



HAL
open science

Sur des techniques déterministes et stochastiques appliquées aux problèmes d'identification

Hélène Dousteysier-Buvat

► **To cite this version:**

Hélène Dousteysier-Buvat. Sur des techniques déterministes et stochastiques appliquées aux problèmes d'identification. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1995. Français. NNT: . tel-00346058

HAL Id: tel-00346058

<https://theses.hal.science/tel-00346058>

Submitted on 11 Dec 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Présentée par

Hélène DOUSTEYSSIER épouse BUVAT

Pour obtenir le titre de

Docteur de l'Université Joseph Fourier - Grenoble I
(arrêtés ministériels du 5 juillet 1984 et du 30 mars 1992)

Spécialité : **Mathématiques Appliquées**

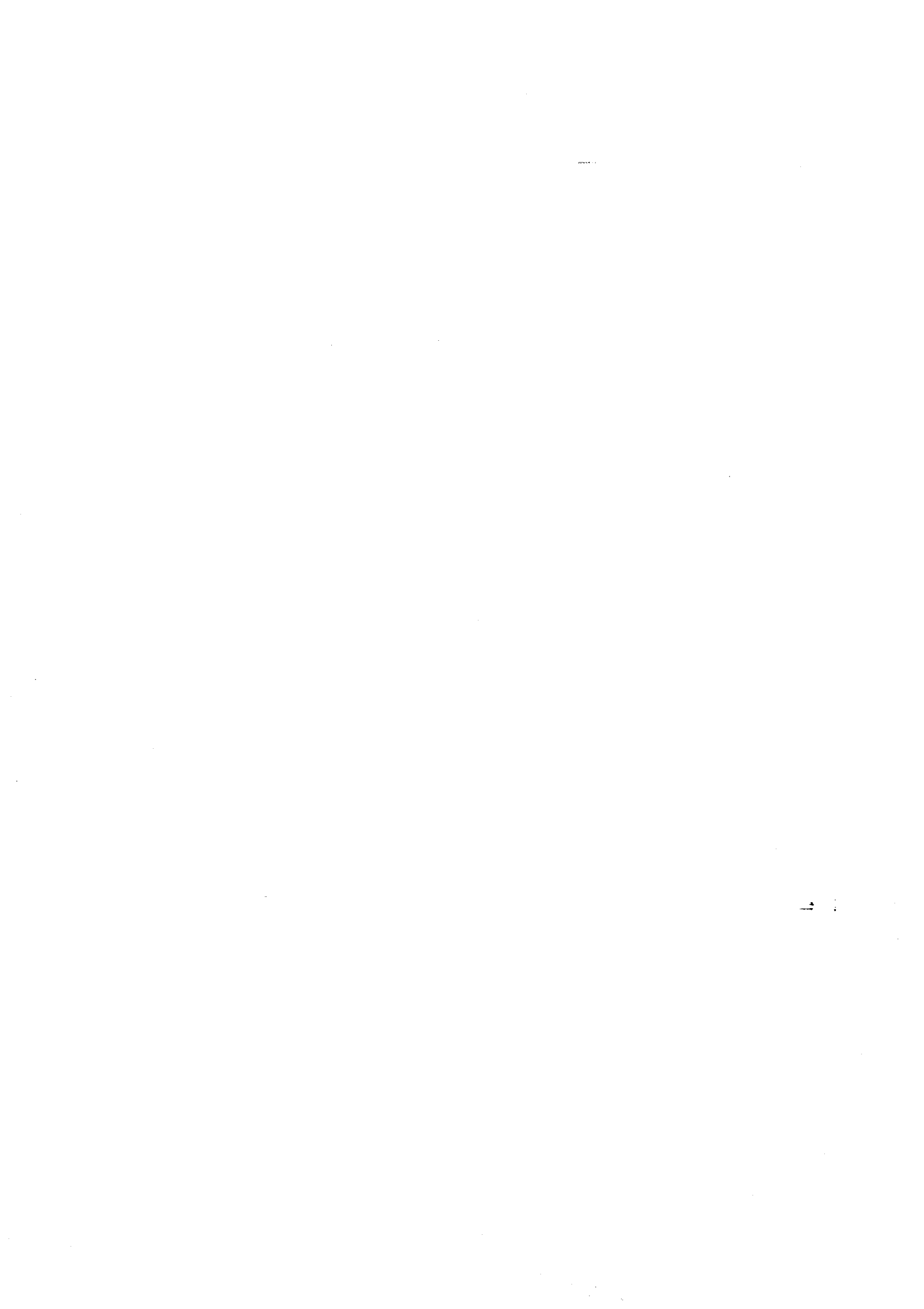
Sur des techniques déterministes et stochastiques appliquées aux problèmes d'identification.

Date de soutenance : 19 septembre 1995

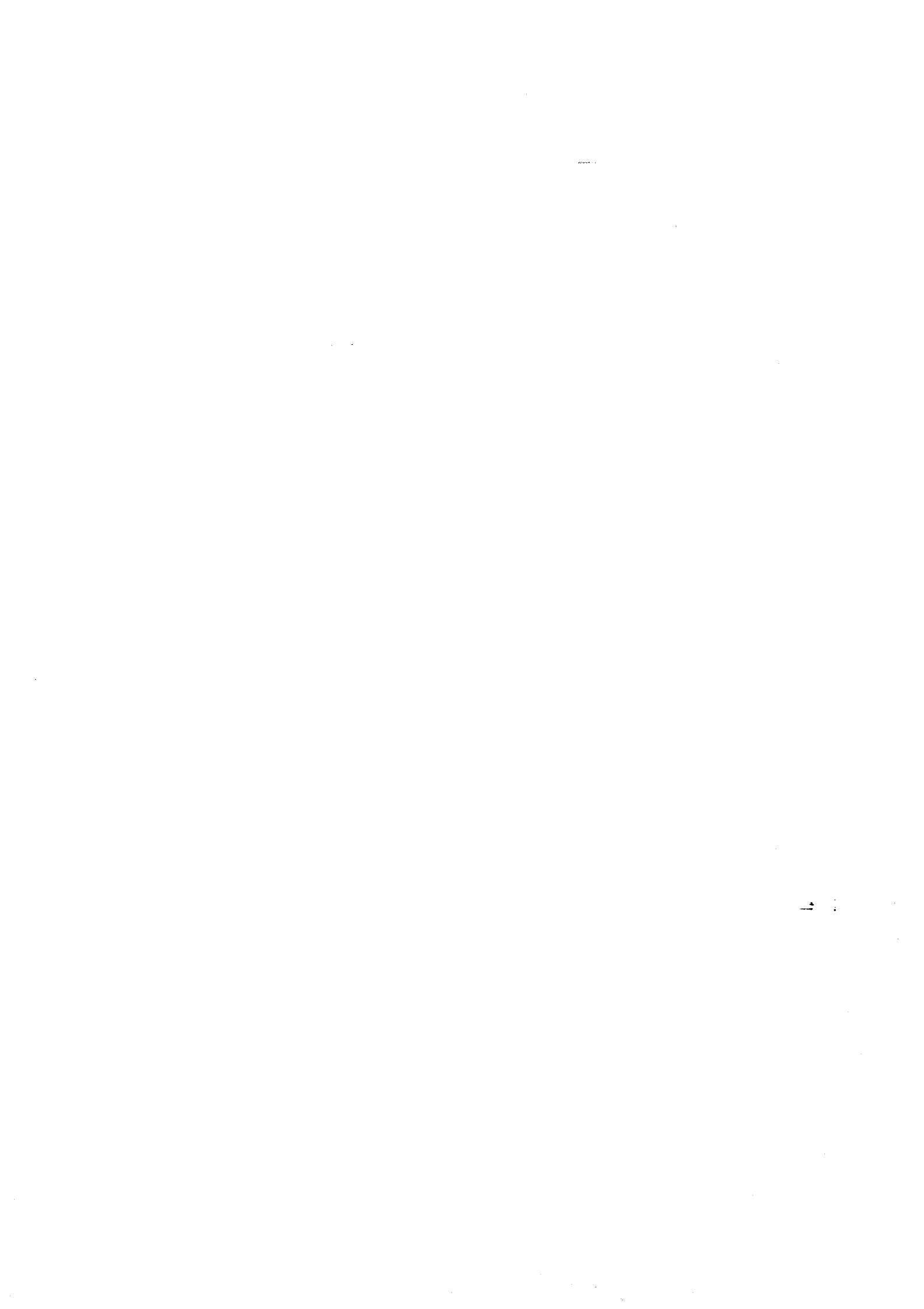
Composition du jury :

M. P. WITOMSKI	Président
M. B. SARAMITO	Rapporteur
M. M. VOGELIUS	Rapporteur
M. A. ANTONIADIS	Examineur
M. J. BLUM	Examineur

Thèse préparée au sein du laboratoire LMC-IMAG
Projet IDOPT (CNRS/INPG/UJF/INRIA)



A Cécile,
à mes parents.



Je tiens tout d'abord à remercier Jacques Blum pour avoir dirigé ce travail et pour ses précieux conseils tout au long de cette thèse. Il a su me donner le goût des problèmes de contrôle optimal et me faire profiter de sa large culture mathématique.

Comment ne pas remercier Anestis Antoniadis qui a toujours su m'encourager avec gentillesse et me faire profiter de sa bonne humeur. Je le remercie d'avoir bien voulu faire partie de ce jury.

Je remercie M. Patrick Witomski qui me fait l'honneur de présider ce jury.

Je remercie MM. Bernard Saramito et Michael Vogelius qui m'ont fait l'honneur d'accepter d'être rapporteurs de cette thèse.

Je remercie également Pierre Vial pour sa disponibilité et la gentillesse dont il a fait preuve en m'aidant à mieux comprendre les ondelettes.

Je tiens tout particulièrement à remercier Bruno Luong, tout d'abord pour ses connaissances mathématiques et son intuition informatique qu'il a su partager et qui m'ont permis de gagner un temps précieux, et surtout pour son amitié.

Je n'oublierai pas les thésards et néanmoins amis du LMC : Emmanuelle, Laurence, Yves, Malek, Emmanuel, Otman et tous les autres ... sans qui l'ambiance au labo n'aurait pas été la même, et sans lesquels je ne connaîtrais toujours pas le surf des neiges.

Merci enfin à Rémi, qui a dû me supporter et m'encourager dans les moments difficiles tout au long de ces années et a su compenser avec succès mon absence à la maison.



Table des matières

Introduction générale	1
Chapitre I	
Le problème général	5
1 Le problème de contrôle	6
2 Quelques notions d'optimisation	7
2.1 Propriétés liées à la convexité	7
2.2 Existence d'un optimum	7
2.3 Caractérisation des optima	8
2.4 Algorithmes de résolution	9
3 Le calcul du gradient	11
3.1 Calcul approché	11
3.2 Calcul exact	11
4 Introduction au contrôle optimal des EDP	13
4.1 Aperçu général	13
4.2 Contrôle optimal des EDP elliptiques	14
4.2.1 Le problème direct	14
4.2.2 Le problème de contrôle	15
4.2.3 Interprétation physique	15
4.2.4 Existence et caractérisation de l'optimum	15
4.2.5 Approche lagrangienne	17
4.3 Contrôle optimal des EDP paraboliques	20
4.3.1 Le problème direct	22
4.3.2 Le problème de contrôle	23
4.3.3 Existence et caractérisation de l'optimum	24
4.3.4 Approche lagrangienne	25
4.4 Résolution numérique	27
5 Introduction d'un terme de régularisation	28
5.1 L'approche "classique" de Tikhonov et Arsenine	28
5.2 Construction d'opérateurs régularisants	29
5.3 L'approche de K. Ito et K. Kunisch	31
Chapitre II	
Identification pour des problèmes inverses non linéaires	33
1 Identification du coefficient de diffusion de la chaleur	34
1.1 Position du problème	34
1.2 Obtention du système d'optimalité	36
1.3 Etude de la base de décomposition	38

1.3.1	Quelques rappels sur les B -splines	38
1.3.2	Application au problème d'identification	40
1.4	Linéarisation du problème	42
1.5	Discrétisation des systèmes d'EDP	44
1.6	Validation du gradient	46
1.7	Résultats numériques	47
1.7.1	Identification sans bruit	47
1.7.2	Identification avec bruit, sans régularisation	48
1.7.3	Régularisation pour l'identification à partir de données bruitées	49
2	Identification de sources non linéaires	51
2.1	Le problème direct	53
2.1.1	Formulation faible	53
2.1.2	"Equivalence" de la formulation faible et du problème direct	54
2.1.3	Linéarisation du problème	55
2.1.4	Méthodes numériques de résolution	61
2.1.5	Résolution sur un domaine quelconque de \mathbb{R}^2	64
2.2	Le problème de contrôle	70
2.2.1	Les paramètres de contrôle	71
2.2.2	Les équations d'état	72
2.2.3	La fonction coût	72
2.2.4	Formulation du problème de contrôle	73
2.3	Introduction du lagrangien. Système d'optimalité	73
2.3.1	Définition du lagrangien	73
2.3.2	Conditions nécessaires d'optimalité	74
2.3.3	Interprétation de l'état adjoint en termes d'EDP	76
2.4	Méthodes numériques de résolution	77
2.4.1	Itérations externes de l'algorithme séquentiel quadratique	77
2.4.2	Itérations internes de gradient conjugué	79
2.4.3	Structure et convergence de l'algorithme	80
2.4.4	Validation du gradient	81
2.4.5	Résultats numériques	82

Chapitre III

	Choix du paramètre de régularisation	89
1	Cas d'un problème inverse linéaire	90
1.1	Choix du paramètre par la méthode du résidu	90
1.2	La méthode de C.L. Mallows	91
1.3	Choix du paramètre par validation croisée	91
1.4	Validation croisée généralisée	93
1.5	Estimation de la trace de la matrice de lissage	93
2	Cas d'un problème inverse non linéaire	95
2.1	Choix du paramètre par GCV	96
2.2	Expression analytique pour $(B'_\epsilon)_z$ dans le cas général	97
2.3	Minimisation des fonctions de validation croisée	99

3	Application à l'identification de la diffusion	100
3.1	Calcul direct des dérivées première et seconde	101
3.2	Calcul de $B'_\epsilon(z)$ par le hessien de la fonction-coût	104
3.3	Calcul de $\xi^t (\text{Hess}J(v))^{-1} \xi$ pour $\xi \in \mathbb{R}^m$	108
3.4	Résultats numériques	108
4	Application à l'identification de sources non linéaires	113
4.1	Calcul de $B'_\epsilon(z)$	113
4.2	Résultats numériques	123
4.2.1	Validation croisée ordinaire	123
4.2.2	Validation croisée généralisée	125

Chapitre IV

	Identification par décomposition sur une base d'ondelettes	129
1	Quelques généralités sur les ondelettes	130
1.1	Analyse multirésolution de $L^2(\mathbb{R})$	130
1.2	Bases orthonormées d'ondelettes de $L^2(\mathbb{R})$	131
1.3	Ondelettes à support compact	133
1.4	Quelques résultats d'approximation	135
1.5	L'algorithme de S. Mallat	136
1.6	Ondelettes sur l'intervalle	138
2	Régression par ondelettes	140
2.1	Régression linéaire par ondelettes	140
2.1.1	Règles de seuillage	141
2.1.2	Choix du seuil	141
2.2	Application aux problèmes inverses non linéaires	142
2.2.1	Calcul de ϕ_N et ψ_N sur une grille diadique	142
2.2.2	Calcul de la transformée en ondelettes discrète des polynômes	144
2.2.3	Mise en œuvre numérique d'une méthode de seuillage	145
3	Identification par changement de base	148
3.1	Changement de base dans l'espace des contrôles	148
3.2	Effet du changement de base sur le gradient	150
3.3	Effet du changement de bases sur le hessien	151
3.4	Changement de base pour l'identification de sources non linéaires	153
3.4.1	Quelques résultats avec les ondelettes dans V_3	153
3.4.2	Résultats pour les ondelettes dans V_0	162
3.4.3	Comparaison avec les splines	165
3.5	Changement de base pour la diffusion de la chaleur	172
3.6	Conclusion	177

Chapitre V

	Mesures de non linéarité	179
1	Introduction	180
2	Aspect géométrique	181

3	Aspect statistique	183
3.1	La méthode de vraisemblance	183
3.2	Test du rapport des vraisemblances maximales	183
3.3	Régions de confiance	184
4	Calcul des courbures	186
5	Quelques résultats sur les courbures	188
	Conclusion	191
	Annexes	195
	Annexe 1 : Coefficients des filtres d'ondelettes	197
	Annexe 2 : Règles de calcul des tableaux 3D	203
	Bibliographie	205

Introduction générale



De nombreux domaines de la physique, de la mécanique, de la biologie sont modélisés par des systèmes à paramètres répartis, régis par des équations aux dérivées partielles qui décrivent le comportement spatio-temporel des inconnues du modèle. Le projet IDOPT de l'INRIA Rhône-Alpes, au sein duquel s'est déroulée cette thèse, s'est fixé pour but d'étudier les deux types de problèmes qui se posent alors : l'identification des systèmes à paramètres répartis et l'optimisation de ces systèmes.

Certains paramètres ou fonctions intervenant dans ces modèles sont inconnus ou plutôt mal connus (coefficients de diffusion dans des équations paraboliques, sources non linéaires dans des équations elliptiques). Nous nous proposons d'identifier ces paramètres ou fonctions à partir d'observations expérimentales : ce sont des problèmes inverses, par opposition à la résolution des équations elles-mêmes qui constitue le problème direct. La résolution de ces problèmes est une aide précieuse pour le physicien qui, en général, possède un modèle de son système, mais avec une large incertitude sur ses paramètres. La résolution du problème inverse lui fournit donc une information primordiale.

Les dispositifs expérimentaux sont pilotés par un physicien qui dispose en général d'un certain nombre de fonctions de contrôle qui lui permettent d'optimiser et éventuellement de stabiliser le système.

Le lien entre les problèmes d'identification et d'optimisation réside dans le fait qu'il s'agit, dans les deux cas, de minimiser une fonctionnelle dépendant de la solution d'une équation aux dérivées partielles. En effet, les problèmes d'identification seront ici formulés comme la minimisation de l'écart quadratique entre les observations expérimentales et les quantités correspondantes calculées par la résolution du système d'équations : les variables de contrôle sont dans ce cas les paramètres ou les fonctions à identifier. La minimisation de telles fonctionnelles relève de la théorie du contrôle optimal des équations aux dérivées partielles, due à J.-L. LIONS, et qui sera largement utilisée tout au long de cette thèse.

Dans le premier chapitre, nous rappelons quelques généralités sur les techniques de contrôle optimal. Nous traitons ensuite la résolution d'un problème type dans le seul but d'exposer la méthodologie que nous allons utiliser pour résoudre les deux problèmes posés au chapitre II.

Les problèmes considérés par la suite étant mal posés au sens de Hadamard, nous rappelons le principe de la méthode de régularisation, visant à rendre le problème plus stable.

Le deuxième chapitre concerne la modélisation mathématique et la résolution de deux problèmes inverses : l'identification du coefficient de diffusion dans

l'équation de la chaleur, et l'identification de sources non linéaires dans une équation aux dérivées partielles elliptique non linéaire. Nous mettrons en avant la difficulté supplémentaire qui existe dans ce type de problèmes par rapport au cadre du chapitre I.

Nous conduisons les calculs jusqu'à l'obtention du système d'optimalité, que nous résolvons par une méthode de gradient conjugué.

Dans le troisième chapitre, nous abordons le problème du choix du paramètre de régularisation pour les deux problèmes inverses précédemment cités. Plusieurs méthodes de choix sont étudiées : méthode du résidu, principe de Mallows, validation croisée ordinaire et généralisée, et nous en donnons ensuite une comparaison. Nous abordons les problèmes de régularité et de dérivabilité du problème de sources non linéaires afin de pouvoir appliquer la validation croisée généralisée.

Le quatrième chapitre est consacré à l'identification par décomposition sur une base d'ondelettes. Nous rappelons la théorie des ondelettes à support compact de I. Daubechies, puis nous montrons comment, à l'aide de telles fonctions, nous pouvons résoudre un problème inverse mal posé sans terme de régularisation dans la fonctionnelle à minimiser. Nous présentons ensuite des résultats numériques.

Nous introduisons des techniques de mesure de la non linéarité. En nous inspirant de techniques de régression non linéaire, nous définissons une surface de réponses associée au modèle, qui a la propriété de variété différentiable. Dans ce contexte, en calculant les courbures de cette surface au voisinage du point identifié, nous avons une information sur le degré de non linéarité du problème, et donc sur la qualité de l'estimateur obtenu.

Chapitre I

Le problème général

De tous les outils dont peuvent disposer les problèmes inverses régis par des équations aux dérivées partielles, la théorie du contrôle optimal développée par J.L. Lions, dont nous donnons ici quelques résultats fondamentaux est certainement l'un des plus importants. L'introduction d'un état adjoint permet un calcul exact du gradient de la fonctionnelle à minimiser, et le système d'optimalité est constructif en ce sens qu'il débouche sur un algorithme naturel d'obtention du point-selle.

Nous introduisons de façon très générale le problème d'optimisation ; nous donnons ensuite une description des méthodes de contrôle optimal, puis nous introduisons ensuite le terme de régularisation afin de rendre le problème stable.

1 Le problème de contrôle

Le problème que nous nous proposons de résoudre dans les chapitres suivants peut se formuler dans toute sa généralité de la façon suivante :

A partir d'observations $z \in E$ (espace des observations), peut-on identifier le contrôle $u \in \mathcal{U}_{\text{ad}}$ (espace des contrôles admissibles) avec $Tu = z$ où T est une application non linéaire de \mathcal{U}_{ad} dans E ?

L'application non linéaire T traduit les équations d'état du problème physique et le lien de la variable de contrôle u à la variable d'état notée par la suite y .

Ce problème inverse non linéaire sera résolu de façon classique par minimisation d'une certaine fonctionnelle J traduisant la fidélité de la solution cherchée u aux observations z :

$$\begin{aligned} J : \mathcal{U}_{\text{ad}} &\rightarrow \mathbb{R} \\ v &\mapsto J(v). \end{aligned}$$

Notre problème d'identification s'écrira donc :

Trouver $u \in \mathcal{U}_{\text{ad}}$ qui vérifie :

(Pb1)

$$J(u) = \min_{v \in \mathcal{U}_{\text{ad}}} J(v).$$

2 Quelques notions d'optimisation

Le problème (Pb1) précédent entre dans le cadre bien connu des problèmes d'optimisation. Nous allons ici en donner rapidement quelques notions de base, en citant les théorèmes d'existence, d'unicité et de caractérisation, sans toutefois les démontrer. Le lecteur pourra, pour plus de précisions, se reporter à [CEA 71] et [FAU 86].

Considérons pour cela :

V : un hilbert sur \mathbb{R} ,

K : un convexe non vide de V ,

J : une fonctionnelle définie sur K :

$$J : K \rightarrow \mathbb{R}$$

$$v \mapsto J(v).$$

Là encore, on cherche $u \in K$ qui vérifie :

$$J(u) = \min_{v \in K} J(v).$$

2.1 Propriétés liées à la convexité

La notion de convexité est essentielle en optimisation, comme nous le montre la propriété suivante :

Propriété 2.1 *Si la fonctionnelle J est convexe, alors tout minimum local est global. Si de plus, J est strictement convexe, alors le minimum, s'il existe, est unique.*

La plupart des algorithmes classiques de descente qui résolvent des problèmes d'optimisation sont des processus itératifs qui nécessitent la donnée d'une estimation initiale. Si la fonctionnelle n'est pas convexe, il y a un risque de convergence vers un minimum local et non global.

2.2 Existence d'un optimum

Le premier théorème d'existence est le suivant :

Théorème 2.2 *Si la fonctionnelle J est convexe continue, si de plus $\lim_{\|v\| \rightarrow \infty} J(v) = +\infty$ ou bien si K est un convexe fermé borné non vide, alors J admet un minimum.*

La condition de continuité peut en fait être remplacée par une condition plus faible de continuité semi-inférieure :

Théorème 2.3 *Si J est α -convexe et semi continue inférieure, alors le problème (Pb1) admet une solution unique.*

2.3 Caractérisation des optima

Avant d'aborder les caractéristiques des optima, nous allons rappeler quelques notions de calcul différentiel. Il existe plusieurs sortes de dérivées d'une fonctionnelle. On distingue tout d'abord la *dérivée directionnelle* de J au point v dans la direction w :

$$J'(v, w) = \lim_{\theta \rightarrow 0^+} \frac{J(v + \theta w) - J(v)}{\theta}.$$

Si de plus, l'application $w \rightarrow J'(v, w)$ est linéaire continue, alors J est *Gâteaux-dérivable* et sa Gâteaux-dérivée $J'(v)$ est donnée par $(J'(v), w) = J'(v, w)$ où (\cdot, \cdot) correspond au produit scalaire dans l'espace dual.

D'autre part, on dit que J est *Fréchet-dérivable* en v si l'on peut écrire :

$$J(v + w) = J(v) + (J'(v), w) + o(w) \text{ avec } \lim_{w \rightarrow 0} \frac{o(w)}{\|w\|} = 0.$$

La forme linéaire $J'(v)$ est alors appelée Fréchet-dérivée de J au point v .

Nous rappelons le lien entre Fréchet et Gâteaux-dérivées :

Propriété 2.4 *Si la fonctionnelle J est Fréchet-dérivable, alors J est Gâteaux-dérivable et les dérivées sont identiques. Si de plus, la Gâteaux-dérivée $J'(v)$ est continue par rapport à v , alors la réciproque est vraie.*

Dans la suite, $J'(v)$ désignera toujours la Gâteaux-dérivée de J . Nous citons à présent le théorème de caractérisation de l'optimum :

Théorème 2.5 *Soit $J : K \rightarrow \mathbb{R}$ une fonctionnelle convexe et Gâteaux-dérivable. Pour que $u \in K$ réalise le minimum de J sur K , il faut et il suffit que $(J'(u), v - u) \geq 0$ pour tout $v \in K$ (inéquation d'Euler). Dans le cas d'une minimisation sans contraintes ($K = V$) l'inéquation devient $J'(u) \equiv 0$ (équation d'Euler).*

Lorsque la fonctionnelle J n'est pas convexe, les inéquation et équation d'Euler sont des conditions nécessaires à l'existence de l'optimum.

2.4 Algorithmes de résolution

La plupart des algorithmes classiques d'optimisation ont de commun qu'ils nécessitent, au cours de leur déroulement, la connaissance a priori d'une direction de descente. Il s'agit de trouver, à partir d'une estimation initiale de la solution v_0 , une direction d_0 dans laquelle, au voisinage de v_0 , le critère J décroît. De par sa définition, la direction opposée au gradient de J au point v_0 : $\nabla J(v_0)$, est un bon candidat. L'intérêt de cette direction est de remplacer un problème d'optimisation de dimension p (généralement grand) par une suite de problèmes de minimisation unidimensionnels.

Nous présentons rapidement la méthode de gradient conjugué que nous utiliserons par la suite :

supposons tout d'abord que la fonctionnelle J à minimiser soit quadratique elliptique :

$$J(v) = \frac{1}{2}(Av, v) - (b, v),$$

A étant une matrice symétrique définie positive $p \times p$, et b une forme linéaire sur \mathbb{R}^p . Partant d'un vecteur initial v_0 arbitraire, posons $d_0 = \nabla J(v_0)$.

Si $\nabla J(v_0) = 0$, l'algorithme est terminé.

Sinon on définit : $r_0 = \frac{(\nabla J(v_0), d_0)}{(Ad_0, d_0)}$

puis $v_1 = v_0 - r_0 d_0$.

Supposant construits de proche en proche les vecteurs $v_1, d_1, \dots, v_k, d_k$, ce qui sous-entend que les gradients $\nabla J(v_\ell)$, $0 \leq \ell \leq k-1$ sont tous non nuls, deux cas se présentent :

ou bien $\nabla J(v_k) = 0$ et l'algorithme est terminé,

ou bien $\nabla J(v_k) \neq 0$, auquel cas on définit :

$$d_k = \nabla J(v_k) + \frac{\|\nabla J(v_k)\|^2}{\|\nabla J(v_{k-1})\|^2} d_{k-1}$$

puis

$$r_k = \frac{(\nabla J(v_k), d_k)}{(Ad_k, d_k)}$$

et $v_{k+1} = v_k - r_k d_k$.

On vérifie aisément qu'un tel choix des nombres r_k entraîne que :

$$J(v_k - r_k d_k) = \inf_{r \in \mathbb{R}} J(v_k - r d_k).$$

D'autre part, les vecteurs $\nabla J(v_\ell)$, $0 \leq \ell \leq k$ sont linéairement indépendants et deux à deux orthogonaux, donc la méthode de gradient conjugué appliquée à une fonctionnelle quadratique elliptique converge en p itérations au plus.

Nous signalons qu'il existe une adaptation de cette méthode qui s'applique aux fonctionnelles non quadratiques : la méthode de gradient conjugué de Polak-Ribière [PR 69]. Dans ce cas de figure, les gradients successifs n'ont aucune raison d'être encore deux à deux orthogonaux, et donc on n'est pas sûr que la méthode converge en un nombre fini d'itérations.

3 Le calcul du gradient

Nous avons vu dans les paragraphes précédents le rôle prépondérant que joue le gradient de la fonctionnelle à minimiser dans la résolution du problème d'optimisation. Nous nous intéressons ici au moyen de l'évaluer.

3.1 Calcul approché

Une première façon de procéder est de calculer ce gradient de manière approchée par différences divisées. Il faut pour cela discrétiser le contrôle v en un certain nombre de paramètres v_i ; la i -ème composante du gradient est approchée par :

$$(\overline{\nabla J})_i = \frac{J(v_1, \dots, v_i + \delta, \dots, v_p) - J(v_1, \dots, v_i, \dots, v_p)}{\delta}$$

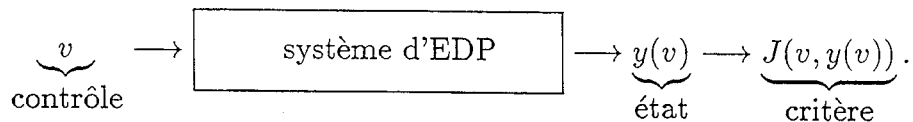
pour δ arbitrairement petit.

Cette méthode, tout en étant relativement efficace, possède de sérieux handicaps. Tout d'abord, elle nécessite d'évaluer autant de fois la fonctionnelle à minimiser qu'il y a de paramètres, soit p . D'autre part, la valeur de δ doit être choisie assez petite pour que le calcul soit une bonne approximation, mais pas trop petite afin d'éviter que la différence finie ne devienne nulle pour l'ordinateur.

On voit donc l'intérêt d'un calcul exact du gradient, qui lui sera indépendant du nombre de paramètres p .

3.2 Calcul exact

Si nous voulons calculer de façon exacte le gradient du critère J , il nous faut utiliser le caractère particulier de notre problème d'optimisation. En effet, la dépendance du contrôle et de la fonctionnelle J n'est pas directe : elle fait intervenir l'état du système qui est fonction du contrôle par l'intermédiaire d'un système d'équations aux dérivées partielles. La situation peut se résumer ainsi :



La fonctionnelle J a donc une double dépendance par rapport au contrôle : dépendance directe et dépendance indirecte par l'intermédiaire de $y(v)$. Elle s'écrit donc sous la forme définitive suivante :

$$(3.1) \quad \begin{aligned} J &: \mathcal{U}_{\text{ad}} \rightarrow \mathbb{R} \\ v &\mapsto J(v, y(v)), \end{aligned}$$

où $y(v)$ est solution des équations d'état du problème.

Le problème d'optimisation (Pb1) s'écrit maintenant comme un problème de contrôle que l'on peut formuler ainsi :

$$(3.2) \quad \text{Trouver } u \in \mathcal{U}_{\text{ad}} \text{ tel que } J(u, y(u)) = \inf_{v \in \mathcal{U}_{\text{ad}}} J(v, y(v)).$$

Si un tel u existe, on l'appellera *contrôle optimal* dans le sens où l'a défini J.L. Lions [LIO 68].

La technique de résolution de ces problèmes, que nous allons essayer d'expliquer dans le paragraphe suivant, aura pour but de caractériser le contrôle optimal. Nous en déduirons alors l'expression exacte du gradient du critère.

4 Introduction au contrôle optimal des EDP

Nous allons nous intéresser dans ce paragraphe à la théorie du contrôle optimal appliqué aux équations aux dérivées partielles (EDP), telle qu'elle a été développée par J.L. Lions. Si nous avons choisi de reprendre le cas des équations elliptiques et paraboliques, c'est tout d'abord pour fixer une fois pour toutes les notations et ensuite parce que nous suivrons dans les chapitres II et III une méthodologie identique à celle que nous allons exposer ci-dessous.

Ce paragraphe est par conséquent destiné aux non initiés qui pourront pour plus de précisions se référer à [LIO 68].

4.1 Aperçu général

Soit Ω un ouvert borné de \mathbb{R}^n , de frontière Γ . Considérons dans cet ouvert un système physique régi par une équation aux dérivées partielles.

Terminologie

On désignera par :

- *état du système* ou *variable d'état*: $y(x) \in V$, $x \in \Omega$, la solution du problème direct,
- *contrôle*: la fonction $v(x) \in \mathcal{U}_{\text{ad}} = \{ \text{contrôles admissibles} \}$ qui sera distribué: $v(x)$, $x \in \Omega$,
ou frontière: $v(x)$, $x \in \Gamma$,
- *équation d'état* ou *système direct*: l'EDP qui lie la variable d'état au contrôle $v \mapsto y(v)$,
- *fonction-coût* ou *critère*: la donnée d'une fonctionnelle sur l'ensemble des contrôles: $J : v \mapsto J(v) = J(v, y(v))$ à valeurs dans \mathbb{R} .
- *contrôle optimal*: le contrôle u vérifiant: $J(u) = \inf_{v \in \mathcal{U}_{\text{ad}}} J(v)$.

4.2 Contrôle optimal des EDP elliptiques

Nous allons dans ce paragraphe expliquer la logique du contrôle optimal sur le cas d'école ci-dessous.

4.2.1 Le problème direct

Considérons l'équation d'état suivante :

$$(4.1) \quad \begin{cases} -\Delta y(x) + y(x) = f(x) & \text{dans } \Omega \\ \frac{\partial y}{\partial n}(x) = v(x) & \text{sur } \Gamma \end{cases}$$

avec :

$f(x)$: fonction donnée, $\in L^2(\Omega)$,

$y(x)$: variable d'état, $\in H^1(\Omega)$,

$v(x)$: contrôle frontière $\in \mathcal{U}_{ad} =$ convexe fermé de $L^2(\Gamma)$.

L'équation d'état (4.1) peut se mettre sous forme variationnelle. Pour cela, on considère une fonction test $\phi \in H^1(\Omega)$. En multipliant (4.1) par ϕ , puis en intégrant par parties sur Ω , on obtient :

$$\int_{\Omega} \nabla y(x) \cdot \nabla \phi(x) dx + \int_{\Omega} y(x) \phi(x) dx = \int_{\Omega} f(x) \phi(x) dx + \int_{\Gamma} v(x) \phi(x) d\sigma,$$

y est donc solution faible de la formulation variationnelle suivante :

$$(4.2) \quad \begin{cases} a(y, \phi) = (f, \phi) + \int_{\Gamma} v \phi d\sigma, & \forall \phi \in H^1(\Omega) \\ \text{avec } a(y, \phi) = \int_{\Omega} \nabla y \cdot \nabla \phi + \int_{\Omega} y \phi \\ \text{et } (f, \phi) = \int_{\Omega} f \phi. \end{cases}$$

Nous pouvons à présent rappeler le théorème d'existence et d'unicité du système (4.2) :

Théorème 4.1 *Pour tout $u \in \mathcal{U}_{ad}$, l'équation variationnelle (4.2) admet une unique solution $y(x) \in H^1(\Omega)$.*

Lemme 4.2 *L'application $v \mapsto y(v)$ est affine et continue de $L^2(\Gamma)$ vers $H^1(\Omega)$.*

4.2.2 Le problème de contrôle

On se donne la fonction coût suivante :

$$(4.3) \quad J(v, y(v)) = \frac{1}{2} \int_{\Gamma} |y(v) - z|^2 d\sigma + \frac{N}{2} \int_{\Gamma} |v|^2 d\sigma$$

où $z \in L^2(\Gamma)$ est une fonction donnée.

Le problème de contrôle sera alors le suivant :

$$(4.4) \quad \boxed{\text{Trouver } u \in \mathcal{U}_{\text{ad}} \text{ tel que } J(u, y(u)) = \inf_{v \in \mathcal{U}_{\text{ad}}} J(v, y(v)).}$$

4.2.3 Interprétation physique

Nous sommes dans le cas d'une observation frontière. Le problème de contrôle (4.4) va nous permettre de trouver une fonction u régulière au sens de la norme L^2 de telle sorte que la trace de la fonction $y(u)$, solution du système (4.2) coïncide sur le bord Γ avec la fonction d'observation z que l'on s'est donnée.

La minimisation du premier terme de la fonctionnelle J nous assure une distance minimale entre $y(u)$ et z (au sens L^2) sur le bord, tandis que la minimisation du second terme nous donne la régularité du contrôle.

4.2.4 Existence et caractérisation de l'optimum

Il est facile de montrer que la fonctionnelle J est continue et strictement convexe. De plus, \mathcal{U}_{ad} est un convexe fermé que l'on supposera non vide. Pour démontrer l'existence et l'unicité de l'optimum, il ne reste plus qu'à montrer que $\lim_{\|v\| \rightarrow \infty} J(v) = \infty$, ce qui est immédiat. Comme nous l'avons vu avec le théorème 2.5, le contrôle optimal est donné par l'inéquation d'Euler :

$$\begin{cases} (J'(u), v - u) \geq 0, & \forall v \in \mathcal{U}_{\text{ad}} \\ J'(u) = 0 & \text{dans le cas sans contrainte.} \end{cases}$$

La suite du paragraphe va concerner le calcul explicite de cette dérivée $J'(u)$ pour notre exemple. Commençons par calculer $J(u + \theta(v - u))$:

$$(4.5) \quad J(u + \theta(v - u)) = \frac{1}{2} \int_{\Gamma} |y(u + \theta(v - u)) - z|^2 d\sigma + \frac{N}{2} \int_{\Gamma} |u + \theta(v - u)|^2 d\sigma.$$

Or, l'application $v \rightarrow y(v)$ est affine, d'où $y(v) = Av + b$ et donc $y(u + \theta(v - u)) = Au + \theta A(v - u) + b$. En reportant cette expression dans (4.5), on obtient :

$$\begin{aligned} J(u + \theta(v - u)) &= J(u) + \theta \left(\int_{\Gamma} (y(v) - y(u))(y(u) - z) d\sigma + N \int_{\Gamma} (v - u, u) d\sigma \right) \\ &\quad + \frac{\theta^2}{2} \left(\int_{\Gamma} |y(v) - y(u)|^2 + N \int_{\Gamma} |v - u|^2 d\sigma \right). \end{aligned}$$

On a utilisé le fait que $Av - Au = y(v) - y(u)$. De ce calcul, nous déduisons le lemme suivant :

$$\text{Lemme 4.3} \quad J'(u, v - u) = \int_{\Gamma} \left((y(v) - y(u))(y(u) - z) + N(v - u)u \right) d\sigma.$$

Le second terme de cette intégrale, à savoir $N \int_{\Gamma} (v - u)u d\sigma$ se met trivialement sous la forme $(j_1(u), v - u)_{L^2(\Gamma)}$ (nous rappelons que le dual de $L^2(\Gamma)$ est lui-même).

Le but de la manipulation qui va suivre est d'écrire le terme restant, à savoir $\int_{\Gamma} (y(v) - y(u))(y(u) - z) d\sigma$, sous une forme équivalente : $(j_2(u), v - u)_{L^2(\Gamma)}$. Dans ces conditions, la Gâteaux-dérivée cherchée sera $J'(u) = j_1(u) + j_2(u)$. Pour cela, nous introduisons, de façon un peu artificielle, la variable adjointe p , solution faible de l'équation ci-dessous :

$$(4.6) \quad a(p, \phi) = \int_{\Gamma} (y(u) - z)\phi d\sigma, \quad \forall \phi \in H^1(\Omega).$$

Cette nouvelle équation, appelée "système adjoint", nous conduit au résultat suivant :

$$\text{Lemme 4.4} \quad j_2(u) = p. \text{ De plus, la Gâteaux-dérivée de } J \text{ au point } u \text{ s'écrit : } J'(u) = p + Nu.$$

Preuve : La démonstration de ce lemme se fait de la façon suivante : on s'intéresse à la formulation variationnelle (4.6) que l'on écrit sous la forme :

$$a(\phi, p) = \int_{\Gamma} (y(u) - z)\phi d\sigma \quad \text{car } a \text{ est symétrique.}$$

Ceci étant vrai pour toute fonction test ϕ , nous prenons successivement $\phi = y(v)$ et $\phi = y(u)$ puis, par différence, on obtient :

$$(4.7) \quad a(y(v) - y(u), p) = \int_{\Gamma} (y(u) - z)(y(v) - y(u)) d\sigma.$$

D'autre part, on s'intéresse à la formulation variationnelle du problème direct (4.2) ; en prenant comme fonction test $\phi = p$ dans le problème avec v puis avec u , et par différence, on obtient :

$$(4.8) \quad a(y(v) - y(u), p) = \int_{\Gamma} (v - u) p d\sigma.$$

En égalant les expressions (4.7) et (4.8) on obtient l'expression de $j_2(u)$ cherchée :

$$(j_2(u), v - u)_{L^2(\Gamma)} = \int_{\Gamma} (y(u) - z)(y(v) - y(u)) d\sigma = \int_{\Gamma} (v - u) p d\sigma,$$

d'où $j_2(u) = p$ et donc $J'(u) = j_1(u) + j_2(u) = p + Nu$. □

Nous avons ainsi établi une caractérisation du contrôle optimal que nous pouvons résumer ainsi :

$$(4.9) \quad \begin{cases} (p + Nu, v - u) \geq 0, & \forall v \in \mathcal{U}_{ad} \\ \text{avec} & a(\psi, p) = \int_{\Gamma} (y(u) - z)\psi d\sigma, & \forall \psi \in H^1(\Omega) \\ \text{et} & a(y(u), \phi) = (f, \phi)_{L^2(\Gamma)} + \int_{\Gamma} u\phi d\sigma, & \forall \phi \in H^1(\Omega). \end{cases}$$

Cette technique de calcul a l'inconvénient d'utiliser la connaissance a priori d'une fonction p solution du système adjoint. Nous allons voir dans le paragraphe suivant une approche un peu différente qui fait apparaître la variable adjointe de façon tout à fait naturelle.

4.2.5 Approche lagrangienne

Nous allons remplacer le problème initial (3.2) par un problème équivalent que nous noterons (Pb2). Il s'agit en fait de supposer, dans un premier temps, que la variable d'état y est indépendante du contrôle v . Dans un deuxième temps,

cette dépendance sera prise en compte dans une contrainte sur le contrôle v . Le nouveau problème se formule ainsi :

$$(Pb2) \quad \left\{ \begin{array}{l} \text{Trouver } (u, \tilde{y}) \in \mathcal{U}_{ad} \times H^1(\Omega) \text{ tel que :} \\ J(u, \tilde{y}) = \inf_{(v, y) \in \mathcal{U}_{ad} \times H^1(\Omega)} J(v, y), \\ \text{sous la contrainte } y = y(v). \end{array} \right.$$

Nous sommes en présence d'un problème d'optimisation à deux paramètres avec contrainte. Nous allons le résoudre en utilisant un multiplicateur de Lagrange q et le lagrangien correspondant :

$$(4.10) \quad \mathcal{L}(v, y, q) = J(v, y) - a(y, q) + (f, q)_{L^2(\Gamma)} + \int_{\Gamma} v q d\sigma.$$

Définition 4.5 *Le point (\tilde{y}, u, p) est point-selle du lagrangien \mathcal{L} si et seulement si $\mathcal{L}(\tilde{y}, u, q) \leq \mathcal{L}(\tilde{y}, u, p) \leq \mathcal{L}(y, v, p)$ pour tout $(y, v, q) \in H^1(\Omega) \times \mathcal{U}_{ad} \times H^1(\Omega)$.*

Théorème 4.6 *Si (\tilde{y}, u, p) est point-selle du lagrangien \mathcal{L} alors on a :*
i) $\tilde{y} = y(u)$, ce qui rétablit la dépendance au contrôle,
ii) u est solution du problème d'optimisation (Pb2).

Preuve : Nous écrivons tout d'abord que (\tilde{y}, u, p) est point-selle de \mathcal{L} :

$$\mathcal{L}(\tilde{y}, u, q) \leq \mathcal{L}(\tilde{y}, u, p) \leq \mathcal{L}(y, v, p), \quad \forall (y, v, q) \in H^1(\Omega) \times \mathcal{U}_{ad} \times H^1(\Omega).$$

Intéressons-nous à l'inégalité de gauche :

$$\mathcal{L}(\tilde{y}, u, q) \leq \mathcal{L}(\tilde{y}, u, p), \quad \forall q \in H^1(\Omega).$$

En remplaçant \mathcal{L} par son expression :

$$J(u, \tilde{y}) - a(\tilde{y}, q) + (f, q) + \int_{\Gamma} u q d\sigma \leq J(u, \tilde{y}) - a(\tilde{y}, p) + (f, p) + \int_{\Gamma} u p d\sigma,$$

d'où :

$$a(\tilde{y}, p - q) + (f, q - p) + \int_{\Gamma} u(q - p) d\sigma \leq 0, \quad \forall q \in H^1(\Omega).$$

Posons alors $\tilde{q} = p - q$:

$$a(\tilde{y}, \tilde{q}) - (f, \tilde{q}) - \int_{\Gamma} u \tilde{q} d\sigma \leq 0, \quad \forall \tilde{q} \in H^1(\Omega).$$

Puis en posant $\tilde{q} = q - p$, nous obtenons l'inégalité inverse, d'où finalement :

$$a(\tilde{y}, \tilde{q}) = (f, \tilde{q}) + \int_{\Gamma} u \tilde{q} d\sigma, \quad \forall \tilde{q} \in H^1(\Omega),$$

ce qui prouve l'assertion i). Considérons ensuite l'inégalité de droite :

$$\mathcal{L}(\tilde{y}, u, p) \leq \mathcal{L}(y, v, p), \quad \forall (y, v) \in H^1(\Omega) \times \mathcal{U}_{\text{ad}}.$$

En développant :

$$J(u, \tilde{y}) - a(\tilde{y}, p) + (f, p) + \int_{\Gamma} u p d\sigma \leq J(v, y) - a(y, p) + (f, p) + \int_{\Gamma} v p d\sigma.$$

D'après i), nous savons que $a(\tilde{y}, p) - (f, p) - \int_{\Gamma} u p d\sigma = 0$. D'où il reste :

$$J(u, \tilde{y}) \leq J(v, y) - a(y, p) + (f, p) + \int_{\Gamma} v p d\sigma, \quad \forall (y, v) \in H^1(\Omega) \times \mathcal{U}_{\text{ad}},$$

ce que l'on peut encore écrire :

$$J(u, \tilde{y}) \leq J(v, y), \quad \forall (y, v) \in H^1(\Omega) \times \mathcal{U}_{\text{ad}} \text{ vérifiant } a(y, p) - (f, p) - \int_{\Gamma} v p d\sigma = 0.$$

Ou, ce qui est encore équivalent :

$$J(u, \tilde{y}) \leq J(v, y), \quad \forall (y, v) \in H^1(\Omega) \times \mathcal{U}_{\text{ad}} \text{ vérifiant } y = y(v).$$

□

On obtient ainsi une caractérisation du point-selle et par conséquent du contrôle optimal u , qui peut s'écrire :

$$\begin{cases} \text{i)} & (\tilde{y}, u) \text{ vérifie le système direct (4.1) ,} \\ \text{ii)} & \mathcal{L}(\tilde{y}, u, p) = \inf_{(y, v) \in H^1 \times \mathcal{U}_{\text{ad}}} \mathcal{L}(y, v, p). \end{cases}$$

La condition ii) s'écrit aussi en termes de conditions d'optimalité d'Euler-Lagrange :

$$\begin{cases} \frac{\partial}{\partial y} \mathcal{L}(y, v, q)|_{(\tilde{y}, u, p)} = 0, \quad \forall v \in \mathcal{U}_{\text{ad}} \\ \left(\frac{\partial}{\partial v} \mathcal{L}(y, v, q)|_{(\tilde{y}, u, p)}, v - u \right) \geq 0, \quad \forall v \in \mathcal{U}_{\text{ad}}. \end{cases}$$

Calculons les deux dérivées en question :

$$\left(\frac{\partial}{\partial y} \mathcal{L}(y, v, q)|_{(\tilde{y}, u, p)}, \psi \right) = \int_{\Gamma} (\tilde{y} - z) \psi d\sigma - a(\psi, p), \quad \forall \psi \in H^1(\Omega).$$

Donc, le système adjoint est donné par :

$$a(p, \psi) = \int_{\Gamma} (\tilde{y} - z) \psi d\sigma, \quad \forall \psi \in H^1(\Omega).$$

D'autre part,

$$\left(\frac{\partial}{\partial v} \mathcal{L}(y, v, q) |_{(\tilde{y}, u, p)}, v - u \right) = N \int_{\Gamma} u(v - u) d\sigma + \int_{\Gamma} p(v - u) d\sigma,$$

d'où :

$$\int_{\Gamma} (p + Nu)(v - u) d\sigma \geq 0, \quad \forall v \in \mathcal{U}_{\text{ad}}.$$

Ceci nous donne l'expression du gradient de la fonctionnelle J au point u :

Lemme 4.7

$$\left(\frac{\partial}{\partial v} \mathcal{L}(y, v, q) |_{(\tilde{y}, u, p)}, v - u \right) = (J'(u), v - u).$$

La caractérisation du point-selle s'écrit maintenant sous la forme d'un système d'optimalité vérifié par le contrôle u , la variable d'état y et l'état adjoint p :

$$(4.11) \quad \begin{cases} a(\tilde{y}, \phi) = (f, \phi) + \int_{\Gamma} u \phi d\sigma, & \forall \phi \in H^1(\Omega) \\ a(p, \psi) = \int_{\Gamma} (\tilde{y} - z) \psi d\sigma, & \forall \psi \in H^1(\Omega) \\ \int_{\Gamma} (p + Nu)(v - u) d\sigma \geq 0, & \forall v \in \mathcal{U}_{\text{ad}} \\ p + Nu = 0 & \text{dans le cas sans contrainte.} \end{cases}$$

4.3 Contrôle optimal des EDP paraboliques

Comme pour les équations elliptiques, nous allons rappeler, dans ce paragraphe, les principaux résultats d'existence et d'unicité dans le cas d'équations aux dérivées partielles paraboliques.

On se donne les espaces de Hilbert V et H avec $V \subset H \subset V'$ où V' est le dual de V ; on identifiera H avec son dual, et on suppose que V est dense dans H avec injection continue. On notera $\| \cdot \|$ la norme sur V et $|\cdot|$ la norme sur H . La variable t désigne le temps : on suppose que $t \in]0, T[$ avec $T < \infty$. On se donne une famille de formes bilinéaires continues sur V :

$$\begin{aligned} (\phi, \psi) &\mapsto a(t; \phi, \psi) \text{ pour tout } t \in]0, T[\\ V \times V &\rightarrow \mathbb{R}. \end{aligned}$$

Relativement à cette famille, on suppose :

$$(4.12) \quad \begin{cases} \forall (\phi, \psi) \in V \times V, \text{ la fonction } t \mapsto a(t; \phi, \psi) \text{ est mesurable sur }]0, T[\\ \text{et } |a(t; \phi, \psi)| \leq C \|\phi\| \times \|\psi\|, \end{cases}$$

et

$$(4.13) \quad \begin{cases} \text{il existe } \lambda \text{ tel que} \\ a(t; \phi, \phi) + \lambda \|\phi\|^2 \geq \alpha \|\phi\|^2, \quad \alpha > 0, \forall \phi \in V, t \in]0, T[. \end{cases}$$

Pour chaque t on peut écrire : $a(t; \phi, \psi) = (A(t)\phi, \psi)_{V', V}$, où $A(t)\phi \in V'$. Nous introduisons maintenant un nouvel espace de fonctions :

Définition 4.8 On notera $L^2(0, T; V)$ l'espace des fonctions $t \mapsto f(t)$ de $]0, T[$ dans V , mesurables et telles que $\int_0^T \|f(t)\|^2 dt < \infty$.

On définit de même $L^2(0, T; V')$ et on note :

$$A(t) \in \mathcal{L}(L^2(0, T; V), L^2(0, T; V')),$$

c'est à dire que si $f \in L^2(0, T; V)$, $A(t)f$ est la fonction $t \mapsto A(t)f(t) \in V'$ dont on vérifie qu'elle est mesurable et continue (d'après (4.12)) :

$$\|A(t)f(t)\|_{V'} \leq C \|f(t)\|.$$

On définit ensuite l'espace des distributions :

Définition 4.9 On notera par $\mathcal{D}'(]0, T[; V)$ l'espace des distributions sur $]0, T[$ à valeurs dans V défini par : $\mathcal{D}'(]0, T[; V) = \mathcal{L}(\mathcal{D}(]0, T[); V)$.

Donc, si $f \in \mathcal{D}'(]0, T[; V)$: pour tout $\phi \in \mathcal{D}(]0, T[)$, $f(\phi) \in V$ et $\phi \mapsto f(\phi)$ est continue de $\mathcal{D}(]0, T[)$ dans V . On écrira :

$$(f, \phi) = \int_0^T f(t)\phi(t)dt, \quad \text{intégrale généralisée à valeurs dans } V.$$

On définit à présent la dérivée distribution d'une fonction f de $L^2(0, T; V)$:

Définition 4.10 Soit $f \in L^2(0, T; V)$, on appelle dérivée distribution de f , que l'on note $\frac{df}{dt}$, l'application linéaire continue de $\mathcal{D}(]0, T[)$ sur V définie par : $(\frac{df}{dt}, \phi) = -(f, \frac{d\phi}{dt})$, $\forall \phi \in \mathcal{D}(]0, T[)$ (donc en particulier $\frac{df}{dt} \in \mathcal{D}'(]0, T[; V)$).

Nous pouvons maintenant définir la convergence dans cet espace des distributions :

Définition 4.11 On dit que f_n converge vers f dans $\mathcal{D}'(]0, T[; V)$ et on note : $f_n \rightarrow f$ si $(f_n, \phi) \rightarrow (f, \phi)$ pour tout $\phi \in \mathcal{D}(]0, T[)$.

On introduit un nouvel espace fonctionnel qui nous sera utile par la suite :

Définition 4.12 $W(0, T) = \{f / f \in L^2(0, T; V), \frac{df}{dt} \in L^2(0, T; V')\}$.

On a alors :

Théorème 4.13 L'espace $W(0, T)$ muni de la norme : $\|f\|_{W(0, T)} = \left(\int_0^T \|f(t)\|^2 dt + \int_0^T \left\| \frac{df}{dt} \right\|_{V'}^2 dt \right)^{\frac{1}{2}}$ est un espace de Hilbert. De plus, toute fonction $f \in W(0, T)$ est continue de $[0, T]$ dans H , soit encore : $W(0, T) \subset C^0([0, T]; H)$.

Nous ne donnons pas la démonstration de ce théorème, mais le lecteur pourra la trouver dans [LIO 68].

4.3.1 Le problème direct

Soient toujours Ω un ouvert borné de \mathbb{R}^n , Γ sa frontière, et soient $V = H^1(\Omega)$ et $H = L^2(\Omega)$. Nous nous intéressons au problème d'évolution suivant :

$$(4.14) \quad \begin{cases} \frac{dy}{dt} + A(t)y = f & \text{dans } \Omega, \forall t \in]0, T[\\ y(x, 0) = y_0(x) & \text{pour } x \in \Omega, \end{cases}$$

où $f \in L^2(0, T; V')$ et $y_0 \in H$ sont donnés. Nous avons le théorème suivant, qui nous assure l'existence et l'unicité d'une solution du problème (4.14) :

Théorème 4.14 Si les hypothèses (4.12) et (4.13) sont vérifiées, alors le problème direct (4.14) admet une solution unique y dans $W(0, T)$. De plus, cette solution y dépend continûment des données :

$$\begin{aligned} & \text{l'application } (f, y_0) \mapsto y \\ & L^2(0, T; V') \times H \rightarrow W(0, T) \text{ est continue.} \end{aligned}$$

Nous renvoyons le lecteur à [LIO 68] pour la démonstration de ce théorème.

4.3.2 Le problème de contrôle

On désigne par \mathcal{U} l'espace de Hilbert des contrôles et on se donne une application linéaire B continue de \mathcal{U} dans $L^2(0, T; V')$. Soient $f \in L^2(0, T; V')$ et $y_0 \in H$ deux fonctions données. Supposant que les hypothèses (4.12) et (4.13) ont lieu, on désigne par $y(v)$ la solution du système :

$$(4.15) \quad \begin{cases} \frac{dy(v)}{dt} + A(t)y(v) = f + Bv \text{ dans } \Omega, \forall t \in]0, T[, \\ y(v)|_{t=0} = y_0 \text{ dans } \Omega, \\ y(v) \in L^2(0, T; V), \end{cases}$$

où $y(v)$ est une fonction $t \mapsto y(v)(t)$ que l'on notera encore $y(x, t; v)$ pour $x \in \Omega$, $t \in]0, T[$. Il s'agit de l'état du système.

On se donne ensuite l'observation $z(v) = Cy(v)$ avec C une application linéaire continue de $L^2(0, T; V)$ dans $L^2(0, T; H)$ et $N \in \mathcal{L}(\mathcal{U}, \mathcal{U})$ coercive : $(Nu, u)_{\mathcal{U}} \geq \nu \|u\|_{\mathcal{U}}^2$ avec $\nu > 0$, et autoadjointe ($N^* = N$).

On considère alors la fonction coût suivante :

$$J(v, y(v)) = \frac{1}{2} \int_0^T |Cy(v) - z_d|^2 dt + \frac{1}{2} \int_0^T (Nv, v)_{\mathcal{U}} dt.$$

On se donne ensuite \mathcal{U}_{ad} un sous-ensemble convexe fermé de \mathcal{U} . Notre problème de contrôle est le suivant :

$$(4.16) \quad \boxed{\text{Trouver } u \in \mathcal{U}_{\text{ad}} \text{ tel que } J(u, y(u)) = \inf_{v \in \mathcal{U}_{\text{ad}}} J(v, y(v)).}$$

Nous énonçons un résultat donnant la continuité de l'état par rapport au contrôle, conséquence du théorème 4.14 :

Propriété 4.15 *L'application $v \mapsto y(v)$ est affine et continue de \mathcal{U} dans $L^2(0, T; V)$.*

La démonstration s'obtient de manière analogue à celle rencontrée dans le cas d'équations elliptiques.

4.3.3 Existence et caractérisation de l'optimum

Commençons tout d'abord par le théorème d'existence et d'unicité :

Théorème 4.16 *Le problème de contrôle (4.16) admet une unique solution $u \in \mathcal{U}_{ad}$ telle que $J(u, y(u)) = \inf_{v \in \mathcal{U}_{ad}} J(v, y(v))$.*

Nous donnons à présent la caractérisation du contrôle optimal en partant de l'inéquation d'Euler : $(J'(u), v - u) \geq 0, \forall v \in \mathcal{U}_{ad}$. Tout d'abord, en utilisant le fait que $y(v)$ est une fonction affine de v , on obtient :

$$(J'(u), v - u) = \int_0^T (Cy(u) - z_d, C(y(v) - y(u)))_H dt + \int_0^T (Nu, v - u)_U dt.$$

De même que précédemment, nous introduisons un état adjoint p solution du problème :

$$(4.17) \quad \begin{cases} -\frac{dp}{dt} + A^*(t)p = C^*(Cy(u) - z_d) \text{ dans } \Omega, \forall t \in]0, T[, \\ p(x, T) = 0 \text{ dans } \Omega. \end{cases}$$

On vérifie facilement que (4.17) est un problème bien posé : en posant $t' = T - t$ on a :

$$\begin{cases} \frac{dp}{dt'} + A^*(t)p = C^*(Cy(u) - z_d) \text{ dans } \Omega, \forall t' \in]0, T[, \\ p(x, 0) = 0 \text{ dans } \Omega, \end{cases}$$

qui est un système parabolique de même type que (4.14). D'autre part, $Cy(u) - z_d \in L^2(0, T; H)$ donc $C^*(Cy(u) - z_d) \in L^2(0, T; V')$. En appliquant le théorème 4.14, l'état adjoint p est bien défini de manière unique par (4.17). Par une méthode analogue à celle du paragraphe 4.2, on déduit :

Théorème 4.17 *J est Gâteaux-dérivable et $J'(u) = B^*p + Nu$.*

Le contrôle optimal u est donc caractérisé par le système d'optimalité :

$$(4.18) \quad \left\{ \begin{array}{l} \begin{cases} \frac{dy}{dt} + A(t)y = f + Bu \\ y(x, 0) = y_0(x) \text{ dans } \Omega \end{cases} \\ \begin{cases} -\frac{dp}{dt} + A^*(t)p = C^*(Cy - z_d) \\ p(x, T) = 0 \text{ dans } \Omega \end{cases} \\ \int_0^T (B^*p + Nu, v - u)_U dt \geq 0, \forall v \in \mathcal{U}_{ad} \\ \text{ou } B^*p + Nu = 0 \text{ dans le cas sans contrainte.} \end{array} \right.$$

4.3.4 Approche lagrangienne

Comme nous l'avons vu précédemment, nous remplaçons maintenant le problème de contrôle initial (4.16) par un problème équivalent : nous supposons dans un premier temps que l'état du système y est indépendant du contrôle v , et le problème devient le suivant :

$$(4.19) \quad \boxed{\begin{array}{l} \text{Trouver } (u, \tilde{y}) \in \mathcal{U}_{\text{ad}} \times L^2(0, T; V) \text{ tel que} \\ J(u, \tilde{y}) = \inf_{(v, y) \in \mathcal{U}_{\text{ad}} \times L^2(0, T; V)} J(v, y). \end{array}}$$

Pour résoudre ce problème d'optimisation à deux paramètres, nous introduisons un multiplicateur de Lagrange $q \in L^2(0, T; V)$ et le lagrangien correspondant :

$$(4.20) \quad \mathcal{L}(v, y, q) = J(v, y) - \int_0^T \left(q, \frac{dy}{dt} + A(t)y - f - Bv \right)_{V, V'} dt,$$

avec $y(0) = y_0$ et

$$J(v, y) = \frac{1}{2} \int_0^T |Cy - z_d|^2 dt + \frac{1}{2} \int_0^T (Nv, v)_U dt.$$

En intégrant par parties, nous pouvons encore écrire ce lagrangien sous la forme :

$$(4.21) \quad \begin{aligned} \mathcal{L}(v, y, q) = & J(v, y) + \int_0^T \left(\frac{dq}{dt}, y \right)_{V', V} dt - (q(T), y(T))_H + (q(0), y_0)_H \\ & - \int_0^T (A^*(t)q, y)_{V', V} dt + \int_0^T (q, f)_{V, V'} + \int_0^T (B^*q, v)_U dt. \end{aligned}$$

On a alors le résultat :

Théorème 4.18 *Si (u, \tilde{y}, p) est point-selle du lagrangien \mathcal{L} alors $\tilde{y} = y(u)$ et u est solution du problème d'optimisation (4.19) .*

Dans la définition du point-selle du lagrangien, la deuxième inégalité se traduit encore en termes de conditions d'Euler-Lagrange :

$$\begin{cases} \frac{\partial}{\partial y} \mathcal{L}(v, y, q)|_{(u, \bar{y}, p)} = 0, \\ \left(\frac{\partial}{\partial v} \mathcal{L}(v, y, q)|_{(u, \bar{y}, p)}, v - u \right) \geq 0, \quad \forall v \in \mathcal{U}_{\text{ad}}. \end{cases}$$

Calculons ces deux dérivées ; pour cela nous utilisons la formulation (4.21) du lagrangien :

$$\begin{aligned} \left(\frac{\partial}{\partial y} \mathcal{L}(v, y, q)|_{(u, \bar{y}, p)}, \phi \right) &= \int_0^T (C\tilde{y} - z_d, C\phi)_H dt \\ &+ \int_0^T \left(\frac{dp}{dt} - A^*(t)p, \phi \right)_{V', V} dt - (p(T), \phi(T))_H \\ &= 0, \quad \forall \phi \in L^2(0, T; V). \end{aligned}$$

Soit encore

$$\begin{aligned} \int_0^T (C^*(C\tilde{y} - z_d), \phi)_{V', V} dt + \int_0^T \left(\frac{dp}{dt} - A^*(t)p, \phi \right)_{V', V} dt \\ - (p(T), \phi(T))_H = 0, \quad \forall \phi \in L^2(0, T; V). \end{aligned}$$

Donc, l'état adjoint sera donné par le système :

$$\begin{cases} -\frac{dp}{dt} + A^*(t)p = C^*(C\tilde{y} - z_d) & \text{dans } \Omega, \quad \forall t \in]0, T[, \\ p(x, T) = 0, & \forall x \in \Omega, \quad t \in]0, T[. \end{cases}$$

D'autre part,

$$\left(\frac{\partial}{\partial v} \mathcal{L}(v, y, q)|_{(u, \bar{y}, p)}, v - u \right) = \int_0^T (Nu, v - u)_U dt + \int_0^T (B^*p, v - u)_U dt,$$

d'où :

$$\int_0^T (B^*p + Nu, v - u)_U dt \geq 0, \quad \forall v \in \mathcal{U}_{\text{ad}},$$

et ceci nous donne l'expression du gradient de la fonctionnelle J au point u . Finalement, la caractérisation du point-selle s'écrit sous la forme du système

d'optimalité :

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} \frac{d\tilde{y}}{dt} + A(t)\tilde{y} = f + Bu \quad \text{dans } \Omega, \forall t \in]0, T[\\ y(x, 0) = y_0 \quad \text{dans } \Omega \end{array} \right. \\ \left\{ \begin{array}{l} -\frac{dp}{dt} + A^*(t)p = C^*(C\tilde{y} - z_d) \quad \text{dans } \Omega, \forall t \in]0, T[\\ p(x, T) = 0 \quad \text{dans } \Omega \end{array} \right. \\ \int_0^T (B^*p + Nu, v - u)_u dt \leq 0, \quad \forall v \in \mathcal{U}_{ad} \\ \text{ou } B^*p + Nu = 0 \quad \text{dans le cas sans contrainte.} \end{array} \right.$$

4.4 Résolution numérique

Numériquement, le système d'optimalité du problème est constructif car il nous permet d'obtenir le point-selle, et par conséquent le contrôle optimal, en utilisant des méthodes itératives de type descente (gradient conjugué). L'algorithme va donc s'écrire sous la forme suivante :

- i) Initialisation $\rightarrow v^0$
- ii) Résolution du problème direct $\rightarrow y^0 = y(v^0)$
 Résolution du système adjoint $\rightarrow p^0 = p(v^0, y^0)$
 Calcul du gradient $\rightarrow \nabla J(v^0)$
 Minimisation par gradient conjugué $\rightarrow v^0 = v^0 - r^0 d^0$
 Si $\|\nabla J(v^0)\| \leq \epsilon \rightarrow$ sortir v^0
 sinon \rightarrow retour en i).

5 Introduction d'un terme de régularisation

Dans le cas d'un problème inverse que l'on modélise sous la forme $Ax = z$, lorsqu'il n'y a pas existence et unicité de la solution x pour toute observation z , ou si l'inverse A^{-1} n'est pas un opérateur continu, on dit que le problème est mal posé. Numériquement, on voit que ceci va nous poser des difficultés pour résoudre ce problème. Dans les cas qui nous intéressent, le problème sera mal posé en raison du manque de stabilité : une faible variation sur les données z risque d'induire des erreurs importantes sur la solution x . Pour pallier à ce type de problèmes, on remplace le problème initial par un problème approché, qui lui, sera stable par rapport aux données : il s'agit de la régularisation.

Dans les paragraphes suivants, nous rappelons les principales techniques de régularisation existantes et leurs propriétés. Ces techniques seront utilisées dans les chapitres ultérieurs afin de résoudre numériquement nos problèmes inverses mal posés.

