générale aux résultats ci-dessous nous n'allons pas utiliser explicitement la coercivité de la forme bilinéaire $a(\cdot, \cdot)$ mais uniquement sa propriété de stabilité résultant du Théorème de Nečas [25] i.e., il existe une constante $c > 0$ telle que

$$\inf_{u \in V} \sup_{v \in V} \frac{a(u, v)}{\|u\|_V \|v\|_V} \geq c. \quad (3.2.1)$$

Notons $V_\sigma \subset V$, l'espace d'approximation correspondant ; signifiant que l'approximation conforme. Dans cas, le problème approché s'écrit sous la forme

$$\begin{aligned} \text{Trouver } u_\sigma \in V_\sigma \text{ tel que} \\ a(u_\sigma, v_\sigma) = (f, v_\sigma), \quad \forall v_\sigma \in V_\sigma. \end{aligned} \quad (3.2.2)$$

3.2.1 Analyse a priori de l'erreur

Lemme 3.2.1. *Soient u et u_σ deux solutions respectivement de (\mathcal{E}_a) et (3.2.2), alors on a*

$$\forall v_\sigma \in V_\sigma, a(u - u_\sigma, v_\sigma) = 0, \quad (3.2.3)$$

qu'on appelle équation de projection par rapport à $a(\cdot, \cdot)$.

Démonstration. Comme on a supposé que

$$\begin{aligned} a(u, v_\sigma) &= (f, v_\sigma) \\ &= a(u_\sigma, v_\sigma), \quad \forall v_\sigma \in V_\sigma, \end{aligned}$$

il vient alors

$$a(u - u_\sigma, v_\sigma) = 0.$$

□

La relation d'orthogonalité (3.2.3) implique

$$u_\sigma = \Pi_\sigma u,$$

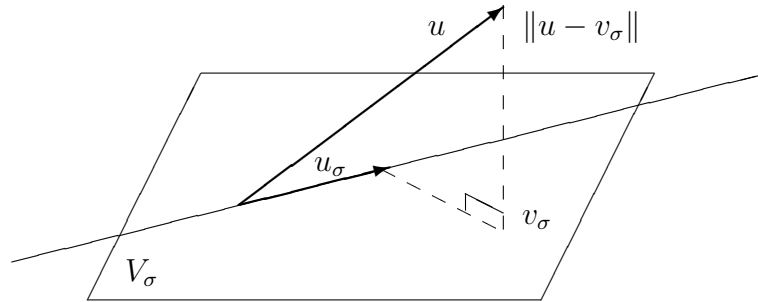


FIG. 3.2 – Représentation de l'approximation de Galerkin.

On dit alors, que la solution approchée u_σ est simplement la projection orthogonale pour le produit scalaire $a(\cdot, \cdot)$ de la solution exacte u sur le sous-espace V_σ . Une estimation d'erreur en norme $\|\cdot\|_V$ est donnée par le résultat suivant.

Remarque 3.2.1. Dans le cas, où on suppose que la forme bilinéaire $a(\cdot, \cdot)$ est continue, symétrique et coercive, alors elle définit un produit scalaire et une norme

$$\|u\|_e = (a(u, u))^{\frac{1}{2}}, \quad \forall u \in V. \quad (3.2.4)$$

appelée norme de l'énergie et équivalente à la norme de V puisque

$$\sqrt{c} \|u\|_V \leq \|u\|_e \leq \sqrt{\|a\|} \|u\|_V, \quad \forall u \in V.$$

Lemme 3.2.2 (Lemme de Céa). Avec les hypothèses données ci-dessus, on a

$$\|u - u_\sigma\|_V \leq \left(1 + \frac{\|a\|}{c_\sigma}\right) \inf_{w_\sigma \in V_\sigma} \|u - w_\sigma\|_V. \quad (3.2.5)$$

Démonstration. Soit $w_\sigma \in V_\sigma$. De la relation d'orthogonalité (3.2.3), nous déduisons

$$a(u_\sigma - w_\sigma, v_\sigma) = a(u - w_\sigma, v_\sigma), \quad \forall v_\sigma \in V_\sigma.$$

En utilisant la propriété (3.2.1) et la continuité de $a(\cdot, \cdot)$, il vient

$$\begin{aligned} c_\sigma \|u_\sigma - w\|_V &\leq \sup_{v_\sigma \in V_\sigma} \frac{a(u_\sigma - w_\sigma, v_\sigma)}{\|v_\sigma\|} = \sup_{v_\sigma \in V_\sigma} \frac{a(u - w_\sigma, v_\sigma)}{\|v_\sigma\|} \\ &\leq \|a\| \|u - w_\sigma\|_V. \end{aligned}$$

Nous concluons avec l'inégalité triangulaire. \square

Corollaire 3.2.1. *Soit $\{V_\sigma\}_\sigma$ une famille de sous espace de dimension finie de V et supposons que*

$$\forall w \in V, \quad \inf_{w_\sigma \in V_\sigma} \|w - w_\sigma\|_V \rightarrow 0 \text{ avec } \sigma \rightarrow \varepsilon^1$$

Remarque 3.2.2. *Dans le cas où la forme bilinéaire $a(\cdot, \cdot)$ est continue et coercive de constante de coercivité c , l'estimation d'erreur (3.2.5) peut être améliorée. La relation d'orthogonalité (3.2.3) implique*

$$a(\mathbf{e}_u, \mathbf{e}_u) = a(\mathbf{e}_u, u_\sigma),$$

où $\mathbf{e}_u = u - u_\sigma$.

En utilisant la coercivité et la continuité de $a(\cdot, \cdot)$, nous obtenons la majoration (3.2.5) avec la constante $\frac{\|a\|}{c}$. Cette majoration peut encore être améliorée si $a(\cdot, \cdot)$ est en plus symétrique. En effet, comme u_σ est la projection orthogonale de u sur V_σ pour le produit scalaire induit par $a(\cdot, \cdot)$, nous avons la relation de Pythagore,

$$\begin{aligned} \forall w_\sigma \in V_\sigma, \quad a(\mathbf{e}_u, \mathbf{e}_u) &= \|\mathbf{e}_u\|_e^2 \\ &\leq \|u - w_\sigma\|_e^2 = a(u - w_\sigma, u - w_\sigma). \end{aligned}$$

Puisque $\forall w_\sigma \in V_\sigma$, nous avons

$$\begin{aligned} c \|\mathbf{e}_u\|_V^2 &\leq a(\mathbf{e}_u, \mathbf{e}_u) \\ &\leq a(u - w_\sigma, u - w_\sigma) \\ &\leq \|a\| \|u - w_\sigma\|_V^2, \end{aligned}$$

ce qui donne finalement

$$\|\mathbf{e}_u\|_V \leq \sqrt{\frac{\|a\|}{c}} \inf_{w_\sigma \in V_\sigma} \|u - w_\sigma\|_V.$$

¹par exemple ε est égale à 0 en éléments finis, et $+\infty$ en méthode spectrale.

Théorème 3.2.1. Soient C et c les deux constantes intervenant dans l'hypothèse de continuité et de coercivité de la forme bilinéaire $a(\cdot, \cdot)$, dans ce cas on a la majoration suivante :

$$\|u - u_\sigma\| \leq \frac{C}{c} \inf_{v_\sigma \in V_\sigma} \|u - v_\sigma\|,$$

si, en outre, $a(\cdot, \cdot)$ est symétrique, on a :

$$a(u - u_\sigma, v_\sigma) = 0 \quad \forall v_\sigma \in V_\sigma, \quad (3.2.6)$$

ce qui signifie que u_σ est la projection de u dans V au sens du produit scalaire $a(\cdot, \cdot)$. Dans ce cas, on a

$$\|u - u_\sigma\| \leq \sqrt{\frac{C}{c}} \inf_{v_\sigma \in V_\sigma} \|u - v_\sigma\|.$$

Dorénavant, le choix d'une méthode d'approximation (éléments finis, méthode psctrle ou autres) s'impose [17].

3.2.2 Analyse a posteriori de l'erreur

Dans les deux dernières décennies, les théories mathématiques et les techniques de calcul scientifique ont été développées pour estimer l'erreur dans l'approximation numérique des solutions des problèmes aux limites et/ou de valeurs initiales. La problématique est généralement connue sous le nom de "**estimation de l'erreur a posteriori**". Fondamentalement, l'erreur a posteriori dépend de la solution numérique trouvée après calcul et des paramètres d'approximation. Elle dépend, de la solution approchée et d'autres quantités connues. Elle fournit quelques informations pour améliorer la qualité de la solution numérique permettant d'évaluer l'erreur de calcul et pourquoi pas la réduire.

Ces informations peuvent se manifester sous plusieurs formes typiques :

- Comme des bornes supérieures et/ou inférieures de l'approximation des erreurs calculées dans des normes appropriées,
- Comme des évaluations quantitatives des erreurs.

Clairement, si l'erreur peut être calculée ou même estimée, alors il devra être possible d'améliorer la précision de calcul et donc de réduire l'erreur c'est-à-dire l'estimation de l'erreur doit en quelque sorte contrôler l'erreur d'approximation. Ce fait est central dans les méthodes adaptatives. Ceci conduit à la modification des paramètres d'approximation qui permettent d'améliorer la précision du calcul.

Suivant la nature du problème, l'erreur de l'approximation \mathbf{e} peut dépendre de la variable spatiale \mathbf{x} et/ou de la variable temporelle \mathbf{t} et elle est définie habituellement par

$$\mathbf{e} = u - u_\sigma, \quad (3.2.7)$$

où σ est le paramètre d'approximation, et u est la solution exacte (la solution continue) d'un certain problème et u_σ est l'approximation (la solution discrète) de u .

Dans certains problèmes, on n'est pas forcément intéressé par l'estimation de la quantité \mathbf{e} (dans une norme convenable), mais parfois par l'estimation de l'erreur réalisée par une certaine application Q (par exemple pour lorsqu'il s'agit du critère du gradient [89]). Donc on a besoin parfois d'évaluer l'erreur \mathcal{E} , définie par

$$\mathcal{E} = Q(u) - Q(u_\sigma).$$

L'application $Q(u)$, peut être linéaire ou non, et dans le cas où l'application Q est linéaire, on a

$$Q(u) - Q(u_\sigma) = Q(\mathbf{e}),$$

ou

$$Q(u) = Q(u_\sigma) + Q(\mathbf{e}).$$

En général, le rôle de la méthode de l'estimation de l'erreur a posteriori est d'évaluer l'erreur $Q(\mathbf{e})$ ou d'obtenir une borne supérieure η_{sup} et si c'est possible avoir aussi une borne inférieure η_{inf} , telle que

$$\eta_{\text{inf}} \leq Q(u) - Q(u_\sigma) \leq \eta_{\text{sup}}.$$

Bien sûr, la solution exacte u , est en général inconnue et parfois impossible à la calculer explicitement. Si la quantité $Q(\mathbf{e})$ est bien évaluée, cela nous permet aussi de bien approximer la quantité $Q(u)$, donc de mieux calculer $Q(u_\sigma)$, ainsi les deux bornes η_{inf} et η_{sup} seront bien évaluées numériquement en fonction de u_σ .

Enfin, le but principal de la méthode d'estimation de l'erreur est de trouver et d'estimer la quantité η qu'on appelle généralement "**l'estimateur**" qui est une majoration de l'erreur \mathbf{e} obtenue dans une norme $\|\cdot\|$ appropriée. Donc le but est de calculer la quantité η qui approche ou qui majore la quantité $\|\mathbf{e}\|$. L'idéal pour donner une estimation de la quantité η est de trouver deux constante C_1 et C_2 , tel que

$$C_1 \|\mathbf{e}\| \leq \eta \leq C_2 \|\mathbf{e}\|. \quad (3.2.8)$$

Par suite lorsque \mathbf{e} est petite, on déduit que η est voisin de zéro, à condition bien sûr que la constante C_2 , ne soit pas trop grande. Inversement si la quantité η est petite, cela entraîne que l'erreur \mathbf{e} est petite, à condition bien sûr que la constante C_1 , ne soit pas trop petite. Bien sûr, il est inutile de préciser qu'en général sauf cas particulier, l'erreur \mathbf{e} est rarement connue explicitement. Tandis que la qualité de certains estimateurs d'erreur η peuvent être évalués expérimentalement et aussi théoriquement mais toujours en fonction du paramètre d'approximation et de la solution approchée u_σ .

3.2.3 Indicateur par résidu

Nous avons établi l'estimation a priori suivante

$$\|u - u_\sigma\|_V \leq \inf_{v_\sigma \in V_\sigma} \|u - v_\sigma\|_V.$$

Cette majoration entre dans le cadre des estimations a priori car elle fait intervenir la quantité inconnue $\|u\|_V$ [17]. Pour obtenir une estimation a posteriori, il nous faut revenir à la propriété de stabilité (3.2.1). Nous avons

$$\begin{aligned} c \|u - u_\sigma\|_{H^1(\Omega)} &\leq \sup_{v \in V} \frac{a(u - u_\sigma, v)}{\|v\|_V} \\ &\leq \sup_{v \in V} \frac{\langle A(u - u_\sigma), v \rangle_{V', V}}{\|v\|_V} \\ &\leq \|f + Au_\sigma\|_{V'}. \end{aligned}$$

Nous venons ainsi de trouver une première estimation a posteriori.

Proposition 3.2.1. *Soient u et u_σ deux solutions respectivement de (\mathcal{E}_a) et de (3.2.2), alors on a*

$$\forall \sigma, \quad \|u - u_\sigma\|_V \leq \frac{1}{c} \|f - Au_\sigma\|_{V'}.$$

Cette estimation entre dans le cadre de l'estimation a posteriori de l'erreur, mais elle n'est pas très intéressante car elle fait intervenir la norme $\|\cdot\|_{V'}$ qui est très difficile à évaluer en pratique. Toutefois, l'idée de l'intégration par parties est bonne car elle fait disparaître la fonction inconnue u dans le second membre, mais il n'est pas très difficile de se débarrasser de la norme $\|\cdot\|_{V'}$.

Donc, pour résumer, puisque $a(u - u_\sigma, v_\sigma) = 0$, alors pour tout $v_\sigma \in V_\sigma$, l'inégalité de stabilité (3.2.1) donne

$$\forall v_\sigma \in V_\sigma, \quad \|u - u_\sigma\|_V \leq \frac{1}{c} \sup_{v \in V} \frac{a(u - u_\sigma, v - v_\sigma)}{\|v\|_V}.$$

Nous pouvons développer le numérateur dans le second membre comme suit :

$$\begin{aligned} a(u - u_\sigma, v - v_\sigma) &= a(u - u_\sigma, v) \\ &= \int_{\Omega} f(v - v_\sigma) dx - a(u_\sigma, v - v_\sigma) \\ &= \int_{\Omega} f(v - v_\sigma) dx - \int_{\Omega} Au_\sigma (v - u_\sigma) dx \\ &= \int_{\Omega} (f - Au_\sigma)(v - v_\sigma) dx. \end{aligned}$$

Ici également, un choix d'une méthode d'approximation (éléments finis, méthodes psctrale ou autres) s'impose (voir par exemple [17]).

Pour l'évaluation à posteriori de l'erreur, on peut aussi faire appel à la théorie de dualité en analyse convexe, (voir annex C).

3.3 Représentation graphique de l'erreurs

Pour mieux comprendre le comportement et l'évolution de l'erreur d'approximation en fonction des paramètres d'approximation, on fait appel à la représentation graphique. Les représentations graphiques, sont représentées dans un repère dans lequel les deux axes (abscisse et ordonnée) sont gradués selon une échelle logarithmique, et cela afin d'avoir plus d'informations sur le plan et en plus si on suppose par exemple qu'une erreur est sous la forme $e = C\sigma^\alpha$, où C est une constante indépendante du paramètre d'approximation σ et α est un réel, alors en passant à l'échelle $\log - \log$, on obtient $\log e = \alpha \log(C\sigma)$, qui est l'équation d'une droite ayant pour pente α , ce qui peut être très utile si on veut par exemple comparer la vitesse de convergence entre deux méthodes différentes, et c'est celle qui a la plus forte pente bien sûr qui converge le mieux.

Chapitre 4

Analyse numérique du problème de commande optimale (\mathcal{P})

4.1 Introduction

Pour la résolution numérique du problème de commande optimale (\mathcal{P}), il y a habituellement deux approches :

1. Une approche indirecte ou après optimisation, où il s'agit d'approximer le système des conditions d'optimalités (\mathcal{S}). On obtient un système d'équations linéaires et/ou non linéaires, qu'on résout par exemple par des méthodes itératives ou directes qui dépendent bien sûr de la nature du système (\mathcal{S}), si il est linéaire ou non.
2. Une approche directe où avant optimisation, on s'agit d'approximer le problème (\mathcal{P}) (approximation de la fonction objectif et l'équation de l'état), écrit sous forme variationnelle ou non (dépendant bien sûr de la méthode d'approximation choisie). Par suite, on obtient un problème d'optimisation qu'on peut résoudre par des techniques habituelles d'optimisations.

Il existe plusieurs méthodes d'approximation pour la résolution numérique des problèmes de commande optimale gouvernés par des équations différentielles [52] pour la méthode des différences finies [45] pour la méthode des éléments finis et [35] pour les méthodes d'approximations spectrales.

L'analyse numérique de la méthode d'approximation par éléments finis pour la résolution des problèmes de commande optimale gouvernés par des équations aux dérivées partielles a eu un grand succès ces dernières années grâce au développement de l'analyse adaptative reposant sur l'estimation de l'erreur a posteriori [5]. Par ailleurs les estimations de l'erreur a priori sont données dans [9].

Malgré le succès des méthodes d'approximation spectrale dans la résolution des équations aux dérivées partielles, leur application pour la résolution numérique des problèmes de commande optimale a été peu développée [86].

4.2 Approximation du problème de commande optimale (\mathcal{P})

Le Chapitre 2, a introduit les propriétés fondamentales des problèmes de Legendre ainsi que la méthode d'approximation de Galerkin. Maintenant, on s'intéresse à la discrétisation du problème (\mathcal{P}) par la méthode de Galerkin utilisant les polynômes de Legendre qu'on appelle la méthode spectrale de Galerkin-Legendre, on développera également les conditions nécessaires et suffisantes d'optimalité du système (\mathcal{S}).

On commence par introduire l'approximation (\mathcal{P}_N) du problème de commande optimale (\mathcal{P}) :

$$\left\{ \begin{array}{l} \min_{u_N \in U_{ad}^N} J(y_N, u_N) = \frac{1}{2} \int_{\Omega} (y_N(u_N) - y_d)^2 d\mathbf{x} + \frac{\alpha}{2} \int_{\Omega} u_N^2 d\mathbf{x}, \\ a(y_N(u_N), w_N) = (f + Bu_N, w_N)_{L^2(\Omega)}, \quad \forall w_N \in V_N. \end{array} \right. \quad (\mathcal{P}_N)$$

En utilisant les mêmes techniques (*Théorème, Corollaire,...*), arguments et hypothèses que dans la section 1.4, on peut toujours démontrer qu'en dimension finie le problème (\mathcal{P}_N) admet une solution unique. On déduit aussi les conditions nécessaire et suffisantes d'optimalité. Ainsi pour que la paire $(y_N(u_N), u_N)$ soit solution optimale du problème (\mathcal{P}) il faut qu'il existe $p_N(u_N)$ l'état adjoint, tel que le triplet $(y_N(u_N), p_N(u_N), u_N)$ satisfait le système des conditions d'optimalité (\mathcal{S}_N) suivant

$$a(y_N(u_N), w_N) = (f + Bu_N, w_N)_{L^2(\Omega)}, \quad \forall w_N \in V_N, \quad (4.2.1a)$$

$$a(p_N(u_N), q_N) = (y_N(u_N) - y_d, q_N)_{L^2(\Omega)}, \quad \forall q_N \in V_N, \quad (4.2.1b)$$

$$(\alpha u_N + B^* p_N(u_N), v_N - u_N)_{L^2(\Omega)} \geq 0, \quad \forall v_N \in U_{ad}^N. \quad (4.2.1c)$$

Dans le cas particulier où $U_{ad}^N = \mathbb{P}_N(\Omega) \cap L^2(\Omega)$ (i.e. l'espace des commandes admissibles U_{ad} est égal à tout l'espace $L^2(\Omega)$), l'équation de projection (4.2.1c) s'écrit sous la forme

$$\alpha u_N + B^* p_N(u_N) = 0.$$

Par suite, le système des conditions d'optimalité (\mathcal{S}_N) s'écrit sous la forme

$$a(y_N(u_N), w_N) = (f + Bu_N, w_N)_{L^2(\Omega)}, \quad \forall w_N \in V_N, \quad (4.2.2a)$$

$$a(p_N(u_N), q_N) = (y_N(u_N) - y_d, q_N)_{L^2(\Omega)}, \quad \forall q_N \in V_N, \quad (4.2.2b)$$

$$\alpha u_N + B^* p_N(u_N) = 0. \quad (4.2.2c)$$

En substituant l'équation de projection ¹ (4.2.2c), dans l'équation d'état (4.2.2a), on obtient le système équivalent suivant

¹Par abus de langage, on préfère encore garder cette terminologie pour cette équation

$$a(y_N(u_N), w_N) = \left(f - \frac{1}{\alpha} B^* B p_N(u_N), w_N\right), \quad \forall w_N \in V_N, \quad (4.2.3a)$$

$$a(p_N(u_N), q_N) = (y_N(u_N) - y_d, q_N), \quad \forall q_N \in V_N. \quad (4.2.3b)$$

Dans le but de faciliter les calculs pour les estimations des erreurs, il convient d'introduire de nouvelles fonctions $y(u_N)$, $p(u_N)$, u_N qu'on appellera fonctions auxiliaires, tel que le triplet $(y(u_N), p(u_N), u_N)$ soit solution du système d'équations suivant qu'on appellera système intermédiaire

$$a(y(u_N), w) = (f + B u_N, w)_{L^2(\Omega)}, \quad \forall w \in H_0^1(\Omega), \quad (4.2.4a)$$

$$a(p(u_N), q) = (y(u_N) - y_d, q)_{L^2(\Omega)}, \quad \forall q \in H_0^1(\Omega), \quad (4.2.4b)$$

$$(\alpha u_N + B^* p(u_N), v - u_N)_{L^2(\Omega)} \geq 0, \quad \forall v \in U_{ad}. \quad (4.2.4c)$$

Dans le cas particulier où $U_{ad} = L^2(\Omega)$, l'équation (4.2.4c) s'écrit sous la forme

$$\alpha u_N + B^* p(u_N) = 0.$$

Par suite, le système des conditions d'optimalité (4.2.4) s'écrit sous la forme

$$a(y(u_N), w) = (f + B u_N, w)_{L^2(\Omega)}, \quad \forall w \in H_0^1(\Omega), \quad (4.2.5a)$$

$$a(p(u_N), q) = (y(u_N) - y_d, q)_{L^2(\Omega)}, \quad \forall q \in H_0^1(\Omega), \quad (4.2.5b)$$

$$\alpha u_N + B^* p(u_N) = 0. \quad (4.2.5c)$$

En substituant l'équation de projection (4.2.5c) dans l'équation d'état (4.2.5a), on obtient le système équivalent suivant

$$a(y(u_N), w) = \left(f - \frac{1}{\alpha} B^* B p(u_N), w\right), \quad \forall w \in H_0^1(\Omega), \quad (4.2.6a)$$

$$a(p(u_N), q) = (y(u_N) - y_d, q), \quad \forall q \in H_0^1(\Omega). \quad (4.2.6b)$$

où

$$a(v_N, w_N) = \int_{\Omega} \nabla v_N \nabla w_N \, d\mathbf{x} \quad \forall v_N, w_N \in V_N.$$

4.2.1 Approximation numérique de l'équation d'état

Dans ce paragraphe, on considère l'approximation numérique de l'équation d'état (4.2.2a) par la méthode dite de *Galerkin-Legendre*. Pour le choix de la base, il faut qu'elle satisfasse les conditions au bord imposées dans le problème (\mathcal{P}) qui sont, bien entendu, les conditions homogènes de type Dirichlet. Il ne faut pas non plus oublier qu'on doit toujours chercher une solution optimale $(y(u), u, p(u))$ sous forme d'une combinaison linéaire de polynômes de Legendre en $2 - D$, comme suit

$$u(\mathbf{x}) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} b_{ij} \Phi_i(x) \otimes \Phi_j(y), \quad (4.2.7)$$

$$y(u)(\mathbf{x}) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_{ij} \Phi_i(x) \otimes \Phi_j(y). \quad (4.2.8)$$

Comme V_N est de dimension finie, et pour des raisons de calcul, on ne considère que la série tronqué d'ordre N de (4.2.7) et (4.2.8) écrit respectivement sous la forme

$$u_N(\mathbf{x}) = \Pi_N u(\mathbf{x}) = \sum_{i=0}^{N-2} \sum_{j=0}^{N-2} b_{ij} \Phi_i(x) \otimes \Phi_j(y), \quad (4.2.9)$$

$$y_N(u)(\mathbf{x}) = \Pi_N y(x, y) = \sum_{i=0}^{N-2} \sum_{j=0}^{N-2} a_{ij} \Phi_i(x) \otimes \Phi_j(y), \quad (4.2.10)$$

$$f_{ij} = (f, \Psi_{ij}(\mathbf{x})). \quad (4.2.11)$$

Il s'agit de trouver u_N et $y_N(u_N)$ qui vérifient le système (4.2.2) tel que $u_N \in V_N$ et $y_N(u_N) \in V_N$; c'est-à-dire, trouver la valeur des inconnues $(a_{ij})_{0 \leq i, j \leq N-2}$ et $(b_{ij})_{0 \leq i, j \leq N-2}$.

$$(\Delta y_N(u)(\mathbf{x}), \Psi_{kl}(\mathbf{x})) + (u_N(\mathbf{x}), \Psi_{kl}(\mathbf{x})) = (f(\mathbf{x}), \Psi_{kl}(\mathbf{x})), \quad (4.2.12)$$

où $\Psi_{kl}(\mathbf{x}) = \Phi_k(x) \Phi_l(y)$ pour $k, l = 0, \dots, N-2$. Il est utile de préciser que les éléments de la famille des polynômes $\{\Psi_{kl}(\mathbf{x})\}_{0 \leq k, l \leq N-2}$ sont les fonctions tests qui vérifient les conditions au bord du problème (\mathcal{P}), où on peut écrire encore que $\Psi_{kl}(\mathbf{x})$ est dans V_N . On remarque, que les conditions de régularité concernant la fonction de commande u et la fonction d'état $y(u)$ sont les mêmes, et c'est pour cette raison qu'on choisit la même base, ce qui n'est pas vrai en règle générale.

Par l'application du produit de tensorisation dans l'équation (4.2.12), cette dernière s'écrit sous la forme

$$\begin{aligned} & \sum_{i=0}^{N-2} \sum_{j=0}^{N-2} a_{ij} \left(\int_{\Omega} \Phi_i''(x) \Phi_k(x) d\mathbf{x} \otimes \int_{\Omega} \Phi_j(y) \Phi_l(y) d\mathbf{x} + \right. \\ & \quad \left. + \int_{\Omega} \Phi_i(x) \Phi_k(x) d\mathbf{x} \otimes \int_{\Omega} \Phi_j''(y) \Phi_l(y) d\mathbf{x} \right) + \\ & \quad + \sum_{i=0}^{N-2} \sum_{j=0}^{N-2} b_{ij} \left(\int_{\Omega} \Phi_i(x) \Phi_k(x) d\mathbf{x} \otimes \int_{\Omega} \Phi_l(y) \Phi_j(y) d\mathbf{x} \right) = \\ & \quad = \int_{\Omega} f(\mathbf{x}) \Phi_k(x) \Phi_l(y), \text{ où } k, l = 0, 1, \dots, N-2. \end{aligned} \quad (4.2.13)$$

La relation (4.2.13) semble compliquée à manipuler, mais dès qu'on passe à l'écriture matricielle, on y voit plus clair. Pour cela, on doit commencer par donner la définition d'une notion très importante et très utile sur laquelle repose le calcul numérique et les méthodes spectrales en 2 – D à savoir le produit tensoriel matricielle.

Définition 4.2.1. Soit $A = (a_{ij})$ et $B = (b_{ij})$, deux matrices de dimension $n \times m$ et $p \times q$, respectivement leur produit matricielle est défini par la matrice

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1m}B \\ a_{21}B & a_{11}B & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & \dots & \dots & a_{nm}B \end{pmatrix}, \tag{4.2.14}$$

de dimension $np \times mq$.

Remarque 4.2.1. Le produit tensoriel matricielle (4.2.14) s'obtient on multipliant chaque coefficient (élément par élément) de la matrice A par la matrice B .

Finalement, on obtient le système d'équations

$$\mathbf{A}\mathbf{y} + \mathbf{B}\mathbf{u} = \mathbf{f}, \tag{4.2.15}$$

où \mathbf{A} est une matrice de dimension $(N + 1) \times (N + 1)$ et \mathbf{B} est une matrice de dimension $(N + 1) \times (N + 1)$, $\mathbf{f} = (f_0, \dots, f_i, \dots, f_{N-2})$, $\mathbf{y} = (a_{00}, \dots, a_{ij}, \dots, a_{N-2N-2})$, et $\mathbf{u} = (b_{00}, \dots, b_{ij}, \dots, b_{N-2N-2})$. Par suite, on définit la matrice des contraintes \mathbf{C} de dimension $(N + 1) \times 2(N + 1)$ sous la forme

$$\mathbf{C} = (\mathbf{A} \mid \mathbf{B}), \tag{4.2.16}$$

qu'on représente schématiquement par

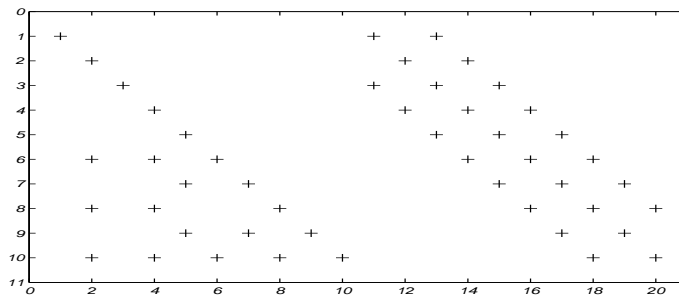


FIG. 4.1 – La matrice de contraintes \mathbf{C} .

où les éléments non nuls sont représentés par des croix. On remarque d'après la représentation de la matrice \mathbf{C} , que la matrice \mathbf{A} est une matrice triangulaire inférieure et que la matrice \mathbf{B} est une matrice tridiagonale.