5.1 L'approche "classique" de Tikhonov et Arsenine

On considère le problème inverse :

$$(5.1) \quad \text{Trouver } x \in X / Ax = z, \quad A : X \rightarrow Y,$$

z étant les observations dans Y connues. On suppose que l'inverse A^{-1} existe sur l'espace $R(A)$ mais n'est pas continu sur cet espace, on suppose de plus que l'espace des solutions Y n'est pas compact. Si l'on a un élément $z_\delta \in Y$ dont l'écart par rapport aux données exactes z_{ex} n'est pas supérieur à δ : $\|z_\delta - z_{ex}\| \leq \delta$, on voit bien que la solution approchée x_δ telle que $Ax_\delta = z_\delta$ ne peut être définie par $x_\delta = A^{-1}z_\delta$. Supposons qu'il existe une solution x_{ex} telle que $Ax_{ex} = z_{ex}$ avec $x_{ex} \in X$, $z_{ex} \in Y$, nous donnons la définition d'une classe d'opérateurs d'après Tikhonov et Arsenine [TA 77], appelés opérateurs régularisants, qui visent à rendre le problème stable par rapport aux données.

Définition 5.1 Un opérateur $R(z, \varepsilon)$, $z \in Y$, $\varepsilon \in \mathbb{R}^+$ est appelé opérateur régularisant pour le problème (5.1) dans le voisinage de $z = z_{ex}$ s'il possède les propriétés suivantes :

- 1) il existe $\delta_1 > 0$, tel que $R(z, \varepsilon)$ soit défini pour tout $\varepsilon > 0$ et pour tout $z \in Y$ tel que $\|z - z_{ex}\| \leq \delta \leq \delta_1$,
- 2) il existe $\varepsilon = \varepsilon(\delta)$ tel que pour tout $\eta > 0$, il existe $\delta(\eta) \leq \delta_1$ tel que si $z_\delta \in Y$ et $\|z_\delta - z_{ex}\| \leq \delta(\eta)$ alors $\|x_\varepsilon - x_\delta\| \leq \eta$ où x_ε est défini par : $x_\varepsilon = R(z_\delta, \varepsilon(\delta))$.

Cette solution x_ε est appelée *solution régularisée* du problème (5.1), le paramètre ε est appelé *paramètre de régularisation*.

Il est évident que tout opérateur régularisant, avec un choix approprié du paramètre de régularisation par rapport à l'erreur entachant les données δ , définit une méthode stable de construction approchée des solutions du problème inverse : si l'on a $\|z_\delta - z_{\text{ex}}\| \leq \delta$, alors, d'après la définition 5.1, on peut choisir $\varepsilon = \varepsilon(\delta)$ tel que $\lim_{\delta \rightarrow 0} \|x_\varepsilon - x_{\text{ex}}\| = 0$.

5.2 Construction d'opérateurs régularisants

Soit $\Omega(x)$ une fonctionnelle positive continue, définie sur un sous-ensemble X_1 de X dense dans X . On suppose de plus :

- a) $x_{\text{ex}} \in X_1$,
 - b) pour tout $d > 0$, $X_{1,d} = \{x \in X_1 / \Omega(x) \leq d\}$ est un compact de X ,
- alors on dit que Ω est une *fonctionnelle stabilisatrice*.

Soit

$$Q_\delta = \{x / \|Ax - z_\delta\| \leq \delta\} = \text{ensemble des solutions possibles.}$$

Cet ensemble est encore trop étendu pour que l'on puisse y choisir arbitrairement une solution approchée : celle-ci ne sera en général pas continue par rapport à δ . On cherche à résoudre :

$$(5.2) \quad \min_{x \in Q_\delta \cap X_1} \Omega(x).$$

L'existence de x vérifiant (5.2) est assurée, on la note x_δ . On pose $\tilde{R}(z_\delta, \delta) = x_\delta$, on peut montrer que \tilde{R} est un opérateur régularisant pour le problème (5.1), et que par conséquent $x_\delta = \tilde{R}(z_\delta, \delta)$ définit bien une solution approchée du problème.

• Méthode variationnelle de construction

On montre facilement que le problème d'optimisation (5.2) est équivalent au suivant :

$$\min_{x \in X_1} J_\varepsilon(x, z_\delta) \quad \text{avec } J_\varepsilon(x, z) = \|Ax - z\|^2 + \varepsilon \Omega(x).$$

Soit x_ε l'élément en lequel J_ε atteint son minimum, alors $R(z_\delta, \varepsilon) = x_\varepsilon$ définit un opérateur régularisant sous la condition $\|Ax_\varepsilon - z_\delta\| = \delta$ par exemple.

Il est à noter que, bien que le problème initial ne soit pas stable, le problème de la minimisation de la fonctionnelle J_ε est stable par rapport à de faibles variations

sur les données z . Cette stabilité est due à l'introduction de la fonctionnelle stabilisatrice $\Omega(x)$. La fonctionnelle $J_\varepsilon(x, z)$ sera appelée la fonctionnelle lissante du problème initial.

Toujours sous l'hypothèse qu'il existe une solution x_{ex} au problème initial, Tikhonov et Arsenine ont démontré le théorème d'existence suivant :

Théorème 5.2 *Soit A un opérateur continu de X dans Y , alors pour tout $z \in Y$ et $\varepsilon > 0$, il existe $x_\varepsilon \in X_1$ tel que :*

$$J_\varepsilon(x_\varepsilon, z) = \inf_{x \in X_1} J_\varepsilon(x, z).$$

Une condition suffisante pour que x_ε soit unique est que l'opérateur A soit linéaire, l'espace X soit un Hilbert et que $\Omega(x)$ soit une fonctionnelle stabilisatrice quadratique. Lorsque A est non linéaire, on n'a plus l'unicité de x_ε pour un ε donné.

D'autre part, en construisant des opérateurs régularisants par minimisation de la fonctionnelle lissante $J_\varepsilon(x, z)$, on n'arrive pas à définir de manière unique ce paramètre ε en fonction de l'erreur sur les données δ . Ce paramètre peut être défini en imposant au résidu, c'est à dire $\|Ax_\varepsilon^\delta - z_\delta\|$ la valeur δ , mais aussi à partir d'autres procédés (voir par exemple [MOR 84]).

- *Obtention de la solution régularisée sous forme d'une série*

On s'intéresse maintenant au cas particulier où A est un opérateur linéaire continu et où les espaces X et Y définis précédemment sont des espaces de Hilbert. Soit X_1 un sous-espace hilbertien de X muni de la norme $\| \cdot \|$ telle que $\forall d > 0$, l'ensemble $\{x \in X_1 / \|x\| \leq d\}$ soit compact dans X . On peut alors prendre $\Omega(x) = \|x\|^2$ comme fonctionnelle stabilisatrice. Dans ce cas, l'équation d'Euler associée à la fonctionnelle lissante $J_\varepsilon(x, z)$ s'écrit :

$$\langle Ax - z, A\phi \rangle + \varepsilon \langle x, \phi \rangle = 0, \quad \forall \phi \in X$$

d'où l'on tire :

$$A^*Ax + \varepsilon x = A^*z.$$

A^*A est un opérateur auto-adjoint. Soit $\{\phi_n\}$ un système de vecteurs propres pour A^*A et $\{\lambda_n\}$ les valeurs propres associées, alors on a :

$$A^*z = \sum_{n=1}^{\infty} c_n \phi_n.$$

Si l'on cherche la solution x sous la forme : $x = \sum_{n=1}^{\infty} x_n \phi_n$, on obtient pour ses coefficients la formule :

$$x_n = \frac{c_n}{\lambda_n + \varepsilon}$$

où le paramètre ε se définit par la formule du résidu.

Le problème de ce type de méthodes dites "classiques" est qu'elles nécessitent l'existence d'une solution exacte pour le problème initial non régularisé.

5.3 L'approche de K. Ito et K. Kunisch

Dans un article assez récent, K. Ito et K. Kunisch ont développé une nouvelle approche dont les résultats nous ont paru très intéressants [IK 90]. Comme on l'a vu au paragraphe précédent, on s'intéresse au problème de minimisation de la fonctionnelle lissante :

$$(5.3) \quad \min_{\substack{e(u,y)=0 \\ \langle Pu,u \rangle \leq \gamma \\ \ell(u) \in K}} \left(|y(u) - z|^2 + \varepsilon \langle Pu, u \rangle \right)$$

où $e(u, y) = 0$ représente l'équation d'état du problème, P est un opérateur linéaire auto-adjoint positif représentant une norme ou une semi-norme sur \mathcal{U}_{ad} . Pour $\varepsilon \geq 0$ et $\gamma \leq +\infty$, soit x_ε^δ la solution de ce problème d'optimisation. Si les données z n'étaient pas bruitées et si nous pouvions résoudre (5.3) avec une précision infinie, alors on pourrait prendre $\varepsilon = 0$ et $\gamma = +\infty$. Puisque ce n'est pas le cas en général, on choisit $\varepsilon > 0$ et $\gamma < +\infty$ pour stabiliser ce problème mal posé. Sous les hypothèses :

- $\exists m > 0$ tel que $\langle Pu, u \rangle \geq m |u|^2$ pour tout $u \in (\text{Ker } P)^\perp$,
- $\forall u \in \mathcal{U}_{\text{ad}} / \ell(u) \in K$, il existe un unique $y(u)$ tel que $(u, y(u))$ soit solution du problème direct,
- pour toute suite (u_n) convergeant faiblement vers u et telle que $\ell(u_n) \in K$, on a : $\liminf |y(u_n) - z|^2 \geq |y(u) - z|^2$ et $\lim_{n \rightarrow \infty} e(u_n, y(u_n)) = e(u, y(u))$,
- si (u_n) est une suite de \mathcal{U}_{ad} avec $\ell(u_n) \in K$ et $\langle Pu_n, u_n \rangle \leq \gamma$ pour tout n , et si $(\|P_1 u_n\|)$ et $(\|P_2 u_n\|)$ sont non bornées (où P_1 est la projection orthogonale sur $\text{Ker } P$ et P_2 sur $(\text{Ker } P)^\perp$), alors (u_n) ne peut pas être une suite minimisante pour le problème (5.3) ,

et si $\varepsilon > 0$ ou $\gamma < +\infty$, alors K. Ito et K. Kunisch ont montré qu'il existe une solution (u^*, y^*) de (5.3) .

Lorsque le problème initial est mal posé, l'introduction d'un terme de régularisation nous assure la stabilité du problème, au sens continuité de la solution par rapport aux données, mais d'un autre côté, cela introduit une nouvelle erreur sur la solution. D'un point de vue numérique, on voit donc l'importance du choix du paramètre de régularisation, qui doit réaliser le meilleur compromis entre la fidélité de la solution aux données et la stabilité du problème. Nous détaillerons au chapitre III plusieurs méthodes de choix de ce paramètre pour lesquels les résultats sont bien connus lorsque le problème inverse est linéaire, moins bien dans le cas non linéaire.

Chapitre II

Identification pour des problèmes inverses non linéaires

Nous nous intéressons dans ce chapitre à l'étude de problèmes d'identification de paramètres, mettant en œuvre les méthodes classiques du contrôle optimal. La difficulté essentielle de tels problèmes réside dans la non-linéarité de la variable d'état par rapport au contrôle, ce qui va nous conduire à modifier un peu notre approche par rapport aux méthodes classiques. Une autre difficulté sera due au fait que ces problèmes inverses sont mal posés au sens d'Hadamard : dans la plupart des cas, la solution n'est pas stable par rapport aux données. Ceci va nous inciter à introduire un terme de régularisation dans la fonctionnelle à minimiser afin de conserver une certaine régularité de la solution cherchée et une certaine stabilité.

Nous rappelons tout d'abord les résultats obtenus par G. Chavent concernant le problème d'identification de la diffusion de la chaleur ; dans un deuxième temps, nous présentons un problème d'identification de sources non linéaires dans une équation aux dérivées partielles elliptique.

1 Identification du coefficient de diffusion de la chaleur

Ce problème très classique a déjà été abondamment abordé dans la littérature : on pourra par exemple consulter [CHA 71], [CL 74], [BCK 83] ou [KS 85]. Cependant, il nous a semblé intéressant de le traiter à notre tour afin de tester notre approche et de la confronter à la théorie déjà existante. D'autre part, au chapitre IV, nous verrons une méthode originale de mesures de la non linéarité de ce problème par le calcul de courbures statistiques. Nous rappelons tout d'abord quelques résultats concernant l'identifiabilité de la diffusion et l'obtention du système d'optimalité du problème.

1.1 Position du problème

Par la suite, Ω sera un ouvert borné de \mathbb{R} ($\Omega = [0, 1]$ pour fixer les idées). Nous considérons les équations d'évolution suivantes :

$$(1.1) \quad \left\{ \begin{array}{l} \frac{\partial y}{\partial t}(x, t) - \frac{\partial}{\partial x} \left(\rho(x) \frac{\partial y}{\partial x}(x, t) \right) = f(x, t), \quad \forall x \in [0, 1], \forall t \in [0, T] \\ y(x, 0) = y_0(x), \quad \forall x \in [0, 1] \\ \rho(0) \frac{\partial y}{\partial x}(0, t) = d_0(t), \quad \rho(1) \frac{\partial y}{\partial x}(1, t) = d_1(t), \quad \forall t \in [0, T]. \end{array} \right.$$

Nous posons $V = H^1(\Omega)$ et $H = L^2(\Omega)$, et nous identifions H avec son dual ; on a donc :

$$V \subset H \subset V',$$

chaque inclusion étant dense et continue.

Les fonctions $f \in L^2(]0, T[\times \Omega) = L^2(0, T; H)$, $y_0 \in H$ et $d_0, d_1 \in L^2(]0, T[)$ sont toutes supposées connues. La fonction $y(x, t) \in L^2(0, T; V)$ représente la température du système à l'instant t , au point x , il s'agit pour nous de la variable d'état ($y = y(\rho)$). La fonction $\rho(x) \in L^\infty(\Omega)$, qui représente physiquement le coefficient de diffusion de la chaleur, sera notre variable de contrôle.

Afin d'obtenir la formulation variationnelle de ce problème, multiplions l'équation d'état (1.1) par une fonction $\phi \in V$ quelconque, et intégrons sur Ω en utilisant les conditions limites :

$$(1.2) \quad \left(\frac{\partial y}{\partial t}(t), \phi \right) + \int_{\Omega} \rho(x) \frac{\partial y}{\partial x}(x, t) \frac{\partial \phi}{\partial x}(x) dx \\ = (f(t), \phi) + d_1(t)\phi(1) - d_0(t)\phi(0).$$

Cette expression a bien un sens puisque l'application $\phi \mapsto d_0(t)\phi(0, t)$ est continue de $L^2(0, T; V)$ dans \mathbb{R} , donc appartient à $L^2(0, T; V')$. Nous définissons $y(t)$ la fonction d'espace par :

$$y(t) : \Omega \rightarrow L^2(0, T; V)$$

$$x \mapsto y(x, t).$$

On vérifie que les intégrales définies ci-dessus ont bien un sens pourvu que l'on choisisse ρ, f, ϕ et d_0, d_1 comme précédemment, et si pour tout $t \in]0, T[, y(t)$ appartient à l'espace V .

Nous nous donnons ensuite une forme bilinéaire a sur V :

$$a : V \times V \rightarrow \mathbb{R}$$

$$(1.3) \quad (\phi, \psi) \mapsto \int_{\Omega} \rho(x) \frac{\partial \phi}{\partial x}(x) \frac{\partial \psi}{\partial x}(x) dx,$$

et une forme linéaire $L \in L^2(0, T; V')$ telle que :

$$\forall \phi \in V, \quad L(t)(\phi) = \int_{\Omega} f(x, t)\phi(x)dx + d_1(t)\phi(1) - d_0(t)\phi(0).$$

Nous pouvons donc reformuler (1.2) , avec la condition initiale :

$$(1.4) \quad \begin{cases} \frac{d}{dt}(y(t), \phi) + a(y(t), \phi) &= L(t)(\phi), \quad \forall \phi \in V, \text{ p.p sur }]0, T[\\ y(0) &= y_0. \end{cases}$$

Ceci constitue la formulation faible du problème direct (1.1) . L'existence d'une solution y de (1.4) est donnée par le théorème suivant :

Théorème 1.1 *Sous les hypothèses précédentes, et pour tout ρ dans $\mathcal{U}_{ad} = \{f \in L^\infty(\Omega) / \exists \alpha > 0 : f(x) \geq \alpha \forall x \in \Omega\}$ et tout $(d_0, d_1) \in (L^2(]0, T[))^2$, le système (1.4) admet une solution unique $y \in W(0, T)$.*

Preuve : Il suffit pour la démonstration de ce résultat de constater que :

$$|a(\phi, \psi)| \leq \|\rho\|_\infty \times \|\phi\| \times \|\psi\|$$

et

$$a(\phi, \phi) = \int_{\Omega} \rho(x) \left(\frac{\partial \phi}{\partial x}(x)\right)^2 dx \geq \alpha \left|\frac{\partial \phi}{\partial x}\right|^2.$$

Donc on a $a(\phi, \phi) + \alpha \|\phi\|^2 \geq \alpha \|\phi\|^2$ et finalement, a est continue et coercive. En appliquant le théorème 4.14 du chapitre I, on obtient le résultat. □

Nous nous donnons à présent des observations $z_j \in L^2(]0, T[)$:

$$z_j(t) = \int_{\Omega} \zeta_j(x)y(x, t)dx + \text{perturbation}$$

localisées aux points $x_j, j = 1, \dots, m$, et où les fonctions ζ_j sont telles que :

$$\begin{cases} \zeta_j(x) \geq 0 \quad \forall x, & \int_{\Omega} \zeta_j(x) dx = 1 \\ \zeta_j(x) = 0 & \text{en dehors d'un petit voisinage de } x_j. \end{cases}$$

Le problème d'identification du coefficient de diffusion de la chaleur ρ se résume à minimiser la fonctionnelle suivante :

$$(1.5) \quad J(\rho) = \frac{1}{2} \sum_{j=1}^m \int_0^T \left(z_j(t) - \int_{\Omega} \zeta_j(x) y(x, t) dx \right)^2 \chi_j(t) dt,$$

où χ_j est une fonction caractéristique, valant 1 si l'on dispose d'une observation au point x_j à l'instant t , et 0 sinon. Nous choisissons maintenant pour ensemble des contrôles admissibles :

$$\mathcal{U}_{ad} = \left\{ \rho / 0 < \alpha_1 \leq \rho(x) \leq \alpha_2, \forall x \in \Omega \right\}$$

et l'on va chercher $\rho^* \in \mathcal{U}_{ad}$ qui vérifie :

$$(1.6) \quad \boxed{J(\rho^*) = \inf_{\rho \in \mathcal{U}_{ad}} J(\rho).}$$

Nous constatons immédiatement que cette fonctionnelle n'est pas convexe du fait de la non linéarité de l'état y par rapport au contrôle ρ .

1.2 Obtention du système d'optimalité

Pour tout j et pour presque tout $t \in]0, T[$, notons

$$e_j(t) = z_j(t) - (\zeta_j, y(t)).$$

Théorème 1.2 *Sous les hypothèses précédentes, si $\rho \in \mathcal{U}_{ad}$ et si les observations $z_d \in L^2(0, T)$, alors J est Fréchet-différentiable sur \mathcal{U}_{ad} , sa dérivée est donnée par : $(\nabla J(\rho), \xi) = \int_{\Omega} \xi(x) \int_0^T \frac{\partial y}{\partial x} \frac{\partial p}{\partial x} dt dx$ où p est la solution du problème :*

$$(1.7) \quad \begin{cases} -\frac{d}{dt}(p(t), \phi) + a(p(t), \phi) = \sum_{j=1}^m \int_{\Omega} \chi_j(t) e_j(t) \phi(x) dx, & \forall \phi \in V \\ p(T) = 0. \end{cases}$$

On sait que ce système admet une solution unique $p \in W(0, T)$. Ceci peut encore s'écrire :

$$\int_{\Omega} \rho(x) \frac{\partial p}{\partial x} \frac{\partial \phi}{\partial x} dx - \left(\frac{\partial p}{\partial t}, \phi \right) = \int_{\Omega} \sum_{j=1}^m \chi_j(t) e_j(t) \phi(x) dx.$$

En intégrant par parties sur Ω , on obtient l'état adjoint comme solution du système suivant :

$$(1.8) \quad \left\{ \begin{array}{l} -\frac{\partial p}{\partial t} - \frac{\partial}{\partial x} \left(\rho^*(x) \frac{\partial p}{\partial x} \right) = \sum_{j=1}^m \chi_j(t) e_j(t) \quad \text{dans } \Omega \times]0, T[\\ p(x, T) = 0 \quad \text{dans } \Omega \\ \rho^*(0) \frac{\partial p}{\partial x}(0, t) = 0, \quad \forall t \in]0, T[\\ \rho^*(1) \frac{\partial p}{\partial x}(1, t) = 0, \quad \forall t \in]0, T[. \end{array} \right.$$

En résumé, le système d'optimalité a donc la forme suivante :

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} \frac{\partial \tilde{y}}{\partial t} - \frac{\partial}{\partial x} \left(\rho^*(x) \frac{\partial \tilde{y}}{\partial x} \right) = f \quad \text{dans } \Omega \times]0, T[\\ \tilde{y}(x, 0) = y_0(x) \quad \text{dans } \Omega \\ \rho^*(0) \frac{\partial \tilde{y}}{\partial x}(0, t) = d_0(t), \quad \forall t \in]0, T[\\ \rho^*(1) \frac{\partial \tilde{y}}{\partial x}(1, t) = d_1(t), \quad \forall t \in]0, T[\end{array} \right. \\ \left\{ \begin{array}{l} -\frac{\partial p}{\partial t} - \frac{\partial}{\partial x} \left(\rho^*(x) \frac{\partial p}{\partial x} \right) = \sum_{j=1}^m \chi_j(t) e_j(t) \quad \text{dans } \Omega \times]0, T[\\ p(x, T) = 0 \quad \text{dans } \Omega \\ \rho^*(0) \frac{\partial p}{\partial x}(0, t) = 0, \quad \forall t \in]0, T[\\ \rho^*(1) \frac{\partial p}{\partial x}(1, t) = 0, \quad \forall t \in]0, T[\end{array} \right. \\ \int_0^T \int_{\Omega} \left(\rho(x) - \rho^*(x) \right) \frac{\partial \tilde{y}}{\partial x} \frac{\partial p}{\partial x} dx dt \geq 0, \quad \forall \rho \in \mathcal{U}_{ad}. \end{array} \right.$$

Faisons tout d'abord quelques remarques concernant l'existence et l'unicité du contrôle optimal pour ce problème. Dans un article [CHA 74], G. Chavent

a remarqué que même dans le cas le plus favorable d'une observation répartie en espace et en temps (c'est à dire que $z(x, t)$ est connue pour tout $x \in \Omega$, et pour tout $t \in]0, T[$), la solution pouvait être unique dans $L^\infty(]0, T[)$ à condition que l'ensemble $\Gamma_-(t) = \{x \in \partial\Omega / \nabla z(t) \cdot \vec{n} < 0\}$ soit vide (ce qui est le cas pour notre problème). Malheureusement, ceci n'empêche pas qu'une infinité de fonctions distributions ne puissent également être solutions. Numériquement, afin d'éviter de tels problèmes d'interférence avec des solutions distributions, nous voyons quel rôle important va jouer la base de décomposition du coefficient de diffusion.

Dans [CHA 79], une définition de l'identifiabilité de paramètres est donnée, ainsi qu'une condition suffisante d'identifiabilité.

1.3 Etude de la base de décomposition

Nous avons choisi une base de B -splines cubiques pour calculer la décomposition de ρ en raison de leur facilité de calcul : ce sont des polynômes de degré 3 par morceaux sur $[0, 1]$, qui sont globalement de classe C^2 sur $[0, 1]$. Ils constituent d'autre part une bonne approximation polynomiale des fonctions de classe C^2 .

1.3.1 Quelques rappels sur les B -splines

On considère des fonctions N_i^k , polynomiales par morceaux de degré au plus $k-1$, définies par la relation de récurrence : soit $\{t_0, t_1, \dots, t_m\}$ une suite croissante d'abscisses, on définit :

$$N_i^1(x) = \begin{cases} 1 & \text{si } t_i \leq x < t_{i+1}, \\ 0 & \text{sinon,} \end{cases}$$

puis, pour $k \geq 2$:

$$N_i^k(x) = \frac{x - t_i}{t_{i+k-1} - t_i} N_i^{k-1}(x) + \frac{t_{i+k} - x}{t_{i+k} - t_{i+1}} N_{i+1}^{k-1}(x).$$

Rappelons quelques propriétés de ces fonctions :

Propriété 1.3

$N_i^k(x) = 0$ en dehors de l'intervalle $[t_i, t_{i+k}]$ appelé support de la B -spline,

$N_i^k(x) \geq 0$ pour tout x ,

Si les nœuds sont tous distincts, alors N_i^k est de classe C^{k-2} .

Dans le cas qui nous intéresse, nous choisissons une base de B -splines de degré 3 (c'est à dire $k = 4$), et de la relation de récurrence ci-dessus, nous déduisons leur expression : soit $\{t_0, t_1, \dots, t_m\}$ une subdivision régulière de $[0, 1]$, de pas h ,

$$N_i^4(x) = \frac{(x - t_i)^3}{6h^3} \quad \text{si } x \in [t_i, t_{i+1}[,$$

$$N_i^4(x) = \frac{(x - t_i)^2(t_{i+2} - x)}{6h^3} + \frac{(x - t_i)(t_{i+3} - x)(x - t_{i+1})}{6h^3} + \frac{(t_{i+4} - x)(x - t_{i+1})^2}{6h^3}, \quad \text{si } x \in [t_{i+1}, t_{i+2}[$$

$$N_i^4(x) = \frac{(t_{i+4} - x)(x - t_{i+1})(t_{i+3} - x)}{6h^3} + \frac{(t_{i+4} - x)^2(x - t_{i+2})}{6h^3} + \frac{(x - t_i)(t_{i+3} - x)^2}{6h^3}, \quad \text{si } x \in [t_{i+2}, t_{i+3}[,$$

$$N_i^4(x) = \frac{(t_{i+4} - x)^3}{6h^3}, \quad \text{si } x \in [t_{i+3}, t_{i+4}[,$$

$$N_i^4(x) = 0 \quad \text{ailleurs.}$$

Nous donnons ici une illustration de ces fonctions :

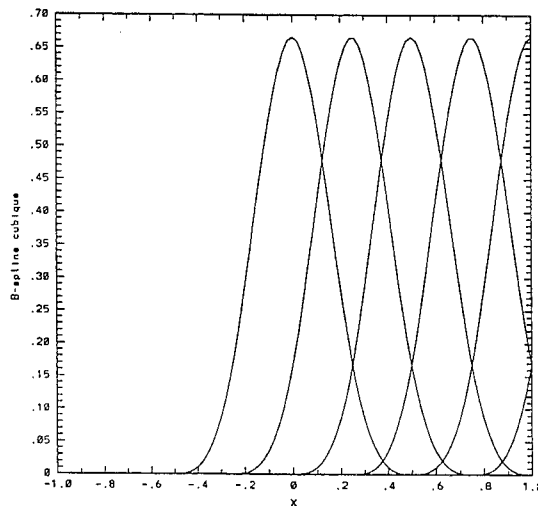


Figure 1.1 Les fonctions B -splines cubiques

Ces N_i^4 sont donc des fonctions polynomiales de degré 3 par morceaux sur chaque intervalle de la subdivision, et globalement de classe C^2 sur $[0, 1]$. La décomposition d'une fonction g de classe C^2 dans cette base s'écrira :

$$g(x) = \sum_{i=-3}^{m-1} \beta_i N_i^4(x)$$

où les $(\beta_{-3}, \beta_{-2}, \dots, \beta_{m-1})$ sont les nombres réels qui réalisent le minimum de la fonctionnelle :

$$\mathcal{J}(\beta_{-3}, \dots, \beta_{m-1}) = \sum_{j=-3}^{m-1} \left(\sum_{i=-3}^{m-1} \beta_i N_i^4(t_j) - g(t_j) \right)^2 + \lambda \int_0^1 \left(\sum_{i=-3}^{m-1} \beta_i (N_i^4)''(x) \right)^2 dx.$$

Le premier terme de cette fonctionnelle indique que g et $\sum_{i=-3}^{m-1} \beta_i N_i^4$ doivent coïncider aux points de contrôle (condition d'interpolation) et le second terme est une condition de régularité. Les coefficients de décomposition $(\beta_{-3}, \dots, \beta_{m-1})$ seront donc les $m+3$ nombres réels vérifiant : $\frac{\partial \mathcal{J}}{\partial \beta_k} = 0$ pour tout $k = -3, \dots, m-1$, soit encore :

$$\sum_{j=-3}^{m-1} \left(\sum_{i=-3}^{m-1} \beta_i N_i^4(t_j) - g(t_j) \right) N_k^4(t_j) + \lambda \int_0^1 \left(\sum_{i=-3}^{m-1} \beta_i (N_i^4)''(x) \right) (N_k^4)''(x) dx = 0$$

pour tout k . On obtient donc un système linéaire vérifié par les coefficients (β_i) : $A\beta = b$ où A est la matrice symétrique donnée par :

$$A_{ij} = \sum_{k=-3}^{m-1} N_i^4(t_k) N_j^4(t_k) + \lambda \int_0^1 (N_i^4)''(x) (N_j^4)''(x) dx.$$

On vérifie très facilement que cette matrice est symétrique et définie positive, d'où l'on déduit :

Propriété 1.4 Pour toute fonction de $C^2([0, 1])$, la décomposition dans la base de B -splines cubiques existe et est unique.

1.3.2 Application au problème d'identification

Afin de nous ramener en dimension finie, nous allons chercher à identifier le coefficient de diffusion de la chaleur dans cette base de B -splines cubiques :

$$(1.9) \quad \rho(x) = \sum_{i=-3}^{m-1} \beta_i N_i^4(x)$$

et notre problème d'optimisation va s'écrire :

$$\text{Trouver } \beta \in \mathbb{R}^{m+3} \text{ tel que } J(\beta) = \inf_{\gamma \in \mathbb{R}^{m+3}} J(\gamma)$$

avec

$$J(\gamma_{-3}, \gamma_{-2}, \dots, \gamma_{m-1}) = \frac{1}{2} \sum_{j=1}^n \int_0^T e_j^2(t) \chi_j(t) dt + \frac{\varepsilon}{2} \int_0^1 \left(\sum_{k=-3}^{m-1} \gamma_k (N_k^4)''(x) \right)^2 dx,$$

y_γ étant la solution du problème direct :

$$(1.10) \quad \left\{ \begin{array}{l} \frac{\partial y_\gamma}{\partial t} - \frac{\partial}{\partial x} \left(\sum_{i=-3}^{m-1} \gamma_i N_i^4(x) \frac{\partial y_\gamma}{\partial x} \right) = f \quad \text{dans } \Omega \times]0, T[\\ y_\gamma(x, 0) = y_0(x) \quad \text{dans } \Omega \\ \sum_{i=-3}^{m-1} \gamma_i N_i^4(0) \frac{\partial y_\gamma}{\partial x}(0, t) = d_0(t), \quad \forall t \in]0, T[\\ \sum_{i=-3}^{m-1} \gamma_i N_i^4(1) \frac{\partial y_\gamma}{\partial x}(1, t) = d_1(t), \quad \forall t \in]0, T[. \end{array} \right.$$

L'état adjoint sera obtenu comme la solution du système :

$$\left\{ \begin{array}{l} -\frac{\partial p}{\partial t} - \frac{\partial}{\partial x} \left(\sum_{i=-3}^{m-1} \gamma_i N_i^4(x) \frac{\partial p}{\partial x} \right) = \sum_{j=1}^n e_j(t) \chi_j(t) \quad \text{dans } \Omega \times]0, T[\\ p(x, T) = 0, \quad \forall x \in \Omega \\ \sum_{i=-3}^{m-1} \gamma_i N_i^4(0) \frac{\partial p}{\partial x}(0, t) = 0, \quad \forall t \in]0, T[\\ \sum_{i=-3}^{m-1} \gamma_i N_i^4(1) \frac{\partial p}{\partial x}(1, t) = 0, \quad \forall t \in]0, T[\end{array} \right.$$

et le gradient de la fonctionnelle s'écrira :

$$\left(\nabla J(\gamma) \right)_i = \int_0^T \int_\Omega N_i^4(x) \frac{\partial y_\gamma}{\partial x} \frac{\partial p}{\partial x} dx dt + \varepsilon \sum_{k=-3}^{m-1} \gamma_k \int_0^1 (N_i^4)''(x) (N_k^4)''(x) dx.$$

Dans la pratique, il s'agit d'une minimisation dans \mathbb{R}^{m+3} ; une méthode de gradient conjugué avec le calcul des $m+3$ directions de descente sera donc utilisée. Afin d'avoir une fonctionnelle quadratique à minimiser, il est cependant nécessaire de travailler sur le problème linéarisé.

1.4 Linéarisation du problème

Notons le problème direct (1.10) sous la forme simplifiée suivante :

$$(1.11) \quad \begin{cases} \frac{\partial y}{\partial t} + L_\gamma y &= f \quad \text{dans } \Omega \times]0, T[\\ y(x, 0) &= y_0(x), \quad \forall x \in \Omega \\ T_\alpha(\gamma, y) &= d_\alpha(t), \quad \forall t \in]0, T[, \quad \alpha = 0, 1 \end{cases}$$

$$\text{avec } L_\gamma y = -\frac{\partial}{\partial x} \left(\sum_{i=-3}^{m-1} \gamma_i N_i^4(x) \frac{\partial y}{\partial x} \right) \text{ et } T_\alpha(\gamma, y) = \sum_{i=-3}^{m-1} \gamma_i N_i^4(\alpha) \frac{\partial y}{\partial x}(\alpha, t).$$

Donnons-nous $\gamma^0 \in \mathbb{R}^{m+3}$, $y^0 \in L^2(0, T; V)$ et écrivons un développement Taylor à l'ordre 1 au voisinage de ces points :

$$L_\gamma y = L_{\gamma^0} y + L'_{\gamma^0}(\gamma - \gamma^0) y^0$$

et

$$\begin{aligned} T_\alpha(\gamma, y) &= T_\alpha(\gamma^0, y^0) + T'_{\alpha, \gamma}(\gamma - \gamma^0)(\gamma^0, y^0) + T'_{\alpha, y}(y - y^0)(\gamma^0, y^0) \\ &= \rho(\alpha) \frac{\partial y^0}{\partial x}(\alpha, t) + \rho^0(\alpha) \frac{\partial y}{\partial x}(\alpha, t) - \rho^0(\alpha) \frac{\partial y^0}{\partial x}(\alpha, t) \\ &= d_\alpha(t) - T_\alpha(\gamma - \gamma^0, y^0) \end{aligned}$$

où ρ et ρ^0 sont les fonctions dont la décomposition correspond aux vecteurs γ et γ^0 . On a donc finalement :

$$(1.12) \quad \begin{cases} \frac{\partial y}{\partial t} + L_{\gamma^0} y &= f - L_{\gamma - \gamma^0} y^0 \\ y(x, 0) &= y_0(x) \\ T_\alpha(\gamma^0, y) &= d_\alpha - T_\alpha(\gamma - \gamma^0, y^0), \quad \alpha = 0, 1. \end{cases}$$

On remarque que pour ρ , ρ^0 , y^0 donnés, ce système admet une unique solution $y_\ell \in L^2(0, T; V)$. Pour trouver la formulation faible de ce problème, nous multiplions les équations (1.12) par une fonction test $\phi \in V$; en intégrant par parties sur Ω , et en utilisant les conditions limites, on déduit :

$$\begin{cases} \left(\frac{\partial y(t)}{\partial t}, \phi \right) + a_{\rho^0}(y(t), \phi) &= a_{\rho - \rho^0}(y^0(t), \phi) + \left(f(t), \phi \right) \\ &+ L_1(t)(\phi) + L_2(t)(\phi, \rho), \quad \forall \phi \in V \\ y(x, 0) &= y_0(x), \end{cases}$$

avec

$$L_1(t)(\phi) = -\phi(0)d_0(t) - 2\phi(0)\rho^0(0)\frac{\partial y^0}{\partial x}(0,t) + \phi(1)d_1(t) + 2\phi(1)\rho^0(1)\frac{\partial y^0}{\partial x}(1,t),$$

$$L_2(t)(\phi, \rho) = -2\phi(1)\rho(1)\frac{\partial y^0}{\partial x}(1,t) + 2\phi(0)\rho(0)\frac{\partial y^0}{\partial x}(0,t).$$

Nous considérons ensuite le lagrangien associé :

$$\begin{aligned} \mathcal{L}_\ell(\rho, y, p) = & J(\rho, y) + \int_0^T \left(\left(\frac{\partial y}{\partial t}(t), p(t) \right) + a_{\rho^0}(y(t), p(t)) \right. \\ & - a_{\rho-\rho^0}(y^0(t), p(t)) - (f(t), p(t)) \\ & \left. - L_1(t)(p(t)) + L_2(t)(p(t), \rho) \right) dt. \end{aligned}$$

Nous écrivons ensuite les conditions d'optimalité d'Euler-Lagrange :

$$\left(\frac{\partial \mathcal{L}_\ell}{\partial y}, \phi \right) = 0, \quad \forall \phi \in L^2(0, T; V).$$

Soit encore, après une intégration par parties sur Ω :

$$\begin{aligned} & - \int_0^T \int_\Omega \phi \frac{\partial p}{\partial t} dx dt - \int_0^T \int_\Omega \frac{\partial}{\partial x} \left(\rho^0(x) \frac{\partial p}{\partial x} \right) \phi dx dt \\ & + \int_0^T \rho(1) \frac{\partial p}{\partial x}(1,t) \phi(1,t) dt - \int_0^T \rho(0) \frac{\partial p}{\partial x}(0,t) \phi(0,t) dt \\ & + \int_\Omega \phi(x, T) p(x, T) dx = \sum_{j=1}^m \int_0^T \int_\Omega \chi_j(t) e_j(t) dx dt, \quad \forall \phi \in L^2(0, T; V). \end{aligned}$$

On obtient donc l'état adjoint correspondant au système linéarisé autour de ρ^0 et y^0 :

$$(1.13) \quad \left\{ \begin{array}{l} -\frac{\partial p}{\partial t} - \frac{\partial}{\partial x} \left(\rho^0(x) \frac{\partial p}{\partial x} \right) = \sum_{j=1}^n \chi_j(t) e_j(t) \\ p(x, T) = 0 \\ \rho^0(0) \frac{\partial p}{\partial x}(0, t) = 0 \\ \rho^0(1) \frac{\partial p}{\partial x}(1, t) = 0 \end{array} \right.$$

avec y la solution du système (1.12). Ce système, tout comme celui de l'état linéarisé admet une unique solution $p \in L^2(0, T; V)$, et le gradient de la fonctionnelle sera donné par :

$$(1.14) \quad (\nabla J(\rho), \xi) = \left(\frac{\partial \mathcal{L}_\ell}{\partial \rho}, \xi \right) = \varepsilon \int_0^1 \rho'' \xi'' dx - \int_0^T \int_\Omega \xi(x) \frac{\partial y^0}{\partial x} \frac{\partial p}{\partial x} dx dt.$$

L'algorithme de minimisation va donc s'écrire :

Choix de ρ^0 .

Calcul de y^0 solution du problème direct (1.1) en ρ^0 .

Calcul du point-selle (ρ, y, p) de \mathcal{L}_ℓ .

Tant que $\|y - y^0\| \geq \varepsilon \|y^0\|$ faire :

$$\rho^0 \leftarrow \rho$$

$$y^0 \leftarrow y$$

Calcul du point-selle (ρ, y, p) de \mathcal{L}_ℓ .

Fin tant que.

$\rho =$ la solution cherchée.

Il s'agit d'une méthode séquentielle quadratique : à chaque pas de la linéarisation, on cherche l'optimum de la fonctionnelle quadratique correspondant à la linéarisation de l'état par rapport au contrôle. La boucle "tant que" correspond à la linéarisation du problème d'optimisation ; ses itérations constituent les *itérations externes* de l'algorithme. La recherche du minimum de la fonctionnelle quadratique constitue les *itérations internes* de cet algorithme.

1.5 Discrétisation des systèmes d'EDP

Pour inverser les systèmes (1.12) et (1.13), nous nous proposons d'utiliser un schéma aux différences finies de type Crank-Nicholson, ce qui nous donnera une meilleure stabilité qu'avec un schéma implicite seulement (on pourra éventuellement consulter [RT 88]).

Soit $\{t_0, \dots, t_M\}$ une discrétisation de l'espace de temps $[0, T]$ et $\{x_0, \dots, x_N\}$ une discrétisation en espace de l'intervalle $[0, 1]$. Soit y_h la solution approchée du système (1.12) et soient Δ_t et Δ_h les pas de discrétisation en temps et en espace. Notons $y_i^k = y_h(x_i, t_k)$, on obtient le système approché :

$$\rho(x_{j-\frac{1}{2}}) \frac{(y_j^{k+1} - y_{j-1}^{k+1} + y_j^k - y_{j-1}^k)}{2\Delta_h^2} - \rho(x_{j+\frac{1}{2}}) \frac{(y_{j+1}^{k+1} - y_j^{k+1} + y_{j+1}^k - y_j^k)}{2\Delta_h^2} + \frac{y_j^{k+1} - y_j^k}{\Delta_t} = \frac{f(x_j, t_k) + f(x_j, t_{k+1})}{2}, \quad \forall j = 0, \dots, N, \quad \forall k = 0, \dots, M-1,$$

avec la condition initiale :

$$y_j^0 = y_0(x_j), \quad \forall j = 0, \dots, N$$

et les conditions au bord :

$$\rho(x_{-\frac{1}{2}}) \frac{y_0^k - y_{-1}^k}{\Delta h} = 2d_0(t_k) - \rho(x_{\frac{1}{2}}) \frac{y_1^k - y_0^k}{\Delta h}, \quad \forall k = 0, \dots, M$$

et

$$\rho(x_{N+\frac{1}{2}}) \frac{y_{N+1}^k - y_N^k}{\Delta h} = 2d_1(t_k) - \rho(x_{N+\frac{1}{2}}) \frac{y_N^k - y_{N-1}^k}{\Delta h}, \quad \forall k = 0, \dots, M.$$

Pour chaque instant t_i , nous résolvons ce système en inversant la matrice triangulaire \mathcal{A} :

$$\mathcal{A} = \begin{pmatrix} b_0 & c_1 & 0 & \dots & 0 \\ a_0 & b_1 & c_2 & \dots & 0 \\ 0 & a_1 & b_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & b_N \end{pmatrix}$$

avec

$$\begin{cases} b_0 = \frac{1}{\Delta t} + \frac{\rho_{\frac{1}{2}}}{\Delta h} \\ b_i = \frac{1}{\Delta t} + \frac{\rho_{i-\frac{1}{2}} + \rho_{i+\frac{1}{2}}}{2\Delta h^2} \text{ pour } i = 1, \dots, N-1, \\ b_N = \frac{1}{\Delta t} + \frac{\rho_{N-\frac{1}{2}}}{\Delta h^2} \end{cases}$$

et

$$\begin{cases} a_i = -\frac{\rho_{i+\frac{1}{2}}}{2\Delta h^2} \text{ pour } i = 0, \dots, N-2 \\ a_{N-1} = -\frac{\rho_{N-\frac{1}{2}}}{\Delta h^2} \end{cases}$$

et

$$\begin{cases} c_1 = -\frac{\rho_{\frac{1}{2}}}{\Delta h^2} \\ c_i = -\frac{\rho_{i-\frac{1}{2}}}{2\Delta h^2} \text{ pour } i = 2, \dots, N-1 \end{cases}$$

et où

$$\rho_{\frac{k}{2}} = \rho(x_{\frac{k}{2}}) = \rho\left(\frac{k\Delta h}{2}\right).$$

Nous en déduisons les $y_h(x_j, t_i)$ avec $i = 0, \dots, M$ et $j = 0, \dots, N$ puis, en résolvant le système transposé avec le second membre approprié, nous obtenons les $p_h(x_j, t_i)$. La matrice \mathcal{A} est inversée en la rendant triangulaire supérieure.

1.6 Validation du gradient

Afin de vérifier le gradient obtenu par (1.14), nous considérons le rapport des différences divisées :

$$\frac{J(\rho + h\nabla J(\rho)) - J(\rho)}{h \|\nabla J(\rho)\|^2},$$

qui, pour un pas h arbitrairement petit, devrait tendre vers la valeur 1. Nous représentons numériquement la fonction

$$h \mapsto 1 - \frac{J(\rho + h\nabla J(\rho)) - J(\rho)}{h \|\nabla J(\rho)\|^2}.$$

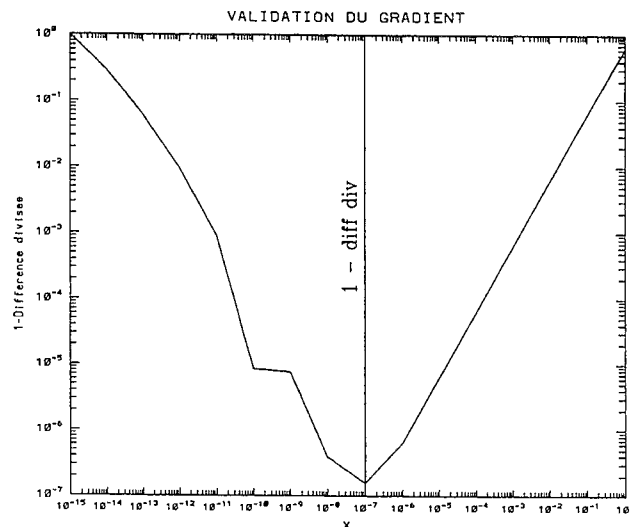


Figure 1.2 Validation du gradient par différence divisée.

On constate que la précision obtenue est d'autant meilleure que h est petit, mais lorsque h devient inférieur à 10^{-7} , ce rapport "explose" en raison des erreurs numériques d'arrondi. Il s'agit donc d'un bon test du gradient lorsqu'on connaît une valeur exacte de celui-ci, mais cette méthode en elle-même est à déconseiller pour un calcul direct du gradient en raison de ces instabilités numériques.

1.7 Résultats numériques

1.7.1 Identification sans bruit

Nous présentons quelques résultats d'identification de la diffusion. Les fonctions à identifier sont choisies non linéaires, positives et de classe C^2 sur l'intervalle $[0,1]$. Tout d'abord, les résultats sont donnés pour des données simulées non bruitées, sans terme de régularisation (i.e. $\varepsilon = 0$).

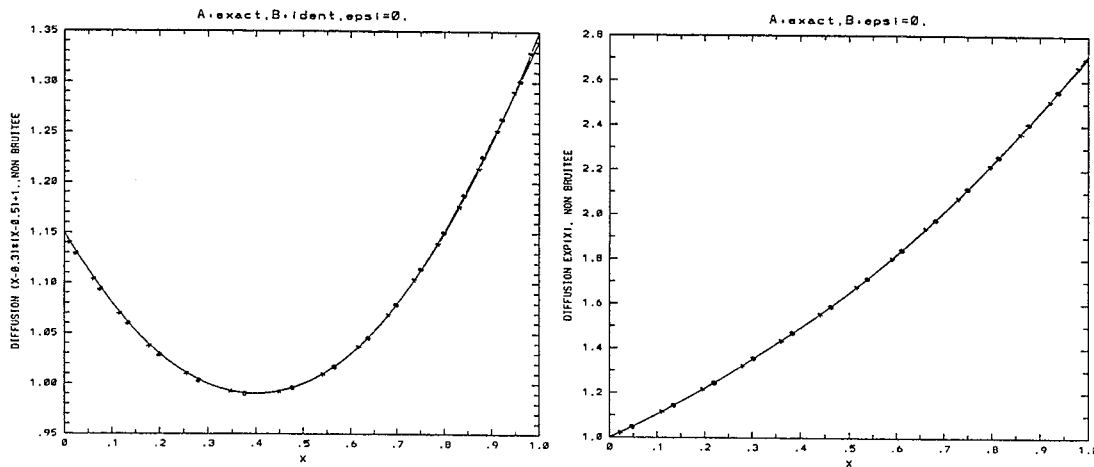


Figure 1.3 A gauche, $\rho(x) = (x - 0.5)(x - 0.3) + 1$, à droite, $\rho(x) = e^x$.

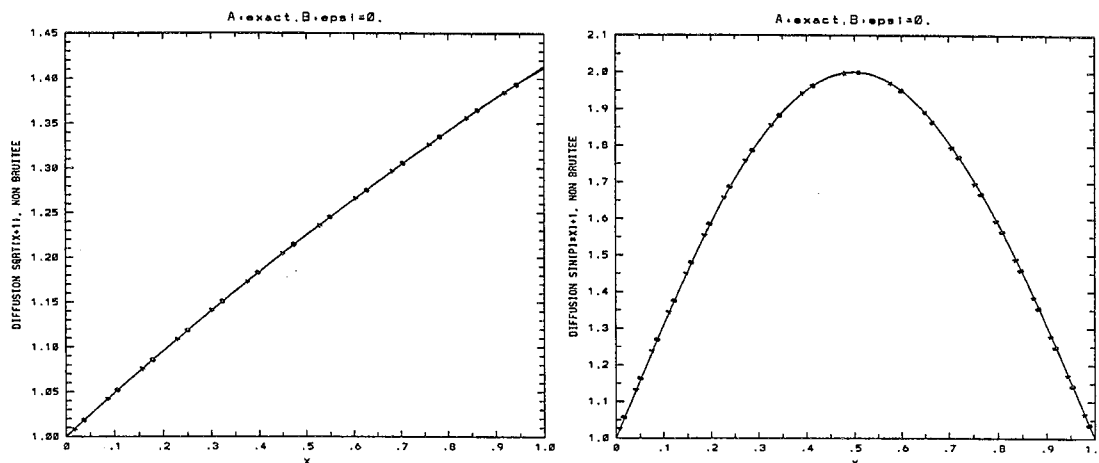


Figure 1.4 A gauche, $\rho(x) = \sqrt{x+1}$, à droite, $\rho(x) = \sin(\pi x) + 1$.

Nous voyons que la méthode utilisée nous permet d'obtenir d'assez bons résultats sur la diffusion que l'on souhaite identifier. Il est à noter que l'algorithme converge rapidement, en moyenne 4 ou 5 itérations externes suffisent, et la méthode d'optimisation à l'intérieur des itérations internes converge en 6 itérations (la décomposition de la diffusion utilise ici 7 fonctions B -splines).

1.7.2 Identification avec bruit, sans régularisation

Lorsque les observations sont bruitées à 2%, c'est à dire que l'on rajoute aux observations un terme pseudo-aléatoire de loi $\mathcal{N}(0, \sigma)$ où l'écart-type σ est pris égal à 2% de l'énergie totale des observations exactes, alors si l'on n'introduit pas de régularisation, l'algorithme a des difficultés à identifier la bonne fonction. Lorsque la proportion de bruit augmente, l'identification sera d'autant plus difficile à réaliser. Nous représentons ci-dessous les fonctions identifiées à partir de données simulées bruitées (à gauche: 2% de bruit, à droite: 5% de bruit) sans terme de régularisation dans la fonctionnelle à minimiser. (La courbe A représente la fonction exacte, x^2 ici, B la fonction identifiée.)

Pour seulement 2% de bruit, l'algorithme converge en une trentaine d'itérations externes, ce qui représente environ 10 secondes CPU sur un CRAY C90. L'augmentation de bruit se traduit généralement par une augmentation du temps de calcul (pour 5% de bruit, 120 itérations externes sont nécessaires).

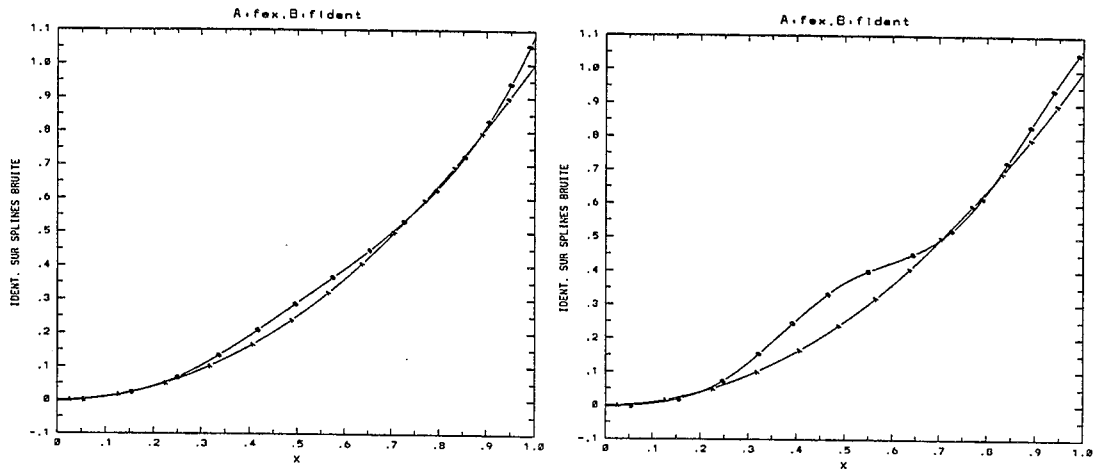


Figure 1.5 Identification de la diffusion à partir de données bruitées.

1.7.3 Régularisation pour l'identification à partir de données bruitées

L'effet de la régularisation est illustré par les deux figures suivantes pour un paramètre $\varepsilon = 10^{-2}$, avec les mêmes proportions de bruit sur les observations :

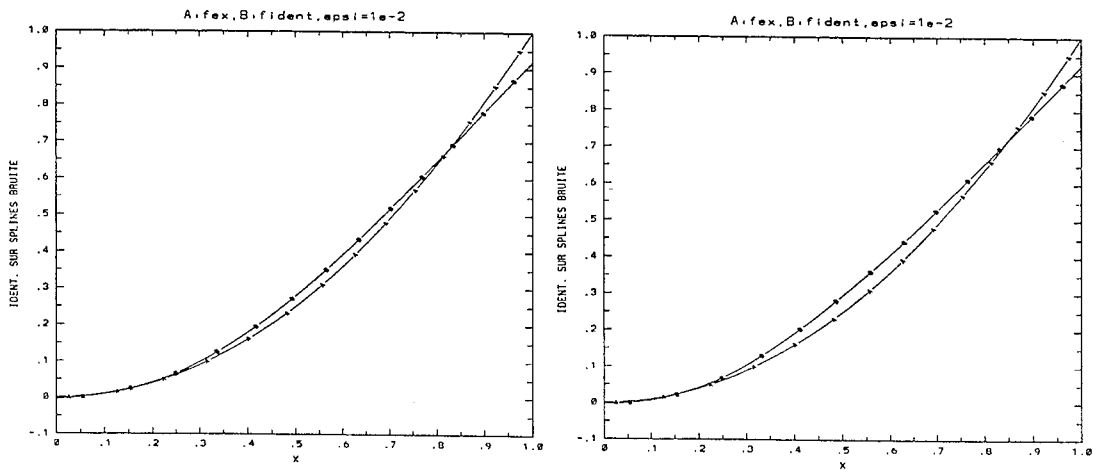


Figure 1.6 Régularisation pour l'identification à partir de données bruitées.

Les fonctions identifiées sont donc beaucoup plus proches de la réalité que dans le cas de la figure 1.5 . Il est aussi à noter que les temps d'exécution ont diminué: 5 itérations externes pour 2% de bruit (environ 1,5 secondes CPU) et 7 itérations externes pour 5% de bruit (soit 1,8 secondes CPU). Le choix du paramètre de régularisation est donc crucial dans ce problème. Nous l'étudierons en détail au chapitre III.

2 Identification de sources non linéaires

Nous nous intéressons ici à un problème inverse non linéaire provenant, sous une forme très simplifiée, du problème de l'interprétation de mesures expérimentales dans le plasma (gaz ionisé) d'un Tokamak (dispositif expérimental visant à confiner le plasma dans un champ magnétique, dont une analyse a été faite par J. Blum dans [BLU 89] et [BLO 90]). L'équilibre axisymétrique du plasma est régi par une EDP elliptique non linéaire qui s'écrit : $A\psi = rf_1(\psi) + \frac{1}{r}f_2(\psi)$ avec $A = -\operatorname{div}\left(\frac{1}{r}\operatorname{grad}\right)$ (équation de Grad-Shafranov) dans le plan de section méridienne du tore de coordonnées (r, z) . L'inconnue $\psi(r, z)$ est le flux du champ magnétique, l'opérateur A est elliptique linéaire. Le second membre de l'équation représente la densité de courant du plasma. Il s'agit d'identifier les fonctions $f_1(\psi)$ et $f_2(\psi)$ qui ne peuvent être mesurées directement dans le plasma. Pour ce faire, on dispose des informations suivantes :

- les mesures expérimentales du flux ψ et de $\frac{\partial\psi}{\partial n}$ sur le bord du domaine,
- la connaissance des intégrales sur un certain nombre de cordes de la composante verticale du champ magnétique, à savoir $\frac{1}{r}\frac{\partial\psi}{\partial r}$.

Le problème mathématique de l'identifiabilité de f_1 et f_2 à partir des mesures est un problème ouvert.

Nous allons nous restreindre ici au cas cylindrique : lorsque le rapport $\frac{R}{r}$ des rayons du tore tend vers $+\infty$, le tore tend vers un cylindre infini et l'opérateur A devient l'opérateur laplacien. L'équation d'état devient $-\Delta\psi = f(\psi)$ et nous allons étudier le problème d'identification de f à partir de la donnée $\frac{\partial\psi}{\partial n}$ sur le bord du domaine.

Il s'agit d'identifier la non linéarité du problème à partir de la condition de Neumann $\frac{\partial y}{\partial n}$ sur le bord de Ω ; pour cela, nous nous donnons une fonctionnelle à minimiser :

$$J(f) = \frac{1}{2} \int_{\partial\Omega} \left(\frac{\partial y(f)}{\partial n} - z \right)^2 d\sigma + \frac{\varepsilon}{2} \int_0^1 f''(x)^2 dx$$

où $y(f)$ est solution de :

$$\begin{cases} -\Delta y = f(\bar{y}) & \text{dans } \Omega \\ y = 0 & \text{sur } \partial\Omega \end{cases}$$

avec $\bar{y} = \frac{y}{\max y}$.

Le premier terme traduit la fidélité de la solution aux données et le second est un terme de régularisation de Tikhonov. Il est à noter que la formulation du second membre de l'équation d'état nous permet d'avoir à identifier la fonction f sur un intervalle connu, ce qui ne serait pas le cas si on avait considéré $g(y)$.

Ce problème est resté ouvert durant une assez longue période, sauf dans le cas où Ω est un disque, où y devient radiale, et sa dérivée normale étant constante sur le bord du domaine, n'apporte pas suffisamment d'informations sur le profil de la fonction f (cf [GNN 79]). Nous voyons donc l'importance jouée par la forme du domaine Ω sur l'identification.

Lorsque $\partial\Omega$ possède au moins un angle $\theta_0 \in]\frac{\pi}{2}, 2\pi[\setminus \{\frac{3\pi}{2}, \pi\}$, E. Beretta et M. Vogelius [BV 91] ont montré que les hypothèses :

- f_1 et f_2 sont continues sur $\Omega \subset \mathbb{R}^2$ et infiniment différentiables au voisinage de 0,
- y_1 et y_2 sont solutions classiques des problèmes :

$$\begin{cases} -\Delta y_i = f_i(y_i) \text{ dans } \Omega \\ y_i = 0 \text{ sur } \partial\Omega \end{cases}$$

avec $f_i(y_i(x)) \geq 0$ dans Ω ,

- $\frac{\partial y_1}{\partial n} = \frac{\partial y_2}{\partial n}$ sur un segment d'extrémité le point où l'on a l'angle θ_0 ,

entraînent

$$f_1^{(k)}(0) = f_2^{(k)}(0), \quad \forall k > 0.$$

Comme f_1 et f_2 sont analytiques au voisinage de 0, on en déduit que f_1 et f_2 coïncident sur ce voisinage. Si d'autre part, le domaine ne contient qu'un seul coin, il est immédiat que ces deux fonctions f_1 et f_2 sont identiques sur tout le domaine.

Il est à noter que les deux auteurs viennent d'étendre ce résultat au cas où l'angle $\theta_0 \in [0, 2\pi] \setminus \{\pi\}$ (cf [BV 95]) sous les hypothèses suivantes : Ω est un domaine borné de \mathbb{R}^2 , polygonal curvilinéaire de classe $\mathcal{C}^{2,\alpha}$, possédant au moins un angle "propre" θ_0 (c'est à dire $\theta_0 \neq \pi$) dont les extrémités sont des segments du bord. Alors, si $\frac{\partial y_1}{\partial n} = \frac{\partial y_2}{\partial n}$ sur $\partial\Omega \setminus \{\theta_0\}$ on a :

- si $f_1(0) = 0$ ou $f_2(0) = 0$, alors $f_1 = f_2$,
- si $f_1(0) \neq 0$ ou $f_2(0) \neq 0$, alors $f_1 = f_2$ pourvu que $\theta_0 \geq \frac{\pi}{2}$.

Lorsque $\partial\Omega$ ne possède pas de coin, par exemple dans le cas d'un ouvert régulier, le seul résultat (existant à notre connaissance) a été établi par M. Vogelius [VOG 94] : il a montré que pour un domaine de classe \mathcal{C}^3 de \mathbb{R}^2 borné, et pour une fonction $\psi \in \mathcal{C}^1(\partial\Omega)$ non dégénérée (i.e. non constante et ayant au plus un ensemble dénombrable de zéros), il existe seulement un nombre fini de couples

(c_k, d_k) de nombres réels tels que le problème de Cauchy :

$$\begin{cases} -\Delta y = c_k y + d_k & \text{dans } \Omega, \\ y = 0 & \text{sur } \partial\Omega, \\ \frac{\partial y}{\partial n} = \psi & \text{sur } \partial\Omega, \end{cases}$$

admette une solution. Le problème de l'identification de f est donc partiellement résolu lorsque f est une fonction affine. Dans le cas plus général, ce problème est ouvert. On pourra éventuellement consulter [BFV 92] pour des expériences numériques menées par M. Vogelius et al.

Nous allons par la suite nous placer dans le cadre le plus général de ce problème, et nous considérerons un domaine quelconque $\Omega \subset \mathbb{R}^2$ régulier et simplement connexe.

2.1 Le problème direct

Nous considérons le problème suivant :

Trouver y tel que :

$$(2.1) \quad \begin{cases} -\Delta y = f(\bar{y}) & \text{dans } \Omega \\ y = 0 & \text{sur } \partial\Omega \end{cases}$$

où Ω est un ouvert borné de \mathbb{R}^2 de classe \mathcal{C}^2 , f est une fonction positive de classe \mathcal{C}^2 définie sur $[0, 1]$ à valeurs dans \mathbb{R} , et $\bar{y} = \frac{y}{\max_{\Omega} y}$. Le principe du maximum nous assure que y est positif dans $\bar{\Omega}$, et que \bar{y} est à valeurs dans $[0, 1]$.

2.1.1 Formulation faible

Nous considérons le problème direct : trouver y vérifiant (2.1). Nous rappelons la définition de l'espace de Sobolev $W^{1,p_0}(\Omega)$:

$$W^{1,p_0}(\Omega) = \{\phi \in L^{p_0}(\Omega) / \nabla \phi \in (L^{p_0}(\Omega))^2\}$$

avec $\nabla \phi = \left(\frac{\partial \phi}{\partial x_1}, \frac{\partial \phi}{\partial x_2} \right)$. On notera

$$(2.2) \quad V^{p_0}(\Omega) = \{\phi \in W^{1,p_0}(\Omega) / \phi = 0 \text{ sur } \partial\Omega\}.$$

Cet espace $V^{p_0}(\Omega)$ est contenu dans $W^{1,p_0}(\Omega)$. Soit \bar{p}_0 le scalaire conjugué de p_0 au sens :

$$\frac{1}{p_0} + \frac{1}{\bar{p}_0} = 1.$$

Pour tout couple $(\phi, \psi) \in V^{p_0} \times V^{\bar{p}_0}$, nous définissons la forme bilinéaire :

$$a(\phi, \psi) = \int_{\Omega} \nabla \phi \cdot \nabla \psi dx.$$

Puisque $\nabla \phi \in (L^{p_0}(\Omega))^2$ et $\nabla \psi \in (L^{\bar{p}_0}(\Omega))^2$, nous avons $\nabla \phi \cdot \nabla \psi \in L^1(\Omega)$, ce qui nous assure que la forme bilinéaire a est bien définie. La formulation faible du problème direct s'écrit alors : étant donnée une fonction f appartenant à $C^2([0, 1])$,

Trouver $y \in V^{p_0}(\Omega)$ avec $p_0 > 2$, tel que :

$$(2.3) \quad a(y, \psi) = \int_{\Omega} f(\bar{y}) \psi dx, \quad \forall \psi \in V^{\bar{p}_0}(\Omega).$$

Remarque 2.1 : le scalaire p_0 a été choisi strictement supérieur à 2, de telle façon que toute fonction de $V^{p_0}(\Omega)$ soit continue sur $\bar{\Omega}$ (théorème de Rellich-Kondrachov [BRE 87]). Cette continuité va nous permettre par la suite de pouvoir considérer $\max_{\Omega} y = y(M_0)$ qui sera donc bien défini.

2.1.2 “Equivalence” de la formulation faible et du problème direct

Proposition 2.2 Soit $y \in V^{p_0}(\Omega)$ solution du problème (2.1) avec $p_0 > 2$. Si de plus, ∇y appartient à $(W^{1,p_0}(\Omega))^2$, alors y est solution de la formulation faible (2.3). Réciproquement, si $y \in V^{p_0}(\Omega)$ est solution de la formulation faible, avec $p_0 > 2$, et si ∇y appartient à $(W^{1,p_0}(\Omega))^2$ alors y est solution du problème (2.1).

La démonstration de cette proposition réside dans un lemme dû à J. Nečas [NEC 67] donnant la formule d'intégration par parties :

Lemme 2.3 Soit Ω un ouvert borné de \mathbb{R}^2 , de frontière régulière Γ . Soient $u \in (W^{1,p_0}(\Omega))^2$ et $v \in W^{1,p'_0}(\Omega)$ où p'_0 est le conjugué scalaire de p_0 . Alors, on a la formule suivante d'intégration par parties :

$$(2.4) \quad \int_{\Omega} v \operatorname{div}(u) dx = - \int_{\Omega} u \cdot \nabla v dx + \int_{\Gamma} (u \cdot \vec{n}) v d\sigma$$

où \vec{n} est le vecteur normal unitaire de Γ .

Preuve de la proposition : Soit y solution du problème direct, avec $y \in V^{p_0}(\Omega)$ et soit ψ une fonction de $V^{p'_0}(\Omega)$. En multipliant (2.1) par ψ et en intégrant sur Ω :

$$\int_{\Omega} -\Delta y \psi dx = \int_{\Omega} f(\bar{y}) \psi dx.$$

Si ∇y est dans $(W^{1,p_0}(\Omega))^2$, on peut appliquer la formule (2.4) avec $u = \nabla y$ et $v = \psi$, on obtient alors :

$$\int_{\Omega} \nabla y \cdot \nabla \psi dx - \int_{\Gamma} (\nabla y \cdot \vec{n}) \psi d\sigma = \int_{\Omega} f(\bar{y}) \psi dx$$

et puisque $\psi \in V^{p'_0}(\Omega)$, ψ est nul sur la frontière $\partial\Omega$:

$$a(y, \psi) = \int_{\Omega} f(\bar{y}) \psi dx$$

et donc y est bien solution faible. Réciproquement, si $y \in V^{p_0}(\Omega)$ est solution faible, alors on considère $\psi \in \mathcal{D}(\Omega)$:

$$\int_{\Omega} \nabla y \cdot \nabla \psi dx = \int_{\Omega} f(\bar{y}) \psi dx.$$

En appliquant la formule d'intégration par parties précédente avec $u = \nabla y$ et $v = \psi$:

$$- \int_{\Omega} \psi \Delta y dx + \int_{\partial\Omega} (\psi \cdot \vec{n}) d\sigma = \int_{\Omega} f(\bar{y}) \psi dx, \quad \forall \psi \in \mathcal{D}(\Omega)$$

d'où l'on déduit que y est bien solution du problème direct (2.1) . \square

2.1.3 Linéarisation du problème

Les problèmes (2.1) et (2.3) contiennent une non linéarité due au terme de source $f(\bar{y})$. La linéarisation du problème faible (2.3) est plus facilement

réalisable du point de vue numérique que celle du problème direct. Pour cela, on considère l'application :

$$(2.5) \quad \begin{aligned} b : V^{p_0}(\Omega) &\rightarrow V^{p_0}(\Omega \rightarrow [0, 1]) \\ \phi &\mapsto \bar{\phi} = \frac{\phi}{\max_{\Omega} \phi}. \end{aligned}$$

Nous faisons ensuite les hypothèses suivantes sur y solution du problème faible :

H1) $\max_{\Omega} y$ est atteint en un unique point M_0 intérieur à Ω .

H2) y est de classe C^2 localement dans un voisinage de M_0 , et M_0 est un point elliptique non dégénéré.

On remarque immédiatement que si $y \in W^{1,p_0}(\Omega)$ avec $p_0 > 2$ alors $y(M_0)$ est bien défini car y est une fonction continue. Considérons ensuite une fonction f dans l'ensemble des paramètres admissibles \mathcal{U}_{ad} , on définit \tilde{y} comme la dérivée au sens de Gâteaux de y par rapport à f , dans la direction \tilde{f} :

$$\tilde{y} = \lim_{\theta \rightarrow 0} \frac{y(f + \theta \tilde{f}) - y(f)}{\theta}.$$

Nous pouvons alors établir la proposition suivante :

Proposition 2.4 *Si les hypothèses H1 et H2 sont vérifiées par y solution de (2.3) , et si \tilde{y} défini précédemment existe, alors \tilde{y} est solution de :*

$$(2.6) \quad a(\tilde{y}, \psi) - \int_{\Omega} f'(\bar{y}) \frac{y(M_0)\tilde{y} - y\tilde{y}(M_0)}{y(M_0)^2} \psi dx = \int_{\Omega} \tilde{f}(\bar{y}) \psi dx, \quad \forall \psi \in V^{p'_0}(\Omega).$$

Pour établir cette propriété, nous utilisons un premier lemme, dû à J. Blum [BLU 89] :

Lemme 2.5 *Si les hypothèses H1 et H2 sont satisfaites, alors l'application $y \mapsto \max_{\Omega} y$ définie sur $V^{p_0}(\Omega)$ est Gâteaux-dérivable, et sa dérivée est l'application qui à $\phi \in V^{p_0}(\Omega)$ associe la quantité $\phi(M_0)$.*

Preuve : Soit M un point dans un voisinage de M_0 . Puisque $\nabla y(M_0)$ est nul, on peut écrire $y(M) = y(M_0) + \frac{1}{2} X^t \mathcal{H} X + o(\| \overrightarrow{M_0 M} \|^2)$ avec $X = \overrightarrow{M_0 M}$ et où \mathcal{H} désigne le hessien de y en M_0 :

$$\mathcal{H} = \begin{pmatrix} \frac{\partial^2 y}{\partial x_1^2}(M_0) & \frac{\partial^2 y}{\partial x_1 x_2}(M_0) \\ \frac{\partial^2 y}{\partial x_2 x_1}(M_0) & \frac{\partial^2 y}{\partial x_2^2}(M_0) \end{pmatrix}$$

M_0 étant un point elliptique non dégénéré, la forme quadratique associée à \mathcal{H} est définie négative, d'où :

$$(2.7) \quad \exists C > 0, \forall M \text{ dans un voisinage de } M_0 : y(M) < y(M_0) - C \|\overrightarrow{M_0 M}\|^2.$$

D'autre part, d'après le lemme de Morrey [BRE 87], puisque $y \in W^{1,p_0}(\Omega)$, nous avons :

$$(2.8) \quad |y(M) - y(M_0)| \leq C_2 \|\overrightarrow{M_0 M}\|^\alpha, \text{ avec } \alpha = 1 - \frac{2}{p_0}.$$

Or, pour θ assez petit, $\max_{\bar{\Omega}}(y + \theta\phi)$ est atteint en M'_0 assez proche de M_0 : en effet, soit une suite de points (M_n) dans $\bar{\Omega}$ telle que $y(M_n) \rightarrow y(M_0)$ quand $n \rightarrow +\infty$. Alors les valeurs d'adhérence de la suite (M_n) ont leur image par y qui tend vers $y(M_0)$ (par continuité de y), et donc nécessairement $M_n \rightarrow M_0$. Comme la fonction $y + \theta\phi$ est bornée sur $\bar{\Omega}$, nous avons : $y + \theta\phi \rightarrow y$ lorsque θ tend vers 0. Par continuité : $\max(y + \theta\phi) \rightarrow \max y$, soit encore : $y(M'_0) + \theta\phi(M'_0) \rightarrow y(M_0)$. Comme ϕ est bornée, nous en déduisons : $y(M'_0) \rightarrow y(M_0)$. D'après ce qui précède, $M'_0 \rightarrow M_0$ lorsque $\theta \rightarrow 0$. D'autre part,

$$(y + \theta\phi)(M'_0) > (y + \theta\phi)(M_0)$$

soit encore :

$$y(M'_0) + \theta\phi(M'_0) > y(M_0) + \theta\phi(M_0)$$

donc, en utilisant (2.7) :

$$\theta\phi(M_0) < -C \|\overrightarrow{M_0 M'_0}\|^2 + \theta\phi(M'_0).$$

En combinant ceci avec (2.8) :

$$\|\overrightarrow{M_0 M'_0}\|^2 < \tilde{C}\theta \|\overrightarrow{M_0 M'_0}\|^\alpha,$$

soit

$$\|\overrightarrow{M_0 M'_0}\| < C_3 \theta^{\frac{1}{2-\alpha}}.$$

Comme d'autre part

$$|y(M'_0) - y(M_0)| \leq C \|\overrightarrow{M_0 M'_0}\|^2 < C_4 \theta^{\frac{2}{2-\alpha}},$$

nous avons finalement :

$$\frac{|y(M'_0) - y(M_0)|}{\theta} \leq C_4 \theta^{\frac{\alpha}{2-\alpha}}$$

et $\frac{\alpha}{2-\alpha} = \frac{p_0-2}{p_0+2} > 0$ car $p_0 > 2$, donc

$$(2.9) \quad \lim_{\theta \rightarrow 0} \frac{|y(M'_0) - y(M_0)|}{\theta} = 0$$

et

$$\max_{\Omega} (y + \theta\phi) - \max_{\Omega} y = y(M'_0) + \theta\phi(M'_0) - y(M_0),$$

donc

$$\frac{\max_{\Omega} (y + \theta\phi) - \max_{\Omega} y}{\theta} = \phi(M'_0) + \frac{y(M'_0) - y(M_0)}{\theta}.$$

En prenant la limite de ceci lorsque $\theta \rightarrow 0$, et en utilisant (2.9) il ne reste plus à droite que $\phi(M_0)$, car M'_0 tend vers M_0 et car ϕ est une fonction continue, ce qui achève la démonstration du lemme. \square

Nous démontrons à présent la dérivabilité de l'application b définie précédemment par (2.5) :

Lemme 2.6 *Sous les hypothèses H1 et H2, l'application b définie sur $V^{p_0}(\Omega)$, qui à y associe $\bar{y} = \frac{y}{\max_{\Omega} y}$ est Gâteaux-dérivable et sa dérivée est l'application : $b'(y)(\phi) = \frac{y(M_0)\phi - y\phi(M_0)}{y(M_0)^2}$ pour toute fonction ϕ appartenant à $V^{p_0}(\Omega)$.*

Preuve : Comme précédemment, donnons-nous $\phi \in V^{p_0}(\Omega)$, on considère M'_0 tel que :

$$(y + \theta\phi)(M'_0) = \max_{\Omega} (y + \theta\phi)$$

qui existe pour θ assez petit (compacité de $\bar{\Omega}$), et donc $M'_0 \rightarrow M_0$ lorsque $\theta \rightarrow 0$. Nous avons alors :

$$\begin{aligned} b(y + \theta\phi) - b(y) &= \frac{y + \theta\phi}{y(M'_0) + \theta\phi(M'_0)} - \frac{y}{y(M_0)} \\ &= \frac{y(M_0)y - yy(M'_0)}{y(M_0)(y(M'_0) + \theta\phi(M'_0))} + \theta \frac{\phi y(M_0) - y\phi(M'_0)}{y(M_0)(y(M'_0) + \theta\phi(M'_0))}, \end{aligned}$$

d'où :

$$(2.10) \quad \begin{aligned} \frac{b(y + \theta\phi) - b(y)}{\theta} &= \frac{y}{y(M_0)(y(M'_0) + \theta\phi(M'_0))} \times \frac{y(M_0) - y(M'_0)}{\theta} \\ &\quad + \frac{\phi y(M_0) - y\phi(M'_0)}{y(M_0)(y(M'_0) + \theta\phi(M'_0))}. \end{aligned}$$

Puisque les fonctions ϕ et y sont continues, et que comme on l'a vu précédemment

$$\frac{y(M_0) - y(M'_0)}{\theta} \xrightarrow{\theta \rightarrow 0} 0,$$

par passage à la limite dans (2.10), on obtient le résultat cherché. \square

Nous pouvons désormais démontrer la proposition 2.4 :

Preuve de la proposition : Nous allons utiliser les deux lemmes précédents. Considérons la formulation faible au point $y(f)$:

$$a(y(f), \psi) = \int_{\Omega} f ob(y(f)) \psi dx, \quad \forall \psi \in V^{p_0}(\Omega)$$

et au point $y(f + \theta \tilde{f})$:

$$a(y(f + \theta \tilde{f}), \psi) = \int_{\Omega} (f + \theta \tilde{f}) ob(y(f + \theta \tilde{f})) \psi dx, \quad \forall \psi \in V^{p_0}(\Omega).$$

Prenons la différence de ces deux expressions, divisons par θ et posons

$$\tilde{y}_{\theta} = \frac{y(f + \theta \tilde{f}) - y(f)}{\theta} \in V^{p_0}(\Omega),$$

alors :

$$(2.11) \quad a(\tilde{y}_{\theta}, \psi) = \int_{\Omega} \left(\frac{f ob(y(f + \theta \tilde{f})) - f ob(y(f))}{\theta} \right) \psi dx + \int_{\Omega} \tilde{f} ob(y(f + \theta \tilde{f})) \psi dx.$$

Or, si M_{θ} est tel que : $(y + \theta \tilde{y}_{\theta})(M_{\theta}) = \max_{\Omega} (y + \theta \tilde{y}_{\theta})$, pour θ assez petit, et en posant $y = y(f)$, nous avons :

$$\begin{aligned} b(y + \theta \tilde{y}_{\theta}) &= \frac{y + \theta \tilde{y}_{\theta}}{y(M_{\theta}) + \theta \tilde{y}_{\theta}(M_{\theta})} \\ &= b(y) + \frac{y + \theta \tilde{y}_{\theta}}{y(M_{\theta}) + \theta \tilde{y}_{\theta}(M_{\theta})} - \frac{y}{y(M_0)} \\ &= b(y) + \frac{y(M_0) - y(M_{\theta})}{y(M_0)(y(M_{\theta}) + \theta \tilde{y}_{\theta}(M_{\theta}))} y + \theta \frac{\tilde{y}_{\theta} y(M_0) - y \tilde{y}_{\theta}(M_{\theta})}{y(M_0)(y(M_{\theta}) + \theta \tilde{y}_{\theta}(M_{\theta}))}, \end{aligned}$$

d'où nous déduisons : $f ob(y + \theta \tilde{y}_{\theta}) = f(b(y) + T)$ où T est un terme qui tend vers 0 lorsque $\theta \rightarrow 0$, donc puisque f est de classe C^2 sur $[0, 1]$:

$$\begin{aligned} \frac{f(b(y + \theta \tilde{y}_{\theta})) - f(b(y))}{\theta} &= \left(\frac{y(M_0) - y(M_{\theta})}{\theta} \times \frac{y}{y(M_0)(y(M_{\theta}) + \theta \tilde{y}_{\theta}(M_{\theta}))} \right. \\ &\quad \left. + \frac{\tilde{y}_{\theta} y(M_0) - y \tilde{y}_{\theta}(M_{\theta})}{y(M_0)(y(M_{\theta}) + \theta \tilde{y}_{\theta}(M_{\theta}))} \right) f'(b(y)) + O(\theta). \end{aligned}$$

Le premier terme tend vers 0, car $\frac{y(M_0) - y(M_\theta)}{\theta} \rightarrow 0$, et le second :

$$\begin{aligned} \tilde{y}_\theta y(M_0) - y\tilde{y}_\theta(M_\theta) - (\tilde{y}y(M_0) - y\tilde{y}(M_0)) \\ &= (\tilde{y}_\theta - \tilde{y})y(M_0) + y(\tilde{y}(M_0) - \tilde{y}_\theta(M_\theta)) \\ &= (\tilde{y}_\theta - \tilde{y})y(M_0) + y(\tilde{y}(M_0) - \tilde{y}(M_\theta) + \tilde{y}(M_\theta) - \tilde{y}_\theta(M_\theta)) \end{aligned}$$

et donc :

$$\begin{aligned} &\| \tilde{y}_\theta y(M_0) - y\tilde{y}_\theta(M_\theta) - (\tilde{y}y(M_0) - y\tilde{y}(M_0)) \| \\ &\leq \| \tilde{y}_\theta - \tilde{y} \| \| y \|_\infty + \| y \| (|\tilde{y}(M_0) - \tilde{y}(M_\theta)| + \| \tilde{y} - \tilde{y}_\theta \|_\infty), \end{aligned}$$

soit encore, puisque la norme $\| \cdot \|_\infty$ est majorée par la norme dans $W^{1,p_0}(\Omega)$:

$$\begin{aligned} &\| \tilde{y}_\theta y(M_0) - y\tilde{y}_\theta(M_\theta) - (\tilde{y}y(M_0) - y\tilde{y}(M_0)) \| \\ &\leq C \| \tilde{y}_\theta - \tilde{y} \| \| y \| + \| y \| (|\tilde{y}(M_0) - \tilde{y}(M_\theta)| + C \| \tilde{y} - \tilde{y}_\theta \|). \end{aligned}$$

Puisque \tilde{y} est continu, et que $M_\theta \rightarrow M_0$, nous déduisons de ceci que :

$$\lim_{\theta \rightarrow 0} (\tilde{y}_\theta y(M_0) - y\tilde{y}_\theta(M_\theta)) = \tilde{y}y(M_0) - y\tilde{y}(M_0)$$

et donc :

$$\begin{aligned} \lim_{\theta \rightarrow 0} \frac{f(b(y + \theta\tilde{y}_\theta)) - f(b(y))}{\theta} &= f'(b(y)) \frac{\tilde{y}y(M_0) - y\tilde{y}(M_0)}{y(M_0)^2} \\ &= f'(b(y))(b'(y)\tilde{y}). \end{aligned}$$

En utilisant ce résultat, et en prenant la limite dans (2.11) lorsque θ tend vers 0, nous obtenons bien la formulation annoncée :

$$a(\tilde{y}, \psi) - \int_{\Omega} f'(b(y))(b'(y)\tilde{y})\psi dx = \int_{\Omega} \tilde{f}(b(y))\psi dx, \quad \forall \psi \in V^{p_0}(\Omega).$$

□

Remarque 2.7 : La linéarisation du problème direct (2.1) s'écrit :

$$(2.12) \quad \begin{cases} -\Delta \tilde{y} - f'(\bar{y}) \frac{y(M_0)\tilde{y} - \tilde{y}(M_0)y}{y(M_0)^2} = \tilde{f}(\bar{y}) & \text{dans } \Omega \\ \tilde{y} = 0 & \text{sur } \partial\Omega \end{cases}$$

et la formulation (2.6) correspond en fait à une formulation "faible" de (2.12).

2.1.4 Méthodes numériques de résolution

- *Approximation par éléments finis*

On introduit une triangulation \mathcal{T}_h de Ω par des triangles de diamètre inférieur ou égal à h :

$$\bar{\Omega} = \bigcup_{T \in \mathcal{T}_h} T$$

avec la condition que si T_1 et T_2 sont deux triangles distincts de la triangulation, alors :

- soit $T_1 \cap T_2 = \emptyset$,
- soit T_1 et T_2 ont un côté en commun,
- soit T_1 et T_2 ont un sommet en commun.

On définit l'espace associé à cette triangulation :

$$V_h = \{ \phi \in C^0(\bar{\Omega}) / \forall T \in \mathcal{T}_h, \phi|_T \in P_1 \text{ et } \phi = 0 \text{ sur } \partial\Omega \}$$

où P_1 représente l'espace vectoriel des polynômes de degré 1 à coefficients réels constants en x_1 et x_2 . Cet espace V_h approche l'espace $V^{p_0}(\Omega)$ pour $p_0 > 2$ (consulter par exemple [CR 72]).

Soit γ_h l'ensemble des points de la triangulation appartenant au bord $\partial\Omega$ et ω_h son complémentaire dans l'ensemble des points de la triangulation \mathcal{T}_h . Soit $N_h = \text{Card}(\omega_h)$, tout élément ϕ de V_h est déterminé par ses valeurs aux N_h sommets M_j de ω_h . Pour tout sommet M_j de ω_h , on définit la i -ème fonction $\phi_i \in V_h$ par :

$$\phi_i(M_j) = \delta_{ij}, \quad 1 \leq i, j \leq N_h.$$

L'ensemble de ces N_h fonctions ϕ_i constitue une base de V_h .

La discrétisation du problème faible (2.3) va s'écrire sous la forme :

Trouver $y_h \in V_h$ qui satisfait le système discret :

$$(2.13) \quad a(y_h, \phi) = \int_{\Omega} f(\bar{y}_h) \phi dx, \quad \forall \phi \in V_h.$$

La méthode d'intégration utilisée pour résoudre (2.13) consiste à prendre pour chaque intégrale sur un triangle la mesure de ce triangle multipliée par la valeur de la fonction à intégrer au centre de gravité du triangle :

$$\int_{T_k} g(x) dx \simeq \text{mes}(T_k) g(G_k) \quad \text{où } G_k = \text{centre de gravité de } T_k.$$

Cette méthode est compatible avec les éléments P_1 qui sont linéaires sur chaque triangle.

La difficulté essentielle pour la résolution de (2.13) réside dans la non linéarité $f(\bar{y}_h)$. Nous allons donc remplacer le problème (2.13) par une suite de problèmes linéaires. Nous étudions un algorithme de point fixe (Picard) puis un autre de type Newton-Raphson.

• *Algorithme de type Picard*

Supposons que le problème à résoudre s'écrive $\mathcal{L}\psi = F(\psi)$ où \mathcal{L} est un opérateur linéaire et F est non linéaire. Alors l'algorithme de point fixe va s'écrire :

- Choix de ψ^0 arbitraire,
- Calcul de ψ^{n+1} par : $\mathcal{L}(\psi^{n+1}) = F(\psi^n)$.

Si l'opérateur dérivé de F existe, alors une condition nécessaire de convergence est que le spectre de l'opérateur $\Phi(\psi) = \mathcal{L}^{-1}F'(\psi)$ soit contenu dans la boule unité.

Dans notre cas, \mathcal{L} est l'opérateur "laplacien", et $F = f \circ b$ dont la dérivée existe au sens de Gâteaux.

• *Algorithme de Newton-Raphson*

Supposons que le problème à résoudre s'écrive $\mathcal{L}\psi = F(\psi)$ où \mathcal{L} est un opérateur linéaire et F est non linéaire. Alors l'algorithme de Newton s'écrit à l'étape n :

$$(\mathcal{L} - F'(\psi^n))\psi^{n+1} = F(\psi^n) - F'(\psi^n)\psi^n.$$

On devra donc utiliser la linéarisation du problème discret (2.13) ; pour réaliser ceci, nous allons procéder de la même façon que pour la linéarisation du problème au paragraphe 2.1.3 . Pour pouvoir appliquer l'algorithme de Newton, il est nécessaire que la linéarisation de (2.13) existe et soit inversible pour tout itéré ψ^n . Tout d'abord, regardons ce que deviennent les hypothèses H1 et H2 dans le cas du problème discret. On utilise le fait que y_h est linéaire et ∇y_h constant sur chaque triangle :

H') $\max_{\Omega} y_h$ est atteint en un unique point M_h de la triangulation \mathcal{T}_h intérieur à Ω .

Supposons que la solution y_h du problème discret (2.13) satisfasse l'hypothèse H'), et que la matrice du système linéaire soit inversible. Dans ce cas, à cause de la méthode d'éléments finis choisie, ces hypothèses sont encore satisfaites dans un voisinage de la solution y_h dans l'espace V_h , et donc l'opérateur linéarisé existe et est inversible dans ce voisinage. Ceci n'est pas vrai dans le cas du problème continu, où l'on a la dérivabilité en la solution du problème mais pas dans un voisinage.

L'algorithme de Newton va donc s'écrire :

- Choix de $y^0 \in V_h$,
- y^n étant donné dans V_h , soit $M^n = M_{i_n}$ le point où y^n atteint son maximum. On calcule y^{n+1} dans V_h par :

$$(2.14) \quad \begin{aligned} a(y^{n+1}, \phi) - \int_{\Omega} f(\bar{y}^n) \frac{y^{n+1}y^n(M^n) - y^n y^{n+1}(M^n)}{y^n(M^n)^2} \phi dx \\ = \int_{\Omega} f(\bar{y}^n) \phi dx, \quad \forall \phi \in V_h. \end{aligned}$$

Cette équation est donc équivalente à un système linéaire dont les inconnues sont les N_h valeurs de y^{n+1} aux points de la triangulation. Soit K^n la matrice associée, on peut la décomposer en une somme $K_0^n + K_1^n$ où K_0^n est la partie symétrique et K_1^n contient la partie antisymétrique de K^n :

$$(K_0^n)_{ij} = a(\phi_i, \phi_j) - \int_{\Omega} f'(\bar{y}^n) \frac{\phi_i \phi_j}{y^n(M^n)} dx$$

$$(K_1^n)_{ij} = \begin{cases} \int_{\Omega} f'(\bar{y}^n) \frac{y^n \phi_i}{y^n(M^n)^2} dx, & \text{si } j = i_n \\ 0 & \text{sinon.} \end{cases}$$

On voit donc que K_1^n est une matrice très "creuse". Pour inverser K^n , on va donc utiliser la formule de Woodbury (cf [HOU 75]) :

$$(K_0 + uv^t)^{-1} = \left(I - (1 + v^t K_0^{-1} u)^{-1} K_0^{-1} uv^t \right) K_0^{-1}$$

où K_0 est une matrice symétrique, u et v sont des vecteurs colonnes de \mathbb{R}^{N_h} . Pour résoudre le système (2.14), nous écrivons donc $K_1^n = u_n v_n^t$ et nous allons résoudre successivement les systèmes linéaires :

$$\begin{cases} K_0^n z_n = b_n \\ K_0^n w_n = u_n \text{ puis nous aurons :} \\ y^{n+1} = \left(I - \frac{w_n v_n^t}{1 + v_n^t w_n} w_n v_n^t \right) z_n. \end{cases}$$

Les deux premiers systèmes linéaires sont résolus en utilisant la décomposition de Crout LDL^t de K_0^n . De la sorte, le coût de l'algorithme à chaque itération est principalement celui de la décomposition LDL^t .

2.1.5 Résolution sur un domaine quelconque de \mathbb{R}^2

Afin d'avoir un programme indépendant du domaine Ω de \mathbb{R}^2 considéré, nous avons tout d'abord supposé qu'il était régulier, simplement connexe, de frontière $\partial\Omega$ de classe \mathcal{C}^2 . Ainsi, l'existence et l'unicité de la transformation conforme bijective de Riemann R , qui envoie Ω sur le disque unité $D(0,1)$ de \mathbb{R}^2 et le point $(0,0)$ sur lui-même avec $R'(0) > 0$ sont assurées. Lorsque Ω n'a pas une forme géométrique simple (cercle, ellipse, ...), le calcul de R peut se révéler un peu ardu.

Nous utilisons ici les travaux réalisés par N. Kerzman et M. Trummer [KT 86], dont nous rappelons les points essentiels, qui vont nous permettre de faire ce calcul pour tout domaine de \mathbb{R}^n dont la frontière est au moins de classe \mathcal{C}^2 , pourvu que l'on dispose d'une paramétrisation de sa frontière.

Rappelons tout d'abord la définition du noyau de Cauchy :

$$H(w, z) = \frac{1}{2i\pi} \frac{\tau(z)}{z - w}, \quad \forall z \in \partial\Omega, \forall w \in \Omega,$$

où $\tau(z)$ est le vecteur unitaire tangent à $\partial\Omega$ au point z . A ce noyau est associé la transformation de Hilbert :

$$\mathbb{H}u(w) = \int_{\partial\Omega} H(w, z)u(z)ds, \quad \forall w \in \Omega, \forall u \in L^2(\partial\Omega)$$

où s désigne l'abscisse curviligne. On définit ensuite $\gamma(s)$ la paramétrisation de $\partial\Omega$ par l'abscisse curviligne, et pour tout $\varepsilon > 0$ assez petit :

$$\gamma_\varepsilon(s) = \gamma(s) + \varepsilon i\gamma'(s),$$

$i\gamma'(s)$ étant le vecteur normal unitaire rentrant à $\partial\Omega$. Soit $\mathcal{H}(\Omega)$ l'espace des fonctions holomorphes sur Ω muni de la norme de la convergence uniforme sur les compacts, on peut montrer :

Propriété 2.8 Pour toute fonction f de $\mathcal{H}(\Omega)$, il y a équivalence entre les propriétés suivantes :

- i) f_ε converge dans $L^2(\partial\Omega)$ vers une fonction f_0
- ii) $\|f_\varepsilon\|_{L^2(I)}$ reste bornée lorsque ε tend vers 0
- iii) Il existe une suite ε_p de nombres strictement positifs tendant vers 0 telle que la suite $\|f_{\varepsilon_p}\|_{L^2(I)}$ soit bornée
- iv) Il existe $u \in L^2(\partial\Omega)$ tel que $f = \mathbb{H}u$

On a alors $f = \mathbb{H}f_0$.

Cette propriété, dont nous ne donnons pas la démonstration qui se trouve dans [KT 86], nous permet de rappeler la définition de l'espace de Hardy $\mathcal{H}^2(\partial\Omega)$ qui est l'ensemble des fonctions de $\mathcal{H}(\Omega)$ vérifiant l'une de ces propriétés équivalentes i), ii), iii), iv), muni de la norme :

$$\|f\|_{\mathcal{H}^2} = \|f_0\|_{L^2(\partial\Omega)}.$$

Le projecteur de Szegö est, par définition, le projecteur orthogonal \mathcal{S} de $L^2(\partial\Omega)$ sur $\mathcal{H}^2(\partial\Omega)$. Le noyau S qui lui est associé est appelé noyau de Szegö de Ω . N. Kerzman a montré que l'on peut retrouver \mathcal{S} à partir du projecteur oblique \mathcal{H} : comme \mathcal{S} et \mathcal{H} coïncident avec l'identité sur $\mathcal{H}^2(\partial\Omega)$, on a :

$$\mathcal{S}\mathcal{H} = \mathcal{H}, \quad \mathcal{H}\mathcal{S} = \mathcal{S}.$$

Comme $\mathcal{S}^* = \mathcal{S}$, on trouve $\mathcal{S}\mathcal{H}^* = \mathcal{S}$ et :

$$\mathcal{S}(\mathcal{H}^* - \mathcal{H}) = \mathcal{S} - \mathcal{H},$$

soit encore :

$$\mathcal{S}(\mathcal{I} - \mathcal{A}) = \mathcal{H}, \quad \text{avec } \mathcal{A} = \mathcal{H}^* - \mathcal{H}.$$

On montre ensuite que l'opérateur $\mathcal{I} - \mathcal{A}$ est un isomorphisme de $L^2(\partial\Omega)$ et on peut calculer \mathcal{S} par la formule :

$$\mathcal{S} = \mathcal{H}(\mathcal{I} - \mathcal{A})^{-1},$$

soit ensuite S le noyau associé :

$$(2.15) \quad S(w, z) = H(w, z) + \int_{\partial\Omega} S(w, \xi) A(\xi, z) dx(\xi), \quad w \in \Omega, \quad z \in \partial\Omega$$

et sur le disque, on a : $\mathcal{S}^D = \mathcal{H}^D$.

Par le changement de variable $z = R(w)$, on obtient :

$$\int_{\partial D} |u(z)|^2 |dz| = \int_{\partial\Omega} |u \circ R(w)|^2 |R'(w)| |dw|, \quad \forall u \in L^2(\partial\Omega).$$

On voit donc qu'on a une isométrie

$$\begin{aligned} L^2(\partial D) &\rightarrow L^2(\partial\Omega) \\ u &\mapsto u \circ R(R')^{\frac{1}{2}}, \end{aligned}$$

où $(R')^{\frac{1}{2}}$ est la détermination de la racine carrée complexe telle que $R'(0)^{\frac{1}{2}} > 0$ (on notera que R' ne s'annule pas sur $\partial\Omega$). L'isométrie envoie $\mathcal{H}(D) \cap C^0(\bar{D})$ sur

$\mathcal{H}(\Omega) \cap C^0(\bar{\Omega})$ et par densité, puisque l'ensemble des fonctions holomorphes sur un voisinage de $\bar{\Omega}$ est dense dans $\mathcal{H}^2(\partial\Omega)$, elle envoie $\mathcal{H}^2(\partial D)$ sur $\mathcal{H}^2(\partial\Omega)$.

Si (ψ_j) est une base orthonormée de $\mathcal{H}^2(\partial D)$ alors $(\psi_j \circ R(R')^{\frac{1}{2}})$ est une base orthonormée de $\mathcal{H}^2(\partial\Omega)$ et par conséquent :

$$S(w, z) = S^D(R(w), R(z)) R'(w)^{\frac{1}{2}} \overline{R'(z)^{\frac{1}{2}}}$$

où S^D est le noyau de Szegö du disque unité D . Soit encore :

$$(2.16) \quad S(w, z) = \frac{1}{2\pi} \frac{R'(w)^{\frac{1}{2}} \overline{R'(z)^{\frac{1}{2}}}}{1 - R(w)R(z)}.$$

On peut donc en déduire les formules suivantes, en prenant $z = 0$ dans (2.16)

$$(2.17) \quad \begin{cases} R'(w) = 2\pi \frac{S(w, 0)^2}{S(0, 0)}, & \forall w \in \bar{\Omega} \\ R(w) = -i\tau(w) \frac{S(w, 0)^2}{|S(w, 0)|^2}, & \forall w \in \partial\Omega. \end{cases}$$

• *Calcul numérique de l'application conforme*

La formule intégrale (2.15) donne :

$$S(0, w) = H(0, w) + \int_{\partial\Omega} S(0, \xi) A(\xi, w) ds(\xi), \quad \forall w \in \partial\Omega.$$

Comme $\overline{A(z, w)} = -A(w, z)$, on trouve, après conjugaison :

$$(2.18) \quad S(w, 0) + \int_{\partial\Omega} A(w, z) S(z, 0) ds = \overline{H(0, w)}, \quad \forall w \in \partial\Omega,$$

ce qui permet d'obtenir $S(w, 0)$ en fonctions des noyaux H et A explicitement connus. On utilise ensuite les formules (2.17) pour calculer l'application de Riemann R à partir de S .

Les intégrales $\int_{\alpha}^{\beta} f(t) dt$ mises en jeu seront évaluées par la méthode des trapèzes relative à une subdivision $t_i = \alpha + ih, 0 \leq i \leq N$, de pas constant $h = \frac{\beta - \alpha}{N}$:

$$\int_{\alpha}^{\beta} f(t) dt \simeq h \left(\frac{f(\alpha)}{2} + f(t_1) + \dots + f(t_{N-1}) + \frac{f(\beta)}{2} \right).$$

L'erreur est en général de l'ordre de $O(h^2) = O(N^{-2})$. Comme les intégrales à évaluer sont des intégrales de fonctions périodiques, la convergence de la méthode des trapèzes est extrêmement rapide, tout au moins dans le cas des fonctions C^∞ . Cette remarque montre que l'on n'a pas intérêt à calculer R par intégration de la première formule de (2.17) mais plutôt à partir de la seconde et de la formule de Cauchy :

$$R(w) = \frac{1}{2i\pi} \int_{\partial\Omega} \frac{R(z)dz}{z-w}.$$

Supposons donnée une paramétrisation quelconque $t \in [\alpha, \beta] \mapsto z(t)$ de la courbe $\partial\Omega$. En multipliant (2.18) par $|z'(t)|^{\frac{1}{2}}$ après avoir substitué $z(t)$ à w et $z(u)$ à z , on obtient la relation équivalente :

$$\sigma(t) + \int_{\alpha}^{\beta} a(t, u)\sigma(u)du = g(t)$$

avec les notations :

$$\sigma(t) = |z'(t)|^{\frac{1}{2}} S(z(t), 0),$$

$$g(t) = |z'(t)|^{\frac{1}{2}} \overline{H(0, z(t))},$$

$$a(t, u) = |z'(t)|^{\frac{1}{2}} |z'(u)|^{\frac{1}{2}} A(z(t), z(u)).$$

Cette écriture a l'avantage de préserver le caractère hermitien antisymétrique du noyau $a(t, u)$. Les fonctions $a(t, u)$ et $g(t)$ sont données par les formules explicites :

$$g(t) = \frac{1}{2i\pi} \left(\frac{z'(t)}{-z(t)} \right) |z'(t)|^{-\frac{1}{2}},$$

$$h(t, u) = \frac{1}{2i\pi} \frac{z'(u)}{z(u) - z(t)} |z'(t)|^{\frac{1}{2}} |z'(u)|^{-\frac{1}{2}},$$

$$a(t, u) = \begin{cases} \overline{h(u, t)} - h(t, u) & \text{si } t \neq u, \\ 0 & \text{si } t = u. \end{cases}$$

L'utilisation de la méthode des trapèzes conduit à résoudre le système linéaire suivant :

$$\sigma(t_i) + h \sum_{0 \leq j < N} a(t_i, t_j)\sigma(t_j) = g(t_j), \quad 0 \leq i < N.$$

Ce système est de rang N car la matrice antisymétrique $a(t_i, t_j)$ a toutes ses valeurs propres imaginaires. La résolution du système fournit les valeurs $\sigma(t_j)$ cherchées, par exemple à l'aide d'une méthode itérative de Gauss-Seidel. Une fois les valeurs $\sigma(t_j)$ connues, on obtient :

$$R(z(t_j)) = -i \frac{z'(t_j)}{|z'(t_j)|} \frac{\sigma(t_j)^2}{|\sigma(t_j)^2|},$$

$$R(w) \simeq \frac{h}{2i\pi} \sum_{0 \leq j < N} \frac{z'(t_j)}{z(t_j) - w} R(\bar{z}(t_j)).$$

Tous ces calculs sont immédiats dès lors que la paramétrisation $z(t)$ de $\partial\Omega$ est connue aux points $t = t_j$ ainsi que sa dérivée $z'(t)$. L'application de Riemann inverse

$$Q = R^{-1}, \quad Q : D \rightarrow \Omega, \quad Q(0) = 0$$

peut être évaluée comme suit : la formule des résidus implique :

$$Q'(w) = \frac{1}{R'(R^{-1}(w))} = \frac{1}{2i\pi} \int_{\partial\Omega} \frac{dz}{R(z) - w},$$

$$Q(w) = \int_0^w Q'(v) dv = -\frac{1}{2i\pi} \int_{\partial\Omega} \log(1 - w\overline{R(z)}) dz,$$

le développement en série entière du logarithme fournit alors :

$$Q(w) = \sum_{k=0}^{+\infty} c_k w^k$$

avec $c_0 = 0$ et

$$\begin{aligned} c_k &= \frac{1}{2i\pi} \int_{\alpha}^{\beta} \frac{1}{k} \overline{R(z(t))}^k z'(t) dt \\ &\simeq \frac{h}{2i\pi k} \sum_{0 \leq j < N} \overline{R(z(t_j))}^k z'(t_j), \quad k \geq 1. \end{aligned}$$

En pratique, on calculera les coefficients c_k pour $0 \leq k < N$ et le polynôme $\sum_{0 \leq k < N} c_k w^k$ servira à évaluer $Q(w)$, tout au moins si $|w|$ n'est pas trop voisin de

1. Pour $w = e^{i\theta} \in \partial D$, il est préférable de chercher $Q(w) = z(t)$ grâce à une interpolation sur quelques points de la correspondance $\theta_j \rightarrow t_j$ où $R(z(t_j)) = e^{i\theta_j}$.

Notre problème direct (2.1) , moyennant le changement de variable $y = w \circ R$ va devenir :

$$-\Delta y = -\Delta(w \circ R)$$

et

$$\begin{aligned} \Delta(w \circ R)(x_1, x_2) &= \Delta w(X_1, X_2) \left(\left(\frac{\partial R_1}{\partial x_1} \right)^2 + \left(\frac{\partial R_2}{\partial x_2} \right)^2 \right) \\ &+ 2 \frac{\partial^2 w}{\partial X_1 \partial X_2} \left(\frac{\partial R_1}{\partial x_1} \frac{\partial R_2}{\partial x_1} + \frac{\partial R_1}{\partial x_2} \frac{\partial R_2}{\partial x_2} \right) + \frac{\partial w}{\partial X_1} \Delta R_1 + \frac{\partial w}{\partial X_2} \Delta R_2. \end{aligned}$$

Or, d'après les relations de Cauchy, nous avons :

$$\begin{cases} \frac{\partial R_1}{\partial x_1} = \frac{\partial R_2}{\partial x_2} \\ \frac{\partial R_1}{\partial x_2} = -\frac{\partial R_2}{\partial x_1} \end{cases}$$

et donc en particulier, les composantes R_1 et R_2 sont harmoniques :

$$\Delta R_1 = 0, \quad \Delta R_2 = 0.$$

Finalement :

$$\Delta(w \circ R)(x_1, x_2) = \|\nabla R_1\|^2 \Delta w.$$

Notons C_Ω^2 le carré de la norme du gradient de Q_1 ou Q_2 :

$$C_\Omega^2 = \|\nabla Q_1\|^2 = \|\nabla Q_2\|^2 = \frac{1}{\|\nabla R_1\|^2} = \frac{1}{\|\nabla R_2\|^2},$$

nous obtenons alors :

$$(2.19) \quad \begin{cases} -\Delta w = C_\Omega^2 f(\bar{w}) & \text{dans } D(0, 1) \\ w = 0 & \text{sur } \mathcal{C}(0, 1). \end{cases}$$

Nous pouvons de la même façon calculer la dérivée normale de w sur le cercle unité. Pour cela, nous nous donnons $t \mapsto (x_1(t), x_2(t))$ une paramétrisation de $\partial\Omega$. Le vecteur unitaire normal s'écrit alors :

$$\vec{n}(t) = \frac{1}{\sqrt{x_1'(t)^2 + x_2'(t)^2}} \begin{pmatrix} -x_2'(t) \\ x_1'(t) \end{pmatrix}.$$

D'autre part, nous écrivons :

$$\begin{aligned} \nabla y(x_1, x_2) &= \begin{pmatrix} \frac{\partial R_1}{\partial x_1} & \frac{\partial R_2}{\partial x_1} \\ \frac{\partial R_1}{\partial x_2} & \frac{\partial R_2}{\partial x_2} \end{pmatrix} \nabla w(X_1, X_2) \\ &= \frac{1}{\|\nabla Q_1\|^2} \begin{pmatrix} \frac{\partial Q_1}{\partial X_1} & \frac{\partial Q_1}{\partial X_2} \\ -\frac{\partial Q_1}{\partial X_2} & \frac{\partial Q_1}{\partial X_1} \end{pmatrix} \nabla w(X_1, X_2) \end{aligned}$$

en vertu des relations de Cauchy. Soit $\vec{N}(t)$ le vecteur unitaire normal au cercle unité de \mathbb{R}^2 . Si $(X_1(t), X_2(t))$ est une paramétrisation du cercle, nous avons :

$$\vec{N}(t) = \frac{1}{\sqrt{X_1'(t)^2 + X_2'(t)^2}} \begin{pmatrix} -X_2'(t) \\ X_1'(t) \end{pmatrix}.$$

En écrivant que

$$\begin{cases} x_1'(t) &= \frac{\partial Q_1}{\partial X_1} X_1'(t) + \frac{\partial Q_1}{\partial X_2} X_2'(t) \\ x_2'(t) &= -\frac{\partial Q_1}{\partial X_2} X_1'(t) + \frac{\partial Q_1}{\partial X_1} X_2'(t) \end{cases}$$

nous en déduisons :

$$\vec{n}(t) = \frac{1}{\|\nabla Q_1\|} \times \frac{1}{\sqrt{X_1'(t)^2 + X_2'(t)^2}} \begin{pmatrix} \frac{\partial Q_1}{\partial X_2} X_1'(t) - \frac{\partial Q_1}{\partial X_1} X_2'(t) \\ \frac{\partial Q_1}{\partial X_1} X_1'(t) + \frac{\partial Q_1}{\partial X_2} X_2'(t) \end{pmatrix}.$$

Et donc :

$$\nabla y(x_1, x_2) \cdot \vec{n}(x_1, x_2) = \frac{1}{\|\nabla Q_1\|^2} \vec{n}^t \begin{pmatrix} \frac{\partial Q_1}{\partial X_1} & \frac{\partial Q_1}{\partial X_2} \\ -\frac{\partial Q_1}{\partial X_2} & \frac{\partial Q_1}{\partial X_1} \end{pmatrix} \nabla w(X_1, X_2).$$

D'où finalement :

$$\begin{aligned} \nabla y(x_1, x_2) \cdot \vec{n}(x_1, x_2) &= \frac{1}{\|\nabla Q_1\|} \vec{N}(X_1, X_2)^t \nabla w(X_1, X_2) \\ &= \frac{1}{\|\nabla Q_1\|} \nabla w(X_1, X_2) \cdot \vec{N}(X_1, X_2) \end{aligned}$$

et donc

$$\frac{\partial y}{\partial n}(x_1, x_2) = \frac{1}{\|\nabla Q_1\|} \frac{\partial w}{\partial N}(X_1, X_2).$$

Nous considérerons donc toujours une triangulation \mathcal{T}_h du disque $D(0, 1)$ et, par le seul calcul de la fonction C_Ω^2 , le problème direct sera résolu sur un ouvert Ω quelconque, de classe C^2 de \mathbb{R}^2 . Nous donnons une illustration du maillage utilisé avec la figure 2.7 .

2.2 Le problème de contrôle

Il s'agit d'identifier la source non linéaire f sur l'intervalle $[0, 1]$ à partir d'observations sur le bord du domaine: $z = \left(\frac{\partial y}{\partial n}\right)_{/\partial\Omega}$ où y est lié à f par les équations d'état (2.1) . Si l'on se ramène au disque unité D , les observations deviennent $z_D = (\nabla w \cdot \vec{N})_{/C} = \frac{1}{C_\Omega} \frac{\partial w}{\partial N}_{/C} \in L^2(C)$.

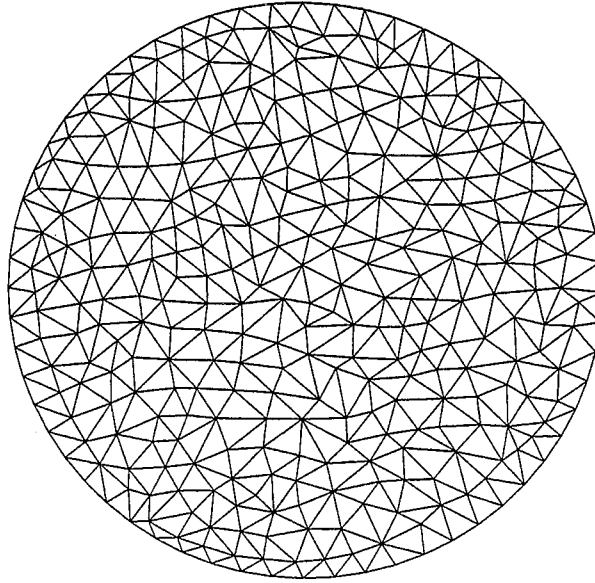


Figure 2.7 Maillage du disque unité

2.2.1 Les paramètres de contrôle

Afin de nous ramener en dimension finie, nous supposons que la source f est assez "régulière", de sorte à avoir :

$$f(x) = \sum_{i=-3}^{m-1} \alpha_i N_i^4(x)$$

où les N_i^4 sont les B -splines cubiques définies au paragraphe 1.3.1. Il s'agit donc d'identifier le vecteur $\alpha \in \mathbb{R}^{m+3}$ qui déterminera la source de façon unique dans cette base de B -splines. Le vecteur α est donc le vecteur de contrôle soumis à la contrainte suivante, qui nous assure que la source est nulle sur le bord du domaine :

$$\begin{aligned} f(0) = 0 &\iff \sum_{i=-3}^{m-1} \alpha_i N_i^4(0) = 0 \\ &\iff \alpha_{-3} N_{-3}^4(0) + \alpha_{-2} N_{-2}^4(0) + \alpha_{-1} N_{-1}^4(0) = 0. \end{aligned}$$

On peut donc exprimer α_{-3} en fonction de α_{-2} et α_{-1} :

$$\frac{\alpha_{-3}}{6} = \frac{-2}{3} \alpha_{-2} - \frac{1}{6} \alpha_{-1}$$

d'où :

$$f(x) = \sum_{i=-2}^{m-1} \alpha_i N_i^4(x) - \frac{2}{3} \alpha_{-2} N_{-2}^4(x) - \frac{1}{6} \alpha_{-1} N_{-1}^4(x).$$

On définit alors de nouvelles fonctions de base : les $(\Phi_i)_{i=-2, \dots, m-1}$ par :

$$\begin{cases} \Phi_{-2} &= N_{-2}^4 - \frac{2}{3} N_{-3}^4 \\ \Phi_{-1} &= N_{-1}^4 - \frac{1}{6} N_{-3}^4 \\ \Phi_i &= N_i^4 \quad \forall i = 0, \dots, m-1. \end{cases}$$

Et donc

$$f(x) = \sum_{i \in I} \beta_i \Phi_i(x)$$

avec $I = \{-2, -1, 0, \dots, m-1\}$ et l'ensemble des contrôles admissibles \mathcal{U}_{ad} sera le sous-espace convexe fermé de \mathbb{R}^{m+2} des combinaisons linéaires de ces fonctions Φ_i .

2.2.2 Les équations d'état

Le vecteur d'état est $y \in V^{p_0}(\Omega)$, où $V^{p_0}(\Omega)$ est défini par (2.2). Nous écrivons que $y = w \circ Q$ avec $w \in V^{p_0}(D)$, où D est le disque unité de \mathbb{R}^2 ; les équations d'état (2.3) s'écrivent sur D en termes dépendant des paramètres de contrôle β_i :

$$(2.20) \quad a(w, \phi) = \sum_{i \in I} \beta_i \int_D C_{\Omega}^2 \Phi_i(\bar{w}) \phi dx, \quad \forall \phi \in V^{p_0}(D)$$

avec $a(w, \phi) = \int_D \nabla w \nabla \phi dx$.

2.2.3 La fonction coût

Les observations $z_D \in L^2(\mathcal{C})$ étant données, la fonctionnelle à minimiser s'écrit sous la forme :

$$(2.21) \quad J(\beta, w) = \frac{1}{2} \int_{\mathcal{C}} \left(\frac{\partial w}{\partial N} - z_D \right)^2 d\sigma + \frac{\varepsilon}{2} \beta^t \Lambda \beta$$

où Λ est la matrice carrée d'ordre $m+2$ définie par : $\Lambda_{ij} = \int_0^1 \Phi_i''(x) \Phi_j''(x) dx$ pour $i, j = -2, -1, \dots, m-1$, et où w est lié à β par (2.20).

Remarque 2.9: (Régularité de w). L'opérateur Δ est elliptique linéaire sur D , donc si le second membre $C_{\Omega}^2 \sum_{i \in I} \beta_i \Phi_i(\bar{w})$ est dans $L^2(D)$ pour tout $w \in V^{p_0}(D)$, alors la solution w est dans $H^2(D)$ [LM 68], ce qui nous assure que $\frac{\partial w}{\partial N}$ est défini dans $L^2(C)$, et donc que la fonctionnelle J est bien définie à son tour. Nous supposerons cette hypothèse vérifiée par la suite.

Remarque 2.10: Lorsque nous avons traité les équations d'état, nous cherchions y solution dans $V^{p_0}(\Omega)$, ce qui revient à chercher w dans l'espace $V^{p_0}(D)$. Ceci n'est pas incompatible avec la remarque 2.9 puisque (cf [BRE 87]) lorsque Ω est un domaine borné de classe C^1 de \mathbb{R}^2 , nous savons que $W^{1,2}(\Omega) \subset L^q(\Omega)$ pour tout $q \in [1, +\infty[$. Donc, si $y \in H^2(\Omega)$, alors $y \in L^q(\Omega)$ et $\partial y \in W^{1,2}(\Omega) \subset L^q(\Omega)$, d'où: $y \in L^q(\Omega)$, $\partial y \in L^q(\Omega) \Rightarrow y \in W^{1,q}(\Omega)$ avec q quelconque dans $[1, +\infty[$, donc en particulier dans $W^{1,p_0}(\Omega)$ avec $p_0 \in [2, +\infty[$.

2.2.4 Formulation du problème de contrôle

Le problème de contrôle peut désormais se formuler de la façon suivante:

Trouver $\beta \in \mathbb{R}^{m+2}$ et w vérifiant les équations d'état (2.20) tels que :

$$(P) \quad \begin{cases} J(\beta, w) = \inf_{\gamma \in \mathbb{R}^{m+2}} J(\gamma, w') \\ w' \in V^{p_0}(D) \end{cases}$$

où w' est lié à γ par les équations d'état (2.20) et où J est donnée par (2.21) .

2.3 Introduction du lagrangien. Système d'optimalité

2.3.1 Définition du lagrangien

Puisque l'équation d'état (2.20) est non linéaire, la fonctionnelle J n'est pas convexe. Nous allons donc définir un lagrangien pour pouvoir en déduire formellement les conditions nécessaires d'optimalité du problème (P). Nous allons donc considérer les équations (2.20) comme une contrainte et introduire un multiplicateur de Lagrange $p \in W^{1,p_0}(D)$ par le lagrangien \mathcal{L} défini par :

$$(2.22) \quad \begin{cases} \mathcal{L}(\beta, w, p) = J(\beta, w) - \int_D \Delta w p \, dx - \sum_{i \in I} \beta_i \int_D C_{\Omega}^2 \Phi_i(\bar{w}) p \, dx \\ \forall \beta \in \mathbb{R}^{m+2}, w \in V^{p_0}(D), p \in W^{1,p_0}(D) \end{cases}$$

ou encore :

$$(2.23) \quad \begin{cases} \mathcal{L}(\beta, w, p) = J(\beta, w) + a(w, p) - \sum_{i \in I} \beta_i \int_D C_{\Omega}^2 \Phi_i(\bar{w}) p \, dx - \int_c \frac{\partial w}{\partial N} p \, d\sigma \\ \forall \beta \in \mathbb{R}^{m+2}, w \in V^{p_0}(D), p \in W^{1,p'_0}(D). \end{cases}$$

Une condition nécessaire et suffisante pour que $(\beta, w) \in \mathbb{R}^{m+2} \times V^{p_0}(D)$ soit solution du problème (P) est que (β, w, p) soit un point-selle pour \mathcal{L} , c'est à dire :

$$(2.24) \quad \begin{aligned} \mathcal{L}(\beta, w, q) &\leq \mathcal{L}(\beta, w, p) \leq \mathcal{L}(\gamma, w', p), \\ \forall (\gamma, w', q) &\in \mathbb{R}^{m+2} \times V^{p_0}(D) \times W^{1,p'_0}(D). \end{aligned}$$

Comme ce lagrangien n'est pas convexe, l'existence d'un point-selle n'est pas garantie, mais nous allons déduire de (2.24) des conditions nécessaires d'optimalité du problème (P).

2.3.2 Conditions nécessaires d'optimalité

Proposition 2.11 *Si les hypothèses H1 et H2 du paragraphe 2.1.3 sont satisfaites, alors une condition nécessaire et suffisante pour que (β, w, p) appartenant à $\mathbb{R}^{m+2} \times V^{p_0}(D) \times W^{1,p'_0}(D)$ soit un point-selle de \mathcal{L} est que (β, w) satisfasse les équations d'état (2.20) et que*

$$(2.25) \quad \begin{aligned} a(\phi, p) - \sum_{i \in I} \beta_i \int_D C_{\Omega}^2 \Phi'_i(\bar{w}) \frac{\phi w(M_0) - w \phi(M_0)}{w(M_0)^2} p \, dx \\ = \int_c \left(z_D - \frac{\partial w}{\partial N} \right) \frac{\partial \phi}{\partial N} \, d\sigma, \quad \forall \phi \in V^{p_0}(D) \end{aligned}$$

$$(2.26) \quad \varepsilon(\Lambda\beta)_i - \int_D C_{\Omega}^2 \Phi_i(\bar{w}) p \, dx = 0, \quad \forall i \in I$$

où M_0 est le point où w atteint son maximum sur D .

Preuve : L'inégalité de gauche de (2.24) peut s'écrire :

$$\begin{aligned} a(w, q) - \sum_{i \in I} \beta_i \int_D C_{\Omega}^2 \Phi_i(\bar{w}) q \, dx - \int_c \frac{\partial w}{\partial N} q \, d\sigma \\ \leq a(w, p) - \sum_{i \in I} \beta_i \int_D C_{\Omega}^2 \Phi_i(\bar{w}) p \, dx - \int_c \frac{\partial w}{\partial N} p \, d\sigma, \quad \forall q \in W^{1,p'_0}(D). \end{aligned}$$

Soit encore :

$$a(w, q') - \sum_{i \in I} \beta_i \int_D C_\Omega^2 \Phi_i(\bar{w}) q' dx - \int_C \frac{\partial w}{\partial N} q' d\sigma = 0, \quad \forall q' \in W^{1,p_0}(D).$$

En prenant q' dans $V^{p_0}(D) \subset W^{1,p_0}(D)$, on retrouve le fait que (β, w) vérifie la formulation variationnelle (2.20). De l'inégalité de droite de (2.24), nous déduisons les conditions d'Euler-Lagrange :

$$\begin{cases} \left(\frac{\partial \mathcal{L}}{\partial w}(\beta, w, p), \phi \right) = 0, & \forall \phi \in V^{p_0}(D) \\ \left(\frac{\partial \mathcal{L}}{\partial \beta_i}(\beta, w, p), \beta' - \beta \right) \geq 0, & \forall i \in I, \forall \beta' \in \mathbb{R}^{m+2} \end{cases}$$

or

$$\begin{aligned} \left(\frac{\partial \mathcal{L}}{\partial w}(\beta, w, p), \phi \right) &= \int_C \left(\frac{\partial w}{\partial N} - z_D - p \right) \frac{\partial \phi}{\partial N} d\sigma \\ &\quad + a(\phi, p) - \sum_{i \in I} \beta_i \int_D C_\Omega^2 \Phi_i'(\bar{w}) (b'(w)\phi) p dx. \end{aligned}$$

D'après l'expression de $b'(w)\phi$ calculée au lemme 2.6, nous retrouvons bien l'équation (2.25) vérifiée par l'état adjoint p . Puis, en dérivant \mathcal{L} par rapport au contrôle :

$$\frac{\partial \mathcal{L}}{\partial \beta_i}(\beta, w, p) = \varepsilon(\Lambda\beta)_i - \int_D C_\Omega^2 \Phi_i(\bar{w}) p dx, \quad \forall i \in I.$$

Nous obtenons donc bien l'expression (2.26). □

En réalité, puisque les contraintes sur β ont été prises en compte dans la formulation de la base de décomposition, nous nous trouvons dans un cas sans contraintes et (2.26) devient en fait une égalité :

$$\varepsilon(\Lambda\beta)_i - \int_D C_\Omega^2 \Phi_i(\bar{w}) p dx = 0.$$

Remarque 2.12 : L'existence et l'unicité de p sont assurées dès que la Gâteaux-dérivée $\tilde{y} = \lim_{\theta \rightarrow 0} \frac{y(f + \theta \tilde{f}) - y(f)}{\theta}$ de y par rapport au contrôle existe, puisque p est solution d'une équation adjointe de celle vérifiée par \tilde{y} . Nous supposons pour la suite cette condition vérifiée.

Remarque 2.13 : Des propriétés générales du lagrangien et de cette proposition, nous déduisons que le gradient de la fonction coût est le vecteur de \mathbb{R}^{m+2} de composantes :

$$\left(\nabla J(\beta) \right)_i = \varepsilon(\Lambda\beta)_i - \int_D C_\Omega^2 \Phi_i(\bar{w}) p dx, \quad i = -2, -1, \dots, m-1.$$

La fonctionnelle J est donc Gâteaux-dérivable dès que la Gâteaux-dérivée \tilde{y} de y existe. Ce gradient sera utilisé dans la méthode numérique d'optimisation de J .

2.3.3 Interprétation de l'état adjoint en termes d'EDP

Proposition 2.14 *Si la variable p définie par (2.25) existe et appartient à $W^{1,p'_0}(D)$, et si de plus ∇p appartient à $(W^{1,p'_0}(D))^2$, alors p vérifie les équations aux dérivées partielles suivantes :*

$$(2.27) \quad \begin{cases} -\Delta p = \sum_{i \in I} \beta_i C_\Omega^2 \Phi'_i(\bar{w})(b'(w)^t p) & \text{dans } D \\ p = \frac{\partial w}{\partial N} - z_D & \text{sur } \mathcal{C}. \end{cases}$$

Preuve : Considérons (2.25) appliquée à une fonction ϕ de $\mathcal{D}(D)$, et appliquons le lemme de J. Nečas (lemme 2.3) avec $u = \nabla p$ et $v = \phi$ (en inversant p_0 et p'_0) :

$$(2.28) \quad \begin{aligned} \int_D -\Delta p \phi \, dx + \int_{\mathcal{C}} \frac{\partial p}{\partial N} \phi \, d\sigma - \sum_{i \in I} \beta_i \int_D C_\Omega^2 \Phi'_i(\bar{w})(b'(w)\phi) p \\ = \int_{\mathcal{C}} \left(z_D - \frac{\partial w}{\partial N} + p \right) \frac{\partial \phi}{\partial N} \, d\sigma. \end{aligned}$$

Nous obtenons donc :

$$-\Delta p = \sum_{i \in I} \beta_i \int_D C_\Omega^2 \Phi'_i(\bar{w})(b'(w)^t p) \quad \text{dans } D.$$

Multiplions cette équation par $\phi \in V^{p_0}(D)$ et intégrons sur D :

$$\int_D -\Delta p \phi \, dx - \sum_{i \in I} \beta_i \int_D C_\Omega^2 \Phi'_i(\bar{w})(b'(w)^t p) \phi \, dx = 0.$$

En utilisant à nouveau la formule de J. Nečas :

$$\int_D \nabla p \cdot \nabla \phi \, dx - \sum_{i \in I} \beta_i \int_D C_\Omega^2 \Phi'_i(\bar{w})(b'(w)^t p) \phi \, dx = 0, \quad \forall \phi \in V^{p_0}(D). \quad \rightarrow$$

En soustrayant ceci avec (2.25) , il vient :

$$\int_{\mathcal{C}} \left(\frac{\partial w}{\partial N} - z_D - p \right) \frac{\partial \phi}{\partial N} \, d\sigma = 0, \quad \forall \phi \in V^{p_0}(D),$$

d'où nous déduisons la condition de type Neumann non homogène :

$$p = \frac{\partial w}{\partial N} - z_D \quad \text{sur } \mathcal{C}.$$

□

Remarque 2.15 : Nous pouvons noter dans la formulation variationnelle (2.25) (respectivement dans les équations aux dérivées partielles (2.27)) que l'état adjoint p est aussi l'état adjoint pour la formulation variationnelle linéarisée (2.6) (respectivement pour le système linéarisé (2.12)). Ceci s'explique facilement par le fait que l'état adjoint p est défini comme annulant la dérivée du lagrangien par rapport au vecteur d'état w et ceci apporte une nouvelle contrainte sur les équations d'état.

Remarque 2.16 : Dans la dualité entre w dans $V^{p_0}(D)$ et p dans $W^{1,p'_0}(D)$, nous avons choisi $p_0 > 2$ pour avoir la continuité de w ; ceci nous permettait de définir l'application b et d'avoir sa Gâteaux-dérivabilité. L'état adjoint p appartient à un espace de fonctions moins régulières ($p'_0 < 2$) et donc nous n'avons pas continuité de p .

2.4 Méthodes numériques de résolution

Le problème d'optimisation (P) est un problème de contrôle optimal avec des équations d'état non linéaires et une fonctionnelle non quadratique. Nous allons donc le résoudre par une méthode séquentielle quadratique consistant à résoudre une suite de problèmes de contrôle aux équations linéaires avec une fonctionnelle quadratique. Cette méthode séquentielle quadratique est dérivée de la méthode de Newton qui était la plus efficace pour résoudre le problème direct. La différence essentielle avec la méthode classique est que le contrôle β est modifié à chaque itération de façon à minimiser une fonctionnelle proche de J .

Pour commencer, précisons que nous allons résoudre le problème de contrôle pour les équations d'état (2.13) pour le problème discret dans l'espace V_h . L'algorithme consiste en itérations externes, chacune d'elles exécutant la résolution d'un problème linéaire quadratique de contrôle par une suite d'itérations internes de gradient conjugué. Plus précisément, la n -ème itération externe de l'algorithme consiste à définir pour le problème linéarisé en $w^n \in V_h$ un problème de contrôle (P^n) avec une fonctionnelle quadratique J^n "proche" de J . Ce problème linéaire quadratique de contrôle (P^n) sera lui-même résolu par une succession d'itérations internes de gradient conjugué et (β^{n+1}, w^{n+1}) sera obtenu comme solution du problème d'optimisation (P^n).

2.4.1 Itérations externes de l'algorithme séquentiel quadratique

Les itérations externes sont définies comme suit :

- Soit $(\beta^0, w^0) \in \mathbb{R}^{m+2} \times V_h$
- n -ème itération externe : on considère le problème linéarisé en (β^n, w^n) pour le système (2.13) ; alors l'état $w \in V_h$ est lié au contrôle $\beta \in \mathbb{R}^{m+2}$ par les équations :

$$(2.29) \quad a(w, \phi) - \sum_{i \in I} (\beta^n)_i \int_D C_{\Omega}^2 \Phi'_i(\overline{w^n})(b'(w^n)w) \phi = \sum_{i \in I} \beta_i \int_D C_{\Omega}^2 \Phi_i(\overline{w^n}) \phi.$$

Nous définissons alors la fonctionnelle :

$$J^n(\beta, w) = \frac{1}{2} \int_C \left(\frac{\partial w}{\partial N} - z_D \right)^2 d\sigma + \frac{\varepsilon}{2} \beta^t \Lambda \beta$$

où w est lié à β par (2.29). Cette fonction coût est bien quadratique, et nous allons résoudre le problème suivant :

Trouver $(\beta^{n+1}, w^{n+1}) \in \mathbb{R}^{m+2} \times V_h$ satisfaisant (2.29) et tel que :

$$(P^n) \quad J^n(\beta^{n+1}, w^{n+1}) = \inf_{(\beta, w) \in \mathbb{R}^{m+2} \times V_h} J^n(\beta, w)$$

où (β, w) satisfait (2.29).

Nous pouvons facilement prouver :

Proposition 2.17 *Le système d'optimalité pour le problème linéaire quadratique de contrôle (P^n) peut s'écrire :*

$$\left\{ \begin{array}{l} a(w^{n+1}, \phi) - \sum_{i \in I} (\beta^n)_i \int_D C_{\Omega}^2 \Phi'_i(\overline{w^n})(b'(w^n)w^{n+1}) \phi \\ \quad \quad \quad = \sum_{i \in I} (\beta^{n+1})_i \int_D C_{\Omega}^2 \Phi_i(\overline{w^n}) \phi \\ a(\phi, p^{n+1}) - \sum_{i \in I} (\beta^n)_i \int_D C_{\Omega}^2 \Phi'_i(\overline{w^n})(b'(w^n)^t \phi) p^{n+1} \\ \quad \quad \quad = \int_C \left(z_D - \frac{\partial w^{n+1}}{\partial N} + p^{n+1} \right) \frac{\partial \phi}{\partial N} d\sigma \\ \varepsilon \left(\Lambda \beta^{n+1} \right)_i - \int_D C_{\Omega}^2 \Phi_i(\overline{w^n}) p^{n+1} dx = 0, \quad \forall i \in I \end{array} \right.$$

avec $(\beta^{n+1}, w^{n+1}, p^{n+1}) \in \mathbb{R}^{m+2} \times V_h \times V_h$ et pour toute fonction $\phi \in V_h$.

2.4.2 Itérations internes de gradient conjugué

Pour résoudre le problème (P^n), nous utilisons un algorithme de gradient conjugué (cf [CIA 88] ou Chapitre I).

- $\beta_0^n = \beta^n$,
- j -ème itération interne :
 - ★ calcul de $w_j^n \in V_h$ par :

$$(2.30) \quad a(w_j^n, \phi) - \sum_{i \in I} (\beta^n)_i \int_D C_{\Omega}^2 \Phi'_i(\bar{w}^n) (b'(w^n) w_j^n) \phi = \sum_{i \in I} (\beta_j^n)_i \int_D C_{\Omega}^2 \Phi_i(\bar{w}^n) \phi$$

- ★ calcul de $p_j^n \in V_h$ par :

$$(2.31) \quad a(\phi, p_j^n) - \sum_{i \in I} (\beta^n)_i \int_D C_{\Omega}^2 \Phi'_i(\bar{w}^n) (b'(w^n)^t \phi) p_j^n = \int_C \left(z_D - \frac{\partial w_j^n}{\partial N} + p_j^n \right) \frac{\partial \phi}{\partial N} d\sigma$$

- ★ calcul du gradient $G_j^n \in \mathbb{R}^{m+2}$ par :

$$(2.32) \quad (G_j^n)_i = \varepsilon (\Lambda \beta_j^n)_i - \int_D C_{\Omega}^2 \Phi_i(\bar{w}^n) p_j^n dx$$

- ★ calcul des directions de descente $d_j^n \in \mathbb{R}^{m+2}$ par :

$$(2.33) \quad \begin{cases} d_0^n &= G_0^n \\ d_j^n &= G_j^n + \frac{\|G_j^n\|^2}{\|G_{j-1}^n\|^2} d_{j-1}^n \end{cases}$$

- ★ calcul de l'état $\xi_j^n \in V_h$ correspondant à la direction de descente d_j^n par :

$$(2.34) \quad a(\xi_j^n, \phi) - \sum_{i \in I} (\beta^n)_i \int_D C_{\Omega}^2 \Phi'_i(\bar{w}^n) (b'(w^n) \xi_j^n) \phi = \sum_{i \in I} (d_j^n)_i \int_D C_{\Omega}^2 \Phi_i(\bar{w}^n) \phi$$

- ★ calcul du taux optimal de descente $r_j^n \in \mathbb{R}^+$ par :

$$(2.35) \quad r_j^n = \frac{(G_j^n, d_j^n)}{\int_C \left(\frac{\partial \xi_j^n}{\partial n} \right)^2 + \varepsilon (d_j^n)^t \Lambda d_j^n}$$

* calcul du contrôle suivant :

$$\beta_j^{n+1} = \beta_j^n - r_j^n d_j^n.$$

Comme le nombre de paramètres de contrôle est $m+2$, l'algorithme de gradient conjugué converge en au plus $m+2$ itérations et donc :

$$\begin{cases} \beta_{m+2}^n & \rightarrow \beta^{n+1} \\ w_{m+2}^n & \rightarrow w^{n+1}. \end{cases}$$

Alors, on recalcule la matrice du système (2.29) pour la prochaine itération externe.