4.2.2 Approximation numérique de la fonction objectif

Pour l'approche directe, on est obligé de considérer l'approximation de la fonction objectif $J(\cdot)$. En substituant les deux relations (4.2.10) et (4.2.9) dans la fonction objectif $J(\cdot)$, il vient

$$J_N(y_N, u_N) = \frac{1}{2} \int_{\Omega} \left(\sum_{i,j=0}^{N-2} a_{ij} \Phi_{ij}(\mathbf{x}) - y_d \right)^2 d\mathbf{x} + \frac{\alpha}{2} \int_{\Omega} \left(\sum_{i,j=0}^{N-2} b_{ij} \Phi_{ij}(\mathbf{x}) \right)^2 d\mathbf{x}. \quad (4.2.17)$$

Des techniques simples et élémentaires de calcul donnent

$$\int_{\Omega} y_N^2(\mathbf{x}) d\mathbf{x} = \sum_{i,j=0}^{N-2} \sum_{k,l=0}^{N-2} \Phi_{ij} \otimes \Phi_{kl} a_{ij} a_{ji}, \quad (4.2.18)$$

et

$$\int_{\Omega} u_N^2(\mathbf{x}) d\mathbf{x} = \sum_{i,j=0}^{N-2} \sum_{k,l=0}^{N-2} \Phi_{ij} \otimes \Phi_{kl} b_{ij} b_{ji}. \quad (4.2.19)$$

Enfin, l'approximation $J_N(\cdot)$ de la fonction objectif $J(\cdot)$ s'écrit sous la forme

$$J_N(a_{ij}, b_{ij}) = \underbrace{\frac{1}{2} \left\{ \left(\sum_{i,j=0}^{N-2} \sum_{k,l=0}^{N-2} \Phi_{ij} \otimes \Phi_{kl} (a_{ij} a_{kl} + \alpha b_{ij} b_{kl}) \right) \right\}}_{\frac{1}{2} \zeta^t \mathbf{H} \zeta} - \underbrace{\int_{\Omega} \sum_{i,k=0}^{N-2} a_{ik} y_d \Phi_{ik}(\mathbf{x}) d\mathbf{x}}_{\zeta^t \mathbf{c}} + \underbrace{\frac{1}{2} \int_{\Omega} y_d^2 d\mathbf{x}}_{\mathbf{d}}. \quad (4.2.20)$$

Dans ce cas, la fonction objective $J_N(\cdot)$ s'écrit sous la forme matricielle comme suit

$$J_N(a_{ij}, b_{ij}) = \frac{1}{2} \zeta^t \mathbf{H} \zeta + \zeta^t \mathbf{c} + \mathbf{d}, \quad (4.2.21)$$

où \mathbf{H} est la matrice hessienne qui est une matrice symétrique définie positive de dimension $2(N+1) \times 2(N+1)$, \mathbf{c} est le coût final qui est un vecteur $2(N+1)$ composantes, \mathbf{d} une constante, la matrice \mathbf{H} peut s'écrire aussi la forme

$$\left(\begin{array}{c|c} \mathbf{G} & \mathbf{O} \\ \hline \mathbf{O} & \alpha \mathbf{G} \end{array} \right),$$

où la matrice \mathbf{G} est une matrice symétrique définie positive de dimension $(N+1) \times (N+1)$ et $\zeta = \begin{pmatrix} \mathbf{y} \\ \mathbf{u} \end{pmatrix}$. La matrice \mathbf{H} est représentée schématiquement par

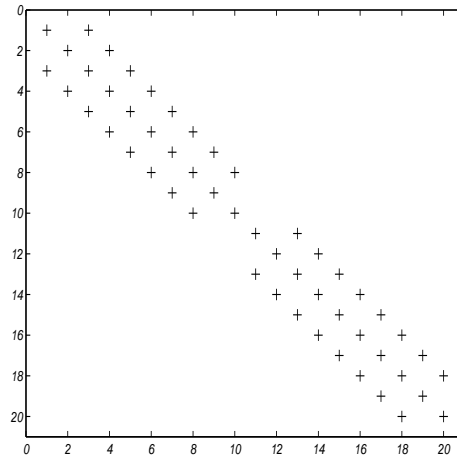


FIG. 4.2 – La matrice globale \mathbf{H} .

où les croix représentent les éléments non nuls.

On remarque que les deux matrices \mathbf{H} et \mathbf{G} sont deux matrices tridiagonale.

Finalement, on obtient le problème de programmation quadratique suivant

$$\begin{cases} \min \frac{1}{2} \zeta^t \mathbf{H} \zeta + \zeta^t \mathbf{c} + \mathbf{d}, \\ \mathbf{C} \zeta = \mathbf{f}, \end{cases} \quad (\mathcal{Q})$$

qu'on peut résoudre par les techniques habituelles d'optimisation. Les résultats sont obtenus à l'aide du code **E04NFF** de la bibliothèque Nag Mark 19 (pour plus de détail sur le code **E04NFF** et sur l'algorithme utilisé, on peut se référer au manuel d'utilisation de la bibliothèque Nag Mark 19), où il faut savoir que le choix de la méthode d'optimisation n'est pas notre soucis principal.

Résultats numériques

Pour l'application numérique, on considère $\Omega = [0, 1] \times [0, 1]$, $y_d = 1$, $f = 0$, $A = -\Delta$ et $B = I$ (l'identité), $\alpha = 1$. Pour l'utilisation des polynômes de Legendre, on doit transformer le domaine $[0, 1] \times [0, 1]$ au domaine $[-1, 1] \times [-1, 1]$, en utilisant le changement de variable classique

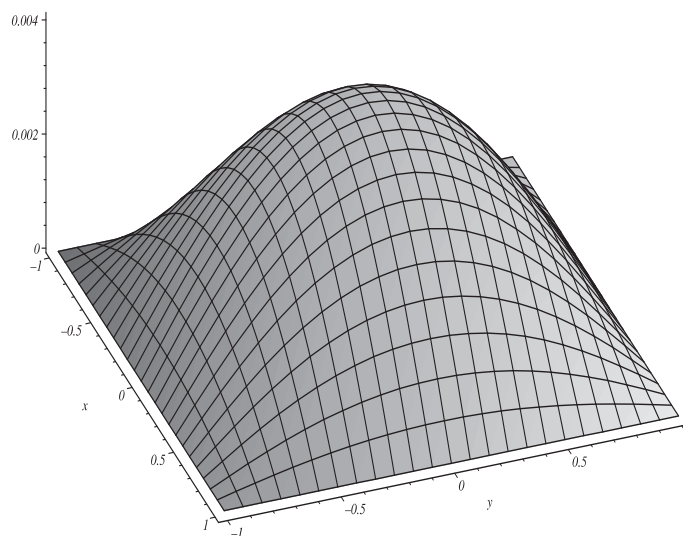
$$\tau(X, Y) = (x, y) \text{ où } x = \frac{1}{2}(1 + X) \text{ et } y = \frac{1}{2}(1 + Y),$$

tel que, le couple (x, y) est dans $[0, 1] \times [0, 1]$ alors le couple (X, Y) est dans $[-1, 1] \times [-1, 1]$. Pour ne pas encombrer la notation, on préfère conserver la même notation du domaine Ω , dans ce cas on obtient les résultats numériques suivant le Tableau 4.1

Degré N	Nbr. Var. ¹	Nbr. Cont. ²	Valeurs de J_N
2	18	9	0.4991508446
3	32	16	0.4991508446
4	50	25	0.4991509031
5	72	36	0.4991509031
6	98	49	0.4991509038
7	128	64	0.4991509038
8	162	81	0.4991509038
9	200	100	0.4991509038

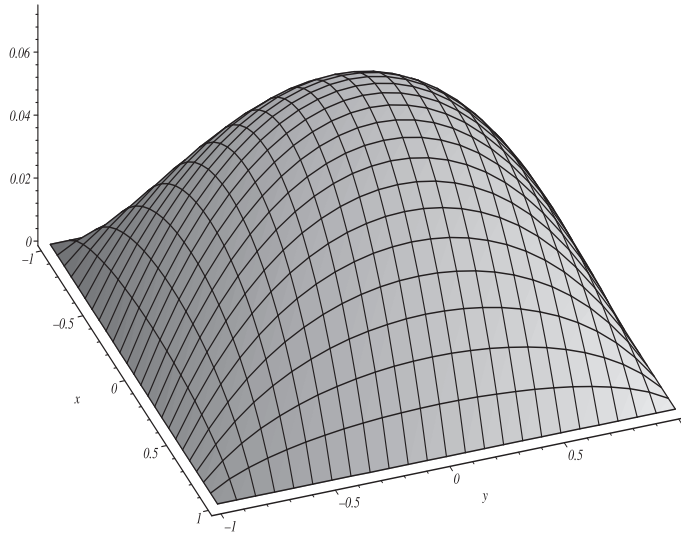
TAB. 4.1 – Approximation numérique de la fonction objectif J .

Par suite, on obtient les représentations graphiques suivantes

FIG. 4.3 – Etat optimal y_N .

¹Nombre de variables.

²Nombre de contraintes.

FIG. 4.4 – Contrôle optimal u_N .

Les représentations graphiques des approximations des fonctions contrôle et commande données par les Figures 4.3 et 4.4 sont régulières sont parfaitement lisses et ne représentent aucune irrégularités sur leurs surfaces respectives surtout sur le bord et les coins du domaine Ω .

4.3 Estimation d'Erreur a Posteriori

Après avoir calculé la solution approchée $(y_N(u_N), u_N)$ du problème (\mathcal{P}_N) , nous souhaitons maintenant savoir si cette solution est une bonne approximation de la solution du problème (\mathcal{P}) . Si ce n'est pas le cas, nous souhaiterions adapter le degré N du polynôme de *Legendre* afin de pouvoir calculer une solution approchée plus précise et avoir une précision de calcul meilleure. Pour ainsi faire, on doit estimer l'erreur uniquement en fonction des quantités connues, i.e. le degré du polynôme de Legendre N et la solution approchée $(y_N(u_N), u_N)$.

4.3.1 Estimation d'Erreur sur l'état et la commande

Nous nous intéressons maintenant à l'estimation d'erreurs a posteriori sur l'état et la commande simultanément, i.e. on donne une estimation de l'erreur de type $\|u - u_N\|_{H^1(\Omega)}^2 + \|y - y_N\|_{H^1(\Omega)}^2$, qu'on appellera *estimation couplée*, où dans la suite on suppose que $A = -\Delta$, $B = I$ et $\alpha = 1$.

Théorème 4.3.1. Soient $(y(u), p(u))$ et $(y_N(u_N), p_N(u_N))$ les solutions optimales des systèmes (1.4.21) et (4.2.6) respectivement tels que f et y_d sont dans $L^2(\Omega)$ alors, on obtient

$$\|y(u) - y_N(u_N)\|_{H^1(\Omega)}^2 + \|p(u) - p_N(u_N)\|_{H^1(\Omega)}^2 \leq CN^{-2} (\eta_y^2 + \eta_u^2), \quad (4.3.1)$$

où η_y^2 et η_u^2 , sont respectivement les estimateurs relatifs à la fonction d'état y et la fonction de commande u , telle que

$$\eta_y^2 = \|f - p_N(u_N) + \Delta y_N(u_N)\|_{L^2(\Omega)}^2, \quad (4.3.2)$$

et

$$\eta_u^2 = \|\Delta p_N(u_N) - y_d - y_N(u_N)\|_{L^2(\Omega)}^2, \quad (4.3.3)$$

où C est une constante indépendante de N .

Démonstration. D'après l'équation d'état (4.2.2a), on peut écrire

$$a(y(u), w_N) = (f - p(u), w_N)_{L^2(\Omega)}, \quad \forall w_N \in V_N,$$

et

$$a(y_N(u_N), w_N) = (f - p_N(u_N), w_N)_{L^2(\Omega)}, \quad \forall w_N \in V_N,$$

Une soustraction, donne

$$a(y(u) - y_N(u_N), w_N) = -(p(u) - p_N(u_N), w_N)_{L^2(\Omega)}, \quad \forall w_N \in V_N. \quad (4.3.4)$$

D'après l'équation d'état adjoint (4.2.2b), on peut écrire

$$a(p(u), q_N) = (y(u) - y_d, q_N)_{L^2(\Omega)}, \quad \forall q_N \in V_N,$$

et

$$a(p_N(u_N), q_N) = (y_N(u_N) - y_d, q_N)_{L^2(\Omega)}, \quad \forall q_N \in V_N.$$

Par soustraction, on obtient

$$a(p(u) - p_N(u_N), q_N) = (y(u) - y_N(u_N), q_N)_{L^2(\Omega)}, \quad \forall q_N \in V_N. \quad (4.3.5)$$

Puisque $a(\cdot, \cdot)$ est coercive et symétrique, alors on a

$$\underbrace{c \|y(u) - y_N(u_N)\|_{H^1(\Omega)}^2}_{\|e_y\|_{H^1(\Omega)}^2} \leq a(y(u) - y_N(u_N), y(u) - y_N(u_N)),$$

et

$$c \underbrace{\|p(u) - p_N(u_N)\|_{H^1(\Omega)}^2}_{\|\mathbf{e}_p\|_{H^1(\Omega)}^2} \leq a(p(u) - p_N(u_N), p(u) - p_N(u_N)).$$

Par addition, on obtient

$$\begin{aligned} c \underbrace{\|y(u) - y_N(u_N)\|_{H^1(\Omega)}^2}_{\|\mathbf{e}_y\|_{H^1(\Omega)}^2} + c \underbrace{\|p(u) - p_N(u_N)\|_{H^1(\Omega)}^2}_{\|\mathbf{e}_p\|_{H^1(\Omega)}^2} &\leq \\ &\leq a(y(u) - y_N(u_N), y(u) - y_N(u_N)) + a(p(u) - p_N(u_N), p(u) - p_N(u_N)). \end{aligned}$$

D'après l'équation d'état (4.2.2a), on peut écrire

$$a(y(u), \mathbf{e}_y - \Pi_N^{1,0} \mathbf{e}_y) + (p(u), \mathbf{e}_y - \Pi_N^{1,0} \mathbf{e}_y)_{L^2(\Omega)} = (f, \mathbf{e}_y - \Pi_N^{1,0} \mathbf{e}_y)_{L^2(\Omega)},$$

où il est clair que $\mathbf{e}_y - \Pi_N^{1,0} \mathbf{e}_y$ est dans $H_0^1(\Omega)$.

D'après l'équation d'état adjoint, on peut écrire

$$a(p(u), \mathbf{e}_p - \Pi_N^{1,0} \mathbf{e}_p) - (y(u), \mathbf{e}_p - \Pi_N^{1,0} \mathbf{e}_p)_{L^2(\Omega)} = - (y_d, \mathbf{e}_p - \Pi_N^{1,0} \mathbf{e}_p)_{L^2(\Omega)},$$

avec $\mathbf{e}_p - \Pi_N^{1,0} \mathbf{e}_p$ est dans $H_0^1(\Omega)$.

Dans ce qui précède, on a $\mathbf{e}_y = y(u) - y_N(u)$ et $\mathbf{e}_p = p(u) - p_N(u_N)$, d'où :

$$\begin{aligned} c \|\mathbf{e}_y\|_{H^1(\Omega)}^2 + c \|\mathbf{e}_p\|_{H^1(\Omega)}^2 &\leq a(y(u) - y_N(u_N), y(u) - y_N(u_N)) + \\ &\quad + a(p(u) - p_N(u_N), p(u) - p_N(u_N)). \end{aligned}$$

Dans l'inégalité précédente, on a ajouté et on a retranché la quantité $(p(u) - p_N(u_N), y(u) - y_N(u_N))_{L^2(\Omega)}$, Ainsi, on a

$$\begin{aligned} c \|\mathbf{e}_y\|_{H^1(\Omega)}^2 + c \|\mathbf{e}_p\|_{H^1(\Omega)}^2 &\leq a(y(u) - y_N(u_N), \mathbf{e}_y) + (p(u) - p_N(u_N), \mathbf{e}_y)_{L^2(\Omega)} + \\ &\quad + (p(u) - p_N(u_N), y(u) - y_N(u_N))_{L^2(\Omega)} - (p(u) - p_N(u_N), y(u) - y_N(u_N))_{L^2(\Omega)} + \\ &\quad + a(\mathbf{e}_p, p(u_N) - p_N(u_N)) - (y(u_N) - y_N(u_N), \mathbf{e}_p)_{L^2(\Omega)}. \end{aligned} \quad (4.3.6)$$

Dans l'équation (4.3.4), si on pose $w_N = \Pi_N^{1,0} \mathbf{e}_y$, on obtient

$$a(y(u) - y_N(u_N), \Pi_N^{1,0} \mathbf{e}_y) + (p(u) - p_N(u_N), \Pi_N^{1,0} \mathbf{e}_y)_{L^2(\Omega)} = 0. \quad (4.3.7)$$

D'après l'équation (4.3.5), on peu écrire

$$a(p(u) - p_N(u_N), \Pi_N^{1,0} \mathbf{e}_p) - (y(u) - y_N(u_N), \Pi_N^{1,0} \mathbf{e}_p)_{L^2(\Omega)} = 0. \quad (4.3.8)$$

En substituant, les deux équations (4.3.7) et (4.3.8), dans l'inégalité (4.3.6), on obtient

$$\begin{aligned} c \|\mathbf{e}_y\|_{H^1(\Omega)}^2 + c \|\mathbf{e}_p\|_{H^1(\Omega)}^2 &\leq a(y(u) - y_N(u_N), \mathbf{e}_y) + (p(u) - p_N(u_N), \mathbf{e}_y)_{L^2(\Omega)} + \\ &+ (p(u) - p_N(u_N), y(u) - y_N(u_N))_{L^2(\Omega)} - (p(u) - p_N(u_N), y(u) - y_N(u_N))_{L^2(\Omega)} + \\ &\quad + a(\mathbf{e}_p, p(u_N) - p_N(u_N)) - (y(u_N) - y_N(u_N), \mathbf{e}_p)_{L^2(\Omega)} - \\ &\quad - a(y(u) - y_N(u_N), \Pi_N^{1,0} \mathbf{e}_y) - (p(u) - p_N(u_N), \Pi_N^{1,0} \mathbf{e}_p)_{L^2(\Omega)} - \\ &\quad - a(p(u) - p_N(u_N), \Pi_N^{1,0} \mathbf{e}_p) + (y(u) - y_N(u_N), \Pi_N^{1,0} \mathbf{e}_p)_{L^2(\Omega)}. \end{aligned} \quad (4.3.9)$$

Dans ce cas, l'inégalité (4.3.9), peut s'écrire sous la forme

$$\begin{aligned} c \|\mathbf{e}_y\|_{H^1(\Omega)}^2 + c \|\mathbf{e}_p\|_{H^1(\Omega)}^2 &\leq a(y(u) - y_N(u_N), \mathbf{e}_y - \Pi_N^{1,0} \mathbf{e}_y) + \\ &\quad + (p(u) - p_N(u_N), \mathbf{e}_y - \Pi_N^{1,0} \mathbf{e}_p)_{L^2(\Omega)} + a(\mathbf{e}_p - \Pi_N^{1,0} \mathbf{e}_p, p(u) - p_N(u_N)) - \\ &\quad - (y(u) - y_N(u_N), \mathbf{e}_p - \Pi_N^{1,0} \mathbf{e}_p)_{L^2(\Omega)}. \end{aligned} \quad (4.3.10)$$

Par suite d'après l'équation d'état (1.4.21a), il vient

$$\begin{aligned} a(y_N(u_N), \mathbf{e}_y - \Pi_N^{1,0} \mathbf{e}_y) &= a(y(u), \mathbf{e}_y - \Pi_N^{1,0} \mathbf{e}_y) + (p(u) - p_N(u_N), \mathbf{e}_y - \Pi_N^{1,0} \mathbf{e}_p)_{L^2(\Omega)} \\ &= (f - p_N(u_N), \mathbf{e}_y - \Pi_N \mathbf{e}_y)_{L^2(\Omega)}, \end{aligned}$$

et d'après l'équation d'état adjoint (1.4.21b) on a

$$\begin{aligned} a(p(u), \mathbf{e}_p - \Pi_N^{1,0} \mathbf{e}_p) &= a(p_N(u_N), \mathbf{e}_p - \Pi_N^{1,0} \mathbf{e}_p) + (y(u) - y_N(u_N), \mathbf{e}_p - \Pi_N^{1,0} \mathbf{e}_p)_{L^2(\Omega)} \\ &= -(y_d + y_N(u_N), \mathbf{e}_p - \Pi_N^{1,0} \mathbf{e}_p)_{L^2(\Omega)}. \end{aligned}$$

D'après les relations (4.3.10), on écrit

$$\begin{aligned} c \|\mathbf{e}_y\|_{H^1(\Omega)}^2 + c \|\mathbf{e}_p\|_{H^1(\Omega)}^2 &\leq \int_{\Omega} (f - p_N(u_N) + \Delta y_N(u_N)) (\mathbf{e}_y - \Pi_N^{1,0} \mathbf{e}_y) \, d\mathbf{x} + \\ &\quad + \int_{\Omega} (\Delta p_N(u_N) - y_d - y_N(u_N)) (\mathbf{e}_p - \Pi_N^{1,0} \mathbf{e}_p) \, d\mathbf{x}, \end{aligned}$$

et grâce à l'inégalité de Cauchy-Scharwz, on obtient

$$\begin{aligned} \|\mathbf{e}_y\|_{H^1(\Omega)}^2 + \|\mathbf{e}_p\|_{H^1(\Omega)}^2 &\leq \frac{1}{c} \|f - p_N(u_N) + \Delta y_N(u_N)\|_{L^2(\Omega)} \|\mathbf{e}_y - \Pi_N^{1,0} \mathbf{e}_y\|_{L^2(\Omega)} + \\ &\quad + \frac{1}{c} \|\Delta p_N(u_N) - y_d - y_N(u_N)\|_{L^2(\Omega)} \|\mathbf{e}_p - \Pi_N^{1,0} \mathbf{e}_p\|_{L^2(\Omega)}. \end{aligned}$$

En appliquant le Théorème 2.5.1 on a

$$\begin{aligned} \|\mathbf{e}_y\|_{H^1(\Omega)}^2 + \|\mathbf{e}_p\|_{H^1(\Omega)}^2 &\leq \frac{CN^{-1}}{c} \|f - p_N(u_N) + \Delta y_N(u_N)\|_{L^2(\Omega)} \|y(u) - y_N(u_N)\|_{L^2(\Omega)} + \\ &\quad + \frac{CN^{-1}}{c} \|\Delta p_N(u_N) - y_d - y_N(u_N)\|_{L^2(\Omega)} \|p(u) - p_N(u_N)\|_{L^2(\Omega)}. \end{aligned}$$

L'inégalité de Hölder, donne

$$\begin{aligned} \|\mathbf{e}_y\|_{H^1(\Omega)}^2 + \|\mathbf{e}_p\|_{H^1(\Omega)}^2 &\leq \frac{CN^{-2}}{2} \|f - p_N(u_N) + \Delta y_N(u_N)\|_{L^2(\Omega)}^2 + \frac{1}{2} \|y(u) - y_N(u_N)\|_{L^2(\Omega)}^2 \\ &\quad + \frac{CN^{-2}}{2} \|\Delta p_N(u_N) - y_d - y_N(u_N)\|_{L^2(\Omega)}^2 + \frac{1}{2} \|p(u) - p_N(u_N)\|_{L^2(\Omega)}^2. \end{aligned}$$

Finalement, on déduit

$$\begin{aligned} \|y(u) - y_N(u_N)\|_{H^1(\Omega)}^2 + \|p(u) - p_N(u_N)\|_{H^1(\Omega)}^2 &\leq \\ &\leq CN^{-2} \left(\|f - p_N(u_N) + \Delta y_N(u_N)\|_{L^2(\Omega)}^2 + \|\Delta p_N(u_N) - y_d - y_N(u_N)\|_{L^2(\Omega)}^2 \right), \end{aligned}$$

où, C est une constante qui ne dépend pas de N .

□

4.3.2 Estimation d'Erreur sur l'état, le contrôle, la fonction objectif

Ci-dessous, on donne des majorations sur les deux erreurs $\|u - u_N\|_{L^2(\Omega)}$ et $\|y - y_N\|_{H^1(\Omega)}$, séparément, qu'on appellera *estimations non-couplées*, où dans la suite on suppose que $A = -\Delta$ et $\alpha = 1$.

Lemme 4.3.1. *Soient $(u, p(u))$ et $(u_N, p_N(u_N))$, deux solutions optimales qui vérifient respectivement les deux systèmes d'équations (1.4.21) et (4.2.2), alors on obtient*

$$\|p(u) - p_N(u_N)\|_{H^1(\Omega)} \leq C \|u - u_N\|_{H^1(\Omega)} + \|p(u_N) - p_N(u_N)\|_{H^1(\Omega)}, \quad (4.3.11)$$

où C est une constante indépendante de N .

Démonstration. Par soustraction des deux équations d'état adjoint (1.4.20b) et (4.2.5b), on obtient

$$a(p(u) - p(u_N), v) = (y(u) - y(u_N), v)_{L^2(\Omega)}, \quad \forall v \in V.$$

Dans le cas, où on choisit de prendre $v = p(u) - p(u_N)$, on obtient

$$a(p(u) - p(u_N), p(u) - p(u_N)) = (y(u) - y(u_N), p(u) - p(u_N))_{L^2(\Omega)}.$$

Ci-dessous, on pose $\mathbf{e}_{\hat{p}} = p(u) - p(u_N)$ et en tenant compte que la forme bilinéaire $a(\cdot, \cdot)$ est coercive, on peut écrire

$$\begin{aligned} c \|\mathbf{e}_{\hat{p}}\|_{H^1(\Omega)}^2 &\leq a(\mathbf{e}_{\hat{p}}, \mathbf{e}_{\hat{p}}) \\ &\leq (y(u) - y(u_N), \mathbf{e}_{\hat{p}})_{L^2(\Omega)}. \end{aligned}$$

A l'aide de l'inégalité de Cauchy-Schwarz, on obtient

$$c \|\mathbf{e}_{\hat{p}}\|_{H^1(\Omega)}^2 \leq \|y(u) - y(u_N)\|_{L^2(\Omega)} \|\mathbf{e}_{\hat{p}}\|_{L^2(\Omega)},$$

et par injection continue de $H^1(\Omega)$ dans $L^2(\Omega)$, on a

$$c \|\mathbf{e}_{\hat{p}}\|_{H^1(\Omega)}^2 \leq \|y(u) - y(u_N)\|_{L^2(\Omega)} \|\mathbf{e}_{\hat{p}}\|_{H^1(\Omega)}.$$

Finalement, on obtient

$$\|p(u) - p(u_N)\|_{H^1(\Omega)} \leq C \|y(u) - y(u_N)\|_{L^2(\Omega)}.$$

Par injection continue de $H^1(\Omega)$ dans $L^2(\Omega)$, on obtient aussi

$$\|p(u) - p(u_N)\|_{H^1(\Omega)} \leq C \|y(u) - y(u_N)\|_{H^1(\Omega)}, \quad (4.3.12)$$

où C est une constante qui ne dépend pas de N .

Par soustraction des deux équations d'états (1.4.20a) et (4.2.5a), on obtient

$$a(y(u) - y(u_N), v) + (B(u - u_N), v)_{L^2(\Omega)} = 0, \quad \forall v \in V.$$

Dans ce cas, on choisit $v = y(u) - y(u_N)$, et l'on a :

$$a(y(u) - y(u_N), y(u) - y(u_N)) = -(B(u - u_N), y(u) - y(u_N))_{L^2(\Omega)},$$

Ci-après, on pose $\mathbf{e}_{\hat{y}} = y(u) - y(u_N)$ en tenant compte que la forme bilinéaire $a(\cdot, \cdot)$ est coercive, on a

$$\begin{aligned} c \|\mathbf{e}_{\hat{y}}\|_{H^1(\Omega)}^2 &\leq a(\mathbf{e}_{\hat{y}}, \mathbf{e}_{\hat{y}}) \\ &\leq -(B(u - u_N), \mathbf{e}_{\hat{y}})_{L^2(\Omega)}. \end{aligned}$$

Grâce à l'inégalité de Cauchy-Schwarz, on obtient

$$c \|\mathbf{e}_{\hat{y}}\|_{H^1(\Omega)}^2 \leq \|B\|_{L^2(\Omega)} \|u - u_N\|_{L^2(\Omega)} \|\mathbf{e}_{\hat{y}}\|_{L^2(\Omega)}.$$

De même l'injection continue de $H^1(\Omega)$ dans $L^2(\Omega)$, donne

$$c \|\mathbf{e}_{\hat{y}}\|_{H^1(\Omega)}^2 \leq \|B\|_{L^2(\Omega)} \|u - u_N\|_{L^2(\Omega)} \|\mathbf{e}_{\hat{y}}\|_{H^1(\Omega)}.$$

D'où :

$$\|y(u) - y(u_N)\|_{H^1(\Omega)} \leq C \|u - u_N\|_{L^2(\Omega)}.$$

L'injection continue de $H^1(\Omega)$ dans $L^2(\Omega)$, donne aussi

$$\|y(u) - y(u_N)\|_{H^1(\Omega)} \leq C \|u - u_N\|_{H^1(\Omega)}. \quad (4.3.13)$$

D'autre par, on a

$$\|p(u) - p_N(u_N)\|_{H^1(\Omega)} \leq \|p(u) - p(u_N)\|_{H^1(\Omega)} + \|p(u_N) - p_N(u_N)\|_{H^1(\Omega)}.$$

Les inégalités (4.3.12) et (4.3.13) donnent respectivement

$$\|p(u) - p_N(u_N)\|_{H^1(\Omega)} \leq C \|y(u) - y(u_N)\|_{H^1(\Omega)} + \|p(u_N) - p_N(u_N)\|_{H^1(\Omega)},$$

et

$$\|p(u) - p_N(u_N)\|_{H^1(\Omega)} \leq C \|u - u_N\|_{H^1(\Omega)} + \|p(u_N) - p_N(u_N)\|_{H^1(\Omega)},$$

□

Lemme 4.3.2. *Soient (u, p) et (u_N, p_N) deux solutions optimales qui vérifient respectivement les deux systèmes d'équations (1.4.19) et (4.2.1), alors on obtient*

$$\|u - u_N\|_{L^2(\Omega)} \leq C \|p(u_N) - p_N(u_N)\|_{H^1(\Omega)}, \quad (4.3.14)$$

où C est une constante indépendante de N .

Démonstration. On peut écrire $\|u - u_N\|_{L^2(\Omega)}^2$, sous la forme

$$\|u - u_N\|_{L^2(\Omega)}^2 = (u, u - u_N)_{L^2(\Omega)} - (u_N, u - u_N)_{L^2(\Omega)}.$$

Ci-dessous, on pose $\mathbf{e}_u = u - u_N$ et dans l'égalité de droite, on ajoute et on retranche les deux quantités $(B^*p_N(u_N), u - u_N)_{L^2(\Omega)}$ et $(Bp(u_N), u - u_N)_{L^2(\Omega)}$. Dans ce cas, on a

$$\begin{aligned} \|\mathbf{e}_u\|_{L^2(\Omega)}^2 &= (u, u - u_N)_{L^2(\Omega)} - (u_N, u - u_N)_{L^2(\Omega)} + \\ &\quad + (B^* p_N(u_N), u - u_N)_{L^2(\Omega)} - (B^* p_N(u_N), u - u_N)_{L^2(\Omega)} + \\ &\quad + (Bp(u_N), u - u_N)_{L^2(\Omega)} - (Bp(u_N), u - u_N)_{L^2(\Omega)}. \end{aligned}$$

D'après l'équation de projection (1.4.20c), on déduit

$$\begin{aligned} \|\mathbf{e}_u\|_{L^2(\Omega)}^2 &= - (B^* p(u), u - u_N)_{L^2(\Omega)} - (u_N, u - u_N)_{L^2(\Omega)} + \\ &\quad + (B^* p_N(u_N), u - u_N)_{L^2(\Omega)} - (B^* p_N(u_N), u - u_N)_{L^2(\Omega)} + \\ &\quad + (Bp(u_N), u - u_N)_{L^2(\Omega)} - (Bp(u_N), u - u_N)_{L^2(\Omega)}. \end{aligned}$$

De même l'équation de projection (4.2.5c), donne

$$\|\mathbf{e}_u\|_{L^2(\Omega)}^2 = (B^* (p_N(u_N) - p(u_N)), u - u_N)_{L^2(\Omega)} + (B^* (p(u_N) - p(u)), u - u_N)_{L^2(\Omega)}.$$

Enfin, l'équation de projection (4.2.1c), implique

$$\begin{aligned} \|\mathbf{e}_u\|_{L^2(\Omega)}^2 &= (B^* (p_N(u_N) - p(u_N)), u - u_N)_{L^2(\Omega)} - (u - u_N, u - u_N)_{L^2(\Omega)} \\ &= (B^* (p_N(u_N) - p(u_N)), u - u_N) - \|u - u_N\|_{L^2(\Omega)}^2. \end{aligned}$$

Comme le second membre de l'inégalité de droite est négatif, on déduit facilement que

$$\|\mathbf{e}_u\|_{L^2(\Omega)}^2 = (B^* (p_N(u_N) - p(u_N)), u - u_N).$$

Grâce à l'inégalité de Cauchy-Schwarz, on déduit

$$\|\mathbf{e}_u\|_{L^2(\Omega)}^2 \leq \|B^*\|_{L^2(\Omega)} \|p_N(u_N) - p(u_N)\|_{L^2(\Omega)} \underbrace{\|u - u_N\|_{L^2(\Omega)}}_{\|\mathbf{e}_u\|_{L^2(\Omega)}}.$$

Finalement, on obtient

$$\|u - u_N\|_{L^2(\Omega)} \leq C \|p_N(u_N) - p(u_N)\|_{L^2(\Omega)},$$

et par injection continue de $H^1(\Omega)$ dans $L^2(\Omega)$, on déduit

$$\|u - u_N\|_{L^2(\Omega)} \leq C \|p_N(u_N) - p(u_N)\|_{H^1(\Omega)}.$$

□

Lemme 4.3.3. Soient (y_N, p_N) et $(y(u_N), p(u_N))$ un couple de solutions qui vérifient respectivement les deux systèmes (4.2.1) et (4.2.4), alors on obtient

$$\|p(u) - p_N(u_N)\|_{H^1(\Omega)} \leq C \|p(u_N) - p_N(u_N)\|_{H^1(\Omega)}, \quad (4.3.15)$$

où C est une constante qui ne dépend pas de N .