2.4.3 Structure et convergence de l'algorithme

L'algorithme global est résumé par le diagramme suivant :

$n \leftarrow 0$; Initialisation de β^0 , w^0 , $e \leftarrow 1$

Tant que $e > \varepsilon$ faire :

Calcul et factorisation de la matrice du système (2.29)

$j \leftarrow 0$

$\beta_0^n \leftarrow \beta^n$

Calcul de w_0^n par (2.30)

Tant que $j < m+2$ faire :

Calcul de p_j^n par (2.31)

Calcul du gradient G_j^n par (2.32)

Si $j = 0$ alors $d_0^n \leftarrow G_0^n$ sinon calcul de d_j^n par (2.33)

Calcul de ξ_j^n par (2.34)

Calcul de r_j^n par (2.35)

$\beta_{j+1}^n \leftarrow \beta_j^n - r_j^n d_j^n$

Calcul de w_j^n par (2.30)

$j \leftarrow j + 1$

Fin tant que

$\beta^{n+1} \leftarrow \beta_{m+2}^n$; $w^{n+1} \leftarrow w_{m+2}^n$; $n \leftarrow n + 1$

$e \leftarrow \frac{\|w^{n+1} - w^n\|}{\|w^n\|}$

Fin tant que

Fin.

2.4.4 Validation du gradient

Afin de vérifier le gradient obtenu par (2.32), nous considérons le rapport des différences divisées :

$$\frac{J(\beta + h\nabla J(\beta)) - J(\beta)}{h \|\nabla J(\beta)\|^2},$$

qui, pour un pas h arbitrairement petit, devrait tendre vers la valeur 1. Nous représentons numériquement la fonction

$$h \mapsto 1 - \frac{J(\beta + h\nabla J(\beta)) - J(\beta)}{h \|\nabla J(\beta)\|^2}.$$

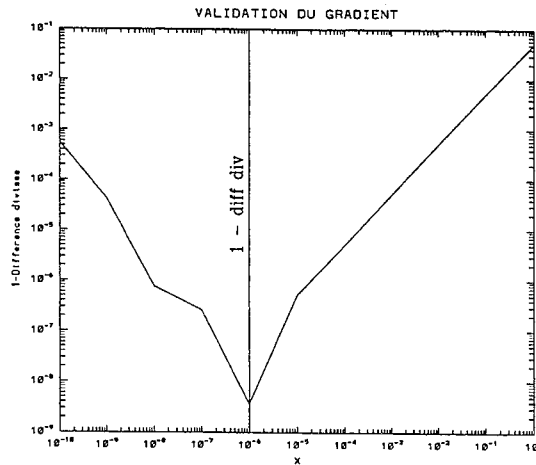


Figure 2.8 Validation du gradient par différence divisée.

On constate que la précision obtenue est d'autant meilleure que h est petit, mais lorsque h devient inférieur à 10^{-6} , ce rapport "explose" en raison des erreurs numériques d'arrondi. Il s'agit donc d'un bon test du gradient lorsqu'on connaît une valeur exacte de celui-ci, mais cette méthode en elle-même est à déconseiller pour un calcul direct du gradient en raison de ces instabilités numériques.

2.4.5 Résultats numériques

- *Identification sans bruit*

Dans la pratique, cet algorithme converge généralement au bout de 5 ou 6 itérations externes lorsqu'on n'a pas de bruit sur les données. Au vu du nombre assez restreint de paramètres de contrôle utilisés, la majorité du temps d'exécution du programme est consacrée au calcul et à la factorisation de la matrice du système (2.29) ; les itérations internes nécessitent seulement une partie infime du temps de calcul. Le temps d'exécution d'un tel algorithme séquentiel quadratique est donc à peine plus long que celui de la méthode de Newton classique pour résoudre le problème direct. Cela justifie notre choix d'un tel algorithme pour résoudre le problème d'optimisation, plutôt qu'une méthode de type "steepest descent" ou "gradient conjugué non linéaire" (de Fletcher-Reeves ou Polak-Ribière) qui elles, nécessiteraient des résolutions numériques d'équations d'état non linéaires pour déterminer le paramètre de descente optimale.

Nous donnons ensuite quelques fonctions identifiées à l'aide de cet algorithme séquentiel quadratique sur des données simulées non bruitées.

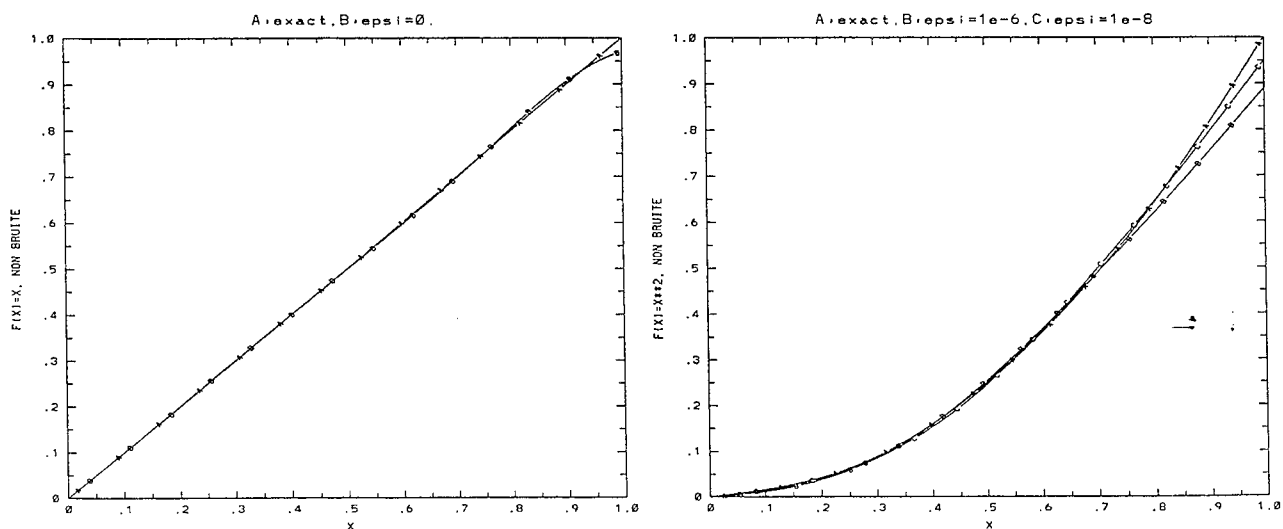


Figure 2.9 A gauche, $f(x) = x$, à droite $f(x) = x^2$.

On constate (figure 2.9) qu'en l'absence de bruit sur les données, l'identification donne des résultats corrects, et ne nécessite pas de terme de régularisation dans la fonction coût. Pour la fonction x^2 (à droite de la figure 2.9), nous avons réalisé l'identification avec deux valeurs distinctes du paramètre de régularisation, afin d'avoir une idée de la forme de la solution régularisée dans ce cas. On peut constater que ces valeurs de ε restent assez faibles, et si l'on avait pris ε plus grand, le résultat de l'identification aurait été une droite.

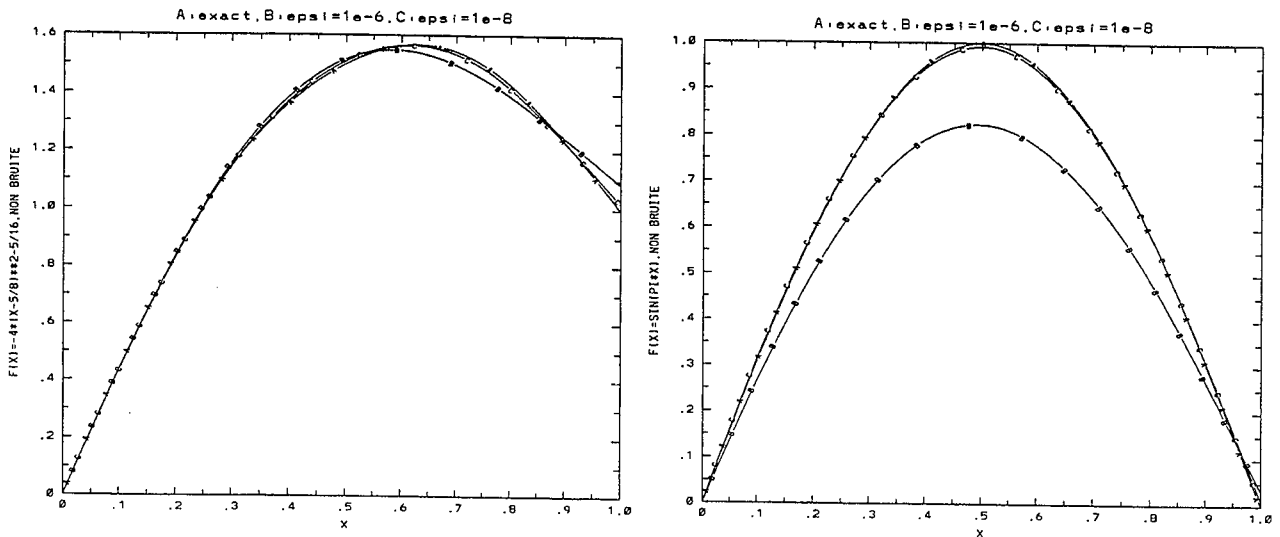


Figure 2.10 A gauche, $f(x) = -4x^2 - 5x$, à droite $f(x) = \sin(\pi x)$.

Sur ces deux dernières fonctions identifiées, (figure 2.10), on constate déjà que la non linéarité a pour conséquence qu'un mauvais choix du paramètre ε , entraîne une mauvaise identification de la fonction. En particulier, pour la fonction sinus, lorsque ε est choisi trop grand, la fonction identifiée tend vers l'axe des x à cause de la condition imposée $f(0) = 0$.

Il est à noter que le domaine choisi dans ces quatre cas de figure est la surface de \mathbb{R}^2 délimitée par l'ellipse de centre 0, de petit axe 1 et de grand axe 2. Lorsque l'on choisit comme frontière du domaine un cercle, on constate que la fonction d'état devient radiale, comme l'illustre la figure 2.11, donc les observations $\frac{\partial y}{\partial n}$ deviennent constantes, et l'on n'arrive plus à identifier correctement la source.

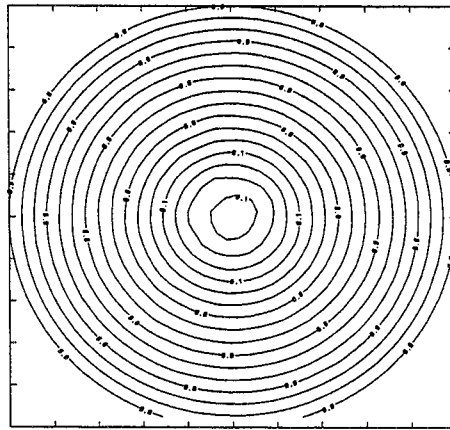


Figure 2.11 Isovaleurs de y lorsque le domaine est un cercle.

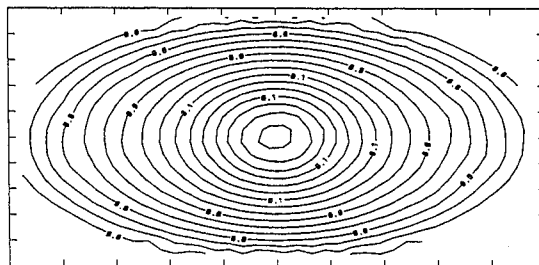


Figure 2.12 Isovaleurs de y lorsque le domaine est une ellipse.

- *Identification pour des données bruitées*

Nous illustrons ici l'identification de sources non linéaires pour différents domaines de \mathbb{R}^2 .

Les figure 2.13 et figure 2.14 illustrent l'identification d'un polynôme de degré 2 ($-4x^2 + 5x$) lorsque l'on fait tendre le domaine Ω de l'intérieur de l'ellipse

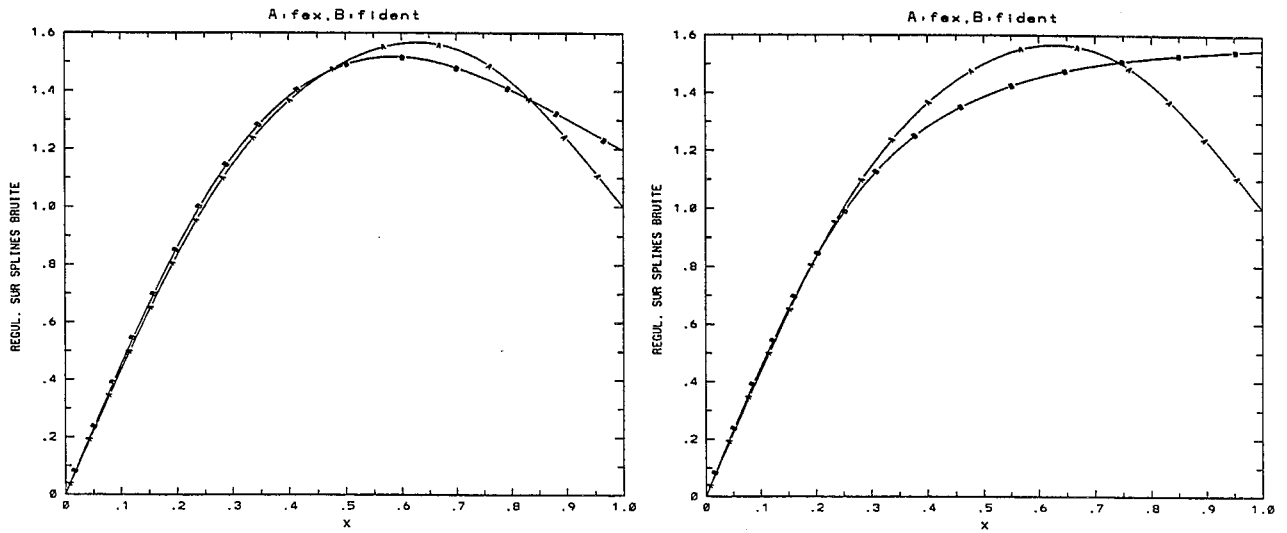


Figure 2.13 Identification sur ellipse $x^2 + \frac{y^2}{2^2} = 1$ à gauche et sur $x^2 + \frac{y^2}{1.1^2} = 1$ à droite.

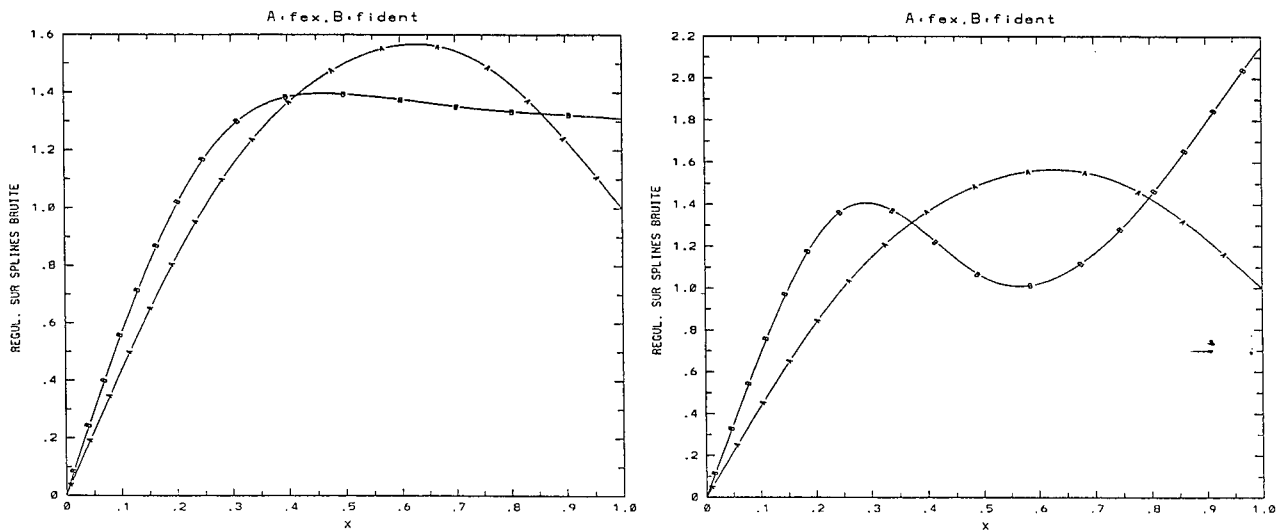


Figure 2.14 Identification sur ellipse $x^2 + \frac{y^2}{1.01^2} = 1$ à gauche et sur $x^2 + \frac{y^2}{1.0001^2} = 1$ à droite.

d'axes 1 et 2 vers le disque unité de \mathbb{R}^2 . Dans le cas du disque unité, nous n'avons pu représenter de résultat, car le programme ne converge pas.

Nous avons utilisé ici des données bruitées à 5%, et pour chaque courbe représentée, nous avons choisi le paramètre de régularisation qui donnait la meilleure identification au vu de la courbe obtenue. Nous avons ensuite représenté la fonction identifiée avec ce ε . Les temps de calculs sont les suivants: pour l'ellipse(1,2): 4 itérations externes sont nécessaires, soit environ 7,6 secondes CPU d'un CRAY C90 avec $\varepsilon = 10^{-5}$. Pour l'ellipse(1,1.1): 5 itérations externes sont nécessaires, soit environ 8.3 secondes CPU, avec $\varepsilon = 10^{-7}$. Pour l'ellipse(1,1.01): 8 itérations externes sont nécessaires, soit environ 11.3 secondes CPU, avec $\varepsilon = 10^{-8}$, et pour l'ellipse(1,1.0001): 19 itérations externes sont nécessaires, soit environ 27.8 secondes CPU, avec $\varepsilon = 10^{-9}$.

Comme on pouvait s'y attendre, plus le domaine s'approche de la boule unité de \mathbb{R}^2 et plus l'identification s'avère délicate. Dans la pratique, il est généralement admis que pour une ellipse d'élongation inférieure à 1.2, l'identification est quasi impossible à réaliser à partir de données bruitées.

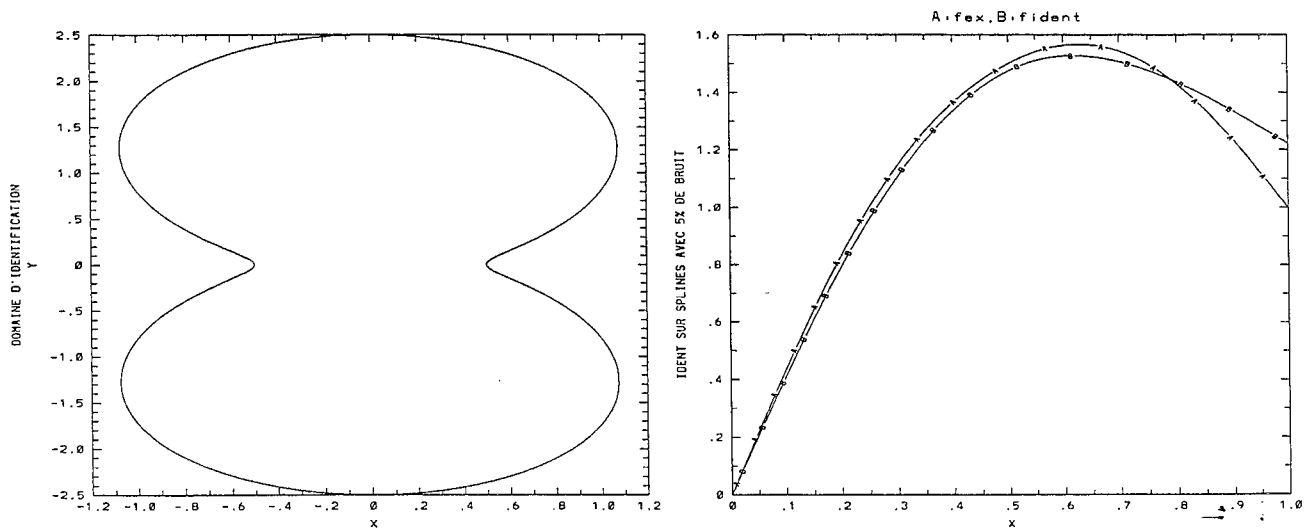


Figure 2.15 A gauche, domaine utilisé, à droite identification avec bruit

Dans le cas de ce domaine particulier (figure 2.15), dont la forme est relativement éloignée de celle d'un cercle, l'identification a été réalisée toujours avec 5% de bruit, l'algorithme a convergé en 4 itérations externes pour $\varepsilon = 10^{-6}$ choisi par le même procédé que précédemment (environ 6,8 secondes CPU).

Nous nous fixons ensuite pour domaine l'intérieur de l'ellipse d'axes 1 et 2, et nous donnons les résultats de l'identification pour la source $\sin(\pi x)$ pour différents paramètres ε avec un niveau de bruit fixé à 5%.

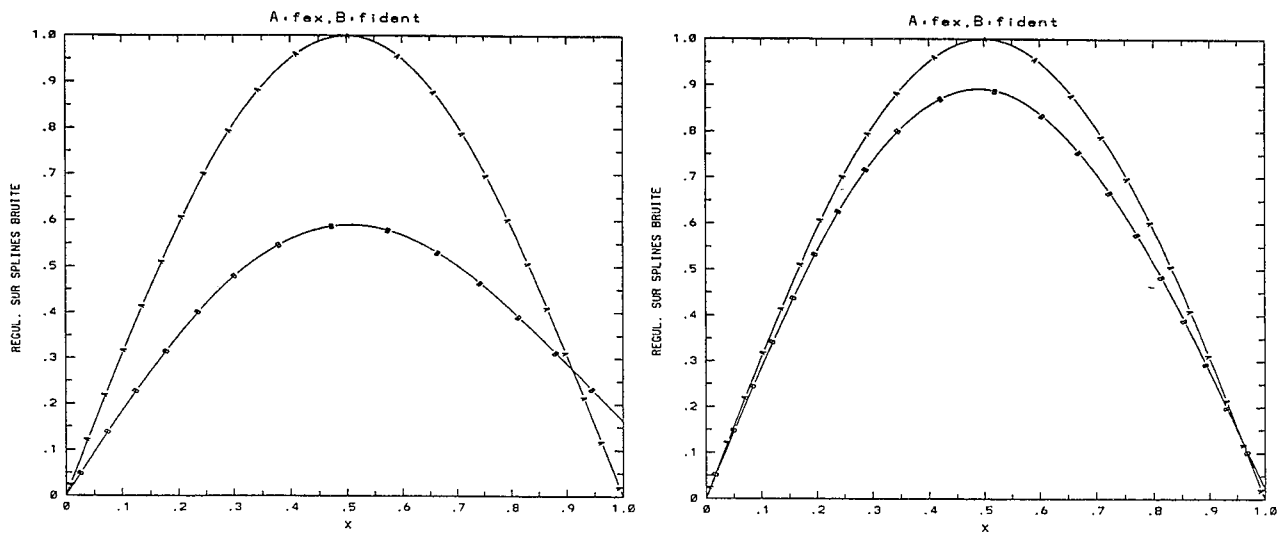


Figure 2.16 Identification de $\sin(\pi x)$ pour $\varepsilon = 10^{-5}$ à gauche et 10^{-6} à droite.

Sur les courbes obtenues figure 2.16 , nous voyons que les paramètres ε employés sont trop grands, car la fonction identifiée a tendance à se rapprocher de l'axe des x . Par contre, figure 2.17 à droite, ε est trop petit, ce qui engendre une fonction identifiée légèrement instable. Le paramètre de régularisation idéal se situe donc dans ce cas entre les valeurs 10^{-6} et 10^{-7} . Cet exemple illustre donc bien l'importance jouée par ε lors de l'identification.

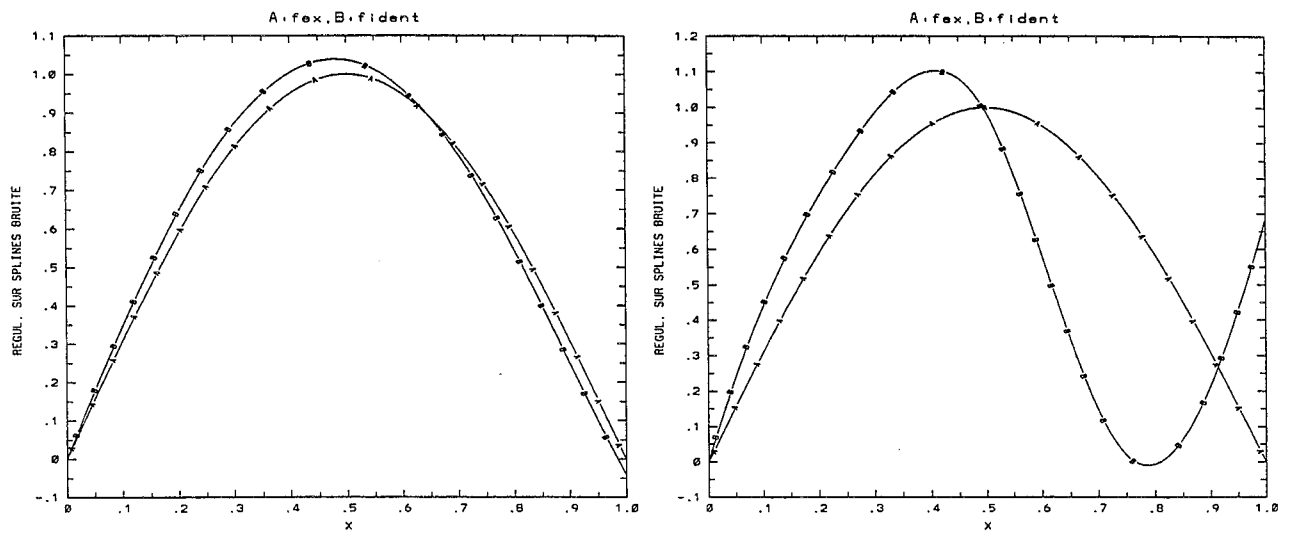


Figure 2.17 Identification de $\sin(\pi x)$ pour $\varepsilon = 10^{-7}$ à gauche et 10^{-9} à droite.

Chapitre III

Choix du paramètre de régularisation

Nous avons vu au chapitre II l'importance du choix du paramètre de régularisation : un coefficient trop grand conduit à l'estimation d'une droite, tandis qu'un coefficient trop petit laisse apparaître des instabilités numériques.

Nous rappelons dans un premier temps les techniques de choix de ce paramètre dans le cas d'un problème inverse linéaire : la méthode du résidu, utile lorsqu'on a une idée du niveau de bruit entachant les données, les méthodes de validation croisée ordinaire et généralisée, qui elles, sont mieux adaptées aux problèmes concrets où l'on ne connaît pas la proportion de bruit dans les observations.

Nous verrons dans un deuxième temps comment adapter ces méthodes pour les deux problèmes inverses traités au chapitre précédent : identification du coefficient de diffusion de la chaleur et identification de sources non linéaires. Pour ces problèmes inverses non linéaires, il est nécessaire de connaître la jacobienne de l'opérateur de lissage afin d'estimer la fonction de validation croisée généralisée. Une justification théorique de son existence sera donnée, ainsi qu'une méthode "adjointe" pour son calcul dans les deux cas cités ci-dessus.

Enfin, des résultats numériques seront présentés.

1 Cas d'un problème inverse linéaire

Dans tout le paragraphe on considère le modèle suivant :

$$z = Tv + \nu$$

où T est une application linéaire connue de \mathbb{R}^m dans \mathbb{R}^n , v est inconnu et ν un vecteur aléatoire de loi $\mathcal{N}(0, \sigma^2 I)$ représentant un bruit blanc sur les données. On s'intéresse au problème suivant :

$$(P) \quad \text{Trouver } v \in \mathbb{R}^m \text{ tel que } \sum_{i=1}^n \alpha_i^2 |(Tv)_i - z_i|^2 \text{ soit minimale,}$$

où les α_i^2 sont des poids positifs connus, provenant de la discrétisation du problème. On se place dans le cas où le problème inverse est mal posé à cause de la non-continuité de l'inverse de T . On introduit une régularisation de Tikhonov et on considère le problème linéarisé :

$$(P_\varepsilon) \quad \text{Trouver } v \in \mathbb{R}^p \text{ tel que } \sum_{i=1}^n \alpha_i^2 |(Tv)_i - z_i|^2 + \varepsilon v^t \Lambda v \text{ soit minimale.}$$

où Λ est une matrice $m \times m$ symétrique semi-définie positive. Par dérivation, on obtient qu'une estimation régularisée \hat{v} de v doit vérifier :

$$\langle D(T\hat{v} - z), T\phi \rangle + \varepsilon \langle \Lambda \hat{v}, \phi \rangle = 0, \quad \forall \phi \in \mathbb{R}^p$$

avec $D = \text{diag}(\alpha_1^2, \dots, \alpha_n^2)$, d'où :

$$\hat{v} = (T^*DT + \varepsilon\Lambda)^{-1}T^*Dz.$$

On pose $M_\varepsilon = T(T^*DT + \varepsilon\Lambda)^{-1}T^*D$, matrice $n \times n$ de lissage correspondant à ε : $T\hat{v} = M_\varepsilon z$.

1.1 Choix du paramètre par la méthode du résidu

On définit tout d'abord le résidu associé à ε par :

$$S(\varepsilon) = \frac{1}{n} \sum_{i=1}^n \alpha_i^2 (z_{i,\varepsilon} - z_i)^2$$

où z_ε représente le vecteur des données lissées, c'est à dire que $z_\varepsilon = T(v_\varepsilon)$ avec v_ε solution de (P_ε) .

Puisque le bruit ν suit une loi normale $\mathcal{N}(0, \sigma^2 I)$, si l'on connaît la variance du bruit σ^2 , il semble tout à fait naturel de choisir le paramètre ε de sorte à avoir un écart pondéré entre les données exactes et les données lissées qui soit de même nature que le bruit, c'est à dire :

$$(1.1) \quad \text{Choix de } \varepsilon \text{ tel que } S(\varepsilon) = \sigma^2.$$

Dans [REI 67], C. Reinsch a obtenu dans ce cas des résultats numériques satisfaisants. Il est facile de montrer que $S(\varepsilon)$ est une fonction croissante de ε lorsque M_ε est bien définie, et donc l'existence et l'unicité de ε vérifiant (1.1) est assurée dès que $\sigma^2 \in [S(0), \lim_{\varepsilon \rightarrow \infty} S(\varepsilon)]$.

1.2 La méthode de C.L. Mallows

Elle consiste à considérer le résidu $S(\varepsilon)$, que l'on peut encore écrire :

$$S(\varepsilon) = \frac{1}{n} \| D^{\frac{1}{2}}(M_\varepsilon z - \bar{z}) \|^2$$

avec \bar{z} l'espérance mathématique de z et $z_\varepsilon = T(v_\varepsilon) = M_\varepsilon z$. L'espérance de cette erreur "vraie" se calcule facilement :

$$(1.2) \quad E\left(S(\varepsilon)\right) = \frac{1}{n} \| D^{\frac{1}{2}}(M_\varepsilon \bar{z} - \bar{z}) \|^2 + \frac{2\sigma^2}{n} \text{Tr}(M_\varepsilon)$$

et on prendra pour ε la valeur qui minimise $E\left(S(\varepsilon)\right)$.

L'inconvénient majeur de ces deux méthodes est qu'elles nécessitent la connaissance a priori de la variance du bruit σ^2 , qui n'est jamais connue dans la pratique. Ceci explique la popularité de méthodes telles que la validation croisée, qui elles, ne demandent pas la connaissance de σ^2 .

1.3 Choix du paramètre par validation croisée

L'idée est de séparer les observations en deux sous-ensembles, le premier étant destiné à calculer, à ε fixé, l'estimation de v et le second à évaluer si ε est bien choisi.

On oublie temporairement la k -ème donnée z_k , et on note $z_\varepsilon^{(k)}$ le lissage de la solution associée à ε :

$$z_\varepsilon^{(k)} = Tv_\varepsilon^{(k)}$$

avec $v_\varepsilon^{(k)}$ solution du problème :

$$(P_\varepsilon^{(k)}) \quad \text{Trouver } v \in \mathbb{R}^m \text{ tel que } \sum_{\substack{i=1 \\ i \neq k}}^n \alpha_i^2 \left((Tv)_i - z_i \right)^2 + \varepsilon v^t \Lambda v \text{ soit minimal.}$$

On peut alors mesurer la validité de ε par l'erreur de prédiction sur la k -ème donnée oubliée : $z_\varepsilon^{(k)} - z_k$. Le principe de la validation croisée est de choisir ε qui minimise la moyenne quadratique de ces erreurs pondérées pour tout k :

$$(1.3) \quad \text{Choix de } \varepsilon \text{ tel que } \mathcal{V}_0(\varepsilon) = \frac{1}{n} \sum_{k=1}^n \alpha_k^2 (z_{\varepsilon,k}^{(k)} - z_k)^2 \text{ soit minimal.}$$

Heureusement, l'évaluation de $\mathcal{V}_0(\varepsilon)$ ne nécessite pas la résolution de n sous problèmes $(P_\varepsilon^{(k)})$ mais seulement la connaissance de $M_\varepsilon z$. Pour le montrer, il suffit de voir que $v_\varepsilon^{(k)}$ est aussi solution d'un problème de même type que (P_ε) avec pour k -ème donnée $\left(Tv_\varepsilon^{(k)} \right)_k$ au lieu de z_k . On écrit ensuite que la k -ème valeur prédite par $Tv_\varepsilon^{(k)}$ est en fait la k -ème ligne de $M_\varepsilon \tilde{z}_\varepsilon^{(k)}$ avec $\tilde{z}_\varepsilon^{(k)} = (z_1, \dots, z_{k-1}, (Tv_\varepsilon^{(k)})_k, z_{k+1}, \dots, z_n)$ et donc :

$$\left(Tv_\varepsilon^{(k)} \right)_k = \sum_{\substack{i=1 \\ i \neq k}}^n (M_\varepsilon)_{ki} z_i + (M_\varepsilon)_{kk} \left(Tv_\varepsilon^{(k)} \right)_k$$

d'où :

$$\left(1 - (M_\varepsilon)_{kk} \right) \left(Tv_\varepsilon^{(k)} \right)_k = \sum_{i=1}^n (M_\varepsilon)_{ki} z_i - (M_\varepsilon)_{kk} z_k.$$

Soit encore :

$$z_{\varepsilon,k}^{(k)} = \frac{1}{1 - (M_\varepsilon)_{kk}} \left((M_\varepsilon z)_k - (M_\varepsilon)_{kk} z_k \right)$$

d'où :

$$z_{\varepsilon,k}^{(k)} - z_k = \frac{1}{1 - (M_\varepsilon)_{kk}} \left((M_\varepsilon z)_k - z_k \right);$$

en reportant ceci dans (1.3) , on obtient :

$$(1.4) \quad \mathcal{V}_0(\varepsilon) = \frac{1}{n} \sum_{k=1}^n \alpha_k^2 \frac{\left((M_\varepsilon z)_k - z_k \right)^2}{\left(1 - (M_\varepsilon)_{kk} \right)^2}.$$

1.4 Validation croisée généralisée

Lorsque toutes les données interviennent de façon symétrique, c'est à dire qu'une permutation circulaire des z_k ne modifie le vecteur lissé $M_\varepsilon z$ que par cette même permutation (M_ε est circulante), alors on a :

$$1 - (M_\varepsilon)_{kk} = \frac{1}{n} \text{Tr}(I_n - M_\varepsilon).$$

Cette expression donne alors une moyenne arithmétique des termes diagonaux de M_ε . En remplaçant cette approximation dans (1.4), on obtient l'expression de la validation croisée généralisée (GCV) $\mathcal{V}(\varepsilon)$:

$$(1.5) \quad \mathcal{V}(\varepsilon) = \frac{\frac{1}{n} \| D^{\frac{1}{2}}(I_n - M_\varepsilon)z \|^2}{\left(\frac{1}{n} \text{Tr}(I_n - M_\varepsilon) \right)^2}.$$

Pour obtenir un bon choix du paramètre de régularisation ε , on sera amené à minimiser la fonction $\mathcal{V}(\varepsilon)$ sur l'ensemble des réels positifs. Cette modification de la fonction de validation croisée ordinaire est justifiée théoriquement par P. Craven et G. Wahba dans [CW 79]. Les résultats de convergence pour le paramètre donné par GCV et pour l'estimateur de C.L. Mallows sont démontrés par K.C. Li dans [LI 85], [LI 86] dans le cas où $\Lambda = I_m$ (appelé aussi "ridge regression" en statistique).

1.5 Estimation de la trace de la matrice de lissage

La difficulté essentielle dans les calculs d'estimation du paramètre de régularisation ε réside dans le calcul de la trace de M_ε . Dans le cas où $\Lambda = I_m$, un algorithme classique consiste à décomposer T en valeurs singulières, ce qui permet ensuite de calculer en $O(n)$ opérations aussi bien les valeurs propres de M_ε que le résidu $S(\varepsilon)$ pour tout ε (voir par exemple [GHW 79]). Un autre algorithme utilise à la place de la décomposition en valeurs singulières la bidiagonalisation de T pour un coût total un peu plus faible (voir les travaux de V.A. Morozov [MOR 84]).

Pour le cas général $\Lambda \neq I_m$, on peut se ramener au précédent par certaines décompositions de matrices préalables (consulter par exemple les travaux de D.M. Bates et G. Wahba [BWb 82]), mais en l'absence de symétries particulières, toutes ces méthodes demandent $O(n^3)$ opérations, même pour des matrices bandes, et ne sont donc pas réalisables pour des problèmes de grande taille.

D. Girard a proposé une approximation de la trace pour n assez grand ([GIR1 87], [GIR2 87]) utilisant un estimateur de Monte Carlo. On considère pour cela M vecteurs pseudo-aléatoires w_1, \dots, w_M de \mathbb{R}^n de loi normale $\mathcal{N}(0, I)$, on calcule pour chacun des vecteurs Aw_i où A est une matrice $n \times n$ dont on veut calculer la trace, et on prend alors comme estimateur :

$$\frac{1}{n} \text{Tr}(A) \simeq \frac{1}{M} \sum_{i=1}^M \frac{w_i^t A w_i}{w_i^t w_i}.$$

Pour des propriétés théoriques et une justification théorique, voir [GIR2 87]. Des résultats de convergence asymptotique pour le paramètre obtenu par GCV et par l'estimateur de C.L. Mallows (pour $M = 1$) sont démontrés dans [GIR 89]. En particulier, pour $M = 1$, D. Girard a montré que l'estimateur $\frac{w^t A w}{w^t w}$ est sans biais, d'écart relatif $\sqrt{\frac{2}{n+2}}$ fois la dispersion relative des valeurs propres.

2 Cas d'un problème inverse non linéaire

Nous considérons dans ce paragraphe le modèle :

$$z_i = T_i(v) + \nu_i, \quad i = 1, \dots, n$$

où les T_i , $i = 1, \dots, n$ sont certaines fonctionnelles non linéaires connues et ν_1, \dots, ν_n sont n variables aléatoires indépendantes de loi $\mathcal{N}(0, \sigma^2)$, où la variance σ^2 est généralement inconnue. Nous cherchons une estimation de v à partir des données z bruitées. Pour cela, nous voulons trouver $v_{\varepsilon, z}$ qui minimise la fonctionnelle :

$$(2.1) \quad \sum_{i=1}^n \alpha_i^2 (T_i(v) - z_i)^2 + \varepsilon \int_0^1 v''(t)^2 dt$$

pour $v \in H^2(]0, 1[)$ et $\varepsilon > 0$ donné. Le paramètre de régularisation ε contrôle le compromis entre la fidélité aux données mesurée par $\sum_{i=1}^n \alpha_i^2 (T_i(v) - z_i)^2$ et le

caractère lisse de la solution mesuré par $\int_0^1 v''(t)^2 dt$. Il s'agit d'une formulation non quadratique classique pour estimer une fonction bornée à partir de données erronées. En pratique, le choix de ε est tout à fait crucial. Numériquement, v sera cherchée dans une base de B -splines cubiques de telle sorte que : $\int_0^1 v''(t)^2 dt = v^t \Lambda v$ par abus de langage, en notant encore par v le vecteur de \mathbb{R}^m tel que $v(x) = \sum_{i=1}^m v_i N_i^4(x)$ et avec $\Lambda_{ij} = \int_0^1 N_i^{4''}(x) N_j^{4''}(x) dx$ matrice symétrique positive. Nous allons donc chercher $v_{\varepsilon, z} \in \mathbb{R}^m$ tel que :

$$(2.2) \quad \sum_{i=1}^n \alpha_i^2 (T_i(v) - z_i)^2 + \varepsilon v^t \Lambda v \quad \text{soit minimale.}$$

Ce problème peut être résolu par une méthode de Gauss-Newton ou par une méthode séquentielle quadratique décrite au paragraphe 2.4 du chapitre II.

Le paragraphe 2.1 est consacré au choix du paramètre de régularisation par validation croisée généralisée qui fait intervenir la dérivée de l'opérateur non linéaire de lissage par rapport à z de façon tout à fait formelle. La méthode de minimisation des fonctions de validation croisée sera détaillée en 2.2. Nous verrons ensuite dans les sections 3 et 4 une expression explicite de cette dérivée dans le cas des problèmes d'identification de la diffusion et de sources non linéaires, évoqués

au chapitre II. Nous donnerons pour ces deux problèmes une série de résultats numériques.

2.1 Choix du paramètre par GCV

Nous considérons le problème (2.2) et nous voulons choisir le paramètre de régularisation ε par validation croisée. L'idée de base est la même que dans le cas classique linéaire. Soit $v_\varepsilon^{(k)}$ la solution calculée à partir de toutes les données exceptée la k -ème z_k , alors $v_\varepsilon^{(k)}$ réalise le minimum de :

$$\sum_{\substack{i=1 \\ i \neq k}} \alpha_i^2 (T_i(v) - z_i)^2 + \varepsilon v^t \Lambda v.$$

On voit facilement que $v_\varepsilon^{(k)}$ réalise également le minimum de la fonctionnelle :

$$(2.3) \quad \sum_{\substack{i=1 \\ i \neq k}} \alpha_i^2 (T_i(v) - z_i)^2 + \alpha_k^2 (T_k(v) - T_k(v_\varepsilon^{(k)}))^2 + \varepsilon v^t \Lambda v.$$

Considérons l'opérateur de lissage B_ε défini par :

$$\text{si } v_{\varepsilon, z} \text{ est solution de (2.2), alors } T(v_{\varepsilon, z}) = B_\varepsilon(z).$$

Puisque (2.3) est de même nature que (2.2) pour des observations différentes, nous en déduisons que :

$$B_\varepsilon(z_1, \dots, z_{k-1}, T_k(v_\varepsilon^{(k)}), z_{k+1}, \dots, z_n) = T(v_\varepsilon^{(k)})$$

et donc :

$$(2.4) \quad T_k(v_\varepsilon^{(k)}) = [B_\varepsilon(z + \delta_k e_k)]_k, \quad \rightarrow$$

où e_k est le k -ème vecteur de la base canonique de \mathbb{R}^n et $\delta_k = T_k(v_\varepsilon^{(k)}) - z_k$. Or, la fonction de validation croisée est définie par :

$$\mathcal{V}_0(\varepsilon) = \frac{1}{n} \sum_{k=1}^n \alpha_k^2 (T_k(v_\varepsilon^{(k)}) - z_k)^2.$$

Posons

$$a_{kk}(\varepsilon) = \frac{[B_\varepsilon(z + \delta_k e_k)]_k - [B_\varepsilon(z)]_k}{\delta_k},$$

nous avons alors :

$$\mathcal{V}_0(\varepsilon) = \frac{1}{n} \sum_{k=1}^n \alpha_k^2 \delta_k^2$$

or,

$$\begin{aligned} \delta_k &= T_k(v_\varepsilon^{(k)}) - z_k \\ &= [B_\varepsilon(z + \delta_k e_k)]_k - [B_\varepsilon(z)]_k + T_k(v_{\varepsilon,z}) - z_k \end{aligned}$$

d'où :

$$\delta_k = \frac{T_k(v_{\varepsilon,z}) - z_k}{1 - a_{kk}(\varepsilon)},$$

et finalement :

$$\mathcal{V}_0(\varepsilon) = \frac{1}{n} \sum_{k=1}^n \alpha_k^2 \frac{(T_k(v_{\varepsilon,z}) - z_k)^2}{(1 - a_{kk}(\varepsilon))^2}.$$

Par analogie avec le cas classique, on remplace $a_{kk}(\varepsilon)$ par $\frac{1}{n} \sum_{j=1}^n a_{jj}(\varepsilon)$ pour obtenir la GCV, donnée par :

$$\mathcal{V}(\varepsilon) = \frac{\frac{1}{n} \sum_{k=1}^n \alpha_k^2 (T_k(v_{\varepsilon,z}) - z_k)^2}{\left(\frac{1}{n} (n - \sum_{k=1}^n a_{kk}(\varepsilon)) \right)^2}.$$

Lorsque ε est bien choisi, la quantité $\delta_k = T_k(v_\varepsilon^{(k)}) - z_k$ devrait être petite, puisqu'il s'agit de l'erreur de prédiction de la k -ème donnée oubliée z_k , ainsi $a_{kk}(\varepsilon)$ peut être approché par $\frac{\partial (B_\varepsilon(z))_k}{\partial z_k}$ et la fonction GCV par :

$$\mathcal{V}(\varepsilon) = \frac{\frac{1}{n} \sum_{k=1}^n \alpha_k^2 (T_k(v_{\varepsilon,z}) - z_k)^2}{\left(\frac{1}{n} \text{Tr}(I - (B'_\varepsilon)_z) \right)^2}.$$

2.2 Expression analytique pour $(B'_\varepsilon)_z$ dans le cas général

Nous utilisons pour le calcul de cette dérivée, de manière tout à fait formelle, les travaux de L.N. Deshpande [DES 90]. Posons :

$$J_{\varepsilon,z}(v) = \frac{1}{2} \sum_{i=1}^n \alpha_i^2 (T_i(v) - z_i)^2 + \frac{\varepsilon}{2} v^t \Lambda v,$$

nous considérons le problème :

$$(2.5) \quad \min_{v \in \mathbb{R}^m} J_{\varepsilon, z}(v)$$

dont la solution (si elle existe) $v_{\varepsilon, z}$ sera notée par la suite v_z . Le gradient de cette fonctionnelle s'écrit :

$$\nabla_v J_{\varepsilon, z}(v) = \sum_{i=1}^n \alpha_i^2 (T_i(v) - z_i) T_i'(v) + \varepsilon \Lambda v.$$

Nous avons donc, comme condition nécessaire à (2.5) :

$$\nabla J_{\varepsilon, z}(v_z) = 0, \quad \forall z.$$

Nous allons faire maintenant l'hypothèse suivante :

Quelles que soient les observations z , si v_z est solution du problème de minimisation (2.5) alors $\text{Hess}_v J_{\varepsilon, z}(v_z)$ existe et il est défini positif.

Alors, d'après le théorème des fonctions implicites, il existe un voisinage \mathcal{V}_z de z tel que sur ce voisinage $\nabla_z v_z$ existe et soit égal à :

$$\begin{aligned} \nabla_z v_z &= \left(\text{Hess}_v J_{\varepsilon, z}(v_z) \right)^{-1} DT'(v_z)^t \\ &= \left(\varepsilon \Lambda + D(T'(v_z)^t T'(v_z) + (T(v_z) - z)^t T''(v_z)) \right)^{-1} DT'(v_z)^t \end{aligned}$$

où D est comme avant, la matrice diagonale composée des poids $\alpha_1^2, \dots, \alpha_n^2$. D'autre part, nous avons :

$$B_\varepsilon(z) = T(v_z),$$

donc B_ε est dérivable par rapport à z sur le voisinage \mathcal{V}_z et

$$B'_\varepsilon(z) = T'(v_z) \nabla_z v_z.$$

Et donc,

$$B'_\varepsilon(z) = T'(v_z) (\varepsilon \Lambda + W)^{-1} D(T'(v_z))^t$$

où $T'(v)$ désigne la matrice jacobienne $n \times m$ de T :

$$(T'(v))_{ij} = \frac{\partial T_i}{\partial v_j}(v),$$

et où W est la matrice $m \times m$ dont les éléments w_{kl} vérifient :

$$w_{kl} = \sum_{i=1}^n \alpha_i^2 \left(\frac{\partial T_i}{\partial v_k}(v_z) \frac{\partial T_i}{\partial v_l}(v_z) + (T_i(v_z) - z_i) \frac{\partial^2 T_i}{\partial v_k \partial v_l}(v_z) \right).$$

Ceci bien sûr, si l'on sait calculer les dérivées première et seconde de T et si elles existent.

2.3 Minimisation des fonctions de validation croisée

Dans la pratique, il s'agit de trouver le minimum d'une fonction de validation croisée \mathcal{V} dont nous ne connaissons pas l'expression analytique, mais dont nous pouvons calculer les valeurs qu'elle prend. Nous avons donc choisi d'employer la méthode de la *section dorée*, qui possède l'avantage de ne pas demander trop de valeurs de \mathcal{V} . Dans la pratique, nous supposons que ε appartient à l'intervalle $[0, 1]$, et au lieu de chercher un minimum de $\mathcal{V}(\varepsilon)$, nous allons minimiser $\tilde{\mathcal{V}}(\alpha) = \mathcal{V}(10^\alpha)$ avec α dans l'intervalle $[-15, 0]$. L'algorithme est le suivant :

```

 $\tau = \frac{\sqrt{5}-1}{2}$ 
 $a = -15; b = 0$ 
 $x_1 = b - \tau(b - a); x_2 = a + \tau(b - a)$ 
Tant que  $b - a > 0.05$  faire :
  si  $\tilde{\mathcal{V}}(x_2) > \tilde{\mathcal{V}}(x_1)$  alors
     $b = x_2; x_2 = x_1$ 
     $x_1 = b - \tau(b - a)$ 
  sinon
     $a = x_1; x_1 = x_2$ 
     $x_2 = a + \tau(b - a)$ 
  fin si
Fin tant que
 $\varepsilon = 10^{-\frac{a+b}{2}}$ 
Fin.
```

Cette méthode est intéressante du point de vue numérique puisqu'à chaque itération, elle consiste en une comparaison de quatre valeurs de $\tilde{\mathcal{V}}(\alpha)$, dont il n'y en a qu'une seule à calculer, les autres l'ayant été précédemment. Nous avons constaté d'autre part que le fait de travailler sur l'exposant α plutôt que sur ε accélérât les calculs et tendait à rendre la fonction à minimiser plus convexe. Bien qu'étant une recherche locale, comme nous le verrons par la suite, les résultats obtenus sont plutôt satisfaisants.

3 Application à l'identification de la diffusion

Dans ce cas (cf chapitre II, paragraphe 1) notre modèle s'écrit :

$$z = T(v) + \nu$$

avec $T(v) = y(v)$ où $y(v)$ est la solution de :

$$\left\{ \begin{array}{l} \frac{\partial y}{\partial t}(x, t) - \frac{\partial}{\partial x} \left(v(x) \frac{\partial y}{\partial x}(x, t) \right) = f(x, t), \quad \forall x \in]0, 1[, \forall t \in]0, T[\\ y(x, 0) = y_0(x), \quad \forall x \in]0, 1[\\ v(0) \frac{\partial y}{\partial x}(0, t) = d_0(t), \quad v(1) \frac{\partial y}{\partial x}(1, t) = d_1(t), \quad \forall t \in]0, T[\end{array} \right.$$

et la fonctionnelle lissante :

$$J_{\varepsilon, z}(v) = \frac{1}{2} \| D^{\frac{1}{2}}(y(v) - z) \|^2 + \frac{\varepsilon}{2} v^t \Lambda v.$$

Puisque v_z minimise $J_{\varepsilon, z}$, on a nécessairement :

$$\partial_v J_{\varepsilon, z}(v_z) = 0, \quad \forall z.$$

Or,

$$\partial_v J_{\varepsilon, z}(v) = D(\partial_v y(v))^t (y(v) - z) + \varepsilon \Lambda v$$

et donc, formellement :

$$\text{Hess} J_{\varepsilon, z}(v) = D(\text{Hess}_v y(v)) (y(v) - z) + D(\partial_v y(v))^t \partial_v y(v) + \varepsilon \Lambda.$$

On voit donc que la matrice $\varepsilon \Lambda + W$ précédente n'est rien d'autre que le hessien de la fonctionnelle à minimiser au point v_z , et donc :

$$B'_\varepsilon(z) = (\partial_v y(v_z)) \left(\text{Hess}_v J_{\varepsilon, z}(v_z) \right)^{-1} D(\partial_v y(v_z))^t.$$

Le calcul de cette dérivée est donc envisagé sous deux aspects. Dans la suite, nous prendrons toujours les poids égaux à 1, de telle sorte que $D = I_n$.

3.1 Calcul direct des dérivées première et seconde

On considère pour cela l'opérateur linéaire L_v défini pour $v \in \mathcal{U}_{\text{ad}}$ avec $\mathcal{U}_{\text{ad}} = \{v \in \mathcal{C}^1 / v \text{ et } v' \text{ bornées, et } \exists \alpha, \beta > 0 \text{ t.q. } 0 < \alpha \leq v(x) \leq \beta, \forall x \in [0, 1]\}$ de $L^2(0, T; V)$ dans $L^2(0, T; V')$ par :

$$L_v = \frac{\partial}{\partial t} - \frac{\partial}{\partial x} \left(v(x) \frac{\partial}{\partial x} \right).$$

Soit y_v la solution du problème :

$$\left\{ \begin{array}{l} L_v y_v = f \quad \text{dans }]0, 1[\times]0, T[\\ y_v(x, 0) = y_0(x), \quad \forall x \in]0, 1[\\ v(0) \frac{\partial y_v}{\partial x}(0, t) = d_0(t), \quad \forall t \in]0, T[\\ v(1) \frac{\partial y_v}{\partial x}(1, t) = d_1(t), \quad \forall t \in]0, T[. \end{array} \right.$$

Nous considérons ensuite une petite variation $\delta_k = \delta N_k^4$ dans la k -ème direction de base, nous supposons que $v + \delta_k \in \mathcal{U}_{\text{ad}}$ et soit $y_{v+\delta_k}$ la solution de :

$$\left\{ \begin{array}{l} L_{v+\delta_k} y_{v+\delta_k} = f \\ y_{v+\delta_k}(x, 0) = y_0(x) \\ (v + \delta_k)(0) \frac{\partial y_{v+\delta_k}}{\partial x}(0, t) = d_0(t) \\ (v + \delta_k)(1) \frac{\partial y_{v+\delta_k}}{\partial x}(1, t) = d_1(t) \end{array} \right.$$

Considérons ensuite $h_{v,k}(\delta) = \frac{y_{v+\delta_k} - y_v}{\delta}$; en remarquant que

$$L_{v+\delta_k} = L_v - \delta \frac{\partial}{\partial x} \left(N_k^4(x) \frac{\partial}{\partial x} \right),$$

nous obtenons $h_{v,k}(\delta)$ comme la solution de :

$$\left\{ \begin{array}{l} L_v h_{v,k}(\delta) = \frac{\partial}{\partial x} \left(N_k^4(x) \frac{\partial y_{v,k}(\delta)}{\partial x} \right) \\ h_{v,k}(\delta)(x, 0) = 0 \\ v(0) \frac{\partial h_{v,k}(\delta)}{\partial x}(0, t) = -N_k^4(0) \frac{\partial y_{v+\delta_k}}{\partial x}(0, t) \\ v(1) \frac{\partial h_{v,k}(\delta)}{\partial x}(1, t) = -N_k^4(1) \frac{\partial y_{v+\delta_k}}{\partial x}(1, t) \end{array} \right.$$

ou encore sous la forme faible suivante :

$$\begin{cases} \left(\frac{dh_{v,k}(\delta)(t)}{dt}, \phi \right) + a_v(h_{v,k}(\delta)(t), \phi) = -a_{N_k^4}(y_{v+\delta_k}, \phi), \forall \phi \in V \\ h_{v,k}(\delta)(0) = 0 \end{cases}$$

avec $a_v(\phi, \psi) = \int_0^T \int_{\Omega} v(x) \frac{\partial \phi}{\partial x} \frac{\partial \psi}{\partial x} dx dt$. Pour étudier la limite de $h_{v,k}(\delta)$ lorsque δ tend vers 0, nous rappelons un théorème dû à G. Chavent et P. Lemonnier [CL 74] : soit y_v la solution de :

$$(3.1) \quad \begin{cases} \left(\frac{dy(t)}{dt}, \phi \right) + a_v(y(t), \phi) = (f(t), \phi), \forall \phi \in V, \text{ p.p sur }]0, T[\\ y(0) = y_0, \end{cases}$$

alors ils ont montré le résultat suivant :

Théorème 3.1 Soit y_v la solution de (3.1) , si $f \in L^\infty(0, T; H)$ et si $y_0 \in L^\infty(\Omega)$ alors y_v appartient à $L^\infty(0, T; H)$ et $v \mapsto y_v$ est une application continue de \mathcal{U}_{ad} dans $L^2(0, T; V)$ fortement. Si de plus, au point v considéré, $\frac{\partial y}{\partial x} \in L^\infty(\Omega \times]0, T[)$, alors y_v est Fréchet-dérivable au point v dans les mêmes espaces.

Nous déduisons de ce résultat que si $f \in L^\infty(0, T; H)$ et si $y_0 \in L^\infty(\Omega)$ alors $v \mapsto y_v$ est continue de \mathcal{U}_{ad} dans $L^2(0, T; V)$ et donc, lorsque $\delta \rightarrow 0$, $h_{v,k}(\delta)$ tend vers $h_{v,k}$, la dérivée de y_v par rapport à v_k , qui est définie comme étant la solution de :

$$(3.2) \quad \begin{cases} \left(\frac{dh_{v,k}(t)}{dt}, \phi \right) + a_v(h_{v,k}(t), \phi) = -a_{N_k^4}(y_v(t), \phi), \forall \phi \in V, \text{ p.p sur }]0, T[\\ h_{v,k}(0) = 0, \end{cases}$$

d'où l'on déduit :

$$\begin{cases} L_v h_{v,k} = \frac{\partial}{\partial x} \left(N_k^4(x) \frac{\partial y_v}{\partial x} \right) \text{ dans }]0, 1[\times]0, T[\\ h_{v,k}(x, 0) = 0 \\ v(0) \frac{\partial h_{v,k}}{\partial x}(0, t) = -N_k^4(0) \frac{\partial y_v}{\partial x}(0, t) \\ v(1) \frac{\partial h_{v,k}}{\partial x}(1, t) = -N_k^4(1) \frac{\partial y_v}{\partial x}(1, t). \end{cases}$$

Considérons $h_{v+\delta_l, k}$ solution du système :

$$\left\{ \begin{array}{l} L_{v+\delta_l} h_{v+\delta_l, k} = \frac{\partial}{\partial x} \left(N_k^4(x) \frac{\partial y_{v+\delta_l}}{\partial x} \right) \\ h_{v+\delta_l, k}(x, 0) = 0 \\ (v + \delta_l)(0) \frac{\partial h_{v+\delta_l, k}}{\partial x}(0, t) = -N_k^4(0) \frac{\partial y_{v+\delta_l}}{\partial x}(0, t) \\ (v + \delta_l)(1) \frac{\partial h_{v+\delta_l, k}}{\partial x}(1, t) = -N_k^4(1) \frac{\partial y_{v+\delta_l}}{\partial x}(1, t). \end{array} \right.$$

Comme précédemment, nous posons

$$w_{v, k, l}(\delta) = \frac{h_{v+\delta_l} - h_{v, k}}{\delta}$$

alors $w_{v, k, l}(\delta)$ est obtenu comme la solution de :

$$\left\{ \begin{array}{l} L_v w_{v, k, l}(\delta) = \frac{\partial}{\partial x} \left(N_l^4(x) \frac{\partial h_{v+\delta_l, k}}{\partial x} \right) + \frac{\partial}{\partial x} \left(N_k^4(x) \frac{\partial h_{v, l}}{\partial x} \right) \\ w_{v, k, l}(\delta)(x, 0) = 0 \\ v(0) \frac{\partial w_{v, k, l}(\delta)}{\partial x}(0, t) = -N_l^4(0) \frac{\partial h_{v+\delta_l, k}}{\partial x}(0, t) - N_k^4(0, t) \frac{\partial h_{v, l}(\delta)}{\partial x}(0, t) \\ v(1) \frac{\partial w_{v, k, l}(\delta)}{\partial x}(1, t) = -N_l^4(1) \frac{\partial h_{v+\delta_l, k}}{\partial x}(1, t) - N_k^4(1, t) \frac{\partial h_{v, l}(\delta)}{\partial x}(1, t), \end{array} \right.$$

d'où la formulation faible :

$$\left\{ \begin{array}{l} \left(\frac{dw_{v, k, l}(\delta)}{dt}, \phi \right) + a_v(w_{v, k, l}(\delta), \phi) = -a_{N_l^4}(h_{v+\delta_l, k}, \phi) - a_{N_k^4}(h_{v, l}(\delta), \phi) \\ w_{v, k, l}(\delta)(0) = 0. \end{array} \right.$$

Lorsque $\delta \rightarrow 0$, $h_{v, k}(\delta) \rightarrow h_{v, k}$ dans $L^2(0, T; V)$. Supposons que $v \mapsto (h_{v, k})_k$ soit dérivable, et posons

$$w_{v, k, l} = \lim_{\delta \rightarrow 0} w_{v, k, l}(\delta)$$

alors, $w_{v, k, l}$, dérivée seconde de y_v par rapport à v_k et v_l , est obtenue comme la solution de :

$$\left\{ \begin{array}{l} \left(\frac{dw_{v, k, l}}{dt}, \phi \right) + a_v(w_{v, k, l}(\delta), \phi) = -a_{N_l^4}(h_{v, k}, \phi) - a_{N_k^4}(h_{v, l}, \phi) \\ w_{v, k, l}(0) = 0. \end{array} \right.$$

Et donc, la dérivée seconde de y par rapport à v_k, v_l , si elle existe, s'obtiendra comme solution de :

$$\left\{ \begin{array}{l} L_v w_{v,k,l} = \frac{\partial}{\partial x} \left(N_k^4(x) \frac{\partial h_{v,l}}{\partial x} + N_l^4(x) \frac{\partial h_{v,k}}{\partial x} \right) \\ w_{v,k,l}(x, 0) = 0 \\ v(0) \frac{\partial w_{v,k,l}}{\partial x}(0, t) = -N_k^4(0) \frac{\partial h_{v,l}}{\partial x}(0, t) - N_l^4(0) \frac{\partial h_{v,k}}{\partial x}(0, t) \\ v(1) \frac{\partial w_{v,k,l}}{\partial x}(1, t) = -N_k^4(1) \frac{\partial h_{v,l}}{\partial x}(1, t) - N_l^4(1) \frac{\partial h_{v,k}}{\partial x}(1, t). \end{array} \right.$$

Il est donc possible, sous certaines hypothèses, de calculer explicitement les dérivées de y par rapport au contrôle, cependant, lorsqu'on a beaucoup de paramètres, il faut alors inverser m systèmes pour obtenir la dérivée première et m^2 systèmes pour la dérivée seconde. On propose donc une autre approche, qui elle ne nécessite que le calcul de $T'(v)\phi$ et de $\psi^t T''(v)\phi$ qui sont nettement moins coûteuses.

3.2 Calcul de $B'_\varepsilon(z)$ par le hessien de la fonction-coût

Nous avons vu précédemment que :

$$B'_\varepsilon(z) = \partial_v y(v_z) \left(\text{Hess}_v J_{\varepsilon,z}(v_z) \right)^{-1} (\partial_v y(v_z))^t$$

et que le calcul explicite des composantes de $\partial_v y(v)$ et de $\frac{\partial^2 y(v)}{\partial v^2}$ est très coûteux. Dans l'évaluation de la fonction de validation croisée généralisée $\mathcal{V}(\varepsilon)$, on a seulement besoin de calculer $\text{Tr}(B'_\varepsilon(z))$ et donc de $w^t B'_\varepsilon(z) w$ où w est un vecteur aléatoire de \mathbb{R}^n si l'on utilise une méthode de Monte-Carlo :

$$w^t B'_\varepsilon(z) w = \left((\partial_v y(v_z))^t w \right)^t \left(\text{Hess} J_{\varepsilon,z}(v_z) \right)^{-1} (\partial_v y(v_z))^t w.$$

Il suffit donc de calculer $\xi = (\partial_v y(v_z))^t w$, puis $\xi^t \left(\text{Hess} J_{\varepsilon,z}(v_z) \right)^{-1} \xi$. Pour cette dernière expression, il est nécessaire de savoir calculer $\left(\text{Hess} J_{\varepsilon,z}(v_z) \right) \eta$, puis on calcule ensuite $\xi \left(\text{Hess} J_{\varepsilon,z}(v_z) \right)^{-1} \xi$ par un algorithme de gradient conjugué.

Proposition 3.2 Soit $w \in \mathbb{R}^n$, alors $(\partial_v y(v_z))^t w$ est obtenu par :

$$\forall i = 1, \dots, m, \quad \left[(\partial_v y(v_z))^t w \right]_i = \int_{\Omega} N_i^4(x) \left(\frac{\partial y}{\partial x} \frac{\partial \tilde{p}}{\partial x} - \frac{\partial p}{\partial x} \frac{\partial \tilde{y}}{\partial x} \right) dx$$

avec $y = y(v_z)$, p est l'état-adjoint associé et où \tilde{y} et \tilde{p} sont les solutions de :

$$(3.3) \quad \begin{cases} \left(\frac{d\tilde{y}}{dt}(t), \phi \right) + a_v(\tilde{y}(t), \phi) = -a_{\tilde{v}}(y(t), \phi), & \forall \phi \in V \\ \tilde{y}(0) = 0; \end{cases}$$

$$(3.4) \quad \begin{cases} \left(-\frac{d\tilde{p}}{dt}(t), \phi \right) + a_v(\tilde{p}(t), \phi) = -a_{\tilde{v}}(p(t), \phi) - (w, \phi), & \forall \phi \in V \\ \tilde{p}(T) = 0. \end{cases}$$

Preuve : On remarque facilement que \tilde{y} vérifiant (3.3) est la variation obtenue sur l'état lorsqu'on a une variation de \tilde{v} sur le contrôle :

$$v \rightarrow v + \tilde{v}, \quad y \rightarrow y + \tilde{y}$$

et donc, on peut écrire : $\tilde{y} \simeq (\partial_v y(v))\tilde{v}$. Remplaçons dans (3.4) la fonction arbitraire ϕ par $\tilde{y}(t)$, il vient après intégration par parties sur $]0, T[$:

$$\left(\tilde{p}(t), \frac{d\tilde{y}}{dt}(t) \right) + a_v(\tilde{p}(t), \tilde{y}(t)) + a_{\tilde{v}}(p(t), \tilde{y}(t)) = -(w, \tilde{y}(t))$$

soit encore, en écrivant que \tilde{y} est solution de (3.3) :

$$a_{\tilde{v}}(y(t), \tilde{p}(t)) - a_{\tilde{v}}(p(t), \tilde{y}(t)) = (w, \tilde{y}(t)).$$

En remarquant que $\tilde{v}(x) = \sum \tilde{v}_i N_i^4(x) = \tilde{v}^t N$, on trouve finalement :

$$\begin{aligned} (w, \tilde{y}(t)) &= \langle \tilde{v}, a_N(y(t), \tilde{p}(t)) - a_N(p(t), \tilde{y}(t)) \rangle_{\mathbb{R}^m} \\ &= \langle w, (\partial_v y(v))\tilde{v} \rangle_{\mathbb{R}^n} \\ &= \langle (\partial_v y(v))^t w, \tilde{v} \rangle_{\mathbb{R}^m} \end{aligned}$$

d'où finalement :

$$\left[(\partial_v y(v))^t w \right]_i = \int_{\Omega} N_i^4(x) \left(\frac{\partial y}{\partial x} \frac{\partial \tilde{p}}{\partial x} - \frac{\partial p}{\partial x} \frac{\partial \tilde{y}}{\partial x} \right) dx, \quad i = 1, \dots, m. \quad \square$$

Remarque 3.3 : Puisque \tilde{y} et \tilde{p} sont définis de manière unique, il en découle que $(\partial_v y(v))^t w$ est bien défini pour tout $w \in \mathbb{R}^n$.