Démonstration. L'estimation (4.3.15) se déduit facilement par applications directes du Lemme 4.3.1 et du Lemme 4.3.2. □

Lemme 4.3.4. Soient $(y(u), p(u))$ et $(y_N(u), p_N(u))$ et $p_N(u_N)$ des solutions qui vérifient respectivement les trois systèmes (1.4.19), (4.2.1) et (4.2.4), alors on obtient

$$\|p(u_N) - p_N(u_N)\|_{H^1(\Omega)} \leq C \|y(u_N) - y_N(u_N)\|_{H^1(\Omega)}, \quad (4.3.16)$$

où C est une constante qui ne dépend pas de N .

Démonstration. Puisque la forme bilinéaire $a(\cdot, \cdot)$ est coercive, dans ce cas, on a

$$\begin{aligned} c \|p(u_N) - p_N(u_N)\|_{H^1(\Omega)}^2 &\leq a(p(u_N) - p_N(u_N), p(u_N) - p_N(u_N)) \\ &\leq a(p(u_N), p(u_N) - p_N(u_N)) - a(p_N(u_N), p(u_N) - p_N(u_N)). \end{aligned}$$

Soit $\mathbf{e}_{\tilde{p}} = p(u_N) - p_N(u_N)$, et d'après les deux équations de l'état (4.2.2a) et (4.2.5a), on obtient

$$\begin{aligned} c \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}^2 &\leq (y(u_N) - y_d, \mathbf{e}_{\tilde{p}})_{L^2(\Omega)} - (y_N(u_N) - y_d, \mathbf{e}_{\tilde{p}})_{L^2(\Omega)} \\ &\leq (y(u_N) - y_N(u_N), \mathbf{e}_{\tilde{p}})_{L^2(\Omega)}. \end{aligned}$$

L'inégalité de Cauchy-Schwarz, donne

$$c \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}^2 \leq \|y(u_N) - y_N(u_N)\|_{L^2(\Omega)} \|\mathbf{e}_{\tilde{p}}\|_{L^2(\Omega)}.$$

Par injection continue de $H^1(\Omega)$ dans $L^2(\Omega)$, on obtient

$$c \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}^2 \leq \|y(u_N) - y_N(u_N)\|_{L^2(\Omega)} \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}.$$

Finalement, on déduit

$$\|p(u_N) - p_N(u_N)\|_{H^1(\Omega)} \leq C \|y(u_N) - y_N(u_N)\|_{L^2(\Omega)}.$$

Par injection continue de $H^1(\Omega)$ dans $L^2(\Omega)$, on obtient aussi

$$\|p(u_N) - p_N(u_N)\|_{H^1(\Omega)} \leq C \|y(u_N) - y_N(u_N)\|_{H^1(\Omega)}.$$

□

Lemme 4.3.5. Soient $y(u)$, $y(u_N)$ et $y_N(u_N)$, des solutions vérifiant respectivement les systèmes d'équations (1.4.19), (4.2.1) et (4.2.4), alors on obtient

$$\|y(u) - y_N(u_N)\|_{H^1(\Omega)} \leq C \|y(u_N) - y_N(u_N)\|_{H^1(\Omega)}, \quad (4.3.17)$$

où, C est une constante indépendante de N .

Démonstration. Dans l'équation d'état (1.4.20a), on choisit de prendre $w = y(u) - y_N(u_N)$. Dans ce cas, on obtient

$$a(y(u), y(u) - y_N(u_N)) + (Bu, y(u) - y_N(u))_{L^2(\Omega)} = (f, y(u) - y_N(u_N))_{L^2(\Omega)}.$$

Si dans l'équation d'état discrète (1.4.20a), on pose $w_N = y(u) - y_N(u_N)$, on obtient

$$a(y_N(u_N), y(u) - y_N(u_N)) + (Bu_N, y(u) - y_N(u_N))_{L^2(\Omega)} = (f, y(u) - y_N(u_N))_{L^2(\Omega)}.$$

Par soustraction des deux dernières équations, on trouve

$$a(y(u) - y_N(u_N), y(u) - y_N(u_N)) = -(B(u - u_N), y(u) - y_N(u_N))_{L^2(\Omega)}.$$

Soit $\mathbf{e}_y = y(u) - y_N(u_N)$ et en tenant compte que la forme bilinéaire $a(\cdot, \cdot)$ est coercive, alors on écrit

$$\begin{aligned} c \|\mathbf{e}_y\|_{H^1(\Omega)}^2 &\leq a(\mathbf{e}_y, \mathbf{e}_y) \\ &= -(B(u - u_N), \mathbf{e}_y)_{L^2(\Omega)}. \end{aligned}$$

L'inégalité de Cauchy-Schwarz, implique

$$c \|y(u) - y_N(u_N)\|_{H^1(\Omega)}^2 \leq \|B\|_{L^2(\Omega)} \|u - u_N\|_{L^2(\Omega)} \|\mathbf{e}_y\|_{H^1(\Omega)}.$$

Enfin, on obtient

$$\|y(u) - y_N(u_N)\|_{H^1(\Omega)} \leq C \|u - u_N\|_{L^2(\Omega)},$$

Le Lemme 4.3.2, implique

$$\|y(u) - y_N(u_N)\|_{H^1(\Omega)} \leq C \|p_N(u_N) - p(u_N)\|_{H^1(\Omega)}.$$

Enfin, on conclut à l'aide de l'utilisaton du Lemme 4.3.3

$$\|y(u) - y_N(u_N)\|_{H^1(\Omega)} \leq C \|y(u_N) - y_N\|_{H^1(\Omega)}.$$

□

Lemme 4.3.6. Soient $y(u_N)$ et $(y_N(u_N), u_N)$ deux solutions vérifiant respectivement des deux systèmes (4.2.4) et (4.2.1), alors on obtient l'estimation suivante

$$\|y(u_N) - y_N(u_N)\|_{H^1(\Omega)} \leq CN^{-1} \|f + Bu_N + \Delta y_N(u_N)\|_{L^2(\Omega)}. \quad (4.3.18)$$

Démonstration. Puisque que la forme bilinéaire $a(\cdot, \cdot)$ est symétrique, continue et coercive, alors on peut écrire que

$$c \|y(u_N) - y_N(u_N)\|_{H^1(\Omega)}^2 \leq a(y(u_N) - y_N(u_N), y(u_N) - y_N(u_N)).$$

Soit $\mathbf{e}_{\tilde{y}} = y(u_N) - y_N(u_N)$. D'après l'équation de projection (3.2.3) relative à la forme bilinéaire $a(\cdot, \cdot)$, on obtient

$$c \|\mathbf{e}_{\tilde{y}}\|_{H^1(\Omega)}^2 \leq a(y(u_N) - y_N(u_N), \mathbf{e}_{\tilde{y}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{y}}),$$

ou encore

$$c \|\mathbf{e}_{\tilde{y}}\|_{H^1(\Omega)}^2 \leq \int_{\Omega} -\Delta(y(u_N) - y_N(u_N)) (\mathbf{e}_{\tilde{y}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{y}}) \, d\mathbf{x}.$$

Dans ce cas, on a

$$c \|\mathbf{e}_{\tilde{y}}\|_{H^1(\Omega)}^2 \leq \int_{\Omega} (-\Delta y(u_N) (\mathbf{e}_{\tilde{y}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{y}}) + \Delta y_N(u_N) (\mathbf{e}_{\tilde{y}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{y}})) \, d\mathbf{x}.$$

D'après l'équation d'état (4.2.2a), on obtient

$$c \|\mathbf{e}_{\tilde{y}}\|_{H^1(\Omega)}^2 \leq \int_{\Omega} (f + Bu_N) (\mathbf{e}_{\tilde{y}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{y}}) \, d\mathbf{x} + \int_{\Omega} \Delta y_N(u_N) (\mathbf{e}_{\tilde{y}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{y}}) \, d\mathbf{x}.$$

D'où :

$$c \|\mathbf{e}_{\tilde{y}}\|_{H^1(\Omega)}^2 \leq \int_{\Omega} (f + Bu_N + \Delta y_N(u_N)) (\mathbf{e}_{\tilde{y}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{y}}) \, d\mathbf{x},$$

par application de l'inégalité de Cauchy-Schwarz, on déduit que

$$c \|\mathbf{e}_{\tilde{y}}\|_{H^1(\Omega)}^2 \leq \|f + Bu_N + \Delta y_N(u_N)\|_{L^2(\Omega)} \|\mathbf{e}_{\tilde{y}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{y}}\|_{L^2(\Omega)}.$$

Par injection continue de $H^1(\Omega)$ dans $L^2(\Omega)$, on obtient

$$c \|\mathbf{e}_{\tilde{y}}\|_{H^1(\Omega)}^2 \leq \|f + Bu_N + \Delta y_N(u_N)\|_{L^2(\Omega)} \|\mathbf{e}_{\tilde{y}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{y}}\|_{H^1(\Omega)}.$$

Par application du Théorème 2.5.1, on a

$$c \|\mathbf{e}_{\tilde{y}}\|_{H^1(\Omega)}^2 \leq CN^{-1} \|f + Bu_N + \Delta y_N(u_N)\|_{L^2(\Omega)} \|\mathbf{e}_{\tilde{y}}\|_{H^1(\Omega)}.$$

Finalement, on obtient

$$\|y(u_N) - y_N(u_N)\|_{H^1(\Omega)} \leq CN^{-1} \|f + Bu_N + \Delta y_N(u_N)\|_{L^2(\Omega)},$$

où C est une constante qui ne dépend pas de N . \square

Lemme 4.3.7. Soient, $p(u_N)$ et $(y_N(u_N), p_N(u_N))$ des solutions vérifiant, respectivement les deux systèmes (4.2.4) et (4.2.1), alors on obtient l'estimation suivante

$$\|p_N(u_N) - p(u_N)\|_{H^1(\Omega)} \leq CN^{-1} \|-\Delta p_N(u_N) - y_N(u_N) + y_d\|_{L^2(\Omega)}. \quad (4.3.19)$$

où, C est une constante indépendante de N .

Démonstration. D'après la coercivité de la forme bilinéaire $a(\cdot, \cdot)$, on a

$$c \|p(u_N) - p_N(u_N)\|_{H^1(\Omega)}^2 \leq a(p(u_N) - p_N(u_N), p(u_N) - p_N(u_N)).$$

Soit $\mathbf{e}_{\tilde{p}} = p(u_N) - p_N(u_N)$. Dans l'inégalité de droite on ajoute et on retranche la quantité $a(\Pi_N^{1,0} \mathbf{e}_{\tilde{p}}, \mathbf{e}_{\tilde{p}})$, pour obtenir

$$c \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}^2 \leq a(\mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}, \mathbf{e}_{\tilde{p}}) + a(\Pi_N^{1,0} \mathbf{e}_{\tilde{p}}, \mathbf{e}_{\tilde{p}}).$$

Par soustraction des deux équations d'état adjoint (4.2.2b) et (4.2.5b), on obtient

$$a(p(u_N) - p_N(u_N), \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}) = (y_N(u_N) - y(u_N), \Pi_N^{1,0} \mathbf{e}_{\tilde{p}})_{L^2(\Omega)}.$$

En substituant dans l'inégalité (4.3.2), on a

$$c \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}^2 \leq \int_{\Omega} -\Delta(p_N(u_N) - p(u_N)) (\mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}) d\mathbf{x} + (y_N(u_N) - y(u_N), \Pi_N^{1,0} \mathbf{e}_{\tilde{p}})_{L^2(\Omega)},$$

ou encore

$$c \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}^2 \leq \int_{\Omega} -\Delta p_N(u_N) (\mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}) d\mathbf{x} - \int_{\Omega} \Delta p(u_N) (\mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}) d\mathbf{x} + (y_N(u_N) - y(u_N), \Pi_N^{1,0} \mathbf{e}_{\tilde{p}})_{L^2(\Omega)}.$$

D'après l'équation d'état adjoint (4.2.2b), on déduit

$$c \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}^2 \leq \int_{\Omega} -\Delta p_N(u_N) (\mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}) d\mathbf{x} - \int_{\Omega} (y(u_N) - y_d) (\mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}) d\mathbf{x} + (y_N(u_N) - y(u_N), \Pi_N^{1,0} \mathbf{e}_{\tilde{p}})_{L^2(\Omega)}.$$

Dans le second membre de l'inégalité de droite, on ajoute et on retranche la quantité $y_N(u_N)$, pour obtenir

$$\begin{aligned} c \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}^2 &\leq \int_{\Omega} -\Delta p_N(u_N) (\mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}) \, d\mathbf{x} - \\ &\quad - \int_{\Omega} ((y(u_N) - y_N(u_N) + y_N(u_N) - y_d) (\mathbf{e}_{\tilde{p}} - \Pi_N \mathbf{e}_{\tilde{p}})) \, d\mathbf{x} - \\ &\quad - (y(u_N) - y_N(u_N), \Pi_N^{1,0} \mathbf{e}_{\tilde{p}})_{L^2(\Omega)}. \end{aligned}$$

Dans ce cas, on obtient

$$\begin{aligned} c \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}^2 &\leq \int_{\Omega} -\Delta p_N(u_N) (\mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}) \, d\mathbf{x} - \\ &\quad - (y(u_N) - y_N(u_N), \mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}})_{L^2(\Omega)} - (y(u_N) - y_N(u_N), \Pi_N^{1,0} \mathbf{e}_{\tilde{p}})_{L^2(\Omega)} - \\ &\quad - \int_{\Omega} (y_N(u_N) - y_d) (\mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}) \, d\mathbf{x}. \end{aligned}$$

Par soustraction des deux équations de l'état adjoint (4.2.2b) et (4.2.5b), on obtient

$$a(\mathbf{e}_{\tilde{p}}, q) = (y(u_N) - y_N(u_N), q)_{L^2(\Omega)}, \quad \forall q \in H_0^1(\Omega).$$

Dans ce cas, si on choisit de prendre $q = \mathbf{e}_{\tilde{p}}$, il vient

$$a(\mathbf{e}_{\tilde{p}}, \mathbf{e}_{\tilde{p}}) = (y(u_N) - y_N(u_N), \mathbf{e}_{\tilde{p}})_{L^2(\Omega)}.$$

Il est alors facile de remarquer que

$$(y(u_N) - y_N(u_N), \mathbf{e}_{\tilde{p}})_{L^2(\Omega)} > 0. \quad (4.3.20)$$

Par suite, on a

$$\begin{aligned} c \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}^2 &\leq \int_{\Omega} -\Delta p_N(u_N) (\mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}) \, d\mathbf{x} - \int_{\Omega} (y_N(u_N) - y_d) (\mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}) \, d\mathbf{x} - \\ &\quad - (y(u_N) - y_N(u_N), \mathbf{e}_{\tilde{p}})_{L^2(\Omega)} \\ &\leq \int_{\Omega} (-\Delta p_N(u_N) - y_N(u_N) + y_d) (\mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}) \, d\mathbf{x} - (y(u_N) - y_N(u_N), \mathbf{e}_{\tilde{p}})_{L^2(\Omega)}. \end{aligned}$$

D'après l'inégalité (4.3.20), on obtient

$$c \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}^2 \leq \int_{\Omega} (-\Delta p_N(u_N) - y_N(u_N) + y_d) (\mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}) \, d\mathbf{x}.$$

A l'aide de l'inégalité de Cauchy-Schwarz, on a

$$c \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}^2 \leq \|-\Delta p_N(u_N) - y_N(u_N) + y_d\|_{L^2(\Omega)} \|\mathbf{e}_{\tilde{p}} - \Pi_N^{1,0} \mathbf{e}_{\tilde{p}}\|_{L^2(\Omega)},$$

et par application du Théorème 2.5.1, il vient

$$c \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}^2 \leq CN^{-1} \|-\Delta p_N(u_N) - y_N(u_N) + y_d\|_{L^2(\Omega)} \|\mathbf{e}_{\tilde{p}}\|_{H^1(\Omega)}.$$

Finalement, on déduit

$$\|p_N(u_N) - p(u_N)\|_{H^1(\Omega)} \leq CN^{-1} \|-\Delta p_N(u_N) - y_N(u_N) + y_d\|_{L^2(\Omega)}.$$

□

Théorème 4.3.2. *Soient $y(u)$, $y_N(u_N)$ et u_N , des solutions vérifiant respectivement les systèmes d'équations (4.2.1) et (4.2.4), alors on obtient l'estimation suivante*

$$\|y(u) - y_N(u_N)\|_{H^1(\Omega)} \leq CN^{-1} \|f + Bu_N + \Delta y_N(u_N)\|_{L^2(\Omega)}, \quad (4.3.21)$$

où, η_y est l'estimateur relatif à la fonction d'état y , tel que

$$\eta_y = CN^{-1} \|f + Bu_N + \Delta y_N(u_N)\|_{L^2(\Omega)}, \quad (4.3.22)$$

où C est une constante indépendante de N .

Démonstration. L'estimation (4.3.21), se déduit facilement par applications directe des Lemmes 4.3.5 et 4.3.6. □

Théorème 4.3.3. *Soient u , u_N , $p(u_N)$ et $p_N(u_N)$ des solutions qui vérifient respectivement les systèmes d'équations (1.4.19), (4.2.1) et (4.2.4), alors on obtient l'estimation suivante*

$$\|u - u_N\|_{L^2(\Omega)} \leq \eta_u, \quad (4.3.23)$$

où, η_u est l'estimateur relatif à la fonction de contrôle u , tel que

$$\eta_u = CN^{-1} \|-\Delta p_N(u_N) - y_N(u_N) + y_d\|_{L^2(\Omega)}, \quad (4.3.24)$$

où C est une constante indépendante de N .

Démonstration. L'estimation (4.3.23), se déduit facilement par applications directes des Lemmes 4.3.7 et 4.3.2. □

Théorème 4.3.4. *Soient $(y(u), u)$ et $(y_N(u_N), u_N)$, deux solutions optimales qui vérifient respectivement les deux systèmes d'équations (1.4.19) et (4.2.1), alors on obtient l'estimation suivante*

$$|J(y(u), u) - J(y_N(u_N), u_N)| \leq CN^{-1} (\eta_y + \eta_u), \quad (4.3.25)$$

où η_y et η_u , sont donnés respectivement par le Théorème 4.3.2 et le Théorème 4.3.3, où C est une constante indépendante de N .

Démonstration. Par soustraction des fonctions $J(y(u), u)$ et $J(y_N(u_N), u_N)$, on obtient

$$\begin{aligned} \mathbf{e}_J &= \left| \frac{1}{2} \left(\|y(u) - y_d\|_{L^2(\Omega)}^2 - \|y_N(u_N) - y_d\|^2 \right) + \frac{1}{2} \left(\|u\|_{L^2(\Omega)}^2 - \|u_N\|_{L^2(\Omega)}^2 \right) \right| \\ &= \left| \frac{1}{2} (y(u) - y_d, y(u) + y_N(u_N) - 2y_d)_{L^2(\Omega)} + \frac{1}{2} (u - u_N, u + u_N)_{L^2(\Omega)} \right|, \end{aligned} \quad (4.3.26)$$

où $\mathbf{e}_J = |J(y(u), u) - J_N(y_N(u_N), u_N)|$.

Dans le premier membre de l'égalité de droite, on ajoute et on retranche $y(u)$ et dans le second on ajoute et on retranche u , pour obtenir :

$$\begin{aligned} \mathbf{e}_J &= \left| \frac{1}{2} (y(u) - y_N(u_N), y_N(u_N) - y(u) + 2(y(u) - y_d))_{L^2(\Omega)} + \right. \\ &\quad \left. + \frac{1}{2} (u - u_N, u + u_N - u + u)_{L^2(\Omega)} \right|, \end{aligned}$$

grâce à l'inégalité de Cauchy-Schwarz, on déduit que

$$\begin{aligned} \mathbf{e}_J &\leq \frac{1}{2} \|y(u) - y_N(u_N)\|_{L^2(\Omega)} \left(\|y_N(u_N) - y(u)\|_{L^2(\Omega)} + 2 \|y(u) - y_d\|_{L^2(\Omega)} \right) + \\ &\quad + \frac{1}{2} \|u - u_N\|_{L^2(\Omega)} \left(\|u - u_N\|_{L^2(\Omega)} + 2 \|u\|_{L^2(\Omega)} \right). \end{aligned}$$

En supposant que les deux solutions optimale $y(u)$ et u , sont bornées et en supposant l'existence de deux constante c_1 et c_2 , indépendantes de N , réalisant $\|y(u)\|_{L^2(\Omega)} \leq c_1$ et $\|u\|_{L^2(\Omega)} \leq c_2$, on écrit alors :

$$\begin{aligned} \mathbf{e}_J &\leq \left(\frac{1}{2} \|y_N(u_N) - y(u)\|_{L^2(\Omega)}^2 + c_1 \|y(u) - y_N(u_N)\|_{L^2(\Omega)} \right) + \\ &\quad + \left(\frac{1}{2} \|u - u_N\|_{L^2(\Omega)}^2 + c_2 \|u - u_N\|_{L^2(\Omega)} \right). \end{aligned}$$

Dans ce cas, il vient

$$\begin{aligned} \mathbf{e}_J &\leq \max \left\{ \frac{1}{2}, c_1 \right\} \left(\|y_N(u_N) - y(u)\|_{L^2(\Omega)}^2 + \|y(u) - y_N(u_N)\|_{L^2(\Omega)} \right) + \\ &\quad + \max \left\{ \frac{1}{2}, c_2 \right\} \left(\frac{1}{2} \|u - u_N\|_{L^2(\Omega)}^2 + \|u - u_N\|_{L^2(\Omega)} \right), \end{aligned}$$

soient $C_1 = \max \left\{ \frac{1}{2}, c_1 \right\}$ et $C_2 = \max \left\{ \frac{1}{2}, c_2 \right\}$, alors on a :

$$\begin{aligned} \mathbf{e}_J &\leq C_1 \left(\|y_N(u_N) - y(u)\|_{L^2(\Omega)}^2 + \|y(u) - y_N(u_N)\|_{L^2(\Omega)} \right) + \\ &\quad + C_2 \left(\|u - u_N\|_{L^2(\Omega)}^2 + \|u - u_N\|_{L^2(\Omega)} \right), \end{aligned}$$

ou encore

$$e_J \leq \max \{C_1, C_2\} \left(\left(\|y_N(u_N) - y(u)\|_{L^2(\Omega)}^2 + \|y(u) - y_N(u_N)\|_{L^2(\Omega)} \right) + \left(\|u - u_N\|_{L^2(\Omega)}^2 + \|u - u_N\|_{L^2(\Omega)} \right) \right).$$

Soit $C = \max \{C_1, C_2\}$, alors on a

$$e_J \leq C \left(\left(\|y_N(u_N) - y(u)\|_{L^2(\Omega)}^2 + \|y(u) - y_N(u_N)\|_0 \right) + \left(\|u - u_N\|_{L^2(\Omega)}^2 + \|u - u_N\|_{L^2(\Omega)} \right) \right),$$

on obtient alors

$$e_J \leq C \left(\max \left\{ \|y_N(u_N) - y(u)\|_{L^2(\Omega)}^2, \|y(u) - y_N(u_N)\|_{L^2(\Omega)} \right\} + \max \left\{ \|u - u_N\|_{L^2(\Omega)}^2, \|u - u_N\|_{L^2(\Omega)} \right\} \right).$$

Par injection continue de $H^1(\Omega)$ dans $L^2(\Omega)$, on a

$$e_J \leq C \left(\max \left\{ \|y_N(u_N) - y(u)\|_{H^1(\Omega)}^2, \|y(u) - y_N(u_N)\|_{H^1(\Omega)} \right\} + \max \left\{ \|u - u_N\|_{L^2(\Omega)}^2, \|u - u_N\|_{L^2(\Omega)} \right\} \right).$$

Pour la suite, si on suppose que $\|y(u) - y_N(u_N)\|_{H^1(\Omega)} < 1$ et $\|u - u_N\|_{L^2(\Omega)} < 1$, et par application directe des Théorèmes 4.3.21 et 4.3.23, on déduit facilement l'estimation suivante

$$\begin{aligned} |J(y(u), u) - J(y_N(u_N), u_N)| &\leq \\ &\leq CN^{-1} \left(\|-\Delta p_N(u_N) - y_N(u_N) + y_d\|_{L^2(\Omega)} + \|f + Bu_N + \Delta y_N(u_N)\|_{L^2(\Omega)} \right). \end{aligned}$$

□

4.3.3 Evaluations des estimateurs d'erreurs

Pour l'évaluation de l'estimateur (4.3.22), donné par

$$\begin{aligned} \|y - y_N\|_{H^1(\Omega)} &\leq CN^{-1} \left[\int_{\Omega} |\Delta y_N + u_N + f|^2 d\mathbf{x} \right]^{\frac{1}{2}} \\ &\leq CN^{-1} \left[\int_{\Omega} (\Delta y_N + u_N + f)^2(\mathbf{x}) d\mathbf{x} \right]^{\frac{1}{2}} \\ &\leq CN^{-1} \left[\int_{\Lambda} \int_{\Lambda} (\Delta y_N + u_N + f)^2(x, y) dx dy \right]^{\frac{1}{2}}, \end{aligned}$$

on appliquera la formule de quadrature (2.6.2) d'abord par rapport à la variable x , ensuite par rapport à la variable y , pour obtenir successivement

$$\begin{aligned} \|y - y_N\|_{H^1(\Omega)} &\leq CN^{-1} \left[\int_{\Lambda} \left(\sum_{i=1}^N (\Delta y_N + u_N + f)^2(x_i, y) \omega_i \right) dy \right]^{\frac{1}{2}} \\ &\leq CN^{-1} \left[\sum_{j=1}^N \left(\sum_{i=1}^N (\Delta y_N + u_N + f)^2(x_i, y_j) \omega_i \right) \omega_j \right]^{\frac{1}{2}} \\ &\leq CN^{-1} \left[\sum_{i=1}^N \sum_{j=1}^N (\Delta y_N + u_N + f)^2(x_i, y_j) \omega_i \omega_j \right]^{\frac{1}{2}}, \end{aligned}$$

et

$$\|y - y_N\|_{H^1(\Omega)} \leq CN^{-1} \left[\sum_{i=1}^N \sum_{j=1}^N (\Delta y_N + u_N + f)^2(x_i, y_j) \omega_i \omega_j \right]^{\frac{1}{2}}. \quad (4.3.27)$$

D'après l'équation (4.2.1c) (avec $\alpha = 1$ et $B = I$), on déduit que

$$u_N = -p_N. \quad (4.3.28)$$

En substituant l'équation (4.2.1c) dans l'inégalité (4.3.23), on obtient

$$\|u - u_N\|_{L^2(\Omega)} \leq CN^{-1} \|\Delta u_N - y_N + y_d\|_{L^2(\Omega)}. \quad (4.3.29)$$

Pour l'évaluation de l'estimateur (4.3.24), on fait appel de nouveau à la formule de quadrature (2.6.2).

De la même façon, l'estimation ci-dessous, s'écrit

$$\begin{aligned}
 \|u - u_N\|_{L^2(\Omega)} &\leq CN^{-1} \left[\int_{\Omega} |\Delta u_N - y_N + y_d|^2 d\mathbf{x} \right]^{\frac{1}{2}} \\
 &\leq CN^{-1} \left[\int_{\Omega} (\Delta u_N - y_N + y_d)^2(\mathbf{x}) d\mathbf{x} \right]^{\frac{1}{2}} \\
 &\leq CN^{-1} \left[\int_{\Lambda} \int_{\Lambda} (\Delta u_N - y_N + y_d)^2(x, y) dx dy \right]^{\frac{1}{2}}.
 \end{aligned} \tag{4.3.30}$$

On appliquera la formule de quadrature (2.6.2) d'abord par rapport à la variable x , ensuite par arpport à la variable y , on obtient successivement

$$\begin{aligned}
 \|u - u_N\|_{L^2(\Omega)} &\leq CN^{-1} \left[\int_{\Lambda} \left(\sum_{i=1}^N (\Delta u_N - y_N + y_d)^2(x_i, y) \omega_i \right) dy \right]^{\frac{1}{2}} \\
 &\leq CN^{-1} \left[\sum_{j=1}^N \left(\sum_{i=1}^N (\Delta u_N - y_N + y_d)^2(x_i, y_j) \omega_i \right) \omega_j \right]^{\frac{1}{2}}, \\
 &\leq CN^{-1} \left[\sum_{i=1}^N \sum_{i=1}^N (\Delta u_N - y_N + y_d)^2(x_i, y_j) \omega_i \omega_j \right]^{\frac{1}{2}}
 \end{aligned}$$

et

$$\|u - u_N\|_{L^2(\Omega)} \leq CN^{-1} \left[\sum_{i=1}^N \sum_{i=1}^N (\Delta u_N - y_N + y_d)^2(x_i, y_j) \omega_i \omega_j \right]^{\frac{1}{2}}. \tag{4.3.31}$$

Résultats numériques

En utilisant les estimateurs (4.3.27) et (4.3.31), on obtient les résultats suivants

Degré N	Err. Cont. ³	Err. Eta. ⁴	Err. Eta. Cont ⁵
2	0.1179088032e - 0	0.1696372214e - 2	0.1196051754e - 0
3	0.4822312565e - 1	0.1061585829e - 2	0.4928471148e - 1
4	0.2526545882e - 1	0.6986595698e - 4	0.2533532478e - 1
5	0.1505884182e - 1	0.5457961204e - 4	0.1511342143e - 1
6	0.9837472942e - 2	0.1099574404e - 4	0.9848468686e - 2
7	0.6878230143e - 2	0.8481069131e - 5	0.6886711212e - 2
8	0.4963911268e - 2	0.2560023336e - 5	0.4966471292e - 2
9	0.3810036117e - 2	0.1992205088e - 5	0.3812028624e - 2

TAB. 4.2 – Evaluation des estimateurs d'erreurs.

D'après les résultat du Tableau 4.2, on obtient les représentations graphiques suivantes concernant les estimateurs (4.3.31) et (4.3.27).

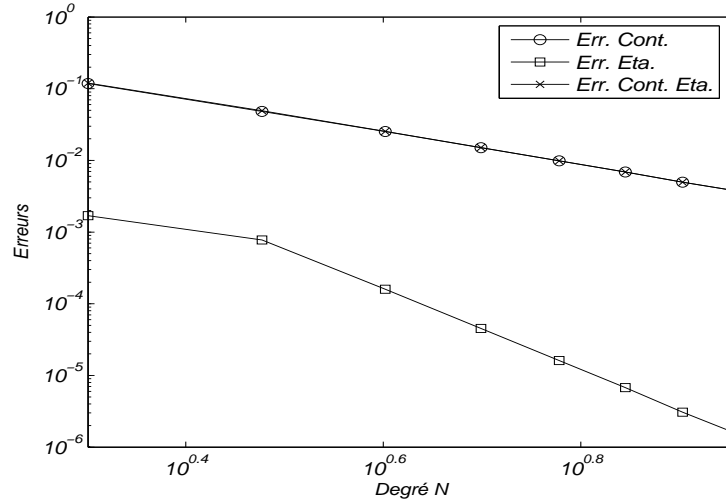


FIG. 4.5 – Courbes des estimateurs d'erreurs.