Lorsqu'on fait subir une variation \tilde{v} à la diffusion, nous avons vu que l'état varie de \tilde{y} solution de (3.3) ; regardons ce que devient l'état-adjoint : $p \rightarrow p + \tilde{p}$. Or $p(t)$ est solution de :

$$\begin{cases} \left(-\frac{dp}{dt}(t), \phi \right) + a_v(p(t), \phi) = (z(t) - y(t), \phi), & \forall \phi \in V \\ p(T) = 0 \end{cases}$$

donc, on obtient \tilde{p} comme la solution de :

$$(3.5) \quad \begin{cases} \left(-\frac{d\tilde{p}}{dt}(t), \phi \right) + a_v(\tilde{p}(t), \phi) = -a_{\tilde{v}}(p(t), \phi) - (\tilde{y}(t), \phi), & \forall \phi \in V \\ \tilde{p}(T) = 0. \end{cases}$$

Si l'on écrit de même que le gradient subit une variation $\tilde{\nabla} J$, alors puisque

$$(\nabla J(v))_i = \varepsilon(\Lambda v)_i + a_{N_i^4}(y(t), p(t)),$$

on en déduit :

$$(\tilde{\nabla} J(v))_i = \varepsilon(\Lambda \tilde{v})_i + a_{N_i^4}(y, \tilde{p}) + a_{N_i^4}(\tilde{y}, p).$$

D'autre part, puisque ∇J est une fonction de v uniquement, nous avons :

$$(3.6) \quad (\tilde{\nabla} J(v))_i = [(\text{Hess} J(v))^t \tilde{v}]_i = \varepsilon(\Lambda \tilde{v})_i + a_{N_i^4}(y, \tilde{p}) + a_{N_i^4}(\tilde{y}, p).$$

Puisque \tilde{y} et \tilde{p} sont définis de façon unique, $(\text{Hess} J(v))^t \tilde{v}$ l'est aussi. Nous savons donc à présent calculer $(\text{Hess} J(v))^t \tilde{v}$. Il ne reste plus qu'à transposer ceci afin d'avoir l'expression de $(\text{Hess} J(v))\eta$ pour η vecteur de \mathbb{R}^m .

Proposition 3.4 Soit $\eta \in \mathbb{R}^m$, alors $(\text{Hess} J(v))\eta$ est donné par :

$$(3.7) \quad [(\text{Hess} J(v))\eta]_i = \varepsilon(\Lambda \eta)_i + a_{N_i^4}(y, Q) + a_{N_i^4}(p, Y)$$

où Y et Q sont donnés par :

$$(3.8) \quad \begin{cases} \left(\frac{dY}{dt}(t), \phi \right) + a_v(Y(t), \phi) = -a_\eta(y(t), \phi), & \forall \phi \in V, \\ Y(0) = 0; \end{cases}$$

$$(3.9) \quad \begin{cases} \left(-\frac{dQ}{dt}(t), \phi \right) + a_v(Q(t), \phi) = -a_\eta(p(t), \phi) - (Y(t), \phi), & \forall \phi \in V, \\ Q(T) = 0. \end{cases}$$

Preuve : Ecrivons (3.9) en prenant comme fonction $\phi = \tilde{y}(t)$ solution de (3.3) et (3.8) avec $\phi = \tilde{p}(t)$ solution de (3.5), alors nous avons :

$$\begin{aligned} & \left(\frac{dY}{dt}(t), \tilde{p}(t) \right) + a_v(Y(t), \tilde{p}(t)) + a_\eta(y(t), \tilde{p}(t)) \\ & + \left(-\frac{dQ}{dt}(t), \tilde{y}(t) \right) + a_v(Q(t), \tilde{y}(t)) + a_\eta(p(t), \tilde{y}(t)) + (Y(t), \tilde{y}(t)) = 0, \end{aligned}$$

soit encore, en intégrant par parties sur $]0, T[$:

$$\begin{aligned} & \left(-\frac{d\tilde{p}}{dt}(t), Y(t) \right) + a_v(Y(t), \tilde{p}(t)) + a_\eta(y(t), \tilde{p}(t)) \\ & + \left(\frac{d\tilde{y}}{dt}(t), Q(t) \right) + a_v(\tilde{y}(t), Q(t)) + a_\eta(p(t), \tilde{y}(t)) + (Y(t), \tilde{y}(t)) = 0. \end{aligned}$$

Comme $\tilde{p}(t)$ est solution de (3.5) et $\tilde{y}(t)$ de (3.3), il reste :

$$-a_{\tilde{v}}(p(t), Y(t)) - a_{\tilde{v}}(y(t), Q(t)) + a_\eta(y(t), \tilde{p}(t)) + a_\eta(p(t), \tilde{y}(t)) = 0,$$

d'où l'on tire :

$$(3.10) \quad a_\eta(y(t), \tilde{p}(t)) + a_\eta(p(t), \tilde{y}(t)) = a_{\tilde{v}}(p(t), Y(t)) + a_{\tilde{v}}(y(t), Q(t)).$$

Effectuons ensuite le produit scalaire de (3.6) par η :

$$\langle \eta, (\text{Hess}J(v))^t \tilde{v} \rangle = \langle \eta, \varepsilon \Lambda \tilde{v} \rangle + a_\eta(y, \tilde{p}) + a_\eta(\tilde{y}, p)$$

d'après (3.10), on déduit que :

$$\langle \text{Hess}J(v)\eta, \tilde{v} \rangle = \langle \varepsilon \Lambda \eta, \tilde{v} \rangle + \langle \tilde{v}, a_N(p, Y) + a_N(y, Q) \rangle. \quad \square$$

Remarque 3.5 : Les variables (Y, Q) peuvent être interprétées en termes de variables adjointes du couple (\tilde{p}, \tilde{y}) .

Remarque 3.6 : Le hessien du critère est symétrique :

$$(\text{Hess}J(v))^t \tilde{v} = (\text{Hess}J(v))\tilde{v}.$$

En effet, pour $\eta = \tilde{v}$, on a : $Y = \tilde{y}$ et $Q = \tilde{p}$.

3.3 Calcul de $\xi^t(\text{Hess}J(v))^{-1}\xi$ pour $\xi \in \mathbb{R}^m$

On considère pour cela la fonctionnelle :

$$\mathcal{J}(\tilde{v}) = \frac{1}{2}(\tilde{v}, \text{Hess}J(v)\tilde{v}) - (\xi, \tilde{v});$$

son gradient est donné par :

$$\nabla\mathcal{J}(\tilde{v}) = (\text{Hess}J(v))\tilde{v} - \xi.$$

\mathcal{J} étant une fonctionnelle quadratique, le minimum v^* réalisera la condition :

$$v^* = (\text{Hess}J(v))^{-1}\xi.$$

En résumé, on peut écrire :

$$w^t B'_\varepsilon(z)w = \xi^t v^* \text{ avec :}$$

$$\xi = a_N(y, \tilde{p}) - a_N(p, \tilde{y}),$$

$$\text{où } \tilde{y} \text{ est solution de (3.3) , } \tilde{p} \text{ est solution de (3.4) .}$$

Cette méthode est nettement plus avantageuse que la précédente pour le calcul de $\text{Tr}(B'_\varepsilon(z))$ puisqu'elle demande seulement la résolution des systèmes vérifiés par $y, p, \tilde{y}, \tilde{p}$, puis la résolution du problème de minimisation à m composantes. Le coût sera donc essentiellement celui de la résolution de m systèmes au lieu de $m^2 + m$ avec la méthode précédente.

Il est à noter que les systèmes vérifiés par y et \tilde{y} admettent la même matrice de rigidité A , et p et \tilde{p} ont pour matrice A^t ; pour inverser ces quatre systèmes, nous ferons un seul calcul de factorisation de la matrice A . C'est donc cette méthode que l'on préférera utiliser, d'autant plus que le nombre m de paramètres est plus élevé.

3.4 Résultats numériques

Le niveau de bruit sur les données est de l'ordre de 2%. La méthode de validation croisée ordinaire nécessite environ 1 heure de calcul sur un CRAY lorsqu'on enlève les observations une par une. Dans la pratique, nous procédons de la façon suivante : pour chaque instant, nous enlevons toutes les observations spatiales (ce qui représente des paquets de 10), et ceci diminue considérablement

le temps de calcul (20 minutes CPU). Il s'agit donc d'une validation croisée en temps.

La GCV quant à elle, converge généralement en moins de 3 minutes. C'est cette méthode que l'on préférera utiliser, d'autant plus que le nombre d'observations est élevé. Nous donnons une comparaison entre validation croisée ordinaire et validation croisée généralisée pour chaque fonction identifiée. Pour les courbes suivantes, nous avons représenté à gauche la diffusion ρ identifiée par validation croisée ordinaire (OCV) et ρ_{ex} , et à droite la diffusion ρ identifiée par validation croisée généralisée (GCV) et ρ_{ex} .

L'estimation de la trace de la matrice $B'_\varepsilon(z)$ est faite à l'aide d'une dizaine de réalisations de Monte Carlo.

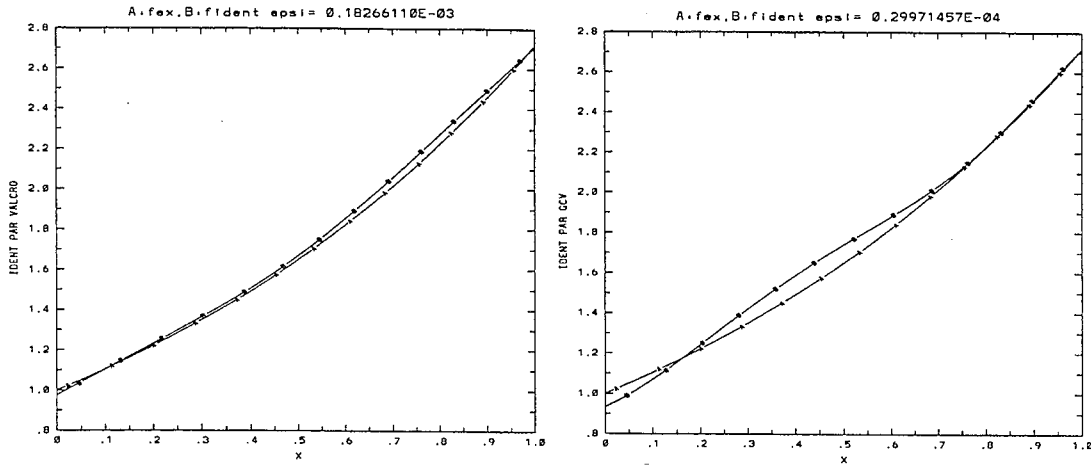


Figure 3.1 Comparaison OCV (à gauche)-GCV (à droite) pour e^x

Dans le cas de la fonction e^x , figure 3.1, les valeurs obtenues pour ε sont : $\varepsilon_{OCV} = 1.8 \times 10^{-4}$ et $\varepsilon_{GCV} = 2.9 \times 10^{-5}$. La valeur obtenue par OCV est donc la meilleure et donne une solution plus régulière.

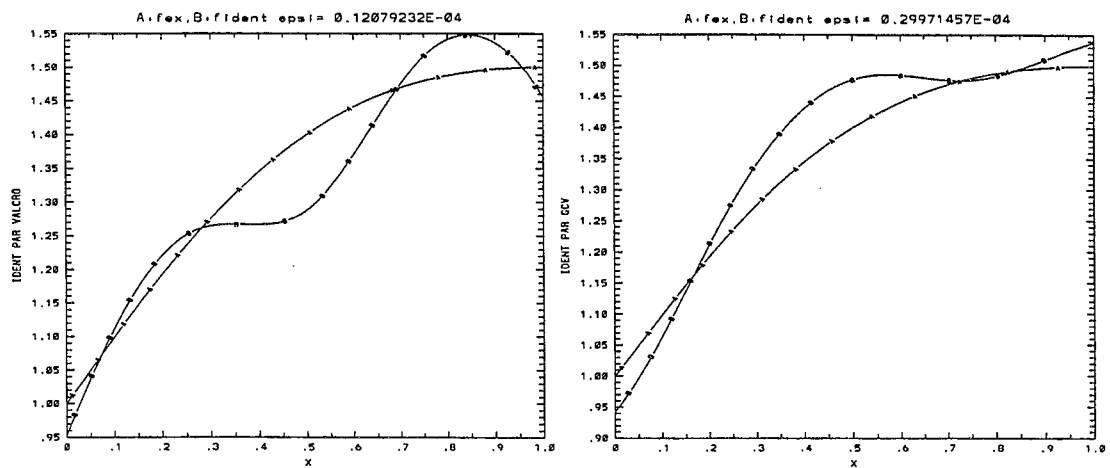


Figure 3.2 Comparaison OCV (à gauche)-GCV (à droite) pour $\frac{x^2+x+1}{x^2+1}$

Pour $\frac{x^2+x+1}{x^2+1}$, figure 3.2, $\varepsilon_{OCV} = 1.2 \times 10^{-5}$ et $\varepsilon_{GCV} = 2.9 \times 10^{-5}$, ce qui explique que la solution obtenue par la GCV soit plus régulière que par OCV. Nous constatons d'autre part, que bien que ces deux valeurs du paramètre de régularisation ne soient pas très éloignées l'une de l'autre, les résultats sur l'identification font apparaître des solutions relativement éloignées : ceci illustre bien à quel point le choix de ε est crucial pour l'identification.

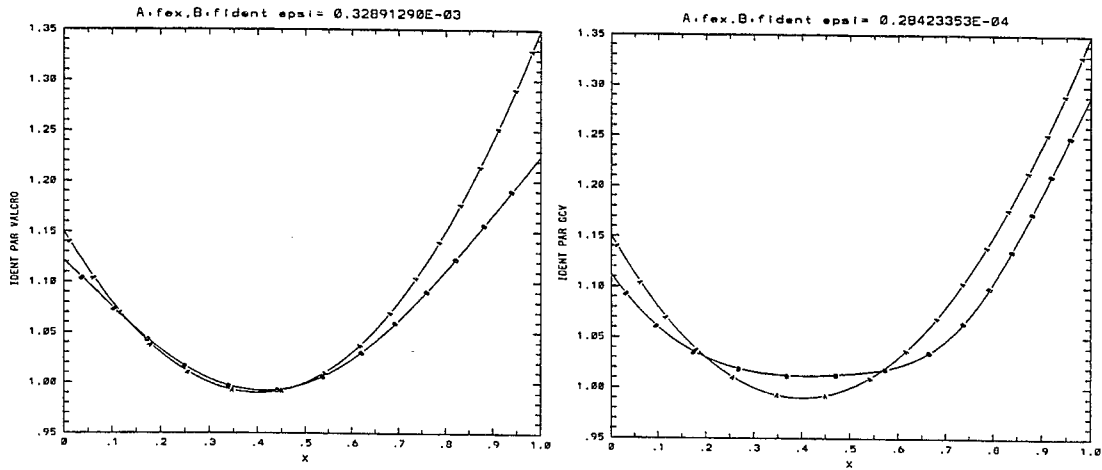


Figure 3.3 Comparaison OCV (à g.)-GCV (à dr.) pour $(x - 0.3)(x - 0.5) + 1$

Dans le cas de ce polynôme de degré 2, figure 3.3, $\varepsilon_{OCV} = 3.3 \times 10^{-4}$ et $\varepsilon_{GCV} = 2.8 \times 10^{-5}$.

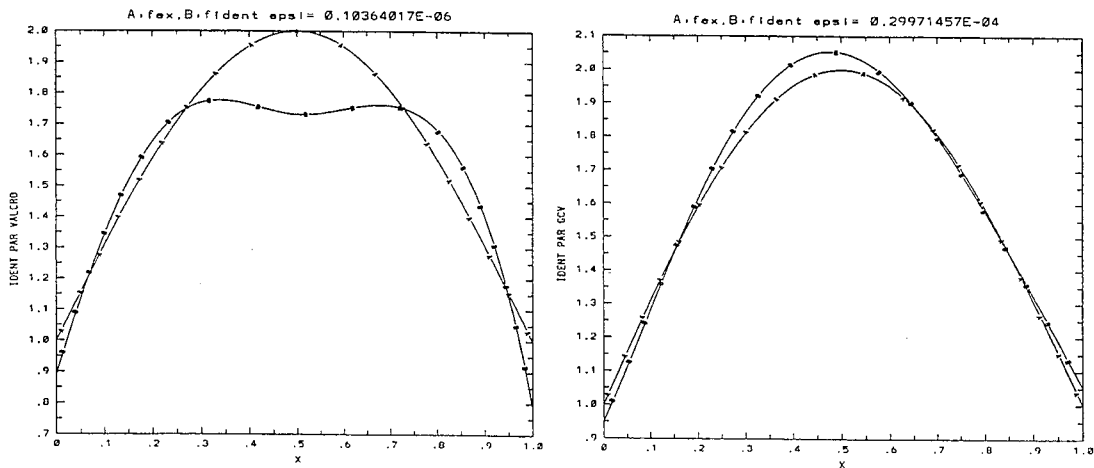


Figure 3.4 Comparaison OCV (à gauche)-GCV (à droite) pour $\sin(\pi x) + 1$

Dans le cas de la fonction $\sin(\pi x) + 1$, figure 3.4, $\varepsilon_{OCV} = 1.0 \times 10^{-7}$, alors que $\varepsilon_{GCV} = 3.0 \times 10^{-5}$. Dans ce cas, nous pouvons remarquer que les deux valeurs

de ε sont assez éloignées l'une de l'autre. Les résultats de l'identification tendent à montrer qu'ici, la validation croisée ordinaire (OCV) ne donne pas une bonne valeur du paramètre: ε_{OCV} est très petit, relativement à ε_{GCV} , d'où une solution identifiée moins "lisse" qu'avec la GCV.

En conclusion, au vu de ces résultats, il n'est donc pas possible d'affirmer que l'une ou l'autre de ces deux méthodes soit la meilleure pour ce problème d'identification. En effet, il semble que ce soit la fonction à identifier qui détermine la méthode la mieux adaptée au problème. Une solution pratique, lorsqu'on ne connaît pas la forme exacte de la diffusion, pourrait être de calculer ε avec ces deux méthodes, et de retenir la solution identifiée avec une valeur de ε correspondant au milieu de l'intervalle $[\min(\varepsilon_{OCV}, \varepsilon_{GCV}), \max(\varepsilon_{OCV}, \varepsilon_{GCV})]$.

4 Application à l'identification de sources non linéaires

Dans ce cas, nous avons vu au chapitre II, paragraphe 2, que notre modèle s'écrit :

$$z = T(v) + \nu$$

avec $T : v \mapsto \left(\frac{\partial y(v)}{\partial n} \right)_{/\partial\Omega}$ où z sont les observations, ν est une variable aléatoire de loi $\mathcal{N}(0, \sigma)$ qui représente le bruit, et où l'état $y(v)$ est donné comme étant solution du système :

$$(4.1) \quad \begin{cases} -\Delta y = v(\bar{y}) = vob(y) & \text{dans } \Omega, \\ y = 0 & \text{sur } \partial\Omega. \end{cases}$$

Nous avons vu que moyennant l'hypothèse : pour tout $y \in V^{p_0}(\Omega)$, $v(\bar{y}) \in L^2(\Omega)$, nous pouvons affirmer que y solution de (4.1) était dans $H^2(\Omega)$, et donc $\frac{\partial y}{\partial n} \in H^{\frac{1}{2}}(\partial\Omega) \subset L^2(\partial\Omega)$. Donc, nous pouvons voir l'application T comme étant définie de \mathcal{U}_{ad} dans $H^{\frac{1}{2}}(\partial\Omega)$.

Pour calculer la fonction de validation croisée généralisée, il va nous falloir calculer la quantité :

$$B'_\varepsilon(z) = T'(v_z)[\text{Hess}J(v_z)]^{-1}(T'(v_z))^t.$$

Comme nous l'avons vu au paragraphe 1.5, si nous savons calculer $T'(v_z)^t w$ pour un vecteur w de \mathbb{R}^n et $(\text{Hess}J(v_z))\eta$ pour η un vecteur de \mathbb{R}^m , alors nous saurons évaluer $\text{Tr}(B'_\varepsilon(z))$ grâce à la méthode de Monte-Carlo.

4.1 Calcul de $B'_\varepsilon(z)$

Le lemme 2.6 du chapitre II nous assurait (sous certaines hypothèses sur y) la dérivabilité de l'application $b(y)$ au sens de Gâteaux. Nous allons utiliser cette propriété afin de montrer que l'on peut calculer $T'(v)^t w$. Nous supposons par la suite que les hypothèses H1 et H2 du chapitre II, paragraphe 2.1.3 que nous rappelons ci-dessous sont vérifiées par y , solution du problème faible :

(H1) $\max_{\Omega} y$ est atteint en un unique point M_0 intérieur à Ω ,

(H2) y est de classe \mathcal{C}^2 localement dans un voisinage de M_0 , et M_0 est un point elliptique non dégénéré.

Proposition 4.1 Soit $w \in \mathbb{R}^n$, alors $T'(v)^t w$ est défini par :

$$\forall i = 1, \dots, m, \quad (T'(v)^t w)_i = \int_{\Omega} (N_i^4)'(\bar{y})(b'(y)^t p) \tilde{y} dx - \int_{\Omega} N_i^4(\bar{y}) \tilde{p} dx$$

avec $y = y(v)$, p l'état adjoint associé, et où \tilde{y} et \tilde{p} sont définis par :

$$(4.2) \quad \begin{cases} -\Delta \tilde{y} - v'(\bar{y})(b'(y)\tilde{y}) = \tilde{v}(\bar{y}) & \text{dans } \Omega, \\ \tilde{y} = 0 & \text{sur } \partial\Omega; \end{cases}$$

$$(4.3) \quad \begin{cases} -\Delta \tilde{p} - v'(\bar{y})(b'(y)^t \tilde{p}) = \tilde{v}'(\bar{y})(b'(y)^t p) & \text{dans } \Omega, \\ \tilde{p} = w & \text{sur } \partial\Omega. \end{cases}$$

Preuve : On constate que (4.2) est une version linéarisée de (4.1), on en déduit donc que :

$$\tilde{y} = y'(v)\tilde{v}.$$

Or, $T(v) = \frac{\partial y(v)}{\partial n} = \gamma_1(y(v))$ et l'on sait (voir par exemple [LM 68]) que γ_1 est une application linéaire continue de $H^2(\Omega)$ vers $H^{\frac{1}{2}}(\partial\Omega)$, donc on peut écrire :

$$T'(v)\tilde{v} = \gamma_1(y'(v)\tilde{v}) = \gamma_1\tilde{y}.$$

Puisque \tilde{p} vérifie (4.3), en faisant le produit par \tilde{y} et en intégrant par parties sur Ω , on obtient :

$$\int_{\Omega} \nabla \tilde{p} \cdot \nabla \tilde{y} dx - \int_{\Omega} v'(\bar{y})(b'(y)^t \tilde{p}) \tilde{y} dx - \int_{\Omega} \tilde{v}'(\bar{y})(b'(y)^t p) \tilde{y} dx = 0,$$

soit encore :

$$\int_{\Omega} -\Delta \tilde{y} \tilde{p} dx + \int_{\partial\Omega} \gamma_1(\tilde{y}) \tilde{p} d\sigma - \int_{\Omega} v'(\bar{y})(b'(y)\tilde{y}) \tilde{p} dx - \int_{\Omega} \tilde{v}'(\bar{y})(b'(y)^t p) \tilde{y} dx = 0.$$

En écrivant que \tilde{y} est solution de (4.2), on obtient :

$$\int_{\Omega} \tilde{v}(\bar{y}) \tilde{p} dx + \int_{\partial\Omega} w \gamma_1(\tilde{y}) d\sigma = \int_{\Omega} \tilde{v}'(\bar{y})(b'(y)^t p) \tilde{y} dx$$

soit encore, en posant $\tilde{v}(x) = \sum \tilde{v}_i N_i^4(x) = \tilde{v}^t N$:

$$(4.4) \quad \langle w, \gamma_1(\tilde{y}) \rangle = \langle \tilde{v}, \int_{\Omega} N'(\bar{y})(b'(y)^t p) \tilde{y} dx - \int_{\Omega} N(\bar{y}) \tilde{p} dx \rangle.$$

D'autre part,

$$(4.5) \quad \langle T'(v)\tilde{v}, w \rangle = \langle \gamma_1(\tilde{y}), w \rangle = \langle \tilde{v}, T'(v)^t w \rangle.$$

En égalant (4.4) et (4.5), on en déduit le résultat annoncé. \square

Remarque 4.2 : L'existence et l'unicité de \tilde{y} sont assurées dès que l'état y est Gâteaux-dérivable par rapport au contrôle, et \tilde{y} coïncide alors avec cette dérivée. Si cette hypothèse est vérifiée, nous en déduisons immédiatement l'existence et l'unicité de \tilde{p} car il vérifie le système adjoint de \tilde{y} .

Pour calculer le hessien du critère appliqué à un vecteur de \mathbb{R}^m , nous allons avoir besoin de résultats supplémentaires concernant la fonction b , en particulier de la Gâteaux-dérivabilité de b à l'ordre deux. Pour cela, nous allons nous placer dans un nouvel espace fonctionnel :

$$W = \{\phi \in W^{3,p_0}(\Omega) / \phi = 0 \text{ sur } \partial\Omega\}$$

beaucoup plus régulier que l'espace $V^{p_0}(\Omega)$ utilisé au chapitre II. Nous supposons que $p_0 > 2$, de telle sorte que $W \subset C^2(\bar{\Omega})$: nous verrons que cette hypothèse est fondamentale pour montrer que b est deux fois Gâteaux-dérivable.

La démarche est la suivante : nous montrerons tout d'abord que pour une fonction y de W vérifiant les hypothèses H1 et H2, lorsqu'on définit $y_\theta = y + \theta\psi$ avec $\psi \in W^{3,p_0}(\Omega)$, alors y_θ vérifie également ces deux hypothèses pourvu que θ soit assez petit. Ceci nous permettra de pouvoir définir $b'(y_\theta)$. Nous étudierons ensuite la limite de rapports $\frac{b'(y+\theta\psi)\phi - b'(y)\phi}{\theta}$ lorsque θ tend vers 0, puis nous prouverons que l'expression obtenue à la limite est continue par rapport à ϕ et ψ dans $W^{3,p_0}(\Omega)$.

Lemme 4.3 Soit $y \in W$ satisfaisant les hypothèses H1 et H2. Pour toute fonction $\psi \in W^{3,p_0}(\Omega)$, soit y_θ défini par $y + \theta\psi$. Si θ est suffisamment petit, alors y_θ atteint son maximum sur Ω en un unique point M_θ . De plus, l'application $\theta \mapsto M_\theta$ est dérivable en 0, de dérivée $\left(\frac{dM_\theta}{d\theta}\right)_{\theta=0} = -(\text{Hess } y(M_0))^{-1} \nabla\psi(M_0)$.

Preuve : Notons par M_0 le point où y atteint son maximum (par hypothèse). Puisque la fonction ψ appartient à $W^{3,p_0}(\Omega)$, elle est de classe C^1 sur $\bar{\Omega}$ qui est compact. Donc ψ atteint son maximum sur $\bar{\Omega}$. On peut en déduire que $y_\theta = y + \theta\psi \in C^1(\bar{\Omega})$ et atteint son maximum sur Ω . Choisissons un relèvement M_θ tel que : $y_\theta(M_\theta) = \max_{\Omega} y_\theta$. Pour que y_θ vérifie l'hypothèse H1, il faut prouver que le maximum de y_θ est atteint en un unique point. Or, par compacité de $\bar{\Omega}$, nous savons qu'un tel point M_θ tend vers M_0 lorsque $\theta \rightarrow 0$. Considérons la fonction F définie sur $\mathbb{R} \times \bar{\Omega}$ par :

$$F : \mathbb{R} \times \bar{\Omega} \rightarrow \mathbb{R}^2$$

$$(\theta, M) \mapsto \nabla(y + \theta\psi)(M).$$

Comme $y + \theta\psi \in W^{3,p_0}(\Omega) \subset C^2(\bar{\Omega})$, la fonction $\nabla(y + \theta\psi) = \nabla y_\theta$ est de classe C^1 sur $\bar{\Omega}$. Pour tout θ assez voisin de 0, nous avons : $\nabla(y_\theta(M_\theta)) = 0$, soit encore :

$$\nabla y + \theta \nabla\psi(M_\theta) = 0, \quad \forall \theta.$$

Nous avons aussi : $F(0, M_0) = \nabla y(M_0) = 0$. Or, l'application qui à M de $\bar{\Omega}$ associe $F(\theta, M)$ pour θ fixé quelconque est de classe \mathcal{C}^1 sur $\bar{\Omega}$, donc :

$$\begin{aligned} M &\mapsto F(\theta, M) \text{ est uniformément continue sur } \bar{\Omega} \\ M &\mapsto \frac{\partial F}{\partial M}(\theta, M) \text{ est uniformément continue sur } \bar{\Omega}. \end{aligned}$$

D'autre part,

$$D_M F(\theta, M) = \text{Hess } y(M) + \theta \text{Hess } \psi(M),$$

donc $D_M F(0, M_0) = \text{Hess } y(M_0)$ qui est inversible par l'hypothèse H2. Le théorème des fonctions implicites nous permet donc d'affirmer qu'il existe un voisinage \mathcal{V} de $\theta = 0$, et une unique fonction g définie sur \mathcal{V} telle que :

$$\begin{aligned} g &: \mathcal{V} \rightarrow \Omega \\ \theta &\mapsto M_\theta. \end{aligned}$$

Donc pour θ assez petit, nous aurons bien unicité du point M_θ en lequel y_θ atteint son maximum. Le théorème des fonctions implicites nous assure également que g est dérivable sur ce voisinage \mathcal{V} et est donnée par :

$$Dg(\theta) = \frac{dM_\theta}{d\theta} = -(D_M F(\theta, M_\theta))^{-1} D_\theta F(\theta, M_\theta), \quad \forall \theta \in \mathcal{V},$$

ce qui achève la preuve. □

Nous montrons à présent que les hypothèses H1 et H2, si elles sont vérifiées pour $y \in W$, le sont aussi pour $y_\theta = y + \theta\psi$ avec $\psi \in W^{3,p_0}(\Omega)$, pourvu que θ soit assez petit.

Lemme 4.4 *Soit $y \in W$ vérifiant H1 et H2. Soit $\psi \in W^{3,p_0}(\Omega)$ et $y_\theta = y + \theta\psi$. Si θ est choisi suffisamment petit, alors y_θ vérifie les hypothèses H1 et H2.* →

Preuve : Nous avons déjà vérifié dans le lemme 4.3 que y_θ satisfait l'hypothèse H1. Appelons M_0 le point où y atteint son maximum, nous avons :

$$\text{Hess } y_\theta = \text{Hess } y + \theta \text{Hess } \psi$$

qui est bien défini puisque $W^{3,p_0}(\Omega) \subset \mathcal{C}^2(\bar{\Omega})$. Comme $\text{Hess } \psi$ est une fonction continue sur un compact $(\bar{\Omega})$, elle est bornée, et donc :

$$\lim_{\theta \rightarrow 0} \text{Hess } y_\theta = \text{Hess } y$$

dans l'espace des fonctions continues sur $\bar{\Omega}$. Par H2, nous savons que $\text{Hess } y(M_0)$ est défini négatif, donc puisque $\text{Hess } y_\theta(M_0)$ tend vers $\text{Hess } y(M_0)$ lorsque $\theta \rightarrow 0$, nous pouvons affirmer que $\text{Hess } y_\theta(M_0)$ est défini négatif lorsque θ est assez voisin de 0. Pour θ assez petit, M_θ et M_0 sont dans un même voisinage, donc $\text{Hess } y_\theta(M_\theta)$ est lui aussi défini négatif. D'autre part, $y_\theta \in W \subset C^2(\bar{\Omega})$, donc y_θ est localement C^2 au voisinage de M_θ . \square

Grâce au lemme 4.4, $b'(y_\theta)$ est bien défini pour θ proche de 0. Nous nous donnons $y \in W$ vérifiant les hypothèses H1 et H2, et nous considérons le rapport :

$$R = \frac{b'(y_\theta)\phi - b'(y)\phi}{\theta} = \frac{b'(y + \theta\phi)\phi - b'(y)\phi}{\theta}$$

où ϕ et ψ sont deux fonctions quelconques de $W^{3,p_0}(\Omega)$, avec toujours $p_0 > 2$. Le lemme 2.6 du chapitre II nous donnait l'expression de la Gâteaux-dérivée de b :

$$b'(y)\phi = \frac{y(M_0)\phi - y\phi(M_0)}{y(M_0)^2}$$

si M_0 est le point où y atteint son maximum. Donc :

$$R = \frac{(y(M_\theta) + \theta\psi(M_\theta))\phi - (y + \theta\psi)\phi(M_\theta)}{\theta y_\theta(M_\theta)^2} - \frac{y(M_0)\phi - y\phi(M_0)}{\theta y(M_0)^2},$$

soit encore, en réduisant au même dénominateur :

$$\begin{aligned} R = & \left(y(M_\theta)^2 y(M_0)^2 \phi - y\phi(M_\theta) y(M_0)^2 - y(M_0) y(M_\theta)^2 \phi + y\phi(M_0) y(M_\theta)^2 \right. \\ & + \theta(\psi(M_\theta)\phi y(M_0)^2 - \psi\phi(M_\theta) y(M_0)^2) \\ & - 2\psi(M_\theta) y(M_\theta) y(M_0)\phi + 2y\phi(M_0)\psi(M_\theta) y(M_\theta) \\ & \left. + \theta^2(\psi(M_\theta)^2 y\phi(M_0) - \psi(M_\theta)^2 y(M_0)\phi) \right) / (\theta y(M_0)^2 y_\theta(M_\theta)^2) \end{aligned}$$

En regroupant les termes :

$$\begin{aligned} R = & \left. \frac{y(M_\theta)\phi}{y(M_0)y_\theta(M_\theta)^2} \times \frac{y(M_0) - y(M_\theta)}{\theta} \right\} \mathcal{T}_1 \\ + & \left. \frac{y}{y(M_0)^2 y_\theta(M_\theta)^2} \times \frac{\phi(M_0)y(M_\theta)^2 - \phi(M_\theta)y(M_0)^2}{\theta} \right\} \mathcal{T}_2 \\ & + \left. 2y(M_\theta)\psi(M_\theta) \frac{y\phi(M_0) - y(M_0)\phi}{y(M_0)^2 y_\theta(M_\theta)^2} \right\} \mathcal{T}_3 \\ & + \left. \frac{\psi(M_\theta)\phi - \psi\phi(M_\theta)}{y_\theta(M_\theta)^2} + O(\theta) \right\} \mathcal{T}_4 \end{aligned}$$

Nous nous intéressons aux limites des quatre quantités $\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3$ et \mathcal{T}_4 lorsque $\theta \rightarrow 0$. En ce qui concerne \mathcal{T}_1 , dans la preuve du lemme 2.6 du chapitre II, nous avons vu que : $\frac{y(M_\theta) - y(M_0)}{\theta} \rightarrow 0$ lorsque $\theta \rightarrow 0$. Comme y est continue et que $M_\theta \rightarrow M_0$, $y(M_\theta)$ tend vers $y(M_0)$ lorsque θ tend vers 0. Et $y_\theta(M_\theta) = y(M_\theta) + \theta\psi(M_\theta)$, et ψ est bornée, donc $y_\theta(M_\theta)$ tend vers $y(M_0)$. D'où :

$$\| \mathcal{T}_1 \|_{3,p_0} \leq C \| \phi \|_{3,p_0} \times \left| \frac{y(M_\theta) - y(M_0)}{\theta} \right|$$

donc $\lim_{\theta \rightarrow 0} \mathcal{T}_1 = 0$.

Nous développons la quantité \mathcal{T}_2 :

$$\begin{aligned} \phi(M_0)y(M_\theta)^2 - \phi(M_\theta)y(M_0)^2 &= \phi(M_0)(y(M_\theta)^2 - y(M_0)^2) \\ &\quad + y(M_0)^2(\phi(M_0) - \phi(M_\theta)), \end{aligned}$$

et donc pour étudier la limite de \mathcal{T}_2 lorsque θ tend vers 0, nous étudierons respectivement les limites de :

$$\frac{y(M_\theta)^2 - y(M_0)^2}{\theta} \text{ et } \frac{\phi(M_0) - \phi(M_\theta)}{\theta}.$$

Toujours grâce au lemme 2.6 du chapitre II, nous avons :

$$\lim_{\theta \rightarrow 0} \frac{y(M_\theta)^2 - y(M_0)^2}{\theta} = 0.$$

Il nous faut à présent étudier le rapport $\frac{\phi(M_0) - \phi(M_\theta)}{\theta}$ pour θ au voisinage de 0. D'après le lemme 4.3, il existe un voisinage \mathcal{V} de 0, tel que si $\theta \in \mathcal{V}$, alors il existe une fonction unique $g : \theta \rightarrow M_\theta$ qui est dérivable en $\theta = 0$. Or, nous pouvons écrire :

$$\frac{\phi(M_\theta) - \phi(M_0)}{\theta} = \frac{\phi \circ g(\theta) - \phi \circ g(0)}{\theta}.$$

Comme g est dérivable en 0 :

$$\lim_{\theta \rightarrow 0} \frac{\phi(M_\theta) - \phi(M_0)}{\theta} = \nabla \phi^t(g(0))g'(0) = -\nabla \phi^t(M_0) \left(\text{Hess } y(M_0) \right)^{-1} \nabla \psi(M_0)$$

d'après l'expression de $g'(0)$. D'où nous tirons :

$$\lim_{\theta \rightarrow 0} \mathcal{T}_2 = \frac{y}{y(M_0)^2} \nabla \phi^t(M_0) \left(\text{Hess } y(M_0) \right)^{-1} \nabla \psi(M_0).$$

Les limites de \mathcal{T}_3 et \mathcal{T}_4 seront quant à elles :

$$\lim_{\theta \rightarrow 0} (\mathcal{T}_3 + \mathcal{T}_4) = \frac{2\psi(M_0)\phi(M_0)y}{y(M_0)^3} - \frac{\psi(M_0)\phi + \psi\phi(M_0)}{y(M_0)^2}$$

Nous déduisons de tout ceci :

$$(4.6) \quad \lim_{\theta \rightarrow 0} \frac{b'(y + \theta\psi)\phi - b'(y)\phi}{\theta} = B(y)(\phi, \psi) = -\frac{\psi(M_0)\phi + \psi\phi(M_0)}{y(M_0)^2} + \frac{2\psi(M_0)\phi(M_0)y}{y(M_0)^3} + \frac{y}{y(M_0)^2} \nabla\phi^t(M_0) \left(\text{Hess } y(M_0) \right)^{-1} \nabla\psi(M_0).$$

Et donc, cette quantité (4.6) est un bon candidat pour être la Gâteaux-dérivée à l'ordre deux de b . Nous pouvons constater qu'il s'agit d'une application bilinéaire symétrique. Il nous reste à prouver que l'application :

$$\begin{aligned} W^{3,p_0}(\Omega) \times W^{3,p_0}(\Omega) &\rightarrow W^{3,p_0}(\Omega) \\ (\phi, \psi) &\mapsto B(y)(\phi, \psi) \end{aligned}$$

qui est bilinéaire sur $W^{3,p_0}(\Omega) \times W^{3,p_0}(\Omega)$, est également continue sur cet espace. Pour cela, nous allons majorer la norme de $B(y)(\phi, \psi)$ dans $W^{3,p_0}(\Omega)$, pour tout $(\phi, \psi) \in W^{3,p_0}(\Omega) \times W^{3,p_0}(\Omega)$:

$$\begin{aligned} \| B(y)(\phi, \psi) \|_{3,p_0} \leq C &\left(|\psi(M_0)| \times \|\phi\|_{3,p_0} + \|\psi\|_{3,p_0} \times |\phi(M_0)| \right. \\ &+ 2 |\psi(M_0)| \times |\phi(M_0)| \times \|y\|_{3,p_0} \\ &\left. + \|y\|_{3,p_0} \times |\nabla\phi(M_0)| \times |\nabla\psi(M_0)| \right) \end{aligned}$$

où C désigne une constante indépendante de ϕ et ψ . Comme $p_0 > 2$, nous savons que $W^{1,p_0}(\Omega) \subset L^\infty(\Omega)$ (cf [BRE 87]), et comme les fonctions ϕ et ψ sont dans $W^{3,p_0}(\Omega)$, elles sont dans $L^\infty(\Omega)$, d'où :

$$|\phi(M_0)| \leq \|\phi\|_\infty \leq C \|\phi\|_{1,p_0} \leq C \|\phi\|_{3,p_0},$$

et l'inégalité similaire pour ψ . Puisque $\nabla\psi \in W^{2,p_0}(\Omega)$, nous avons également :

$$|\nabla\phi(M_0)| \leq \|\nabla\phi\|_\infty \leq C \|\phi\|_{1,p_0} \leq C \|\phi\|_{3,p_0}$$

et l'identique pour ψ . Et donc :

$$\| B(y)(\phi, \psi) \|_{3,p_0} \leq C' \|\phi\|_{3,p_0} \times \|\psi\|_{3,p_0}.$$

Nous avons donc démontré que l'application b était deux fois Gâteaux-dérivable sur l'espace $\mathcal{L}(W, \mathcal{L}(W, W))$. Tous les calculs précédents se résument donc par la proposition suivante :

Proposition 4.5 *L'application $b : y \mapsto \frac{y}{\max_{\Omega} y}$ est deux fois Gâteaux-dérivable sur l'espace $W \subset W^{3,p_0}(\Omega)$ (avec $p_0 > 2$), cette dérivée est l'application $b''(y)$ qui à $(\phi, \psi) \in W^{3,p_0}(\Omega) \times W^{3,p_0}(\Omega)$ associe l'expression (4.6) .*

Lorsqu'on fait subir une petite variation \tilde{v} à la source, l'état admet une variation \tilde{y} vérifiant (4.2) ; regardons formellement ce qu'il en est pour l'état adjoint et le gradient :

$$v \rightarrow v + \tilde{v}, \quad y \rightarrow y + \tilde{y}, \quad p \rightarrow p + \tilde{p}, \quad \nabla J \rightarrow \nabla J + \tilde{\nabla} J.$$

Rappelons que l'état adjoint p vérifie le système :

$$\begin{cases} -\Delta p - v^t N' ob(y) (b'(y)^t p) = 0 & \text{dans } \Omega, \\ p = \frac{\partial y}{\partial n} - z & \text{sur } \partial\Omega, \end{cases}$$

et donc :

$$\begin{cases} -\Delta p - \Delta \tilde{p} - (v + \tilde{v})^t N' ob(y + \tilde{y}) (b'(y + \tilde{y})^t (p + \tilde{p})) = 0 & \text{dans } \Omega \\ p + \tilde{p} = \frac{\partial y}{\partial n} + \frac{\partial \tilde{y}}{\partial n} - z & \text{sur } \partial\Omega. \end{cases}$$

D'où, puisque b est deux fois dérivable :

$$(4.7) \quad \begin{cases} -\Delta \tilde{p} - v^t N' ob(y) (b'(y)^t \tilde{p}) = v^t N' ob(y) (\tilde{y}^t b''(y) p) \\ \quad + v^t N'' ob(y) (b'(y) \tilde{y}) (b'(y)^t p) + \tilde{v}^t N' ob(y) (b'(y)^t p) & \text{dans } \Omega \\ \tilde{p} = \frac{\partial \tilde{y}}{\partial n} & \text{sur } \partial\Omega. \end{cases}$$

A nouveau, l'existence et l'unicité de \tilde{p} vérifiant (4.7) sont assurées dès que y est Gâteaux-dérivable par rapport au contrôle v .

L'expression du gradient s'écrivait :

$$\nabla J(v) = \varepsilon \Lambda v - \int_{\Omega} N ob(y) p dx,$$

d'où l'on tire :

$$\begin{aligned} \nabla J(v) + \tilde{\nabla} J &= \varepsilon \Lambda v + \varepsilon \Lambda \tilde{v} - \int_{\Omega} N ob(y + \tilde{y}) (p + \tilde{p}) dx \\ &= \varepsilon \Lambda v + \varepsilon \Lambda \tilde{v} - \int_{\Omega} N ob(y) p dx - \int_{\Omega} N ob(y) \tilde{p} dx \\ &\quad - \int_{\Omega} N' ob(y) (b'(y) \tilde{y}) p dx, \end{aligned}$$

et donc :

$$\tilde{\nabla} J = \varepsilon \Lambda \tilde{v} - \int_{\Omega} N ob(y) \tilde{p} dx - \int_{\Omega} N' ob(y) (b'(y) \tilde{y}) p dx.$$

Comme on l'a vu précédemment, nous déduisons de ceci la valeur de la transposée du hessien appliqué à \tilde{v} :

$$(4.8) \quad (\text{Hess}J(v))^t \tilde{v} = \varepsilon \Lambda \tilde{v} - \int_{\Omega} N \text{ob}(y) \tilde{p} dx - \int_{\Omega} N' \text{ob}(y) (b'(y) \tilde{y}) p dx.$$

Il nous reste à transposer cette expression pour en déduire le hessien appliqué à un vecteur η de \mathbb{R}^m :

Proposition 4.6 *Soit $\eta \in \mathbb{R}^m$, alors on a :*

$$(4.9) \quad (\text{Hess}J(v))\eta = \varepsilon \Lambda \eta - \int_{\Omega} N \text{ob}(y) Q dx - \int_{\Omega} N' \text{ob}(y) (b'(y)^t p) Y dx,$$

où $y = y(v)$, p désigne l'état adjoint associé, et où Y et Q sont donnés par :

$$(4.10) \quad \begin{cases} -\Delta Y - v^t N' \text{ob}(y) (b'(y) Y) = \eta^t N \text{ob}(y) & \text{dans } \Omega, \\ Y = 0 & \text{sur } \partial\Omega; \end{cases}$$

$$(4.11) \quad \begin{cases} -\Delta Q - v^t N' \text{ob}(y) (b'(y)^t Q) = v^t N' \text{ob}(y) (Y^t b''(y) p) \\ \quad + v^t N'' \text{ob}(y) (b'(y)^t p) (b'(y) Y) + \eta^t N' \text{ob}(y) (b'(y)^t p) & \text{dans } \Omega, \\ Q = \frac{\partial Y}{\partial n} & \text{sur } \partial\Omega. \end{cases}$$

Preuve : Multiplions (4.11) par \tilde{y} solution de (4.2) et (4.10) par \tilde{p} solution de (4.7) ; en intégrant par parties sur Ω et en les additionnant, nous obtenons :

$$\begin{aligned} & \int_{\Omega} \nabla Y \cdot \nabla \tilde{p} - \int_{\partial\Omega} \frac{\partial Y}{\partial n} \tilde{p} d\sigma - \int_{\Omega} v^t N' \text{ob}(y) (b'(y) Y) \tilde{p} - \int_{\Omega} \eta^t N \text{ob}(y) \tilde{p} \\ & + \int_{\Omega} \nabla Q \cdot \nabla \tilde{y} - \int_{\partial\Omega} \frac{\partial Q}{\partial n} \tilde{y} d\sigma - \int_{\Omega} v^t N' \text{ob}(y) (b'(y)^t Q) \tilde{y} \\ & - \int_{\Omega} v^t N' \text{ob}(y) (Y^t b''(y) p) \tilde{y} - \int_{\Omega} v^t N'' \text{ob}(y) (b'(y)^t p) (b'(y)^t Y) \tilde{y} \\ & - \int_{\Omega} \eta^t N' \text{ob}(y) (b'(y)^t p) \tilde{y} = 0. \end{aligned}$$

Soit encore, en intégrant à nouveau par parties, et en transposant :

$$\begin{aligned}
& \int_{\Omega} -\Delta \tilde{p} Y + \int_{\partial\Omega} \frac{\partial \tilde{p}}{\partial n} Y d\sigma - \int_{\partial\Omega} \frac{\partial \tilde{y}}{\partial n} \frac{\partial Y}{\partial n} d\sigma - \int_{\Omega} v^t N' ob(y) (b'(y)^t \tilde{p}) Y \\
& - \int_{\Omega} \eta^t N ob(y) \tilde{p} + \int_{\Omega} -\Delta \tilde{y} Q + \int_{\partial\Omega} \frac{\partial \tilde{y}}{\partial n} \frac{\partial Y}{\partial n} d\sigma - \int_{\Omega} v^t N' ob(y) (b'(y)^t \tilde{y}) Q \\
& - \int_{\Omega} v^t N' ob(y) (\tilde{y}^t b''(y) p) Y - \int_{\Omega} v^t N'' ob(y) (b'(y)^t p) (b'(y)^t \tilde{y}) Y \\
& - \int_{\Omega} \eta^t N' ob(y) (b'(y)^t \tilde{y}) p = 0.
\end{aligned}$$

En écrivant que \tilde{y} est solution de (4.2) et \tilde{p} de (4.7), il reste :

$$\int_{\Omega} \tilde{v}^t N ob(y) Q + \int_{\Omega} \tilde{v}^t N' ob(y) (b'(y)^t p) Y = \int_{\Omega} \eta^t N ob(y) \tilde{p} + \int_{\Omega} \eta^t N' ob(y) (b'(y)^t \tilde{y}) p$$

ou encore :

$$\begin{aligned}
\langle \tilde{v}, \int_{\Omega} N ob(y) Q + \int_{\Omega} N' ob(y) (b'(y)^t p) Y \rangle \\
& = \langle \eta, \int_{\Omega} N ob(y) \tilde{p} + \int_{\Omega} N' ob(y) (b'(y)^t \tilde{y}) p \rangle \\
& = \langle \eta, -(\text{Hess}J(v))^t \tilde{v} + \varepsilon \Lambda \tilde{v} \rangle
\end{aligned}$$

en utilisant (4.8). En transposant encore une fois, nous obtenons :

$$\langle (\text{Hess}J(v))\eta, \tilde{v} \rangle = \langle \tilde{v}, \varepsilon \Lambda \eta - \int_{\Omega} N ob(y) Q - \int_{\Omega} N' ob(y) (b'(y)^t p) Y \rangle,$$

ceci étant valable pour tout \tilde{v} de \mathbb{R}^m , nous obtenons bien l'expression annoncée pour $(\text{Hess}J(v))\eta$. \square

Remarque 4.7 : Dès lors que y est Gâteaux-dérivable par rapport au contrôle, nous aurons l'existence et l'unicité pour Y et Q , donc le hessien de la fonctionnelle J appliqué à un vecteur η de \mathbb{R}^m sera bien défini.

Remarque 4.8 : Comme précédemment, on note que le couple (Y, Q) peut être interprété en terme de variable adjointe du couple (\tilde{p}, \tilde{y}) , solutions de l'équation linéaire tangente linéarisée et de l'équation linéaire tangente du problème.

Remarque 4.9 : Le hessien du critère est symétrique

$$(\text{Hess}J(v))^t \tilde{v} = (\text{Hess}J(v))\tilde{v}$$

En effet, pour $\eta = \tilde{v}$, nous avons : $Y = \tilde{y}$ et $Q = \tilde{p}$.

4.2 Résultats numériques

Nous nous plaçons ici dans un domaine Ω de \mathbb{R}^2 délimité par une ellipse dont le rapport des longueurs des axes vaut 2. Les observations sont bruitées avec un niveau de bruit de 5%.

4.2.1 Validation croisée ordinaire

Nous représentons pour chaque fonction identifiée: à gauche, la fonction identifiée avec ε_{OCV} (la valeur du paramètre obtenue par validation croisée) et à droite les fonctions identifiées avec un paramètre plus petit que ε_{OCV} et l'autre plus grand.

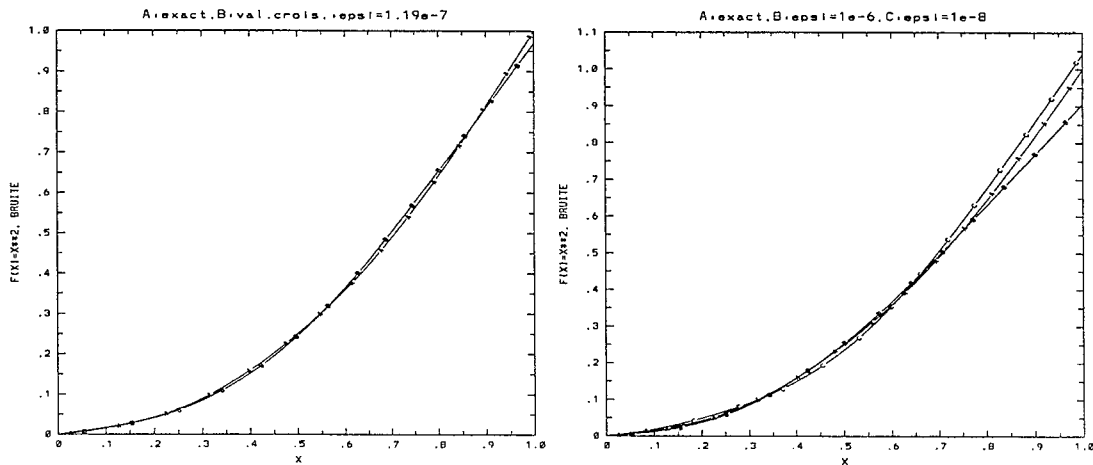


Figure 4.5 Comparaison de ε_{OCV} avec 2 autres valeurs de ε pour x^2

Pour la source x^2 , la valeur de ε_{OCV} est de $1,2 \cdot 10^{-7}$. Afin de valider cette valeur, nous avons comparé le résultat de l'identification OCV avec la source identifiée pour $\varepsilon = 10^{-6}$ et $\varepsilon = 10^{-8}$ (figure 4.5). Le test est positif, puisque les sources identifiées avec ces deux valeurs sensiblement différentes de ε_{OCV} sont plus éloignées de la solution exacte.

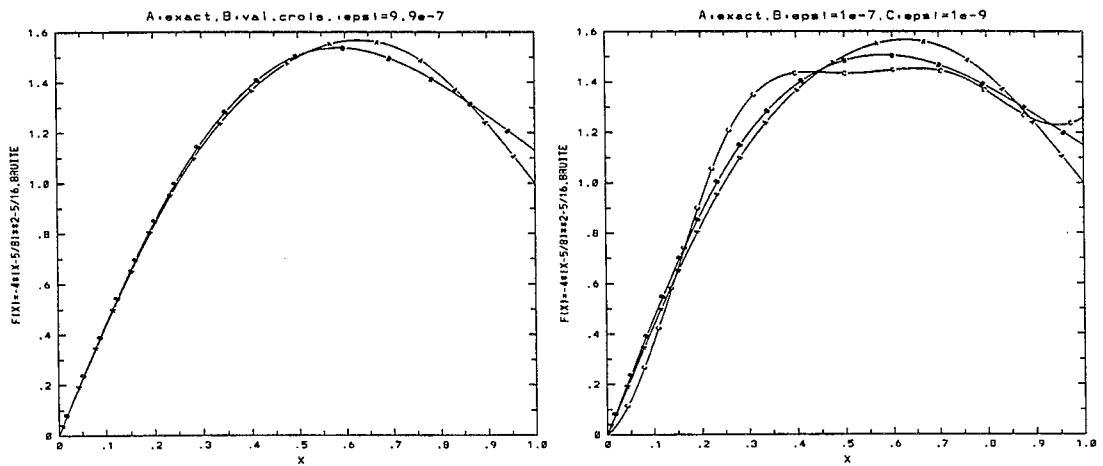


Figure 4.6 Comparaison de ε_{OCV} avec 2 autres valeurs de ε pour $-4x^2 + 5x$

Dans le cas de ce polynôme de degré deux, figure 4.6, $\varepsilon_{OCV} = 10^{-6}$, et nous avons tracé à droite les sources identifiées avec un paramètre de régularisation valant respectivement 10^{-5} et 10^{-7} . Comme pour la courbe précédente, ce test est positif relativement au test de l'optimalité du paramètre ε_{OCV} . Nous constatons d'autre part que 10^{-5} étant une valeur trop élevée du paramètre, l'identification réalisée avec ce coefficient donne une fonction plus "aplatie", tandis que pour la valeur trop faible 10^{-7} , la fonction obtenue n'est plus très régulière.

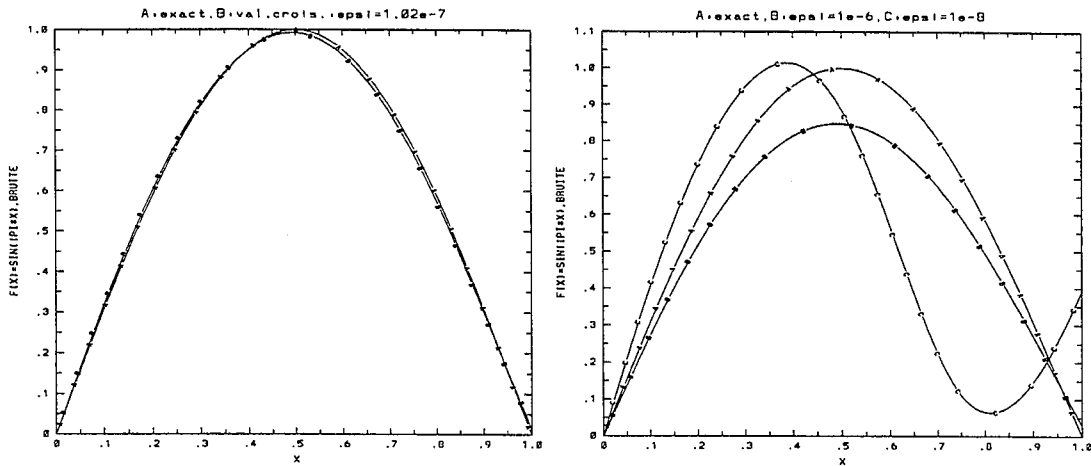


Figure 4.7 Comparaison de ε_{OCV} avec 2 autres valeurs de ε pour $\sin(\pi x)$

Dans le cas de la fonction $\sin(\pi x)$, figure 4.7, $\varepsilon_{OCV} = 10^{-7}$, et nous avons représenté les deux fonctions identifiées avec $\varepsilon = 10^{-6}$ et 10^{-8} . Les mêmes remarques que les précédentes s'appliquent encore : optimalité du paramètre déterminé par OCV, phénomène de "linéarisation" de la solution pour un paramètre trop grand, et apparition d'instabilités pour un ε trop petit.

En conclusion, pour ce problème d'identification de sources non linéaires, la validation croisée ordinaire donne de très bons résultats, et paraît très intéressante du fait qu'elle ne dépend pas du profil de la fonction à identifier. Par contre, elle est assez coûteuse au niveau du temps de calcul nécessaire (environ 40 minutes CPU d'un CRAY C90) ; nous verrons que de ce point de vue, la GCV est beaucoup plus intéressante.

4.2.2 Validation croisée généralisée

Nous représentons pour chaque courbe, la fonction exacte et la fonction identifiée avec ε_{GCV} . Pour estimer la trace de $B'_\varepsilon(z)$, dix réalisations de Monte Carlo ont été utilisées.

Il est à noter, que pour le calcul du hessien du critère, il nous a fallu utiliser une approximation : en effet, dans le calcul de Q , (équation (3.9)) intervient un terme de la forme : $b''(y)$, ce terme faisant intervenir le hessien du vecteur d'état y . Ayant utilisé des éléments finis $P1$ pour le calcul de y , il nous a fallu approcher

le hessien de y par une formule de différences divisées. Cette méthode, on le sait, est très instable numériquement, d'où des erreurs.

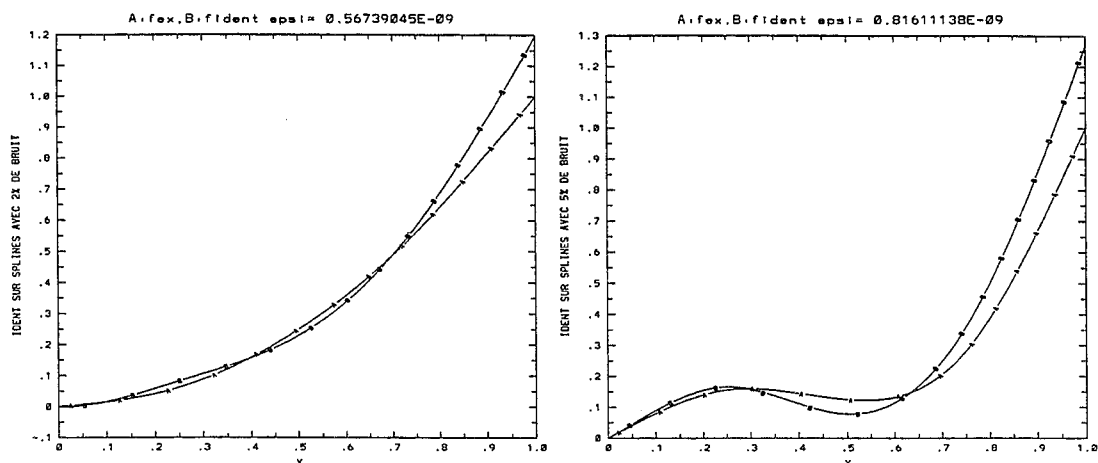


Figure 4.8 Identification avec ε_{GCV} pour x^2 et $x(4x^2 - \frac{9}{2}x + \frac{3}{2})$

Pour la figure 4.8, à gauche pour la fonction x^2 , ε_{GCV} vaut $5,7 \cdot 10^{-10}$ et à droite pour $x(4x^2 - \frac{9}{2}x + \frac{3}{2})$, ε_{GCV} vaut $8,2 \cdot 10^{-10}$. Les fonctions identifiées à l'aide du paramètre GCV sont donc assez satisfaisantes dans ces deux cas de figure.

Dans le cas de la figure 4.9, les valeurs du paramètre obtenues sont $\varepsilon_{GCV} = 5,4 \cdot 10^{-11}$ à gauche, et $\varepsilon_{GCV} = 8,2 \cdot 10^{-10}$ à droite. Il est manifeste que pour la fonction $-4x^2 + 5x$ (à gauche), la valeur de ε_{GCV} est largement sous-estimée, alors que pour $e^x - 1$, le résultat reste acceptable.

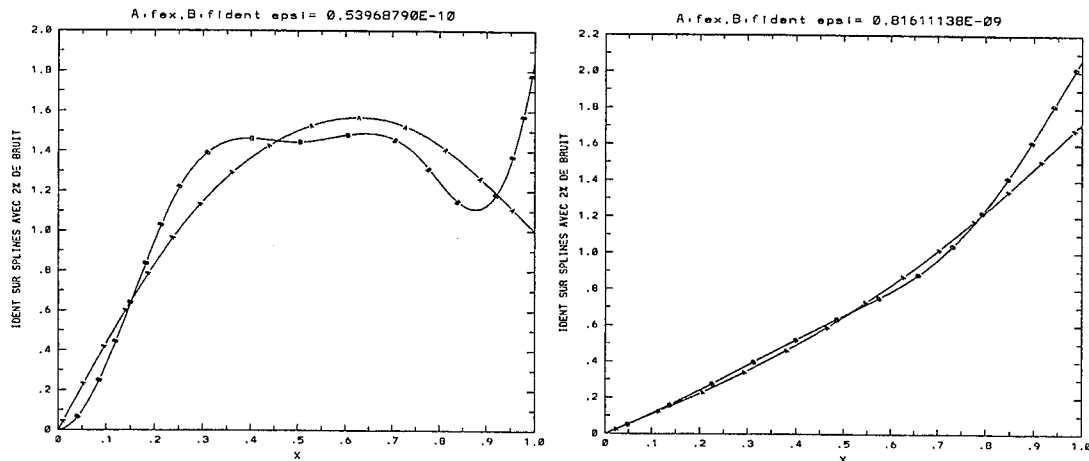


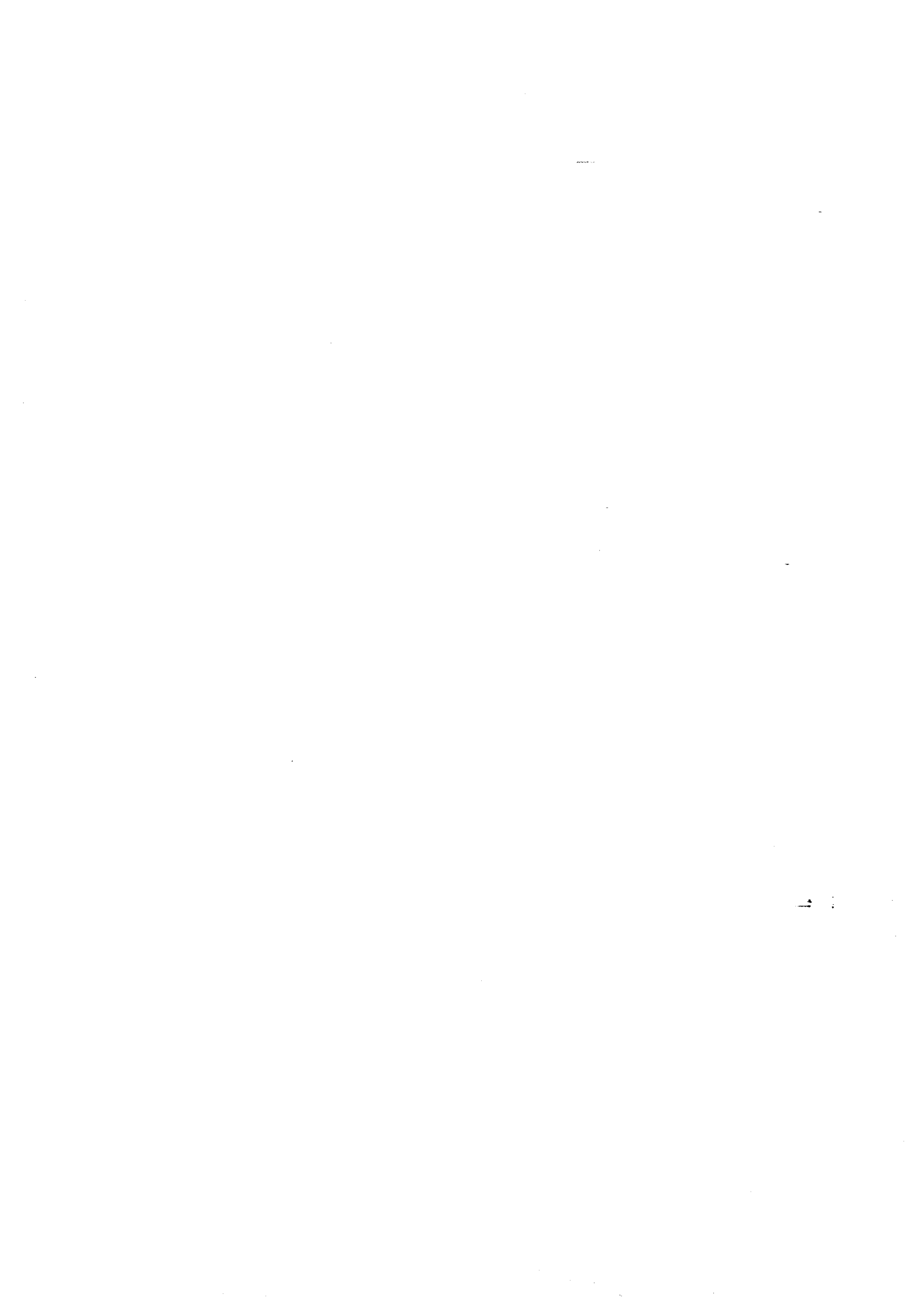
Figure 4.9 Identification avec ε_{GCV} pour $-4x^2 + 5x$ et $e^x - 1$

La GCV semble donc plus sensible à la forme de la fonction à identifier que la validation croisée ordinaire. Son unique avantage réside donc dans le temps de calcul nécessaire : en moyenne 5 minutes CPU contre 40 minutes CPU sur un CRAY C90 pour la validation croisée ordinaire.

Considérant les courbes obtenues précédemment, il semble donc que pour ce problème, la validation croisée ordinaire soit mieux adaptée.

Numériquement, et comme dans le cas de la diffusion de la chaleur, la GCV est nettement moins coûteuse en temps de calcul que la validation croisée ordinaire. Mais contrairement au cas précédent, la validation croisée ordinaire donne de bien meilleurs résultats que la GCV du fait de l'approximation faite sur le hessien de l'état.