La Figure 4.5, représente l'erreur commise sur l'approximation de la fonction contrôle u et la fonction d'état y , cette erreur est représentée dans un repère avec une échelle $\log - \log$. Il est clair que l'erreur est inversement proportionnelle au degré du polynôme de Legendre, on a des droites parfaitement lisses et régulières, où l'erreur commise sur la fonction objectif et de même ordre que l'erreur commise sur la fonction contrôle. La droite qui représente l'erreur commise sur la fonction d'état est plus raide que la droite qui représente l'erreur commise sur la fonction état ce qui veut dire que l'approximation de la fonction état converge plus rapidement que l'approximation de la fonction contrôle.

Après avoir calculé les solutions approchées y_N et u_N et évalué les estimateurs d'erreur (par exemple η_y et η_u), on propose maintenant un algorithme d'adaptation, pour un choix approprié du degré du polynôme de Legendre N correspondant à la précision de calcul voulue ε recherchée.

³Erreurs sur la fonction contrôle.

⁴Erreurs sur la fonction état.

⁵Erreurs sur la fonction objectif.

Algorithme (Algorithme adaptatif relatif au choix de N).

Initialisation : Choisir N (il est souhaitable de choisir N le plus petit possible).

Itération N :

Étape 1. Calcul des solution y_N et u_N .

Étape 2. Calcul de η_y et η_u .

Critère d'arrêt :

si $\max\{\eta_y, \eta_u\} \leq \varepsilon$; **STOP.**
 si **NON**; **faire** : $N \leftarrow N + 1$; **aller à** : *Étape 1.*

Remarque 4.3.1. L'algorithme précédent est relatif au choix de N par rapport à une précision donnée ε et à titre d'exemple, on applique cet algorithme d'adaptation pour l'estimation non-couplée.

Ensuite, on obtient la solution optimale (y_N, u_N) , qui correspond à la précision ε .

4.4 Conclusion et perspectives

La méthode de discrétisation utilisée, ci-dessus, s'applique parfaitement à l'approximation du problème (\mathcal{P}), elle donne des résultats satisfaisants par rapport au volume des calculs, de la précision des résultats et de la vitesse de convergence. Il est clair que l'analyse et l'étude numérique de ce problème ont été réalisées sur tout le domaine Ω .

Notre objectif principal consistait à appliquer cette méthode d'approximation pour ce type de problème puis surmonter les différentes difficultés rencontrées et aussi maîtriser les différentes techniques utilisées pour l'analyse numérique du problème (\mathcal{P}). Nous nous proposons plus tard de généraliser cette étude pour le même problème avec et/ou sans contrainte sur l'état y et/ou sur la commande u en conservant les mêmes hypothèses mais en utilisant la méthode d'approximation de collocation.

Une piste intéressante à poursuivre serait de considérer la discrétisation des problèmes de commande optimale gouvernés par des inéquation variationnelles, modélisées par le problème mathématique abstrait suivant

$$\min \{J(y, u), y = \mathcal{T}(u), u \in \mathcal{U}_{ad} \subset \mathcal{U}, y \in K \subset \mathcal{X}\},$$

où \mathcal{T} est un opérateur qui à u associe y la solution (si elle existe et est unique) de

$$\forall z \in K, \quad \langle A(y, u), z - u \rangle \geq 0.$$

Ces problèmes sont intéressants car ils permettent de contrôler certaines frontières libres via l'inéquation variationnelle associée et débouchent sur des questions d'optimisation de formes. Les résultats d'existence et d'unicité sont nombreux [14]. Concernant les conditions d'optimalité, plusieurs difficultés sont à signaler :

- Même si l'opérateur \mathcal{A} est linéaire, l'opérateur \mathcal{T} est toujours non-linéaire.
- La seconde difficulté est qu'il est impossible d'assurer une quelconque propriété de différentiabilité (même de Gâteaux) de l'application \mathcal{T} . La seule différentiabilité "faible" utilisable est la notion de différentiabilité conique proposée par Mignot dans [64].

Annexe A

Autres techniques pour le calcul des conditions d'optimalité

A.1 Lagrangien et multiplicateur de Lagrange

On peut aussi définir le Lagrangien du problème (\mathcal{P}) sur $H_0^1(\Omega) \times L^2(\Omega) \times H_0^1(\Omega)$ par :

$$\mathcal{L}(y(u), p(u), u) = J(y(u), u) + \langle p(u), -\Delta y(u) + u - f \rangle_{H_0^1, H^{-1}}. \quad (\text{A.1.1})$$

On peut montrer que $(\bar{y}(u), \bar{u})$ est solution du problème de commande optimale si et seulement si¹ $(\bar{y}(u), \bar{p}(u), \bar{u})$ est un point-selle de \mathcal{L} sur $H_0^1(\Omega) \times L^2(\Omega) \times H_0^1(\Omega)$ c'est-à-dire

$$\begin{aligned} \forall (y(u), p(u), u) \in H_0^1(\Omega) \times L^2(\Omega) \times H_0^1(\Omega) \\ \mathcal{L}(\bar{y}(u), p(u), \bar{u}) \leq \mathcal{L}(\bar{y}(u), \bar{p}(u), \bar{u}) \leq \mathcal{L}(y(u), \bar{p}(u), u). \end{aligned} \quad (\text{A.1.2})$$

L'état adjoint $\bar{p}(u)$ est le multiplicateur de Lagrange associé à la contrainte : $-\Delta y(u) + u = f$. Bien entendu si on ajoute d'autres contraintes sur l'état et la commande, la question essentielle est de prouver l'existence de multiplicateurs de Lagrange associés de façon à "incorporer" les contraintes dans le Lagrangien pour obtenir un problème plus simple "sans contrainte" [11],[12].

A.2 Principe du maximum de Pontryagin

A défaut de régularité, on peut envisager une alternative aux conditions nécessaires d'optimalité du premier ordre. Il s'agit du principe du minimum de Pontryagin, que l'on

¹en général la condition de point selle n'est que suffisante :ici le caractère nécessaire et suffisant provient du cadre qui est convexe.

peut énoncer dans l'exemple précédent de la manière suivante : on définit le Hamiltonien du système sur \mathbb{R}^4 par

$$\mathcal{H}(y, p, u) = \underbrace{\frac{1}{2}(y - y_d)^2 + \frac{\alpha}{2}u^2}_{\text{Terme intervenant dans l'intégrale de } J} + p \underbrace{u + f.}_{\text{second membre de l'EDP}} \quad (\text{A.2.1})$$

Appliquons la relation de droite de (A.1.2) à $y(u) = \bar{y}(u)$: on obtient

$$\forall u \in U_{ad} \quad J(\bar{y}, \bar{u}) + \langle \bar{p}, -\Delta \bar{y} - \bar{u} - f \rangle_{H_0^1, H^{-1}} \leq \leq J(\bar{y}, u) + \langle \bar{p}, -\Delta \bar{y} - u - f \rangle_{H_0^1, H^{-1}}, \quad (\text{A.2.2})$$

c'est à dire $\forall u \in U_{ad}$

$$\int_{\Omega} \left(\frac{1}{2} (\bar{y} - y_d)^2 + \frac{\alpha}{2} \bar{u}^2 - \bar{p} (\bar{u} - f) \right) d\mathbf{x} \leq \int_{\Omega} \left(\frac{1}{2} (\bar{y} - y_d)^2 + \frac{\alpha}{2} u^2 - \bar{p} (\bar{u} - f) \right) d\mathbf{x}, \quad (\text{A.2.3})$$

$$\forall u \in U_{ad} \quad \int_{\Omega} \mathcal{H}(\bar{y}, -\bar{p}, \bar{u}) d\mathbf{x} \leq \int_{\Omega} \mathcal{H}(\bar{y}, -\bar{p}, u) d\mathbf{x}. \quad (\text{A.2.4})$$

Ceci entraîne que, pour tout v dans \mathbb{R} tel qu'il existe u dans U_{ad} vérifiant $u = v$,

$$\mathcal{H}(\bar{y}, -\bar{p}, \bar{u}) \leq \mathcal{H}(\bar{y}, v, -\bar{p}) \text{ p.p. dans } \Omega. \quad (\text{A.2.5})$$

Annexe B

Calcul matriciel des nœuds ξ_j

B.1 Calcul des nœuds ξ_j

Une manière simple et efficace de calculer les nœuds ξ_j , $1 \leq j \leq N - 1$, consiste à exhiber une matrice symétrique dont ils sont valeurs propres. Pour cela on pose :

$$NL'_{N+1}(\xi) = (2N + 1)\xi L'_N(\xi) - (N + 1)L_{N-1}(\xi), \quad (\text{B.1.1})$$

$$J_n^* = L'_{n+1} \sqrt{\frac{n + \frac{3}{2}}{(n + 1)(n + 2)}},$$

ce qui signifie que les J_n^* , $0 \leq n \leq N$, forment une base orthonoemée de $\mathbb{P}_N(\Lambda)$ pour le produit scalaire

$$(\varphi, \psi) \mapsto \int_{\Lambda} \varphi(\zeta) \psi(\zeta) (1 - \zeta^2) d\zeta. \quad (\text{B.1.2})$$

La formule de récurrence (B.1.1) s'écrit

$$\begin{aligned} n \sqrt{\frac{(n + 1)(n + 2)}{n + \frac{3}{2}}} J_n^*(\zeta) &= (2n + 1) \sqrt{\frac{n(n + 1)}{n + \frac{1}{2}}} \zeta J_{n-1}^*(\zeta) - \\ &\quad - (n + 1) \sqrt{\frac{n(n - 1)}{n - \frac{1}{2}}} J_{n-2}^*(\zeta), \end{aligned} \quad (\text{B.1.3})$$

ou encore

$$2\zeta J_n^*(\zeta) = (2n + 1) \sqrt{\frac{n(n + 2)}{(n + \frac{1}{2})(n + \frac{3}{2})}} J_n^*(\zeta) + \sqrt{\frac{(n - 1)(n + 1)}{(n - \frac{1}{2})(n + \frac{1}{2})}} J_{n-2}^*(\zeta), \quad (\text{B.1.4})$$

ce qui prouve que les ξ_j , $1 \leq j \leq N - 1$, sont les valeurs propres de la matrice

$$\mathcal{M} = \begin{pmatrix} 0 & \gamma_1 & \cdots & 0 & 0 \\ \gamma_1 & \ddots & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \gamma_{N-2} \\ 0 & 0 & \cdots & \gamma_{N-2} & 0 \end{pmatrix},$$

avec

$$\gamma_n = \frac{1}{2} \sqrt{\frac{n(n+2)}{(n+\frac{1}{2})(n+\frac{3}{2})}}. \quad (\text{B.1.5})$$

La matrice \mathcal{M} est une matrice symétrique tridiagonale à diagonale nulle. Finalement les poids ρ_j $1 \leq j \leq N - 1$ peuvent se calculer soit à partir des vecteurs propres de la matrice \mathcal{M} , soit par la relation (2.6.1).

Annexe C

Estimation de l'erreur a posteriori par dualité

C.1 Introduction

On utilise la théorie de la dualité en analyse convexe pour l'estimation a posteriori de l'erreur. La méthode peut être vue comme étant un développement théorique pour l'obtention d'une majoration de l'estimation de la norme de l'énergie (3.2.4). Il faut savoir qu'il existe plusieurs techniques et théories pour la définition du problème adjoint d'un certain problème variationnel [48]. Dans la théorie de l'analyse convexe. Pour l'estimation a posteriori de l'erreur, on commence d'abord par définir la fonction d'énergie de notre équation (s'il s'agit bien sûr d'une équation différentielle ordinaire ou d'une équation aux dérivées partielles, ou même une inéquation variationnelle). Le problème se transforme en un problème de minimisation avec et / ou sans contraintes. Par exemple, soient V , Q deux espaces normés, et V^* , Q^* leur espace duals respectif, on suppose l'existence d'un opérateur Λ dans $\mathcal{L}(V, Q)$, avec son adjoint Λ^* dans $\mathcal{L}(V^*, Q^*)$. Soit J , une fonction (dans notre cas la fonction d'énergie) de $V \times Q$ dans $\overline{\mathbb{R}}$ on considère le problème de minimisation (qui est équivalent au problème variationnel considéré ci-dessus)

$$\inf_{v \in V} J(v, \Lambda v). \quad (\text{C.1.1})$$

On définit son problème dual par

$$\sup_{q^* \in Q^*} [J^*(\Lambda^* q^*, -q^*)], \quad (\text{C.1.2})$$

où J^* est la fonction conjuguée de J :

$$J^*(v^*, q^*) = \sup_{v \in V, q \in Q} [\langle v, v^* \rangle + \langle q, q^* \rangle - J(v, q)]. \quad (\text{C.1.3})$$

Pour la relation entre les deux problèmes (C.1.2) et (C.1.3), on a le théorème de dualité suivant :

Théorème C.1.1. *On suppose que :*

1. V est un espace de Banach réflexif, Q est un espace normé.
2. $J : V \times Q \rightarrow \overline{\mathbb{R}}$ est une fonction propre, semi continue, convexe.
3. $\exists u_0 \in V$, telles que $J(u_0, \Lambda u_0) < \infty$ et $q \mapsto J(u_0, q)$ est continue à Λu_0 .
4. $J(v, \Lambda v) \rightarrow +\infty$, si $\|v\| \rightarrow \infty$, pour v dans V .

Alors le problème (C.1.2) admet une solution u dans V , et le problème (C.1.3) admet une solution p^* dans Q^* , et

$$J(u, \Lambda u) = -J^*(\Lambda^* p^*, -p^*). \quad (\text{C.1.4})$$

En plus, si J est strictement convexe, alors la solution du problème (C.1.2) est unique.

C.1.1 Estimation d'erreur a posteriori dans un context général abstrait

Soit u dans V une solution du problème de minimisation (C.1.2). On suppose que J est Gâteaux différentiable au point u . Soit v dans V un arbitraire dans V avec $J(v, \Lambda v) < \infty$. On pose :

$$D(u, v) = J(v, \Lambda v) - J(u, \Lambda u) - \langle J'(u, \Lambda u), (v - u, \Lambda v - \Lambda u) \rangle, \quad (\text{C.1.5})$$

où, $\langle \cdot, \cdot \rangle$ est le crochet de dualité entre $(V \times Q)^*$ et $(V \times Q)$.

Théorème C.1.2. *On suppose d'une part que les conditions du Théorème C.1.1, restent l'existence de la solution du problème (C.1.2). Si J est Gâteaux différentiable à u , alors*

$$D(u, v) \leq J(v, \Lambda v) + J^*(\Lambda^* q^*, -q^*), \quad \forall q^* \in Q^*. \quad (\text{C.1.6})$$

En utilisant le Théorème C.1.2, le processus pour déduire la différence entre u et v est décomposé en deux étapes

Etape 1 Trouver une borne inférieure de $D(u, v)$, qui mesure la différence entre u et v . Généralement, cette borne inférieure sera l'erreur dans la norme d'énergie ou une quantité qui dépend de $\|v - u\|$, dans un espace bien défini.

Etape 2 Trouver un q^* adéquat de telle sorte que l'estimation (C.1.6), soit le plus précise possible.

Bibliographie

- [1] A. ABDELLI, *Résolution Numérique de Quelques Problèmes Issus du Contrôle Optimal*, Mémoire de Magister, Université Badji Mokhtar, Annaba, 1995.
- [2] R.A. ADAMS, *Sobolev Spaces*, Academic Press, New-York, San Francisco, London, 1975.
- [3] JULIÁN AGUIRRE ET JUDITH RIVAS, *Hermite pseudospectral approximations. An error estimate*, Journal of Mathematical Analysis and Applications. **304**, (2005), 189-197.
- [4] VIOREL ARNĂUTU ET PEKKA NEITTAANMÄKI, *Discretization estimates for an elliptic control problem*, Numerical Functional Analysis and Optimization. **19**(5 et 6), (1998), 431-464.
- [5] J. BARANGER ET H. EL AMRI, *Estimateurs a posteriori d'erreur pour le calcul adaptatif d'écoulement Quasi-Newtoniens*. Mathematical Modelling and Numerical Analysis. **25**(1), (1991), 31-48.
- [6] V. BARBU, *Optimal control of variational inequalities*, Pitman, London, 1984.
- [7] V. BARBU, *Analysis and control of non linear Infinite Dimensional Systems*, Mathematics in Science and Engineering. 190, Academic Press, San Diego, 1993.
- [8] G. BARLES, *Solutions de viscosité des équations d'Hamilton-Jacobi*, Collection Mathématiques et Applications, **17**, Springer-Verlag, Berlin Heidelberg 1994.
- [9] R. BECKER, H. KAPP ET R. RANNACHER, *Adaptive finite element methods for optimal control of partial differential equations : Basic concept*, SIAM Journal on Control and Optimization, **39**(1), (2000), 2401-2416.
- [10] J.D. BENAMOU, *A domain decomposition method with coupled transmission conditions for the optimal control of systems governed by elliptic partial differential equations*, SIAM Journal on Numerical Analysis, **33**(6), (1996), 2401-2416.
- [11] M.BERGOUNIOUX ET K. KUNISCH, *Lagrange multipliers structure for state constrained optimal control*, Systems and Control Letters. **48**(3-4), (1973), 169-176.
- [12] M.BERGOUNIOUX ET D.TIBA, *General optimality conditions for constrained convex control problems*, SIAM Journal on Control and Optimization. **34**(2), (1996), 698-711.
- [13] M.BERGOUNIOUX ET H.ZIDANI, *Pontryagin principle for problems governed by parabolic variational inequalities*, SIAM Journal on Control and Optimization. **37**(4), (1999), 1273-1290.

- [14] M. BERGOUNIOUX ET S. LENHART, *Optimal control of the bilateral obstacle problems*, SIAM Journal on Control and Optimization. **43**(1), (2004), 240-255.
- [15] C. BERNARDI ET Y. MADAY, *Polynomial interpolation results in Sobolev spaces*, Journal of computational and applied mathematics. **43**, (1992), 53-80.
- [16] C. BERNARDI ET Y. MADAY, *Properties of some weighted Sobolev spaces and application to spectral approximations*, SIAM Journal on Numerical Analysis. **226**, (1989), 769-829.
- [17] C. BERNARDI, Y. MADAY ET F. RAPETTI, *Discrétisations variationnelles de problèmes aux limites elliptiques*, Springer-Verlag, Berlin Heidelberg 2004.
- [18] F. BONNANS ET E. CASAS, *An extension of pontryagin's principle for state-constrained optimal control of semilinear elliptic equations and variational inequalities*, SIAM Journal on Control and Optimization. **33**(1), (1995), 274-298.
- [19] R. BUTT, *Optimal shape design of systems governed by variational inequalities*, Mathematical and Computer Modelling. **25**, (1997), 47-55.
- [20] GIUSEPPE BUTTAZZO ET ALFRED WAGNER, *On the optimal shape of rigid body supported by an elastic membrane*, Nonlinear Analysis. **39**, (2000), 47-63.
- [21] H. BREZIS, *Analyse fonctionnelle*, masson, 1983.
- [22] C. CANUTO, M. Y. HUSSAINI, A. QUARTERONI ET T. A. ZANG, *Spectral Methods in Fluid Dynamics*, Springer, New York, (1988).
- [23] E. CASAS ET F. TRÖLTZSCH, *Error estimates for linear-quadratic elliptic control problems*, Analysis and optimization of differential systems (Constanta, 2002), 89-100, Kluwer Academic Publication, Boston, MA, (2003).
- [24] X. CHEN, *Finite difference smoothing solutions of nonsmooth constrained optimal control problems*, Numerical functional analysis and optimization. **26**, (2005), 49-68.
- [25] P. G. CIARLET ET J. L. LIONS (EDS), *Handbook of numerical analysis*, **I**(1), Elsevier, North-Holland, (1990).
- [26] P. G. CIARLET ET J. L. LIONS (EDS), *Handbook of numerical analysis*, **II**(1), Elsevier, North-Holland, (1991).
- [27] E. A. COUSTIAS, T. HAGSTROM ET D. TORRES, *An efficient spectral method for ordinary differential equations with rational function coefficients*, Mathematics of computation. **65**(214), (1996), 611-635.
- [28] R. DAUTRAY ET J. L. LIONS, *Mathematical analysis and numerical methods for science and technology*, Vol. **2**, Springer-Verlag, Berlin, 1988.
- [29] S. DEHILIS, *Methodes Spectrales pour la Résolution Numérique de Quelques Problèmes de Commande Optimale avec Contraintes d'Égalité et d'Inégalité*, Mémoire de Magister, Université Badji Mokhtar, Annaba, 1999.
- [30] PEDRO DIEZ, JUAN JOSÉ EGOZCUE ET ANTONIO HUERTA, *A posteriori error estimation for standard finite element analysis*, Computer Methods in Applied Mechanics and Engineering. **163**, (1998), 141-157.

- [31] R. DOUGLAS C., *A finite difference technique for solving optimization problems governed by linear functional-differential equations* Journal of Differential Equations. **32**(2), (1979), 193-232.
- [32] R. V. DOOREN ET J. VLASSEN BROECK, *A New Look at the Brachistochrone Problem*, Journal of Applied Mathematics and Physics (ZAMP). **31**, 1980.
- [33] R. V. DOOREN ET J. VLASSEN BROECK, *Chebyshev Series Solution of the Controlled Duffing Oscillator*, Journal of Computational Physics. **47**, (1982), 321-329.
- [34] M. ELKADY ET ELSAYED M.E. ELBARBARY, *A Chebyshev expansion method for solving nonlinear optimal control problems*, Applied Mathematics and Computation. **129**, (2002), 171-182.
- [35] G. N. ELNAGAR ET A. KHAMAYSEH, *On the Optimal Spectral Chebyshev Solution of a Controlled Nonlinear Dynamical System*, IMA Journal of Applied Mathematics. **58**, (1997), 147-157.
- [36] ELSAYED M.E. ELBARBARY, *Legendre expansion method for the solution of the second-and fourth-order elliptic equations*, Mathematics and Computers in Simulation. **59**, (2002), 389-399.
- [37] ELSAYED M.E. ELBARBARY ET M. ELKADY, *Chebyshev finite difference approximation for the boundary value problems*, Applied Mathematics and Computation. **139**, (2003), 513-523.
- [38] G. N. ELNAGAR ET M. RAZZAGHI, *A Chebyshev spectral method for the solution of nonlinear optimal control problems*, Applied Mathematical Modelling. **21**, (1997), 255-260.
- [39] R. S.FALK, *Approximation of a Class of Optimal Control Problems with Order of Convergence Estimates*, Journal of Mathematical Analysis and Applications **44**, (1973), 28-47.
- [40] P. FAURRE, *Analyse numérique, Notes d'Optimisation.*, Ellipses, France, 1988.
- [41] R. GHANEM, *Méthode spectrale de Chebychev pour la résolution numérique de problème de commande optimale*, Mémoire de Magister, Université Badji Mokhtar, Annaba, Algérie, 2000.
- [42] D. GIVOLI, *A direct approach to the finite element solution of elliptic optimal control problem*, Numerical Methods for Partial Differential Equations. **15**, (1999), 371-388.
- [43] R. GLOWINSKI, *Numerical methods for nonlinear variational problems*, Springer, New York, 1984.
- [44] R.GLOWINSKI, J.L.LIONS ET R.TRÉMOLIÈRES, *Analyse Numérique des inéquations variationnelles*, Dunod, 1976.
- [45] C. GOONG ET M. WENDELL H. JR., *Finite elements and terminal penalization for quadratic cost optimal control problems governed by ordinary differential equations*. SIAM Journal on Control and Optimization, **19**(6), (1981), 744-764.

- [46] D. GOTTLIEB ET S.A. ORZAG, *Numerical Analysis of Spectral Methods, Theory and applications*, Siam Publications, Philadelphia, 1977.
- [47] W. HACKBUSCH, *Fast Solution of Elliptic Control Problems.*, Journal of Optimization theory and applications, **31**(4), (1980),565-581.
- [48] W. HAN, *A posteriori Error Analysis via Duality Theory*, Springer, New York, 2005.
- [49] J.HASLINGER ET P.NEITTAANMÄKI, *Finite Element Approximation for Optimal Shape Design*, John Wiley and Sons LTD, New York, 1988.
- [50] D. B.HAIDVOGEL ET T. ZANG, *The Accurate Solution of Poisson's Equation by Expansion in Chebyshev Polynomials.*, Journal of Computational Physics, **30**(2),(1979), 167-180.
- [51] T. S. HORNER, *A Double Chebyshev Series Methods for Elliptic Partial Differential Equations*, Numerical Solution of Partial Differential Equations, J. Noye (editor), 573-590, 1982.
- [52] E. HUNTLEY, *Finite Methods for an Optimal Steady-State Control Problem*, International Journal for Numerical Methods in Engineering **12**, (1978),1375-1382.
- [53] HUSSEIN JADDU, *Direct solution of nonlinear optimal control problems usingquasilinearization and Chebyshev polynomials*, Journal of the Franklin Institute. **339**,(2002), 479-498.
- [54] HUSSEIN JADDU, *Spectral method for constrained linear quadratic optimal control*, Mathematics and Computers in Simulation. **58**,(2002), 159-169.
- [55] T. KÄRKKÄINEM ET T. RÄISÄNEM, *A numerical method for general optimal control problem*, University of Jyväskylä, Departement of mathematics, Laboratory of scientific computing, Report **16**,(1996).
- [56] AHMED K. KHALIFA A, ELSAYED M.E. ELBARBARY ET MOHAMED A. ABD ELRAZEK, *Chebyshev expansion method for solving second and fourth-order elliptic equations*, Applied Mathematics and Computation. **135**,(2003), 307-318.
- [57] J. L. LIONS ET E. MAGENS, *Problèmes aux limites non homogènes et applications, Vol. 1*, Dunod, Paris, 1968.
- [58] J. L. LIONS, *Optimal Control of Systems Governed by Partial Differential Equations*. Springer-Verlag, New York, 1971.
- [59] WENBIN LIU ET NINGNING YAN, *A posteriori error estimates for control problems governed by nonlinear elliptic equations*, Applied Numerical Mathematics. **47**,(2003), 173-187.
- [60] R. LU ET AL., *Adaptive Finite Element Approximation for Distributed Elliptic Optimal Control Problems*, SIAM Journal on Control and Optimization. **41**,(1994),1498-1505.
- [61] Y. MADAY, ET A. QUARTERONI, *Legendre and Chebyshev Spectral Approximations of Burgers' Equation*, Numerische Mathematik. **37**, (1981), 321-332.

- [62] K. MALANOWSKI, *Convergence of Approximations vs. Regularity of Solutions for Convex, Control-Constrained Optimal-Control Problems*, Applied Mathematics and Optimization. **8**, (1981), 69-95.
- [63] R. A. MERIC, *Finite Methods for an Optimal Steady-State Control Problem*, International Journal for Numerical Methods in Engineering. **12**, (1978), 1375-1382.
- [64] F. MIGNOT, *Contrôle dans les inéquations variationnelles elliptiques*, Journal of Functional Analysis. **22**, (1976), 466-476.
- [65] F. MIGNOT ET J.P. PUEL, *Optimal Control in Some Variational Inequalities*, SIAM Journal on Control and Optimization. **22**, (1984), 466-476.
- [66] T. MIYOSHI, *A finite element method for the solutions of fourth order partial differential equations*, Kumamoto Journal of Science. (Mathematics). **9**, (1973), 87-116.
- [67] M.A. NOOR, *Finite difference technique for solving obstacle problems*, Applied Mathematics Letters. **1**(3), (1988), 267-271.
- [68] S. H. OH ET R. LUUS, *Use of Orthogonal Collocation Methods in Optimal Control*, International Journal of Control. **26**(5), (1977), 657-673.
- [69] S. A. ORSZAG, *Comparison of pseudospectral and spectral Approximation*, Studies in Applied Mathematics. **L1**, (1972).
- [70] L. PONTRIAGUINE, V. BOLTJANSKI, R. GAMKRELIDZE ET E. MICHTCHEIKO, *Théorie Mathématique des Processus Optimaux*, Edition Mir, Moscow. 1974.
- [71] MOHSEN RAZZAGHI ET MOHAMMAD HABIBI, *Application of Legendre series to the control problems governed by linear parabolic equations*, Mathematics and Computers in Simulation. **42**, (1996), 77-84.
- [72] R. RANNACHER, *Adaptive Galerkin finite element methods for partial differential equations*, Journal of Computational and Applied Mathematics. **128**, (2001), 205-233.
- [73] ABDUR RASHID, *Error estimation of pseudospectral method for solving the Barotropic vorticity equation*, Novi Sad Journal of Mathematics. **35**, (2005), 137-149.
- [74] M. RAZZAGHI ET G.ELNAGAR, *Numerical Solution of the Controlled Duffing Oscillator by the Pseudospectral Method*, Journal of Computational and Applied Mathematics. **56**(3), (1994), 253-261.
- [75] A. RINCON ET I-SHIH LIU, *On Numerical Approximation of an Optimal Control Problem in Linear Elasticity*, Divulgaciones Matemáticas. **11**, (2003), 91-107.
- [76] R.W.H. SARGENT, *Optimal control*, Journal of Computational and Applied Mathematics. **124**, (2000), 361-371.
- [77] L.SCHWARTZ, *Théorie des distributions*, Hermann, Paris, 1978.
- [78] J.SHEN, *Efficient Spectral-Galerkin Method I. Direct Solvers for the Second and Fourth Order Equations Using Legendre Polynomials*, SIAM Journal of scientific computation. **15**(6), (1994), 1489-1505.

-
- [79] J. SHEN, *Efficient Spectral-Galerkin Method II. Direct Solvers for the Second and Fourth Order Equations Using Chebyshev Polynomials*, SIAM Journal of scientific computation. **16**(1), (1995), 74-87.
- [80] JIE SHEN ET LI-LIAN WANG, *Error analysis for mapped Legendre spectral and pseudospectral methods*, SIAM Journal of Numerical Analysis. **42**, (2004), 326-349.
- [81] G. SZEGO, *Orthogonal Polynomials*, American Mathematical Society, 1991.
- [82] VIDAR THOMÉE, *From finite differences to finite elements A short history of numerical analysis of partial differential equations*, Journal of Computational and Applied Mathematics. **128**, (2001), 1-54.
- [83] D. TIBA ET F. TRÖLTZSCH, *Error estimates for the discretization of state constrained convex control problems*, Numerical Functional Analysis and Optimization. **17**(9-10), (1996), 1005-1028.
- [84] R. TYRRELL ROCKAFELLAR, *Duality and Stability in Extremum Problems Involving Convex Functions*, SIAM Journal on Scientific Computing. **15**, (1967), 167-187.
- [85] A. VIOREL ET P. NEITTAANMAKI, *Discretization estimates for an elliptic control problem*, Numerical Functional Analysis and Optimization. **19**, (1998), 431-464.
- [86] J. VLASSENBROECK, *A Chebyshev Polynomial Method for Optimal Control With State Constraints*, Automatica. **24**, (1988), 499-506.
- [87] P. WILLIAMS, *A Gauss-Lobatto quadrature method for solving optimal control problem*, ANZIAM J. **47**(EMAC2005), (2006), C101-C115.
- [88] H. J. WIRSCHING, *On Finite Element Approximation of Variational Inequalities Arising from Elliptic Control Problems*, Numerical Functional Analysis and Optimization. **19**, (1998), 917-932.
- [89] NINGNING YAN, *A posteriori error estimators of gradient recovery type for FEM of model optimal control problem*, Advances in computational mathematics. **19**, (2003), 323-336.
- [90] KÔSAKU YOSIDA, *Functional analysis*. Springer-Verlag, New York, 1975.
- [91] CHEN ZHIMING, *A posteriori error analysis and adaptive methods for parabolic problems*, Recent progress in computational and applied partial differential equations (Zhangjiajie, 2001), (2002), 145-156, Kluwer/Plenum, New York.
- [92] ZHANG ZHIMIN, *A posteriori error estimates on irregular grids based on gradient recovery. A posteriori error estimation and adaptive computational methods*, Advances in computational mathematics. **15**(1-4), (2001), 363-374.