Un compromis serait de commencer avec une procédure GCV, pour avoir une estimation de la valeur du paramètre optimal, ce qui nous donnerait $\varepsilon \in [\varepsilon_a, \varepsilon_b]^*$, puis de continuer avec une procédure de validation croisée ordinaire en partant de l'intervalle $[\varepsilon_a, \varepsilon_b]$ pour la méthode de la section dorée.



Chapitre IV

Identification par décomposition sur une base d'ondelettes

Le but de ce chapitre est de réaliser l'identification de paramètres à partir de données bruitées, sans introduire de régularisation dans la fonction coût. Cette étude a été motivée par A. Antoniadis, qui dans un papier assez récent [ANT 93], a exhibé un estimateur de fonctions par des ondelettes particulières, les "coifflets", ayant de bonnes propriétés asymptotiques.

Nous rappelons tout d'abord quelques résultats de base sur les ondelettes, dûs en grande partie à I. Daubechies ([DAU 88] et [DAU 92]) ainsi qu'à Y. Meyer [MEY 92]. Nous travaillerons par la suite avec les ondelettes à support compact de I. Daubechies.

Nous présentons ensuite la méthode de régression par ondelettes, ainsi que des techniques utilisées afin de réduire l'influence du bruit sur l'estimateur obtenu. Nous appliquerons ces techniques aux problèmes inverses de l'identification du coefficient de diffusion de la chaleur et de l'identification de sources non linéaires présentés précédemment.

Les méthodes de seuillage n'étant pas très fiables pour notre problème, nous donnons ensuite une méthode de changement de base sur l'espace des contrôles par rapport à la norme d'un espace de Hilbert. Nous verrons que cette méthode, qui ne demande pas de régularisation, a l'avantage d'accélérer considérablement la convergence numérique et ne nécessite pas le choix d'un seuil.

1 Quelques généralités sur les ondelettes

La notion d'analyse multirésolution de $L^2(\mathbb{R})$ a été introduite par Y. Meyer [MEY 87] et S. Mallat [MA 89]. Il s'agit d'approcher chaque fonction f de $L^2(\mathbb{R})$ par une suite de fonctions f_j , $j \in \mathbb{Z}$ approximant de mieux en mieux la fonction f .

1.1 Analyse multirésolution de $L^2(\mathbb{R})$

Définition 1.1 On appelle analyse multirésolution de $L^2(\mathbb{R})$ une suite d'espaces V_j fermés tels que :

$$(1.1) \quad \dots V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \dots,$$

$$(1.2) \quad \overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R}),$$

$$(1.3) \quad \bigcap_{j \in \mathbb{Z}} V_j = \{0\},$$

$$(1.4) \quad \forall j \in \mathbb{Z}, f \in V_j \iff f(2 \cdot) \in V_{j+1},$$

$$(1.5) \quad f \in V_0 \iff f(\cdot - n) \in V_0, \forall n \in \mathbb{Z}.$$

On supposera de plus qu'il existe une fonction ϕ telle que

$$(1.6) \quad \{\phi(x - n), n \in \mathbb{Z}\} \text{ forme une base de Riesz de } V_0.$$

Alors, si pour tous les entiers $j, n \in \mathbb{Z}$ on définit les fonctions

$$\phi_{j,n}(x) = 2^{\frac{j}{2}} \phi(2^j x - n),$$

les conditions (1.4) et (1.6) nous assurent que $\{\phi_{j,n}, n \in \mathbb{Z}\}$ est une base de V_j pour tout $j \in \mathbb{Z}$ (orthonormée si les $\{\phi(x - n), n \in \mathbb{Z}\}$ sont orthonormées).

Remarque 1.2 : V_j peut s'interpréter comme l'espace de tous les approximants possibles à l'échelle 2^{-j} . La condition $V_j \subset V_{j+1}$ traduit le fait que toute information accessible à l'échelle 2^{-j} (c'est à dire la projection sur V_j) l'est aussi à l'échelle $2^{-(j+1)}$: l'information contenue dans le signal est dégradée lorsque j décroît. Les conditions (1.3) et (1.4) indiquent que si $f_j = P_j f$ est la projection de f sur V_j , alors $\lim_{j \rightarrow +\infty} P_j f = f$ et $\lim_{j \rightarrow -\infty} P_j f = 0$ dans $L^2(\mathbb{R})$, donc le signal approché contient de moins en moins d'informations.

Définition 1.3 Une analyse multirésolution V_j , $j \in \mathbb{Z}$ de $L^2(\mathbb{R})$ est dite r -régulière ($r \in \mathbb{N}$) si la fonction d'échelle $\phi \in V_0$ peut être choisie telle que les $\phi^{(k)}$ soient à décroissance rapide pour $k = 1, \dots, r$.

1.2 Bases orthonormées d'ondelettes de $L^2(\mathbb{R})$

Pour tout $j \in \mathbb{Z}$, on considère le supplémentaire orthogonal de V_j dans V_{j+1} : $V_{j+1} = V_j \oplus W_j$. D'après les propriétés des espaces V_j , on vérifie facilement que :

$$(1.7) \quad f \in W_j \iff f(2 \cdot) \in W_{j+1}, \quad \forall j \in \mathbb{Z},$$

$$(1.8) \quad f \in W_0 \iff f(\cdot - k) \in W_0, \quad \forall k \in \mathbb{Z},$$

$$(1.9) \quad W_j \perp W_{j'}, \quad \text{si } j \neq j',$$

$$(1.10) \quad L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j.$$

Dans ce qui suit, nous supposons que $\{\phi(x-n), n \in \mathbb{Z}\}$ est une base orthonormée de V_0 . Alors, $\{\phi_{j,n}, n \in \mathbb{Z}\}$ est une base orthonormée de V_j pour tout j dans \mathbb{Z} . Il s'agit de trouver une fonction ψ telle que si l'on pose

$$\psi_{j,n}(x) = 2^{\frac{j}{2}} \psi(2^j x - n),$$

les $\{\psi_{j,n}, j, n \in \mathbb{Z}\}$ forment une base orthonormée de $L^2(\mathbb{R})$. D'après la propriété (1.10), si $\{\psi(x-k), k \in \mathbb{Z}\}$ est une base orthonormée de W_0 alors $\{\psi_{j,n}, n \in \mathbb{Z}\}$

est une base orthonormée de W_j pour tout $j \in \mathbb{Z}$. Il suffit donc de trouver ψ telle que $\{\psi_{0,n}, n \in \mathbb{Z}\}$ soit une base orthonormée de W_0 .

Puisque $\phi \in V_0 \subset V_1$, on peut écrire :

$$\phi(x) = \sum_{n \in \mathbb{Z}} h_n \phi_{1,n}(x) = \sqrt{2} \sum_{n \in \mathbb{Z}} h_n \phi(2x - n)$$

où les coefficients h_n vérifient :

$$h_n = \langle \phi, \phi_{1,n} \rangle \text{ et } \sum_{n \in \mathbb{Z}} |h_n|^2 = 1.$$

Par passage à la transformée de Fourier :

$$\widehat{\phi}(\xi) = \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}} h_n e^{-in\xi/2} \widehat{\phi}\left(\frac{\xi}{2}\right).$$

En posant

$$(1.11) \quad m_0(\xi) = \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}} h_n e^{-in\xi},$$

où les séries sont convergentes en norme L^2 , nous en déduisons que :

$$\widehat{\phi}(\xi) = m_0\left(\frac{\xi}{2}\right) \widehat{\phi}\left(\frac{\xi}{2}\right).$$

m_0 est une fonction 2π -périodique de $L^2([0, 2\pi])$. Pour calculer l'ondelette ψ , I. Daubechies [DAU 92] a établi le résultat suivant :

Théorème 1.4 Si une suite $(V_j)_{j \in \mathbb{Z}}$ de sous-espaces fermés de $L^2(\mathbb{R})$ satisfait les conditions (1.1) à (1.6) alors il existe une base orthonormée d'ondelettes $\{\psi_{j,k}, j, k \in \mathbb{Z}\}$ de $L^2(\mathbb{R})$ telle que

$$P_{j+1} = P_j + \sum_{k \in \mathbb{Z}} \langle \cdot, \psi_{j,k} \rangle \psi_{j,k}.$$

Une possibilité pour l'ondelette ψ est :

$$\widehat{\psi}(\xi) = e^{i\xi/2} \overline{m_0\left(\frac{\xi}{2} + \pi\right)} \widehat{\phi}\left(\frac{\xi}{2}\right)$$

où m_0 est la fonction définie par (1.11) , ou de façon équivalente :

$$\psi(x) = \sqrt{2} \sum_{n \in \mathbb{Z}} (-1)^n h_{1-n} \phi(2x - n)$$

avec convergence au sens L^2 pour cette dernière série.

Disposant ainsi d'une base orthonormée d'ondelettes de $L^2(\mathbb{R})$, tout signal se décompose en :

$$f = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \langle f, \psi_{j,k} \rangle \psi_{j,k}$$

ce que l'on appelle *transformée en ondelettes discrète* de f . Les $\psi_{j,k}$ permettent ainsi de caractériser à chaque échelle 2^{-j} l'information sur les détails nécessaires pour passer d'un niveau donné (V_j) à l'espace de résolution deux fois plus fine ($V_{j+1} = V_j \oplus W_j$).

1.3 Ondelettes à support compact

Cette nouvelle classe d'ondelettes a été introduite par I. Daubechies (voir par exemple [DAU 88]). Dans le procédé de construction d'une base orthonormée d'ondelettes de $L^2(\mathbb{R})$ à partir d'une analyse multirésolution, l'élément essentiel est la fonction d'échelle ϕ de V_0 telle que $\{\phi(x - k), k \in \mathbb{Z}\}$ forme une base orthonormée de V_0 . A partir de cette fonction ϕ on peut ensuite construire l'ondelette mère associée ψ .

L'idée de base est de construire le filtre H tel que sa réponse impulsionnelle h soit de longueur finie et vérifie les conditions suivantes :

- i) $\sum_{k \in \mathbb{Z}} h_k = \sqrt{2}$,
- ii) $\sum_{k \in \mathbb{Z}} h_k h_{k+2n} = \delta_{0,n}, \forall n \in \mathbb{Z}$,
- iii) $\sum_{k \in \mathbb{Z}} k^n g_k = 0, \forall n = 0, \dots, N-1$ avec $g_k = (-1)^k h_{1-k}$.

A partir de l'équation d'échelle :

$$\phi(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} h_k \phi(2x - k), \quad \forall x \in \mathbb{R},$$

on cherche une fonction ϕ normalisée par $\int_{\mathbb{R}} \phi(x) dx = 1$. Ensuite, on définit l'ondelette ψ par :

$$\psi(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} g_k \phi(2x - k), \quad \forall x \in \mathbb{R}.$$

Remarque 1.5 : La condition ii) sur les h_k n'est rien d'autre que la condition d'orthogonalité des ondelettes, elle se traduit sur la fonction m_0 par :

$$(1.12) \quad |m_0(\xi)|^2 + |m_0(\xi + \pi)|^2 = 1.$$

La condition iii) indique que l'ondelette ψ possède N moments nuls :

$$\int_{\mathbb{R}} x^n \psi(x) dx = 0 \text{ pour } n = 0, \dots, N - 1.$$

Quant à i), il s'agit de la condition de normalisation.

I. Daubechies ([DAU 88]) a donné la forme de la fonction m_0 :

Lemme 1.6 Tout polynôme trigonométrique m_0 solution de (1.12) est de la forme :

$$m_0(\xi) = \left(\frac{1 + e^{i\xi}}{2} \right)^N Q(e^{i\xi})$$

où N est le nombre de moments nuls et où Q est un polynôme à coefficients réels.

Si la suite (h_k) est de longueur finie, la fonction ϕ est à support compact, et par conséquent ψ l'est aussi. Pour $N = 1$, $m_0(\xi) = \frac{1+e^{i\xi}}{2}$ et l'ondelette associée est l'ondelette de Haar. Dans ce cas, $h_0 = h_1 = \frac{1}{\sqrt{2}}$, ϕ est la fonction caractéristique de l'intervalle $[0, 1]$ et $\psi(x) = \chi_{[0, \frac{1}{2}]}(x) - \chi_{[\frac{1}{2}, 1]}(x)$. Pour $N \geq 2$, I. Daubechies a exhibé une famille spéciale de fonctions m_0 :

$$m_{0,N}(\xi) = \frac{1}{\sqrt{2}} \sum_{k=0}^{2N-1} h_{N,k} e^{ik\xi}.$$

On note ϕ_N et ψ_N les fonctions correspondantes :

$$\phi_N(x) = \sqrt{2} \sum_{k=0}^{2N-1} h_{N,k} \phi_N(2x - k),$$

$$\psi_N(x) = \sqrt{2} \sum_{k=2-2N}^1 g_{N,k} \phi_N(2x - k).$$

On remarque facilement que $Supp(\phi_N) = [0, 2N - 1]$ et $Supp(\psi_N) = [-N + 1, N]$. Les coefficients $h_{N,k}$ sont nuls pour $k < 0$ ou $k > 2N - 1$. Une table des coefficients $h_{N,k}$ est donnée dans [DAU 88] pour $N = 2, 3, \dots, 10$. Nous la reproduisons en annexe 1, et nous donnons une illustration de ϕ_7 et ψ_7 .

Lorsque N augmente, ϕ_N et ψ_N deviennent plus régulières et leur support grandit en conséquence. I. Daubechies a montré que pour $N \geq 2$, ϕ_N et ψ_N appartiennent à l'espace de Hölder $C^{(\mu-\epsilon)N}$ avec $\mu \simeq 0.3485$.

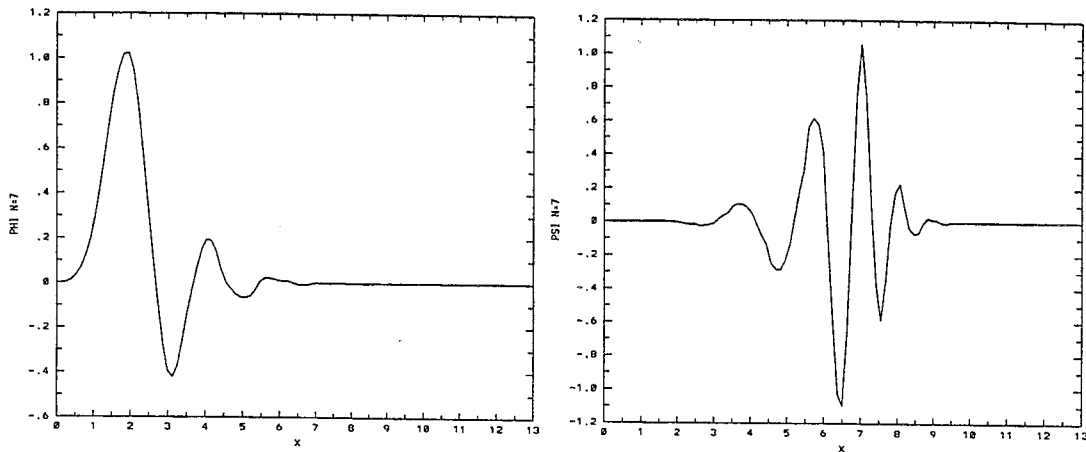


Figure 1.1 A gauche, fonction d'échelle ϕ_7 , à droite ondelette ψ_7 .

1.4 Quelques résultats d'approximation

Tout d'abord, nous donnons une propriété de l'ondelette mère ψ , due à I. Daubechies [DAU 88] :

Propriété 1.7 *Si l'analyse multirésolution est r -régulière, l'ondelette ψ est de classe C^r et ses moments sont nuls jusqu'à l'ordre r : $\int_{\mathbb{R}} x^k \psi(x) dx = 0$, $0 \leq k \leq r$.*

En général, le nombre de moments nuls pour ψ est plus grand que la régularité r de l'analyse multirésolution. Une conséquence importante est que les polynômes de degré inférieur ou égal à r s'expriment comme des combinaisons linéaires des translations de ϕ , i.e. appartiennent à V_0 . D'après la définition de ψ , et comme seulement un nombre fini de coefficients g_k sont non nuls, ψ se réduit à une combinaison linéaire finie de fonctions régulières à support compact, et donc ψ est à support compact, de même régularité que ϕ .

La taille du support de ψ augmente linéairement par rapport au nombre de moments nuls N . Un avantage d'avoir un grand nombre de moments nuls pour ψ est que les coefficients d'ondelettes d'une fonction f , pris à une échelle assez fine, seront nuls si f est régulière. Les résultats suivants sont dûs à Y. Meyer [MEY 92] et S. Mallat [MA 89]. Tout d'abord, si P_j désigne la projection sur l'espace d'approximation V_j , l'erreur commise est donnée par :

Théorème 1.8 Si $f \in H^s(\mathbb{R})$ et si l'analyse multirésolution est r -régulière, alors $\|f - P_j f\| = o(2^{-j \min(s,r)})$ en norme L^2 lorsque $j \rightarrow +\infty$.

Le résultat liant la décroissance des coefficients d'ondelettes à la régularité d'une fonction f est le suivant :

Théorème 1.9 Si f appartient à l'espace de Hölder $C^s(\mathbb{R})$ alors

$$|\langle f, \psi_{j,k} \rangle| \leq C 2^{-j(\min(s,r) + \frac{1}{2})}, \quad \forall j \geq 0, \quad \forall k \in \mathbb{Z}.$$

1.5 L'algorithme de S. Mallat

Il s'agit d'un algorithme de décomposition-reconstruction. Etant donnée une analyse multirésolution de $L^2(\mathbb{R}) : \dots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset \dots$, si f est une fonction de V_j alors on peut l'écrire :

$$(1.13) \quad f = \sum_{k \in \mathbb{Z}} a_k^j \phi_{j,k} \text{ avec } a_k^j = \langle f, \phi_{j,k} \rangle, \quad \forall k \in \mathbb{Z}$$

et comme $V_j = V_{j-1} \oplus W_{j-1}$, nous avons également :

$$(1.14) \quad f = \sum_{\ell \in \mathbb{Z}} a_\ell^{j-1} \phi_{j-1,\ell} + \sum_{\ell \in \mathbb{Z}} c_\ell^{j-1} \psi_{j-1,\ell}$$

avec $a_\ell^{j-1} = \langle f, \phi_{j-1,\ell} \rangle$ et $c_\ell^{j-1} = \langle f, \psi_{j-1,\ell} \rangle$. En effectuant le produit scalaire de (1.13) par $\phi_{j-1,\ell}$:

$$\langle f, \phi_{j-1,\ell} \rangle = a_\ell^{j-1} = \sum_{k \in \mathbb{Z}} a_k^j \langle \phi_{j,k}, \phi_{j-1,\ell} \rangle.$$

Il est facile de voir que $\langle \phi_{j,k}, \phi_{j-1,\ell} \rangle = h_{k-2\ell}$ d'après la définition de h , et donc :

$$a_\ell^{j-1} = \sum_{k \in \mathbb{Z}} h_{k-2\ell} a_k^j.$$

De la même façon, on a :

$$\langle f, \psi_{j-1,\ell} \rangle = c_\ell^{j-1} = \sum_{k \in \mathbb{Z}} a_k^j \langle \phi_{j,k}, \psi_{j-1,\ell} \rangle$$

et $\langle \phi_{j,k}, \psi_{j-1,\ell} \rangle = g_{k-2\ell}$ avec $g_k = (-1)^k h_{1-k}$, donc on obtient

$$c_\ell^{j-1} = \sum_{k \in \mathbb{Z}} a_k^j g_{k-2\ell}.$$

Ainsi, connaissant les coefficients a_k^j , $k \in \mathbb{Z}$ à l'échelle 2^{-j} , on peut en déduire les coefficients à l'échelle 2^{-j+1} par

$$a_\ell^{j-1} = \sum_{k \in \mathbb{Z}} h_{k-2\ell} a_k^j \quad \text{ou } a^{j-1} = H a^j,$$

$$c_\ell^{j-1} = \sum_{k \in \mathbb{Z}} g_{k-2\ell} a_k^j \quad \text{ou } c^{j-1} = G a^j.$$

Ceci constitue l'algorithme de décomposition de S. Mallat, et on le notera par la suite

$$\begin{pmatrix} a^{j-1} \\ c^{j-1} \end{pmatrix} = W a^j.$$

De la même manière, étant donnés les coefficients a^{j-1} , c^{j-1} à une échelle 2^{-j+1} , il est possible de retrouver les coefficients à l'échelle 2^{-j} . En effet, en prenant le produit scalaire de (1.14) contre $\phi_{j,k}$, on obtient :

$$\langle f, \phi_{j,k} \rangle = a_k^j = \sum_{\ell \in \mathbb{Z}} a_\ell^{j-1} \langle \phi_{j,k}, \phi_{j-1,\ell} \rangle + \sum_{\ell \in \mathbb{Z}} c_\ell^{j-1} \langle \phi_{j,k}, \psi_{j-1,\ell} \rangle$$

et puisque $\langle \phi_{j,k}, \phi_{j-1,\ell} \rangle = h_{k-2\ell}$ et $\langle \phi_{j,k}, \psi_{j-1,\ell} \rangle = g_{k-2\ell}$:

$$a_k^j = \sum_{\ell \in \mathbb{Z}} h_{k-2\ell} a_\ell^{j-1} + \sum_{\ell \in \mathbb{Z}} g_{k-2\ell} c_\ell^{j-1}$$

soit encore : $a^j = H^* a^{j-1} + G^* c^{j-1}$, ou

$$a^j = W^* \begin{pmatrix} a^{j-1} \\ c^{j-1} \end{pmatrix},$$

l'opérateur W étant orthogonal.

Remarque 1.10 : les filtres H et G ne dépendent pas de l'échelle considérée, ainsi :

$$a^n = H^* a^{n-1} + G^* c^{n-1} = \dots = (H^*)^n a^0 + (G^*)^n c^0$$

et

$$a^0 = H a^1 = \dots = H^n a^n, \quad c^0 = G a^1 = \dots = G H^{n-1} a^n.$$

Remarque 1.11 : la transition $a^1 \rightarrow a^0, \bar{c}^0$ correspond en fait à un changement de base dans $V_1 : \{\phi_{1,n}, n \in \mathbb{Z}\} \rightarrow \{\phi_{0,n}, \psi_{0,n}, n \in \mathbb{Z}\}$. En raison des propriétés de ces fonctions, la matrice du changement de base est orthogonale.

1.6 Ondelettes sur l'intervalle

Cette revue sur les ondelettes n'aurait pas été complète sans avoir parlé des ondelettes sur l'intervalle, introduites tout d'abord par Y. Meyer puis par I. Daubechies, A. Cohen et P. Vial [DCV 93].

Lorsqu'on désire travailler sur un intervalle de \mathbb{R} ($[0, 1]$ pour fixer les idées), les ondelettes à support compact de I. Daubechies décrites précédemment sont d'emploi un peu plus délicat. Tout d'abord, si l'on considère les restrictions des $\phi_{j,n}$ ou $\psi_{j,n}$ à $[0, 1]$, ces fonctions ne sont plus orthonormales sur $[0, 1]$, sauf dans le cas bien précis de la base de Haar. Ensuite, si l'on ne prend pas un minimum de précautions, on peut voir apparaître des singularités aux extrémités de l'intervalle considéré. Divers choix sont possibles : périodisation, réflexion, ..., nous en énumérons quelques-uns, afin de les comparer et de mettre en évidence les avantages de la nouvelle approche de I. Daubechies, A. Cohen et P. Vial.

Si l'on décidait d'étendre une fonction f en dehors de $[0, 1]$ en la prolongeant par 0, il serait toujours possible de l'analyser en termes d'ondelettes, mais on introduirait alors une discontinuité en $x = 0$ ou en $x = 1$, ce qui se traduirait par des coefficients d'ondelettes $\langle f, \psi_{j,k} \rangle$ très grands aux échelles les plus fines près des deux extrémités, ceci même si f était très régulière sur $[0, 1]$; on ne pourrait donc plus caractériser la régularité de f en 0 ou 1 par la décroissance de ses coefficients d'ondelettes $\langle f, \psi_{j,k} \rangle$ lorsque $j \rightarrow +\infty$.

Une seconde méthode peut consister à "périodiser" les coefficients d'ondelette. Pour cela il suffit de considérer les nouvelles fonctions :

$$\psi_{j,k}^{per}(x) = 2^{\frac{j}{2}} \sum_{\ell \in \mathbb{Z}} \psi(2^j x + 2^j \ell - k) \text{ avec } j > j_0 > 0 \text{ et } 0 \leq k \leq 2^j - 1$$

et une formule analogue pour les fonctions d'échelle $\phi_{j,k}^{per}$. De la sorte, on obtient une base orthonormée de $L^2([0, 1])$ associée à l'analyse multirésolution V_j^{per} où V_j^{per} est engendré par les $\{\phi_{j,k}^{per}, k \in \mathbb{Z}\}$. On remarque facilement que développer une fonction f sur cette base d'ondelettes périodisées revient à écrire f comme une fonction de période 1. Donc, sauf si f était déjà périodique, cette construction introduit à nouveau des discontinuités en $x = 0$ et $x = 1$, et nous amène à lui faire les mêmes critiques que la première.

Une autre solution, souvent adoptée en analyse d'images, consiste à rendre la fonction f symétrique par rapport à $x = 0$ et $x = 1$:

$$\begin{aligned} f(x) &= f(2-x) \text{ si } 1 \leq x \leq 2, \\ f(x) &= f(-x) \text{ si } -1 \leq x \leq 0. \end{aligned}$$

L'opération consiste à répéter ce schéma sur \mathbb{R} tout entier. Si la fonction de départ était continue en 0 et en 1, cette extension reste continue aux valeurs entières, cependant la dérivée n'est plus continue aux valeurs entières. Donc à nouveau, cette construction présente de nombreux inconvénients.

Y. Meyer a proposé une autre construction : partant des ondelettes à support compact de I. Daubechies avec N moments nuls, il a construit, aux échelles les plus fines, les espaces d'approximation $V_j^{[0,1]}$ qui consistent en $2^j - 2N - 2$ fonctions "intérieures" (c'est à dire les $\phi_{j,k}$ habituelles ayant leur support contenu dans $[0,1]$) auxquelles on ajoute $2N - 2$ fonctions à gauche et $2N - 2$ fonctions à droite. Les fonctions du bord doivent être construites explicitement. Les espaces $W_j^{[0,1]}$ sont quant à eux générés par $2^j - 2N - 2$ ondelettes intérieures (les $\psi_{j,k}$ qui ont leur support inclus dans $[0,1]$) et $N - 1$ ondelettes sur chaque bord. Cette construction fournit une base orthonormée d'ondelettes sur $[0,1]$ avec N moments nuls et avec la même régularité que les fonctions originales. Un des problèmes de cette construction est que le processus d'orthonormalisation de la restriction des fonctions d'échelle et des ondelettes est mal conditionné dès que $N > 3$. (On pourra se référer à [DCV 93] pour plus de détails).

La construction de I. Daubechies, A. Cohen et P. Vial ([DCV 93]), la plus récente en la matière à notre connaissance, consiste à garder, comme dans la construction de Y. Meyer, pour fonctions intérieures celles dont le support est inclus dans $[0,1]$, soient $2^j - 2N$ fonctions. Les N fonctions de chaque bord sont obtenues en appliquant une transformation orthogonale sur les fonctions intérieures. L'avantage de cette méthode par rapport à celle de Y. Meyer est qu'à chaque échelle 2^{-j} , la dimension de V_j et W_j est encore 2^j , et que le procédé de calcul est beaucoup plus stable en ce sens que la matrice d'orthonormalisation des fonctions de base est nettement mieux conditionnée que celle de Y. Meyer.

2 Régression par ondelettes

Dans ce paragraphe, nous nous intéressons au problème d'identification d'une fonction f au moyen de sa décomposition dans une base d'ondelettes. Il s'agit donc de trouver des coefficients \hat{a}^0 tels que $\hat{f} = \sum_k \hat{a}_k^0 \phi_{0,k}$ soit un estimateur de la fonction f .

2.1 Régression linéaire par ondelettes

Nous nous plaçons tout d'abord dans le cadre d'un problème inverse linéaire : supposons que l'on doive estimer une fonction f à partir de données bruitées z :

$$z_i = f(t_i) + \nu_i$$

où ν est un vecteur aléatoire de \mathbb{R}^n suivant une loi $\mathcal{N}(0, \sigma^2 I)$. On suppose que les points t_i sont équidistants sur une grille diadique : $t_i = 2^{-n}i$. Un estimateur de régression par ondelettes fonctionne de la manière suivante :

i) à partir des observations $(z_1, z_2, \dots, z_n) = a^0$, on calcule

$$\begin{pmatrix} a^n \\ c^n \end{pmatrix} = W^n a^0 = W^n z,$$

ii) modification des coefficients : $a^n, c^n \rightarrow \tilde{a}^n, \tilde{c}^n$

iii) reconstruction de l'estimateur de f par

$$\hat{f} = (\hat{f}_1, \dots, \hat{f}_n) = (W^n)^* \begin{pmatrix} \tilde{a}^n \\ \tilde{c}^n \end{pmatrix}.$$

Le point crucial de la méthode de régression est le point ii). Une méthode bien connue consiste à faire du seuillage. D.L. Donoho et I. Johnstone [DJ 92] ont établi que les méthodes de seuillage sont appropriées pour des modèles de signal bruité, car d'une part, la transformée en ondelettes discrète du signal est concentrée en seulement un petit nombre de coefficients, et d'autre part la transformée en ondelettes discrète d'un bruit blanc est encore un bruit blanc, et donc se propage sur tous les coefficients (car W est orthogonal).

Dans la pratique, il s'agit donc de déterminer une estimation robuste du niveau de bruit, et ensuite de faire du seuillage sur les coefficients d'ondelettes en fonction de cette estimation.

2.1.1 Règles de seuillage

Il existe essentiellement deux règles de seuillage. La première, est une stratégie de “keep-or-kill” : soit les coefficients sont conservés, soit ils sont supprimés. Ce qui, pour un coefficient d’ondelettes w , est résumé par la valeur dure du seuillage :

$$T_{hard}(w, s) = w\chi_{|w|>s}.$$

Si $|w|$ est plus petit que la valeur du seuil s , on le conserve, sinon on l’annule.

La seconde technique est un peu plus subtile, il s’agit d’une stratégie de “schrink-or-kill” ou encore seuillage mou. Elle consiste à transformer un coefficient w en :

$$T_{soft}(w, s) = \text{sign}(w)(|w| - s)\chi_{|w|>s}.$$

Si $|w|$ est assez grand, w prend la valeur $\text{sign}(w)(|w| - s)$, sinon on l’annule.

Il est à noter que les techniques de seuillage ne seront appliquées que dans les espaces des détails, c’est à dire uniquement sur les coefficients c^j et non sur les a^j . Dans la pratique, se pose alors le problème du choix de l’espace W_j à partir duquel nous commencerons le seuillage.

2.1.2 Choix du seuil

D.L. Donoho et I. Johnstone ont défini une valeur du seuil, appelée *seuil universel*, qui combine une estimation du niveau de bruit sur les données $\hat{\sigma}$ avec le terme $\sqrt{2\log(n)}$: le seuil universel est

$$s_U = \sqrt{2\log(n)}\hat{\sigma}.$$

Ce terme $\sqrt{2\log(n)}$ provient du fait que la probabilité que le maximum des modules d’une suite de bruits ν_i de loi $\mathcal{N}(0, 1)$ soit supérieur à $\sqrt{2\log(n)}$ tend vers 0 lorsque $n \rightarrow +\infty$.

$$\text{Prob}\left\{\max_i |\nu_i| > \sqrt{2\log(n)}\right\} \xrightarrow{n \rightarrow +\infty} 0.$$

Bien qu’il soit asymptotiquement optimal, le seuil universel a l’inconvénient de sous-estimer les données, en ce sens qu’il supprime ou diminue trop de coefficients.

2.2 Application aux problèmes inverses non linéaires

Dans la pratique, nous allons identifier une fonction $f \in \mathcal{C}^2(\mathbb{R})$ au moyen de ses coefficients dans un espace d'approximation V_j , c'est à dire les

$$a_k^j = \langle f, \phi_{j,k} \rangle,$$

avec $\phi = \phi_7$, fonction d'échelle associée à l'ondelette à support compact de I. Daubechies. Nous choisirons donc N au moins égal à 7 pour avoir suffisamment de régularité pour la base d'approximation, et les méthodes employées utilisant des différences finies ou des éléments finis, nous avons besoin de calculer les valeurs de ϕ_N et ψ_N en certains points. Ces calculs peuvent être faits de façon exacte.

2.2.1 Calcul de ϕ_N et ψ_N sur une grille diadique

Etant donnée une subdivision diadique régulière de \mathbb{R} , nous pouvons calculer de manière exacte les valeurs de ϕ_N et ψ_N aux entiers de \mathbb{Z} . Pour cela il suffit d'écrire l'équation d'échelle :

$$(2.1) \quad \phi_N(x) = \sqrt{2} \sum_{k=0}^{2N-1} h_{N,k} \phi_N(2x - k), \quad \forall x \in \mathbb{R}$$

en gardant à l'esprit que $\phi_N(n) = 0$ si $n < 0$ ou $n > 2N - 1$, et l'analogie pour ψ_N :

$$(2.2) \quad \psi_N(x) = \sqrt{2} \sum_{k=2-2N}^1 g_{N,k} \phi_N(2x - k), \quad \forall x \in \mathbb{R}.$$

On ajoute à ceci la condition de normalisation :

$$(2.3) \quad \sum_{n=0}^{2N-1} \phi_N(n) = 1.$$

Les équations (2.1) et (2.3) permettent de calculer les valeurs $\phi_N(n)$ pour tout $n = 0, \dots, 2N - 1$. Les valeurs sur une grille diadique seront données par : supposant calculées les valeurs $\phi_N(\frac{n}{2^k})$ pour $n = 0, \dots, 2^k(2N - 1)$, on calcule les valeurs $\phi_N(\frac{m}{2^{k+1}})$ pour $m = 0, \dots, 2^{k+1}(2N - 1)$ par :

$$\phi_N\left(\frac{m}{2^{k+1}}\right) = \sqrt{2} \sum_{\ell=0}^{2N-1} h_{N,\ell} \phi_N\left(\frac{m}{2^k} - \ell\right) = \sqrt{2} \sum_{\ell=0}^{2N-1} h_{N,\ell} \phi_N\left(\frac{m - 2^k \ell}{2^k}\right).$$

Connaissant toutes les valeurs de ϕ_N sur une grille $t_i = \frac{i}{2^m}$, les valeurs de ψ_N sur cette même grille sont obtenues par l'équation (2.2) prise en des points de la grille diadique :

$$\psi_N\left(\frac{n}{2^m}\right) = \sqrt{2} \sum_{k=2-2N}^1 g_{N,k} \phi_N\left(\frac{n - 2^{p-1}k}{2^{p-1}}\right), \text{ pour } n = 2^m(-N + 1), \dots, 2^m N.$$

Les dérivées ϕ'_N et ψ'_N sont toutes deux à support compact inclus dans $[0, 2N - 1]$ et $[-N + 1, N]$ respectivement. En dérivant (2.1)

$$(2.4) \quad \phi'_N(x) = 2\sqrt{2} \sum_{k=0}^{2N-1} h_{N,k} \phi'_N(2x - k), \quad \forall x \in \mathbb{R},$$

et en dérivant (2.2) :

$$(2.5) \quad \psi'_N(x) = 2\sqrt{2} \sum_{k=2-2N}^1 g_{N,k} \phi'_N(2x - k), \quad \forall x \in \mathbb{R}.$$

Nous avons besoin d'une condition supplémentaire, la dérivation de (2.3) ne nous apportant qu'une condition triviale. Il suffit de remarquer que V_0 contient tous les polynômes de degré $N - 1$, et en particulier :

$$\sum_{n \in \mathbb{Z}} \phi_N(x - n) = x + C$$

d'où par dérivation,

$$(2.6) \quad \sum_{n=0}^{2N-1} n \phi'_N(n) = -1.$$

Les deux équations (2.4) et (2.5) sont du même type que (2.1) et (2.2) à un coefficient multiplicatif près. Ainsi, à l'aide de (2.4), (2.5) et (2.6) nous obtenons deux systèmes linéaires de $2N$ équations à $2N$ inconnues de même type que ceux nécessaires au calcul de ϕ_N et ψ_N sur une grille diadique. Nous obtenons donc de façon similaire les valeurs des dérivées premières sur cette grille diadique.

En dérivant une seconde fois l'équation d'échelle (2.1) et en supposant que $N \geq 7$ de telle sorte que ϕ_N et ψ_N soient de classe \mathcal{C}^2 , nous obtenons :

$$(2.7) \quad \phi''_N(x) = 4\sqrt{2} \sum_{k=0}^{2N-1} h_{N,k} \phi''_N(2x - k), \quad \forall x \in \mathbb{R},$$

et l'analogue pour ψ_N'' :

$$(2.8) \quad \psi_N''(x) = 4\sqrt{2} \sum_{k=2-2N}^1 g_{N,k} \phi_N''(2x-k), \quad \forall x \in \mathbb{R}.$$

La condition supplémentaire nécessaire au calcul sera déduite de :

$$\forall x \in \mathbb{R}, \quad \sum_{n \in \mathbb{Z}} n^2 \phi_N(x-n) = x^2 + ax + b$$

qui nous assure du fait que

$$(2.9) \quad \sum_{n \in \mathbb{Z}} n^2 \phi_N''(n) = 2.$$

Comme précédemment, la résolution des systèmes (2.7) , (2.8) et (2.9) nous permettent d'obtenir les valeurs exactes de ϕ_N'' et ψ_N'' sur une grille diadique.

2.2.2 Calcul de la transformée en ondelettes discrète des polynômes

On se place dans un espace d'approximation à l'échelle 2^{-j} : V_j . Il s'agit de calculer les coefficients $a_k^j = \langle P(x), \phi_{j,k} \rangle$ où $P(x)$ est un polynôme quelconque sur \mathbb{R} . Pour cela, il suffit de calculer ces coefficients pour les monômes :

$$\begin{aligned} a_k^j(x^n) &= \int_{\mathbb{R}} x^n \phi_{j,k}(x) dx = 2^{\frac{j}{2}} \int_{\mathbb{R}} x^n \phi(2^j x - k) dx \\ &= 2^{-j(n+\frac{1}{2})} \int_{\mathbb{R}} (x+k)^n \phi(x) dx. \end{aligned}$$

En développant $(x+k)^n$ par la formule du binôme :

$$a_k^j(x^n) = 2^{-j(n+\frac{1}{2})} \sum_{p=0}^n C_n^p k^{n-p} I_p$$

avec $I_p = \int_{\mathbb{R}} x^p \phi(x) dx$, et $I_0 = 1$. Nous avons alors

Propriété 2.1 Les I_p peuvent être calculées par la formule de récurrence :

$$\begin{cases} I_0 = 1, \\ \forall p \in \mathbb{N}, I_p = 2^{-(p+\frac{1}{2})} \sum_{m=0}^{2N-1} h_m \sum_{k=0}^p C_p^k m^{p-k} I_k. \end{cases}$$

Pour vérifier ceci, il suffit d'appliquer l'équation d'échelle à I_p :

$$\begin{aligned} I_p &= \int_{\mathbb{R}} x^p \sum_{m=0}^{2N-1} h_m \phi(2x - m) dx \\ &= \sqrt{2} \sum_{m=0}^{2N-1} h_m \int_{\mathbb{R}} \left(\frac{u+m}{2} \right)^p \phi(u) \frac{du}{2} \\ &= 2^{-(p+\frac{1}{2})} \sum_{m=0}^{2N-1} h_m \sum_{k=0}^p C_p^k m^{p-k} \int_{\mathbb{R}} u^k \phi(u) du. \end{aligned}$$

2.2.3 Mise en œuvre numérique d'une méthode de seuillage

Nous considérons le modèle non linéaire :

$$z = T(v) + \nu.$$

Nous allons identifier v par un algorithme séquentiel quadratique décrit au chapitre II, paragraphe 2.4. La seule différence sera qu'à la fin de chaque itération interne, où l'on a identifié v solution du problème de contrôle linéarisé, nous allons appliquer une méthode de seuillage sur les coefficients de v . Ne disposant pas d'un estimateur a priori du niveau de bruit, nous prendrons pour seuil

$$s = \sqrt{2 \log(j)} \left(2^{-(j-1)} \sum_{i=1}^{2^{-(j-1)}} \tilde{z}_i^2 \right)^{\frac{1}{2}}$$

où \tilde{z}_i sont les coefficients d'ondelettes de z à l'échelle la plus fine (i.e. les coefficients dans l'espace W_{j-1}). Connaissant les coefficients a_k^j , $k = 1, \dots, 2^j$ de v dans l'espace d'approximation V_j , nous en déduirons les nouveaux coefficients $a_k^{j_p}$ et $c_\ell^{j_p}$ par l'algorithme de décomposition de S. Mallat :

$$\begin{array}{ccccccc} a^j & \rightarrow & a^{j-1} & \rightarrow & a^{j-2} & & a^{j_p+1} & \rightarrow & a^{j_p} \\ & & \searrow & & \searrow & \dots & \searrow & & \\ & & c^{j-1} & & c^{j-2} & & c^{j_p+1} & & c^{j_p}. \end{array}$$

Nous transformons alors les c^{j_p} en \tilde{c}^{j_p} avec :

$$\tilde{c}_k^{j_p} = \begin{cases} c_k^{j_p} & \text{si } |c_k^{j_p}| > s, \\ 0 & \text{sinon.} \end{cases}$$

Il s'agit donc d'une méthode de seuillage dur, puisqu'elle supprime tous les coefficients plus petits que le seuil. Ensuite, à l'aide de l'algorithme de reconstruction de S. Mallat, nous revenons dans l'espace V_j par l'opération inverse de la précédente :

$$\begin{array}{ccccccc}
 a^{j_p} & \rightarrow & \tilde{a}^{j_p+1} & \rightarrow & \tilde{a}^{j+1} & \rightarrow & \tilde{a}^j \\
 & \nearrow & & \nearrow & \dots & \nearrow & \\
 \tilde{c}^{j_p} & & c^{j_p+1} & & c^{j+1} & &
 \end{array}$$

et nous poursuivons les itérations externes avec

$$v = \sum_k \tilde{a}_k^j \phi_{j,k}.$$

Cette méthode s'étant avérée plutôt décevante dans la pratique, nous n'en donnons que quelques résultats pour des données bruitées à 2% dans le problème d'identification de sources non linéaires.

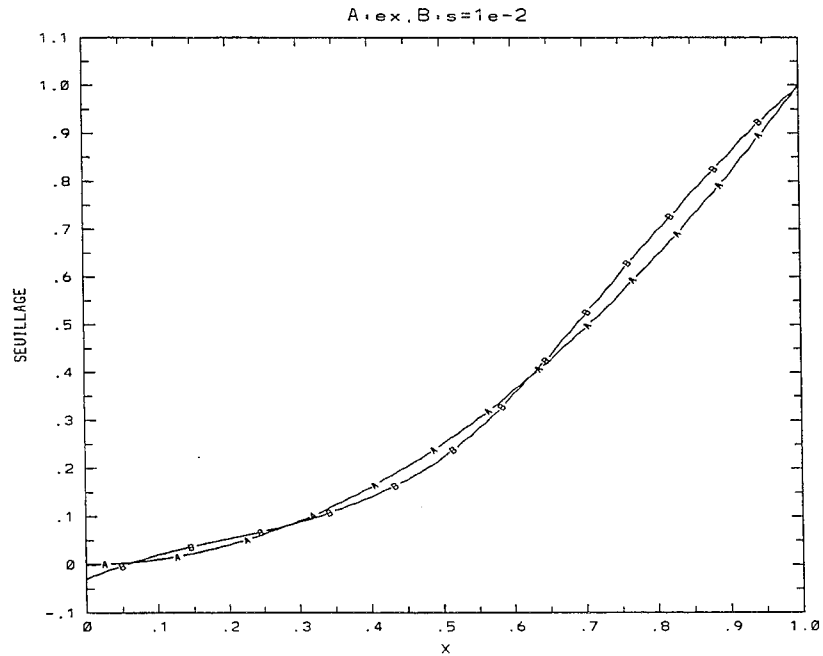


Figure 2.2 Identification de x^2 par seuillage dur

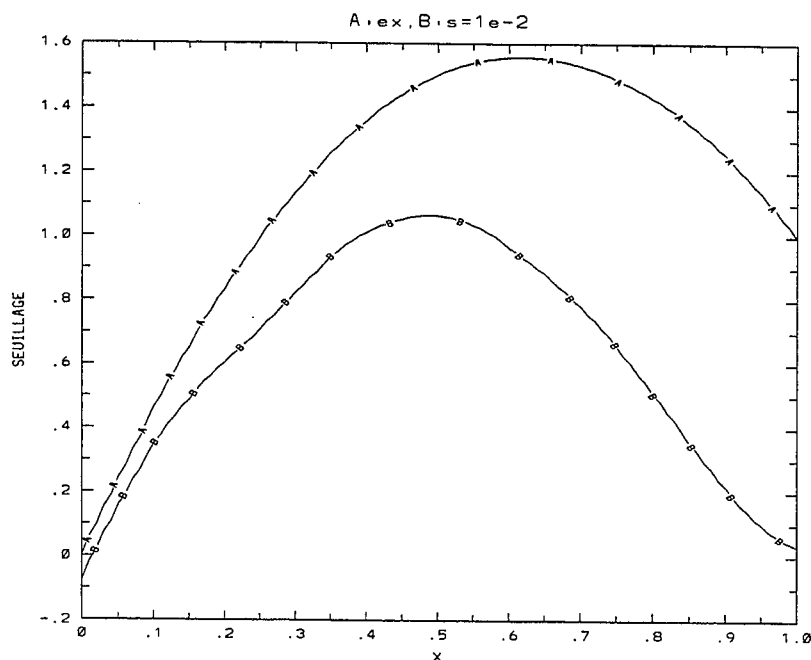


Figure 2.3 Identification de $\sin(\pi x) + x$ par seuillage dur

La figure 2.3 est une bonne illustration d'un lissage trop fort des données : la valeur du seuil choisie est trop importante, et donc nous avons supprimé trop de coefficients.

Nous avons constaté que même pour des données non bruitées, il était nécessaire d'utiliser une méthode de seuillage : sans seuillage, l'algorithme séquentiel quadratique ne converge pas. L'introduction de bruit dans les observations induit de fortes instabilités numériques : le problème inverse devient mal conditionné. Le seuillage ne fournit donc pas un bon estimateur de la non linéarité du problème. Nous allons voir au paragraphe suivant une manière de remédier à ces problèmes par un changement de base dans l'espace des contrôles par rapport à une norme d'un espace de Hilbert.

3 Identification par changement de base

Nous nous plaçons formellement dans le cadre général d'un problème inverse non linéaire. Nous nous donnons tout d'abord un espace de fonctions :

$$\mathcal{U}_{\text{ad}} = \{v \in E / v(x) = \sum_{i=1}^m a_i \phi_i(x), a_i \in \mathbb{R}\},$$

où \mathcal{U}_{ad} est un sous espace de E , espace de Hilbert, et où les fonctions $\{\phi_1, \dots, \phi_m\}$ sont une base fixée de $L^2(\mathbb{R})$. Il s'agit de déterminer v à partir d'observations bruitées

$$z = T(v) + \nu \quad \text{dans } F,$$

où F est un espace de Hilbert et T une application non linéaire de \mathcal{U}_{ad} vers F . Pour cela, nous nous proposons de minimiser la fonction coût :

$$J(v) = \frac{1}{2} \|T(v) - z\|_F^2$$

dans l'espace \mathcal{U}_{ad} . Le cadre est le même que celui des chapitres II et III, la seule différence réside dans le fait que nous n'avons plus de terme de régularisation dans la fonctionnelle à minimiser $J(v)$. Théoriquement, les problèmes que nous considérons étant mal posés, l'absence de régularisation devrait nuire à l'identification. Nous allons voir que grâce à une méthode de changement de base dans un espace hilbertien approprié, l'identification peut être réalisée même à partir de données bruitées.

Nous introduisons la méthode de changement de base dans un cadre général, nous l'expliciterons et la commenterons ensuite pour les problèmes d'identification de sources non linéaires et de la diffusion de la chaleur.

3.1 Changement de base dans l'espace des contrôles

Supposons que nous écrivions le contrôle v dans la base $\{\phi_1, \dots, \phi_m\}$ de \mathcal{U}_{ad} :

$$v = \sum_{j=1}^m X_j \phi_j = X^t \phi.$$

Les fonctions ϕ_i correspondent soit aux fonctions d'échelle : $\phi_i = \phi_{k,i}$ dans un espace d'approximation V_k de $L^2(\mathbb{R})$ fixé, la fonction d'échelle ϕ est alors égale à ϕ_7 , fonction à support compact de I. Daubechies si on choisit de travailler avec

les ondelettes, soit aux B -splines cubiques : $\phi_i = N_i^4$. Parmi ces fonctions, qui sont à support compact inclus dans \mathbb{R} , certaines n'ont pas leur support contenu dans $[0,1]$. Afin de prendre en compte ce fait, soit donc \mathbb{R}^m l'espace des contrôles discrétisé. Nous considérons un espace de Hilbert E , $\langle \cdot, \cdot \rangle_E$ son produit scalaire associé, et soit G la matrice de Gram symétrique définie positive des produits scalaires des fonctions de base $\{\phi_1, \dots, \phi_m\}$ dans cet espace E :

$$G_{i,j} = \langle \phi_i, \phi_j \rangle_E.$$

Dans la base canonique de \mathbb{R}^m , le produit scalaire de l'espace E est donc associé à la matrice G par :

$$\forall u, v \in E, \quad \langle u, v \rangle_E = u^t G v$$

si l'on identifie les fonctions u et v de E à leurs vecteurs de composantes dans la base des $\{\phi_1, \dots, \phi_m\}$.

Considérons $\{\vec{w}_1, \dots, \vec{w}_m\}$ une base de vecteurs propres de G . Cette base, correspondant à des valeurs propres décroissantes $\lambda_1 > \lambda_2 > \dots > \lambda_m > 0$, est orthonormée par rapport au produit scalaire euclidien :

$$G \vec{w}_i = \lambda_i \vec{w}_i, \quad \text{et} \quad \vec{w}_i^t \vec{w}_j = \delta_{ij}.$$

La base définie par les vecteurs $\{\vec{\Psi}_i\}$ avec $\vec{\Psi}_i = \frac{\vec{w}_i}{\sqrt{\lambda_i}}$ est donc orthonormée par rapport à la norme dans E . Notons $\Psi_{i,k}$ les composantes du i -ème vecteur $\vec{\Psi}_i$, et considérons la fonction $\tilde{\phi}_i(x) = \sum_{k=1}^m \Psi_{i,k} \phi_k(x)$ définie sur la base initiale que l'on a choisie ($\phi_k =$ ondelettes, B -splines). Les $\{\tilde{\phi}_i, i = 1, \dots, m\}$ forment une nouvelle base de fonctions de \mathcal{U}_{ad} . Le changement de coordonnées doit être effectué de la façon suivante :

$$f(x) = \sum_{i=1}^m X_i \phi_i(x) = \sum_{i=1}^m \tilde{X}_i \tilde{\phi}_i(x)$$

avec

$$\tilde{X}_i = \lambda_i \sum_{k=1}^m X_k \Psi_{i,k}.$$

Le changement de base inverse est donné par :

$$X_k = \sum_{i=1}^m \tilde{X}_i \Psi_{i,k}.$$

L'intérêt de minimiser la fonctionnelle $J(v)$ dans la base des $\{\tilde{\phi}_i\}$ réside dans le fait qu'effectuer un produit scalaire dans cette nouvelle base de fonctions revient

à faire un produit scalaire euclidien classique sur les composantes dans cette base :

$$\left. \begin{aligned} f &= \sum_i \tilde{X}_i \tilde{\phi}_i \\ g &= \sum_j \tilde{Y}_j \tilde{\phi}_j \end{aligned} \right\} \Rightarrow \langle f, g \rangle_E = \sum_k \tilde{X}_k \tilde{Y}_k = \tilde{X}^t \tilde{Y}.$$

3.2 Effet du changement de base sur le gradient

Nous avons procédé au paragraphe précédent à un changement de base sur l'espace des fonctions en même temps que sur les composantes du contrôle. Nous allons nous intéresser à ce que deviennent les composantes du gradient de la fonctionnelle à minimiser dans la nouvelle base.

Nous aurons besoin de définir la matrice orthogonale P dont les colonnes sont constituées des vecteurs propres de G : $\{\vec{w}_1, \dots, \vec{w}_m\}$. Soit $D = \text{diag}(\lambda_1, \dots, \lambda_m)$ la matrice diagonale des valeurs propres de G . Nous avons alors la relation : $G = PDP^t$. Les changements de coordonnées dans ces bases définies plus haut, peuvent s'exprimer aussi par :

$$(X)_{(\phi)} \leftrightarrow (\tilde{X})_{(\tilde{\phi})} \Leftrightarrow X = PD^{-\frac{1}{2}} \tilde{X} \Leftrightarrow \tilde{X} = D^{\frac{1}{2}} P^t X.$$

Si le contrôle v s'exprime par :

$$v = X^t \phi = \tilde{X}^t \tilde{\phi},$$

le gradient au point v dans l'ancienne base s'écrit :

$$\nabla J(v) = \sum_{i=1}^m S_i \phi_i \text{ avec } S_i = \frac{\partial J}{\partial X_i}(v),$$

et dans la nouvelle base, il sera :

$$\widetilde{\nabla J}(v) = \sum_{i=1}^m \xi_i \tilde{\phi}_i \text{ avec } \xi_i = \frac{\partial J}{\partial \tilde{X}_i}(v).$$

Lemme 3.1 Si S désigne le vecteur des composantes de $\nabla J(v)$ dans la base $\{\phi_1, \dots, \phi_m\}$: $\nabla J(v) = \sum_i S_i \phi_i$ alors le gradient de J dans la nouvelle base $\{\tilde{\phi}_1, \dots, \tilde{\phi}_m\}$ s'écrit, relativement à l'ancienne base $\{\phi_1, \dots, \phi_m\}$:

$$\widetilde{\nabla J}(v) = \sum_i \tilde{S}_i \phi_i \text{ avec } \tilde{S} = G^{-1} S.$$

Preuve : Le gradient dans la base des $\{\phi_1, \dots, \phi_m\}$ s'écrivait :

$$\nabla J(v) = \sum_i S_i \phi_i \quad \text{avec } S_i = \langle T'(v)^t (T(v) - z), \phi_i \rangle$$

dans la base des $\{\tilde{\phi}_1, \dots, \tilde{\phi}_m\}$, il a une expression analogue :

$$\nabla \widetilde{J}(v) = \sum_i \langle T'(v)^t (T(v) - z), \tilde{\phi}_i \rangle \tilde{\phi}_i.$$

Nous remplaçons maintenant $\tilde{\phi}_i$ par son expression en fonction des ϕ_j :

$$\tilde{\phi}_i = \sum_{k=1}^m \Psi_{i,k} \phi_k = \frac{1}{\sqrt{\lambda_i}} \sum_{k=1}^m P_{k,i} \phi_k,$$

et donc

$$\begin{aligned} \nabla \widetilde{J}(v) &= \sum_i \lambda_i^{-1} \sum_k P_{ki} \sum_n P_{ni} \langle T'(v)^t (T(v) - z), \phi_k \rangle \phi_n \\ &= \sum_n (PD^{-1}P^t S)_n \phi_n = \sum_n (G^{-1}S)_n \phi_n, \end{aligned}$$

d'où l'expression attendue. □

Dans la pratique, nous n'aurons pas nécessairement à calculer ces nouvelles fonctions de base $\tilde{\phi}_j$: il suffira de calculer les composantes du gradient dans l'ancienne base, puis de faire leur multiplication par G^{-1} .

Remarque 3.2 : Cette méthode est inspirée en grande partie des méthodes de décomposition en valeurs singulières d'une matrice (cf [BDP 85]). Ici, le problème étant non linéaire, la matrice à considérer sera le hessien de la fonction coût.

3.3 Effet du changement de bases sur le hessien

Nous allons maintenant nous intéresser à l'effet de ces changements de base sur le hessien du critère, au point que l'on a identifié. Dans l'ancienne base $\{\phi_1, \dots, \phi_m\}$, le hessien était représenté par la matrice

$$H = (h_{ij}) \quad \text{avec } h_{ij} = (\phi_j, (T'(v)^t T'(v) + (T(v) - z)^t T''(v)) \phi_i)$$

et $\text{Hess}J(v)\phi_j = \sum_i h_{ij}\phi_i$. Le hessien de J au point v dans la base $\{\tilde{\phi}_i, i = 1, \dots, m\}$, relativement à l'ancienne base $\{\phi_i, i = 1, \dots, m\}$, va s'écrire :

$$\widetilde{\text{Hess}J}(v)\tilde{\phi}_j = \sum_i \tilde{h}_{ij}\tilde{\phi}_i \quad \text{avec } \tilde{h}_{ij} = (\tilde{\phi}_j, (T'(v)^t T'(v) + (T(v) - z)^t T''(v))\tilde{\phi}_i).$$

Nous en déduisons :

$$\begin{aligned} \tilde{h}_{ij} &= \lambda_i^{-\frac{1}{2}} \lambda_j^{-\frac{1}{2}} \sum_{k,\ell} P_{kj} P_{\ell i} (\phi_k, (T'(v)^t T'(v) + (T(v) - z)^t T''(v))\phi_\ell) \\ &= \lambda_i^{-\frac{1}{2}} \lambda_j^{-\frac{1}{2}} \sum_{k,\ell} P_{kj} P_{\ell i} h_{k\ell} \\ &= \lambda_i^{-\frac{1}{2}} \lambda_j^{-\frac{1}{2}} \sum_k P_{kj} \sum_\ell h_{k\ell} P_{\ell i} \\ &= \lambda_i^{-\frac{1}{2}} \lambda_j^{-\frac{1}{2}} \sum_k P_{kj} (HP)_{ki} = \lambda_i^{-\frac{1}{2}} \lambda_j^{-\frac{1}{2}} (P^t HP)_{ji} \\ &= (D^{-\frac{1}{2}} P^t HP D^{-\frac{1}{2}})_{ij} \end{aligned}$$

d'où finalement : le hessien de J au point v dans la nouvelle base est représenté, dans l'ancienne base, par la matrice

$$(3.1) \quad \tilde{H} = D^{-\frac{1}{2}} P^t HP D^{-\frac{1}{2}}.$$

Nous rappelons que le conditionnement d'une matrice A est donné par le rapport :

$$\text{Cond}(A) = \frac{\lambda_{\max}}{\lambda_{\min}}$$

si λ_{\max} est la plus grande valeur singulière de A (c'est à dire que λ_{\max}^2 est valeur propre de $A^t A$) et λ_{\min} la plus petite. Nous avons le résultat suivant, concernant le conditionnement de \tilde{H} :

Lemme 3.3 *Le conditionnement de la matrice \tilde{H} du hessien dans la nouvelle base est le rapport $\frac{\mu_{\max}}{\mu_{\min}}$ où les μ_i sont les valeurs propres de $G^{-1}H$.*

Preuve : Puisque la matrice \tilde{H} est symétrique, ses valeurs singulières coïncident avec ses valeurs propres. Soit μ une valeur propre de \tilde{H} , alors il existe un vecteur \tilde{X} non nul tel que :

$$\tilde{H}\tilde{X} = \mu\tilde{X}.$$

En utilisant la relation (3.1) , nous obtenons :

$$\begin{aligned} H\tilde{X} = \mu\tilde{X} &\iff D^{-\frac{1}{2}} P^t HP D^{-\frac{1}{2}} \tilde{X} = \mu\tilde{X} \\ &\iff P^t HP D^{-\frac{1}{2}} \tilde{X} = \mu D^{\frac{1}{2}} \tilde{X} \\ &\iff HP D^{-\frac{1}{2}} \tilde{X} = \mu P D^{\frac{1}{2}} \tilde{X}, \end{aligned}$$

et le vecteur \tilde{X} s'écrit dans l'ancienne base : $\tilde{X} = D^{\frac{1}{2}}P^t X$ avec X vecteur non nul, puisque D et P sont deux matrices inversibles. Donc

$$\begin{aligned} H\tilde{X} = \mu\tilde{X} &\iff HX = \mu PDP^t X \\ &\iff HX = \mu GX \\ &\iff G^{-1}HX = \mu X \end{aligned}$$

et donc nous avons montré que toute valeur propre de \tilde{H} est valeur propre de $G^{-1}H$ et réciproquement. \square

Remarque 3.4 : Ce changement de bases sur le contrôle se traduit donc par un "préconditionnement" à gauche du hessien par la matrice G^{-1} .

3.4 Changement de base pour l'identification de sources non linéaires

L'algorithme utilise toujours la méthode séquentielle quadratique décrite au chapitre II, paragraphe 2.4, à laquelle nous appliquons le changement de base par rapport à la norme d'un espace de Hilbert E . Dans la pratique, nous utiliserons surtout pour E les espaces : $L^2[0, 1]$, $H^1[0, 1]$, $H^2[0, 1]$, et des espaces que nous noterons $H_{\alpha, \beta, \gamma}^2[0, 1]$, dont la norme est équivalente à celle des espaces précédents selon les valeurs des paramètres α , β , γ , tous positifs (et $\alpha > 0$).

Il est à noter que selon la non linéarité que l'on veut identifier, et selon le pourcentage de bruit sur les données, l'espace de Hilbert E ne sera pas choisi de la même manière. Ainsi, en toute généralité, dans l'espace $H_{\alpha, \beta, \gamma}^2[0, 1]$ défini ci-dessus, la matrice G sera composée des coefficients :

$$(3.2) \quad G_{k\ell} = \alpha \int_0^1 \phi_{j,k} \phi_{j,\ell} dx + \beta \int_0^1 \phi'_{j,k} \phi'_{j,\ell} dx + \gamma \int_0^1 \phi''_{j,k} \phi''_{j,\ell} dx,$$

où les poids α , β , et γ seront déterminés en fonction du problème à résoudre.

3.4.1 Quelques résultats avec les ondelettes dans V_3

Nous nous intéressons tout d'abord à l'effet de changement de base sur la méthode séquentielle quadratique. Pour cela nous représentons le spectre de la matrice $T'(v)^t T'(v)$ au voisinage du point v identifié pour les normes L^2 et $H_{1,1,100}^2$ utilisées sur un espace d'approximation V_j , correspondant aux ondelettes à support compact de I. Daubechies. Dans ce paragraphe, nous choisirons $j = 3$, de façon à travailler dans l'espace V_3 et que la projection d'une fonction quelconque f dans V_3 approche assez bien la fonction f . Nous rappelons que pour ce problème d'identification de sources non linéaires, $T(v)$ est égale à la dérivée normale de l'état $\frac{\partial y}{\partial n}$ sur le bord du domaine.

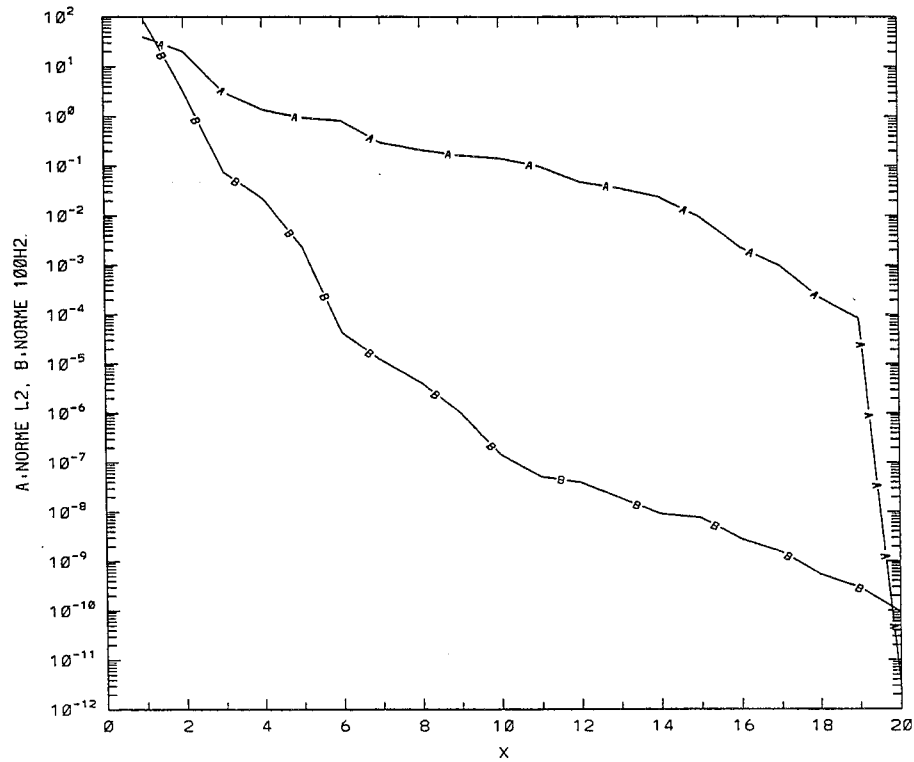


Figure 3.4 Spectres de $T'(v)^t T'(v)$ en norme L^2 (courbe A) et $H^2_{1,1,100}$ (B)

On constate donc que la matrice en norme L^2 est un peu mieux conditionnée qu'en norme $H^2_{1,1,100}$. Nous traçons ci-dessous les six premiers vecteurs propres (i.e. ceux qui correspondent aux valeurs propres les plus grandes) de $T'(v)^t T'(v)$ pour les normes L^2 et $H^2_{1,1,100}$, dans la nouvelle base $\{\tilde{\phi}_1, \dots, \tilde{\phi}_m\}$.

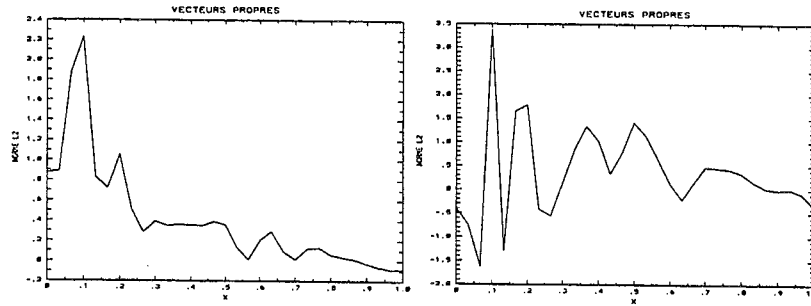


Figure 3.5 Vecteurs propres no 1 et 2 de $T'(v)^t T'(v)$ en norme L^2

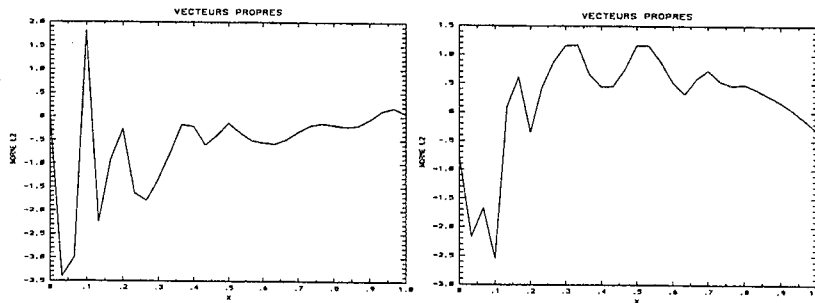


Figure 3.6 Vecteurs propres no 3 et 4 de $T'(v)^t T'(v)$ en norme L^2

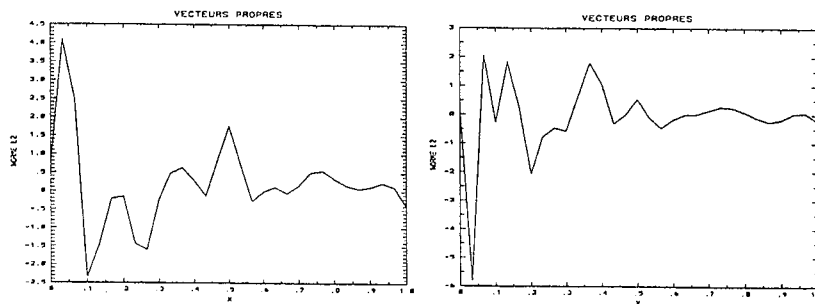


Figure 3.7 Vecteurs propres no 5 et 6 de $T'(v)^t T'(v)$ en norme L^2

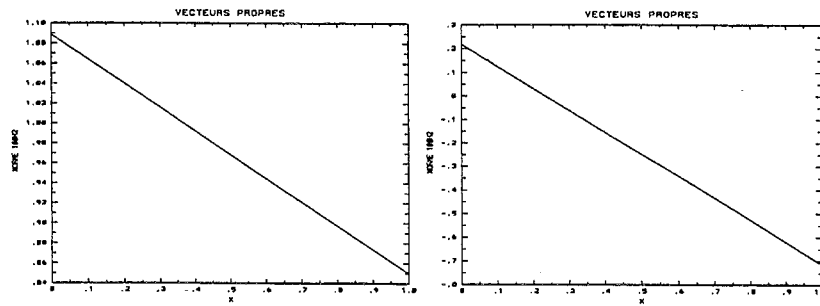


Figure 3.8 Vecteurs propres no 1 et 2 de $T'(v)^t T'(v)$ en norme $H^2_{1,1,100}$

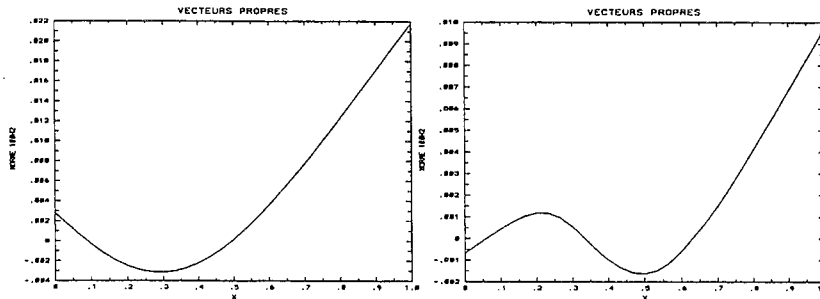


Figure 3.9 Vecteurs propres no 3 et 4 de $T'(v)^t T'(v)$ en norme $H^2_{1,1,100}$

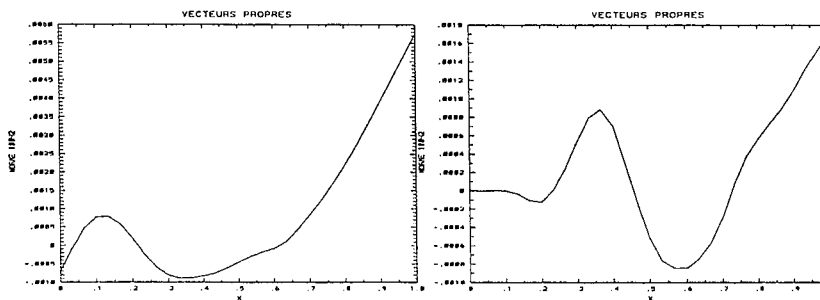


Figure 3.10 Vecteurs propres no 5 et 6 de $T'(v)^t T'(v)$ en norme $H^2_{1,1,100}$

Nous constatons donc que les vecteurs propres de l'opérateur $T'(v)^t T'(v)$ sont nettement plus réguliers dans le cas de la norme $H_{1,1,100}^2$ qu'avec la norme L^2 , ce qui, comme nous le verrons avec les expériences numériques, explique que l'identification est bien mieux réalisée lorsqu'on utilise une norme équivalente à H^2 .

Nous donnons tout d'abord les résultats de l'identification de la fonction x^2 sur l'ellipse d'élongation 2, avec 5% de bruit sur les données dans la base d'ondelettes à support compact de I. Daubechies.

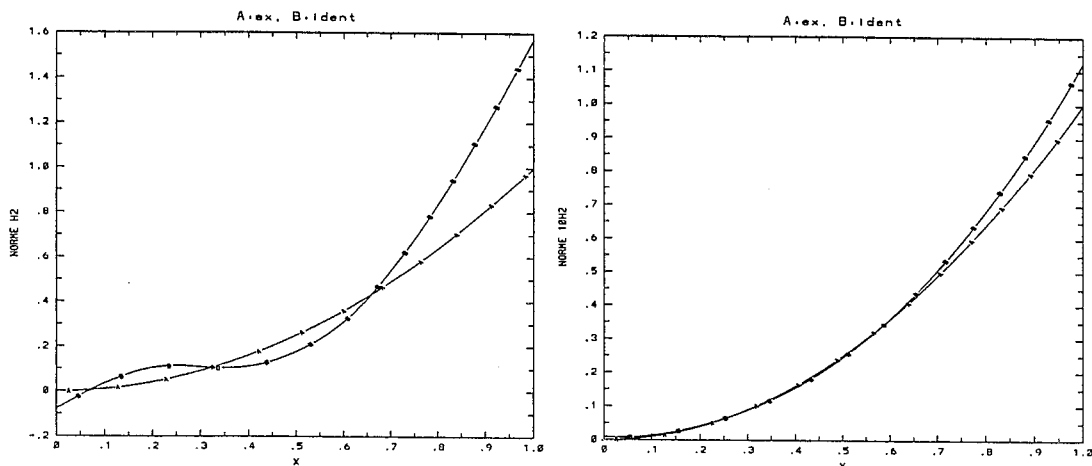


Figure 3.11 A gauche, norme H^2 , à droite norme $H_{1,1,10}^2$ pour x^2

Pour la figure 3.11, on constate que la fonction identifiée avec une norme H^2 ($\alpha = 1$, $\beta = 1$ et $\gamma = 1$) est très peu satisfaisante : numériquement cela se traduit par une mauvaise convergence de la méthode de gradient conjugué : à chaque itération interne, la norme du gradient ne diminue pas beaucoup. Pour une norme $H_{1,1,10}^2$ par contre, les résultats sont nettement meilleurs. La figure 3.12 représente l'identification avec deux normes équivalentes à H^2 : $H_{1,1,100}^2$ et $H_{1,1,1000}^2$, avec un poids très important devant le terme semi-norme H^2 . Pour cette fonction x^2 , nous avons constaté qu'il nous fallait prendre une norme équivalente à la norme de H^2 : pour la norme L^2 , l'algorithme n'a pas convergé. En termes de minimisation, pour les figure 3.11 et figure 3.12, la fonctionnelle qui valait $J(v_0) = 0,82$ au départ, a diminué jusqu'à $1.2 \cdot 10^{-5}$ dans les quatre cas. Le critère de décroissance de cette fonctionnelle n'est donc pas nécessairement synonyme de bonne identification. En termes de temps de calculs, pour la norme H^2 , 10 itérations externes sont nécessaires, soit environ 81 secondes CPU d'un CRAY C90 ; pour la norme $H_{1,1,10}^2$, le même nombre d'itérations externes suffit,

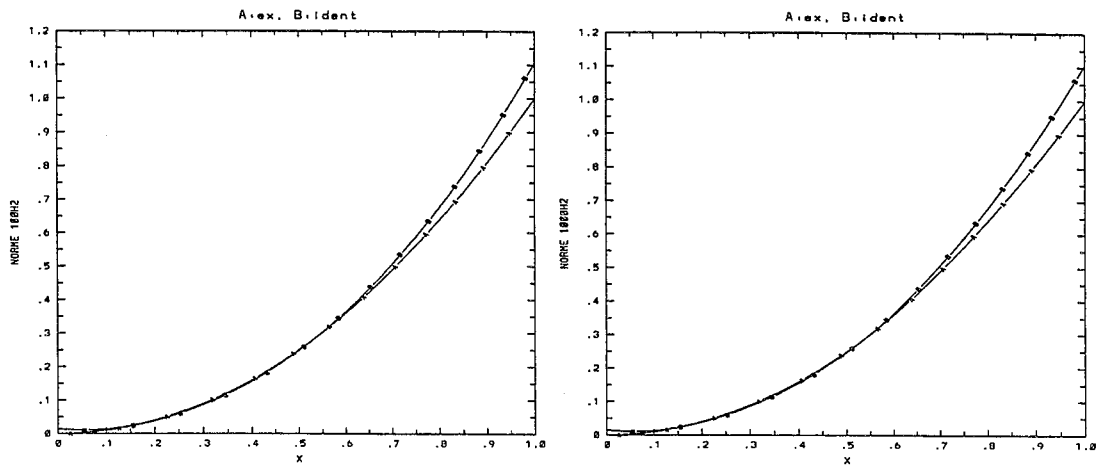


Figure 3.12 A gauche, norme $H^2_{1,1,100}$, à droite $H^2_{1,1,1000}$ pour x^2

et l'algorithme converge en 55 secondes CPU. Pour les deux dernières normes $H^2_{1,1,100}$ et $H^2_{1,1,1000}$, la fonction identifiée est la même, ce sont les temps de calcul qui diminuent : 42 secondes CPU pour $H^2_{1,1,100}$ et 30 secondes CPU pour $H^2_{1,1,1000}$.

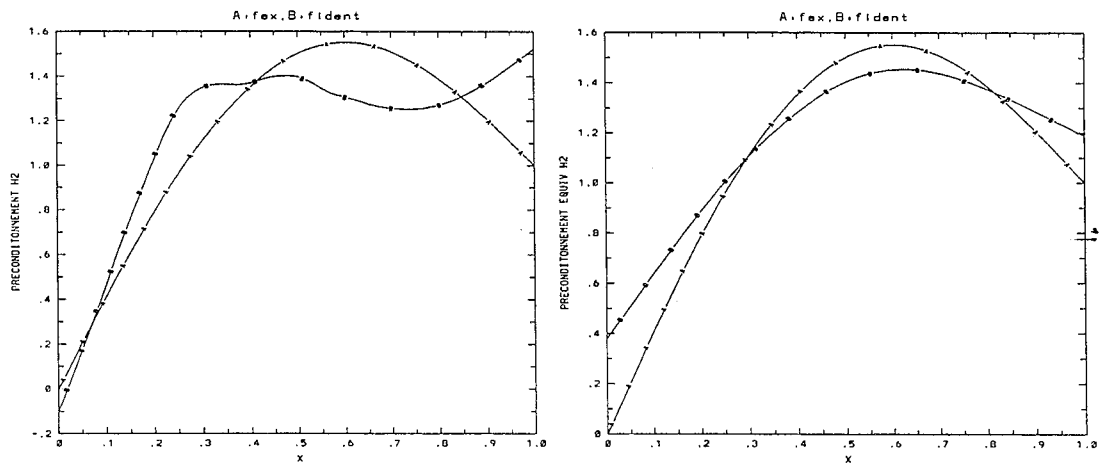


Figure 3.13 A gauche, norme H^2 , à droite norme $H^2_{1,1,10}$ pour $\sin(\pi x) + x$

Nous donnons en figure 3.13 des résultats d'identification de la fonction $\sin(\pi x) + x$. Pour cette fonction, nous avons constaté que le programme ne converge pas pour une norme H^1 ou L^2 . On voit que la norme H^2 ne suffit pas, il est nécessaire de mettre un poids de 10 devant la semi-norme H^2 . Pour la figure de gauche, l'algorithme converge au bout de 6 itérations externes, soit 63 secondes CPU d'un CRAY C90, pour celle de droite, 5 itérations externes suffisent, c'est à dire 42 secondes CPU. Dans les deux cas, la fonction coût J décroît de 0.30 à $4.5 \cdot 10^{-4}$.

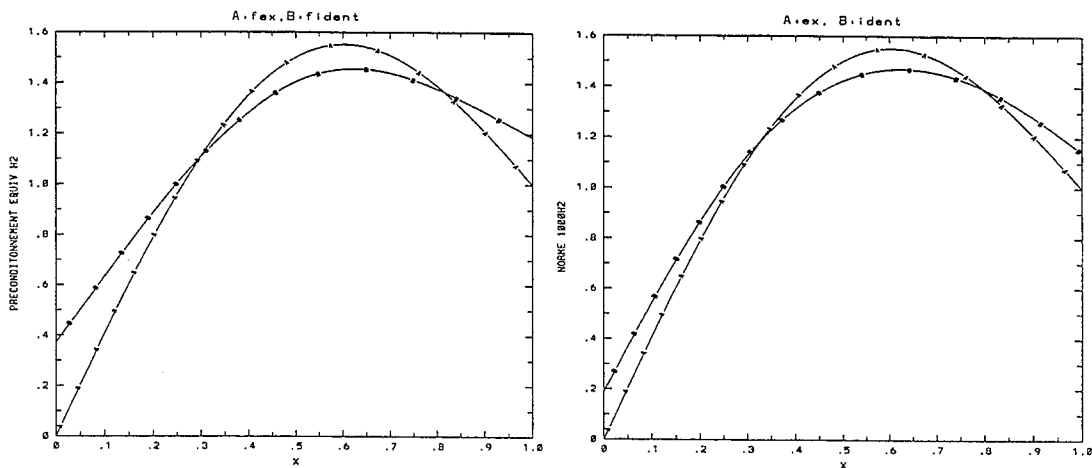


Figure 3.14 A gauche, norme $H_{1,1,100}^2$, à droite $H_{1,1,1000}^2$ pour $\sin(\pi x) + x$

La figure 3.14 représente l'identification de la fonction $\sin(\pi x) + x$ pour des normes plus régulières au sens de la semi-norme H^2 . On constate que plus le poids de la semi-norme H^2 augmente, et mieux l'identification est réalisée.

Nous voyons sur les figure 3.15 et figure 3.16 les résultats de l'identification de la fonction $x(4x^2 - \frac{9}{2}x + \frac{3}{2})$. Pour cette fonction identifiée, la fonctionnelle a diminué à chaque fois de 0.83 à $9.6 \cdot 10^{-6}$. A nouveau, même si le fait d'augmenter le poids de la semi-norme H^2 ne change pas la forme de la fonction identifiée, les temps de calculs diminuent considérablement. Il est à noter que l'algorithme n'arrive pas à converger si l'on prend une norme L^2 ou H^1 .

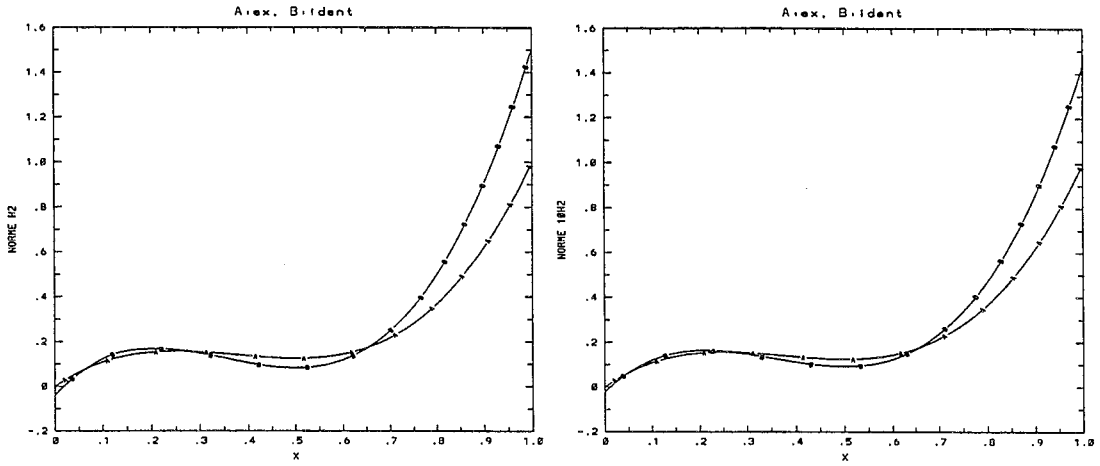


Figure 3.15 Norme H^2 à gauche, $H^2_{1,1,10}$ à droite pour $x(4x^2 - \frac{9}{2}x + \frac{3}{2})$

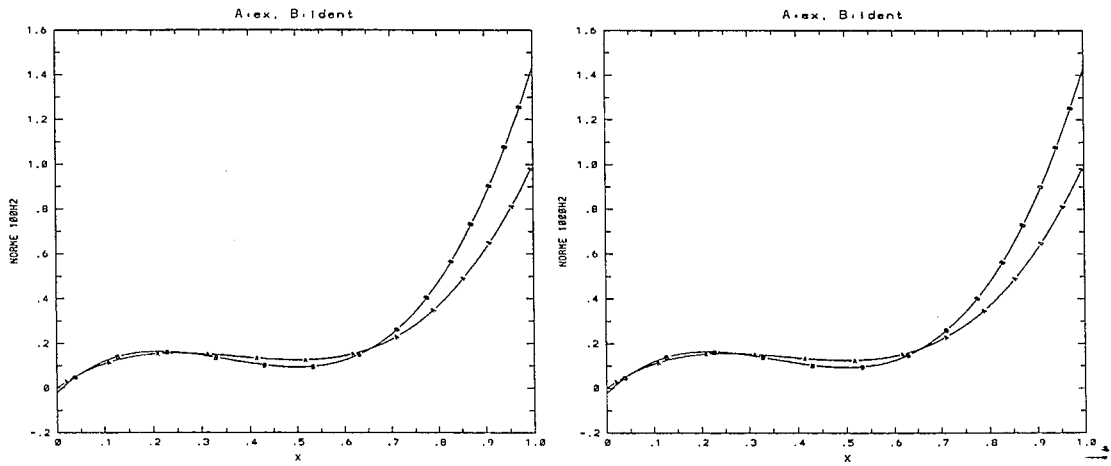


Figure 3.16 Norme $H^2_{1,1,100}$ à gauche, $H^2_{1,1,1000}$ à droite pour $x(4x^2 - \frac{9}{2}x + \frac{3}{2})$

Il s'agit pour les figure 3.17 , figure 3.18 et figure 3.19 de l'identification d'un polynôme de degré 2: $-4x^2 + 5x$. Ces figures illustrent bien l'importance de la norme intervenant dans le changement de base: une norme H^1 ne lisse pas suffisamment les données, alors qu'une norme équivalente à H^2 , avec des poids

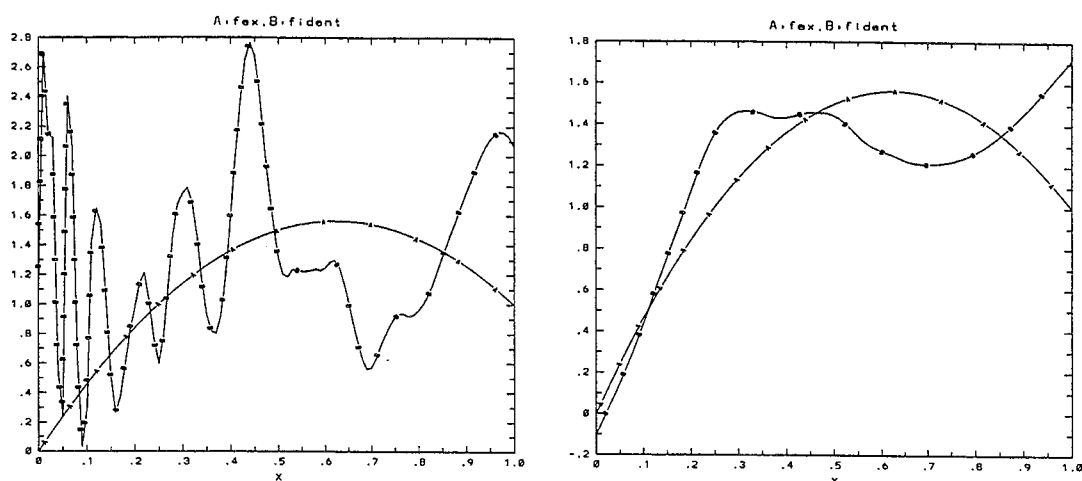


Figure 3.17 A gauche norme H^1 , à droite H^2 pour $-4x^2 + 5x$

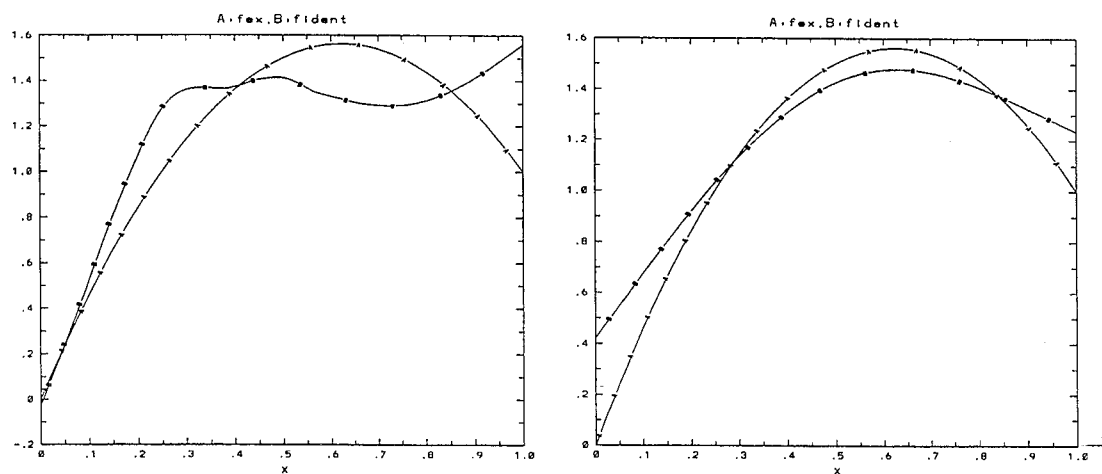


Figure 3.18 Norme $H^2_{1,1,10}$ à gauche et $H^2_{1,1,100}$ à droite pour $-4x^2 + 5x$

judicieusement choisis, donne de bien meilleurs résultats. Dans tous les cas, la fonction J a décréu de la valeur 0.29 à $4.5 \cdot 10^{-4}$. Les temps d'exécution sont plus ou moins élevés selon la norme choisie : 182 secondes CPU (21 itérations externes) pour H^1 , 205 secondes CPU (26 itérations externes) pour H^2 , 76 secondes CPU

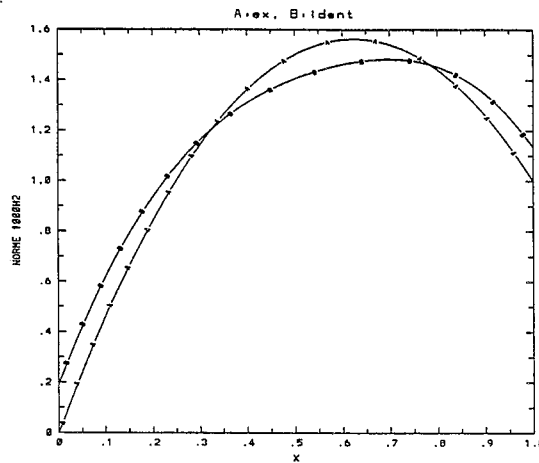


Figure 3.19 Norme $H_{1,1,1000}^2$ pour $-4x^2 + 5x$

(12 itérations externes) pour $H_{1,1,10}^2$, 39 secondes CPU (5 itérations externes) pour $H_{1,1,100}^2$, et 23 secondes CPU pour $H_{1,1,1000}^2$. Ces temps d'exécution sont donc liés à la norme choisie pour réaliser l'identification. Pratiquement, nous constatons que nous obtenons une bonne identification avec les normes $H_{1,1,10}^2$ ou $H_{1,1,100}^2$ selon les cas.

Nous donnons à présent des résultats d'identification dans l'espace d'approximation V_0 pour des polynômes uniquement, puisque si les polynômes appartiennent à V_0 , et donc admettent une représentation exacte dans cette base, pour des fonctions non polynomiales, la projection dans V_0 ne sera certainement plus un bon approximant de cette fonction.

3.4.2 Résultats pour les ondelettes dans V_0

Nous comparons dans ce paragraphe les résultats de l'identification par changement de norme dans les espaces V_0 et V_3 . Il est à noter que le nombre de fonctions de base, et donc de paramètres varie : dans V_3 nous en avons 20 contre 13 dans V_0 .

Nous représentons tout d'abord les résultats obtenus dans V_0 pour le polynôme x^2 en figure 3.20 et figure 3.21. On constate que dans V_0 , l'algorithme parvient à converger en norme L^2 , ce qui n'était pas le cas dans V_3 . Pour les normes H^2 et $H_{1,1,100}^2$, les résultats sont comparables dans les deux espaces V_0 et V_3 : en norme

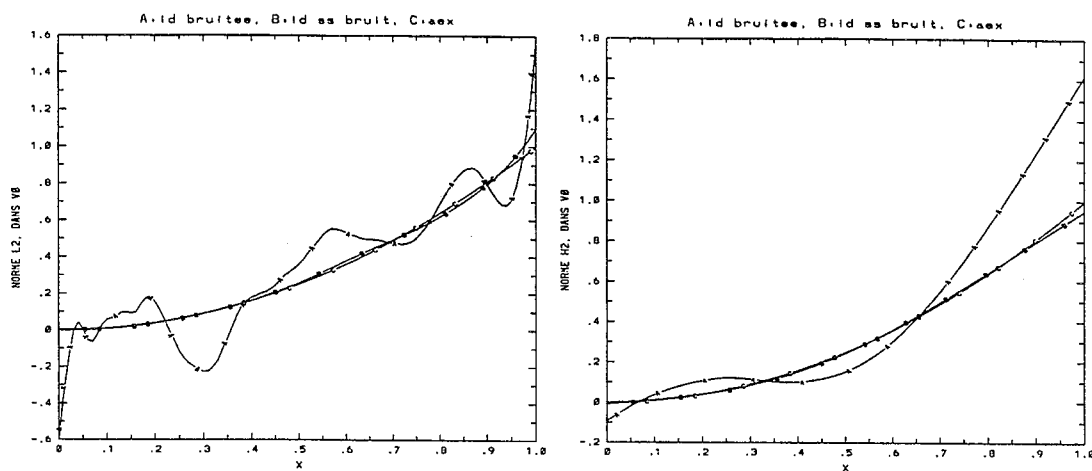


Figure 3.20 A gauche norme L^2 , à droite norme H^2 pour x^2 .

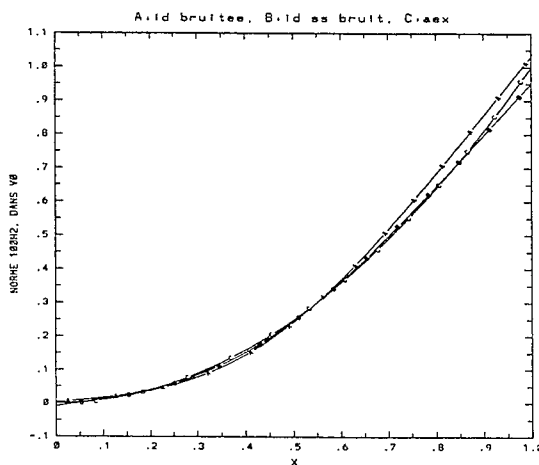


Figure 3.21 Norme $H^2_{1,1,100}$ pour x^2 .

H^2 la fonction n'est pas bien identifiée alors qu'en norme $H^2_{1,1,100}$ on retrouve bien sa forme.

En figure 3.22 les résultats sont présentés pour le polynôme $-4x^2 + 5x$. Pour cette fonction, l'algorithme ne converge pas en norme L^2 . A nouveau, les fonctions

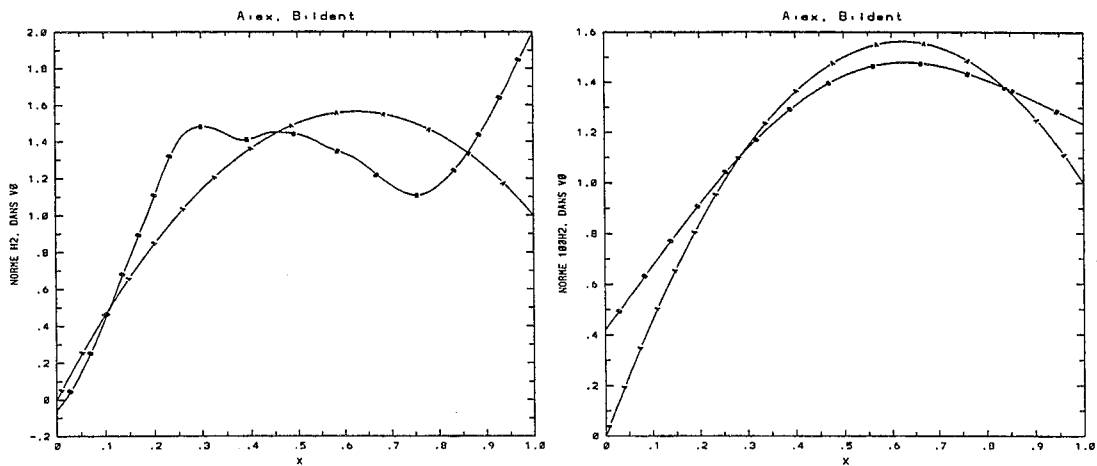


Figure 3.22 A gauche, norme H^2 , à droite norme $H^2_{1,1,100}$ pour $-4x^2 + 5x$.

identifiées en norme H^2 et $H^2_{1,1,100}$ sont comparables à celles obtenues dans V_3 . Les temps de calcul sont nettement significatifs : 27 itérations externes en norme H^2 contre 5 en norme $H^2_{1,1,100}$.

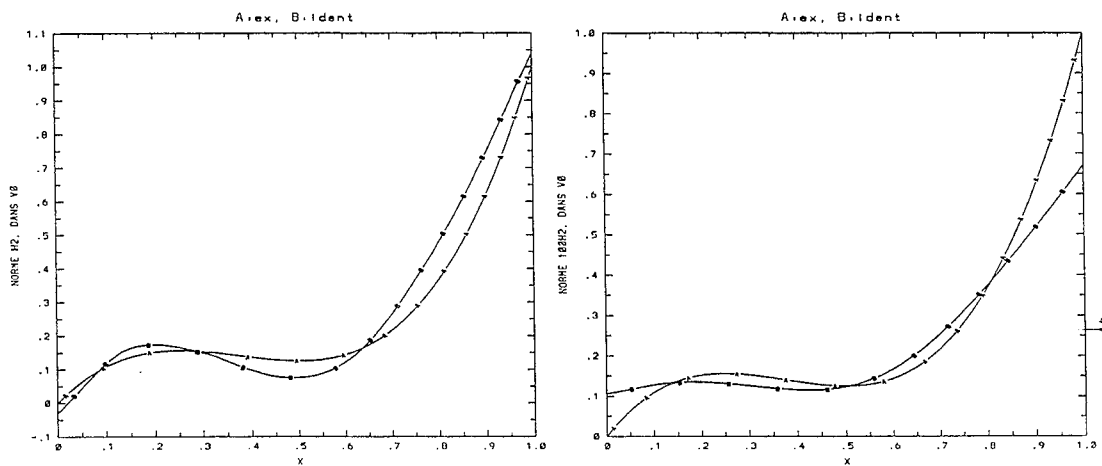


Figure 3.23 A gauche, norme H^2 , à droite norme $H^2_{1,1,100}$ pour $x(4x^2 - \frac{9}{2}x + \frac{3}{2})$.

La figure 3.23 concerne les résultats de l'identification de la fonction

$x(4x^2 - \frac{9}{2}x + \frac{3}{2})$ dans V_0 . En norme H^2 , le résultat est comparable à celui obtenu dans V_3 tandis qu'avec la norme $H^2_{1,1,100}$ la fonction identifiée paraît plus "lisse" que la vraie fonction.

En conclusion, les résultats dans V_0 , quand l'algorithme converge, sont tout à fait comparables à ceux obtenus dans V_3 , sauf pour la dernière fonction $x(4x^2 - \frac{9}{2}x + \frac{3}{2})$. Nous avons constaté également que le critère de décroissance de la fonction coût ne signifiait pas nécessairement une bonne identification ; nous allons voir au chapitre suivant comment, à partir des valeurs des courbures du modèle au point identifié, on peut savoir si l'estimation réalisée est acceptable d'un point de vue statistique. En effet, dans la pratique, la fonction exacte est inconnue, et on ne peut donc la comparer avec la fonction identifiée.

Nous donnons d'abord une comparaison avec une méthode analogue de changement de base appliquée aux fonctions splines.

3.4.3 Comparaison avec les splines

La méthode utilisée est essentiellement la même que précédemment, il n'y a que la matrice G des produits scalaires des fonctions de base qui va changer. Nous présentons tout d'abord quelques résultats d'identification de la source x^2 . Les données sont toujours bruitées à 5%, le domaine d'identification est le domaine de \mathbb{R}^2 délimité par l'ellipse de centre $(0,0)$, d'axes 1 en x et 2 en y .

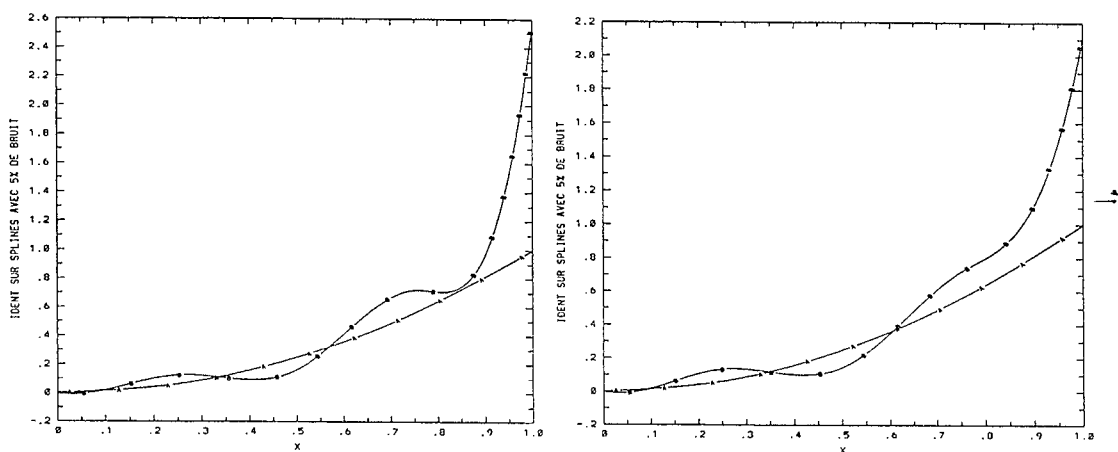


Figure 3.24 A gauche, norme L^2 , à droite H^1 pour x^2

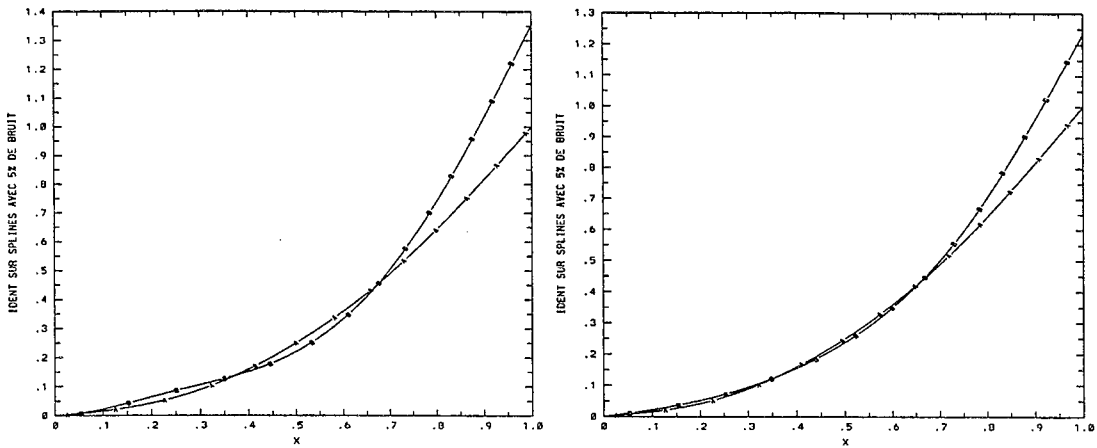


Figure 3.25 A gauche, norme H^2 , à droite $H^2_{1,1,10}$ pour x^2

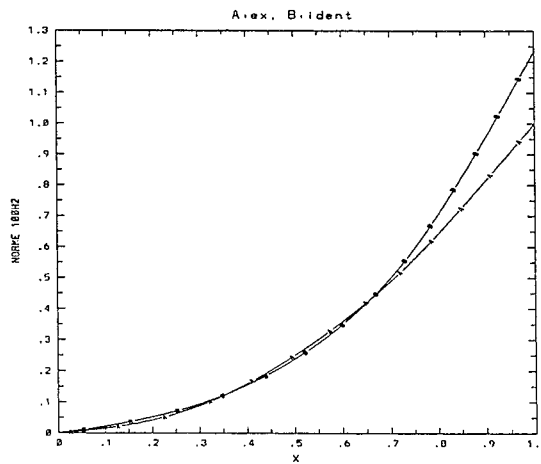


Figure 3.26 Norme $H^2_{1,1,100}$ pour x^2

En ce qui concerne les figure 3.24 , figure 3.25 et figure 3.26 , il s'agit de la source x^2 . La fonction coût décroît de $1.2 \cdot 10^{-3}$ à $1.2 \cdot 10^{-5}$ dans tous les cas. Pour la norme L^2 , le temps d'exécution est de 40 secondes CPU (soit 15 itérations externes) contre 78 secondes CPU (26 itérations externes) pour la norme

H^1 . Lorsqu'on prend une norme H^2 , les temps de calcul diminuent : 29 secondes CPU pour H^2 (12 itérations externes), 8 secondes CPU pour $H_{1,1,10}^2$ (4 itérations externes) et 7 secondes CPU pour $H_{1,1,100}^2$ (4 itérations externes).

On peut tout d'abord constater que les temps de calcul sont nettement inférieurs à ceux obtenus dans le cas des ondelettes. Ceci est dû au fait que pour les ondelettes nous avons 20 fonctions de base, alors que dans le cas des splines, 7 sont suffisantes. Il en résulte que les itérations internes des splines convergent plus vite que dans le cas des ondelettes. D'autre part, pour cette source particulière x^2 , le résultat de l'identification est moins sensible à la norme choisie pour le changement de base que dans le cas des ondelettes (figure 3.11 et figure 3.12).

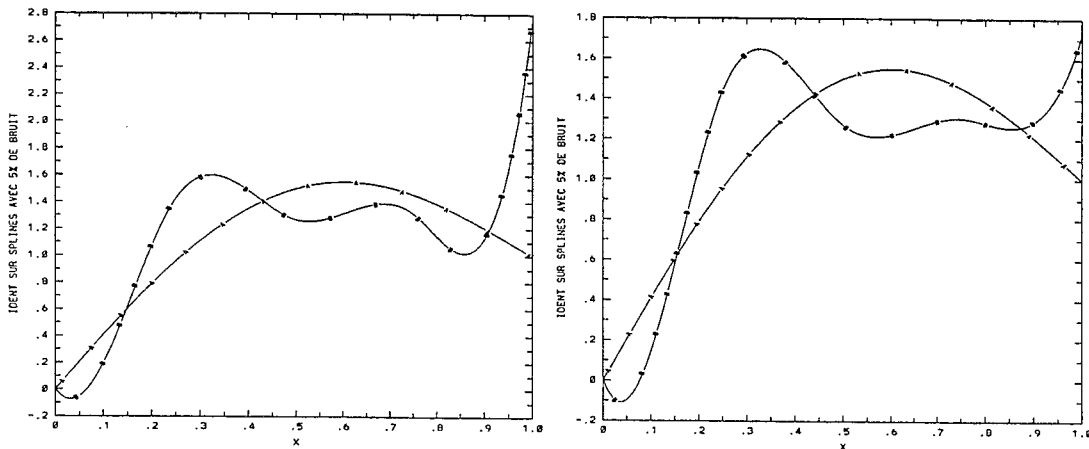


Figure 3.27 A gauche, norme L^2 , à droite H^1 pour $\sin(\pi x) + x$

Nous donnons en figure 3.27, figure 3.28 et figure 3.29 des résultats pour la fonction $\sin(\pi x) + x$. On constate que même avec un fort coefficient devant le terme de la semi-norme H^2 , la fonction n'est pas très bien identifiée. Les temps de calcul sont ici, pour H^2 : 32 secondes CPU, c'est à dire 14 itérations externes, et pour la norme équivalente à H^2 : 12 secondes CPU, soit 7 itérations externes.

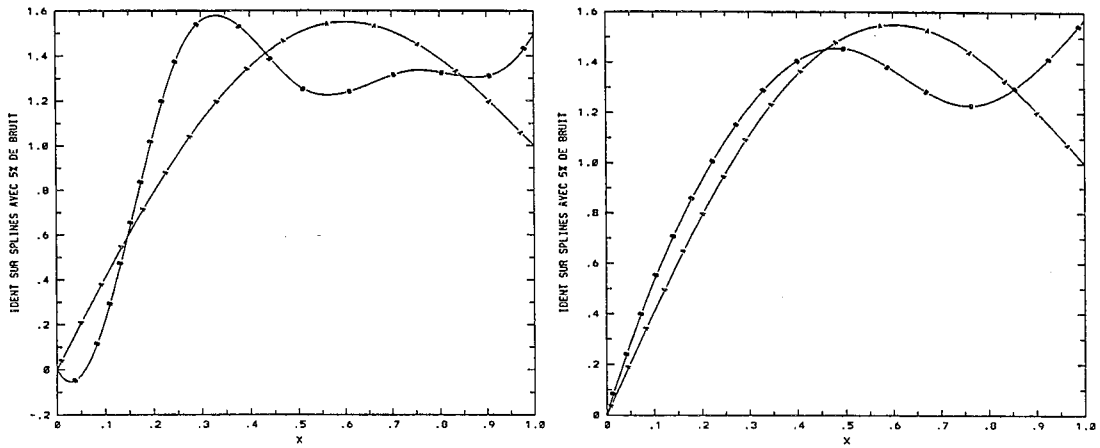


Figure 3.28 A gauche, norme H^2 , à droite $H^2_{1,1,10}$ pour $\sin(\pi x) + x$

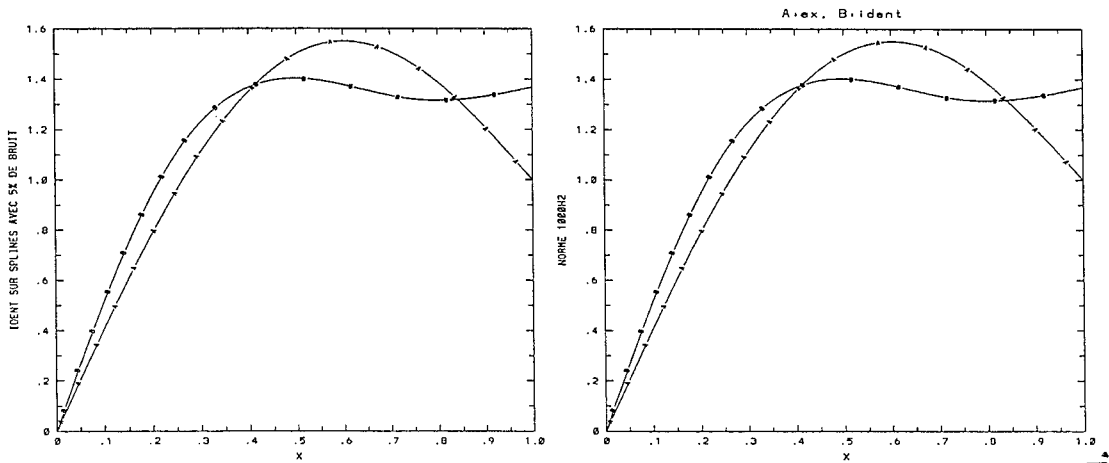


Figure 3.29 A gauche, norme $H^2_{1,1,100}$, à droite $H^2_{1,1,1000}$ pour $\sin(\pi x) + x$

Dans ce cas-là, contrairement à la fonction x^2 , les résultats étaient nettement plus satisfaisants avec les ondelettes (cf figure 3.13). Il est à noter qu'ici, la fonction coût J décroît à chaque fois de la valeur 0.15 à $4.5 \cdot 10^{-4}$.

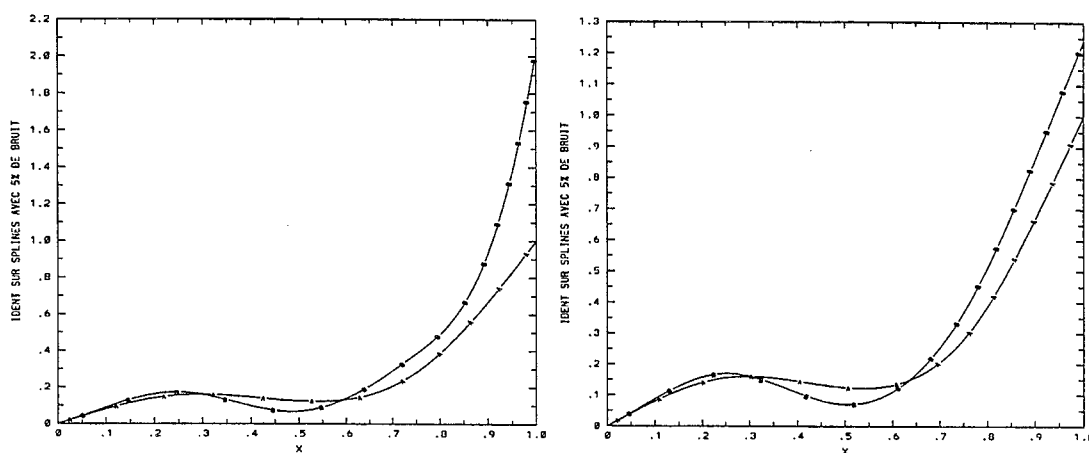


Figure 3.30 A gauche, norme L^2 , à droite H^1 pour $x(4x^2 - \frac{9}{2}x + \frac{3}{2})$

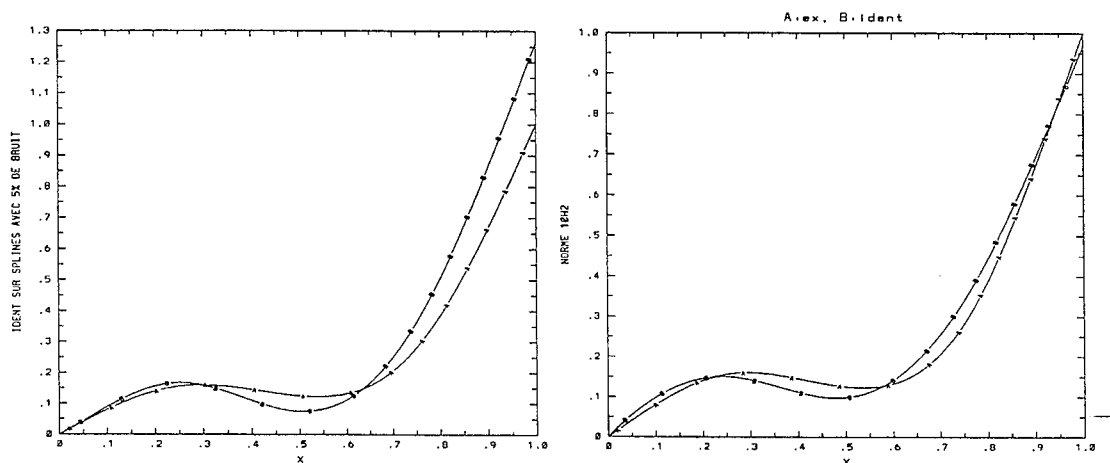


Figure 3.31 A gauche, norme H^2 , à droite $H_{1,1,10}^2$ pour $x(4x^2 - \frac{9}{2}x + \frac{3}{2})$

Nous poursuivons cette comparaison en donnant (figure 3.30 , figure 3.31 et figure 3.32) les résultats obtenus pour la source $x(4x^2 - \frac{9}{2}x + \frac{3}{2})$.

Il est intéressant de noter ici que l'identification obtenue par norme L^2 ou H^1 est acceptable, contrairement au cas des ondelettes, où l'on n'avait même pas

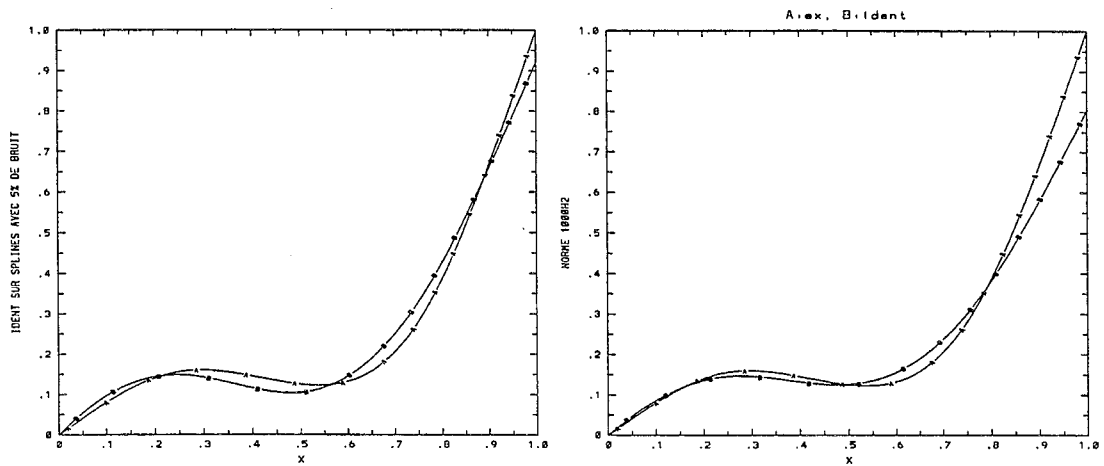


Figure 3.32 A gauche, norme $H^2_{1,1,100}$, à droite $H^2_{1,1,1000}$ pour $x(4x^2 - \frac{9}{2}x + \frac{3}{2})$

convergence de l'algorithme pour une norme H^1 . La fonctionnelle décroît ici de $7 \cdot 10^{-4}$ à $8 \cdot 10^{-6}$. Plus on prend un produit scalaire régulier, et plus les temps de calcul diminuent: 60 secondes pour L^2 , 35 secondes pour H^1 , 24 secondes pour H^2 et 19 secondes pour la norme équivalente à H^2 .

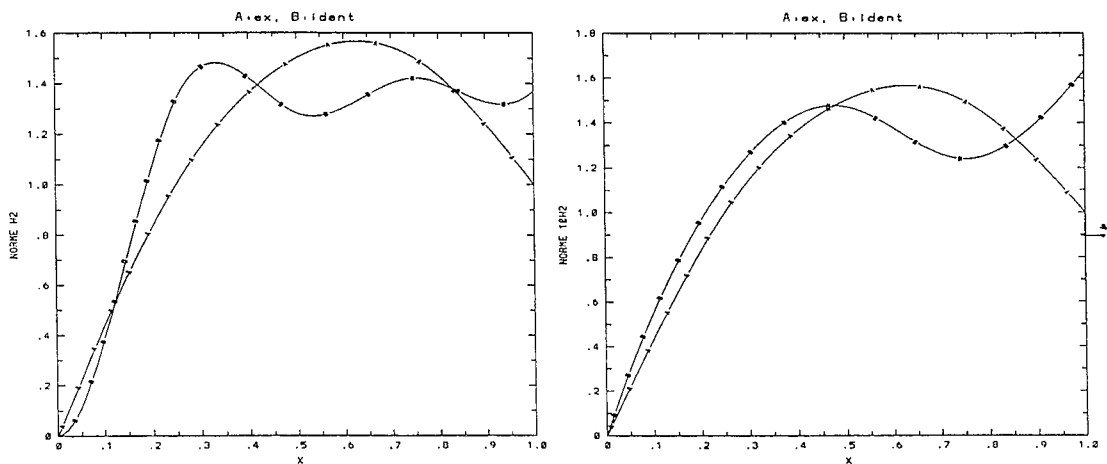


Figure 3.33 A gauche, norme H^2 , à droite $H^2_{1,1,10}$ pour $-4x^2 + 5x$

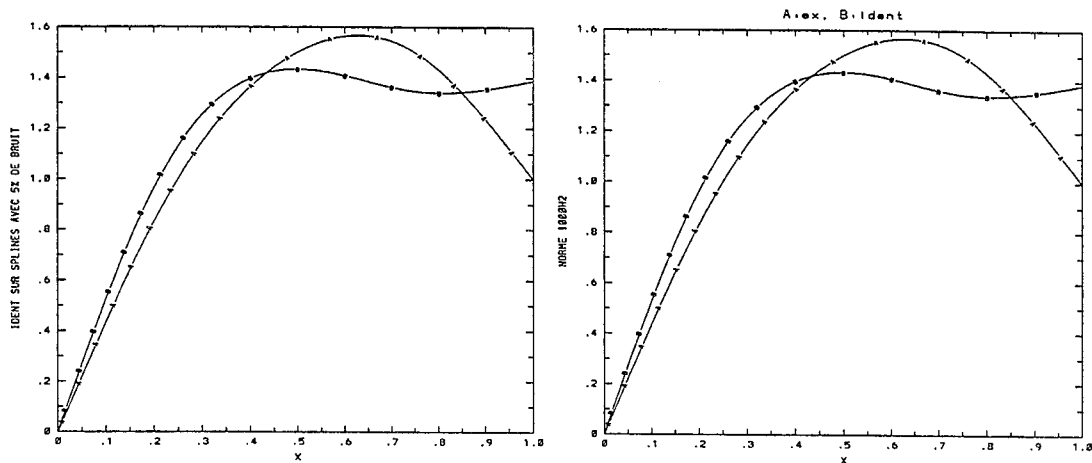


Figure 3.34 A gauche, norme $H_{1,1,100}^2$, à droite $H_{1,1,1000}^2$ pour $-4x^2 + 5x$

Pour finir, nous donnons en figure 3.33 et figure 3.34 des résultats pour l'identification de la source $-4x^2 + 5x$. Nous pouvons remarquer que la fonction identifiée n'est pas très satisfaisante, ceci même si l'on augmente le coefficient poids devant la semi-norme H^2 . Dans le cas des ondelettes, nous avons vu (figure 3.18) que lorsqu'on prenait ce coefficient assez grand, on pouvait obtenir de meilleurs résultats. Le problème se situe surtout autour de la valeur $x = 1$. Or, cette valeur, d'un point de vue physique, correspond à $\bar{y} = 1$, c'est à dire au point maximum de l'état y . Dans le contexte de l'identification de l'intensité de courant dans un Tokamak (cf [BLO 90] par exemple), il est bien connu qu'au voisinage de l'axe magnétique (qui correspond ici au point où y atteint son maximum), et pour une intensité dite à *profil creux*, l'identification est très délicate à réaliser. Malheureusement, on ne sait pas expliquer ce phénomène d'un point de vue strictement mathématique.

En conclusion, il semble que l'identification dans une base d'ondelettes soit plus sensible à la norme choisie pour le changement de base que dans le cas des splines. Ceci s'explique par le nombre de paramètres de contrôle, nettement plus élevé dans le cas des ondelettes que dans celui des splines, ce qui induit un moins bon conditionnement du problème de minimisation.

3.5 Changement de base pour la diffusion de la chaleur

Nous nous intéressons tout d'abord à l'effet du changement de base sur la méthode séquentielle quadratique. Pour cela nous représentons le spectre de la matrice $T'(v)^t T'(v)$ au voisinage du point v identifié pour les normes L^2 et $H_{1,1,100}^2$ utilisées dans l'espace d'approximation V_3 . Ici, $T(v)$ est l'application qui à un contrôle v associe l'état aux points observés.

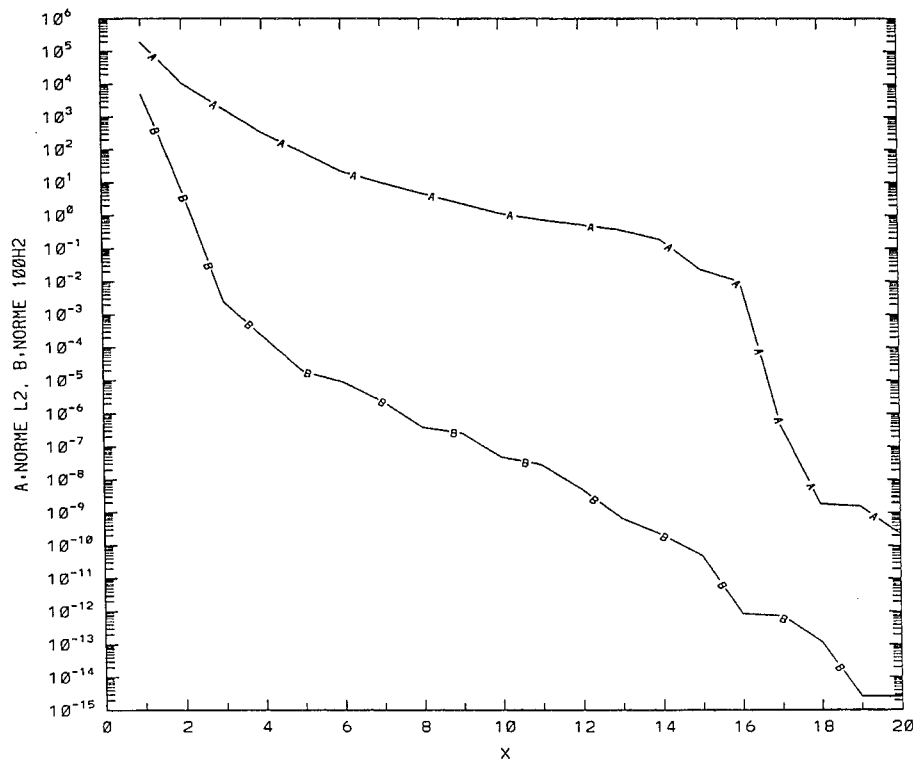


Figure 3.35 Spectres de $T'(v)^t T'(v)$ en norme L^2 (courbe A) et $H_{1,1,100}^2$ (B)

On constate donc que la matrice en norme L^2 est un peu mieux conditionnée qu'en norme $H_{1,1,100}^2$. Nous traçons ci-dessous les six premiers vecteurs propres

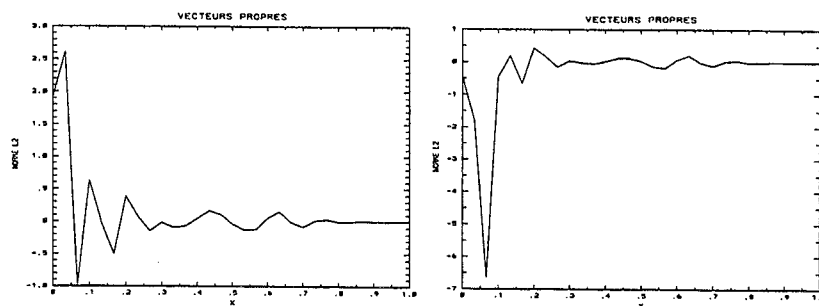


Figure 3.36 Vecteurs propres no 1 et 2 de $T'(v)^t T'(v)$ en norme L^2

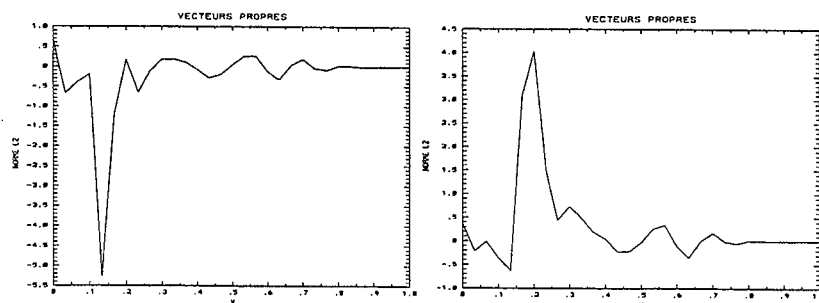


Figure 3.37 Vecteurs propres no 3 et 4 de $T'(v)^t T'(v)$ en norme L^2

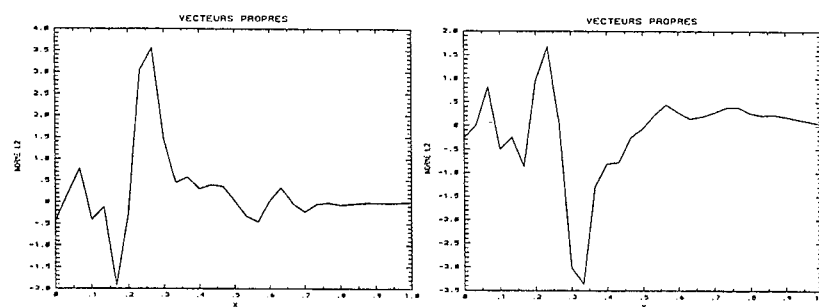


Figure 3.38 Vecteurs propres no 5 et 6 de $T'(v)^t T'(v)$ en norme L^2

(i.e. ceux qui correspondent aux valeurs propres les plus grandes) de $T'(v)^t T'(v)$ pour ces mêmes normes L^2 et $H_{1,1,100}^2$, dans la nouvelle base $\{\tilde{\phi}_1, \dots, \tilde{\phi}_m\}$.

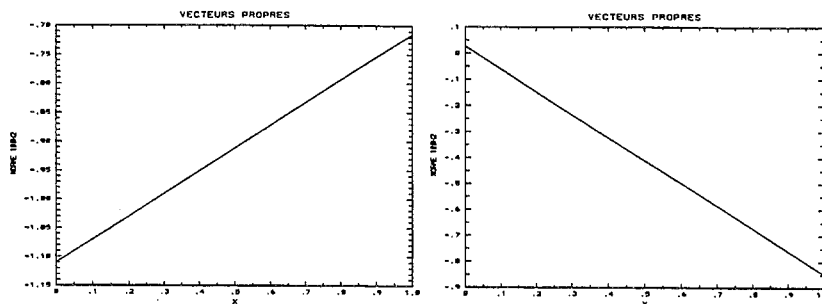


Figure 3.39 Vecteurs propres no 1 et 2 de $T'(v)^t T'(v)$ en norme $H^2_{1,1,100}$

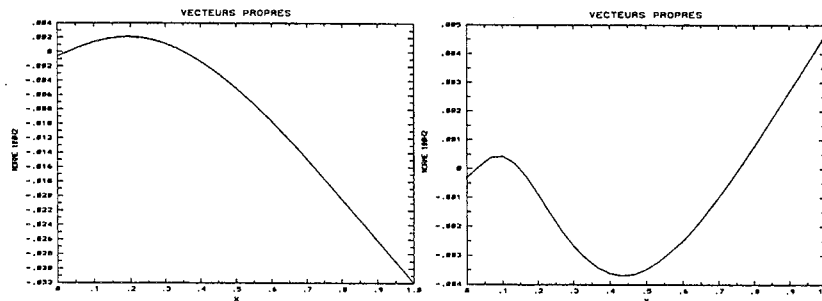


Figure 3.40 Vecteurs propres no 3 et 4 de $T'(v)^t T'(v)$ en norme $H^2_{1,1,100}$

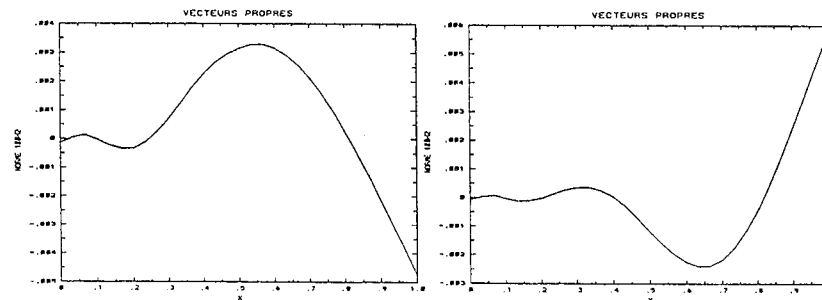


Figure 3.41 Vecteurs propres no 5 et 6 de $T'(v)^t T'(v)$ en norme $H^2_{1,1,100}$

Nous constatons donc que les vecteurs propres de $T'(v)^t T'(v)$ sont nettement plus réguliers par un changement de base par rapport à la norme de $H_{1,1,100}^2$ qu'avec une norme L^2 . Ceci explique donc pourquoi il va être plus intéressant d'identifier par un changement de base dans des espaces de norme équivalente à celle de H^2 que dans L^2 .

Dans cette section, nous avons identifié le coefficient de diffusion de la chaleur par décomposition sur une base d'ondelettes, en utilisant toujours la méthode de changement de base décrite auparavant. Nous présentons tout d'abord quelques résultats d'identification de la diffusion $\rho(x) = (x - 0.3)(x - 0.5) + 1$ pour des données bruitées à 2%.

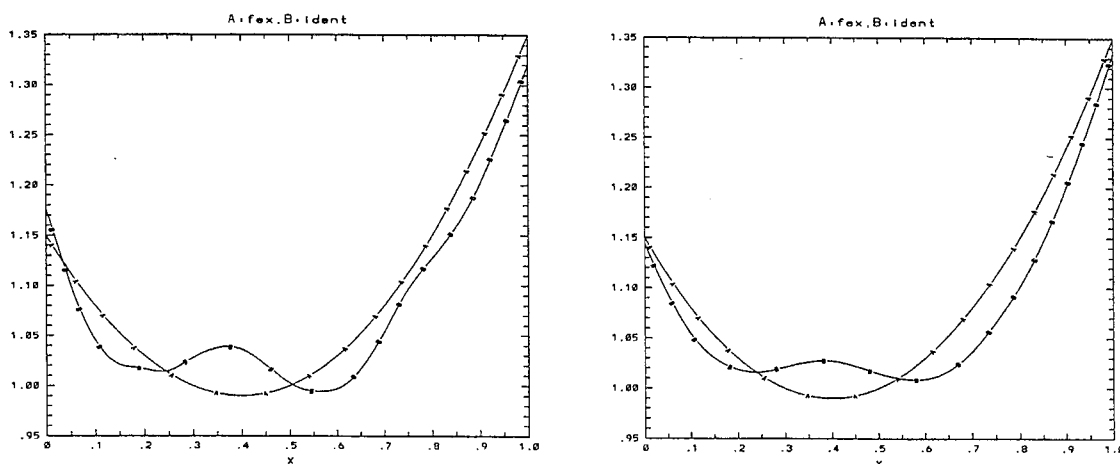


Figure 3.42 A gauche, norme $H_{1,1,10}^2$, à droite $H_{1,1,100}^2$ pour $(x - 0.3)(x - 0.5) + 1$

Les résultats sont présentés dans le sens d'une norme équivalente à H^2 de plus en plus "régulière" (i.e. le coefficient devant les termes en dérivées secondes augmente).

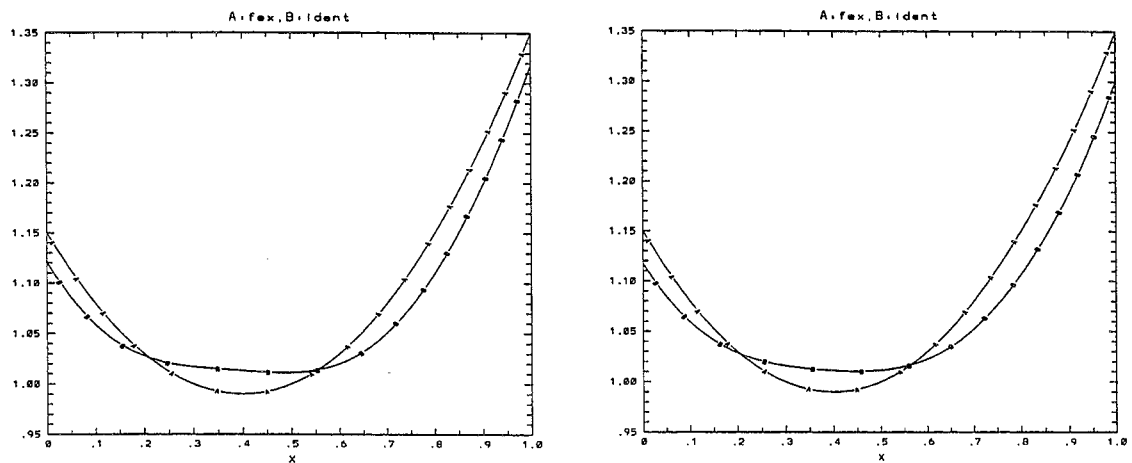


Figure 3.43 A gauche, norme $H^2_{1,1,10^3}$, à droite $H^2_{1,1,10^4}$ pour $(x - 0.3)(x - 0.5) + 1$

Dans les quatre cas présentés, figure 3.42 et figure 3.43, la fonctionnelle décroît de $5.2 \cdot 10^{-1}$ à $2.0 \cdot 10^{-2}$, comme dans le cas de l'identification de sources non linéaires, plus la norme choisie est régulière, et plus les temps de calcul diminuent : 716s CPU pour la première, 370s CPU pour la seconde, 208s CPU pour la troisième et enfin 173s CPU pour la dernière. Une estimation correcte de la diffusion nécessite donc une norme assez régulière.

On remarque en figure 3.43 que les normes choisies dans ce cas ont un coefficient très grand devant le terme de la semi-norme H^2 : 10^3 ou 10^4 . Ce fait, que nous n'avons pas observé dans le problème de l'identification de sources non linéaires tend à montrer que pour une telle diffusion ($\rho(x) = (x - 0.3)(x - 0.5) + 1$), une identification correcte ne peut être réalisée que dans un espace de fonctions très régulières. On pourrait penser a priori qu'il serait plus judicieux d'utiliser directement la semi-norme H^2 , puisque les poids devant ce terme sont importants, mais ceci n'est pas réalisable numériquement : en effet, la matrice de Gram dans ce cas $G_{i,j} = \int_0^1 \phi_i''(x)\phi_j''(x)dx$ n'est plus inversible. Le fait de garder les termes en semi norme L^2 et H^1 permet de rendre la matrice G plus régulière.

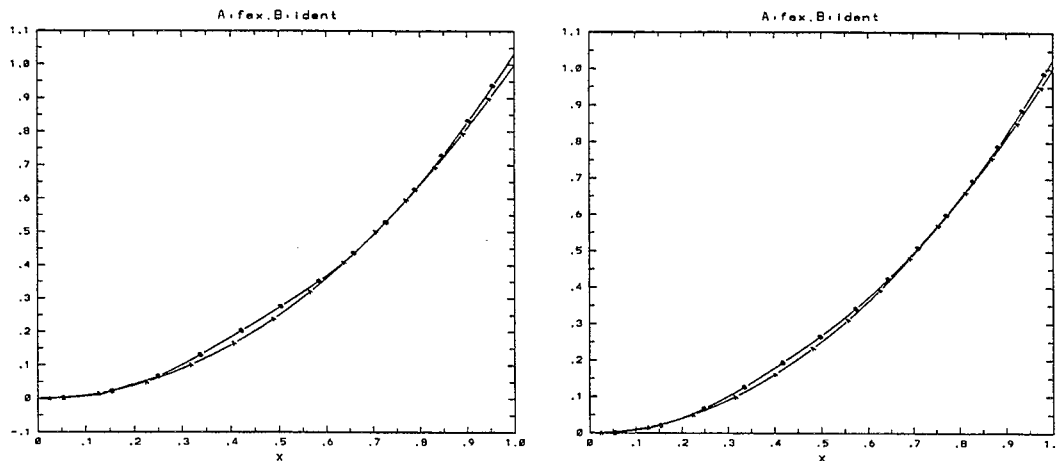


Figure 3.44 A gauche, norme $H^2_{1,1,10}$, à droite $H^2_{1,1,100}$ pour x^2

Pour la diffusion $\rho(x) = x^2$, les résultats obtenus sont représentés par la figure 3.44. Il est à noter que même avec une norme H^2 , les résultats sont satisfaisants : le seul changement intervient à nouveau dans les temps de calcul : 704s CPU à gauche (norme $\int_0^1 fg + \int_0^1 f'g' + 10 \int_0^1 f''g''$) contre 387s CPU à droite (norme $\int_0^1 fg + \int_0^1 f'g' + 100 \int_0^1 f''g''$).

Contrairement à la diffusion précédemment identifiée $((x - 0.3)(x - 0.5) + 1)$, l'identification de $\rho(x) = x^2$ nécessite le choix d'une norme similaire à celles du problème de sources non linéaires, et moins régulière que précédemment. La fonction x^2 étant plus "plate" que $(x - 0.3)(x - 0.5) + 1$, ce phénomène semble donc s'expliquer par une non-linéarité moins "importante" du problème inverse.

3.6 Conclusion

Le changement de base pour l'identification de la diffusion de la chaleur dans une base d'ondelettes donne donc d'assez bons résultats, en ce sens que l'on parvient à l'identifier d'une manière satisfaisante sans introduire de régularisation. Il est à noter que le changement de base, pour les B -splines présente un moindre intérêt ; en effet, l'estimation de la diffusion obtenue par changement de base dans une base de B -splines est très peu sensible à un changement de norme, et donne des résultats très peu satisfaisants, et qui sont similaires à ceux que l'on obtiendrait par une résolution classique à partir de données bruitées, sans régularisation. Pour ce problème, la seule façon de procéder à l'identification sans régularisation semble

donc être d'utiliser une base d'ondelettes à support compact avec une norme appropriée.

Nous avons vu comment, dans le problème d'identification de sources non linéaires, il était possible d'appliquer le changement de base, aussi bien avec les splines comme fonctions de base, qu'avec les ondelettes. Cependant, pour ces deux problèmes, nous retiendrons plus volontiers une décomposition par ondelettes, puisqu'elle est plus sensible à la norme choisie.

Bien sûr, le problème du choix du produit scalaire à utiliser peut sembler, a priori, un peu contraignant. Néanmoins, par rapport à une identification avec régularisation, cette méthode semble plus intéressante: nous avons vu au chapitre III que le choix du paramètre de régularisation était très coûteux du point de vue temps de calculs, de ce fait le changement de base des contrôles nous permet de diviser les temps d'exécution par un facteur 10, voire plus. Cette méthode est donc légèrement moins précise que la régularisation avec un paramètre bien choisi, mais cet inconvénient est compensé par la rapidité des calculs.

Chapitre V

Mesures de non linéarité

Les techniques introduites au chapitre précédent nous permettent de réaliser l'identification de la diffusion ou de sources non linéaires sans terme de régularisation dans la fonction coût.

Dans les problèmes réels, ni la solution exacte, ni le niveau de bruit présent sur les observations dont on dispose ne sont connus. Dans ce cadre, nous essayons maintenant de quantifier la non linéarité du modèle au voisinage du point identifié par des techniques stochastiques faisant intervenir les courbures de la surface des réponses au voisinage de la solution identifiée. Ceci nous permettra de savoir si la méthode de linéarisation utilisée était raisonnable, et si l'estimateur obtenu a de bonnes propriétés asymptotiques.

1 Introduction

Dans cette partie, nous considérons toujours un problème inverse sous la forme

$$z = T(v) + \nu$$

$T(v)$ étant une application de $\mathcal{U}_{\text{ad}} \subset \mathbb{R}^m$ vers une surface \mathcal{M} de dimension m contenue dans \mathbb{R}^n .

$Z \in \mathbb{R}^n$: vecteur des réponses

$v \in \mathbb{R}^m$: vecteur des paramètres inconnus

$T(v) \in \mathbb{R}^n$: vecteur de régression

$\nu \in \mathbb{R}^n$: vecteur des erreurs ou bruit.

En général, les méthodes d'estimation du paramètre v reposent sur une approximation linéaire de la surface des réponses \mathcal{M} au voisinage du bon paramètre. Il est donc essentiel de pouvoir quantifier le degré de non linéarité de la surface des réponses. Nous utiliserons pour cela la notion de courbure d'une surface en un point développée par A. Antoniadis et al. dans [ABC 92].

Il est à noter que G. Chavent dans [CHA1 91], [CHA2 91] et [CHA3 91] a utilisé les courbures dans un problème aux moindres carrés non linéaire d'une manière différente, plus déterministe. Soit le problème $\min_{v \in C} J(v) = \|T(v) - z\|_F^2$ à résoudre sur un sous-espace convexe C d'un espace de Hilbert. G. Chavent a montré que sous certaines hypothèses sur les rayons de courbure des images par T de segments de C , on peut construire un voisinage \mathcal{V} de $T(C)$ dans lequel le problème de minimisation est *quadratiquement bien posé*, i.e.

- pour tout $z \in \mathcal{V}$, il existe une unique solution $\hat{v} \in C$,
- pour tout $z \in \mathcal{V}$, J n'admet pas de minima locaux,
- toute suite minimisante converge vers \hat{v} ,
- l'application $z \mapsto \hat{v}$ est lipschitzienne sur \mathcal{V} .

Grâce à cette définition et aux propriétés géométriques des courbures, il est possible de dégager une étude locale du problème de minimisation.

Notre approche n'est pas tout à fait la même, puisque nous visons à obtenir des informations sur l'estimateur obtenu, afin de pouvoir en dégager de bonnes propriétés asymptotiques. Nous supposons dans tout ce paragraphe que \mathcal{U}_{ad} est d'intérieur non vide, que $T(v)$ est injective et deux fois continûment différentiable en tout point intérieur à \mathcal{U}_{ad} . Nous admettrons de plus que le bruit ν suit une loi normale centrée $\mathcal{N}(0, \sigma^2)$. Disposant d'un espace d'approximation de dimension finie m de \mathcal{U}_{ad} , notre problème inverse dans cet espace se résume à un problème de dimension finie dans \mathbb{R}^m .

2 Aspect géométrique

Un segment de droite de \mathcal{U}_{ad} passant par un point $v_0 \in \mathcal{U}_{ad}$ peut être paramétré de la façon suivante :

$$v(\theta) = v_0 + \theta h, \quad \theta \in \mathbb{R}, \quad h \in \mathbb{R}^m.$$

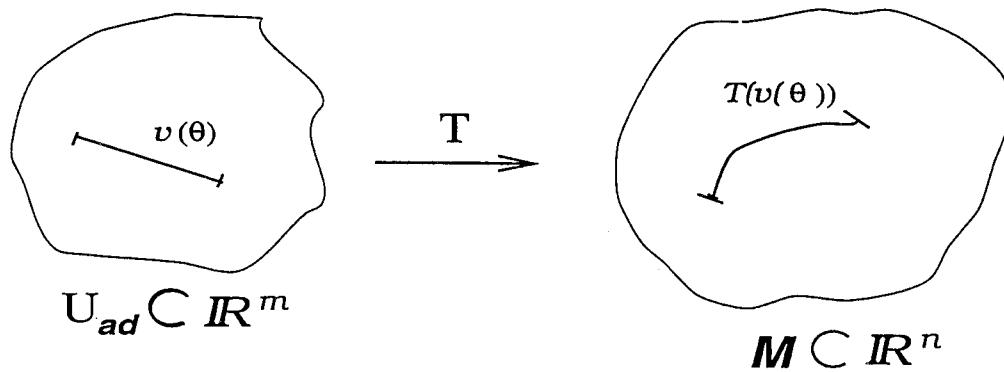


Figure 2.1 Aspect géométrique

L'image de ce segment par T définit une courbe de la surface \mathcal{M} :

$$T_h(\theta) = T(v(\theta)).$$

Pour définir une notion de courbure, nous pouvons interpréter $T_h(\theta)$ comme la position d'un mobile à l'instant θ . Les vitesses et accélération de ce mobile sont données à l'instant 0 par :

$$\dot{T}_h = \left(\frac{dT_h}{d\theta} \right)_{\theta=0} = \sum_{i=1}^m \left(\frac{\partial T}{\partial v_i} \right)_{v=v_0} \left(\frac{dv_i}{d\theta} \right)_{\theta=0} = \sum_{i=1}^m V_i h_i,$$

où $V_i = \left(\frac{\partial T}{\partial v_i} \right)_{v=v_0}$ pour $i = 1, \dots, m$ et

$$\ddot{T}_h = \left(\frac{d^2 T_h}{d\theta^2} \right)_{\theta=0} = \sum_{i=1}^m \sum_{j=1}^m \left(\frac{\partial^2 T}{\partial v_i \partial v_j} \right)_{v=v_0} h_i h_j.$$

Ce que l'on écrit encore :

$$\dot{T}_h = DT(v_0)h \quad \text{et} \quad \ddot{T}_h = h^t D^2 T(v_0)h.$$

Lorsque T est une submersion, la fonction $T(v)$ définit une sous variété différentiable de dimension m de \mathbb{R}^n . La courbe $T_h(v)$ est une courbe paramétrée de cette surface. Au point $v(\theta_0)$, \dot{T}_h est une combinaison linéaire des colonnes de $DT(v_0)$, il appartient donc à l'espace tangent à la surface \mathcal{M} au point $T(v_0)$. Le vecteur des accélérations peut se décomposer suivant :

$$\ddot{T}_h = \ddot{T}_h^N + \ddot{T}_h^T$$

où \ddot{T}_h^N est la composante normale à la surface et \ddot{T}_h^T la composante tangentielle. Il est bien connu que \ddot{T}_h^N ne dépend pas de la paramétrisation choisie, contrairement à \ddot{T}_h^T . Nous pouvons définir la courbure normale à la surface \mathcal{M} en v_0 dans la direction h par :

$$K_h^N = \frac{\|\ddot{T}_h^N\|}{\|\dot{T}_h\|^2}$$

et la courbure tangentielle par :

$$K_h^T = \frac{\|\ddot{T}_h^T\|}{\|\dot{T}_h\|^2}.$$

La quantité K_h^N sera appelée *courbure intrinsèque* et K_h^T *courbure paramétrique* en raison de leurs comportements respectifs vis à vis d'un changement de paramétrisation.

3 Aspect statistique

Nous rappelons tout d'abord quelques notions statistiques de base.

3.1 La méthode de vraisemblance

Soit z la réalisation d'un vecteur aléatoire Z de densité f_v connue au paramètre v près que nous voudrions estimer. Considérons la fonction de vraisemblance de Z pour le paramètre v :

$$L_z(v) = f_v(z_1)f_v(z_2) \dots f_v(z_n).$$

Le principe du maximum de vraisemblance consiste à choisir pour v la valeur qui rend les observations les plus vraisemblables, c'est à dire la valeur de v qui maximise la fonction de vraisemblance L_z . Sous certaines hypothèses que nous supposerons vérifiées ici, l'existence et l'unicité de l'estimation du maximum de vraisemblance sont assurées.

Plus concrètement, si $J(v) = \|z - T(v)\|^2$ représente l'erreur au sens des moindres carrés, et si l'on connaît la variance σ^2 du bruit, alors la fonction de vraisemblance sera donnée par :

$$L_z(v, \sigma^2) = \frac{C}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2} J(v)}$$

et les estimateurs des moindres carrés de v et σ^2 sont des estimateurs du maximum de vraisemblance.

3.2 Test du rapport des vraisemblances maximales

Supposons que nous voulions vérifier une certaine hypothèse, par exemple $(\mathcal{H}_0) : v \in V$, pour cela nous mesurons la quantité :

$$\max\{L_z(v), \quad v \in V\} = L_z(\tilde{v}).$$

Si l'hypothèse (\mathcal{H}_0) est vraie, cette expression ne devrait pas trop différer de la vraisemblance maximale :

$$\max\{L_z(v), \quad v \in V\} = L_z(\hat{v}).$$

Le test du rapport des vraisemblances maximales sera défini par la région critique :

$$D_\alpha = \left\{ z \in \mathbb{R}^n, \quad \frac{L_z(\tilde{v})}{L_z(\hat{v})} < c_\alpha \right\}.$$

Lorsque nous avons un problème de régression linéaire à résoudre, les régions de confiance sont de forme ellipsoïdale dans l'espace des paramètres et l'estimation \hat{v} du maximum de vraisemblance est le centre de cette ellipsoïde.

3.3 Régions de confiance

Les régions de confiance sont définies par l'ensemble des paramètres v qui ne sont pas rejetés, au seuil donné, comme valeurs possibles du paramètre par le test du rapport des vraisemblances maximales. Ce sont donc les régions, dans l'espace des paramètres, formées de l'ensemble des valeurs qui paraissent raisonnables au vu des observations faites.

Dans la suite, nous considérerons la linéarisation

$$(3.1) \quad T(v) = T(\hat{v}) + DT(\hat{v})(v - \hat{v}),$$

effectuée à chaque étape de l'algorithme séquentiel quadratique. Lorsqu'on a obtenu un estimateur \hat{v} par cette méthode, cela signifie que $\hat{T} = T(\hat{v}) \in \mathcal{M}$ est le pied de la perpendiculaire abaissée depuis l'observation z sur le plan tangent à \mathcal{M} . La région de confiance sera définie par :

$$D_\alpha = \{v / \| DT(\hat{v})(v - \hat{v}) \|^2 \leq m\hat{\sigma}^2 F_{m, n-m}(1 - \alpha)\},$$

en approchant localement la surface \mathcal{M} au point $T(\hat{v})$ par le plan tangent en \hat{v} , où la fonction $F_{n,p}$ est la fonction de répartition de la loi de Fisher-Snedecor à n et p degrés de liberté (voir [ABC 92] pour plus de détails), et où $\hat{\sigma}^2$ est l'estimation de σ^2 obtenue par :

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (z_i - T_i(\hat{v}))^2.$$

La linéarisation (3.1) n'est acceptable que si le terme du second ordre dans le développement de Taylor de $T(v)$ au voisinage de \hat{v} est petit devant le terme du premier ordre. Si l'on veut comparer ces deux termes, nous sommes amenés à considérer les courbures :

$$K_h^N(\hat{v}) = \frac{\| h^t D^2 T^N(\hat{v}) h \|}{\| DT(\hat{v}) h \|^2} \quad \text{et} \quad K_h^T(\hat{v}) = \frac{\| h^t D^2 T^T(\hat{v}) h \|}{\| DT(\hat{v}) h \|^2}$$

pour toute direction h .

Pour rendre ces courbures invariantes à tout changement d'échelle, on définit de nouvelles courbures :

$$\gamma_h^N = K_h^N \sqrt{m\hat{\sigma}^2} \quad \text{et} \quad \gamma_h^T = K_h^T \sqrt{m\hat{\sigma}^2}.$$

Une mesure globale de la non linéarité ne doit pas dépendre de la direction h , on pose donc :

$$\gamma^N = \max_h \gamma_h^N \quad \text{et} \quad \gamma^T = \max_h \gamma_h^T.$$

Il nous faut ensuite définir un seuil C au dessous duquel les courbures γ^N et γ^T sont suffisamment petites pour que les calculs reposant sur une approximation linéaire de la surface \mathcal{M} soient valables. Compte tenu de l'interprétation des courbures, il est plausible de comparer l'inverse de ces quantités au rayon r_α d'une sphère de centre $T(\hat{v})$, telle que : $Prob\{|T(\hat{v}) - T(v)| \leq r_\alpha\} = \alpha$. C'est pourquoi, dans [ABC 92], le seuil proposé est :

$$C = \frac{1}{2\sqrt{F_{m,n-m}^{-1}(1-\alpha)}}$$

Lorsque les courbures γ^N et γ^T obtenues sont suffisamment "petites", l'approximation linéaire faite pour réaliser l'identification sera donc justifiée. D'autre part, nous pourrions en déduire qu'autour du point $T(\hat{v})$, la surface des réponses apparaît suffisamment plane pour que si $T(\hat{v})$ est proche de z_{ex} cela implique que \hat{v} soit un "bon" estimateur asymptotique du contrôle exact v_{ex} .

4 Calcul des courbures

Afin de faciliter les calculs de γ^N et γ^T , A. Antoniadis et al. ont proposé une reparamétrisation : ayant supposé que T est une submersion, $DT(\hat{v})$ admet une décomposition QR :

$$DT(\hat{v}) = Q \begin{pmatrix} R \\ 0 \end{pmatrix}$$

où Q est une matrice orthogonale d'ordre n et R une matrice triangulaire supérieure d'ordre m . En écrivant $Q = (Q_m \quad Q_{n-m})$, nous obtenons pour tout h dans \mathbb{R}^m :

$$\dot{T}_h = DT(\hat{v})h = Q_m R h = Q_m d$$

avec $d = R h$. Introduisons une nouvelle paramétrisation $\phi = R v$, alors

$$T(v) = T(v(\phi)) = T(R^{-1}\phi) = g(\phi)$$

et

$$Dg(\hat{\phi}) = DT(\hat{v})R^{-1} = Q_m$$

$$D^2g(\hat{\phi}) = R^{-t} D^2T(\hat{v})R^{-1}.$$

Le segment $v(\theta) = \hat{v} + \theta h$ est reparamétrisé en :

$$\phi(\theta) = R(\hat{v} + \theta h) = \hat{\phi} + \theta R h = \hat{\phi} + \theta d$$

et puisque

$$\dot{g}_d = \dot{T}_h \quad \text{et} \quad \ddot{g}_d = h^t D^2T(\hat{v})h = \ddot{T}_h,$$

les courbures pour cette nouvelle paramétrisation s'écrivent :

$$\Gamma_d^T = \gamma_h^T = \sqrt{m\hat{\sigma}^2} \frac{\|\ddot{g}_d^T\|}{\|Q_m d\|^2} = \sqrt{m\hat{\sigma}^2} \frac{\|\ddot{g}_d^T\|}{\|d\|^2}$$

et

$$\Gamma_d^N = \gamma_h^N = \sqrt{m\hat{\sigma}^2} \frac{\|\ddot{g}_d^N\|}{\|Q_m d\|^2} = \sqrt{m\hat{\sigma}^2} \frac{\|\ddot{g}_d^N\|}{\|d\|^2}.$$

Nous pouvons encore simplifier en remarquant que :

$$\|\ddot{g}_d^N\| = \|Q_m^t \ddot{g}_d\| = \|d^t \ddot{A}^N d\|$$

où \ddot{A}^N est le tableau tri-dimensionnel des $n - m$ dernières faces de \ddot{A} et où \ddot{A} est le tableau tri-dimensionnel obtenu comme :

$$\ddot{A} = [Q^t][R^{-t}D^2T(\hat{v})R^{-1}]$$

appelé *tenseur d'accélération*. Nous rappelons en annexe 2 quelques règles de calcul sur les tableaux tridimensionnels.

Nous utilisons un algorithme de type itératif proposé par D.M. Bates et D.G. Watts (référence [BW 80]) pour la recherche des courbures maximales quand d varie dans la sphère unité de \mathbb{R}^m . Dans la suite, nous considérerons toujours que d est dans cette sphère :

Initialisations

$$d_0 = (0, \dots, 0, 1)^t$$

$$k = 0$$

$$\tilde{r}_0 = r_0 = 4[d_0^t \ddot{A}^N d_0][\ddot{A}^N d_0]$$

tant que $\tilde{r}_k^t d_k < 1 - \varepsilon$ **faire**

début

$$d_{k+1} = \frac{\mu r_k + d_k}{\|\mu r_k + d_k\|}$$

$$r_{k+1} = 4[d_k^t \ddot{A}^N d_k][\ddot{A}^N d_k]$$

$$\tilde{r}_{k+1} = \frac{r_{k+1}}{\|r_{k+1}\|}$$

$$k = k + 1$$

fin

$$\gamma^N = \|\ddot{A}^N d_k\|$$

où μ est un réel fixé dans $[0, +\infty]$. L'algorithme de calcul de γ^T est analogue. La quantité $4[d^t \ddot{A}^N d][\ddot{A}^N d]$ intervenant dans l'algorithme n'est rien d'autre que le gradient de l'application: $d \mapsto (\Gamma_d^N)^2$. Pour calculer le maximum des courbures Γ_d^N , pour $d \in \mathbb{R}^m$, on se déplace donc en chaque point courant dans la direction du gradient, et on normalise, afin de rester dans la sphère unité. \rightarrow

Le coefficient μ détermine l'importance du déplacement selon la direction du gradient. Dans la pratique, nous avons constaté que la convergence de l'algorithme peut dépendre de la valeur de μ .

Une fois ce calcul fait, on vérifie que ces courbures sont "assez petites". Si ce n'est pas le cas, on change de paramétrisation, jusqu'à ce que l'on trouve des courbures satisfaisantes.

5 Quelques résultats sur les courbures

Nous présentons ici les courbures intrinsèque γ^N et paramétrique γ^T calculées par l'algorithme ci-dessus, dans le cas du problème d'identification du coefficient de diffusion de la chaleur $(x-0.3)(x-0.5)+1$ dans une base d'ondelettes à support compact de I. Daubechies.

Norme employée	γ^N	γ^T
L^2	$3,7 \cdot 10^{-2}$	$5,8 \cdot 10^{-1}$
H^1	$7,8 \cdot 10^{-3}$	$7,1 \cdot 10^{-2}$
H^2	$6,5 \cdot 10^{-3}$	$5,8 \cdot 10^{-2}$
$H_{1,1,10}^2$	$6,3 \cdot 10^{-3}$	$5,7 \cdot 10^{-2}$
$H_{1,1,100}^2$	$6,1 \cdot 10^{-3}$	$5,4 \cdot 10^{-2}$

Tableau 1 : Courbures pour $\rho(x) = (x - 0.3)(x - 0.5) + 1$.

Ce tableau est cohérent avec les résultats obtenus sur la diffusion identifiée, puisqu'avec une norme L^2 , où les courbures sont les plus importantes, nous n'avons jamais réussi à obtenir une diffusion correcte. Dans le cas de normes H^2 ou équivalentes, la fonction identifiée était satisfaisante, ce que l'on retrouve ici, avec des valeurs correctes des courbures.

Nous donnons dans le tableau suivant les courbures obtenues lors de l'identification de la diffusion $\rho(x) = x^2$ sur une base d'ondelettes en fonction de la norme utilisée.

Norme employée	γ^N	γ^T
L^2	$9,7 \cdot 10^{10}$	$3,9 \cdot 10^{11}$
H^1	$1,7 \cdot 10^3$	$8,3 \cdot 10^3$
H^2	$6,2 \cdot 10^0$	$2,6 \cdot 10^1$
$H_{1,1,10}^2$	$2,2 \cdot 10^0$	$9,4 \cdot 10^0$
$H_{1,1,100}^2$	$1,7 \cdot 10^0$	$7,2 \cdot 10^0$

Tableau 2 : Courbures pour $\rho(x) = x^2$.

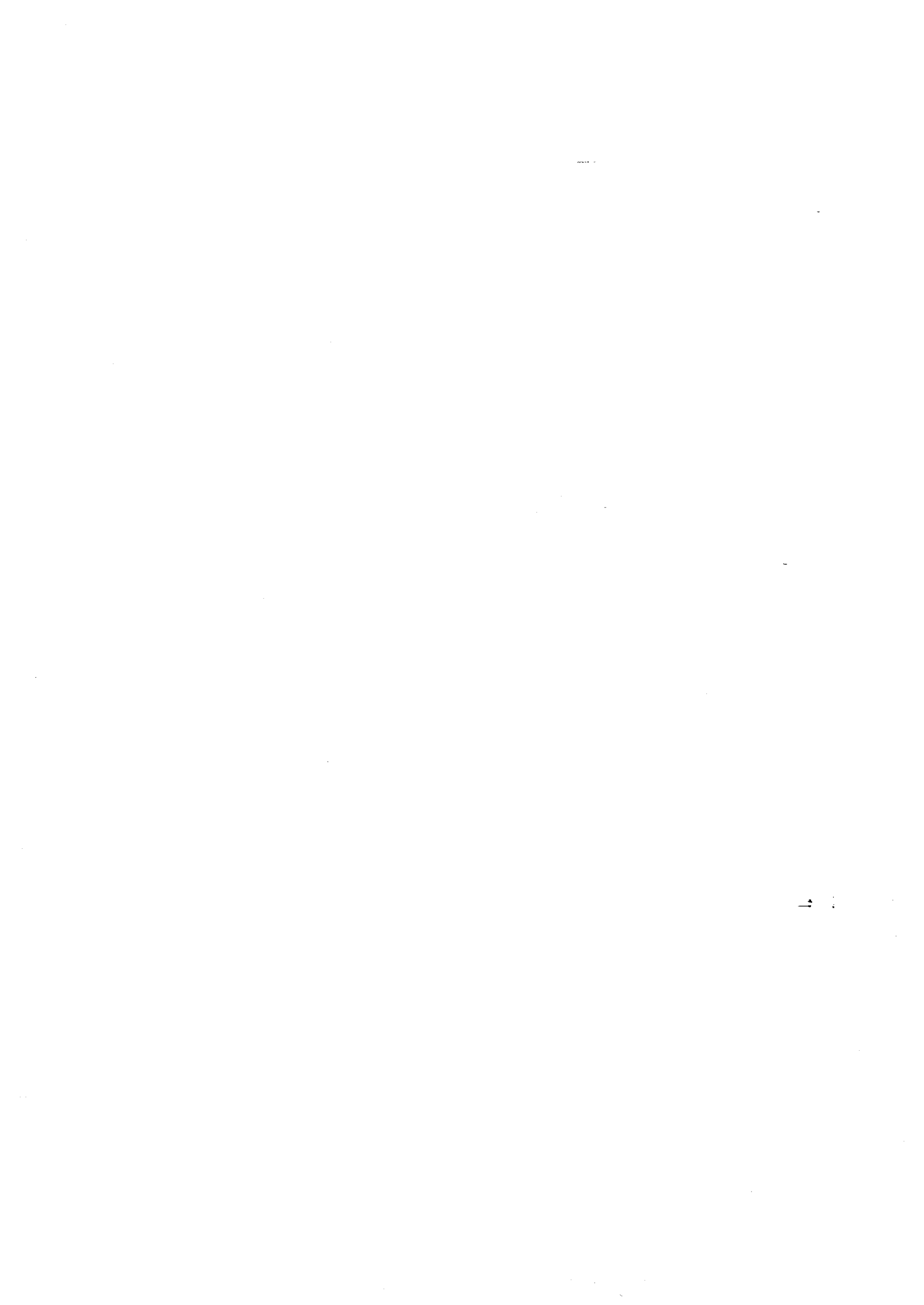
Nous constatons que les courbures sont les plus petites lorsque la norme employée est plus régulière. Par contre, les valeurs obtenues sont beaucoup plus importantes que dans le cas de la diffusion $\rho(x) = (x - 0.3)(x - 0.5) + 1$, ce qui peut s'expliquer par le fait que la fonction x^2 s'annule sur l'intervalle $[0, 1]$, contrairement à la fonction précédente. Pour illustrer ce fait, nous présentons les courbures L^2 obtenues avec la diffusion $\rho(x) = x^2 + C$ lorsqu'on fait tendre la constante C vers 0 :

Constante C	γ^N	γ^T
1	$1, 8 \cdot 10^{-2}$	$2, 4 \cdot 10^{-1}$
10^{-1}	$3, 6 \cdot 10^1$	$3, 7 \cdot 10^2$
10^{-2}	$1, 2 \cdot 10^{10}$	$6, 9 \cdot 10^{10}$
0	$9, 7 \cdot 10^{10}$	$3, 9 \cdot 10^{11}$

Tableau 3 : Courbures L^2 pour $\rho(x) = x^2 + C$.

Dans ce dernier cas, l'augmentation des courbures traduit le fait que le problème est de moins en moins bien posé : intuitivement, lorsqu'on veut identifier la diffusion, le problème se comporte plus ou moins selon $\rho(x)^{-1}$.

En résumé les courbures nous apportent des informations essentielles relatives au problème étudié : lorsque ces courbures sont inférieures au seuil proposé, cela signifie qu'au voisinage de la diffusion estimée, la surface des réponses est relativement plane. Donc une petite variation sur les observations n'entraîne pas une variation très forte sur les contrôles. Par contre, en présence de courbures importantes, on peut se poser les questions suivantes : l'approximation linéaire n'était peut-être pas justifiée au voisinage de l'estimation, ou alors l'estimateur obtenu n'est pas un bon estimateur de la solution exacte.



Conclusion



Le but de cette étude était d'envisager les différents moyens de résolution de problèmes inverses non linéaires. Ces problèmes s'apparentent à des problèmes de type minimisation par le biais d'une fonction coût représentant l'erreur aux moindres carrés du modèle par rapport aux observations. L'approche contrôle optimal nous a permis un calcul exact du gradient du critère à minimiser, rendant ainsi possible l'utilisation d'algorithmes séquentiels quadratiques, dont chaque itération correspond à la minimisation d'une nouvelle fonctionnelle quadratique par une méthode de gradient conjugué.

La mise en œuvre de ces méthodes a été réalisée sur deux types de problèmes inverses : identification du coefficient de diffusion de la chaleur, et identification de sources non linéaires.

En ce qui concerne ce dernier problème, les récentes investigations de M. Vogelius semblent prometteuses, puisqu'il a partiellement résolu le problème de l'identifiabilité lorsque la source est affine, dans un domaine de \mathbb{R}^2 délimité par une ellipse non réduite à un cercle.

Les résultats que nous avons obtenus sur ce problème ont été réalisés dans un domaine quelconque de \mathbb{R}^2 grâce au calcul de la transformée conforme de Riemann.

Les problèmes considérés étant mal posés, nous nous sommes penchés sur le choix du paramètre de régularisation. Les techniques de validation croisée ordinaire et généralisée nous ont permis de déterminer de manière satisfaisante la valeur idéale de ce paramètre.

Dans l'application de la validation croisée généralisée au problème de sources non linéaires, une difficulté a consisté à calculer les dérivées première et seconde du modèle. Pour ce faire, une analyse au second ordre du modèle a été employée, réduisant considérablement les temps de calcul par rapport à un calcul explicite de toutes ces dérivées. Les estimateurs de Monte Carlo introduits par D. Girard ont été utilisés afin de rendre le calcul de la trace de la matrice de lissage le moins coûteux possible.

La validation croisée ordinaire, bien que très simple à mettre en œuvre, est cependant très coûteuse du point de vue temps de calcul, puisqu'elle nécessite au moins autant de fois la résolution du problème d'optimisation qu'il y a d'observations.

Les récents développements de la théorie des ondelettes nous ont conduits à réaliser l'identification de ces fonctions (diffusion, sources non linéaires) par décomposition sur une base d'ondelettes, afin de ne plus introduire de régularisation. En effet, en analyse d'images, l'utilisation des ondelettes offre de nombreuses possibilités : compression des données, débruitage, seuillage, ... Nous avons donc identifié le contrôle au moyen de ses coefficients dans un espace d'approximation fixé. Le seuillage a été expérimenté sur le problème de sources non linéaires, sans apporter les résultats escomptés.

Nous avons alors abordé le problème sous un autre angle, toujours de façon à supprimer le terme de régularisation. Introduisant la matrice des produits scalaires des fonctions de base dans un espace de Hilbert, nous faisons un changement de base par rapport à ses vecteurs propres. Ceci nous permet alors de traiter le problème de contrôle comme un problème de minimisation classique dans \mathbb{R}^m , les produits scalaires de fonctions dans cette base s'exprimant comme les produits scalaires euclidiens de leurs composantes.

D'autre part, si l'espace de Hilbert choisi correspond à des fonctions très régulières (H^2 par exemple), l'identification est réalisée beaucoup plus rapidement et de manière plus satisfaisante. Dans le meilleur des cas, nous avons pu ainsi voir des temps de calcul divisés par un facteur de dix, par rapport à l'identification réalisée avec une régularisation.

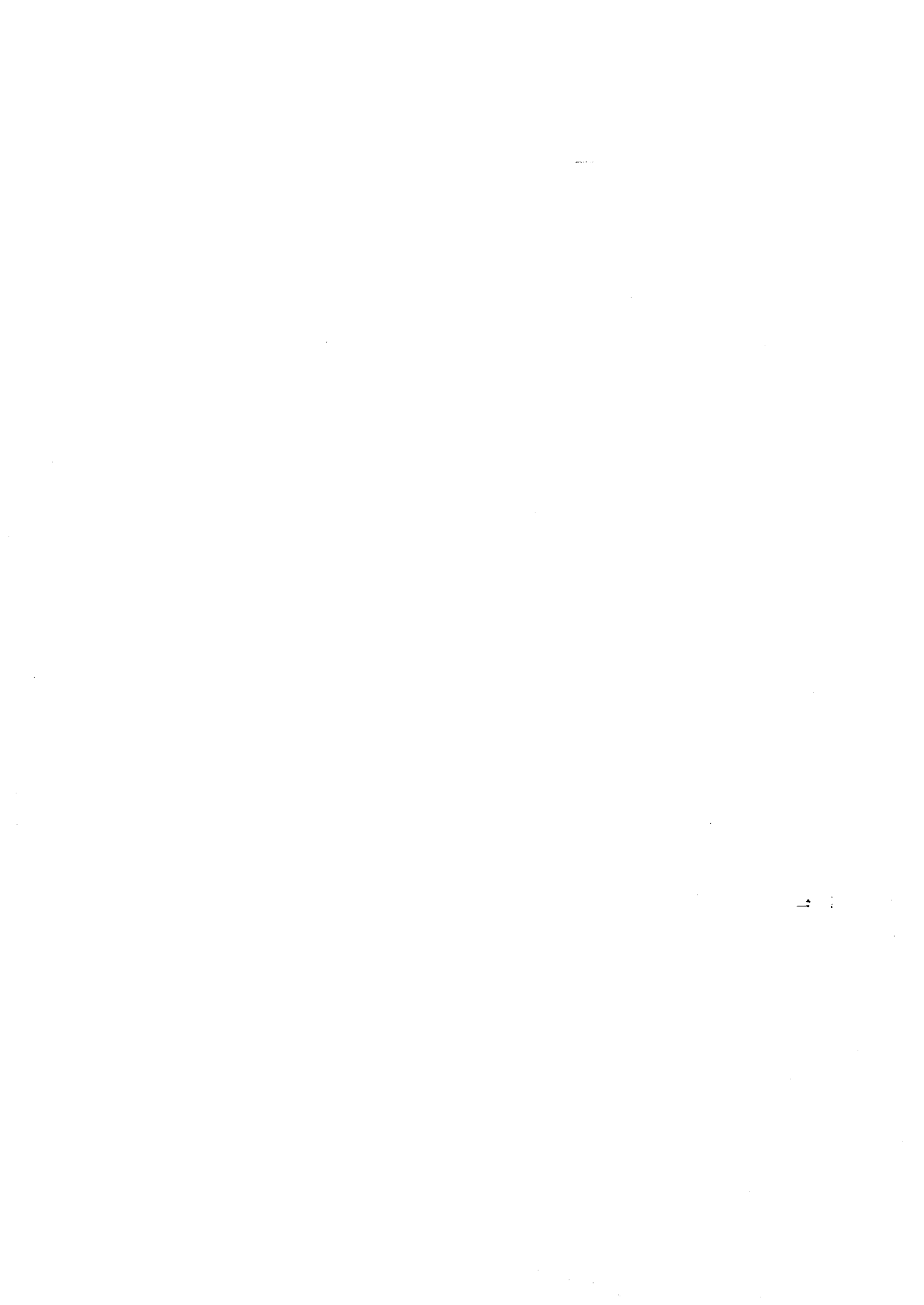
Ces résultats nous semblent donc très encourageants, puisqu'ils pourraient permettre d'éviter à l'avenir l'utilisation de régularisation ; nous avons constaté que pour cette dernière, la détermination du "bon" paramètre pouvait être assez coûteuse.

Disposant d'un estimateur, ou d'un contrôle identifié, et ne connaissant a priori pas la solution exacte du problème inverse, il nous a paru utile de définir un critère qui nous permette de savoir si l'estimation obtenue est correcte. Nous avons vu que grâce aux courbures de la surface des réponses, il était possible d'envisager des mesures de non linéarité. Nous avons également constaté que plus la norme intervenant dans le changement de base était régulière, plus les courbures diminuaient, ce qui conforte le fait que l'identification est d'autant mieux réalisable que cette norme est régulière.

Ces courbures sont donc très intéressantes pour les problèmes réels, où l'on ne connaît ni la solution exacte, ni le niveau de bruit sur les observations.

Une perspective pourrait être d'essayer de lier les performances de l'algorithme aux valeurs des courbures calculées. D'autre part, les résultats obtenus dans le cas de l'identification de sources non linéaires laissent présager que l'identifiabilité doit être possible lorsque le domaine considéré est distinct d'une boule de \mathbb{R}^2 et que la fonction à identifier est assez régulière. →

Annexes



Annexe 1

Coefficients des filtres associés aux ondelettes à support compact de I. Daubechies

Nous donnons ici la liste des coefficients des filtres h_N , $N = 2, \dots, 7, 10$ intervenant dans la construction des ondelettes à support compact de I. Daubechies. N représente le nombre de moments nuls de l'ondelette ψ_N associée au filtre h_N , tandis que $2N$ est le nombre total de coefficients non nuls de h_N . Nous représentons également les fonctions ϕ_N et ψ_N pour $N = 2, \dots, 7, 10$.

	n	$h_N(n)$
$N = 2$	0	0.482962913145
	1	0.836516303738
	2	0.224143868042
	3	-0.129409522551

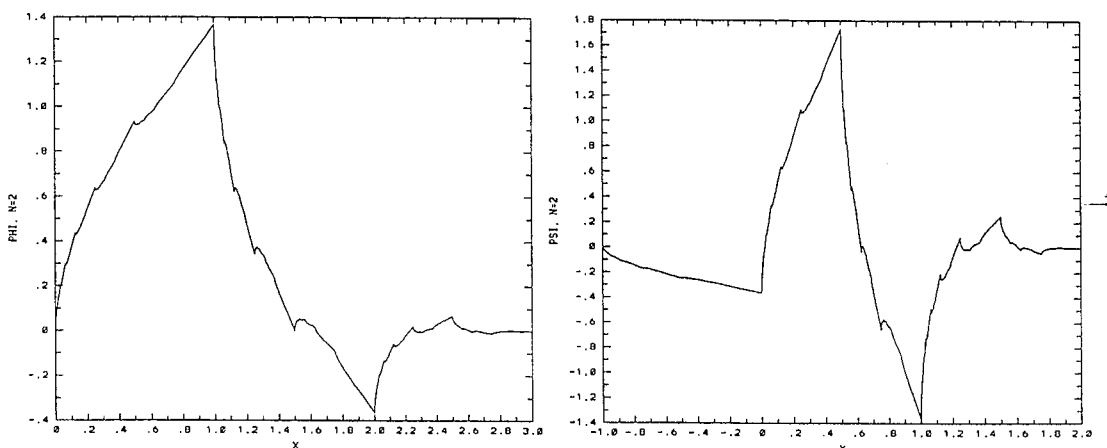


Figure 1 Fonction d'échelle ϕ_2 et ondelette ψ_2 avec 2 moments nuls.

	n	$h_N(n)$
$N = 3$	0	0.332670552950
	1	0.806891509311
	2	0.459877502118
	3	-0.135011020010
	4	-0.085441273882
	5	0.035226291882

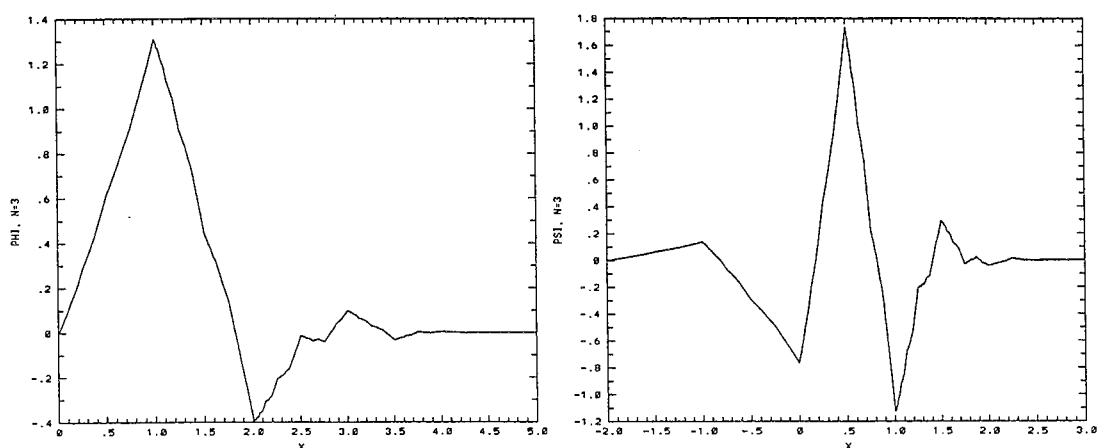


Figure 2 Fonction d'échelle ϕ_3 et ondelette ψ_3 avec 3 moments nuls.

	n	$h_N(n)$	n	$h_N(n)$
$N = 4$	0	0.230377813309	4	-0.187034811719
	1	0.714846570553	5	0.030841381836
	2	0.630880767930	6	0.032883011667
	3	-0.027983769417	7	-0.010597401785
	n	$h_N(n)$	n	$h_N(n)$
$N = 5$	0	0.160102397974	5	-0.032244869585
	1	0.603829269797	6	0.077571493840
	2	0.724308528438	7	-0.006241490213
	3	0.138428145901	8	-0.012580751999
	4	-0.242294887066	9	0.003335725285

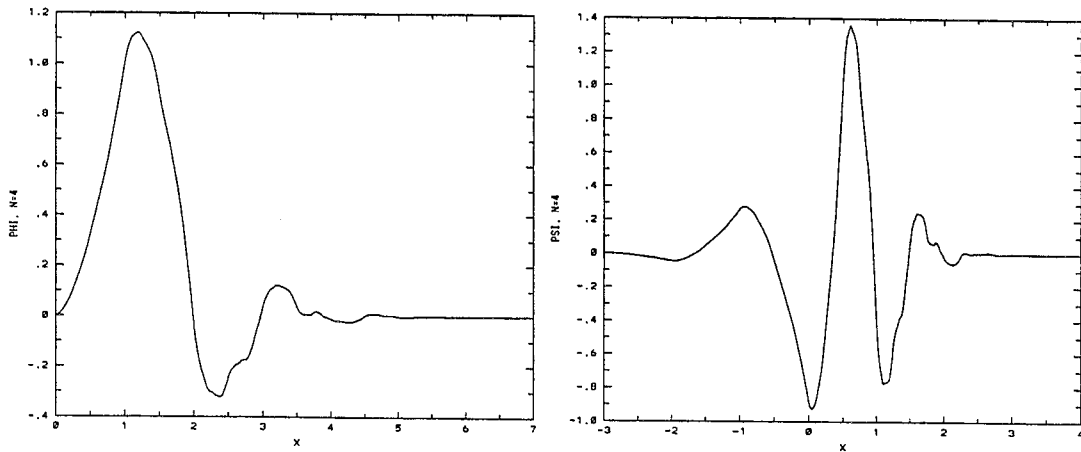


Figure 3 Fonction d'échelle ϕ_4 et ondelette ψ_4 avec 4 moments nuls.

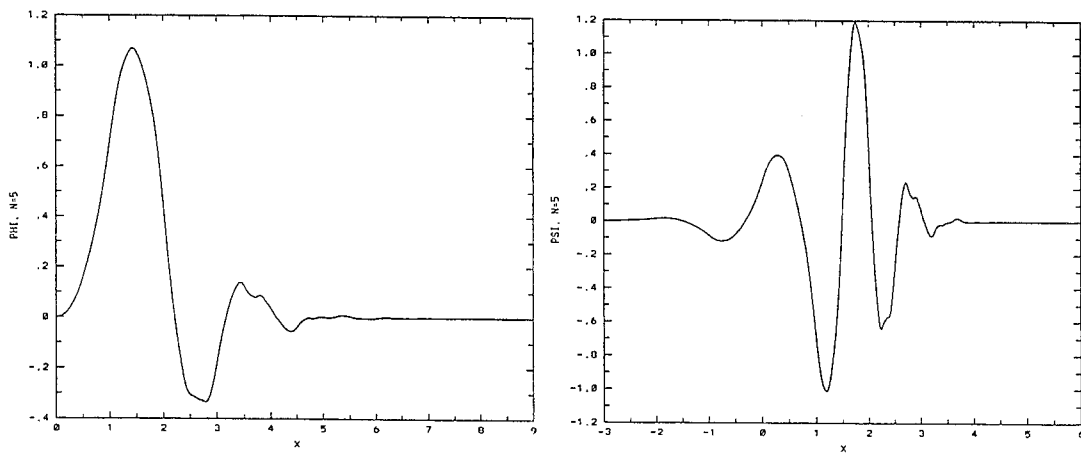


Figure 4 Fonction d'échelle ϕ_5 et ondelette ψ_5 avec 5 moments nuls.

	n	$h_N(n)$	n	$h_N(n)$
$N = 6$	0	0.111540743350	6	0.097501605587
	1	0.494623890398	7	0.027522865530
	2	0.751133908021	8	-0.031582039318
	3	0.315250351709	9	0.000553842201
	4	-0.226264693965	10	0.004777257511
	5	-0.129766867567	11	-0.001077301085

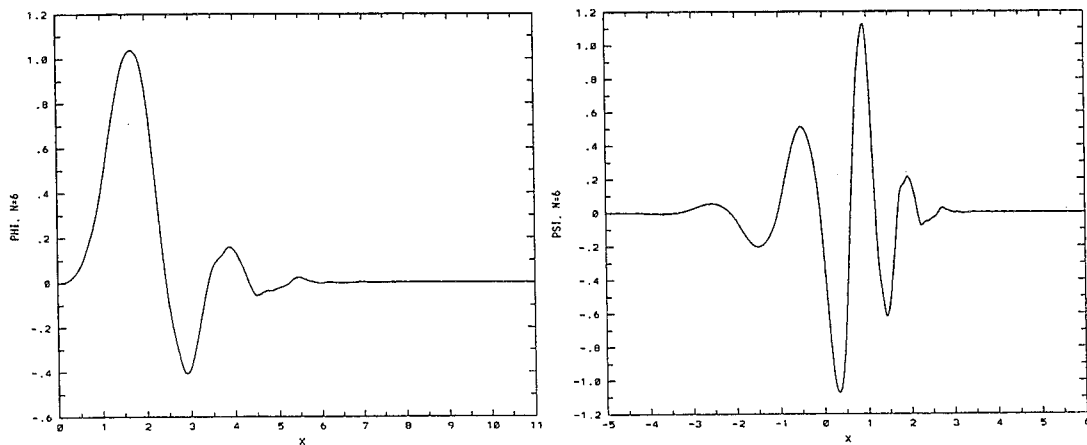


Figure 5 Fonction d'échelle ϕ_6 et ondelette ψ_6 avec 6 moments nuls.

	n	$h_N(n)$	n	$h_N(n)$
$N = 7$	0	0.077852054085	7	0.080612609151
	1	0.396539319482	8	-0.038029936935
	2	0.729132090846	9	-0.016574541631
	3	0.469782287405	10	0.012550998556
	4	-0.143906003929	11	0.000429577973
	5	-0.224036184994	12	-0.001801640704
	6	0.071309219267	13	0.000353713800

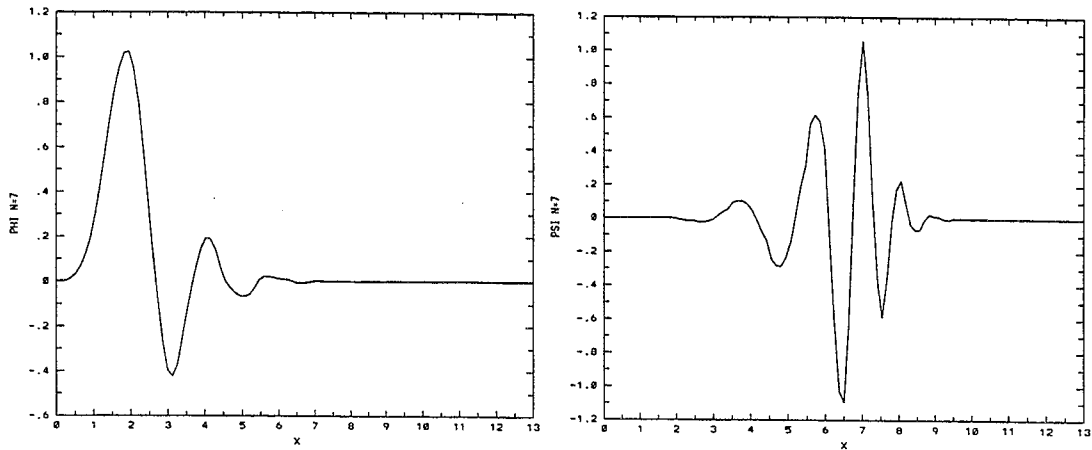


Figure 6 Fonction d'échelle ϕ_7 et ondelette ψ_7 avec 7 moments nuls.

	n	$h_N(n)$	n	$h_N(n)$
$N = 10$	0	0.026670057901	10	-0.029457536822
	1	0.188176800078	11	0.033212674059
	2	0.527201188932	12	0.003606553567
	3	0.688459039454	13	-0.010733175483
	4	0.281172343661	14	0.001395351747
	5	-0.249846424327	15	0.001992405295
	6	-0.195946274377	16	-0.000685856695
	7	0.127369340336	17	-0.000116466855
	8	0.093057364604	18	0.000093588670
	9	-0.071394147166	19	-0.000013264203

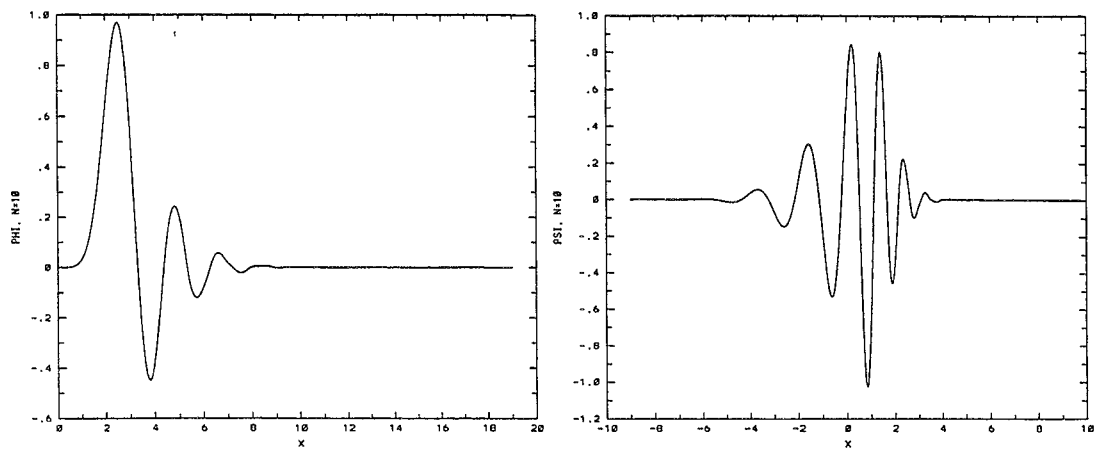


Figure 7 Fonction d'échelle ϕ_{10} et ondelette ψ_{10} avec 10 moments nuls.

Annexe 2

Règles de multiplication de tableaux tridimensionnels

Soit $A = (a_{ij}^k)$ pour $k = 1, \dots, n$, $i = 1, \dots, p$ et $j = 1, \dots, q$ un tableau tridimensionnel d'ordre $n \times p \times q$, obtenu en juxtaposant n matrices A^k d'ordre $p \times q$ appelées faces. Soient X une matrice d'ordre $r \times p$ et Y une matrice d'ordre $q \times s$. Par définition, le produit XAY est le tableau tridimensionnel B d'ordre $n \times r \times s$ dont la k -ème face B^k est le produit XA^kY :

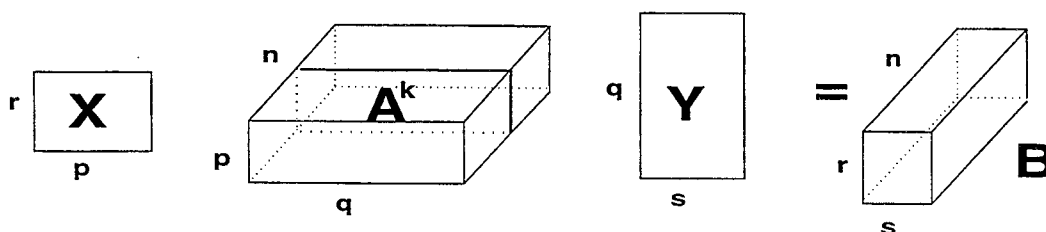


Figure 8 Règle de multiplication d'un tableau 3D par 2 matrices

Si on interprète le tableau tridimensionnel A comme une matrice d'ordre $p \times q$ ayant pour éléments les vecteurs A_{ij} de \mathbb{R}^n , on peut définir le produit d'une matrice D d'ordre $r \times n$ par A comme étant la matrice d'ordre $p \times q$ dont les éléments sont les vecteurs DA_{ij} de \mathbb{R}^r . Ce nouveau tableau tridimensionnel d'ordre $r \times p \times q$ sera noté $[D][A]$:

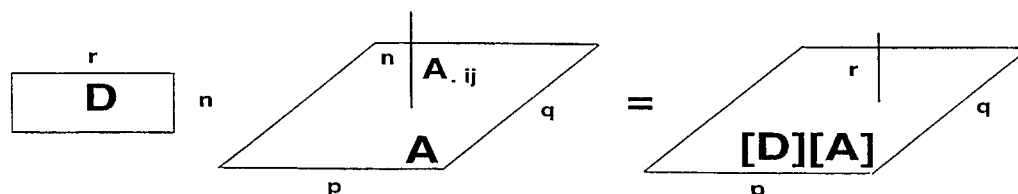


Figure 9 Multiplication matrice \times tableau 3D

On peut facilement prouver la relation suivante :

$$[DX][A] = [D][[X][A]]$$

où D est une matrice $q \times n$, X une matrice $n \times n$ et A un tableau tridimensionnel d'ordre $n \times p \times p$.

Bibliographie



- [ABC 92] A. ANTONIADIS, J. BERRUYER, R. CARMONA
Régression non linéaire et applications
Coll. Esa, ed. Economica, Paris 1992
- [ADA 75] R. A. ADAMS
Sobolev Spaces
Academic Press, New York, (1975)
- [ANT 93] A. ANTONIADIS
Smoothing Noisy Data with Coiflets
Rapport Technique, RT 93, Université de Grenoble 1, janvier 1993
- [BCK 83] H. T. BANKS, J. M. CROWLEY, K. KUNISCH
Cubic Spline Approximation Techniques for Parameter Estimation in Distributed Systems
IEEE Transactions on Automatic Control, Vol AC-28, No 7, July 1983
- [BDP 85] M. BERTERO, C. DE MOL, E. R. PIKE
Linear inverse problems with discrete data. I: General formulation and singular system analysis
Inverse problems, I (1985), pp 301-330
- [BLO 90] J. BLUM, E. LAZZARO, J. O'ROURKE, B. KEEGAN, Y. STEFAN
Problems and methods of self-consistent reconstruction of Tokamak equilibrium profiles from magnetic and polarimetric measurements
Nuclear fusion, Vol 30, No 8 (1990)
- [BLU 89] J. BLUM
Numerical Simulation and Optimal Control in Plasma Physics
Wiley/Gauthier-Villars (1989)
- [BRE 87] H. BREZIS
Analyse fonctionnelle, Théorie et Applications
Masson, 1987
- [BFV 92] E. BERETTA, E. FISCHER, M. VOGELIUS
An inverse problem originating from magnetohydrodynamics. Some numerical experiments.
Ill-posed Problems in Natural Sciences, ed. A. N. Tikhonov VSP, Utrecht (1992), pp. 254-269
- [BV 91] E. BERETTA, M. VOGELIUS
An Inverse Problem Originating from Magnetohydrodynamics
Arch. Rat. Mech. Anal., Vol 115, No 2 (1991), pp 137-152
- [BV 95] E. BERETTA, M. VOGELIUS
An Inverse Problem Originating from Magnetohydrodynamics III. Domains with corners of arbitrary angles.
A paraître dans Asymptotic Analysis.
- [BW 80] D. M. BATES, D. G. WATTS
Relative Curvature Measures of Nonlinearity
J. R. Statist. Soc., B (1980), 42, No 1, pp 1-25

- [BWb 82] D. M. BATES, G. WAHBA
Computational methods for generalized cross-validation with large data sets
Technical Report No 697 (1982), dpt of Statistics, University of Wisconsin-Madison
- [CEA 71] J. CEA
Optimisation : Théorie et Algorithmes
Ed. Dunod 1971
- [CHA 71] G. CHAVENT
Analyse fonctionnelle et Identification de coefficients répartis dans les équations aux dérivées partielles
Thèse Doctorat ès Sciences, Paris, 1971
- [CHA 74] G. CHAVENT
Identification of Functional Parameters in Partial Differential Equations
Proc. of the Joint Automatic Control Conf., ed. by R. Goodson and M. Polis, the Amer. Soc. of Mech. Eng., Austin, Texas, 1974
- [CHA 79] G. CHAVENT
About the Stability of the Optimal Control Solution of Inverse Problems
Proc. of Conf. on Math. and Num. Methods for Inverse and Improperly Posed Problems, ed. by G. Anger, Akademie Verlag, Berlin (GDR) 1979
- [CHA 82] G. CHAVENT
Local Stability of the Output Least Squares, Parameter Estimation Technique
Report RR No 136, INRIA Rocquencourt, june 1982
- [CHA1 90] G. CHAVENT
A New Sufficient Condition for the Well Posedness of Nonlinear Least Square Problems Arising in Identification and Control
in A. Bensoussan and J.-L. Lions editors, Lectures Notes in Control and Information Sciences 144, Springer 1990, pp 452-463
- [CHA2 90] G. CHAVENT
A Geometrical Theory for Nonlinear Least Squares Problems
Proc. IFIP TC 7, Conference on Control Theory of DPS and Applications, Shanghai, May 6-9 1990, Springer
- [CHA1 91] G. CHAVENT
On the Theory and Practice of Nonlinear Least-Squares
Advances in Water Resources, No 2, April 1991
- [CHA2 91] G. CHAVENT
Quasi-Convex Sets and Size \times Curvature Condition, Application to Nonlinear Inversion
Appl. Math. Opt., 24, pp 129-169, 1991
- [CHA3 91] G. CHAVENT
New Size \times Curvature Conditions for Strict Quasiconvexity of Sets
SIAM J. Contr. Opt., Vol 29, No 6, pp 1348-1372, november 1991
- [CIA 88] P. G. CIARLET
Introduction à l'analyse numérique matricielle et à l'optimisation
Masson 1988

- [CL 74] G. CHAVENT, P. LEMONNIER
Identification de la non linéarité d'une équation parabolique quasi-linéaire
Appl. Math. Opt., Vol 1, No 2, Springer Verlag 1974
- [CR 72] P. G. CIARLET, P. A. RAVIART
General Lagrange and Hermite interpolation in \mathbb{R}^n with applications to finite element methods
Arch. Rat. Mech. Anal., Vol 46 (1972), pp 177-199
- [CW 79] P. CRAVEN, G. WAHBA
Smoothing Noisy Data with Spline Functions
Num. Math. 31 (1979), pp 377-403
- [DAU 88] I. DAUBECHIES
Orthonormal Bases of Compactly Supported Wavelets
Comm. Pure Appl. Math., Vol 41 (1988), pp 909-996
- [DAU 92] I. DAUBECHIES
Ten Lectures on Wavelets
Notes for the CBMS Conference (Lowell), SIAM 61 (1992)
- [DCV 93] I. DAUBECHIES, A. COHEN, P. VIAL
Wavelets on the interval and fast wavelet transforms
Applied and Computational Harmonic Analysis, Vol 1, No 1, pp 54-81, 1993
- [DES 90] L. N. DESHPANDE
Des algorithmes rapides pour la validation croisée sur des problèmes d'approximation non linéaires
Thèse de Doctorat, Université de Grenoble 1, 1990
- [DJ 92] D. L. DONOHO, I. JOHNSTONE
Minimax Estimation via Wavelet Shrinkage
Technical Report, Stanford University, 1992
- [EG 88] H. W. ENGL, H. GFRERER
A Posteriori Parameter Choice for General Regularization Methods for Solving Linear Ill-Posed Problems
Appl. Num. Math., 4 (1988), pp 395-417
- [FAU 86] P. FAURE
Note d'optimisation
Ecole Polytechnique, Dpt de Mathématiques Appliquées, Ellipse 1986
- [GHW 79] G. H. GOLUB, M. HEATH, G. WAHBA
Generalized cross-validation as a method for choosing a good ridge parameter
Technometrics 21 (1979), pp 215-224
- [GIR1 87] D. GIRARD
Un algorithme rapide pour le calcul de la trace de l'inverse d'une grande matrice
Rapport RR 665-M, IMAG, Université de Grenoble 1, mai 1987
- [GIR2 87] D. GIRARD
Un algorithme simple et rapide pour la validation croisée généralisée sur des problèmes de grande taille
Rapport de Recherche, RR 669-M, IMAG, Université de Grenoble, mai 1987

- [GIR 89] D. GIRARD
Asymptotic Optimality of the Fast Randomized Versions of GCV and C_L in Ridge Regression and Regularization
Rapport RR 793-M, TIM3-IMAG, Université de Grenoble 1, octobre 1989
- [GNN 79] B. GIDAS, W. N. NI, L. NIRENBERG
Symmetry and related properties via the maximum principle
Comm. Math. Phys. 68 (1979), pp 209-243
- [HJ 92] P. HALL, I. JOHNSTONE
Empirical Functionals and Efficient Smoothing Parameter Selection
J. R. Stat. Soc., B (1992), 54, No 2, pp 475-530
- [HOU 75] A. S. HOUSEHOLDER
The theory of matrix in numerical analysis
Dover ed. (1975)
- [IK 90] K. ITO, K. KUNISCH
On the Choice of the Regularization Parameter in Nonlinear Inverse Problems
Center for Applied Mathematical Sciences #90-8
- [KS 85] C. KRAVARIS, J. H. SEINFELD
Identification of parameters in distributed parameter systems by regularization
SIAM J. Cont. Opt., Vol 23, No 2, mar 1985, pp 217-241
- [KT 86] N. KERZMAN, M. TRUMMER
Numerical conformal mapping via the Szegö kernel
J. Comp. Appl. Math., 14 (1986), pp 111-123
- [KW 86] K. KUNISCH, L. WHITE
Parameter Estimation for Elliptic Equations in Multidimensional Domains with Point and Flux Observations
Nonlinear Analysis Th., Meth. Appl., Vol 10, No 2 (1986), pp 121-146
- [LI 85] K. C. LI
From Stein's unbiased risk estimates to the method of generalized cross-validation
Ann. Statist. 13 (1985), pp 1352-1377
- [LI 86] K. C. LI
Asymptotic optimality of C_L and generalized cross-validation in ridge regression with application to spline smoothing
Ann. Statist. 14 (1986), pp 1101-1112
- [LIO 68] J. -L. LIONS
Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles
Dunod (1968)
- [LIO 69] J. -L. LIONS
Quelques méthodes de résolution de problèmes aux limites non linéaires
Dunod-Gauthiers Villars, Paris 1969
- [LM 68] J. -L. LIONS, E. MAGENES
Problèmes aux limites non homogènes et applications
Paris, Dunod 1968

- [MA 89] S. MALLAT
Multiresolution Approximations and Wavelet Orthonormal Bases of $L^2(\mathbb{R})$
Transactions of the American Math. Soc., vol. 315, No 1, pp 69-87, 1989
- [MAL 73] C. L. MALLOWS
Some concepts on C_p
Technometrics 15 (1973), pp 661-675
- [MEY 87] Y. MEYER
Ondelettes, fonctions splines et analyse graduée
Cahiers du Ceremade, 8703, Université Paris Dauphine, 1987
- [MEY 92] Y. MEYER
Wavelets and operators
Cambridge Studies in Advanced Mathematics 37 (1992)
- [MOR 84] V. A. MOROZOV
Methods for solving incorrectly posed problems
Springer Verlag, New York (1984)
- [NEC 67] J. NEČAS
Les méthodes directes en théorie des équations elliptiques
Masson 1967
- [PR 69] E. POLAK, G. RIBIERE
Sur la convergence de la méthode des gradients conjugués
Revue Française d'Informatique et de Recherche Opérationnelle, 16-R1 (1969)
- [REI 67] C. REINSCH
Smoothing by spline functions
Num. Math. 10 (1967), pp 177-183
- [RT 88] P. A. RAVIART, J. M. THOMAS
Introduction à l'Analyse numérique des équations aux dérivées partielles
Masson, 1988
- [TA 77] A. N. TIKHONOV, V. Y. ARSENIN
Solutions of Ill-posed Problems
Wiley, New York 1977
- [VOG 94] M. VOGELIUS
An inverse problem for the equation $\Delta u = -cu - d$
Annales de l'Institut Fourier, Tome 44, fascicule 4, pp 1181-1209, 1994
- [WAH 90] G. WAHBA
Spline Models for Observational Data
Notes for the CMBS Conference (Colombus), SIAM 59 (1990)
- [WW 75] G. WAHBA, S. WOLD
A completely automatic french curve : fitting spline functions by cross-validation
Comm. Statist. 4 (1975), pp 1-7

